Network Working Group
Request for Comments: 2362
Obsoletes: 2117
Category: Experimental

D. Estrin
USC
D. Farinacci
CISCO
A. Helmy
USC
D. Thaler
UMICH
S. Deering
XEROX
M. Handley
UCL
V. Jacobson
LBL
C. Liu
USC
P. Sharma
USC
L. Wei
CISCO

June 1998

Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification

Status of this Memo

This memo defines an Experimental Protocol for the Internet community. It does not specify an Internet standard of any kind. Discussion and suggestions for improvement are requested. Distribution of this memo is unlimited.

Copyright Notice

1 Introduction

This document describes a protocol for efficiently routing to multicast groups that may span wide-area (and inter-domain) internets. We refer to the approach as Protocol Independent Multicast--Sparse Mode (PIM-SM) because it is not dependent on any particular unicast routing protocol, and because it is designed to support sparse groups as defined in [1][2]. This document describes the protocol details. For the motivation behind the design and a description of the architecture, see [1][2]. Section 2 summarizes PIM-SM operation. It describes the protocol from a network perspective, in particular, how the participating routers interact to create and maintain the multicast distribution tree. Section 3 describes PIM-SM operations from the perspective of a single router implementing the protocol; this section constitutes the main body of the protocol specification. It is organized according to PIM-SM message type; for each message type we describe its contents, its generation, and its processing.

Sections 3.8 and 3.9 summarize the timers and flags referred to throughout this document. Section 4 provides packet format details.

The most significant functional changes since the January ’95 version involve the Rendezvous Point-related mechanisms, several resulting simplifications to the protocol, and removal of the PIM-DM protocol details to a separate document [3] (for clarity).

2 PIM-SM Protocol Overview

In this section we provide an overview of the architectural components of PIM-SM.

A router receives explicit Join/Prune messages from those neighboring routers that have downstream group members. The router then forwards data packets addressed to a multicast group, G, only onto those interfaces on which explicit joins have been received. Note that all routers mentioned in this document are assumed to be PIM-SM capable, unless otherwise specified.

A Designated Router (DR) sends periodic Join/Prune messages toward a group-specific Rendezvous Point (RP) for each group for which it has active members. Each router along the path toward the RP builds a wildcard (any-source) state for the group and sends Join/Prune messages on toward the RP. We use the term route entry to refer to the state maintained in a router to represent the distribution tree. A route entry may include such fields as the source address, the group address, the incoming interface from which packets are accepted, the list of outgoing interfaces to which packets are sent,

timers, flag bits, etc. The wildcard route entry's incoming interface points toward the RP; the outgoing interfaces point to the neighboring downstream routers that have sent Join/Prune messages toward the RP. This state creates a shared, RP-centered, distribution tree that reaches all group members. When a data source first sends to a group, its DR unicasts Register messages to the RP with the source's data packets encapsulated within. If the data rate is high, the RP can send source-specific Join/Prune messages back towards the source and the source's data packets will follow the resulting forwarding state and travel unencapsulated to the RP. Whether they arrive encapsulated or natively, the RP forwards the source's decapsulated data packets down the RP-centered distribution tree toward group members. If the data rate warrants it, routers with local receivers can join a source-specific, shortest path, distribution tree, and prune this source's packets off of the shared RP-centered tree. For low data rate sources, neither the RP, nor last-hop routers need join a source-specific shortest path tree and data packets can be delivered via the shared, RP-tree.

The following subsections describe SM operation in more detail, in particular, the control messages, and the actions they trigger.

2.1 Local hosts joining a group

In order to join a multicast group, G, a host conveys its membership information through the Internet Group Management Protocol (IGMP), as specified in [4][5], (see figure 1). From this point on we refer to such a host as a receiver, R, (or member) of the group G.

Note that all figures used in this section are for illustration and are not intended to be complete. For complete and detailed protocol action see Section 3.

[Figures are present only in the postscript version]
Fig. 1 Example: how a receiver joins, and sets up shared tree

When a DR (e.g., router A in figure 1) gets a membership indication from IGMP for a new group, G, the DR looks up the associated RP. The DR creates a wildcard multicast route entry for the group, referred to here as a (*,G) entry; if there is no more specific match for a particular source, the packet will be forwarded according to this entry.

The RP address is included in a special field in the route entry and is included in periodic upstream Join/Prune messages. The outgoing interface is set to that included in the IGMP membership indication for the new member. The incoming interface is set to the interface used to send unicast packets to the RP.

When there are no longer directly connected members for the group, IGMP notifies the DR. If the DR has neither local members nor downstream receivers, the (*,G) state is deleted.

2.2 Establishing the RP-rooted shared tree

Triggered by the (*,G) state, the DR creates a Join/Prune message with the RP address in its join list and the the wildcard bit (WC-bit) and RP-tree bit (RPT-bit) set to 1. The WC-bit indicates that any source may match and be forwarded according to this entry if there is no longer match; the RPT-bit indicates that this join is being sent up the shared, RP-tree. The prune list is left empty. When the RPT-bit is set to 1 it indicates that the join is associated with the shared RP-tree and therefore the Join/Prune message is propagated along the RP-tree. When the WC-bit is set to 1 it indicates that the address is an RP and the downstream receivers expect to receive packets from all sources via this (shared tree) path. The term RPT-bit is used to refer to both the RPT-bit flags associated with route entries, and the RPT-bit included in each encoded address in a Join/Prune message.

Each upstream router creates or updates its multicast route entry for (*,G) when it receives a Join/Prune with the RPT-bit and WC-bit set. The interface on which the Join/Prune message arrived is added to the list of outgoing interfaces (oifs) for (*,G). Based on this entry each upstream router between the receiver and the RP sends a Join/Prune message in which the join list includes the RP. The packet payload contains Multicast-Address=G, Join=RP,WC-bit,RPT-bit, Prune=NULL.

2.3 Hosts sending to a group

When a host starts sending multicast data packets to a group, initially its DR must deliver each packet to the RP for distribution down the RP-tree (see figure 2). The sender's DR initially encapsulates each data packet in a Register message and unicasts it to the RP for that group. The RP decapsulates each Register message and forwards the enclosed data packet natively to downstream members on the shared RP-tree.

[Figures are present only in the postscript version]
Fig. 2 Example: a host sending to a group

If the data rate of the source warrants the use of a source-specific shortest path tree (SPT), the RP may construct a new multicast route entry that is specific to the source, hereafter referred to as (S,G) state, and send periodic Join/Prune messages toward the source. Note that over time, the rules for when to switch can be modified without

global coordination. When and if the RP does switch to the SPT, the routers between the source and the RP build and maintain (S,G) state in response to these messages and send (S,G) messages upstream toward the source.

The source's DR must stop encapsulating data packets in Registers when (and so long as) it receives Register-Stop messages from the RP. The RP triggers Register-Stop messages in response to Registers, if the RP has no downstream receivers for the group (or for that particular source), or if the RP has already joined the (S,G) tree and is receiving the data packets natively. Each source's DR maintains, per (S,G), a Register-Suppression-timer. The Register-Suppression-timer is started by the Register-Stop message; upon expiration, the source's DR resumes sending data packets to the RP, encapsulated in Register messages.

2.4 Switching from shared tree (RP-tree) to shortest path tree (SP-tree)}

A router with directly-connected members first joins the shared RP-tree. The router can switch to a source's shortest path tree (SP-tree) after receiving packets from that source over the shared RP-tree. The recommended policy is to initiate the switch to the SP-tree after receiving a significant number of data packets during a specified time interval from a particular source. To realize this policy the router can monitor data packets from sources for which it has no source-specific multicast route entry and initiate such an entry when the data rate exceeds the configured threshold. As shown in figure 3, router 'A' initiates a (S,G) state.

[Figures are present only in the postscript version]
Fig. 3 Example: Switching from shared tree to shortest path tree

When a (S,G) entry is activated (and periodically so long as the state exists), a Join/Prune message is sent upstream towards the source, S, with S in the join list. The payload contains Multicast-Address=G, Join=S, Prune=NULL. When the (S,G) entry is created, the outgoing interface list is copied from (*,G), i.e., all local shared tree branches are replicated in the new shortest path tree. In this way when a data packet from S arrives and matches on this entry, all receivers will continue to receive the source's packets along this path. (In more complicated scenarios, other entries in the router have to be considered, as described in Section 3). Note that (S,G) state must be maintained in each last-hop router that is responsible for initiating and maintaining an SP-tree. Even when (*,G) and (S,G) overlap, both states are needed to trigger the source-specific Join/Prune messages. (S,G) state is kept alive by data packets arriving from that source. A timer, Entry-timer, is set for the (S,G)

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.