SANOG 2008
IPv6 Multicast

Luly Motomura

Cisco Systems, Inc.

Why IPv6 Multicast?

Why IPv6 Multicast?

- Customers who are impacted by IPv4 address scarcity
  - Applications require IPv6
- Facilitates deployment of multicast apps to users behind v4 NAT services
- Good chance to simplify and improve the multicast model through use of IPv6
  - Embedded RP addresses, Unicast prefix addresses
  - No fragmentation below 1280 bytes
  - Powerful address scoping (no TTL scoping !)
  - Potentially simpler interdomain solutions
  - No DVMRP, perhaps no need for PIM-DM, and likely no need for MSDP
  - Proven routing technology (PIM), lessons learned from v4 implementations
- Good chance to get new applications to utilize IP multicast
  - A lot of research is spent on just "anything with IPv6"
- Good chance to establish IPv6 multicast as a ubiquitous day-1 service

End-to-End, Fully Integrated Solutions

- IPv6 Multicast deployed as a part of end-end solutions:
  - Initial Target Applications
    - Interactive TV, gaming, mobile services, conferencing
  - IPv6 Stacks & Applications supporting Multicast
  - Network infrastructure
    - IPv6 Multicast services

IPv6 Multicast Terminology

IPv6 Multicast Terminology

Protocol Independent Multicast v2 (PIMv2)

- Provides intradomain multicast forwarding for all underlying unicast routing protocols Independent from any underlying unicast protocol such as Open Shortest Path First (OSPF) or Multiprotocol Border Gateway Protocol (MP -BGP)
- IOS-XR only supports explicit join (a.k.a sparse mode) unlike IOS which supports flood-and-prune (dense mode), or hybrid sparse-dense modes in addition to sparse mode.
- Sparse Mode: relies upon an explicit joining method before attempting to send multicast data to receivers of a multicast group

Multicast Listener Discovery (MLD) v1 & v2

- Protocol used by IPv6 hosts to communicate multicast group membership states to local multicast routers
- Version 2 of MLD adds source awareness to the protocol. This allows the inclusion or exclusion of sources. MLDv2 is required for Source Specific Multicast (SSM)
- Cisco IOS-XR Software also supports the explicit tracking of MLDv2 receivers

IPv6 Multicast Terminology (continue…)

PIM Source Specific Multicast

- SSM forwarding uses only source-based forwarding trees. SSM range is defined for interdomain use, and Cisco IOS-XR Software allows other groups to be configured using the SSM forwarding model.

PIM Bi-dir

- A variant of the PIM, whereby data flows both up and down the same distribution tree
- Bi-directional PIM uses only shared tree forwarding, thereby reducing state creation

IPv6 Multicast Terminology (continue…)

Boot Strap Router (BSR)

- BSR is a mechanism where a PIM router learns the set of group -to-RP mappings required for PIM SM

BSR flooding

- The ability to flood BSR messages without processing them

Embedded Rendezvous Point

- Utilizes unicast based prefix addressing to include within the group address (the Rendezvous Point address)

Static Rendezvous Point

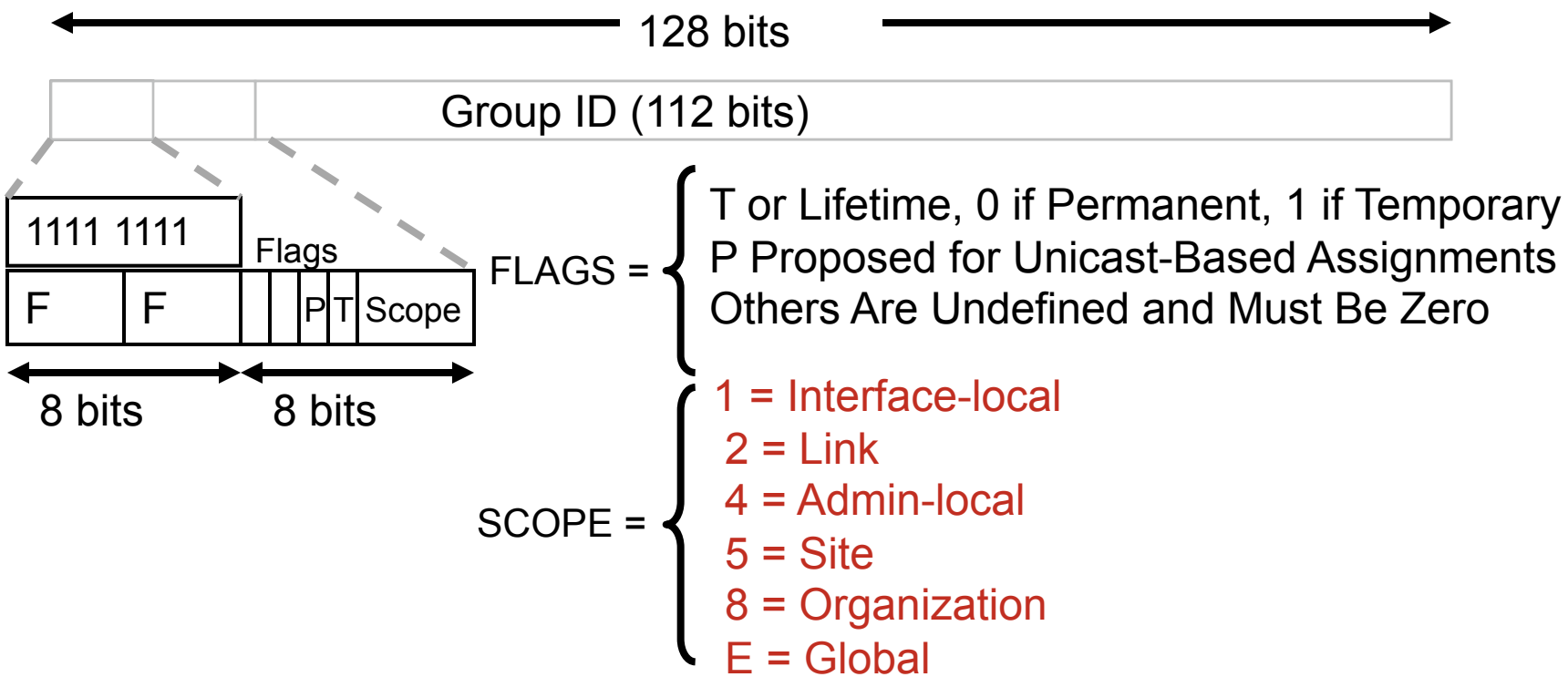
- Allows the manual configuration of the IPv6 PIM SM RP address

IPv4 - IPv6 COMPARISON
ADDRESSING
SERVICE MODELS

IPv4 and IPv6 Multicast Comparison

Service	IPv4 Solution	IPv6 Solution
Addressing Range	32-bit, Class D	128-bit (112-bit Group)
Routing	Protocol Independent, All IGPs and MBGP	Protocol Independent, All IGPs and MBGP with v6 mcast SAFI
Forwarding	PIM-DM, PIM-SM, PIM-SSM, PIM-bidir, PIM -BSR	PIM-SM, PIM-SSM, PIM-bidir, PIM-BSR
Group Management	IGMPv1, v2, v3	MLDv1, v2
Domain Control	Boundary, Border	Scope Identifier
Interdomain Solutions	MSDP across Independent PIM Domains	Single RP within Globally Shared Domains

IPv6 Multicast Addresses (RFC 3513)



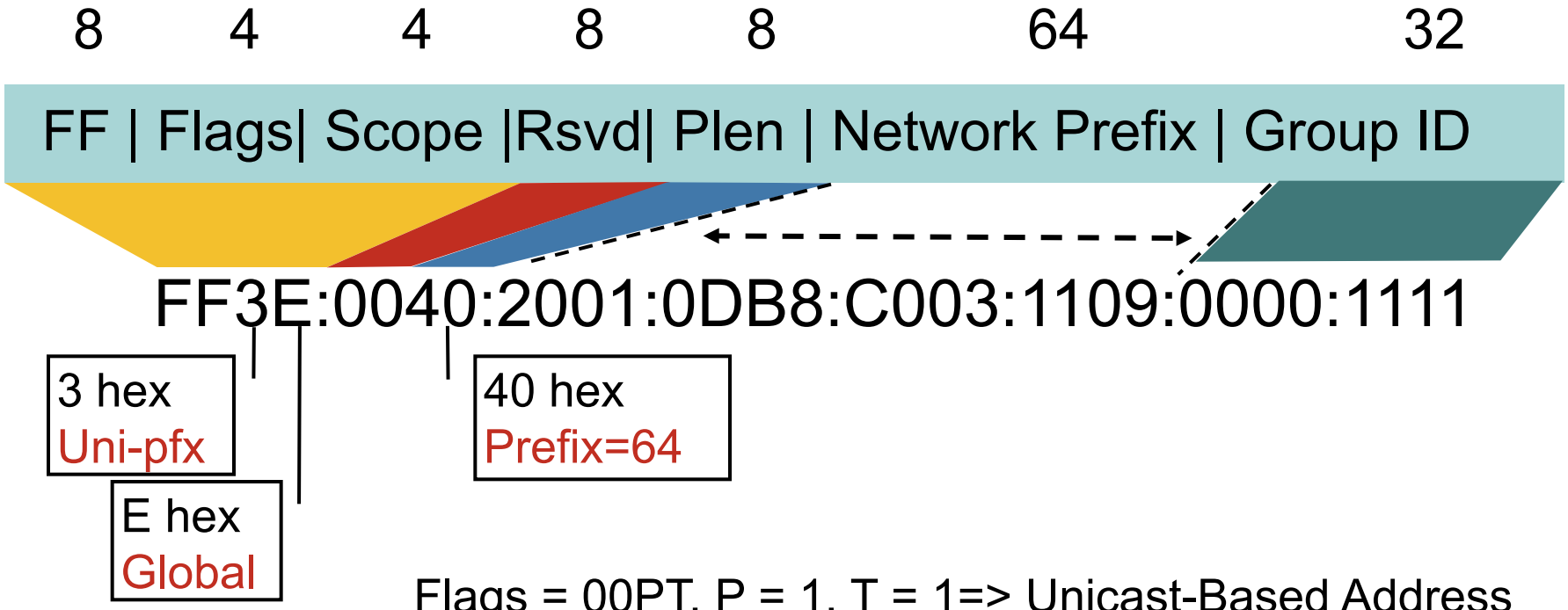
Permanently-Assigned Address Example

The "meaning" of a permanently-assigned multicast address is independent of the scope value. For example, if the "NTP servers group" is assigned a permanent multicast address with a group ID of 101 (hex), then:

- FF01:0:0:0:0:0:0:101 means all NTP servers on the same node as the sender.
- FF02:0:0:0:0:0:0:101 means all NTP servers on the same link as the sender.
- FF05:0:0:0:0:0:0:101 means all NTP servers at the same site as the sender.
- FF0E:0:0:0:0:0:0:101 means all NTP servers in the internet

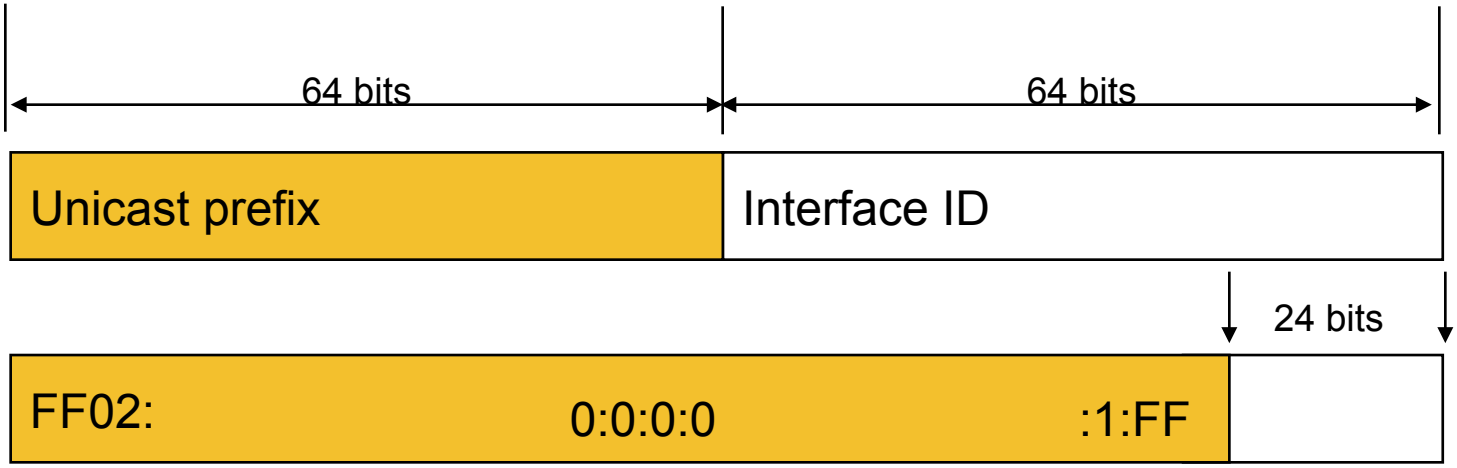
IPv6 Unicast Based Multicast Addresses (RFC3306)

- Solves the old IPv4 address assignment problem:
  How can I get global IPv4 multicast addresses?
- In IPv6, if you own an IPv6 unicast address prefix you implicitly own an RFC3306 IPv6 multicast address prefix:



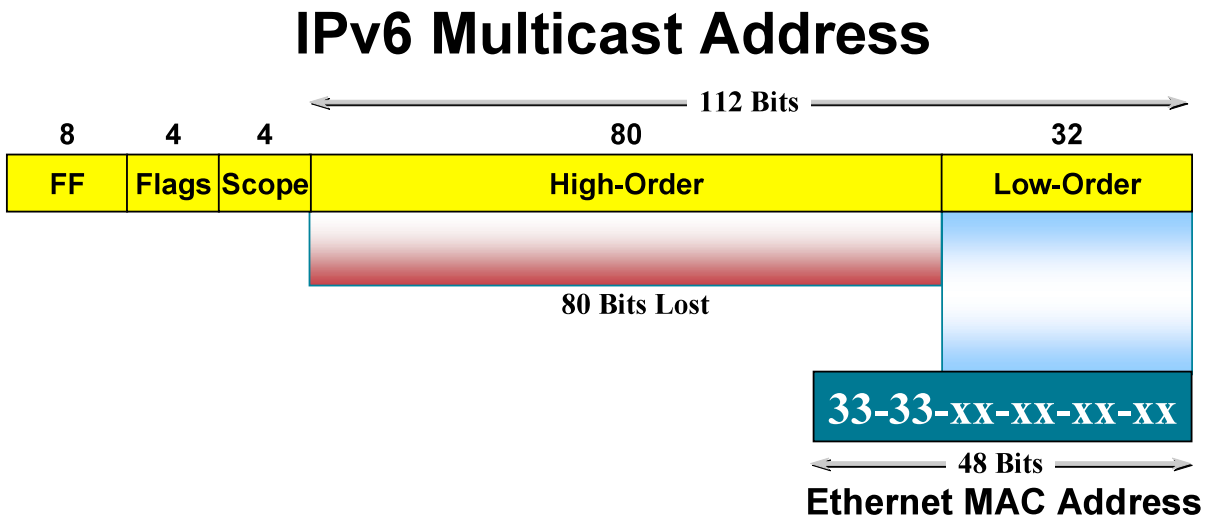
Flags = 00PT, P = 1, T = 1=> Unicast-Based Address

Solicited-Node Multicast Address



- FF02::1:FF00:0000/104 - IPv6 prefix (compressed)
- Consists of the prefix and the low-order 24-bits of the unicast or anycast address
- Link-local address - FE80::20B:45FF:FE94:1C00
- Solicited-node address - FF02::1:FF94:1C00

MAC Address Mapping



RFC2464

Example: FF05:1::5 → 33:33:0:0:0:5

More than 1 IPv6 multicast address will map to the same MAC address (80 bit is lost)

IPv6 Multicast Service Models

- Any Source Multicast (ASM)
  (Traditionally just called IP Multicast)

  MLDv1: RFC2710 or MLDv2: RFC3810

  PIM-Sparse Mode (PIM-SM)

  Bidirectional PIM (PIM-bidir)

  PIM-BSR (Boot Strap Router)

- Source Specific Multicast (SSM)

  MLDv2 required

  PIM-SSM – not a separate protocol, just a subset of PIM-SM

IPv6 Multicast Service Models (Cont.)

Recommendations:

- Use SSM for media-broadcast or interdomain applications due to simplicity and protection from denial of service (DoS) attacks

  Requires moderate amount of application side work

- Use ASM for legacy, dynamic- or many-source multi-party application, try to limit their use to Intradomain:

  Start with PIM-SM and consider Bidir-PIM for many-source applications

  Use PIM-SM with embedded Rendezvous Point for simple and reliable Interdomain ASM

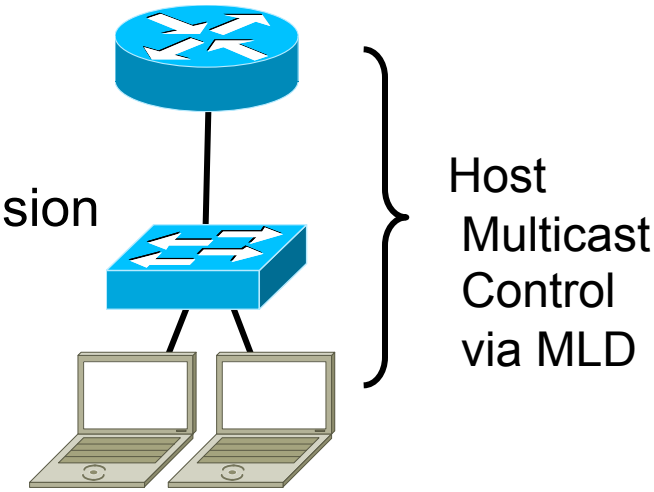
Multicast Listener Discovery: MLD

Multicast Listener Discovery: MLD
Multicast Host Membership Control

- MLD is equivalent to IGMP in IPv4
- MLD messages are transported over ICMPv6
- MLD uses link local source addresses
- MLD packets use “Router Alert” in extension header (RFC2711)
- Version number confusion:

  MLDv1 (RFC2710) like IGMPv2 (RFC2236)

  MLDv2 (RFC3810) like IGMPv3 (RFC3376)
- Only MIB available today is for MLDv1
- MLD snooping



Multicast Listener Discover

- Service Model requirements:

  ASM – MLDv1 sufficient

  SSM – Requires MLDv2

  Cisco IOS XR Software only provides MLDv2 router side:

  Fully backward compatible with MLDv1 hosts

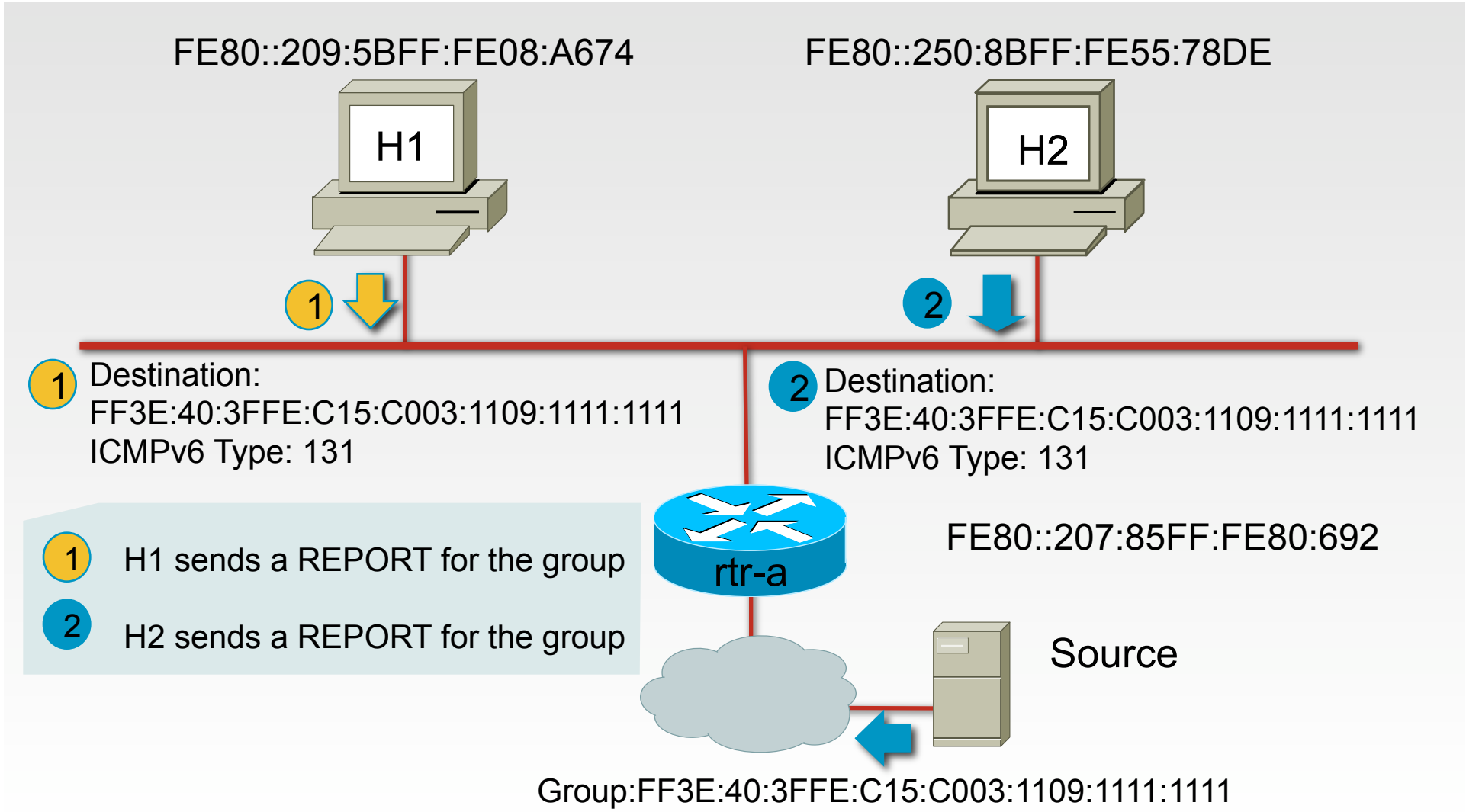
- SSM transition methods

  Cisco IOS XR IPv4 multicast has 3 transition methods for IGMPv3:

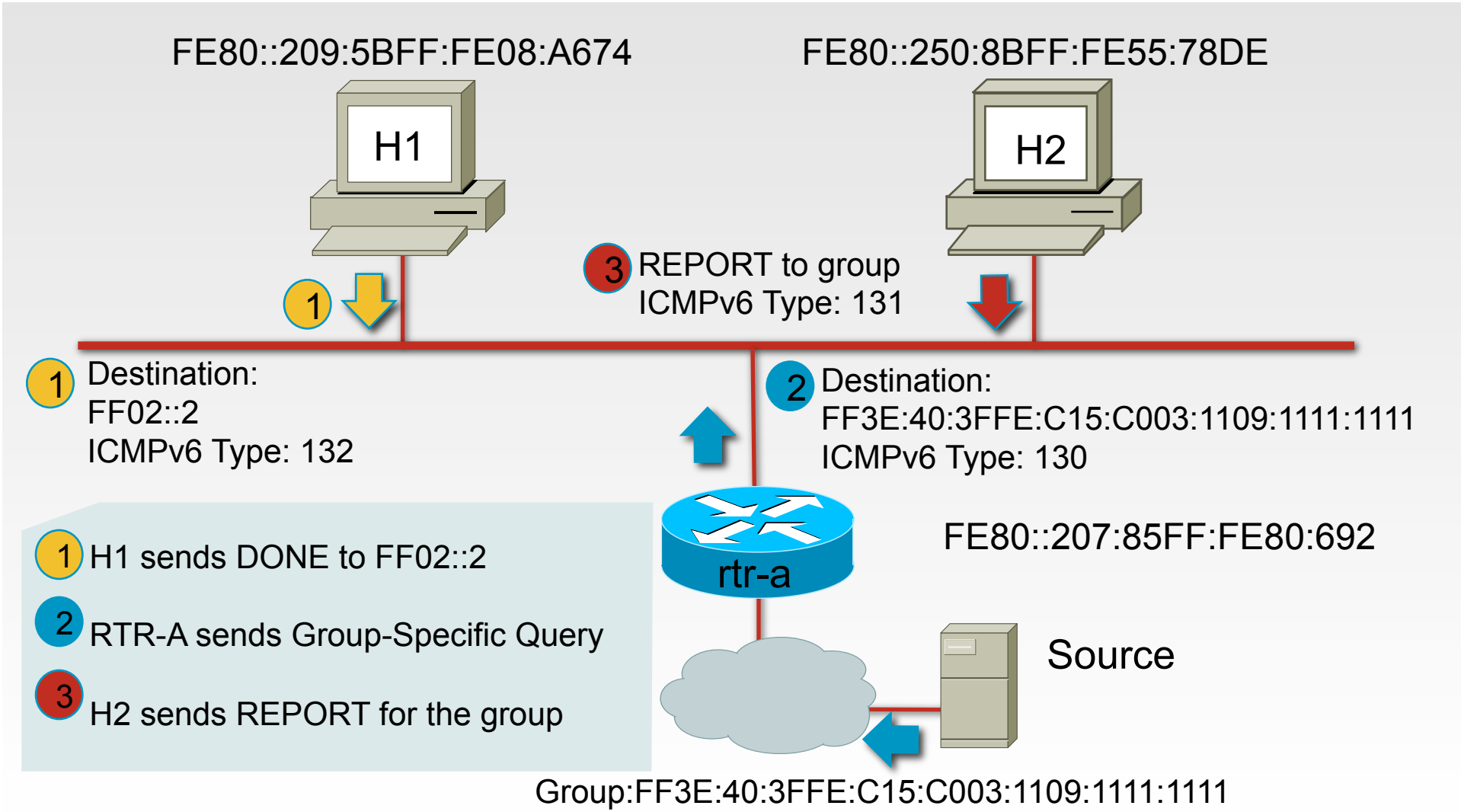
  IGMPv3 lite, URD, SSM-Mapping

  SSM mapping for MLDv1 -> MLDv2 supported in 12.2(18)SXE, 12.4T

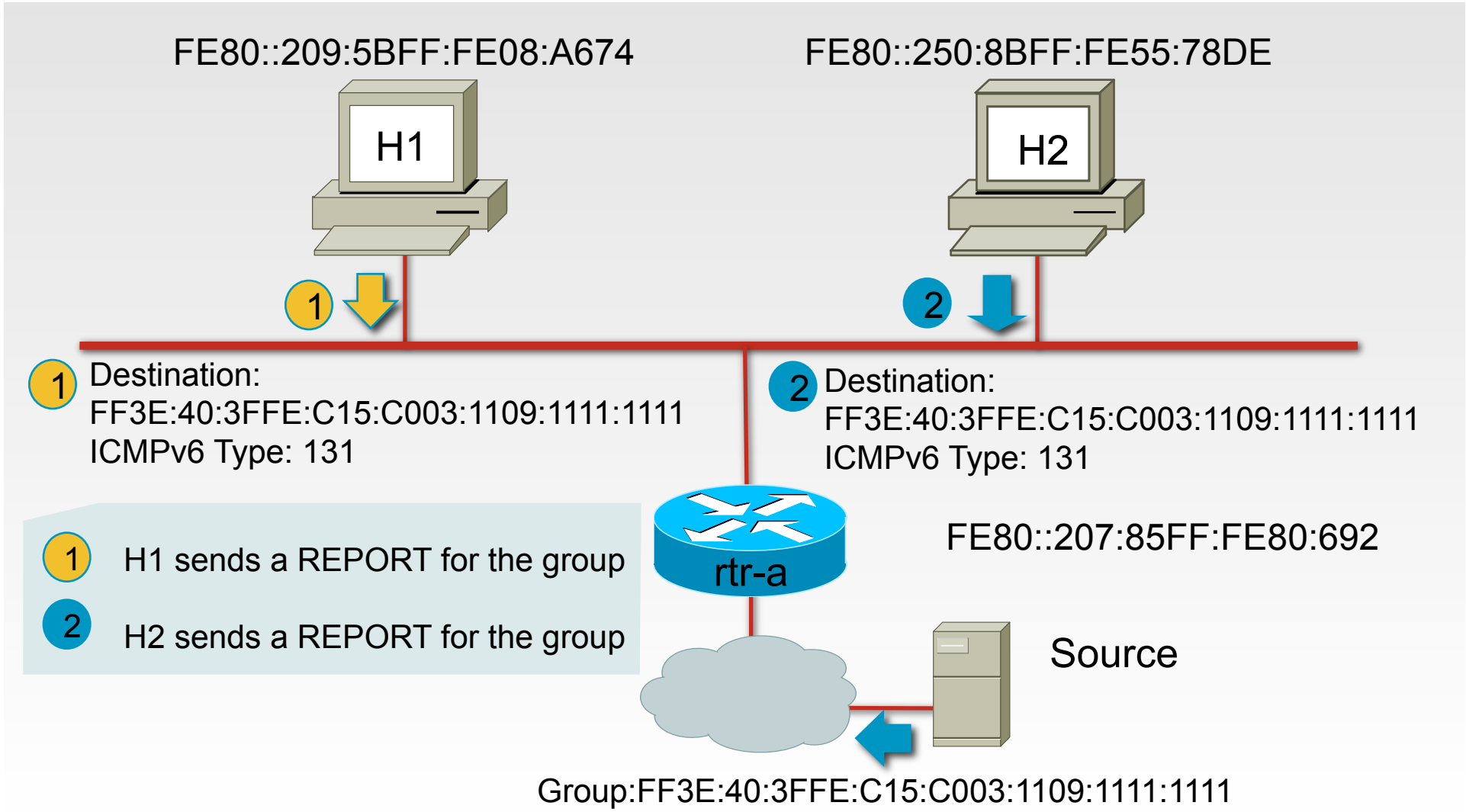
MLDv1 - Joining a Group (REPORT)



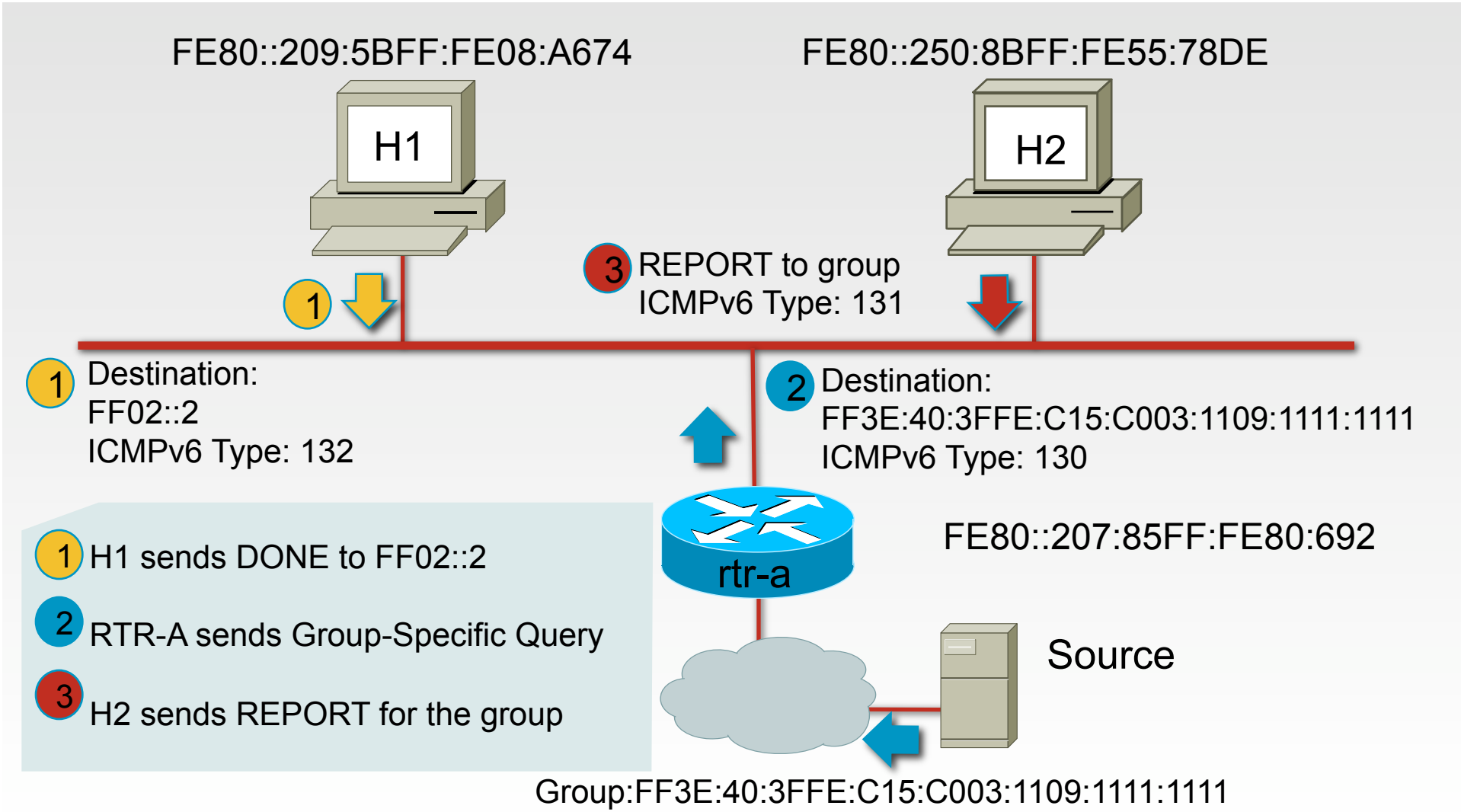
MLDv1 – Host Management (Group-Specific Query)



MLDv2 - Joining a Group (REPORT)



MLDv2 – Host Management (Group-Specific Query)



RPF Selection

Cisco IOS IP Multicast RPF

- IPv4 RPF selection (distance-only)

  Consider the following tables for RPF information:

  1. IPv4 MBGP RIB 10.0.0.0/8
  2. IPv4 static mroutes
  3. IPv4 DVMRP RIB
  4. IPv4 unicast RIB 10.1.1.0/24

  Find route with lowest administrative distance across these tables (within each table a longest-mask-first lookup is done!)

  If two or more tables have routes with identical distance, take first route according to above table order

- *Can select a route with shorter prefix-length over one with a longer prefix*

  eg RPF lookup for 10.1.1.1 will use MBGP route

Cisco IOS IP Multicast RPF

- IPv6 RPF selection

  I. Look up the longest mask route from the available route sources:

  1. static mroutes

  2. MBGP RIB

  3. unicast RIB

  Ia. Exclude IPv6 BGP unicast routes by default

  II. If more than one of these three sources returns a route with the same longest mask then select amongst these routes the one with the lowest (= best) distance

  III. If the distance of more than one of these routes is the same too, then use the first one according to above table (eg: prefer static mroute over MBGP over unicast RIB)

- Same algorithm is used in IP Unicast route selection

Cisco IOS IP Multicast RPF

- IPv6 BGP/MBGP

  Unicast: MBGP, RFC2858/2545, AFI=IPv6, SAFI=1

  SAFI =1 indicates that these prefixes are only usable for IP unicast, but not for IP multicast

  BGP routes in the IPv6 unicast RIB must be ignored in the IPv6 multicast RPF lookup

  Cisco IOS Software does not support SAFI=3 (routes reachable for both multicast and unicast), because IETF has removed SAFI=3

  Multicast: MBGP, RFC2858/2545, AFI=IPv6, SAFI=2

  SAFI=2 – the route is only usable for IP multicast, but not for IP unicast.
  This type of announcement is also used in Cisco IOS IPv4 Multicast

Cisco IOS IP Multicast RPF

- Allow BGP routes to be used for RPF computation

  ipv6 rpf use-bgp

- Some peers do not support multicast SAFI

  address-family ipv6

    neighbor 2001::2 translate-update ipv6 multicast unicast

    neighbor 2001::2 activate

    exit-address-family

  address-family ipv6 multicast

    neighbor 2001::2 activate

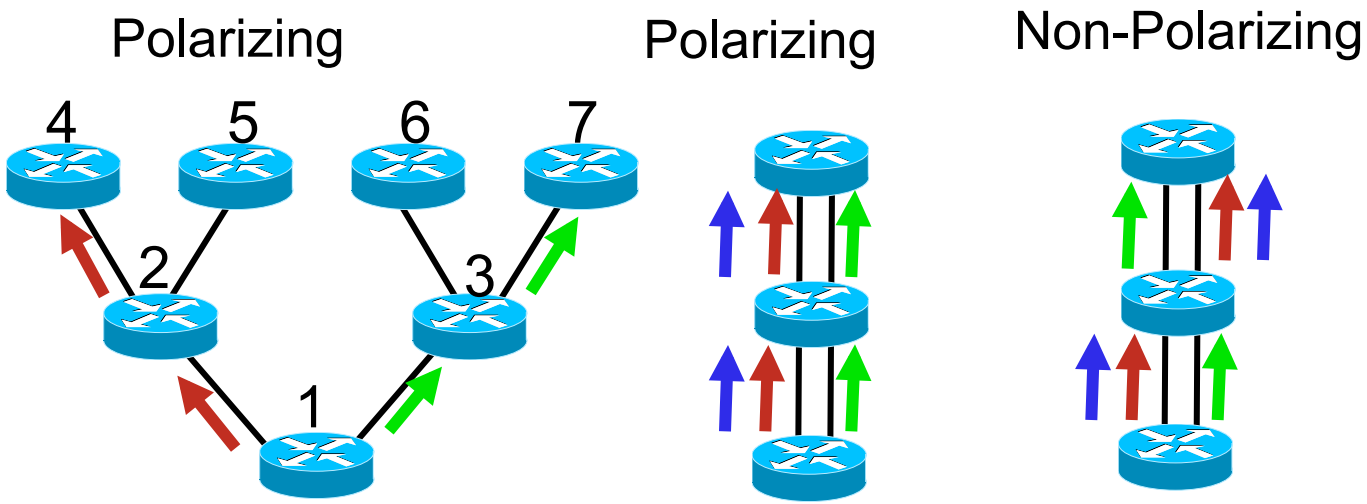
Cisco IOS IP Multicast RPF

- ECMP – Equal Cost Multipath
  - Polarizing / *planable* in IPv4:

    i = (hash(S) % n)
  - Non-polarizing in IPv6:

    i = i | max(hash(S, Nbr-i))
    - Stable in case of link failure. Uses PIMv2 hash function.
    - See *ftp://ftpeng.cisco.com/ipmulticast/config-notes/multipath.txt*

Given *1..n* (eg: 2) ECMPs, if all routers select the same neighbor *I* for a source S, then polarization may happen: A rtr2 will only be joined to by rtr1 for Sources that it's own ECMP would RPF to rtr4, but never to rtr5!



Multicast RPF Selection for different source addresses

Cisco IOS IP Multicast RPF

- Improved static (m)routes

  Cisco IOS IPv4

  ip route *dest mask* [nexthop | *interface*] [*distance*]

  ip mroute *dest mask* [nexthop | *interface*] [*distance*]

  Cisco IOS IPv6

  ipv6 route *dest mask* [*nexthop* | *interface*]
  [*distance*] [*mdistance* | unicast | multicast]

  - Same behavior as ipv4 static routes (except for new options)
  - Support equal-cost multipath mroutes
  - Support unicast only static routes
  - Equal or less than in IPv4 config lines are required

PIM-SM

PIM-SM (RFC 2362) – Overview

- Supports both source and shared trees

  - Assume no hosts wants multicast taffic unless they specifcally ask for it.

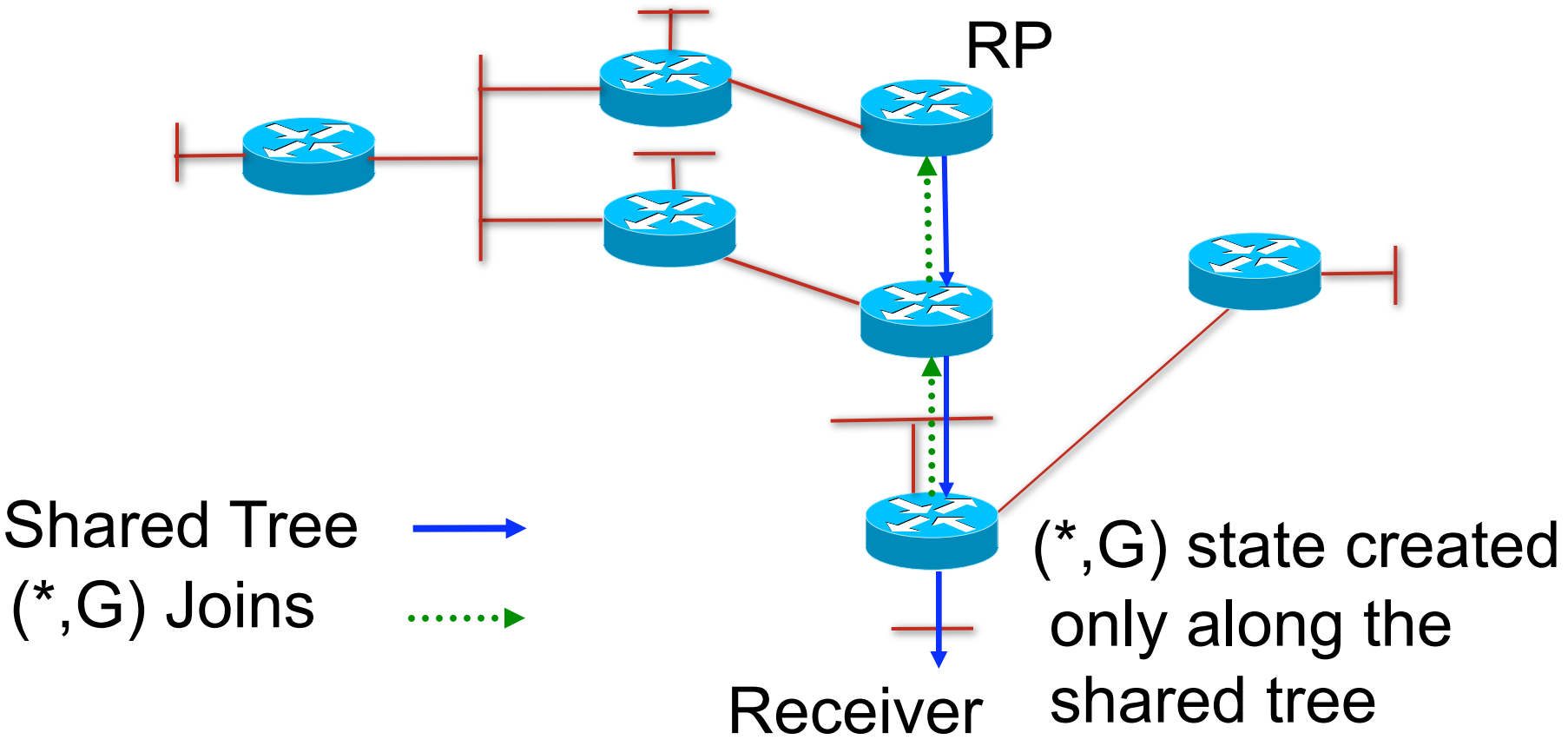
- Uses a Rendezvous point (RP)
  - Senders and receivers “rendezvous” at this point to learn of each other existence.
  - Senders are “registered” with their RP by their first-hop router.
  - Receivers are joined to Shared tree (rooted at the RP) by their local Designated Router (DR)
  - Data flows down the shared tree and goes only to places that needs the data from the sources
  - Last hop routers can join source tree if data rate warrants by sending joins to the source

PIM-SM (RFC 2362) – Overview

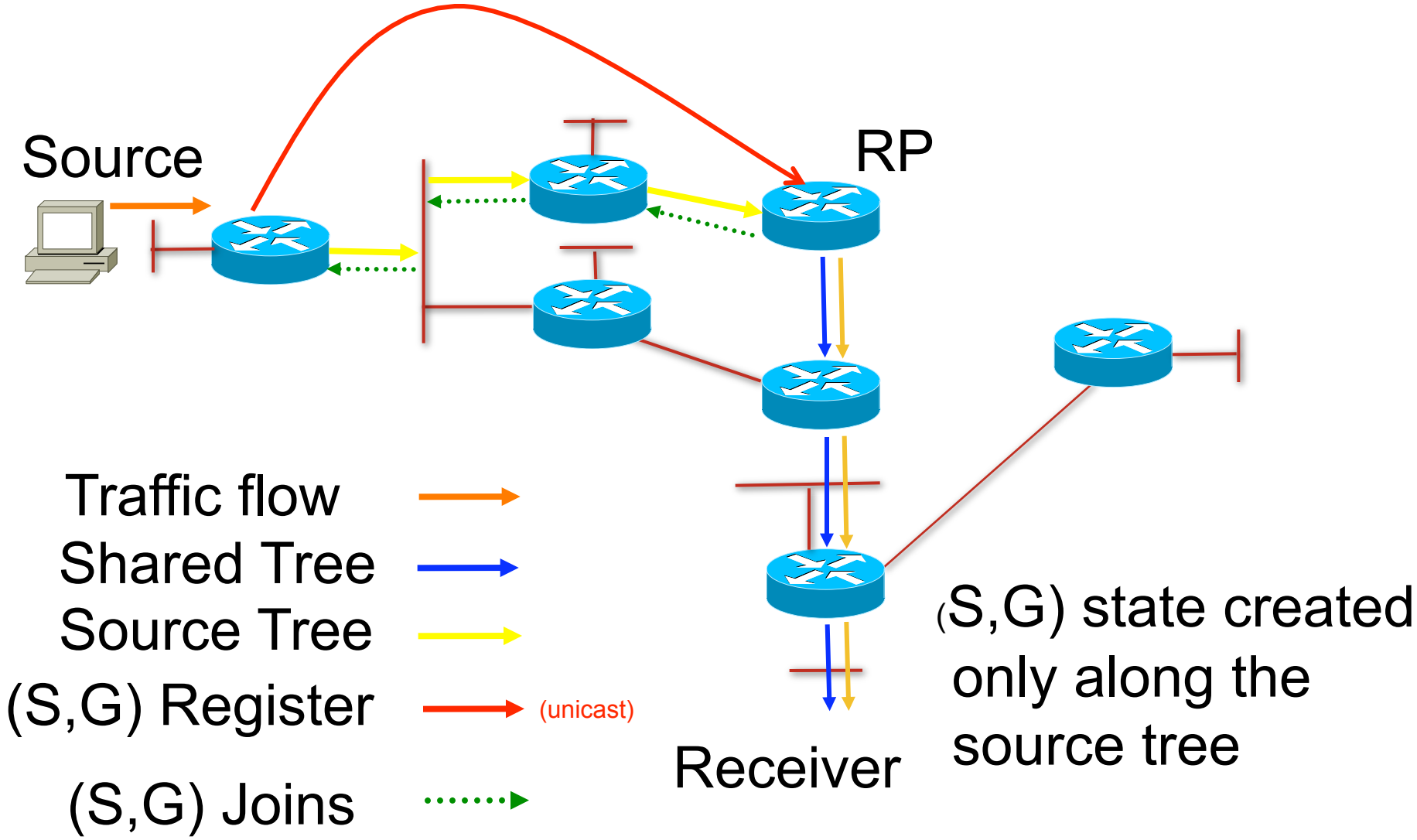
- Only one RP is chosen for a particular group
- RP statically configured or dynamically learned

  (Auto-RP, PIM v2 Candidate RP advertisements)
- Data forwarded based on the source state (S,G) if it exist, otherwise use the shared state (*,G)

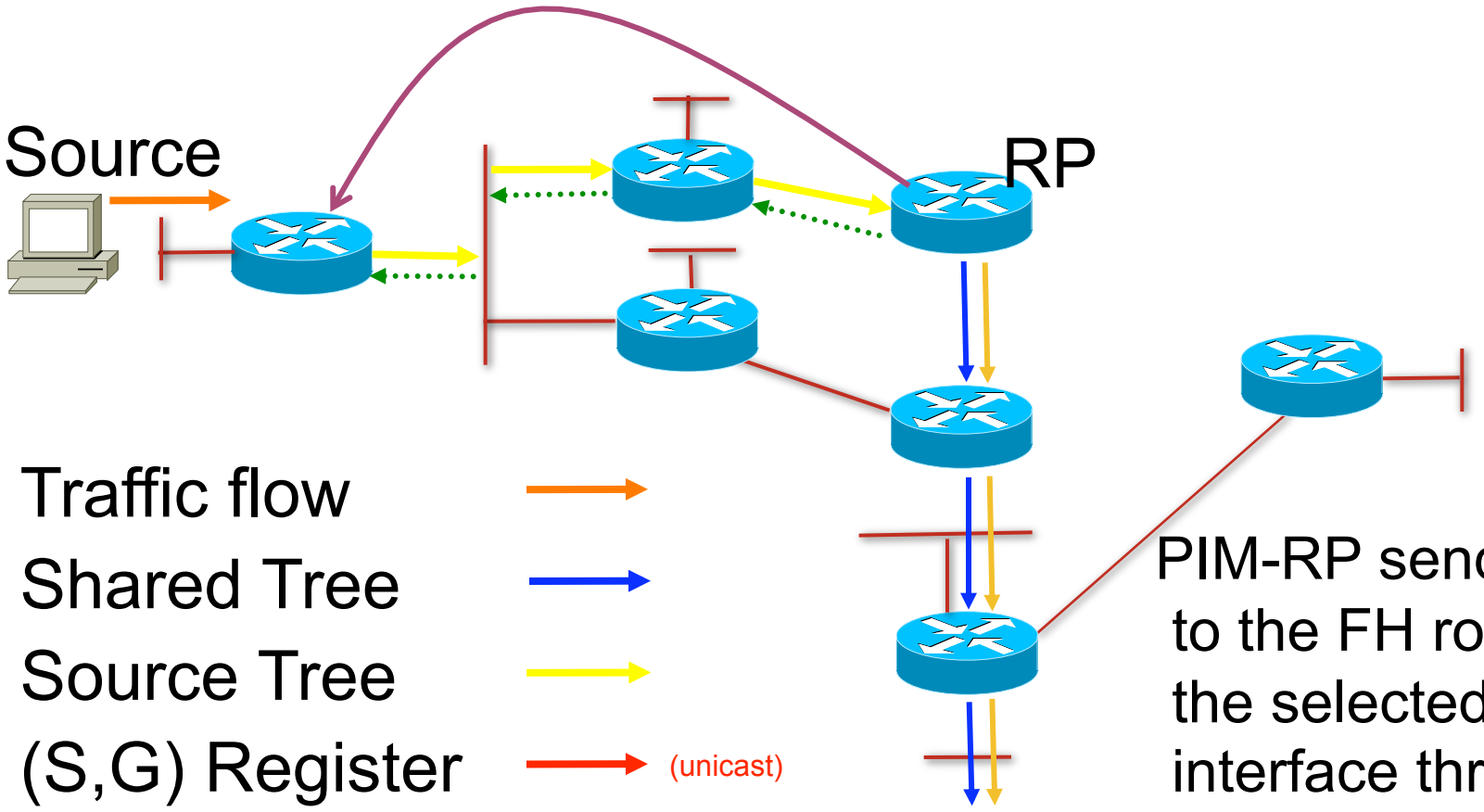
PIM-SM Shared Tree Join



PIM-SM Sender Registration



PIM-SM RP Joins After Sender Registration



Traffic flow

Shared Tree

Source Tree

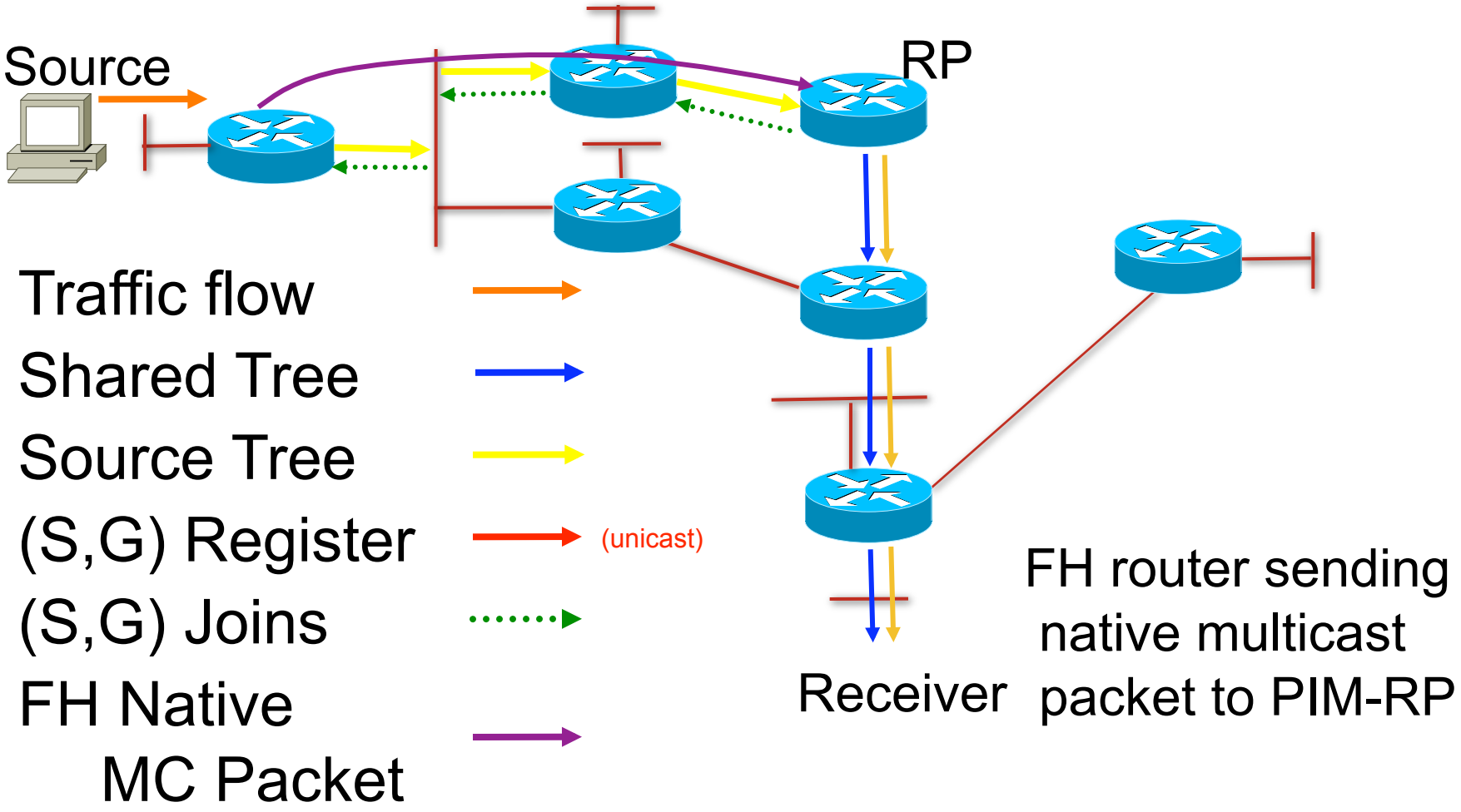
(S,G) Register (unicast)

(S,G) Joins

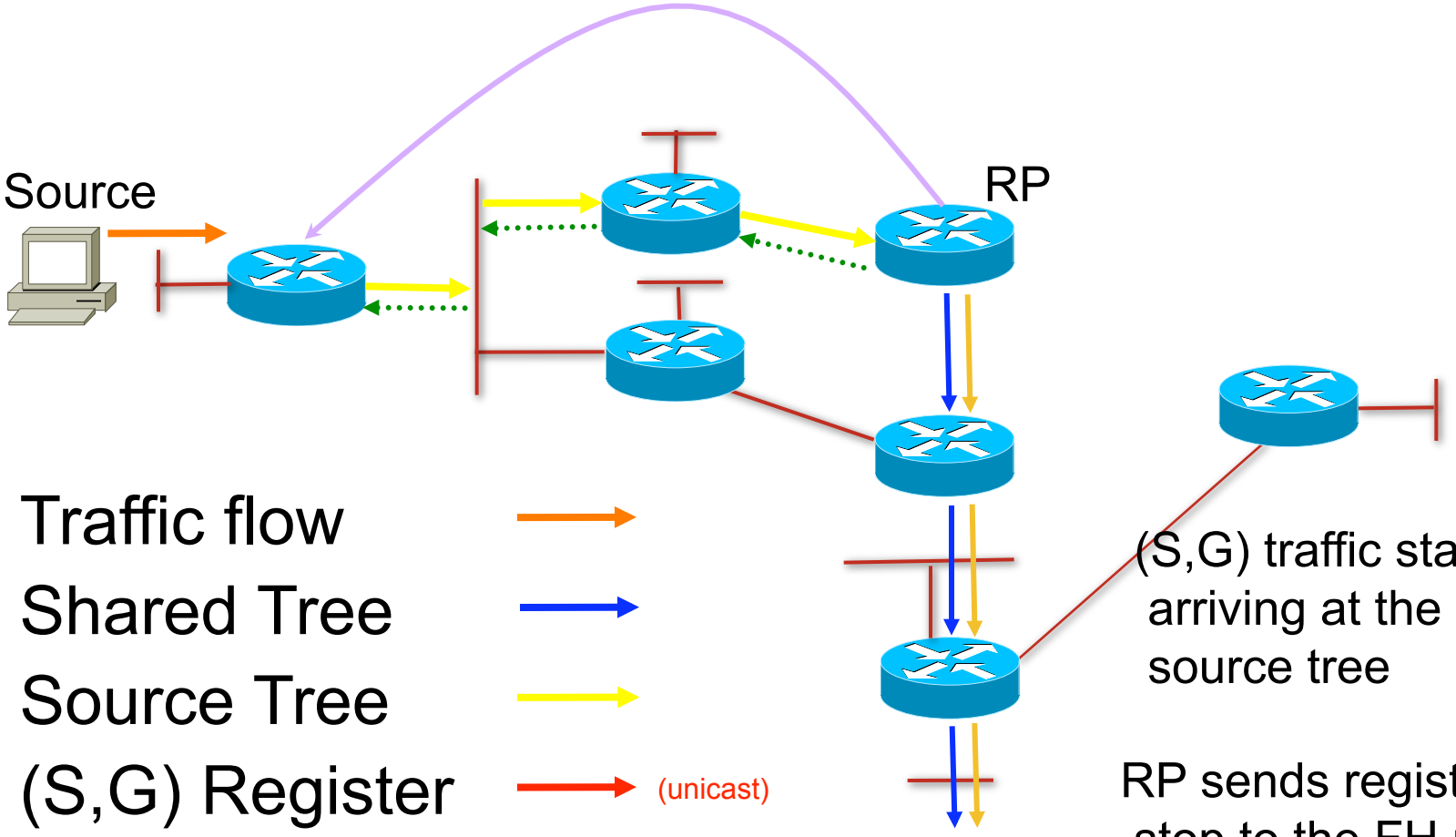
RP Join

PIM-RP sends Joins to the FH router on the selected interface through which source is reachable.

PIM-SM Sender Sending Native MC Packet



PIM-SM RP sends Registration Stop to FH



Traffic flow

Shared Tree

Source Tree

(S,G) Register (unicast)

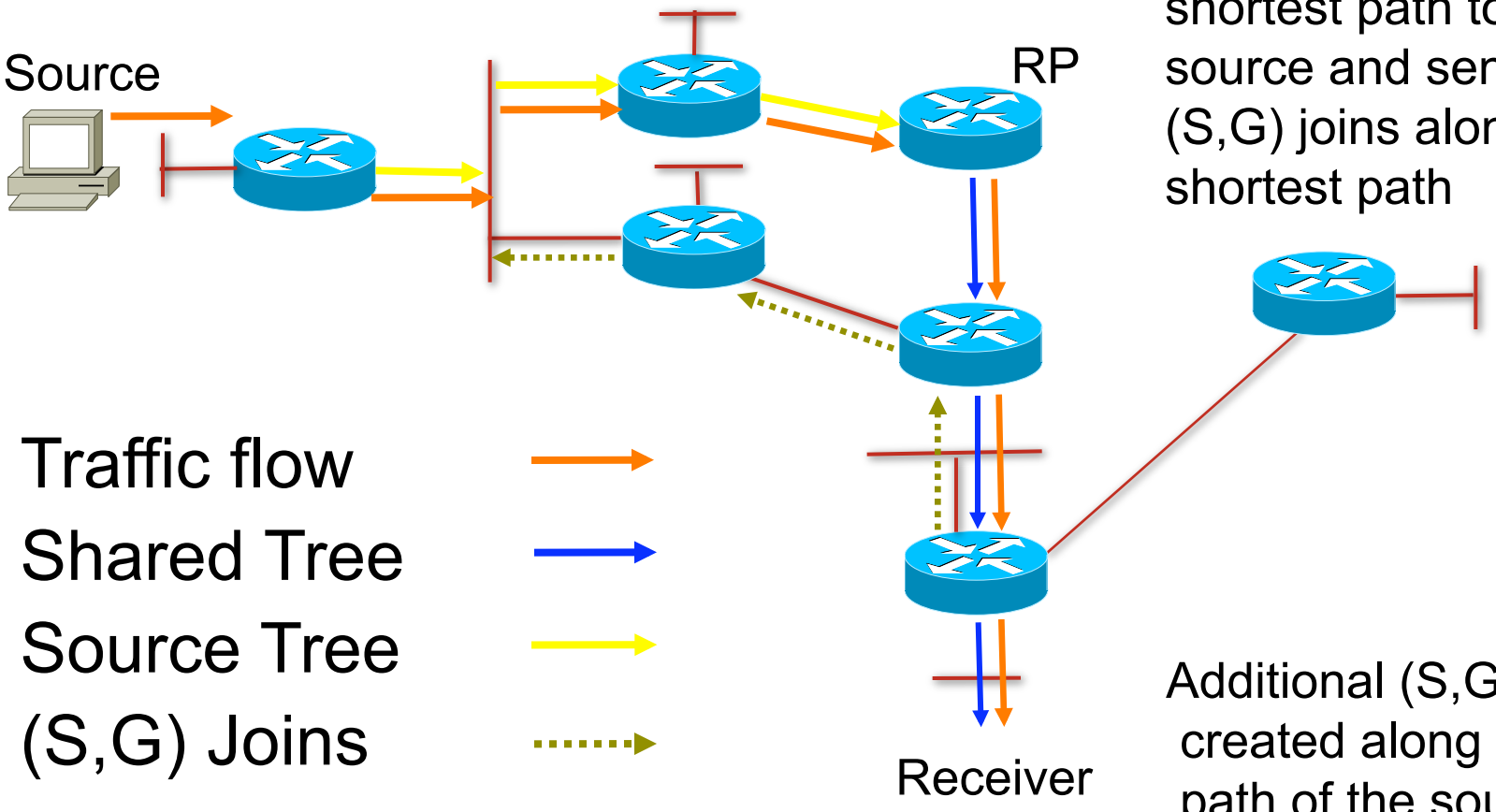
(S,G) Joins

(S,G) Joins

(S,G) traffic starts arriving at the RP via source tree

RP sends registration stop to the FH router to stop the registration process.

PIM-SM Sender Registration



LH router after CEF look up picks up shortest path to the source and sends (S,G) joins along the shortest path

Traffic flow

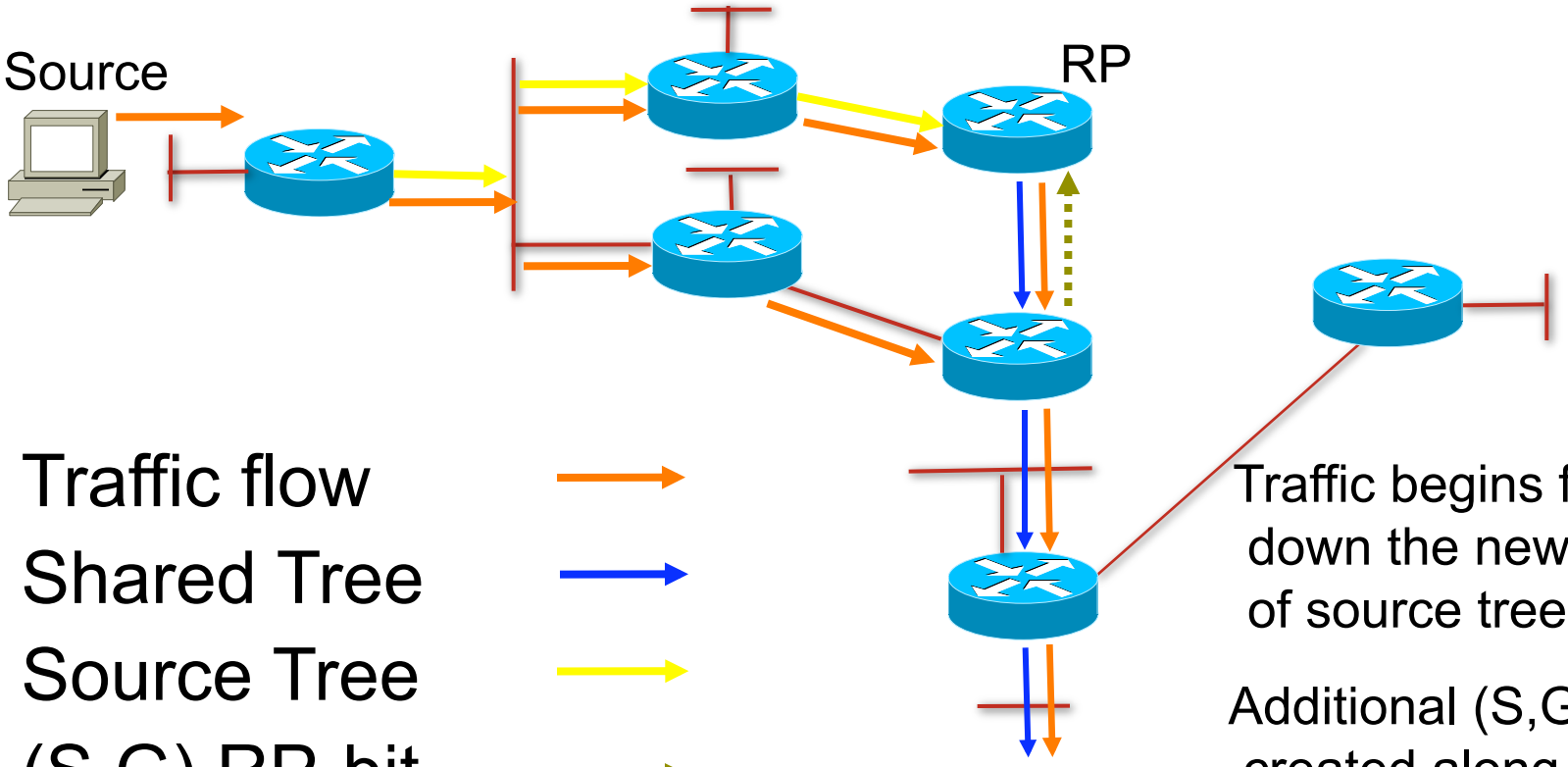
Shared Tree

Source Tree

(S,G) Joins

Additional (S,G) state is created along the new path of the source tree

PIM-SM Sender Registration



Traffic flow

Shared Tree

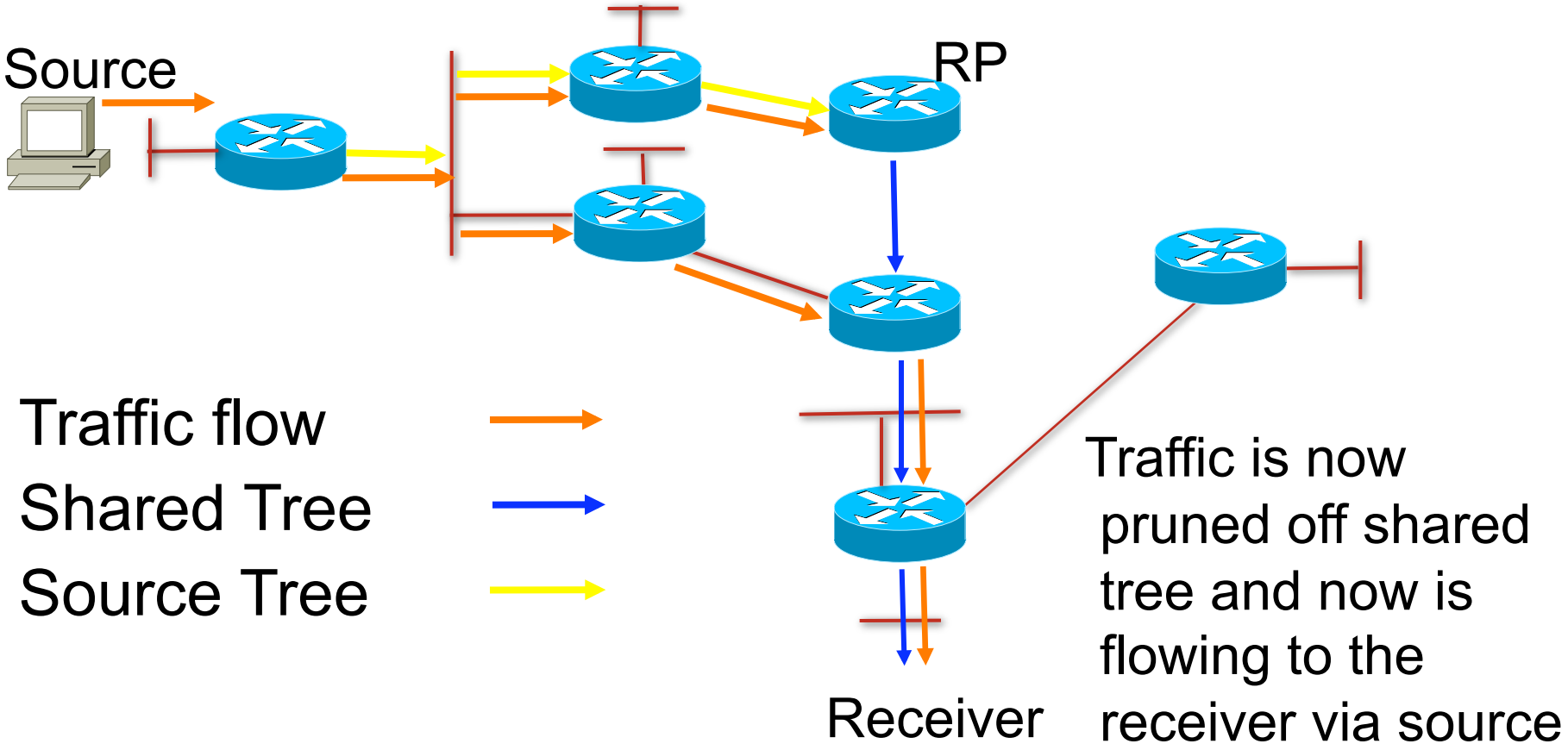
Source Tree

(S,G) RP-bit Prune

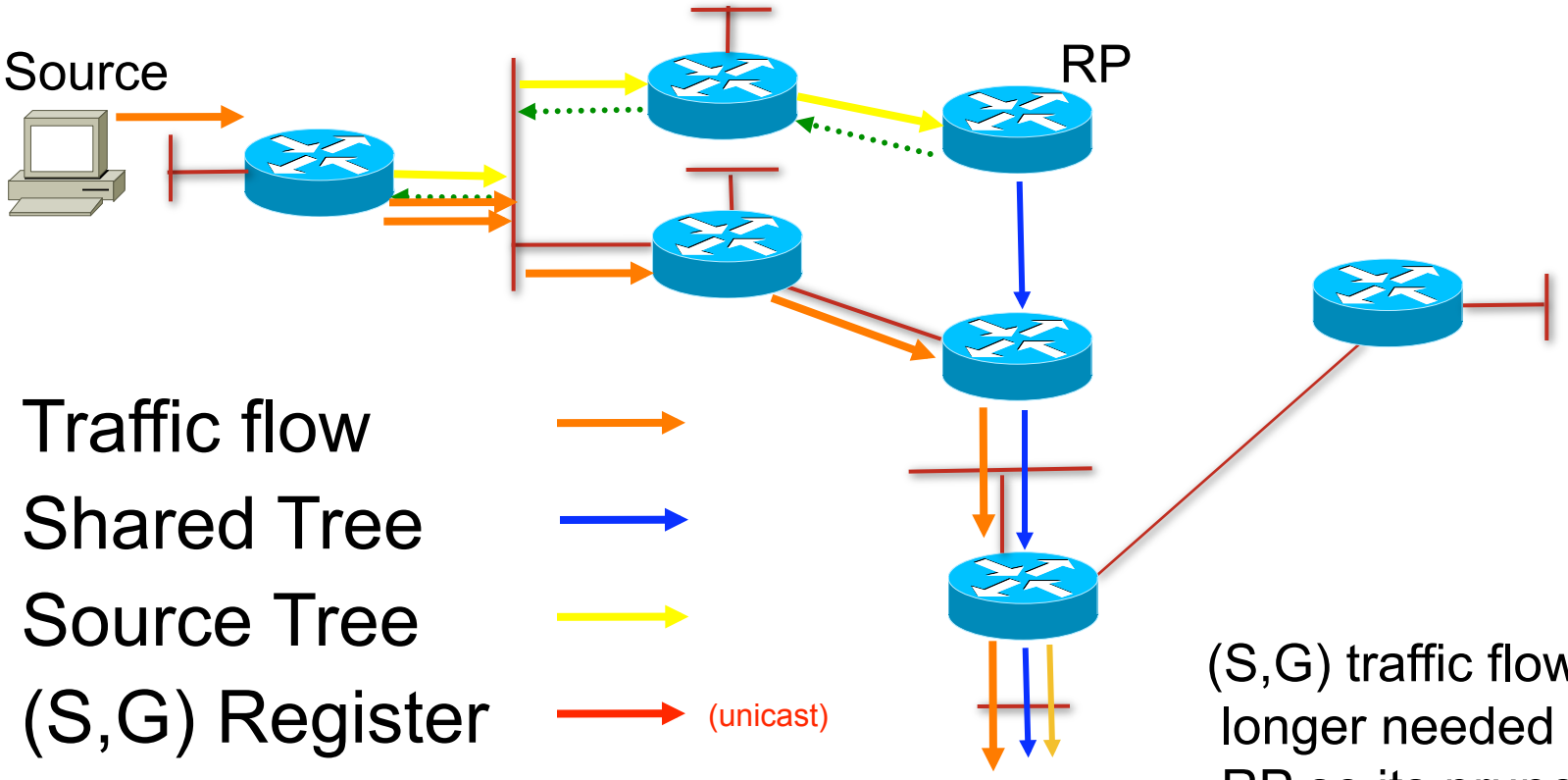
Traffic begins flowing down the new branch of source tree

Additional (S,G) state is created along the shared tree to prune off (S,G) traffic

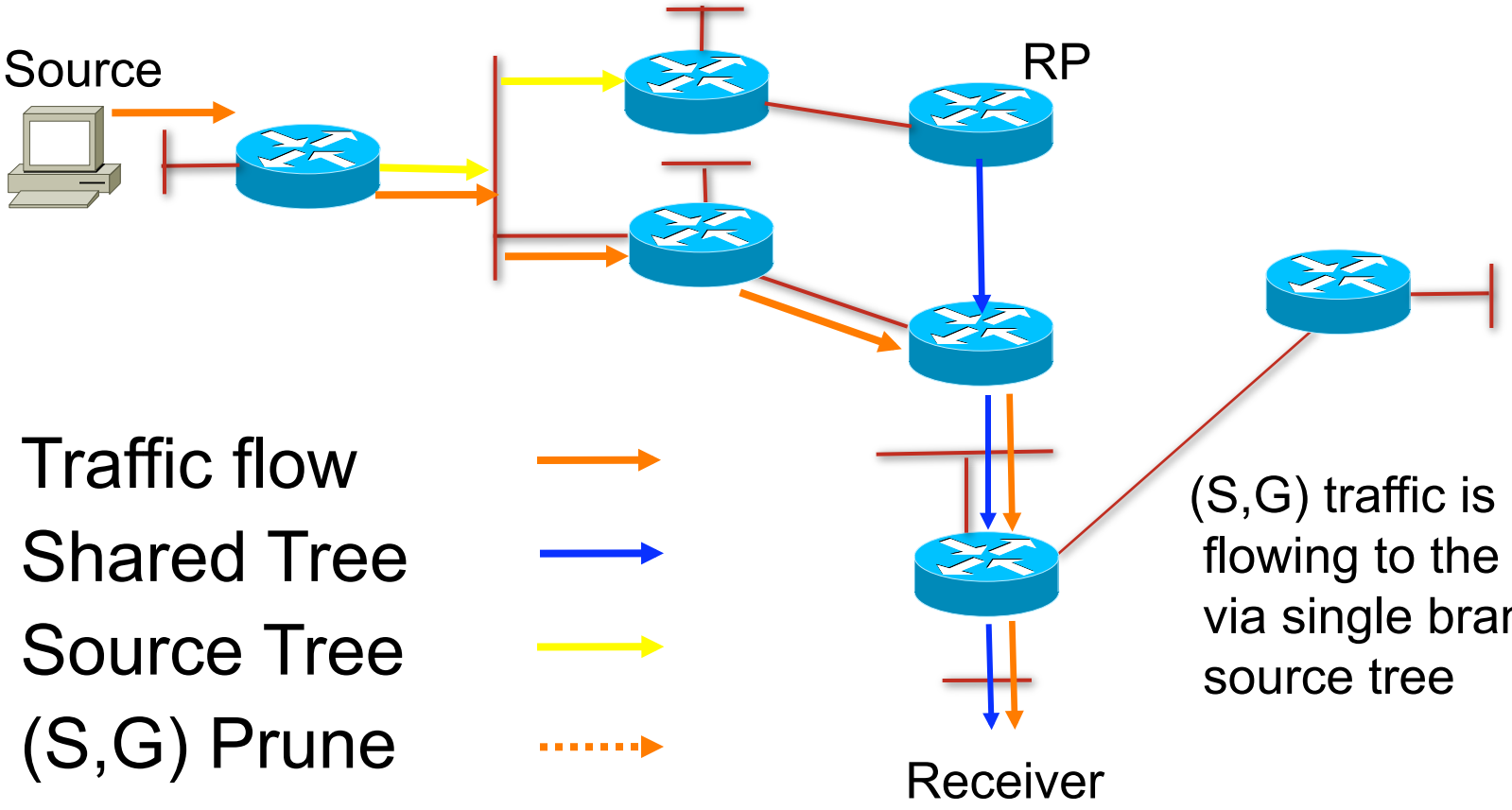
PIM-SM Sender Registration



PIM-SM: RP sends Prune for the Group



PIM-SM : Entire Traffic Switched to Source Tree



PIM-SSM

PIM-SSM – Overview

- Assume a One-To-Many Multicast Model
  - Example : Video/Audio broadcasts, Stock Market data
- Why does PIM-SM need a Shared Tree?
  - So the hosts and First hop routers knows who the active source is for the group.
- What if this was already known?
  - Hosts could use IGMPv3 to signal *exactly* which (S,G) SPT to join.
  - The shared tree and RP wouldn't be necessary.
  - Different sources could share the same Group address and not interfere with each other.
- Result: Source Specific Multicast (SSM)

PIM-SSM – Overview

- Simplifies multicast deployment and eleminates the concept of RP and dependence of MSDP for finding sources.
- Optimised and Reduced Latency for Multicast forwarding in case of one to Many applications
  - So the hosts and First hop routers knows who the active source is for the group.
- Simplifies Address allocation problem for global single source group
- Allows immediate use of shortest forwarding path to a specific source, without need to create shared tree.

PIM-SSM – Reserved Address

- Group range is reserved as per RFC 3306
- Range is allocated as soon as IPv6 multicast is enabled
- SSM group ranges are automatically defined

FF3x::/96

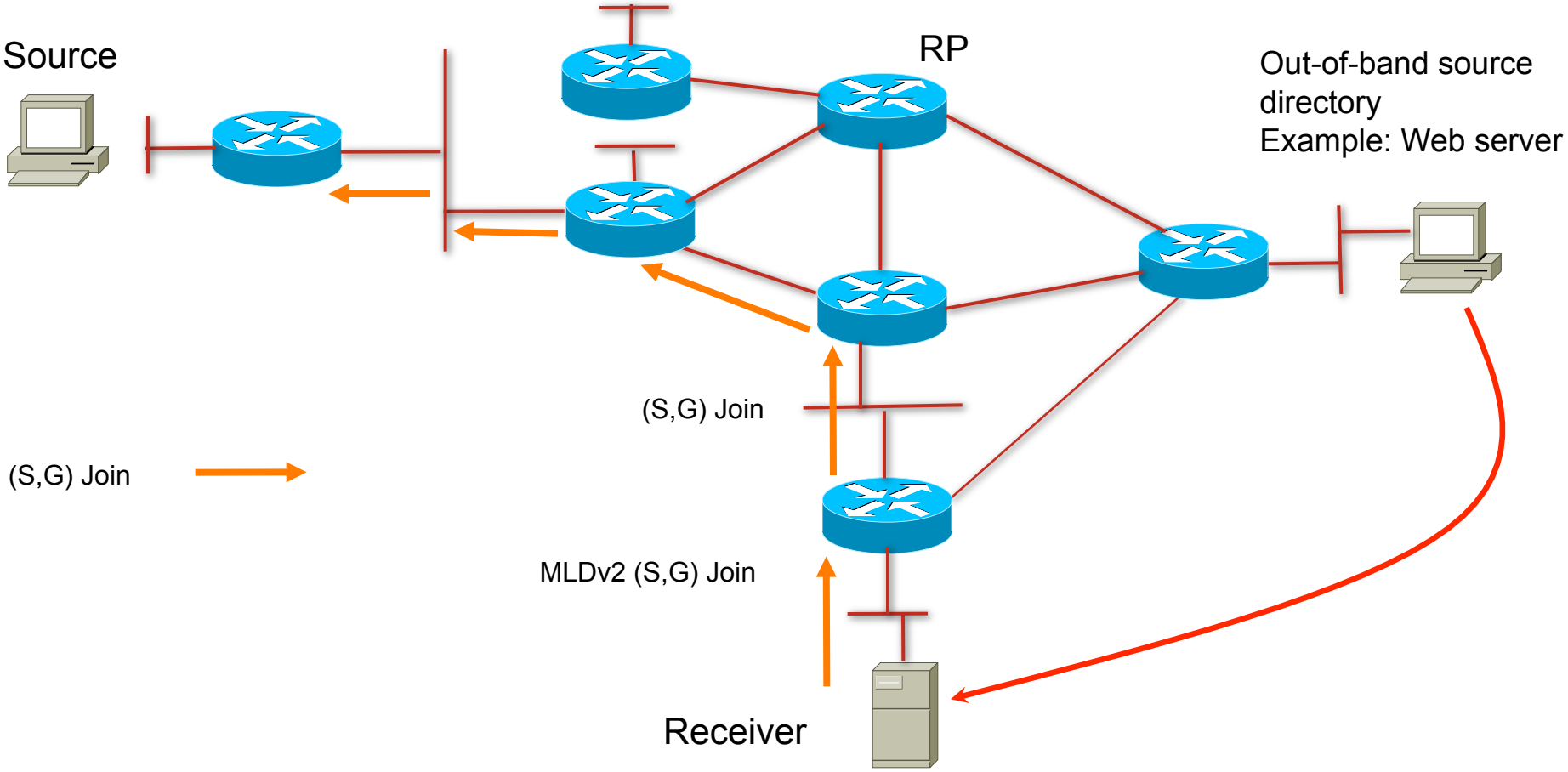
```
Reserved ipv6 pim range-list

  FF33::/32
  FF34::/32
  FF35::/32
  FF36::/32
  FF37::/32
  FF38::/32
  FF39::/32
  FF3A::/32
  FF3B::/32
  FF3C::/32
  FF3D::/32
  FF3E::/32
  FF3F::/32
```

PIM-SSM - Concept

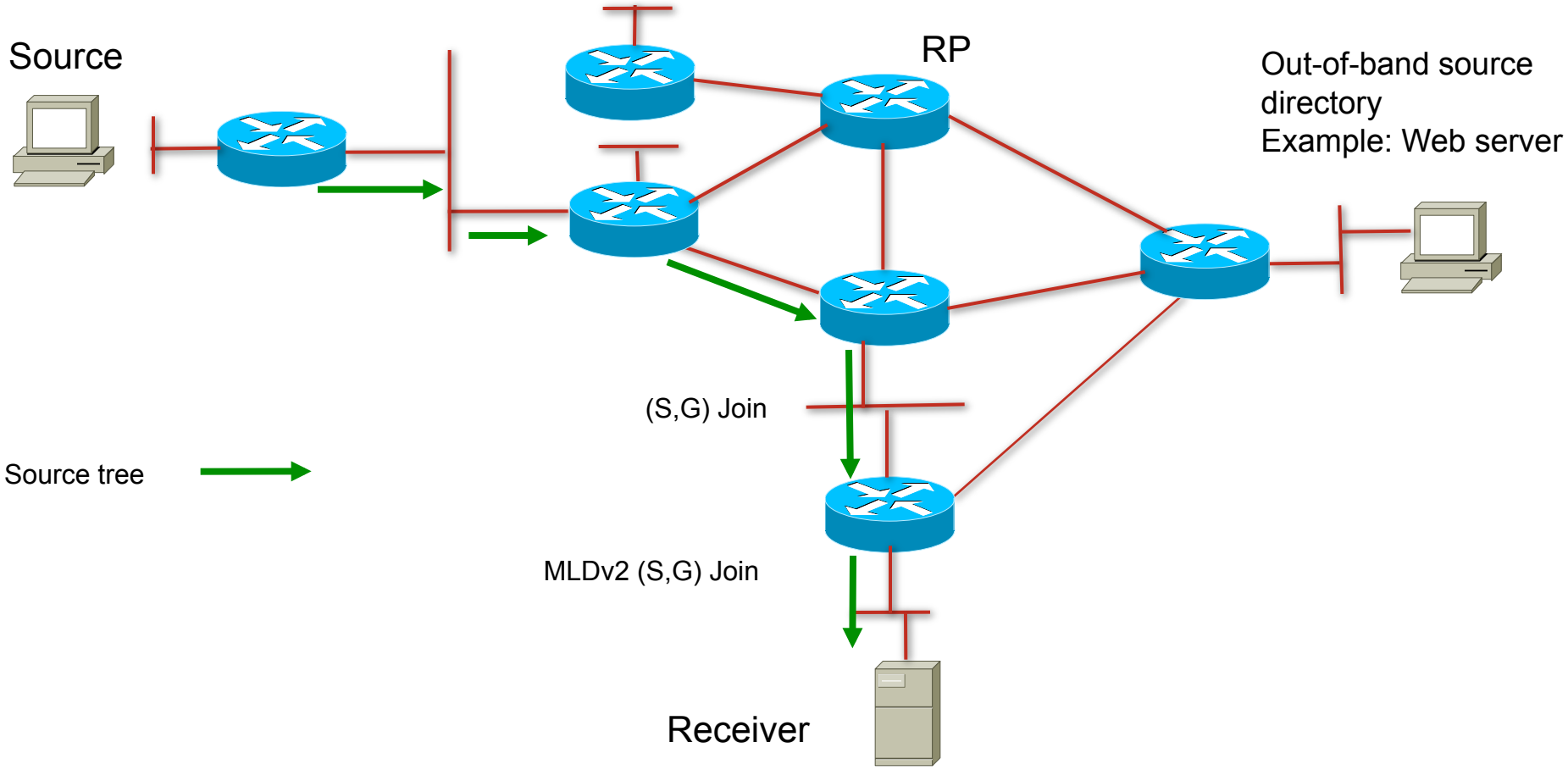
MLDv2 (S,G)
Router sends PIM (S,G)
joins directly towards the
source

Receiver learns of source, group/ports
Receiver sends MLDv2 (S,G)
Router sends PIM (S,G) joins directly towards the source



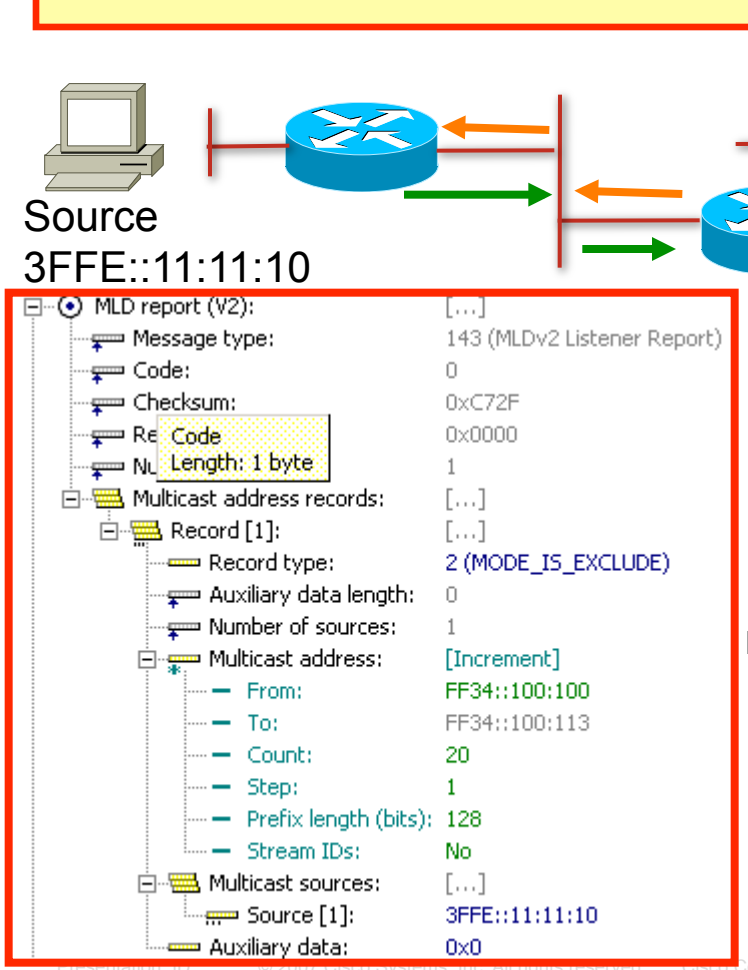
PIM-SSM - Concept

Result: Shortest Path Tree Rooted @ the Source , with no Shared Tree.



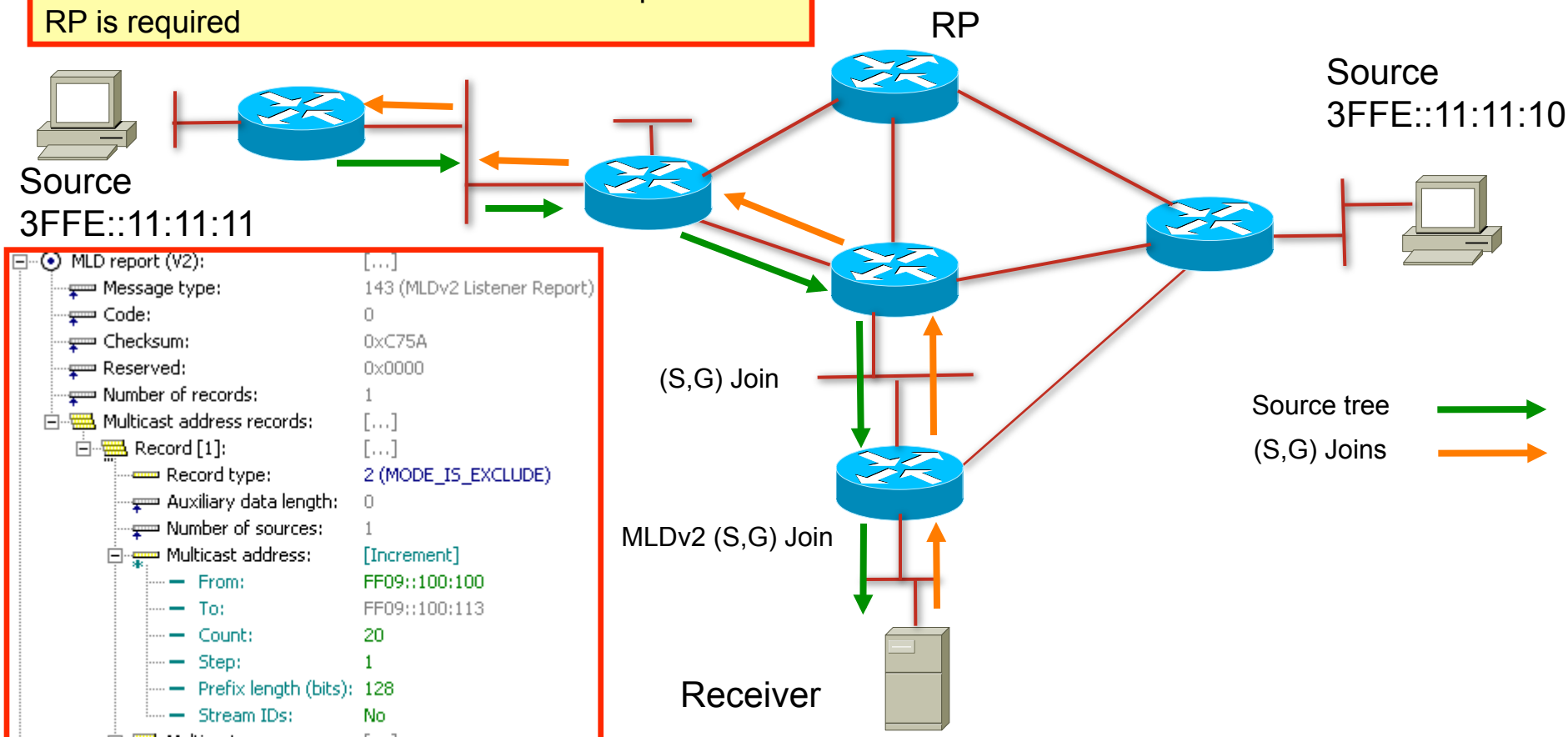
PIM-SSM – Record Type Include

Reserved address FF3x::x/96 should be sent with
INCLUDE_OPTION specifying SOURCES
No need for RP in this case



PIM-SSM – Record type Exclude

EXCLUDE_OPTION option can be used with
groups other then reserved
Receive traffic from the source other then specified
RP is required



MLD report (V2):	[...]
Message type:	143 (MLDv2 Listener Report)
Code:	0
Checksum:	0xC75A
Reserved:	0x0000
Number of records:	1
Multicast address records:	[...]
Record [1]:	[...]
Record type:	2 (MODE_IS_EXCLUDE)
Auxiliary data length:	0
Number of sources:	1
Multicast address:	[Increment]
From:	FF09::100:100
To:	FF09::100:113
Count:	20
Step:	1
Prefix length (bits):	128
Stream IDs:	No
Multicast sources:	[...]
Source [1]:	3FFE::11:11:10
Auxiliary data:	0x0

PIM-SSM – Evaluation

- Ideal for applications with one source sending traffic to many receivers
- Solves multicast address allocation problems.
  - Flows differentiated by both source and group
    - not just by group
- Content providrs can use same group ranges
  - Since each (S,G) flow is different
- Helps prevent certain DoS attacks
  - “Bogus” source traffic:
    - Can’t consume network bandwidth
    - Not received by host applications

Rendezvous Point (RP) Operation

Rendezvous Point (RP) Deployment Types

- Static RP
  - –For PIM-SM
  - –Provides group-to-RP mapping, no RP redundancy
- Boot Strap Router (BSR)
  - Provides group-to-RP mapping AND RP redundancy
- Embedded-RP
  - –Easy to deploy
  - –Group-to-RP mapping only, no RP redundancy
  - –PIM-SM only (today), no Bidir-PIM
- RP redundancy options for static/embedded-RP
  - –MSDP mesh-group, PIM/Anycast ?, Prefixlength/Anycast
  - –Could also be combined with BSR for faster convergence
- AutoRP
  - –No option yet – IPv4 only

IPv6 Multicast Static Rendezvous Point

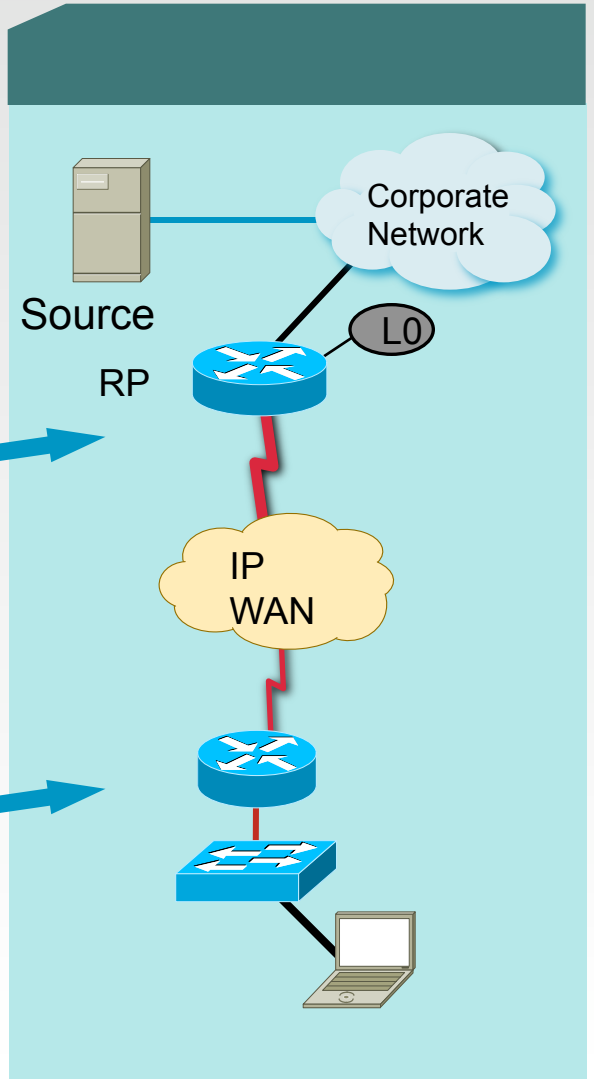
- PIM is enabled when multicast-routing is enabled on interface

```
multicast-routing address-family ipv6

interface Loopback0
 description IPV6 IPmc RP
ipv6 address 3FFE:C15:C003:110A::1/64


router pim address-family ipv6
rp-address 3FFE:C15:C003:110A::1
```

```
ipv6 multicast-routing
!
ipv6 pim rp-address 3FFE:C15:C003:110A::1
```



IPv6 Multicast BSR Overview

- A single Bootstrap Router (BSR) is elected
  - Multiple Candidate BSR‘s (C-BSR) can be configured
    - Provides backup in case currently elected BSR fails.
  - C-RP‘s send C-RP announcements to BSR Solves multicast address allocation problems.
    - C-RP‘s announcements are sent via unicast
    - BSR stores All C-RP announcements in RP-set
  - BSR periodically sends BSR messages to all routers
    - BSR messages contains entire RP-set and and IP address of the BSR
    - Messages are flooded through out the network away from the BSR
  - All routers selects the RP from the RP-set
    - All the routers use same selection algorithm ; select the same RP
- BSR can’t be used with Admin-scoping

IPv6 Multicast BSR Fundamentals

BSR election mechanism

C-BSR's:

- Begin in Candidate-BSR state
  - BSR timeout timer started (150 seconds)
  - If high priority (preferred) BSR message received
    - Restart timer and forward BSR message
    - Copy info to local Group-to-RP mapping cache
    - Otherwise discard BSR message
  - If timer expires transition to Elected-BSR state
- While in the Elected BSR state
  - Periodically originates own BSR messages
    - Include local Group-to-RP mapping cache in message
  - Return to Candidate-BSR state if preferred (High Priority)
    - BSR message is received

IPv6 Multicast BSR Fundamentals

Non C-BSR's (i.e., all other routers)

- Start in Accept-Any state
  - Accepts first BSR message received
  - Save BSR info and forward BSR message
  - Transition to Accept-Preferred state
- While in Accept-Preferred state
  - Starts BSR-timeout timer
  - Only accept and forward preferred BSR message
    - (i.e., BSR messages with priority > current BSR Priority)
  - Otherwise discard BSR messages
  - Return to Accept-Any state if timer expires

IPv6 Multicast BSR Fundamentals

All PIMv2 Routers

- Receive BSR messages
  - Stores in local Group-to-RP Mapping Cache
  - Information use to determine active BSR address
- Selects RP using Hash algorithm
  - Selected from local Group-to-RP Mapping Cache
  - All routers select same RP using same algorithm
  - Permits RP-load balancing across group ranges

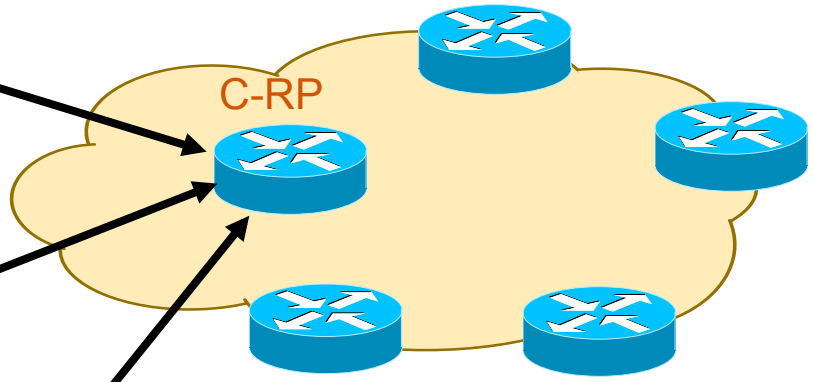
IPv6 Multicast Boot Strap Router Mechanism

Unicast PIMv2 C-RP message to BSR
Learns IP address of BSR from BSR messages
Sent every RP-announce-interval (60s)

C-RP messages contains:
Group range (Default: xxxxx)
Candidate's RP address
Hold Time = 3 x <rp-announce-interval>

Configuration command

```
router pim address-family ipv6
bsr candidate-bsr 3FFE:C15:C003:1116::2
hash-mask-len 128 priority 10
```



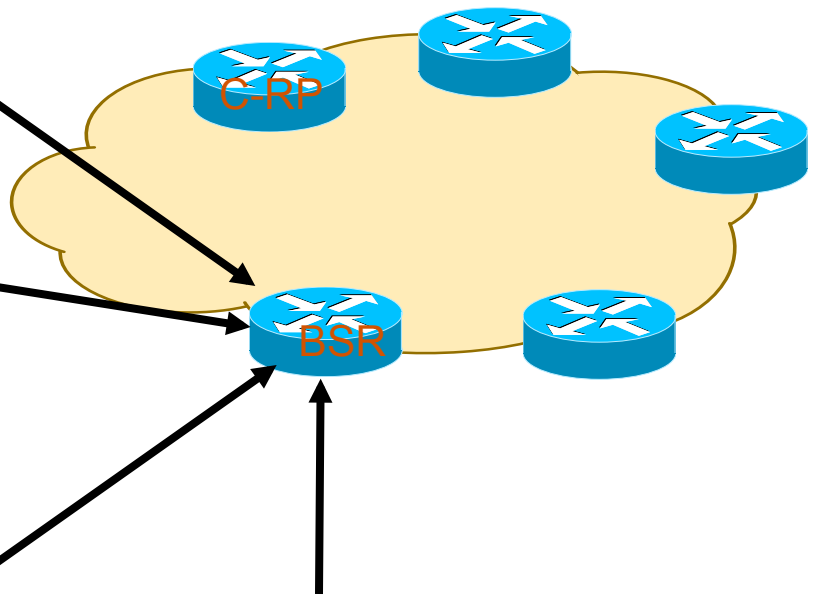
IPv6 Multicast Boot Strap Router Mechanism

Receive C-RP messages
Accepts and stores ALL C-RP messages
Stored in Group-to-RP Mapping cache w/holdtimes

Originates BSR messages
Multicast to ALL-PIM-Routers (224.0.0.13) group
Sent with TTL = 1
Sent out on all interfaces , Propagate hop-by-hop
Sent every 60 seconds or when changes detected

BSR messages contains:
Contents of BSR's Group-To-RP Mapping Cache
IP address of active BSR

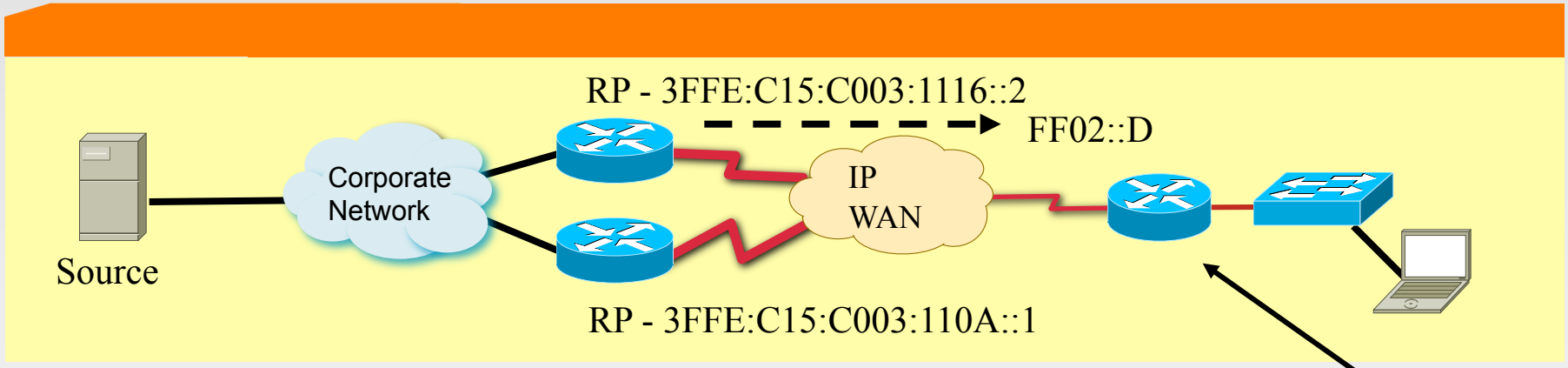
Election:
C-BSR with highest priority elected as BSR
C-BSR IP address used as a tie breaker (Highest IP wins)
The active BSR may be preempted
(New router with higher BSR priority can force election)



Configuration:

```
router pim address-family ipv6
 bsr candidate-bsr 3ffe::10 hash
-mask-len 128 priority 10
```

IPv6 Multicast PIM BSR – Non Rendezvous Points




```
IPv6 BSR: Received BSR message from FE80::230:F2FF:FE15:9C00 for
 3FFE:C15:C003:1116::2, priority 0 hash mask length 126
IPv6 BSR: Skipping interface FastEthernet0/0, no PIM neighbors found
IPv6 BSR: Skipping interface Serial0/2, incoming interface
IPv6 BSR: Received Group range FF00::/8, RP count 2 Fragment RP count2
IPv6 BSR: RP 3FFE:C15:C003:110A::1, Holdtime 150, Priority 192
IPv6 BSR: RP 3FFE:C15:C003:1116::2, Holdtime 150, Priority 192
```

```
BSR Election Information
  Scope Range List: ff00::/8
     BSR Address: 3FFE:C15:C003:1116::2
     Uptime: 01:51:17, BSR Priority: 0, Hash mask length: 126
     RPF: FE80::230:F2FF:FE15:9C00,Serial0/2
     BS Timer: 00:01:13
```

Embedded – Rendezvous Point Addressing Overview

- Why need Embedded RP ?
  - PIM-SM has no way of communicating information about Active Multicast Source to other multicast domains
  - MSDP has deliberately not been specified for IPv6.
    - » ASM model is rendered unusable
  - SSM requires source to be known by receivers
    - » Applications are not SSM aware

Embedded – Rendezvous Point Addressing Overview

Based on RFC 3956 and subset of 3306 IPv6 unicast-prefix based ASM

Group addresses with special encoding rules

- Group addresses carries Randezvous Point address for the group
- For each unicast prefix owned
  - » 15 Rendezvous Point for each of 14 Multicast scopes (210 in total)
  - » 2^32 multicast groups assigned to each Rendezvous Point

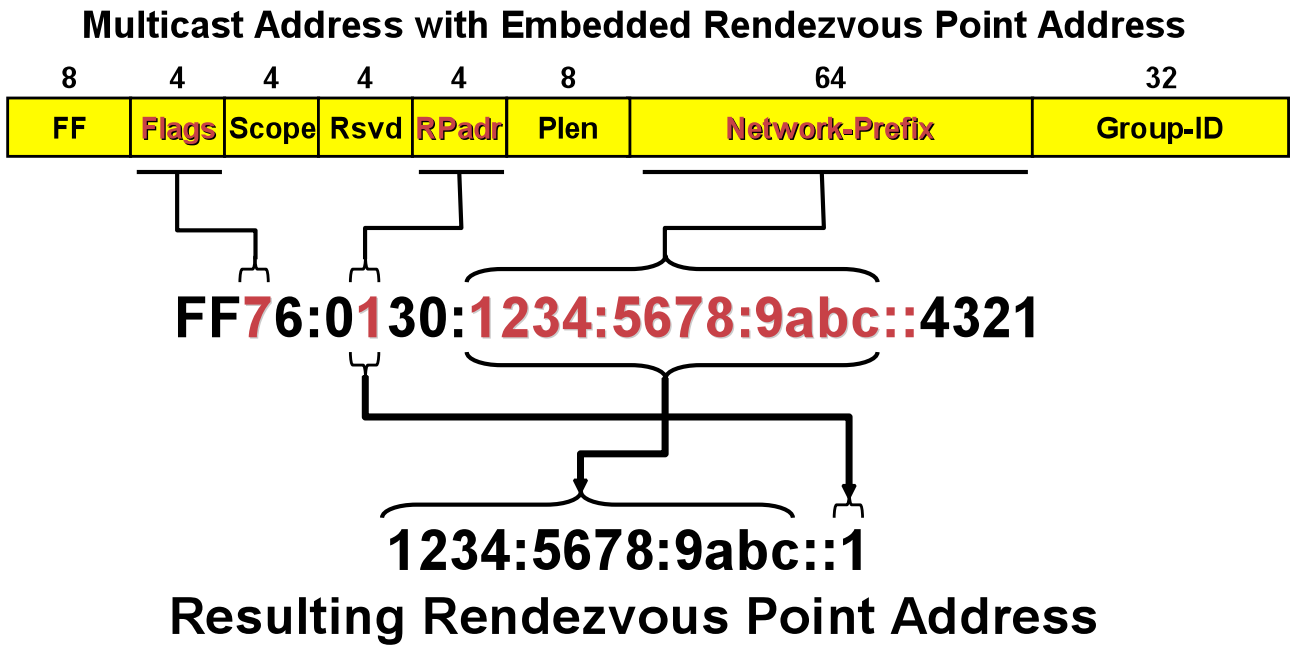
Embedded – Rendezvous Point Addressing Overview

- Flag: 0 RPT
  - R = 1 → Embedded-RP-Address bit
  - P =1 → Prefix based Address
  - T =1 → Temporary Address

FF70::/12

- Scope : Node, Link, Site, Organisation or Global (E) etc ..
- Rsvd: 0
- Plen : Number of bits used from prefix
- Prefix : Bits not coming from prefix should be zero

Embedded – Rendezvous Point Address Example

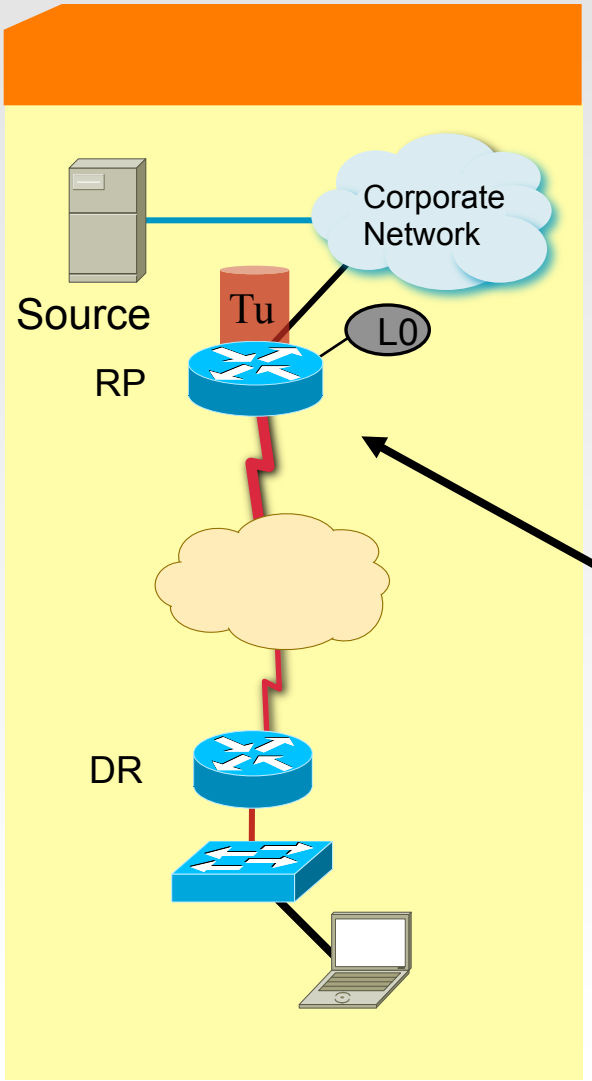


- Rendezvous Point address = network prefix = Rpad
- Sixteen Rendezvous Point addresses per network prefix

Embedded – Rendezvous Point Addressing Benefit

- PIM-SM protocol operations with embedded-Rendezvous Point:
  - No change in PIM-SM protocol operations
    Just an automatic replacement to static Rendezvous Point configuration
  - Can replace BSR for Group-to-RP mapping
  - Method requires large IPv6 addresses -
    No equivalent possible in IPv4
  - Intradomain transition into embedded-Rendezvous Point is easy:
    - Non-supporting routers simply need to be configured statically or via BSR for the embedded-Rendezvous Points

Embedded – Rendezvous Point Configuration Example



- Rendezvous Point used as an Embedded-Rendezvous Point needs to be configured with address/group range
- All other non-Rendezvous Point routers do not require any special configuration

```
ipv6 pim rp-address 3FFE:C15:C003:111D::1 ERP
!
ipv6 access-list ERP
permit ipv6 any FF7E:140:3FFE:C15:C003:111D::/96
```

Embedded Rendezvous Point – Does it work?

```
branch#show ipv6 pim group

FF7E:140:3FFE:C15:C003:111D::/96*
     RP       : 3FFE:C15:C003:111D::1
     Protocol: SM
     Client   : Embedded
     Groups   : 1
     Info     : RPF: Se0/0.1,FE80::210:7FF:FEDD:40
```

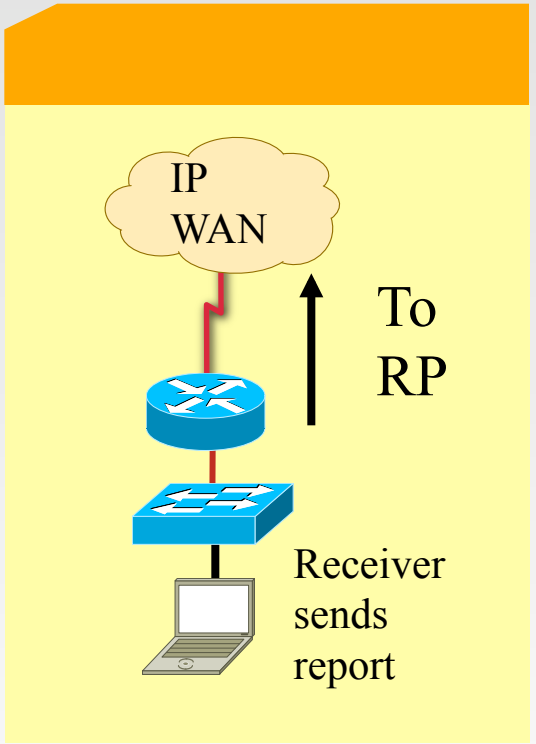
```
branch#show ipv6 mroute active

Active IPv6 Multicast Sources - sending >= 4 kbps
Group: FF75:140:3FFE:C15:C003:111D:0:1112
  Source: 3FFE:C15:C003:1109::2
    Rate: 21 pps/122 kbps(1sec), 124 kbps(last 100 sec)
```

```
branch#show ipv6 pim range | include Embedded

Embedded SM RP: 3FFE:C15:C003:111D::1 Exp: never Learnt from : ::

  FF7E:140:3FFE:C15:C003:111D::/96 Up: 00:00:24
```



Embedded Rendezvous Point – Interdomain Concerns and Answers

- In current IPv4 Multicast:
  - Multicast group is served by typically one set of Rendezvous Points per domain and these Rendezvous Points are interconnected by MSDP
- With IPv6 embedded-Rendezvous Point:
  - There is just one set of anycast-Rendezvous Points globally for a group
- Scalability … flat virtual topology
  - Similar to SSM (with just one added Rendezvous Point – simple != scalable?)
  - No MSDP scalability / reliability / administration concerns
  - Almost arbitrary number of Rendezvous Points that can be used:
  - Each Rendezvous Point may need to serve only very few groups
- Third party Rendezvous Point dependency:
  - Yes – for totally anarchic applications that must not have a single identifiable point of origin
  - No – the majority of IP multicast applications will (for example: web applications) have an identifiable owner. This owner must take care of using an appropriate Rendezvous Point under his control

Appendix

Multicast Applications

- Microsoft Windows Media Server/Player (9 & 10)

  http://www.microsoft.com/windows/windowsmedia/default.aspx

- VideoLAN - www.videolan.org
- DVTS (Digital Video Transport System)

  http://www.sfc.wide.ad.jp/DVTS/

  http://www.dvts.jp/en/dvts.html

- Internet Radio Stations over IPv6

  http://www.ipv6.ecs.soton.ac.uk/virginradio/

  Supported on iTunes 4.5, Windows Media Player, XMMS 1.2.8, etc…

- Many more applications…Google is your friend :-)

RP redundancy

RP Redundancy
Potential RP Redundancy Alternatives

- MSDPv6 ?

  Perfectly well suited to support Anycast-RP (one mesh-group)

  Complex protocol—only a small subset of MSDP needed for Anycast-RP

- Anycast-RP Using PIM (RFC4610)

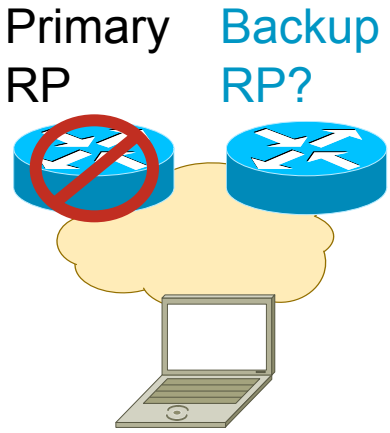
  Most simple protocol doing exactly what MSDP needs to do in one mesh-group: PIM-SM register messages are unicast forwarded between the redundant RPs

  (Almost) no operational differences to MSDP for Anycast-RP

- Prioritycast-RP

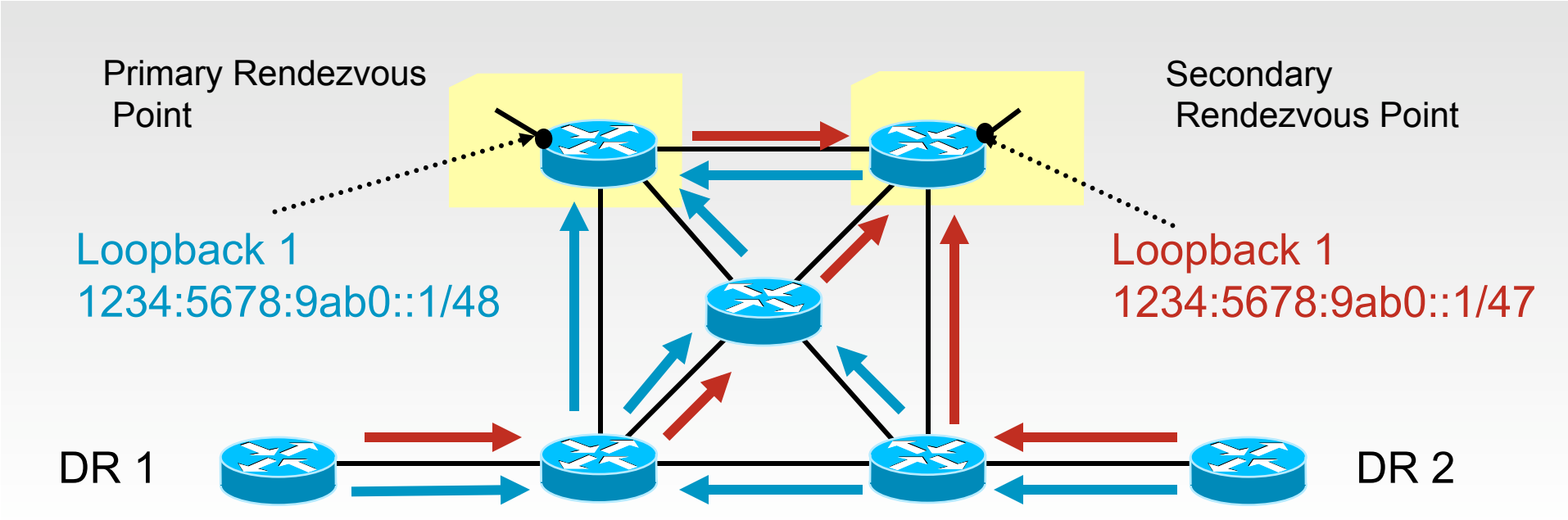
  Solution without any new protocol (in that way similar to embedded-RP)—a.k.a.: most simple solution?

  Could support PIM-SM and Bidir-PIM, IPv4 and IPv6



Now What Do I Do??

RP-Redundancy
Prioritycast Rendezvous Point



- NEW: Designate a primary and a secondary (tertiary and etc are also possible) Rendezvous Point for the anycast group
- NEW: Configure Primary Rendezvous Point with longest subnet mask on the loopback (secondary has longer mask)
- OLD: Distribute loopback interfaces routes into Interior Gateway Protocol (IGP)

Rendezvous Point Redundancy – Prioritycast Rendezvous Point

- Result:
  - All routers will converge on the primary Rendezvous Point, if it is available – because the longer mask route always wins
  - If the primary Rendezvous Point fails, failover is as fast as with the known MSDP /anycast-Rendezvous Point
    - Depends only on the convergence speed of the IGP
  - Because only one Rendezvous Point of the anycast group is active at any time, MSDP (or equivalent) is not needed
- Major difference:
  - No load-sharing between Rendezvous Points
    - Load-sharing is not necessary in IPv4
    - Load –sharing comes for free with the MSDP/anycast-Rendezvous Point redundancy
  - Scalability behavior is also different from MSDP
- No new protocol

Summary: Advantages of SSM

- No RPTs
- No register packets
- No RP mapping required (no RP required!)
- No RP-to-RP source discovery (no MSDP required!)
- No RP means no concentration of traffic towards the RP, and no single point to attack
- Rogue sources cannot easily spoof traffic
- SSM can use entire multicast address space, but FF3x/32 is reserved for SSM exclusively

A Few Notes on Tunnels

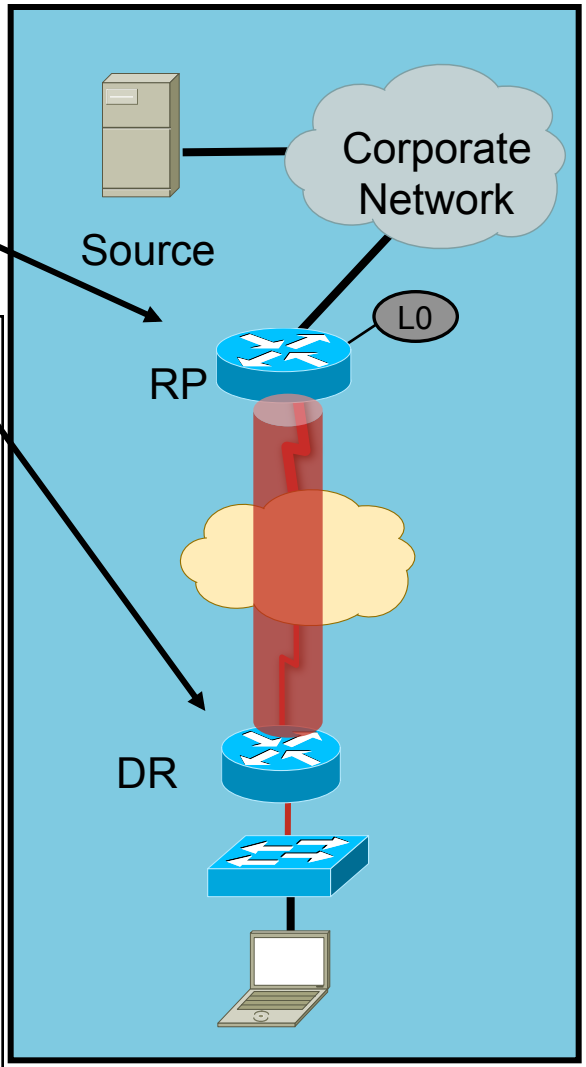
A Few Notes On Tunnels…

- PIM uses tunnels when RPs/Sources are known
- Source registering (on first-hop router)
  - Uses virtual tunnel interface (appear in OIL for (S,G))
  - Created automatically on first-hop router when RP is known
  - IOS keeps tunnel as long as RP is known
  - Unidirectional (transmit only) tunnels
  - PIM Register-Stop messages are sent directly from RP to registering router (not through tunnel!)

PIM Tunnels (DR-to-RP)

```
branch#show ipv6 pim tunnel
Tunnel1*
  Type  : PIM Encap
  RP    : 3FFE:C15:C003:1116::2
  Source: 3FFE:C15:C003:111E::2
```

```
branch#show interface tunnel 1
Tunnel1 is up, line protocol is up
  Hardware is Tunnel
  MTU 1514 bytes, BW 9 Kbit, DLY 500000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 3FFE:C15:C003:111E::2 (Serial0/2), destination
 3FFE:C15:C003:1116::2
  Tunnel protocol/transport PIM/IPv6, key disabled, sequencing
 disabled
  Checksumming of packets disabled
  Tunnel is transmit only
  Last input never, output never, output hang never
  Last clearing of "show interface" counters never
… output truncated…
```



PIM Tunnels (RP)

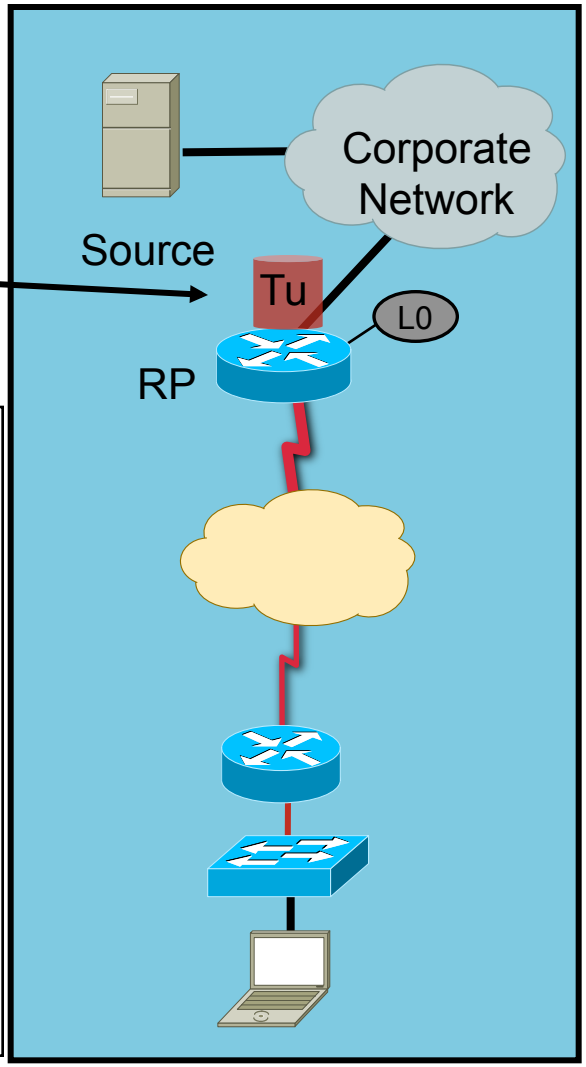
- Source registering (on RP) → 2 virtual tunnels are created
  - 1 transmit only for registering sources locally connected to the RP
  - 1 receive only for decapsulating incoming registers from remote designated routers
  - No one-to-one relationship between virtual tunnels on designated routers and RP!

PIM Tunnels (RP-for-Source)

```
RP-router#show ipv6 pim tunnel
Tunnel0*
  Type  : PIM Encap
  RP    : 3FFE:C15:C003:1116::2
  Source: 3FFE:C15:C003:1116::2
Tunnel1*
  Type  : PIM Decap
  RP    : 3FFE:C15:C003:1116::2
  Source:  -
```

```
RP-router#show interface tunnel 1
Tunnel1 is up, line protocol is up
  Hardware is Tunnel
  MTU 1514 bytes, BW 9 Kbit, DLY 500000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 3FFE:C15:C003:1116::2 (FastEthernet0/0), destination
 3FFE:C15:C003:1116::2
  Tunnel protocol/transport PIM/IPv6, key disabled, sequencing disabled
  Checksumming of packets disabled
  Tunnel is receive only

… output truncated…
```




CISCO