CONFIGURATION GUIDE

Brocade SLX-OS IP Multicast Configuration Guide, 16r.1.01

Supporting the Brocade SLX 9850 Router

53-1004816-01
11 November 2016

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.

The authors and Brocade Communications Systems, Inc. assume no liability or responsibility to any person or entity with respect to the accuracy of this document or any loss, cost, liability, or damages arising from the information contained herein or the computer programs that accompany it.

The product described by this document may contain open source software covered by the GNU General Public License or other open source license agreements. To find out which open source software is included in Brocade products, view the licensing terms applicable to the open source software, and obtain a copy of the programming source code, please visit http://www.brocade.com/support/oscd.

Contents

Preface

Document conventions

The document conventions describe text formatting conventions, command syntax conventions, and important notice formats used in Brocade technical documentation.

Notes, cautions, and warnings

Notes, cautions, and warning statements may be used in this document. They are listed in the order of increasing severity of potential hazards.

NOTE
A Note provides a tip, guidance, or advice, emphasizes important information, or provides a reference to related information.

ATTENTION
An Attention statement indicates a stronger note, for example, to alert you when traffic might be interrupted or the device might reboot.

CAUTION
A Caution statement alerts you to situations that can be potentially hazardous to you or cause damage to hardware, firmware, software, or data.

DANGER
A Danger statement indicates conditions or situations that can be potentially lethal or extremely hazardous to you. Safety labels are also attached directly to products to warn of these conditions or situations.

Text formatting conventions

Text formatting conventions such as boldface, italic, or Courier font may be used to highlight specific words or phrases.

Format	Description
bold text	Identifies command names. Identifies keywords and operands. Identifies the names of GUI elements. Identifies text to enter in the GUI.
italic text	Identifies emphasis. Identifies variables. Identifies document titles.
`Courier font`	Identifies CLI output.

Format	Description
	Identifies command syntax examples.

Command syntax conventions

Bold and italic text identify command syntax components. Delimiters and operators define groupings of parameters and their logical relationships.

Convention	Description
bold text	Identifies command names, keywords, and command options.
italic text	Identifies a variable.
value	In Fibre Channel products, a fixed value provided as input to a command option is printed in plain text, for example, **--show** WWN.
[ ]	Syntax components displayed within square brackets are optional. Default responses to system prompts are enclosed in square brackets.
{ **x** | **y** | **z** }	A choice of required parameters is enclosed in curly brackets separated by vertical bars. You must select one of the options. In Fibre Channel products, square brackets may be used instead for this purpose.
x | **y**	A vertical bar separates mutually exclusive elements.
< >	Nonprinting characters, for example, passwords, are enclosed in angle brackets.
...	Repeat the previous element, for example, *member*[*member*...].
\	Indicates a "soft" line break in command examples. If a backslash separates two lines of a command input, enter the entire command at the prompt without the backslash.

Brocade resources

Visit the Brocade website to locate related documentation for your product and additional Brocade resources.

White papers, data sheets, and the most recent versions of Brocade software and hardware manuals are available at www.brocade.com. Product documentation for all supported releases is available to registered users at MyBrocade.
Click the **Support** tab and select **Document Library** to access documentation on MyBrocade or www.brocade.com You can locate documentation by product or by operating system.

Release notes are bundled with software downloads on MyBrocade. Links to software downloads are available on the MyBrocade landing page and in the Document Library.

Document feedback

Quality is our first concern at Brocade, and we have made every effort to ensure the accuracy and completeness of this document. However, if you find an error or an omission, or you think that a topic needs further development, we want to hear from you. You can provide feedback in two ways:

- Through the online feedback form in the HTML documents posted on www.brocade.com
- By sending your feedback to documentation@brocade.com

Provide the publication title, part number, and as much detail as possible, including the topic heading and page number if applicable, as well as your suggestions for improvement.

Contacting Brocade Technical Support

As a Brocade customer, you can contact Brocade Technical Support 24x7 online, by telephone, or by e-mail. Brocade OEM customers should contact their OEM/solution provider.

Brocade customers

For product support information and the latest information on contacting the Technical Assistance Center, go to www.brocade.com and select **Support**.

If you have purchased Brocade product support directly from Brocade, use one of the following methods to contact the Brocade Technical Assistance Center 24x7.

Online	Telephone	E-mail
Preferred method of contact for non-urgent issues: • Case management through the MyBrocade portal. • Quick Access links to Knowledge Base, Community, Document Library, Software Downloads and Licensing tools	Required for Sev 1-Critical and Sev 2-High issues: • Continental US: 1-800-752-8061 • Europe, Middle East, Africa, and Asia Pacific: +800-AT FIBREE (+800 28 34 27 33) • Toll-free numbers are available in many countries. • For areas unable to access a toll-free number: +1-408-333-6061	support@brocade.com Please include: • Problem summary • Serial number • Installation details • Environment description

Brocade OEM customers

If you have purchased Brocade product support from a Brocade OEM/solution provider, contact your OEM/solution provider for all of your product support needs.

- OEM/solution providers are trained and certified by Brocade to support Brocade® products.
- Brocade provides backline support for issues that cannot be resolved by the OEM/solution provider.
- Brocade Supplemental Support augments your existing OEM support contract, providing direct access to Brocade expertise. For more information, contact Brocade or your OEM.
- For questions regarding service levels and response times, contact your OEM/solution provider.

About This Document

Supported hardware and software

In those instances in which procedures or parts of procedures documented here apply to some devices but not to others, this guide identifies exactly which devices are supported and which are not.

Although many different software and hardware configurations are tested and supported by Brocade Communications Systems, Inc. for SLX-OS Release 16r.1.01, documenting all possible configurations and scenarios is beyond the scope of this document.

The following hardware platforms are supported by this release:

- Brocade SLX 9850-4 router
- Brocade SLX 9850-8 router

To obtain information about other Brocade OS versions, refer to the documentation specific to that version.

IP Multicast Overview

IP multicast overview

Multicast protocols allow a group or channel to be accessed over different networks by multiple stations (clients) for the receipt and transmission of multicast data. Distribution of stock quotes, video transmissions such as news services and remote classrooms, and video conferencing are all examples of applications that use multicast routing.

Brocade devices support the Protocol-Independent Multicast (PIM) protocol, along with the Internet Group Management Protocol (IGMP).

The Internet Group Management Protocol (IGMP) is used by IP hosts to report their multicast group memberships to any immediately-neighboring multicast routers.

PIM is a broadcast and pruning multicast protocol that delivers IP multicast datagrams. This protocol employs reverse path lookup check and pruning to allow source-specific multicast delivery trees to reach all group members. PIM builds a different multicast tree for each source and destination host group.

IPv4 Multicast Traffic Reduction

IGMP snooping overview

The forwarding of multicast control packets and data through a Layer 2 device configured with VLANs is most easily achieved by the Layer 2 forwarding of received multicast packets on all the member ports of the VLAN interfaces. However, this simple approach is not bandwidth efficient, because only a subset of member ports may be connected to devices interested in receiving those multicast packets. In a worst-case scenario, the data would get forwarded to all port members of a VLAN with a large number of member ports, even if only a single VLAN member is interested in receiving the data. Such scenarios can lead to loss of throughput for a device that gets hit by a high rate of multicast data traffic.

Internet Group Management Protocol (IGMP) snooping is a mechanism by which a Layer 2 device can effectively address this issue of inefficient multicast forwarding to VLAN port members. Snooping involves "learning" forwarding states for multicast data traffic on VLAN port members from the IGMP control (join/leave) packets received on them. The Layer 2 device also provides for a way to configure forwarding states statically through the CLI.

Multicast routing and IGMP snooping

Multicast routers use IGMP snooping to learn which groups have members on each of their attached physical networks. A multicast router keeps a list of multicast group memberships for each attached network, and a timer for each membership.

NOTE
"Multicast group memberships" means that at least one member of a multicast group on a given attached network is available.

There are two ways that hosts join multicast routing groups:

- By sending an unsolicited IGMP join request.
- By sending an IGMP join request as a response to a general query from a multicast router.

In response to the request, the device creates an entry in its Layer 2 forwarding table for that VLAN. When other hosts send join requests for the same multicast, the device adds them to the existing table entry. Only one entry is created per VLAN in the Layer 2 forwarding table for each multicast group.

VLANs can be configured as snooping only or routing with snooping. When Layer 3 multicast routing is enabled on a particular VE, snooping for the underlying VLAN is enabled implicitly. Explicit snooping can be enabled on a VLAN in addition to implicit snooping. Implicit snooping is by default IGMP snooping. With routing enabled on a VE, when explicit snooping is disabled, snooping reverts back to implicit snooping. This does not change the functionality in any way, but only removes the configuration. When routing is disabled on a VE where explicit snooping is configured, the routing side of the programming stops and the snooping side programming takes over. When routing is enabled, the Layer 3 IGMP querier takes precedence on that VLAN. When routing is disabled, and if the snooping querier is configured, then the snooping querier takes effect.

PIM multicast router presence detection

The PIM hello-based multicast router presence detection feature scans the network traffic for incoming PIM hellos.

This feature is enabled when multicast routing or snooping is enabled.

When a PIM hello is detected, that port is marked for the presence of a multicast router and the information is saved. This prevents unnecessary flooding if the PIM designated router (DR) goes offline, as IGMP reports are forwarded to the multicast routers and not only the snooping-enabled router.

Enabling IGMP snooping

Use the following procedure to enable IGMP snooping on a VLAN.

1. Enter the **configure terminal** command to access global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the VLAN configuration mode.

   ```
   device(config)# vlan 1
   device(config-vlan-1)
   ```

3. Enable IGMP snooping.

   ```
   device(config-vlan-1)# ip igmp snooping enable
   ```

Configuring the IGMP snooping querier

If your multicast traffic is not routed because Protocol-Independent Multicast (PIM) is not configured, use the IGMP snooping querier in a VLAN.

The IGMP snooping querier sends out IGMP queries to trigger IGMP responses from devices that are to receive IP multicast traffic. The IGMP snooping querier listens for these responses to map the appropriate forwarding addresses.

Use the following procedure to configure the IGMP snooping querier.

1. Enter the **configure terminal** command to access global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **vlan** command with the VLAN number.

   ```
   device(config)# vlan 25
   ```

3. Set the IGMP query interval for the VLAN.

   ```
   device(config-Vlan-25)# ip igmp snooping query-interval 125
   ```

   The valid range is from 1 through 18000 seconds. The default is 125 seconds.

4. Set the last member query interval.

   ```
   device(config-Vlan-25)# ip igmp snooping last-member-query-interval 1000
   ```

   The valid range is from 1000 through 25500 milliseconds. The default is 1000 milliseconds.

5. Set the Maximum Response Time.

   ```
   device(config-Vlan-25)# ip igmp snooping query-max-response-time 10
   ```

   The valid range is 1 through 25 seconds. The default is 10 seconds.

6. Configure the static Mrouter port.

   ```
   device(config-Vlan-25)# ip igmp snooping mrouter interface ethernet 3/2
   ```

7. Configure a static IGMP group.

   ```
   device(config-vlan-25)# ip igmp snooping static-group 225.0.0.1 interface ethernet 6/15
   ```

8. Configure the IGMP version.

   ```
   device(config-vlan-25)# ip igmp snooping version v2
   ```

9. Activate the IGMP snooping querier functionality for the VLAN.

   ```
   device(config-Vlan-25)# ip igmp snooping querier enable
   ```

   NOTE
   IGMP snooping querier and the static Mrouter can be configured together on a VLAN interface.

Monitoring IGMP snooping

Monitoring the performance of your IGMP traffic allows you to diagnose any potential issues on your device. This helps you utilize bandwidth more efficiently by setting the device to forward IP multicast traffic only to connected hosts that request multicast traffic.

Use the following commands to monitor IGMP snooping on the device; the commands do not need to be entered in any specific order.

1. Enter the **show ip igmp groups** command to display all information on IGMP multicast groups for the device. Use this command to display the IGMP database, including configured entries for either all groups on all interfaces, all groups on specific interfaces, or specific groups on specific interfaces.

   ```
   device# show ip igmp groups
   Total Number of Groups: 2
   IGMP Connected Group Membership
   Group Address  Interface Uptime      Expires      Last Reporter   Version
   225.1.1.1      vlan25    00:05:27     00:02:32     25.1.1.1202
   Member Ports: eth 2/24
   ```

2. Enter the **show ip igmp snooping** command specifying the VLAN ID to view snooping configuration such as snooping querier enable, snooping query-interval, IGMPv2 or v3, PIM snooping configuration, and IGMP snooping configuration.

   ```
   device# show ip igmp snooping vlan 45
   Vlan ID: 45
    Multicast Router ports:  eth3/2
    Querier - Enabled,
    IGMP Operation mode: IGMPv2
    Is Fast-Leave Enabled : Disabled
    Max Response time = 10
    Last Member Query Interval = 1
    Query interval = 125
    Number of Multicast Groups: 1
     Group: 225.0.0.1
     Member Ports:  eth4/22 eth6/15
     Mapped MAC address: 0100.5e00.0001
   ```

3. Enter the **show ip multicast snooping mcache** command to view snooping configuration and PIM snooping configuration information.

```
device# show ip multicast snooping mcache
Flags : V2|V3 : IGMP Receiver, P_G : PIM (*,G) Join, P_SG: PIM (S,G) Join
VlanID : 25
-------------
1(*, 225.1.1.1 )00:02:15NumOIF: 1
Outgoing Ports:
eth2/24       Flags: 0x14 ( V2)  00:02:15/126s
```

4. Enter the **show ip multicast snooping mcache** command specifying the VLAN ID to view combined group membership (IGMP or PIM Snooping) information with member port flags indicating how member ports are learned (host port, pimjoin, igmpreport etc).

```
device# show ip multicast snooping mcache
Flags : V2|V3 : IGMP Receiver, P_G : PIM (*,G) Join, P_SG: PIM (S,G) Join
VlanID : 25
-------------
1(*, 225.1.1.1 )00:02:15NumOIF: 1
Outgoing Ports:
eth2/24 Flags: 0x14 ( V2) 00:02:15/126s
```

5. Enter the **show ip igmp statistics interface** command to display the IGMP statistics for a VLAN or interface.

```
device# show ip igmp statistics interface vlan 1

IGMP packet statistics for all interfaces in vlan 1:
IGMP Message type     Edge-Received   Edge-Sent   Edge-Rx-Errors   ISL Received
Membership Query                  0           0                0              0
V1 Membership Report              0           0                0              0
V2 Membership Report              0           0                0              0
Group Leave                       0           0                0              0
V3 Membership Report              0           0                0              0
PIM hello                         0           0                0              0

IGMP Error Statistics:
Unknown types            0
Bad Length               0
Bad Checksum             0
```

6. Enter the **show ip igmp interface** command to display the Layer 3 IGMP interface configuration information.

```
device# show ip igmp interface
Interface Ve100
IGMP enabled
IGMP query interval 30 seconds
IGMP other-querierinterval 65 seconds
IGMP query response time 10 seconds
IGMP last-member query interval 1 seconds
IGMP immediate-leave disabled
IGMP querier100.0.0.1(this system)
IGMP version 2
```

7. Use the **show ip igmp snooping mrouter vlan** command to display mrouter port-related information.

```
device# show ip igmp snooping mrouter vlan 10
Vlan     Interface    Expires (Sec)
10       eth1/4       250
10       eth1/1       238
```

8. Use the **show ip igmp ssm-map** command to display the SSM mapping with the prefix list name and source address details.

```
device# show ip igmp ssm-map

+------------------------------------+-------------------+
|         PrefixList Name            |   Source Address  |
+------------------------------------+-------------------+
   ssm-map-230-to-232                 203.0.0.10
  ssm-map-233-to-234                 204.0.0.11
```

When you have reviewed the IGMP statistics for the device, refer to Enabling IGMP snooping on page 14 or Configuring the IGMP snooping querier on page 14 to make any needed corrections.

IPv4 Multicast Routing

IGMP

The Internet Group Management Protocol (IGMP) allows an IPv4 system to communicate IP multicast group membership information to its neighboring routers. The routers, in turn, limit the multicast of IP packets with multicast destination addresses to only those interfaces on the router that are identified as IP multicast group members.

In IGMPv2, when a router sends a query to the interfaces, the clients on the interfaces respond with a membership report of multicast groups to the router. The router can then send traffic to these groups, regardless of the traffic source. When an interface no longer needs to receive traffic from a group, it sends a leave message to the router, which in turn sends a group-specific query to that interface to see if any other clients on the same interface are still active.

In contrast, IGMPv3 provides selective filtering of traffic based on the traffic source. A router running IGMPv3 sends queries to every multicast-enabled interface at the specified interval. These general queries determine if any interface wants to receive traffic from the router. The are three variants of the query message:

- A "General Query" is sent by a multicast router to learn the complete multicast reception state of the neighboring interfaces. In a General Query, both the Group Address field and the Number of Sources (N) field are zero.
- A "Group-Specific Query" is sent by a multicast router to learn the reception state, with respect to a *single* multicast address, of the neighboring interfaces. In a Group-Specific Query, the Group Address field contains the multicast address of interest, and the Number of Sources (N) field contains zero.
- A "Group-and-Source-Specific Query" is sent by a multicast router to learn if any neighboring interface desires reception of packets sent to a specified multicast address, from any of a specified list of sources. In a Group-and-Source-Specific Query, the Group Address field contains the multicast address of interest, and the Source Address [i] fields contain the source addresses of interest.

The interfaces respond to these queries by sending a membership report that contains one or more of the following records that are associated with a specific group:

- The current-state record indicates from which sources the interface wants to receive and not receive traffic. The record contains the source address of the interfaces and whether or not traffic will be received or included (IS_IN) or not received or excluded (IS_EX) from that source.
- The filter-mode-change record indicates that if the interface changes its current state from IS_IN to IS_EX, a TO_EX record is included in the membership report. Likewise, if the interface changes its current status from IS_EX to IS_IN, a TO_IN record appears in the membership report.
- The IGMPv2 Leave report is equivalent to a TO_IN (empty) record in IGMPv3. This record indicates that no traffic from this group will be received regardless of the source.

- The IGMPv2 group report is equivalent to an IS_EX (empty) record in IGMPv3. This record indicates that all traffic from this group will be received regardless of the source.
- The source-list-change record indicates that If the interface wants to add or remove traffic sources from its membership report, the membership report can have an ALLOW record, which contains a list of new sources from which the interface wishes to receive traffic. It can also contain a BLOCK record, which lists current traffic sources from which the interface wants to stop receiving traffic.

In response to membership reports from the interfaces, the router sends a Group-Specific Query or a Group-and-Source Specific Query to the multicast interfaces. For example, a router receives a membership report with a source-list-change record to block old sources from an interface. The router sends Group-and-Source Specific Queries to the source-group pair (S,G) identified in the record. If none of the interfaces is interested in the (S,G), it is removed from the (S,G) list for that interface on the router.

Each IGMPv3-enabled router maintains a record of the state of each group and each physical port within a virtual routing interface. This record contains the group, group-timer, filter mode, and source records information for the group or interface. Source records contain information on the source address of the packet and source timer. If the source timer expires when the state of the group or interface is in include mode, the record is removed.

Default IGMP version

IGMP v2 is enabled by default only when snooping or multicast routing are enabled on the system.

Also, you can specify what version of IGMP you want to run on a device on a per-VLAN basis. You can change the IGMP version for router ports, but not for Ve interfaces. If you do not specify an IGMP version, IGMPv2 is used.

Compatibility with IGMPv1 and IGMPv2

Different multicast groups, interfaces, and routers can run their own versions of IGMP. The version of IGMP is reflected in the membership reports that the hosts send to the router. Routers and interfaces must be configured to recognize the version of IGMP you want them to process.

An interface or router sends the queries and reports that include its IGMP version specified on it. The interface may recognize a query or report that has a different version. For example, an interface running IGMPv2 can recognize IGMPv3 packets, but cannot process them. When the router sends out IGMP queries over an IGMPv2 interface, the equal or lower version of reports are supported, but higher version of reports are not supported.

Reports sent by interfaces to routers that contain different versions of IGMP do not trigger warning messages; however, you can see the versions of the packets by using the **show ip igmp traffic** command.

The version of IGMP can be specified per interface (physical port or virtual routing interface), and per physical port within a virtual routing interface.

The IGMP version set on a Layer 3 physical interface or under a VLAN of the virtual routing interface supersedes the version set on a physical or virtual routing interface.

Likewise, the version on a physical or virtual routing interface supersedes the version set globally on the device.

Enabling the IGMP version

You can enable or change the IGMP version per interface or VLAN setting.

1. Enter global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the interface configuration mode.

   ```
   device(config)# interface ethernet 1/5
   ```

3. Enter the **ip igmp version** command.

   ```
   device(config-if-1/5)# ip igmp version 3
   ```

Configuring RA option disable

RA (router alert) option disable can be configured at the global level.

The router alert disable option disables the snooping check for the presence of the router alert option. By default, IGMP snooping checks for the presence of the router alert option in the IP packet header of the IGMP message. Packets that do not include this option are dropped.

1. Enter the global configuration mode.

   ```
   device# configure terminal
   ```

2. To disable the RA option, enter the **router-alert-check-disable** command.

   ```
   device(config)# ip igmp router-alert-check-disable
   ```

Configuring IGMPv2 SSM mapping

The **ip ssm-map** commands can be used to enable the IGMPv2 mapping feature and to define the maps between IGMPv2 group addresses and multicast source addresses.

The PIM-SSM feature requires all IGMP hosts to send IGMPv3 reports. Where you have an IGMPv2 host, this can create a compatibility problem. In particular, the reports from an IGMPv2 host contain a group multicast address but do not contain source addresses. The IGMPv3 reports contain both the group multicast address and one or more source addresses. This feature converts IGMPv2 reports into IGMPv3 reports through use of the ip igmp ssm-map commands and a configured prefix list.

The prefix list used with this feature filters for the group multicast address. The prefix list is then associated with one or more source addresses. When the **ip igmp ssm-map enable** command is configured, IGMPv3 reports are sent for IGMPv2 hosts.

1. Enter global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **ip igmp ssm-map enable** command to enable the IGMPv2 mapping.

   ```
   device(config)# ip igmp ssm-map enable
   ```

The following example configures the SSM map at the global configuration level.

```
device(config)# ip igmp ssm-map enable
device(config)# ip igmp ssm-map ssm-map-230-to-232 203.0.0.10
device(config)# ip igmp ssm-map ssm-map-233-to-234 204.0.0.10
```

The following example configures the prefix list for an SSM range.

```
device(config)# ip prefix-list ssm-map-230-to-232 seq 5 permit 230.0.0.0/8
device(config)# ip prefix-list ssm-map-230-to-232 seq 10 permit 231.0.0.0/8
device(config)# ip prefix-list ssm-map-230-to-232 seq 15 permit 232.0.0.0/8

device(config)# ip prefix-list ssm-map-233-to-234 seq 5 permit 233.0.0.0/8
device(config)# ip prefix-list ssm-map-233-to-234 seq 10 permit 234.0.0.0/8
device(config)# ip prefix-list ssm-map-230-to-232 seq 15 permit 232.0.0.0/8
```

Protocol-Independent Multicast overview

The Protocol-Independent Multicast (PIM) protocol is a family of IPv4 multicast protocols. PIM does not rely on any particular routing protocol for creating its network topology state. Instead, PIM uses routing information supplied by other traditional routing protocols, such as Open Shortest Path First, Border Gateway Protocol, and Multicast Source Discovery Protocol.

PIM messages are sent encapsulated in an IP packet with the IP protocol field set to 103. Depending on the type of message, the packet is either sent to the PIM All-Router-Multicast address (224.0.0.13) or sent as unicast to a specific host.

As with IP multicast, the main use case of PIM is for the source to be able to send the same information to multiple receivers by using a single stream of traffic. This helps minimize the processing load on the source, as the source needs to maintain only one session irrespective of the number of actual receivers. It also minimizes the load on the IP network, because the packets are sent only on links that lead to an interested receiver.

Several types of PIM exist, but Brocade supports only PIM sparse mode (PIM-SM, and PIM-SSM). PIM-sparse explicitly builds unidirectional shared trees rooted at a rendezvous point (RP) per group, and optionally creates shortest-path trees per source.

Enabling PIM on a router

Use the following procedure to enable PIM globally.

1. Enter global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **router pim** command to enter the PIM router configuration mode and configure a variety of options.

   ```
   device(config)# router pim
   device(config-pim-router)#
   ```

Configuring PIM

Once you enable PIM on a device, you can configure a variety of options in the router PIM configuration mode.

1. Enter the **hello-interval** command to configure the PIM hello timeout.

   ```
   device(config-pim-router)# hello-interval 40
   ```

2. Enter the **nbr-timeout** command to configure the PIM neighbor timeout.

   ```
   device(config-pim-router)# nbr-timeout 160
   ```

3. Enter the **bsr-candidate** command to configure the BSR candidate.

   ```
   device(config-pim-router)# bsr-candidate interface loopback 11 mask 32
   ```

4. Enter the **prune-wait** command to configure the PIM prune pending timeout.

   ```
   device(config-pim-router)# prune-wait 5
   ```

5. Enter the **message-interval** command to configure the PIM join or prune interval.

   ```
   device(config-pim-router)# message-interval 180
   ```

6. Enter the **spt-threshold** command to configure the PIM Shortest Path Tree (SPT) threshold.

   ```
   device(config-pim-router)# spt-threshold 10
   ```

7. Enter the **ssm-enable range** command to set the multicast address range to use for SSM.

   ```
   device(config-pim-router)#  ssm-enable range PL_ssm_range-230-to-234
   ```

   Entering only the **ssm-enable** command applies the 232.0.0.0/8 default SSM range. This default range is displayed in the **show ip pim settings** output.

PIM-sparse overview

PIM-sparse is most effective in large networks sparsely populated with hosts interested in multicast traffic, with most hosts not interested in all multicast data streams.

PIM-sparse devices are organized into domains. A PIM-sparse domain is a contiguous set of devices that all implement PIM and are configured to operate within a common boundary.

PIM-sparse creates unidirectional shared trees that are rooted at a common node in the network called the rendezvous point (RP). The RP acts as the messenger between the source and the interested hosts or routers. There are various ways of identifying an RP within a network. An RP can be configured either statically per PIM router, or by means of a bootstrap router (BSR). Within a network, the RP must always be upstream from the destination hosts.

Once the RP is identified, interested hosts and routers send join messages to the RP for the group in which they are interested. To reduce the number of Join messages incoming to an RP, the local network selects one of its upstream routers as the designated router (DR). All hosts below a DR send IGMP join messages to the DR. The DR sends only one join message to the RP on behalf of all its interested hosts.

PIM-sparse also provides the option of creating a source-based tree rooted at a router adjacent to the tree. This provides the destination hosts with an option of switching from the shared tree to the source-based tree if the latter has a shorter path between the source and the destination.

PIM-sparse device types

Devices configured with PIM-sparse interfaces also can be configured to fill one or more of the following roles:

- Bootstrap router (BSR): A router that distributes rendezvous point (RP) information to the other PIM-sparse devices within the domain. Each PIM-sparse domain has one active BSR. For redundancy, you can configure ports on multiple devices as candidate BSRs. The PIM-sparse protocol uses an election process to select one of the candidate BSRs as the BSR for the domain. The BSR with the highest BSR priority (a user-configurable parameter) is elected. If the priorities result in a tie, then the candidate BSR interface with the highest IP address is elected.

  The BSR must be configured as part of the Layer 3 core network.

- Rendezvous point (RP): The meeting point for PIM-sparse sources and receivers. A PIM-sparse domain can have multiple RPs, but each PIM-sparse multicast group address can have only one active RP. PIM-sparse devices learn the addresses of RPs and the groups for which they are responsible from messages that the BSR sends to each of the PIM-sparse devices.

  The RP must be configured as part of the Layer 3 core network.

  > NOTE
  > Brocade recommends that you configure the same ports as candidate BSRs and RPs.

- PIM designated router (DR): Once the RP has been identified, each interested host or router sends join messages to the RP for the group in which they are interested. The local network selects one of its upstream routers as the DR. All hosts below a DR send IGMP join messages to the DR. The DR sends only one join message to the RP on behalf of all its interested hosts. The RP receives the first few packets of the multicast stream, encapsulated in the PIM register message, from the source hosts. These messages are sent as a unicast to the RP. The RP de-encapsulates these packets and forwards them to the respective DRs.

  > NOTE
  > DR election is based first on the router with the highest configured DR priority for an interface (if DR priority has been configured), and based next on the router with the highest IP address. To configure DR priority, use the **ip pim dr-priority** command.

Enabling PIM-sparse on routed interfaces

The following procedure enables PIM-sparse and other options on supported interfaces.

You must enable PIM globally before enabling PIM sparse on the interface.

1. To enable PIM-sparse on an interface (Ethernet, loopback, or VE), enter the global configuration mode.

   ```
   device# configure terminal
   ```

2. In global configuration mode, specify an interface.

   ```
   device(config)# interface ethernet 1/1
   ```

3. Enter the **ip pim-sparse** command in the interface configuration mode.

   ```
   device(conf-if-eth-1/1)# ip pim-sparse
   ```

4. (Optional) To change the designated router (DR) priority from the default, enter the **ip pim dr-priority** command in interface subtype configuration mode and specify a non-default value:

   ```
   device(conf-if-eth-1/1# ip pim dr-priority 200
   ```

5. (Optional) To set the TTL threshold, enter the **ip pim ttl-threshold** command in the interface configuration mode.

   ```
   device(conf-if-eth-1/1# ip pim ttl-threshold 50
   ```

Configuring PIM RP

You can use the PIM Sparse protocol's RP election process so that a backup RP can automatically take over if the active RP router becomes unavailable.

However, if you do not want the RP to be selected by the RP election process but want to explicitly identify the RP by address, use the **rp-address** command.

If you explicitly specify the RP, the device uses the specified RP for all group-to-RP mappings and overrides the set of candidate RPs supplied by the BSR.

1. Enter the **configure terminal** command to enter global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **router pim** command to enter the router PIM configuration mode.

   ```
   device(config)# router pim
   ```

3. Enter the **rp-address** command to specify the IP address of the RP.

   ```
   device(config-pim-router)# rp-address 4.4.4.4
   ```

   The command in this example identifies the device interface at IP address 4.4.4.4 as the RP for the PIM-sparse domain. The device uses the specified RP and ignores group-to-RP mappings received from the BSR.
4. For static RP configuration with specific group ranges, enter the following commands.

   ```
   device(config-pim-router)# rp-address 4.4.4.4 static-rp-list
   device(config)# ip prefix-list static-rp-list permit 225.1.1.0/24
   ```

   The following commands configure the RP candidate.

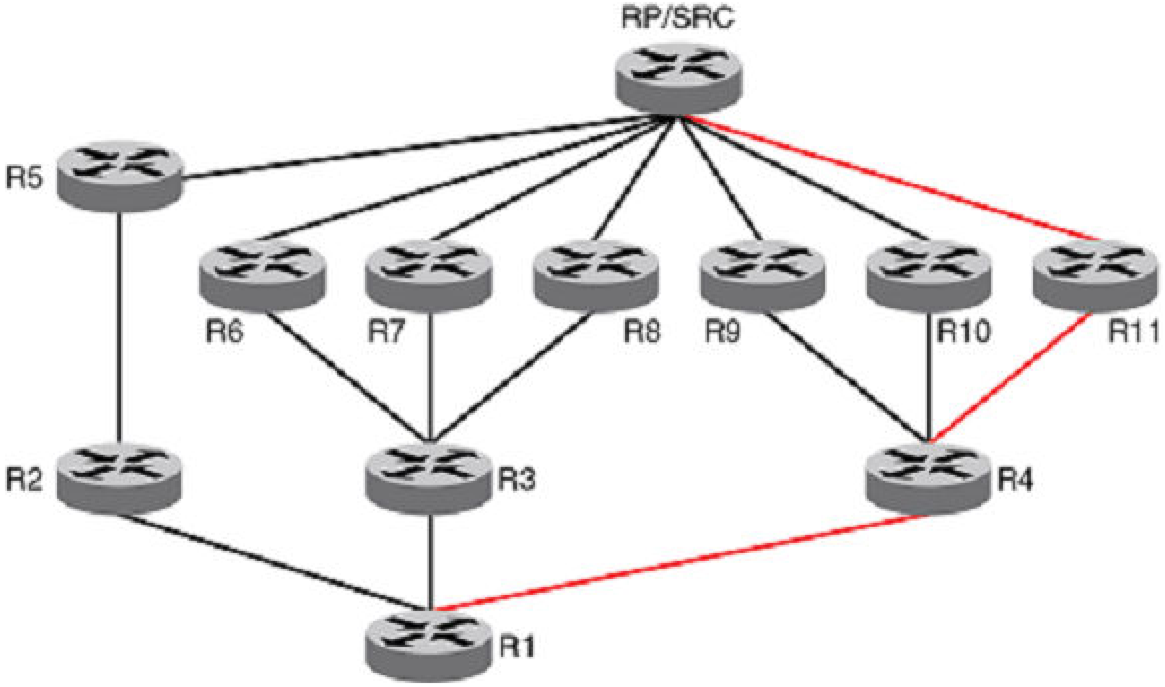
   ```
   device(config-pim-router)# rp-candidate interface loopback 11
   device(config-pim-router)# rp-candidate prefix my-rp-cand-list
   device(config)# ip prefix-list my-rp-cand-list permit 226.1.1.0/24
   device(config)# ip prefix-list my-rp-cand-list permit 228.1.1.0/24
   ```

Multicast ECMP support

If there are multiple equal cost paths between PIM routers to reach the source or the RP, the multicast RPF algorithms distribute the load across available paths to take advantage of those paths.

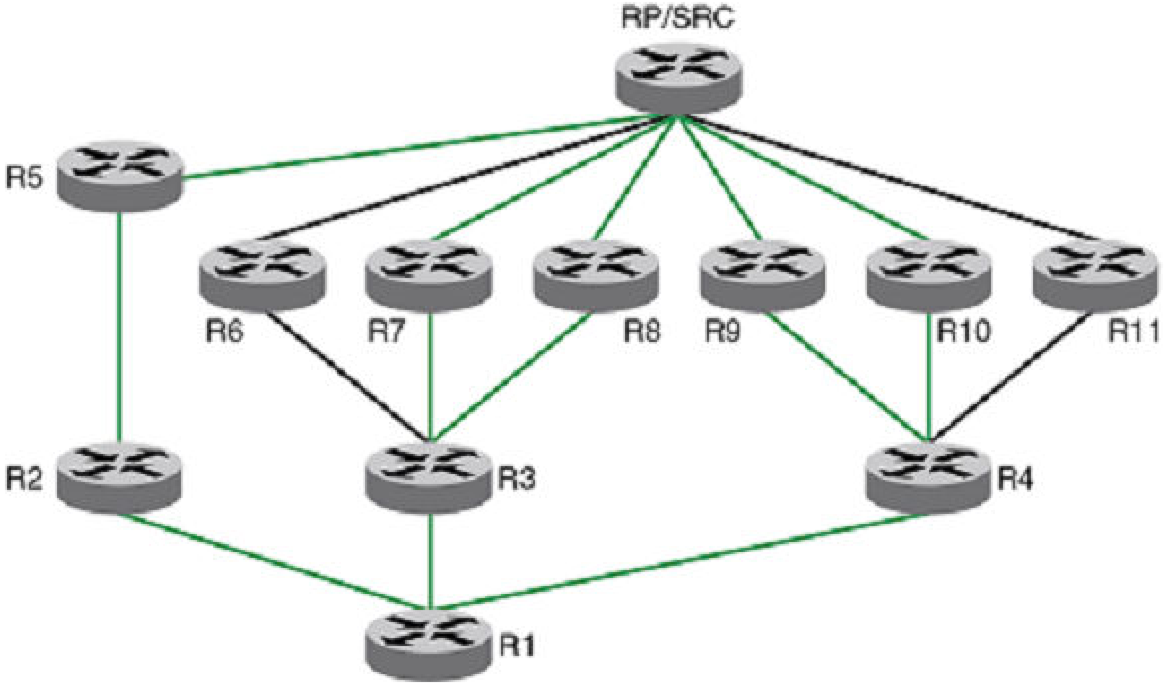
Figure 1 shows a topology in which R1 through R11 have IP addresses in ascending order (R1 having the lowest IP address and R11 having the highest). All the routers are PIM-enabled routers. The links emanating from each router are equal-cost multi-path (ECMP) links. The existing behavior path utilization is indicated in red. With the highest IP address neighbor chosen for the ECMP paths available, the multicast cache entries utilize only the R1-R4-R11-SRC/RP path.

FIGURE 1 Path utilization without multicast ECMP support



In the following figure, with the ECMP support turned on, the multicast entries will be distributed among the equal-cost next hops as indicated in green for better utilization of the available paths.

FIGURE 2 Path utilization with multicast ECMP support



The load distribution is achieved by distributing the multicast cache entries (*,G or S,G) to the available paths, and thus distributing the traffic. Two different methods are widely used to achieve this distribution:

- Hash based - Load splitting
- Least used path based - Load balancing

Brocade devices support the Hash based method of load distribution for multicast ECMP.

Hash based load distribution

The hash based load distribution depends on a hash function to distribute the multicast cache entries. The S, G, next-hop addresses are hash function based. This method splits the cache entries by choosing a different RPF neighbor and splits the traffic. Load balancing is based on the distribution of the keys S, G, next-hop. This method of distribution is the least disruptive as the hashing redistributes only those cache entries that are affected during link flaps. Some paths may not be utilized for the distribution of the multicast entries. For example, for the ECMP paths from R3 to R6, R7 and R8, only paths R3 to R7 and R3 to R8 are being utilized.

Deleting a path

When an ECMP path goes down, all the multicast entries using that path get redistributed among the other available paths.

Adding a path

When a new path is added to the ECMP set, there is no redistribution (default behavior without rebalance option) of the cache entries. Here optimal utilization of the paths is traded off in favor of not disturbing the existing flow. This method also requires a full branch setup towards the source or RP of the multicast distribution tree sometimes. When a path flaps (goes down and comes back up), the multicast entries which had been using this path would not be using this path anymore and it becomes worse if a subset of paths go down and come back up one by one, resulting in only the paths that did not flap to carry all entries.

Dynamic rebalancing

This option rebalances the traffic immediately on a new next-hop or path addition and helps in both new next-hop and path addition and path flap cases. There is least disruption in existing flows by using the hashing method.

Limitations and prerequisites

The following limitations and prerequisites apply to the configuration of ECMP path load balancing:

- The hash method is a load splitting method and hence traffic load balancing is not supported.
- S-based and S,G based hashing is not supported.
- The hash method is a load splitting method and not a load balancing method and hence the load balancing effect due to load splitting the multicast entries is only a best effort and the splitting is actually based on the number of S, G flows and the number of next-hops and the actual distribution of the S,G and the next-hop addresses.
- If the rebalancing is not configured, then link flap results in sub-optimal utilization of the ECMP links.
- The number of paths supported by multicast ECMP would be the same as unicast ECMP which is 32.

Enabling ECMP dynamic rebalance

Enabling ECMP dynamic rebalance configures the hash based distribution among the ECMP paths.

The **rebalance** option enables redistributing the load when a new next-hop is added. The redistribution is based on the hash function.

1. Enter the **configure terminal** command to enter the global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **router pim** command to enter the router PIM configuration mode.

   ```
   device(config)# router pim
   ```

3. Enter the **rpf ecmp** command to enable ECMP load sharing.

   ```
   device(config-pim-router)# rpf ecmp
   ```

4. Enter the **rpf ecmp rebalance** command to enable load sharing with dynamic rebalance.

   ```
   device(config-pim-router)# rpf ecmp rebalance
   ```

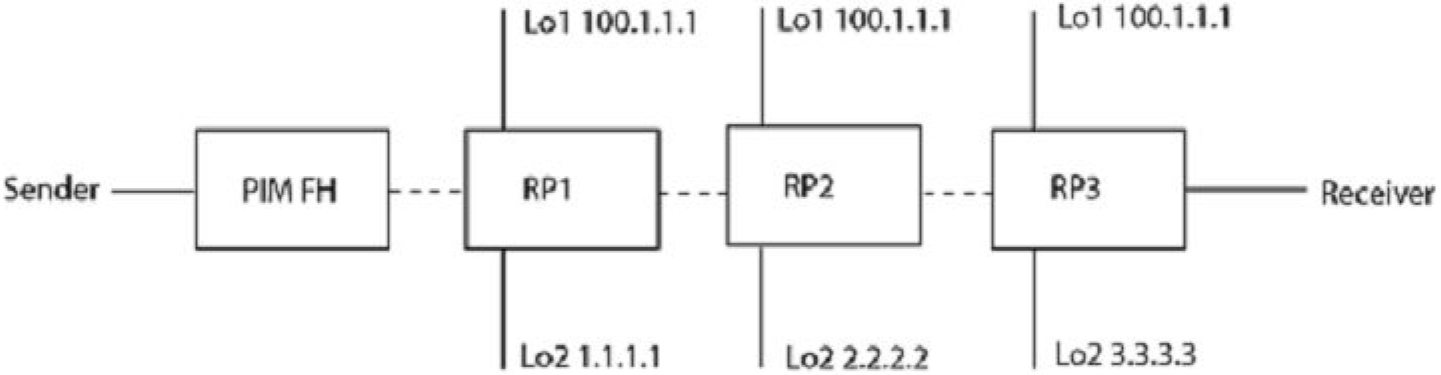
PIM Anycast RP

PIM Anycast RP is a method of providing load balancing and fast convergence to PIM RPs in an IPv4 multicast domain. The RP address of the Anycast RP is a shared address used among multiple PIM routers, known as PIM RP. The PIM RP routers create an Anycast RP set. Each router in the Anycast RP set is configured using two IP addresses; a shared RP address in their loopback address and a separate, unique ip address. The loopback address must be reachable by all PIM routers in the multicast domain. The separate, unique ip address is configured to establish static peering with other PIM routers and communication with the peers.

When the source is activated in a PIM Anycast RP domain**,** the PIM First Hop (FH) will register the source to the closet PIM RP. The PIM RP follows the same MSDP Anycast RP operation by decapsulating the packet and creating the (s,g) state. If there are external peers in the Anycast RP set, the router will re-encapsulate the packet with the local peering address as the source address of the encapsulation. The router will unicast the packet to all Anycast RP peers. The re-encapsulation of the data register packet to Anycast RP peers ensures source state distribution to all RPs in a multicast domain.

The example shown in the figure below is a PIM Anycast-enabled network with 3 RPs, 1 PIM-FH router connecting to its active source and local receiver. Loopback 1 in RP1, RP2, and RP3 have the same IP addresses 100.1.1.1. Loopback 2 in RP1, RP2, and RP3 each have separate IP addresses configured to communicate with their peers in the Anycast RP set.

FIGURE 3 Example of a PIM Anycast RP network



Configuring PIM Anycast RP

The PIM CLI specifies mapping of the RP and the Anycast RP peers.

1. Enter the **configure terminal** command to enter the global configuration mode.

   ```
   device# configure terminal
   ```

2. Enter the **router pim** command to enter the router PIM configuration mode.

   ```
   device(config)# router pim
   ```

3. Enter the **rp-address** command followed by the IP address to be configured as the RP for the PIM Sparse domain.

   ```
   device(config-pim-router)# rp-address 100.1.1.1
   ```

4. Enter the **anycast-rp** command followed by the RP address and the **anycast-rp-set** parameter, which specifies a host based simple prefix list name used to specify the address of the Anycast RP set, including a local address.

   ```
   device(config-pim-router)# anycast-rp 100.1.1.1 anycast-rp-set
   ```

The following example is a configuration of PIM Anycast RP 100.1.1.1. The example avoids using loopback 1 interface when configuring PIM Anycast RP because the loopback 1 address could be used as a router-id. A PIM First Hop router will register the source with the closest RP. The first RP that receives the register will re-encapsulate the register to all other Anycast RP peers.

The RP shared address 100.1.1.1 is used in the PIM domain. IP addresses 1.1.1.1, 2.2.2.2, and 3.3.3.3 are listed in the ACL that forms the self inclusive Anycast RP set. Multiple anycast-rp instances can be configured on a system; each peer with the same or different Anycast RP set.

```
device(config)# interface loopback 2
device(config-lbif-2)# ip address 100.1.1.1/24
device(config-lbif-2)# ip pim-sparse
device(config-lbif-2)# interface loopback 3
device(config-lbif-3)# ip address 1.1.1.1/24
device(config-lbif-3)# ip pim-sparse
device(config-lbif-3)# router pim
device(config-pim-router)# rp-address 100.1.1.1
device(config-pim-router)# anycast-rp 100.1.1.1 anycast-rp-set
device(config)# ip prefix-list anycast-rp-set permit 1.1.1.1/32
device(config)# ip prefix-list anycast-rp-set permit 2.2.2.2/32
device(config)# ip prefix-list anycast-rp-set permit 3.3.3.3/32
```

Displaying PIM information

You can use several show commands to view information about PIM.

Use one of the following commands to view PIM information. The commands do not need to be entered in the specified order.

1. Enter the **show ip pim settings** command.

```
device# show ip pim settings
  Maximum mcache              : 24576     Current Count                : 0
  Hello interval              : 30        Neighbor timeout             : 105
  Join/Prune interval         : 60        Inactivity interval          : 180
  Hardware drop enabled       : 1         Prune wait interval          : 3
  Register Suppress Time      : 60        Register Probe Time          : 10
  Register Stop Delay         : 0         Register Suppress interval   : 0
  SSM Enabled                 : No        SPT Threshold                : 1
  Route Precedence            : uc-non-default uc-default none
```

2. Enter the **show ip pim mcache** command.

```
device# show ip pim mcache 50.1.1.101 230.1.1.1
IP Multicast Mcache Table
Entry Flags    : sm  - Sparse Mode, ssm - Source Specific Multicast
                 RPT - RPT Bit, SPT - SPT Bit, LSrc - Local Source
                 LRcv - Local Receiver, RegProbe - Register In Progress
                 RegSupp - Register Suppression Timer, Reg - Register Complete
                 needRte - Route Required for Src/RP
Interface Flags: IM - Immediate, IH - Inherited, WA - Won Assert
                 MJ - Membership Join, BR - Blocked RPT, BA - Blocked Assert
                 BF - Blocked Filter
Total entries in mcache: 8
1    (50.1.1.101, 230.1.1.1) in Ve 40, Uptime 00:03:29
     Sparse Mode, RPT=0 SPT=1 Reg=0 RegSupp=0 RegProbe=0 LSrc=0 LRcv=1
     upstream neighbor=40.1.1.3
     num_oifs = 2
           Ve 2(00:03:29/181) Flags: IM
           Ve 10(00:03:29/0) Flags: MJ
     Flags (0x400784d1)
         sm=1 ssm=0 needRte=0
```

   The output of this command displays the multicast Mcache table.

3. Enter the **show ip pim traffic** command to display IPv4 traffic statistics.

```
device# show ip pim traffic
Port      |HELLO  |JOIN  |PRUNE  |ASSERT |GRAFT/REGISTER |REGISTER-STOP |BSR-MSGS   |RPC-MSGS
          |Rx     |Rx    |Rx     |Rx     |Rx             |Rx            |Rx         |Rx
------------+---------+---------+---------+---------+---------+---------+---------+---------+
Ve10         54       0       0       0        0           0              0          0
Lo 1         0        0       0       0        0           0              0          0

device# show ip pim traffic
Port        |HELLO |JOIN  |PRUNE |ASSERT |GRAFT/REGISTER |REGISTER-STOP |BSR-MSGS |RPC-MSGS
            | Tx   |Tx    | Tx   |Tx     |Tx             |Tx
------------+---------+---------+---------+---------+---------+---------+---------+---------+---
Ve10         29         0     0        0         0            0             0         0
Lo 1         28         0     0        0         0            0             0         0
```

   The output of this command displays the Protocol Independent Multicast (PIM) traffic statistics categorized by each PIM enabled interface.

4. Enter the **show ip pim neighbor** command to display PIM neighbor information.

```
device(config)# show ip pim neighbor
-----+--------+------------+--------+---+---------+---------+-----+----------+-----------+----
Port |PhyPort |Neighbor    |Holdtime|T  |PropDelay|Override |Age  |UpTime    |VRF        |Prio
     |        |            |sec     |Bit|msec     |msec     |sec  |          |           |
-----+--------+------------+--------+---+---------+---------+-----+----------+-----------+----
v2     e1/1     2.1.1.2     105      1   500       3000      0     00:44:10   default-vrf  1
v4     e1/2     4.1.1.2     105      1   500       3000      10    00:42:50   default-vrf  1
v5     e1/1     5.1.1.2     105      1   500       3000      0     00:44:00   default-vrf  1
v22    e1/1     22.1.1.1    105      1   500       3000      0     00:44:10   default-vrf  1
Total Number of Neighbors : 4
```

5. Enter the **show ip pim bsr** command to display the bootstrap router information.

```
device# show ip pim bsr
PIMv2 Bootstrap information for Vrf Instance : default-vrf
------------------------------------------------------------------------------
  This system is the Elected BSR
  BSR address: 1.51.51.1. Hash Mask Length 32. Priority 255.
  Next bootstrap message in 00:01:00
  Configuration:
    Candidate loopback 2 (Address 1.51.51.1). Hash Mask Length 32. Priority 255.
Next Candidate-RP-advertisment in 00:01:00
  RP: 1.51.51.1
    group prefixes:
    224.0.0.0 / 4
  Candidate-RP-advertisement period: 60
```

6. Enter the **show ip pim rp-candidate** to display the rendezvous point (RP) information.

```
device# show ip pim rp-candidate
Next Candidate-RP-advertisement in 00:00:10
  RP: 207.95.7.1
    group prefixes:
    224.0.0.0 / 4
  Candidate-RP-advertisement period: 60
```

PIM multinet

Brocade devices support PIM over secondary addresses in an IPv4 environment by configuring an IPv4 address with a secondary keyword.

Whenever a secondary address is configured on a interface, all the secondary addresses configured on the interface are sent out on the PIM Hello using the secondary address option.

Whenever a receiver uses a secondary address as its source and sends a IGMP group report, the PIM join and prunes are propagated up the network.

Whenever a secondary address is configured as a RP, the packets are processed appropriately

Displaying the secondary address

In this example the PIM neighbor on Ve10 has multiple IP addresses configured on the interface.

```
device# show ip pim neighbor
Total Number of Neighbors : 1
Port         Phy_Port     Neighbor        Holdtime Age            UpTime  Priority
                                          sec      sec      Dd HH:MM:SS
Ve10         Ve10         10.10.10.17     105      10          00:26:10         1
                         +20.20.20.21
```

Mtrace overview

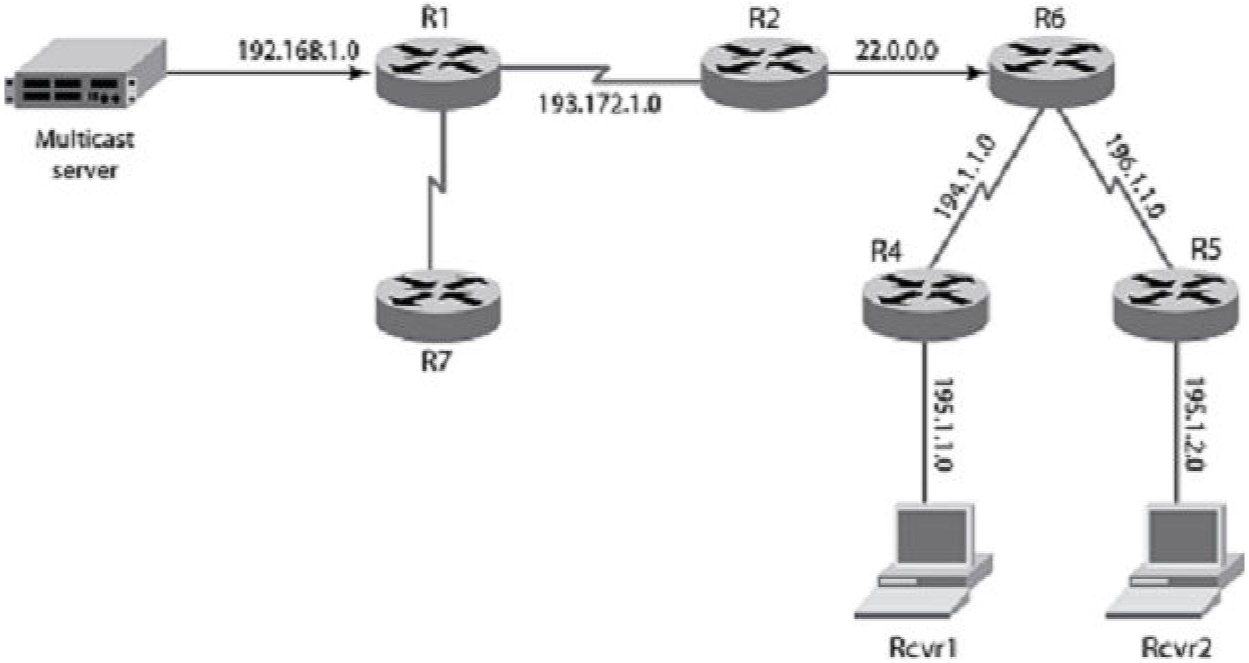
Mtrace is a diagnostic tool to trace the multicast path from a specified destination to a source for a multicast group. It runs over IGMP protocol. Mtrace uses any information available to it to determine a previous hop to forward the trace towards the source.

There are three main components in an Mtrace implementation. They are mtrace query, mtrace request, and mtrace response.

The unicast traceroute program allows the tracing of a path from one machine to another. The key mechanism for unicast traceroute is the ICMP TTL exceeded message, which is specifically excluded as a response to multicast packets. The multicast traceroute facility allows the tracing of an IP multicast routing path. Multicast traceroute also requires special implementations on the part of routers.

Multicast traceroute uses any information available to it in the router to determine a previous hop to forward the trace towards the source. Multicast routing protocols vary in the type and amount of state they keep; multicast traceroute endeavors to work with all of them by using whatever is available. For example, if a PIM-SM router is on the (*,G) tree, it chooses the parent towards the RP as the previous hop. In these cases, no source/group-specific state is available, but the path may still be traced.

FIGURE 4 Network topology



Mtrace components

There are 3 main components in a multicast traceroute implementation. They are:

1. Mtrace Query
2. Mtrace Request
3. Mtrace Response

- Mtrace Query

The party requesting the traceroute sends a traceroute query packet to the last-hop multicast router for the given destination. The query and request have the same opcode, the receiving router can distinguish between a query and a request by checking the size of the packet. A query is a request packet with none of the response fields filled up.

- Mtrace Request

  The last-hop router turns the Query packet into a Request packet by adding a response data block containing its interface addresses and packet statistics, and then forwards the Request packet via unicast to the router that it believes is the proper previous hop for the given source and group. Each hop adds its response data to the end of the Request packet, then unicast forwards it to the previous hop.

- Mtrace Response

  The first hop router (the router that believes that packets from the source originate on one of its directly connected networks) changes the packet type to indicate a Response packet and sends the completed response to the response destination address. The response may be returned before reaching the first hop router if a fatal error condition such as "no route" is encountered along the path.

Configuring mtrace

The mtrace can be started on any router on the network.

Assume that the destination is 195.1.2.1, source is 192.168.1.1 and group is 225.1.1.1. The mtrace query is initially sent from R7. The initial header is not to be modified by any of the routers. R5 adds a response block based on the (S, G) or the (*, G) entry and adds its incoming interface, outgoing interface and other information specified in the draft and sends it to its upstream neighbor which is R6. R6 similarly adds a response block and sends it to its upstream neighbor R2, likewise till it reaches R1. Once it reaches R1, R1 determines that it is the first hop router and completes the response block and sends the response back to R7. R7 now reads the information from the packet and prints it out.

Enter Privileged EXEC mode and enter the **mtrace** command followed by the source, destination and group IP address.

```
device# mtrace source 20.1.1.2 destination 155.1.1.1 group 225.0.0.1
```

The following output displays:

```
Mtrace handle query from src 20.1.1.2  to dest 155.1.1.1  through group 225.0.0.1

Collecting Statistics, waiting time 5 seconds.....

Type Control-c to abort 0 12::1 PIM thresh^ 1 MTRACE_NO_ERR 1 13::1 PIM thresh^ 1 MTRACE_NO_ERR 2 102::2
PIM thresh^ 1 MTRACE_REACHED_RP
```