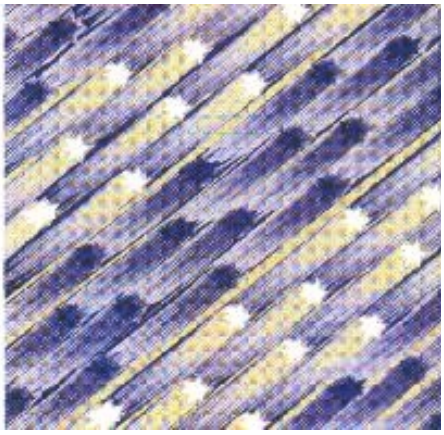

IP Multicast and IBM Networking Products Made Easy

September 9, 1999

Peter Guest

IBM Networking Hardware Division
Product Installation Support
Cary, North Carolina

Date Last Changed: September 8,1999

Abstract

This document will assist the reader to understand the concepts within IP Multicast and with implementation of multicast techniques in a network design.

Preface

The intent of this document is to help people tie together the different concepts within IP Multicast, understand the functionality of IBM Networking products in this area, and to enable you to design and implement IP Multicast within a network of IBM products. This document does not try to explain all of the concepts in detail as there are a number of good references available already. If you are new to IP Multicast, it is suggested you read through this paper once to get a basic understanding. Then after you read some additional information to better understand the technology, read this paper again to better position the concepts and understand how the IBM products use this technology.

References

IBM Redbook, TCP/IP Tutorial and Technical Overview, GG24-3376-05, (draft was October 1998), http://www.redbooks.ibm.com
The IP Multicast Initiative (IPMI), several good white papers on IP Multicast, http://www.ipmulticast.com

Introduction

From a network perspective, IP Multicast is fairly straight forward. Any device that would like to send data to an IP Multicast group sends it to the group's MAC address and IP address. The MAC address is in a reserved range and is based on the group IP address. Because this MAC address is always a destination address and never a source address, all layer two switches (that are not IP Multicast aware) will flood the data on all ports. The group IP address is in a reserved range of Class D addresses. Routers that are not running a Multicast Routing Protocol will not forward this data onto other subnets. To experiment with IP multicast within a single subnet, no special networking equipment is required.

IP Multicast is being delivered over the Internet today using the MBONE (Multicast Backbone). The first multicast session to run across the MBONE was done in 1992. Today there are hundreds of sited connected to it.

IP Multicast is often associated with video streams, but can also be used for data applications. One example is using a multicast FTP program to download new pricing information to a large number of retail stores. As a matter of fact, if you are running OSPF you are already using IP Multicast in your network for the routers to exchange topology information.

First Virtual Corporation offers IP multicastvideo applications that IBM co-markets.

Acronym Soup

There are a large number of protocols you will hear in reference to IP multicast. Most of these can be grouped into three different functions:

- How IP Hosts talk to a Multicast Router - IGMP (Internet Group Management Protocol)
- How the routers talk to one another - Multicast Routing Protocols
- How switches deal with Multicast traffic - IGMP snooping, VLANs, 802.1p (GARP/GMRP)

How IP Hosts talk to a Multicast Router

The Internet Group Management Protocol (**IGMP**) is used by hosts that want to join or leave a multicast host group. There are two types of IGMP messages; a Host Membership Query is sent by a multicast router, and a Host Membership Report is send by a host. The messages are addressed to the multicast group of interest.

How the routers talk to one another

Multicast Routing Protocols are all IETF RFCs. As with standard routing protocols, they fall into IGP (Interior Gateway Protocol) and EGP (Exterior Gateway Protocol) categories. Here is a quick summary of the different Multicast Routing Protocols:

- **IGPs** (Interior Gateway Protocols)
  - **Broadcast and Prune** implementations - Periodically broadcast multicast traffic out all ports and prune back ports where it is not needed
    - **DVMRP** (Distance Vector Multicast Routing Protocol) - A very old multicast routing protocol that is in widespread use on the Internet and well-understood. Becomes a separate routing topology process from what you are using for unicast traffic.
    - **PIM-DM** (Protocol Independent Multicast - Dense Mode) - Similar to DVMRP but uses the same topology database as for the unicast traffic (RIP, OSPF etc.)
  - **Source Tree** implementations - Through routing advertisements every router has the knowledge of the multicast groups each router is interested in receiving. They all calculate the same path for each multicast group. CPU intensive.
    - **MOSPF** (Multicast extensions to Open Shortest Path First) - By adding a group-membership-LSA (Link State Advertisement), the location of all multicast group members is pinpointed in the database. The path of the multicast datagram can then be calculated by building a shortest-path tree (SPT) rooted at the datagram's source.
- **EGPs** (Exterior Gateway Protocols)
  - **Shared Tree** implementations - Using a protocol one shared tree is built for all IP Multicast traffic. All senders and receivers attach to the shared tree. Works well for sparse group membership.
    - **PIM-SM** (Protocol Independent Multicast - Sparse Mode) - A Designated Router (DR) sends Join/Prune messages toward a group specific Rendezvous Point (RP). Multicast data is always blocked unless a downstream router explicitly asks for it.
    - **CBT** (Core-Based Trees) - Some improvements in scaleability from PIM-SM, but not enough that it is likely to be widely used.

•**BGMP** (Border Gateway Multicast Protocol) - Was initially called Grand Unified Multicast (GUM). Uses BGP as the routing protocol. Will scale much better than PIM-SM or CBT, and may be the primary protocol used across the Internet.

How switches deal with IP Multicast traffic

IP Multicast traffic is flooded on all ports of a layer two switch, causing all stations to see this traffic. Devices that are not listening to any multicast addresses have this traffic filtered by the network interface hardware, but it still wastes bandwidth. There are three approaches that are used to try to eliminate this problem:

- **IGMP Snooping** - snoop all multicast traffic for IGMP queries and reports, as well as multicast routing protocols, to determine which devices are interested in what multicast traffic. Some implementations use this information to create a VLAN for each group.
- **VLANs** - VLAN implementations that do not use IGMP snooping are generally proprietary solutions consisting of port or MAC address based VLANs associated with each group.
- **GMRP** (GARP Multicast Registration Protocol) - A required component of the IEEE 802.1p standard that uses Generic Attribute Registration Protocol (GARP) to register and propagate multicast membership information throughout the switching domain. Obviously the switches and the network adapter device driver would both need to support 802.p. Works fine with non-802.1p switches in the network, but multicast traffic is flooded on all ports of these switches.

Other IP Multicast Concepts

Reliable Multicast

Due to the one-to-many nature of Multicast traffic, having all of the receivers acknowledge the receipt of the data is not realistic. Acknowledgments are not required for voice or video traffic as late delivery is of no use. For file transfers, however, a reliable transport is necessary. In current implementations, any acknowledgment of data is done in a proprietary nature.

There is an Internet draft (submitted by Cisco) being well refined for Pragmatic General Multicast (**PGM**), which is a reliable transport protocol. It is sometimes referred to as Pretty Good Multicast. It is one of several approaches of reliable multicast, with its central design goal as simplicity of operation. It is targeted to applications in which members may join and leave at any time, as opposed to applications that depend upon acknowledged delivery to a known group of recipients. Being a transport protocol, IGMP and an IP Multicast routing protocol would still be used.

RMTP (Reliable Multicast Transport Protocol) is another Internet draft being worked on. Once there is a Reliable Multicast Protocol standard, IBM will support it in our products.

Multicast VLANs

There are a variety of approaches to limit Multicast traffic to only those devices interested in this traffic. A standards based approach is to use 802.1p/Q VLANs. Several IBM products that are announced over the next year will support this standard, however for this to be of use the network adapter device drivers need to be able to support GARP and have it turned on. There are other vendor specific approaches that are currently available. Most of them add devices to these

VLANs based on either port or MAC addresses configured on the switch, or they use IGMP snooping.
The IBM 8371 Layer 3 Ethernet switch (which will GA 1Q/99) creates IP Multicast VLANs by IGMP snooping. Regardless of what subnet the devices are on, they are all put into the same IP Multicast VLAN, which reduces the workload on the routers and improves performance.
There is an even better story if there is an ATM backbone. The multicast VLANs on the 8371 Ethernet switches, as well as the Multicast VLANs on the 8270 Token Ring switches with MSS Clients, are all joined to IP Multicast ELANs in the MSS Server. This can span across all of the subnets in the network.

MARS (Multicast Address Resolution Server)
MARS is an RFC that enables IP Multicast to operate over a Classical IP (RFC 1577) Logical IP Subnet. A primary requirement for this was because there are no broadcast capabilities in Classical IP. In it's current implementation MARS works within a single IP subnet. They are looking at extending this standard for multiple subnets.

MPOA (MultiProtocol Over ATM)
MPOA was defined by the ATM Forum to extend LANE to a Layer 3 cut-through environment. In its current implementation, IP Multicast traffic is simply sent to the BUS. The ATM Forum is watching the work going on for MARS to see how they can use it for Multicast traffic.

IBM Networking Products that support IP Multicast

Products Based on IBM's Multiprotocol Services Software (Common Code)

The **IBM 2210, 2212, and 2216 routers as well as the MSS** (Multiprotocol Switched Services) Server share common Multicast Routing software. The current release (CC 3) supports DVMRP and MOSPF. In CC 5 (1Q/99) they will support PIM-DM, and in CC 6 (3Q/99) PIM-SM and some IP Multicast SNMP mibs.

The MSS Server lags the other common code products by a few months. It also currently supports DVMRP and MOSPF. In MSS Release 3 it will also support auto-creation of IP Multicast VLANs. This allows edge devices that support IBM's IP Multicast VLANs (Such as the 8231 Ethernet switch and the MSS Client for the 8270 Token Ring switch) to share multicast traffic without going through a router, regardless of what IP subnet they are connected to.

IBM 8274 RouteSwitch

The 8274 supports Multicast VLANs in release 3.2. These are created manually by port or MAC address. Any device (regardless of the port it is connected to) can send traffic to a given multicast group and it will only be sent within that Multicast VLAN. ???????TR SUPPORT??????????

In Release 4, which will be GA in 1H/99, the 8274 will be able to use IGMP snooping to create Multicast VLANs.

From a routing perspective, the IBM 8274 also supports DVMRP.

IBM 8275--416 Ethernet Layer 2 Switch

The 8275-416 will be GA in 1Q/99. It will support IEEE 802.1p, which includes support for GARP/GMRP.

IBM 8371 Ethernet Layer 3 Switch

The 8371 will be GA in 1Q/99. It will support Multicast VLANs. These are created through the use of IGMP snooping. When being used in an ATM environment, these Multicast VLANs join an ATM Multicast ELAN and are not required to go through any routers across the ATM backbone.

In Release 2 of this product (2Q/99) the 8371 will support IEEE 802.1p, which includes support for GARP/GMRP. It will also be able to do DVMRP and MOSPF routing in hardware.

IBM 8270 Token Ring Switch with MSS Client

In 1Q/99 the MSS Client will have the capability to do IGMP snooping between domains on the Token Ring switch. The Multicast traffic will then be connected to the appropriate Multicast VLAN (ELAN) in the MSS Server. Because ELANs need to be either Token Ring or Ethernet you would have a Token Ring Multicast ELAN and an Ethernet Multicast ELAN in a mixed environment. The MSS Server could then route the IP Multicast traffic between these two ELANs.

Designing Multicast Networks with IBM Networking Products

From a Layer 2 Perspective

Virtually any Campus network can accommodate a little IP Multicast traffic. The amount of bandwidth used for most existing Multicast applications is not tremendously great, and end stations that are not looking for multicast traffic simply ignore this traffic as they would any unicast traffic (because it is not a broadcast). However, as IP Multicast applications become more widely deployed, and the type of content consumes more bandwidth, some thought must be given as to how to design your network to best accommodate this traffic.

If most of the end station network adapter device drivers support GARP/GMRP and can be enabled to use it, then switches that will support this standard should be deployed. If it is uncertain that this will be the case, at least the collapse point switches should support IGMP snooping.

If the network will have an ATM backbone, IBM's Multicast VLANs (ELANs) is an excellent way to isolate this traffic from the backbone and router hops.

If the network has Token Ring, it should be noted that there is no equivalent to the Ethernet group addresses within Token Ring. What is used for IP Multicast traffic over Token Ring is one functional address. IBM routers do not currently use this functional address, they use the defacto standard of putting IP Multicast traffic onto a Token Ring as a broadcast. At the time of writing, the Token Ring functional address would be recognized by the MSS Client function to allow this traffic to enter the IP Multicast VLANs, however our routing code does not currently recognize this functional address so this traffic would not be routed. If applications begin to use this standard, IBM will enhance the router functionality to support it.

From a Layer 3 Perspective

As with standard IP routing, most private networks will use an Interior Gateway Protocol for routing their IP Multicast traffic. We earlier mentioned 3 different IGP Multicast routing protocols; DVMRP, MOSPF, and PIM-DM. Both DVMRP and PIM-DM are broadcast and prune architectures. MOSPF uses Link State Advertisements so all the routers are aware of the IP Multicast groups. Therefore if you have a need to conserve your bandwidth (such as on a WAN) and you are already using OSPF, MOSPF is a much better solution.

There are a few reasons not to use MOSPF. The most likely case would be interoperating with routers that do not support MOSPF. Cisco routers currently support PIM-DM, and most routers support DVMRP. We can run DVMRP and MOSPF in the same router and bring the DVMRP traffic into our MOSPF network (such as from the Internet). Our implementation of MOSPF does not support being a wild card receiver, therefore we can not send Multicast data from our MOSPF network into a DVMRP network.

One advantage to DVMRP is the fact that it runs its own routing topology database. Therefore you could separate your traffic onto different links by weighting them differently for your multicast traffic versus your unicast traffic. The advantage of PIM-DM (which we will support in 1Q/99) is that it uses the same routing table (RIP or OSPF) so it is more straight forward.
As for Exterior Gateway Protocols, our common code products will support PIM-SM in 3Q/99. There does not seem to be enough benefits to CBT for it to be widely used. It is still uncertain if BGMP will be scaleable enough to be used for the Internet backbone. Once this is determined, IBM will consider providing additional EGP Multicast protocols.

Configuring IBM products for Multicast

Products Based on IBM's Multiprotocol Services Software (Common Code)

The IBM **2210**, **2212** and **2216 Routers** all have a consistent method of configuring IP Multicast from either the command line or the configuration tool. On the IBM **8210 MSS Server** the same command line configuration can be used, but the Configuration Tool does not currently support configuring DVMRP.

consider Monitoring and PD...
DVMRP, OSPF and MFC ELS messages....

IBM 8274 RouteSwitch

IBM IP Multicast VLANs

In 1Q/99 the IBM **8210 MSS Server** (version 3), **8371 Ethernet switch** and **8270 Token Ring switch with MSS Client** will all support IP Multicast VLANs. At that time a Solutions whitepaper will be produced that outlines design considerations and detailed configuration steps.

Conclusion

IBM Networking products currently have strong support for IP Multicast. There are a number of customers that are currently using this capability. IBM will continue to improve our solutions to handle the increase in IP Multicast traffic that is occurring. IP Multicast is one of the many functions that should be considered when doing any significant redesign of a network.

Standards

IETF RFC 1112	IGMP (Interior Gateway Management Protocol)
IETF RFC	IGMP version 2
IETF RFC 1075	DVMRP
IETF RFC	DVMRP version 3
IETF RFC 1469	IP Multicast over Token Ring
IETF RFC 1577	Classical IP over ATM
	MARS
	MPOA
IEEE 802.1p	Supplement to 802.1D bridge standard for time critical and multicast apps (GARP and Priority labeling)
IEEE 802.1Q	VLAN standard - Rev 1 is explicit 2 byte label. The standard will also support implicit VLANs by filtering.
IEEE 802.3ac	VLAN Tagging of Ethernet frames to accommodate 802.1Q
IEEE 802.3z	Gigabit Ethernet standard

IETF RFC Drafts .... PGM, RMTP ....