



Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

This module describes how to optimize Protocol Independent Multicast (PIM) sparse mode for a large deployment of IP multicast. You can set a limit on the rate of PIM register messages sent in order to limit the load on the designated router and RP, you can reduce the PIM router query message interval to achieve faster convergence, and you can delay or prevent the use of the shortest path tree.

Module History

This module was first published on May 2, 2005, and last updated on May 2, 2005.

Finding Feature Information in This Document

Not all features may be supported in your Cisco IOS software release. Use the [“Feature Information for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment”](#) to find information about feature support and configuration.

Contents

- [Prerequisites for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment, page 1](#)
- [Information About Optimizing PIM Sparse Mode in a Large IP Multicast Deployment, page 2](#)
- [How to Optimize PIM Sparse Mode in a Large IP Multicast Deployment, page 5](#)
- [Configuration Examples for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment, page 7](#)
- [Additional References, page 7](#)
- [Feature Information for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment, page 8](#)

Prerequisites for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

This module assumes you have met the following prerequisites:



Americas Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

© 2007 Cisco Systems, Inc. All rights reserved.

- You have PIM sparse mode running in your network.
- You understand the concepts in the “IP Multicast Technology Overview” module.
- If you plan to use a group list to control which groups the shortest-path tree (SPT) threshold applies to, you have configured your access list before performing the task.

Information About Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

Before optimizing PIM sparse mode for a large deployment, you should understand the following concepts:

- [PIM Registering Process, page 2](#)
- [PIM Designated Router, page 3](#)
- [PIM Sparse-Mode Register Messages, page 3](#)
- [Preventing Use of Shortest-Path Tree to Reduce Memory Requirement, page 3](#)

PIM Registering Process

IP multicast sources do not use a signaling mechanism to announce their presence. Sources just send their data into the attached network, as opposed to receivers that use Internet Group Management Protocol (IGMP) to announce their presence. If a source sends traffic to a multicast group configured in PIM sparse mode (PIM-SM), the Designated Router (DR) leading toward the source must inform the rendezvous point (RP) about the presence of this source. If the RP has downstream receivers that want to receive the multicast traffic (natively) from this source and has not joined the shortest path leading toward the source, then the DR must send the traffic from the source to the RP. The PIM registering process, which is individually run for each (S, G) entry, accomplishes these tasks between the DR and RP.

The registering process begins when a DR creates a new (S, G) state. The DR encapsulates all the data packets that match the (S, G) state into PIM register messages and unicasts those register messages to the RP.

If an RP has downstream receivers that want to receive register messages from a new source, the RP can either continue to receive the register messages through the DR or join the shortest path leading toward the source. By default, the RP will join the shortest path, because delivery of native multicast traffic provides the highest throughput. Upon receipt of the first packet that arrives natively through the shortest path, the RP will send a register-stop message back to the DR. When the DR receives this register-stop message, it will stop sending register messages to the RP.

If an RP has no downstream receivers that want to receive register messages from a new source, the RP will not join the shortest path. Instead, the RP will immediately send a register-stop message back to the DR. When the DR receives this register-stop message, it will stop sending register messages to the RP.

Once a routing entry is established for a source, a periodic reregistering takes place between the DR and RP. One minute before the multicast routing table state times out, the DR will send one dataless register message to the RP each second that the source is active until the DR receives a register-stop message from the RP. This action restarts the timeout time of the multicast routing table entry, typically resulting in one reregistering exchange every 2 minutes. Reregistering is necessary to maintain state, to recover from lost state, and to keep track of sources on the RP. It will take place independently of the RP joining the shortest path.

PIM Version 1 Compatibility

If an RP is running PIM Version 1, it will not understand dataless register messages. In this case, the DR will not send dataless register messages to the RP. Instead, approximately every 3 minutes after receipt of a register-stop message from the RP, the DR encapsulates the incoming data packets from the source into register messages and sends them to the RP. The DR continues to send register messages until it receives another register-stop message from the RP. The same behavior occurs if the DR is running PIM Version 1.

When a DR running PIM Version 1 encapsulates data packets into register messages for a specific (S, G) entry, the entry is process-switched, not fast-switched or hardware-switched. On platforms that support these faster paths, the PIM registering process for an RP or DR running PIM Version 1 may lead to periodic out-of-order packet delivery. For this reason, we recommend upgrading your network from PIM Version 1 to PIM Version 2.

PIM Designated Router

Routers configured for IP multicast send PIM hello messages to determine which router will be the designated router (DR) for each LAN segment (subnet). The hello messages contain the router's IP address, and the router with the highest IP address becomes the DR.

The DR sends Internet Group Management Protocol (IGMP) host query messages to all hosts on the directly connected LAN. When operating in sparse mode, the DR sends source registration messages to the rendezvous point (RP).

By default, multicast routers send PIM router query messages every 30 seconds. By enabling a router to send PIM hello messages more often, the router can discover unresponsive neighbors more quickly. As a result, the router can implement failover or recovery procedures more efficiently. It is appropriate to make this change only on redundant routers on the edge of the network.

PIM Sparse-Mode Register Messages

Dataless register messages are sent at a rate of one message per second. Continuous high rates of register messages might occur if a DR is registering bursty sources (sources with high data rates) and if the RP is not running PIM Version 2.

By default, PIM sparse-mode register messages are sent without limiting their rate. Limiting the rate of register messages will limit the load on the DR and RP, at the expense of dropping those register messages that exceed the set limit. Receivers may experience data packet loss within the first second in which packets are sent from bursty sources.

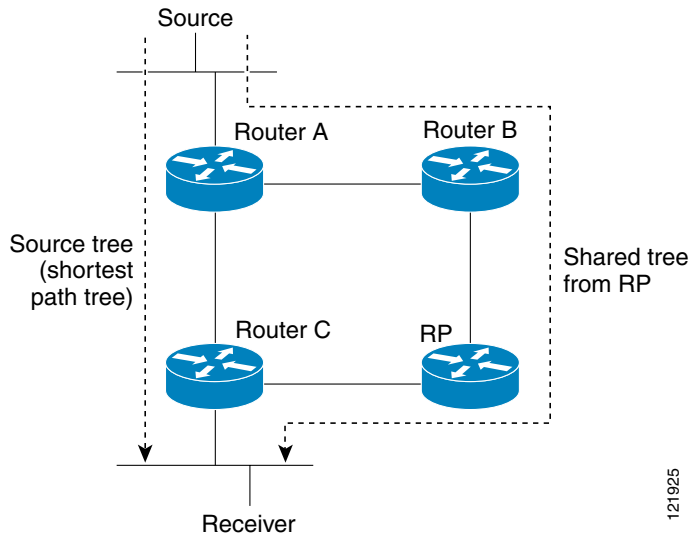
Preventing Use of Shortest-Path Tree to Reduce Memory Requirement

Understanding PIM shared tree and source tree will help you understand how preventing the use of the shortest-path tree can reduce memory requirements.

PIM Shared Tree and Source Tree (Shortest-Path Tree)

By default, members of a multicast group receive data from senders to the group across a single data distribution tree rooted at the rendezvous point (RP). This type of distribution tree is called shared tree, as shown in [Figure 1](#). Data from senders is delivered to the RP for distribution to group members joined to the shared tree.

Figure 1 Shared Tree versus Source Tree (Shortest-Path Tree)



If the data rate warrants, leaf routers on the shared tree may initiate a switch to the data distribution tree rooted at the source. This type of distribution tree is called a shortest-path tree (SPT) or source tree. By default, the software switches to a source tree upon receiving the first data packet from a source.

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

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

Join and Prune messages are sent for sources and RPs. They are sent hop-by-hop and are processed by each PIM router along the path to the source or RP. Register and register-stop messages are not sent hop-by-hop. They are sent by the designated router that is directly connected to a source and are received by the RP for the group.

Multiple sources sending to groups use the shared tree.

Benefit of Preventing or Delaying the Use of the Shortest-Path Tree

The switch from shared to source tree happens upon the arrival of the first data packet at the last hop router (Router C in [Figure 1](#)). This switch occurs because the **ip pim spt-threshold** command controls that timing, and its default setting is 0 kbps.

The shortest-path tree requires more memory than the shared tree, but reduces delay. You might want to prevent or delay its use to reduce memory requirements. Instead of allowing the leaf router to move to the shortest-path tree immediately, you can prevent use of the SPT or specify that the traffic must first reach a threshold.

You can configure when a PIM leaf router should join the shortest-path tree for a specified group. If a source sends at a rate greater than or equal to the specified *kbps* rate, the router triggers a PIM Join message toward the source to construct a source tree (shortest-path tree). If the **infinity** keyword is specified, all sources for the specified group use the shared tree, never switching to the source tree.

How to Optimize PIM Sparse Mode in a Large IP Multicast Deployment

This section contains the following procedure:

- [Optimizing PIM Sparse Mode in a Large Deployment, page 5](#) (optional)

Optimizing PIM Sparse Mode in a Large Deployment

Consider performing this task if your deployment of IP multicast is large.

Steps 3, 5, and 6 in this task are independent of each other and are therefore considered optional. Any one of these steps will help optimize PIM sparse mode. If you are going to perform Step 5 or 6, you must perform Step 4. Step 6 applies only to a designated router; changing the PIM query interval is only appropriate on redundant routers on the edge of the PIM domain.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. **ip pim register-rate-limit** *rate*
4. **ip pim spt-threshold** {*kbps* | **infinity**} [**group-list** *access-list*]
5. **interface** *type number*
6. **ip pim query-interval** *period* [msec]

DETAILED STEPS

	Command or Action	Purpose
Step 1	<p><code>enable</code></p> <p>Example: Router> enable</p>	<p>Enables privileged EXEC mode.</p> <ul style="list-style-type: none"> Enter your password if prompted.
Step 2	<p><code>configure terminal</code></p> <p>Example: Router# configure terminal</p>	<p>Enters global configuration mode.</p>
Step 3	<p><code>ip pim register-rate-limit rate</code></p> <p>Example: Router(config)# ip pim register-rate-limit 10</p>	<p>(Optional) Sets a limit on the maximum number of PIM sparse mode register messages sent per second for each (S, G) routing entry.</p> <ul style="list-style-type: none"> Use this command to limit the number of register messages that the designated router (DR) will allow for each (S, G) entry. By default, there is no maximum rate set. Configuring this command will limit the load on the DR and RP at the expense of dropping those register messages that exceed the set limit. Receivers may experience data packet loss within the first second in which register messages are sent from bursty sources.
Step 4	<p><code>ip pim spt-threshold {kbps infinity} [group-list access-list]</code></p> <p>Example: Router(config)# ip pim spt-threshold infinity group-list 5</p>	<p>(Optional) Specifies the threshold that must be reached before moving to the shortest-path tree.</p> <ul style="list-style-type: none"> The default value is 0, which causes the router to join the SPT immediately upon the first data packet it receives. Specifying the infinity keyword causes the router never to move to the shortest-path tree; it remains on the shared tree. This keyword applies to a multicast environment of “many-to-many” communication. The group list is a standard access list that controls which groups the SPT threshold applies to. If a value of 0 is specified or the group list is not used, the threshold applies to all groups. In the example, group-list 5 is already configured to permit the multicast groups 239.254.2.0 and 239.254.3.0: <pre>access-list 5 permit 239.254.2.0 0.0.0.255 access-list 5 permit 239.254.3.0 0.0.0.255</pre>

	Command or Action	Purpose
Step 5	interface <i>type number</i> Example: Router(config)# interface ethernet 0	Configures an interface. <ul style="list-style-type: none"> If you do not want to change the default values of the PIM SPT threshold or the PIM query interval, do not perform this step; you are done with this task.
Step 6	ip pim query-interval <i>period [msec]</i> Example: Router(config-if)# ip pim query-interval 1	(Optional) Configures the frequency at which multicast routers send PIM router query messages. <ul style="list-style-type: none"> Perform this step only on redundant routers on the edge of a PIM domain. The default query interval is 30 seconds. The <i>period</i> argument is in seconds unless the msec keyword is specified. Set the query interval to a smaller number of seconds for faster convergence, but keep in mind the trade-off between faster convergence and higher CPU and bandwidth usage.

Configuration Examples for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

This section provides the following configuration example:

- [Optimizing PIM Sparse Mode in a Large IP Multicast Deployment: Example, page 7](#)

Optimizing PIM Sparse Mode in a Large IP Multicast Deployment: Example

The following example shows how to:

- Set the query interval to 1 second for faster convergence.
- Configure the router to never move to the SPT but to remain on the shared tree.
- Set a limit of 10 PIM sparse mode register messages sent per second for each (S, G) routing entry.

```
interface ethernet 0
 ip pim query-interval 1
 .
 .
 !
 ip pim spt-threshold infinity
 ip pim register-rate-limit 10
 !
```

Additional References

The following sections provide references related to optimizing PIM sparse mode.

Related Documents

Related Topic	Document Title
IP multicast concepts and tasks	Cisco IOS IP Multicast Configuration Guide , Release 12.4
IP multicast commands: complete command syntax, command mode, command history, defaults, usage guidelines, and examples	Cisco IOS IP Multicast Command Reference , Release 12.4

MIBs

MIB	MIBs Link
—	To locate and download MIBs for selected platforms, Cisco IOS releases, and feature sets, use Cisco MIB Locator found at the following URL: http://www.cisco.com/go/mibs

Technical Assistance

Description	Link
Technical Assistance Center (TAC) home page, containing 30,000 pages of searchable technical content, including links to products, technologies, solutions, technical tips, and tools. Registered Cisco.com users can log in from this page to access even more content.	http://www.cisco.com/public/support/tac/home.shtml

Feature Information for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

Table 1 lists the features in this module and provides links to specific configuration information. Only features that were introduced or modified in Cisco IOS Releases 12.2(1) or later appear in the table.

Not all commands may be available in your Cisco IOS software release. For details on when support for specific commands was introduced, see the command reference documents.

Cisco IOS software images are specific to a Cisco IOS software release, a feature set, and a platform. Use Cisco Feature Navigator to find information about platform support and Cisco IOS software image support. Access Cisco Feature Navigator (<http://www.cisco.com/go/fn>). You must have an account on Cisco.com. If you do not have an account or have forgotten your username or password, click Cancel at the login dialog box and follow the instructions that appear.

Table 1 Feature Information for Optimizing PIM Sparse Mode in a Large IP Multicast Deployment

Feature Name	Releases	Feature Configuration Information
This table is intentionally left blank because no features were introduced or modified in Cisco IOS Release 12.2(1) or later. This table will be updated when feature information is added to this module.	—	—

CCVP, the Cisco logo, and Welcome to the Human Network are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networkers, Networking Academy, Network Registrar, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0711R)

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

© 2007 Cisco Systems, Inc. All rights reserved.

