IPv6 Multicast Tutorial

Senior Architect, APAC Advanced Consulting Engineering
Yasuo Kashimura
Agenda

- IPv6 Basics: header format/addressing
- What is multicast ?
- MLDv1/v2
- Multicast Forwarding and PIM
  - PIM-SM (ASM: Any Source Multicast)
  - PIM-SSM (SSM: Source Specific Multicast)
- MLD Snooping
- IPv6 Multicast Configuration and Show commands,
  Trouble shooting
- Failure case study
- Consideration in deploying IPv6 multicast to SP network
IPv6 Basics
### IPv6 Header Format

<table>
<thead>
<tr>
<th>Version</th>
<th>Traffic Class</th>
<th>Flow Label</th>
<th>Payload Length</th>
<th>Next Header</th>
<th>Hop Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Address</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Newly defined field in IPv6
- Field name is changed in IPv6

# header length is fixed in IPv6
IPv6 Address notation

◆ IPv4 address notation
Binary digit notation (32 bits)

11000000 10101000 00000000 00000001

- For every 8 bits, display in decimal number and separate with ".".

192.168.0.1

◆ IPv6 address notation
Binary digit notation (128 bits)

0010000000000001 0000110110111000 1011111011101111 1100101011111110
0000000000000000 0000000000000000 0000000000000000 0001001000110100

- For every 16 bits, display in Hex number and separate with ":".

2001:0db8:beef:cafe:0000:0000:0000:1234

- Leading zero in 16 bits field can be abbreviated.

2001:db8:beef:cafe:0:0:0:1234

- Compress the zeros with "::".

2001:db8:beef:cafe::1234

#please refer also “draft-kawamura-ipv6-text-representation”
IPv6 Address

◆ IPv6 Address Architecture

<table>
<thead>
<tr>
<th>Global Routing Prefix</th>
<th>Subnet ID</th>
<th>Interface ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Identification</strong></td>
<td>64 bits</td>
<td><strong>Node Identification</strong></td>
</tr>
<tr>
<td>128 bits</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

● **Prefix**
  Upper 64 bits with Global Routing Prefix + Subnet ID

◆ IPv6 Address categories

● **Unicast Address**  one to one communication
  the Address for each network interface
  Global Address, Link-local Address, Unique Local Address

● **Multicast Address**  one(or Many) to Many communication
  receiving hosts are identified by Group Address
  Also used as alternative way of IPv4’s broadcast as well

● **Anycast Address**  one to one of many communication
  can be configured on multiple interfaces/nodes. Assigned for “feature”.
## Unicast Address Format

### Global Unicast Address

<table>
<thead>
<tr>
<th>Bit</th>
<th>Description</th>
<th>Length</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Global Routing Prefix</td>
<td>48</td>
<td>3 bits, 48 bits</td>
</tr>
<tr>
<td></td>
<td>Subnet ID</td>
<td>16</td>
<td>16 bits</td>
</tr>
<tr>
<td></td>
<td>Interface ID</td>
<td>64</td>
<td>64 bits</td>
</tr>
</tbody>
</table>

- Global address (ex) 2001:db8::1

### Link-local Unicast Address

<table>
<thead>
<tr>
<th>Bit</th>
<th>Description</th>
<th>Length</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111111010</td>
<td>0</td>
<td>Interface ID</td>
<td>64 bits</td>
</tr>
<tr>
<td>10</td>
<td>54</td>
<td>54</td>
<td>54 bits</td>
</tr>
</tbody>
</table>

Unique within one Link (fe80::/10)
It is used for the communication within the link.

### Unique Local Unicast Address (ULA)  [RFC4193]

<table>
<thead>
<tr>
<th>Bit</th>
<th>Description</th>
<th>Length</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111110</td>
<td>L</td>
<td>Global ID</td>
<td>40 bits</td>
</tr>
<tr>
<td></td>
<td>16bits</td>
<td>Subnet ID</td>
<td>16 bits</td>
</tr>
<tr>
<td></td>
<td>Interface ID</td>
<td>64</td>
<td>64 bits</td>
</tr>
</tbody>
</table>

L bit:0 future use
1 Locally assigned pseudo-random Global ID
Multicast Address Format

**Multicast Address**

<table>
<thead>
<tr>
<th>Flag</th>
<th>scope</th>
<th>Group ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>11111111</td>
<td>0RPT</td>
<td>scope</td>
</tr>
</tbody>
</table>

8bits 4bits 4bits 112bits

**Flag**

- **T Flag**
  - 0: permanently-assigned ("well-known") multicast address, assigned by IANA
  - 1: non-permanently-assigned multicast address

- **P Flag**
  - 1: Unicast-Prefix-based multicast address (RFC3306) #when P=1, T must be 1

- **R Flag**
  - 1: multicast address that RP address is embedded (RFC3956) #when R=1, P/T must be 1

**Scope: limit the scope of the multicast group**

<table>
<thead>
<tr>
<th>Scope</th>
<th>0000(0)</th>
<th>0001(1)</th>
<th>0010(2)</th>
<th>0100(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>reserved</td>
<td>interface-local scope</td>
<td>link-local scope</td>
<td>admin-local scope</td>
</tr>
<tr>
<td></td>
<td>0101(5)</td>
<td>1000(8)</td>
<td>1110(E)</td>
<td>1111(F)</td>
</tr>
<tr>
<td></td>
<td>site-local scope</td>
<td>organizational-local scope</td>
<td>global scope</td>
<td>reserved</td>
</tr>
</tbody>
</table>

◆ **ex):**

- PIM-SM(ASM) multicast address FF15::1234
- PIM-SSM multicast address FF38::abcd
Multicast Address Format

◆ permanently-assigned multicast address: mainly used for control packets

<table>
<thead>
<tr>
<th>Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF02:0:0:0:0:0:0:1</td>
<td>Link local All IPv6 Nodes</td>
</tr>
<tr>
<td>FF02:0:0:0:0:0:0:2</td>
<td>Link local All IPv6 routers</td>
</tr>
<tr>
<td>FF02:0:0:0:0:0:0:C</td>
<td>Link local DHCP server/relay agent</td>
</tr>
<tr>
<td>FF02:0:0:0:0:1:FFxx:xxxx</td>
<td>Solicited node multicast address (xx:xxxx represents lower 24 bits of node’s unicast/anycast address.)</td>
</tr>
</tbody>
</table>

◆ RFC3306 Unicast-Prefix-based multicast address

If the user has global unicast IPv6 prefix, the user can have globally unique multicast prefix that global unicast prefix is embedded in multicast prefix.

<table>
<thead>
<tr>
<th>11111111</th>
<th>Scope</th>
<th>Reserve</th>
<th>Prefix Len</th>
<th>Network Prefix</th>
<th>Group ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8bits</td>
<td>4bits</td>
<td>4bits</td>
<td>8bits</td>
<td>8bits</td>
<td>32bits</td>
</tr>
</tbody>
</table>

◆ ex):

```
0040:2001:0123:c001:1100:0000:abcd
```

When the User has Unicast Prefix = `2001:0123:c001:1100::/64`. 
Multicast MAC Address for IPv6


**33-33-00-00-ab-cd**

What is Multicast?
Difference between Unicast and Multicast

- **Unicast:** one to one communication

  ![Unicast Diagram]

  - Send to B
  - Src IP = A
  - Dst IP = B

- **Multicast:** one(or Many) to Many that joined to the group

  ![Multicast Diagram]

  - Send to whoever want Group1 information
  - Src IP = A
  - Dst IP = Group1 (Multicast)
Advantage of Multicast

◆ Reduce the server load while sending same contents to many receivers
◆ Effective use of network bandwidth

6Mbps x 1000 receivers => 6Gbps Total

Server just sends 1 x 6Mbps Multicast stream => Less server load

The stream is replicated on network if needed. => Effective use of bandwidth
Multicast applicability

◆ Residential Broadband
  IP/TV, Live Streaming
  Gaming, Contents download

◆ Enterprise network
  E-Learning, broadcasting
  Application delivery
  Multipoint conference
  Sensor network in Factory

◆ Financial
  Hoot and Holler
  Financial systems (Stock exchange, etc)

◆ Public
  Video Surveillance for river, highway
  information broadcasting (voice, disaster information)
  information distribution in local town
IPv6 multicast service example

◆ IPv6 multicast broadcasting system (i-InproV6)
  ● Remote class in preparatory schools
    Cost is 1/10 to compare with satellite broadcasting
    - Initial cost: $several million => about $200k
    - running cost: $100k/month => $10k/month
    Popular teacher’s class is broadcasted to all areas.
    same quality in all areas, more per class profit

◆ Emergency Earthquake information distribution service (OCN)
  ● Emergency Earthquake information/Alert from Meteorological biz support center is distributed to users with urgency/realtime/efficiency.

◆ contents distribution for kiosk hosts at convenience stores. (FamilyMart)
  ● 6,000 stores are dual-stack-ed
  ● change from satellite to broadband with multicast
  ● Distributing large volumes of data by multicast like New product Add/manual for employee

http://becare.co.jp/service/case01.html
Remote Class
http://www.ntt.com/jishinsokuho/index.html
Receiver Application
http://www.ntt.com/jishinsokuho/index.html
Kiosk Host (Fami-port)
Multicast Protocols

MLD: for Signaling between Receiver and Router
PIM: for Signaling between Routers to build multicast distribution tree
Terminologies

• Source/Sender
  – The host/server sending multicast traffic

• Receiver/Listener
  – The host that receives multicast traffic

• Upstream
  – The direction traffic comes in

• Downstream
  – The direction traffic goes out

• RP (Rendezvous Point)
  – In PIM-SM, The router that “rendezvous” source and receiver info

• First Hop Router (FHR)
  – The router that source is connected

• Last Hop Router (LHR)
  – The router that receiver is connected

• (*,G)
  – *:any source, G:group address

• (S,G)
  – S:source address of multicast traffic, G:group address
**Terminologies**

- **Multicast Group Address**
  - Destination address for multicast traffic

- **Multicast Group**
  - The group consist of sender and receiver

- **MDT (Multicast Distribution Tree)**
  - The tree used to distribute multicast traffic on the routers
  - Shortest Path Tree(Source Tree), Shared Tree(RP tree)

- **Join**
  - Joining to the multicast group to receive the traffic

- **Leave(MLD)/Prune(PIM)**
  - Leave from the multicast group to stop receiving the traffic

- **ASM (Any Source Multicast)**
  - Multicast service that only specifies group address (does not specify source address)

- **SSM (Source Specific Multicast)**
  - Multicast service that specifies group and source address
MLD (Multicast Listener Discovery) v1/ v2
MLDv1 (RFC2710)

- Used for signaling between Multicast Listener (Receiver) and First Hop Router to inform/confirm the existence of multicast listener.
- Receiver sends “Report” to the router when join to the group.
- Receiver sends “Done” to the router when leave from the group.
- Router periodically sends MLD General Query to receiver side, and confirm the existence of listener.
- When the listener leave from the group, router sends Multicast Address Specific Query to confirm no any other listener is there.
- MLD is subset of ICMPv6 protocol
### MLDv1 Packet Format

#### Type Field:
- **130**: Multicast Listener Query
  - General Query
  - Multicast-Address-Specific Query
- **131**: Multicast Listener Report
- **132**: Multicast Listener Done

#### Multicast Address Field:
- Report: Target Multicast Address
- General Query: zero
- Multicast-Address-Specific Query: Target Multicast Address

#### Table Format:

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Response Delay</td>
<td>Reserved</td>
<td></td>
</tr>
</tbody>
</table>


Multicast Address
## Message Types

<table>
<thead>
<tr>
<th>Message Types</th>
<th>IPv6 destination Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Query</td>
<td>link-scope all-nodes (FF02::1)</td>
</tr>
<tr>
<td>Multicast-Address-Specific Query</td>
<td>Target multicast address</td>
</tr>
<tr>
<td>Report</td>
<td>Target multicast address</td>
</tr>
<tr>
<td>Done</td>
<td>link-scope all-routers (FF02::2)</td>
</tr>
</tbody>
</table>

## Default Timer/Variable

<table>
<thead>
<tr>
<th>Default Timer/Variable</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Query Interval</td>
<td>125 Sec</td>
</tr>
<tr>
<td>Maximum Response Delay</td>
<td>10000mSec</td>
</tr>
<tr>
<td>Multicast Listener Interval (expire timer on router side)</td>
<td>[Query Interval] * 2 + 10Sec = 260 Sec</td>
</tr>
<tr>
<td>Other querier Present interval</td>
<td>[Query Interval] * 2 + 5Sec = 255 Sec</td>
</tr>
<tr>
<td>Unsolicited Report Interval</td>
<td>10Sec</td>
</tr>
<tr>
<td>Last Listener Query Count(Robustness Variable)</td>
<td>2</td>
</tr>
<tr>
<td>Last Listener Query Interval</td>
<td>1Sec</td>
</tr>
</tbody>
</table>
MLDv1 Joining to the group

- Receiver sends MLD Report to Router for joining to the group.

IPv6 Src=Host’s Link-Local
IPv6 Dst=FF18::1:1
ICMPv6/MLD Report
Multicast-Address=FF18::1:1
MLD Querier election

- When MLD routers become online, all routers start sending Query.
- If the router receives the query from other router that has more smaller source address, the router stop sending query (become non-querier).
- The router that has the most smallest IPv6 address become MLD querier.

IPv6 Src=Router’s Link-Local
IPv6 Dst=FF02::1
ICMPv6/MLD Query
1. Host1 sends MLD-done to “all-routers”.
2. MLD Querier sends Multicsat-Address-Specific-Query to target Group Address.
3. If Host2 still joining to the group, Host2 must send MLD report within ”Last Listener Query Interval”.
1. Host2 sends MLD-done to “all-routers”

2. MLD Querier sends Multicast-Address-Specific-Query to target Group Address. (send 2*queries with Timeout=1sec)

3. After timeout, routers delete MLD entry.
MLDv2 (RFC 3810)

- MLDv2 enables host join/leave to Source and Group (for PIM-SSM)
- Adding Include/Exclude Source-List
- It has backward compatibility with MLDv1.
- All MLDv2 packets use same destination IPv6 address “FF02::16” in all types of messages.
### MLDv2 Report Packet Format

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type=143</th>
<th>Reserved</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>#ofMcast Address Records[M]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multicast Address Records [1]</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Multicast Address Records [2]</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>....</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Multicast Address Records [M]</th>
</tr>
</thead>
</table>
MLDv2 Multicast Address Record Format

<table>
<thead>
<tr>
<th>Record Type</th>
<th>Aux Data Len</th>
<th>Number of Sources (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multicast Address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source Address [1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source Address [N]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Record Type:
1. MODE_IS_INCLUDE - IS_IN ({S},G)
2. MODE_IS_EXCLUDE - IS_EX ({S},G)
3. CHANGE_TO_INCLUDE_MODE - TO_IN ({S},G)
4. CHANGE_TO_EXCLUDE_MODE - TO_EX ({S},G)
5. ALLOW_NEW_SOURCES - ALLOW ({S},G) [join (S,G)]
6. BLOCK_OLD_SOURCES - BLOCK ({S},G) [Leave (S,G)]
MLDv2 Join to (*,G)/(S,G)

IPv6 Src=Host’s Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=ALLOW_NEW_SOURCES
Multicast-Address=FF38::1:1
Source-Address=

IPv6 Src=Host’s Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=CHANGE_TO_EXCLUDE
Multicast-Address=FF18::1:1
Source-Address=2001:10::1
MLDv2 Leave from (*,G)/(S,G)

Leave from (*,G)
IPv6 Src=Host’s Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=CHANGE_TO_INCLUDE
Multicast-Address=FF18::1:1
Source-Address=

Leave from (S,G)
IPv6 Src=Host’s Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=BLOCK_OLD_SOURCES
Multicast-Address=FF38::1:1
Source-Address=2001:10::1
MLDv2 Reply to the Query

IPv6 Src=Host's Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=MODE_IS_EXCLUDE
Multicast-Address=FF18::1:1
Source-Address=

IPv6 Src=Host’s Link-Local
IPv6 Dst=FF02::16
ICMPv6/MLDv2 Report
Record Type=MODE_IS_INCLUDE
Multicast-Address=FF38::1:1
Source-Address=2001:10::1
Multicast Forwarding and PIM (Protocol Independent Multicast)
Multicast Distribution Tree (MDT)

Multicast Traffic is forwarded over the distribution tree that is built by PIM, from Upstream to Downstream.

- **Shortest Path Tree/Source Tree/(S,G) Tree**
  - The top of Shortest Path Tree is Source
  - (S,G) based forwarding
  - shortest path from Receiver to Source

- **Shared Tree / RP Tree / (*,G) Tree**
  - The top of Shared Tree is RP
  - (*,G) based forwarding
  - shortest path from Receiver to RP
Reverse Path Forwarding (RPF)

• **RPF interface (Incoming interface)**
  
  Each router selects the interface as upstream interface for Source/RP. Selection of the upstream interface is based on Unicast Routing information or routing information only for Multicast like BGP-mcast-addr-family/static-mroute.

• **Outgoing interface (List)**
  
  Downstream interface that received Join.OIF/OIL

• **RPF Neighbor**
  
  Next hop router(address) on RPF interface (upstream side) towards source/RP. Each (*,G)/(S,G) entry will select each RPF neighbor. PIM Join/Prune must be sent out to RPF Neighbor.

• **RPF Check**
  
  If the multicast packet for (*,G)/(S,G) is received on RPF interface, that packet is forwarded to OIFs.

  If the multicast packet is received on non-RPF interface, that packet is discarded.
RPF Check

(2001:10::10, FF38::1:1)
RPF interface = E0
RPF Neighbor = FE80::2
Outgoing Interface = E1
# PIM Header Packet Format

IP source address = Router Link-Local address
IP destination address = depends of message type
IP next header = 103 (0x67)

<table>
<thead>
<tr>
<th>Ver=2</th>
<th>Type</th>
<th>Reserved</th>
<th>Checksum</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Message Types</th>
<th>IPv6 destination Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = Hello</td>
<td>ALL-PIM-ROUTERS(FF02::D)</td>
</tr>
<tr>
<td>1 = Register</td>
<td>RP Address</td>
</tr>
<tr>
<td>2 = Register-Stop</td>
<td>First Hop Router(source of register)</td>
</tr>
<tr>
<td>3 = Join/Prune</td>
<td>ALL-PIM-ROUTERS(FF02::D)</td>
</tr>
<tr>
<td>4 = Bootstrap</td>
<td>ALL-PIM-ROUTERS(FF02::D)</td>
</tr>
<tr>
<td>5 = Assert</td>
<td>ALL-PIM-ROUTERS(FF02::D)</td>
</tr>
<tr>
<td>8 = Candidate-RP-Advertisement</td>
<td>BSR address</td>
</tr>
</tbody>
</table>
PIM Hello/PIM Neighbor

- PIM Routers periodically send PIM hello packet on the link, and each Router recognize other routers as PIM Neighbor.
- If RPF neighbor is recognized as PIM Neighbor, router can send PIM Join/Prune to RPF neighbor.
- When there are multiple PIM routers on same Subnet(Link), one PIM DR(Designated-Router) is selected on the link based on PIM-DR-Priority (if routers have same DR-priority, biggest address is selected as DR.)
The router that has most highest DR-priority is elected as DR on that link. (If DR-priority is same, most biggest address wins).

Router periodically sends PIM Hello to “FF02::D” (All-PIM-Routers)

Default Timer Value
Hello Interval = 30 Sec
Holdtime = 30 x 3.5 = 105 sec

When DR is Timeout, new DR is elected.

Only DR can do the PIM Join/Prune activity by receiving MLD Join/Leave. (Non-DR can not start any PIM action)
Requires RP (Rendezvous Point)

Source information is registered on RP (by first hop router). Join request is sent toward to RP by Hop-by-Hop. RP is managing (S,G) information in that multicast domain.

Effective for “one to many” or “many to many” communication.
2. First Hop Router sends PIM Register to RP in Unicast.
   Original multicast packet is encapsulated in PIM packet.
3. RP receives Register, and create (S,G) info on RP.
4. RP sends back PIM Register-Stop to First Hop Router.
   First Hop Router stop sending PIM Register.
5. Receiver sends MLD Report for the group (*,FF18::1::1).
6. Last Hop Router sends PIM (*,G) Join towards RP.
   #Shared-Tree(RP-Tree) is built.
7. From RP towards source, router sends PIM(S,G) Join.
8. First Hop Router receives (S,G) Join and traffic forwarding is started from Source to Receiver through RP.
9. By receiving source traffic, (S,G) entry is created on LHR. In above case, (*,G) RPF-IF = E0
   (S,G) RPF-IF = E1
10. LHR sends (S,G) Join towards RPF interface for Source.
11. Router sends (S,G) Join towards RPF interface for Source to build Shortest-Path-Tree.
12. Start traffic forwarding along Shortest-Path-Tree.
13. LHR start receiving traffic on Shortest-Path-Tree, LHR sends (S,G,rpt)prune towards RP because the traffic via RP is not necessary any more. 

(*,G)RPF-IF ≠ (S,G) RPF-IF


15. (S,G) prune is sent from RP towards source, and traffic from source to RP is stopped.
Finally traffic is forwarded only along shortest-path-tree. (SPT switchover)
After SPT is built, Shared-Tree is maintained but not used for traffic forwarding.
Defining RP

◆ Static-RP: Statically configuring RP address. All routers need configuration.

◆ BSR(Bootstrap Router): Candidate-RPs information is distributed to all routers, and RP is automatically elected from Candidate-RPs based on priority/hash.

◆ Embedded-RP
RP address is embedded in Multicast Address. mainly used for PIM-SM Inter-domain connection.
Candidate-BSR(C-BSR) floods Bootstrap Message(BSM) to all routers by Hop-by-Hop. C-BSR that has most highest BSR Priority is elected as BSR.
Candidate-RPs (C-RP) send C-RP-Advertisement to BSR in Unicast. (C-RP-Advertisement includes C-RP-Addr, Group-range, RP-Priority information.)
BSR floods all C-RPs information by BSM. All other routes receive this BSM, and elect RP.

BSR priority
255 > 0
RP priority
0 > 255
Embedded-RP(RFC3956)

- Based on “Unicast-Prefix based Multicast Address”.
- RP address is embedded in Multicast Address.
- Router can know RP address from Multicast Address.
- Mainly used for PIM-SM inter-domain multicast.

**Example:**

```
```

*Unicast prefix groupID*

Embedded RP Address = `2001:0123:c001:1100::1`
PIM SSM (RFC4601/3569/4607)

◆ No RP required
◆ Receiver join to (S,G)
◆ more simpler than PIM-SM
◆ subset of PIM-SM
2. Receiver send MLDv2 report for the Group/Source. (2001:1::10,FF38::1:1)
3. Last Hop Router sends PIM (S,G)Join towards source. Shortest-Path-Tree is built.
4. Multicast traffic is forwarded from Source to receiver along Shortest-Path-Tree.
MLD Snooping
Necessity of MLD Snooping

◆ On Non-MLD-aware L2-switch, all multicast traffic is flooded to all ports in the vlan.
◆ With MLD-snooping enabled, multicast traffic is forwarded only to the ports that receiver joined.
MLD Snooping: Join to the Group

The switch snoops MLD Packet, and create L2 forwarding table based on the request of MLD packet. After that, MLD packet is forwarded to the router.
MLD Snooping: Maintaining MLD-snooping Entry

MLD General Query from the router is flooded to all ports.
Joined hosts send back MLD report and each MLD snooping entry
is maintained. Only one MLD report is forwarded to the router.
MLD Snooping: Leave from the Group (1)

When one host leaves from the group, switch sends multicast-address-specific-query to that port, and after timeout of the query, switch deletes that port from MLD snooping entry. When still there is any other host joining to that group, switch does not send MLD done to the router.
MLD Snooping: Leave from the Group

When the last listener leaves the group, the switch sends multicast-Addr-Specific-Query to that port. After the query timeout, the switch deletes the MLD entry for that group and sends the MLD done to the router.
MLD Snooping: others

◆ Router Port identification:
  MLD Snooping does work correctly only when Router Port exists or there is MLD Querier on the vlan. Router Port can be identified automatically by receiving MLD Query/PIM Hello packet, or need static configuration.(depends on switch’s implementation.)

◆ The role of Router Port:
  Receiver side: MLD packet is forwarded to Router Port
  Sender side: Multicast Traffic is forwarded to Router Port

◆ MAC address duplication:
  If the switch identify the multicast group based on MAC address only, the switch may not be able to differentiate multiple groups that has same MAC address.(source address identification of MLDv2 snooping has same issue.)
MLD Snooping: immediately Stopping traffic

In IP/TV or video-surveillance environment, listener frequently Join/Leave to the group, router should shorten the time to stop forwarding traffic after receiving leave.

◆ MLD Fast-Leave:
When the router/switch receives the leave, immediately stop the traffic without sending multicast-address-specific query. In this case, it must be 1host/1port.

◆ MLD Host-tracking:
Router/switch is tracking all listener’s address that joining to the group, and last listener leave from the group on that port, immediately stop the traffic.
IPv6 multicast configuration and show command, Trouble shooting
Enable MLD router and confirm MLD router status

ALU# configure router mld interface "e-1/3/1"

ALU# show router mld status
===============================================================================
MLD Status
===============================================================================
Admin State                       : Up
Oper State                        : Up
Query Interval                    : 125
Last Listener Query Interval      : 1
Query Response Interval           : 10
Robust Count                      : 2
===============================================================================

ALU# show router mld interface
===============================================================================
MLD Interfaces
===============================================================================
<table>
<thead>
<tr>
<th>Interface</th>
<th>Adm</th>
<th>Oper</th>
<th>Cfg/Opr</th>
<th>Num</th>
<th>Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-1/3/1</td>
<td>Up</td>
<td>Up</td>
<td>2/2</td>
<td>0</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FE80::216:4DFF:FE56:25DB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interfaces</td>
<td></td>
<td></td>
<td>: 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
===============================================================================

Interfaces : 1
Enable PIMv6 on interfaces

ALU # configure router pim
ALU >config>router>pim# no ipv6-multicast-disable
ALU >config>router>pim# interface ge-1/1/1
ALU >config>router>pim>if# exit
ALU >config>router>pim# interface ge-1/1/2
ALU >config>router>pim>if# exit
ALU >config>router>pim# interface e-1/3/1
ALU >config>router>pim>if# exit
Confirm PIM interfaces status

```
ALU # show router pim interface ipv6

PIM Interfaces ipv6

<table>
<thead>
<tr>
<th>Interface</th>
<th>Adm</th>
<th>Opr</th>
<th>DR Prty</th>
<th>Hello Intvl</th>
<th>Mcast Send</th>
</tr>
</thead>
<tbody>
<tr>
<td>ge-1/1/1</td>
<td>Up</td>
<td>Up</td>
<td>1</td>
<td>30</td>
<td>auto</td>
</tr>
<tr>
<td>FE80::216:4DFF:FE56:25DB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ge-1/1/2</td>
<td>Up</td>
<td>Up</td>
<td>1</td>
<td>30</td>
<td>auto</td>
</tr>
<tr>
<td>FE80::216:4DFF:FE9B:A0FA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e-1/3/1</td>
<td>Up</td>
<td>Up</td>
<td>1</td>
<td>30</td>
<td>auto</td>
</tr>
<tr>
<td>FE80::216:4DFF:FE56:25DB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interfaces : 3
```

Transforming communications
for a world that's always on.
ALU # `show router pim neighbor ipv6`

---

PIM Neighbor ipv6
---

<table>
<thead>
<tr>
<th>Interface</th>
<th>Nbr Address</th>
<th>Nbr</th>
<th>DR</th>
<th>Prty</th>
<th>Up Time</th>
<th>Expiry Time</th>
<th>Hold Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>ge-1/1/1</td>
<td>105 FE80::216:4DFF:FE1D:D1B8</td>
<td>1</td>
<td></td>
<td></td>
<td>0d 06:07:32</td>
<td>0d 00:01:43</td>
<td>105</td>
</tr>
<tr>
<td>ge-1/2/1</td>
<td>105 FE80::216:4DFF:FE9B:A0FA</td>
<td>1</td>
<td></td>
<td></td>
<td>0d 05:54:21</td>
<td>0d 00:01:24</td>
<td>105</td>
</tr>
</tbody>
</table>

Neighbors : 2
---
Configuring static RP for PIM-ASM

```
ALU # configure router pim
ALU >config>router>pim# rp ipv6
ALU >config>router>pim>rp>ipv6# static address 2001::2
<Configure RP address as “2001::2” statically>
ALU >config>router>pim>rp>ipv6>static>address# override
<prioritize static-RP configuration over BSR RP info>
ALU >config>router>pim>rp>ipv6>static>address# group-prefix FF18::/16
<configure group-range for RP=2001::2>
```
ALU # `show router pim rp ipv6`

```
PIM RP Set ipv6

<table>
<thead>
<tr>
<th>Group Address</th>
<th>RP Address</th>
<th>Type</th>
<th>Prio</th>
<th>Hold</th>
<th>Expiry</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF18::/16</td>
<td>2001::2</td>
<td>Static</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
```

Group Prefixes : 1
Configuring PIM-SSM

ALU # configure router pim
ALU >config>router>pim# ssm-groups
ALU >config>router>pim>ssm# group-range FF38::/16
<configure FF38::/16 as group-address-range for PIM-SSM>
Configuring MLD-Snooping

ALU# configure service vpls 3
ALU >config>service>vpls# mld-snooping no shutdown
<Enable MLD-snooping on VPLS 3(VLAN3)>

Transforming communications for a world that's always on.
Confirm MLD Snooping status

ALU # `show service id 10 mld-snooping base`

```
MLD Snooping Base info for service 10

Admin State : Up
Querier : FE80::216:4DFF:FE56:25DB on SAP 1/8/2

<table>
<thead>
<tr>
<th>Sap/Sdp Id</th>
<th>Oper State</th>
<th>MRtr</th>
<th>Send Queries</th>
<th>Max Num Groups</th>
<th>MVR From-VPLS</th>
<th>Num Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>sap:1/8/2</td>
<td>Up</td>
<td>Yes</td>
<td>Disabled</td>
<td>No Limit</td>
<td>Local</td>
<td>0</td>
</tr>
<tr>
<td>sap:1/8/3</td>
<td>Up</td>
<td>No</td>
<td>Disabled</td>
<td>No Limit</td>
<td>Local</td>
<td>0</td>
</tr>
</tbody>
</table>
```
Confirm MLD Entry (MLDv1)

MLD Entryの確認

ALU # show router mld group

MLD Groups

(*,FF18::1:1)
  Up Time : 0d 00:00:14
  Fwd List : e-1/8/1

(*,G)/(S,G) Entries : 1
## Confirm PIM group Entry (PIM-SM)

```
ALU # show router pim group ipv6
```

<table>
<thead>
<tr>
<th>Group Address</th>
<th>Source Address</th>
<th>Type</th>
<th>Spt Bit</th>
<th>Inc Intf</th>
<th>No.Oifs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF18::1:1</td>
<td>*</td>
<td>(*,G)</td>
<td>ge-1/1/1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FF18::1:1</td>
<td>2001::2</td>
<td>spt</td>
<td>ge-1/2/1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2001:1:10::10</td>
<td>2001::2</td>
<td>spt</td>
<td>ge-1/2/1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Groups : 2
Confirm RPF

ALU # `show router route-table 2001::2`

IPv6 Route Table (Router: Base)

<table>
<thead>
<tr>
<th>Dest Prefix</th>
<th>Type</th>
<th>Proto</th>
<th>Age</th>
<th>Pref</th>
<th>Next Hop[Interface Name]</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001::2/128</td>
<td>Remote</td>
<td>OSPF3</td>
<td>01d00h39m</td>
<td>10</td>
<td>FE80::216:4DFF:FE1D:D1B8-&quot;ge-1/1/1&quot;</td>
<td>101</td>
</tr>
</tbody>
</table>

No. of Routes: 1

ALU# `show router route-table 2001:1:10::10`

IPv6 Route Table (Router: Base)

<table>
<thead>
<tr>
<th>Dest Prefix</th>
<th>Type</th>
<th>Proto</th>
<th>Age</th>
<th>Pref</th>
<th>Next Hop[Interface Name]</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001:1:10::/64</td>
<td>Remote</td>
<td>OSPF3</td>
<td>01d00h37m</td>
<td>10</td>
<td>FE80::216:4DFF:FE9B:A0FA-&quot;ge-1/2/1&quot;</td>
<td>201</td>
</tr>
</tbody>
</table>

No. of Routes: 1
ALU #
[ifIndex 4] V2 PDU: :: -> FF02::16 pduLen 28
  Type: V2 REPORT maxrespCode 0x0 checkSum 0x6e74
  Num Group Records: 1
  Group Record 0
    Type: CHG_TO_EXCL, AuxDataLen 0, Num Sources 0
    Mcast Addr: FF18::1:1
    Source Address List

51 2009/11/07 16:24:45.49 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSendJoinPrunePdu
pimSendJoinPrunePdu: if 2, adj FE80::216:4DFF:FE1D:D1B8"

52 2009/11/07 16:24:45.49 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: encoding groupset for group FF18::1:1, numJoinedSrcs 1, numPrunedSrcs 0"

53 2009/11/07 16:24:45.49 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: Encoding Join for (*,G) RP = 2001::2"
Debug example at Last Hop Router (PIM-SM)

Send (S,G) Join to the RPF Neighbor for source

57 2009/11/07 16:24:45.53 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSendJoinPrunePdu
pimSendJoinPrunePdu: if 3, adj FE80::216:4DFF:FE9B:A0FA"

58 2009/11/07 16:24:45.53 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: encoding groupset for group FF18::1:1, numJoinedSrcs 1, numPrunedSrcs 0"

59 2009/11/07 16:24:45.54 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: Encoding Join for source 2001:1:10::10"

60 2009/11/07 16:24:45.54 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base]
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: num joined srcs 1, num pruned srcs 0"
Debug example at Last Hop Router (PIM-SM)

Send (S,G)RPT-Prune to the RPF neighbor for RP

64 2009/11/07 16:24:45.55 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base] 
"PIM[Instance 1 Base]: pimSendJoinPrunePdu
pimSendJoinPrunePdu: if 2, adj FE80::216:4DFF:FE1D:D1B8"

65 2009/11/07 16:24:45.55 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base] 
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: encoding groupset for group FF18::1:1, numJoinedSrcs 0, numPrunedSrcs 1"

66 2009/11/07 16:24:45.55 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base] 
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: Encoding RPT Prune for source 2001:1:10::10"

67 2009/11/07 16:24:45.55 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base] 
"PIM[Instance 1 Base]: pimSGEncodeGroupSet
pimEncodeGroupSet: num joined srcs 0, num pruned srcs 1"

68 2009/11/07 16:24:45.56 JST MINOR: DEBUG #2001 Base PIM[Instance 1 Base] 
"PIM[Instance 1 Base]: pimSendJoinPrunePdu
pimSendJoinPrunePdu2: sending JP PDU with 1 groups."
**Confirm MLD entry (MLDv2)**

```
ALU# show router mld group ff38::2:1

MLD Groups

(2001:1:10::10,FF38::2:1)
  Up Time : 0d 00:00:21
  Fwd List : e-1/8/1

(*,G)/(S,G) Entries : 1
```
**Confirm PIM group entry (PIM-SSM)**

```bash
ALU # show router pim group ff38::2:1

<table>
<thead>
<tr>
<th>Group Address</th>
<th>Type</th>
<th>Spt Bit</th>
<th>Inc</th>
<th>Intf</th>
<th>No.Oifs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF38::2:1</td>
<td>(S,G)</td>
<td>ge-1/2/1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2001:1:10::10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Groups : 1
```
**Useful command**

```
ALU # show router pim group ff38::2:1 detail
<snip>
Group Address      : FF38::2:1
Source Address     : 2001:1:10::10
<snip>
Curr Fwding Rate   : 6920.7 kbps
Forwarded Packets  : 71223              Discarded Packets : 0
Forwarded Octets   : 98002848           RPF Mismatches     : 0
Spt threshold      : 0 kbps             ECMP opt threshold : 7
Admin bandwidth    : 1 kbps

Groups : 1
```

---
### Confirm MLD snooping entry

**ALU # show service id 3 mld-snooping proxy-db**

<table>
<thead>
<tr>
<th>Group Address</th>
<th>Mode</th>
<th>Up Time</th>
<th>Num Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF18::1:1</td>
<td>exclude</td>
<td>0d 00:11:40</td>
<td>0</td>
</tr>
<tr>
<td>FF38::2:1</td>
<td>include</td>
<td>0d 00:11:40</td>
<td>1</td>
</tr>
</tbody>
</table>

Number of groups: 2

**ALU # show service id 3 mld-snooping port-db sap 1/8/3**

<table>
<thead>
<tr>
<th>Group Address</th>
<th>Mode</th>
<th>Type</th>
<th>From-VPLS</th>
<th>Up Time</th>
<th>Expires</th>
<th>Num MC Src</th>
<th>MC Stdby</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF18::1:1</td>
<td>exclude</td>
<td>dynamic local</td>
<td></td>
<td>0d 00:11:41</td>
<td>143s</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>FF38::2:1</td>
<td>include</td>
<td>dynamic local</td>
<td></td>
<td>0d 00:11:41</td>
<td>0s</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Number of groups: 2
Useful command: static MLD join

Configure joined status statically on the interface

```
ALU # configure router mld
ALU>config>router>mld# interface e-1/3/1
ALU>config>router>mld>if# static group ff18::1:1 starg
<static join for (*,G)>
ALU>config>router>mld>if# static group ff38::2:1 source 2001:1:10::10
<static join for (S,G)>
```

By configuring static join, Router emulates MLD joined status statically.
Useful command: SSM Mapping

Configure the mapping information for group and source. With this configuration PIM-SSM can be used even if there is only MLDv1 client.

#need to configure mapping information on all edge routers.

Example command:
```
ALU # configure router mld
ALU >config>router>mld# ssm-translate grp-range ff38::2:1 ff38::2:10 source 2001:1:10::10
```

MDLv1 report
(*, FF38::2:1)

PIM (S,G) Join
(2001:1:10::10, FF38::2:1)

FF38::2:1-10 group range is mapped to source 2001:1:10::10
Useful command: static route only for multicast

Static route configuration used only for RPF in multicast.
It does not affect unicast routing information.
Useful when change RPF neighbor without unicast routing change.

MLD report
Join(*,FF18::1:1)

Best unicast route for source
RPF for source

static-mroute(mcast-ipv6 route) for 2001:1:10::10/128.

ALU # configure router static-route 2001:1:10::10/128 next-hop 2001:2:10::2 mcast-ipv6
Trouble shooting for multicast

First, check the source is sending multicast traffic correctly.
- Hop-limit(TTL) is not “1”, have enough value?
- Source/Group Address is correct?

Check the status from Receiver side towards Source/RP
- LHR has MLD Entry correctly?
- LHR recognizes PIM Neighbor correctly?
- LHR has PIM Entry(mroute) correctly?
- LHR’s RPF Neighbor/Interface is expected one?
- LHR is sending Join/Prune to RPF Neighbor?
  ↓
- Upstream router has PIM Entry correctly?
  - Upstream router’s RPF Neighbor/Interface is expected one?
  - Upstream router is sending PIM Join/Prune to RPF Neighbor?
    ↓
  Repeat those check towards Source
Debug commands

ALU # debug router pim jp
ALU # debug router pim packet
ALU # debug router pim db
ALU # debug router mld packet
Failure Case Study
User expects that Router2 as Primary Multicast Forwarder on that segment, but Router1 has become DR unexpectedly and Router1 become forwarder.

=> Need to configure DR-Priority or Link-Local Address
In above topology and uplink on Router1 become down, Router1 need to know alternative route via Router2. (generally LAN segment is configured as “passive”) In some PIM implementation, above failover scenario does not work.
PIM-Neighbor

When the router finds RPF Neighbor based on unicast routing info, Router cannot send PIM-Join/Prune to the RPF neighbor if that RPF Neighbor is not recognized as PIM-Neighbor.

=> PIM should be enabled on all links in the network

Router can not send PIM Join/Prune because RPF-Neighbor is not PIM-Neighbor.
For downstream router, VRRP virtual address is RPF Neighbor, but it is not PIM-Neighbor. So router cannot send any PIM messages.
L2 Switch only environment

There is no MLD querier in the vlan, MLD-snooping does not work without router port.
=> Need to connect MLD Router, or need to configure MLD querier
Considerations in deploying IPv6 multicast to SP network
Considerations in deploying IPv6 multicast(1)

1. How many groups (multicast streams)?
2. Each stream’s Bandwidth?
3. How much bandwidth available in access network?
Considerations in deploying IPv6 multicast (2)

4. How many subscribers on one L2-edge-device?
5. Max join rate (# of MLD joins/sec) L2-edge-device receiving?
6. How many subscribers on one L3-edge-device?
7. Max join rate (# of MLD joins/sec) L3-edge-device receiving?
8. Max # of replication (Fan out) on L2/L3-device?

(VLAN per user / VLAN is shared by users / multicast over PPP?)
Considerations in deploying IPv6 multicast (3)

9. What is the acceptable Join delay?
Considerations in deploying IPv6 multicast (3)

9. What is the required failover time?

1. Network down detected
2. Unicast Routing convergence
3. Multicast routing detects unicast change, RPF Neighbor is changed
4. Send PIM join to new RPF neighbor
5. PIM Join towards to source
6. Traffic forwarded on new Tree
Hierarchical deployment limits impact of packet loss and reduces retransmission traffic

Linear per-channel scaling yields superior performance with far lower traffic overhead
Distributed & integrated fast channel change support

FCC service for all linear TV channels

Channel change request

FCC unicast burst

Video pictures of new channel delivered via unicast burst starting with an I-frame

FCC UC burst

Hierarchical deployment optimizes efficiency for both popular and less popular content

Integrated FCC functions yield superior performance with far lower traffic overhead

5910 RET/FCC STB SW client

Optionally the FCC service can be enabled here as well

Regional video office

No more FCC traffic!

After the STB buffer is filled by the FCC unicast burst, it joins the regular multicast stream

MC stream