"Tech Session"

IPv6 Multicast Primer

Tim Martin
CCIE #2020
Solutions Architect
Spring 2014
Agenda

- IPv6 Multicast Addressing
- Multicast Listener Discovery (MLD)
- IPv6 Zeroconf
- IPv6 Multicast Impact on 802.11
- Final Thoughts
Why IPv6 and Multicast are important?

- Healthcare
- Energy Efficiency
- Predictive maintenance
- Car 2 Car
- Research & Discovery
- Smart Cities
- Industrial Automation
- Smart Grid
- Smart Home
- Agriculture
- Defense
- Asset tracking
- Intelligent Building

© 2012 Cisco and/or its affiliates. All rights reserved.
IPv6 Addressing

*IPv6 does not use broadcast addressing
## IPv6 over Ethernet

<table>
<thead>
<tr>
<th>Destination Ethernet Address</th>
<th>Source Ethernet Address</th>
<th>0x0800</th>
<th>IPv4 Header and Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>Destination Ethernet Address</td>
<td>Source Ethernet Address</td>
<td>0x86DD</td>
<td>IPv6 Header and Payload</td>
</tr>
</tbody>
</table>

- IPv6 has a specific Ethernet Protocol ID
- IPv6 relies heavily on Multicast

I bit = Local Admin, L bit = Multicast/Broadcast
Multicast Mapping over Ethernet (RFC 2464)

IPv6 Temporary Multicast Address

**FF3E:0040:2001:0DB8:CAFE:0001:11D7:4CD3**

Corresponding Ethernet Address

**33:33:**

IPv6 Well Known Multicast Address

**FF02:0000:0000:0000:0000:0000:0000:0000:0000:0000:0000:0000:0000:0001**

Corresponding Ethernet Address

**33:33:**
IPv6 Multicast Addressing
IPv6 Multicast Address (RFC 4291) (T bit)

- Prefix **FF00::/8**

<table>
<thead>
<tr>
<th>8-bit</th>
<th>4-bit</th>
<th>4-bit</th>
<th>112-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111 1111</td>
<td>0 R P T</td>
<td>Scope</td>
<td>Variable format</td>
</tr>
</tbody>
</table>

**Flags**

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>Reserved</td>
</tr>
<tr>
<td>R = 0</td>
<td>No embedded RP</td>
</tr>
<tr>
<td>R = 1</td>
<td>Embedded RP</td>
</tr>
<tr>
<td>P = 0</td>
<td>Without Prefix</td>
</tr>
<tr>
<td>P = 1</td>
<td>Address based on Prefix</td>
</tr>
<tr>
<td>T = 0</td>
<td>Well Known Address (IANA assigned)</td>
</tr>
<tr>
<td>T = 1</td>
<td>Temporary address (local assigned)</td>
</tr>
</tbody>
</table>

**Scope**

<table>
<thead>
<tr>
<th>Scope</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Node</td>
</tr>
<tr>
<td>2</td>
<td>Link</td>
</tr>
<tr>
<td>3</td>
<td>Subnet</td>
</tr>
<tr>
<td>4</td>
<td>Admin</td>
</tr>
<tr>
<td>5</td>
<td>Site</td>
</tr>
<tr>
<td>8</td>
<td>Organization</td>
</tr>
<tr>
<td>E</td>
<td>Global</td>
</tr>
</tbody>
</table>
IPv6 Multicast Address – Unicast Based (RFC 3306) (P bit)

- Every Unicast prefix can build custom multicast addresses
- Last **32 bits** of unicast address mapped into Group ID (112 Bits)

<table>
<thead>
<tr>
<th>8 Bits</th>
<th>4 Bits</th>
<th>4 Bits</th>
<th>8 Bits</th>
<th>8 Bits</th>
<th>64 Bits</th>
<th>32 Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111 1111</td>
<td>0 0 1 1</td>
<td>1110</td>
<td>Rsvd</td>
<td>plen</td>
<td>Unicast Prefix</td>
<td>Group ID</td>
</tr>
</tbody>
</table>

**Example**
- `plen`: `40 = 64 bits`
- `Prefix`: `2001:db8:cafe:1::`
- `Group ID`: `11d7:4cd3`

**IPv6 Embedded RP Multicast Address (RFC 3956) (R bit)**

- Static mapping of RP into Multicast group
- Solves MSDP and scaling issues

<table>
<thead>
<tr>
<th>8 Bits</th>
<th>4 Bits</th>
<th>4 Bits</th>
<th>4 Bits</th>
<th>4 Bits</th>
<th>8 Bits</th>
<th>64 Bits</th>
<th>32 Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111 1111</td>
<td>0 1 1 1</td>
<td>1110</td>
<td>Rsrd</td>
<td>RPid</td>
<td>plen</td>
<td>Unicast Prefix</td>
<td>Group ID</td>
</tr>
</tbody>
</table>

**Example**

- Rsrcd/RPd: 0000 | 0101
- Prefix: 2001:db8:cafe:1::
- Group ID: 645

**Embedded RP Group**

\[
\text{Embedded RP Group} = \text{FF7E:} \text{540:} \text{2001:db8:cafe:1:} : \text{645}
\]

**RP address**

\[
\text{RP address} = \text{2001:db8:cafe:1:} : \text{0000:0645}
\]
# Well Known Multicast Addresses

<table>
<thead>
<tr>
<th>Address</th>
<th>Scope</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF01::1</td>
<td>Node-Local</td>
<td>This Node</td>
</tr>
<tr>
<td>FF05::2</td>
<td>Site-Local</td>
<td>All Routers</td>
</tr>
<tr>
<td>FF02::1</td>
<td>Link-Local</td>
<td>All Nodes</td>
</tr>
<tr>
<td>FF02::2</td>
<td>Link-Local</td>
<td>All Routers</td>
</tr>
<tr>
<td>FF02::5</td>
<td>Link-Local</td>
<td>OSPFv3 Routers</td>
</tr>
<tr>
<td>FF02::6</td>
<td>Link-Local</td>
<td>OSPFv3 DR Routers</td>
</tr>
<tr>
<td>FF02::9</td>
<td>Link-Local</td>
<td>RIPng</td>
</tr>
</tbody>
</table>

- FF02, is a permanent address and has link scope
- Link Operations, Routing Protocols, Streaming Services
Solicited Node Multicast Address
### Duplicate Address Detection (DAD)

Node A can start using address A

<table>
<thead>
<tr>
<th>ICMP Type</th>
<th>135 NS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IPv6 Source</strong></td>
<td>UNSPEC = ::</td>
</tr>
<tr>
<td><strong>IPv6 Dest.</strong></td>
<td>A Solicited Node Multicast FF02::1:FF00:A</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>FE80::A</td>
</tr>
<tr>
<td><strong>Query</strong></td>
<td>Anyone using A?</td>
</tr>
</tbody>
</table>

- ICMPv6 runs on top of IPv6, etype = 86DD, Layer 3.14 :’)
- Probe neighbors to verify address uniqueness
Solicited-Node Multicast Address

• For each Unicast and Anycast address configured there is a corresponding solicited-node multicast

• Solicited-node multicast consists of
  \textbf{FF02::1:FF}/104 \{lower 24 bits from IPv6 Unicast interface ID\}

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FE80</td>
<td>0000</td>
<td>0000</td>
<td>0000</td>
<td>1234</td>
<td>5678</td>
<td>9ABC</td>
<td>FC0F</td>
</tr>
<tr>
<td>FF02</td>
<td>0000</td>
<td>0000</td>
<td>0000</td>
<td>0000</td>
<td>0001</td>
<td>FFBC</td>
<td>FC0F</td>
</tr>
</tbody>
</table>

Ethernet Multicast
Uses last 32 bits
IPv6 Interface Example

R1#sh ipv6 int e0
Ethernet0 is up, line protocol is up
   IPv6 is enabled, link-local address is FE80::200:CFF:FE3A:8B18
   Global unicast address(es):
       2001:DB8:0:1234::1 subnet is 2001:DB8:0:1234::/64
   Joined group address(es):
       FF02::1
       FF02::2
       FF02::1:FF00:1
       FF02::1:FF3A:8B18
   MTU is 1500 bytes
   ICMP error messages limited to one every 100 milliseconds
   ICMP redirects are enabled
   ND DAD is enabled, number of DAD attempts: 1
   ND reachable time is 30000 milliseconds
   ND advertised reachable time is 0 milliseconds
   ND advertised retransmit interval is 0 milliseconds
   ND router advertisements are sent every 200 seconds

*If EUI format is used then the 1rst solicited node mcast addr is used for both the LL & GU

Solicited-Node Multicast Address*
Neighbor Solicitation & Advertisement

**Table:**

<table>
<thead>
<tr>
<th>ICMP Type</th>
<th>135 NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6 Source</td>
<td>FE80::A</td>
</tr>
<tr>
<td>IPv6 Destination</td>
<td>B Solicited Node Multicast FF02::1:FF00:B</td>
</tr>
<tr>
<td>Target Address</td>
<td>2001:db8:1:46::B</td>
</tr>
<tr>
<td>Code</td>
<td>0 (need link layer)</td>
</tr>
<tr>
<td>Query</td>
<td>What is B link layer address?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ICMP Type</th>
<th>136 NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6 Source</td>
<td>FE80::B</td>
</tr>
<tr>
<td>IPv6 Destination</td>
<td>FE80::A</td>
</tr>
<tr>
<td>Target</td>
<td>Type 2</td>
</tr>
<tr>
<td>Data</td>
<td>Link Layer address of B</td>
</tr>
<tr>
<td>*Flags</td>
<td>R = Router S = Response to Solicitation O = Override cache information</td>
</tr>
</tbody>
</table>

- ARP replacement, Map’s L3 to L2.
- Node B will add node A to it’s neighbor cache during this process w/o sending NS
- Multicast for resolution (new), Unicast for reachability (cache)
Multicast Listener Discovery (MLD)
IPv6 Multicast Listener Discovery (MLD)

- MLD uses LL source addresses
- 3 msg types: Query, Report, Done
- MLD packets use “Router Alert” in HBH
- MLDv1 = (*,G) shared, MLDv2 = (S,G) source

<table>
<thead>
<tr>
<th>MLD</th>
<th>IGMP</th>
<th>Message Type</th>
<th>ICMPv6 Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLDv1</td>
<td>IGMPv2</td>
<td>Listener Query</td>
<td>130</td>
<td>Used to find out if there are any multicast listeners</td>
</tr>
<tr>
<td></td>
<td>(RFC 2710)</td>
<td>Listener Report</td>
<td>131</td>
<td>Response to a query, joins a group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Listener Done</td>
<td>132</td>
<td>Sent by node to report it has stopped listening</td>
</tr>
<tr>
<td>MLDv2</td>
<td>IGMPv3</td>
<td>Listener Query</td>
<td>130</td>
<td>Used to find out if there are any multicast listeners</td>
</tr>
<tr>
<td>(RFC 3810)</td>
<td>(RFC 3376)</td>
<td>Listener Report</td>
<td>143</td>
<td>Enhanced reporting, multiple groups and sources</td>
</tr>
</tbody>
</table>

MLD snooping
Hosts send MLD report to alert router they wish to join a multicast group.

Router then joins the tree to the source or RP.
MLDv1 Example Leaving a Group (Query)

I wish to leave ff38::276

MLD Done (A)

fe80::209:5bff:fe08:a674

Hop-by-Hop Header

IPv6 Source fe80::209:5bff:fe08:a674

Hop Limit 1

Group Address ff38::276

IPv6 Destination FF02::2 (All routers)

ICMP Type 132

MLD Query (C)

Hop-by-Hop Header

IPv6 Destination FF38::276

Hop Limit 1

IPv6 Source fe80::207:85ff:fe80:692

ICMP Type 130

Group Address ff38::276

Hop-by-Hop Header

Router Alert Yes

MLD Report (B)

fe80::250:8bff:fE55:78de

Hop-by-Hop Header

IPv6 Destination FF38::276

Hop Limit 1

IPv6 Source fe80::250:8bff:fE55:78de

ICMP Type 131

Group Address ff38::276

Router Alert Yes

I am watching ff38::276
MLDv2 Joining a Group (Report)

I wish to receive FF38:4000:BA11

MLD Report (A)

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICMP Type</td>
<td>143</td>
</tr>
<tr>
<td>IPv6 Source</td>
<td>fe80::209:5bff:fe08:a674</td>
</tr>
<tr>
<td>IPv6 Destination</td>
<td>FF02::16</td>
</tr>
<tr>
<td>Hop Limit</td>
<td>1</td>
</tr>
<tr>
<td># of Records</td>
<td>Include/exclude</td>
</tr>
<tr>
<td>Group Address</td>
<td>FF38::4000:BA11</td>
</tr>
<tr>
<td>Hop-by-Hop Header</td>
<td></td>
</tr>
<tr>
<td>Router Alert</td>
<td>Yes</td>
</tr>
</tbody>
</table>

fe80::209:5bff:fe08:a674

MLD Report

Source for multicast FF38::4000:BA11
MLD Multicast Maintenance (Query)

- General Query
  FF02::1
  Group list empty, who’s listening?

- Group Specific Query
  FF38::4000:BA11
  Anyone still interested in this stream?

- Group & Source Specific Query
  2001:DB8:CAFÉ::1, FF38::4000:BA11

- Filter Mode, Change Record

- Multiple routers on link
  Lowest address value assumes Querier role
Zero Conf & Service Discovery
What is this about?

- BYOD: Massive influx of consumer devices to be placed on Enterprise networks
- Consumer devices are typically located within a single Layer 2 domain in the home
- Users may expect to have the same type of services in the Enterprise / Campus but also across L3 boundaries
- Device types include mobile devices (iOS, Android), printers, cameras, PCs etc.
Zeroconf over IPv6

- Apple (Bonjour) has a light weight approach, adopted quicker
- FF02::FB – Multicast DNS – mDNS
- Microsoft (Rally) has a more robust, heavier implementation, has moved slower
- FF02::C – Simple Service Discovery Protocol – SSDP, UPnP
- FF02::1:3 – Link Local Multicast Name Resolution – LLMNR (File Sharing enabled)

Personal Computer Operating Systems
- Windows
- Mac OS X
- Linux

Appliances & Networking
- Printers
- Access Points
- Switches
- Routers

AV Equipment
- Speakers
- Cameras
- Displays
- AV Receivers
Service Discovery vs. Access Control

• **Service Discovery**
  
  Is your Phone Book. Tell me, where I can reach Mr. Printer
  
  Doesn’t necessarily mean that you can actually reach / talk to Mr. Printer

• **Access Control**
  
  Is like caller screening
  
  Even if a person is not listed in the phone book, you might call that person because you know the number
  
  “I know Mr. Printer is at 1.2.3.4, let’s call him even if I don’t see him in the phone book”

• **Better Together**
  
  use the phone book for easy lookup (Service Discovery)
  
  use the caller screening for security (ACL / SGT / SGACL …)
IPv6 Multicast over 802.11 Networks
Problem areas

- All Multicast messages require ALL hosts on link to stay awake (1-10ms)
  Even if the message is not for them
- Multicast [on WiFi, battery/sleepy nodes)
  Wasting bandwidth
  Waking up hosts unnecessarily
  Looking at ND (RS, RA, DAD, address resolution)
- Duplicate address detection
  Currently requires always-on to defend address
- Related DAD issues
  Not robust against packet loss
Wi-Fi Multicast Background

- Radio is a shared media.
- Only unicast frames are **acknowledged and retransmitted**
- 10% packet loss appears to be common, both Unicast & Multicast
- AP transmits bcast/mcast frames at the **lowest possible rate** to ensure reception
- Broadcast/Multicast up to 10x more time in air

IEEE 802.11a mcast: 6 Mbps, ucast up to 54 Mbps

IEEE 802.11n mcast: 15 Mbps, ucast up to 150 Mbps
IPv6 ND Chatter

- **DAD**
  - 1 packet per IP address configured on the network
- **RS**
  - 1 packet per host that joins the network
- **RA**
  - Periodic: 1 packet every X seconds
  - Solicited: 1 packet for every host that joins the network
- **NS**
  - 1 packet for every new host/host pair
- **MLD**
  - 1 packet for solicited node multicast group

- What about broken stuff (i.e. MLD packets sent to Solicited Node Mcast)
Not all implementations are the same…

RFC4861, 6.2.6. Processing Router Solicitations

In addition to sending periodic, unsolicited advertisements, a router sends advertisements in response to valid solicitations received on an advertising interface. A router **MAY choose to unicast** the response directly to the soliciting host's address (if the solicitation's source address is not the unspecified address), but the **usual case is to multicast the response to the all-nodes group.**
RA Throttle & ND Multicast Suppression

- Scaling the 802.11 multicast reliability issues
- NDP process is multicast “chatty”, consumes airtime
- Controller rate limits the period RA’s, while allowing RS to flow
- Caching allows the Controller to “proxy” the NA, based on gleaning
IPv6 “Off-link” clients

- By itself doesn’t fix router->host multicast NS
- Need to disable redirects on first-hop router

- ICMPv6 Option (Prefix information: 2001:470:73cd:df03::/64)
  - Type: Prefix information (3)
  - Length: 4 (32 bytes)
  - Prefix Length: 64
  - Flag: 0x40
  0... .... = On-link flag(L): Not set
  .1. .... = Autonomous address-configuration flag(A): Set
  ..0. .... = Router address flag(R): Not set
  ...0 0000 = Reserved: 0
  Valid Lifetime: 86400
  Preferred Lifetime: 1800
  Reserved
Final Thoughts
Malformed Packets

- Intel 1217-LM NIC’s
  - Power management settings
- Dell, HP, Toshiba. Maybe others
- MLD Query packets
  - From the host appear to be cause of the storm
- Some were disabling IPv6, reconfiguring BIOS
  - Download 19.0 drivers
- More issues will appear
  - Know how protocol works

destmac 33.33.00.00.00.01, srcmac C8.1F.66.A8.72.7A, protocol 86DD protocol ipv6: version 6, flow 1610612736, payload 32, nexthdr 0, hoplt 1, class 0, src FE80::CA1F:66FF:FEA8:727A, dst FF02::1
Key Take Away

• Applications We Haven't Even Built Yet
• Large Privately Owned Multicast Address Space
• No NAT required
• Multicast is Foundational in IPv6
• Wake and power issues need resolution
• Invest in your future - IPv6, the future is now