Enterprise IP Multicast Design and Troubleshooting

Part 1

Cisco Advanced Services
Maurice Flint, CCIE #9003
June 3rd, 2009
Multicast Components – Topics for Discussion

<table>
<thead>
<tr>
<th>Topic</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Multicast</td>
<td>PIM SM, DM, Auto RP, IGMP v2, CGMP</td>
</tr>
<tr>
<td>One-to-Many</td>
<td>SSM &amp; IGMP v3</td>
</tr>
<tr>
<td>Many-to-Many</td>
<td>PIM – Bi-Directional</td>
</tr>
<tr>
<td>Interdomain Multicast</td>
<td>MBGP, MSDP, Anycast RP, RGMP, BSR</td>
</tr>
<tr>
<td>Reliable Multicast</td>
<td>PGM</td>
</tr>
<tr>
<td>Multicast VPN</td>
<td>MVPN</td>
</tr>
<tr>
<td>Multicast MPLS</td>
<td>mLDP and P2MP, RSVP-TE</td>
</tr>
<tr>
<td>Multicast Security</td>
<td></td>
</tr>
<tr>
<td>Multicast OAM</td>
<td></td>
</tr>
<tr>
<td>High Availability</td>
<td></td>
</tr>
<tr>
<td>Wireless Multicast</td>
<td></td>
</tr>
</tbody>
</table>

- **Campus Multicast**
  - IGMP Snooping
  - Multicast Source X
  - DR

- **Interdomain Multicast**
  - Multicast Source Y
  - DR
  - RP
  - MSDP
  - MBGP

- **Core**
  - PIM-SM
  - PIM-SSM
  - MVPN
Agenda

- Multicast Essentials
- Enterprise Multicast Service Model
  - Any Source Multicast / Internet Standard Multicast
  - Source Specific Multicast
  - Case Study
- Multicast Campus Network
  - IGMP Snooping
  - PIM Snooping
IP Multicast Essentials
IP Multicast Essentials
Unicast vs Multicast

One-to-One

Host

Number of Streams

Multicast

Router

One-to-Many

Router
## Multicast Applications

<table>
<thead>
<tr>
<th>Multimedia</th>
<th>Real Time</th>
<th>Non-Real Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock Quotes</td>
<td>Stock Quotes</td>
<td>Information Delivery</td>
</tr>
<tr>
<td>News Feeds</td>
<td>Stock Quotes</td>
<td>Information Delivery</td>
</tr>
<tr>
<td>White boarding</td>
<td>Stock Quotes</td>
<td>Information Delivery</td>
</tr>
<tr>
<td>Interactive Gaming</td>
<td>Live Video (IPTV)</td>
<td>Information Delivery</td>
</tr>
<tr>
<td></td>
<td>Video conferencing</td>
<td>Information Delivery</td>
</tr>
<tr>
<td></td>
<td>Live Internet Audio</td>
<td>Information Delivery</td>
</tr>
<tr>
<td></td>
<td>Hoot &amp; Holler</td>
<td>Information Delivery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data-only</th>
<th>Multicast Applications</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Live Video (IPTV)</td>
<td>Video, Web servers</td>
</tr>
<tr>
<td></td>
<td>Video conferencing</td>
<td>Video, Web servers</td>
</tr>
<tr>
<td></td>
<td>Live Internet Audio</td>
<td>Video, Web servers</td>
</tr>
<tr>
<td></td>
<td>Hoot &amp; Holler</td>
<td>Video, Web servers</td>
</tr>
</tbody>
</table>

### Non-Real Time

- Replication
  - Video, Web servers
  - Kiosks
- Content delivery
- Information Delivery
  - Server-to-Server, Server-to-Desktop
- Database replication
- Software distribution
Cisco Applications that use Multicast: 
It’s more than just the technology...

<table>
<thead>
<tr>
<th></th>
<th>Application</th>
<th>Feature</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Scientific Atlanta</td>
<td>Commercial Quality IP Video</td>
<td><a href="www.in.cisco.com/WWSales/wwops/wwssp/sai/archives/scientificatlanta">www.in.cisco.com/WWSales/wwops/wwssp/sai/archives/scientificatlanta</a></td>
</tr>
<tr>
<td>2</td>
<td>Video 2.0 (VQE-S,VQE-C)</td>
<td>Entertainment Grade IP/TV</td>
<td><a href="www.in.cisco.com/sp/messaging/ipngn7.shtml">www.in.cisco.com/sp/messaging/ipngn7.shtml</a></td>
</tr>
<tr>
<td>5</td>
<td>Cisco Digital Media System</td>
<td>Unified Communications</td>
<td><a href="www.cisco.com/emtg/dm2bu/">www.cisco.com/emtg/dm2bu/</a></td>
</tr>
</tbody>
</table>
Why Multicast
Multicast Advantages

- Supports **One-to-Many Applications**: Streaming multimedia, Music on Hold, etc.
- Supports **Periodic data delivery-“push” technology**: Finance, Entertainment, Resource Apps
- Enhanced **Efficiency & Performance**: Eliminates traffic redundancy & reduces server CPU loads

Challenges: Why isn’t IP Multicast widely deployed?

- Protocol complexity & support for a widely-deployed implementation
- Security – Denial of Service Attacks (Intranet & Internet)
- Best Effort Delivery – UDP-based with no Congestion-Avoidance
  * Pragmatic General Multicast (PGM)

**Example: Audio Streaming**
All Clients Listening to the Same 8 Kbps Audio

![Graph showing traffic comparison between Unicast and Multicast]
IP Multicast
Technical Overview:
Building a Solid Foundation
Rube Goldberg (1883-1970)
Inventor / Cartoonist

What’s Multicast Routing Have Anything to do with a “Goldberg Machine?”

A “Rube Goldberg Machine” is an extremely complicated device that executes a very simple task in a complex, indirect way.

Traditional Multicast is our Rube Goldberg Machine!

http://www.rube-goldberg.com
Multicast Address Range
IPv4 Assignment – Class D Address Space

- Link-Local Address Range
  224.0.0.0/24 – local subnet (TTL=1)

- Global Address Range
  224.0.1.0 – 238.255.255.255 – (Globally scoped to/from Internet)
  232.0.0.0/8 – Source Specific Multicast (SSM)
  233.0.0.0/8 - GLOP (ASN Registered)
    - AS number is inserted in middle two octets.
    - Remaining low-order octet (233.x.x.0/24) used for group assignment

  Extended GLOP Addresses (EGLOP – RFC 3180)
    - Make use of private AS numbers
    - Assigned by a Registration Authority

- Administratively Scoped Address Range
  239.0.0.0 – 239.255.255.255 (Private address range – similar to RFC 1918)
Gotcha:
Caveat exists concerning multicast address overlap.

• Multicast represented by MAC 0x01005e
• 5 bits lost – only 1 OUI purchased IEEE
• Creates 32:1 address overlap
• 32 L3 addresses can map to the same MAC!
Multicast Address Assignment
Administratively Scoped Address Range

- **Address Range: 239.0.0.0/8**
  - Private multicast address space
  - Similar to RFC1918 private unicast address space

- **RFC 2365 Administratively Scoped Zones**

  **Organization-Local Scope (239.192/14)**
  - Largest scope within the Enterprise network
  - Regional or global applications that are used within a private enterprise network.

  **Local Scope (239.255/16)**
  - Smallest possible scope within the Enterprise network
  - Expands downward in address range
  - Other scopes may be equal but not smaller
  - Targets local applications that are isolated within a site/region & blocked on defined boundaries.
Scope Relative Example—Local Scope
Top 256 Addresses of Every Admin. Scope Range are reserved.

- 239.0.0.0
- 239.255.0.0
- 239.255.255.0
- 239.255.255.255

Local Scope

Local Scope Relative
(Not to Scale)

<table>
<thead>
<tr>
<th>Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>239.255.255.247</td>
<td>MBUS</td>
</tr>
<tr>
<td>239.255.255.248</td>
<td>AAP</td>
</tr>
<tr>
<td>239.255.255.249</td>
<td>DHCPv4</td>
</tr>
<tr>
<td>239.255.255.250</td>
<td>SSDP</td>
</tr>
<tr>
<td>239.255.255.251</td>
<td>Multicast Discovery of DNS Services</td>
</tr>
<tr>
<td>239.255.255.252</td>
<td>MZAP Protocol</td>
</tr>
<tr>
<td>239.255.255.253</td>
<td>SLPv2 Protocol</td>
</tr>
<tr>
<td>239.255.255.254</td>
<td>MADCAP Protocol</td>
</tr>
<tr>
<td>239.255.255.255</td>
<td>SAP Session Announcement Protocol (SDR)</td>
</tr>
</tbody>
</table>
Scope Relative Example— Org-Local Scope
Top 256 Addresses of Every Admin. Scope Range are reserved.

<table>
<thead>
<tr>
<th>Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>239.195.255.247</td>
<td>MBUS</td>
</tr>
<tr>
<td>239.195.255.248</td>
<td>AAP</td>
</tr>
<tr>
<td>239.195.255.249</td>
<td>DHCPv4</td>
</tr>
<tr>
<td>239.195.255.250</td>
<td>SSDP</td>
</tr>
<tr>
<td>239.195.255.251</td>
<td>Multicast Discovery of DNS Services</td>
</tr>
<tr>
<td>239.195.255.252</td>
<td>MZAP Protocol</td>
</tr>
<tr>
<td>239.195.255.253</td>
<td>SLPv2 Protocol</td>
</tr>
<tr>
<td>239.195.255.254</td>
<td>MADCAP Protocol</td>
</tr>
<tr>
<td>239.195.255.255</td>
<td>SAP Session Announcement Protocol (SDR)</td>
</tr>
</tbody>
</table>
Multicast Address Assignment
Address Ranges to Avoid

- Avoid ranges that map to a MAC address of 0x0100-5E00-00xx!
  i.e. 239.128.0/24 and 239.0.0/24

Are the same as 224.0.0.0/24 - Link Local

These addresses are always flooded by Layer 2 switches!
Multicast Address Design
Administrative Scoping Example – Any Source Multicast

- Develop an Enterprise Multicast Address Assignment Design
  - Global Scope – 239.192.0.0/16
  - Regional Scope
    - US/CALA – 239.195.0.0/16
    - EMEA - 239.196.0.0/16
    - APAC - 239.197.0.0/16
  - Site-Local Scope – 239.255.0.0/16
  - Internet - 233.0.0.0/8: GLOP (RFC 2770)
    - ASN registered & have /24 per AS number.

- Develop an Enterprise Multicast Application Assignment Policy
  - IP/TV or DMS, MoH, Others
  - High-Rate and Low-Rate Streams

Guidelines for Enterprise IP Multicast Address Allocation:
Multicast Forwarding
Administrative Boundaries /Scoping

Administrative Boundary = 239.0.0.0/8

- Configured using the ‘ip multicast boundary <acl>’ interface command
Administratively-Scoped Zones
Multicast Boundary Filters

- Multicast Boundary filters block multicast traffic in both directions on the configured interface.
- The use of TTL filters are no longer recommended.

Interface Serial0
ip multicast boundary 10
access-list 10 deny 239.195.0.0 0.0.255.255
access-list 10 permit any

Interface Serial1
ip multicast boundary 10
access-list 10 deny 239.193.0.0 0.0.255.255
access-list 10 permit any
Multicast Boundary Extensions

- **Boundary Extensions Supported**
  - `ip multicast boundary <acl> [ in | out | filter-autorp ]`
  - `access-list 10 deny 239.192.0.0 0.0.255.255`
  - `access-list 10 permit any`

- **in** – filters source traffic coming into interface
- **out** – prevents state from being created on interface
  - IGMP reports and PIM Joins will not create state
  - Interface will not be added to OIL

- **More that one boundary command is allowed on interface but only one instance of in, out or filter-autorp**

- **Available in IOS** (12.2 and above)

- **Will be available on 6500 in 12.2SXI** (Whitney 2)
IP Multicast Essentials – Terms
Distribution Tree(s)

- **Source Tree**
  - Rooted at the Source
  - Represented by (S,G) entry

- **Shared Tree**
  - Rooted at the Rendevous Point
  - Represented by (*.G) entry (Sparse Mode)
Router#sh ip mroute 224.1.2.3
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned
   R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT
   M - MSDP created entry, X - Proxy Join Timer Running
   A - Advertised via MSDP
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*) , 224.1.2.3), 00:04:28/00:01:32, RP 171.68.28.140, flags: SC
   Incoming interface: Serial1, RPF nbr 171.68.28.140,
   Outgoing interface list:
      Ethernet0, Forward/Sparse, 00:00:30/00:02:30

(10.10.10.1/32, 224.1.2.3), 00:04:28/00:01:32, flags: CT
   Incoming interface: Serial0, RPF nbr 171.68.28.190
   Outgoing interface list:
      Serial1, Forward/Sparse, 00:04:28/00:01:32
      Ethernet0, Forward/Sparse, 00:00:30/00:02:30
Reverse Path Forwarding (RPF)

- **What is RPF?**
  A router forwards a multicast datagram only if received on the upstream interface to the source (i.e. it follows the distribution tree).

- **The RPF Check**
  - The routing table used for multicasting is checked against the “source” address in the multicast datagram.
  - If the datagram arrived on the interface specified in the routing table for the source address; then the RPF check succeeds. This becomes the “Incoming” or RPF Interface.
  - Otherwise, the RPF Check fails.
A closer look: **RPF Check Fails**

Multicast Packet from Source 151.10.3.21

Packet Arrived on Wrong Interface! Discard Packet!
IP Multicast Essentials
IP Multicast Routing/ Multicast Forwarding

A closer look: **RPF Check Fails**

R1#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
5 routes using 3052 bytes of memory
3 groups, 0.66 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 1, Packets forwarded: 92, Packets received: 92
RP-tree: Forwarding: 92/0/100/0, Other: 92/0/0
Source: 10.4.1.6/32, Forwarding: 0/0/0/0, Other: 0/0/0

---

Packet Arrived on Wrong Interface!
Discard Packet!

<table>
<thead>
<tr>
<th>Network</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>151.10.0.0/16</td>
<td>S1</td>
</tr>
<tr>
<td>198.14.32.0/24</td>
<td>S0</td>
</tr>
<tr>
<td>204.1.16.0/24</td>
<td>E0</td>
</tr>
</tbody>
</table>
A closer look: **RPF Check Succeeds**

Multicast Packet from Source 151.10.3.21

RPF Check Succeeds!

<table>
<thead>
<tr>
<th>Unicast Route Table</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>151.10.0.0/16</td>
<td>S1</td>
</tr>
<tr>
<td>198.14.32.0/24</td>
<td>S0</td>
</tr>
<tr>
<td>204.1.16.0/24</td>
<td>E0</td>
</tr>
</tbody>
</table>

Packet Arrived on Correct Interface! Forward out all outgoing interfaces. (i.e. down the distribution tree)
We'll just use the spare 56K line for the IP Multicast traffic and not the T1.

T1/E1 line has best metric to source

no ip pim sparse-mode

We’ll just use the spare 56K line for the IP Multicast traffic and not the T1.

RPF to disabled link!!!!!

T1/E1 56K/64K

src

rcvr
We'll just use the spare 56K line for the IP Multicast traffic and not the T1.

no ip pim sparse-mode
ip pim sparse-mode

T1/E1 line has best metric to source

no ip pim sparse-mode
T1/E1
56K/64K
ip pim sparse-mode

Static multicast routing (static mroutes) can be to resolve this design requirement.

Ip mroute 10.77.86.75 255.255.255.255 10.2.255.72 (RP = 10.77.86.75)
Ip mroute 10.77.76.0 255.255.255.0 10.2.255.72 (Source Subnet = 10.77.76.0/24)
**IP Multicast Essentials**

**Configure Multicast on Every Router**

Classic Partial Multicast Cloud Mistake #2

Highest next-hop IP address used for RPF when equal cost paths exist. Complies with RFC 2362!

We’ll just keep multicast traffic off of certain routers in the network.
Two Options exist for “load-sharing” multicast across equal cost multiple paths (ECMP).

- GRE Tunnels
- IP Multicast Multipath Feature (12.4T)

**IP Multicast Multipath**

- Randomly distribute (*, G) and (S, G) traffic streams across multiple equal-cost reverse path forwarding (RPF) paths.
- Does not necessarily result in balanced IP multicast traffic loads. Networks where there are many traffic streams that utilize approximately the same amount of bandwidth benefit the most.
- Can be configured to load split based on Source, (S,G), or (S,G) and next-hop address.
- All participating interfaces still require PIM.

“Joins” create the OIL
IP Multicast Essentials
Types of Multicast Routing Protocols

- DVMRPv3 (Internet-draft)
- MOSPF (RFC 1584)
- CBT (Internet-draft)
- PIM-DM (Internet-draft)
- PIM Sparse Mode (RFC 2362)
- Source Specific Multicast (SSM)
- Bi-directional PIM (Bidir)
- Multicast VPN (MVPN)

Only PIM-DM is supported in this box
All protocols are supported in this box
Multicast Service Models
Multicast Service Model
IP Multicast Enabled Network

- There are two kinds of multicast-enabled networks available.
  - **Any Source Multicast** – ASM is the original multicast service model as defined in RFC 1112 [Deering]. In this model, a receiver simply joins the group and does not need to know the identity of the source(s). (Also called Internet Standard or Traditional Multicast)
  - **Source Specific Multicast** – RFC 3569 (2003)
    SSM modifies the original service such that in addition to knowing the group address, a receiver must know the relevant source(s). It becomes the application's responsibility to know what kind of multicast capability the network provides.
Multicast Service Model
Essential Protocol Components

- An IP multicast enabled network requires two essential protocol components:
  
  - **Internet Group Membership Protocol** – IGMP is the IP host-based protocol that allows a receiver application to notify a local router(s) that it has joined the group, and initiate the data flow from all sender(s) within the scope.
  
  - **Protocol Independent Multicast** - PIM is an IP router-based protocol that allows routers with multicast group members (receivers) on the local networks to communicate with other routers to ensure that all datagrams sent to the group address are forwarded to all receivers within the intended scope.
Any Source Multicast
Multicast Service Model
Any Source Multicast (ASM) – RFC 1112

- IGMPv2 (Host to Router Signaling)
  
  RFC 2236
  Membership reports used to “Join” multicast groups > 224.0.0.1

- Classic (original) PIMv2 Sparse Mode (Router Signaling)
  
  Current PIMv2 specification is RFC 4601

- Uses both Shared and Source Path Trees
  Requires a Rendezvous Point (RP) and Shared Tree for network-based Source discovery.

- Complex to Troubleshoot

- Susceptible to Denial of Service Attacks
Any Source Multicast – Host-Router Signaling
IGMPv2—Joining a Group

- Joining member sends reports for the multicast group that they want to Join. Called “unsolicited” reports.
Any Source Multicast – Host-Router Signaling
IGMPv2—Verify the Group

IGMP State in “rtr-a”

Group 239.192.1.1 is active
Any Source Multicast – Host-Router Signaling
IGMPv2 – Maintaining the Group

Group 224.1.1.1 Example

- Router sends periodic queries to 224.0.0.1 @ 60-120s
- One member per group per subnet reports
- Other members suppress reports
Any Source Multicast – Host-Router Signaling
IGMPv2 – Leaving a Group

Leaving a Group (224.1.1.1)

- Host sends leave message to 224.0.0.2
- Router sends group-specific query for Group 224.1.1.1
- No IGMP report is received within ~ 3 seconds
- Group 224.1.1.1 times out
Any Source Multicast – Host-Router Signaling
IGMPv2—Querier Election

- Initially all routers send out a query @ 224.0.0.1
- Router with lowest IP address “elected” querier
- Other routers become “non-queriers”
IGMPv2—Querier Election
Determining Which Router is the IGMP Querier

rtr-a>show ip igmp interface e0
Ethernet0 is up, line protocol is up
  Internet address is 1.1.1.1, subnet mask is 255.255.255.0
    IGMP is enabled on interface
    Current IGMP version is 2
    CGMP is disabled on interface
    IGMP query interval is 60 seconds
    IGMP querier timeout is 120 seconds
    IGMP max query response time is 10 seconds
    Inbound IGMP access group is not set
    Multicast routing is enabled on interface
    Multicast TTL threshold is 0
    Multicast designated router (DR) is 1.1.1.1 (this system)
    IGMP querying router is 1.1.1.1 (this system)
    Multicast groups joined: 224.0.1.40 224.2.127.254

IGMP is automatically enabled when PIM is enabled under the interface
IGMPv2

Designated Router (DR)

- If there are multiple routers on a LAN, a designated router (DR) must be elected to avoid duplicating multicast traffic.

- PIM routers follow an election process to select a DR - the PIM router with the highest Interface IP address becomes the DR.

- The DR is responsible for:
  - Sending PIM register and PIM Join and Prune messages toward the rendezvous point (RP) to inform it about host group membership.
  - Sending PIM Join messages for the Shortest Path Tree.

- Gotcha: When troubleshooting PIM, make sure that you identify the DR. The non-DRs will show (*,G) and (S,G) state, but take no action.

- DR Failover - 3 x <query-interval>
  - default interval 30s
Any Source Multicast – Router to Router Signaling
Activating PIM-SM (Sparse-Mode)

- **Interface configuration command** - `ip pim sparse-mode`
  
  Enables multicast forwarding on the interface
  
  Controls the **interface's** mode of operation
  
  Separate control and data plane > Unicast routing protocol-independent

- **Explicit Join Model**
  
  Receivers must first be “connected” to the tree before traffic begins to flow
  
  Traffic only sent down “joined” branches

- **Must configure an RP**
  
  Used to map the source multicast group address to the IPv4 unicast address of the source.

- **Uses both Shared and Shortest Path Distribution Trees**
  
  Starts out using Shared Tree, then switches over to the source-tree
  
  Very complex to troubleshoot
Any Source Multicast (ASM)
PIM-SM with IGMPv2 Operation

Source
10.4.1.6/24

Source transmits the stream @ 239.192.1.1, but isn’t responsible to determine where to send the packets.

Receiver
10.1.1.5/24

Receiver wants to join 239.192.1.1, but doesn’t know the unicast IP address of the source.
Any Source Multicast (ASM)
PIM-SM with IGMPv2 > Map the Group to the Source IP

Any Source Multicast requires a control plane which is anchored by the Rendezvous Point to “join” the Source and Receivers.

The Rendezvous Point is responsible for multicast Group to Source IP address mapping.
Any Source Multicast (ASM)
PIM-SM Shared Tree Join Process

1. Receiver sends IGMPv2 “Join”
2. Last hop PIM router forwards the “Join” to the RP for 239.192.1.1

(*, G) State created via the Shared Tree. The Outgoing Interface List (OIL) is created.
The receiver has joined the group, but is waiting for the source.
Any Source Multicast (ASM)
PIM-SM Shared Tree Join Process – Case Study

R5(config-if)# ip igmp join-group 239.192.1.1

R6# ping 239.192.1.1

R5
Receiver
10.1.1.5/24
1. Receiver sends IGMPv2 “Join”
2. Last hop PIM router forwards the “Join” toward the RP for 239.192.1.1
Any Source Multicast (ASM)
PIM-SM Shared Tree Join Process – Case Study

3. All PIM routers in the path forwards the “Join” to the RP for 239.192.1.1
Any Source Multicast (ASM)

PIM-SM Shared Tree Join Process – Case Study

R3#sh ip mroute 239.192.1.1
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
L - Local, P - Pruned, R - RP-bit set, F - Register flag,
T - SPT-bit set, J - Join SPT, M - MSDP created entry,
X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
U - URD, I - Received Source Specific Host Report,
Z - Multicast Tunnel, z - MDT-data group sender,
y - Joined MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

Group: 239.192.1.1, Source count: 0, Packets forwarded: 0, Packets received: 0

R3#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
5 routes using 3830 bytes of memory
3 groups, 0.66 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 0, Packets forwarded: 0, Packets received: 0

(*, G) State created via the Shared Tree. The Outgoing Interface List (OIL) is created.

4. RP receives the “Join” for 239.192.1.1 and creates the (*,G) Tree
5. Use the “show ip mroute x.x.x.x “count” keyword to verify traffic for the *,G tree
Any Source Multicast (ASM)  
PIM-SM Source Registration Process

Building the PIM-SM Shared Tree is Complete!

5. Source comes “online”. The 1st hop router builds the (*,G) and (S,G) entries
6. The 1st hop router sends Register with unicast stream to the RP for 239.192.1.1
7. Multicast (S, G) traffic arrives at the RP (encapsulated in unicast IPv4 packets). The RP forwards the traffic based on the Outgoing Interface List (OIL). Note: No RPF for Receive Joins.
8. RP sends a “Join” toward the source.
9. The Shortest Path Tree from the RP to the Source is built.
10. RP sends a Register-Stop back to the first-hop router to complete the Register process.
Any Source Multicast (ASM)
PIM-SM Shared Tree Source Registration – Case Study

4. Source comes “online”. The 1st hop router sends Register to the RP for 239.192.1.1
Any Source Multicast (ASM)
PIM-SM Shared Tree Source Registration – Case Study

R3 (RP)# debug ip pim
*May 30 11:50:39.064: PIM(0): Received v2 Register on Ethernet0/0 from 10.34.1.4 for 10.4.1.6, group 239.192.1.1
*May 30 11:50:39.064: PIM(0): Insert (10.4.1.6,239.192.1.1) join in nbr 10.34.1.4’s queue
*May 30 11:50:39.064: PIM(0): Forward decapsulated data packet for 239.192.1.1 on Ethernet0/2
*May 30 11:50:39.064: PIM(0): Building Join/Prune packet for nbr 10.34.1.4
*May 30 11:50:39.064: PIM(0): Adding v2 (10.4.1.6/32, 239.192.1.1), S-bit Join
*May 30 11:50:39.064: PIM(0): Send v2 join/prune to 10.34.1.4 (Ethernet0/0)
*May 30 11:50:40.032: PIM(0): Received v2 Register on Ethernet0/0 from 10.34.1.4 for 10.4.1.6, group 239.192.1.1
*May 30 11:50:40.032: PIM(0): Send v2 Register-Stop to 10.34.1.4 for 10.4.1.6, group 239.192.1.1
*May 30 11:53:54.644: PIM(0): Received v2 Join/Prune on Ethernet0/2 from 10.32.1.2, to us
*May 30 11:53:55.916: PIM(0): Received v2 Join/Prune on Ethernet0/2 from 10.32.1.2, to us
*May 30 11:53:55.916: PIM(0): Update Ethernet0/2/10.32.1.2 to (*, 239.192.1.1), Forward state, by PIM *G Join
*May 30 11:53:55.916: PIM(0): Update Ethernet0/2/10.32.1.2 to (10.4.1.6, 239.192.1.1), Forward state, by PIM *G Join

R3#sh ip mroute 239.192.1.1
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
= L - Local, P - Pruned, R - RP-bit set, F - Register flag,
T - SPT-bit set, J - Join SPT, M - MSDP created entry,
X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
U - URD, I - Received Source Specific Host Report,
Z - Multicast Tunnel, Z - MDT-data group sender,
Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.192.1.1), 00:00:59/stopped, RP 10.3.3.3, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
   Ethernet0/2, Forward/Sparse, 00:00:59/00:02:34

(10.4.1.6, 239.192.1.1), 00:00:03/00:02:59, flags: T
Incoming interface: Ethernet0/0, RPF nbr 10.34.1.4
Outgoing interface list:
   Ethernet0/2, Forward/Sparse, 00:00:03/00:02:56 > Automatically populates based on (*,G) OIL > no RPF

4. Source comes “online”.  The 1st hop router sends Register to the RP for 239.192.1.1
Any Source Multicast (ASM)
PIM-SM Shared Tree Traffic Flow – Case Study

Building the PIM-SM Shared Tree is Complete!

R3#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
6 routes using 4340 bytes of memory
3 groups, 1.00 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 1, Packets forwarded: 5, Packets received: 5
RP-tree: Forwarding: 1/0/100/0, Other: 1/0/0
Source: 10.4.1.6/32, Forwarding: 4/1/100/0, Other: 4/0/0
Any Source Multicast (ASM)
PIM-SM Shared Tree Traffic Flow – Case Study

Building the PIM-SM Shared Tree is Complete!

R2#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
5 routes using 3194 bytes of memory
3 groups, 0.66 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 0, Packets forwarded: 5, packets received: 5
RP-tree: Forwarding: 5/1/100/0, Other: 5/0/0
Any Source Multicast (ASM)
PIM-SM Shared Tree Traffic Flow – Case Study

Building the PIM-SM Shared Tree is Complete!

R1#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
5 routes using 3052 bytes of memory
3 groups, 0.66 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 1, Packets forwarded: 5, Packets received: 5
RP-tree: Forwarding: 5/0/100/0, Other: 5/0/0
Source: 10.4.1.6/32, Forwarding: 0/0/0, Other: 0/0/0
Any Source Multicast (ASM)

PIM-SM Shared Tree Traffic Flow – Case Study

Building the PIM-SM Shared Tree is Complete!

R6#ping
Protocol [ip]:
Target IP address: 239.192.1.1
Repeat count [1]: 5
Datagram size [100]:
Timeout in seconds [2]:
Extended commands [n]:
Sweep range of sizes [n]:
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echo to 239.192.1.1, timeout is 2 seconds:

Reply to request 0 from 10.1.1.5, 20 ms
Reply to request 1 from 10.1.1.5, 4 ms
Reply to request 2 from 10.1.1.5, 4 ms
Reply to request 3 from 10.1.1.5, 8 ms
Reply to request 4 from 10.1.1.5, 4 ms
11. Once multicast traffic arrives at the last hop PIM router via the Shared Tree, a “Join” request is sent toward the source.

12. Traffic flows down the Shortest Path Tree

13. The Shared Tree Path is Pruned toward the RP.

14. If no other Receivers exist, the RP Prunes the Source Path Tree.
Any Source Multicast (ASM)  
PIM-SM SPT Switchover Process – Case Study  

Building the PIM-SM Shared Tree is Complete!

1. Packets arrive at R1 via the Shared (*,G) tree – the purpose of the Shared Tree is Complete!  
2. R1 performs PIM SPT switchover by send (S,G) Join.

The purpose of the Shared Tree Join is complete > the Last-hop Router has learned the IP address of the Source!
Any Source Multicast (ASM)
PIM-SM SPT Switchover Process – Case Study

Building the PIM-SM Shared Tree is Complete!

3. Packets begin to flow down the Source (S,G) Tree to R1. R1 also sends a “Prune” message toward the RP.
Any Source Multicast (ASM)
PIM-SM SPT Switchover Process – Case Study

Building the PIM-SM Shared Tree is Complete!

R1#sh ip mroute 239.192.1.1 count
IP Multicast Statistics
4 routes using 2794 bytes of memory
2 groups, 1.00 average sources per group
Forwarding Counts: Pkt Count/Pkts(neg(-) = Drops) per second/Avg Pkt Size/Kilobits per second
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 239.192.1.1, Source count: 1, Packets forwarded: 5, Packets received: 5
RP-tree: Forwarding: 1/1/100/0, Other: 1/0/0
Source: 10.4.1.6/32, Forwarding: 4/0/100/0, Other: 4/0/0

3. Packets begin to flow down the Source (S,G) Tree to R1. R1 also sends a “Prune” message toward the RP.
Any Source Multicast (ASM)
Issue concerning Multiple Sources

Issue: A 2nd Source appears in the network streaming the same multicast group

Solution: Accept-Register filters to prevent unwanted Multicast flows

ip pim accept-register list 10
access-list 10 permit 10.4.1.6

Eventually, (S, G) Traffic flow creates an issue as two sources exist.
Any Source Multicast (ASM)
How Does the Network Know About the RP?

- **Static configuration**
  Manually on every router in the PIM domain

- **AutoRP**
  Routers learn RP automatically
  Option exists to provide Load-Balancing & Redundancy

- **BSR** (Bootstrap Router)
  draft-ietf-pim-sm-bsr

PIMv2 for Sparse-mode (RFC 2362) defines a Bootstrap mechanism that permits all PIM-SM routers within a domain to dynamically learn all Group-to-RP mappings and avoid any manual RP configurations > AutoRP & BSR
Static RPs

- Hard-configured RP address
  - When used, must be configured on every router
  - All routers must have the same RP address
  - RP failover not possible
    - Exception: if anycast RPs are used

- Command

  \[ \text{ip pim rp-address <address> [group-list <acl>] [override]} \]

  Optional group list specifies group range
  - Default: range = 224.0.0.0/4 (\text{includes auto-RP groups!!!})

  Override keyword “overrides” auto-RP information
  - Default: auto-RP learned info takes precedence
Auto-RP
Dynamic Group to RP Mapping

RP-Announcements Multicast to the Cisco Announce (224.0.1.39) Group

```
ip pim send-rp-announce loopback0 scope 255 group-list 20 access-list 20 permit 239.192.0.0 0.0.255.255
```
Auto-RP
Dynamic Group to RP Mapping

The active RP for each Multicast group range is selected based on the highest Loopback IP address (router ID) from the candidate RPs.

RP Mapping Agents store the candidate-RPs in a Group-to-RP mapping cache. Each entry in the Group-to-RP mapping cache has an expiration timer that is initialized to the holdtime value (3X) in the received RP-Announce message (60s).

Once the timer expires, the Mapping Agent selects a new RP from its Group-to-RP mapping cache and sends out an RP-Discovery message with the updated Group-to-RP mapping. Failover is less than 3 minutes for any NEW connections (Established Source Path Trees don’t need RPs).

RP-Announcements Multicast to the Cisco Announce (224.0.1.39) Group

```
ip pim send-rp-announce loopback0 scope 255 group-list 20 access-list 20 permit 239.192.0.0 0.0.255.255```

Announce
Dynamic Group to RP Mapping
Auto-RP—From 10,000 Feet

Discovery

RP-Discoveries Multicast to the Cisco Discovery (224.0.1.40) Group

ip pim send-rp-discovery loopback0 scope 255
Dynamic Group to RP Mapping
Auto-RP—From 10,000 Feet

RP announcements and RP discovery messages occur every 60 seconds by default with holdtime of 180 seconds. If no RP is found, the next search is done locally on each router for a static RP mapping. If no static RP mapping is configured, the router defaults to dense mode.

Mapping Agents function independently, multicasting identical Group-to-RP mapping information to all routers in the network. Based on this functionality, each device configured as a mapping agent will advertise identical Group-to-RP mapping information based on the fact that they are using the same selection algorithm of highest IP address to select the active RP. This methodology is recommended as Auto-RP Discovery packets are sent unreliably because the protocol has no provision to detect missed packets and no way to request retransmission. MA updates are sent every 60s or when changes are detected.

RP-Discoveries Multicast to the Cisco Discovery (224.0.1.40) Group

```
ip pim send-rp-discovery loopback0 scope 255
```
Any Source Multicast (ASM)
Auto-RP Failover

- Auto-RP failover time (Default)
  Function of ‘Holdtime’ in C-RP Announcement
  \[
  \text{Holdtime} = 3 \times \text{<rp-announce-interval>}
  \]
  Default \(<\text{rp-announce-interval}>\) = 60 seconds
  Default Failover \(~3\) minutes

- Tuning Auto-RP failover
  Tune candidate RPs using the ‘interval’ clause to control failover times
  \[
  \text{ip pim send-rp-announce <intfc> scope <ttl>}
  \begin{array}{l}
  \text{[group-list acl]} \text{ interval <seconds>}
  \end{array}
  \]
  Smaller intervals = faster RP failover + increased amount of RP traffic
  Increase is usually insignificant
  Total RP failover time reduced > Min. failover \(~3\) seconds
Any Source Multicast (ASM)
PIM-SM AutoRP – Case Study

R1#sh ip pim rp map
PIM Group-to-RP Mappings

Group(s) 239.192.0.0/16
RP 10.3.3.3 (?), v2v1
  Info source: 10.3.3.3 (?), elected via Auto-RP
  Uptime: 16:30:44, expires: 00:02:00
Acl: 10, Static
  RP: 10.3.3.3 (?)

R1#sh access-lists 10
Standard IP access list 10
  10 permit 239.192.1.1

Traffic Flow
Shared Tree
Source Tree
Any Source Multicast
PIM-SM AutoRP – Case Study

R1#sh ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
L - Local, P - Pruned, R - RP-bit set, F - Register flag,
T - SPT-bit set, J - Join SPT, M - MSDP created entry,
X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
U - URD, I - Received Source Specific Host Report,
Z - Multicast Tunnel, z - MDT-data group sender,
Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 224.0.1.39), 00:14:44/stopped, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
Ethernet0/2, Forward/Sparse, 00:14:44/00:00:00
Ethernet0/1, Forward/Sparse, 00:14:44/00:00:00

(10.3.3.3, 224.0.1.39), 00:02:44/00:00:28, flags: PT
Incoming interface: Ethernet0/1, RPF nbr 10.21.1.2
Outgoing interface list:
Ethernet0/2, Prune/Sparse, 00:02:44/00:00:25

(*, 224.0.1.40), 00:18:09/stopped, RP 0.0.0.0, flags: DCL
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
Ethernet0/2, Forward/Sparse, 00:18:10/00:00:00
Ethernet0/1, Forward/Sparse, 00:18:10/00:00:00
Ethernet0/0, Forward/Sparse, 00:18:10/00:00:00

(10.3.3.3, 224.0.1.40), 00:17:42/00:02:20, flags: LT
Incoming interface: Ethernet0/1, RPF nbr 10.21.1.2
Outgoing interface list:
Ethernet0/2, Prune/Sparse, 00:17:42/00:00:00

Traffic Flow
Shared Tree
Source Tree
Fast RP Failover = Anycast RP
Multicast Source Discovery Protocol

When a source registers with one RP, a Source-Active (SA) message will be sent to the other RPs informing them that there is an active source for a particular multicast group.
Fast RP Failover = Anycast RP
Multicast Source Discovery Protocol

By default, RP failover is 3-180s.
Anycast RP can be used to reduce failover interval.
Anycast RP

**Diagram:**
- **Src 239.193.1.1** from **RP1** (10.1.1.1)
- **Src 239.194.1.1** from **RP2** (10.1.1.1)
- **Rec** from **RP1** (10.1.1.1)
- **Rec** from **RP2** (10.1.1.1)
Anycast RP Configuration
With Static RP

Interface loopback 0
  description Anycast RP
  ip address 10.0.0.1 255.255.255.255

Interface loopback 1
  ip address 10.0.0.2 255.255.255.255
  !
  ip msdp peer 10.0.0.3 connect-source loopback 1
  ip msdp originator-id loopback 1

MSDP

RP1

RP2

A

B

C

D

ip pim rp-address 10.0.0.1

Interface loopback 0
  description Anycast RP
  ip address 10.0.0.1 255.255.255.255

Interface loopback 1
  ip address 10.0.0.2 255.255.255.255
  !
  ip msdp peer 10.0.0.3 connect-source loopback 1
  ip msdp originator-id loopback 1

ip pim rp-address 10.0.0.1
Any Source Multicast
Design Recommendations

- Use PIM-SM on interfaces in conjunction with the IP PIM Auto-RP Listener command. This feature permits the two AutoRP groups 224.0.1.39 and 224.0.1.40 to be flooded across interfaces operating in PIM sparse mode. (Available 12.3(4)T, 12.2(28)S)

- As described, Mapping Agents operate independently using the same selection algorithm. Depending on when checked, the mroute table will reflect whichever update it received first. This provides Mapping Agent Failover design.

- Using a “Catch-All” RP (224.0.0.0/4) on the network promotes unauthorized multicast states (applications that have multicast enabled by default) > some MPLS service providers limit the number of IP mroutes permitted on the network. Filtering is on option, but hard to manage.

- A caveat exists with PIM such that multicast group and interface states are treated separately. Although no dense-mode flooding will occur on sparse-mode interfaces, the mroute group cache can be impacted (DM-Fallback) resulting in broken SPT state during RP outages. Use the no ip pim dm-fallback global command or RP of Last Resort design to maintain the sparse-mode state. (DM-Fallback is enabled by default)
Any Source Multicast
Design Recommendation – Filter RPs

- Mapping Agent Security - Caveat for filtering RPs - CSCdv79987

The `ip pim rp-announce-filter rp-list <acl> group-list <acl>` command is insufficiently documented. The filter should be configured as described in the example below.

This following filter allows the rogue rp ip-address to pass the rp filter and then be filtered by the group-list filter which denies all groups. The valid RPs (10.1.1.1 & 10.1.1.2) are exempt.

**Documentation:**

```
access-list 14 permit 10.1.1.1 (RP #1)
access-list 14 permit 10.1.1.2 (RP #2)
access-list 14 deny any

access-list 15 permit 224.0.0.0 15.255.255.255
```

**Actual Configuration**

```
! ip pim rp-announce-filter rp-list 14 group-list 15
! access-list 14 deny 10.1.1.1
access-list 14 deny 10.1.1.2
access-list 14 permit any

access-list 15 deny 224.0.0.0 15.255.255.255
```
Any Source Multicast
RP Placement

- Q: “Where do I put the RP?”
  A: “Generally speaking, it’s not critical”

- Shortest Path Tree’s (SPT’s) are normally used by default
  RP is a place for source and receivers to meet
  Traffic does not normally flow through the RP
  RP is therefore not a bottleneck

- Exception: \textit{SPT-Threshold = Infinity}
  Default = 0
  Traffic stays on the shared tree
  RP \textit{could} become a bottleneck
RP Resource Requirements

- Each (*,G) entry requires 380 bytes + outgoing interface list (OIL) overhead.
- Each (S,G) entry requires 220 bytes + outgoing interface list overhead.
- The outgoing interface list overhead is 150 bytes per OIL entry.
- For example, if there are 10 groups with 6 sources per group and 3 outgoing interfaces:
  - # of (*,G)s x (380 + (# of OIL entries x 150)) = 10 x (380 + (3 x 150)) = 8300 bytes for (*,G)
  - # of (S,G)s x (220 + (# of OIL entries x 150)) = 60 x (220 + (3 x 150))= 40,200 bytes for (S,G)
- A total of 48,500 bytes of memory is required for the mroute table.
General RP Recommendations

- Use Auto-RP
  - When minimum configuration is desired and/or
  - When maximum flexibility is desired

- Pros
  - Most flexible method
  - Easiest to maintain

- Cons
  - Increased RP Failover times vs Anycast
  - Special care needed to avoid DM Fallback
  - Some methods greatly increase configuration

Use PIM Sparse-Mode with IP PIM AutoRP Listener!
General RP Recommendations

- Use Anycast RP’s:
  - When network must connect to Internet or
  - When rapid RP failover is critical

- Pros
  - Fastest RP Convergence method
  - Required when connecting to Internet

- Cons
  - Requires more configuration
  - Requires use of MSDP between RP’s

Remember: RPs are only used for new Connections
Multicast Service Model
Any Source Multicast (ASM) Evaluation

- Uses both Shared Trees and Source Path Trees
  Requires RP and Shared Tree for network-based Source Discovery
  Group to RP Mapping must be consistent in the PIM domain

- Pros:
  Traffic only sent down “joined” branches
  Can switch to optimal source-trees for high traffic sources dynamically
  Unicast routing protocol-independent

- Cons:
  Need some form of RP Failover mechanism – No Single Pt. of Failure
  Dense Mode Fallback can be a problem
  Shared to Source Tree switchover complexities
Source Specific Multicast
Multicast Service Model

Source Specific Multicast (SSM) – RFC 4607 (2006)

- IGMPv3 (Host to Router Signaling) – RFC 3376
  - Adds Include/Exclude Source Lists
  - New IGMPv3 stack required in the O/S
  - Apps must be rewritten to use IGMPv3 Include/Exclude features

- Receivers subscribe to the SSM Channel (S,G)
  - Hosts responsible for source discovery (learning the (S,G) information)
  - Uses out-of-band mechanism to learn the source (web page, content server)
  - Hosts uses IGMPv3 to join specific (S,G) instead of (*,G)

- PIM-SM (Router to Router Signaling)
  - No RPs or Shared Trees > Uses Source Trees Only
  - Only the specified (S,G) flow is delivered to host
  - Data and control planes are decoupled
  - Eliminates possibility of Multicast Content Jammers

- Simplifies Address Allocation
  - IANA - 232.0.0.0/8
  - Private SSM Range – 239.232.0.0/16 recommended
  - Different sources can use the same multicast group address
SSM – Host to Router Signaling
IGMPv3 Source Discovery Example

Source = 1.1.1.1
Group = 224.1.1.1

H1—Member of 224.1.1.1

Source = 2.2.2.2
Group = 224.1.1.1

H1 wants to receive only S = 1.1.1.1 and no other.

With IGMP, specific sources can be joined. S = 1.1.1.1 in this case

IGMPv3: Join 224.1.1.1
Include: 1.1.1.1
SSM – Host to Router Signaling
IGMPv3—Joining a Group

- Joining member sends IGMPv3 Report to 224.0.0.22 immediately upon joining
SSM – Host to Router Signaling
IGMPv3—Joining Specific Source(s)

- IGMPv3 report contains desired source(s) in the Include list
- Only “Included” source(s) are joined
SSM – Host to Router Signaling
IGMPv3—Excluding Specific Source(s)

- IGMPv3 report contains undesired source(s) in the Exclude list
- All sources except “Excluded” source(s) are joined
SSM – Host to Router Signaling
IGMPv3—Maintaining State

- Router sends periodic queries
- All IGMPv3 members respond
  Reports contain multiple Group state records
SSM – Router to Router Signaling
PIM Source Specific Mode (PIM-SSM)

Receiver learns of source, group/port
Receiver sends IGMPv3 (S,G) Join
First-hop sends PIM (S,G) Join directly toward Source

Source

Receiver

Out-of-band source directory, example: web page, content server, etc.

PIM-SM (S, G) Join

IGMPv3 (S, G) Join
SSM – Router to Router Signaling
PIM Source Specific Mode (PIM SSM)

Result: Shortest path tree rooted at the source, with no shared tree.

Out-of-band source directory, example: web page, content server, etc.

Source

B

E

F

Receiver

It doesn’t get any simpler than this!
There are 2 options for PIM SSM Addresses

1. Use IANA assigned SSM group range 232.0.0.0/8 or
2. Use ACL to specify multicast address from 224.0.0.0 through 239.255.255.255 > Cisco recommends 239.232.0.0/16

   R2(config)# ip pim ssm ?
   default  Use 232/8 group range for SSM
   range    ACL for group range to be used for SSM

Use “ssm range ACL” for PIM-SSM

   ip pim ssm range ssm_group_map
   ip access-list standard ssm_group_map
   permit 232.0.0.0 0.255.255.255
   permit 239.232.0.0 0.0.255.255
PIM-SSM

What if the Application does not support IGMPv3?

- Why isn’t Everyone using it?
  Multicast Applications must support IGMPv3
  Application support to learn the Source Out-of-Band

- Option - IGMPv2 Mapping
  1. Static mapping
  2. DNS mapping

- Static SSM Mapping
  
  ip igmp ssm-map enable
  ip igmp ssm-map static <group-range-ACL> <source-1 IP address>
  
  * Only Last-hop PIM routers require the static mapping.

- DNS SSM Mapping
  
  ip igmp ssm-map enable
dns
  ip domain multicast "domain-prefix"
ip name-server server-address1 [server-address2...]
SSM Mapping – DNS Example

**IGMPv2 join**

Receiver

PIM (S,G) join

DNS Record Format:

3.2.1.232 IN A 172.23.20.70

**Reverse DNS lookup for group G**

**DNS response:**

Group G -> Source S
SSM—Summary

- Uses Source Trees only
  
  Hosts are responsible for source and group discovery
  Hosts must use IGMPv3 to signal which (S,G) to join

- Pros:
  
  IP Multicast Address Management Simplified
  Denial of Service Attacks from Unwanted Sources Inhibited
  Easy to Troubleshoot and Manage
  Mechanism provided to migrate from Any Source Multicast.

- Cons:
  
  Requires IGMPv3 support on host or SSM Mapping
  Hosts can create unlimited (S,G) state for non-existent sources
  L2 Multicast Mgmt Protocols (IGMPv3 Snooping support required)
IP Multicast at Layer 2
L2 Multicast Frame Switching
IGMP Snooping

**Problem**: Older L2 switches treat multicast traffic as unknown or broadcast and “flood” the frame to every port

- Today’s L3 aware switches implement IGMP Snooping without suffering performance degradation using L3 ASIC’s and the TCAM.
- IGMP packets intercepted by the NMP or by special hardware ASICs.
- Switch examines content of IGMP messages to determine which ports want what traffic
  - IGMP membership reports
  - IGMP leave messages
IGMP Snooping
L3 Aware Switches

The CPU populates the CAM Table with a wildcard MAC address that matches on any IGMP packets. Frames that match, will be forwarded to the CPU. This prevents the switch from being overloaded.
Hosts join multicast groups either by sending an unsolicited IGMP join message or by sending an IGMP join message in response to a general query from a multicast router.

The second entry tells the switching engine to send frames addressed to the 0x0100.5E01.0203 multicast MAC address that are not IGMP packets (!IGMP) to the multicast router and to the host that has joined the group.
IGMP Snooping
L3 Aware Switches

IGMP snooping suppresses all but one of the host join messages per multicast group and forwards this one join message to the multicast router.

<table>
<thead>
<tr>
<th>MAC Address</th>
<th>L3</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0100.5exx.xxxx</td>
<td>IGMP</td>
<td>0</td>
</tr>
<tr>
<td>0100.5e01.0203</td>
<td>IGMP</td>
<td>1,2,5</td>
</tr>
</tbody>
</table>
IGMP Snooping
L3 Aware Switches

Packets in the stream match on the second CAM Table entry and are switched to ports 2 and 5. The CPU is not burdened with this multicast stream.
IGMP Snooping Caveat
Issues when no Mrouter exists

- **IGMP Snooping “Gotcha”**
  - By default, IGMP Snooping is enabled on Catalyst Switches.
  - Expects to “listen” for IGMP messages to build the Snooping table map.
  - Only forwards some of the IGMP reports to the mrouter.
  - This mechanism "breaks down" in the absence of an mrouter port.

- **Two scenarios exist where multicast is impacted:**
  - L2 Multicast applications within a VLAN that do not use IGMP.
  - Daisy-chained switch “absorbs” the IGMP report from its local receiver.

- **Solutions**
  1. Enable PIM on the VLAN interface
  2. Enable the IGMP Querier feature
  3. Configure a static mrouter port “ip igmp snooping vlan 1 mrouter int fa 1/0/33”
  4. Configure static multicast MAC entries
  5. Disable IGMP Snooping – risk of L2 Flooding

In a routed multicast network, the PIM router acts as the querier, but when multicast routing is not needed, the IGMP Snooping Querier functionality can be used as a way to trigger membership reports.

The feature allows the Layer 2 switch to proxy for a multicast router and send out periodic IGMP queries in that network. This action causes the switch to consider itself an mrouter port. The remaining switches in the network simply define their respective mrouter ports as the interface on which they received this IGMP query.

Configuration – per-VLAN

Interface vlan 1
ip igmp snooping querier
IGMP Snooping Caveat
Flooding Multicast to the Mrouters

- By default, routers do not send IGMP Membership Reports for desired multicast flows - They use PIM control messages.
- IGMP Snooping can constrain multicast on host ports, but has no ability to control mrouter ports.
- **PIM Snooping** is configured to enable the switch to listen to PIM control messages and only forwards multicast flows to the mrouter that need it.
- Works with IGMP Snooping
Constraining Multicast
IGMP & PIM Snooping

No Snooping. Flood on all ports

Traffic in

IGMP Snooping only. Flood only on multicast router ports

IGMP and PIM Snooping. No flooding
Enterprise Multicast Design & Troubleshooting

Summary

- IP Multicast is technology that still provides the best solution for one-to-many communication over an IP infrastructure. Understanding IP multicast design and how to troubleshoot the control-plane permits network engineers to effectively plan and support multicast applications in concert with the rest of their business-critical infrastructure.

- Source Specific Multicast significantly reduces network infrastructure complexity as compared to traditional multicast.

- Configuring interfaces to support PIM Sparse-Mode in conjunction with IP PIM AutoRP listener reduces the impact of dense-mode flooding within a traditional multicast deployment.

- Use IGMP Snooping to reduce flooding in switched infrastructure. It is recognized that IGMP Snooping cannot solve all L2 flooding issues.
IP Multicast Essentials
Disabling Multicast Groups— New Method

- New global command extension

  ```
  ip multicast-routing [group-range <acl>]
  ```

  Router drops all control packets (PIM, IGMP) for denied groups
  Router drops all data packets for denied groups
  No IGMP or PIM state created for denied groups
  IPv4 support ships in 12.5(1st)T and 12.2XSI—Whitney 2
Thank You!

Q and A
More Information

- White papers
- Web and mailers
- Cisco Press

CTO Multicast page:
http://www.cisco.com/go/ipmulticast

Questions:
cs-ipmulticast@cisco.com

Customer support mailing list:
tac@cisco.com

RTFB = “Read the Fine Book”