

Multicast & LANs

Gorry Fairhurst, University of Aberdeen.

Abstract

The UK academic network offers an IP Multicast service, virtually all Ethernet adaptors already support IP multicast, and commercial multicast applications are emerging. So what are the issues in using multicast in a large Ethernet network? This presentation gives a guided tour of IP multicasting an Ethernet LAN. It will tell you how IGMP works, what it is used for, and how it is evolving,. It also explains some options for significantly improving multicast support in switched LANs, and identifies ten thorny issues that are best understood by those deploying and using IP multicast in their LANs.

1 Introduction

Multicasting is “the networking technique of delivering the same packet simultaneously to a group of clients using a single local transmit operation”. Many computers now ship with multicast support as standard, and many routers support multicast forwarding or routing. Although multicast delivery is not yet common in service networks, and today represents only a small proportion of total network traffic, this situation is poised to change.

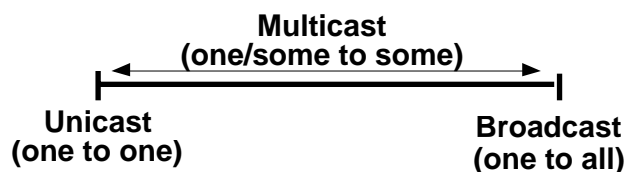


Figure 1 Multicast – a service between unicast and broadcast

Multicast occupies the middle ground between the one-to-all delivery provided by broadcasting (IP packets delivered to all end hosts within the same IP network) and the one-to-one delivery employed in unicast transfers. Multicast allows simultaneous delivery to a set of client computers (known as “end hosts”), but in contrast to broadcast, multicast is selective. The number of senders (servers) sending data to a multicast group is usually determined by the type of application being used (e.g. multimedia, file transfer, database access). In many current cases, there is only one sender to each group. The network uses multicast routing protocols to ensure copies of the packets associated with a group reach all LANs where there are end hosts wishing to receive them.

The key features of multicast are:

- (i) End hosts sending packets to a multicast group are not necessarily aware whether there are any end hosts receiving them. This means the “cost” of sending a packet is the same for one receiver, as for millions. If the sender needs to know which clients are receiving the information, they must use a higher layer protocol.

- (ii) The network undertakes to send packets to all end hosts who need them, and tries to filter packets which end hosts do not need. This reduces the “cost” of processing packets by an end host that are not required. It may also reduce network load and provides automatic forwarding of multicast packets between end hosts in different IP networks.
- (iii) A scheme must be used to assign a multicast group address to each multicast flow, and to inform senders and receivers of the set of multicast group addresses in use.

Multicast delivery is useful if a set of end hosts require a common set of data at the same time, or when the end hosts are able to store (cache) common data until they need it. Multicast transmission may provide significant bandwidth savings (up to 1/N of the bandwidth for N separate unicast clients). Most importantly, it can also eliminate server bottlenecks resulting from an increasing number of clients. Multicast is an optional service in IPv4 networks and a standard service in IPv6 networks.

2 IP Multicast

The format of an IPv4 multicast packet (RFC 1112) is identical to that of an IPv4 unicast packet and is distinguished only by a special class of destination address (class D, i.e., addresses in the range from 224.0.0.0 to 239.255.255.255). The address space allocated for multicast addresses has the high order four bits set to 1110, the remaining 28 bits are used to identify a specific IP multicast. The current assignments are shown below:

224.0.0.0 - 224.0.0.255	(224.0.0/24)	Local Network Control
224.0.1.0 - 224.0.1.255	(224.0.1/24)	Internetwork Control
224.0.2.0 - 224.0.255.0		AD-HOC Block
224.1.0.0 - 224.1.255.255	(224.1/16)	ST Multicast Groups
224.2.0.0 - 224.2.255.255	(224.2/16)	SDP/SAP
224.252.0.0 - 224.255.255.255		DIS Transient
225.0.0.0 - 231.255.255.255		RESERVED
232.0.0.0 - 232.255.255.255	(232/8)	Source Specific Multicast
233.0.0.0 - 233.255.255.255	(233/8)	GLOP Block
234.0.0.0 - 238.255.255.255		RESERVED
239.0.0.0 - 239.255.255.255	(239/8)	Administratively Scoped

Table 1: Allocation of blocks of multicast addresses for types of multicast service

The lowest set of addresses (224.0.0.0-224.0.0.255) are reserved for well-known group addresses that should never be forwarded outside the LAN by routers. Specifically, the group address 224.0.0.1 is assigned to the permanent group of all IP multicast enabled devices in a LAN (hosts/computers and gateways/routers within the IP subnet) and the address 224.0.0.2 to the group of all IP routers in a LAN.

224.0.0.1	All multicast systems on this LAN
224.0.0.2	All Routers on this LAN
224.0.0.4	DVMRP
224.0.0.5	MOSPF
224.0.0.6	MOSPF
224.0.0.9	RIP2 Routers
224.0.0.13	PIM Routers
224.0.0.22	IGMPv3 Membership Reports for this LAN

Table 2: Assignment of multicast addresses within the local network control block

routers connected to a LAN (there may be more than one) that one (or more) end hosts wish to receive packets with a specified group address. The first version of IGMP, IGMPv1 [RFC1112] was little used. Most current end hosts use IGMPv2 [RFC2236]. IGMPv3 is about to be standardised and some vendor implementations are available. A MIB has also been defined [RFC2571].

Initially, when a end host application is launched, it contacts the operating system with a request to receive a specific group (or groups). This is translated into an *IGMP Group Membership Report* message that is sent using the multicast address of the group it wishes to join. The report, often colloquially called a “Join” or “Register message”, is actually sent several times, to increase the chance of reception by multicast routers. A router receiving this message translates the *Report* into a router table. This is used to determine which multicast packets should be forwarded to which LAN interface. It may also trigger a multicast routing protocol to ask upstream routers to send a “router join” message requesting forwarding of this group by the upstream network.



Figure 3: Receivers indicate which groups they wish to receive using IGMP, routers then identify the senders to this group and use a routing protocol to ask for the multicast packets.

To validate the set of currently active groups, one multicast router per LAN periodically (e.g. every 125 seconds for IGMPv2) sends a multicast packet containing an *IGMP Group Membership Query* to all end hosts (i.e. using the IPv4 address 224.0.0.1, MAC address 0x01:00:5E:00:00:01). The router (or layer 3 switch) that sends the *Query* is known as the *Querier*. There is only one *Querier* active within a LAN at any one time. In IGMPv2 (and IGMPv3), the elected *Querier* is the one the highest IP source address. (This doesn't have to be the default/designated multicast router which actually forwards packets to and from the upstream network.) The *Query* message is received by all systems in the LAN that have multicast enabled. In end hosts that are not configured to use IP multicast, this is filtered by the Ethernet NIC, or operating system driver.

Each end host sets a random timer for each group it wishes to receive. When the timer expires, it responds with a message containing an *IGMP Group Membership Report* for the group. Since the purpose of the *Report* is to tell the LAN multicast router(s) to forward the specified group, there is no need for the router(s) to receive more than one *Report* per group address. End hosts therefore suppress their own *Report*, if they receive the same *Report* sent by another end host before the timer expires.

When the *Querier* sends several successive *Queries* (the *group membership interval*) and the multicast routers receive no *Report* for a specific group, the routers stop forwarding multicast packets with this group address. Transmission resumes when an end host next issues a group

membership *Report* indicating it now wishes to receive a suppressed group. This is the basic protocol used in IGMPv1.

The drawback of IGMPv1 is that it may take a considerable time for a multicast router to discover there are no longer any end hosts interested in a group that it is forwarding. While it is discovering this, unwanted packets belonging to the group are sent by the upstream network and forwarded on to the LAN. A refinement in IGMPv2 (RFC2236), allows end hosts to explicitly *Leave* a group they previously joined. To do this, a *Leave* message indicating the group address which the end host wishes to leave is sent to address 224.0.0.2.

Reception of a *Leave* message is however not sufficient to determine that there are no longer any end hosts interested in a group. The IGMPv2 *Querier* therefore responds by sending a *Query* to request responses only from end hosts interested in the group for which the *Leave* was received. This allows multicast routers to more quickly stop transmission of multicast packets to groups for which there are no longer any active receivers

The latest version of the protocol, IGMPv3, allows an end host to report an interest in receiving not only a specific group address, but also from a specific set of IP source addresses (or all except a specific set of IP source addresses). This allows finer control over the multicast packets sent to a LAN. This may conserve bandwidth, preventing overloading of LAN switches and end hosts, especially when a client switches from receiving one multicast group to another. Such changes are common in some scenarios (e.g. viewer channel-hopping between multicast TV channels).

Like IGMPv2, end hosts that receive an IGMPv3 *Query* start a random timer. The maximum value is specified as a parameter in the *Query*, allowing the LAN manager to select an appropriate scaling factor based on the anticipated size of the group. One significant change in IGMPv3, is that end hosts always send *Reports* following a *Query* to indicate the groups they wish to receive (i.e., there is no suppression). This may increase the volume of IGMP messages within the LAN, but allows routers (and LAN switches) to provide “explicit tracking” of which end hosts request each group. Explicit tracking allows routers to discover immediately when the last group member leaves, and suspend forwarding of the group.

IGMPv2 is the protocol used by multicast-enabled end hosts in the current Internet. It is widely expected that networks will soon start to transition to IGMPv3. The next generation Internet built using IPv6 does not use IGMP, but instead uses a similar protocol called Multicast Listener Discovery, MLD, which forms part of ICMPv6. This was developed by the same IETF working group that developed IGMP, it is no surprise that MLD closely resembles IGMPv2. A protocol called MLDv2 also mirrors the functions of IGMPv3.

5. Multicast in Switched Ethernet LANs

Many simple Ethernet switches handle multicast frames in the same way as broadcast frames. When a multicast frame reaches a layer 2 switch, the switch forwards it to all active interfaces, effectively flooding the LAN. This ensures all clients receive the multicast information, but has the drawback that every LAN segment carries all multicast traffic, even when end hosts do not require it. This eliminates most of the advantages of switching for

multicast traffic. Without appropriate controls, multicast traffic can easily overload LAN segments with unnecessary traffic (especially 10 Mbps interfaces!)

A multicast-enabled switch attempts to control this flooding. It may selectively forward multicast frames only to LAN segments where end hosts wish to receive a multicast group address. This may be achieved using a set of filters at the output ports of the switch. For historical reasons, the filter table is sometimes known as a Content Addressable Memory, CAM. This filtering by switches resembles the processing used to implement Virtual LANs (VLANs), and may be performed using the same control processor.

A layer 2 switch forwards frames based on their MAC destination address, not their IP address. Therefore a filter table entry for 224.1.2.3 would also forward 239.129.2.3, due to the multicast address overlap (both correspond to the MAC destination address of 01:00:5E:01:02:03). In contrast, a layer 3 switch, uses the IP group address to control multicast forwarding, eliminates this overlap traffic.

While manual configuration of the filter table may suffice for simple applications such as multicast file transfer, or multicast distribution for network news/web cache clients, identifying multicast addresses in use and maintaining manual tables is very error-prone and would be excessively time consuming.

There is therefore a need for an automatic procedure to introduce the correct multicast entries in the filter table. How does the processor figure out which packets to send to which of its interface ports? – There are essentially six approaches to maintaining the filter table:

- (i) **Flood multicast (like broadcast), and manually configure filter tables.** The network manager determines which Ethernet segments should receive which multicast frames (or packets), in the same manner that VLAN membership is configured. For example, she/he may prevent multicast traffic reaching certain parts of a LAN. Most multicast applications select their group addresses dynamically as the application executes, making this an unsuitable approach.
- (ii) **Implement IGMP Snooping.** A control processor inside a switch can "Snoop" multicast frames passing through the switch and observe the IGMP messages sent by end hosts. These messages contain the layer 3 group addresses requested by end hosts. Switches that are able to monitor/emulate IGMP messages may use this information to configure the switch filters dynamically. Extracting and processing IGMPv2 messages is non-trivial, since IGMP messages have the same multicast address as data packets belonging to the group. Hardware support is desirable (usually in the form of an ASIC) to allow the switch to identify the IGMP messages and only forward these to the control processor. It is simple to add filter entries when *Reports* are received, and to remove them when there are no longer any group members. To allow a switch to remove old entries when an end host leaves, requires more processing. Some processors also originate IGMP *Query* messages to verify that all end hosts connected via a port have actually left.

- (iii) **Implement IGMP Proxy.** Since the control processor in a switch would normally receive all IGMP packets, some manufacturers have gone one stage further, and implement the IGMP protocol within the switch. Such a switch can then behave as if it were a multicast router, and execute an IGMP *Querier* for each interface port. This allows it to collect group membership information, which can be directly loaded into the switch multicast filter table. The processor must also use the group membership information to respond to *Queries* received from upstream routers acting as *Queriers*. In this way, the switch acts as a proxy. An IGMPv3 proxy may also implement “proxy reporting”, where the proxy only reports which sources and groups need to be forwarded, but does not explicitly report all the end hosts that requested each group. This can significantly reduce the number of IGMP *Reports* passed to the LAN routers. A proxy acts as host for IGMP *Queries* received from an a port which is downstream from a router and must be the *Querier* for down-stream LAN ports with end hosts connected.
- (iv) **Arrange for routers to configure switch filter tables.** Each IP multicast router receives copies of all IGMP *Reports* and must track the multicast membership of each LAN. Using this information and noting the MAC source address used in each *Report*, a router can learn the groups requested by each end host and the end host’s MAC address. A protocol may then down-load an entry to the multicast address filter table in all switches along the path to the specified end host MAC address. Messages may also be sent to delete the entry when an end host leaves the group. Most CISCO switches support the proprietary CISCO Group Management Protocol (CGMP) that provides this function.
- (v) **Arrange for end hosts to configure switch filter tables.** Another approach is for end hosts to signal the set of MAC group addresses that they wish to receive in a special control message sent to their local Ethernet switch. The control processor in the switch may then use this information to make / delete entries in their filter tables. Such a control protocol has been defined by the IEEE using 802.1p frames and is known as Generic Multicast Registration Protocol, GMRP. This appears to not be widely used.
- (vi) **Add multicast routing support to switches.** There are some who argue that (ii)-(v) violate the layering of protocols, and that Ethernet Switches, operating at layer 2 of the OSI model, have no need (or right!) to meddle with network layer multicast, which is a layer 3 function. Such people suggest Ethernet switches should be fast, cheap and simple, and advocate the increased use of multicast routing within LANs, rather than using layer 2 switches.

The above techniques all assist in controlling the flooding of multicast traffic in a switched LAN, and are now available from equipment vendors. The implementations of the techniques vary in their ability to remove previously requested groups, and their scalability with increasing multicast traffic. Ten thorny issues have emerged that should be understood by those deploying and using IP multicast in their LAN:

- (i) An increased level of multicast traffic may impact performance. Some implementations are very processor intensive. Also be aware that even simple schemes can lead to bizarre outcomes following a hardware or software fault!
- (ii) Ethernet switches may carry non-IP multicast traffic, flooding of frames sent to these MAC multicast groups will not be controlled by IGMP-based switches.
- (iii) Packets with certain multicast addresses **MUST** be forwarded to all switch interface ports, others may wish to be blocked. A minimum solution is to forward all frames with a destination address in the range 01:00:5E:00:00:01 to 01:00:5E:00:00:FF. Unfortunately, not all switches do this right, and some adjust the filter table based on IGMP messages concerning this group, preventing forwarding of these packets to other end hosts.
- (iv) It is also important to realise that some systems send IGMP messages with the IP Router Alert option. This uses an optional header after the normal 20B IP header. Snooping and proxy switches must correctly interpret this, although some switches have been reported to not support this correctly.
- (v) Multicast Router Discovery (MRD). It is important that a switch recognises which switch ports are connected to routers – these ports **MUST** receive all multicast packets.
- (vi) Switches must allow end hosts on the LAN to send multicast packets. In IP multicast, such sources do not need to be members of the group to which they send. These packets must always be forwarded towards the router(s). Some multicast switches also flood these packets to all ports on the LAN until an end host joins the group causing a filter table entry to be created.
- (vii) How robust is the solution? – There are potentially complex interactions (e.g. with spanning tree, after switch restart, and when end hosts receiving multicast move between switch ports). In general improper handling will cause additional flooding and possibly a short interruption to the multicast service.
- (viii) An IGMP proxy may modify the source MAC address of an IGMP *Report* sent by an end host. This should not impact the IGMP *Querier* or multicast routers. It may appear as an anomaly to a network analyser. It can impact other protocols, such as CGMP, that use the MAC source address.
- (ix) Future proofing? Watch out for IGMPv3, this complicates things by changing the way IGMP works. The Ethernet addresses for IPv6 multicast also differ from those used for IPv4. The way in which MLD and MLDv2 use ICMPv6 significantly complicates the design of IPv6 multicast snooping.
- (x) How do you know the multicast service is working? – Operational staff must be able to see the contents of the switch filter tables. To debug fully, they probably also need network information about multicast routing.

6. Conclusions

Computer operating systems and Ethernet adaptors are usually already multicast-enabled. Multicast support is available in many standard routers (and certainly all high-end routers) and could easily be enabled in service networks, if there was a desire to supply the service. There are however important issues for operators of large switched LANs. Although most Ethernet LANs will readily support multicast, there is a need for operators to understand the way in which the multicast frames are handled, the role of the Internet Group Management Protocol (IGMP), and the way in which this protocol is evolving.

As the layer of multicast traffic increases, it becomes increasingly important to control the parts of a LAN to which IP multicast packets are delivered and to develop appropriate procedures to verify correct operation of the multicast service. Many Ethernet switches provide some of the needed support by interacting with layer 3 IGMP messages. These switches allow a LAN to automatically adapt forwarding of frames, constraining multicast packets to only those parts of the LAN where they are needed. To achieve this, a variety of techniques are used – each with strengths and weaknesses. Operators are therefore encouraged to find out what features are actually available in equipment, to understand their limitations, and to deploy!

7. Some References

- K. Clark & K. Hamilton 'CISCO LAN Switching' ISBN 1-57870-094-9, 1999
 - B. Wilson, 'Developing IP Multicast Networks', CISCO Press, ISBN 1-57870-077-9, 2000.
 - IPMI, 'The IP Multicast Initiative', <http://www.ipmulticast.com>
 - S. Deering, 'Host Extensions for IP Multicasting', RFC 1112, October 1989
 - W. Fenner, 'Internet Group Management Protocol, Version 2', RFC2236, November 1997
 - M. Crawford, 'Transmission of IPv6 Packets over Ethernet Networks', RFC2464, 1998.
 - C. Partridge, & A. Jackson, 'IPv6 Router Alert Option', RFC 2711, 1999
 - S. Deering, et al 'Multicast Listener Discovery (MLD) for IPv6', RFC 2710, 1999.
 - K. McCloghrie, et al, 'Internet Group Management Protocol MIB', RFC2933, 2000.
 - B. Haberman & R. Worzella, 'IP Version 6 MIB for the MLD Protocol', RFC 3019, 2001
 - S. Biswas, B. Haberman & B. Cain, 'IGMP Multicast Router Discovery', draft-ietf-idmr-igmp-mr-disc-XX.txt, WORK IN PROGRESS (IETF MAGMA WG), 2002.
 - R. Vida, et al, 'Multicast Listener Discovery Version 2 for IPv6' <draft-vida-ml-d-v2-xx.txt> WORK IN PROGRESS (IETF IPv6 WG), 2002.
 - B. Cain, S. E. Deering, and A. Thyagarajan, 'Internet Group Management Protocol, Version 3', , <draft-ietf-idmr-igmp-v3-xx.txt> WORK IN PROGRESS (IETF MAGMA WG), 2002.
- Multicast address assignments, <http://www.iana.org/assignments/multicast-addresses>

8. Acknowledgements

The Electronics Research Group (<http://www.erg.abdn.ac.uk>) in the Department of Engineering at the University of Aberdeen thank the European Commission for funding of the GEOCAST Project (<http://www.geocastsatellite.com>) under the IST Programme.