"Tech Session"

IPv6 Multicast Primer

Tim Martin

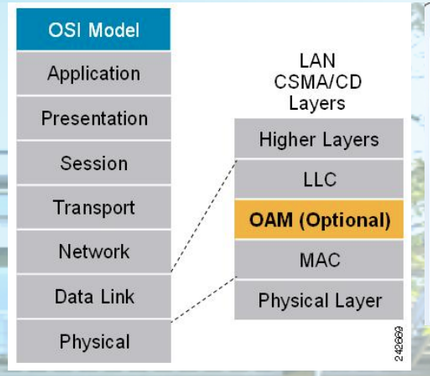
CCIE #2020

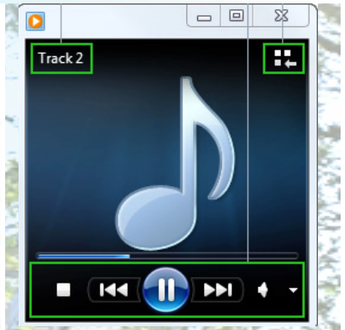
Solutions Architect

Summer 2013

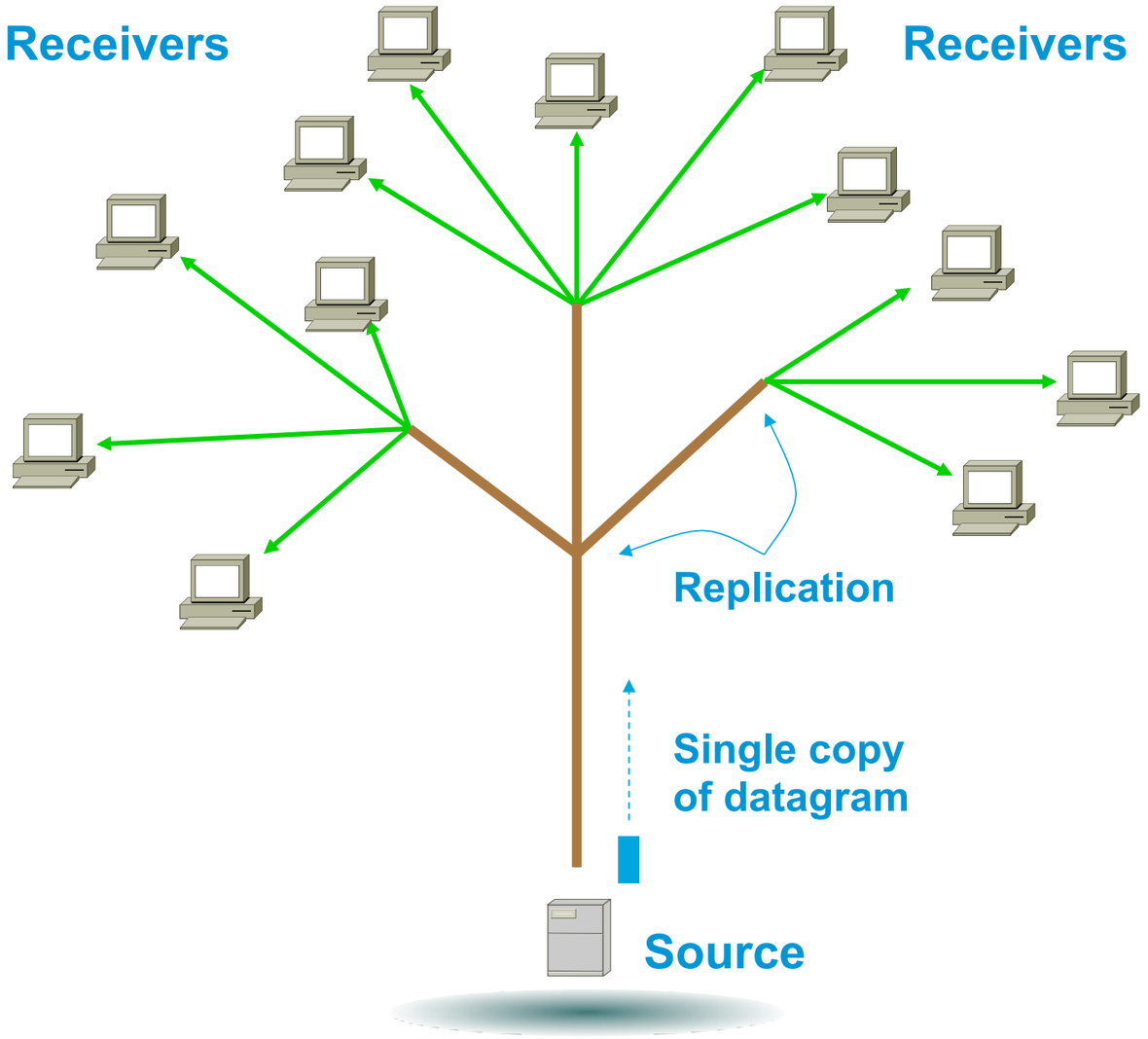
Why Multicast?

- Link Operations
- Routing Protocols
- Distance Learning
- Surveillance
- Metering
- Broadcast Video Services
- Efficient Delivery





It's All About Trees



- Mechanism for transmitting information from a single source (root) to many receivers (leaves)
- Single copy of a datagram is sent from the source and replicated through the tree to receivers
- No restriction on physical or geographical boundary

IPv6 Multicast Trees and Protocols

- Multicast is a normal **IPv6 packet Destination**
- An IPv6 multicast group address always starts with the prefix **FF00::/8 (1111 1111)**
- Multicast Listener Discovery **(MLD)**
- Multicast traffic is forwarded along a multicast tree which can be either a
  - **Source Tree (S, G)**
  - **Shared Tree (*, G)**
- IPv6 supports Protocol Independent Multicast **(PIM)** routing protocols only
  - PIM creates the trees that multicast streams are forwarded on
  - PIM operation is the same in IPv6 as IPv4 (RFC 4601 specifies operation over IPv4 and IPv6)
  - PIM identified by the **IPv6 next header 103** (same protocol type as IPv4)

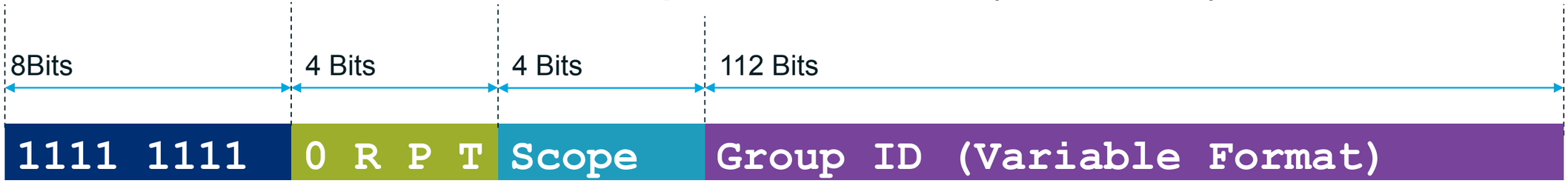
Types of Multicast Groups

- General Any Source Multicast (ASM)
  - PIM-SM, PIM-BiDir
  - Default for generic multicast and unicast prefix-based multicast
  - Start with FF3x::/12
- Source Specific Multicast (SSM)
  - Used by PIM-SSM
  - FF3x::/32 is allocated for SSM by IANA
  - However, at present prefix and plen must be zero so FF3x::/96 is usable as SSM
- Embedded RP groups
  - PIM-SM, PIM-BiDir
  - Start with FF70::/12

IPv6 Multicast Addressing

IPv6 Multicast Address Format (RFC 4291)

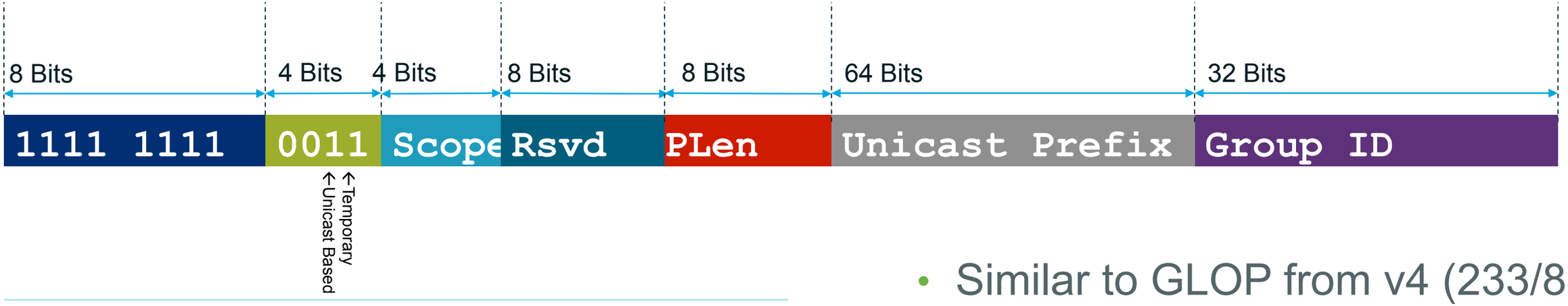
- An IPv6 multicast address has the prefix FF00::/8 (1111 1111)



Flags	
R = 0 R = 1	No embedded RP Embedded RP
P = 0 P = 1	Not based on unicast Based on unicast
T = 0 T = 1	Permanent address (IANA assigned) Temporary address (local assigned)

Scope	
1	Node
2	Link
3	Subnet
4	Admin
5	Site
8	Organization
E	Global

IPv6 Unicast Based Multicast Address (RFC 3306)



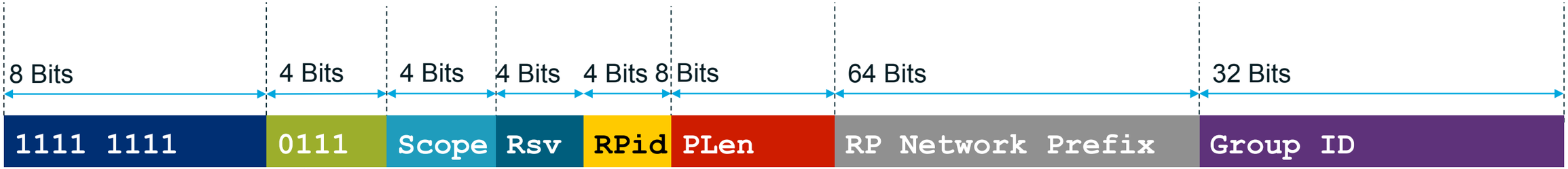
Example	
Unicast Prefix	2001:db8:cafe:1::/64
Flags	No RP, Unicast, Temporary
Scope	E (Global)
Group ID	11ff:11ee

ff3e:40:2001:db8:cafe:1:11ff:11ee

Unique Multicast Group Address

- Similar to GLOP from v4 (233/8) where the ASN is inserted in middle octets.
- Allows Multicast service from organizations that may not have BGP/ASN
- Provides a mechanism of creating globally unique multicast groups

IPv6 Embedded RP Multicast Address (RFC 3956)



Example	
RP Prefix	2001:db8:cafe:1::/64
Flags	RP, Unicast, Temporary
Scope	8 (Organisational)
RP ID	5
Group ID	645

Embedded RP Group

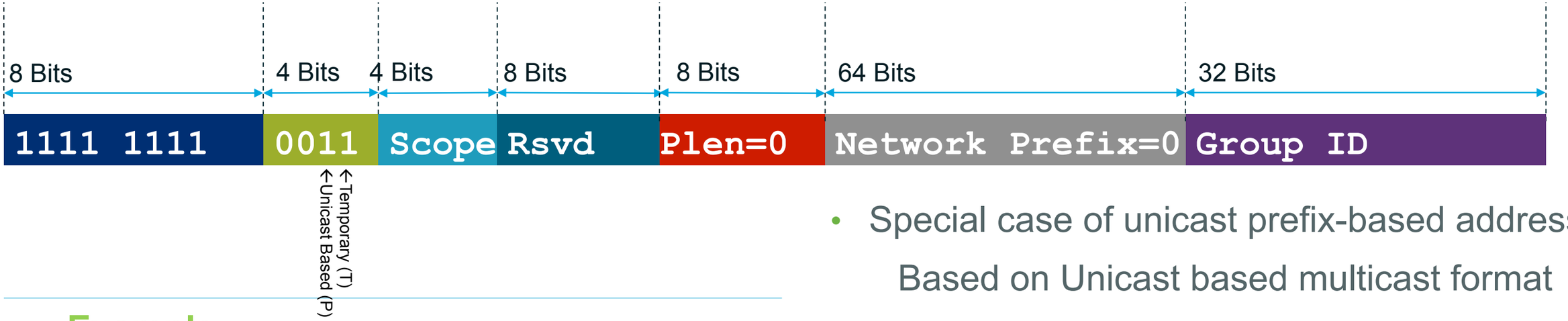
ff78:540:2001:db8:cafe:1::645

2001:db8:cafe:1::5

RP address

- Provides a static RP to multicast group mapping mechanism
- Solves interdomain multicast problem as there is no MSDP for IPv6
- Rpid field can be 1 - 15 (1-F hex)
- There can be 15 RPs per scope per prefix
  - Total of 256 RP addresses per unicast prefix
  - 2^32 groups per RP
- Embedded RP begin with FF7x::/12

IPv6 Source Specific Multicast Address (RFC 3306)



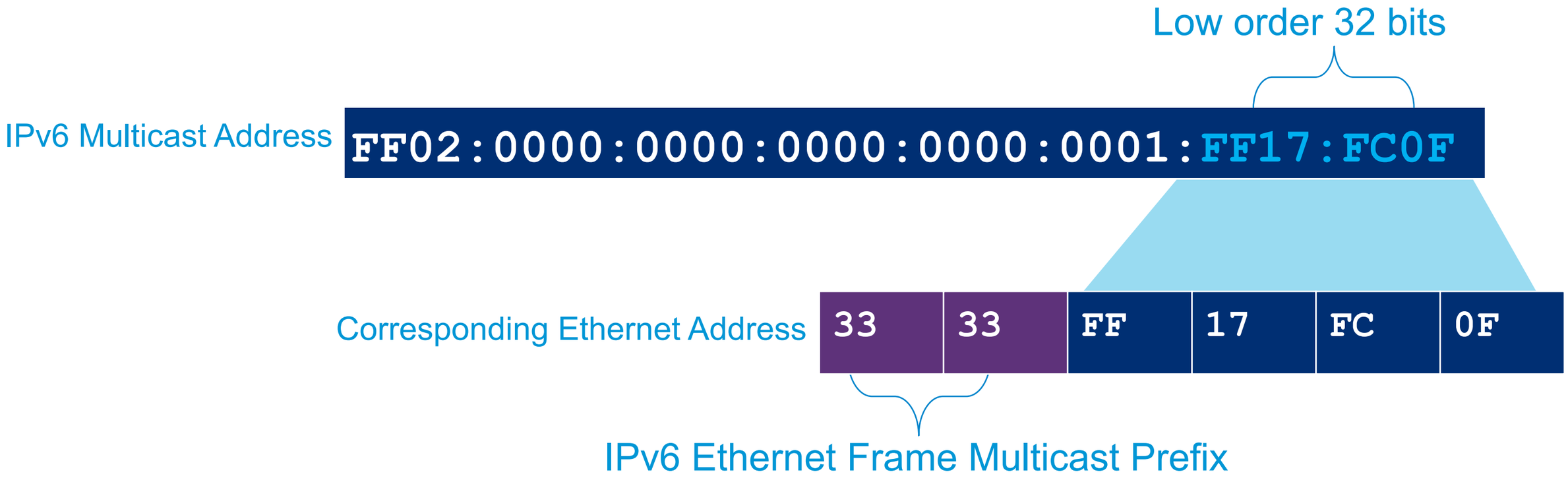
Example	
Unicast Prefix	0::
Flags	No RP, Unicast, Temporary
Scope	8 (Organisation)
Group ID	8000:247

ff38::8000:247 SSM Address

- Special case of unicast prefix-based address
  Based on Unicast based multicast format
- Prefix Len=0, Network Prefix=0
- FF3x::/32 pool is reserved for SSM addresses
  FF3x::/96 initial block allocated from this pool

Range	Usage
FF3x::4000:0001 - FF3x::7FFF:FFFF	IANA allocation
FF3x::8000:0000 - FF3x::FFFF:FFFF	Dynamic allocation
FF3x::0000:0000 - FF3x::3FFF:FFFF	Invalid for IPv6 SSM

Multicast Mapping over Ethernet (RFC 2464)



- IPv6 multicast address to Ethernet mapping

  33:33:{Low Order 32 bits of the IPv6 multicast address}

Well Known Multicast Addresses

Address	Scope	Meaning
FF01::1	Node-Local	All Nodes
FF05::2	Site-Local	All Routers
FF02::1	**Link-Local**	**All Nodes**
FF02::2	**Link-Local**	**All Routers**
FF02::A	Link-Local	EIGRP
FF02::C	Link-Local	SSDP – MSFT
FF02::FB	Link-Local	MDNS - Apple

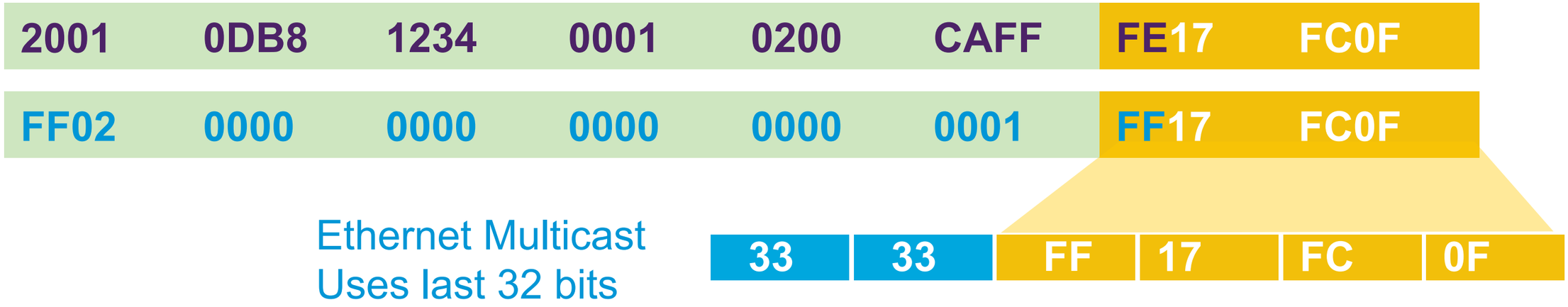
- FF02, is a permanent address and has link scope
- Rather “Chatty” and running in your network now!

Solicited Node Multicast Address

Solicited-Node Multicast Address

- For each Unicast and Anycast address configured there is a corresponding solicited-node multicast
- Used in neighbor solicitation (NS) messages
- Solicited-node multicast consists of

  FF02::1:FF/104 {lower 24 bits from IPv6 Unicast interface ID}



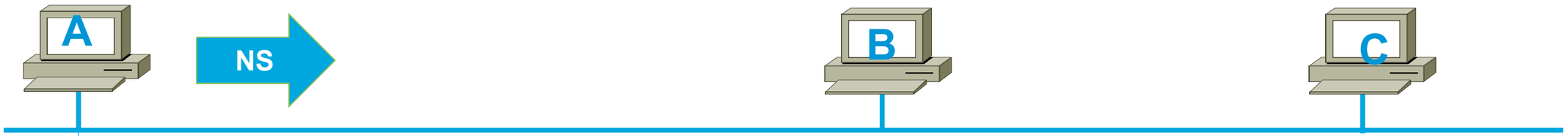
Neighbor Solicitation & Advertisement

ICMP Type	135 NS
IPv6 Source	A Unicast
IPv6 Destination	B Solicited Node Multicast FF02::1:FF00:B
Data	FE80:: address of A
Code	**0** (need link layer)
Query	What is B link layer address?

ICMP Type	136 NA
IPv6 Source	B Unicast
IPv6 Destination	A Unicast
ICMP Option	Type 2 (Target response)
Data	Link Layer address of B
*Flags	R = Router S = Response to Solicitation O = Override cache information

- Local Link only, Not Routed
- ARP replacement, Map's L3 to L2.

Duplicate Address Detection (DAD)



ICMP Type	135 NS
IPv6 Source	UNSPEC = ::
IPv6 Dest.	A Solicited Node Multicast FF02::1:FF00:A
Data	FE80:: address of A
Query	Anyone using A?

Node A can start using address A

- Probe neighbors to verify address uniqueness

IPv6 Interface Example

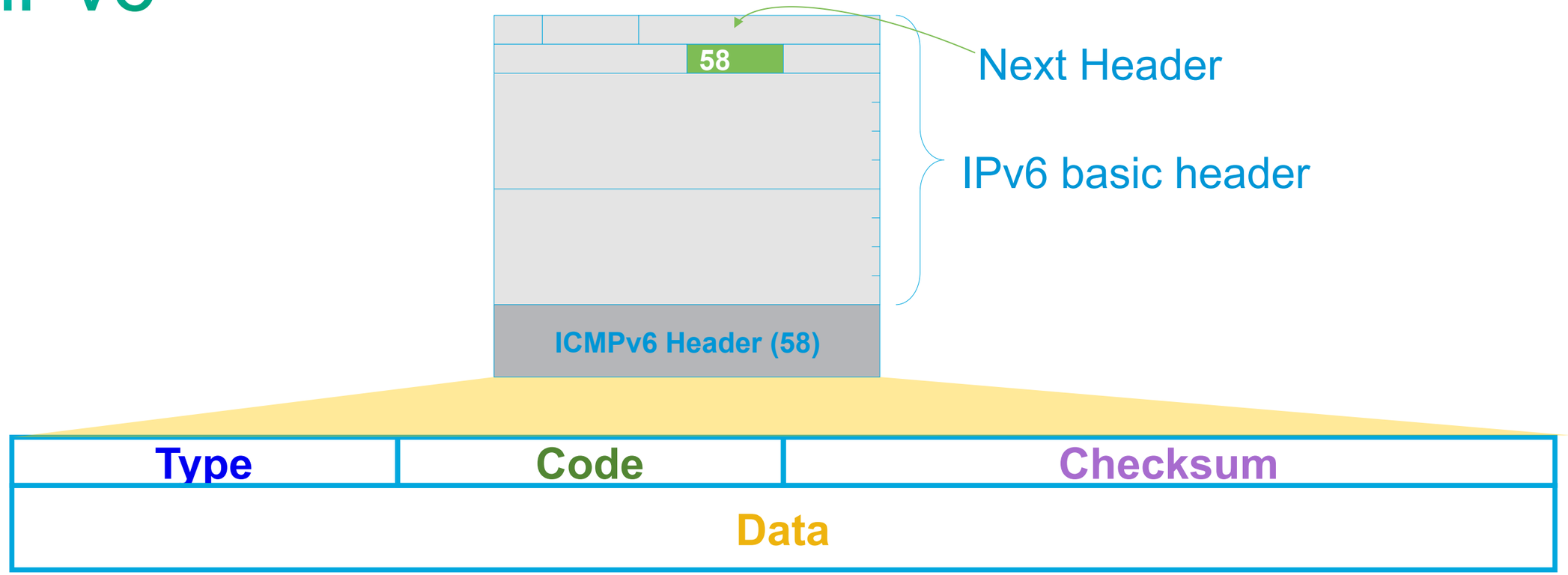
```
R1#sh ipv6 int e0
Ethernet0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::200:CFF:FE3A:8B18
  Global unicast address(es):
       2001:DB8:0:1234::1 subnet is 2006:1::/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:1
    FF02::1:FF3A:8B18
  MTU 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
```

Solicited-Node Multicast Address*

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

Multicast Listener Discovery (MLD)

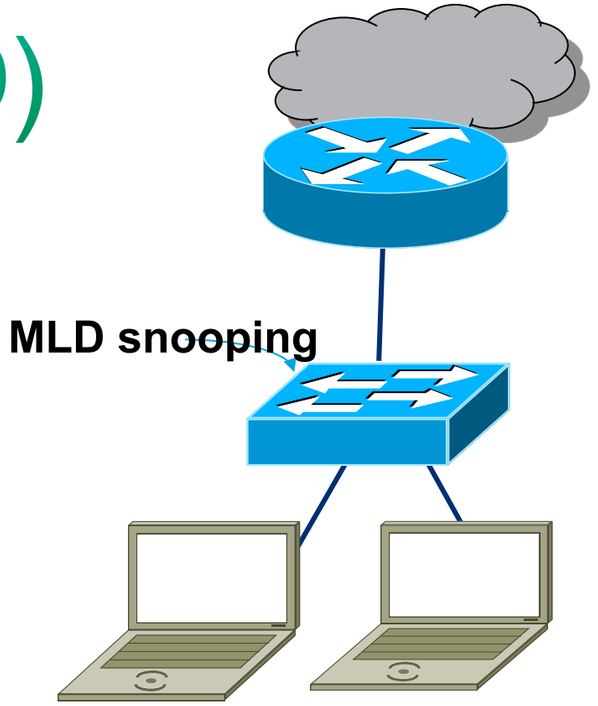
ICMPv6



- Neighbor Discovery, Router Discovery, Path MTU Discovery and (MLD)
  - **Type – (1-127) = Error Messages, (128-255) = Informational Messages**
  - **Code – More Granularity within the Type**
  - **Checksum – computed over the entire ICMPv6**
  - **Data – Diagnostic Information Relative to Packet Processing**

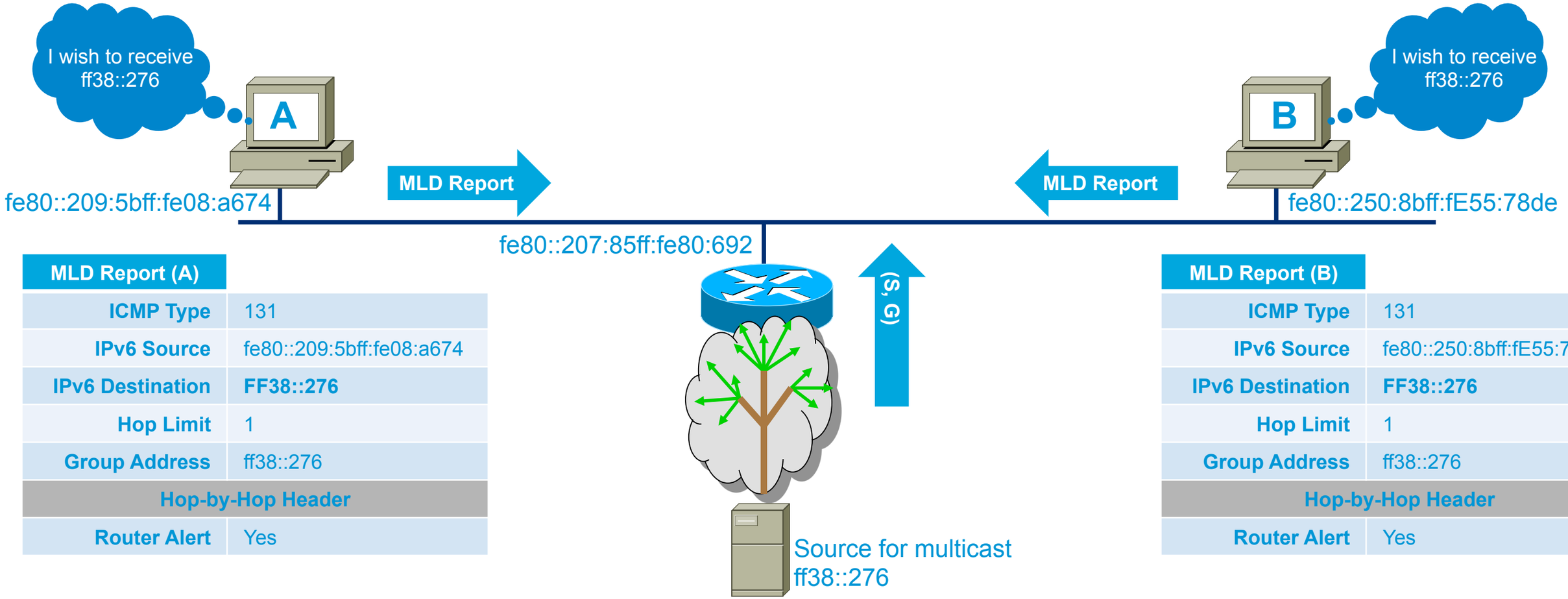
IPv6 Multicast Listener Discovery (MLD)

- MLD uses LL source addresses
- 3 msg types: Query, Report, Done
- MLD packets use “Router Alert” in HBH
- Snooping for efficient delivery at L2 boundary



MLD	IGMP	Message Type	ICMPv6 Type	Function
MLDv1 (RFC2710)	IGMPv2 (RFC 2236)	Listener Query	130	Used to find out if there are any multicast listeners
		Listener Report	131	Response to a query, joins a group
		Listener Done	132	Sent by node to report it has stopped listening
MLDv2 (RFC 3810)	IGMPv3 (RFC 3376)	Listener Query	130	Used to find out if there are any multicast listeners
		Listener Report	143	Enhanced reporting, multiple groups and sources

MLDv1 Example Joining a Group (REPORT)

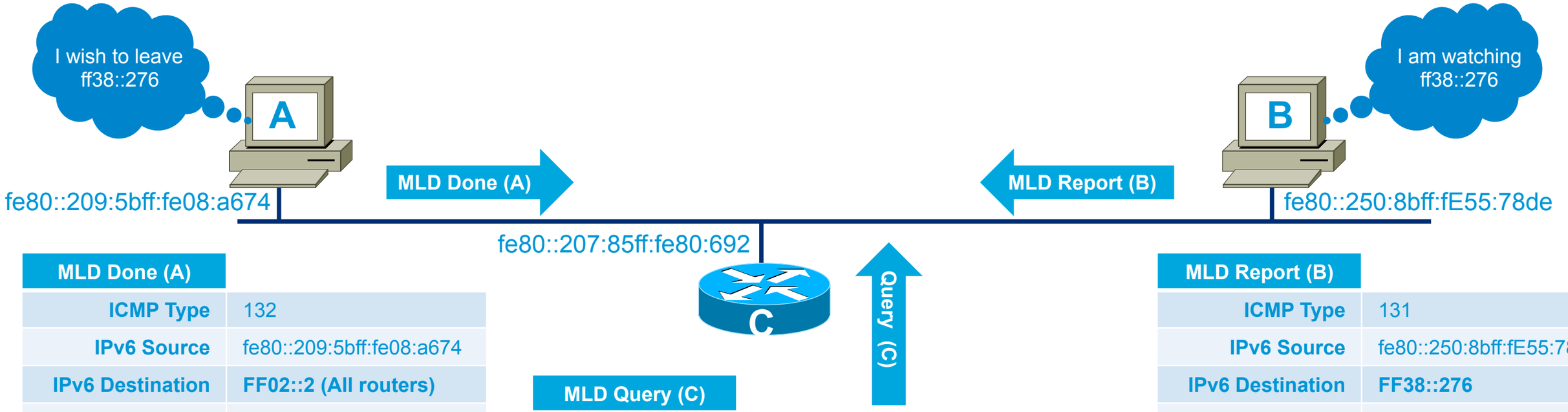


MLD Report (A)	
ICMP Type	131
IPv6 Source	fe80::209:5bff:fe08:a674
IPv6 Destination	**FF38::276**
Hop Limit	1
Group Address	ff38::276
Hop-by-Hop Header	
Router Alert	Yes

MLD Report (B)	
ICMP Type	131
IPv6 Source	fe80::250:8bff:fE55:78de
IPv6 Destination	**FF38::276**
Hop Limit	1
Group Address	ff38::276
Hop-by-Hop Header	
Router Alert	Yes

- 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)

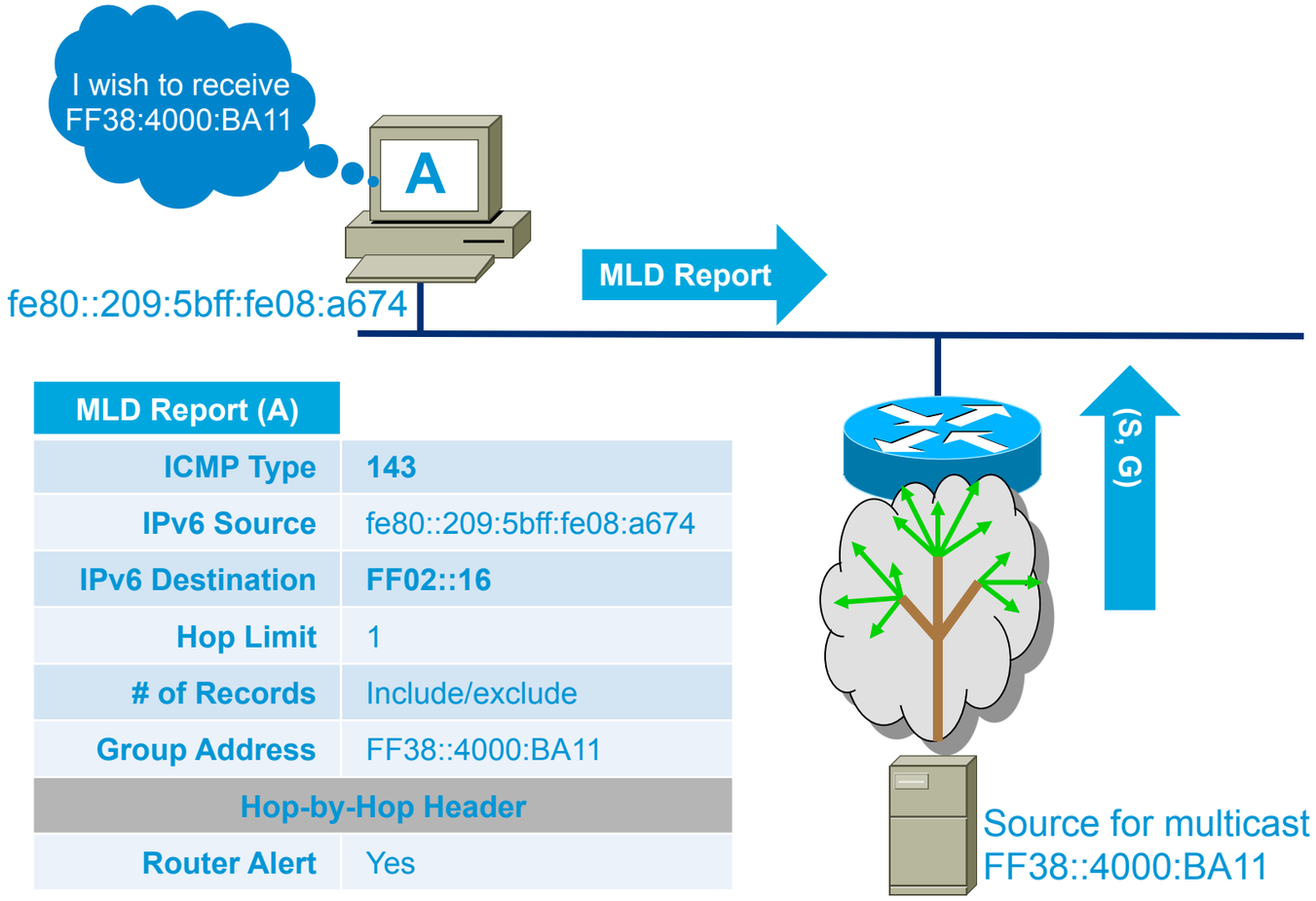


MLD Done (A)	
ICMP Type	132
IPv6 Source	fe80::209:5bff:fe08:a674
IPv6 Destination	FF02::2 (All routers)
Hop Limit	1
Group Address	ff38::276
Hop-by-Hop Header	
Router Alert	Yes

MLD Query (C)	
ICMP Type	130
IPv6 Source	fe80::207:85ff:fe80:692
IPv6 Destination	FF38::276
Hop Limit	1
Hop-by-Hop Header	
Router Alert	Yes

MLD Report (B)	
ICMP Type	131
IPv6 Source	fe80::250:8bff:fE55:78de
IPv6 Destination	FF38::276
Hop Limit	1
Group Address	ff38::276
Hop-by-Hop Header	
Router Alert	Yes

MLDv2 Example (Report & Query)



MLD Report (A)	
ICMP Type	143
IPv6 Source	fe80::209:5bff:fe08:a674
IPv6 Destination	FF02::16
Hop Limit	1
# of Records	Include/exclude
Group Address	FF38::4000:BA11
Hop-by-Hop Header	
Router Alert	Yes

- Query Format MLDv1, MLDv2
- General Query (~125 Seconds)

  FF02::1
- Group Specific Query

  FF38::4000:BA11
- Group & Source Specific Query

  2001:DB8:CAFÉ::1, FF38::4000:BA11
- Leaving a Group MLDv2

  Ignore Query (silent)

  Filter mode Change Record (report)

Protocol Independent Multicast

PIM and IPv6 (RFC 4601)

- PIM is Join and Prune or PULL mode protocol, and transparent to the IP version
  - It is the only multicast protocol supported for IPv6 and uses next header type 103
- PIM Sparse-Mode (PIM-SM) - RP is required
  - Sparse-Mode for many-to-many applications (Multiple sources, single group)
  - Uses shared tree initially but may switch to source tree
- Bi-directional PIM (PIM-BiDir) - RP is required
  - Bi-Directional many-to-many (hosts can be sources and receivers)
  - Like PIM-SM but uses a BiDIR shared tree for all traffic
- PIM Source-Specific Multicast (PIM-SSM) - No RP is required
  - For one-to-many applications (Single source, single group)
  - Always uses a (S, G) source tree –
  - (S) is learnt somehow or known out of band

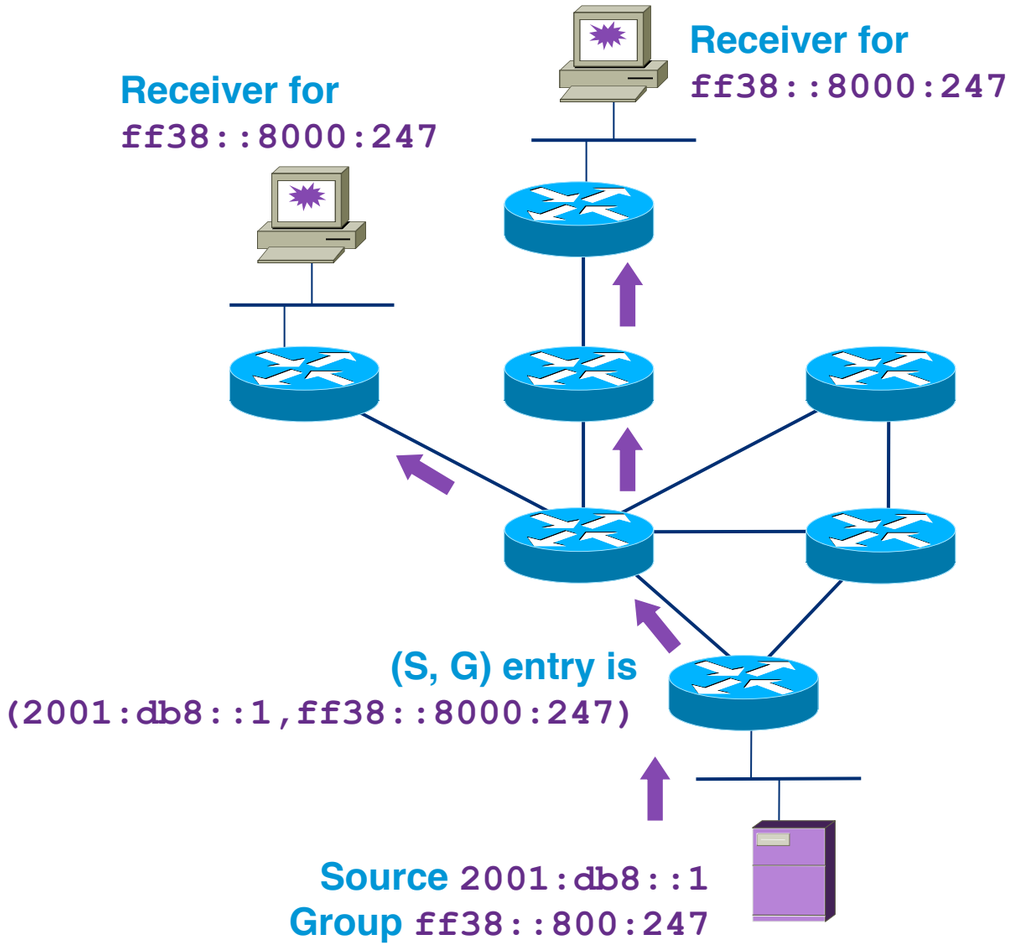
25

Multicast States (S, G) (*, G)

- Provides the forwarding entries for packet distribution down a tree
- Consists of the Source Address (S) and the Destination Group (G) of the multicast stream
- Expressed as (S, G) for Source Trees
  - Means an explicit source for a multicast group
  - More Memory, Optimal Paths, Less Delay
- Expressed as (*, G) for Shared Trees
  - Means ALL sources for a multicast group
  - Less Memory, Sub Optimal paths, Extra Delay

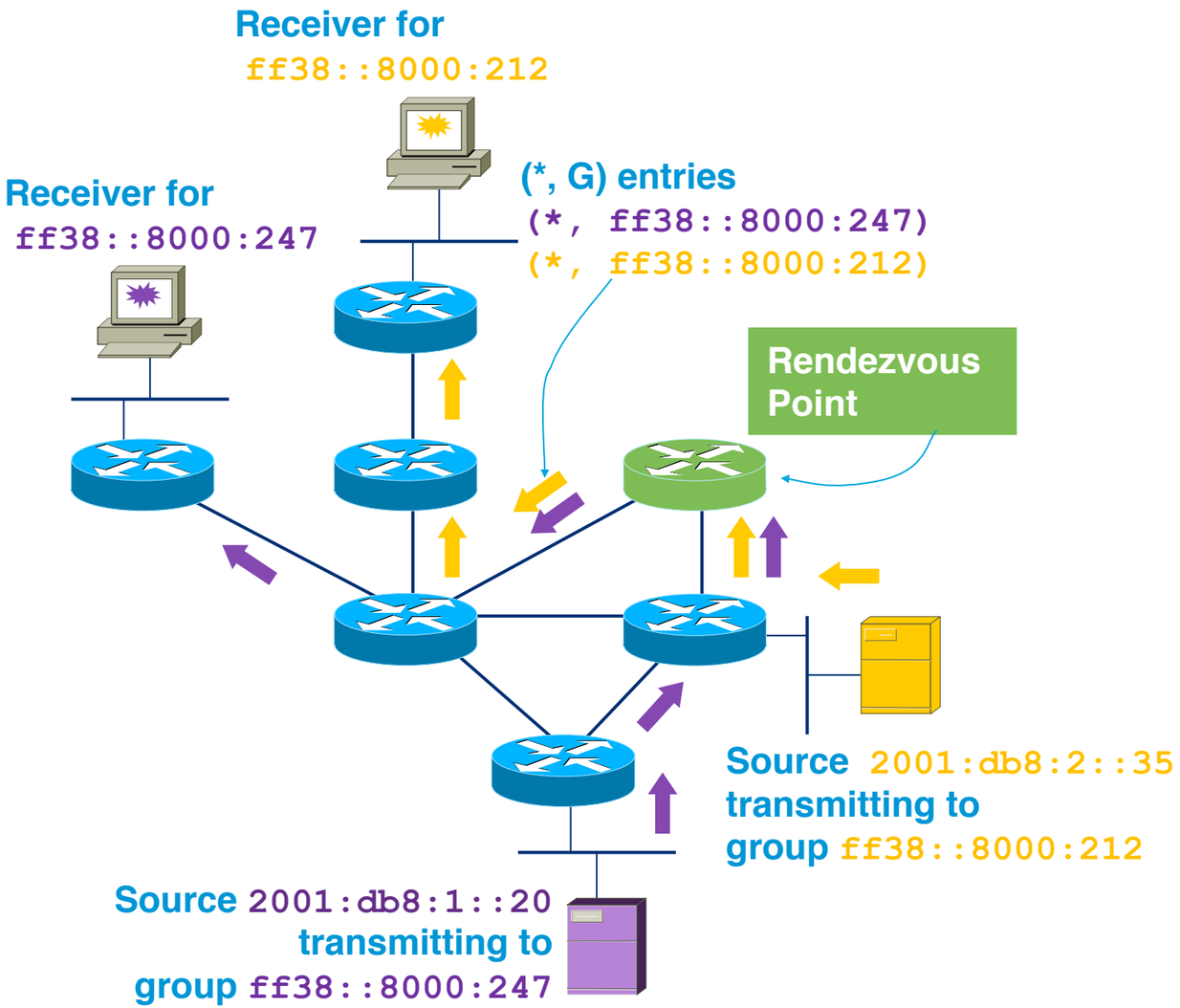
26

Multicast Source Tree



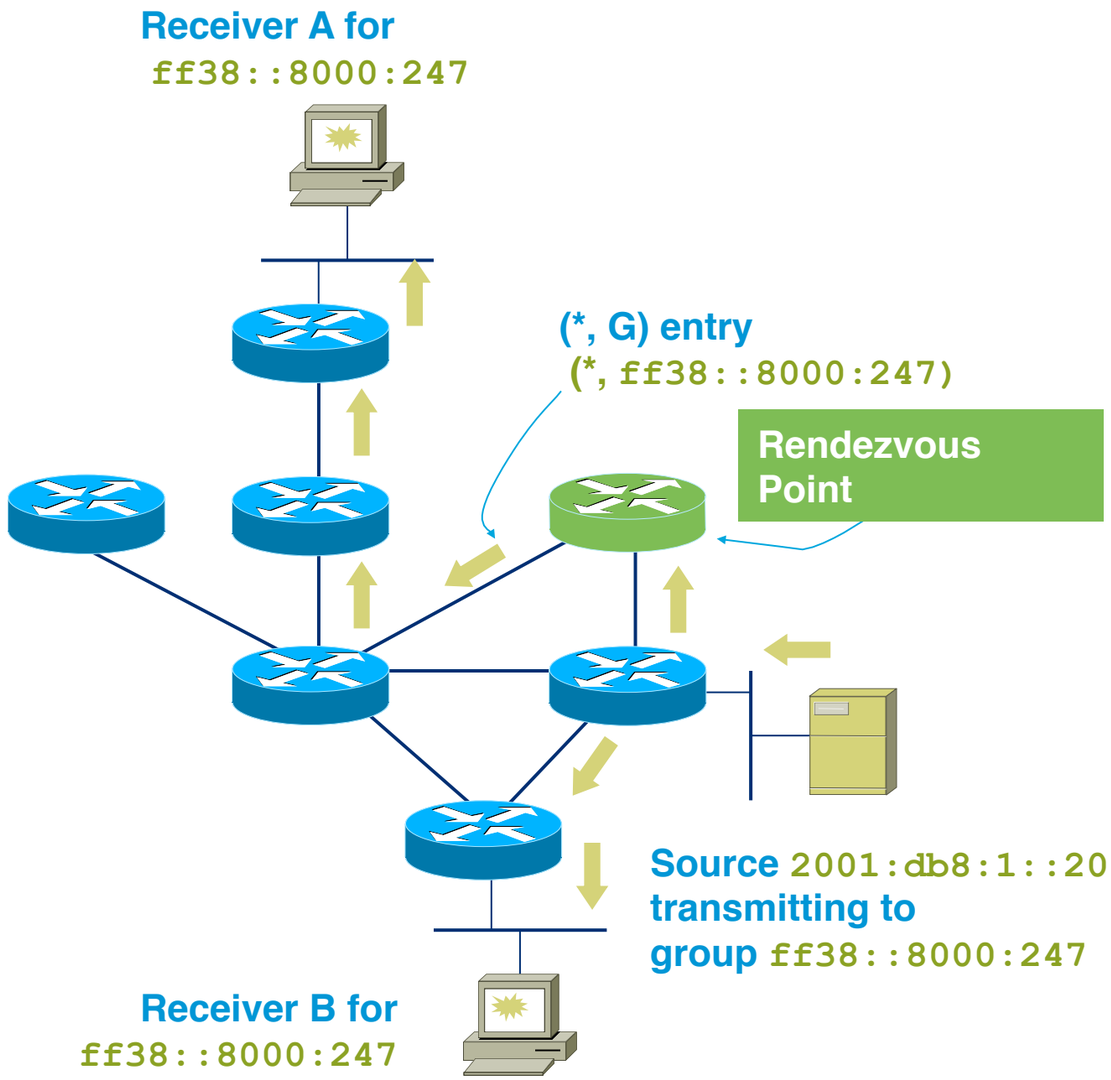
- Simplest form of tree
  - Receiver requires knowledge of source
- Traffic from source (root) to receivers (leaves)
- Shortest path taken
- Packets replicated at branch point
- Forwarding entry states represented as (S, G)
- Provides Optimal routing
  - At the expense of more state (S, G)
- Service model is SSM or ASM that has moved to an SPT

Multicast Shared Tree



- Root is a common point
  - Rendezvous Point
  - Many multicast groups at RP
- Receivers join RP
  - To learn of sources
- Sources only transmit to RP
  RP forward to receivers
- Forwarding represented as (*, G)
- Less state required
  - At expense of optimal routing
- Service model is ASM

Bi-Directional Shared Tree



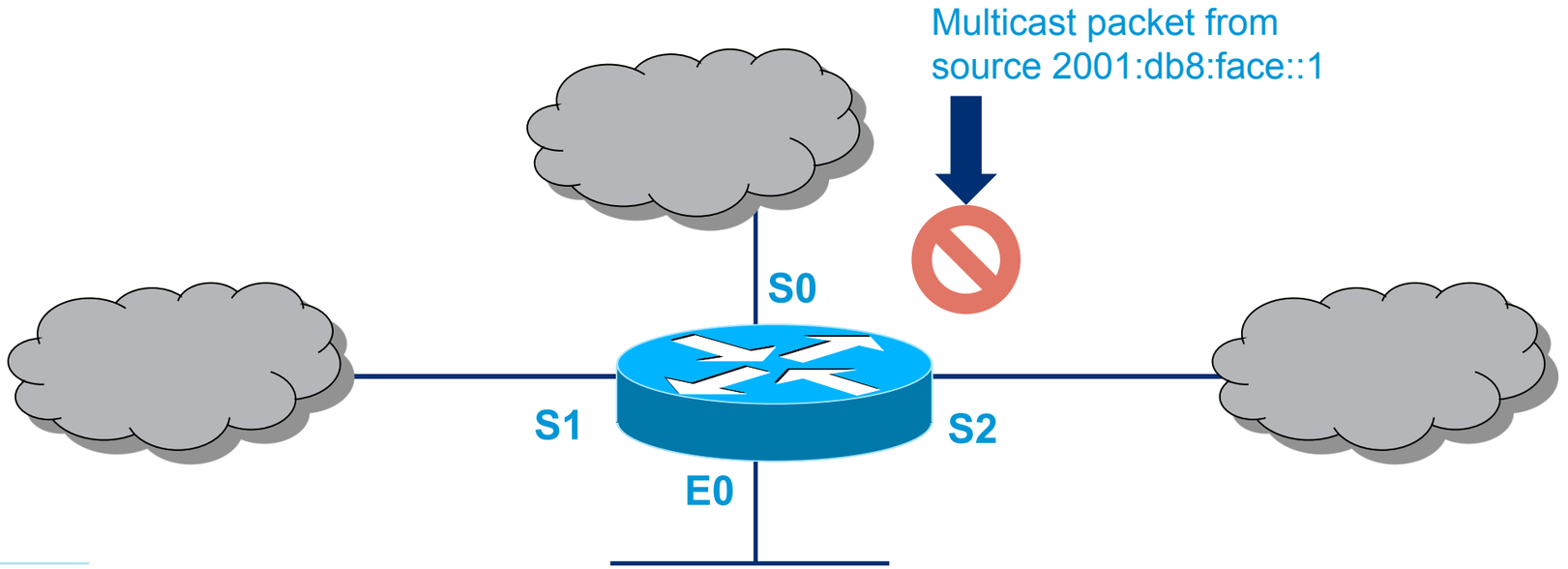
- Traffic can travel in both directions
  - Up and Down the tree
- Source packets do not necessarily have to travel via the RP
- Forwarding entries represented as (*, G)
- Offers improved routing optimality than uni-directional shared tree
- Service model is ASM

Multicast Forwarding

- Multicast forwarding is the opposite of Unicast forwarding
  - Unicast is concerned about where the packet is going
  - Multicast is concerned about where the packet came from
- Multicast uses Reverse-Path Forwarding (RPF)
  - Checks if arriving packet is on reverse path back to source
  - If successful, packets is forwarded, otherwise dropped
- RPF procedure for PIM uses unicast routing table to find source

Multicast Forwarding RPF Check

RPF Check Fails

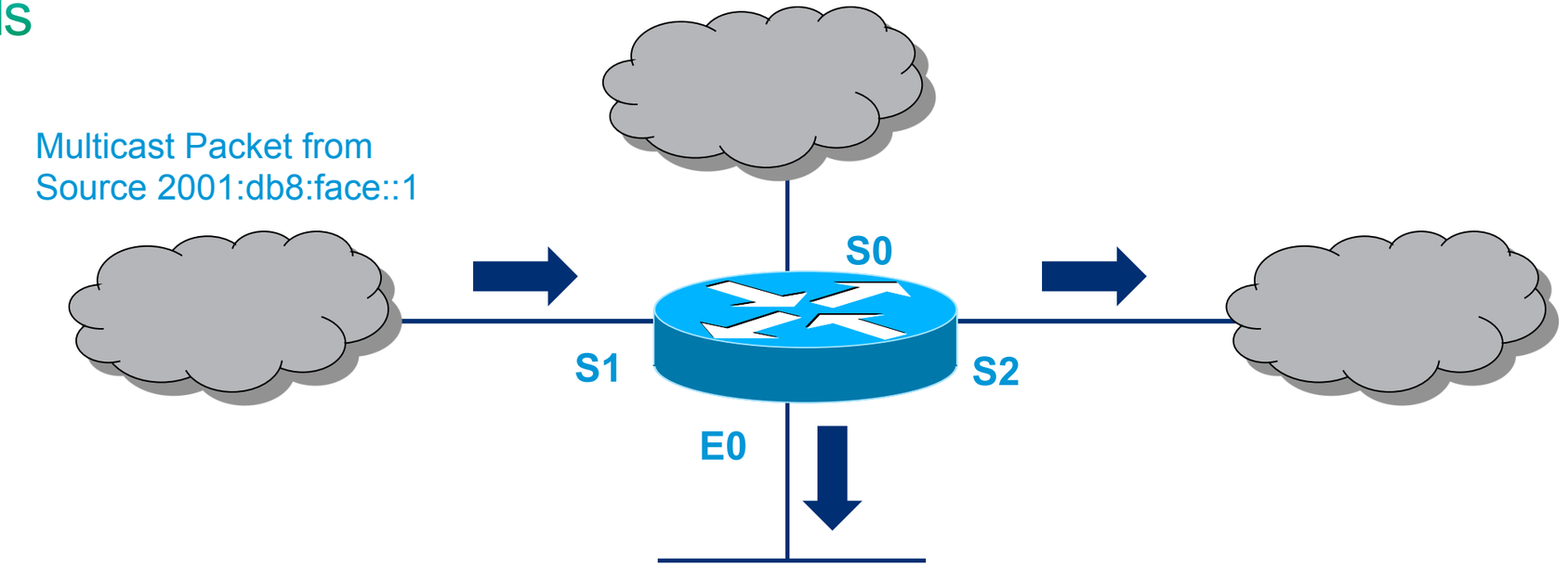


Unicast Route Table	
Network	Interface
2001:db8:face::/48	S1
2001:db8:beef::/48	S0
2001:db8:f00d::/48	E0

Packet has arrived on wrong Interface
Discard the packet!

Multicast Forwarding RPF Check

RPF Check Succeeds



Unicast Route Table	
Network	Interface
2001:db8:face::/48	S1
2001:db8:beef::/48	S0
2001:db8:f00d::/48	E0

Packet arrived on correct interface!
Forward via all outgoing interfaces
(i.e. down the distribution tree)

Final Thoughts

Key Take Away

- Applications We Haven't Even Built Yet
- Large Privately Owned Multicast Address Space
- Built-in Scoping
- No NAT required
- Embedded RP, Anycast, Etc..
- Multicast is Foundational in IPv6
- Invest in your future - IPv6, the future is now

34