

The Basics of Multicasting for the Professional AV Market

White Paper

Preface

The growing demand of the Internet of Things (IoT) has resulted in IP networks becoming a fundamental part of enterprise business, entertainment, and home industries. As the AV industry moves towards IP-based systems, the typical matrix switcher is now being replaced with Ethernet switches for the flexibility of higher resolutions and increased stream bandwidth. A typical Audio/Video over IP (AVoIP) environment consists of an Ethernet switch with a number of encoders (transmitters) and decoders (receivers). Many AVoIP vendors provide solutions which can encode/decode 4K and higher video resolution streams today. With the growing demand of 4K and potentially 8K resolution, the need for 10G and above Ethernet switching infrastructure to provide the proper transport of these high-quality multicast streams is imperative. The NETGEAR M4300 and M4500 switches provide up to 100G switching in a preconfigured, works-out-of-the-box experience for true multicast AV with NETGEAR IGMP Plus™ network configuration¹.

Ethernet Video Streaming Background

There are 3 different types of data traffic that traverse through an Ethernet switch: unicast, broadcast, and multicast. Unicast is a one-to-one type of data transmission which means one sender will only send to one subscriber. Broadcast is a one-to-all type of data transmission which means one sender will send to all users on the local area network (LAN) - even if they do not want to see the video. Lastly, multicast is a one-to-many type of data traffic.

Multicast is one of the many popular methods of sending video to a group interested in receiving the video. The easiest way to stream multicast traffic to all the receivers is by flooding the multicast stream to the entire LAN. This simple method of multicast streaming causes bandwidth delays and unwanted data to other hosts.

A better alternative is to take advantage of Layer 2 Internet Group Management Protocol (IGMP). IGMP is a widely used multicast protocol that is supported at the network layer of the OSI stack². IGMP is used by hosts and their adjacent routers to allow hosts to inform the router of the desire to receive, continue receiving, or stop receiving a multicast stream.

The last alternative is Layer 3 Protocol Independent Multicast (PIM). It is normally designed for a more complex network where the senders and receivers are located on different subnets of a network. PIM multicast routing is generally more complicated to setup than IGMP, but allows for greater control over routing between subnets.

¹ NETGEAR IGMP Plus is available on all M4500s and on M4300s starting with version 12.0.8.8

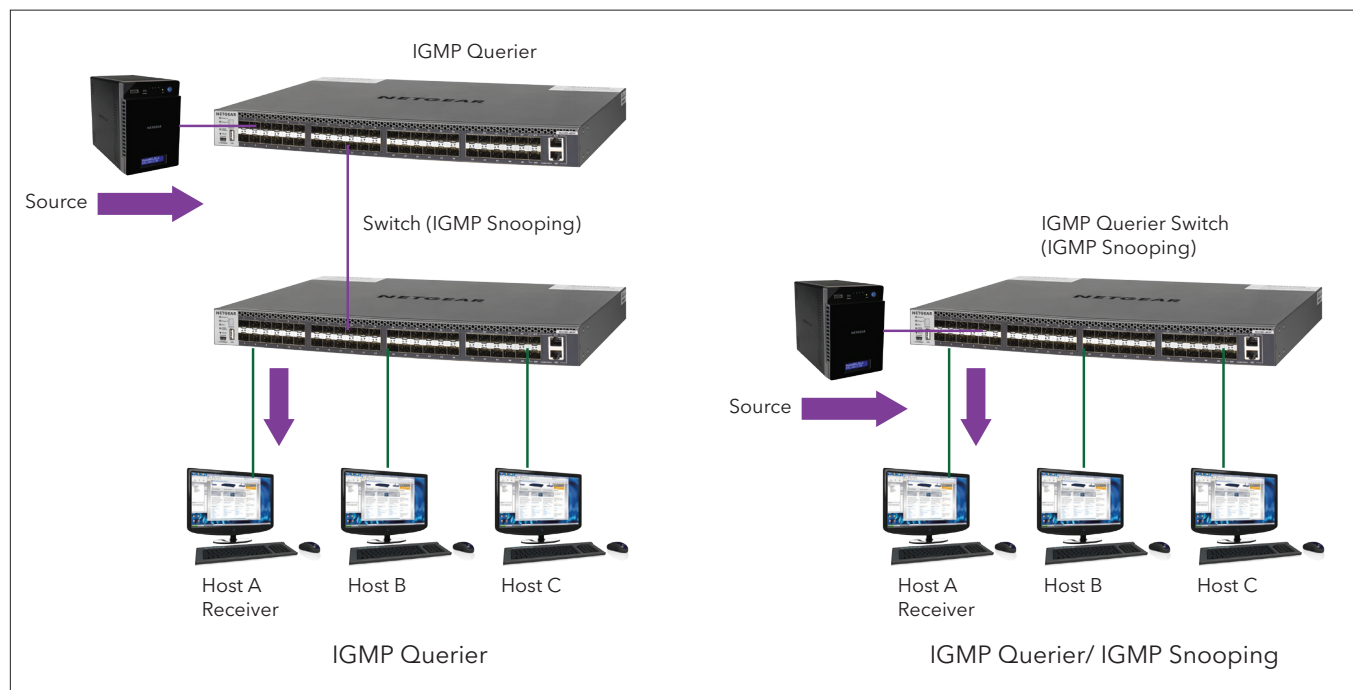
² OSI (Open System Interconnection) stack defines a computer networking framework to implement protocols in seven layers.

Theory of Operation with an L2/L3 Pro AV Multicast Infrastructure

There are 2 types of multicast infrastructures to stream multicast. The first type is Layer 2 multicast where IGMP is used in a single, flat network environment. The second type of network infrastructure is Layer 3 multicast where PIM is used on a routed environment with different subnets and routers.

Layer 2 IGMP is used on a single, configured virtual LAN (VLAN). When the AVoIP endpoint encodes the video stream to a multicast stream, the multicast stream is then transported to the 3rd (Network) layer of the OSI stack with a defined multicast group. With IGMPv2 (RFC 2236), the stream is transported to the IGMP snooping querier. The querier's responsibility is to send out IGMP group membership queries on a timed interval, to retrieve IGMP membership reports from active members, and to allow updating of the group membership tables. There is one active IGMP querier per subnet/VLAN. In general, the IGMP snooping querier takes the responsibility of a multicast router⁴ (mrouter) when one is not available.

One of the other IGMP querier tasks is to provide IGMP snooping for the VLAN. IGMP snooping optimizes bandwidth and listens for IGMP messages. The querier then forwards the multicast packets from the VLAN only to the Ethernet ports that are sources of IGMP membership reports and keeps a cache of the entries in a table on the switch.



³ A group membership is any client that has joined a multicast group. A group member responds to IGMP querier by sending a membership report to the group multicast address of all groups it is a member of for a general query and a membership report to the group multicast address for a group query, to keep the tables updated.

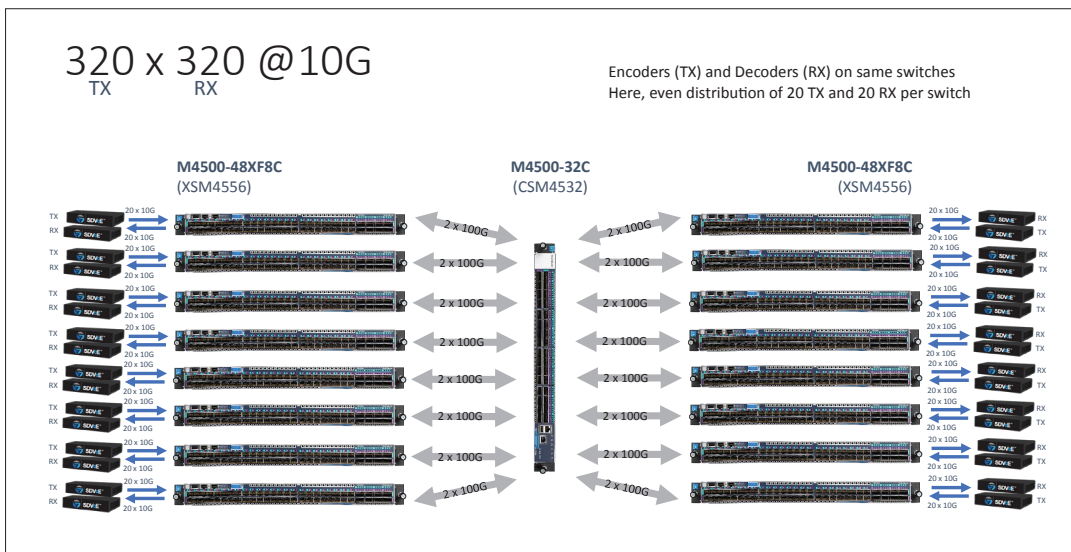
⁴ An mrouter is essentially the multicast router that blocks all known and unknown multicast data streams. The mrouter aims to forward IGMP PDU's and behaves as any other port. It also forwards multicast data streams only if an IGMP membership (v1/v2) has been received through this port.

Basically, when the AV decoder desires to join a multicast stream (multicast group), the IGMP querier receives the 'JOIN' request via IGMP snooping and places the entry into the cache table. This allows the multicast stream request to be transported to the decoder and the video to be streamed to its output.

The IGMP querier also manages the 'LEAVE' request of the hosts. The 'LEAVE' request is a feature of IGMP v2 (RFC 2236). When a group member requests to leave the multicast group, it sends out a 'LEAVE' request to the IGMP querier, indicating its desire to leave the multicast group. The IGMP querier receives the 'LEAVE' request, removes the entry from the IGMP cache table and stops the multicast traffic. This method may not be immediate, but the traffic will be stopped eventually through the use of timers. With the fast-leave functionality on the switch, this process happens significantly quicker. The device can be immediately removed from the Layer 2 forwarding table upon receiving a leave message for a multicast group. This eliminates the need to send out MAC-based general queries.

Multicast can become very bandwidth intensive. With the use of IGMP, IGMP Snooping, and IGMP fast-leave, the bandwidth can be optimized. However, due to the nature of multicast, unknown or unwanted streams can also bombard the IGMP querier and cause delays or loss of bandwidth. The solution to this dilemma is to be sure the multicast flooding control feature is set to block all known (hosts that are not requesting the stream) and unknown multicast data streams.

Many switch manufacturers do not make Layer 2 IGMP configurations as simple as NETGEAR does. Most other 10G / 100G switches allow AV vendors to use their switch in the ProAV environment - but that comes at a price. Switch configuration with other vendors is much more complicated and may lead to user error. The NETGEAR M4300 and M4500 switches come standard with the IGMP Plus feature. NETGEAR IGMP Plus allows the ProAV engineers to configure any Layer 2 VLAN on their network infrastructure to send and receive multicast streams with a single command via the command line interface (CLI) or simple check box in the web-based GUI. For instance, in the illustration below, IGMP Plus can quickly configure a semi-large ProAV spine and leaf network architecture with only using a single command on all 17 M4500 switches.



In a Layer 3 environment, network engineers often use Protocol Independent Multicast (PIM) routing protocol. PIM is normally used in a routed network environment where the streaming host is located on a different router and across networks. PIM builds multicast routes which ensures the shortest path is taken and suppresses loops. There are two types of common PIM modes: PIM dense mode (PIM-DM) and PIM sparse mode (PIM-SM). PIM-DM uses a flood then prune method. The multicast traffic is flooded to all of the hosts, but then pruned for hosts that are not requesting the multicast stream. PIM-SM uses an entirely different method. In PIM-SM the multicast router can take a role as either the rendezvous point (RP)⁵ or designated router (DR)⁶.

PIM-DM (RFC 3973) uses the existing unicast routing table and join, prune, and graft mechanisms to build a multicast tree. PIM-DM creates source-based, shortest-path distribution trees making use of reverse path forwarding (RPF). PIM-DM cannot be used to build a shared distribution tree, as PIM-SM can. PIM-DM assumes that when a sender starts sending data, all downstream routers and hosts want to receive a multicast datagram. PIM-DM initially floods multicast traffic throughout the network. Routers that do not have any downstream neighbors prune back the unwanted traffic.

Apart from the prune messages, PIM-DM also uses two additional messages (graft and assert). Graft messages are used whenever a new host wants to join the group. Assert messages are used to shut off duplicate flows onto the same multi-access network.

To minimize the repeated flooding of packets and subsequent pruning associated with a particular source and group (written as 'S,G') address pair, PIM-DM uses a state refresh message. This message is sent by the routers directly connected to the source and is propagated throughout the network. When received by a router on its RPF interface, the state refresh message causes an existing prune state to be refreshed. State refresh messages are generated periodically by the router directly attached to the source.

PIM-SM uses shared trees by default and implements source-based trees for efficiency. PIM-SM assumes that no hosts want the multicast traffic unless they specifically ask for it and creates a shared distribution tree centered on a defined rendezvous point (RP). Traffic from the source is relayed to the receivers. Senders first send the multicast data to the RP, which in turn sends the data down the shared tree to the receivers.

Shared trees centered on an RP do not necessarily provide the shortest, most optimal path. In such cases PIM-SM provides a means to switch to more efficient source-specific trees. A data threshold rate is defined for toggling between trees. PIM-SM uses a bootstrap router (BSR), which advertises information to other multicast routers about the RP. In a given network, a set of routers can be administratively enabled as candidate bootstrap routers. If it is not apparent which router should be the BSR, the candidates flood the domain with advertisements. The router with the highest priority is elected. If all the priorities are equal, then the candidate with the highest IP address becomes the BSR. PIM-SM is defined in RFC 4601.

⁵ Rendezvous Point (RP) is responsible for keeping track of multicast sources and building trees to distribute multicast to other routers. The RP router discovers the sources of all multicast groups and forwards multicast packets to the designated router requesting them.

⁶ Designated Router (DR) is any multicast router in a network that is not the RP. The function of the DR is to send and forward PIM 'Join' messages to the RP to initiate a multicast to its networks and if attached to a multicast source, forward encapsulated multicast packets to the RP for distribution.

Conclusion

Selecting the correct multicast vehicle (Layer 2 versus Layer 3) can be a daunting decision. The decision usually depends on a few factors: how the network infrastructure is designed and where on the network the multicast is expected to stream. In most ProAV installations, it is recommended that the AV network be isolated to its own network for seamless, uninterrupted, and high bandwidth Layer 2 multicast streaming.

The key aspects to keep in mind are the selection of AV encoders/decoders, the compatibility of the Ethernet switch and the amount of configuration required to implement a ProAV network. With NETGEAR's IGMP Plus feature on the M4300 and M4500 switches, the customer gets a quick, out-of-the-box solution to a complex IGMP configuration. It eliminates the headache of command line configuration and speeds up the process of implementing an AV over IP solution on a complex ProAV network topology.

For more information on NETGEAR and their Pro AV offered solutions, please visit:
<https://www.netgear.com/landings/ProAvCampaign/>

Appendix

For more information on Multicast group addresses and RFC's. Please visit the links below.

Multicast addressing:

https://en.wikipedia.org/wiki/Multicast_address

Multicast RFC Links:

RFC 2236 Internet Group Management Protocol, Version 2 - <https://tools.ietf.org/html/rfc2236>

RFC 3973 Protocol Independent Multicast - Dense Mode - <https://tools.ietf.org/html/rfc3973>

RFC 4601 Protocol Independent Multicast - Sparse Mode - <https://tools.ietf.org/html/rfc4601>