A Brief Overview of IP Multicast

by Joseph Porrovecchio, University of Toronto (ENFM) January 1997

How IP Multicast Works

In the IP multicast scheme, each individual multicast group -- that is, a group of hosts that have all subscribed to a particular application -- can be identified by a particular class D IP address. Each host can register itself as a member of selected multicast groups through use of the Internet Group Management Protocol (IGMP).

To understand the benefits of multicasting, consider an MPEG-based video server. Playback of an MPEG stream requires approximately 1.5 Mbps per client viewer. In an unicast environment, the video server sends 1.5 x n Mbps of traffic to the network (where n = number of client viewers). With a 10-Mbps pipe from the server, approximately 6 to 7 streams can be supported before the network runs out of bandwidth. In a multicast environment, the video server needs to send only one video stream to a multicast address; any number of clients can listen to the multicast address and receive the video stream. In this scenario, the server requires only 1.5 Mbps of bandwidth and leaves the rest of the bandwidth free for other uses.

The set of hosts listening to a particular IP multicast address is called a host group. A host group can span multiple networks. Membership in a group is dynamic - hosts may join and leave host groups. Here is a list of common multicast addresses:

Well-Known Class D Address	**Purpose**
224.0.0.1	All hosts on a subnet
224.0.0.2	All routers on a subnet
224.0.0.4	All DVMRP routers
224.0.0.5	All MOSPF routers
224.0.0.9	RIP Version 2
224.0.1.1	Network Time Protocol (NTP)
224.0.1.2	SGI Dogfight
224.0.1.7	Audio news
224.0.1.11	IETF audio
224.0.1.12	IETF video

The concept of a process joining a multicast group on a given interface on a host is fundamental to multicasting. As stated above, membership in a multicast group on a given interface is dynamic (that is, it changes over time as processes join and leave the group). Thus end users can dynamically join multicast groups based on the applications they execute.

IGMP (Internet Group Management Protocol) messages are used by multicast routers to keep track of group membership on each of the router's physically attached networks.

The following rules apply:

1. A host sends an IGMP report when the first process joins a group. If multiple processes on a given host join the same group, only one report is sent, the first time a process joins the group. This report is sent out on the same interface on which the process joined the group.

2. A host does not send a report when processes leave a group, even when the last process leaves a group. The host knows that there are no members in a given group, so when it receives the next query, it won't report the group.

3. A multicast router sends an IGMP query at regular intervals to see if any hosts still have processes belonging to any groups. The router must send one query out on each interface. The group address in the query is 0 since the router expects one response from a host for every group that contains one or more members on the host.

4. A host responds to an IGMP query by sending one IGMP report for each group that still contains at least one process.

Using these queries and reports, a multicast router keeps a table of which of its interfaces have one or more hosts in a multicast group. When the router receives a multicast datagram to forward, it forwards the datagram (using the corresponding multicast link layer address) out only the interfaces that still have hosts with processes belonging to that group.

Routers keep track of these groups dynamically and build distribution "trees" that chart paths from each sender to all receivers. When a router receives traffic for a multicast group, it refers to the specific tree that it has built for the sender. Currently, the IP standards bodies are discussing several routing protocols that can build the distribution trees for multipoint routing: Distance Vector Multicast Routing Protocol (DVMRP), Multicast Open Shortest Path First (MOSPF), and Protocol-Independent Multicast (PIM). Cisco Systems implemented PIM in the Cisco Internetwork Operating System[tm] (Cisco IOS) Release 10.2 as a pivotal routing component of its IP multicast support.

DVMRP (RFC 1075)

DVMRP uses a technique known as Reverse Path Forwarding. When a router receives a packet, it floods the packet out of all paths except the one that leads back to the packet's source. Doing so allows a datastream to reach all LANs (possibly multiple times). If a router is attached to a set of LANs that do not want to receive a particular multicast group, the router can send a "prune" message back up the distribution tree to stop subsequent packets from traveling where there are no members.

DVMRP periodically refloods in order to reach any new hosts that want to receive a particular group. There is a direct relationship between the time it takes for a new receiver to get the datastream and the frequency of flooding.

DVMRP implements its own unicast routing protocol in order to determine which interface leads back to the source of the datastream. This unicast routing protocol is very much like Routing Information Protocol (RIP) and is based purely on hop counts. As a result, the path that the multicast traffic follows may not be the same as the path that the unicast traffic follows.

DVMRP has significant scaling problems because of the necessity to flood frequently, a limitation that is exacerbated by the fact that early implementations of DVMRP did not implement pruning. As a result, DVMRP typically uses tunneling mechanisms to control flooding and, in some cases, the lack of pruning.

DVMRP has been used to build the MBONE---a multicast backbone across the public Internet---by building tunnels between DVMRP-capable machines. The MBONE is used widely in the research community to transmit the proceedings of various conferences and to permit desktop conferencing. In the near future, the MBONE will move away from DVMRP, opting to use PIM instead because of PIM's greater efficiency. Refer to the section on PIM for a discussion of the benefits of PIM over DVMRP.

PIM

Unlike DVMRP, which has inherent scaling problems, PIM offers two different types of multipoint traffic distribution patterns to address multicast routing scalability: dense mode and sparse mode.

PIM dense mode is most useful when:

- Senders and receivers are in close proximity to one another
- There are few senders and many receivers
- The volume of multicast traffic is high
- The stream of multicast traffic is constant

PIM sparse mode is most useful when:

- There are few receivers in a group
- Senders and receivers are separated by WAN links
- The type of traffic is intermittent

Dense-Mode PIM

Dense-mode PIM was designed for applications that send traffic to high concentrations of LANs. Examples of these applications include LAN TV, corporate broadcasts, and financial broadcasts. Using a technique known as reverse-path forwarding, dense-mode PIM works by flooding incoming multicast traffic out of every interface except the one through which the source can be reached. Routers that have no need for a specific data stream will reply with a "prune" message, which causes the router that performed the flooding to "prune" the relevant interface of the replying router from its flood list.

Designed for use with many receiving LANs, dense-mode PIM determines routes for multicast traffic using a two-step process. First, it floods traffic out all interfaces except for the packet's point of arrival. Second, the receiving routers respond with prune messages if they have no need for the data stream.

Sparse-Mode PIM

Sparse-mode PIM addresses the needs of environments where there might be many data streams at a given moment, but each stream goes to a relatively small number of the LANs in the internetwork. Some examples of applications suited for sparse-mode PIM include desktop conferencing and collaborative computing. For this type of traffic, a flood-and-prune approach would waste bandwidth. Instead, sparse-mode PIM works by designating a router as a rendezvous point. When a sender needs to transmit data, it first registers with the rendezvous point. Likewise, a would-be receiver must first join the group by registering with the rendezvous point.

As the data stream begins to flow from sender to rendezvous point to receiver, the routers along the way optimize the path automatically to remove any unnecessary hops. Routers continue sending to the rendezvous point, however, in case any new senders or receivers become active.

With sparse-mode PIM, routers can locate optimal paths for small, dispersed groups of receivers by naming a rendezvous point. Both sending and receiving routers register with the rendezvous point, allowing the sender to find each receiver.

As traffic begins to flow between them, routers along the way dynamically optimize the path.

Common IP Multicast / MBONE Acronyms and Terms

MBONE

multicast backbone. The multicast backbone of the Internet. MBONE is a virtual multicast network composed of multicast LANs and the point-to-point tunnels that interconnect them.

IGMP

Internet Group Management Protocol. Used by IP hosts to report their multicast group memberships to an adjacent multicast router. See also multicast router.

multicast router

Router used to send IGMP query messages on their attached local networks. Host members of a multicast group respond to a query by sending IGMP reports noting the multicast groups to which they belong. The multicast router takes responsibility for forwarding multicast datagrams from one multicast group to all other networks that have members in the group. See also IGMP.

DVMRP

Distance Vector Multicast Routing Protocol. Internetwork gateway protocol, largely based on RIP, that implements a typical dense mode IP multicast scheme. DVMRP uses IGMP to exchange routing datagrams with its neigh bors. See also IGMP.

PIM

Protocol Independent Multicast. Multicast routing architecture that allows the addition of IP multicast routing on existing IP networks. PIM is unicast routing protocol independent and can be operated in two modes: dense mode and sparse mode. See also PIM dense mode and PIM sparse mode.

PIM dense mode

One of the two PIM operational modes. PIM dense mode is data-driven and resembles typical multicast routing protocols. Packets are forwarded on all outgoing interfaces until pruning and truncation occurs. In dense m ode, receivers are densely populated, and it is assumed that the downstream networks want to receive and will probably use the datagrams that are forwarded to them. The cost of using dense mode is its default flooding behavior. Sometimes called dense mode PIM or PIM DM. Contrast with PIM sparse mode. See also PIM.

PIM sparse mode

One of the two PIM operational modes. PIM sparse mode tries to constrain data distribution so that a minimal number of routers in the network receive it. Packets are sent only if they are explicitly requested at the RP (rendezvous point). In sparse mode, receivers are widely distributed, and the assumption is that downstream networks will not necessarily use the datagrams that are sent to them. The cost of using sparse mode is its reliance on the periodic refreshing of explicit join messages and its need for RPs. Sometimes called sparse mode PIM or PIM SM. Contrast with PIM dense mode. See also PIM and rendezvous point. Router specified in PIM sparse mode implementations to track membership in multicast groups and to forward messages to known multicast group addresses. See also PIM sparse mode.

PIM NBMA

PIM Non-Broadcast Multi-Access mode allows the IP multicast forwarding engine to replicate packets for each neighbor on the NBMA network. Traditionally, cisco routers replicate multicast/broadcast packets to all "broadcast" configured neighbors. This can be inefficient when not all neighbors want packets for certain multicast groups. NBMA-mode allows one to reduce bandwidth on links leading into the NBMA network as well as CPU cycles in switches and attached neighbors.

RPF

Reverse Path Forwarding. Multicasting technique in which a multicast datagram is forwarded out of all but the receiving interface if the receiving interface is one used to forward unicast datagrams to the source of the multicast datagram.

MARS

Multicast Address Resolution Server. A mechanism for supporting IP multicast. A MARS serves a group of nodes (known as a cluster); each node in the cluster is configured with the ATM address of the MARS. The MARS supports multicast through multicast messages of overlaid point-to-multipoint connections or through multicast servers.

References

Cisco Systems IP Multicast Sites

- ftp://ftpeng.cisco.com/ipmulticast.html
- ftp://ftpeng.cisco.com/ipmulticast/