Chapter 2

Introduction to PIM

The predominant multicast routing protocol in use on the Internet today is Protocol Independent Multicast, or PIM. The type of PIM used on the Internet is PIM sparse mode. PIM sparse mode is so accepted that when the simple term “PIM” is used in an Internet context, some form of sparse mode operation is assumed.

This chapter provides an overview of the features and capabilities of PIM.

This chapter discusses the following topics:

PIM Background

PIM emerged as an algorithm to overcome the limitations of dense-mode protocols such as the Distance Vector Multicast Routing Protocol (DVMRP), which was efficient for dense clusters of multicast receivers, but did not scale well for the larger, sparser, groups encountered on the Internet. The Core Based Trees (CBT) protocol was intended to support sparse mode as well, but CBT, with its all-powerful core approach, made placement of the core critical, and large conference-type applications (many-to-many) resulted in a bottlenecked core. PIM was designed to avoid the dense-mode scaling issues of DVMRP and the potential performance problems of CBT at the same time.

PIM is one of the most rapidly evolving specifications on the Internet today. Since its introduction in 1995, PIM has already seen two major revisions to its packet structure (PIM version 1 [PIMv1] and PIM version 2 [PIMv2]), two major RFCs (RFC 2362 obsoleted RFC 2117), and numerous drafts describing major components of PIM, such as many-to-many trees and source-specific multicast (SSM). Long-lasting RFCs are not a feature of PIM, and virtually all of PIM must be researched, understood, and implemented directly from Internet drafts. In fact, no current RFC describes PIMv1 at all: The drafts have all expired, and PIMv1 was never issued as an official RFC.

PIM itself is not nonstandard or unstable, however. PIM has been a promising multicast routing protocol since its inception, especially PIM sparse mode, the first real sparse-mode multicast routing protocol. Work continues on PIM in a number of areas, from bidirectional trees to network management, and the rapid pace of development makes drafts essential for PIM.

PIMv1 and PIMv2 can coexist on the same router or even on the same interface. The main difference between PIMv1 and PIMv2 is the packet format. PIMv1 messages use Internet Group Management Protocol (IGMP) packets, whereas PIMv2 has its own IP protocol number (103) and packet structure. All routers connecting to an IP subnet such as a LAN must use the same PIM version. Some PIM implementations can recognize PIMv1 packets and automatically switch the router interface to PIMv1. Because the difference between PIMv1 and PIMv2 involves the message format, but not the meaning of the message or how the router processes the PIM message, a router can easily mix PIMv1 and PIMv2 interfaces.

Basic PIM Network Components

PIM dense mode requires only a multicast source and series of multicast-enabled routers running PIM dense mode to allow receivers to obtain multicast content. Dense mode makes sure that everything gets everywhere by periodically flooding the network with multicast traffic, and relies on prune messages to make sure that subnets where all receivers are uninterested in that particular multicast group stop receiving packets.

PIM sparse mode is more complicated, and requires the establishment of special routers called *rendezv ous points (RPs)* in the network core. These routers are where upstream join messages from interested receivers meet downstream traffic from the source of the multicast group content. A network can have many RPs, but PIM sparse mode allows only one RP to be active for any multicast group.

If there is only one RP in a routing domain, the RP and adjacent links might become congested and form a single point of failure for all multicast traffic. So multiple RPs are the rule, but the issue then becomes how other multicast routers find the RP that is the source of the multicast group the receiver is trying to join. This RP-to-group mapping is controlled by a special *bootstrap router* running the PIM *bootstrap router (BSR)* mechanism. There can be more than one bootstrap router as well, also for single-point-of-failure reasons.

The bootstrap router does not have to be an RP itself, although this is a common implementation. The bootstrap router's main function is to manage the collection of RPs and allow interested receivers to find the source of their group's multicast traffic.

PIM SSM can be seen as a subset of a special case of PIM sparse mode and requires no specialized equipment other than that used for PIM sparse mode (and IGMP version 3).

PIM Modes of Operation

PIM operates in two basic modes: sparse mode and dense mode. In addition, PIM can operate in sparse-dense mode, with some multicast groups configured as dense mode (flood-and-prune, (S,G) state) and others configured as sparse mode (explicit join to rendezvous point (RP), (*, G) state).

PIM drafts also establish a mode known as PIM source-specific mode, or PIM SSM. In PIM SSM there is only one specific source for the content of a multicast group within a given domain.

PIM Dense Mode

PIM dense mode is less sophisticated than PIM sparse mode. PIM dense mode is useful for multicast LAN applications, the main environment for all dense mode protocols.

PIM dense mode implements the same flood-and-prune mechanism that DVMRP and other dense mode routing protocols employ. The main difference between DVMRP and PIM dense mode is that PIM dense mode introduces the concept of protocol independence. PIM dense mode can use the routing table populated by any underlying unicast routing protocol to perform reverse-path-forwarding (RPF) checks.

Internet Service Providers (ISPs) typically appreciate the ability to use any underlying unicast routing protocol with PIM dense mode because they need not introduce and manage a separate routing protocol just for RPF checks. Unicast routing protocols extended as multiprotocol Border Gateway Protocol (MBGP) and multitopology routing for Intermediate System-to-Intermediate System (M-ISIS) were later employed to build special tables to perform RPF checks, but PIM dense mode does not require them.

PIM dense mode can use the unicast routing table populated by Open Shortest Path First (OSPF), IS-IS, BGP, and so on, or PIM dense mode can be configured to use a special multicast RPF table populated by MBGP or M-ISIS when performing RPF checks.

PIM Sparse Mode

These are the major characteristics of PIM sparse mode:

Routers with downstream receivers join a PIM sparse-mode tree through an explicit join message.

PIM sparse-mode RPs are the routers where receivers meet sources.

Senders announce their existence to one or more RPs, and receivers query RPs to find multicast sessions.

Once receivers get content from sources through the RP, the last-hop router (the router closest to the receiver) can optionally remove the RP from the shared distribution tree (*, G) if the new source-based tree (S,G) is shorter. Receivers then get content directly from the source.

This transitional aspect of PIM sparse mode from shared to source-based tree is one of the major attractions of PIM. This feature prevents overloading the RP or surrounding core links.

There are related issues regarding source, RPs, and receivers when sparse mode multicast is used:

Sources must be able to send to all RPs.

RPs must all know each other.

Receivers must send explicit join messages to a known RP.

Receivers initially need to know only one RP (they later learn about others).

Receivers can explicitly prune themselves from a tree.

Receivers that never transition to a source-based tree are effectively running CBT.

PIM sparse mode has standard features for all of these issues.

PIM SSM

RFC 1112, the original multicast RFC, supported both many-to-many and one-to-many models. These came to be known collectively as any-source multicast (ASM) because ASM allowed one or many sources for a multicast group’s traffic. However, an ASM network must be able to determine the locations of *all* sources for a particular multicast group whenever there are interested listeners, no matter where the sources might be located in the network. In ASM, the key function of *source discovery* is a required function of the network itself.

Multicast source discovery appears to be an easy process, but in sparse mode it is not. In dense mode, it is simple enough to flood traffic to every router in the whole network so that every router knows the source address of the content for that multicast group. However, the flooding presents scalability and network resource use issues and is not a viable option in sparse mode.

PIM sparse mode (like any sparse mode protocol) achieves the required source discovery functionality without flooding at the cost of a considerable amount of complexity. The RP routers must be added and are responsible for knowing all multicast sources, and complicated shared distribution trees must be built to the RPs.

In an environment where many sources come and go, such as for a videoconferencing service, ASM makes perfect sense. However, by ignoring the many-to-many model and focusing attention on the one-to-many SSM model, several commercially promising multicast applications, such as television channel distribution over the Internet, might be brought to the Internet much more quickly and efficiently than if full ASM functionality were required of the network.

PIM SSM is simpler than PIM sparse mode because only the one-to-many model is supported. Initial commercial multicast Internet applications are likely to be available to *subscribers* (that is, receivers that issue join messages) from only a single source (a special case of SSM covers the need for a backup source). PIM SSM therefore forms a subset of PIM sparse mode. PIM SSM builds shortest-path trees (SPTs) rooted at the source immediately because in SSM, the router closest to the interested receiver host is informed of the unicast IP address of the source for the multicast traffic. That is, PIM SSM bypasses the RP connection stage through shared distribution trees, as in PIM sparse mode, and goes directly to the source-based distribution tree.

PIM SSM introduces new terms for many of the concepts in PIM sparse mode. PIM SSM can technically be used in the entire 224/4 multicast address range, although PIM SSM operation is guaranteed only in the 232/8 range (232.0.0/24 is reserved). The new SSM terms are appropriate for Internet video applications and are summarized in Table 3.

Table 3: ASM and SSM Terminology

Term	Any-Source Multicast	Source-Specific Multicast
Address identifier	G	S,G
Address designation	group	channel
Receiver operations	join, leave	subscribe, unsubscribe
Group address range	224/4 excluding 232/8	224/4 (guaranteed only for 232/8)

Although PIM SSM describes receiver operations as *subscribe* and *unsubscribe*, the same PIM sparse mode join and leave messages are used by both forms of the protocol. The terminology change distinguishes ASM from SSM even though the receiver messages are identical.

Mixing Modes

It is possible to mix PIM dense mode, PIM sparse mode, and PIM SSM on the same network, the same router, and even the same interface. This is because modes are effectively tied to multicast groups, an IP multicast group address must be unique for a particular group's traffic, and scoping limits enforce the division between potential or actual overlaps.

A multicast router employing sparse-dense mode is a good example of mixing PIM modes on the same network or router or interface. Dense modes are easy to support because of the flooding, but the scaling issues make dense modes inappropriate for Internet use beyond very restricted uses.

PIM sparse mode was capable of forming SPTs already. Changes to PIM sparse mode to support PIM SSM mainly involved defining behavior in the SSM address range, because shared tree behavior is prohibited for groups in the SSM address range.