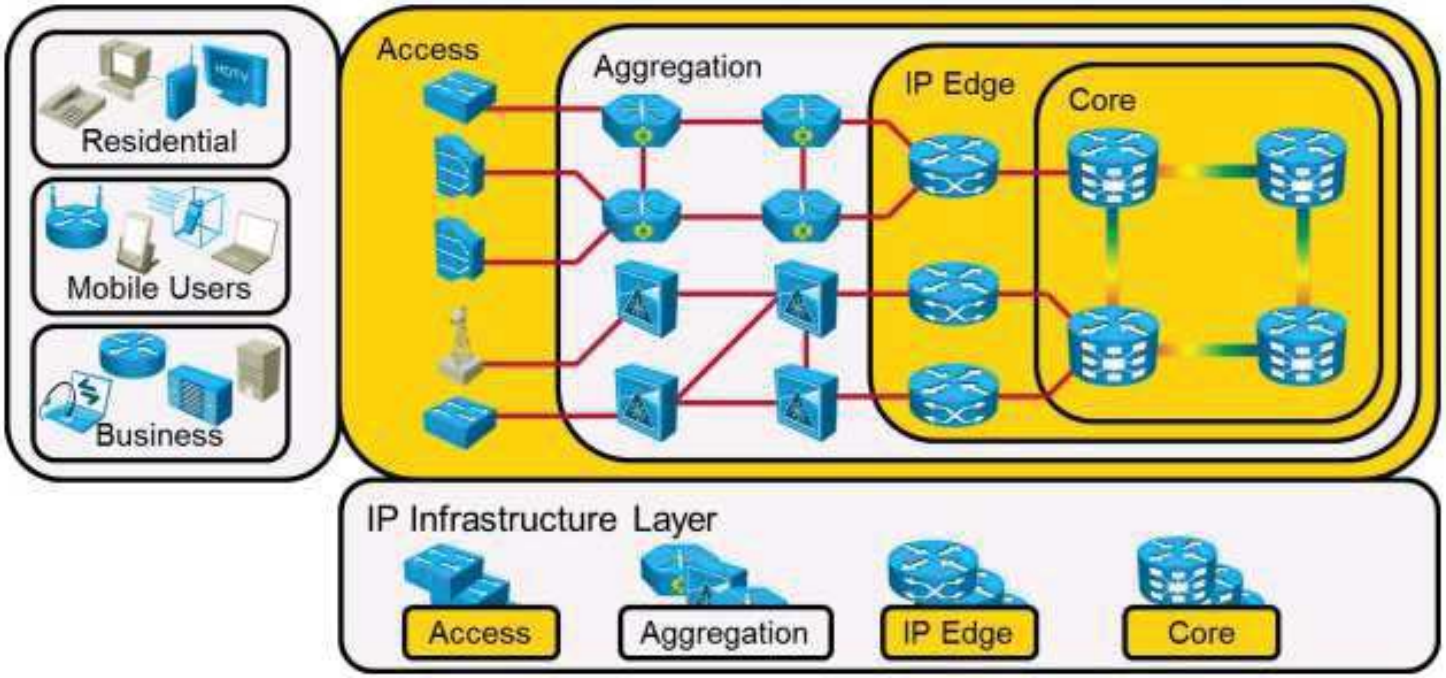
Introducing IP Multicast

Multicast Overview

IP Multicast Benefits and Caveats

IP multicast key benefits and caveats:

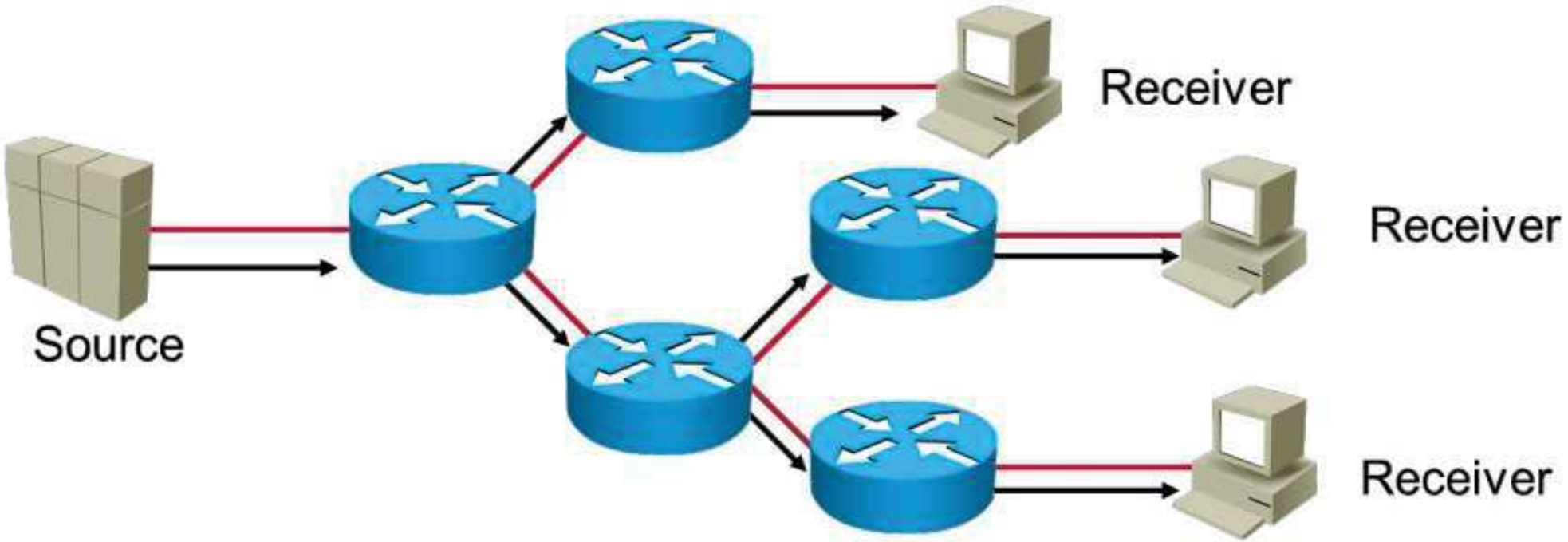
- Provide better bandwidth utilization by sending the same data to multiple receivers.
- Facilitate less host and router processing.
- Accommodate traffic when receiver addresses are unknown.
- Simultaneously deliver data for a group of receivers (simulcast).



Multicast Operations High-Level Overview

The working steps for IP multicast:

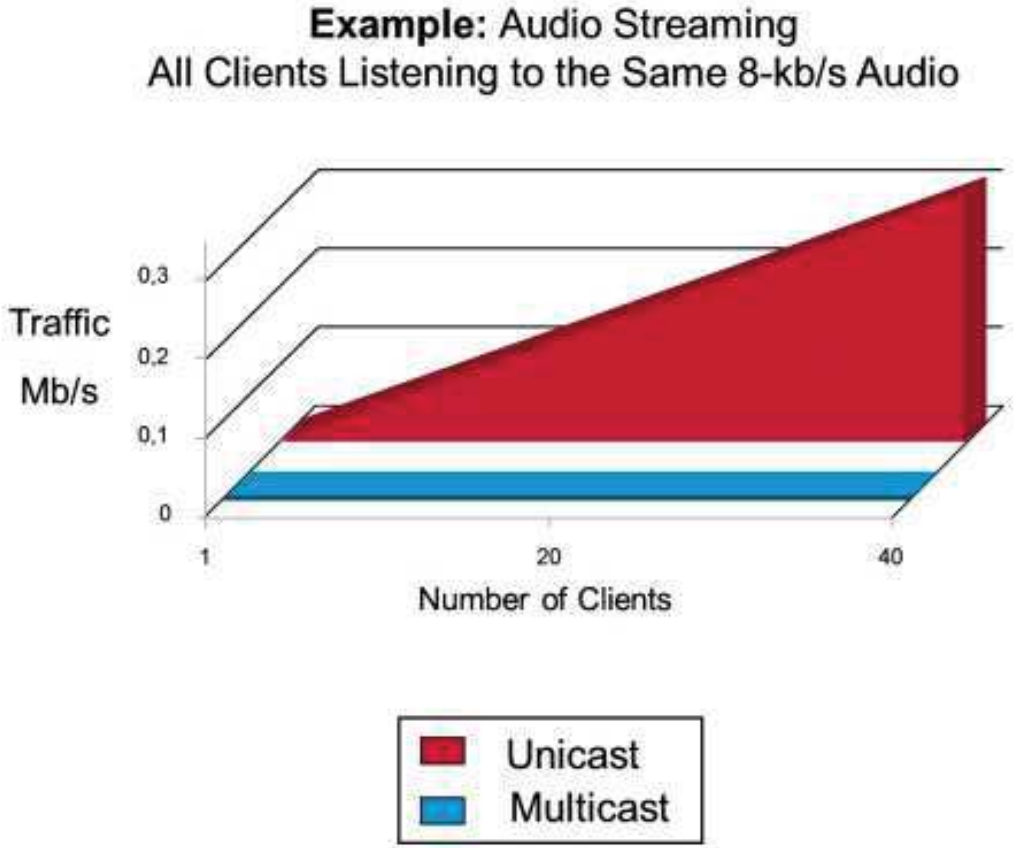
- The sender (source) sends one copy of a single packet addressed to a group of receivers—a multicast group.
- Multicast routers replicate and forward the packet to all the branches where receivers may exist.
- Receivers express their interest in multicast traffic by sending control messages to routers.



Multicast Advantages and Disadvantages

Multicast advantages:

- Enhanced efficiency: Controls network traffic and reduces server and CPU loads.
- Optimized performance: Eliminates traffic redundancy.
- Distributed applications: Makes multipoint applications possible.
- Fewer resources required forbandwidth and host processing power.
- Almost simultaneous delivery is assured.
- Foundation for new applications that was not possible in the past.



Multicast Advantages and Disadvantages (Cont.)

Multicast disadvantages:

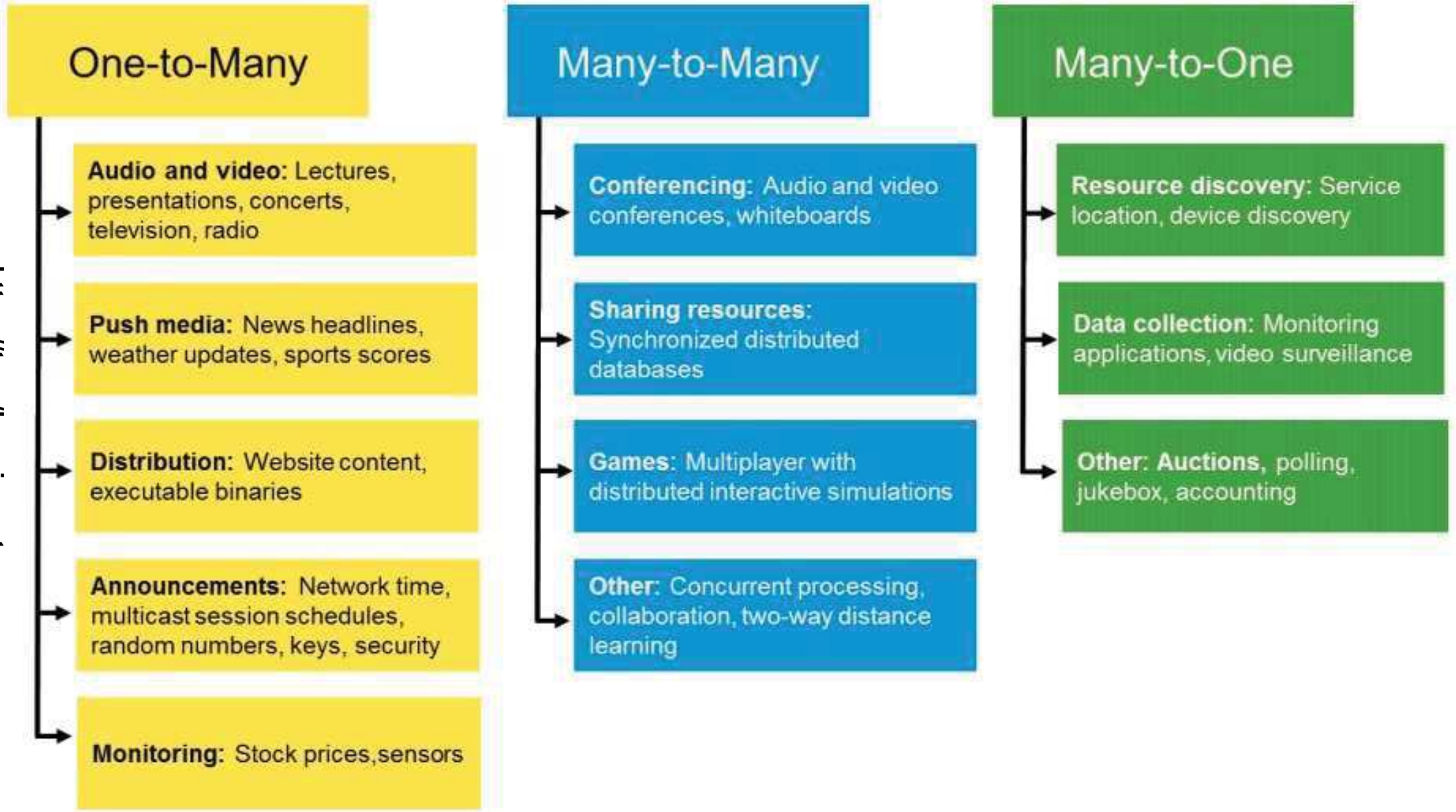
- Best-effort delivery: Multicast applications cannot assure reliable delivery of data and should be designed accordingly.
- No congestion avoidance: Lack of TCP windowing and slow-start mechanisms can result in network congestion.
- Duplicates: Some multicast protocol mechanisms result in the occasional generation of duplicate packets.
- Out-of-sequence delivery: Network topology changes affect the order of delivery.
- Reliability is a special issue not addressed in original IP multicast research.
- Security is another area in IP multicast that has not been sufficiently solved.

Multicast Application Types

There are different types of multicast applications:

- One-to-many: A single host sending to two or more receivers.
- Many-to-many: Any number of hosts sending to the same multicast group—hosts are also members of the group (senders are receivers).
- Many-to-one: Any number of receivers sending data back to a source (via unicast or multicast).

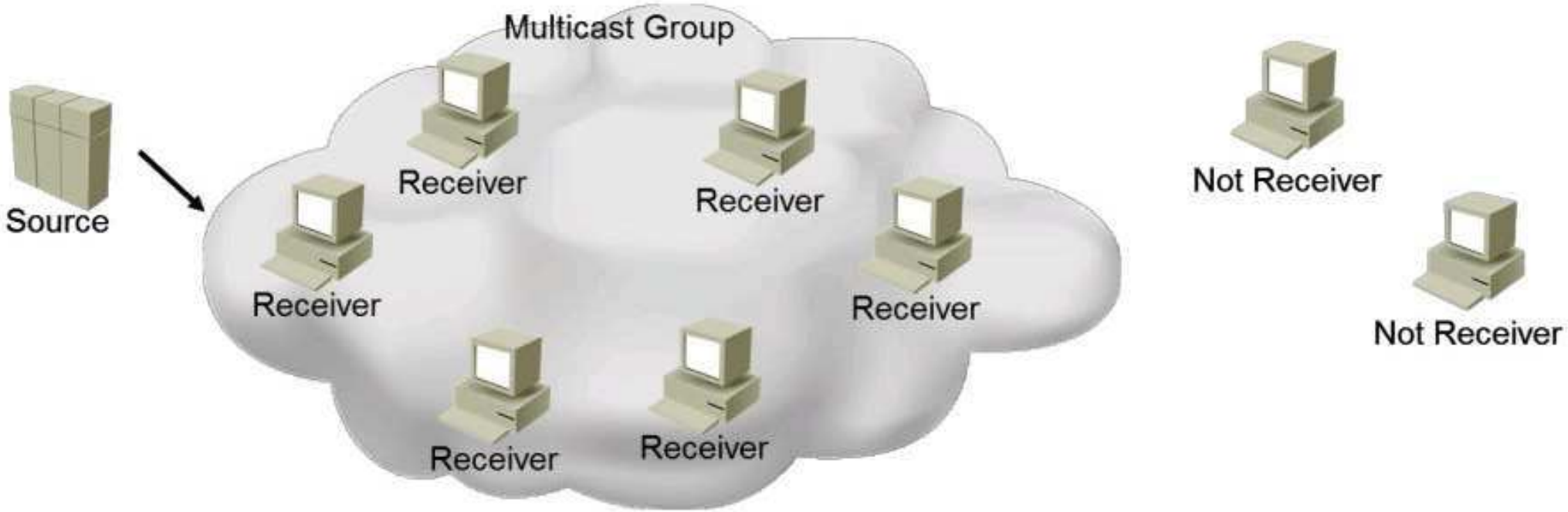
Multicast Application Types (Cont.)



IP Multicast Group Address

IP multicast group address characteristics:

- Multicast group is 32-bit IP address derived from Class D (now RFC 3171).
- Source sends stream to the multicast group with destination address equal to the multicast group.
- Receivers join the multicast group to receive stream from source.



IP Multicast Basic Addressing

IP multicast basic addressing scheme:

- IP group addresses
  - Class D address—high-order 4 bits are set
  - Range from 224.0.0.0 through 239.255.255.255
- Well-known link-local addresses assigned by IANA
  - Reserved use of 224.0.0.0 through 224.0.0.255

IP Address	Description
224.0.0.1	All multicast systems on subnet
224.0.0.2	All routers on subnet
224.0.0.4	All DVMRP routers
224.0.0.13	All PIMv2 routers
224.0.0.5, 224.0.0.6, 224.0.0.9, 224.0.0.10	Used by unicast routing protocols

IP Multicast Basic Addressing (Cont.)

- Transient addresses, assigned and reclaimed dynamically:
  - Global range: 224.0.1.0– 238.255.255.255
  - Limited (local) scope: 239.0.0.0/8
  - Part of a global scope recently used for new protocols and temporary usage.

RFC 2770 and SSM Addressing

Static group address assignment for interdomain multicast:

- Temporary method to meet immediate needs.
- Group range: 233.x.x.0–233.x.x.255
- Autonomous system number is inserted in middle two octets (x.x).
- Remaining low-order octet used for group assignment within a domain.

SSM group address assignment for interdomain multicast:

- Used exclusively for globally known sources and source-specific distribution trees (across domains).
- Group range: 232.0.0.0/8

Multicast Session Directory

Dynamic multicast addressing accomplished using SDR application (multicast backbone):

- Sessions and groups announced over well-known multicast groups (224.2.127.254).
- Address collisions detected and resolved at time of session creation—addresses looked up in an SDR cache.
- Not scalable.

Multicast Session Directory (Cont.)

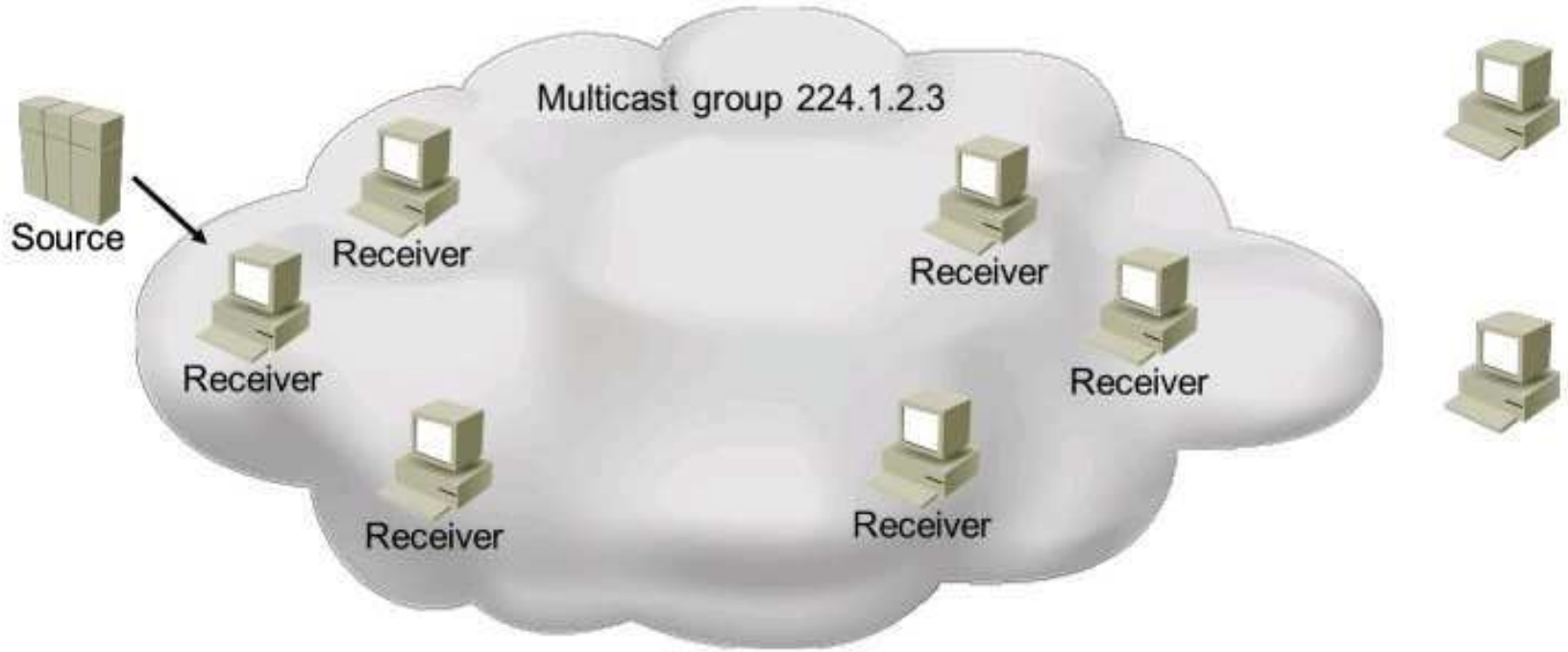
Learning about multicast sessions:

- Potential receivers have to learn about available multicast streams and sessions before a multicast application is launched.
- Possibilities:
  - Another multicast application sending to a well-known group whose members are all potential receivers.
  - Directory services.
  - Web page, email, and so on.
- Multicast backbone uses session directory and an enhanced version, SDR.
- SDR is a session description protocol and transport mechanism.

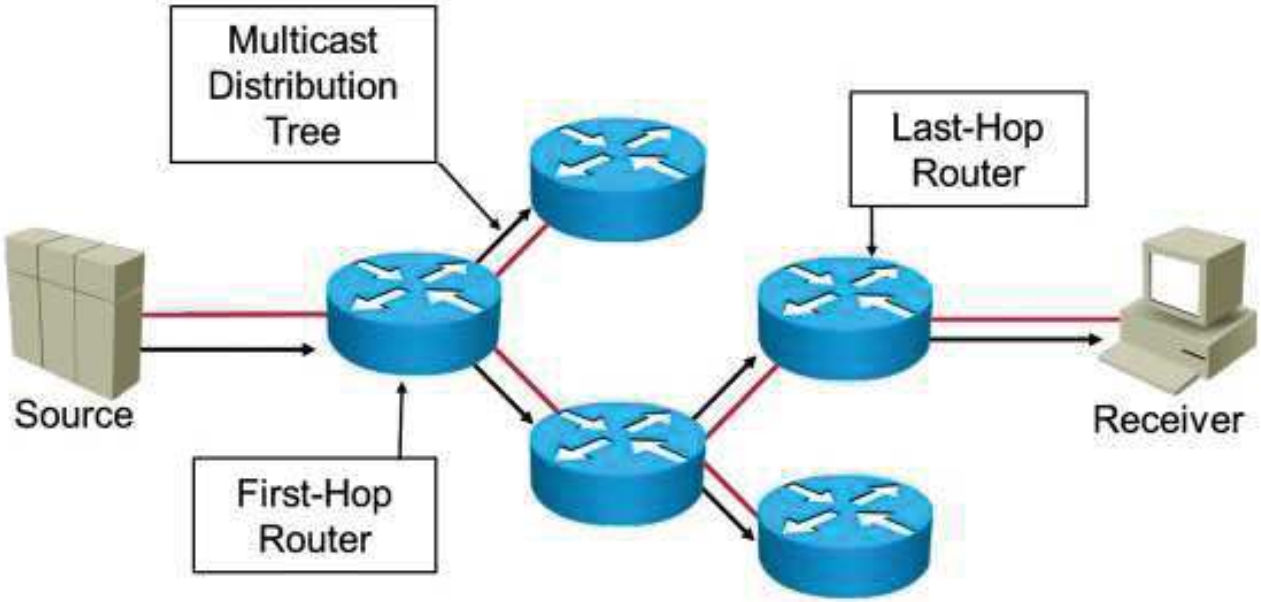
IP Multicast Service Model

IP Multicast service model characteristics:

- RFC 1112—Host Extensions for IP Multicasting.
- Each multicast group is identified by a Class D IP address.
- Members join and leave the group and indicate this to the routers.
- Routers listen to all multicast addresses and use multicast routing protocols to manage groups.



IP Multicast Service Model (Cont.)



- Multicast network routers are distinct from source and receiver segments.
- Sources simply start sending data without any indication and first-hop routers forward data.
- Receivers report their membership to last-hop routers.
- Last-hop (leaf) routers communicate group membership to the network.

Functions of a Multicast Network

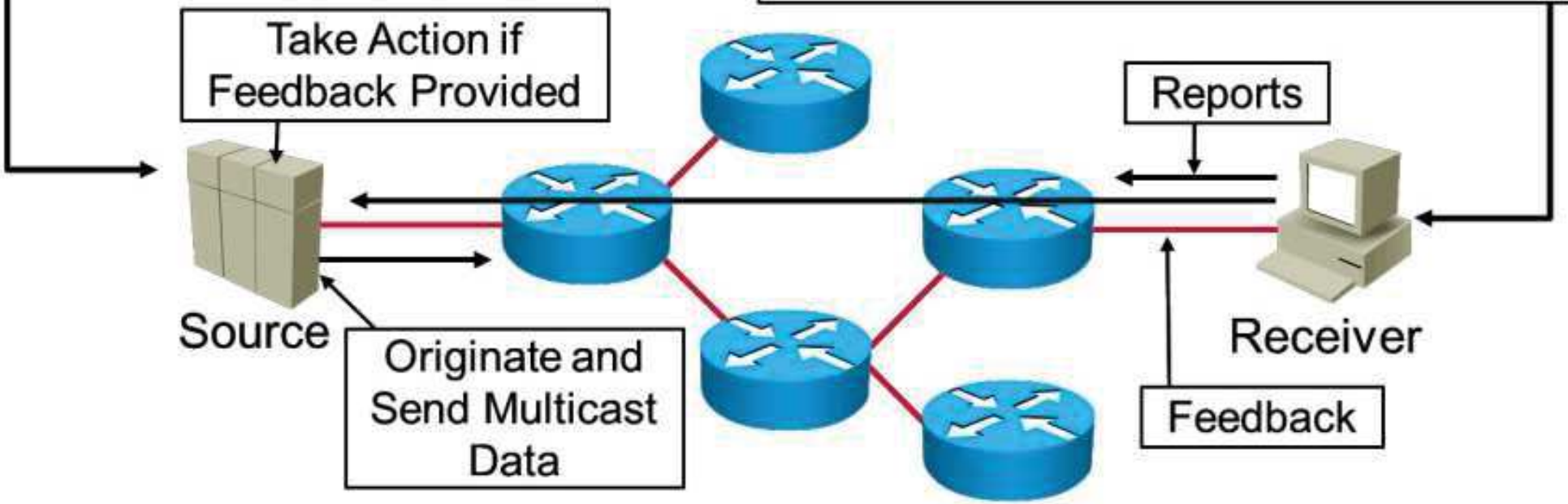
Key steps for a properly functioning multicast network:

- Learn about multicast group members and build an appropriate distribution tree.
- Identify multicast streams and forward them according to a distribution tree.
- Maintain:
  - Group state at leaf segments.
  - Distribution trees in the whole network.
- Prevent loops and apply scoping and filtering.

Multicast Sources and Receivers

- Create a session or group and announce it via session announcement.
- Originate multicast data and send it to a multicast group .
- Apply proper actions if feedback information is available.

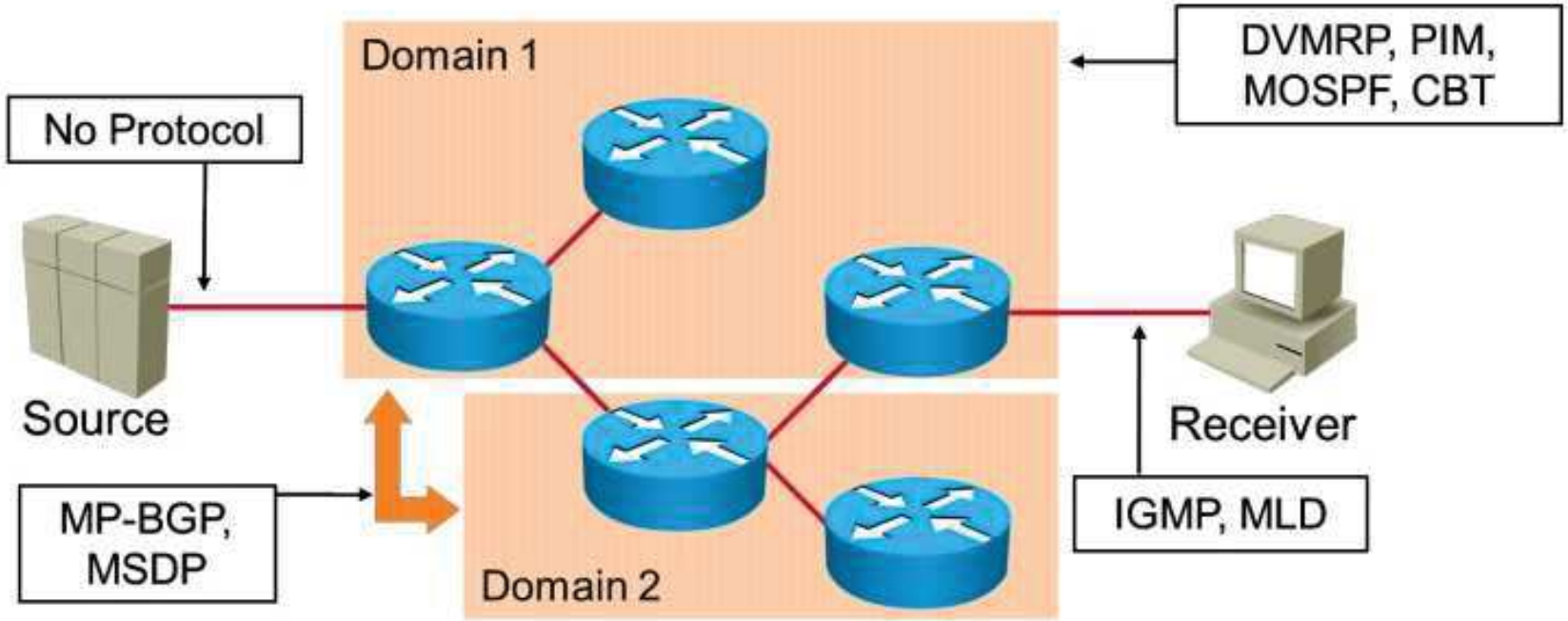
- Listen for session announcements or use some other mechanism to learn about available sessions.
- Report their interest in a certain group by sending messages to the routers.
- Receive multicast data and provide feedback if needed.
- Maintain their group membership and leave the group when necessary.



Multicast Protocols

Multicast protocol types:

- No control protocols spoken at source segments.
- Multicast routing protocols used in a multicast network:
  - Intradomain (DVMRP, PIM and variants, MOSPF, and CBT)
  - Interdomain (MP-BGP and MSDP)
- Receiver segments use IGMP or MLD



Multicast Forwarding and RPF Check

Multicast forwarding and RPF check key characteristics:

- Multicast routing works the opposite way of unicast routing:
  - Unicast routing is concerned with where the packet is going.
  - Multicast routing is concerned with where the packet comes from.
- The routing table for unicast is checked against the source address in the multicast datagram.
- If the datagram arrived on the interface specified in the routing table for the source address:
  - The RPF check succeeds, and the datagram is forwarded.
  - Otherwise, the RPF check fails, and the datagram is silently discarded.
- When a datagram is forwarded, it is sent out of each interface in the OIL.
- The packet is never sent back out of the incoming interface (RPF interface).

Multicast Scoping

What is a TTL threshold ?

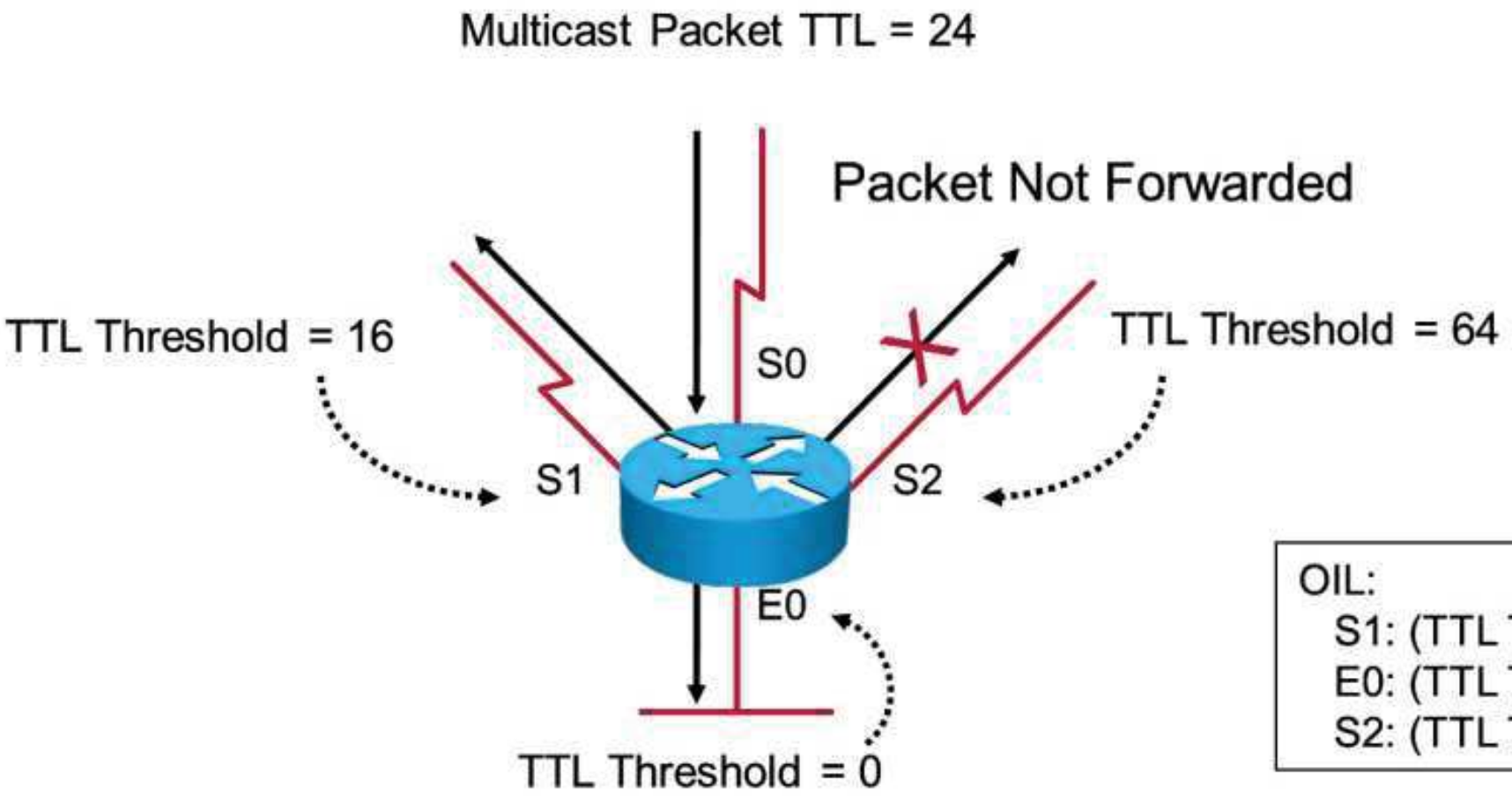
- A TTL threshold may be set on a multicast router interface to limit the forwarding of multicast traffic to outgoing packets with TTLs greater than the threshold.

The TTL threshold check:

- All incoming IP packets first have their TTL decremented by 1. If the TTL is less than or equal to 0, the packets are dropped.
- If a multicast packet is to be forwarded out of an interface with a nonzero TTL threshold, its TTL is checked against the TTL threshold.
- If the TTL of the packet is less than the specified threshold, the packet is not forwarded out of the interface.

Multicast Scoping (Cont.)

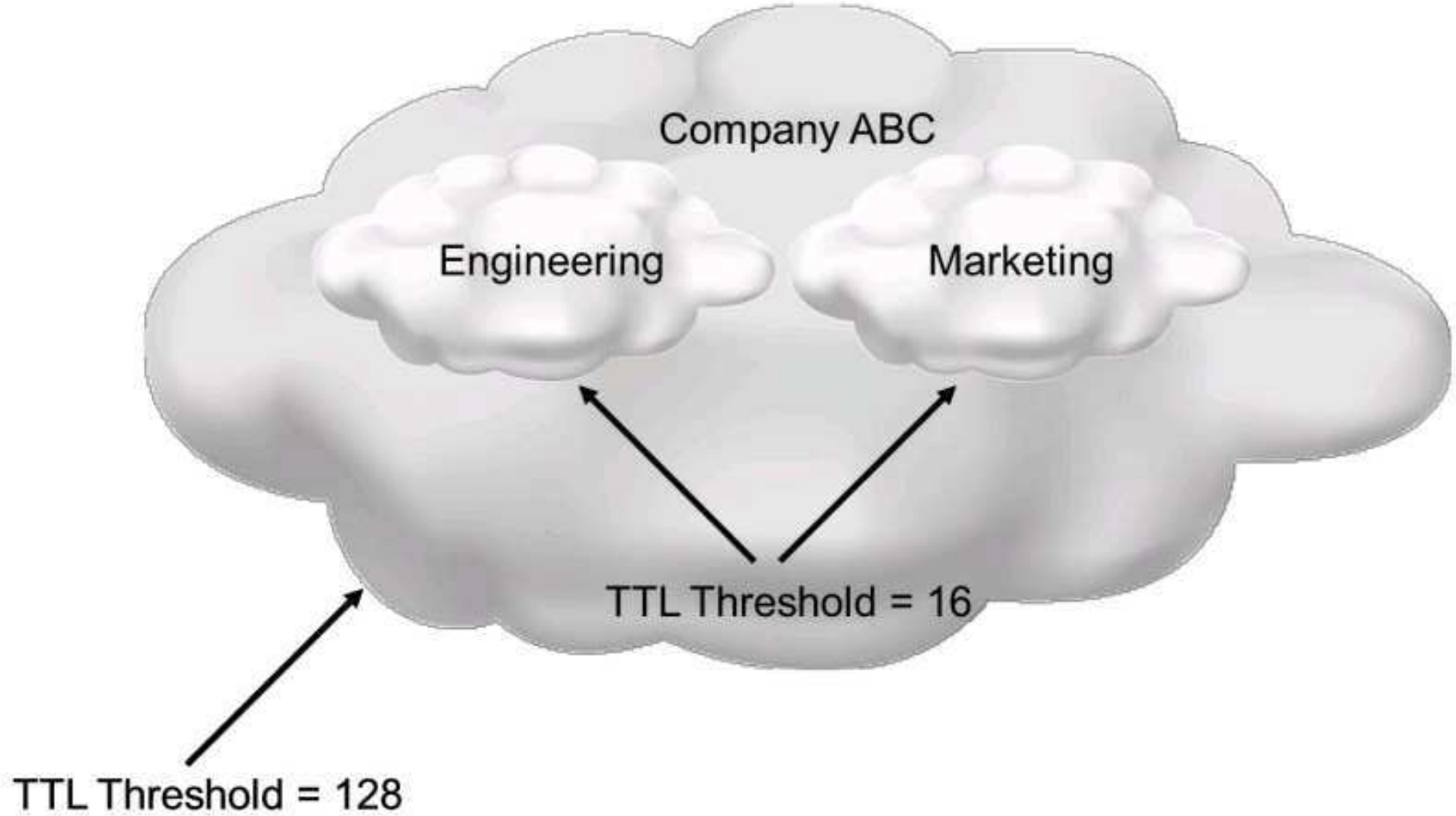
TTL Thresholds Example:



OIL:
S1: (TTL Threshold = 16)
E0: (TTL Threshold = 0)
S2: (TTL Threshold = 64)

Multicast Scoping (Cont.)

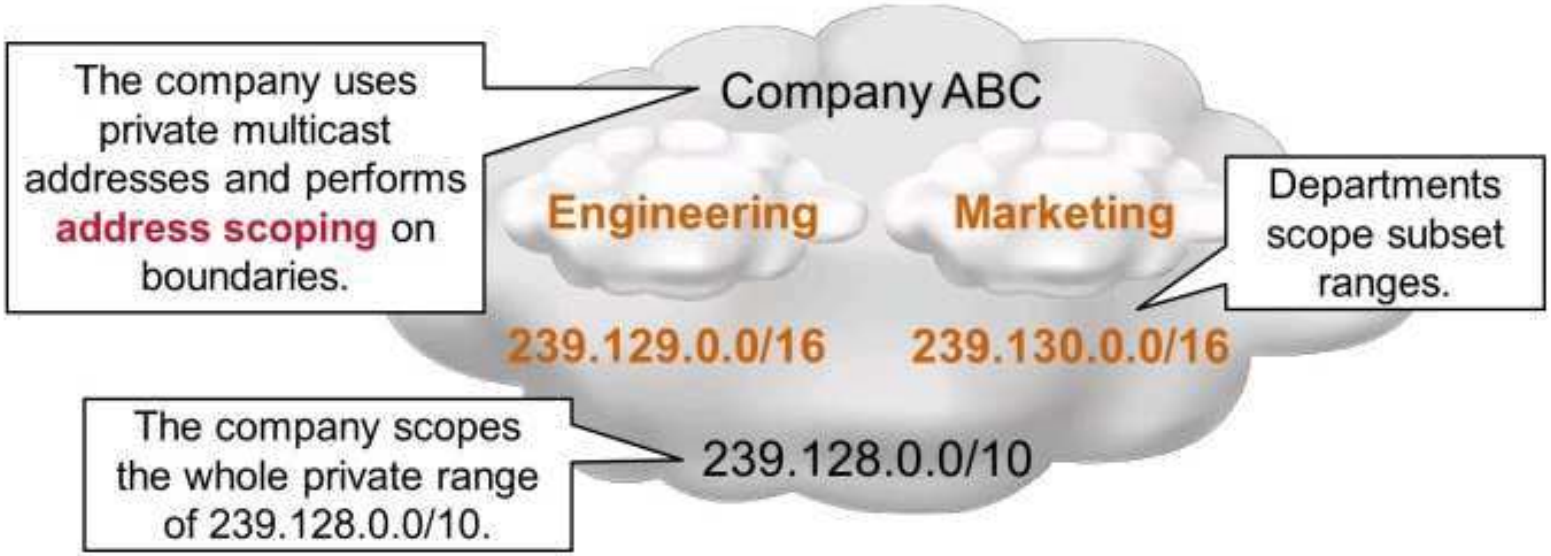
TTL Boundaries:



Multicast Scoping (Cont.)

TTL scoping characteristics:

- TTL scoping depends on the settings of the TTL, which are sometimes unknown or unpredictable.
- An alternative is address scoping, which allows multicast boundaries to be established per group address.
- Traffic that does not match the address is not accepted on an incoming interface or not forwarded to an outgoing interface.

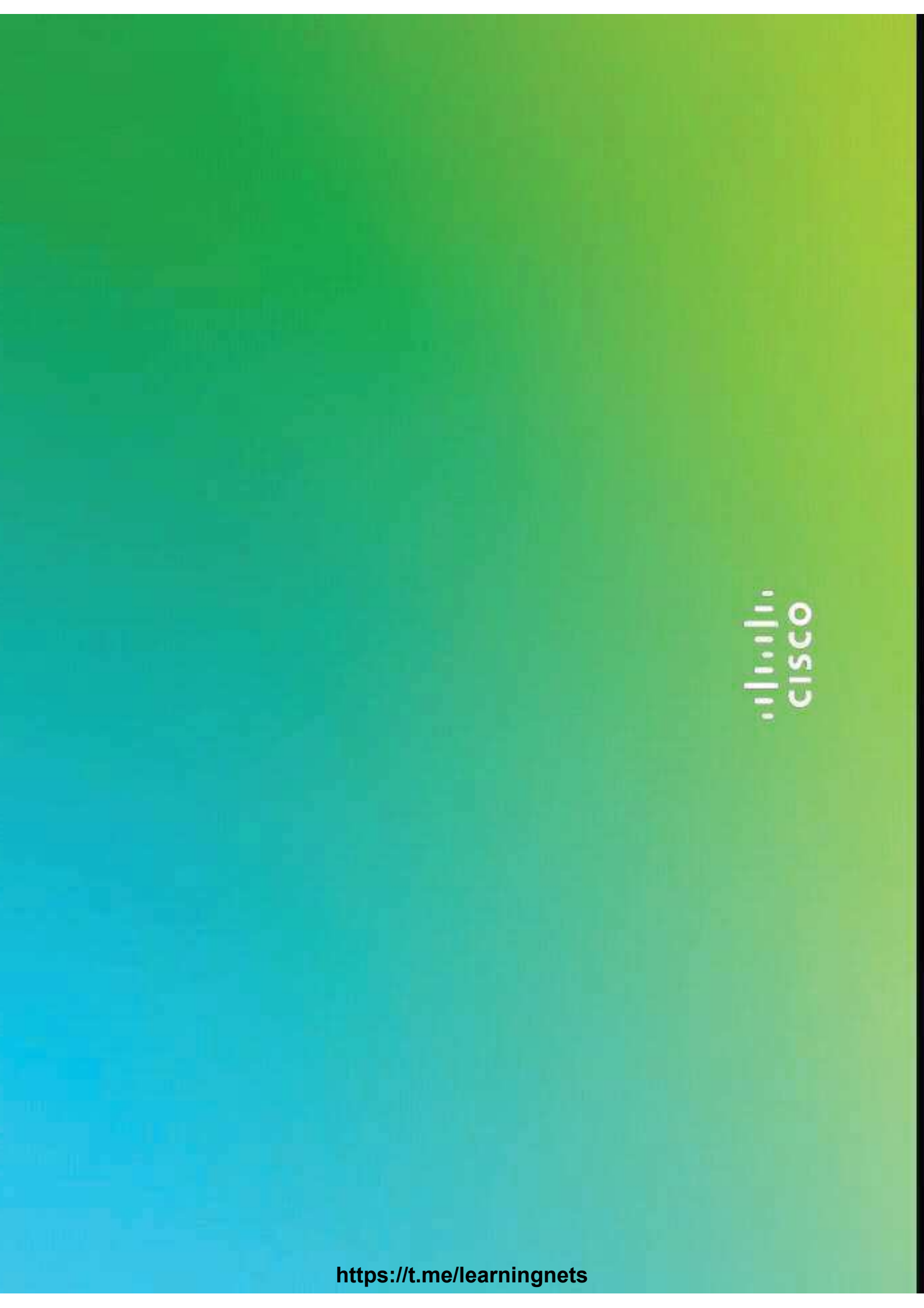


Summary

- Multicast sends a single packet to multiple receivers, thus reducing network load closer to the source.
- Two of the most common multicast application models are one-to-many and many-to-many.
- Multicast reduces load on the network, but adds complexity when compared with unicast.
- Applications may learn about the sessions in several ways:
  - By joining a well-known group.
  - By using directory services.
  - By using a URL to launch an application.
- Multicast IP addresses use the Class D address space.
- SDR is a session description protocol and transport mechanism.
- In a multicast network, there are sources, receivers and routers.

Summary (Cont.)

- A multicast network learns about multicast group members and multicast streams, and builds an appropriate distribution tree between sources and receivers.
- Sources are the source of multicast streams, receivers are the destination.
- Various multicast routing protocols are used inside the multicast network.
- The multicast network performs RPF checks to prevent multicast traffic loops.
- Address scoping or TTL scoping is used to restrain multicast traffic to specific network segments.



Defining Multicast Distribution Trees and Forwarding

Multicast Overview

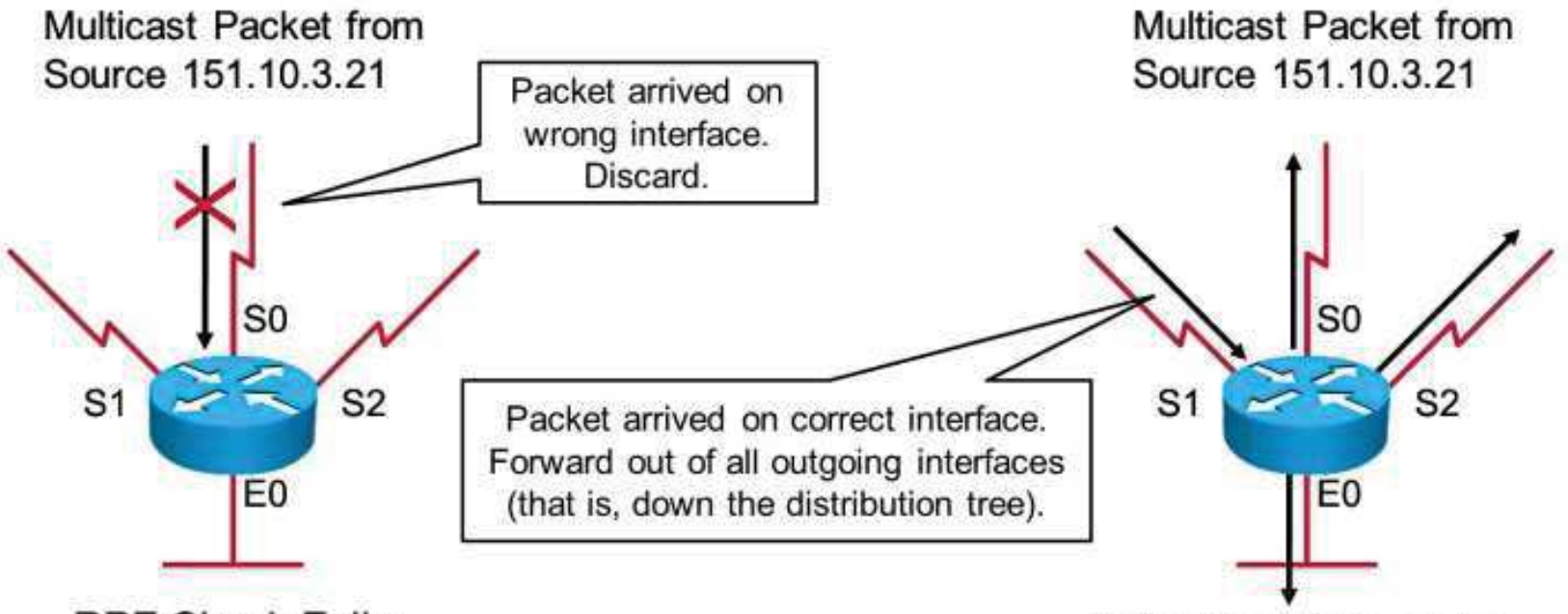
RPF Check

RPF check characteristics:

- RPF check is performed with respect to the RPF interface:
  - The interface that is closest to the source.
  - Determined from any unicast or dedicated multicast table (DVMRP, MP-BGP).
- Periodic recheck of the RPF interface and triggered by unicast routing table change.

RPF Check (Cont.)

Fails and Succeeds:



Unicast Route Table	
Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Unicast Route Table	
Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Types of Multicast Distribution Trees

Types of multicast distribution trees:

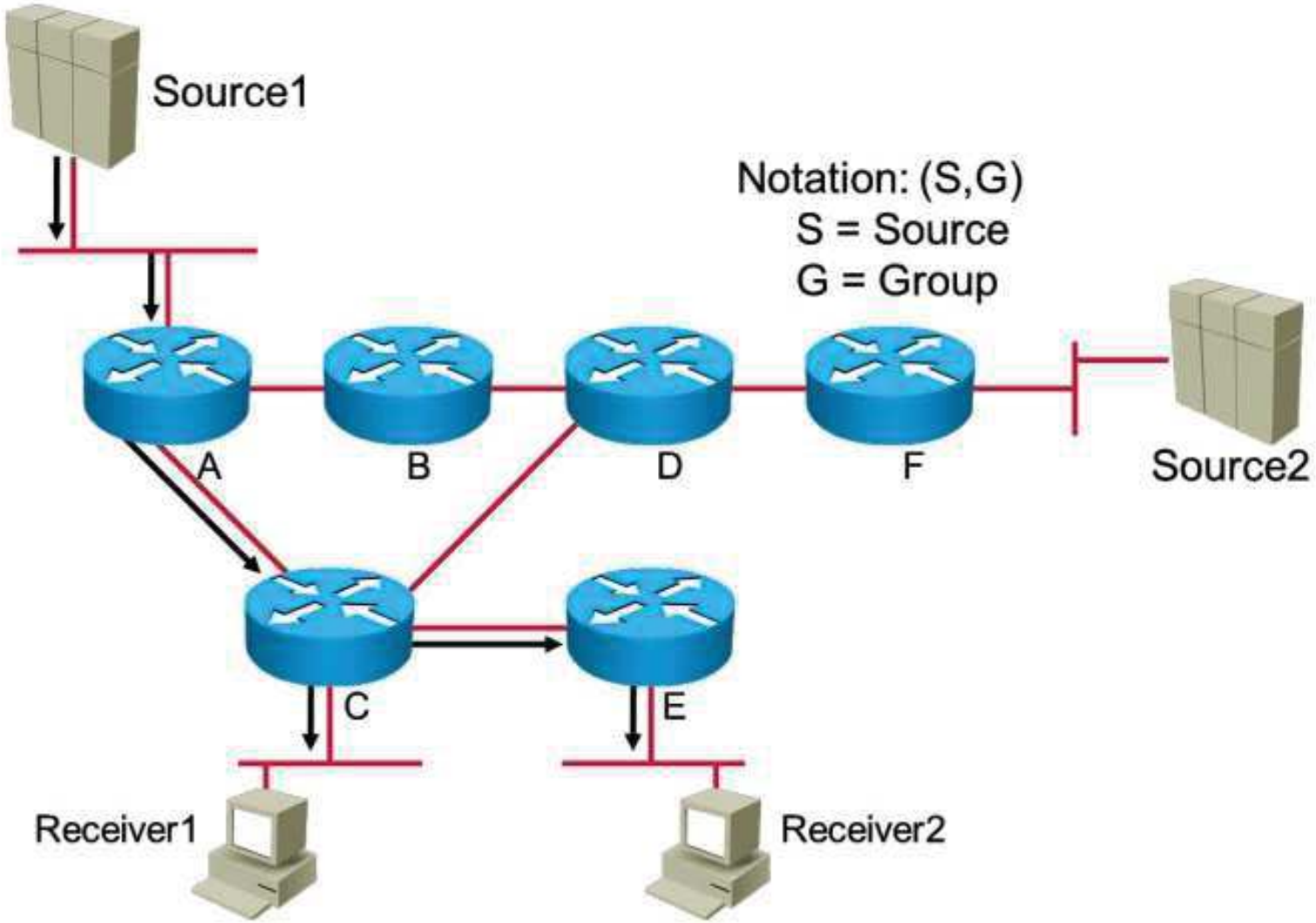
- Source-rooted: SPT
- Rooted at a meeting point in the network: shared trees
  - RP

Types of multicast protocols:

- Dense mode
- Sparse mode

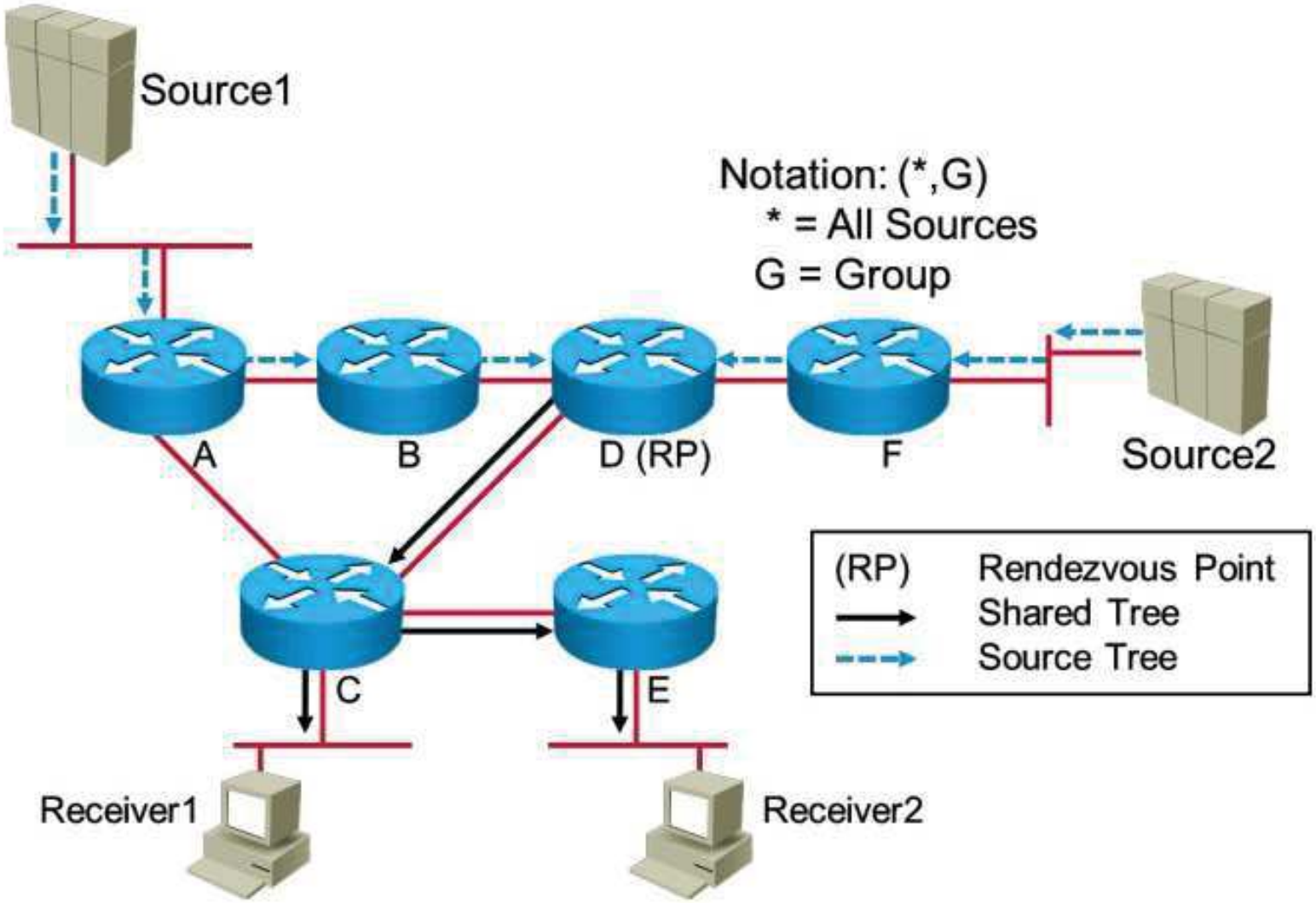
Types of Multicast Distribution Trees (Cont.)

Source Distribution Tree:



Types of Multicast Distribution Trees (Cont.)

Shared Distribution Tree:



Multicast Distribution Trees Identification

(S,G) entries:

- For this particular source sending to this particular group.
- Traffic is forwarded via the shortest path from the source.
- Source or shortest path trees use more memory, but you may get optimal paths from the source to all receivers. Minimizes delay.

(*,G) entries:

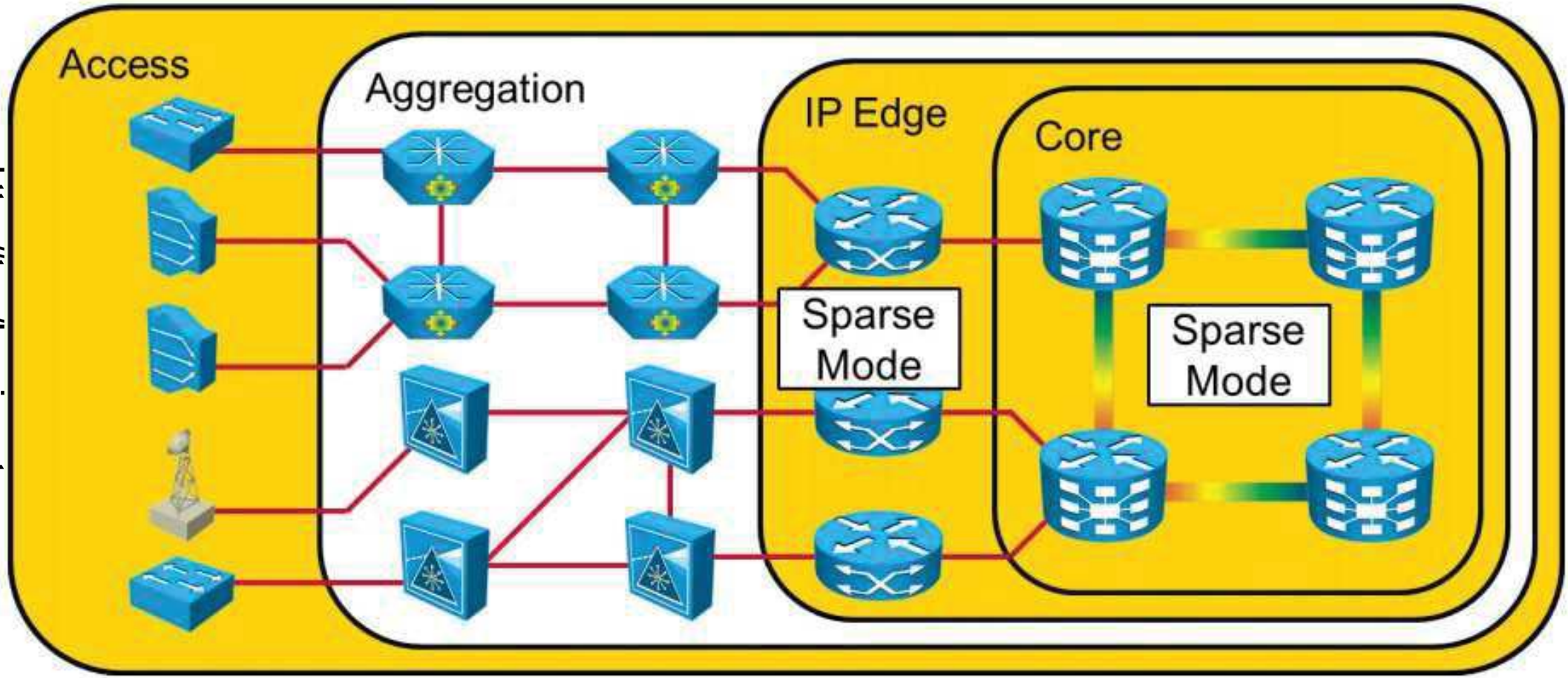
- For any (*) source sending to this group.
- Traffic is forwarded via a meeting point for this group.
- Shared trees use less memory, but you may get suboptimal paths from the source to all receivers. May introduce extra delay.

Multicast Protocols Overview

Multicast Protocols in the Cisco IP NGN Infrastructure Layer

- Dense mode (DVMRP, MOSPF, PIM-DM):
  - Uses the push model.
  - Initial traffic flooded to all branches of the distribution tree.
  - Branches without receivers get pruned.
  - Displays flood-and-prune behavior (typically every 3 minutes).
- Sparse mode (PIM-SM, CBT):
  - Uses the pull model (join behavior).
  - Branches without receivers never get the traffic.
  - Last-hop routers pull the traffic from the meeting point or from the source.

Multicast Protocols Overview (Cont.)



PIM Dense Mode and Sparse Mode High-Level Overview

PIM-DM overview:

- Supports all underlying unicast routing protocols, including static, RIP, EIGRP, IS-IS, OSPF, and BGP.
- Uses flood-and-prune mechanism:
  - Floods network and prunes back based on multicast group membership.
  - Uses assert mechanism to prune off redundant flows on multiaccess networks.
- Appropriate for smaller implementations and pilot networks.

PIM Dense Mode and Sparse Mode High-Level Overview (Cont.)

PIM-SM Overview:

- Works with any of the underlying unicast routing protocols.
- Supports both source trees and shared trees.
- Based on an explicit pull model and uses an RP:
  - Senders are registered with RP by first-hop router.
  - Receivers are joined to the shared tree (rooted at the RP) by last-hop router.
- Appropriate for:
  - Large-scale deployment for both densely and sparsely populated groups in the enterprise.
  - Optimal choice for all production networks, regardless of size and membership density.
- Optimizations and derivatives:
  - BIDIR-PIM, SSM

Intradomain Multicast Routing Protocols

Intradomain multicast routing protocols types:

- PIM:
  - Uses an existing unicast routing table plus a join-prune-graft mechanism.
  - Sparse mode (RFC 4601), dense mode (RFC 3973).
- DVMRP:
  - Uses the DVMRP routing table plus a special poison-reverse mechanism.
  - v2, v3 (Internet draft); v1 (RFC 1075) is obsolete
- MOSPF:
  - Uses an extension of OSPF link-state mechanism.
  - RFC 1584
- CBT:
  - Uses an existing unicast routing table plus a join-prune-graft mechanism.
  - RFC 2189

Interdomain Multicast Routing

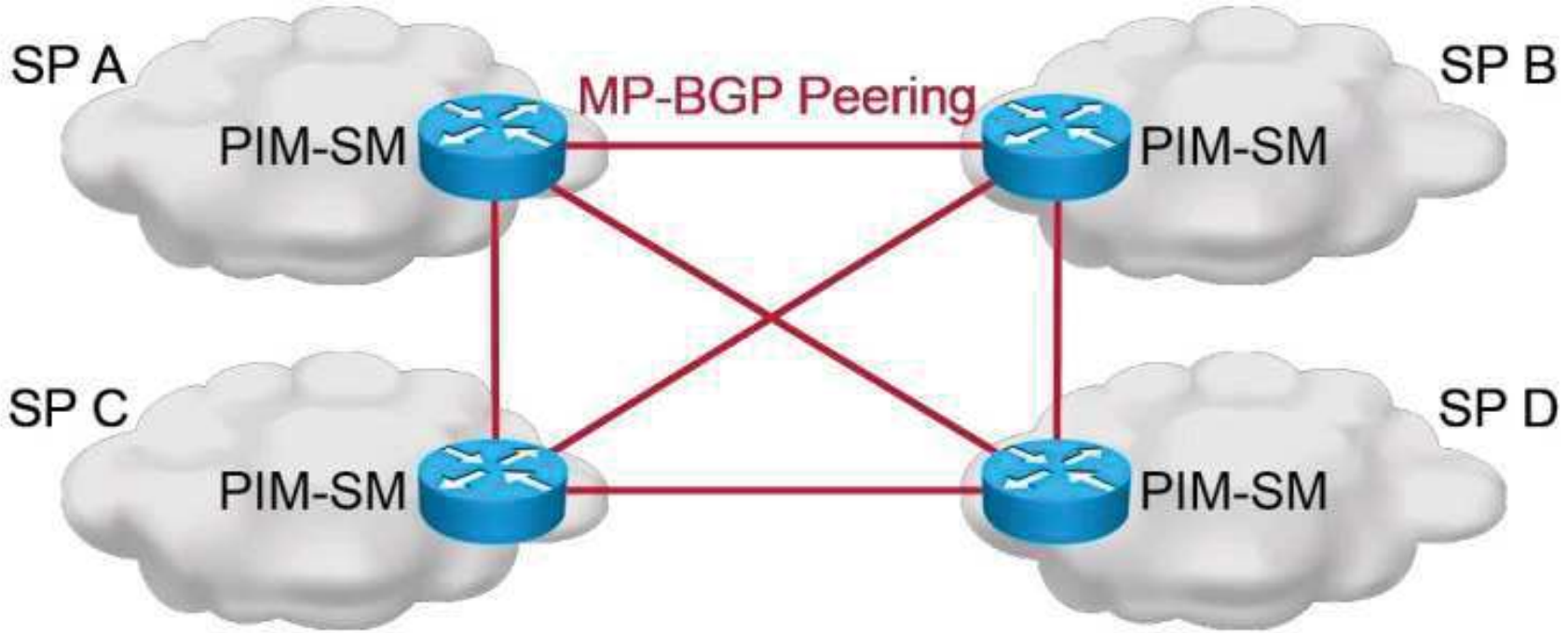
Interdomain multicast routing protocols characteristics:

- Working solution:
  - MP-BGP for RPF information.
  - MSDP to learn about sources.
- Several other attempts as interim solutions.

Interdomain Multicast Routing Protocols (Cont.)

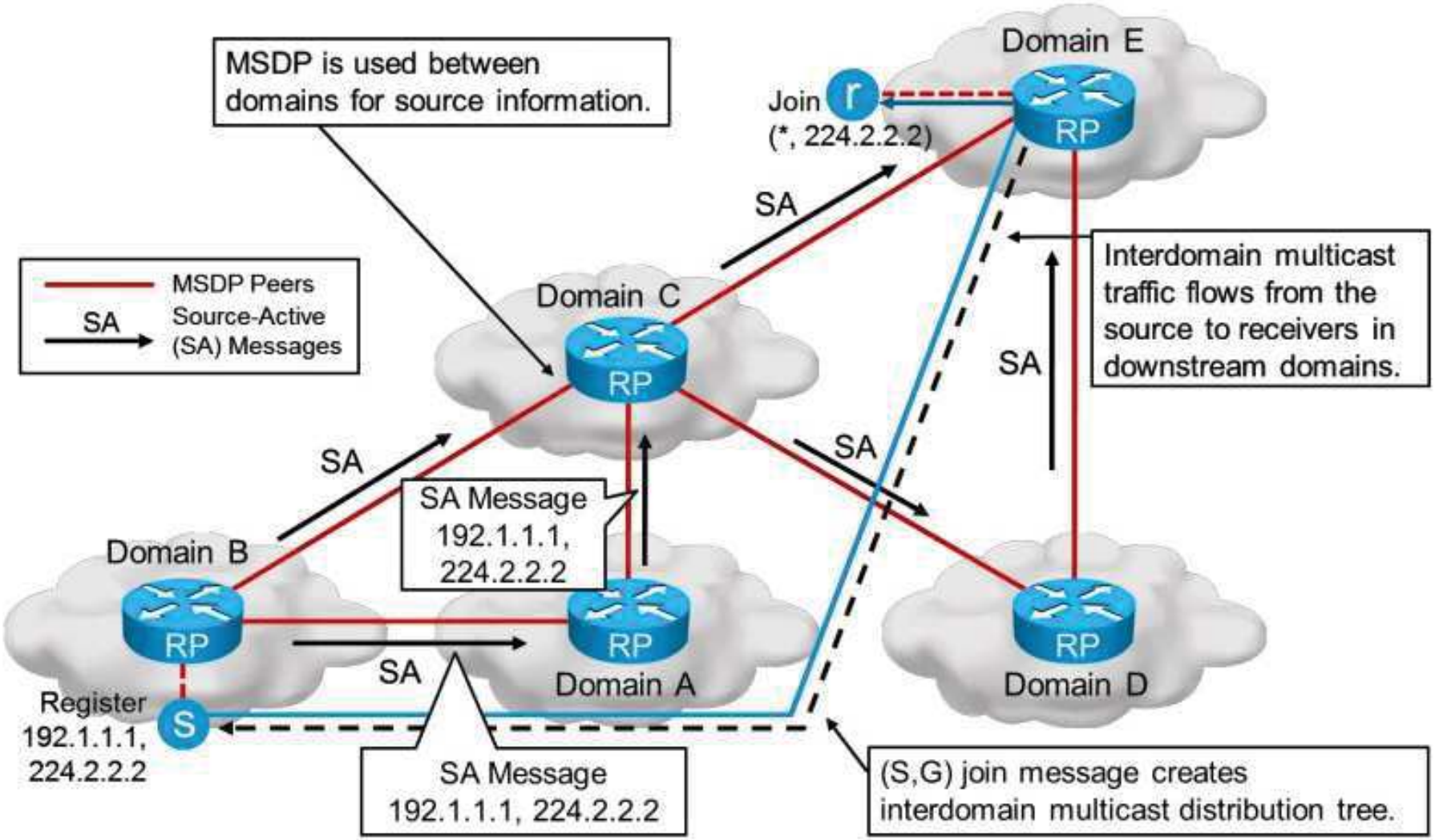
MP-BGP:

- PIM-SM used within and across domains.
- MP-BGP used between domains for source network information (RPF checks).



Interdomain Multicast Routing Protocols (Cont.)

MSDP:



Multicast High-Availability Options

Rendezvous point high availability:

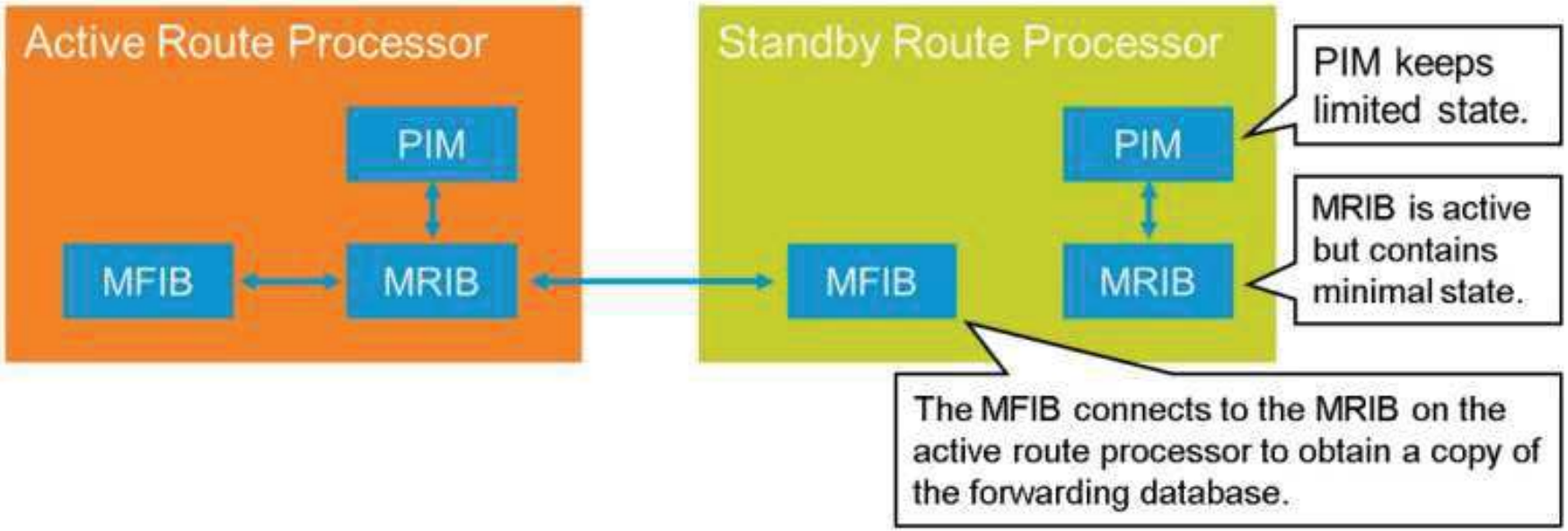
- **Auto-RP:** A mechanism to automate distribution of RP information in a multicast network.
- **The Bootstrap Router:** BSR is a mechanism for a router to learn RP information. It ensures that all routers in the PIM domain have the same RP cache as the BSR.
- **Anycast RP:** Used to define redundant and load-balanced RPs. Anycast-RP has two implementations:
  - Using MSDP
  - Using PIM

Multicast High-Availability Options (Cont.)

Features that are available for dual route processor platforms:

- **Multicast NSF with SSO:** Allows forwarding of multicast traffic during the switchover.
- **PIM triggered joins:** Improves PIM convergence after a switchover by triggering adjacent PIM neighbors to resend join messages.

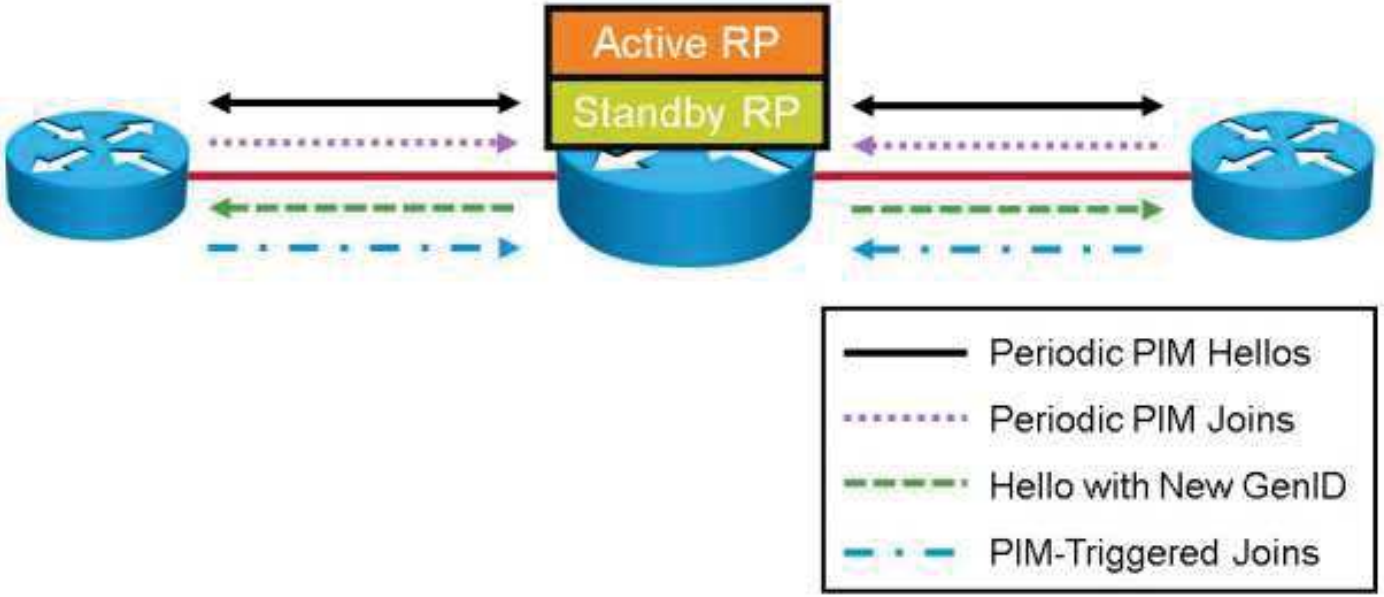
Multicast NSF with Stateful Failover



Multicast NSF with SSO:

- NSF allows synchronization of MFIBs between active and standby route processors.
- During the switchover, MFIB on standby RP is frozen and used to forward multicast traffic.
- After PIM recovers on the new active RP, PIM updates MRIB, and MRIB updates MFIB if needed.

PIM Triggered Joins



PIM triggered join characteristics:

- Triggers resending of PIM joins from neighbors after a switchover.
- Modifies GenID values in the PIM hello packets sent from the new active RP.
- Modified GenID value is a mechanism that alerts neighbors that PIM forwarding on an interface has been lost.
- Neighbors resend PIM joins for joined groups as a result of received modified GenID.

IGMP Overview

IGMP characteristics:

- IGMP is the way that hosts tell routers about group membership.
- Routers solicit group membership from directly connected hosts.
- RFC 1112 specifies the first version of IGMP.
- RFC 2236 specifies the second version of IGMP.
- RFC 3376 specifies the current (third) version of IGMP.

IGMPv1 Overview

RFC 1112, Host Extensions for IP Multicasting:

- **Membership queries:**
  - Querier sends IGMP query messages to 224.0.0.1 with TTL = 1.
  - No special mechanism by which a host can leave a group.
  - Query interval is 60 to 120 seconds.
- **Membership reports:**
  - IGMP report sent by one host suppresses sending by others.
  - Restricted to one report per group per LAN.
  - Unsolicited reports sent by host when it first joins the group.

IGMPv2 Overview

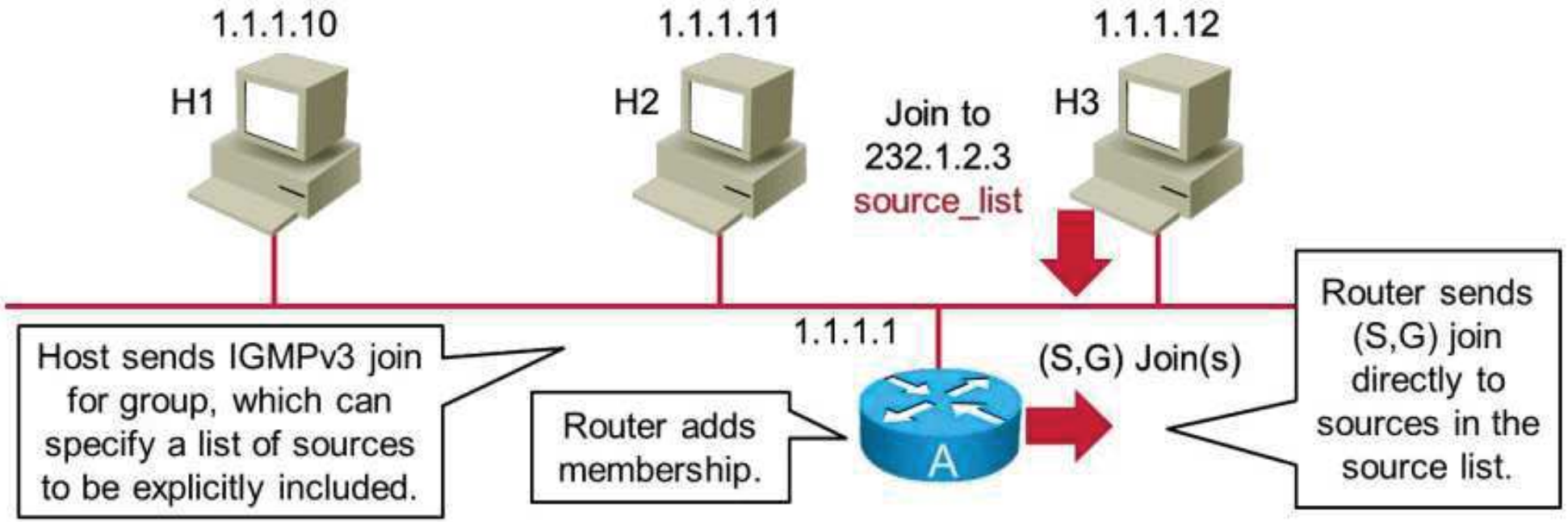
RFC 2236:

- Group-specific query:
  - Router sends group-specific query to make sure there are no members present before stopping the forwarding of data for the group.
- Leave-group message:
  - Host sends leave message if it leaves the group.
  - When router receives a leave message, it queries the network segment if there are any remaining multicast group receivers.
- Querier election mechanism:
  - On multiaccess networks, an IGMP querier router is elected based on the lowest IP address. Only the querier router sends queries.
- Query-interval response time:
  - General queries specify the maximum response time, which informs hosts of the maximum time within which a host must respond to a general query.
- Backward-compatible with IGMPv1.

IGMPv3 Overview

RFC no.3376 specifies:

- Enables hosts to listen only to a specified subset of the hosts sending to the group.
- Allows routers in sparse mode to build source distribution trees directly (avoiding RPs entirely).

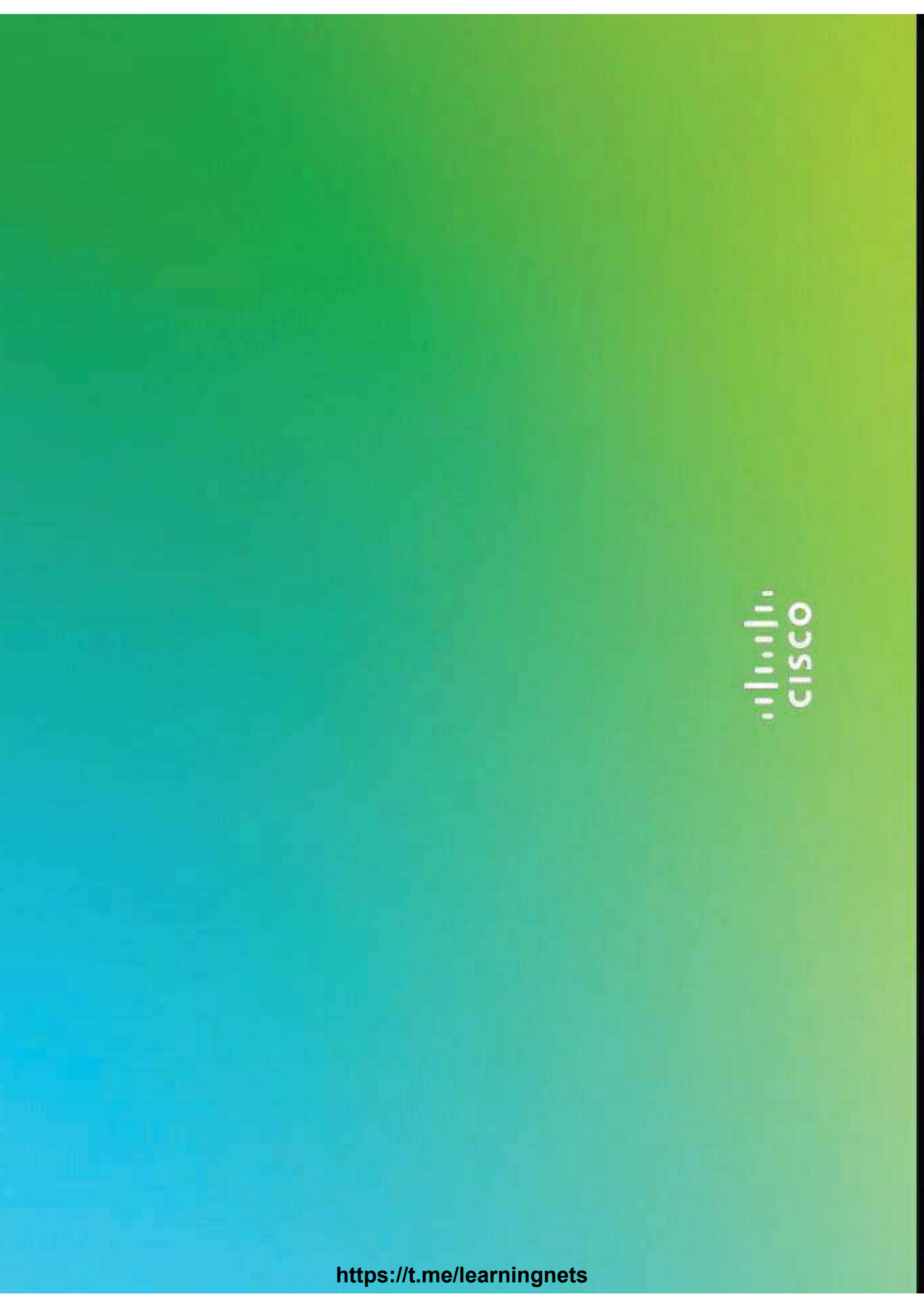


Summary

- RPF is used in multicast to ensure loop-free routing.
- Multicast distribution trees can be rooted at the source or at rendez-vous point.
- There are two notations in multicast for source – group pairings:
  - (S,G) source routed trees group one particular source with one group.
  - (*,G) RP routed trees group any source with a particular group.
- Different protocols can be used for routing, but PIM is most widely used.
- PIM can be used in dense or in sparse mode.
- PIM among other protocols is used for intradomain routing.
- Interdomain routing is accomplished with MP-BGP or MSDP.

Summary (Cont.)

- In addition to regular high availability enhancements, multicast requires RP redundancy.
- NSF and SSO also support multicast forwarding.
- Modified GenID value after a switchover triggers resending of PIM joins from neighbors.
- Between receivers and designated routers IGMP is used to join groups.
- Multicast routers periodically send membership queries to the all-hosts multicast address.
- A group-specific query that was added in IGMPv2 allows the router to query its members only in a single group instead of all groups.
- The main intention of IGMPv3 is to allow hosts to indicate that they only want to receive traffic from a particular source within a multicast group.



Defining Multicast on the LAN

Multicast Overview

Mapping Multicast IP Addresses to MAC Addresses

Layer 3 Multicast Addressing

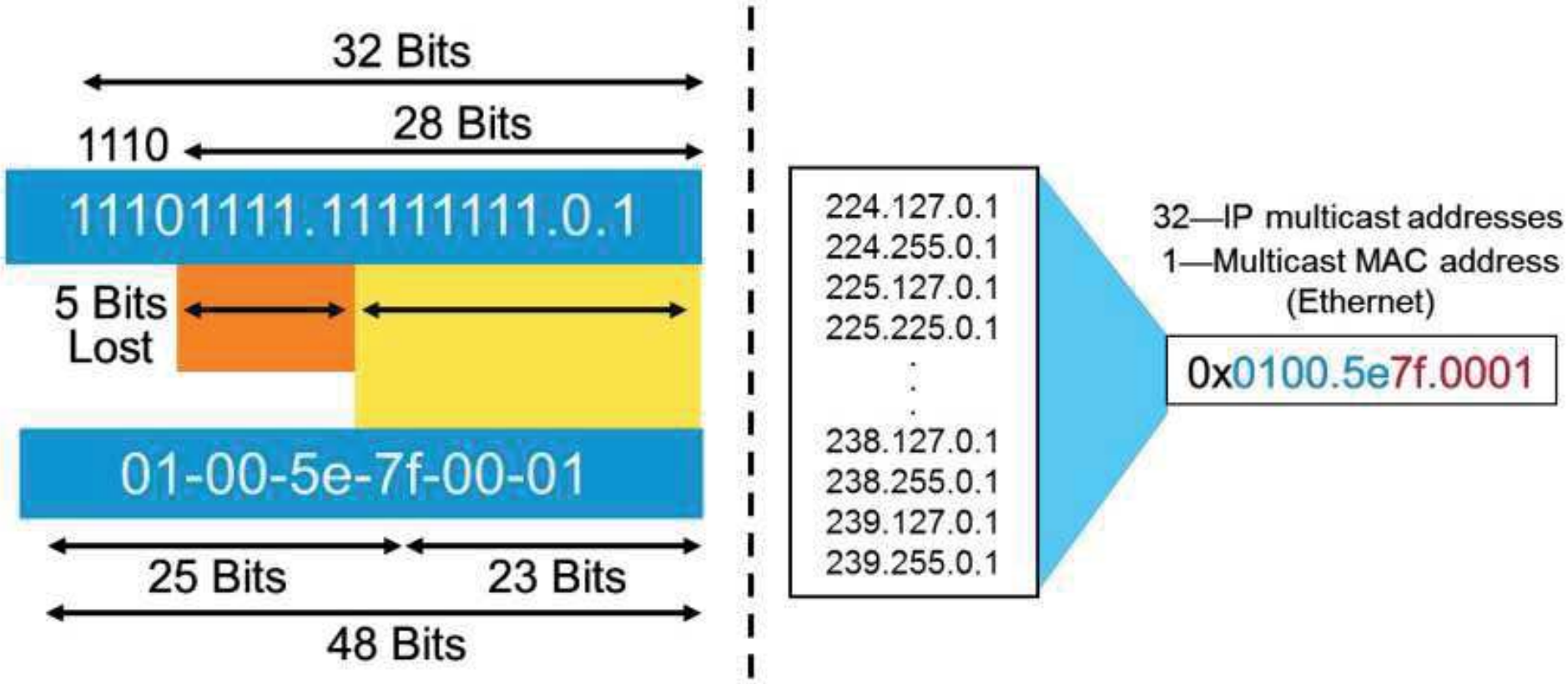
- IP Class D group addresses 224.0.0.0 to 239.255.255.255.
- High-order bits of 1110 (224.0.0.0/4).
- Special reserved group addresses 224.0.0.0 to 224.0.0.255 (TTL = 1).

Address	Description
224.0.0.1	All systems on this subnet
224.0.0.2	All routers on this subnet
224.0.0.4	DVMRP routers
224.0.0.5	OSPF all routers
224.0.0.6	OSPF designated routers
224.0.0.13	PIMv2 routers

Mapping Multicast IP Addresses to MAC Addresses (Cont.)

Layer 2 Multicast Addressing:

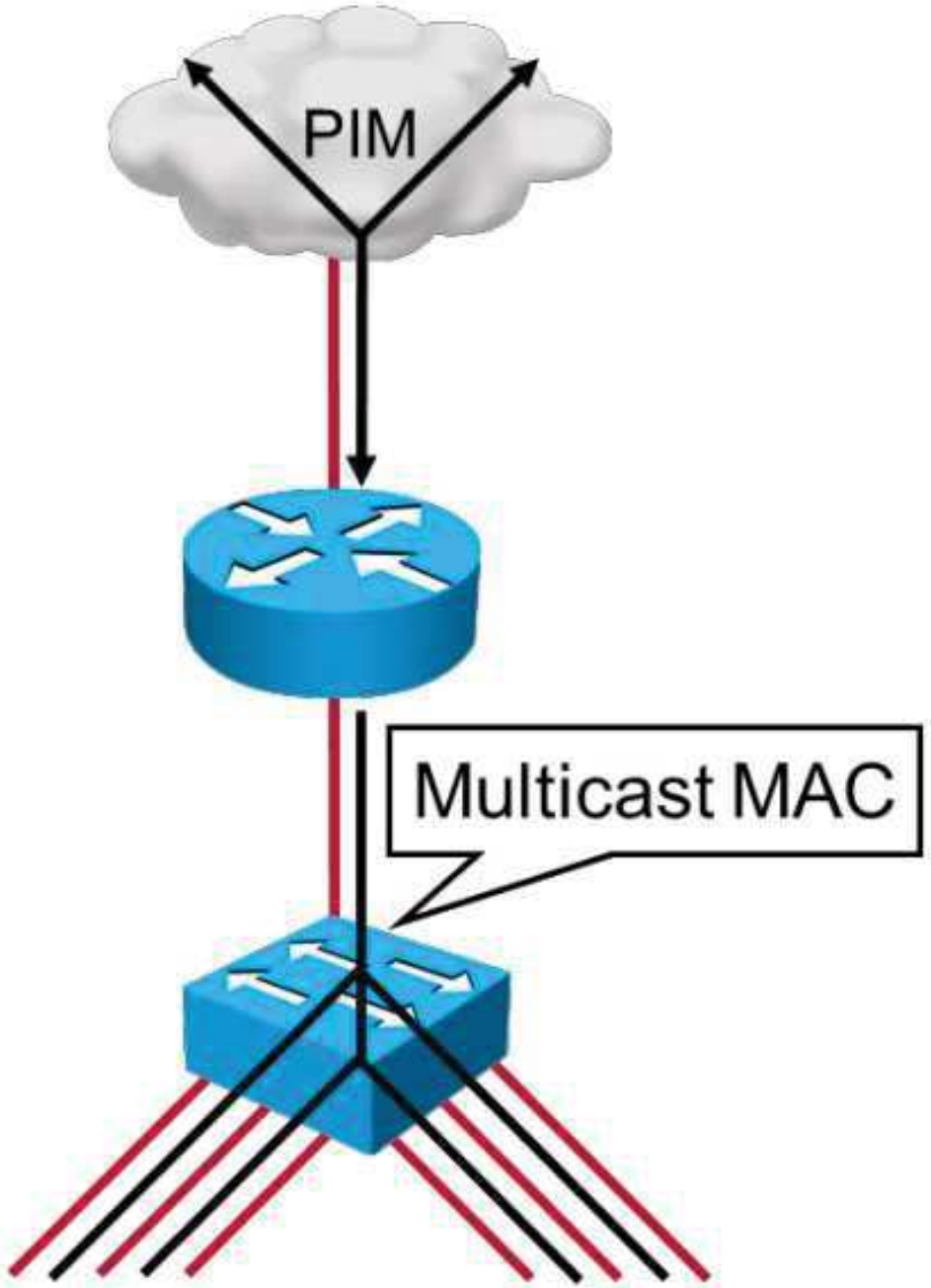
- IP multicast MAC address mapping (Ethernet)
- Be aware of the 32:1 address overlap



Layer 2 Multicast Frame Switching

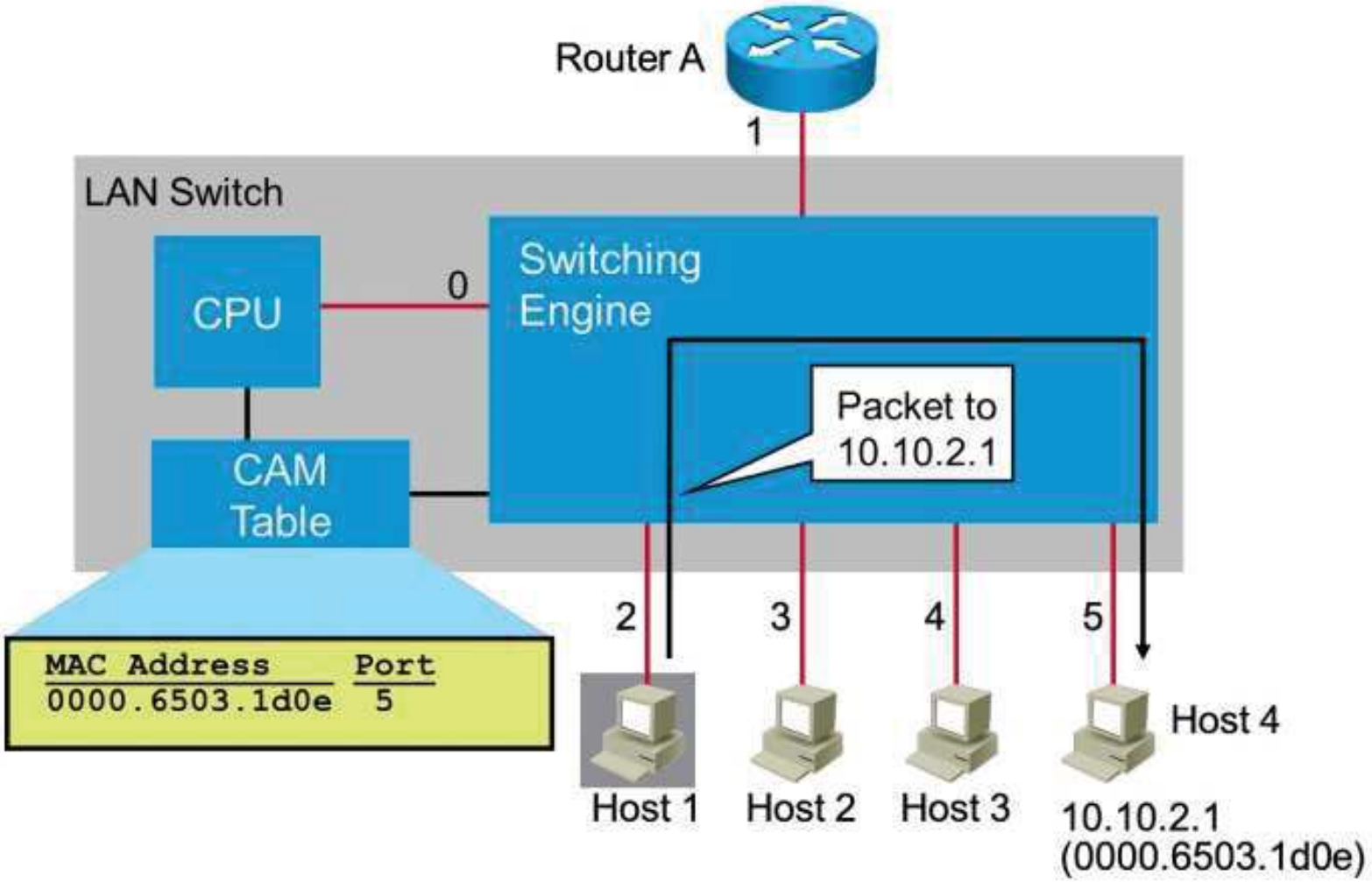
Layer 2 flooding of multicast frames:

- Typical Layer 2 switches treat multicast traffic as unknown or broadcast and must flood the frame out to every port.
- Static entries may sometimes be set to specify which ports must receive which group(s) of multicast traffic.
- Dynamic configuration of these entries may reduce user administration.



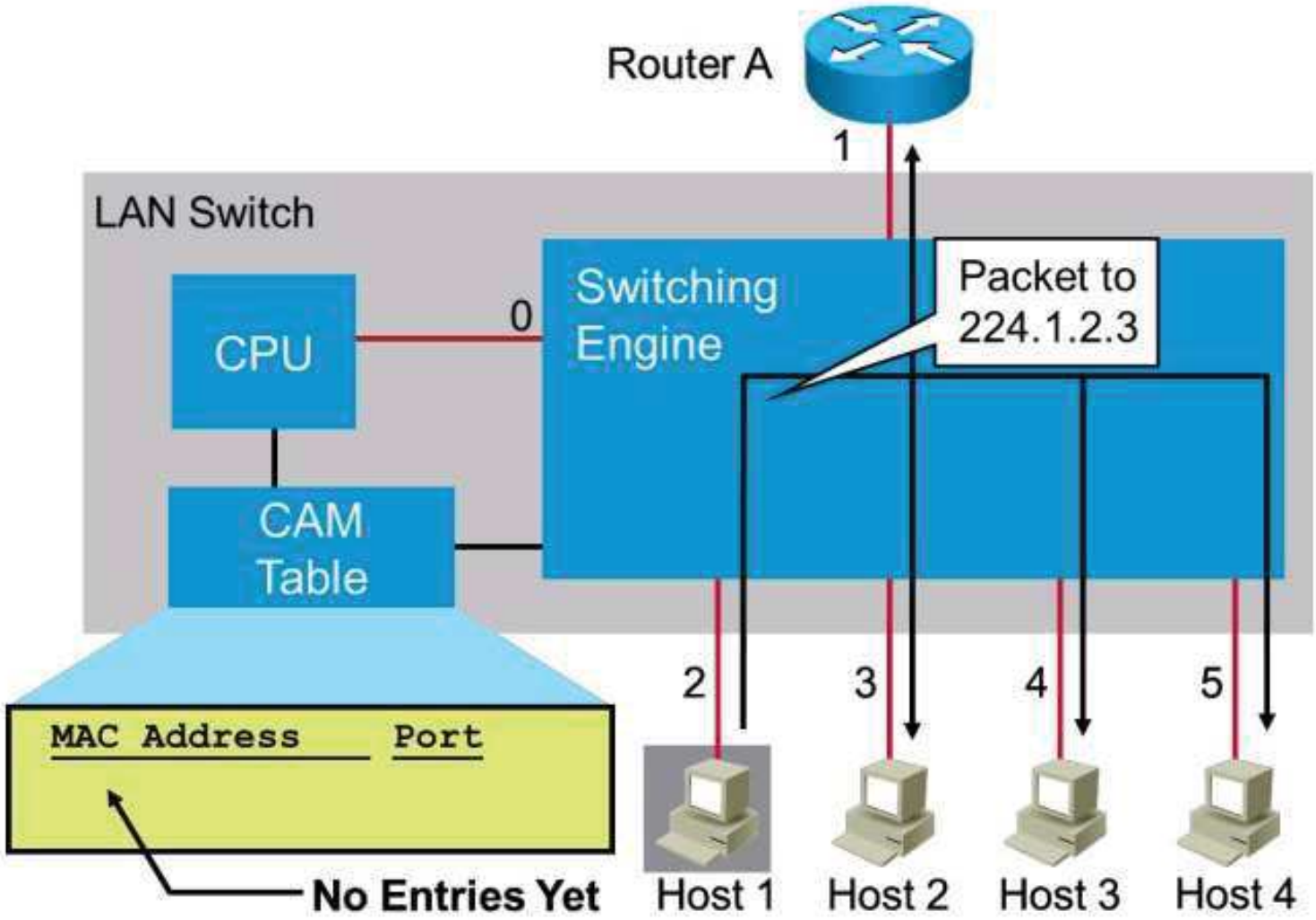
Layer 2 Multicast Frame Switching (Cont.)

Layer 2 Unicast Forwarding:



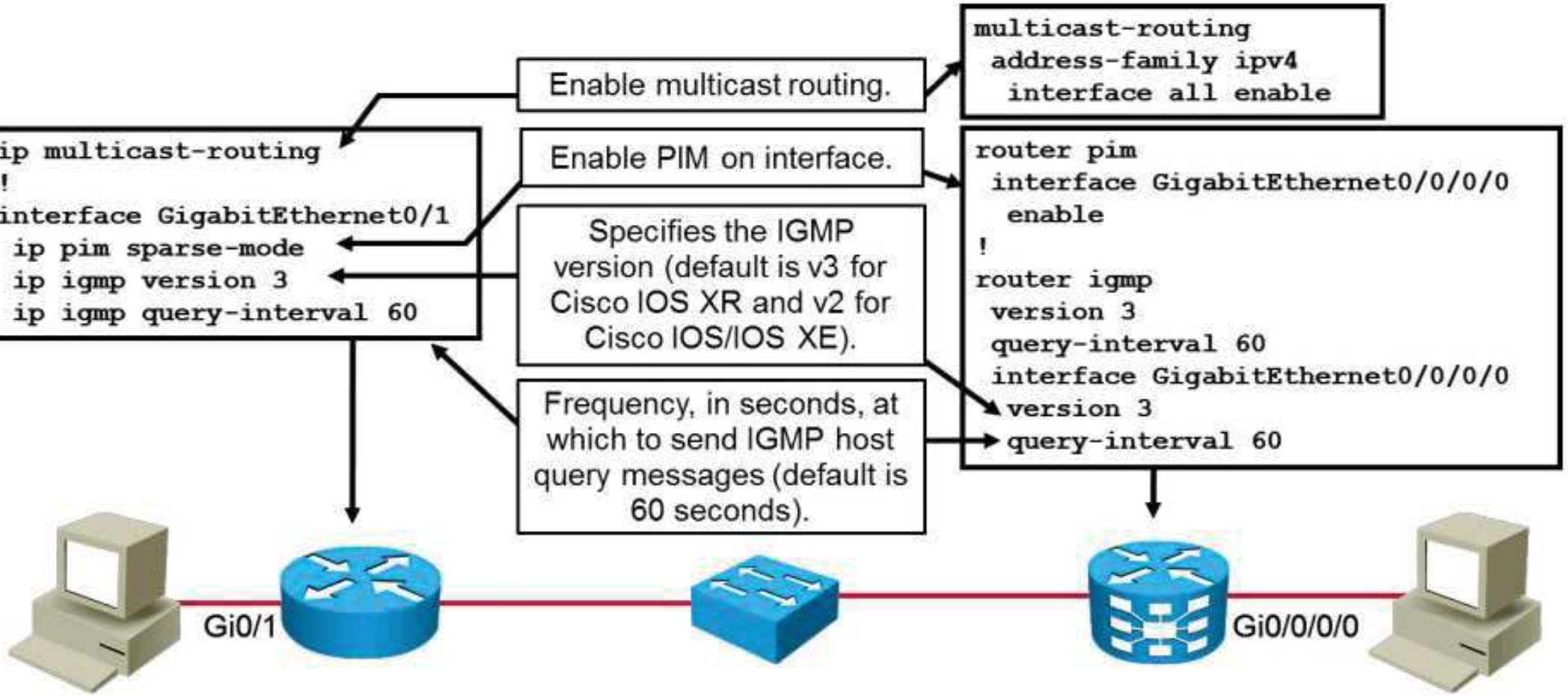
Layer 2 Multicast Frame Switching (Cont.)

Layer 2 Multicast Forwarding:



Implementing IGMP

IGMP Configuration:



Implementing IGMP (Cont.)

```
RP/0/RSP0/CPU0:PE7# show igmp interface GigabitEthernet0/0/0/0
GigabitEthernet0/0/0/0 is up, line protocol is up
  Internet address is 192.168.107.70/24
  IGMP is enabled on interface
  Current IGMP version is 3
  IGMP query interval is 60 seconds
  IGMP querier timeout is 125 seconds
  IGMP max query response time is 10 seconds
  Last member query response interval is 1 seconds
  IGMP activity: 12 joins, 7 leaves
  IGMP querying router is 192.168.107.70 (this system)
```

Displays IGMP interface settings.

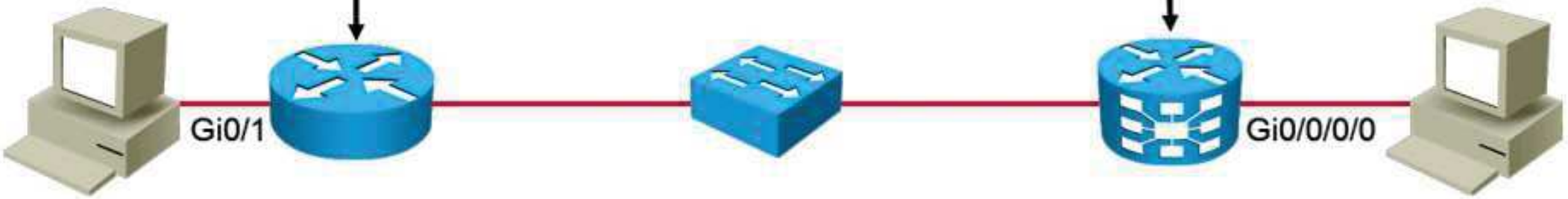
```
RP/0/RSP0/CPU0:PE7# show igmp groups
IGMP Connected Group Membership
Group Address   Interface                Uptime    Expires   Last Reporter
224.0.0.2       GigabitEthernet0/0/0/0   2d19h      never     10.7.1.1
```

Displays the IGMP groups that are registered on the router.

IGMP join-group and static-group

```
interface GigabitEthernet0/1
 ip igmp join-group group-address
 ip igmp static-group group-address
```

```
router igmp
 interface GigabitEthernet0/0/0/0
  join-group group-address
  static-group group-address
```



join-group command:

- Router joins multicast group.
- Populates IGMP cache.
- Sends IGMP report.
- Router joins a group.
- CPU receives data.

static-group command:

- Traffic forwarded on the interface.
- Populates IGMP cache.
- PIM join only if configured on the designated router.
- No CPU impact.

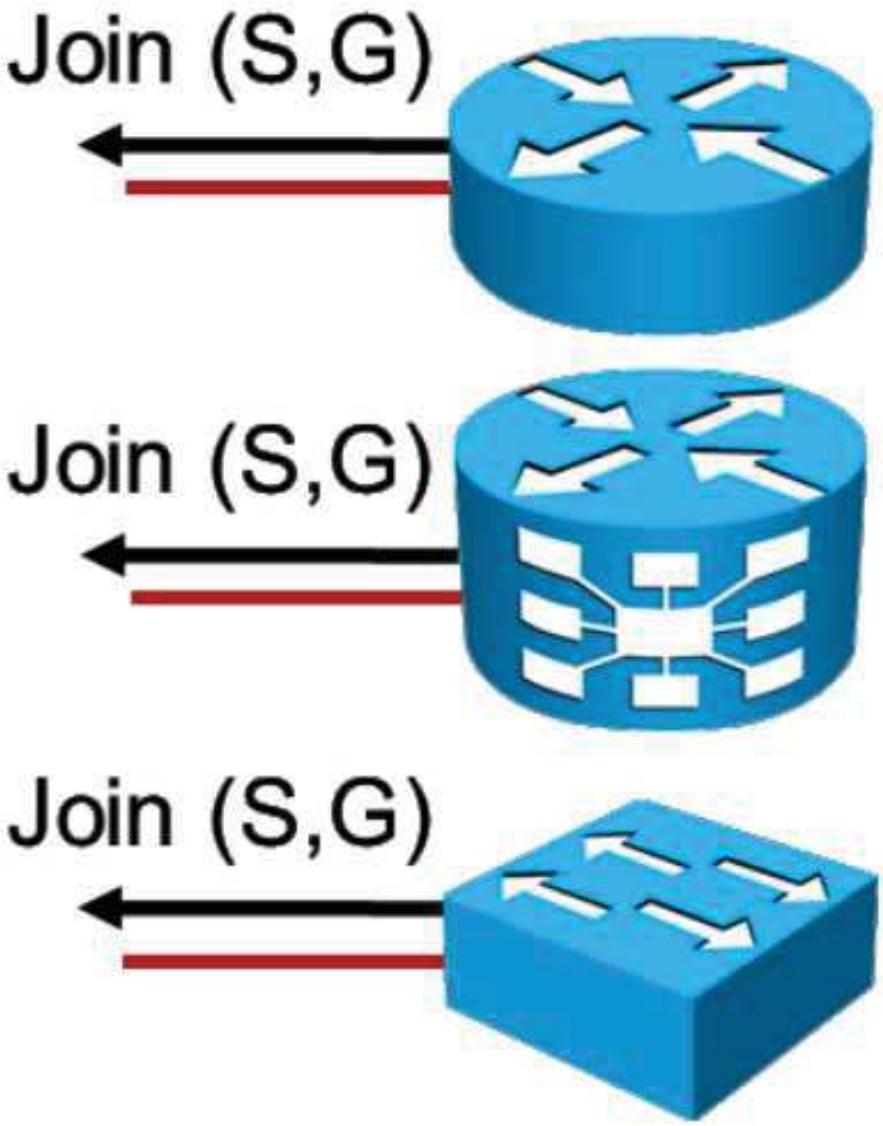
IGMPv3 Host Stack Feature

IGMPv3 Host Stack characteristics:

- Enables routers and switches to function as multicast network endpoints or hosts.
- Adds INCLUDE mode capability to the IGMP version 3 host stack for SSM groups.

Restriction:

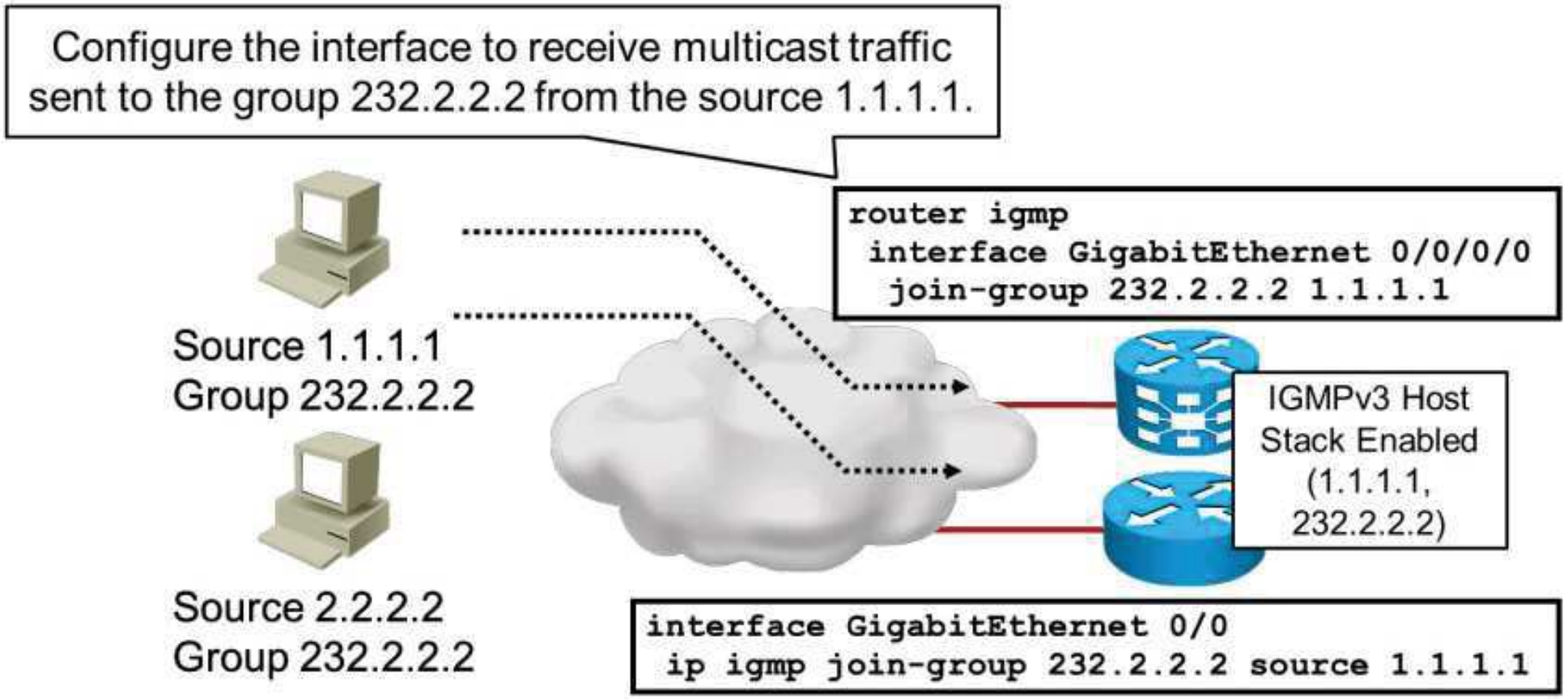
- IGMPv3 must be enabled.



IGMPv3 Host Stack Feature (Cont.)

Enabling the IGMPv3 Host Stack:

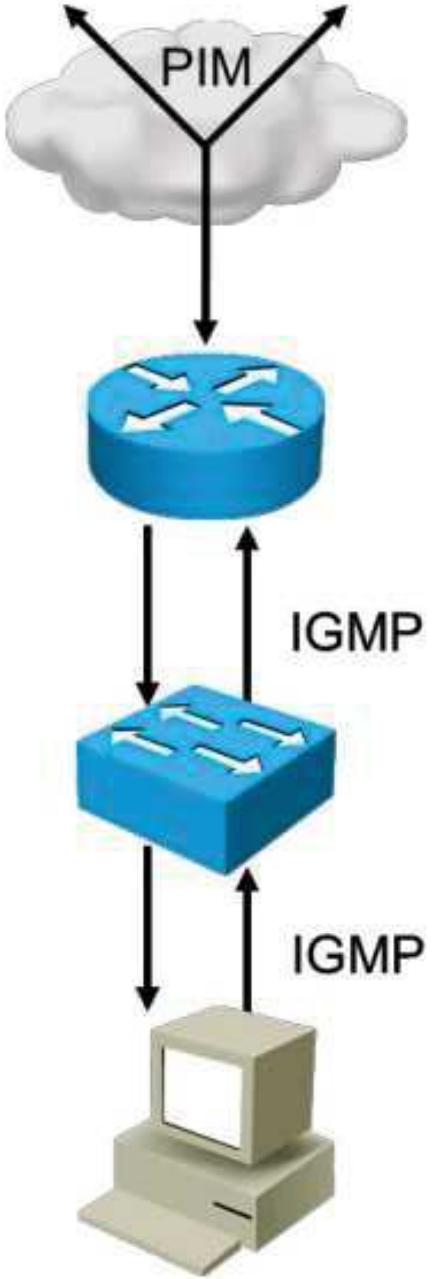
- Applications can leverage SSM as the preferred method.
- Assists in troubleshooting.



Configuring IGMP Snooping

IGMP snooping characteristics:

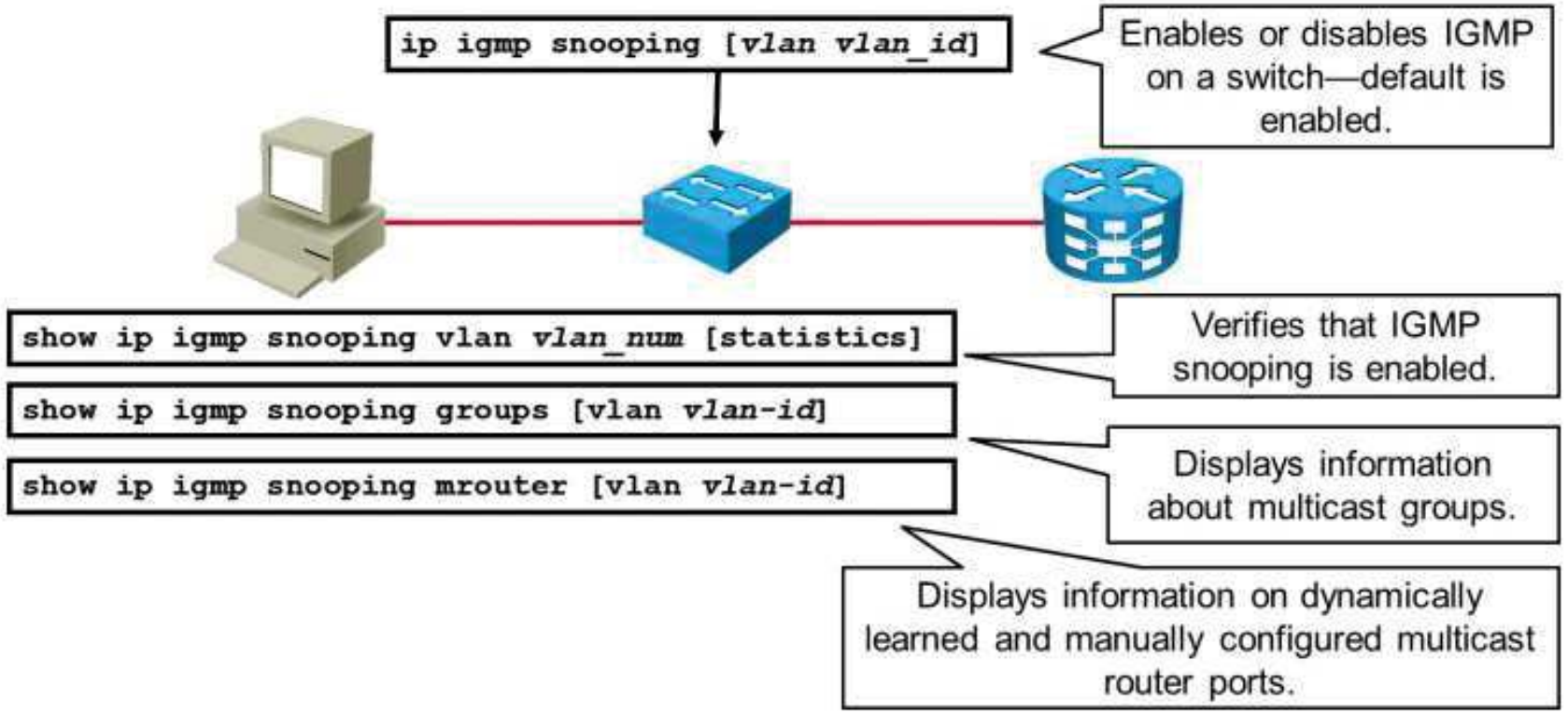
- Switches become IGMP-aware (IGMP packets intercepted).
- The switch must examine contents of IGMP messages to determine which ports want which kind of traffic:
  - IGMP membership reports.
  - IGMP leave messages.
- Effect on switch:
  - Must process all Layer 2 multicast packets.
  - Administration load increases with multicast traffic load.
  - Requires special hardware to maintain throughput.



Configuring IGMP Snooping (Cont.)

Configuring IGMP Snooping:

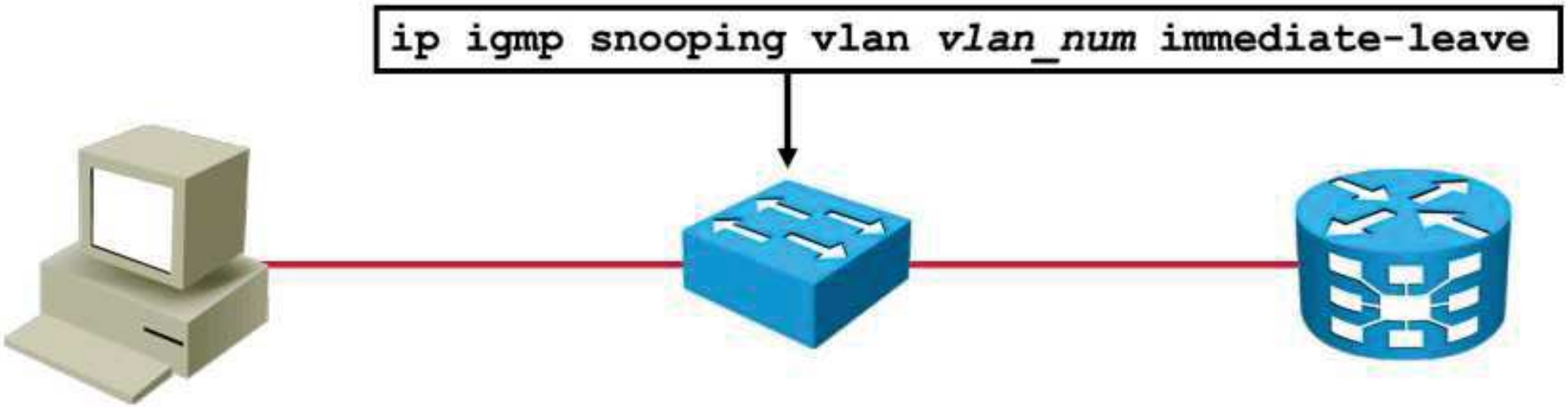
- IGMP snooping in switches enables automatic detection of multicast groups.
- Configuration is needed on switches only: transparent to routers and multicast hosts.



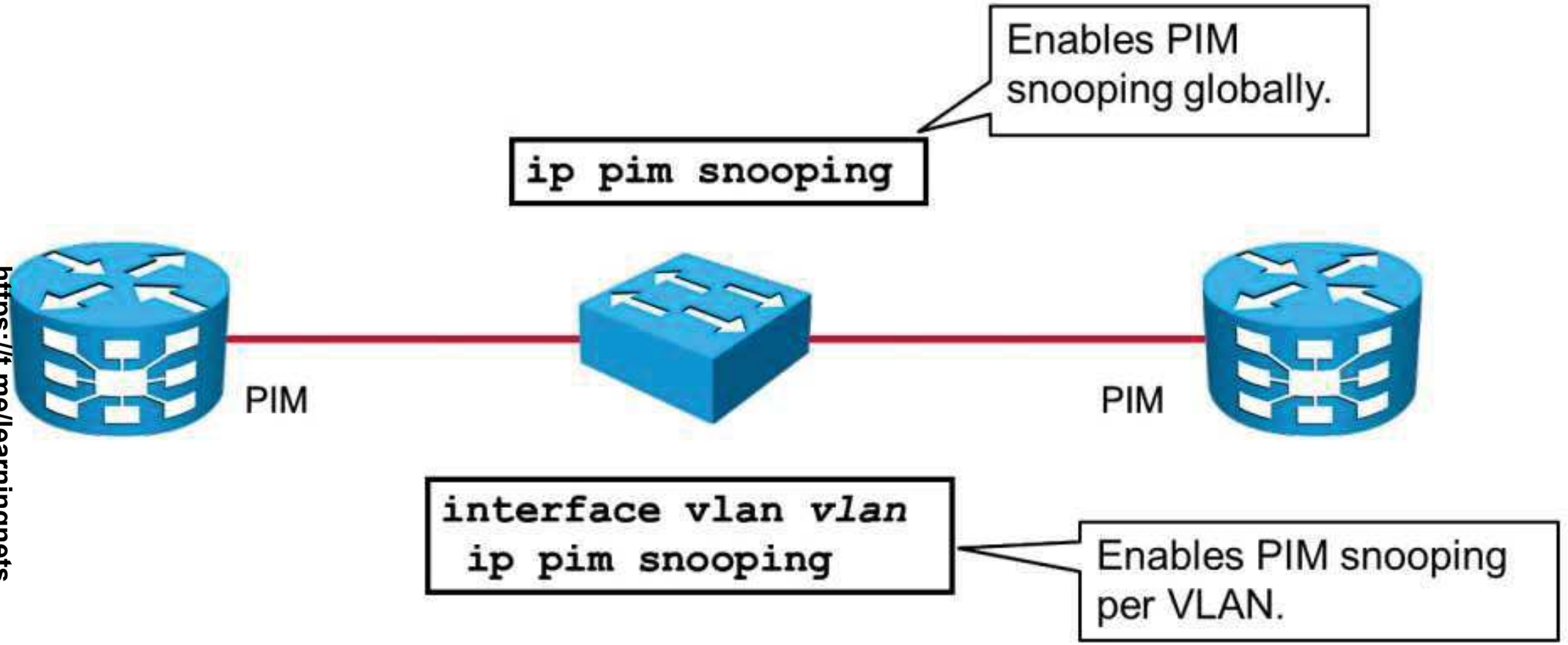
IGMP Fast-Leave in a Switch

IGMP fast-leave in a switch:

- IGMP snooping fast-leave processing is disabled by default.
- IGMP snooping fast-leave processing allows IGMP snooping to remove an interface from the forwarding table entry without first sending out an IGMP query on the interface.
- Enable IGMP snooping fast-leave processing only on VLANs where only one host is connected to each switch port.



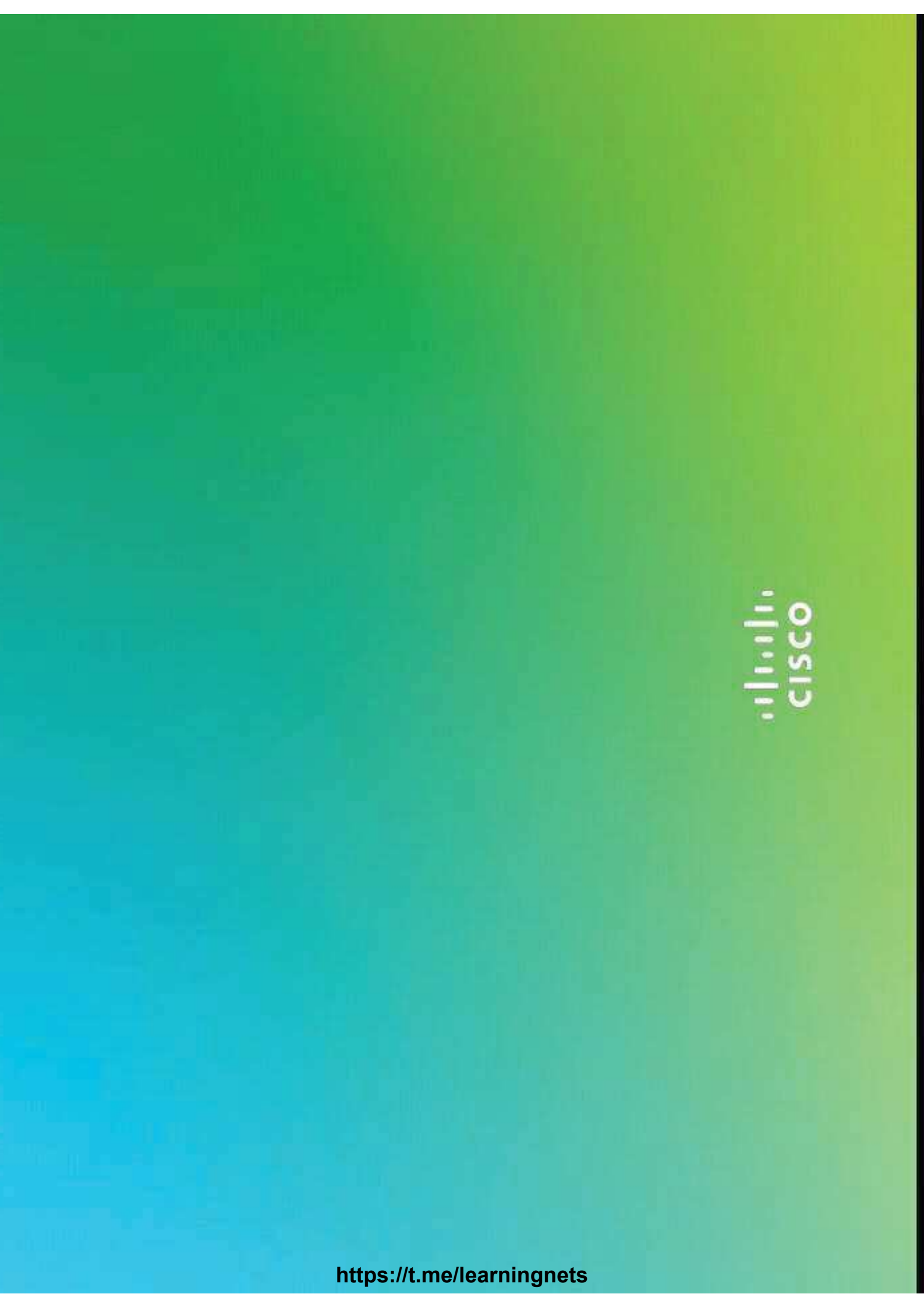
PIM Snooping



With PIM snooping, the LAN switch learns which multicast traffic the PIM routers are willing to receive.

Summary

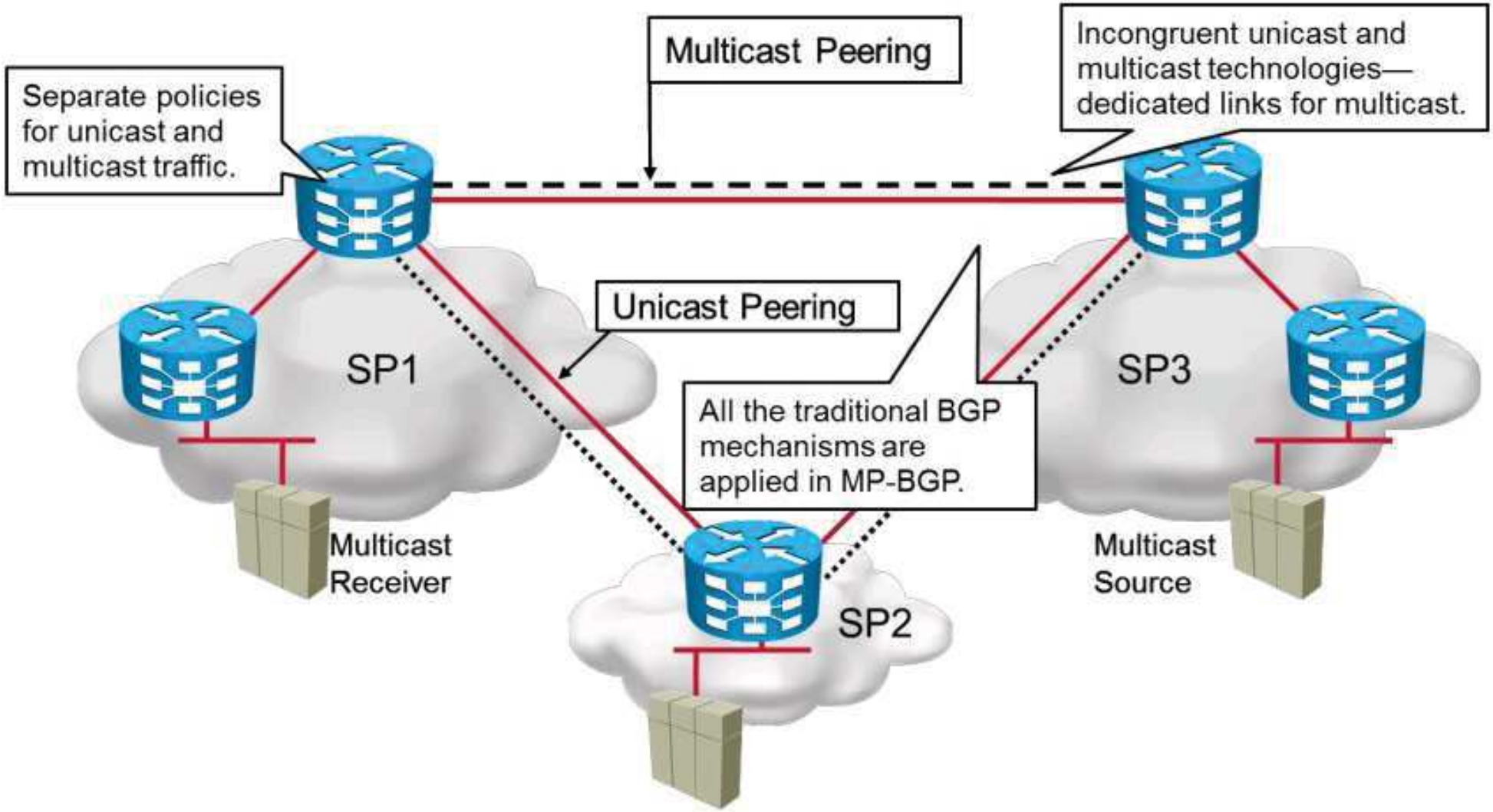
- Mapping from Layer 3 to Layer 2 multicast addresses can cause various problems.
- Multicast traffic is similar to broadcast traffic on layer 2.
- IGMP is used by multicast receivers to join and leave multicast groups.
- Join-group and static-group commands can be used to make a router to be a member of a group.
- The IGMPv3 host stack feature enables routers and switches to function as multicast network endpoints or hosts.
- IGMP snooping is transparent to the multicast hosts.
- With PIM snooping, the LAN switch learns which multicast traffic routers are willing to receive.



Populating the Mroute Table

Multicast Overview

The Mroute Table

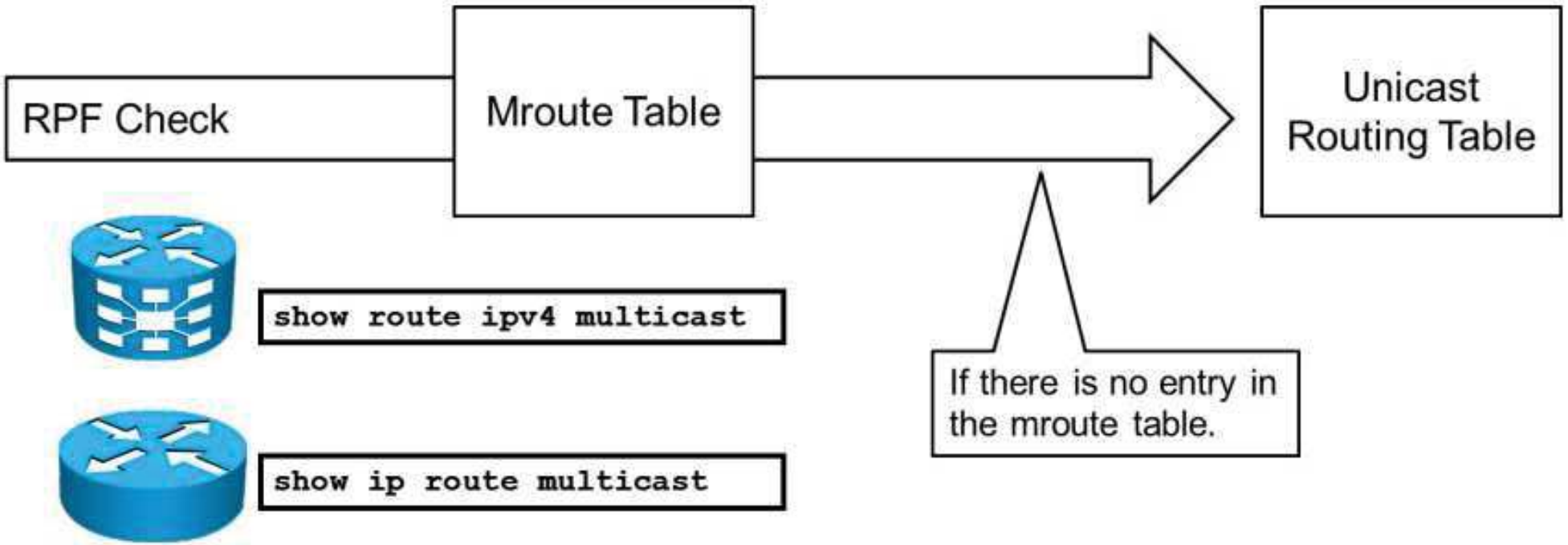


- Service providers SP1 and SP3 exchange multicast traffic.
- SP2 has no need for multicast traffic.

The Mroute Table (Cont.)

Mroute table characteristics:

- RPF check uses mroute table.
- Routers still require PIM to build the multicast distribution trees.
- Mroute table = source multicast routing table.



Multiprotocol BGP

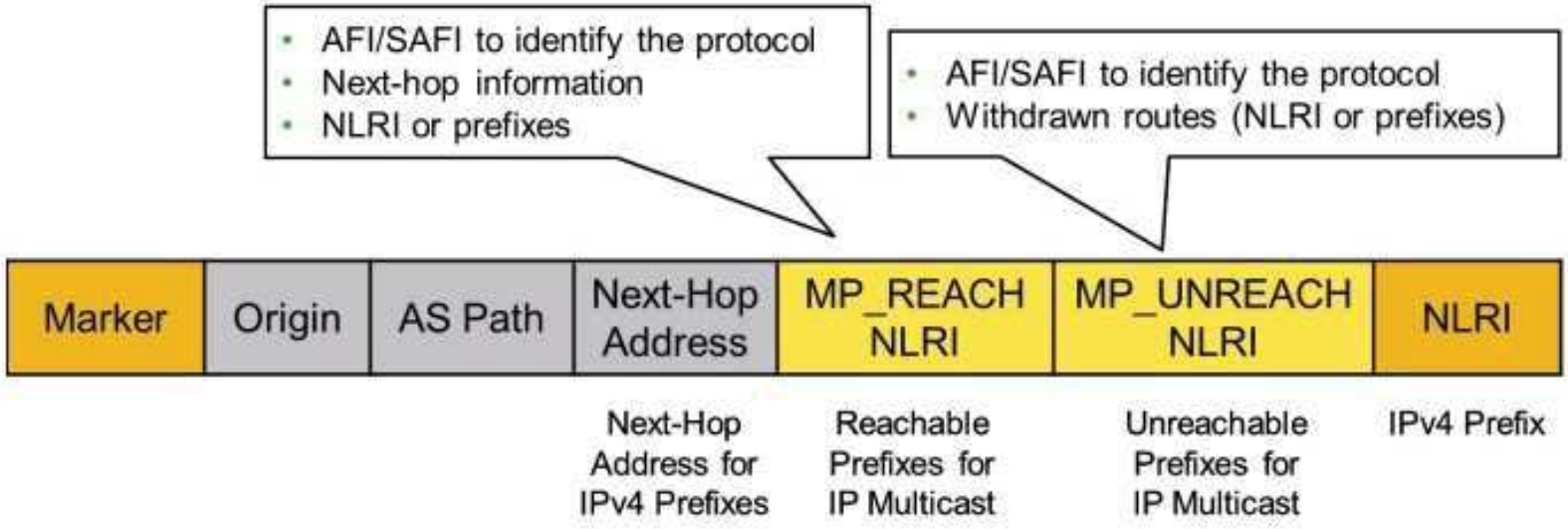
Multiprotocol Extensions to BGP:

- Multiprotocol extensions to BGP (MP-BGP, RFC 4760) allow for different protocols to be carried across the same BGP session.
- Address family is a network layer protocol identifier.
- AFI is a 16-bit value.
- MP-BGP uses an additional SAFI, which is an 8-bit value.
- Address family values used with MPBGP and IP multicast today:
  - 1/1 IPv4 unicast
  - 1/2 IPv4 multicast
  - 1/3 IPv4 unicast and multicast
- The address families are treated separately (as completely different protocols).
- The separate treatment of unicast and multicast allows for separate topologies and policies.

Multiprotocol BGP (MP-BGP) (Cont.)

MP-BGP and IP Multicast:

- MP-BGP can carry routing information for any Layer 3 protocol.
- MP-BGP for IP multicast:
  - MBGP
  - The unicast routes carried in updates are used for multicast purposes—for example, RPF checks to multicast sources or RPs
- Two new path attributes in MP-BGP: MP_REACH_NLRI and MP_UNREACH_NLRI



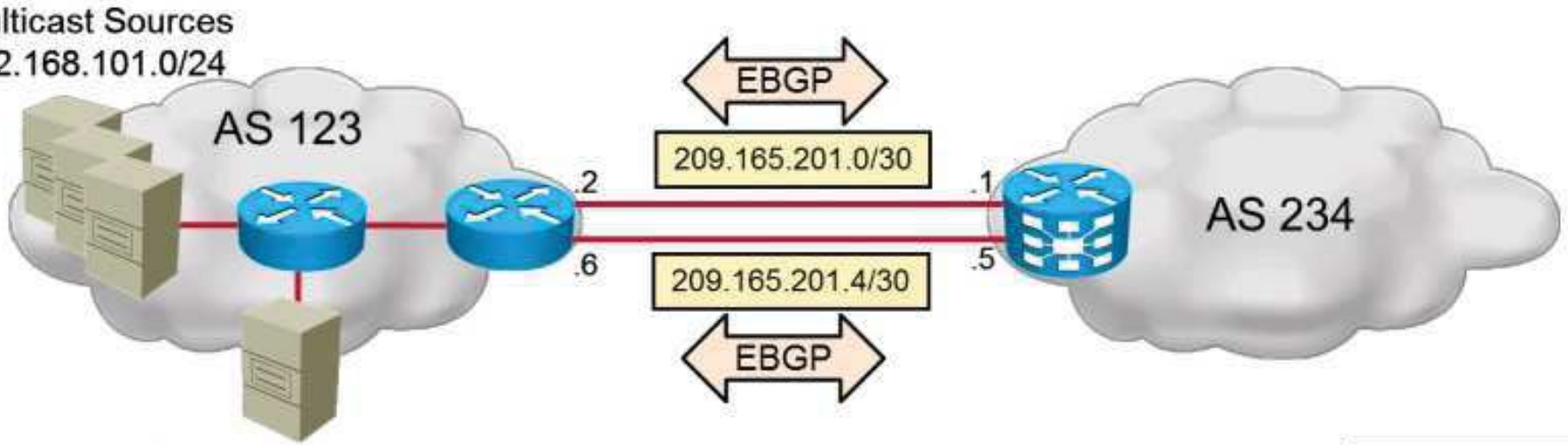
MP-BGP Capabilities Negotiation

BGP peers negotiate about their capabilities before exchanging prefix information:

- A BGPv4 session between neighbors starts with an exchange of open messages.
- Multiprotocol extensions are negotiated as part of open messages.
- An optional parameter is used for negotiation of capabilities.
- Only those capabilities supported by both routers are used.
- If one of the routers does not understand the capabilities parameter:
  - The session may be terminated (RFC 4271).
  - Cisco IOS/IOS XE/IOS XR Software backs off and reopens with no capability parameters.

MP-BGP Multicast Configuration

MP-BGP Multicast Configuration Example:




```
router bgp 123
 neighbor 209.165.201.1 remote-as 234
 neighbor 209.165.201.5 remote-as 234
 address-family ipv4 unicast
  neighbor 209.165.201.1 activate
  no neighbor 209.165.201.5 activate
!
 address-family ipv4 multicast
  neighbor 209.165.201.5 activate
```

Activate address family for neighbor.

```
router bgp 234
 address-family ipv4 unicast
 address-family ipv4 multicast
 !
 neighbor 209.165.201.2
  remote-as 123
  address-family ipv4 unicast
 neighbor 209.165.201.6
  remote-as 64501
  address-family ipv4 multicast
```

Activate address family for BGP.

Activate address family for neighbor.

MP-BGP Multicast Verification

```
RP/0/RSP0/CPU0:PE1# show bgp ipv4 all summary
For address family: IPv4 Unicast
<… output omitted …>
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
209.165.201.2   4   123      38      37        2    0    0 00:35:12          1
For address family: IPv4 Multicast
<… output omitted …>
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
209.165.201.6   4   123      42      41        2    0    0 00:34:23          1
```

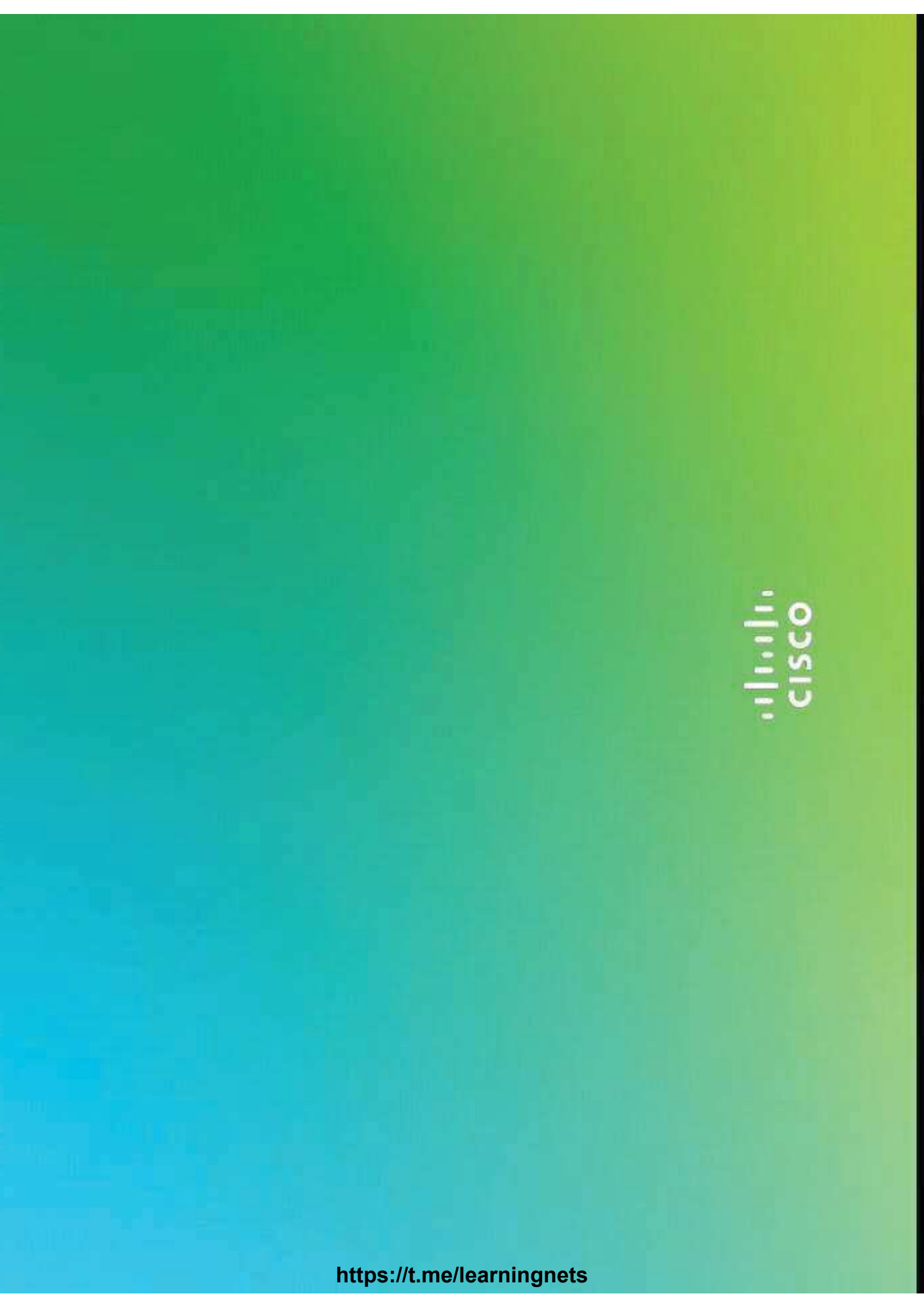
- Displays BGP neighbors for IPv4 unicast and multicast address families.

```
RP/0/RSP0/CPU0:PE1# show bgp ipv4 multicast
BGP table version is 2, local router ID is 209.165.201.5
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
     Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.101.0    209.165.201.6            0             0 123 i
```

- Displays BGP routes for IPv4 multicast address family.

Summary

- Multicast and unicast topology do not need to be congruent. The RPF check uses the mroute table. If there is no mroute table entry, the unicast routing table is used instead.
- In unicast and multicast incongruent topologies, MP-BGP is necessary.
- MP-BGP will negotiate neighbor capabilities when establishing peering session.
- Multicast in MP-BGP is configured under separate address family.



Module Summary

- The biggest benefit of multicasting is sending a single packet to multiple receivers. Inside the multicast network, various multicast routing protocols are used.
- Routers perform RPF checks when multicast packets arrive.
- IGMP and MLD are group reporting protocols and can be snooped by switches.
- MP-BGP is needed to provide proper information for RPF checks and similar IP multicast-related operations. In incongruent topologies, MP-BGP is necessary.

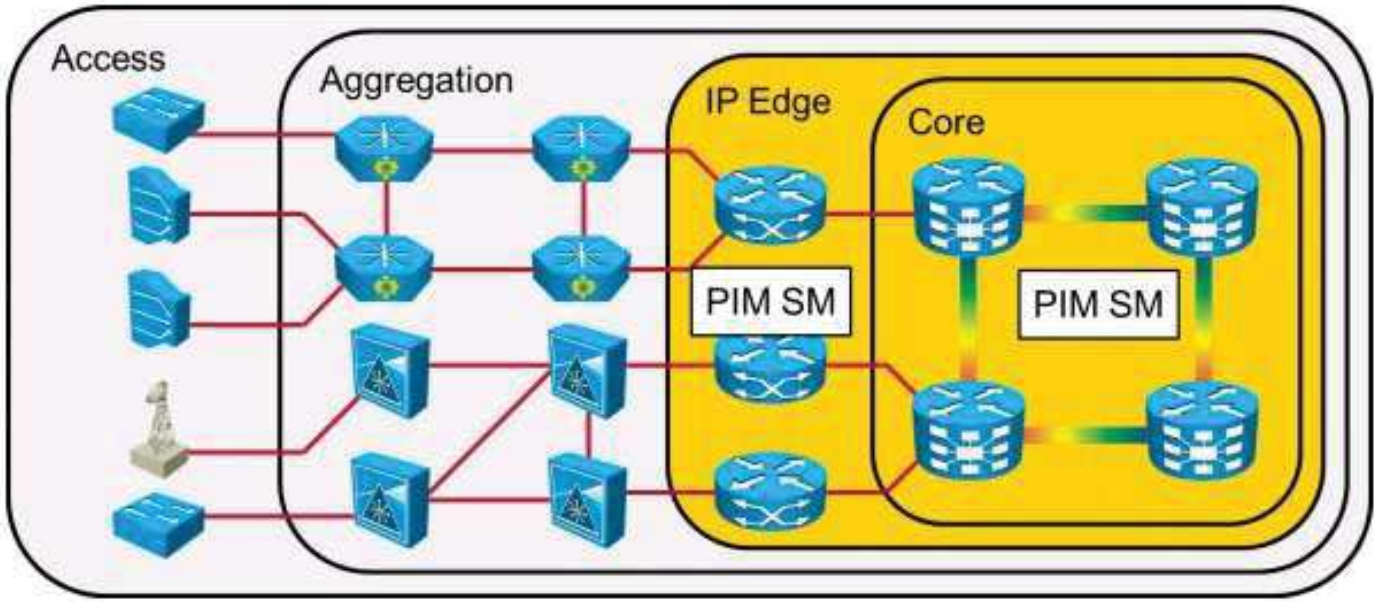
Intradomain and Interdomain Multicast Routing

Deploying Cisco Service Provider Advanced Network Routing (SPADVROUTE v1.2)

Introducing PIM-SM Protocol

Intradomain and Interdomain Multicast Routing

PIM-SM in the Cisco IP NGN Infrastructure Layer



PIM-SM characteristics:

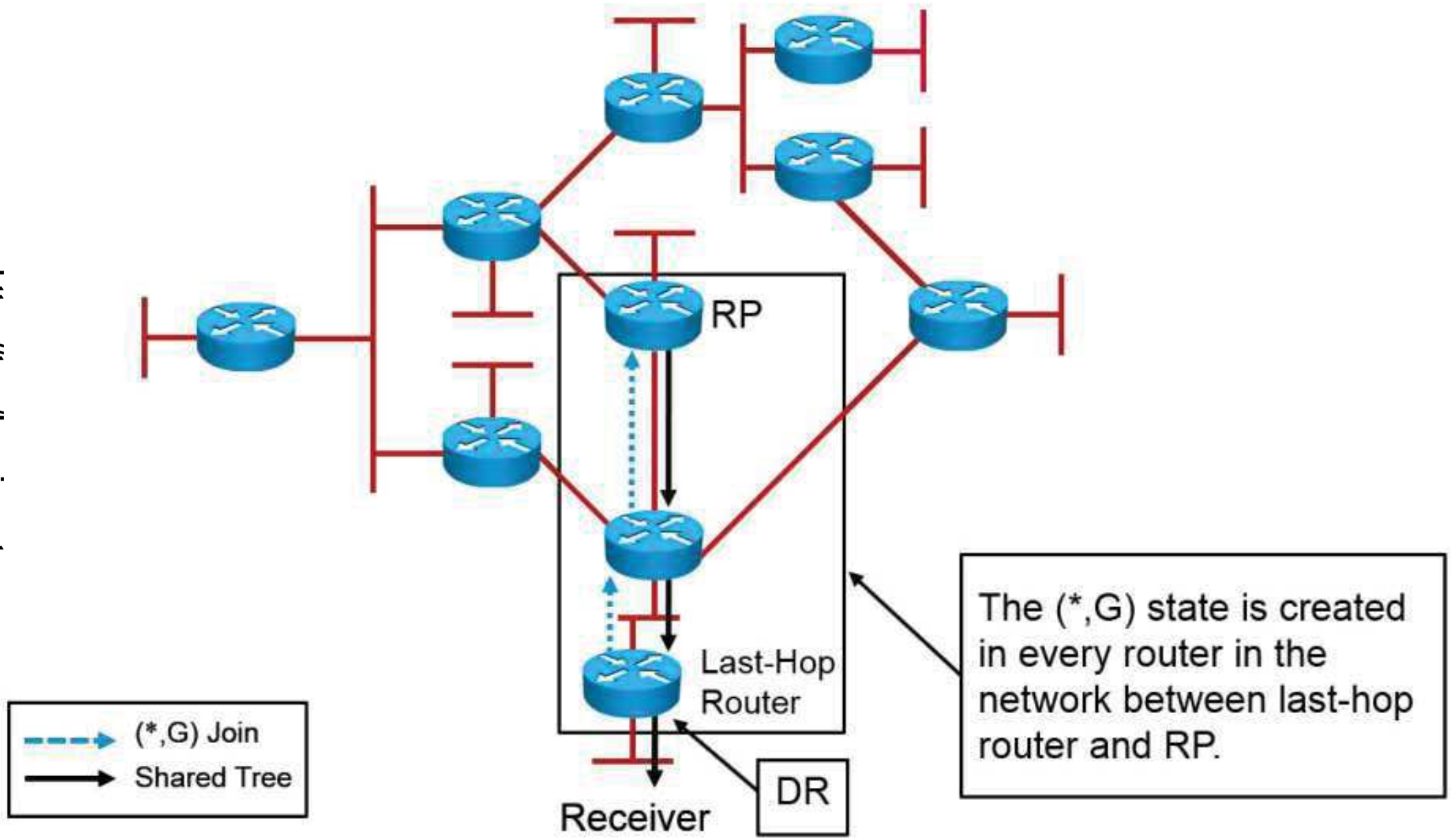
- Explicit join model.
- Receivers join to the RP.
- Senders register with the RP.
- Data flows down the shared tree and goes only to places that need the data from the sources.
- Last-hop routers can join source tree if the data rate exceeds threshold.

PIM-SM in the Cisco IP NGN Infrastructure Layer (Cont.)

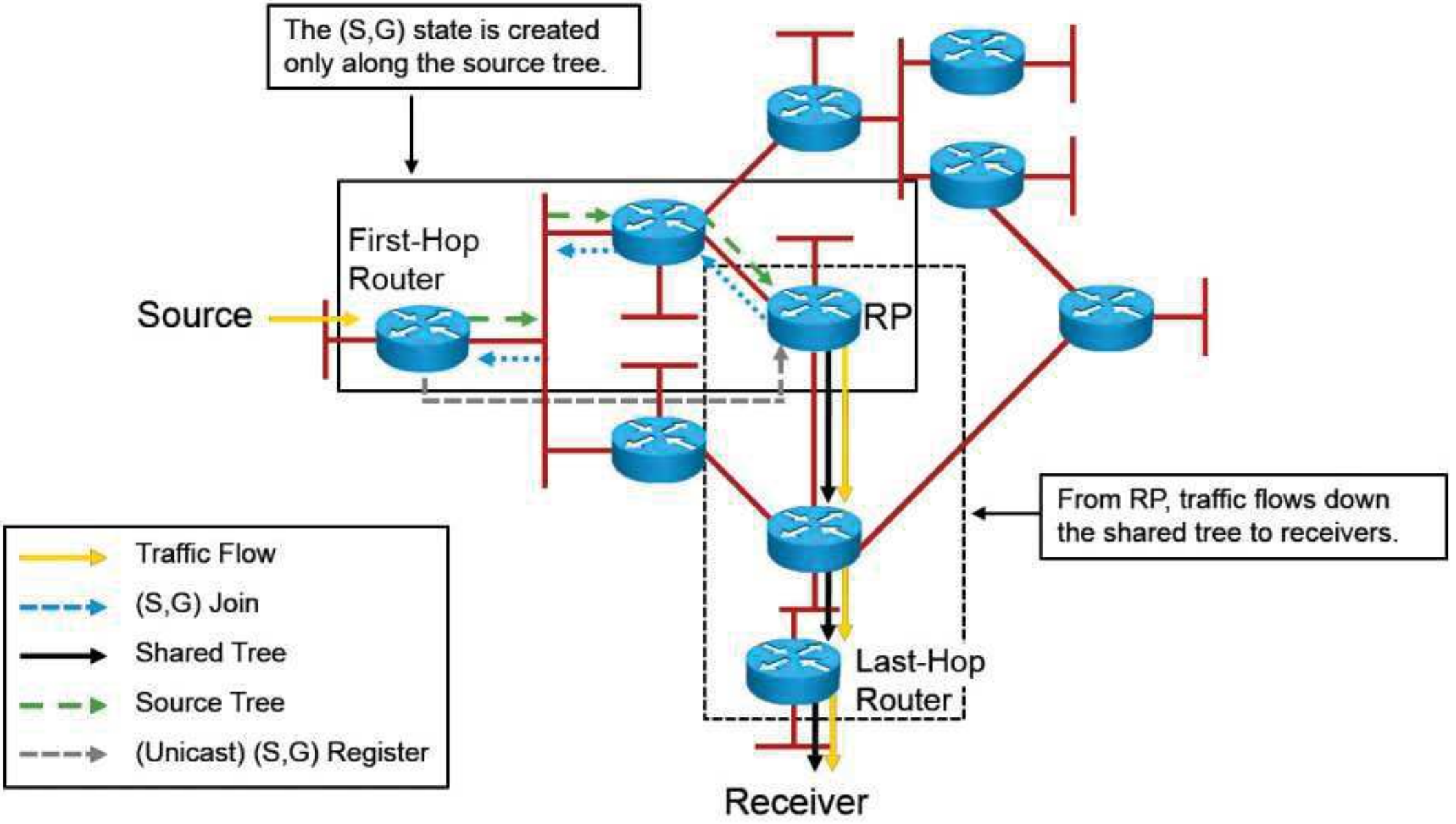
RPF check depends on tree type:

- For shared trees, uses RP address.
- For source trees, uses source address.
- Only one RP is chosen for a particular group.
- RP statically configured or dynamically learned:
  - Auto-RP
  - PIMv2 bootstrap mechanism
- Data forwarded based on the source state (S,G) if it exists; otherwise, use the shared state (*,G).

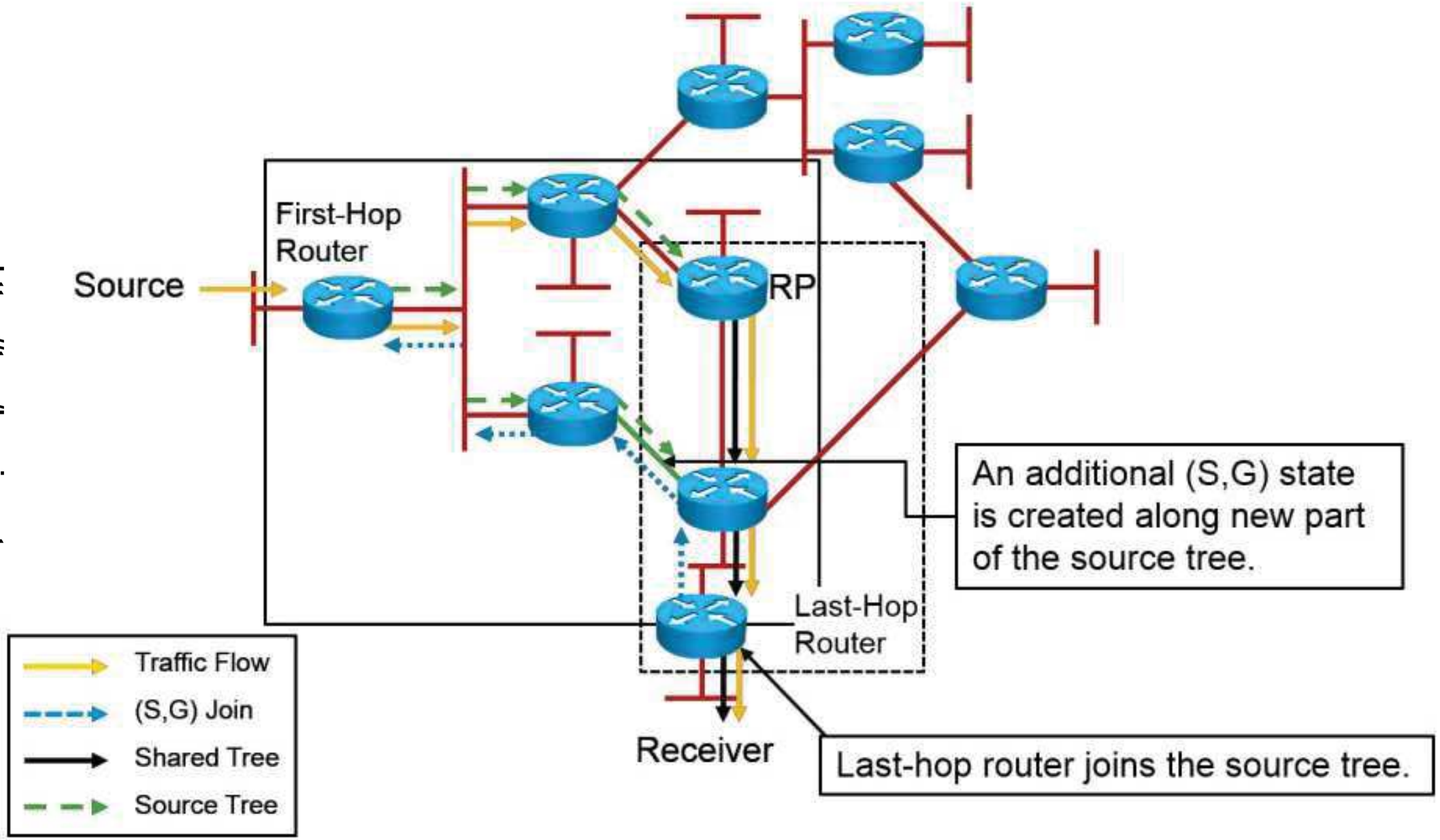
PIM-SM Shared Tree Join



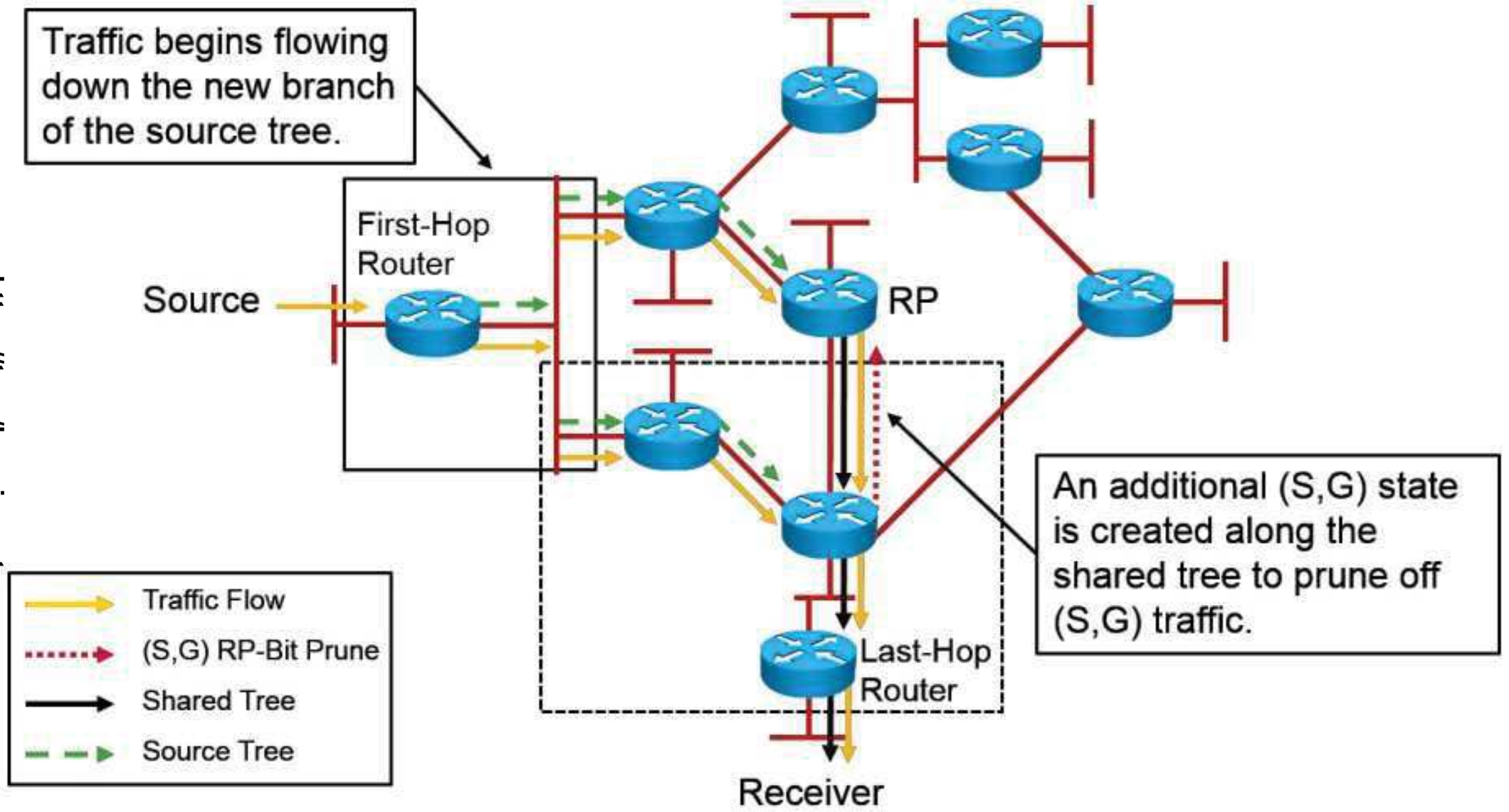
PIM-SM Sender Registration



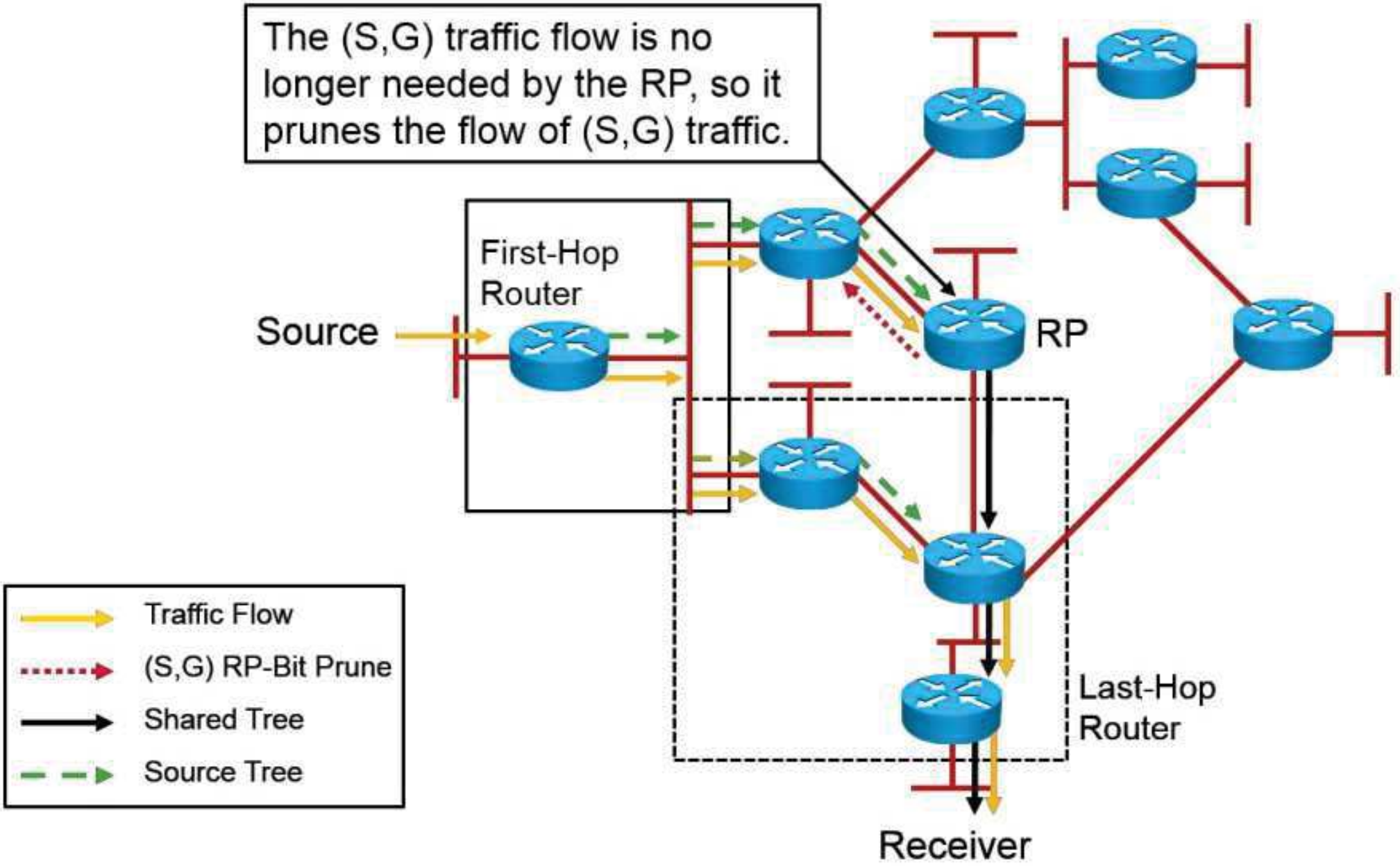
PIM-SM SPT Switchover



PIM-SM SPT Switchover (Cont.)



PIM-SM SPT Switchover (Cont.)



PIM-SM Packets

These are various PIM-SM control packets:

- PIM hello and PIM query in PIMv1.
- PIM join and prune.
- PIM register and register-stop.
- RP announcements:
  - PIMv2 bootstrap mechanism: Bootstrap router and candidate-RP advertisement.
  - Auto-RP mechanism: Cisco announce and Cisco discovery messages (Cisco addition to PIMv1).
- RP reachability (specific to Cisco).

PIM-SM State Information

PIM-SM state information characteristics:

- Describes the state of the multicast distribution trees as understood by the router at this point in the network.
- Represented by entries in the multicast routing table:
  - Used to make multicast traffic forwarding decisions.
  - Composed of (*,G) and (S,G) entries.
  - Each entry contains RPF information:
    - Incoming (such as RPF) interface
    - RPF neighbor (upstream)
  - Each entry contains an OIL:
    - OIL may be null

PIM-SM State Maintenance

PIM-SM state maintenance characteristics:

- Periodic hellos are sent to all PIM neighbors.
- Periodic joins refresh interfaces in a PIM neighbor OIL.
- Received multicast packets reset (S,G) entry expiration timers.

Multicast Routing Table

```
RP/0/RSP0/CPU0:PE1# show mrib ipv4 route
IP Multicast Routing Information Base
Entry flags: L - Domain-Local Source, E - External Source to the Domain,
    C - Directly-Connected Check, S - Signal, IA - Inherit Accept,
    IF - Inherit From, D - Drop, MA - MDT Address, ME - MDT Encap,
    MD - MDT Decap, MT - MDT Threshold Crossed, MH - MDT interface handle
    CD - Conditional Decap, MPLS - MPLS Decap, MF - MPLS Encap, EX - Extranet
    MoFE - MoFRR Enabled, MoFS - MoFRR State
Interface flags: F - Forward, A - Accept, IC - Internal Copy,
<... output omitted ...>
(*,234.1.1.1) RPF nbr: 0.0.0.0 Flags: C
  Up: 1d22h
  Outgoing Interface List
    GigabitEthernet0/0/0/0 Flags: F NS LI, Up: 1d22h
```

```
CE1# show ip mroute
<... output omitted ...>
(*, 224.1.1.1), 20:41:16/00:02:25, RP 1.1.1.1, flags: SJCL
  Incoming interface: GigabitEthernet0/0, RPF nbr 192.168.101.10
  Outgoing interface list:
    Loopback0, Forward/Sparse, 20:41:10/00:02:25
(*, 224.0.1.40), 20:41:11/00:02:19, RP 1.1.1.1, flags: SJCL
  Incoming interface: GigabitEthernet0/0, RPF nbr 192.168.101.10
  Outgoing interface list:
    Loopback0, Forward/Sparse, 20:41:10/00:02:19
```

PIM-SM (S,G) State Rules

The PIM-SM (S,G) state rules characteristics:

- (S,G) creation:
  - By receipt of an (S,G) join or prune.
  - By receipt of traffic from a directly connected source.
  - By register process.
  - When a parent (*,G) is created (if it does not already exist).
- (S,G) reflects forwarding of S to G:
  - IIF is the RPF interface toward the source:
    - RPF toward RP if RP bit is set.
  - OIL initially contains a copy of (*,G) OIL minus IIF.
- (S,G) deletion:
  - By normal (S,G) entry timeout.

PIM-SM OIL Rules

PIM-SM OIL rules characteristics:

- Interfaces in OIL added:
  - By receipt of join message:
    - Interfaces added to (*,G) are added to all (S,G) OILs.
- Interfaces in OIL removed:
  - By receipt of prune message:
    - Interfaces removed from (*,G) are removed from all (S,G) OILs.
- Interface expire timer counts down to 0:
  - Timer reset (to 3 minutes) by receipt of periodic join.

    or
  - By IGMP membership report.

PIM-SM State Flags

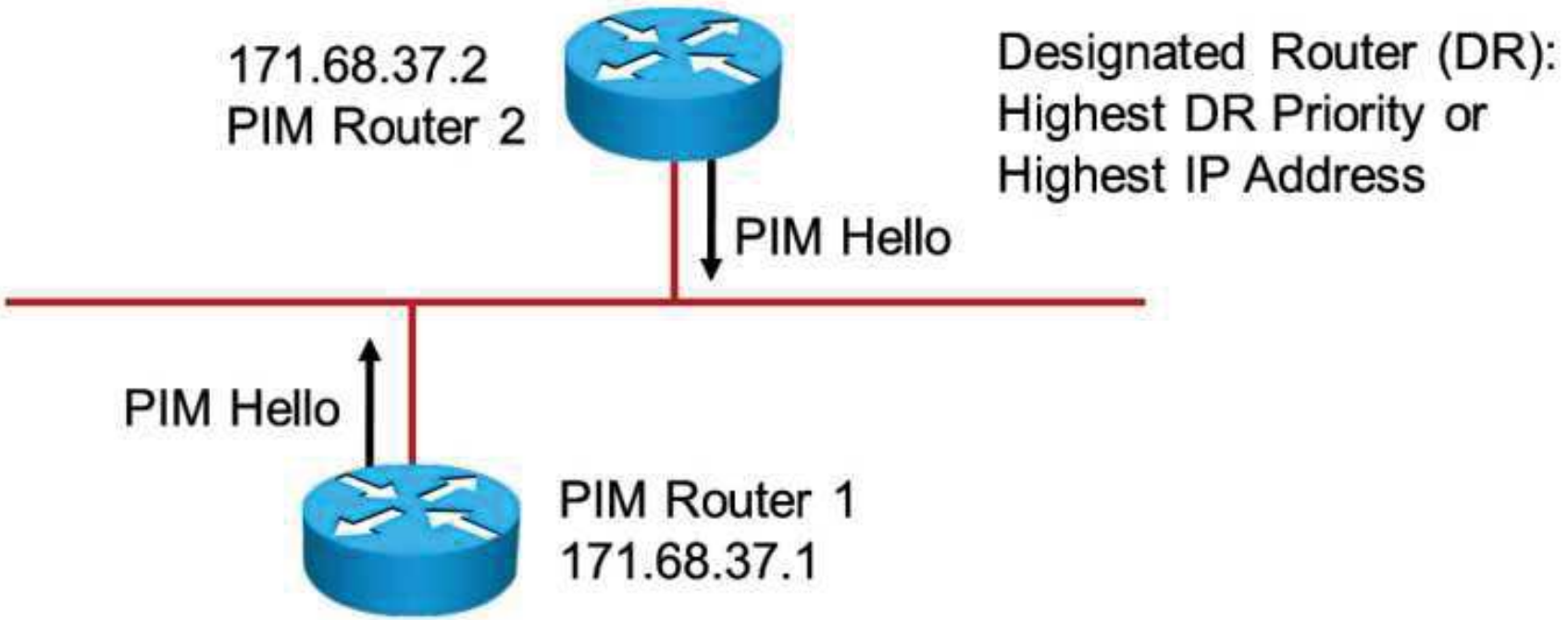
PIM-SM state flags:

- S = sparse mode
- C = directly connected host
- L = local (router is a member)
- P = pruned (all interfaces in OIL are pruned)
- T = forwarding via SPT
- J = join SPT
- F = register
- R = RP bit

PIM-SM Neighbor Discovery

PIM-SM neighbor discovery characteristics:

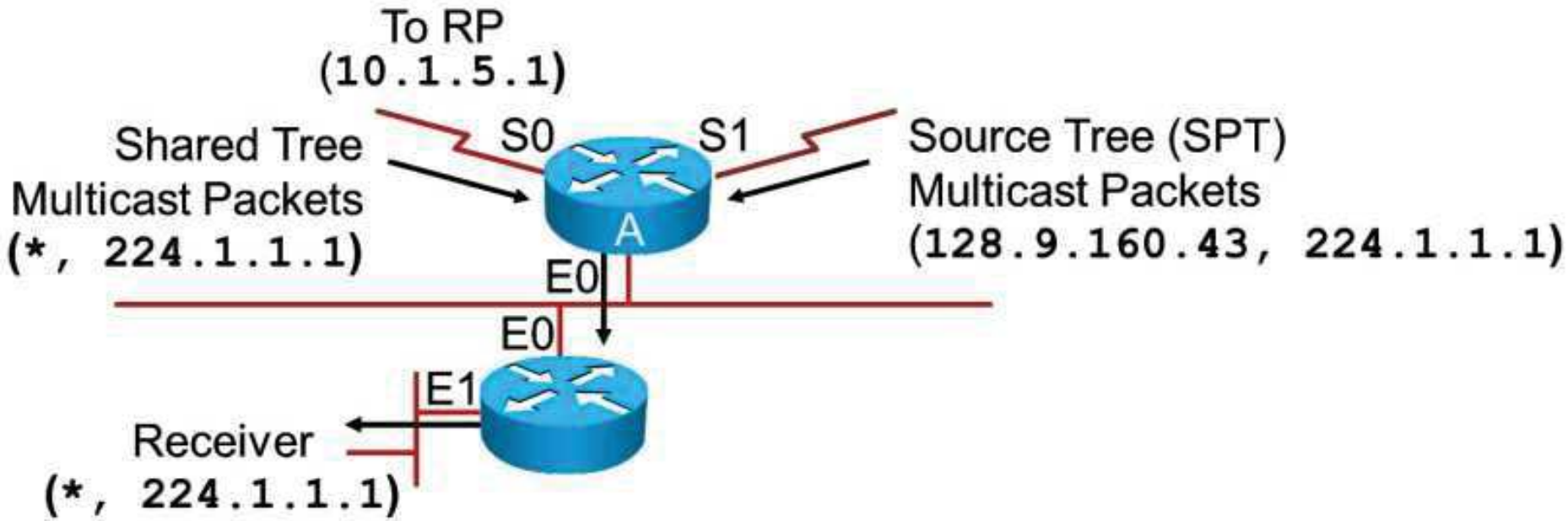
- PIMv2 hellos are multicast periodically (default is every 30 seconds) to the all-PIM-routers (224.0.0.13) group address.
- If the DR times out, a new DR is elected.
- The DR is responsible for sending all join and register messages for any receivers or senders on the network.



PIM-SM Forwarding

PIM-SM forwarding characteristics:

- Packets are forwarded out all interfaces in the OIL.
- Interfaces are placed on the OIL for a multicast group if:
  - A PIM neighbor joins the group on this interface.
  - A host on this interface has joined the group.
  - An interface has been manually configured to join the group.



PIM-SM Joining

PIM-SM joining characteristics:

- Leaf routers send a (*,G) join toward RP:
  - Joins sent hop by hop along path toward RP.
- Each router along path creates (*,G) state:
  - If no (*,G) state:
    - Create it and send a join toward RP.
  - Otherwise:
    - Join process is complete; reached the shared tree.

PIM-SM Registering

PIM-SM registering characteristics:

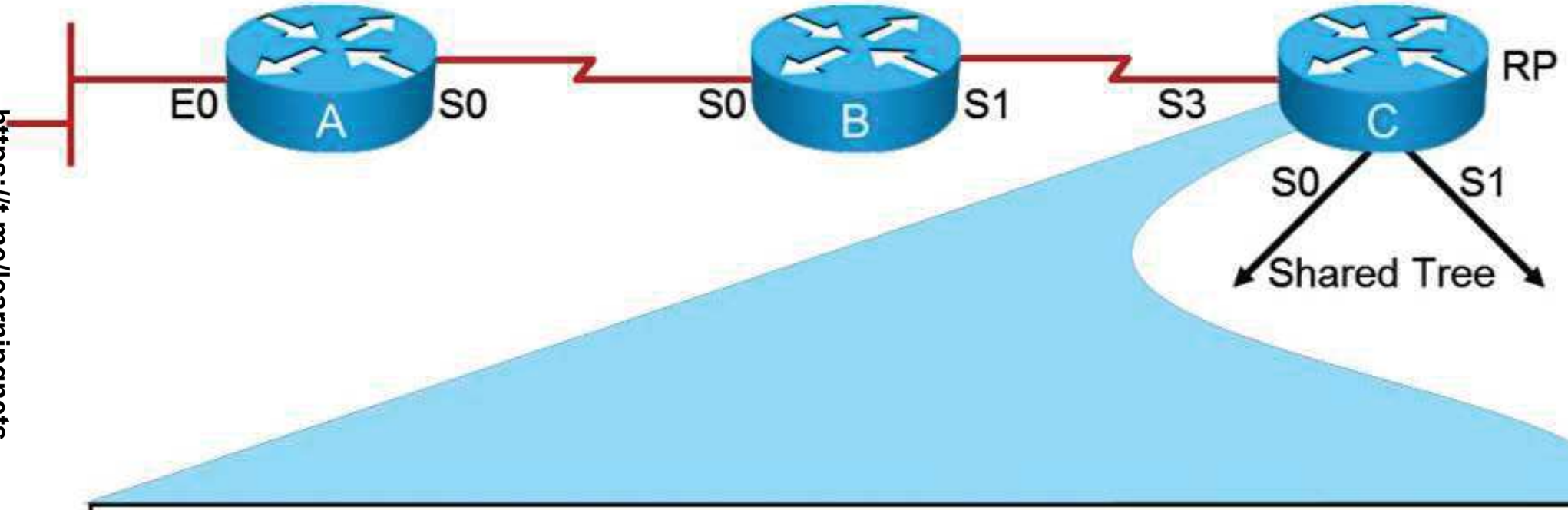
1. Senders begin sourcing multicast traffic.
2. First-hop router unicasts registers to RP:
   - A multicast packet is encapsulated in each register message.
   - Register messages follow unicast path to RP.
3. RP receives register messages:
   - De-encapsulates multicast packet inside register message.
   - Forwards multicast packet down shared tree.
   - Sends (S,G) join toward source and first-hop router to build an (S,G) SPT between source and RP.

PIM-SM Registering (Cont.)

4. First-hop router receives (S,G) join:
   - SPT between source and RP now built.
   - Begins forwarding traffic down (S,G) SPT to RP.
   - (S,G) traffic temporarily flowing down two paths to RP.
5. RP receives traffic down native (S,G) SPT:
   - Sends a register-stop message to source and first-hop router.
6. First-hop router receives register-stop:
   - Stops encapsulating traffic in register messages.
   - (S,G) traffic now flowing down single SPT to RP.

PIM-SM Registering: Receiver Joins First Scenario

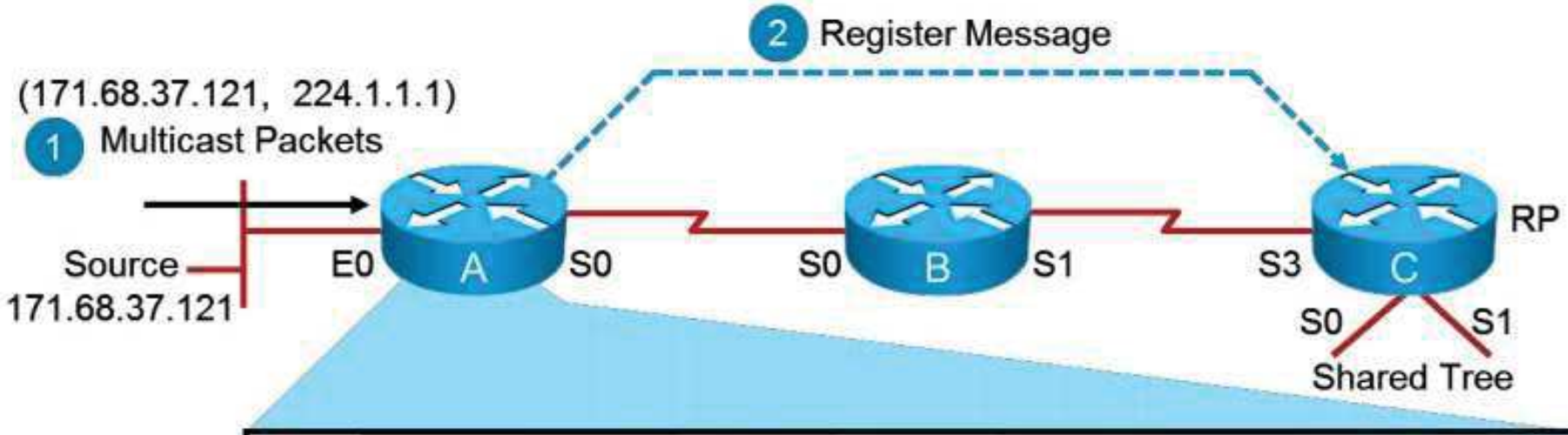
State in RP before any source registers (receivers on shared tree); empty states for group 224.1.1.1 in routers A and B.




```
(*, 224.1.1.1), 00:00:03/00:02:56, RP 171.68.28.140, flags:S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:03:14/00:02:59
    Serial1, Forward/Sparse, 00:03:14/00:02:59
```

PIM-SM Registering: Receiver Joins First Process

1. Source begins sending group G traffic
2. Router A encapsulates packets in registers, unicasts to RP

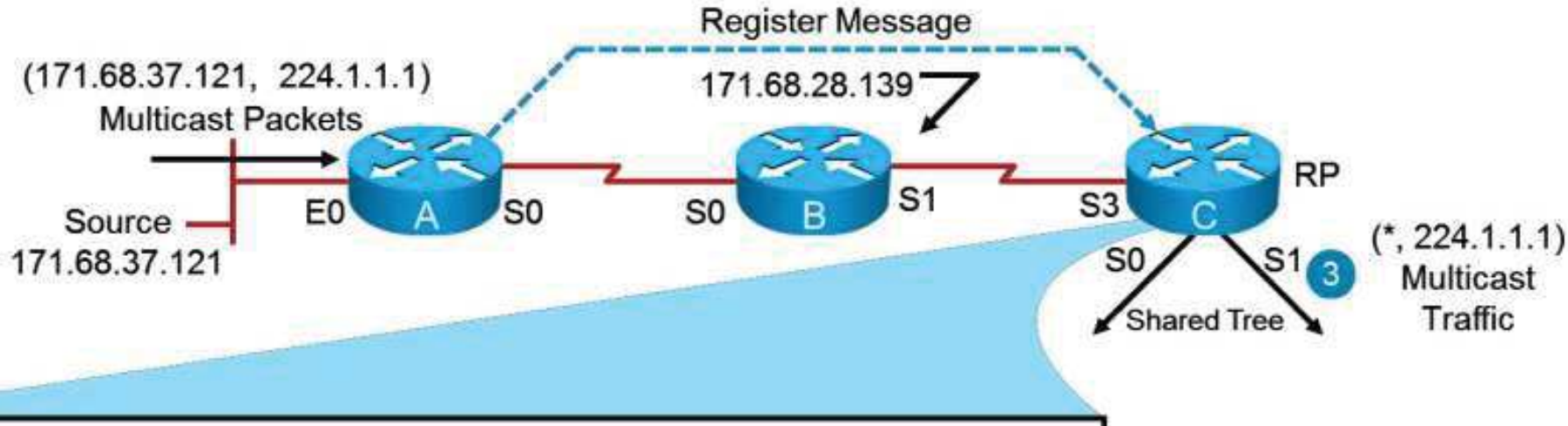


Router A creates (S,G) state for the source (after automatically creating [*,G] entry)

```
(*, 224.1.1.1), 00:00:03/00:02:56, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:00:03/00:02:56, flags: FPT
  Incoming interface: Ethernet0, RPF nbr 0.0.0.0, Registering
  Outgoing interface list: Null
```

PIM-SM Registering: Receiver Joins First Process (Cont.)

3. Router C (RP) de-encapsulates packets, forwards down shared tree

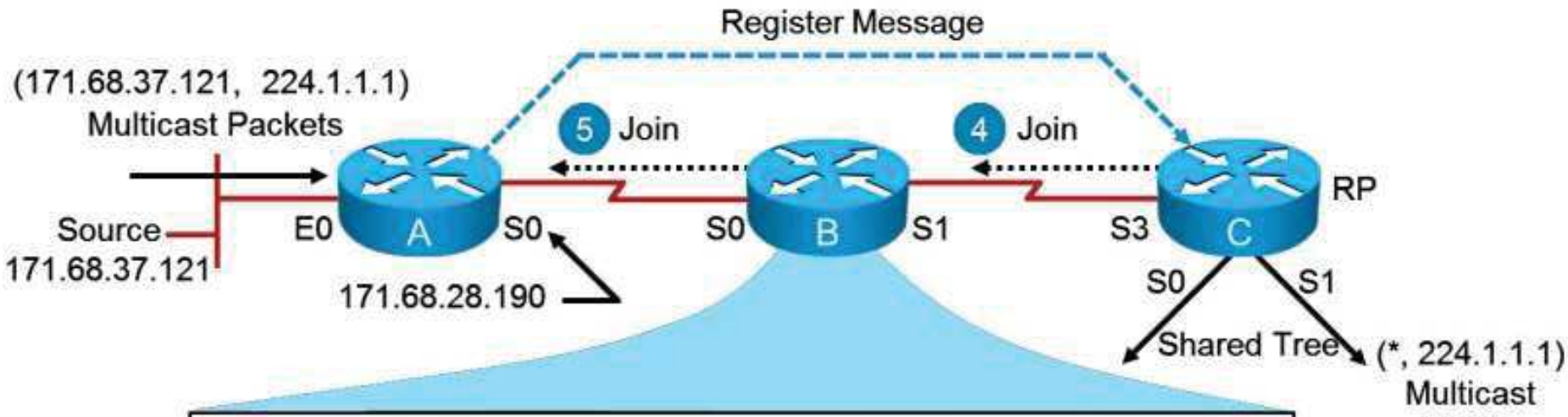



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:09:21/00:02:38
    Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121, 224.1.1.1, 00:01:15/00:02:46, flags:
  Incoming interface: Serial3, RPF nbr 171.68.28.139,
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:00:49/00:02:11
    Serial1, Forward/Sparse, 00:00:49/00:02:11
```

RP processes register; creates (S,G) state

PIM-SM Registering: Receiver Joins First Process (Cont.)

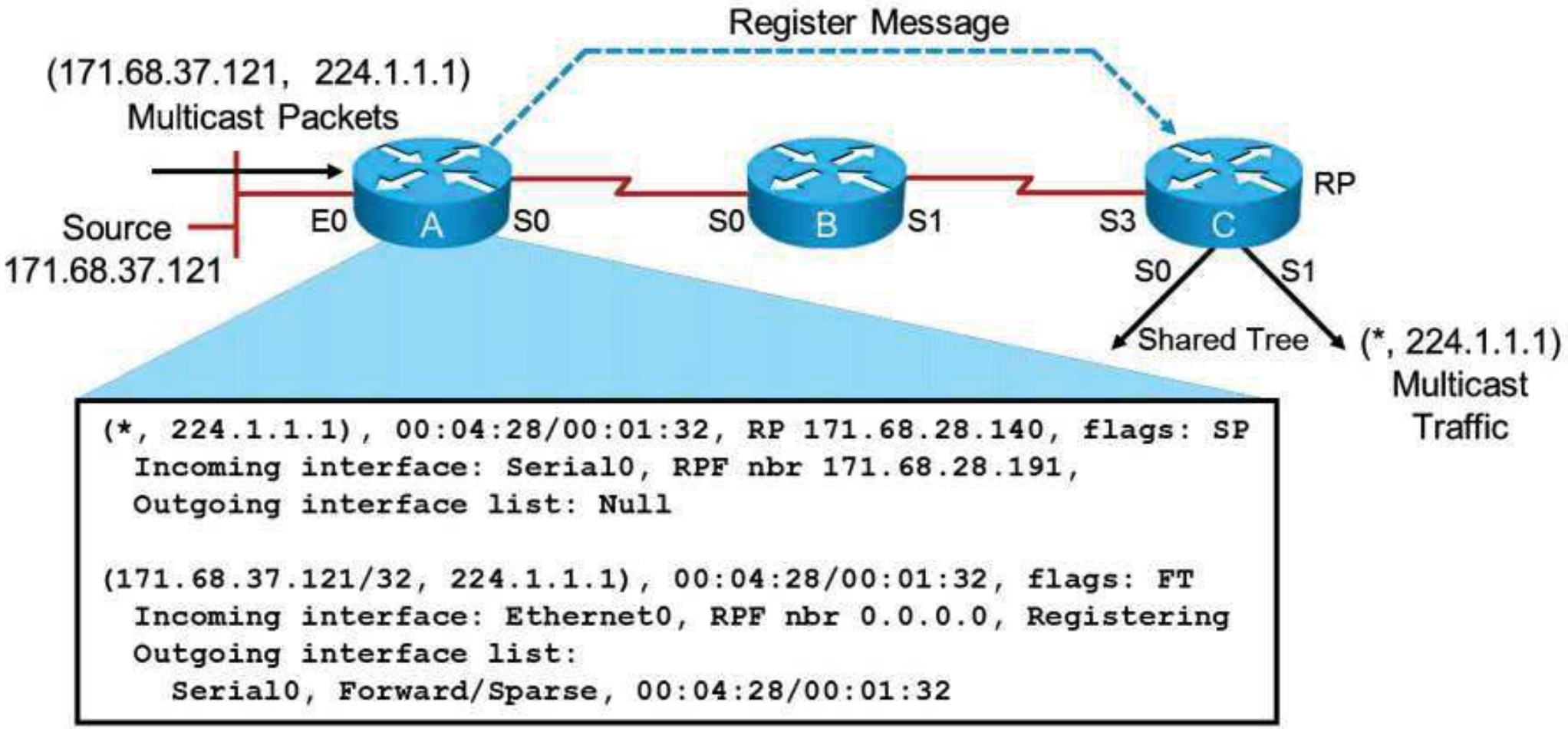
4. RP sends (S,G) join toward source to build SPT
5. Router B sends (S,G) join toward source to continue building SPT



Router B processes join, creates (S,G) state (after automatically creating the [*,G] entry)

```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags:
  Incoming interface: Serial0, RPF nbr 171.68.28.190
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:04:28/00:01:32
```

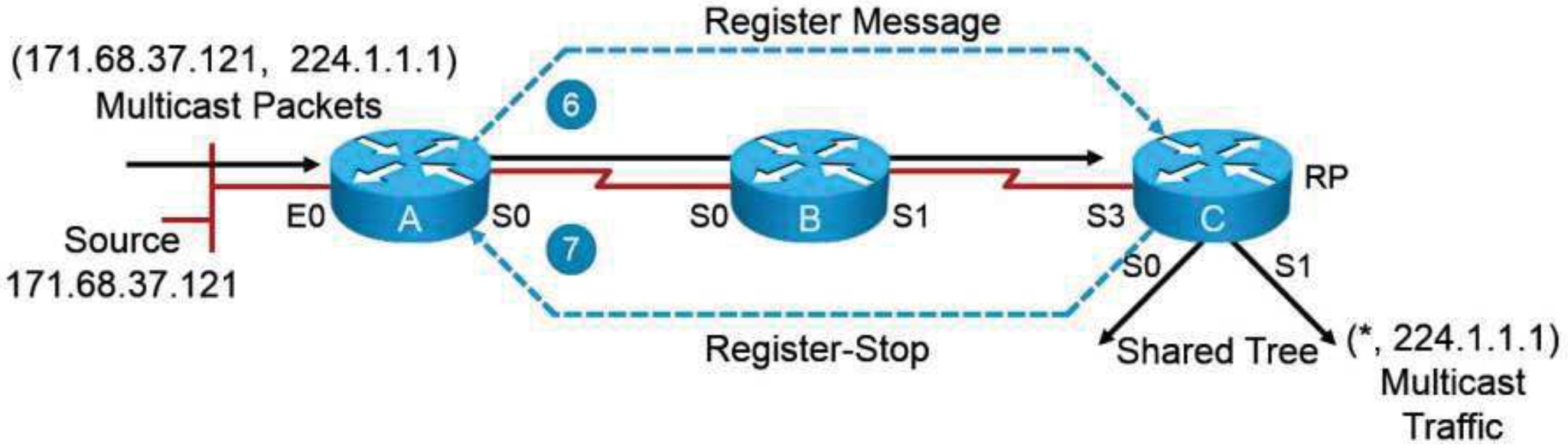
PIM-SM Registering: Receiver Joins First Process (Cont.)



Router A processes the (S,G) join; adds Serial0 to OIL

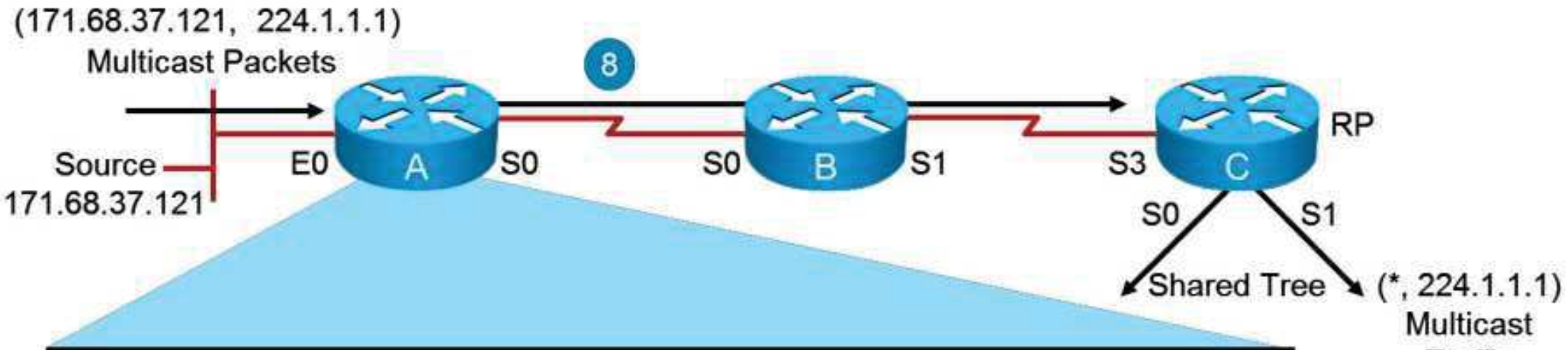
PIM-SM Registering: Receiver Joins First Process (Cont.)

6. RP begins receiving (S,G) traffic down SPT
7. RP sends register-stop to router A



PIM-SM Registering: Receiver Joins First Process (Cont.)

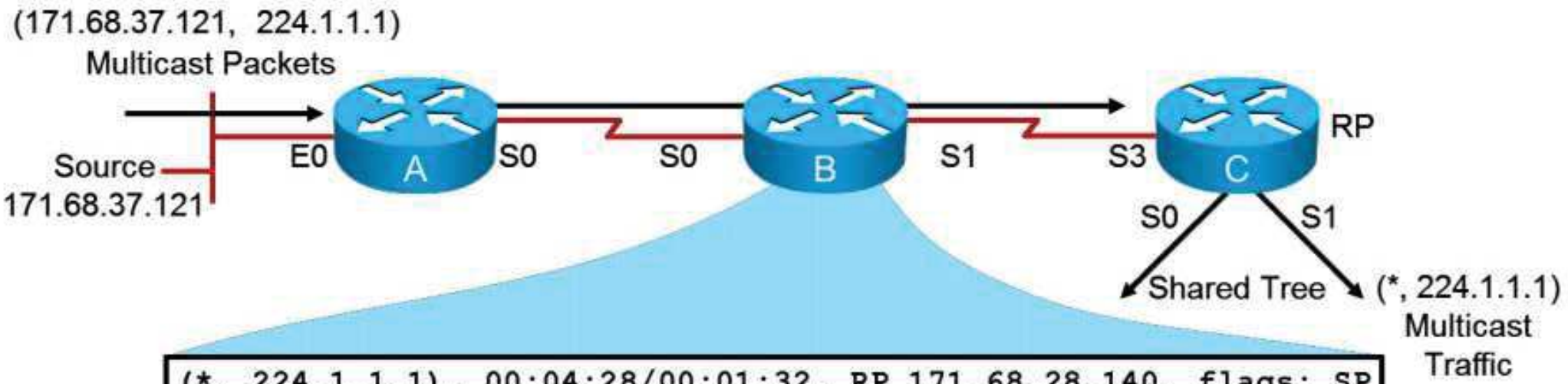
8. (S,G) traffic now flowing down a single path (SPT) to RP




```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT
  Incoming interface: Ethernet0, RPF nbr 0.0.0.0,
  Outgoing interface list:
     Serial0, Forward/Sparse, 00:04:28/00:01:32
```

Router A stops sending register messages (final state in router A)

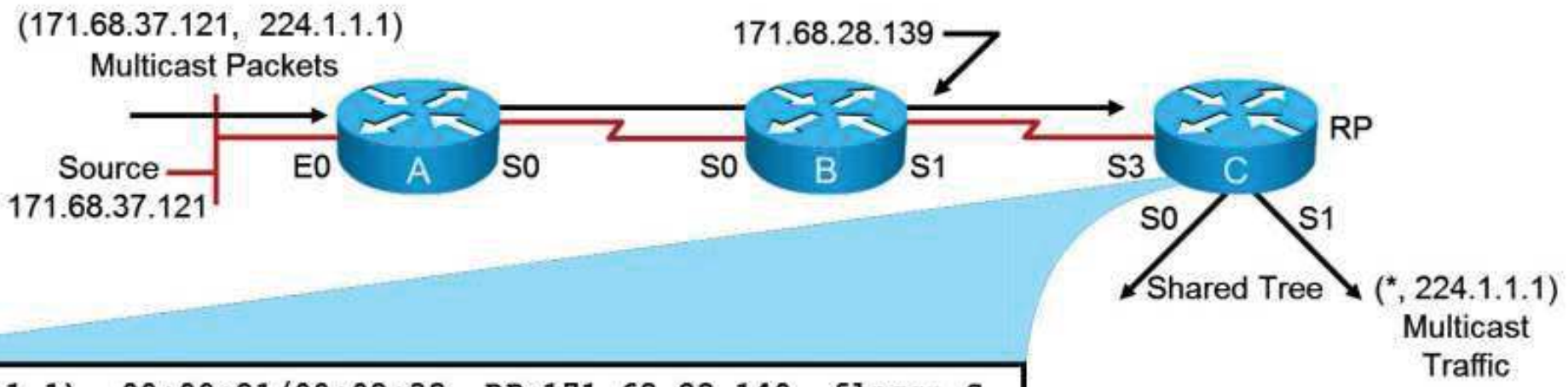
PIM-SM Registering: Receiver Joins First Process (Cont.)




```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
  Incoming interface: Serial0, RPF nbr 171.68.28.190
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:04:28/00:01:32
```

Final state in router B

PIM-SM Registering: Receiver Joins First Process (Cont.)

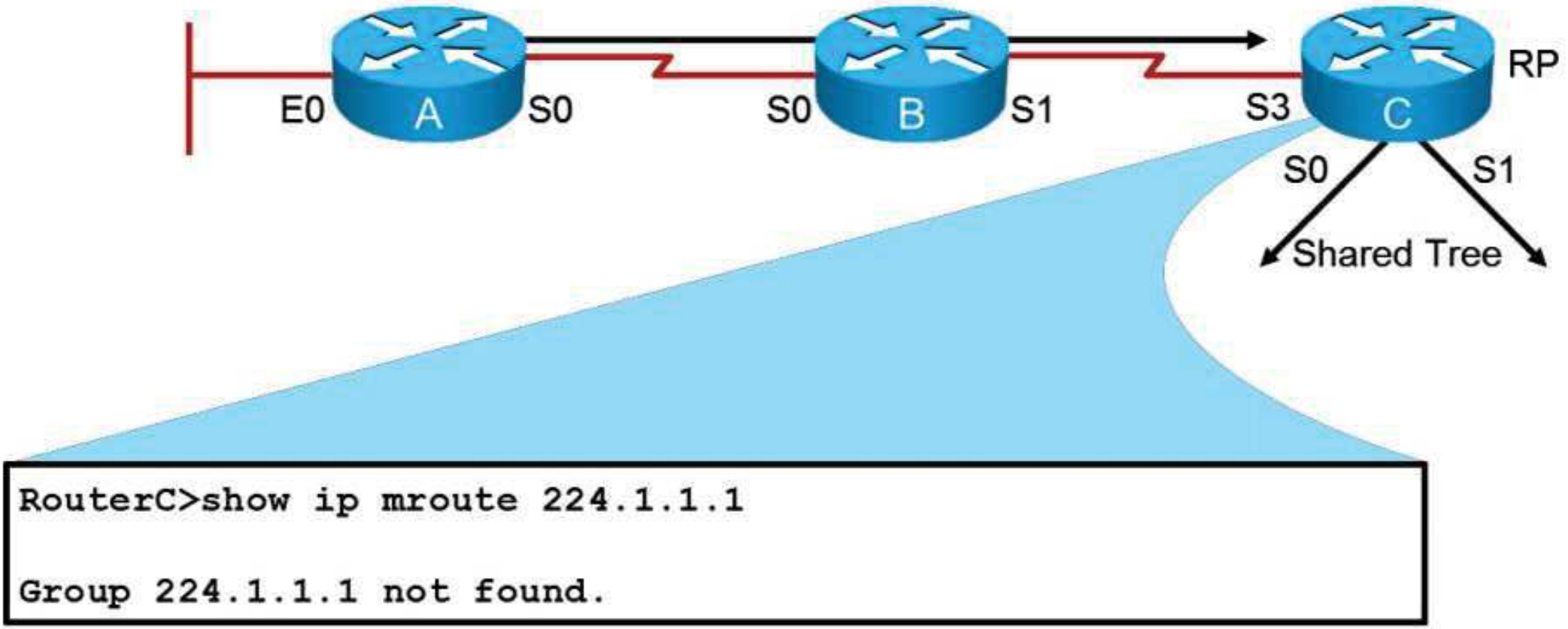



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:09:21/00:02:38
    Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121, 224.1.1.1, 00:01:15/00:02:46, flags: T
  Incoming interface: Serial3, RPF nbr 171.68.28.139,
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:00:49/00:02:11
    Serial1, Forward/Sparse, 00:00:49/00:02:11
```

Final state in the RP (with receivers on shared tree)

PIM-SM Registering: Source Starts First Scenario

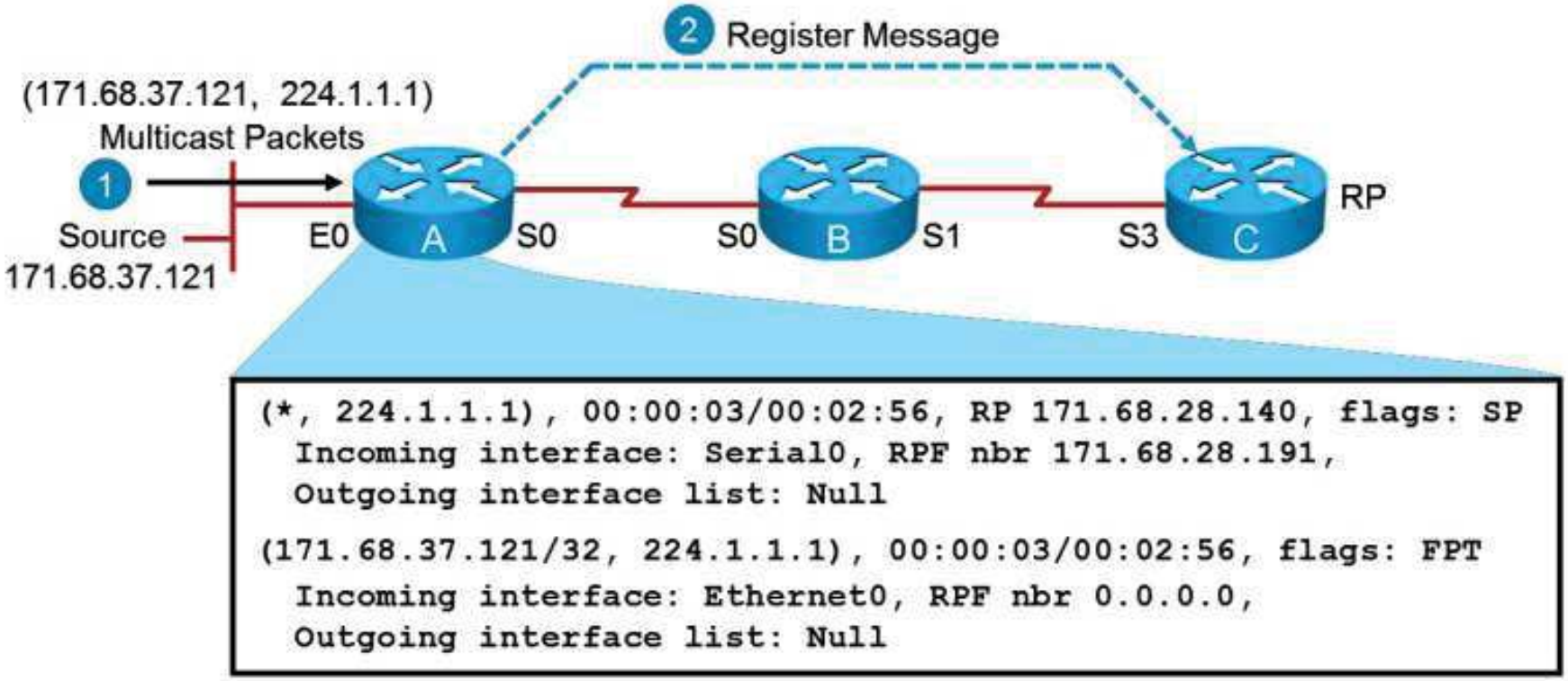
Empty state in RP before any source registers (no receivers on shared tree); empty states for group 224.1.1.1 in routers A and B




```
RouterC>show ip mroute 224.1.1.1

Group 224.1.1.1 not found.
```

PIM-SM Registering: Source Starts First Process

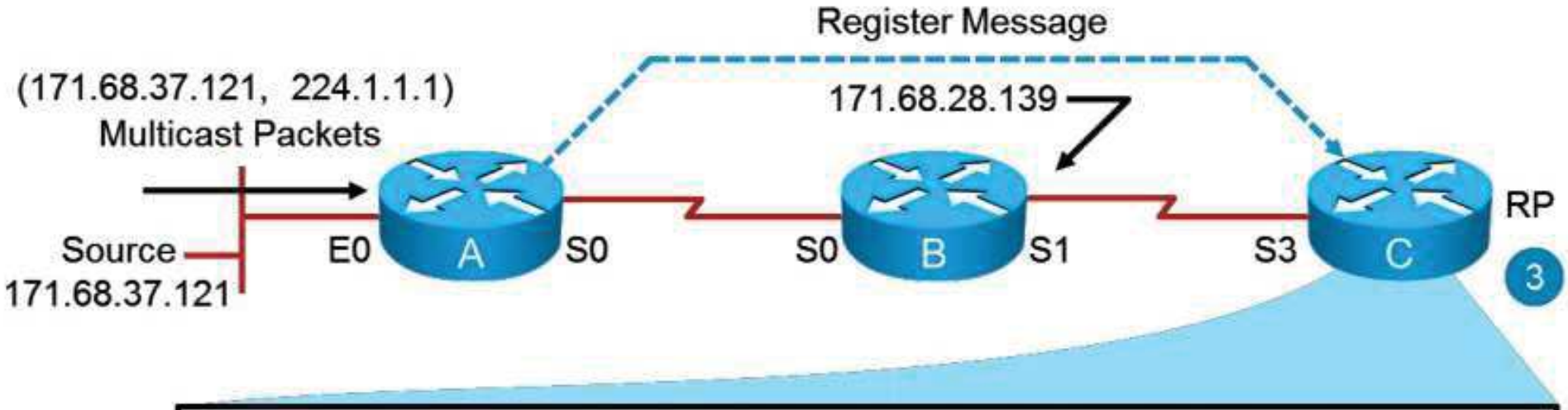
1. Source begins sending group G traffic
2. Router A encapsulates packets in registers, unicasts to RP



Router A creates (S,G) state for source (after automatically creating a [*,G] entry)

PIM-SM Registering Source Starts First (Cont.)

3. Router C (RP) has no receivers on shared tree, discards packet

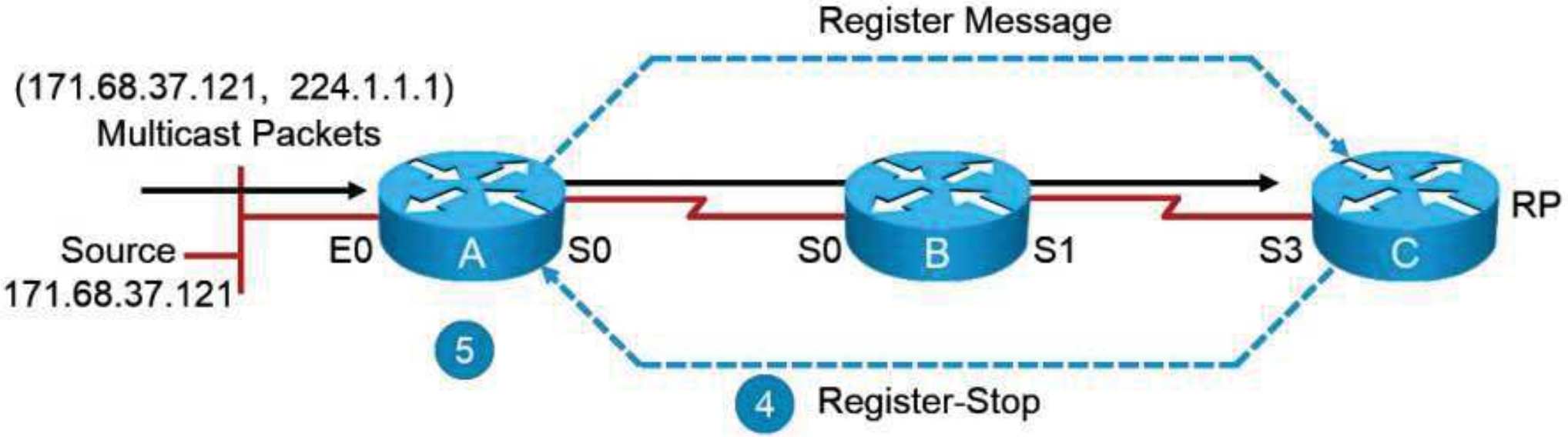



```
(*, 224.1.1.1), 00:01:15/00:01:45, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list: Null

(171.68.37.121, 224.1.1.1), 00:01:15/00:01:45, flags: P
  Incoming interface: Serial3, RPF nbr 171.68.28.139,
  Outgoing interface list: Null
```

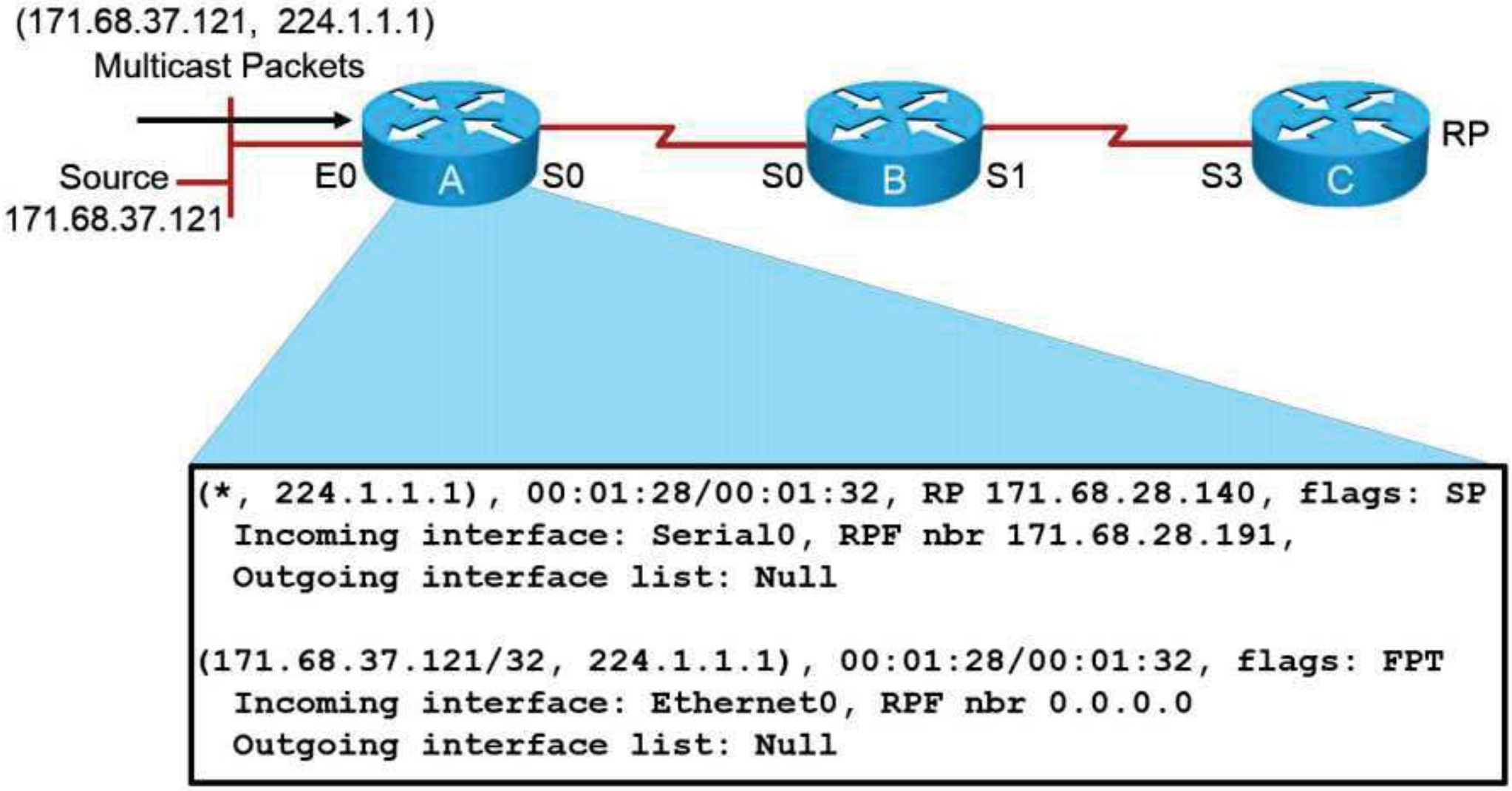
RP processes register; creates (S,G) state (after automatically creating the [*,G] entry)

PIM-SM Registering Source Starts First (Cont.)

4. RP sends register-stop to router A
5. Router A stops encapsulating traffic in register messages, drops packets from source



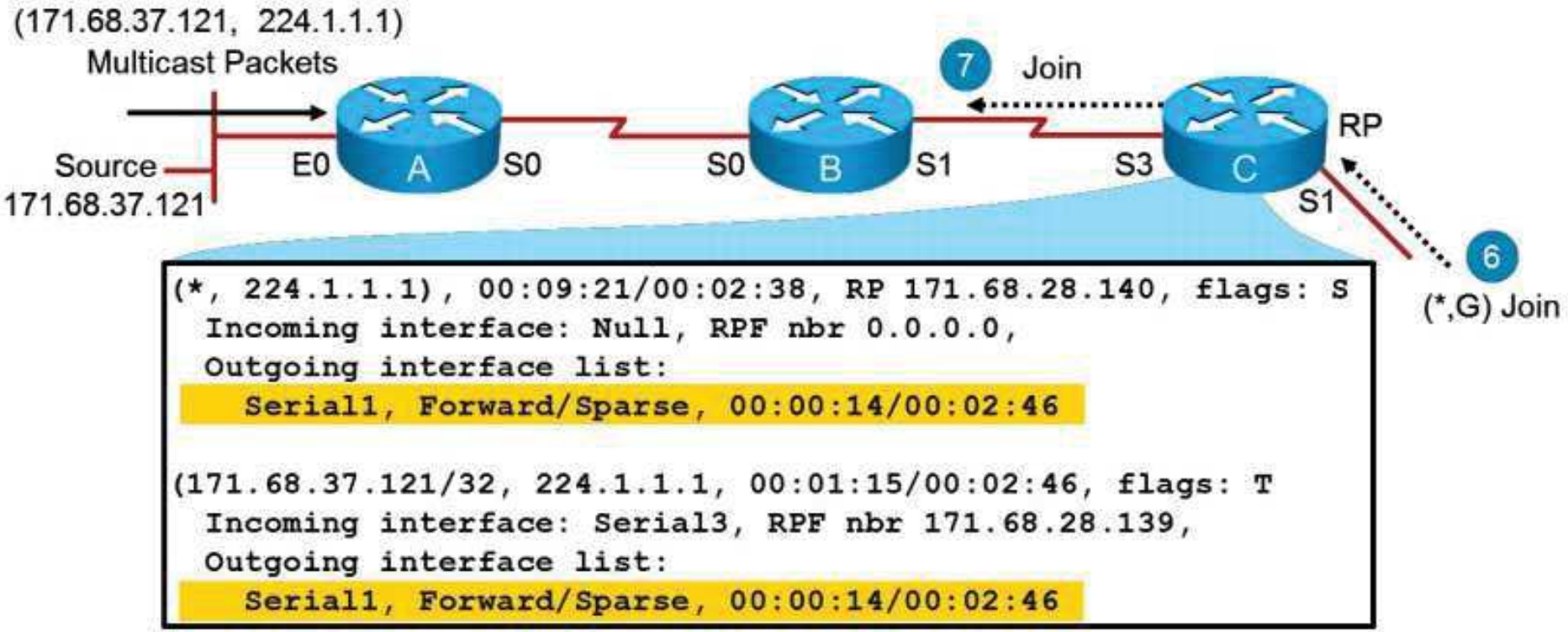
PIM-SM Registering Source Starts First (Cont.)



State in router A after registering (without receivers on shared tree)

PIM-SM Registering Source Starts First (Cont.)

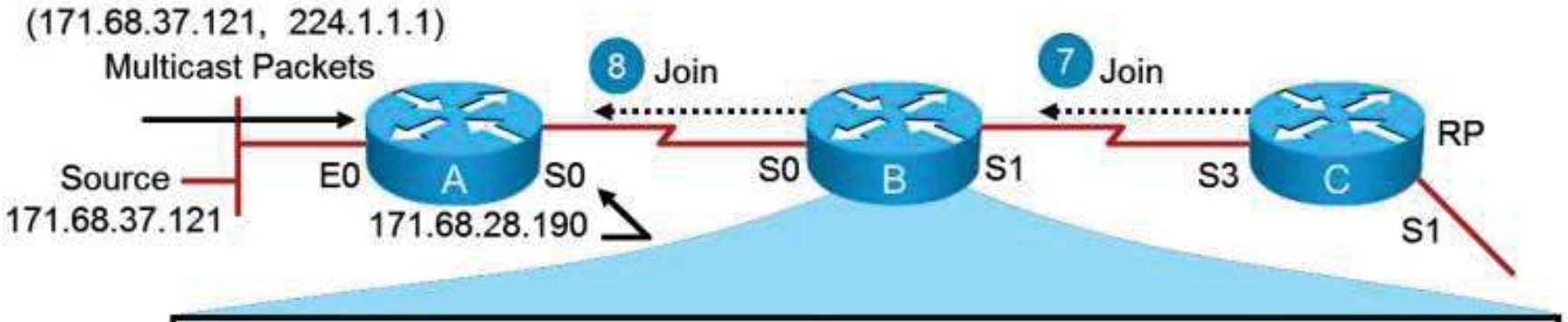
6. RP (router C) receives (*,G) join from a receiver on shared tree.
7. RP sends (S,G) joins for all known sources in group.



RP processes (*,G) join (adds Serial1 to OILs)

PIM-SM Registering Source Starts First (Cont.)

8. Router B sends (S,G) join toward source to continue building SPT.

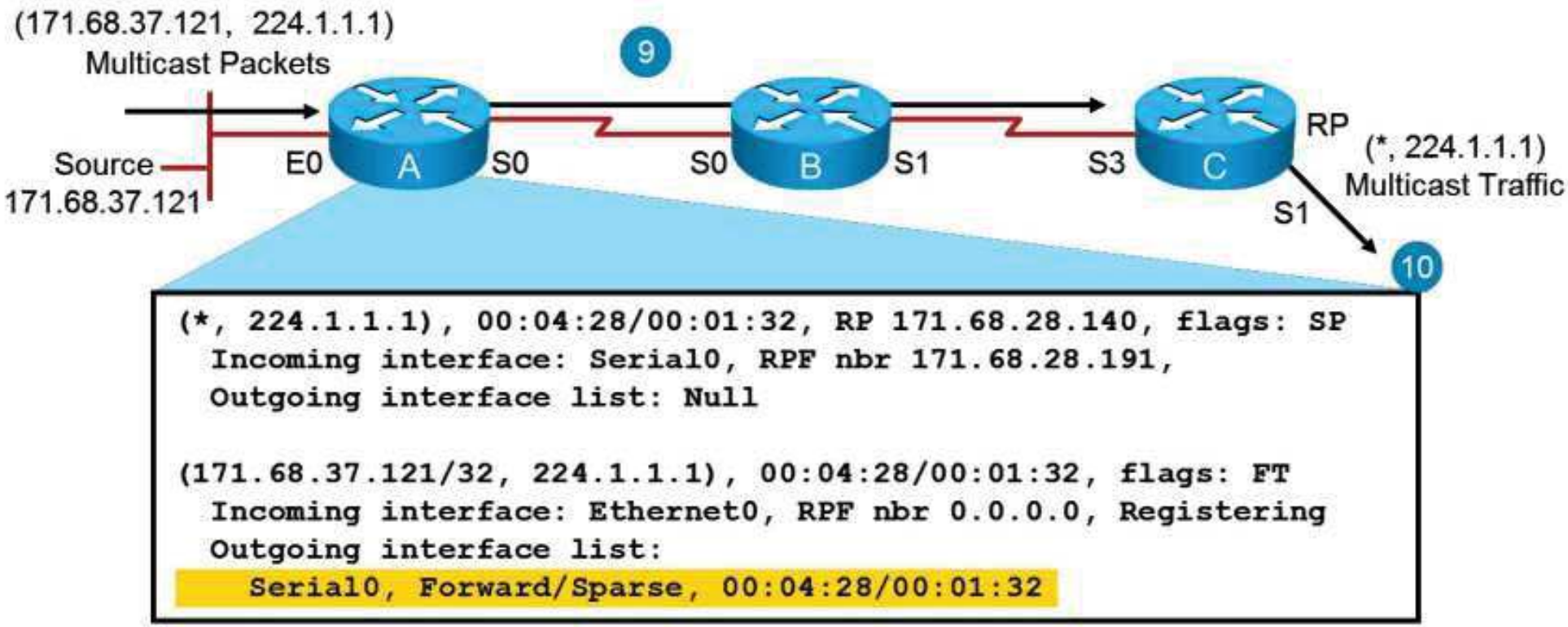



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags:
  Incoming interface: Serial0, RPF nbr 171.68.28.190
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:04:28/00:01:32
```

Router B processes join; creates (S,G) state (after automatically creating the [*,G] entry)

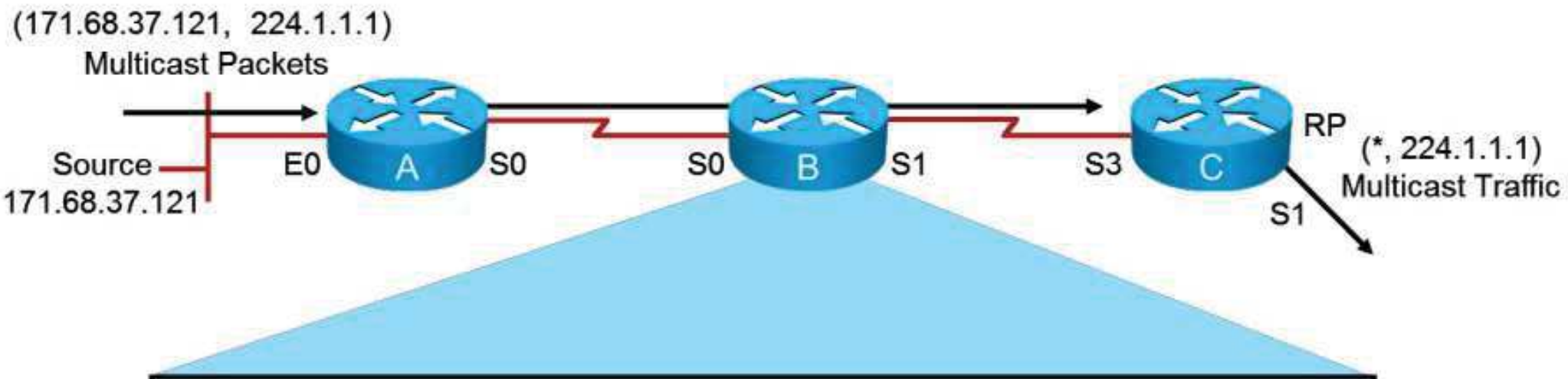
PIM-SM Registering Source Starts First (Cont.)

9. RP begins receiving (S,G) traffic down SPT.
10. RP forwards (S,G) traffic down shared tree to receivers.



Router A processes the (S,G) join; adds Serial0 to OIL

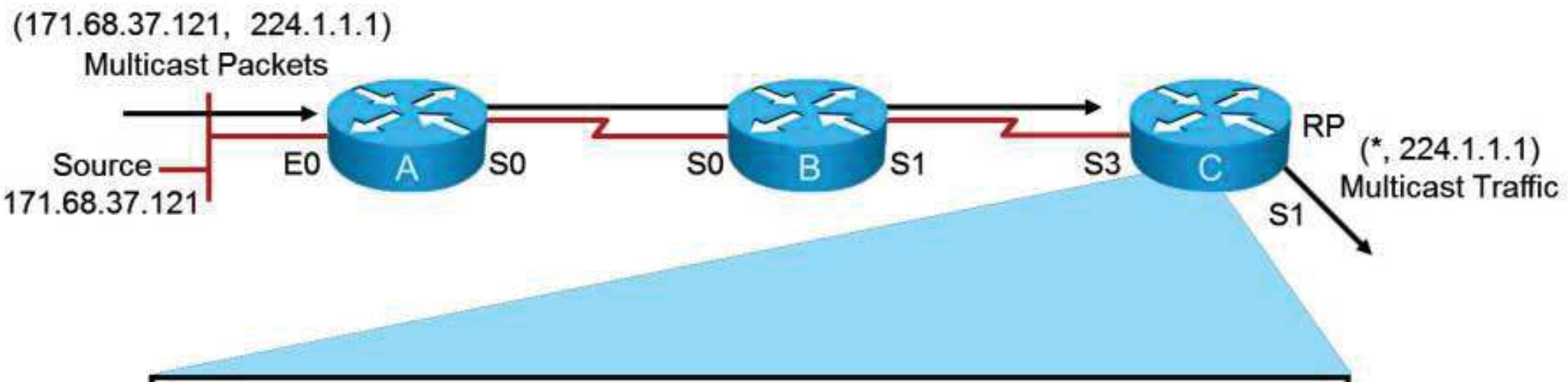
PIM-SM Registering: Receivers Along the SPT Scenario




```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
  Incoming interface: Serial0, RPF nbr 171.68.28.190
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:04:28/00:01:32
```

Current state in router B

PIM-SM Registering: Receivers Along the SPT Scenario (Cont.)

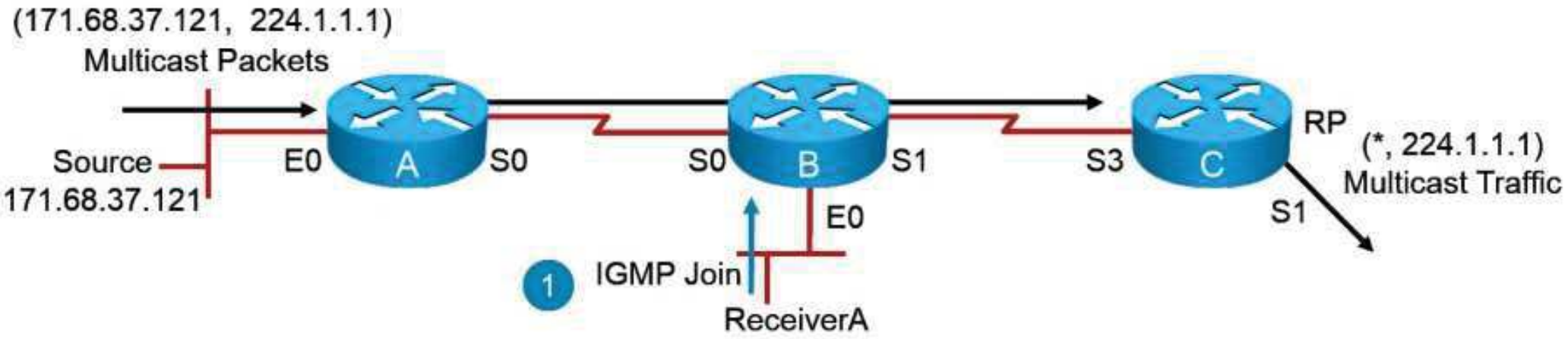



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
  Incoming interface: Serial3, RPF nbr 171.68.28.139,
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:00:49/00:02:11
```

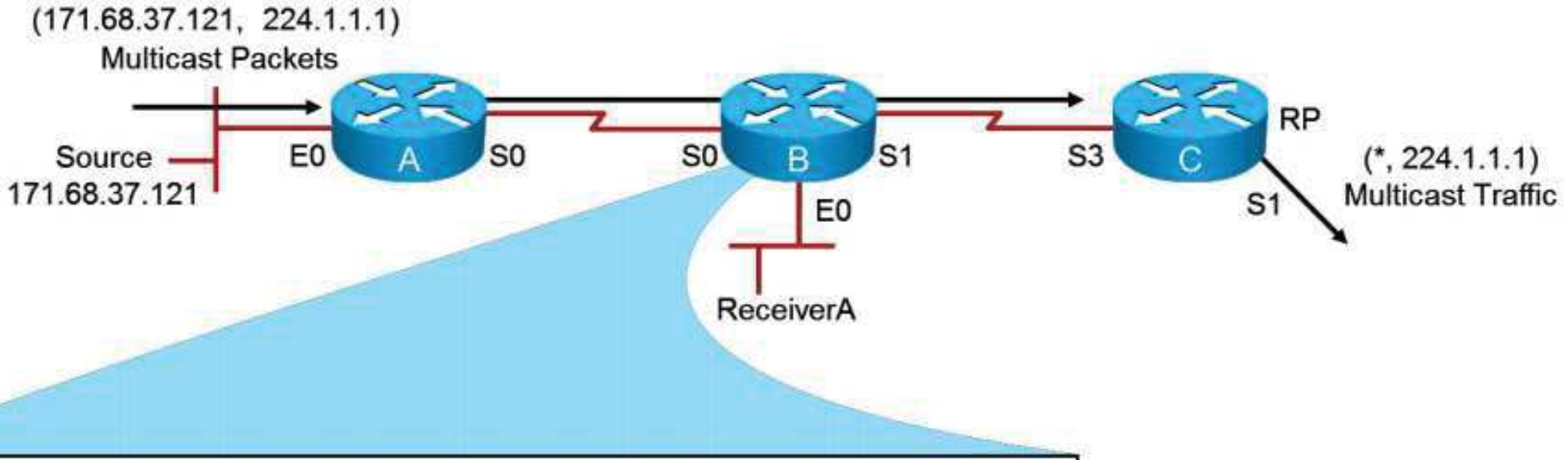
Current state in the RP

PIM-SM Registering: Receivers Along the SPT Process

1. ReceiverA wants to receive group G traffic, sends IGMP join for G



PIM-SM Registering: Receivers Along the SPT Process (Cont.)



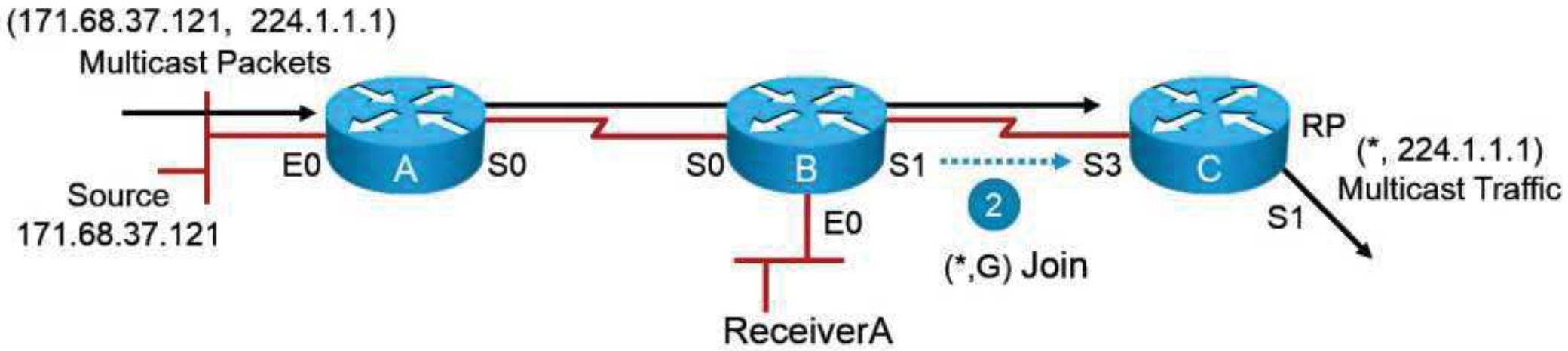

```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SC
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:00:30/00:02:30

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: CT
  Incoming interface: Serial0, RPF nbr 171.68.28.190
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:04:28/00:01:32
    Ethernet0, Forward/Sparse, 00:00:30/00:02:30
```

Added Interfaces

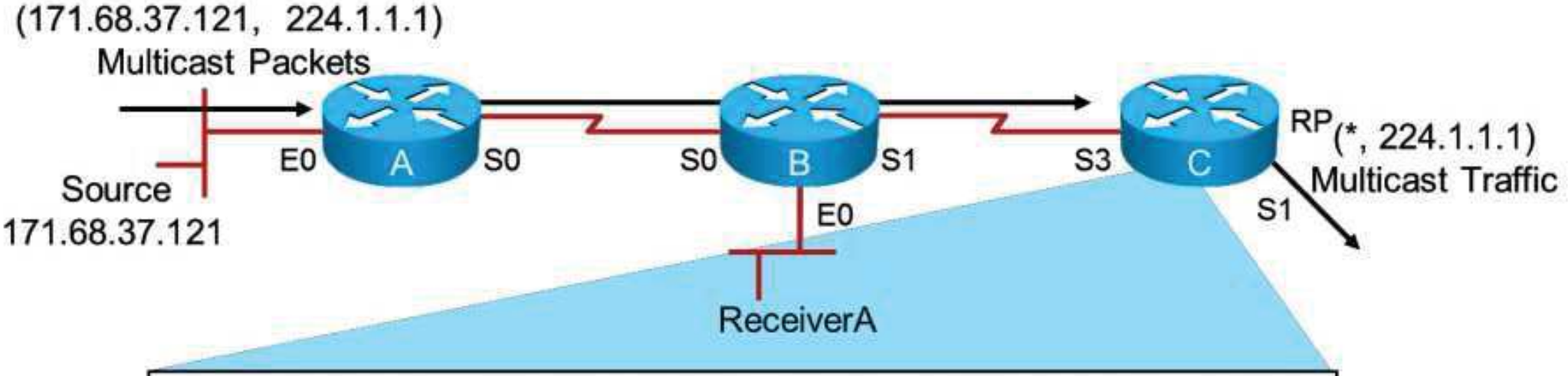
State in router B after ReceiverA joins group

PIM-SM Registering: Receivers Along the SPT (Cont.)

2. Router B triggers a (*,G) join to join the shared tree



PIM-SM Registering: Receivers Along the SPT Process (Cont.)

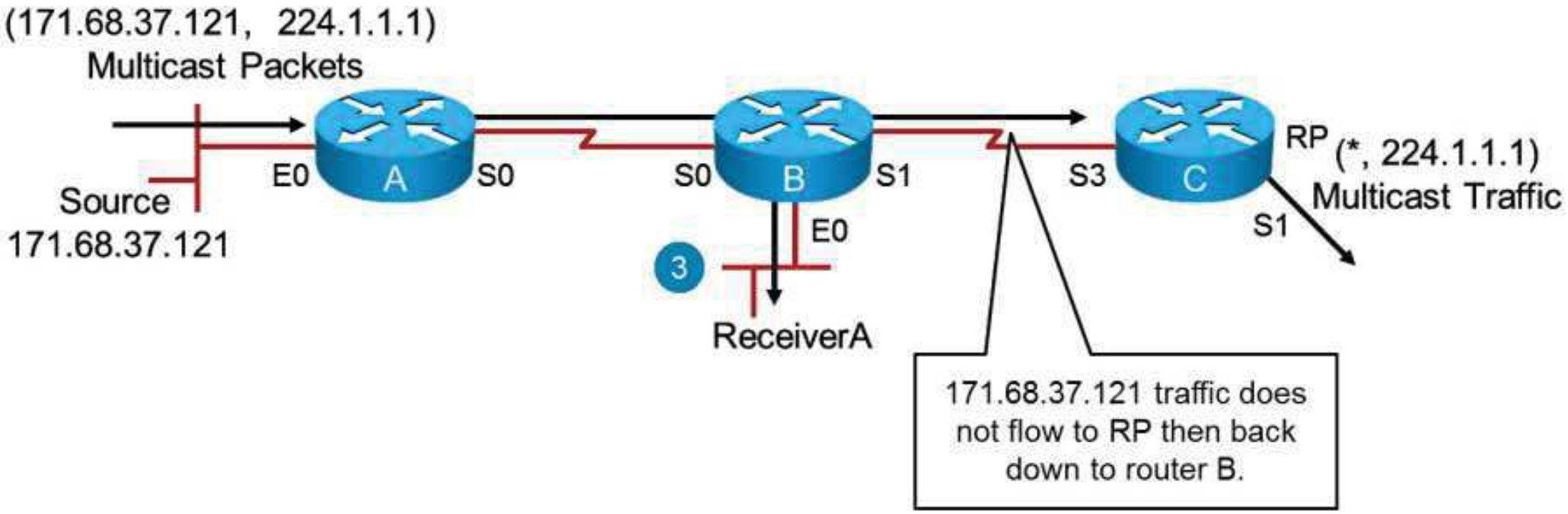



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:03:14/00:02:46
    Serial3, Forward/Sparse, 00:00:10/00:02:50

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
  Incoming interface: Serial3, RPF nbr 171.68.28.139,
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:00:49/00:02:11
```

State in RP after router B joins shared tree

PIM-SM Registering: Receivers Along the SPT Process (Cont.)

3. Group G traffic begins to flow to ReceiverA



PIM-SM SPT Switchover Overview

PIM-SM SPT switchover characteristics:

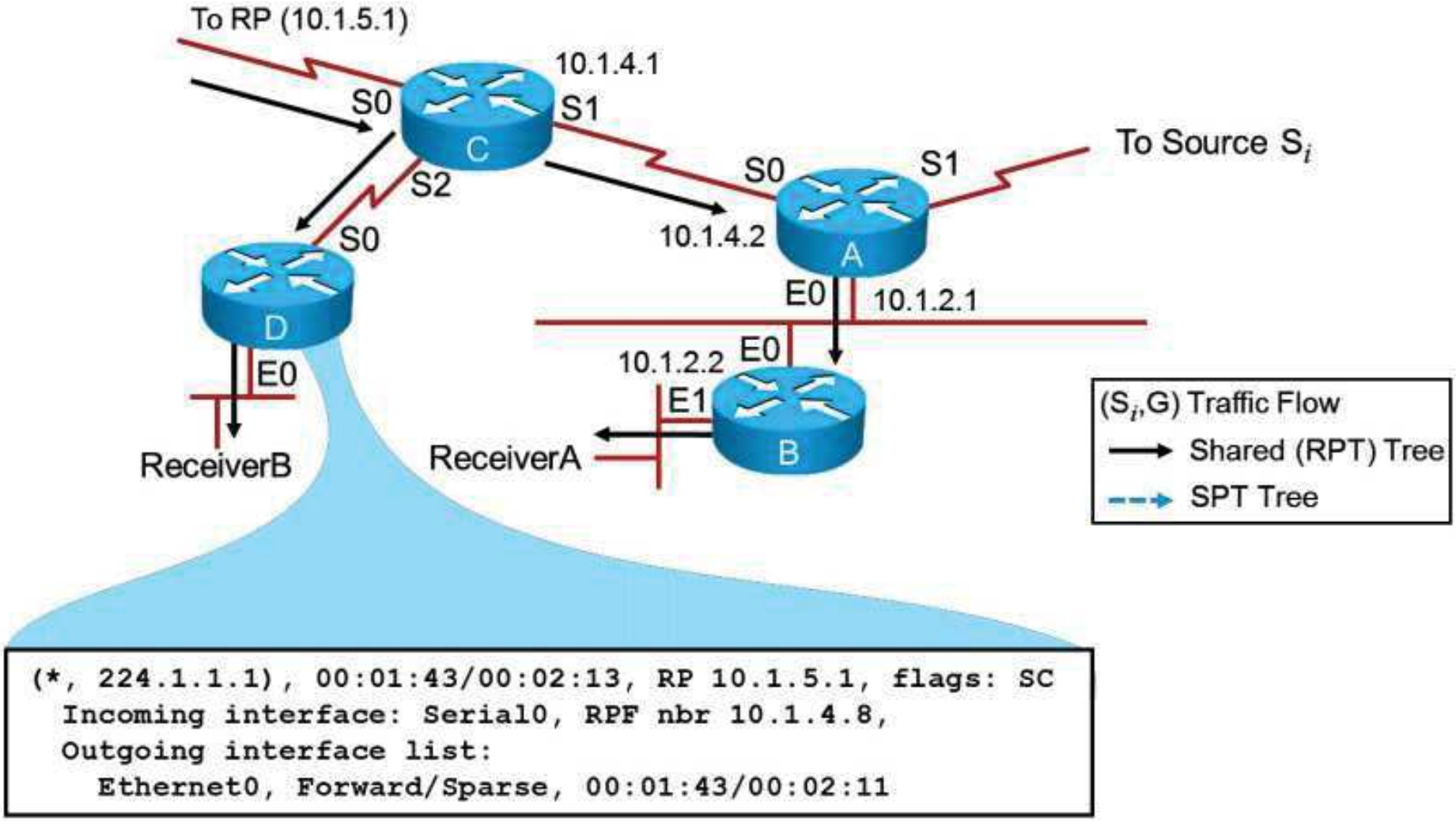
- SPT thresholds may be set for any group:
  - Access lists may be used to specify which groups.
  - Default threshold = 0 kb/s (i.e., immediately join SPT).
  - Threshold of infinity means never join SPT.
- Threshold triggers join of source tree:
  - Pro: Reduces network latency.
  - Con: More (S,G) state must be stored in the routers.

PIM-SM SPT Switchover Overview (Cont.)

SPT switchover mechanism:

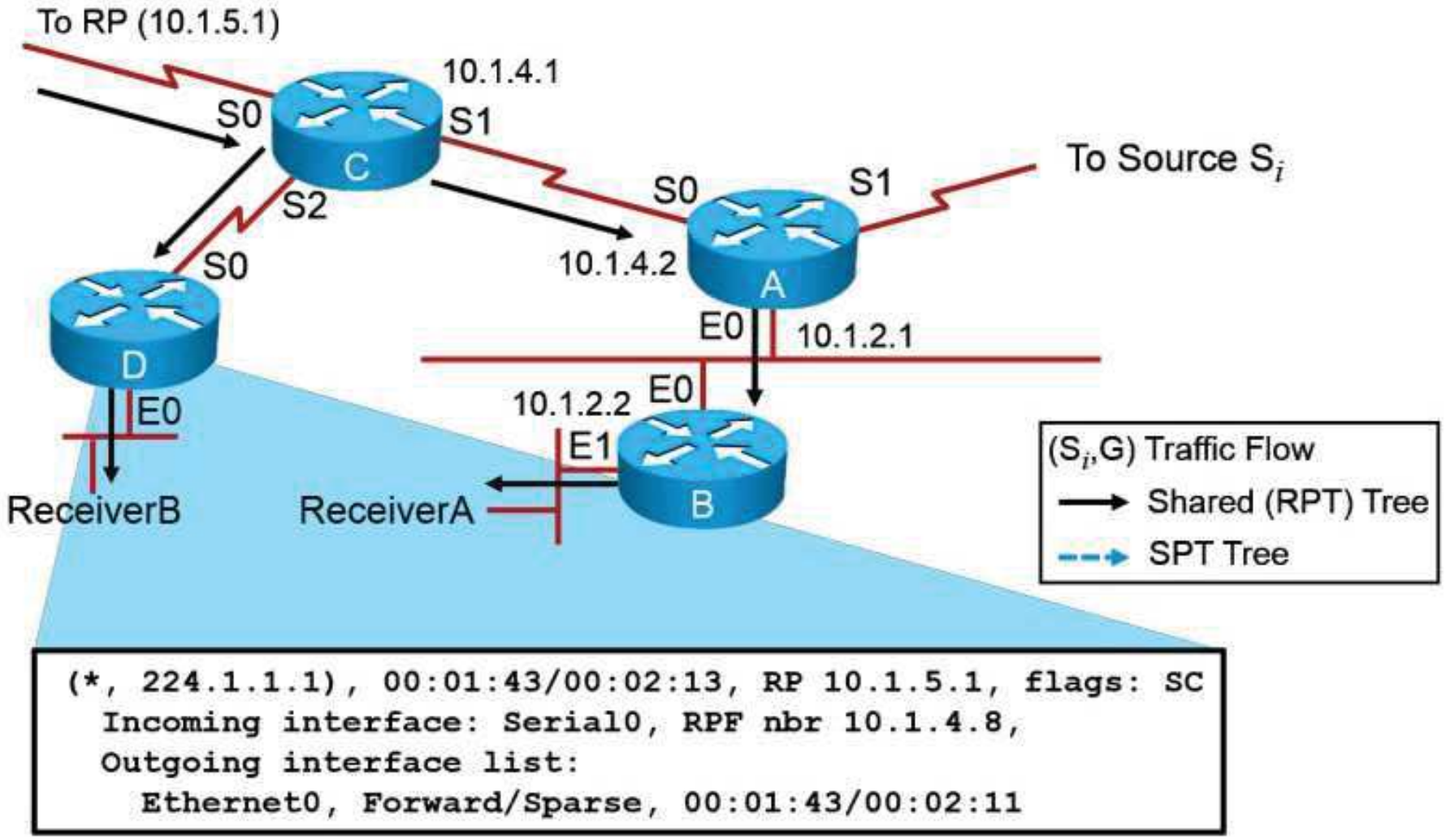
- Once each second:
  - Compute new (*,G) traffic rate
  - If threshold exceeded, set J flag in (*,G)
- For each (S_i,G) packet received:
  - If J flag set in (*,G):
    - Join SPT for (S*i*,G)
    - Mark (S_i,G) entry with J flag
    - Clear J flag in (*,G)

PIM-SM SPT Switchover Overview (Cont.)



State in router C before switch; similar state on router A also.

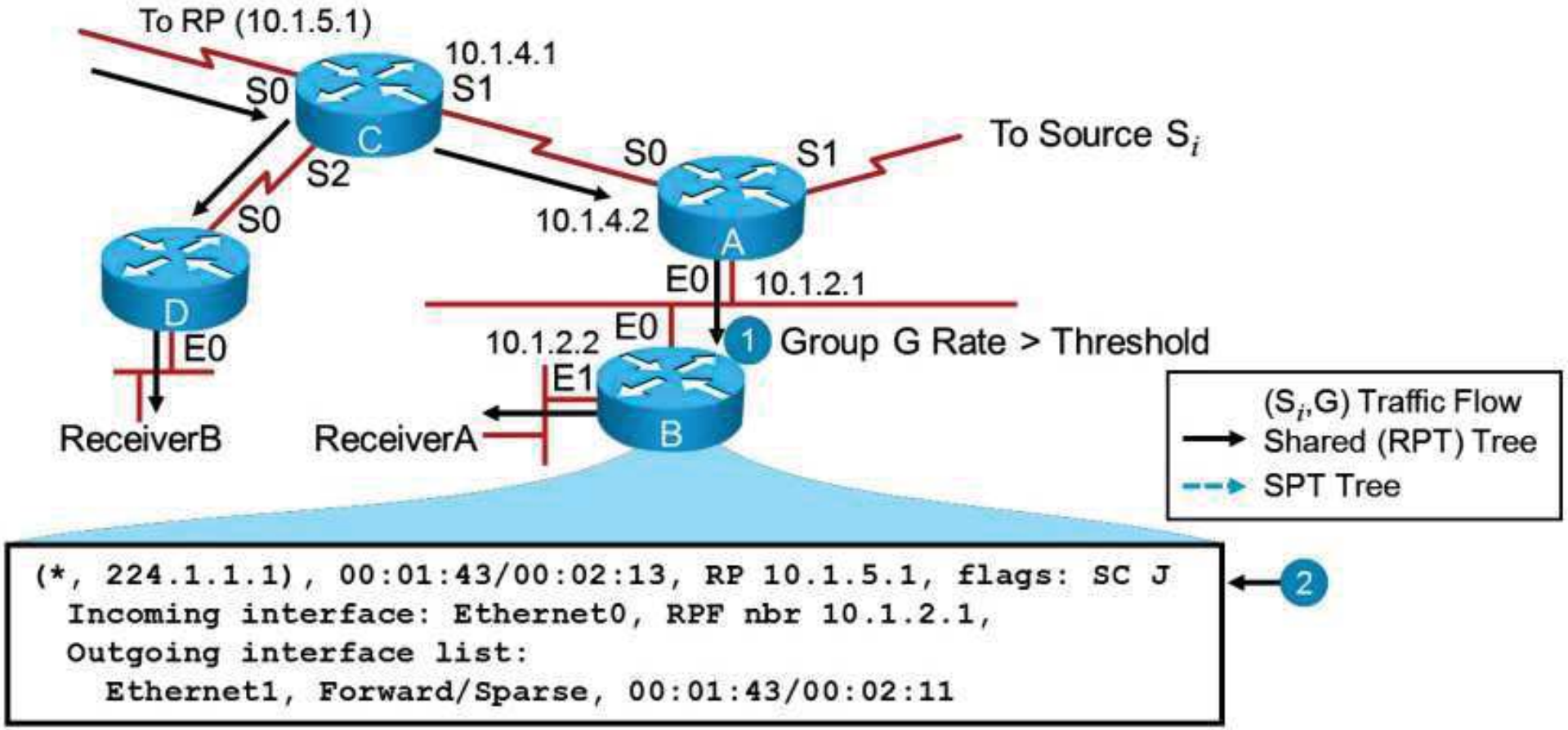
PIM-SM SPT Switchover Overview (Cont.)



State in router D before switch; similar state on router B also.

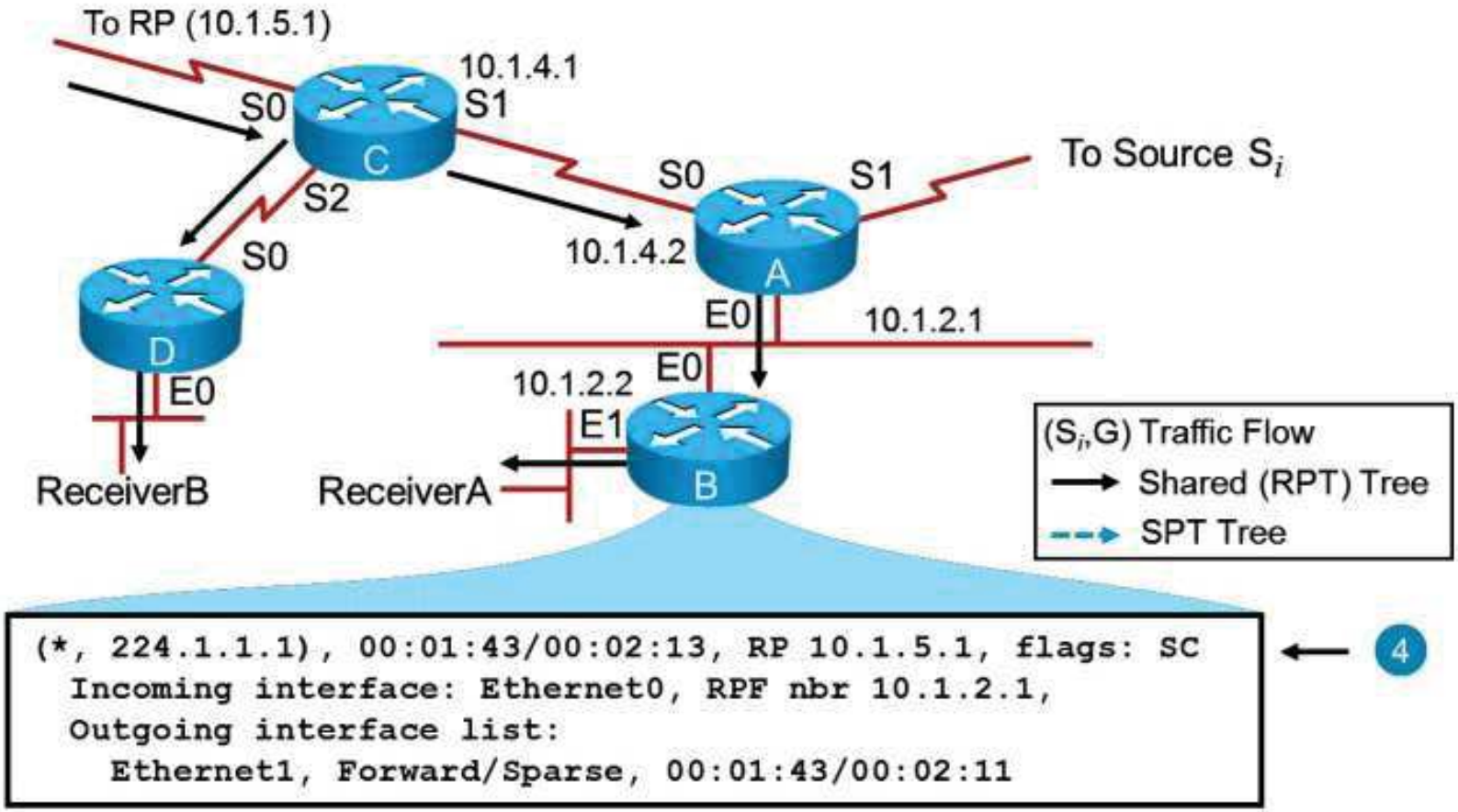
PIM-SM SPT Switchover Process

1. Group G rate exceeds SPT threshold at router B
2. Set J flag in (*,G) and wait for next (S_i,G) packet

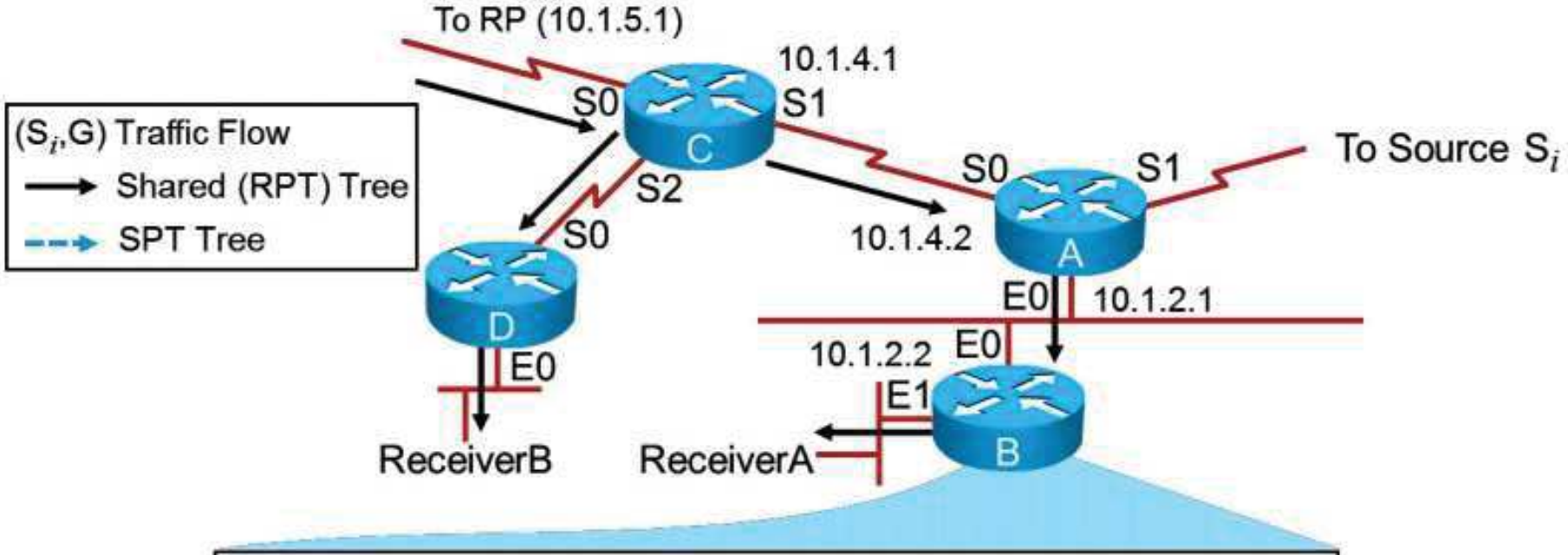


PIM-SM SPT Switchover Process (Cont.)

3. (S_i,G) packet arrives down shared tree
4. Clear J flag in the (*,G) and create (S_i,G) state



PIM-SM SPT Switchover Process (Cont.)

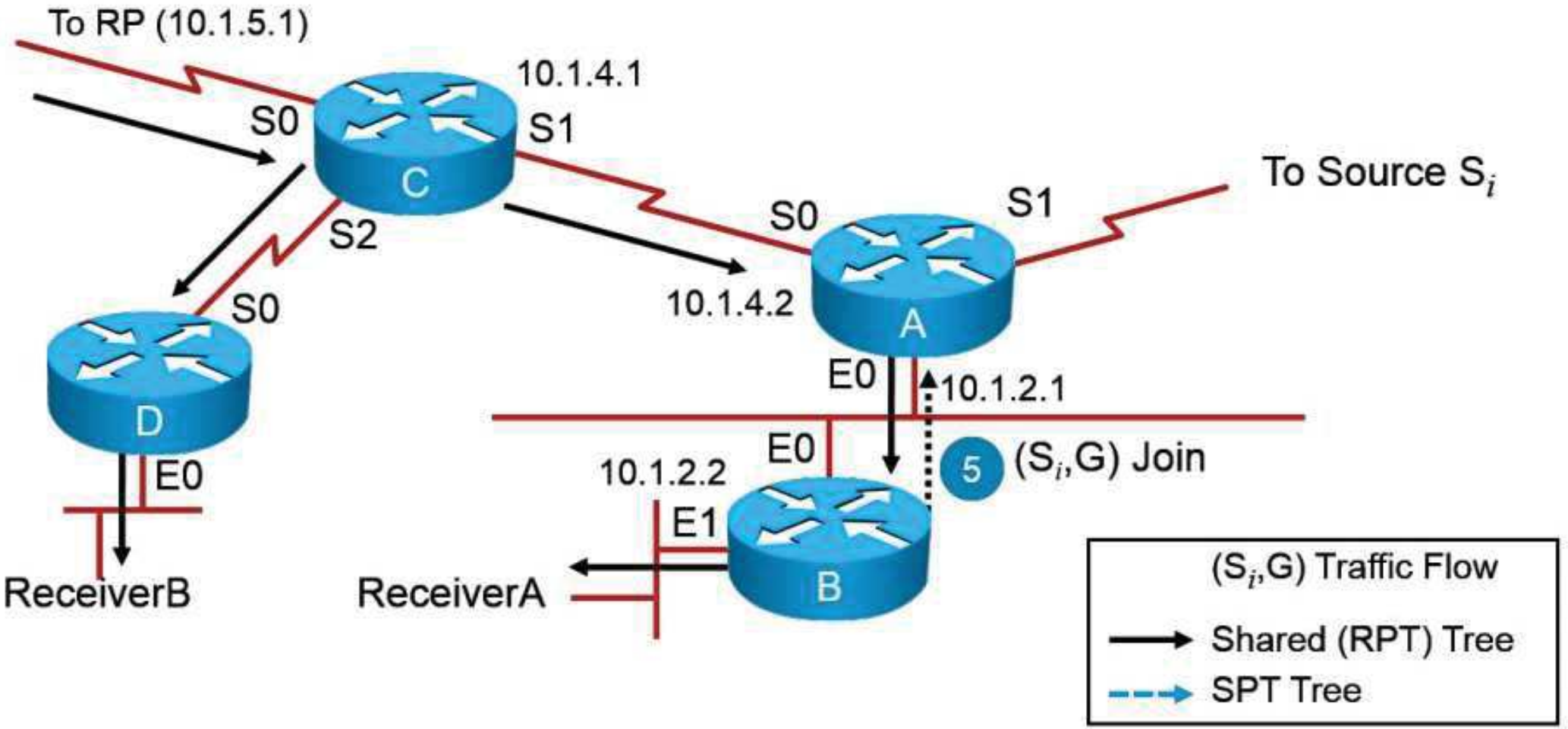



```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: SC
  Incoming interface: Ethernet0, RPF nbr 10.1.2.1,
  Outgoing interface list:
    Ethernet1, Forward/Sparse, 00:01:43/00:02:11

(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: CJT
  Incoming interface: Ethernet0, RPF nbr 10.1.2.1
  Outgoing interface list:
    Ethernet1, Forward/Sparse, 00:13:28/00:02:53
```

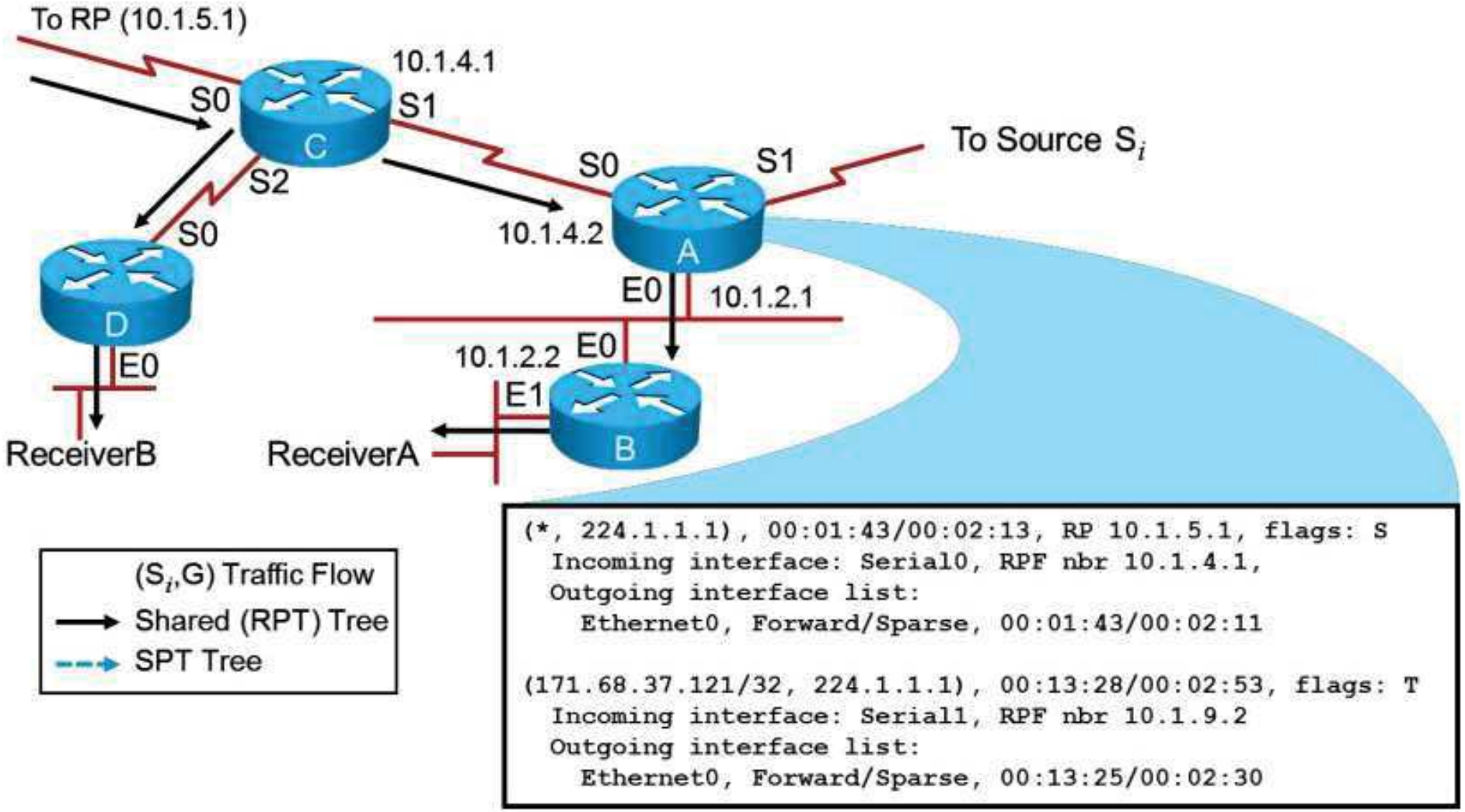
New state in router B.

PIM-SM SPT Switchover Process (Cont.)

5. (S_i,G) join is sent toward Si



PIM-SM SPT Switchover Process (Cont.)

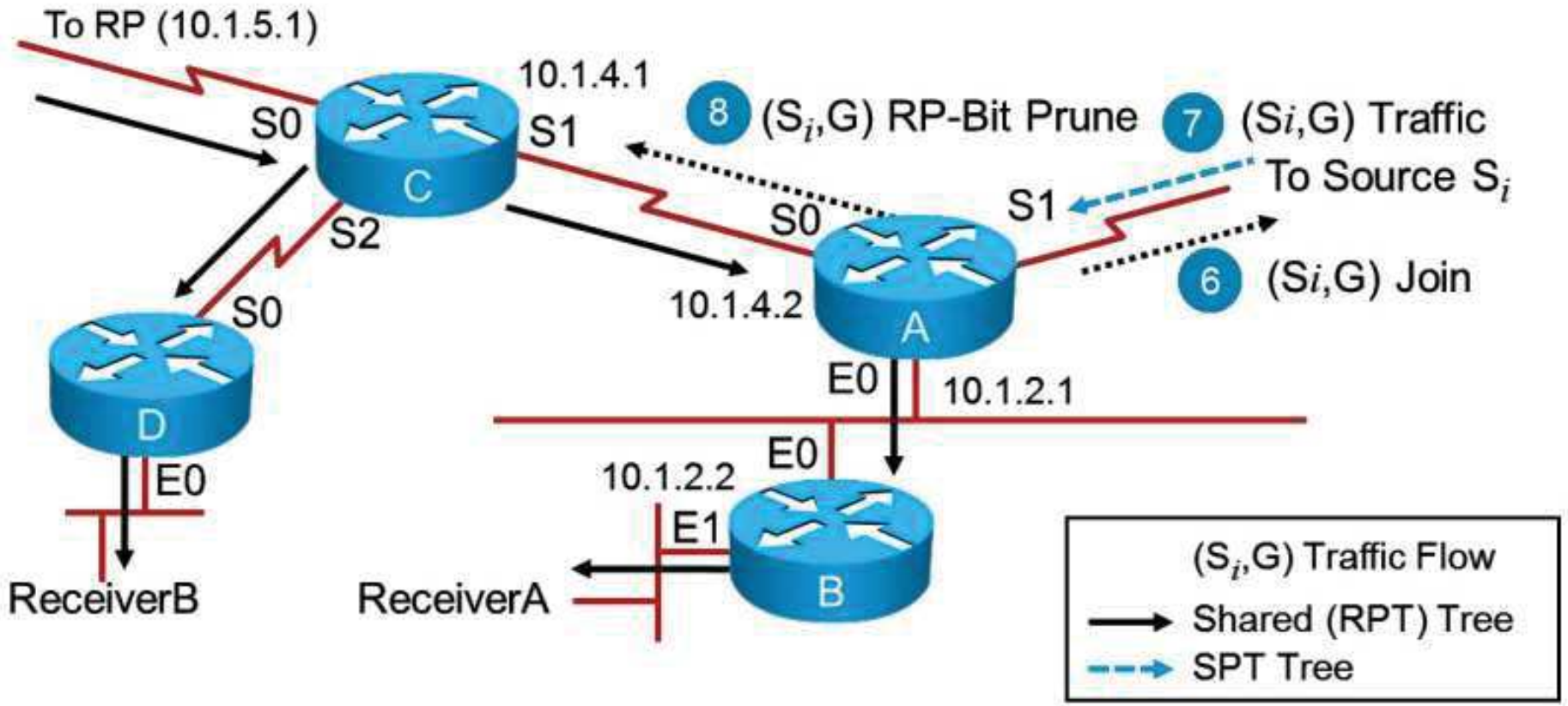



```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
  Incoming interface: Serial0, RPF nbr 10.1.4.1,
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:01:43/00:02:11

(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: T
  Incoming interface: Serial1, RPF nbr 10.1.9.2
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:13:25/00:02:30
```

New state in router A

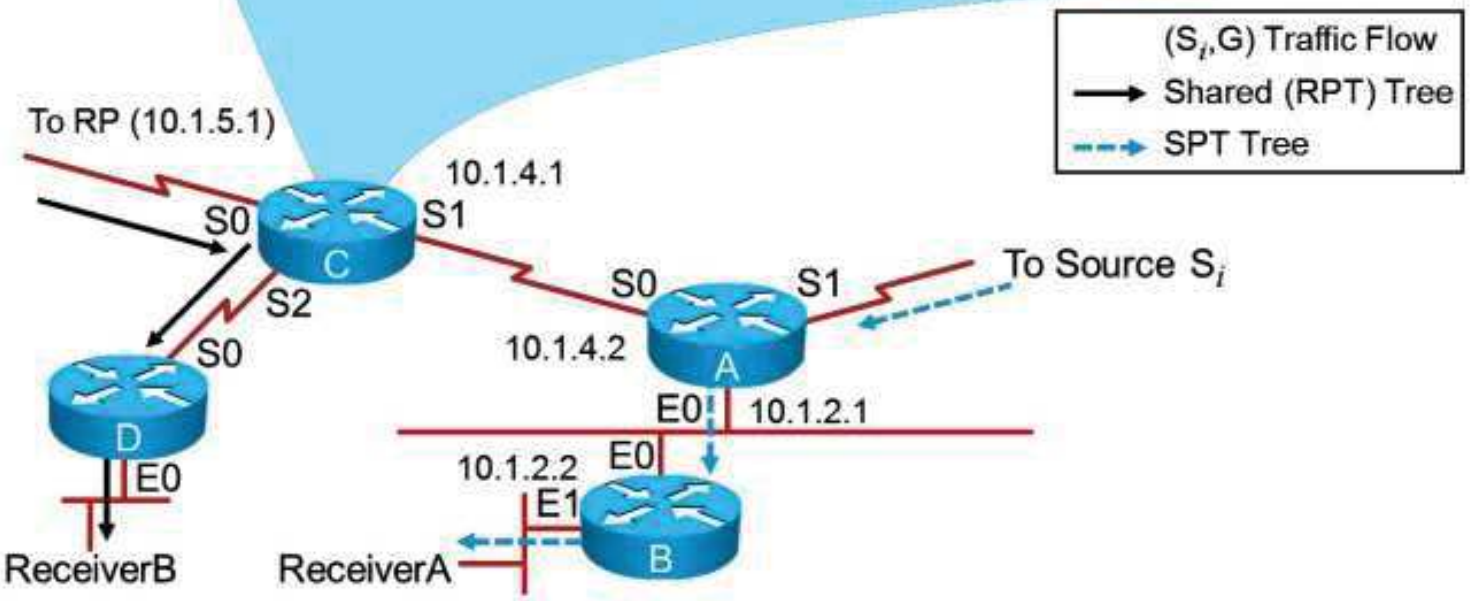
PIM-SM SPT Switchover Process (Cont.)

6. Router A forwards (S_i,G) join toward Si
7. Packets start arriving down the source tree.
8. SPT and shared tree diverge, triggering (S_i,G) RP-bit prunes toward RP



PIM-SM SPT Switchover Process (Cont.)

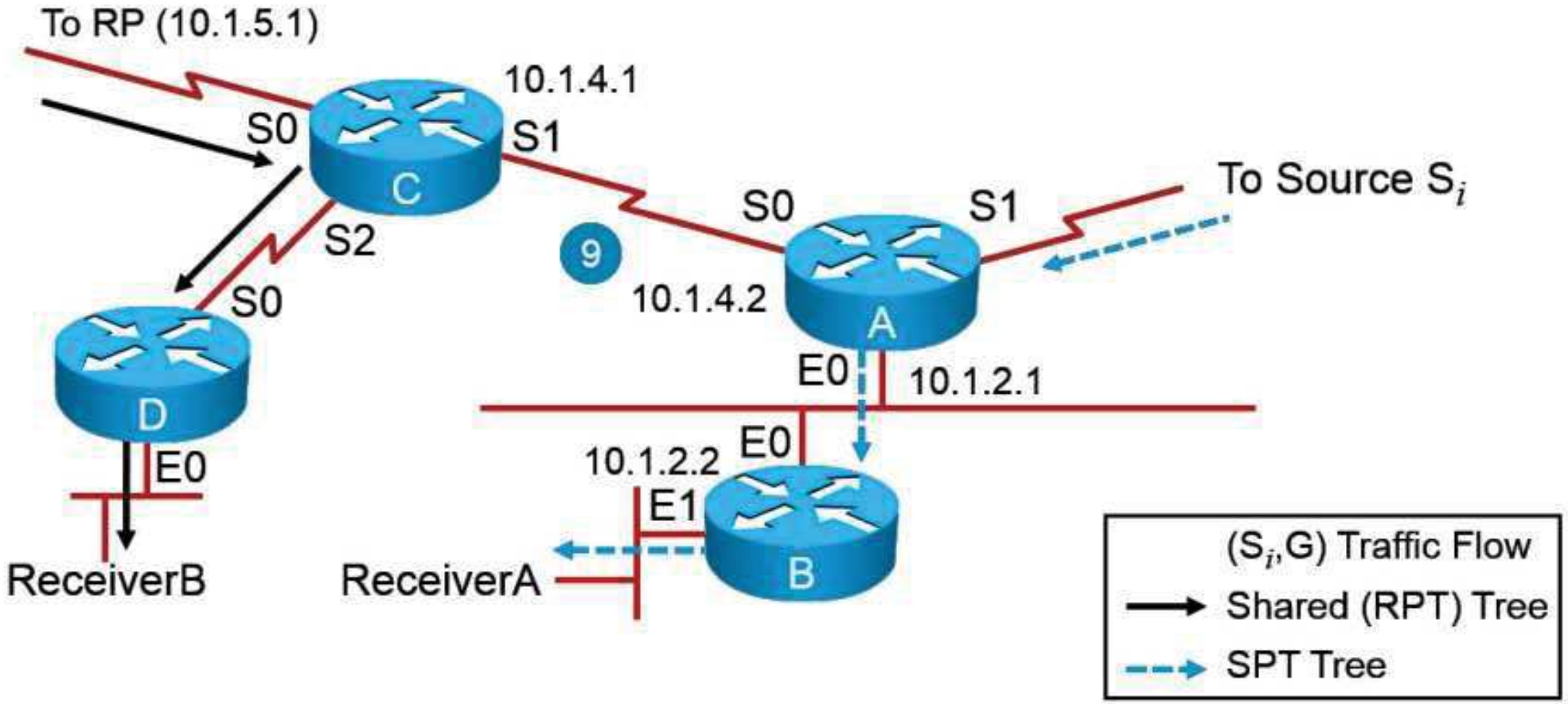
```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
  Incoming interface: Serial0, RPF nbr 10.1.5.1,
  Outgoing interface list:
    Serial1, Forward/Sparse, 00:01:43/00:02:11
    Serial2, Forward/Sparse, 00:00:32/00:02:28

(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: R
  Incoming interface: Serial0, RPF nbr 10.1.5.1
  Outgoing interface list:
    Serial2, Forward/Sparse, 00:00:32/00:02:28
```



State in router C after receiving the (S*i*,G) RP-bit prune.

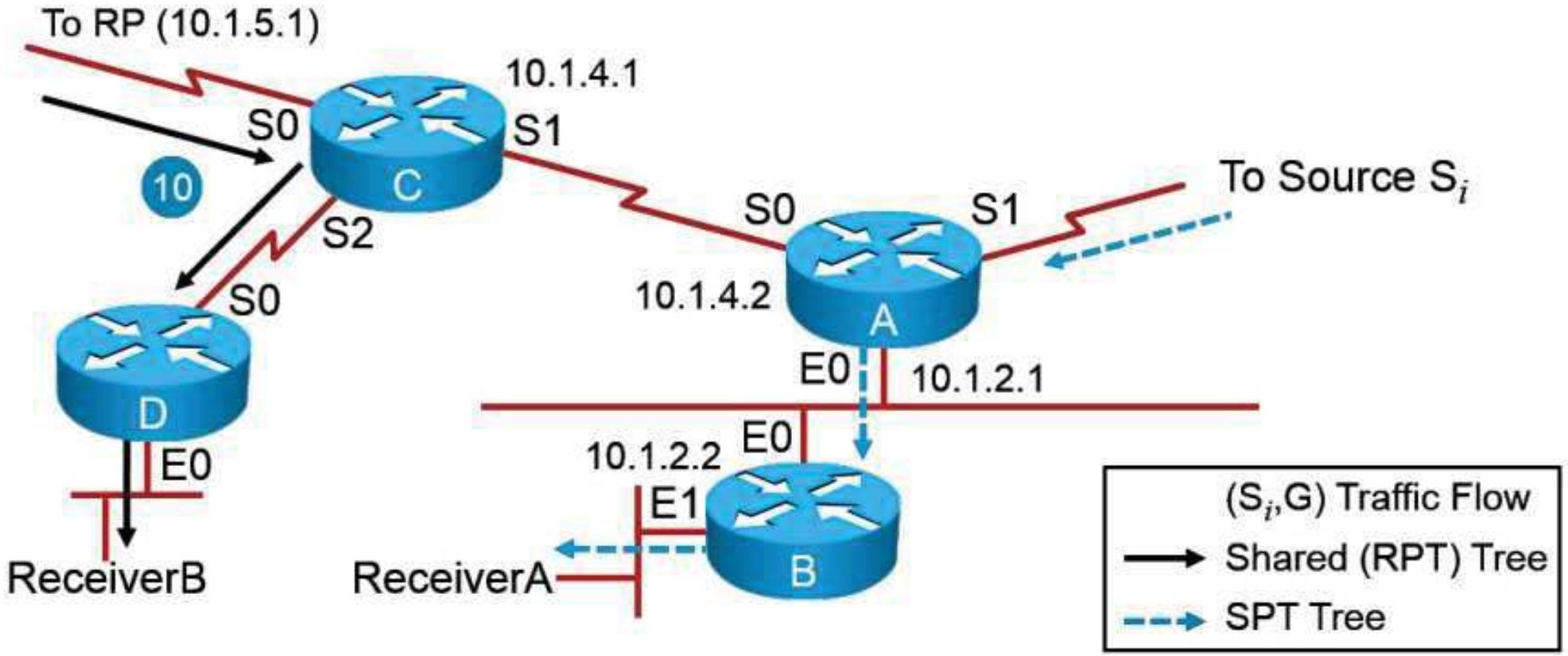
PIM-SM SPT Switchover Process (Cont.)

9. Unnecessary (S_i,G) traffic is pruned from the shared tree.



PIM-SM SPT Switchover Process (Cont.)

10. (S_i,G) traffic still flows via other branches of the shared tree.

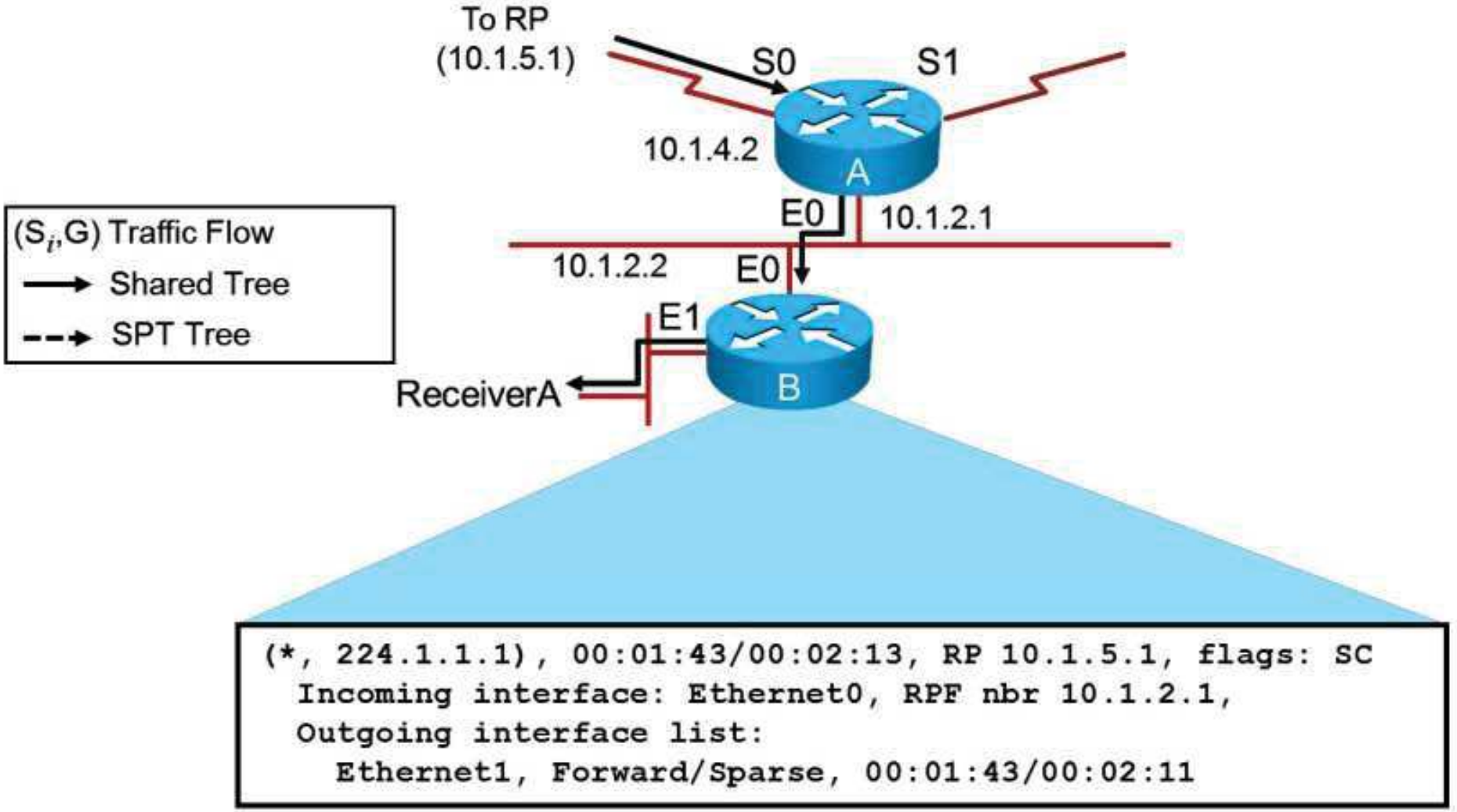


PIM-SM Shared Tree Pruning Overview

PIM-SM shared tree pruning characteristics:

- IGMP group times out and last host sends leave.
- Interface removed from all (*,G) and (S,G) entries:
  - If all interfaces in the OIL for (*,G) are pruned, then prunes are sent up shared tree toward RP.
  - Any (S,G) state is allowed to time out.

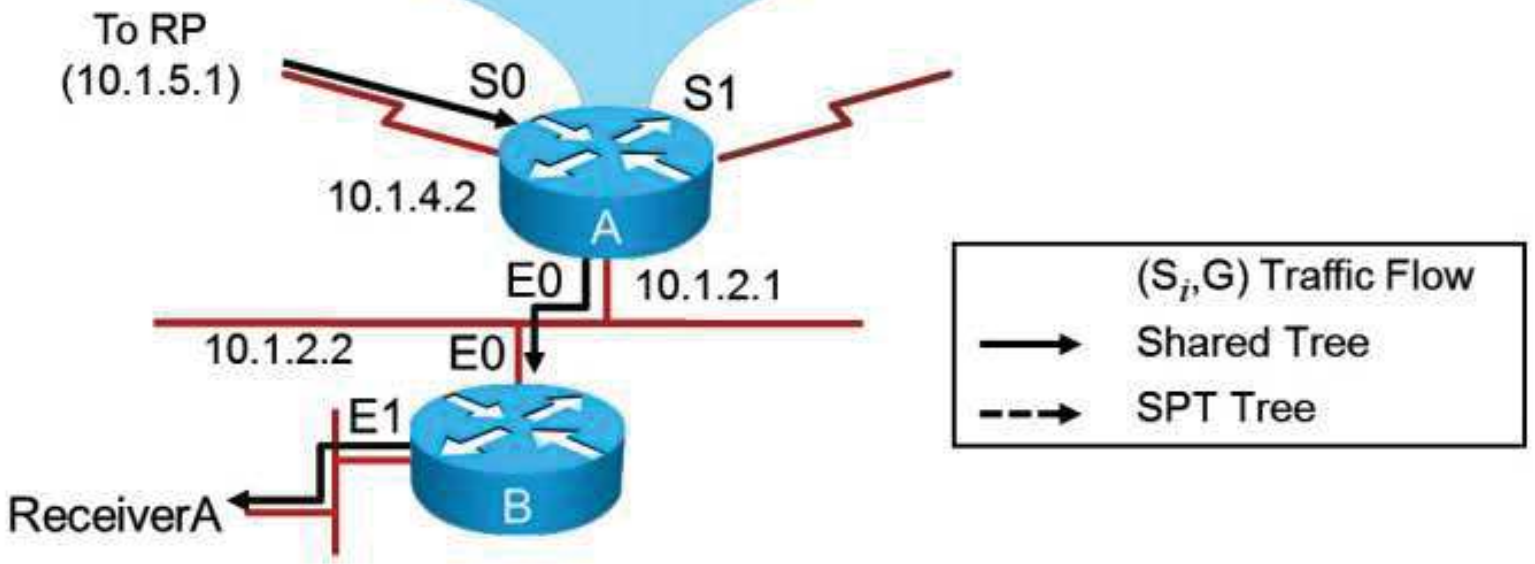
PIM-SM Shared Tree Pruning Overview (Cont.)



State in router B before pruning.

PIM-SM Shared Tree Pruning Overview (Cont.)

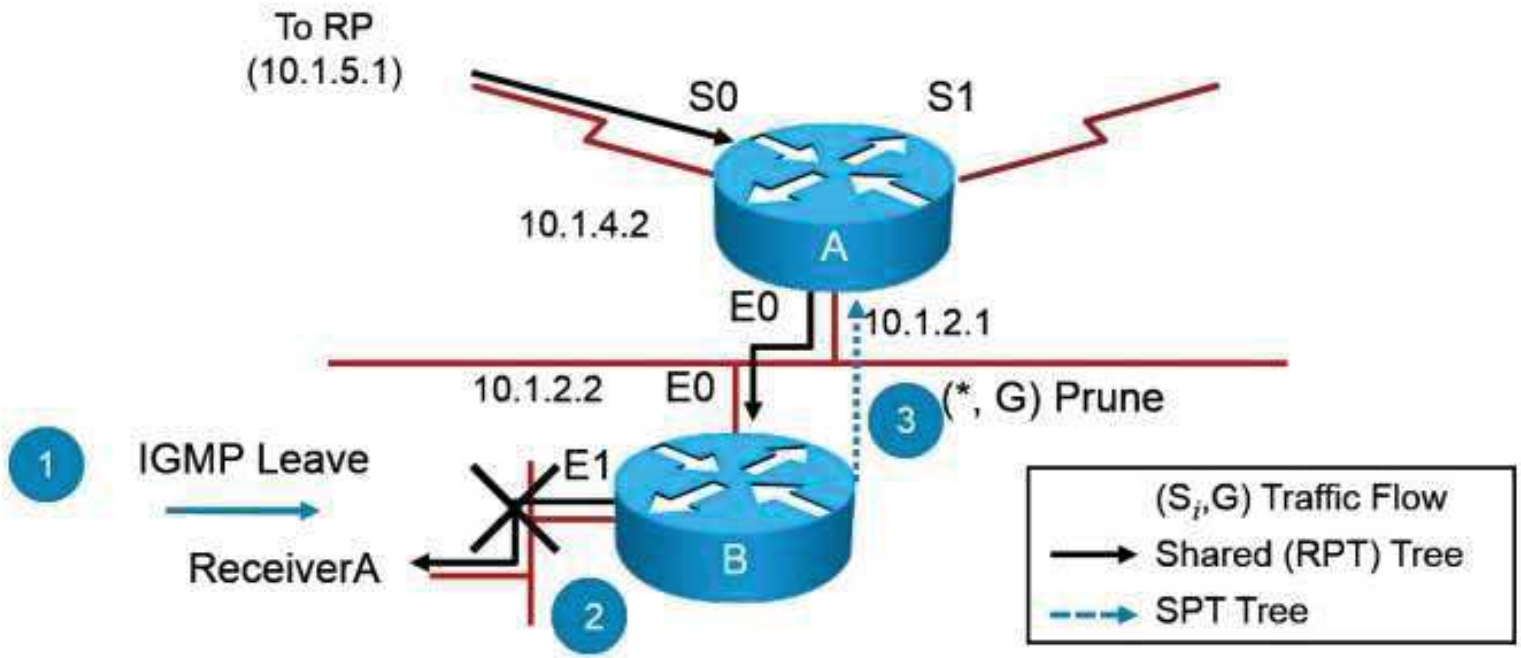
```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
  Incoming interface: Serial0, RPF nbr 10.1.4.1,
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:01:43/00:02:11
```



State in router A before pruning.

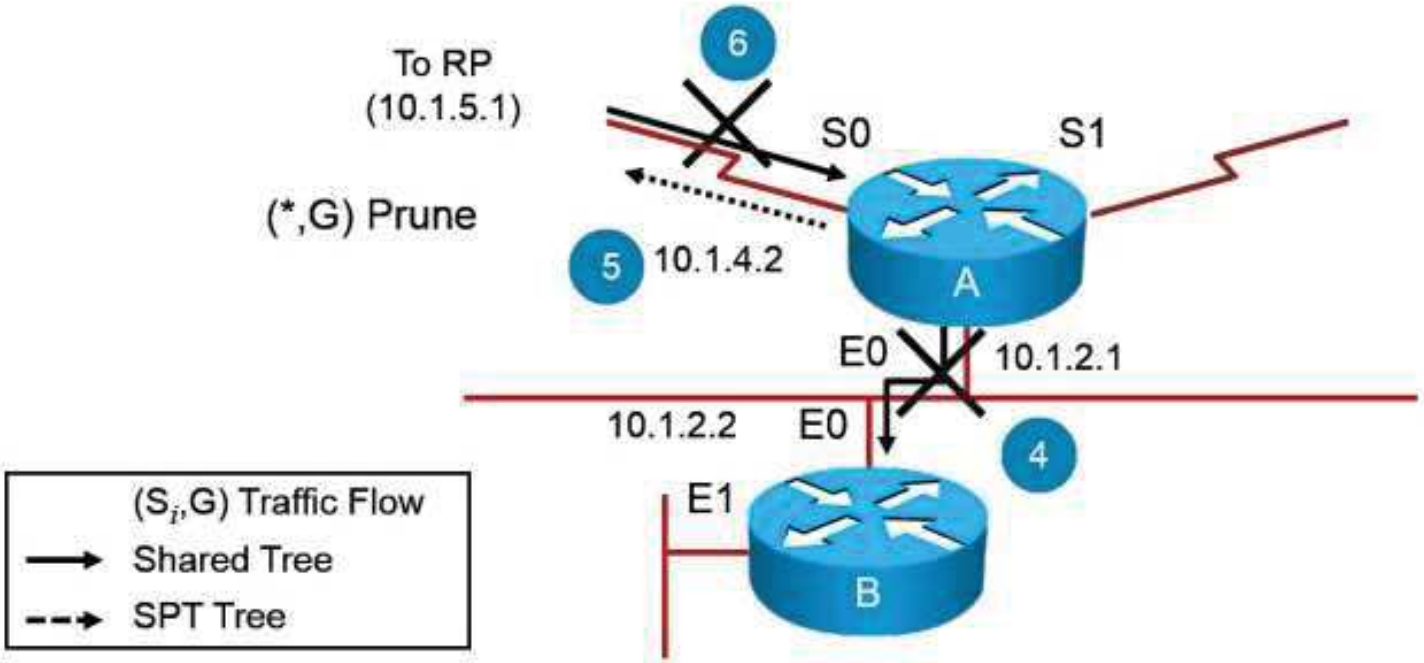
PIM-SM Pruning Shared Tree Process

1. Router B is a leaf router; last host (ReceiverA) leaves group G
2. Router B removes E1 from (*,G) and any (S_i,G) OIL
3. Router B (*,G) OIL now empty; sends (*,G) prune toward RP

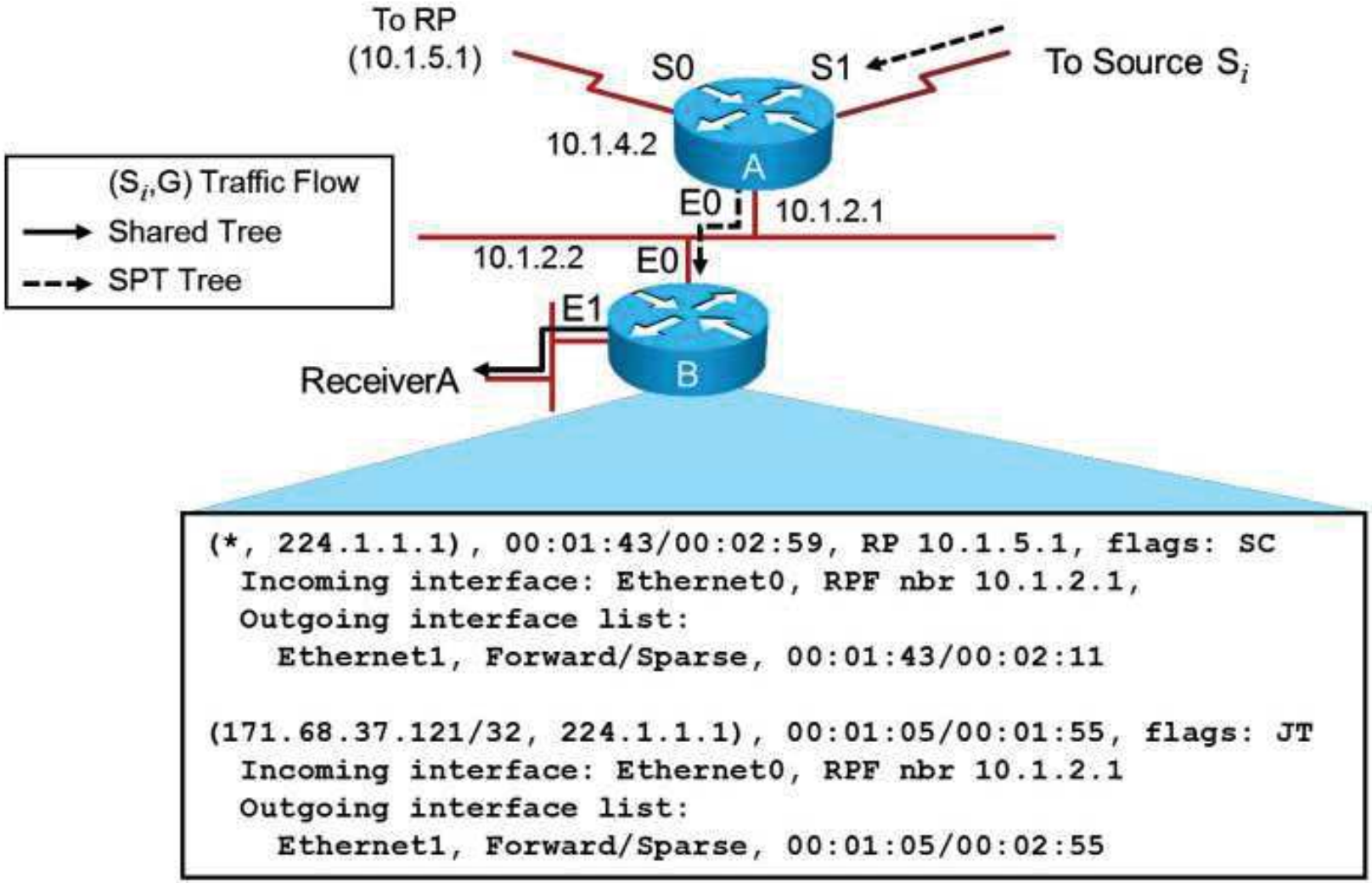


PIM-SM Pruning Shared Tree Process (Cont.)

4. Router A receives prune, removes E0 from (*,G) OIL (after the 3-second multiaccess network prune delay)
5. Router A (*,G) OIL now empty; sends (*,G) prune toward RP
6. Pruning continues back toward RP



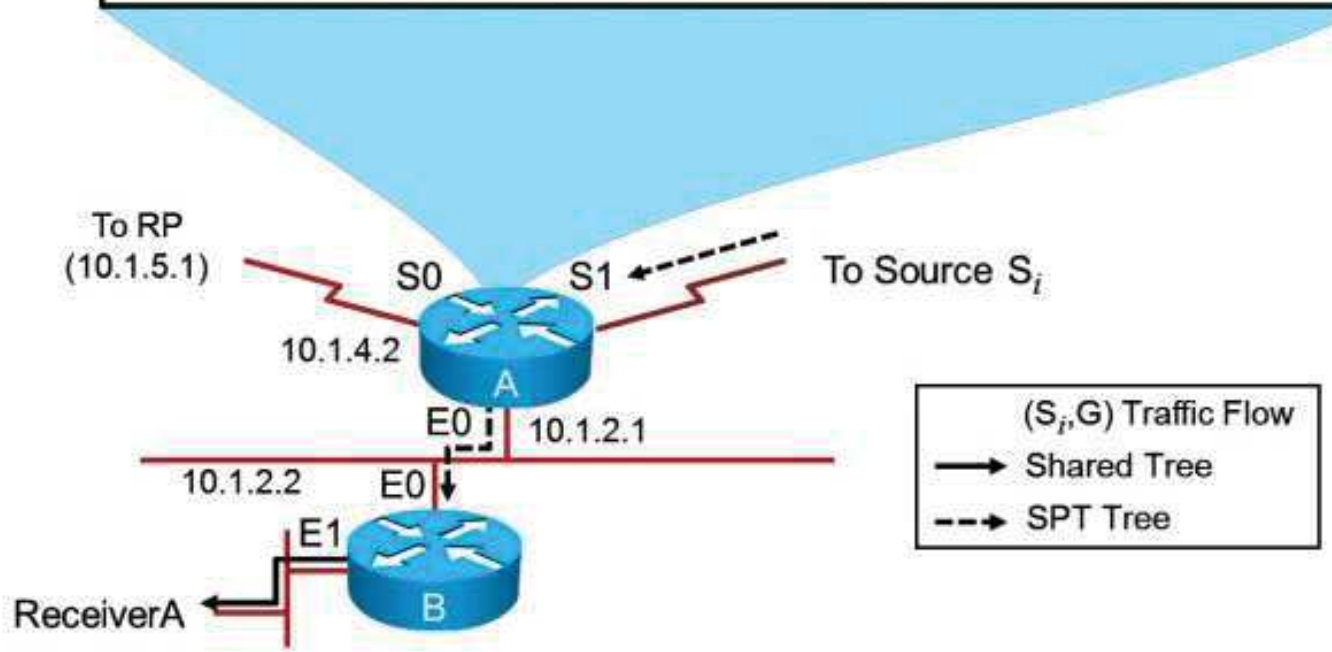
PIM-SM SPT Pruning Overview



State in router B before pruning.

PIM-SM SPT Pruning Overview (Cont.)

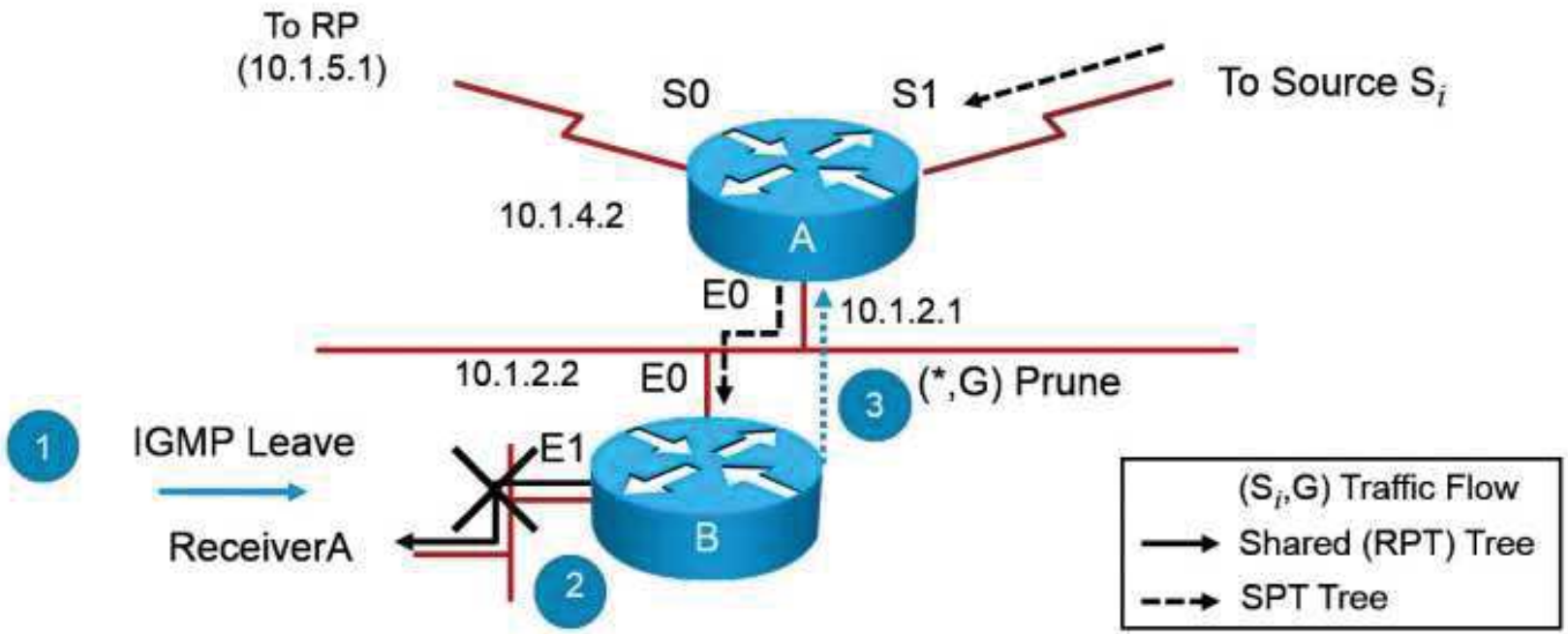
```
(*, 224.1.1.1), 00:01:43/00:02:59, RP 10.1.5.1, flags: S
  Incoming interface: Serial0, RPF nbr 10.1.4.1,
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:01:43/00:02:11

(171.68.37.121/32, 224.1.1.1), 00:01:05/00:01:55, flags: T
  Incoming interface: Serial1, RPF nbr 10.1.9.2
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:01:05/00:02:55
```



State in router A before pruning.

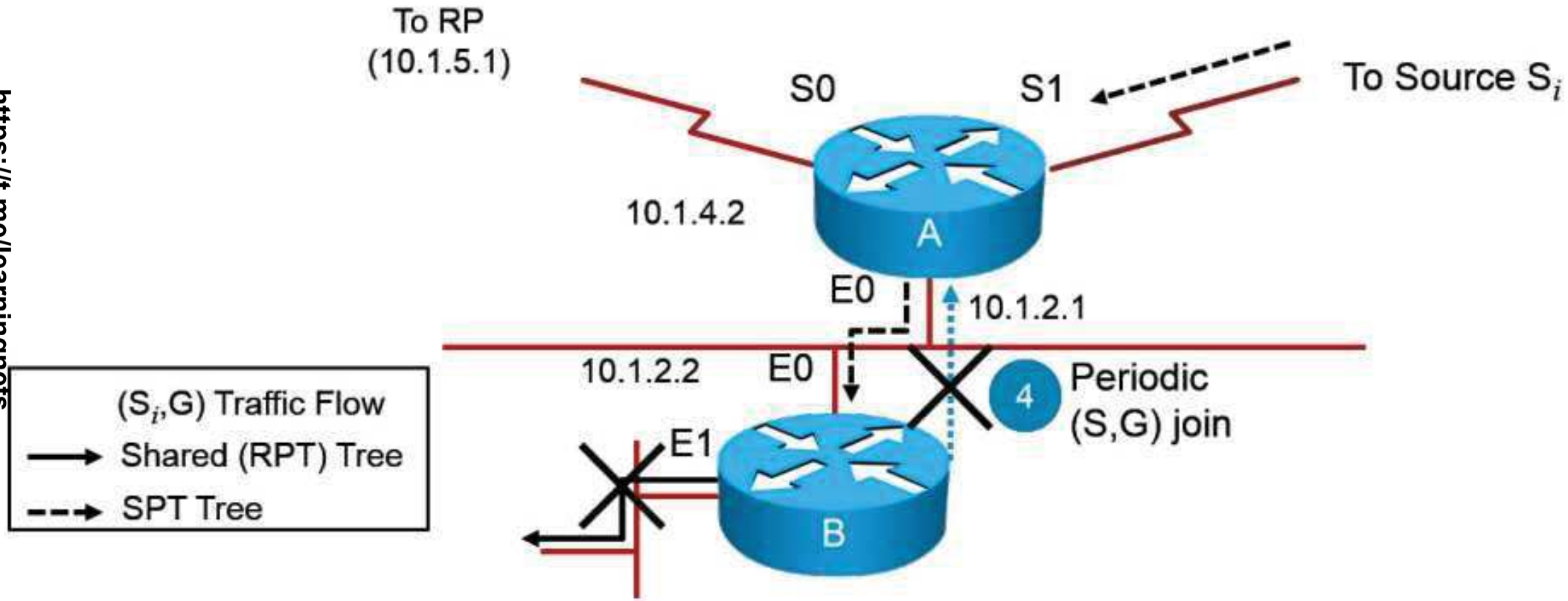
PIM-SM Pruning SPT Process

1. Router B is a leaf router; last host (ReceiverA) leaves group G .
2. Router B removes E1 from (*,G) and any (S_i,G) OIL.
3. Router B (*,G) OIL now empty; sends (*,G) prune toward RP.



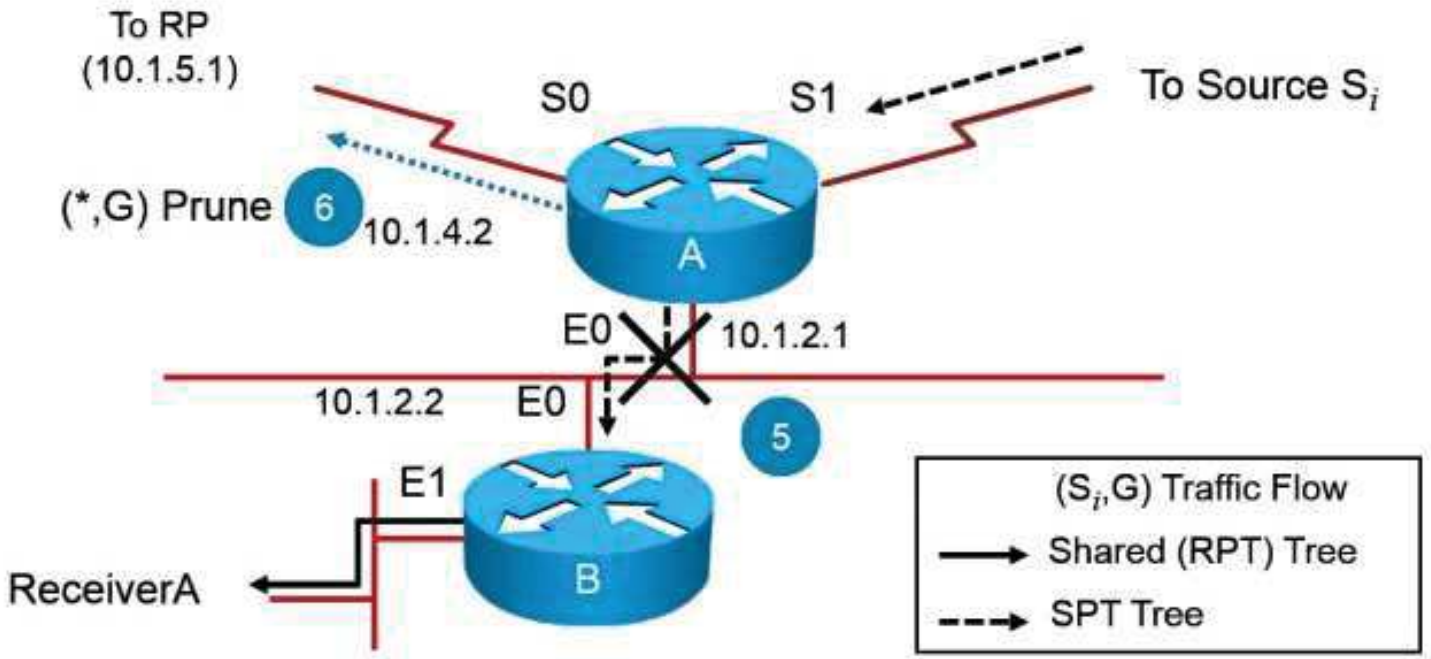
PIM-SM Pruning SPT Process (Cont.)

4. Router B stops sending periodic (S,G) joins.

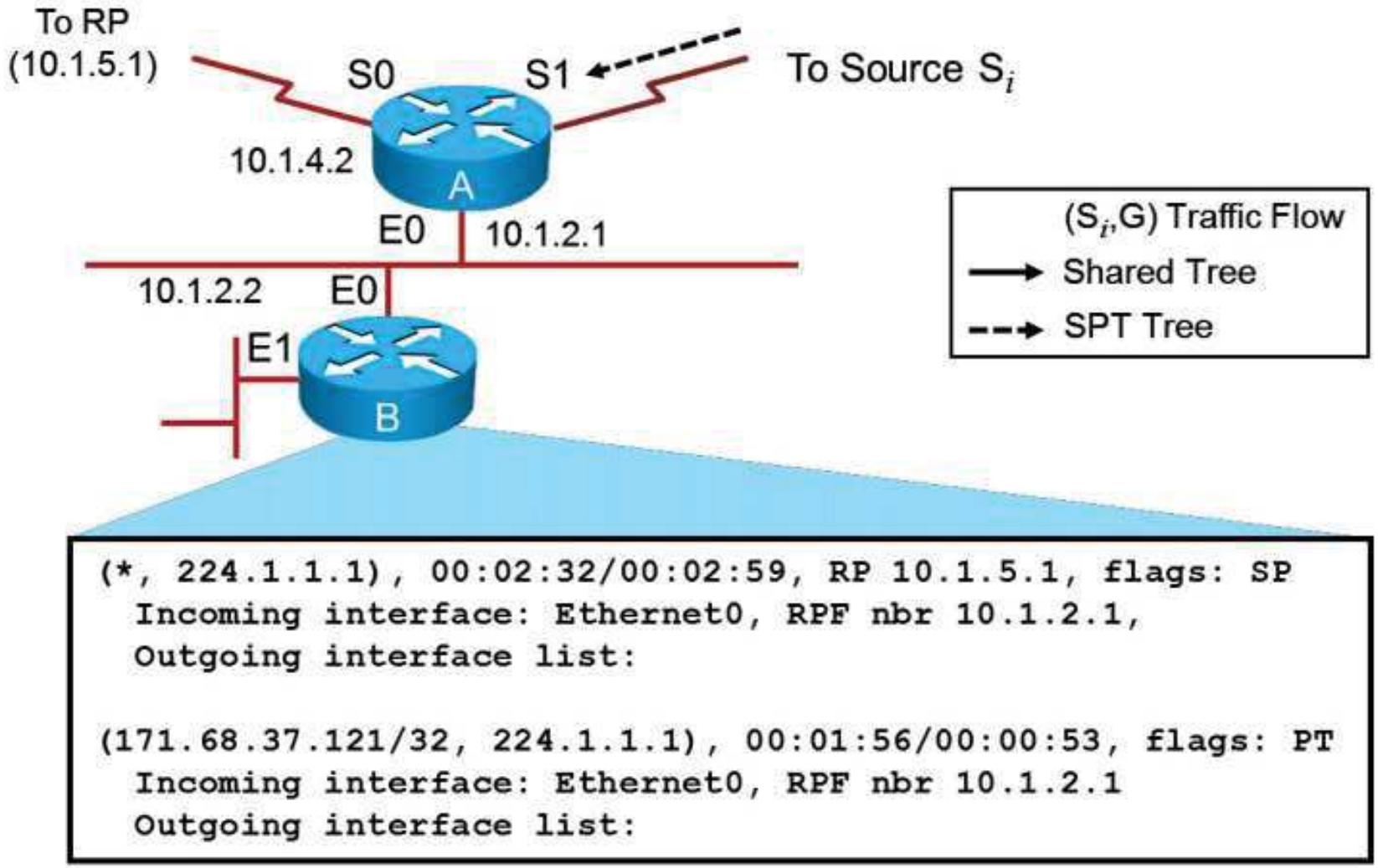


PIM-SM Pruning SPT Process (Cont.)

5. Router A receives prune; removes E0 from (*,G) OIL (after the 3-second multiaccess network prune delay).
6. Router A (*,G) OIL now empty; sends (*,G) prune toward RP.



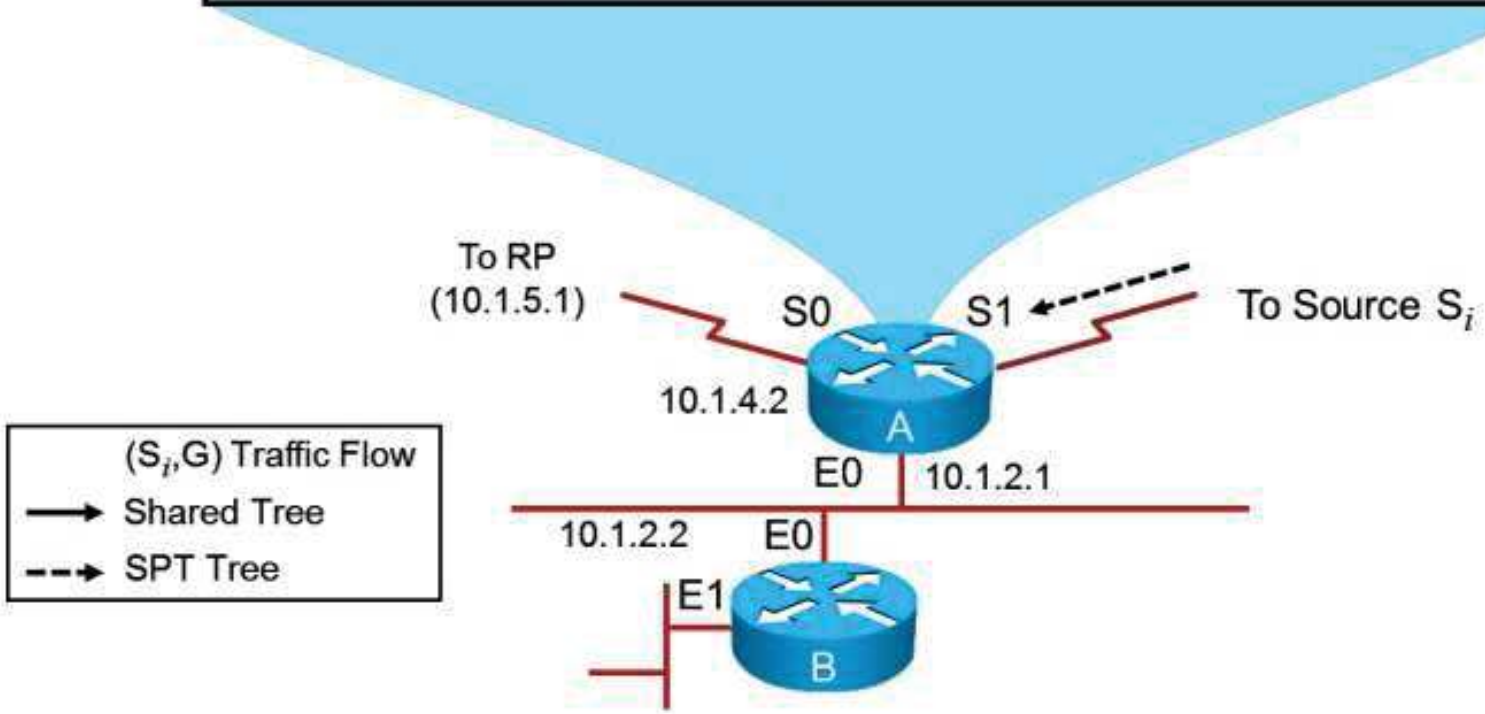
PIM-SM Pruning SPT Process (Cont.)



State in router B after pruning.

PIM-SM Pruning SPT Process (Cont.)

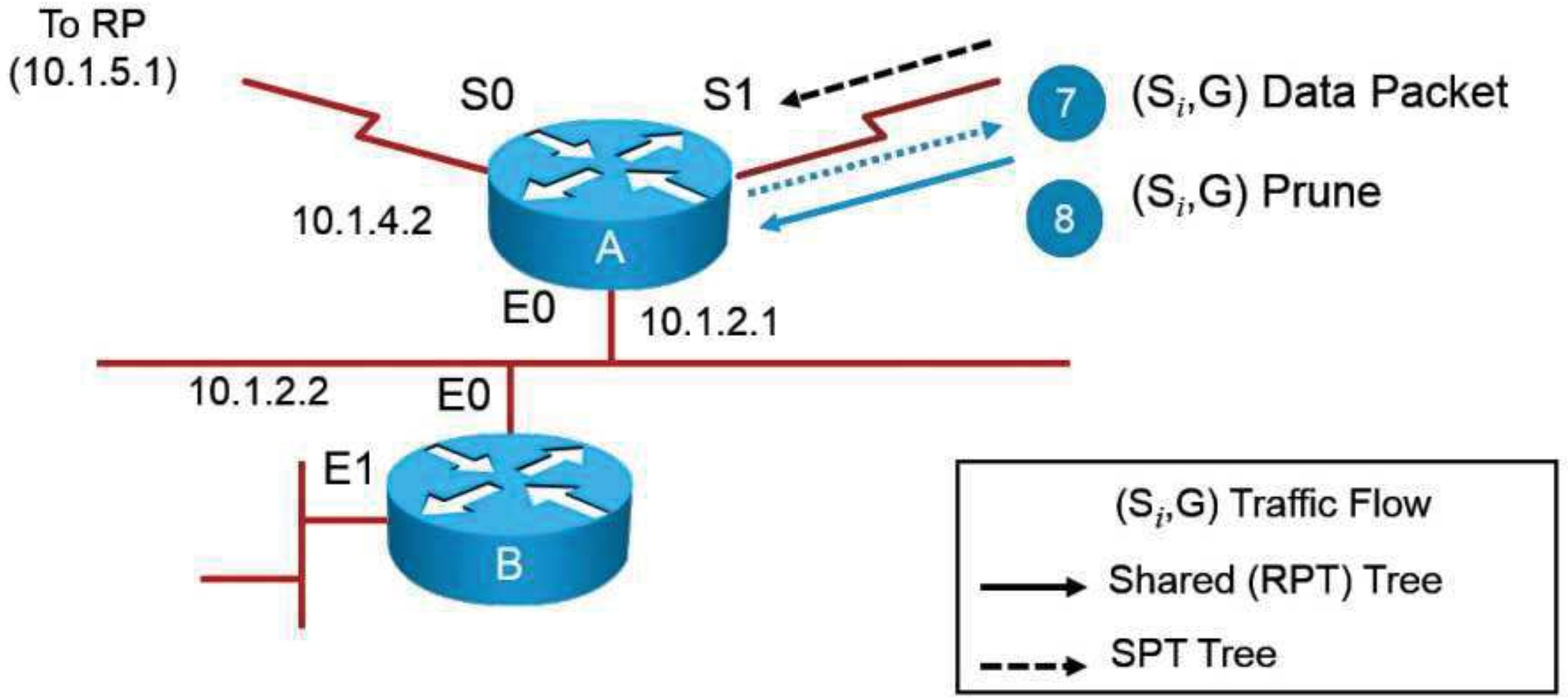
```
(*, 224.1.1.1), 00:02:32/00:02:59, RP 10.1.5.1, flags: SP
  Incoming interface: Serial0, RPF nbr 10.1.4.1,
  Outgoing interface list:

(171.68.37.121/32, 224.1.1.1), 00:01:56/00:00:53, flags: PT
  Incoming interface: Serial1, RPF nbr 10.1.9.2
  Outgoing interface list:
```



State in router A after pruning.

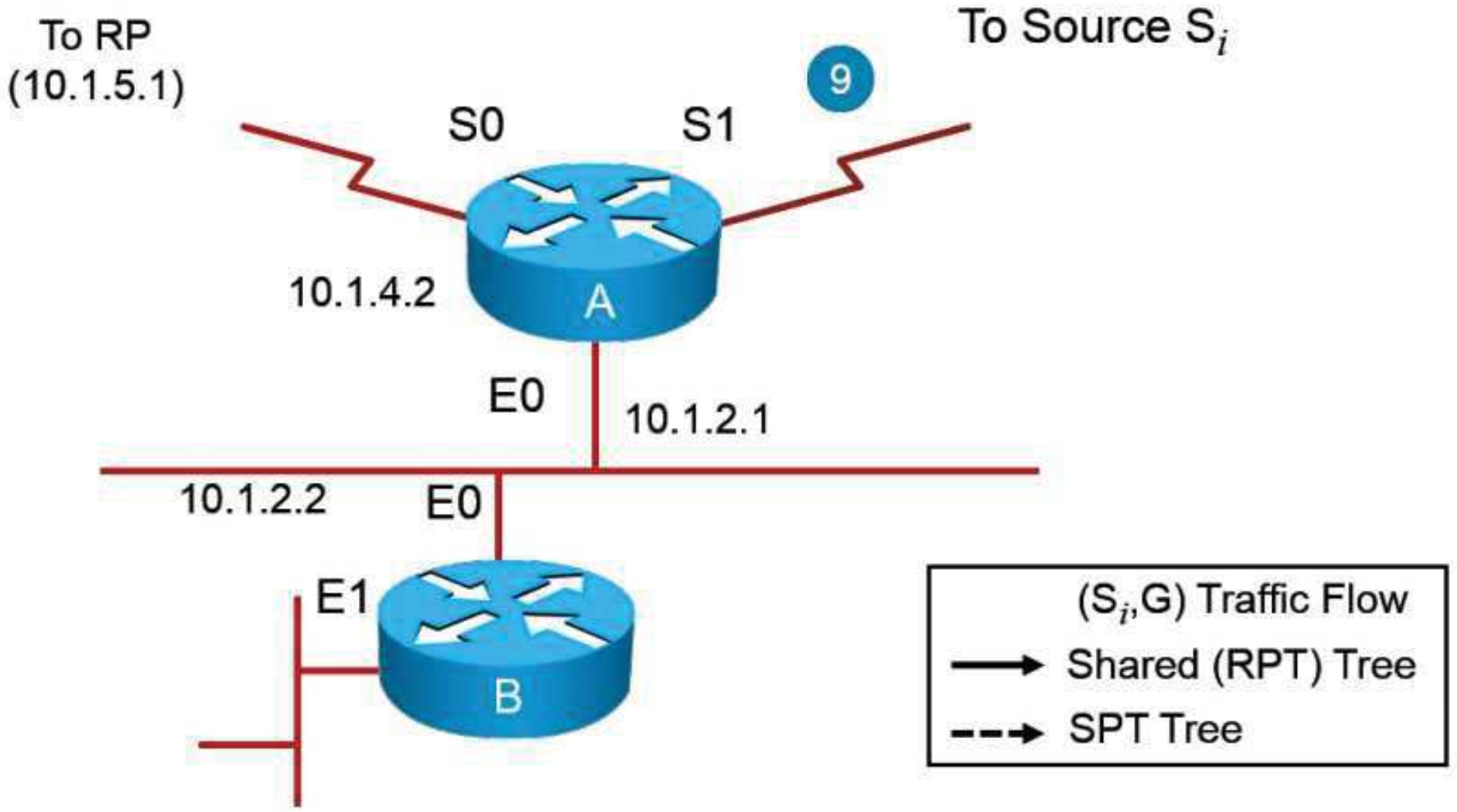
PIM-SM Pruning SPT Process (Cont.)

7. Another (S_i,G) data packet arrives via Serial1.
8. Router A responds by sending an (Si,G) prune toward source.



PIM-SM Pruning SPT Process (Cont.)

9. (S_i,G) traffic ceases flowing down SPT.



Enable PIM-SM

Enables multicast routing

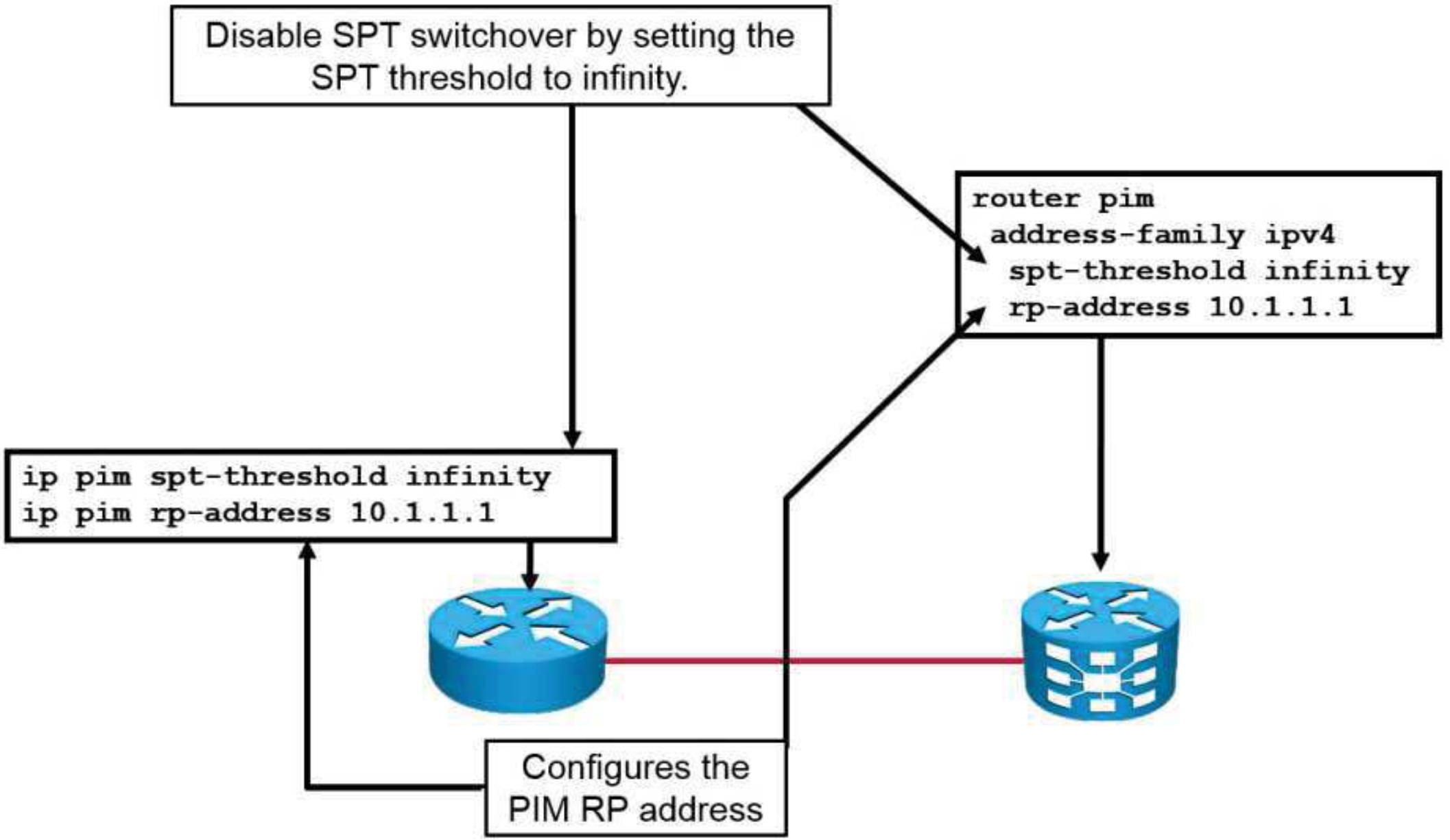
```
ipv4 access-list PIM_FILTER
 permit host 10.1.1.2
!
multicast-routing
 address-family ipv4
  interface GigabitEthernet0/0/0/0
   enable
!
router pim
 neighbor-filter PIM_FILTER
 interface GigabitEthernet 0/0/0/0
  enable
```

Enables PIM filtering

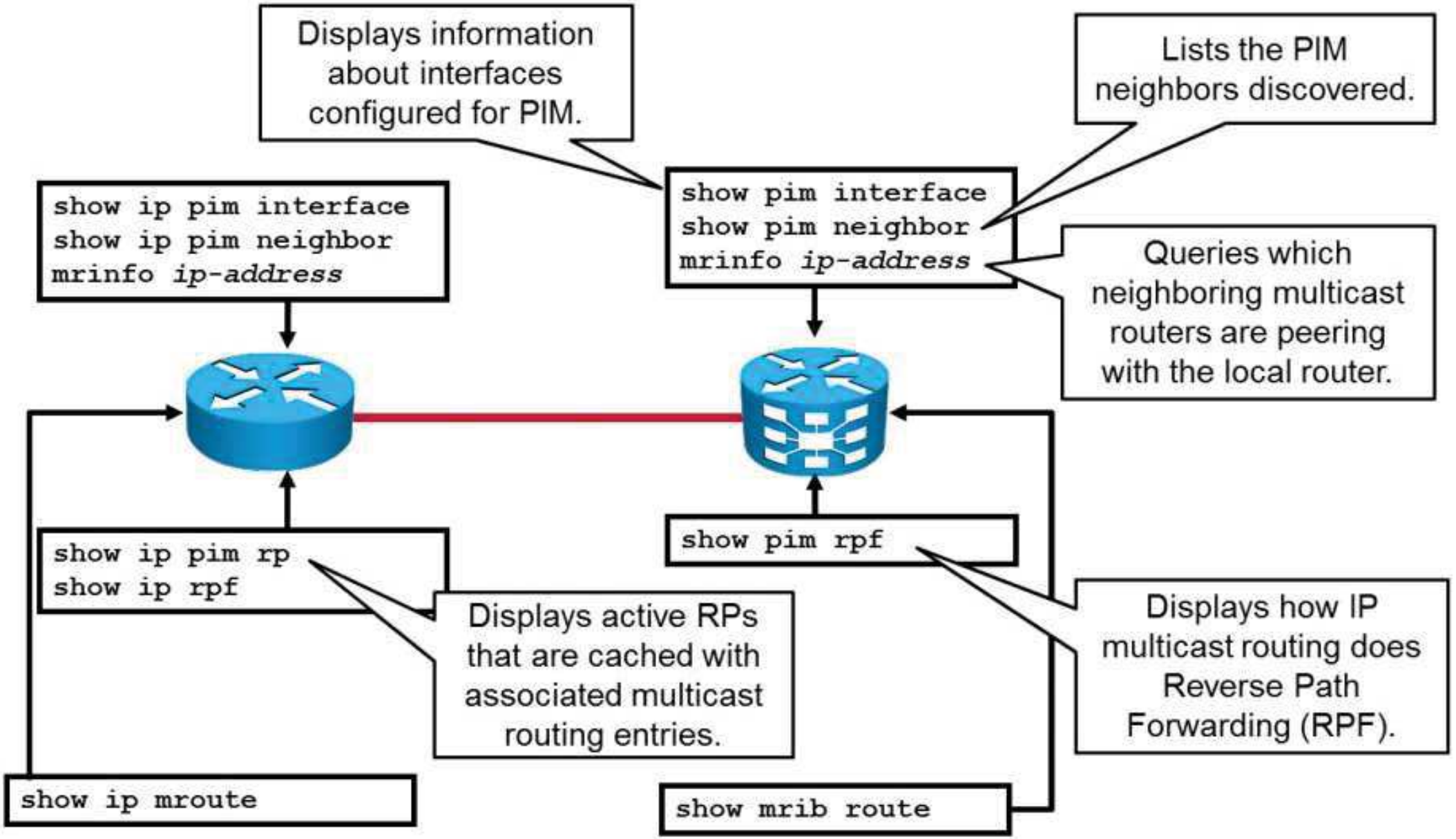
```
access-list 1 permit host 10.1.1.1
!
ip multicast-routing
ip pim neighbor-filter 1
!
interface GigabitEthernet0/0
 ip pim sparse-mode
```

Enables PIM-SM on an interface

Static RP



Finding PIM Neighbors and Checking RP Information



Troubleshooting PIM-SM Guidelines

Troubleshooting PIM-SM guidelines:

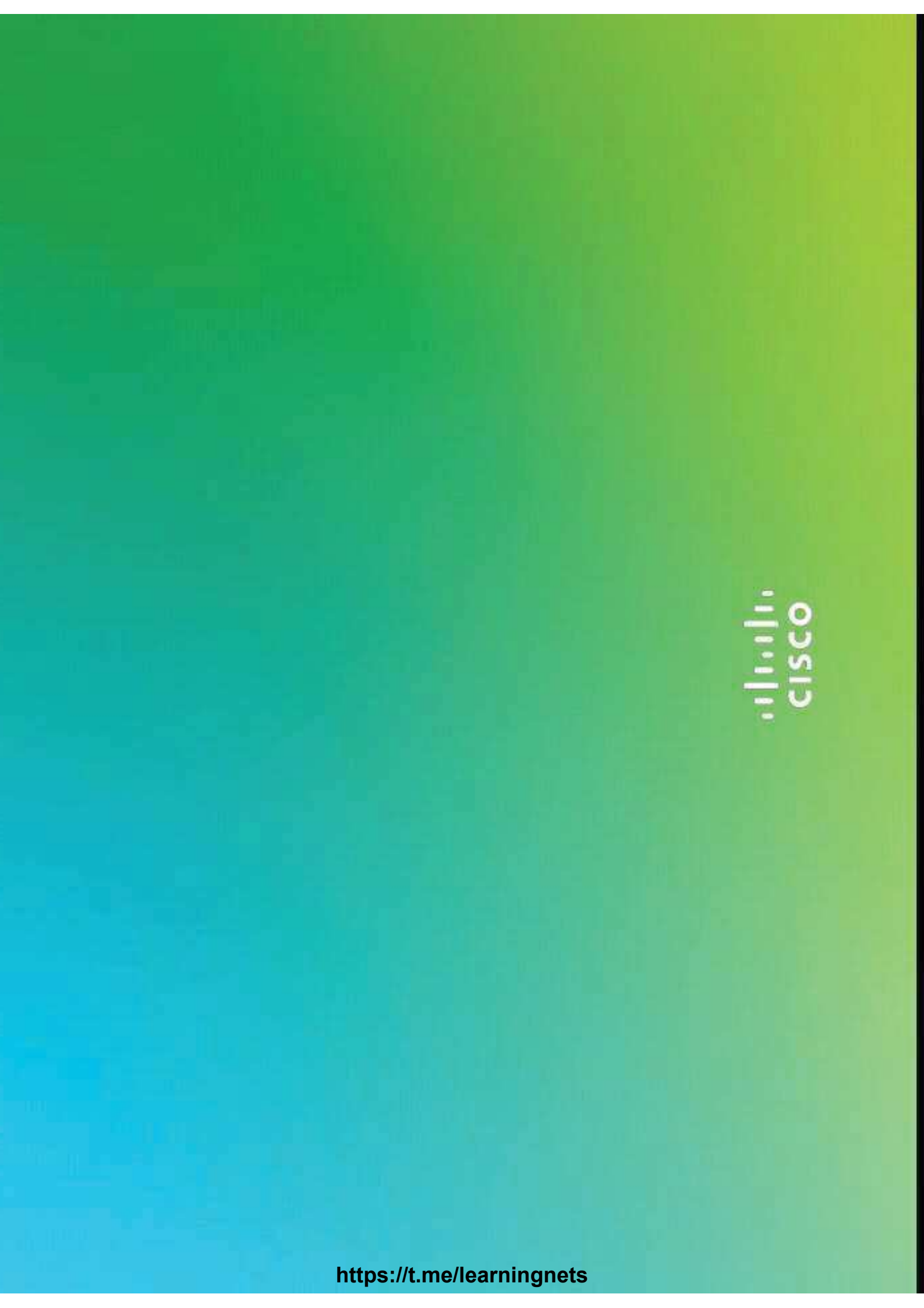
- Check whether receivers are active.
- Check whether a group is present at the last-hop router and the RP.
- Check whether sources are active.
- Check whether a group is present at the first-hop router.
- Check the RP configuration on the routers between the source and the RP.
- Check the RP configuration on the routers between the RP and the receivers.
- Debug the creation of distribution trees.

Summary

- PIM-SM operates with shared tree.
- DR joins the shared tree at RP when a receiver joins a group.
- First hop router builds source tree towards RP.
- Multicast tree can bypass the RP in certain situations.
- PIM uses control packets to send control-type information.
- PIM state is represented by entries in the multicast routing table.
- State is maintained by periodic updates.
- CLI commands show the contents of the multicast routing table.
- Different rules govern creation, deletion and maintenance of:
  - (*,G) states
  - (S,G) states
  - OIL
- Each state is described by several different flags.

Summary (Cont.)

- PIM-SM control messages are passed between multicast routers.
- In the process of creating full multicast distribution tree:
  - the receiver can join first.
  - the source can join first.
  - receivers can be between first hop router and the RP.
- SPT switchover is triggered by exceeding the threshold value.
- When PIM-SM shared tree is pruned, interface is removed from all (*,G) and (S,G) entries.
- In case of SPT pruning only (S,G) entries are affected.
- The easiest way of configuring RP is to configure static RP; there is little redundancy in such setup.
- To successfully troubleshoot multicast implementation, check all components one by one.



Implementing PIM-SM Enhancements

Intradomain and Interdomain Multicast Routing

Source Specific Multicast

Simplified solution for well-known sources, particularly in cases where there is a single source sending to a given group:

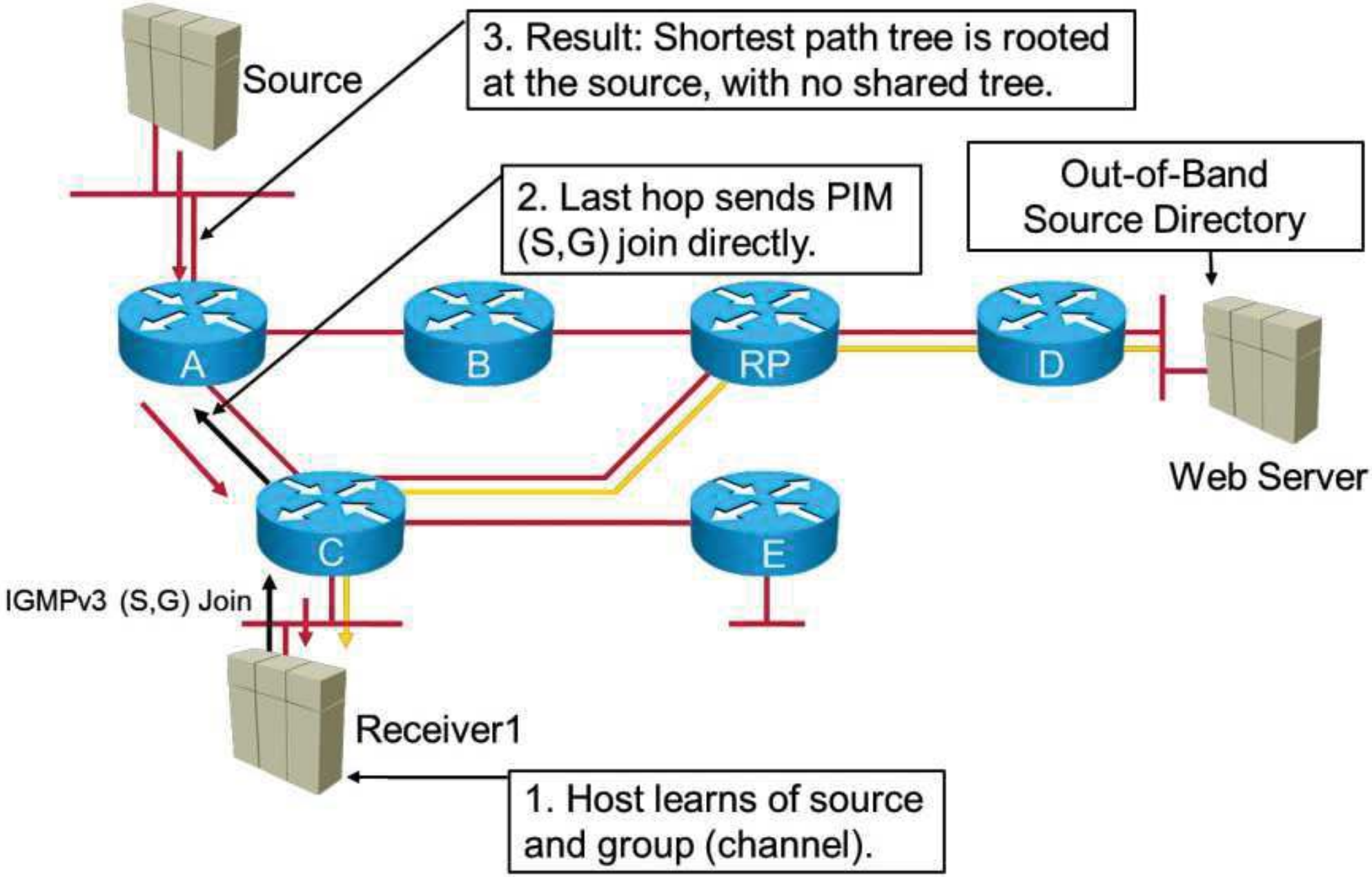
- Allows immediate use of shortest forwarding path from a specific source to the receivers without the need to create a shared tree
- Eliminates dependence on MSDP for finding sources
- Simplifies address allocation for global, single-source groups when combined with the elimination of shared trees (232.0.0.0/8)

Source Specific Multicast (Cont.)

SSM for well-known sources characteristics:

- Allows last-hop router to send (S,G) join directly to source without the creation of a shared tree
- Allows first-hop router to respond to receiver-initiated join requests for specific sources within a group
- Supports elimination of shared-tree state in 232.0.0.0/8, simplifying address allocation

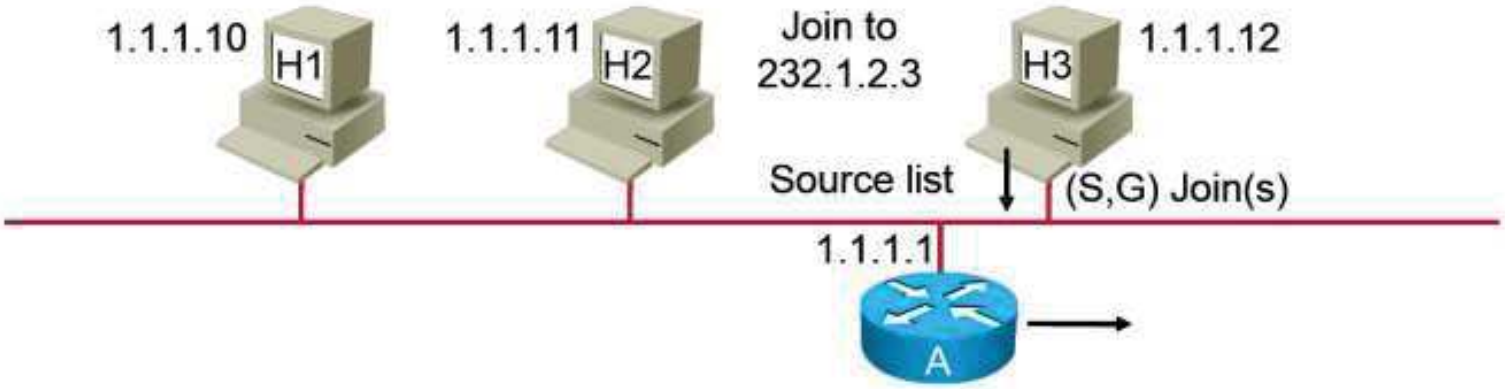
SSM Scenario



SSM with IGMPv3

SSM with IGMPv3 characteristics:

- Host sends IGMPv3 join for group and sources (IGMPv2 reports group only).
- Router adds membership.
- Router sends (S,G) join directly to sources in the source list, and is not required to send (*,G) join to RP (and must not, in 232.0.0.0/8).
- IGMPv3 is defined in the RFC 3376.
- IGMPv3 is backward compatible.

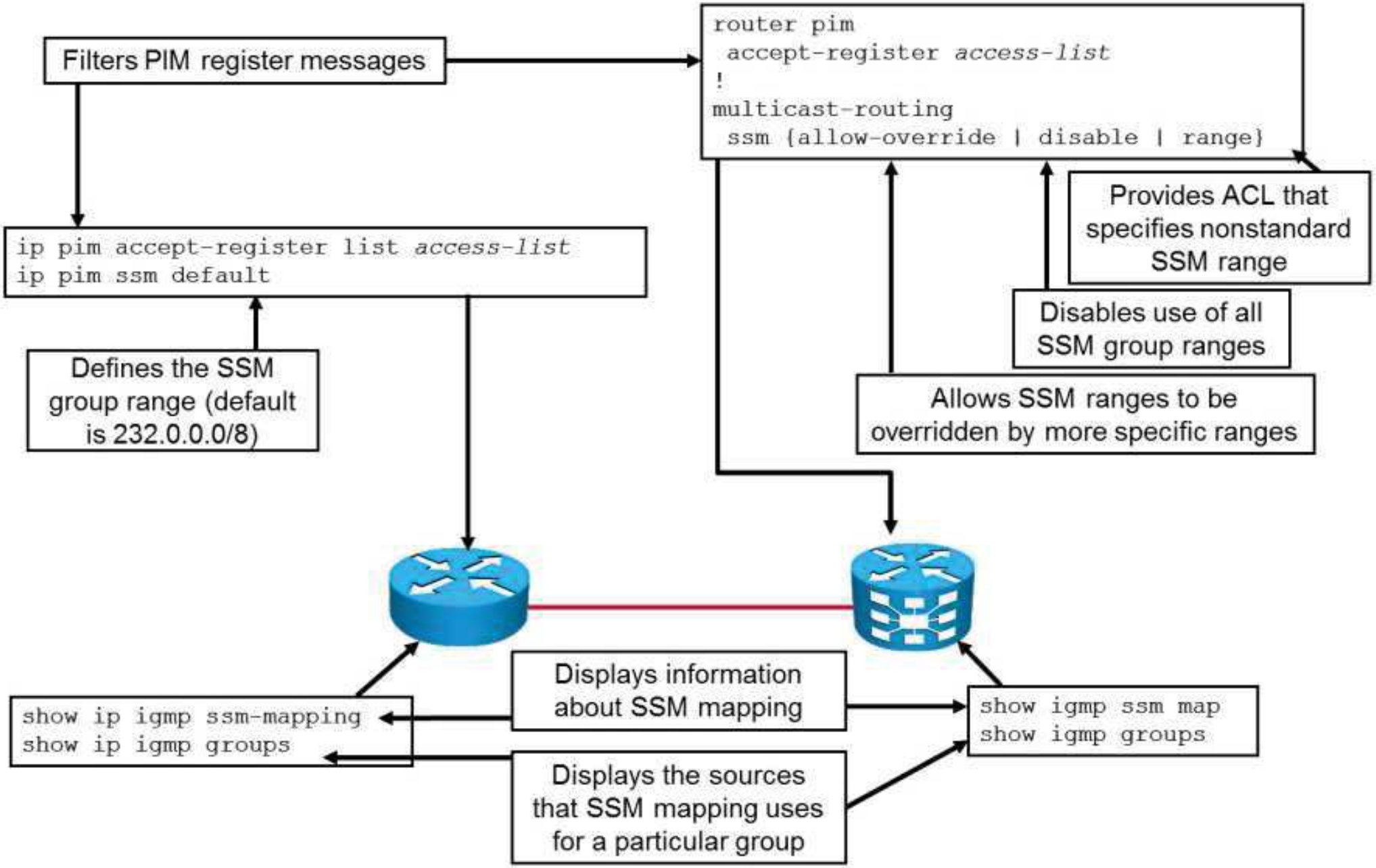


SSM Mapping

SSM mapping characteristics:

- Supports SSM transition where IGMPv3 is not available.
- Enables you to leverage SSM:
  - Video delivery to legacy set-top boxes.
  - No support for IGMPv3.
  - Applications not using IGMPv3 host stack.

Configuring SSM



Bidirectional PIM

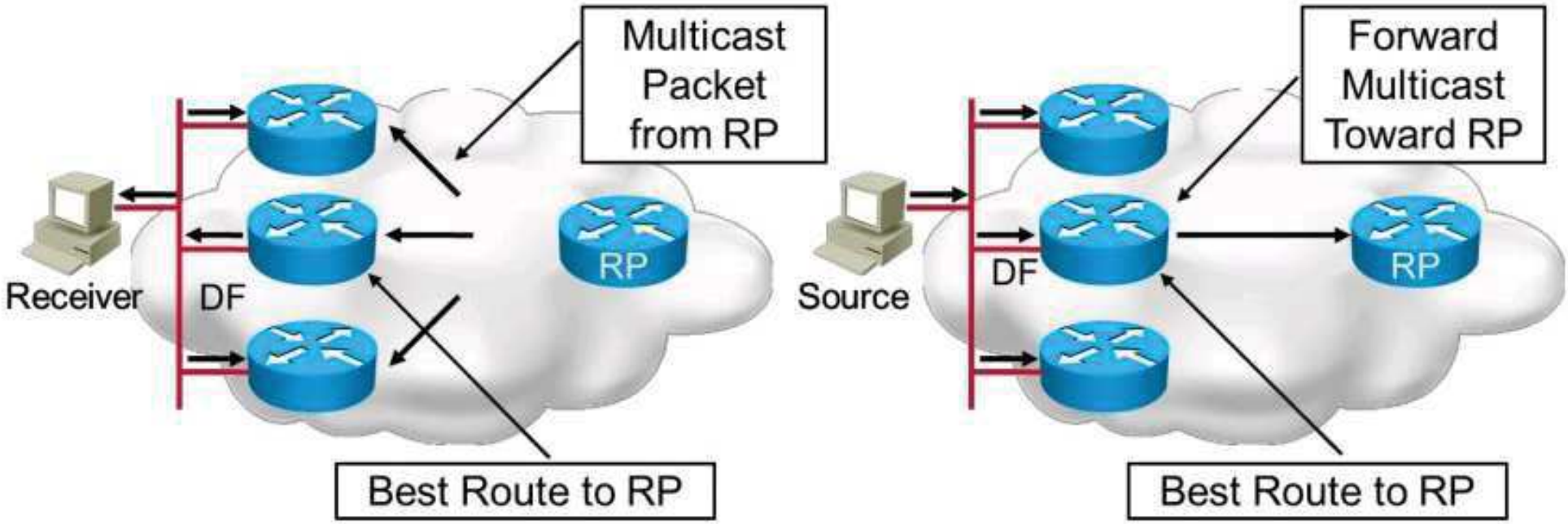
BIDIR-PIM characteristics:

- Idea:
  - Use the same tree for traffic from sources toward RP and from RP to receivers.
- Benefits:
  - Less state in routers (many sources for the same group produce one [*,G] only).
  - Traffic from sources to receivers follows the same path if on the same branch of the RP.

Bidirectional PIM (Cont.)

PIM modifications for bidirectional operation:

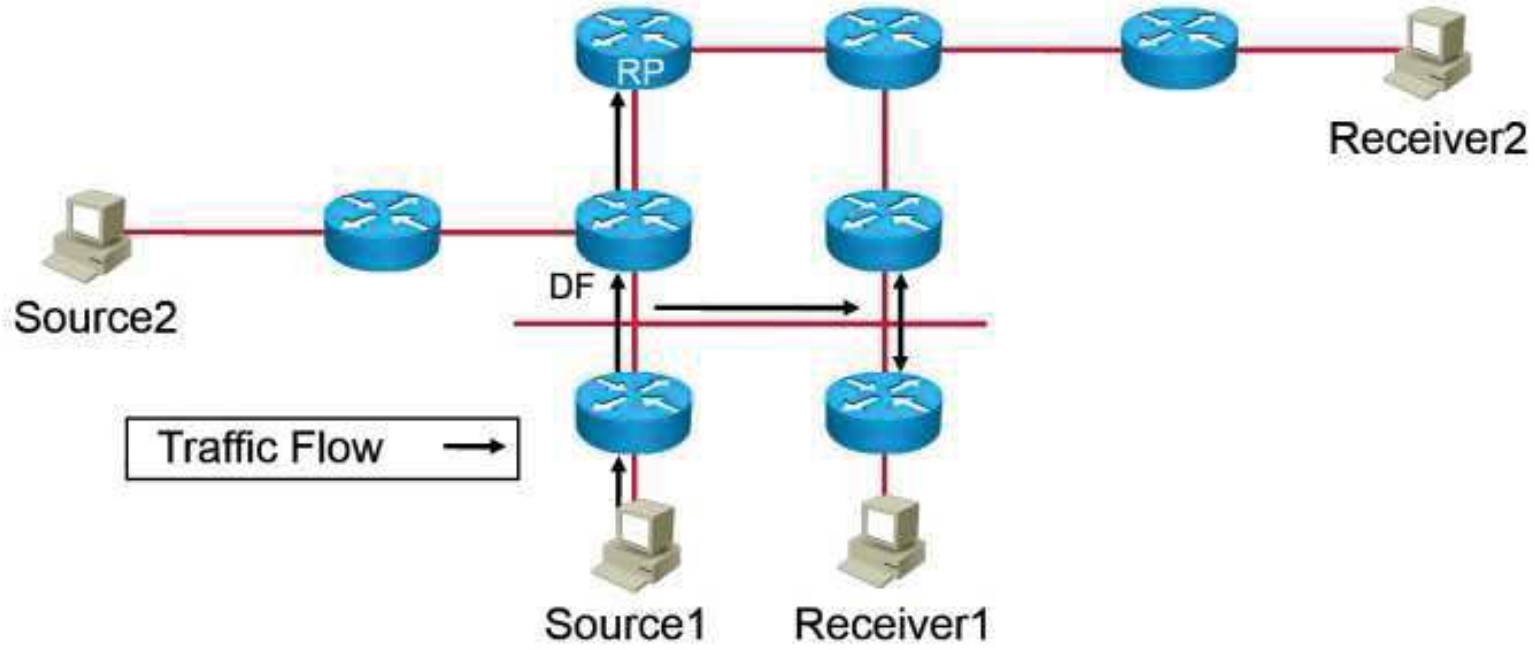
- On each link, the router with the best path to the RP is elected to be the DF.
- The DF is responsible for forwarding upstream toward the RP, as well as for downstream from the RP.
- No registration needed for local sources like in PIM-SM.



Bidirectional PIM Sources and Receivers

Bidirectional PIM sources characteristics:

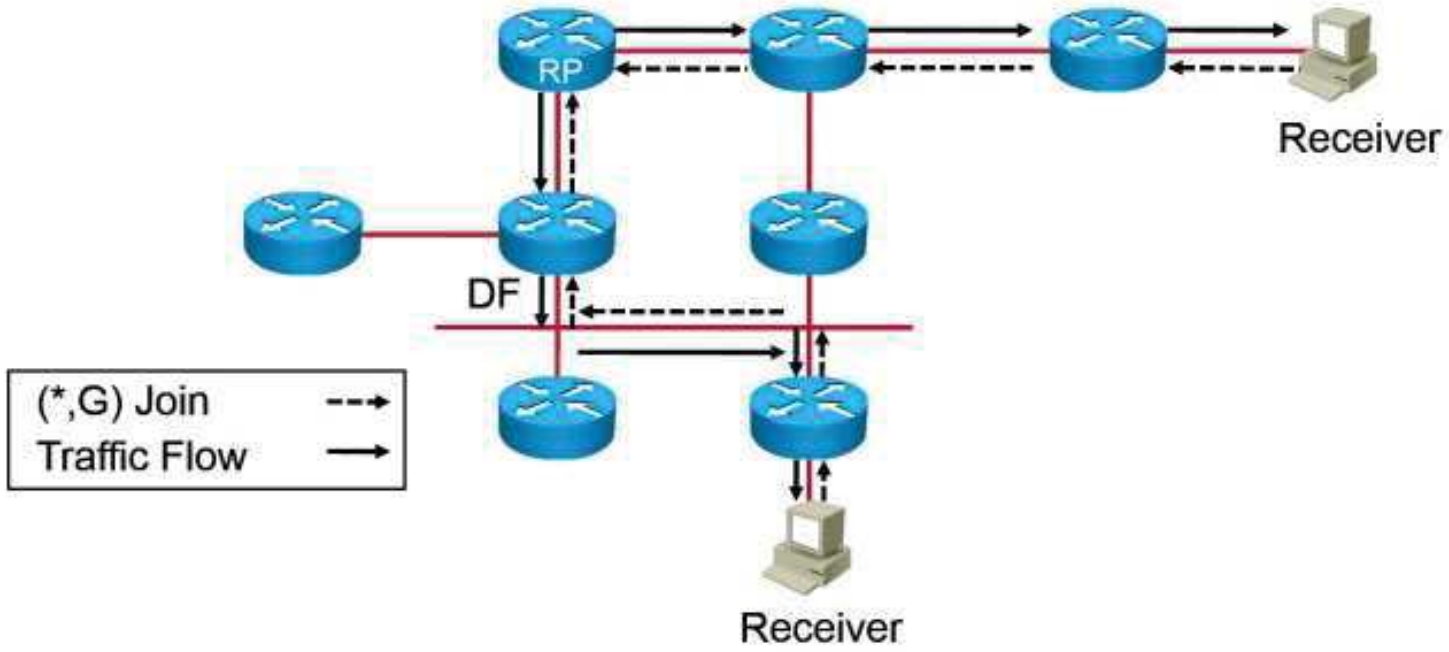
- RP identified for bidirectional groups (statically or dynamically).
- Traffic forwarded natively (hop by hop) toward RP rather than registered (using DF).



Bidirectional PIM Sources and Receivers

Bidirectional PIM receivers characteristics:

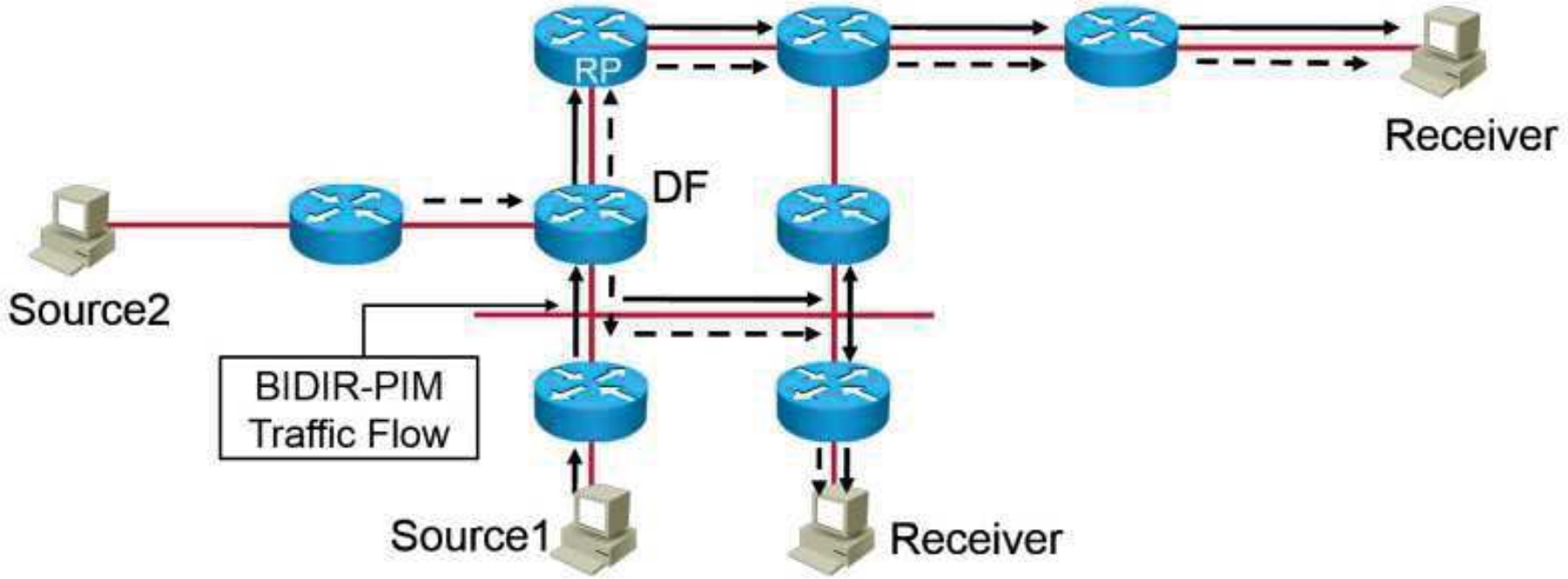
- Receivers join toward RP.
- Forwarding state is created with RPF toward RP as the IIF.



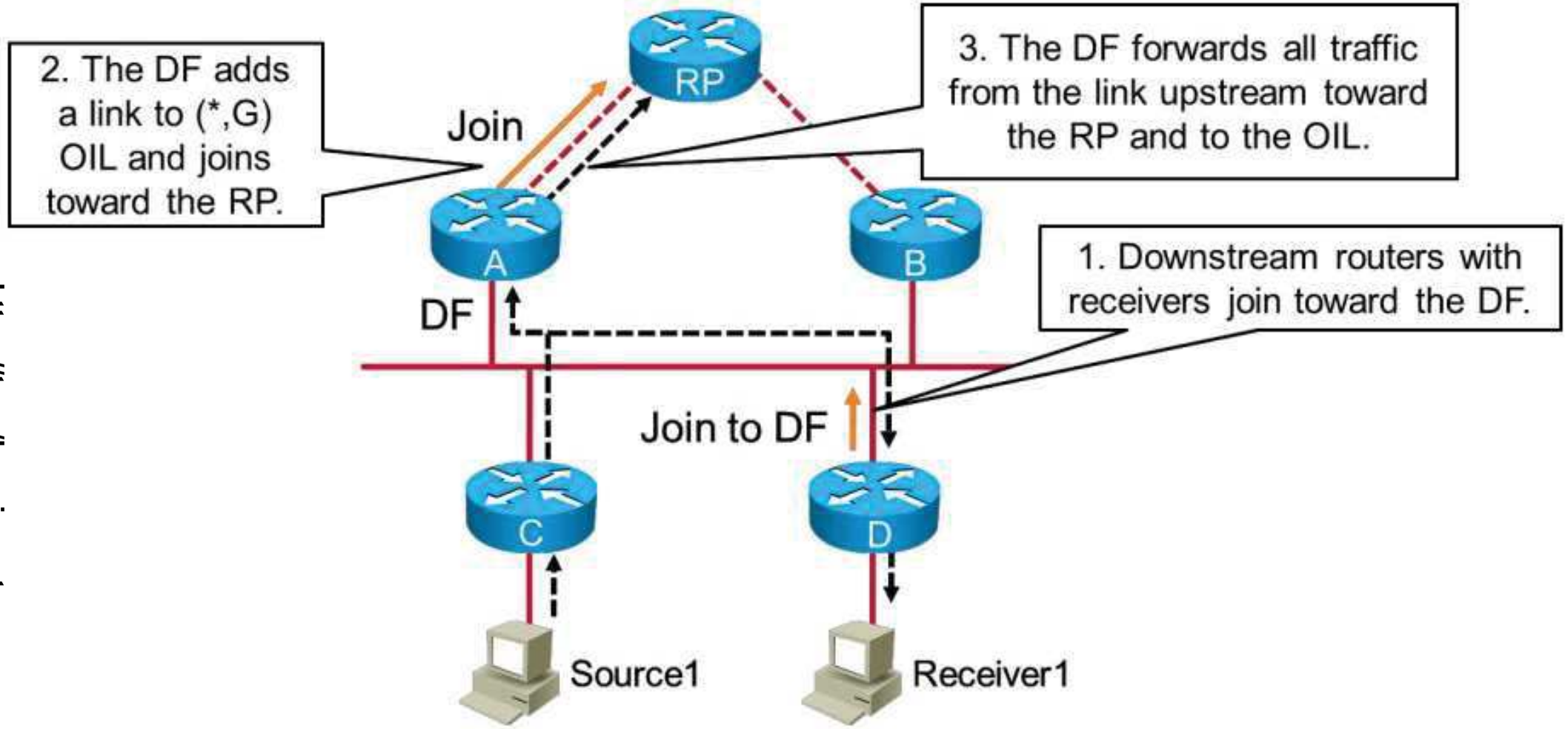
Bidirectional PIM Traffic Flow

BIDIR-PIM traffic flow characteristics:

- Branches of bidirectional distribution tree are established.
- Traffic flows natively toward RP (forwarded by DF) and can be forwarded directly on a branch toward interested receivers without first reaching RP.
- Traffic for all sources in group G is forwarded based on the same (*,G) entry.

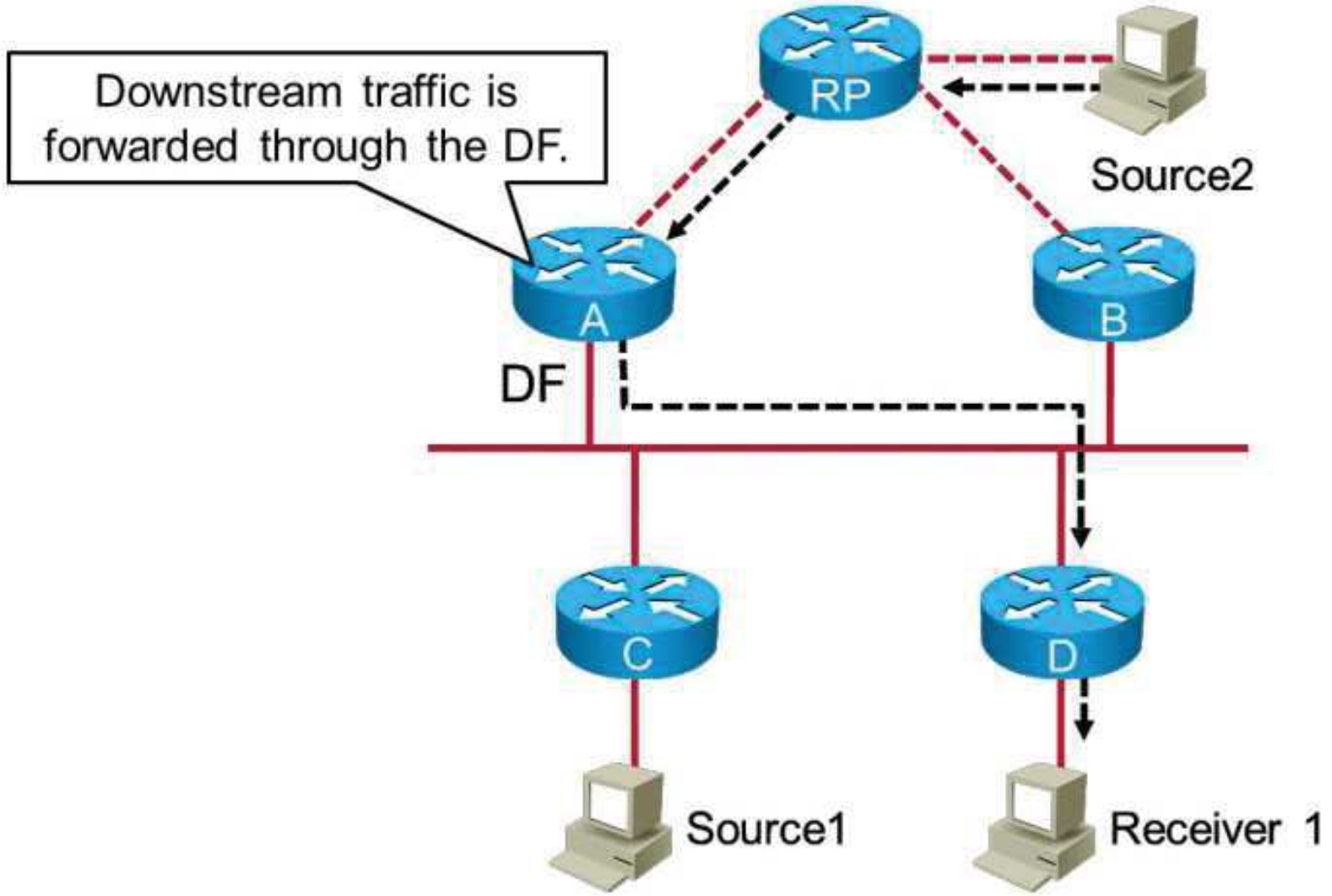


Forwarding and Tree Building Process



Forwarding and Tree Building Process (Cont.)

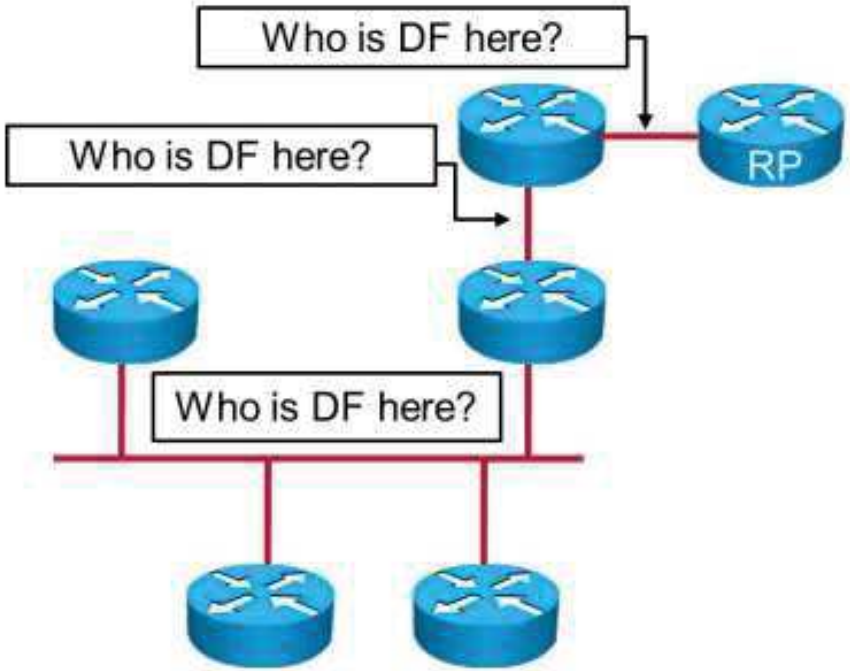
Forwarding and Tree Building—Second Source



DF Election Process

DF election process characteristics:

- Elects the router on the link with the best path to the RP.
- Ensures all routers on the link have a consistent view of the identity and metrics of the winner.
- Uses unicast routing metrics and assert comparison rules to decide between paths through different routers.

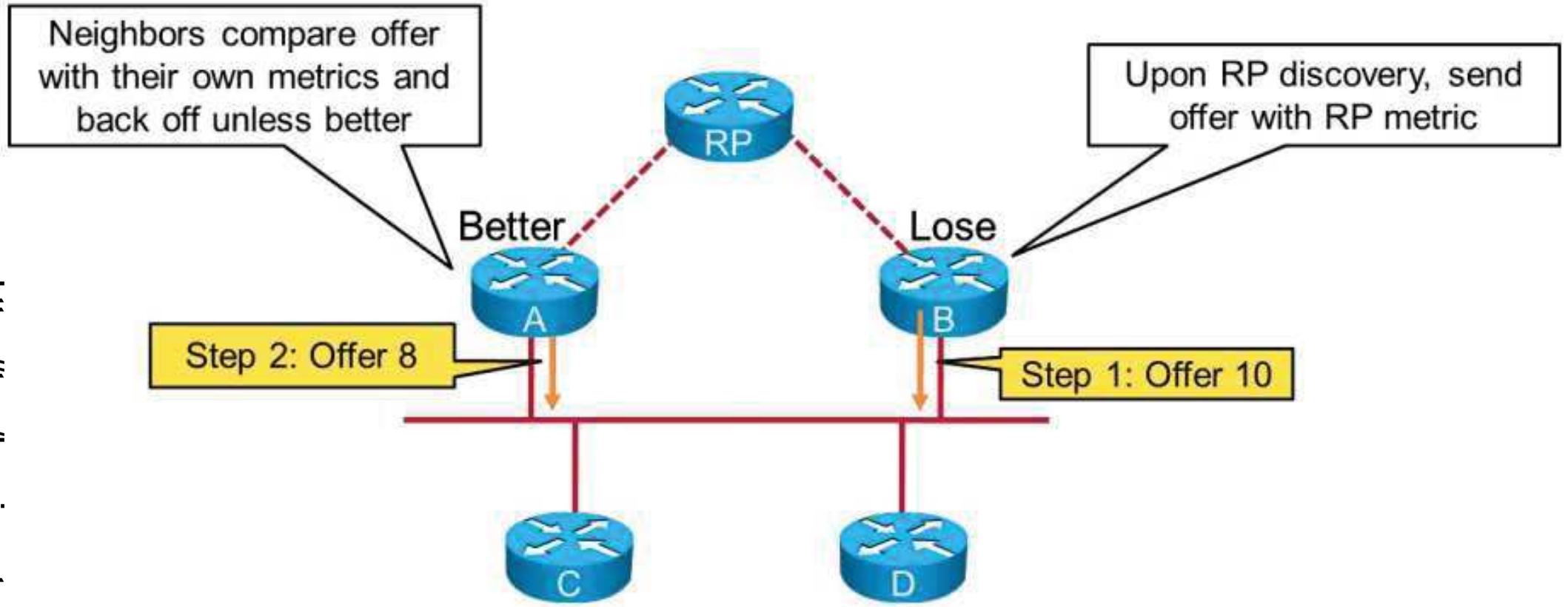


DF Election Messages

DF election messages characteristics:

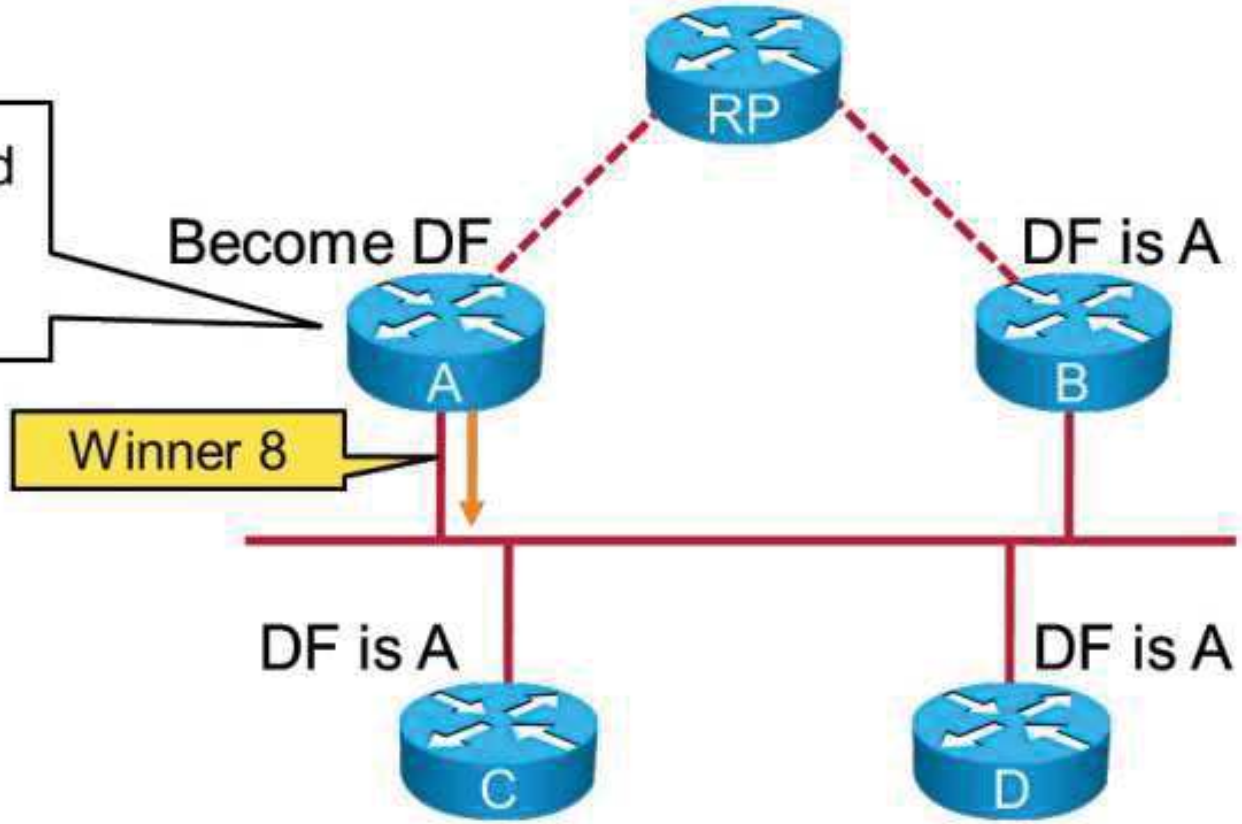
- **Offer:** Used to advertise local metrics to reach the RP.
- **Winner:** Used by a DF announcing or reasserting its status.
- **Backoff:** Used by a DF upon receipt of a better offer.
- **Pass:** Used by an acting DF to pass its responsibility to a better candidate.

Initial Election

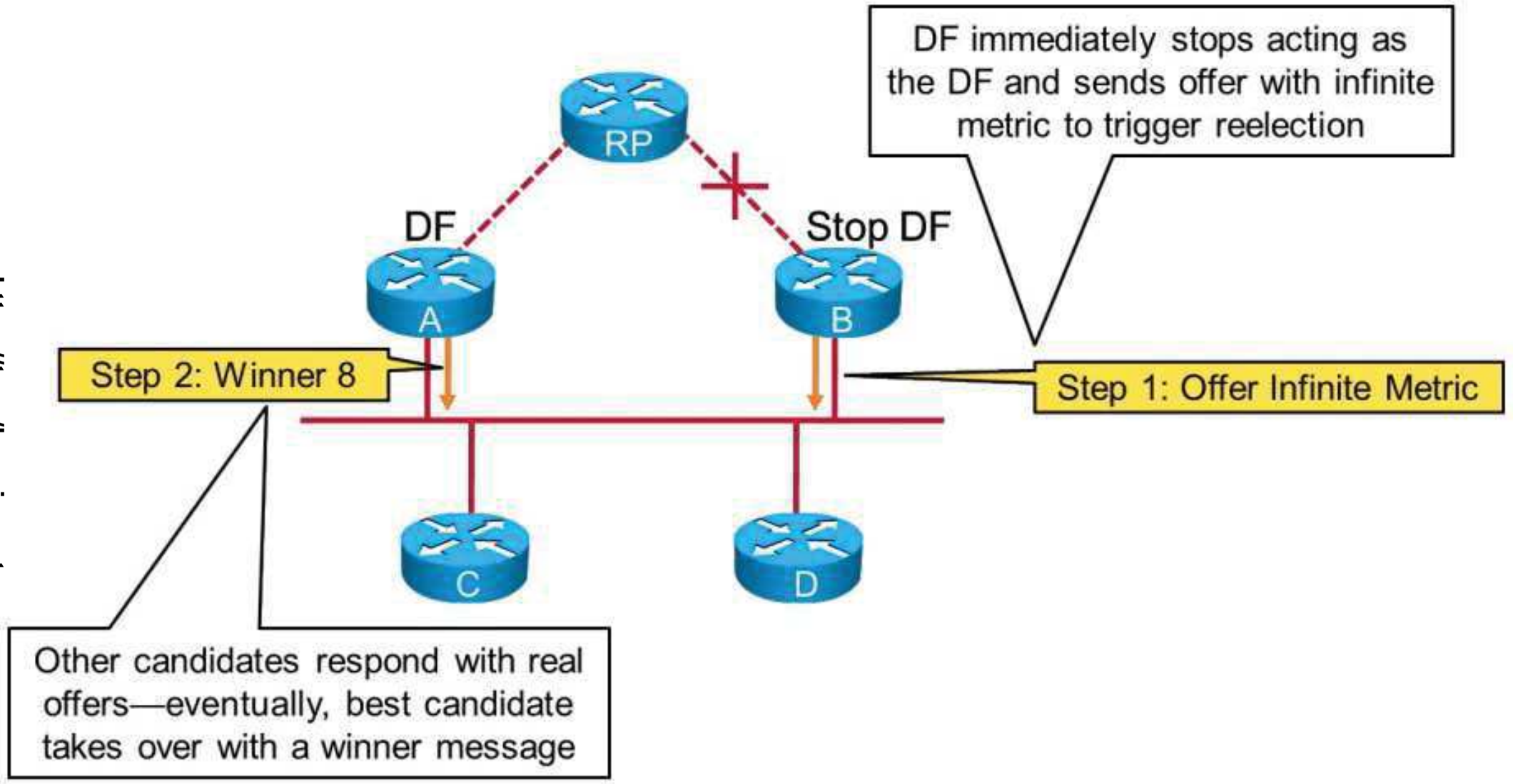


Winning the Election as DF

After repeating three uncontested offers, send a winner message and assume role as DF



DF Loses Path to the RP



DF Dies and Other Metric Changes

When the DF dies:

- Downstream routers notice a change in the RPF information provided by unicast routing.
- Downstream routers trigger a reelection.
- If no downstream routers are available, the PIM neighbor timeout triggers a reelection.

DF Dies and Other Metric Changes (Cont.)

Other metric changes characteristics:

- When the RP metric at a non-DF router changes to a value that is worse than that of the acting DF, then no action is taken.
- When the metric at the DF improves, a winner message may be sent to update information in neighboring routers.
- When the metric at the DF becomes worse, three winner messages are sent to give a better candidate the opportunity to respond with an offer.

Configuring Bidirectional PIM

Globally enable BIDIR-PIM on the router

```
ip pim bidir-enable
```

Enabled by default

Static RP with BIDIR-PIM enabled on every router

```
ip pim rp-address rp-address bidir
```

```
router pim
 rp-address rp-address bidir
```

Auto-RP: BIDIR-PIM enabled on candidate-RP routers only

```
ip pim send-rp-announce interface bidir
```

```
router pim
 auto-rp candidate-rp interface bidir
```

BSR: BIDIR-PIM enabled on candidate-RP routers only

```
ip pim rp-candidate interface bidir
```

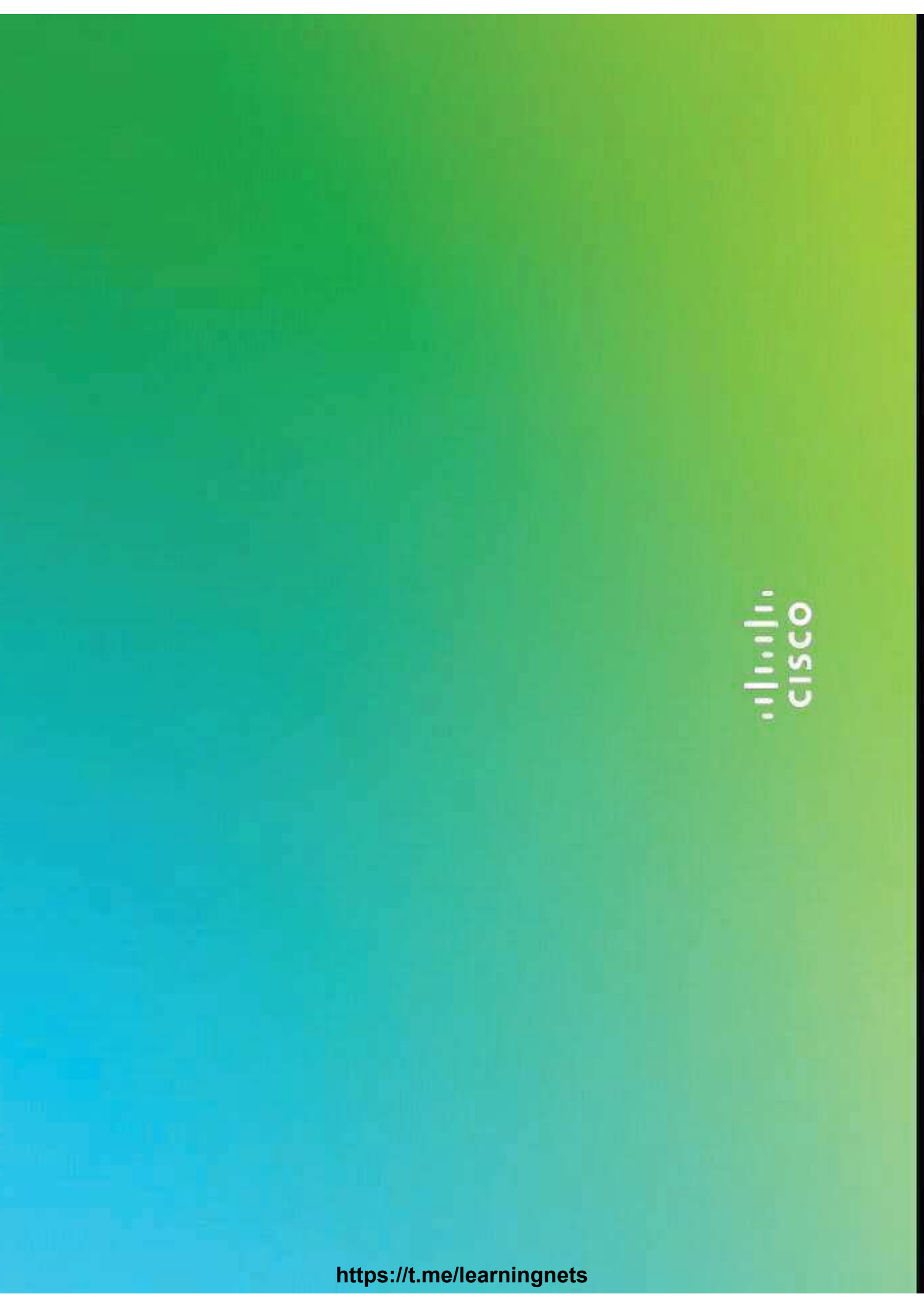
Not supported

Summary

- SSM is a simplified solution for well-known sources
- Best case for SSM is a case where there is a single source sending to a given group
- Host sends IGMPv3 join for group and sources
- SSM mapping supports SSM transition where IGMPv3 is not available
- SSM is configured differently on IOS and IOS XR platforms
- Bidirectional multicast uses the same tree for traffic from sources toward RP and from RP to receivers

Summary (Cont.)

- Source traffic is forwarded natively, while receivers still join towards the RP
- Traffic flows natively toward RP, forwarded by DF
- DF election elects the router on the link with the best path to the RP
- When the DF dies or metric changes, downstream routers notice a change in the RPF information provided by unicast routing
- Bidirectional PIM is enabled by default on IOS XR



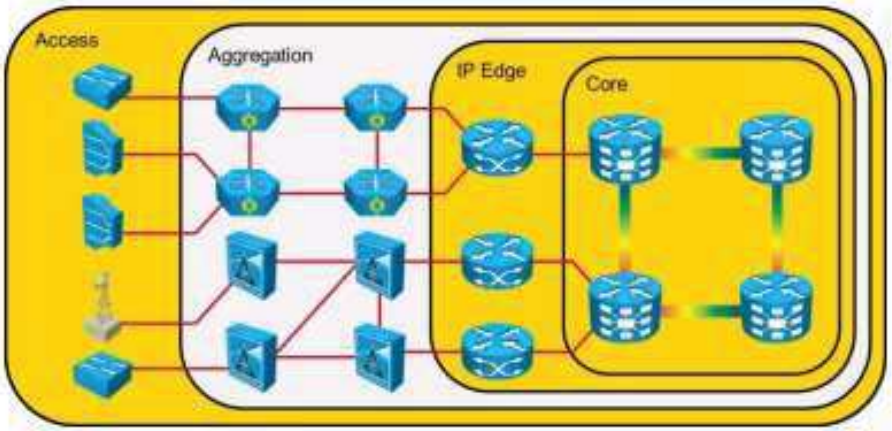
Implementing Interdomain IP Multicast

Intradomain and Interdomain Multicast Routing

Service Provider Multicast Requirements

Multicast in the Service Provider characteristics:

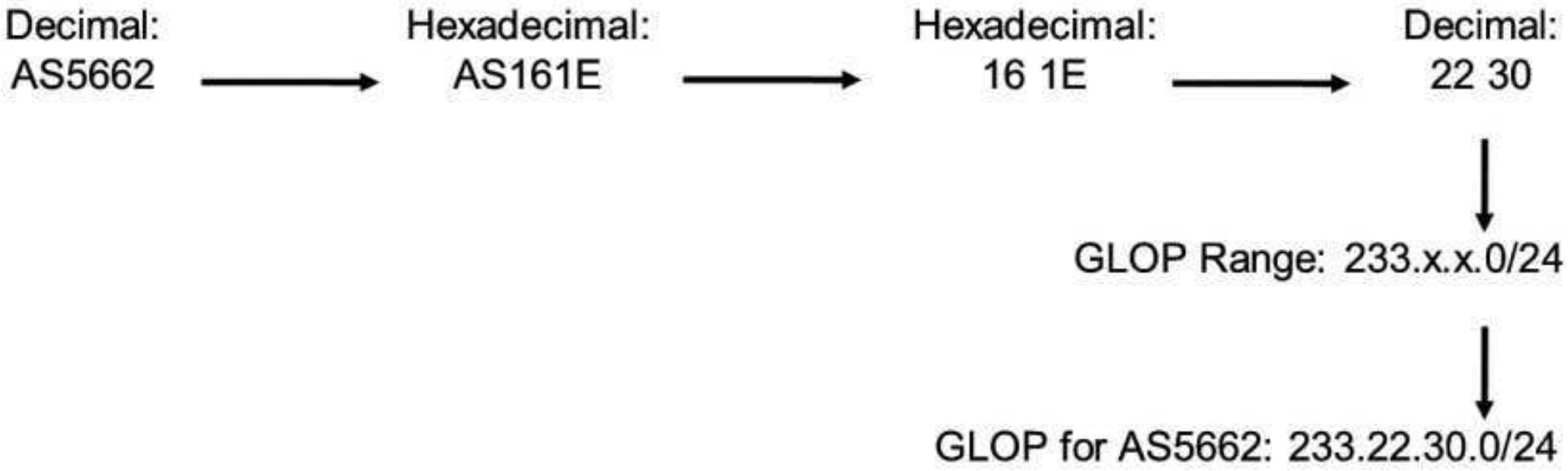
- PIM-SM is run in a service provider network.
- RPF check is done from the MP-BGP tables.
- Service providers want flexibility for RP:
  - Will not share RP with competitors—third-party resource dependency.
  - Want to be flexible in RP placement—not necessary to place RP at the interconnect.
- Problem: How to learn about sources in other domains?



GLOP—Static Allocation of 233/8

GLOP characteristics:

- Temporary allocation of 233.0.0.0/8:
  - RFC 2770
- Statically assigned by mapping AS number into middle octets.
- Provides each AS with /24 addresses to use while waiting for another solution.



SSM Role in Interdomain IP Multicast

SSM characteristics:

- Uses source trees only—eliminates need for RP and shared trees.
- Assumes one-to-many model:
  - Most Internet multicast fits this model.
  - IPTV also fits this model.
- Hosts responsible for source discovery:
  - Typically via some out-of-band mechanism (web page, content server, etc.).
  - Eliminates need for interdomain multicast.

SSM Role in Interdomain IP Multicast (Cont.)

- Hosts join a specific source within a group:
  - Content identified by specific (S,G) instead of (*,G).
  - Hosts responsible for learning (S,G) information.
- Last-hop router sends (S,G) join toward source:
  - Shared tree is never joined or used.
  - Eliminates possibility of content jammers.
  - Only specified (S,G) flow is delivered to host.
- Simplifies address allocation:
  - Dissimilar content sources can use same group without fear of interfering with one another.

MSDP Role in Interdomain IP Multicast

MSDP characteristics:

- Service providers wanting an explicit join protocol for efficiency: PIM-SM with RPs.
- Service providers using existing (unicast) operation model: MP-BGP.
- SSM not yet widely available, and dedicated for 232.0.0.0/8 range only.
- MSDP for interdomain requirement.
- Receivers in a domain join to their local RP only—internal sources are only known inside a domain.
- PIM-SM run on interdomain links as well—scalability.
- MSDP announces active sources to the neighboring domains—messages propagated throughout the Internet.
- SPTs are built across domains—interconnected RPs.

Multicast Service Provider Requirements

Service provider requirements:

- Wants an explicit join protocol for efficiency:
  - PIM-SM
- Uses existing (unicast) operation model:
  - MP-BGP
- Will not share RP with competitors:
  - MSDP
- Wants flexibility regarding RP placement:
  - MSDP

MSDP Protocol

MSDP protocol characteristics:

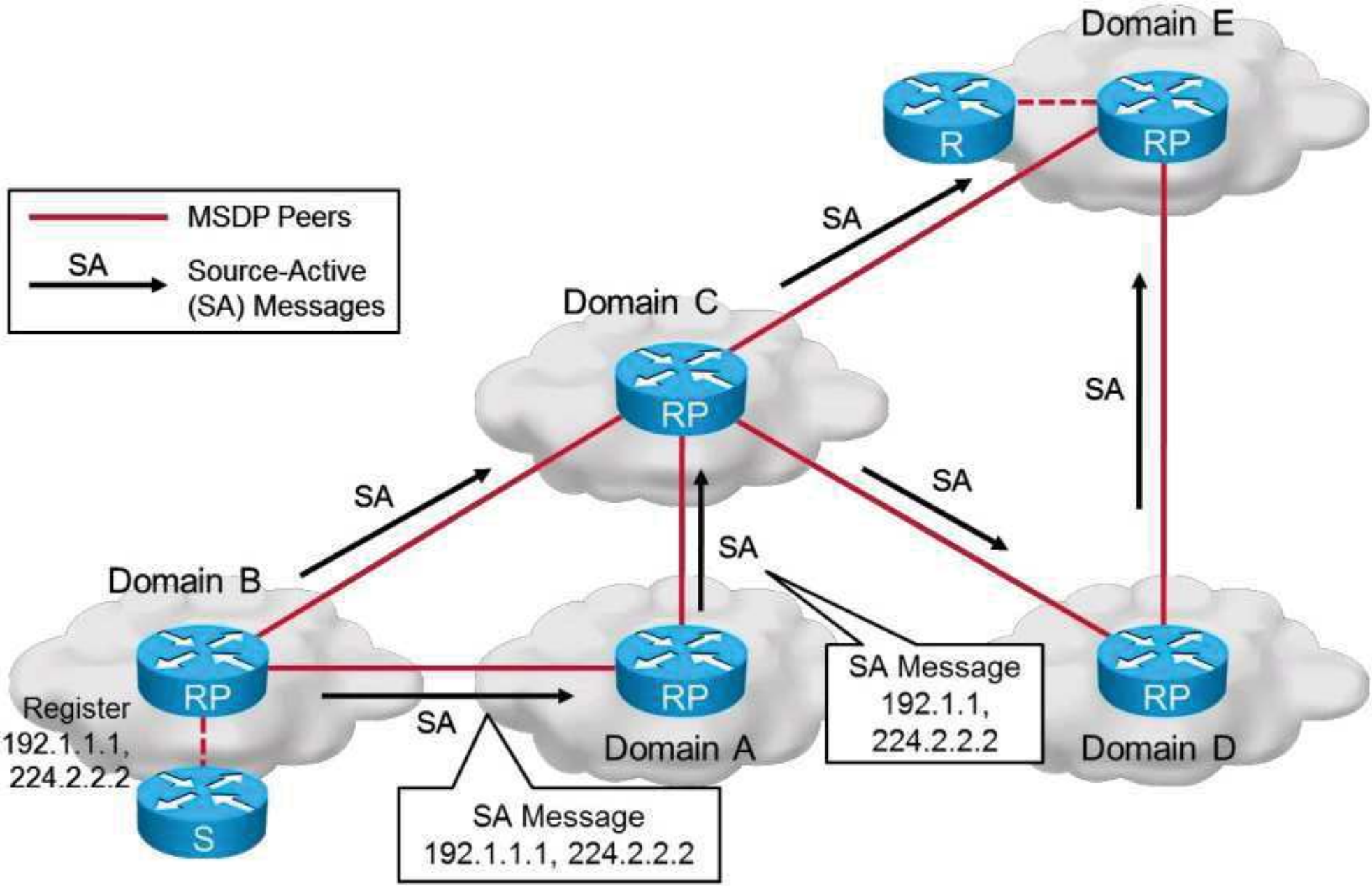
- Use interdomain source trees.
- Reduces problem of locating active sources.
- Allows RP or last-hop receiver to join interdomain source tree.
- RPs know about all sources in a domain:
  - Sources cause a PIM register to the RP.
  - Can tell RPs in other domains of its sources (MSDP SA messages).
- RPs know about receivers in a domain:
  - Receivers cause a (*,G) join to be sent to the RP.
  - RP can join the source tree in the peer domain:
    - Via normal PIM (S,G) joins.
    - Only necessary if there are receivers for the group.

MSDP Concepts

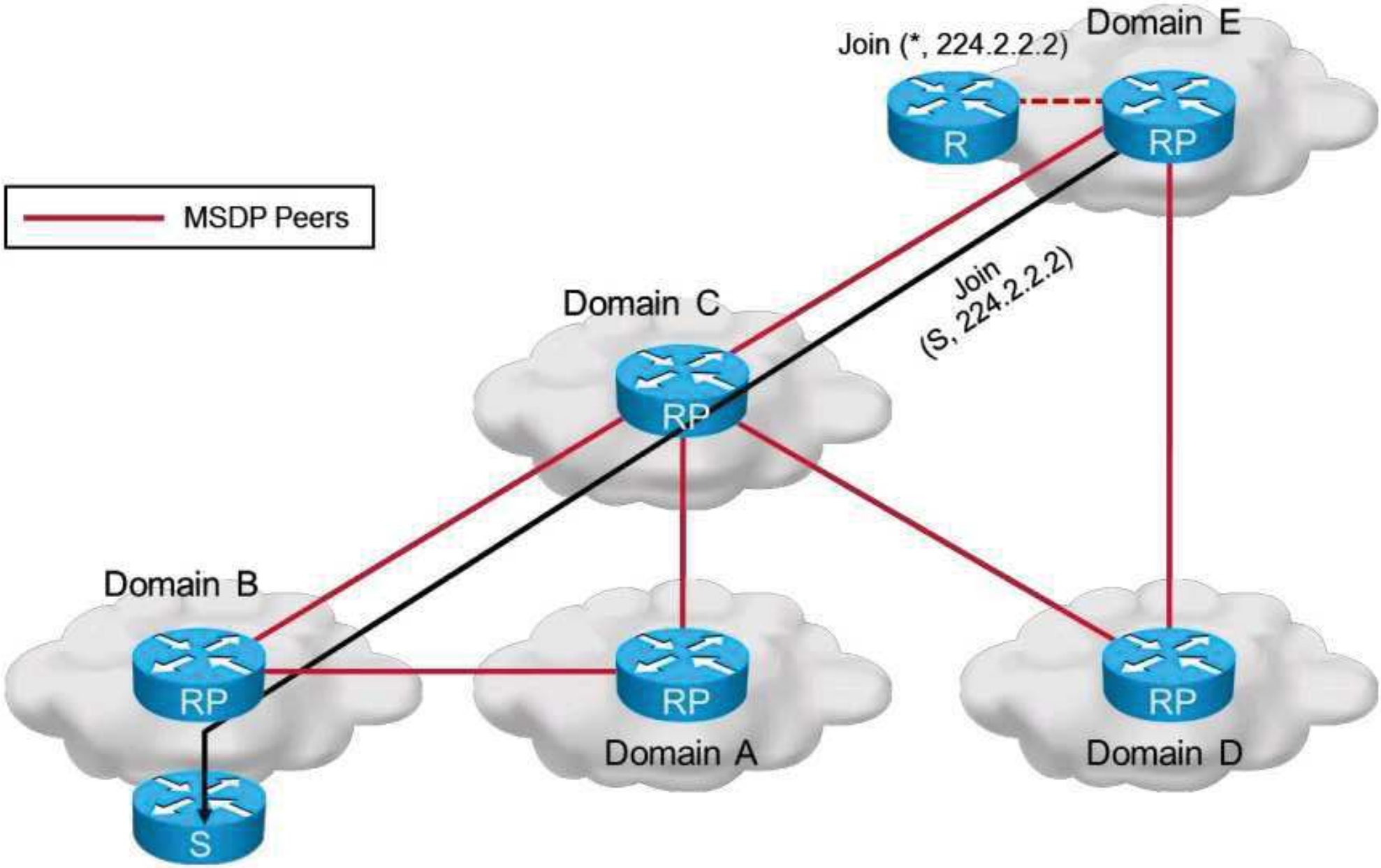
MSDP concepts characteristics:

- MSDP peers talk via TCP connections:
  - UDP encapsulation option.
- MSDP peers should be RPs also.
- SA messages:
  - Using peer RPF forwarding, forwarded to prevent loops:
    - RPF check on AS path back to the peer RP.
    - If successful, flood SA message to other peers.
    - Stub sites accept all SA messages, since they have only one exit.
  - MSDP speaker may cache SA messages:
    - Other MSDP speakers can query for active sources.
    - Reduces join latency:
      - No need to wait for periodic SA messages.

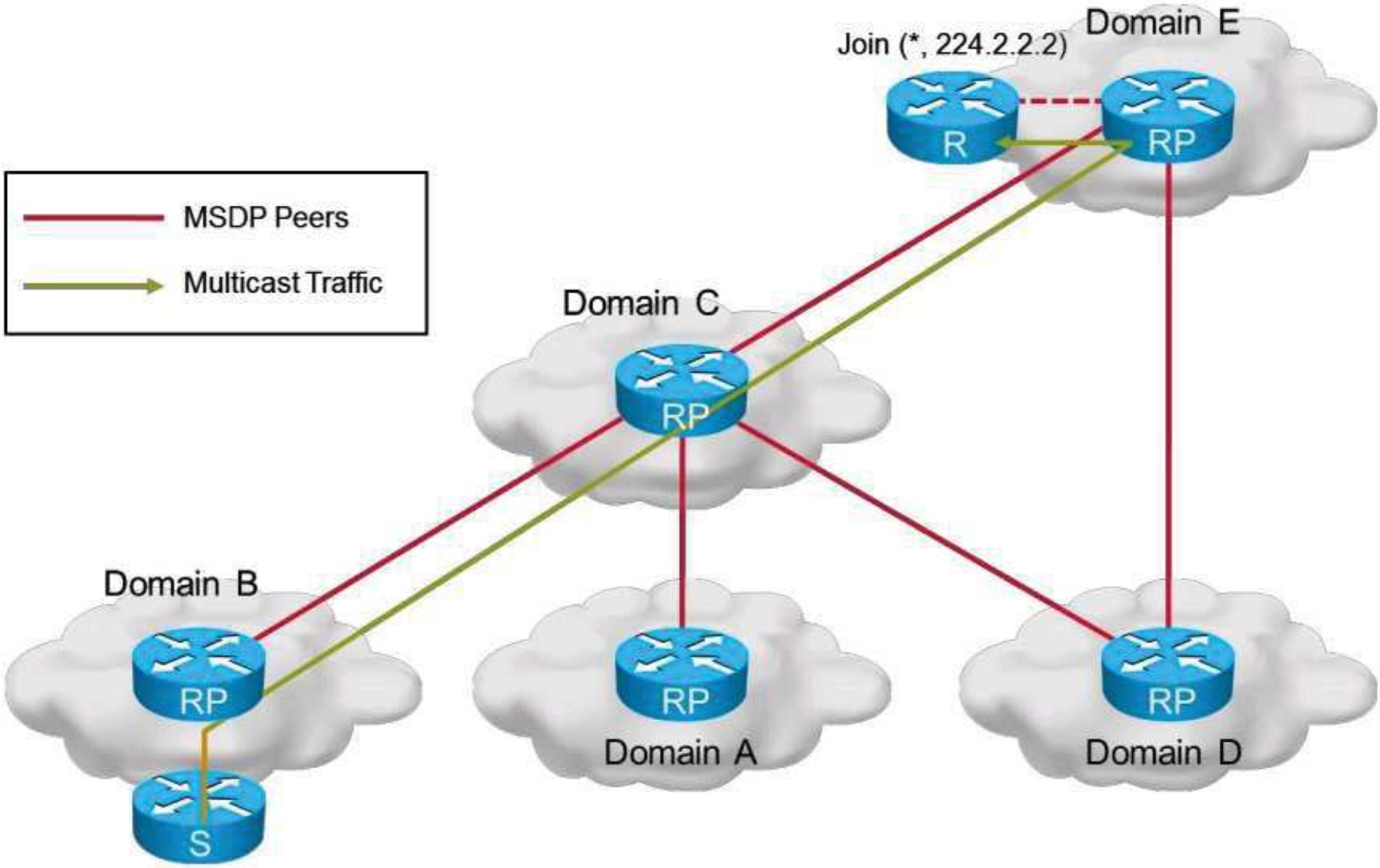
MSDP Concepts (Cont.)



MSDP Concepts (Cont.)



MSDP Concepts (Cont.)



MSDP Neighbor Relationship

MSDP neighbor relationship characteristics:

- Peers connect using TCP port 639.
  - Lower-address peer initiates connection.
  - Higher-address peer waits in listen state.
- Peers send keepalives every 60 seconds (fixed).
- Peer connection is reset after 75 seconds if no MSDP packets or keepalives are received.
- MSDP peers should exchange routing information using BGP:
  - BGP is used to perform an RPF check of arriving SA messages. May use MRIB, URIB, or both.
- Exceptions:
  - When peering with only a single MSDP peer.
  - When using an MSDP mesh group.

MSDP Messages

One or more messages (in TLV format):

- Keepalives
- SA messages
- SA request messages
- SA response messages

SA messages characteristics:

- Used to advertise active sources in a domain
- Can also carry initial multicast packet from source
- SA message contents:
  - IP address of originating RP
  - Number of (S,G) pairs being advertised
  - List of active (S,G) pairs in the domain
  - Encapsulated multicast packet

MSDP SA Message Processing

MSDP SA message processing:

- Check multicast routing table for joined members:
  - Check whether a (*,G) entry with non-null OIL exists.
- If so, create (S,G) state (if it does not already exist).
- Send join toward source.
- Flood SA messages to all other MSDP peers except:
  - The RPF peer.
  - Any MSDP peers that are in the same MSDP mesh group.
- Note: SA messages are saved if SA caching has been enabled.

MSDP SA Message Origination

Originating MSDP SA messages characteristics:

- The MSDP SA messages are triggered when any new source in the local domain goes active.
- Initial multicast packet is encapsulated in an SA message:
  - This is an attempt at solving the issue of bursty sources.

Local sources role in the MSDP SA messages:

- A source is local if:
  - The router received a register for (S,G).
  - The source is directly connected to the RP.
- SA messages are originated only for local sources:
  - Denoted by the A flag on an (S,G) entry.
- Other conditions may suppress SA messages from being originated for local sources.

MSDP SA Message Origination (Cont.)

Encapsulating initial multicast packets:

- Can bypass TTL thresholds:
  - Original TTL is inside of data portion of SA message.
  - SA messages sent via unicast with a TTL of 255.

Requires special command to control:

- **ttl-threshold**
- Encapsulated multicast packets with a TTL lower than the TTL value for the specific MSDP peer are not forwarded or originated.

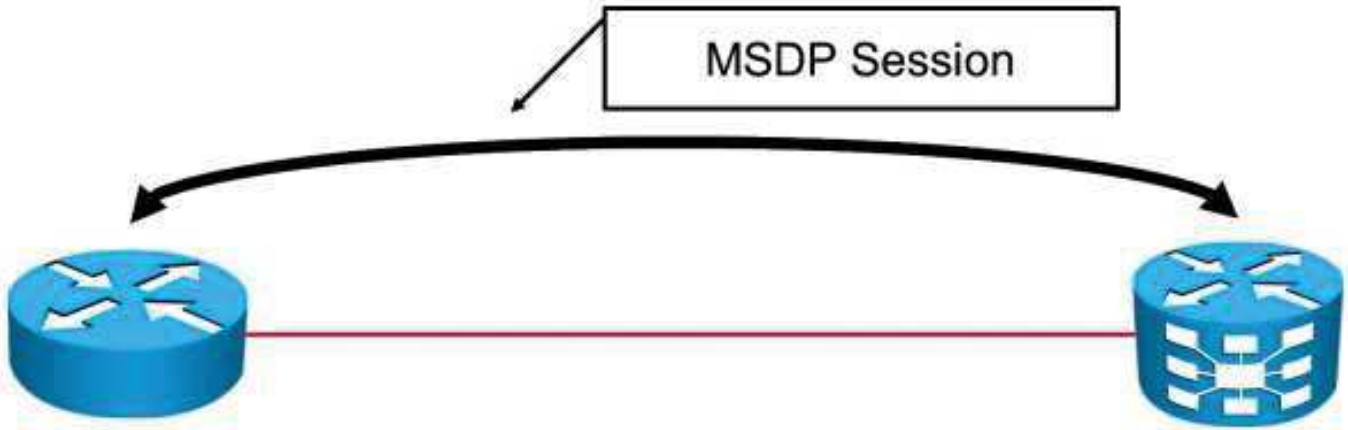
MSDP MD5 Password Authentication

MSDP MD5 password authentication characteristics:

- Supports MD5 hashing algorithm.
- Protects MSDP against the threat of spoofed TCP segments.
- MD5 authentication must be configured with the same password on both MSDP peers.

MSDP MD5 password authentication benefits:

- Prevents spoofed TCP segments.
- Improves reliability and security.



Configuring MSDP


```
ip access-list extended MSDP_ACL
 deny ip any host 224.0.1.39
 deny ip any host 224.0.1.40
!
ip msdp peer 10.2.1.1 connect-source Loopback0
ip msdp originator-id Loopback0
ip msdp ttl-threshold 10.2.1.1 64
ip msdp sa-filter in 10.2.1.1 list MSDP_ACL
ip msdp sa-filter out 10.2.1.1 list MSDP_ACL
ip msdp password peer 10.2.1.1 C!sc()
```

Configures an MSDP peer

Enables MSDP authentication

```
ipv4 access-list MSDP_ACL
 deny ip any host 224.0.1.39
 deny ip any host 224.0.1.40
!
router msdp
 originator-id Loopback0
 peer 10.1.1.1
  connect-source Loopback0
  ttl-threshold 64
  password C!sc()
 sa-filter in list MSDP_ACL
 sa-filter out MSDP_ACL
```

Sets RP ID in SA messages

Sets minimum TTL for encapsulated multicast packets

Configures filtering of SA messages

Verifying MSDP

```
RP/0/RSP0/CPU0:PE5# show msdp summary
Out of Resource Handling Enabled
Maximum External SA s Global : 20000
Current External Active SAs : 0
MSDP Peer Status Summary
Peer Address AS  State   Uptime/     Reset  Peer   Active  Cfg.Max     TLV
                         Downtime    Count  Name   SA Cnt  Ext.SAs     recv/sent
10.6.1.1     0   Up      00:23:32    1      ?      0       0           30/48
```

- Displays summary of MSDP peers.

Verifying MSDP (Cont.)

```
RP/0/RSP0/CPU0:PE5# show msdp sa-cache
MSDP Flags:
E - set MRIB E flag , L - domain local source is active,
EA - externally active source, PI - PIM is interested in the group,
DE - SAs have been denied. Timers age/expiration, Cache Entry:
(192.168.156.60, 226.1.1.1), RP 10.6.1.1, MBGP/AS 0, 00:00:04/00:02:25
  Learned from peer 10.6.1.1, RPF peer 10.6.1.1
  SAs recvd 1, Encapsulated data received: 0
    grp flags: none, src flags: EA
(10.6.1.1, 226.1.1.1), RP 10.6.1.1, MBGP/AS 0, 00:00:04/00:02:25
  Learned from peer 10.6.1.1, RPF peer 10.6.1.1
  SAs recvd 1, Encapsulated data received: 0
    grp flags: none, src flags: EA
```

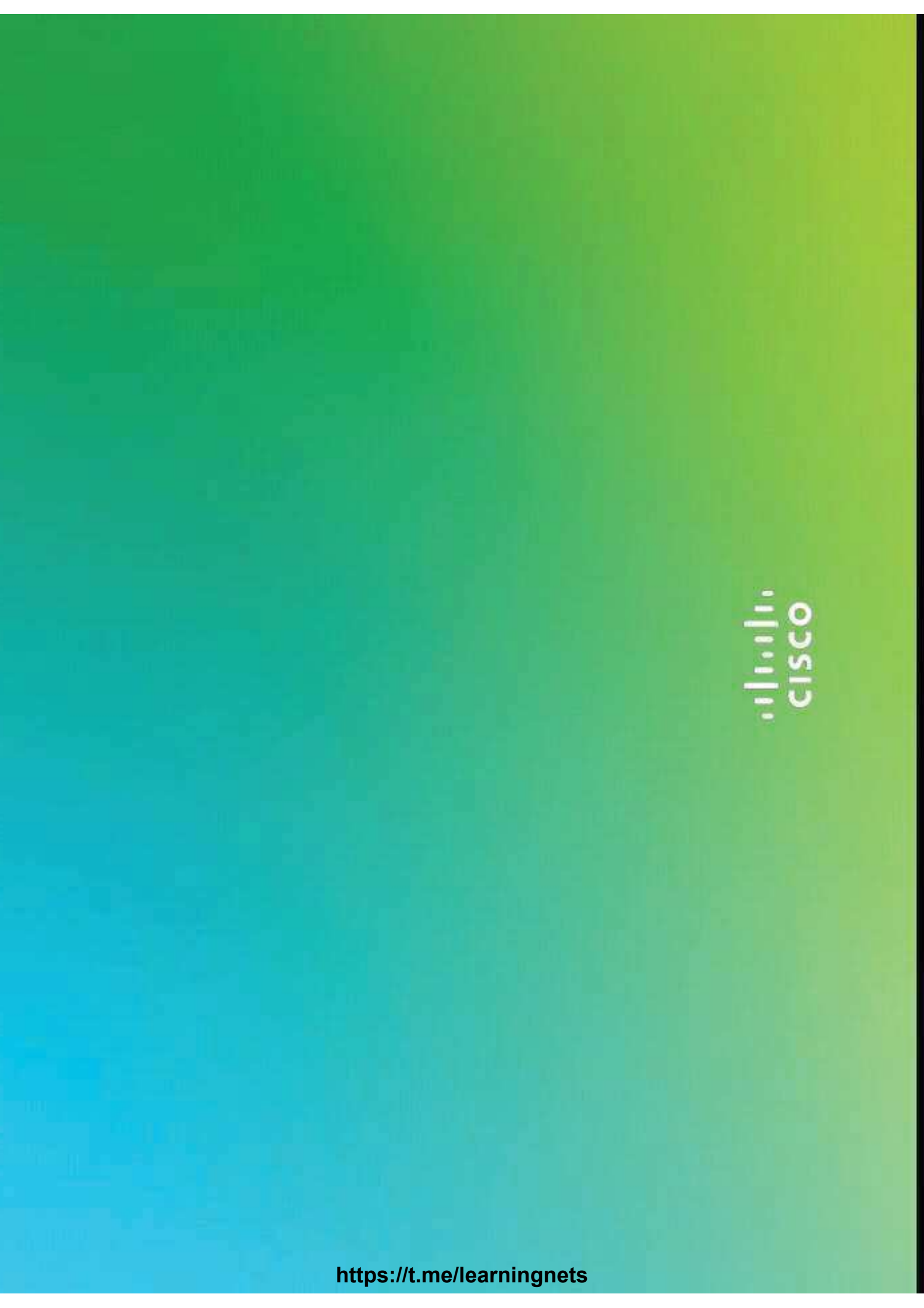
- Displays states learned from MSDP peers.

Summary

- Interdomain multicast traffic has different requirements as intradomain forwarding.
- GLOP is based on temporary allocation of 233.0.0.0/8.
- SSM in interdomain scenario eliminates the need for RPs and shared trees.
- MSDP announces active sources to the neighboring domains.
- MSDP uses interdomain source trees.

Summary (Cont.)

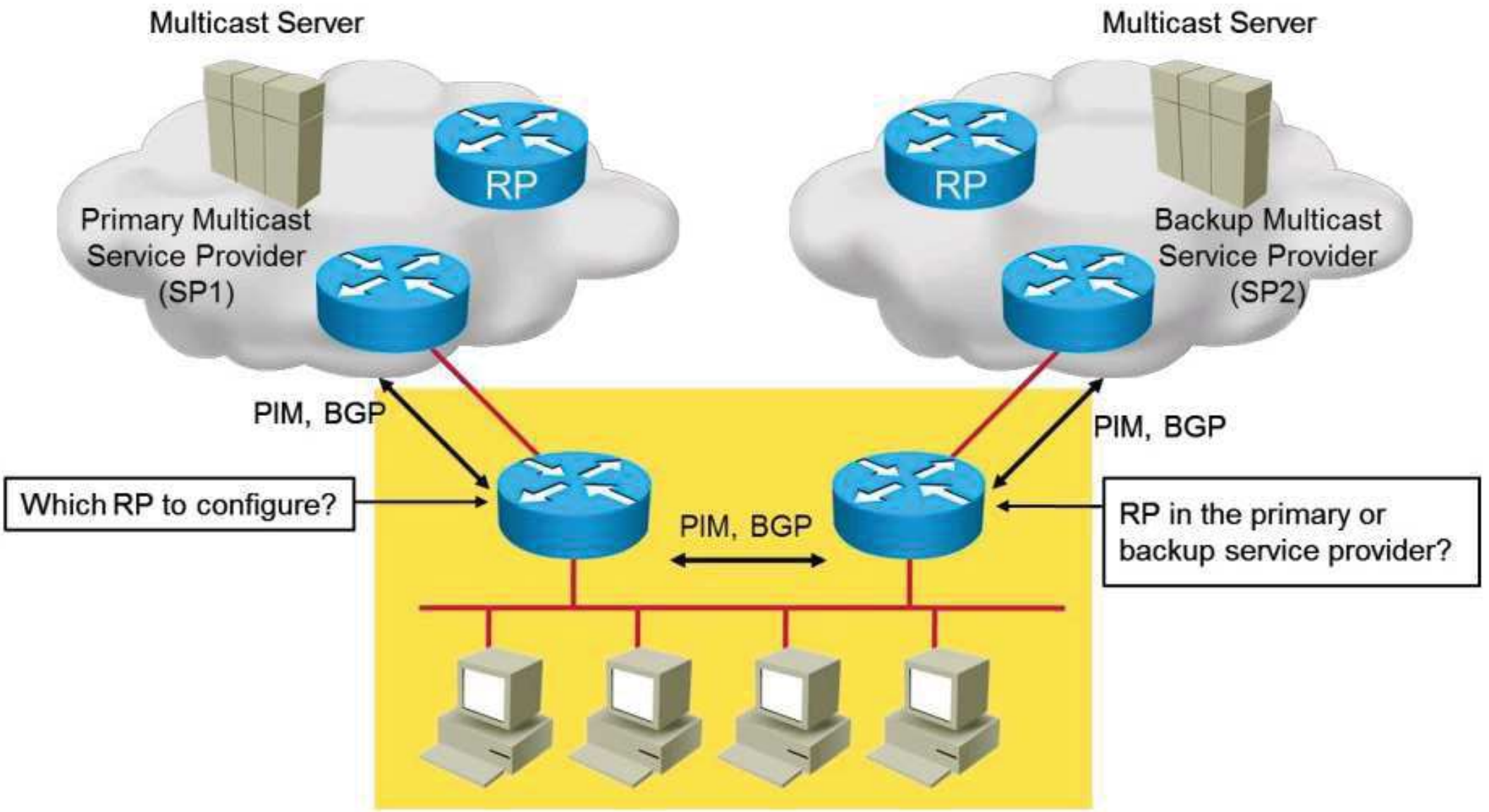
- MSDP uses TCP to exchange information.
- SA messages are used to advertise active sources in a domain.
- They are triggered when any new source in the local domain goes active.
- MSDP messages can be authenticated using MD5 authentication.
- MSDP is configured separately from other multicast configuration.



Identifying Rendezvous Point Distribution Solutions

Intradomain and Interdomain Multicast Routing

Static RP Disadvantages



Dynamic RP Discovery Mechanisms

Dynamic RP discovery mechanisms characteristics:

- Manual configuration:
  - RP failover generally not possible (with some exceptions)
  - All routers must be enabled
- Dynamic distribution:
  - Provided through Auto-RP or BSR mechanism
  - Better scalability
  - Complex environments
- Consistent view needed for all routers:
  - Same groups mapped to the same RP

RP Placement

RP placement characteristics:

- RP can be almost anywhere:
  - Where sources and receivers meet.
- Cisco IOS, IOS XE, and IOS XR Software: default SPT threshold is 0:
  - Immediate switch to SPT.
  - Traffic itself does not flow via RP.
- The issue occurs only when the SPT threshold is set to infinity:
  - Placement closer to the source is better.
  - The RP can become a congestion point.

Auto-RP

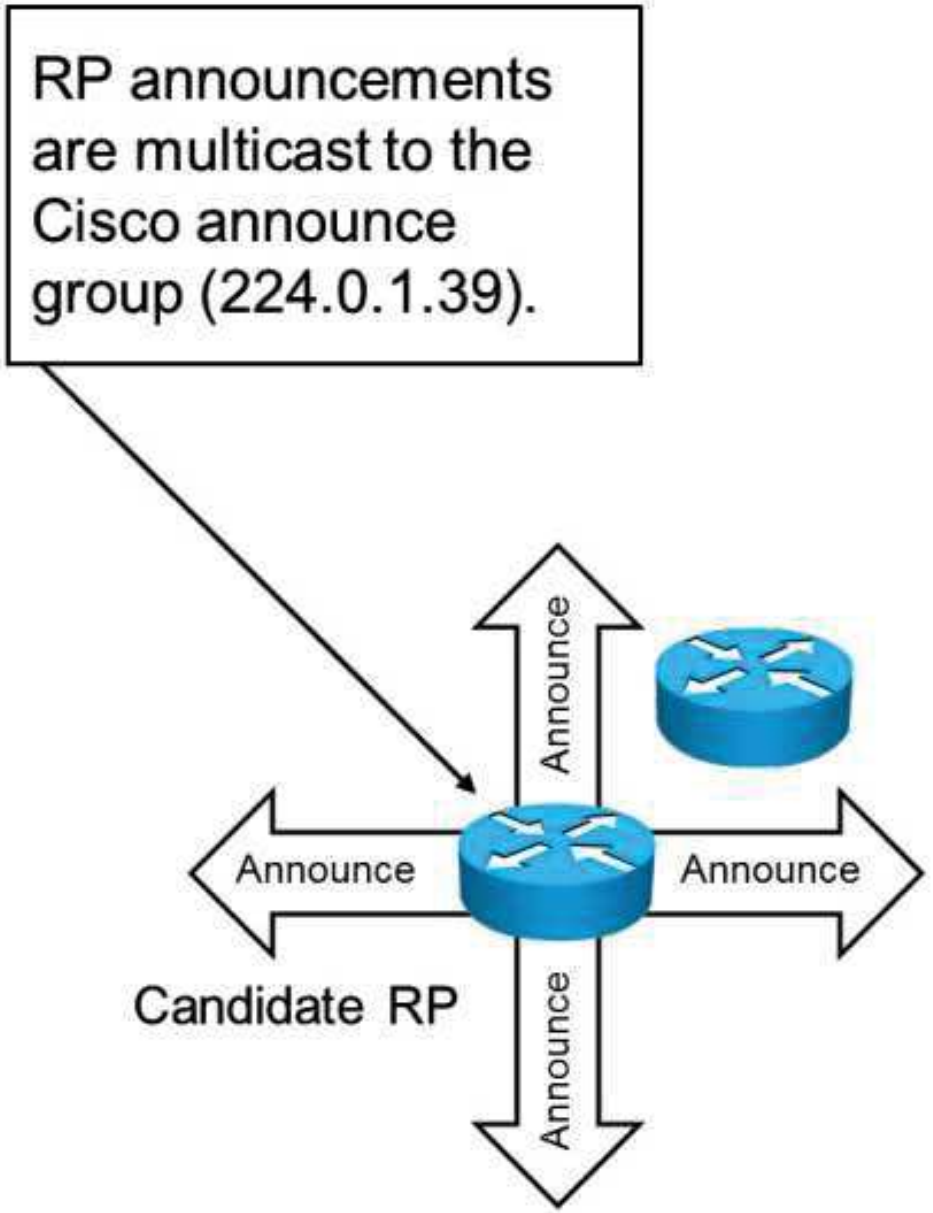
Auto-RP characteristics:

- All routers automatically learn RP address.
- Configuration is necessary only on candidate RP and mapping agent.
- Multicasts are used to distribute RP information:
  - Forwarded using PIM dense mode.
  - Cisco announce: 224.0.1.39.
  - Cisco discovery: 224.0.1.40.
- Backup RPs can be configured.
- It can be used with administratively scoped zones.

Auto-RP Candidate RPs

Candidate RPs advertise themselves:

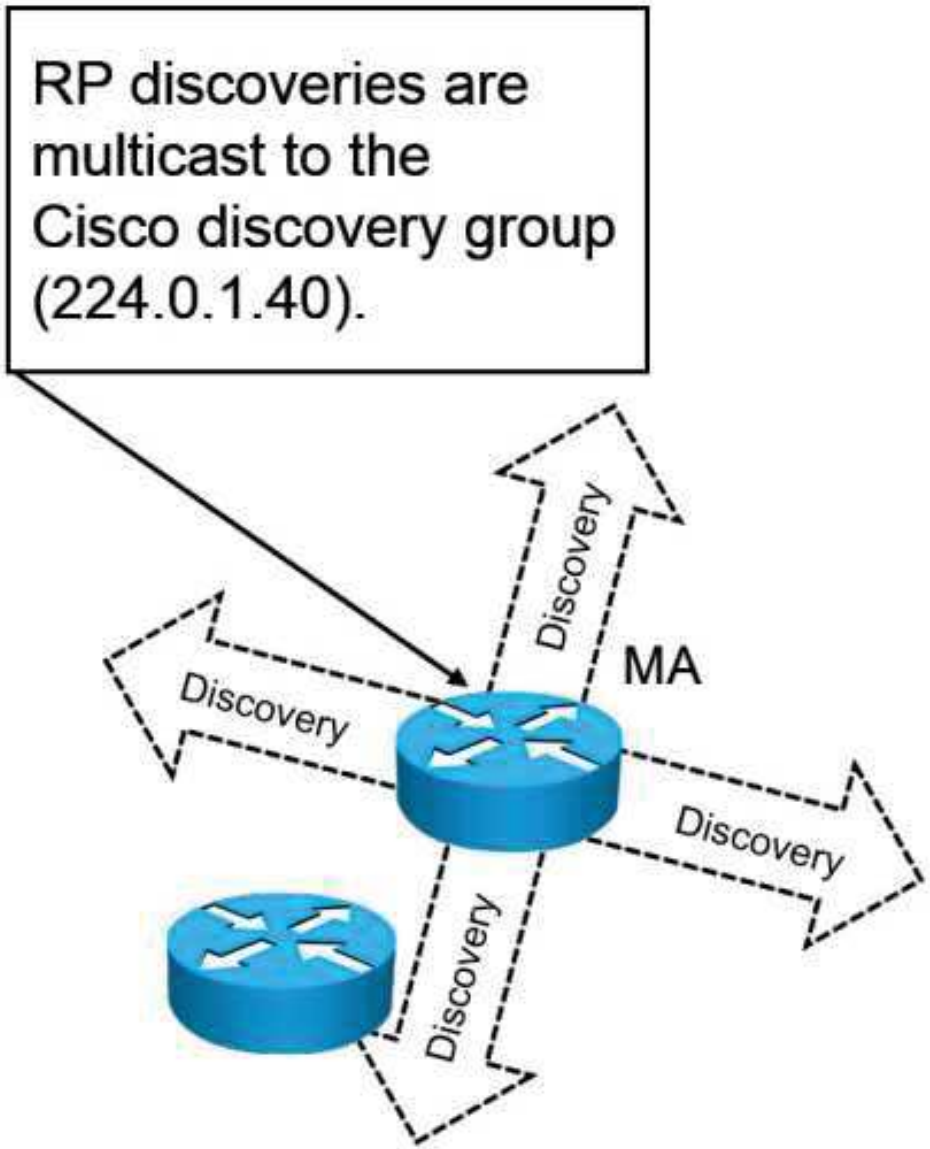
- Using multicast RP-announcement messages.
- To Cisco announce group (224.0.1.39).
- With RP announcements sent at intervals (default is 60 seconds).
- RP announcements contain:
  - Group range (default is 224.0.0.0/4).
  - Candidate-RP address.
  - Hold time, which is three times the duration of the RP-announce interval.



Auto-RP Mapping Agents

Mapping agents act as relays:

- Receive RP announcements.
- Store the group-to-RP mapping in cache:
  - Including the hold times.
- Elect the highest candidate-RP address as RP for group range.
- Advertise RP-discovery messages:
  - As multicast to Cisco discovery group (224.0.1.40).
  - Every 60 seconds, or when changes detected.
- RP-discovery messages contain:
  - Elected RPs from mapping agent group-to-RP mapping cache.



Auto-RP Other Routers

All Cisco routers:

- Join Cisco discovery group (224.0.1.40):
  - Automatic
  - No configuration necessary.
- Receive RP-discovery messages:
  - Stored in local group-to-RP mapping cache.
  - Used to determine RP for group range.

Auto-RP Configuration

Auto-RP configuration process:

- Configure candidate RPs
- Configure mapping agents
- Optional tasks:
  - Advertisement filtering
  - Failover tuning
  - RP fallback
  - Administrative scoping
- Verify and troubleshoot.

Auto-RP Configuration (Cont.)

Configure the candidate on the router.

```
ip pim send-rp-announce interface
```

```
router pim
 auto-rp candidate-rp interface
```

Configure the mapping agent on the router.

```
ip pim send-rp-discovery interface
```

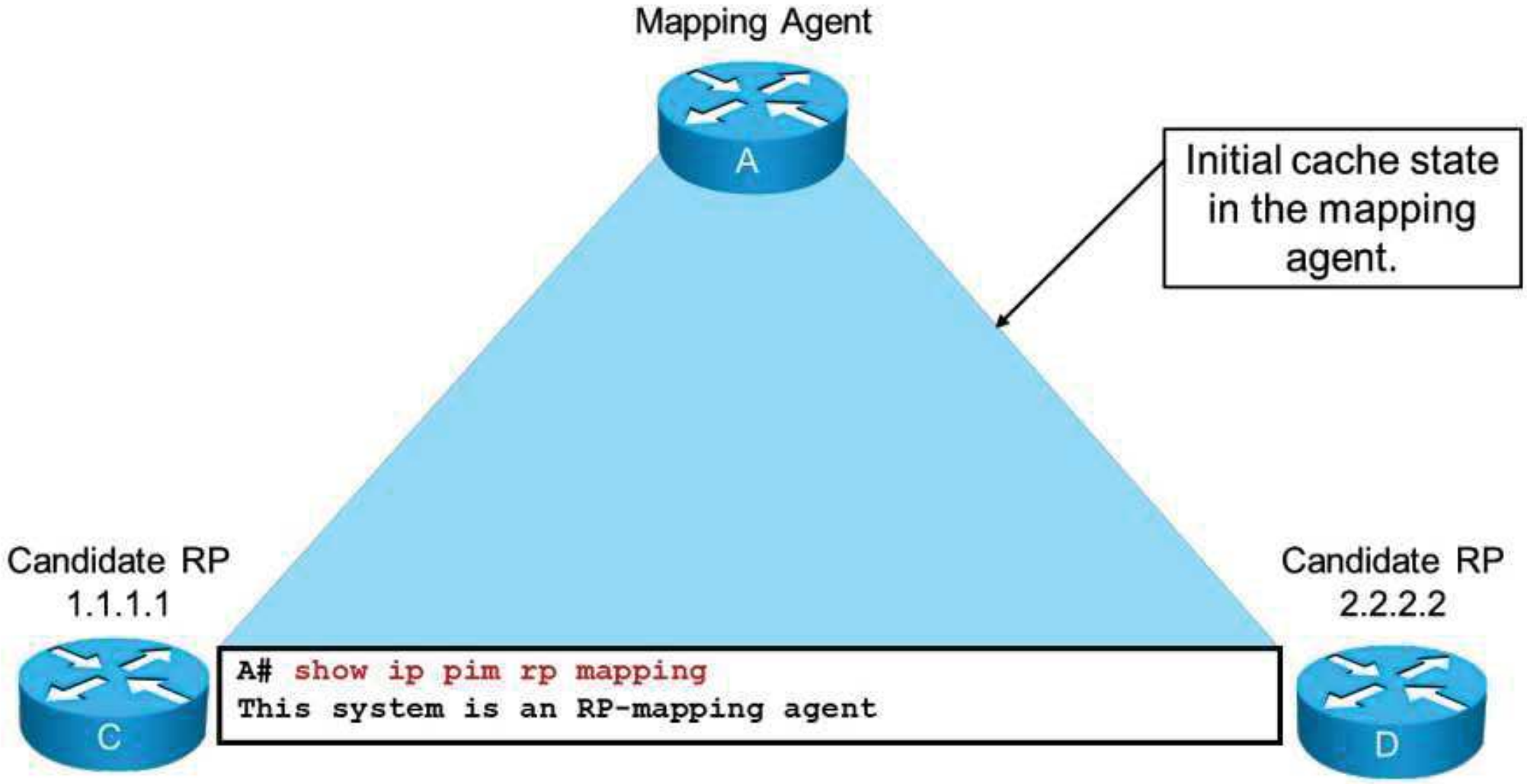
```
router pim
 auto-rp mapping-agent interface
```

Filter RP announcements on the mapping agent.

```
ip pim rp-announce-filter rp-list
access-list group-list access-list
```

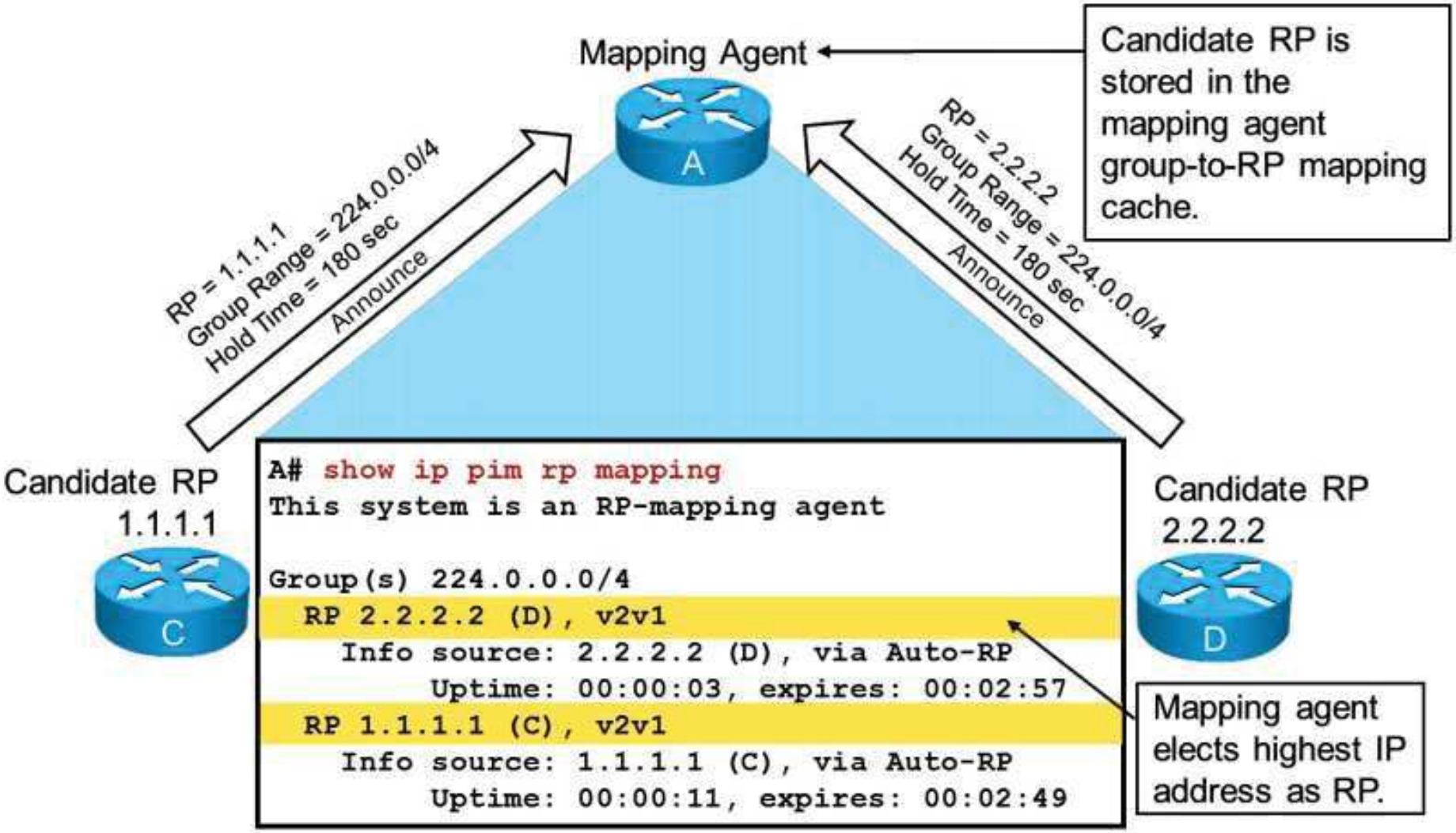
Auto-RP Troubleshooting

Step 1:



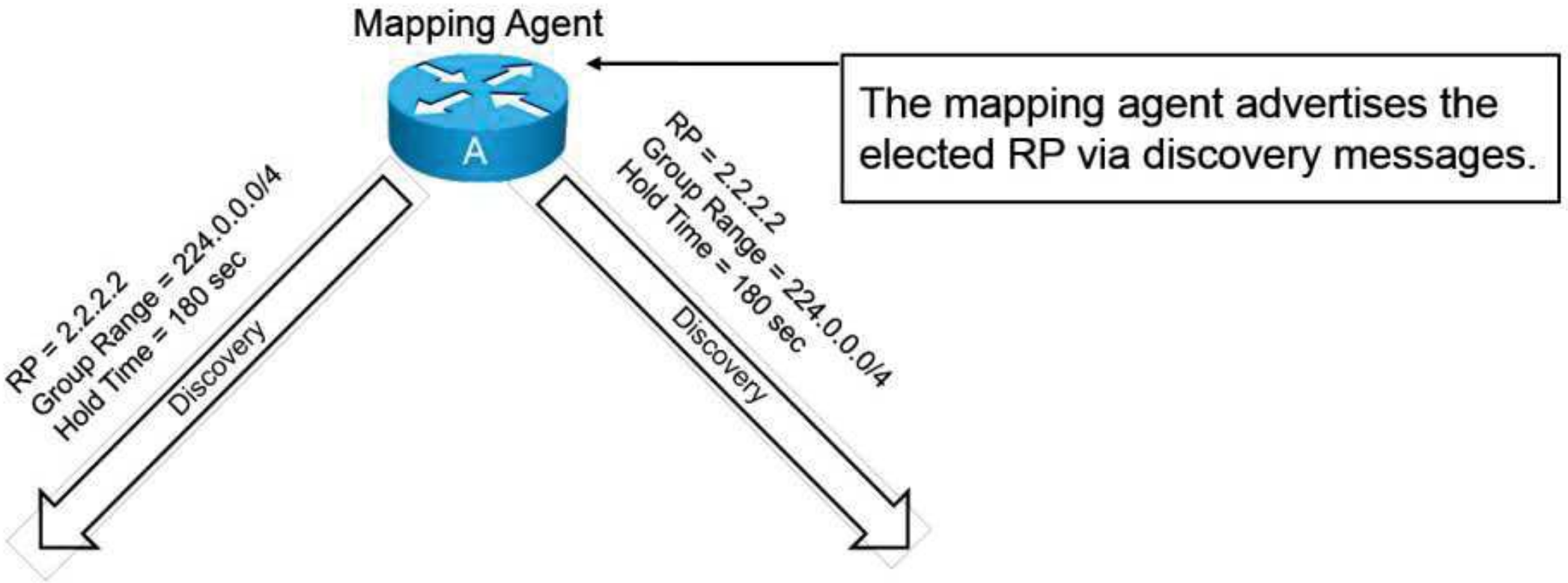
Auto-RP Troubleshooting (Cont.)

Step 2:



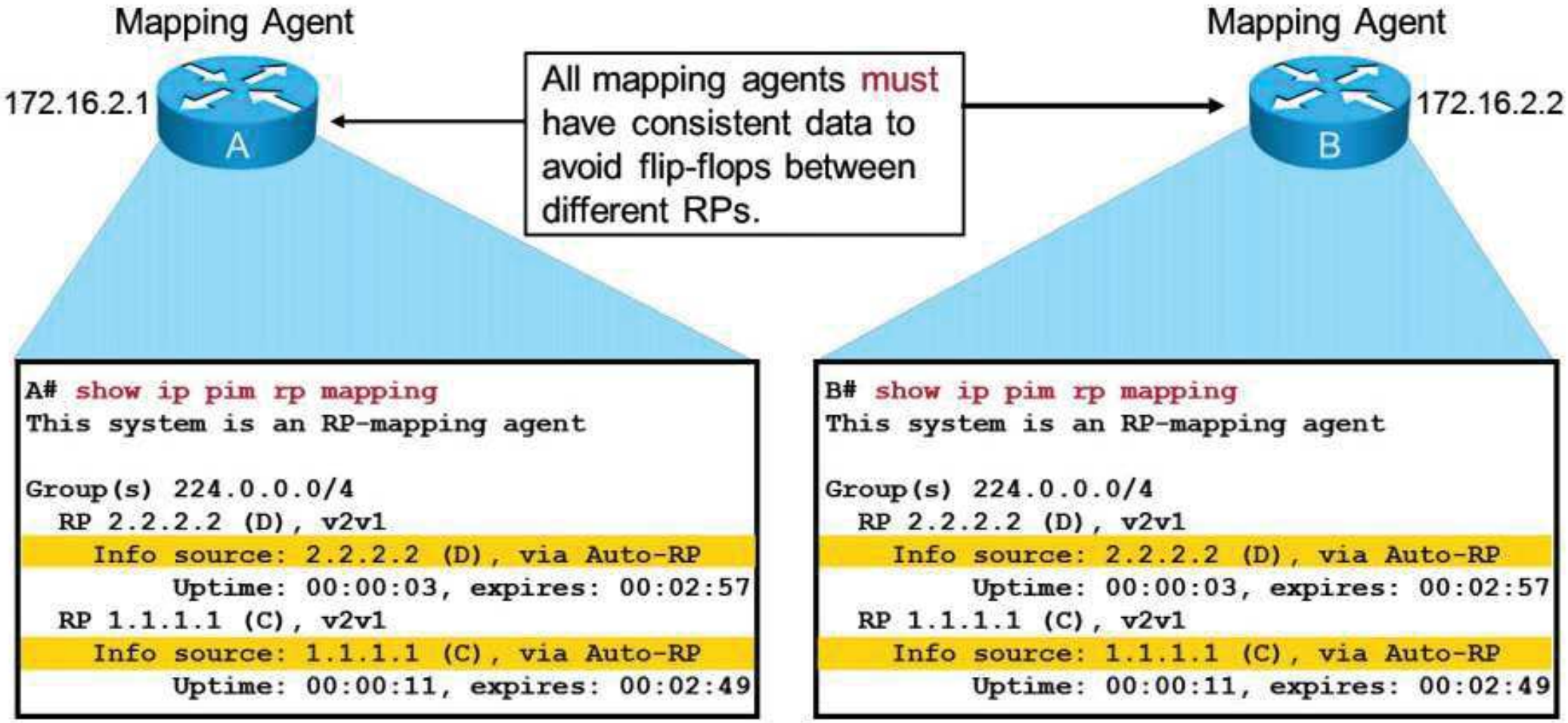
Auto-RP Troubleshooting (Cont.)

Step 3:



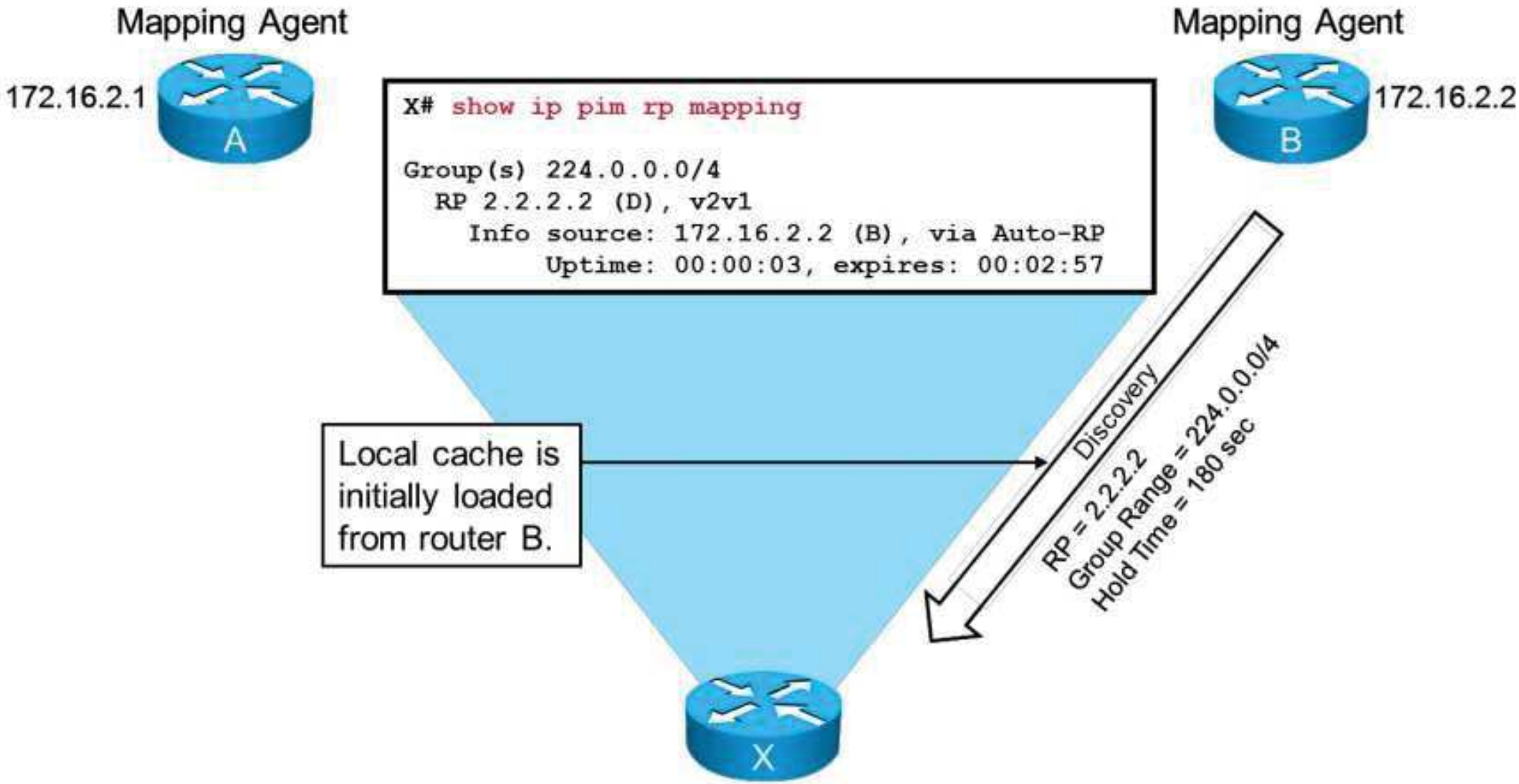
Auto-RP Troubleshooting (Cont.)

Step 4:



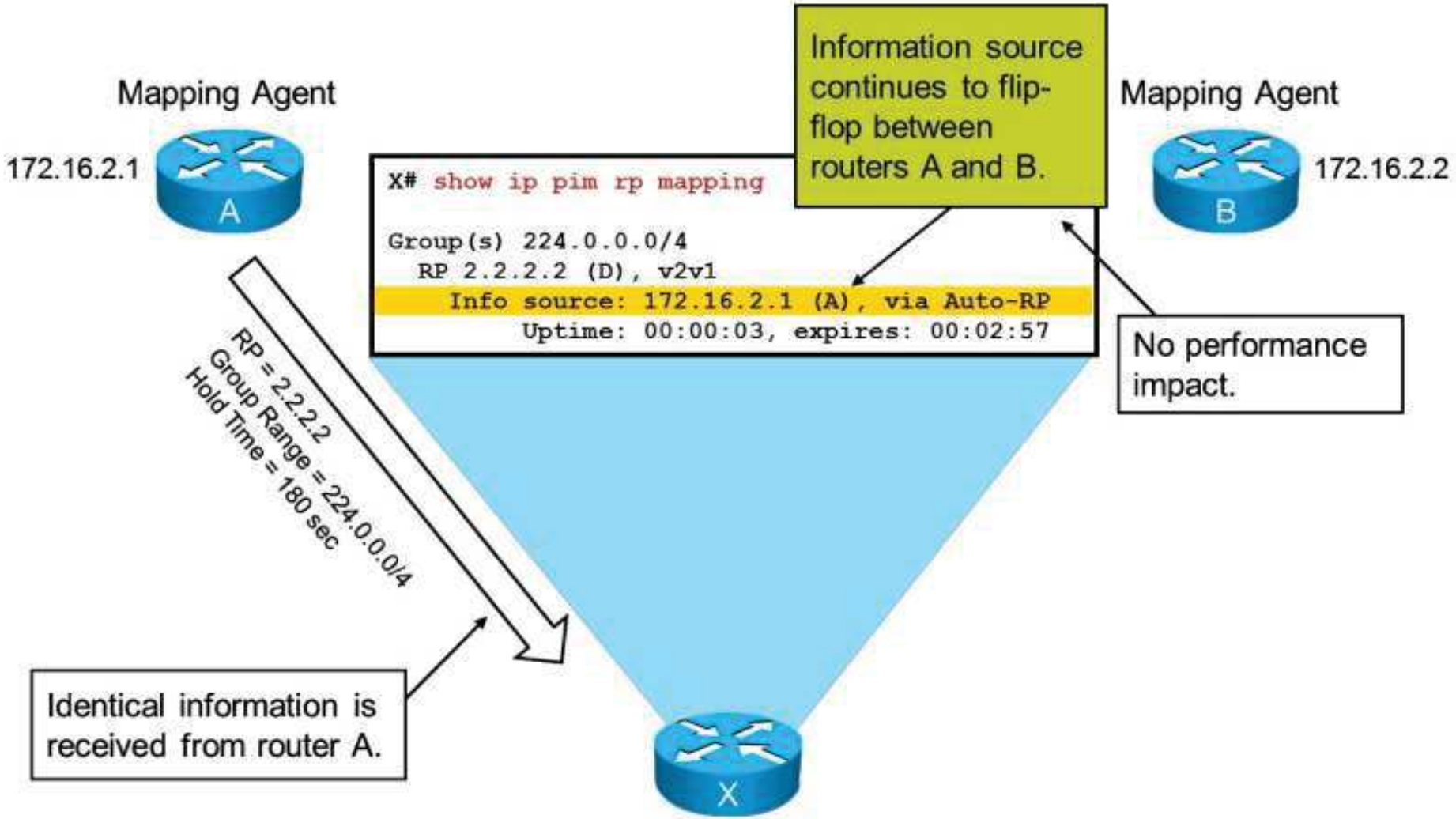
Auto-RP Troubleshooting (Cont.)

Step 5:



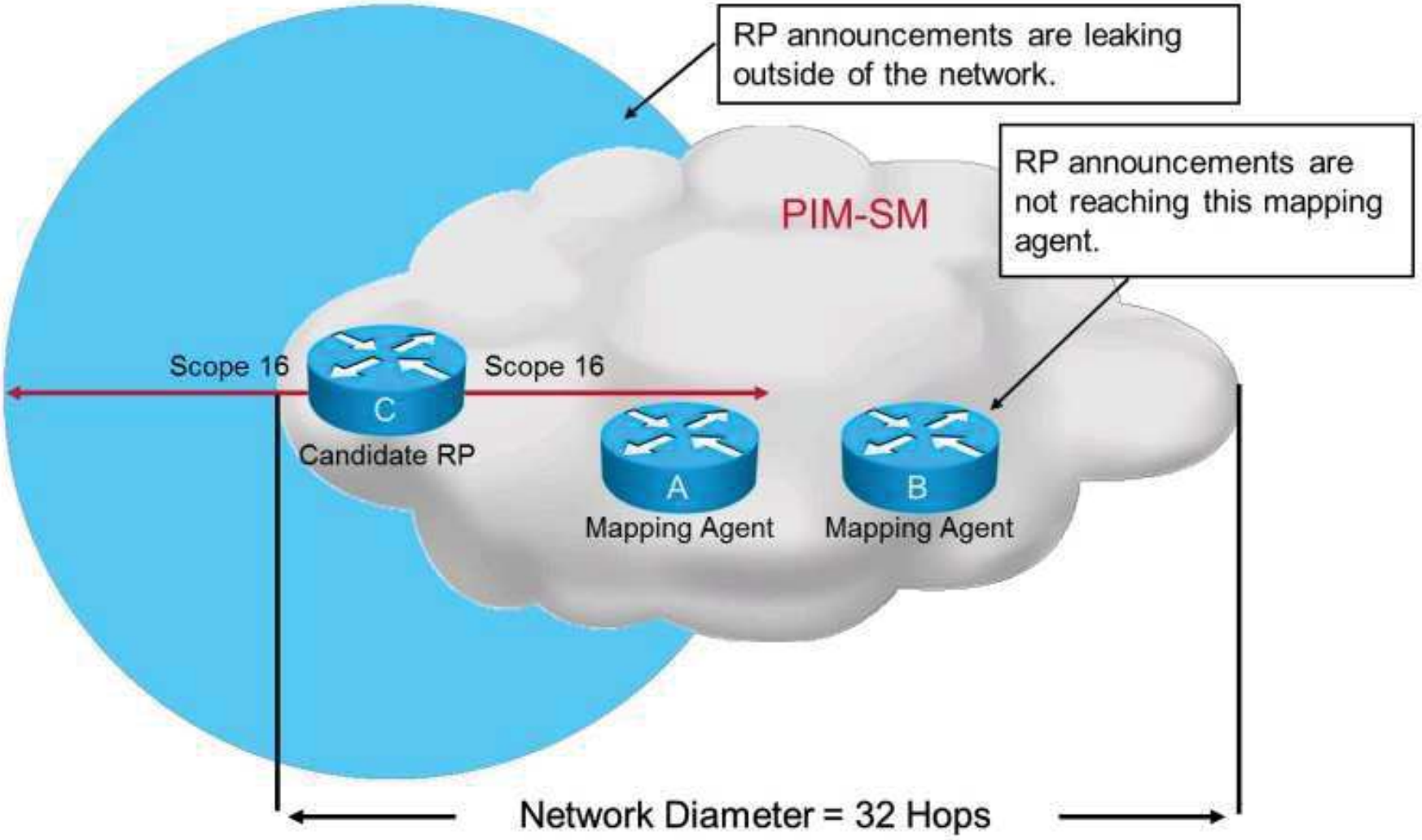
Auto-RP Troubleshooting (Cont.)

Step 6:



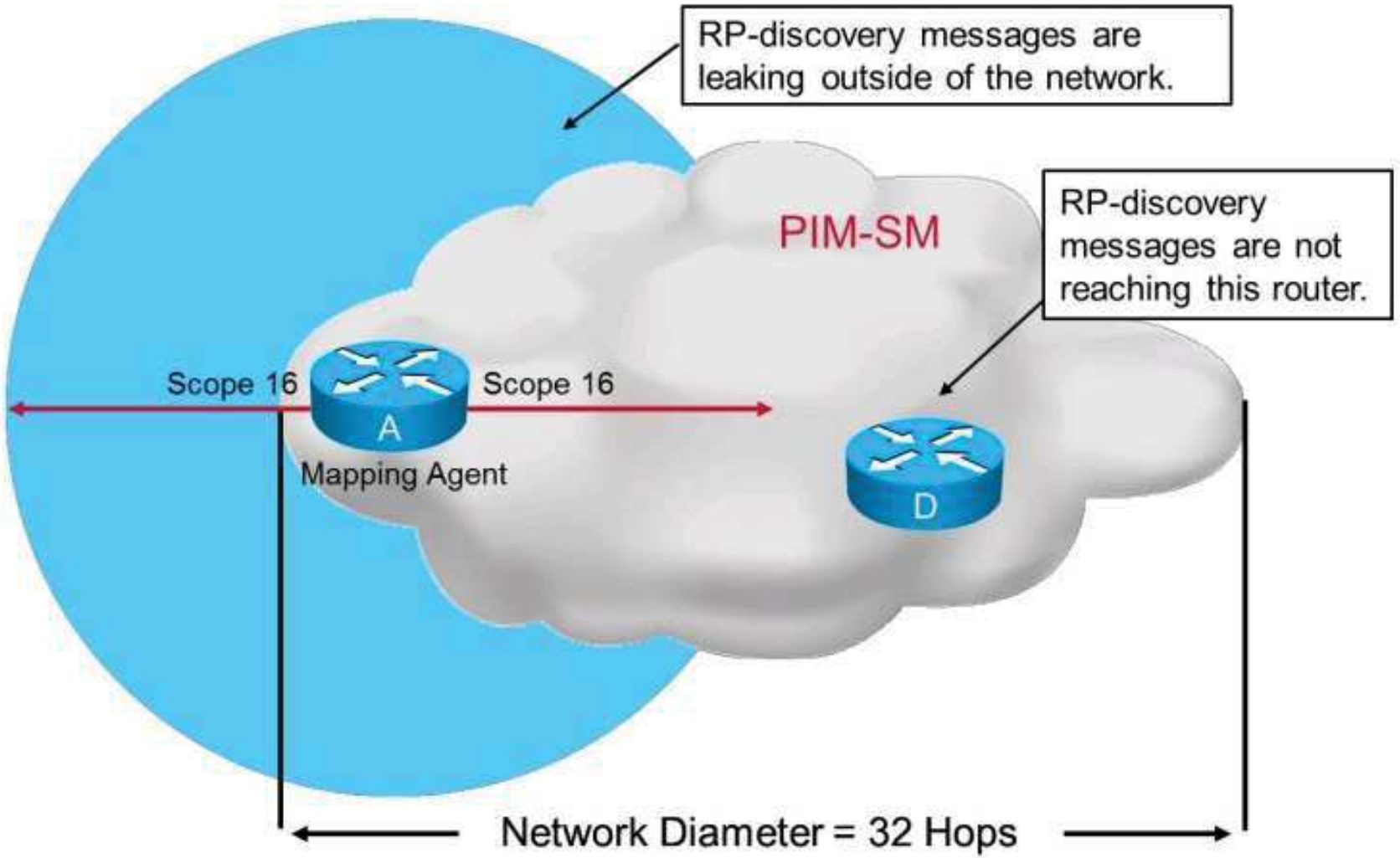
Auto-RP Scoping

Candidate-RP Problem:

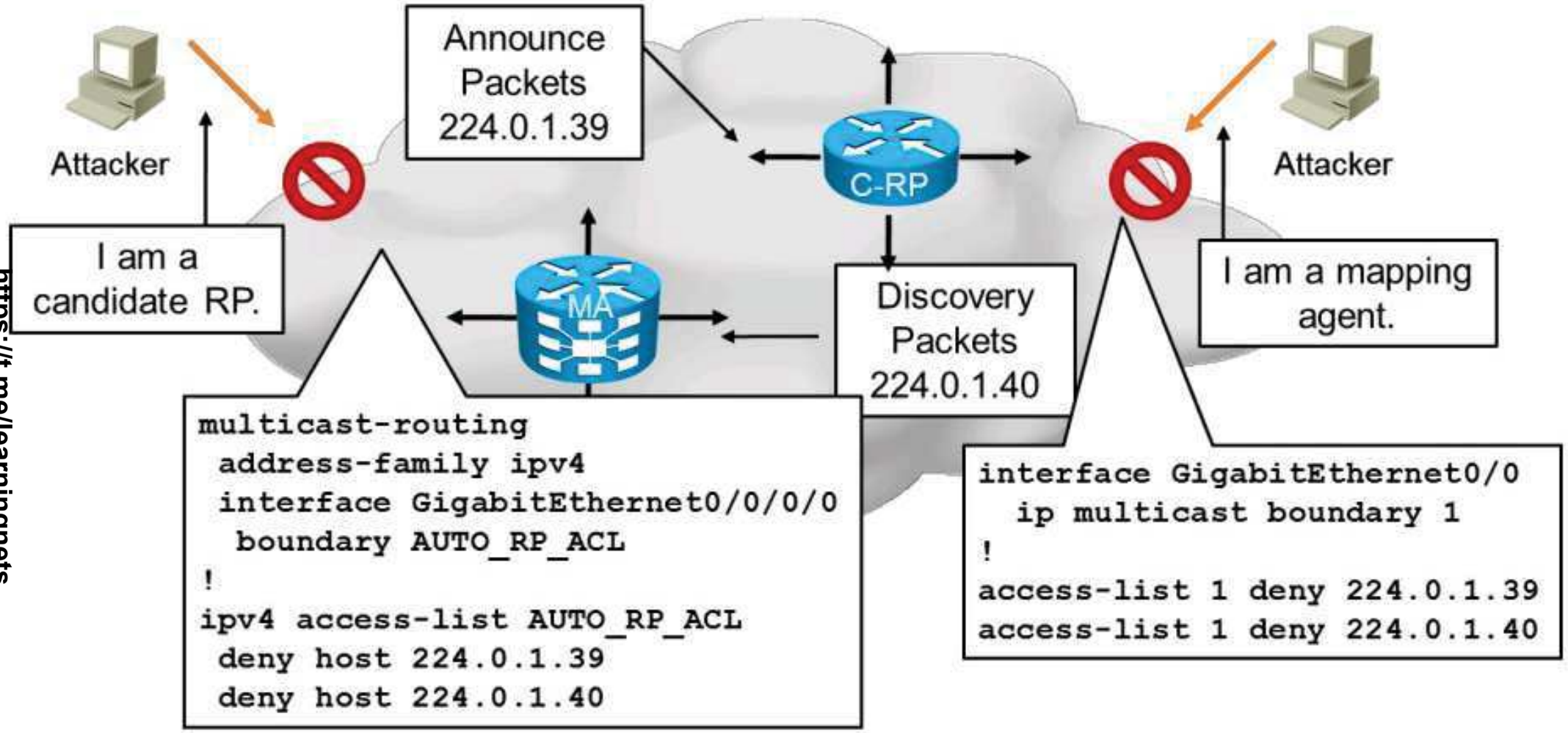


Auto-RP Scoping (Cont.)

Mapping Agent Problem:



Securing Auto-RP Using a Boundary

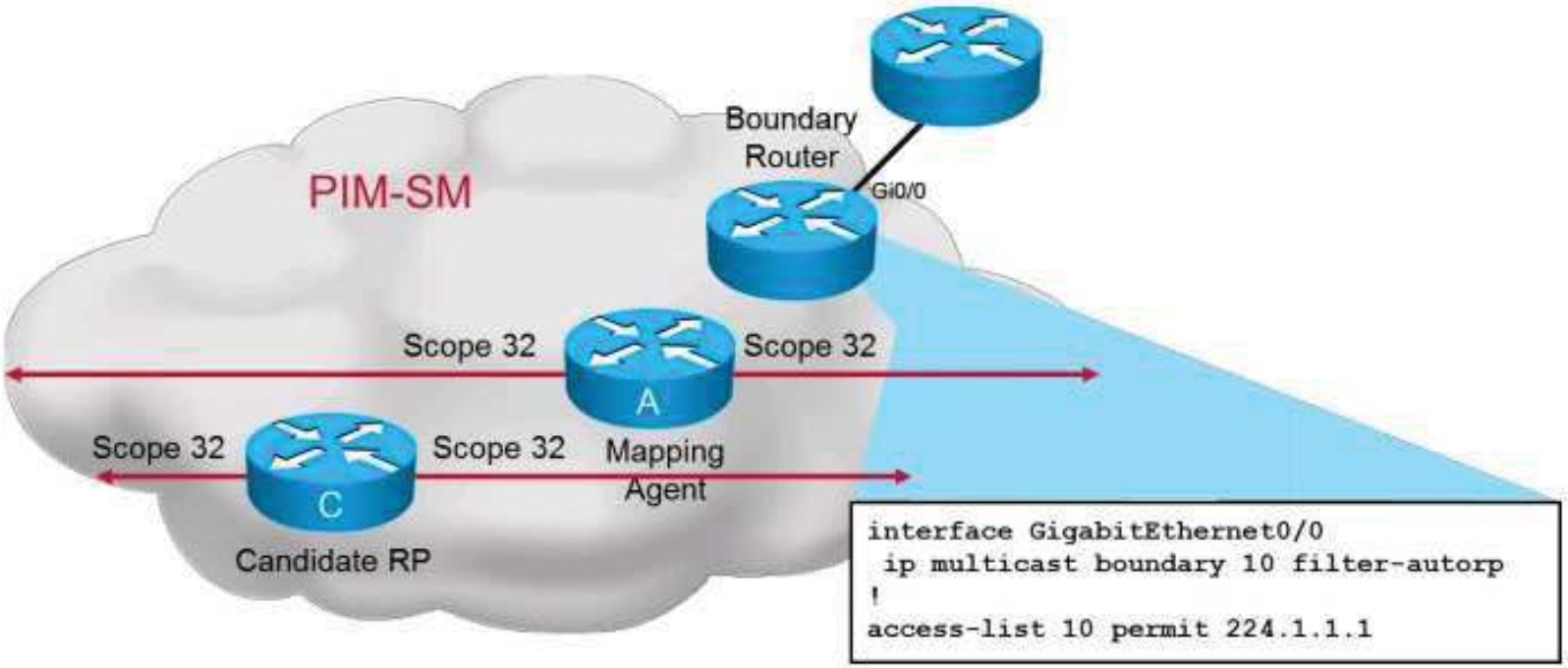


- Filter 224.0.1.39 and 224.0.1.40 at edge.
- **ip multicast boundary** is bidirectional.

Securing Auto-RP Using a Boundary (Cont.)

Securing Auto-RP using a boundary characteristics:

- The **filter-autorp** option filters Auto-RP discovery and announcement messages.
- It permits an Auto-RP group range announcement only if all addresses in the Auto-RP group range are permitted by access list.
- This is important, because RP designation can be configured only on RPs, not on the leaf routers.



PIMv2 Bootstrap Router

PIMv2 BSR characteristics:

- A single BSR is elected. Multiple candidate BSRs can exist.
- Candidate RPs send candidacy announcements to the BSR:
  - Candidate-RP announcements are sent via unicast.
  - BSR stores all candidate-RP announcements in the RP set.
- BSR periodically sends BSR messages to all routers:
  - With the entire RP set and IP address of the BSR.
  - Which are flooded hop by hop throughout the network.
- All routers select the RP from the RP set.

PIMv2 BSR Candidate RPs

PIMv2 BSR Candidate RPs characteristics:

- Unicast PIMv2 candidate-RP messages to BSR.
  - Learn IP address of BSR from BSR messages.
  - Sent at announcement intervals (default is 60 seconds).
- Candidate-RP messages contain:
  - Group range (default is 224.0.0.0/4).
  - Candidate-RP address.
  - Hold time, which is three times the RP-announcement interval.

PIMv2 BSRs

- BSR receives candidate-RP messages:
  - Accepts and stores all candidate-RP messages.
  - Stored in group-to-RP mapping cache with hold time.
- BSR advertises BSR messages:
  - To all-PIM-routers (224.0.0.13) multicast group.
  - With TTL of 1.
  - Via all interfaces, propagated hop by hop.
  - Every 60 seconds, or when changes are detected.
- BSR messages contain:
  - Contents of BSR group-to-RP mapping cache.
  - IP address of active BSR.
- Candidate BSR with highest priority elected as BSR.
  - Candidate-BSR IP address used as tiebreaker (highest wins).
- The active BSR may be preempted:
  - New router with higher BSR priority forces new election.

PIMv2 BSR Election

PIMv2 BSR election characteristics:

- Begin in candidate-BSR state:
  - BSR-timeout timer started.
  - If higher-priority BSR message received:
    - Restart timer and forward BSR message.
    - Copy information to local group-to-RP mapping cache.
  - If other-priority BSR message received, discard it.
  - If timer expires, transition to elected-BSR state.
- While in elected-BSR state:
  - Periodically originate own BSR messages (include local group-to-RP mapping cache).
  - Return to candidate-BSR state if higher-priority BSR message received.

PIMv2 BSR Election (Cont.)

BSR selection for all other routers characteristics:

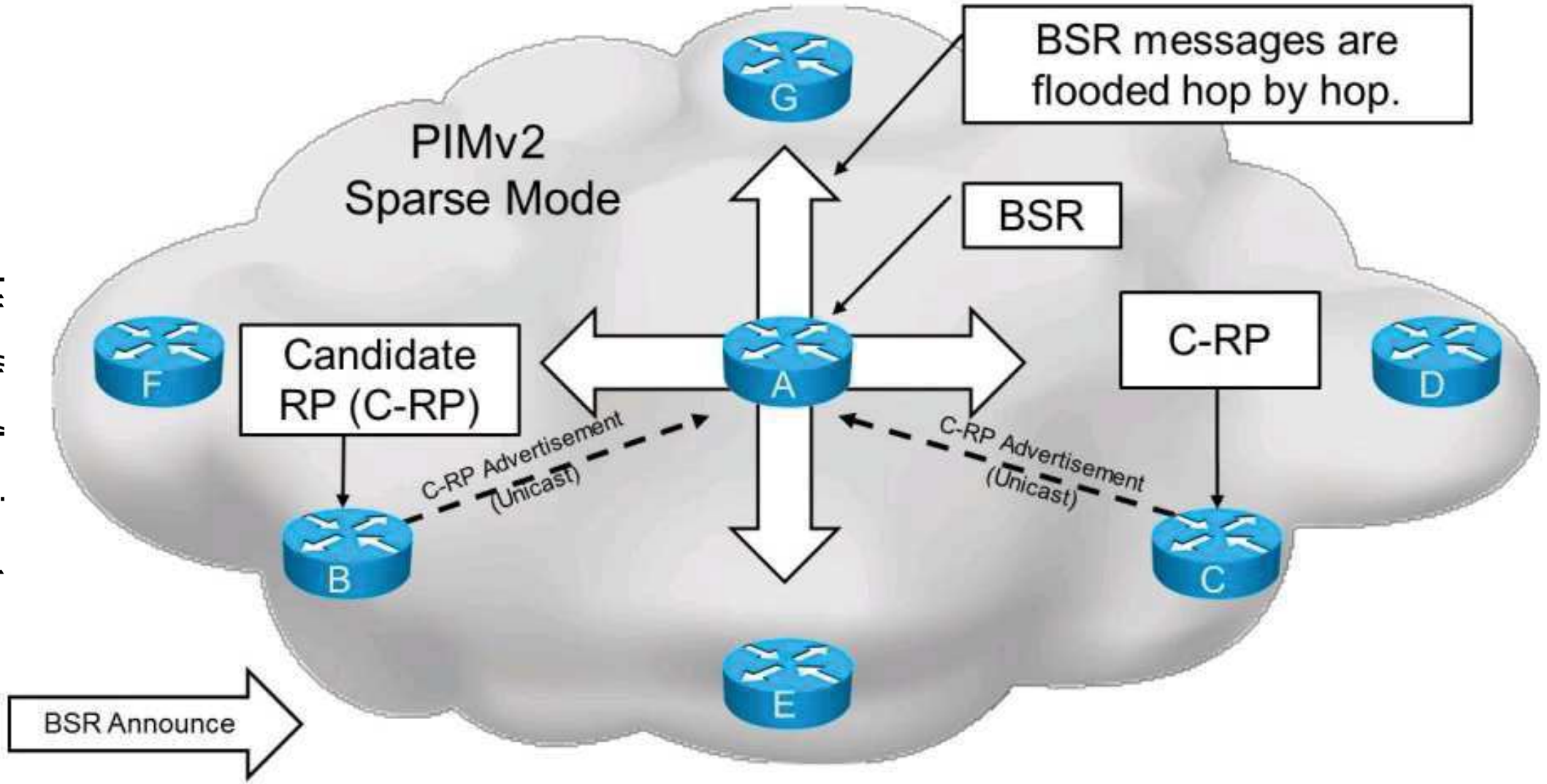
- Start in accept-any state:
  - Accept first BSR message received.
  - Save BSR information, and forward BSR message.
  - Transition to accept-preferred state.
- While in accept-preferred state:
  - Start BSR-timeout timer.
  - Only accept and forward preferred BSR messages with priority higher than current BSR priority.
  - Discard nonpreferred BSR messages.
  - Return to accept-any state if timer expires.

Other PIMv2 BSR Routers

Other PIMv2 BSR routers characteristics:

- Receive BSR messages:
  - Stored in local group-to-RP mapping cache.
  - Information used to determine active BSR address.
- Select 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 range.

PIMv2 BSR Advertisement Process



PIMv2 BSR Configuration

Configure candidate RP on the router.

```
ip pim rp-candidate interface
```

```
router pim
 bsr candidate-rp ip_address
```

Configure candidate BSR on the router.

```
ip pim bsr-candidate interface
```

```
router pim
 bsr candidate-bsr ip_address
```

PIMv2 BSR Verification

```
RP/0/RSP0/CPU0:PE5# show pim bsr election
PIM BSR Election State

Cand/Elect-State      Uptime     BS-Timer    BSR                  C-BSR

Elected/Accept-Pref   00:35:09   00:00:05    10.5.1.1 [1, 30]   10.5.1.1 [1,
30]
```

- Displays candidate election information for BSR.

```
RP/0/RSP0/CPU0:PE5# show pim bsr rp-cache
PIM BSR Candidate RP Cache

Group(s) 224.0.0.0/4, RP count 1
RP-addr    Priority   Holdtime(s)  Uptime   Expires
10.6.1.1    0           150          00:35:34  00:01:55
```

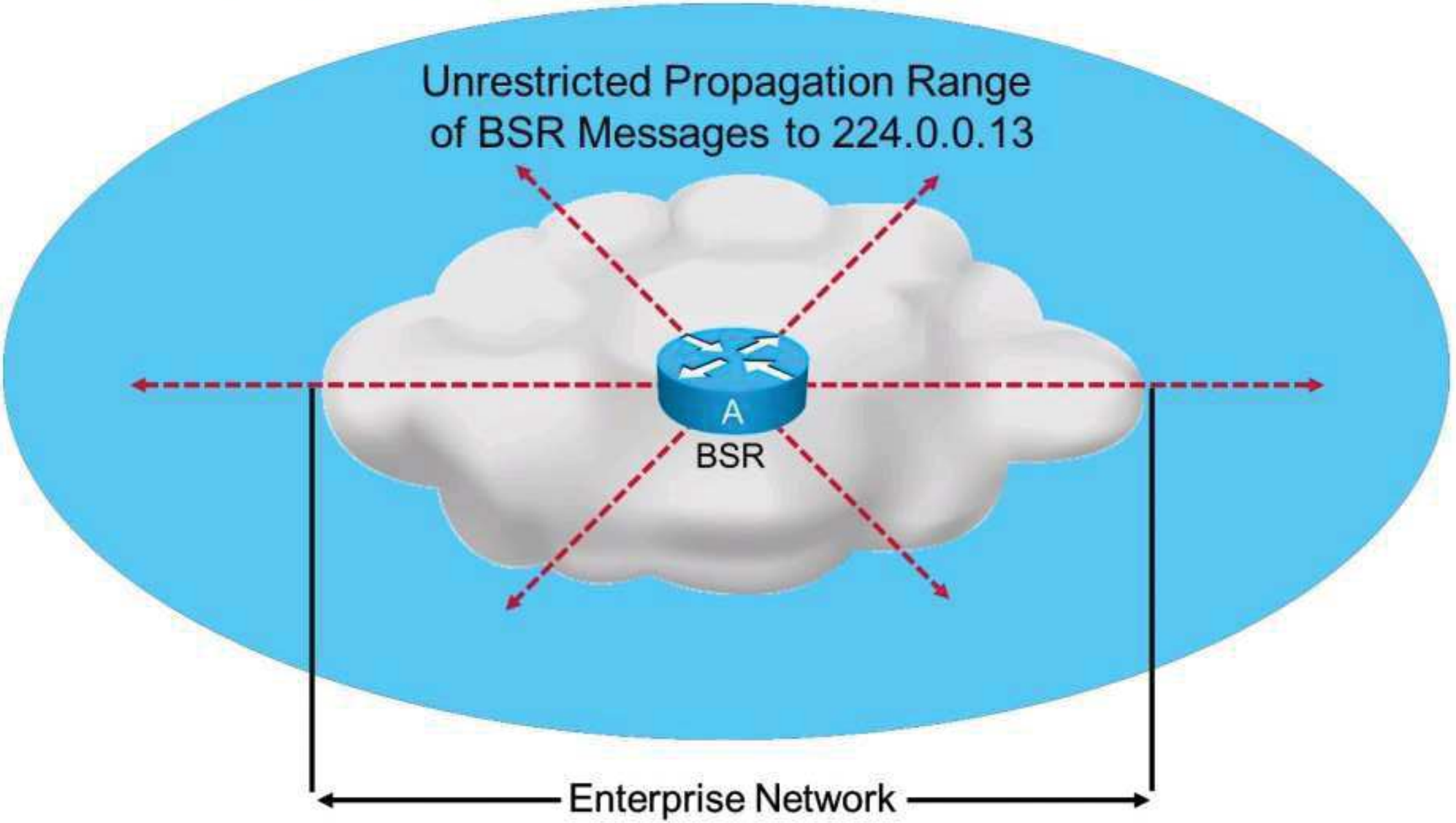
- Displays RP cache information for BSR.

PIMv2 BSR Troubleshooting

This is BSR troubleshoot process:

1. Verify BSR election.
2. Verify group-to-RP mapping caches.
   - First on the BSR:
     - Other routers will learn group-to-RP mapping information from this router.
       - If it is not correct, use **debug** commands to see what is wrong.
   - Then, on other routers:
     - If the information does not match the BSR, there is a problem distributing the information.
     - Use **show** and **debug** commands to find where the break is.

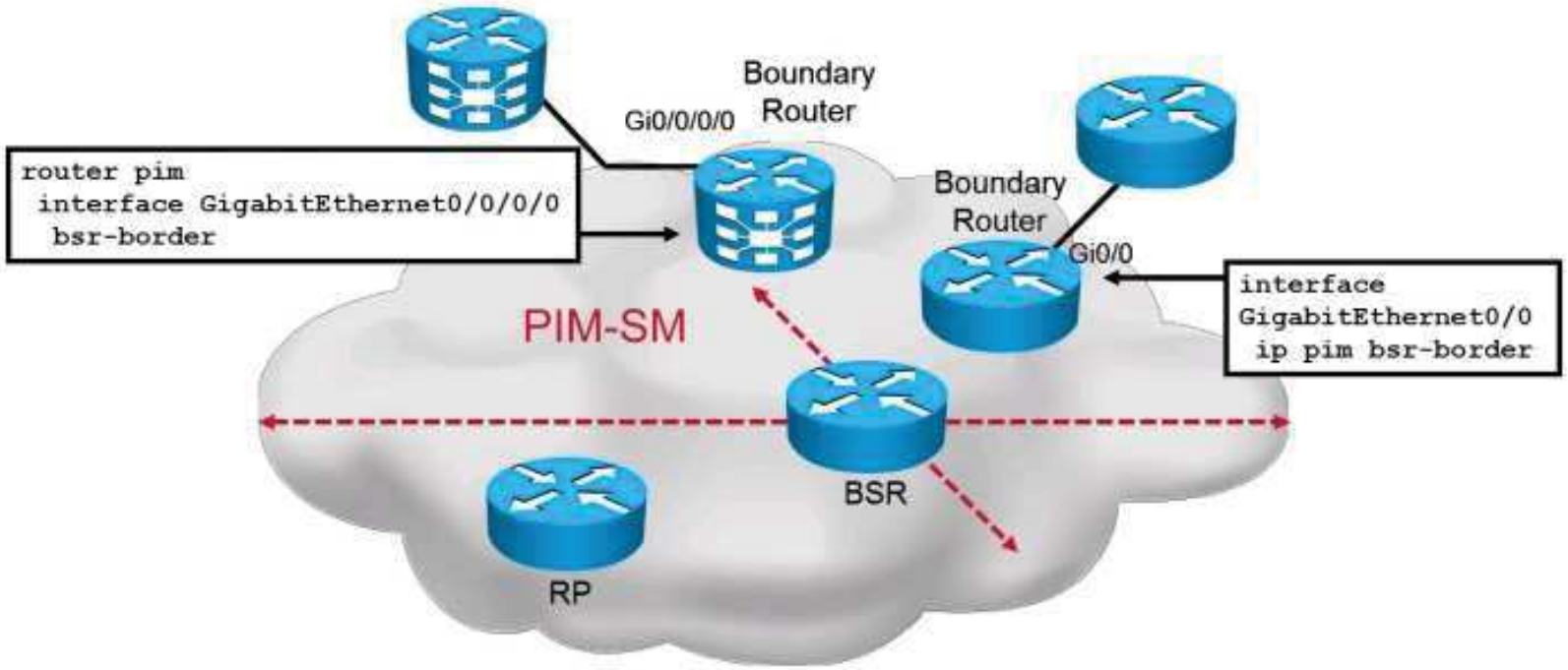
BSR Hop-by-Hop Flooding



Constraining BSR Messages

Constraining BSR messages characteristics:

- Stops BSR messages on an interface.
- Stops inbound and outbound packets.
- Does not set up boundary for other multicasts.



Anycast RP

Anycast RP in the Cisco IP NGN Infrastructure Layer:

- Within a PIM-SM domain, deploy more than one RP for the same group range.
- Each router uses the closest RP.
- RPs use MSDP to inform one another about active sources in their part of the domain.

Anycast RP (Cont.)

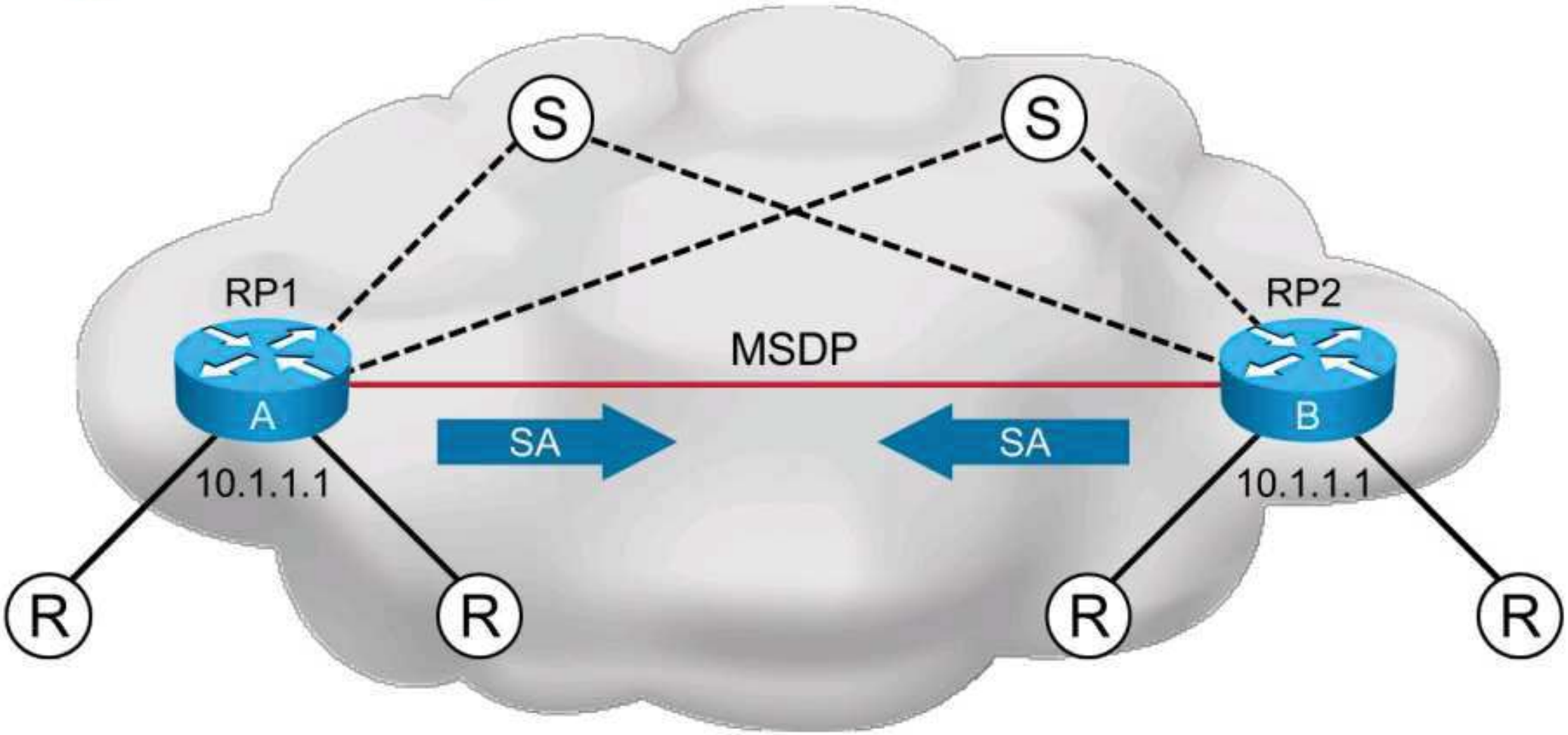
Anycast RP features benefits:

- RP backup without using Auto-RP or BSR.
- RP failover at speed of unicast routing protocol.

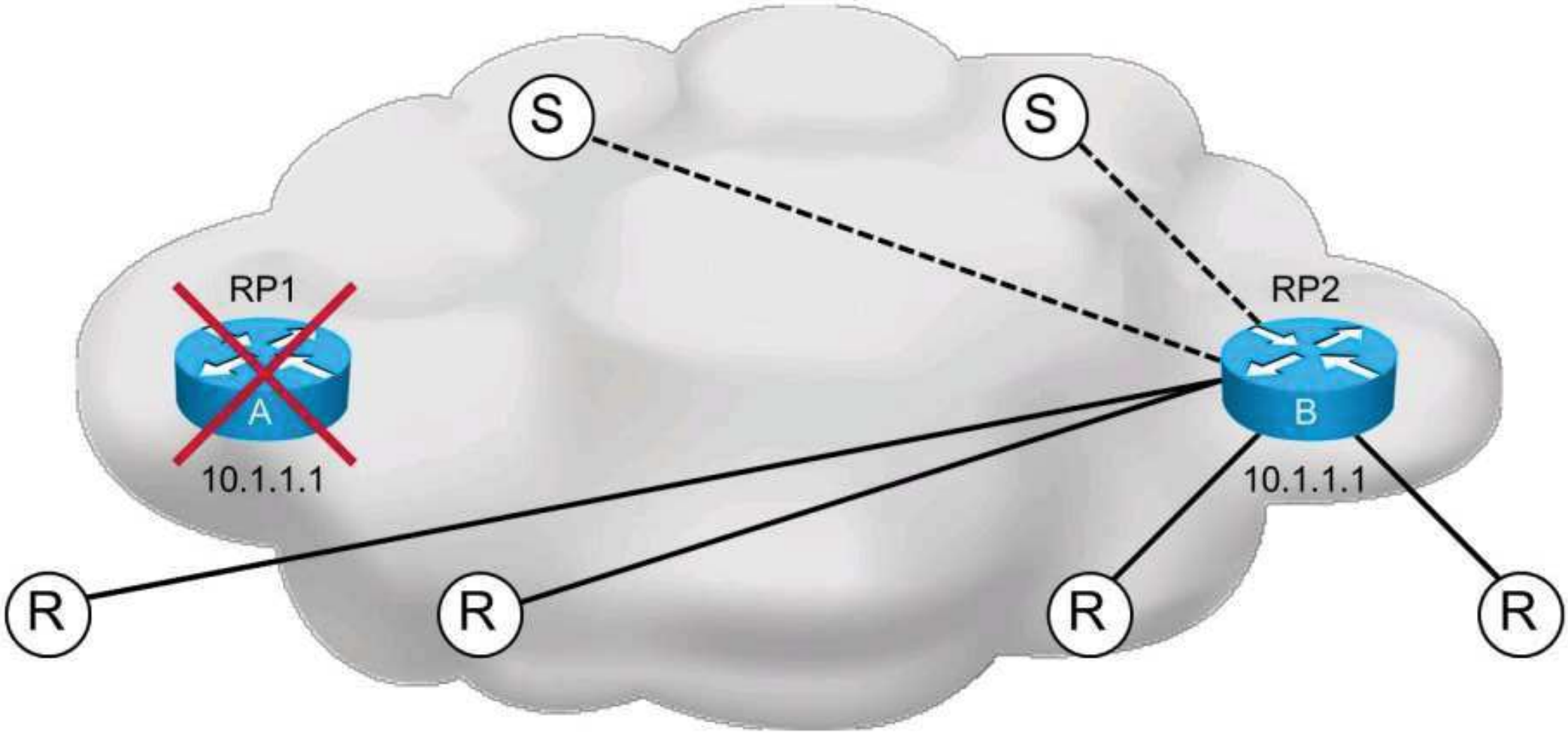
Anycast RP features requirements:

- Use only one IP address for all your RPs.
- RPs advertise this address as a host route.
- MSDP is used between the RP routers.
- **originator-id** command disambiguates which RP originated the source address message.

Anycast RP Example



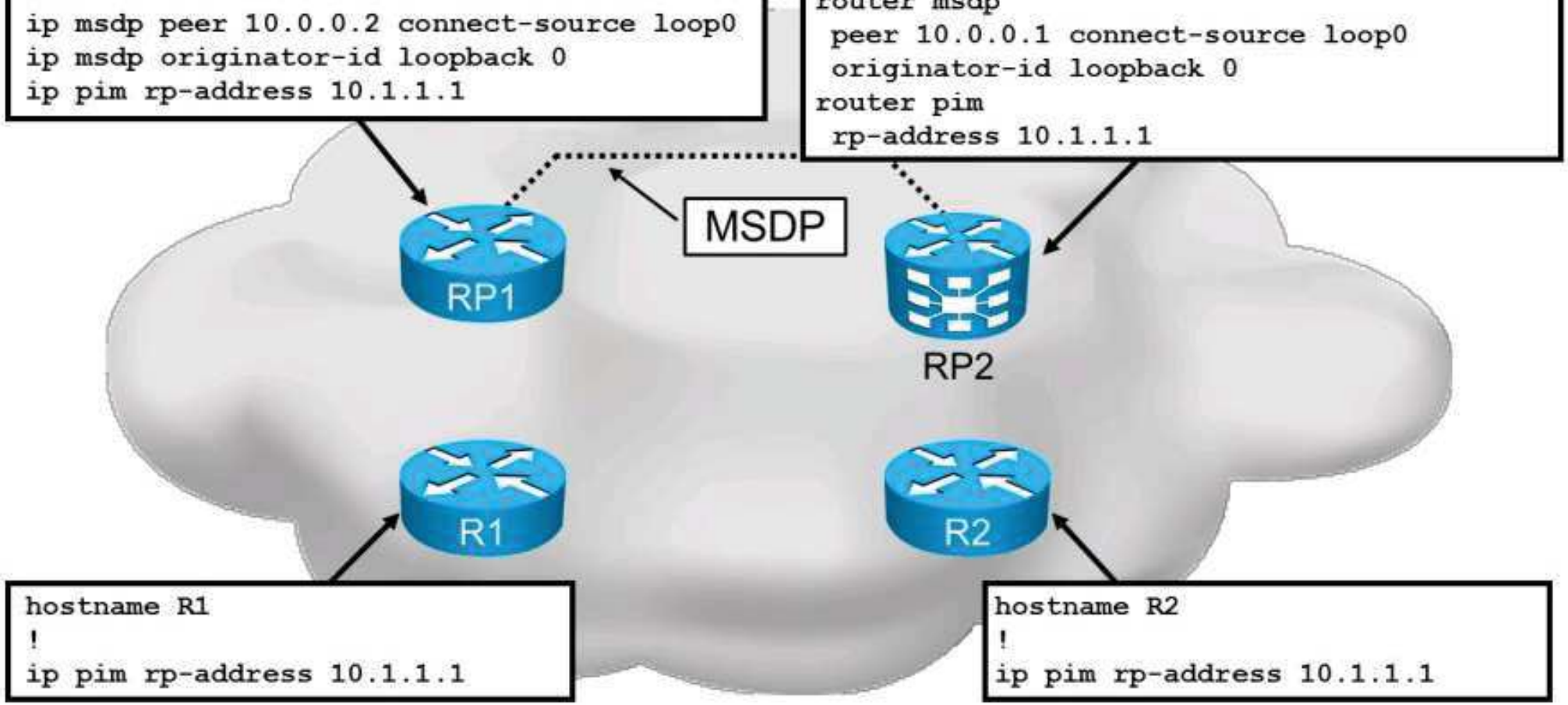
Anycast RP Example (Cont.)



Anycast RP Configuration

```
hostname RP1
!
interface loopback 0
 ip address 10.0.0.1 255.255.255.255
!
interface loopback 1
 ip address 10.1.1.1 255.255.255.255
!
ip msdp peer 10.0.0.2 connect-source loop0
ip msdp originator-id loopback 0
ip pim rp-address 10.1.1.1
```

```
hostname RP2
!
interface loopback 0
 ip address 10.0.0.2 255.255.255.255
!
interface loopback 1
 ip address 10.1.1.1 255.255.255.255
!
router msdp
 peer 10.0.0.1 connect-source loop0
 originator-id loopback 0
router pim
 rp-address 10.1.1.1
```




```
hostname R1
!
ip pim rp-address 10.1.1.1
```

```
hostname R2
!
ip pim rp-address 10.1.1.1
```

Anycast RP Configuration Guidelines

Avoid Anycast RP and router ID conflicts:

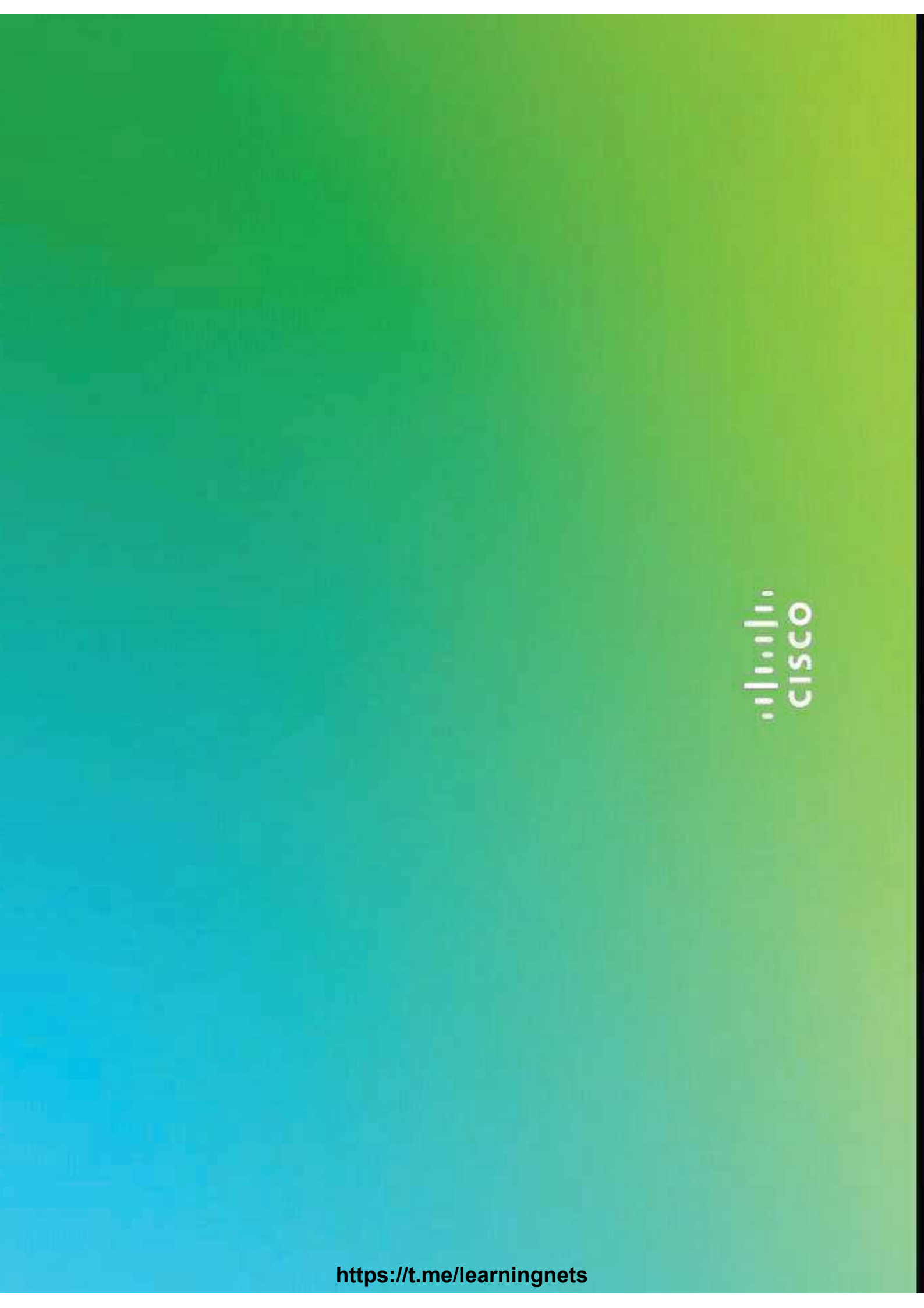
- Ensure that the loopback address used for Anycast RP address is not accidentally used as a router ID:
  - This will break OSPF and BGP.
- How to avoid a conflict with the router ID:
  - Configure the Anycast RP address as the lowest IP address.
  - Use a secondary IP address on the loopback for the Anycast IP address.
  - Use the **router-id** command in OSPF and BGP to statically configure the router ID.

Summary

- Static RP offers no fault tolerance.
- There are two protocols for automatic RP discovery: auto-RP and BSD.
- RP can be anywhere in the network, but smart placing can benefit scalability.
- With auto-RP, all routers learn RP address.
- Candidate RPs will advertise themselves.
- Mapping angents act as relays.
- All other routers will receive and forward auto-RP messages.
- To configure auto-RP, configure candidates and mapping agents.
- Troubleshooting auto-RP can be done in steps.
- Reach of auto-RP can be to great or too smal.
- Boundaries can be put into place to restrict the scope of auto-RP.
- BSR periodically sends relevant RP information to all routers.
- There is only one BSR, but there may be more BSR candidates.

Summary (Cont.)

- BSR receives and forwards all candidate RP information.
- BSR is elected based on priority.
- If a router is neither BSR nor candidate BSR it will still process BSR messages.
- BSR messages are forwarded hop by hop.
- BSR is configured under PIM.
- When troubleshooting BSR, check election process and group-to-RP mapping cache.
- BSR messages may be flooded outside of safe bounds.
- Boundary can be configured to restrain BSR messages.
- Anycast RP can be used to provide RP redundancy without BSR or auto-RP.
- Anycast RP can coexist with MSDP.
- Anycast addresses must be configured for Anycast RP.
- Avoid conflicts when configuring anycast.



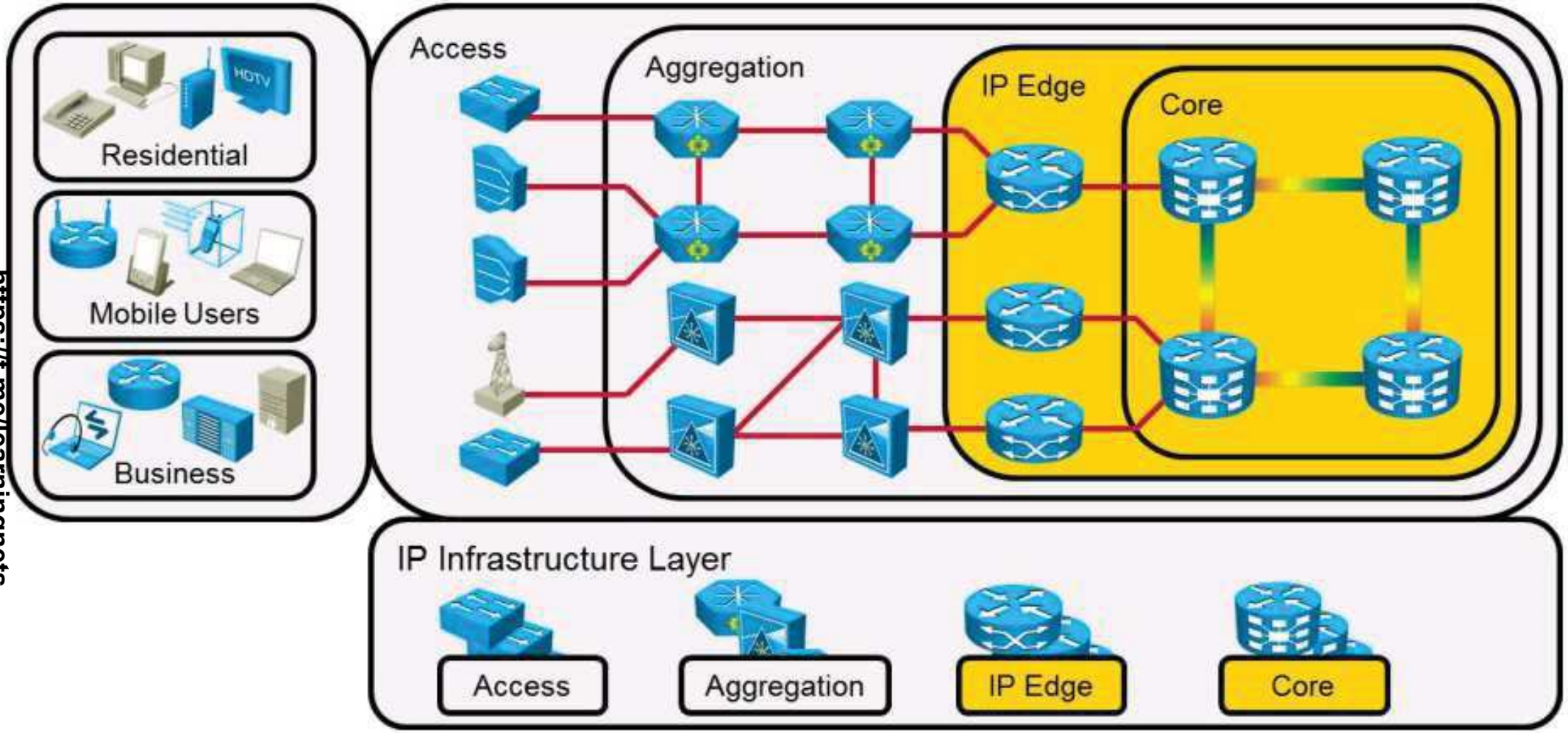
Module Summary

- PIM-SM uses the explicit join model. PIM hello messages are used for discovering PIM neighbors and maintaining neighbor adjacencies.
- SSM supports broadcast applications and builds SPTs only. BIDIR-PIM dispenses with both encapsulation and the (S,G) state.
- Manual RP information configuration does not scale. Two dynamic RP discovery mechanisms—Auto-RP and BSR—are available to support larger configurations.
- MSDP allows RPs to exchange information about active multicast sources. Anycast RP uses MSDP to provide RP redundancy.

Introducing IPv6 Services

Service Provider IPv6 Transition Implementations

Multicast in the Cisco IP NGN Infrastructure Layer

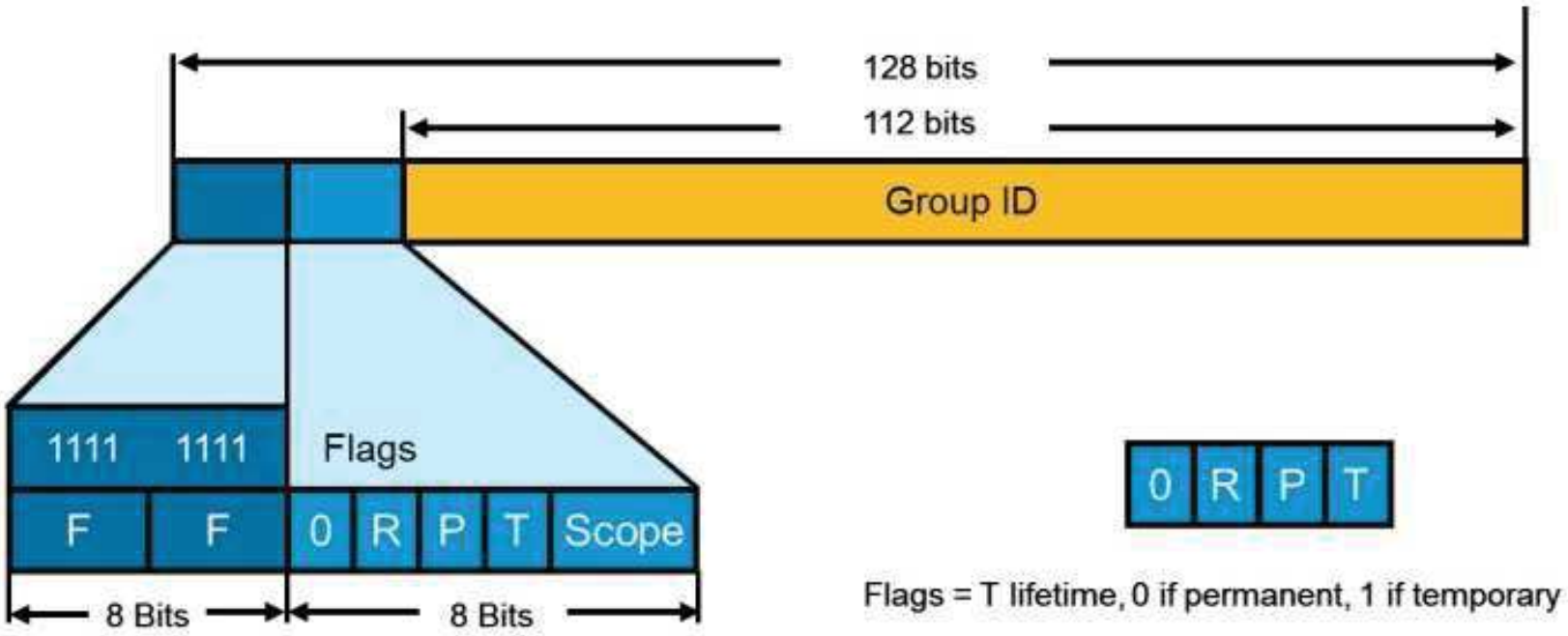


- On the IP infrastructure layer of the Cisco IP NGN.
- On the service provider core and IP edge devices.

IPv6 Multicast Address Format

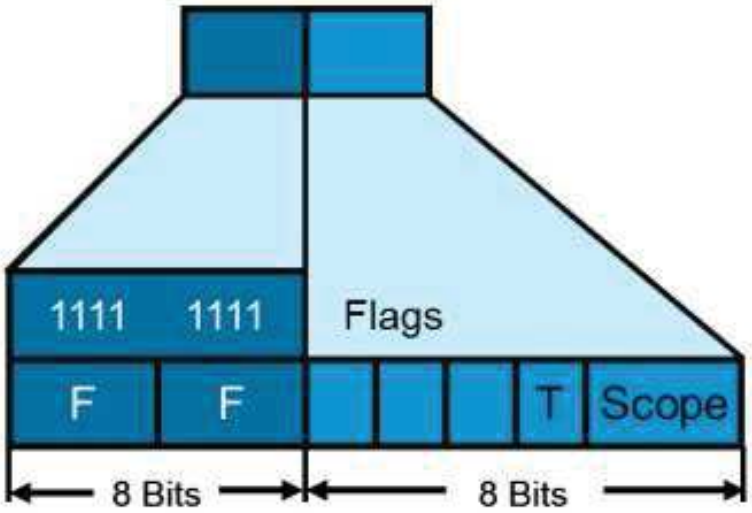
IPv6 multicast address format characteristics:

- IPv6 multicast addresses are distinguishable from unicast and anycast addresses.
- The prefix used is “FF” followed by flags and scope.
- IPv6 multicast address flags define the lifetime of the address.
- IPv6 multicast addresses are temporary or permanent addresses.



IPv6 Multicast Address Scope

- IPv6 multicast address scope defines the reach of the multicast address group.
- Scope defines if group traffic should be forwarded across links, routers, domains, and so on.



Scope	
1 = Interface-Local	0 0 0 1
2 = Link-Local	0 0 1 0
3 = Reserved	0 0 1 1
4 = Admin-Local	0 1 0 0
5 = Site-Local	0 1 0 1
6 = Unassigned	0 1 1 0
7 = Unassigned	0 1 1 1
8 = Organization-Local	1 0 0 0
9 – D = Unassigned	1 0 0 1
	1 1 0 1
E = Global	1 1 1 0
F = Reserved	1 1 1 1

IPv6 Solicited-Node Multicast Address Format

IPv6 solicited-node multicast address format characteristics:

- One solicited-node multicast group for every configured or automatic IPv6 address.
- Multicast address with a link-local scope.
- Formed by a prefix and the rightmost 24 bits of every unicast and anycast address.
- Used to resolve the link-layer address from an IPv6 address.

Solicited-Node Address Format:

FF02	0000	0000	0000	0000	0001	FF	XX	XXXX

IPv6 Address:

2001	0DB8	0001	0000	0001	0800	200E	8C6C

Solicited-Node Multicast Address:

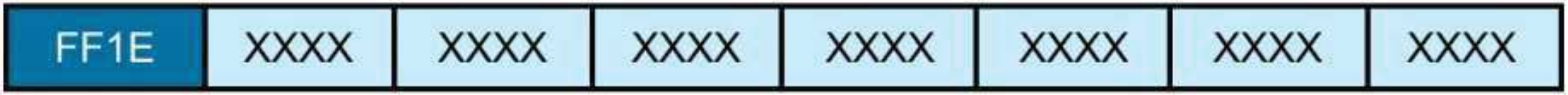
FF02	0000	0000	0000	0000	0001	FF	0E	8C6C

IPv6 Multicast Address with a Global Scope

New IPv6 multicast application address with global scope:

- "FF" specifies multicast.
- "1" specifies a temporary address.
- "E" specifies a global scope.
- "X" is any hexadecimal value.

Temporary multicast address with global scope:



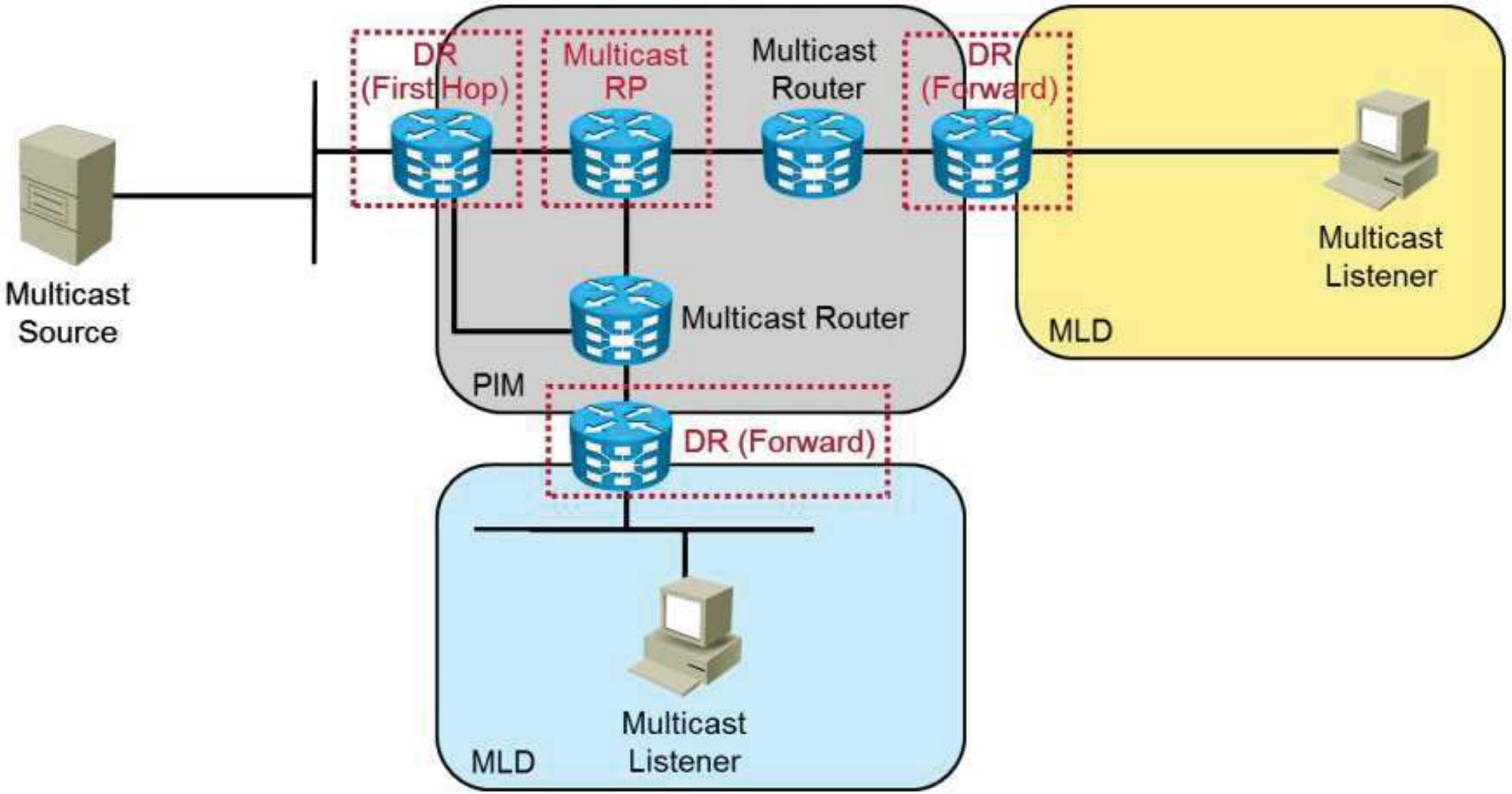
Used by applications transmitting data to a group with active listeners over the IPv6 Internet.

IPv4 and IPv6 Multicast Comparison

IP Service	IPv4 Solution	IPv6 Solution
Addressing Range	32-bit, Class D	128-bit (112-bit group)
Routing	Protocol independent, all IGPs and MP-BGP	Protocol independent, all IGPs and MP-BGP with IPv6 multicast address family
Forwarding	PIM-DM, PIM-SM, PIM-SSM, BIDIR-PIM	PIM-SM, PIM-SSM, BIDIR-PIM
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

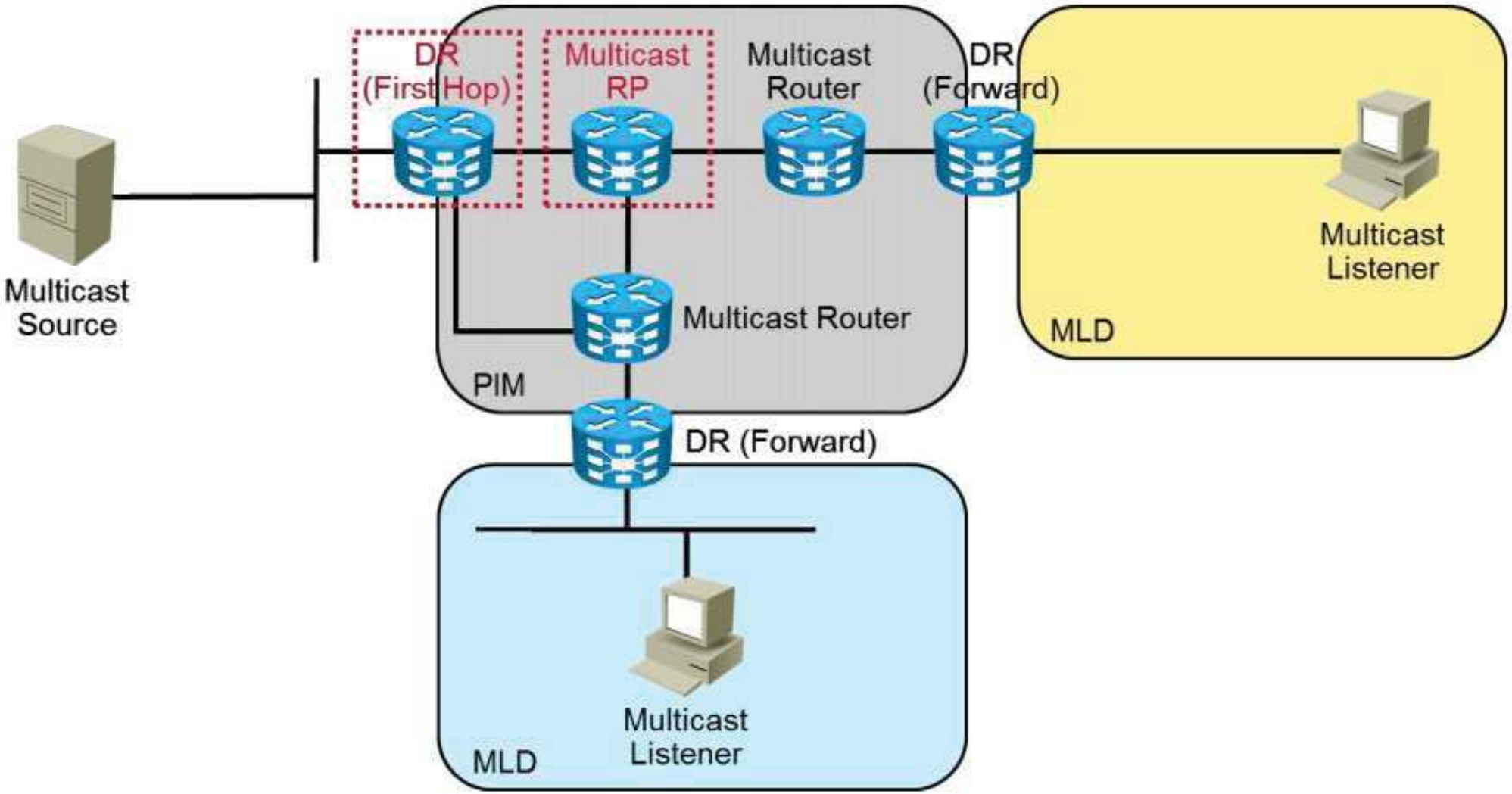
IGP = interior gateway protocol; BIDIR-PIM = bidirectional PIM; MSDP = Multicast Source Discovery Protocol

PIMv6 Overview



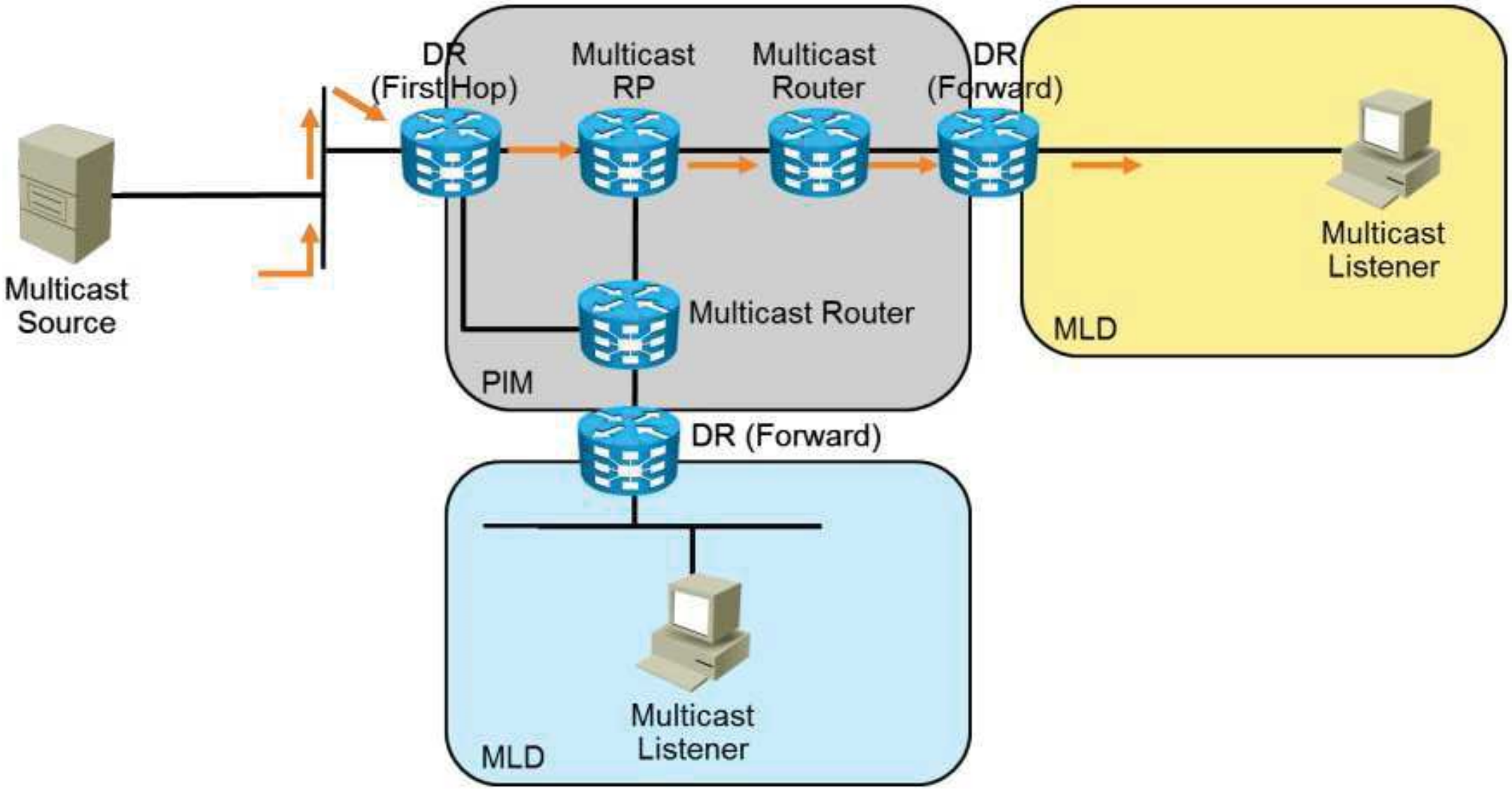
- Protocol independent: MRIB is built on underlying unicast routing table.
- DR and RP routers discover optimal paths to server group traffic.

PIMv6 Phase One Process



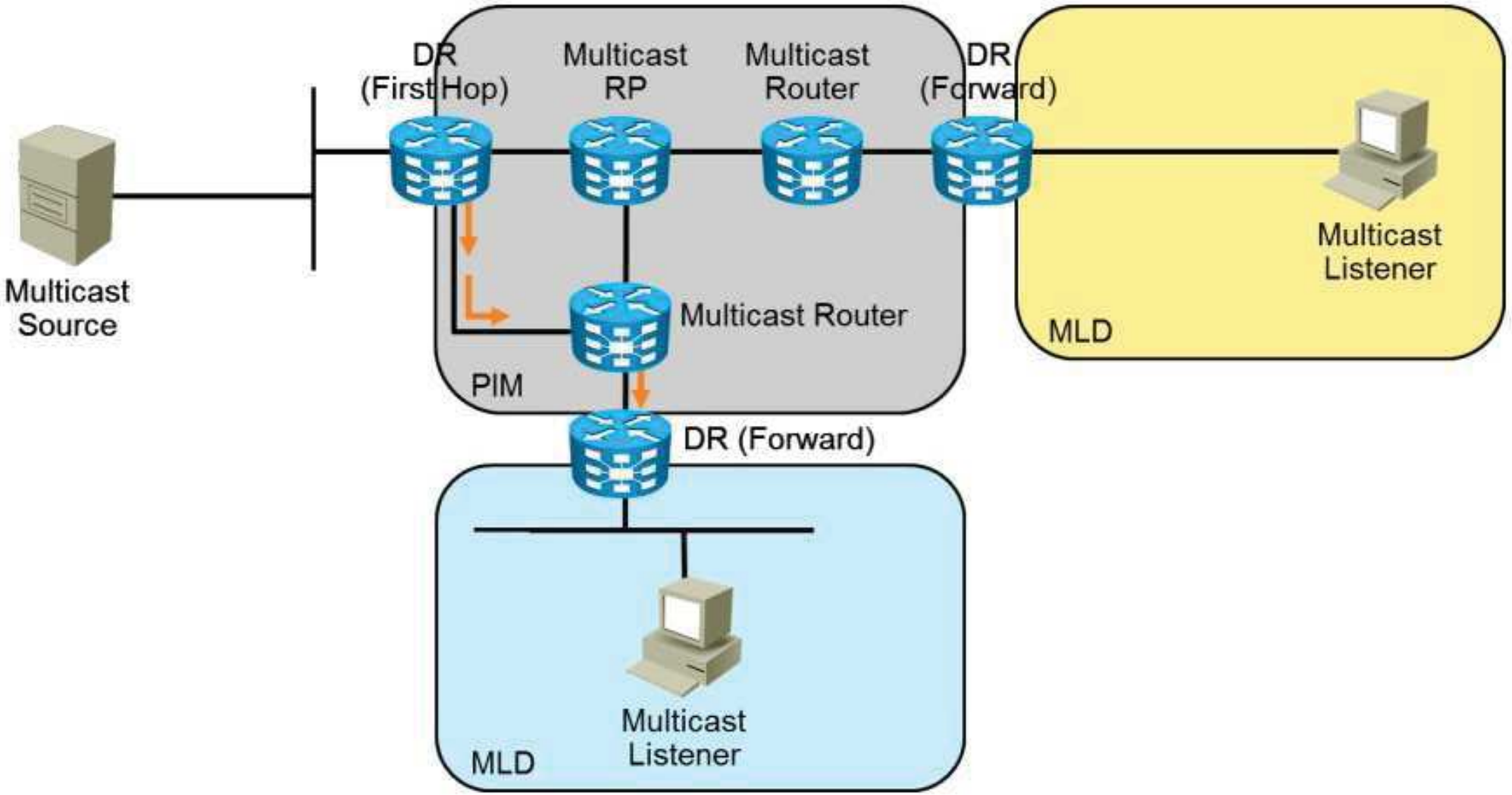
- DR (first-hop) router is the first router after source.
- RP routers consolidate groups from different sources.

PIMv6 Phase Two Process



- DR (first-hop) router is the first router after source.
- RP routers consolidate groups from different sources.

PIMv6 Phase Three Process



- The path through RP may not be optimal.
- The transfer from a shared tree to a source-specific shortest path tree is for lower latency and (possibly) better bandwidth utilization.

Embedding the RP Address in an IPv6 Multicast Address

Embedding the RP address in an IPv6 multicast address characteristics:

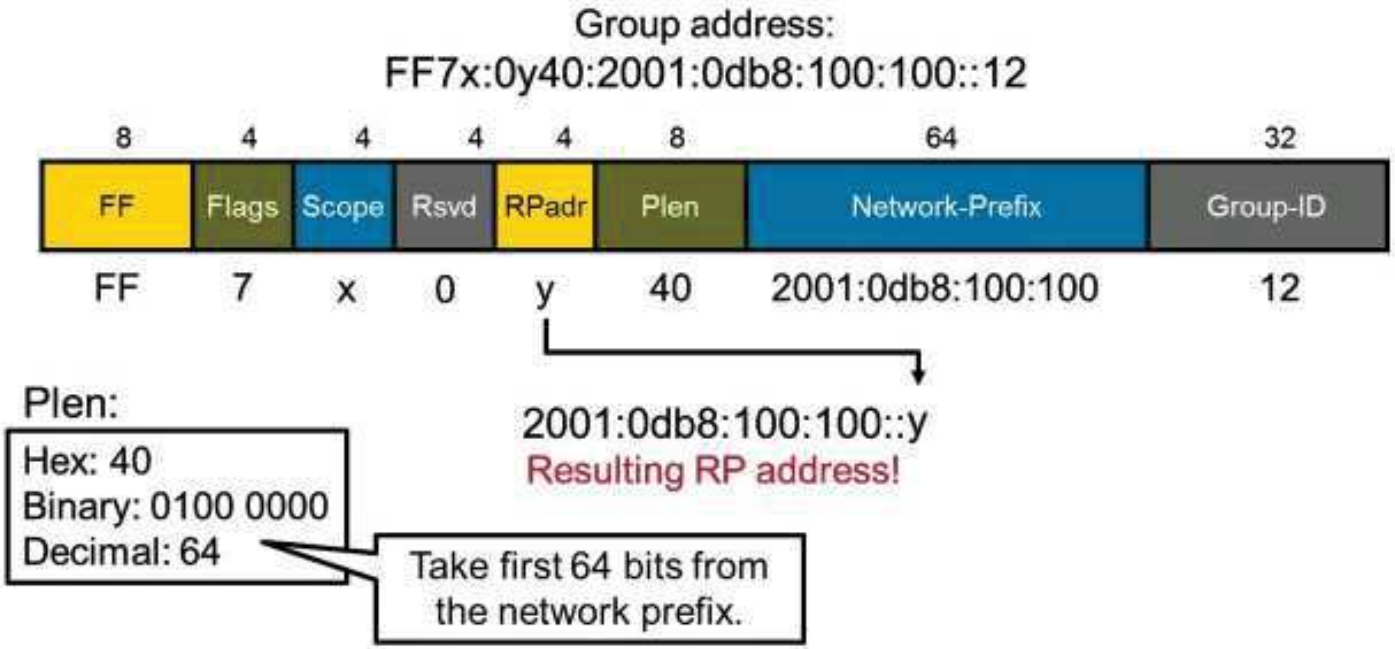
- The RP address is embedded into a multicast group address.
- The embedded RP redefines what was an 8-bit reserved field into a 4-bit reserved field and a 4-bit RP field:
  - The RP field allows the provision of 16 RPs on an embedded address.
  - A 32-bit group ID field provides for 232 multicast groups per RP.
- Plen specifies which part of the network prefix from the group address should be used for the RP address.

Flags = 0RPT, R = 1, P = 1, T = 1 = RP Address Embedded

Embedding the RP Address in an IPv6 Multicast Address (Cont.)

Embedded Rendezvous Points example:

- 16 RP addresses per network prefix.
- 232 multicast groups per RP.
- Guaranteed to be unique because enterprise-assigned /64 network used in address.

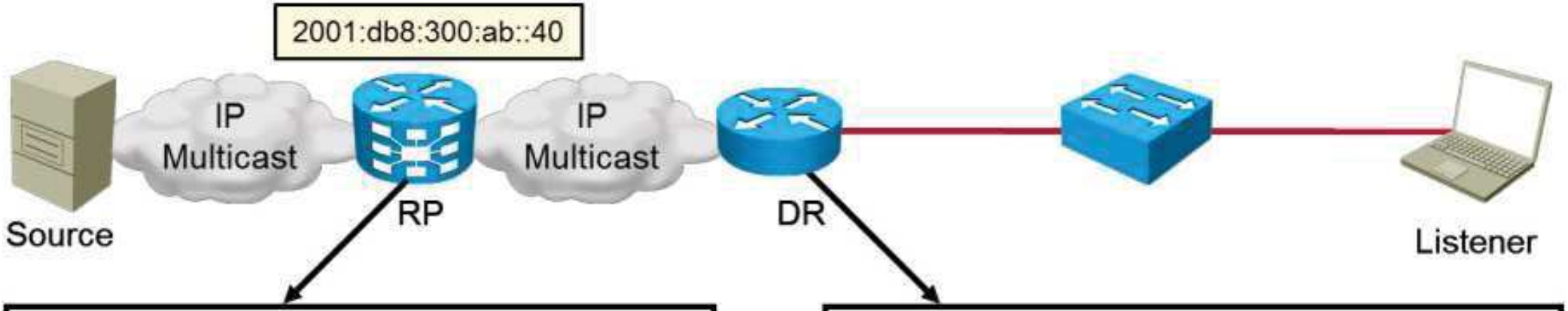


Embedding the RP Address in an IPV6 Multicast Address (Cont.)

Use of Embedded RP:

- Embedded RP can be considered an automatic replacement to static RP configuration.
- Routers that do not support embedded RPs can be configured statically or via BSR.
- Embedded RP does not provide RP redundancy as BSR or anycast RP can.

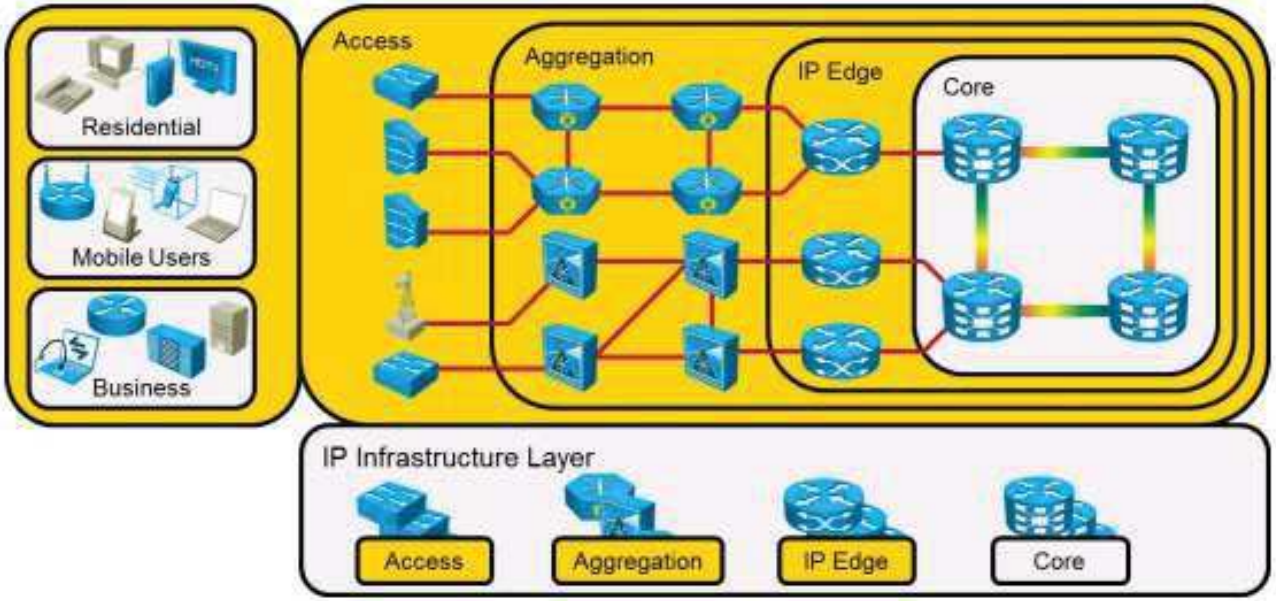
IPv6 Multicast Routing Configuration




```
multicast-routing
 address-family ipv6
  interface all enable
!
router pim
 address-family ipv6
  rp-address 2001:db8:300:ab::40
  interface GigabitEthernet0/0/0/0
   enable
  !
  interface GigabitEthernet0/0/0/1
   enable
!
interface Loopback0
 ipv6 address 2001:db8:300:ab::40/64
```

```
ipv6 multicast-routing
!
ipv6 pim rp-address 2001:db8:300:ab::40
```

IPv6 Multicast Listener Discovery

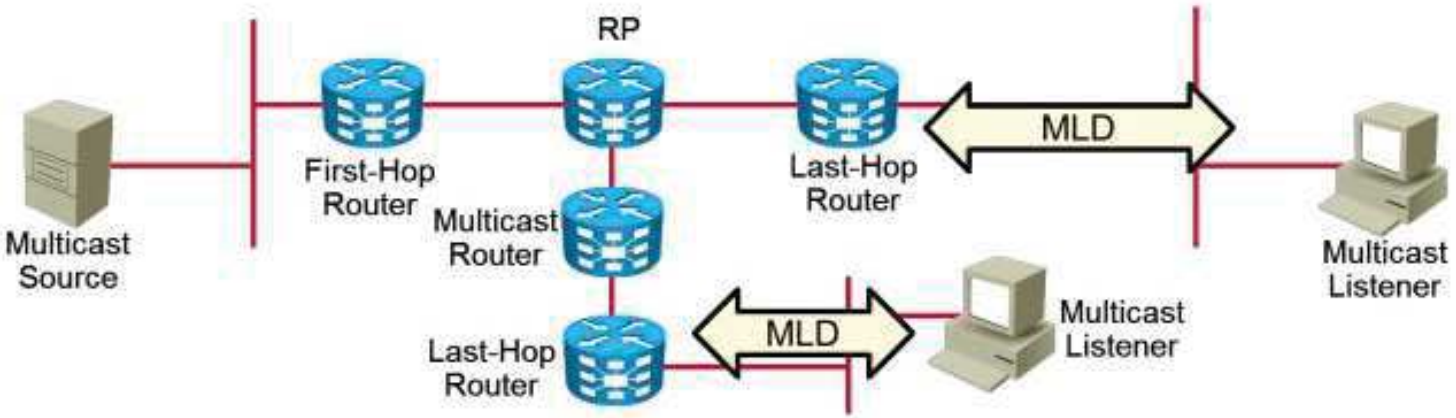


- MLD is used on the IP infrastructure layer of the Cisco IP NGN.
- MLD is used between customer and service provider edge devices.
- MLD snooping is used on access and aggregation devices.

IPv6 Multicast Listener Discovery (Cont.)

IPv6 MLD characteristics:

- To implement multicast, routers should know which users should receive the multicast traffic.
- IPv4: IGMP
- IPv6: MLD
  - MLDv1 is similar to IGMPv2
  - MLDv2 is similar to IGMPv3
- MLD handles join and leave processes on the access segment between the listener and the first multicast router.



MLDv1 Messages

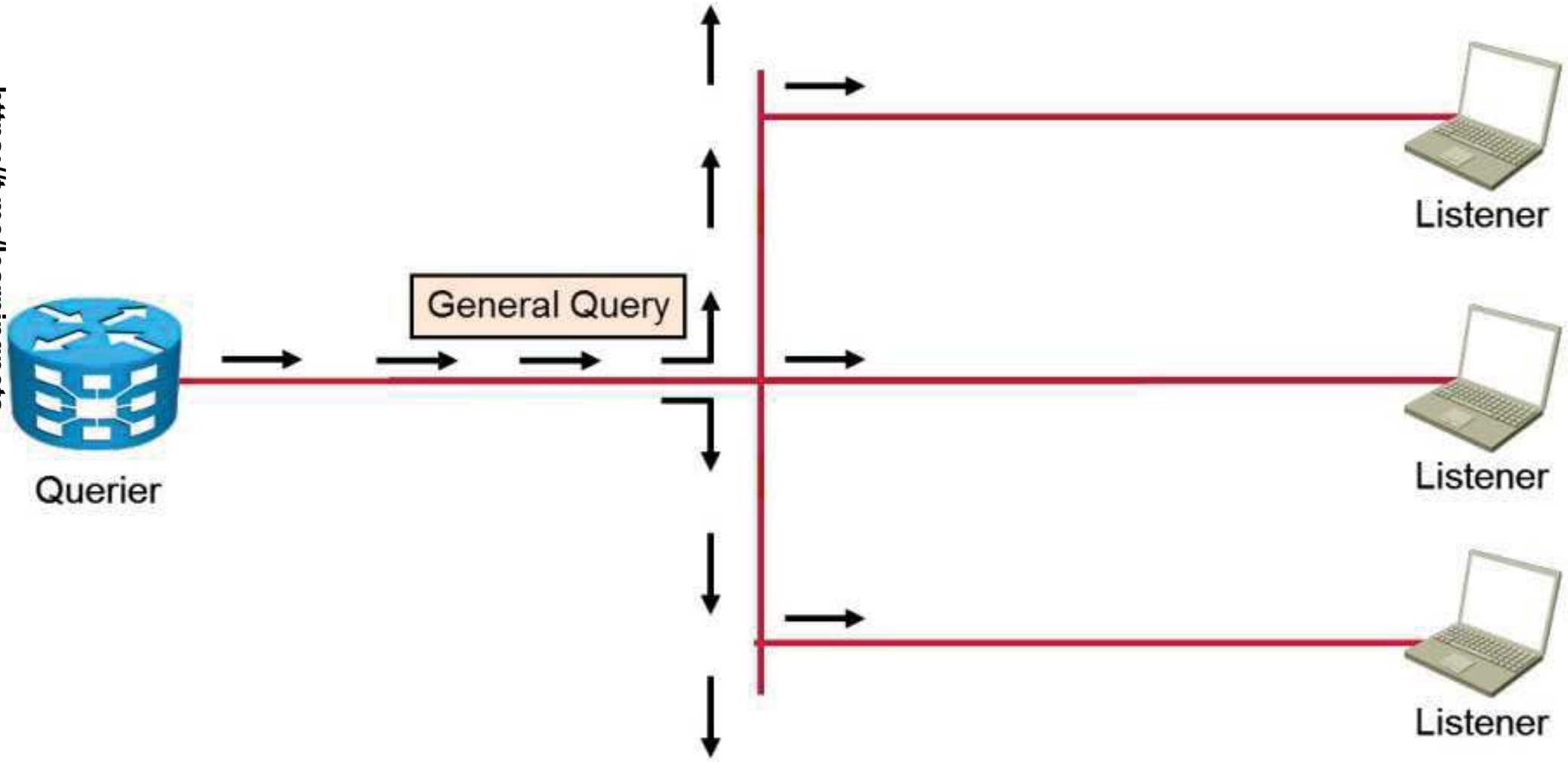
MLDv1 messages characteristics:

- MLDv1 uses three types of messages:
  - Query:
    - General
    - Group-specific
    - Multicast-address-specific
  - Report
  - Done

MLDv1 General Query Message

General query message characteristics:

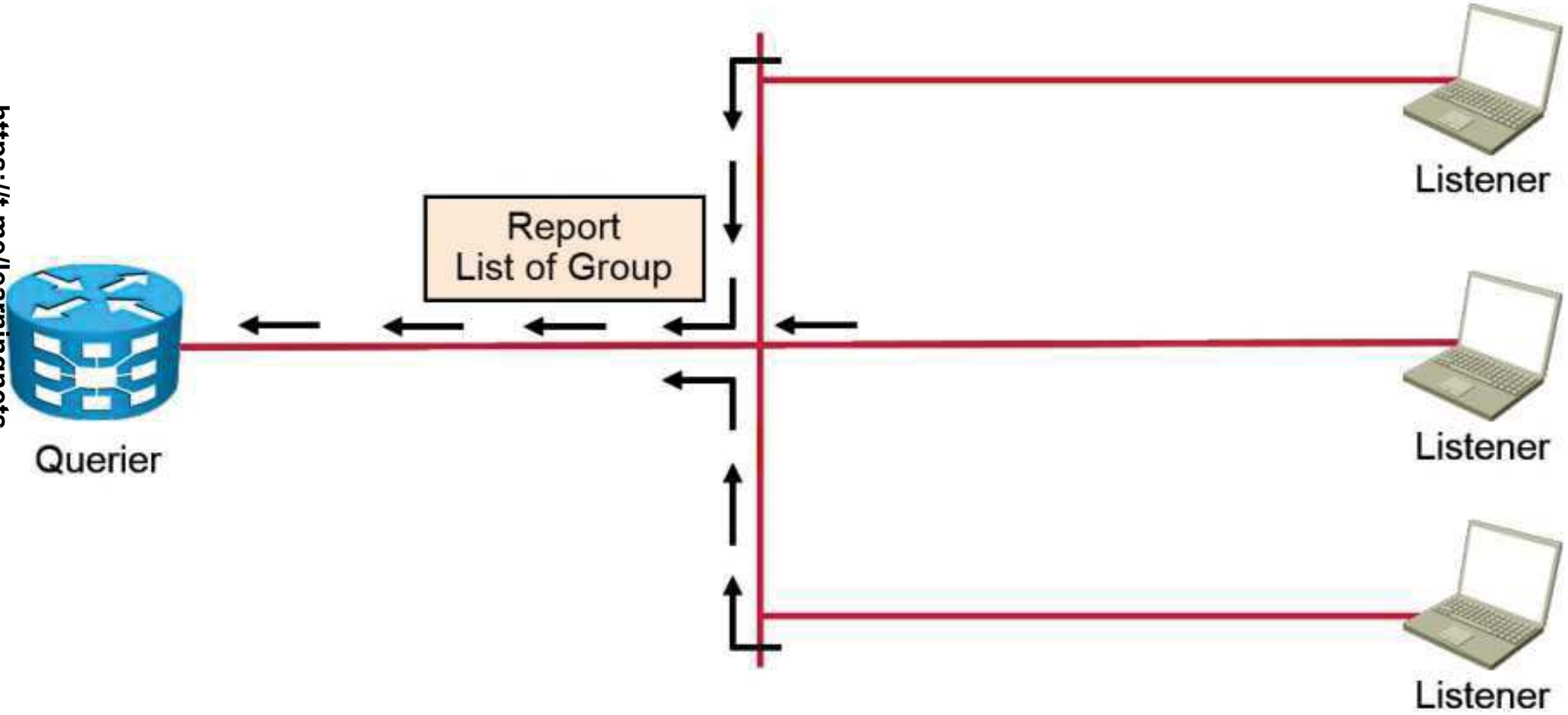
- "Which multicast addresses have listeners on the link?"



MLDv1 Report Message

Report message characteristics:

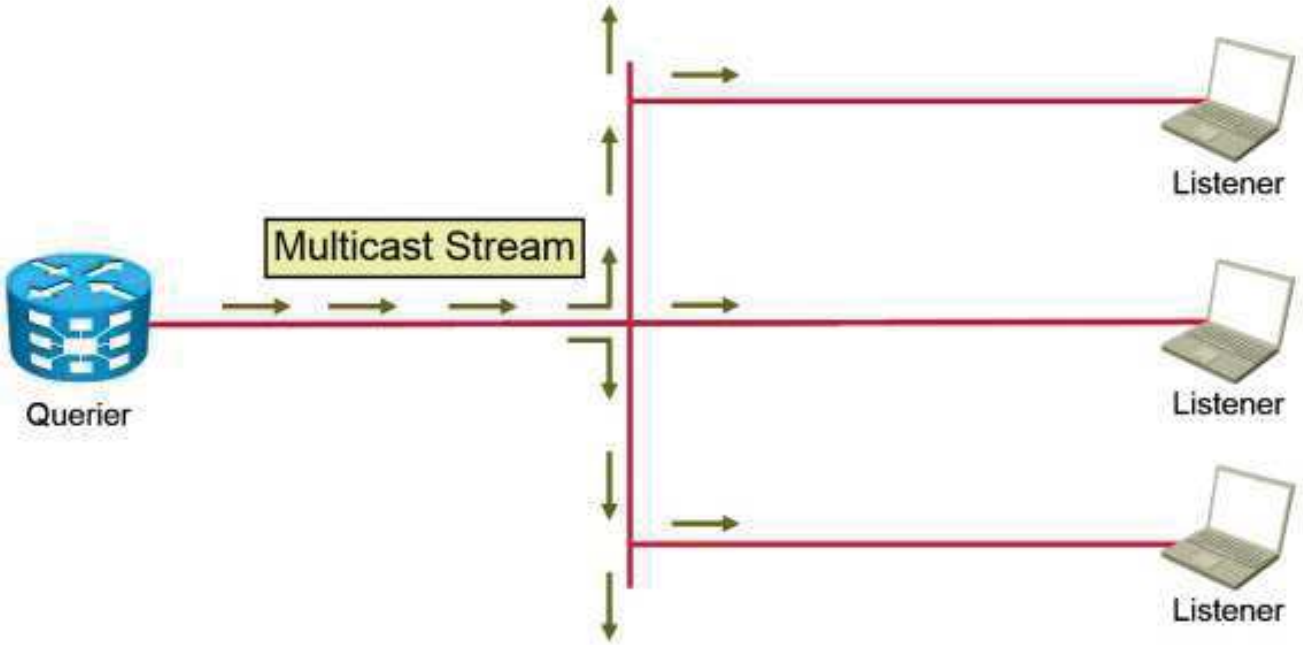
- Nodes reply with the list of multicast groups they are receiving.



MLDv1 Report Message (Cont.)

General report MLDv1 process:

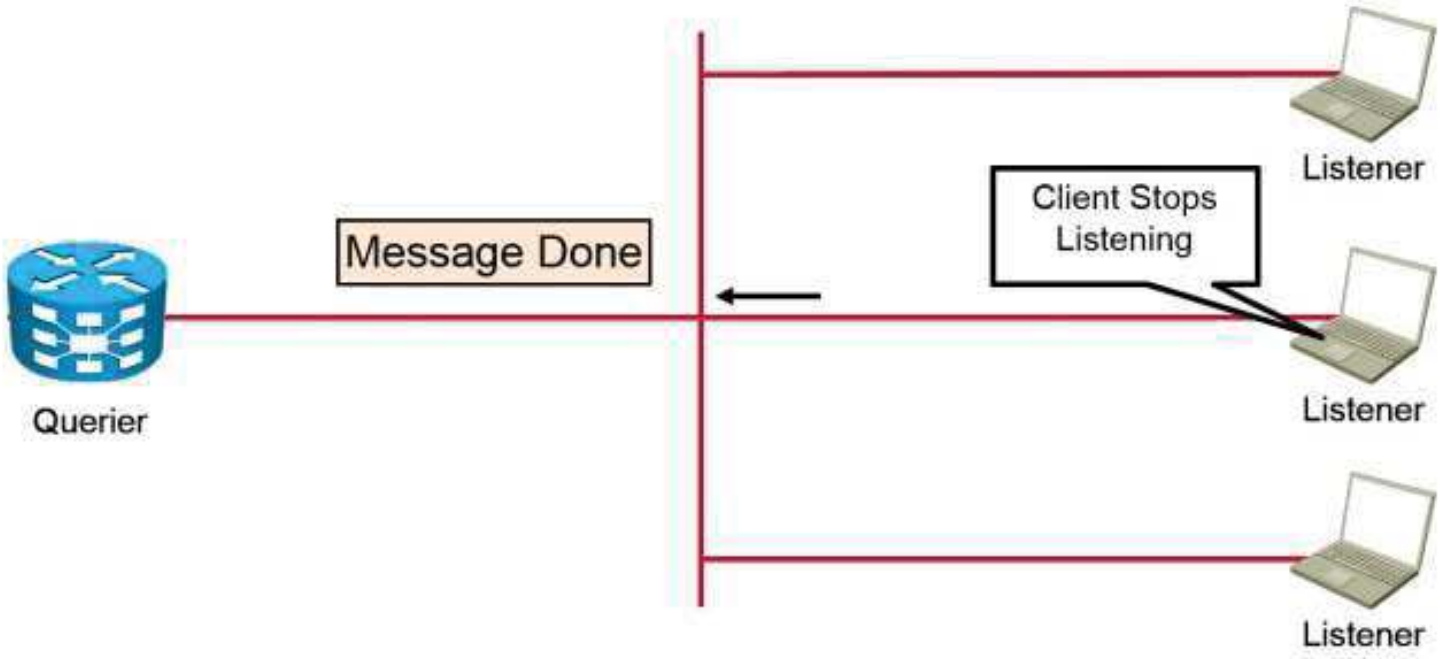
- In response to a MLD report, the router procures itself the traffic for a multicast group (for example, using a PIM join message).
- The router starts transmitting the group data to the listener on the link.



MLDv1 Done Message

General done MLDv1 process:

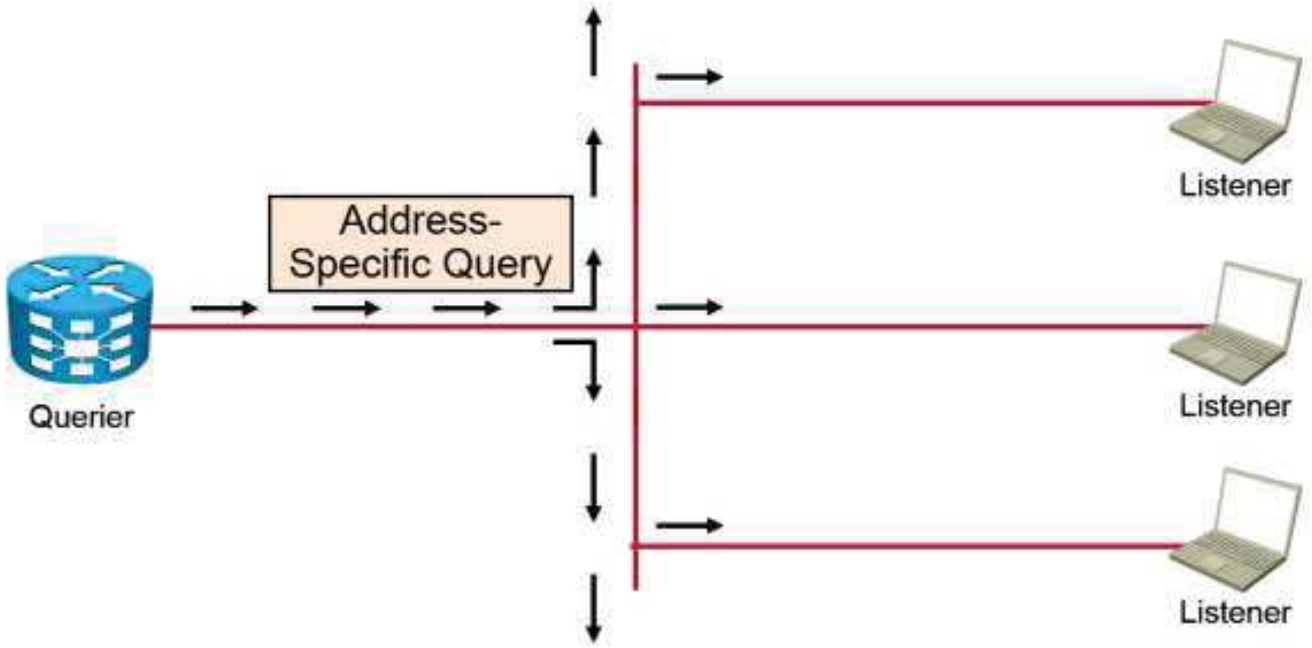
- A done message is a node leaving a group.
- If there are other known listeners for this group, this message might get suppressed.



MLDv1 Address-Specific Query Message

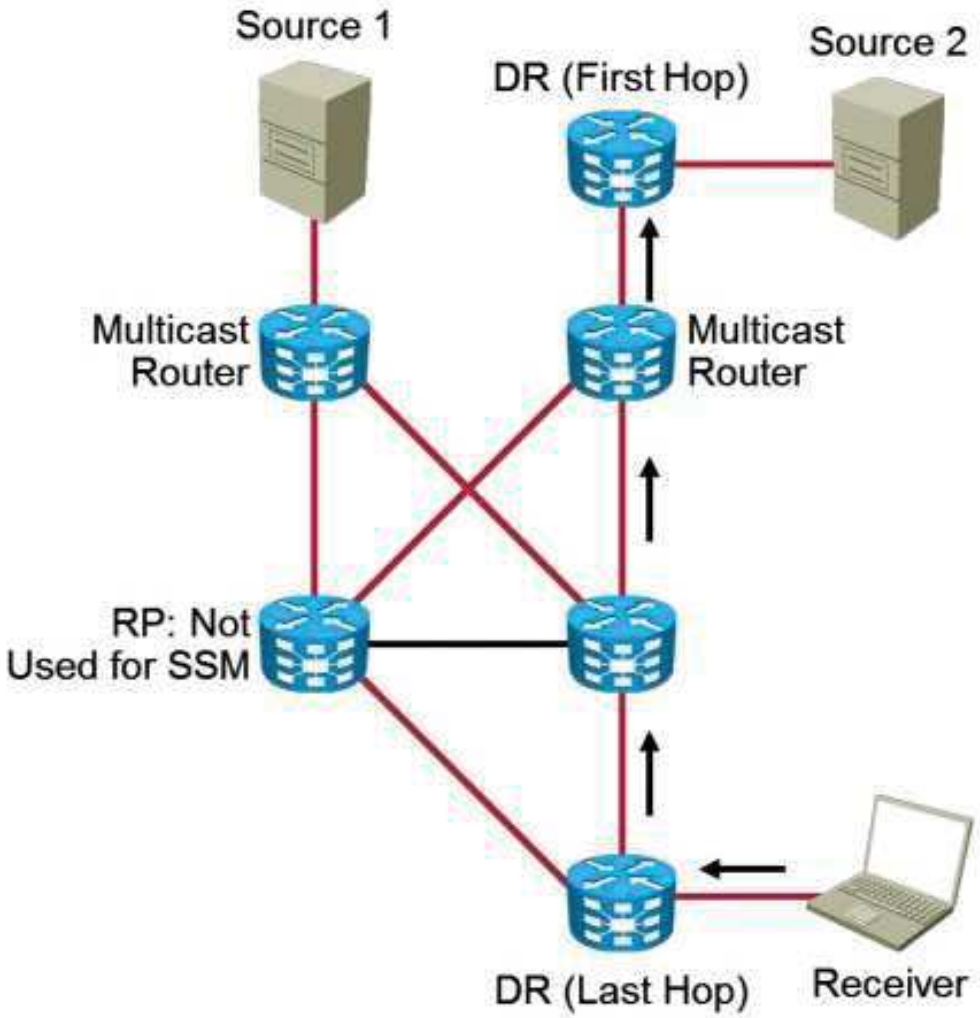
General multicast address specific query MLDv1 process:

- The querier sends an address-specific query message.
- In response to an MLD leave, the router asks if there are any others still interested in the group.



MLDv2 Protocol

- Support for source filtering (a node can select a source for a specific group)
- Interoperable with MLDv1.
- Analogous functionality to IGMPv3 for IPv4.
- Defines new messages.

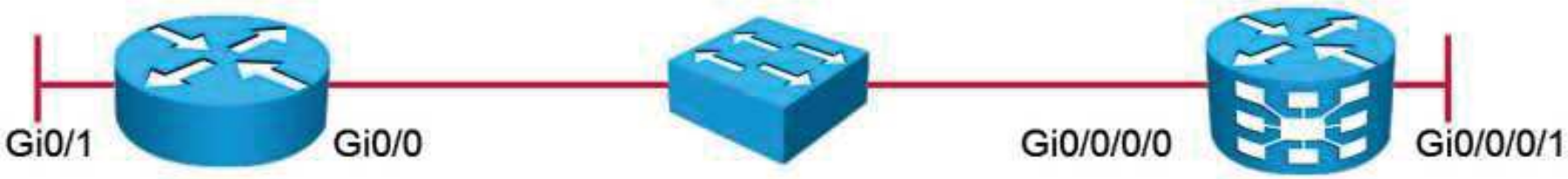


MLD Access Groups and Group Limits

MLD access groups and group limits characteristics:

- Access groups enable access control for multicast group receivers on a MLD router:
  - Limits the list of groups that a receiver can join.
  - Denies or allows sources that are used to join SSM channels.
- Group limits control a number of groups on an interface and can be implemented:
  - Globally
  - Per interface
- Membership reports exceeding the limit are ignored.

MLD Configuration



Enable IPv6 multicast routing.

```
ipv6 multicast-routing
interface GigabitEthernet0/0
 ipv6 mld query-interval 60
```

```
multicast-routing
 address-family ipv6
  interface all enable
router mld
 interface GigabitEthernet0/0/0/0
  router enable
  query-interval 60
  version 2
```

MLD Verification

- Displays MLD interface settings.

```
RP/0/RSP0/CPU0:PE7# show mld interface GigabitEthernet 0/0/0/0
GigabitEthernet0/0/0/0 is up, line protocol is up
  Internet address is fe80::4255:39ff:fe2f:40a8
  MLD is enabled on interface
  Current MLD version is 2
  MLD query interval is 125 seconds
  MLD querier timeout is 255 seconds
  MLD max query response time is 10 seconds
  Last member query response interval is 1 seconds
  MLD activity: 5 joins, 0 leaves
  MLD querying router is fe80::4255:39ff:fe2f:40a8 (this system)
```

- Displays MLD groups that are registered on the router.

```
RP/0/RSP0/CPU0:PE7# show mld groups
MLD Connected Group Membership
GigabitEthernet0/0/0/0
Group Address   : ff02::2
Last Reporter   : fe80::b000:ff:fe00:fb00
       Uptime   : 01:56:39
      Expires   : never
```

MLD Join-Group and Static-Group

```
interface GigabitEthernet0/1
 ipv6 mld join-group group-address
 ipv6 mld static-group group-address
```

```
router mld
 interface GigabitEthernet0/0/0/0
  join-group group-address
  static-group group-address
```

join-group command:

- Router joins multicast group.
- Router populates MLD cache.
- Router sends MLD report.
- Results in:
  - Router joining a group.
  - CPU receives data.

static-group command:

- Configured on the router to forward traffic on the interface.
- Populates MLD cache.
- Results in:
  - PIM join only if configured on the designated router.
  - No CPU impact.

MLD Snooping

MLD snooping characteristics:

- The default behavior for multicast switched traffic is to flood it to all Layer 2 interfaces (in the same VLAN).
- A high volume of multicast traffic loads the switch data plane and control plane.
- Snooping of MLD packets is used to determine and store the information that switch ports receive data for multicast groups.
- Layer 2 ports that do not have any listeners do not transmit multicast traffic.
- If a multicast group has only sources and no receivers in a VLAN, MLD snooping constrains the multicast traffic to only the multicast router ports.

MLD Snooping (Cont.)

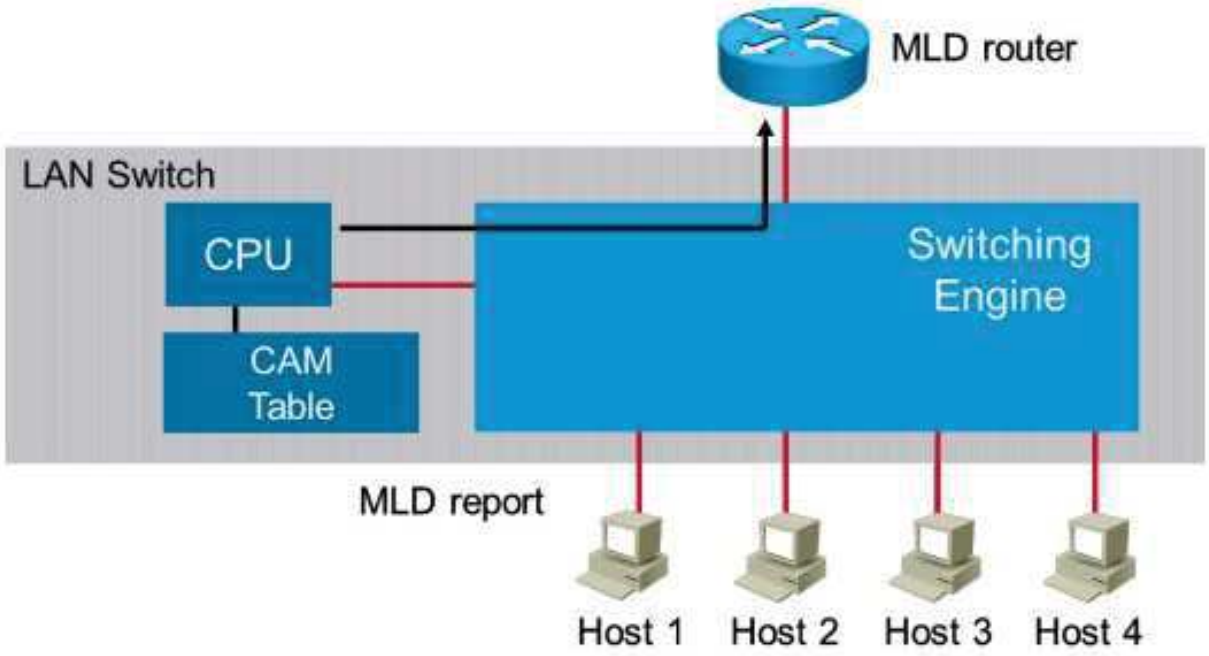
MLD snooping characteristics:

- MLD snooping is available on switches only.
- MLD snooping is available for MLDv1 and MLDv2, depending on the switch.
- MLD snooping monitors Layer 3 MLD traffic.
- MLD snooping proxy reporting: The first report to the MLD router is sent, while all other reports for the same group are suppressed.

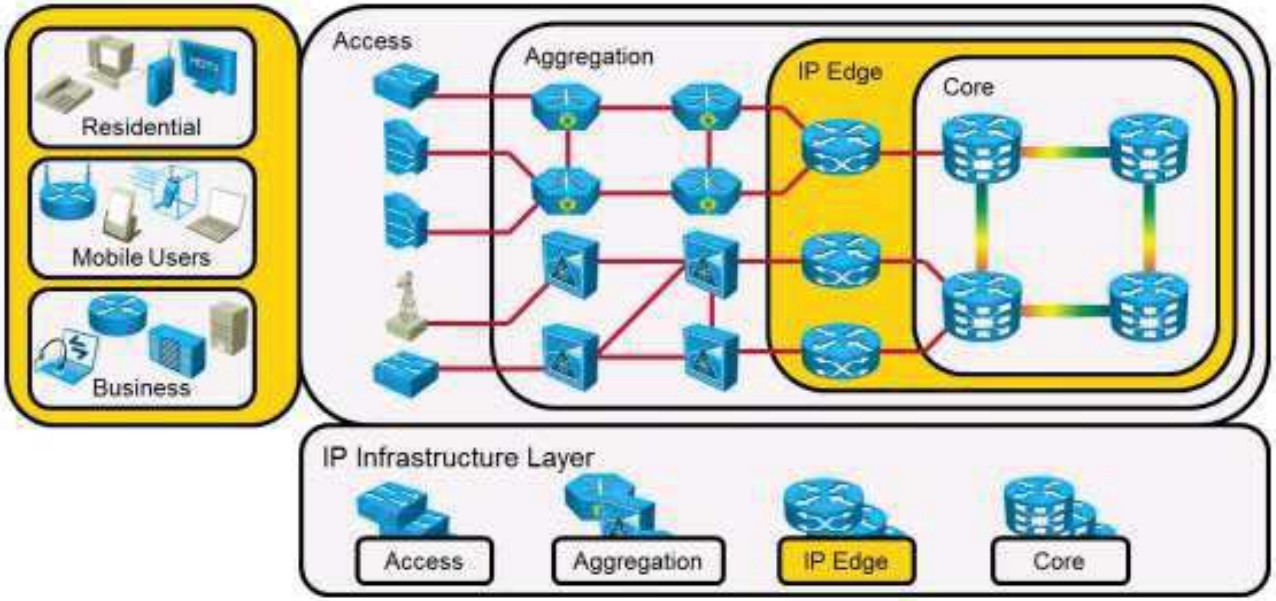
MLD Snooping (Cont.)

MLD snooping group join characteristics:

- Hosts join a multicast group.
- Switch creates a Layer 2 FIB entry for the multicast group in this VLAN.
- When additional hosts want to join the group, the switch adds its port to the Layer 2 FIB for the group.



DNS and DHCPv6 in the Cisco IP NGN Infrastructure Layer



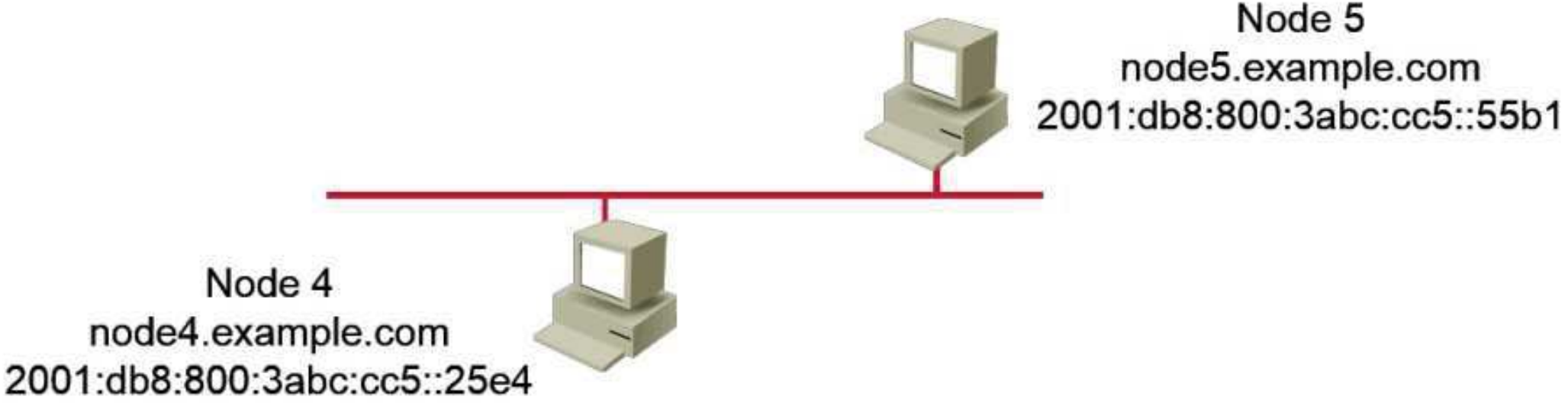
- DNS and DHCP are used on the services layer of the Cisco IP NGN.
- DNS is used on customer end devices.
- DHCP is used between customer and service provider IP edge devices.

DNS IPv6 Support

DNS Supported Objects:

- Two DNS issues exist for IPv6:
  - IPv6 record support.
  - IPv6 transport support.
- Several types of DNS objects exist:
  - AAAA, A, PTR, MX, etc.
- Forward lookups:
  - DNS uses AAAA records for forward IPv6 lookups.
  - PTR records are used for reverse lookups.

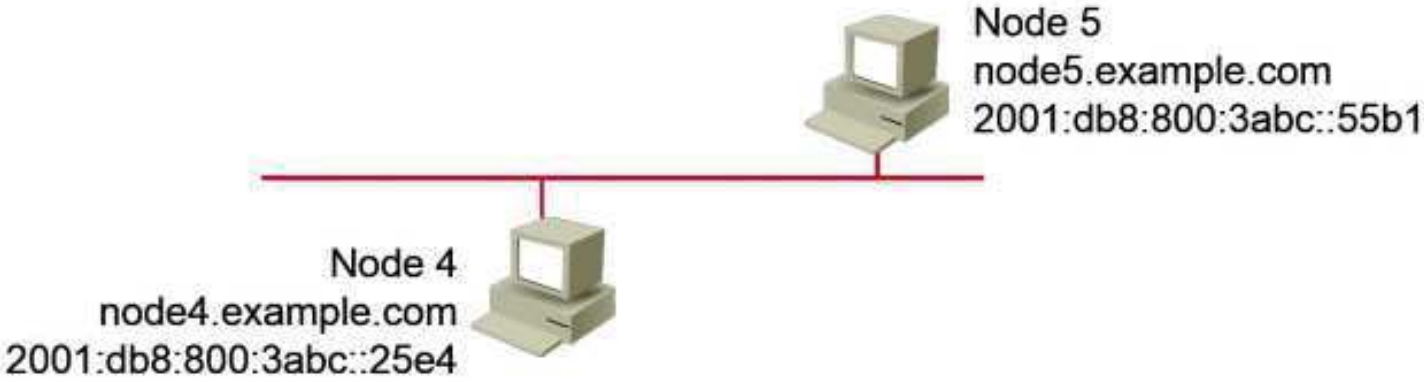
DNS IPv6 Support (Cont.)



Examples of AAAA and A records:

```
node5.example.com.     IN  AAAA     2001:db8:800:3abc:cc5::55b1
node5.example.com.     IN    A          193.77.119.33
```

DNS IPv6 Support (Cont.)



Reverse lookups:

- IPv6 uses PTRs for reverse lookups, similar to IPv4, but with the new nibble format.
- Examples of Nibble-Formatted Records:

```
$ORIGIN c.b.a.3.0.0.8.0.8.b.d.0.1.0.0.2.ip6.arpa.
4.e.5.2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0 14400 IN PTR node4.example.com.
1.b.5.5.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0 14400 IN PTR node5.example.com.
```

DNS IPv6 Support (Cont.)

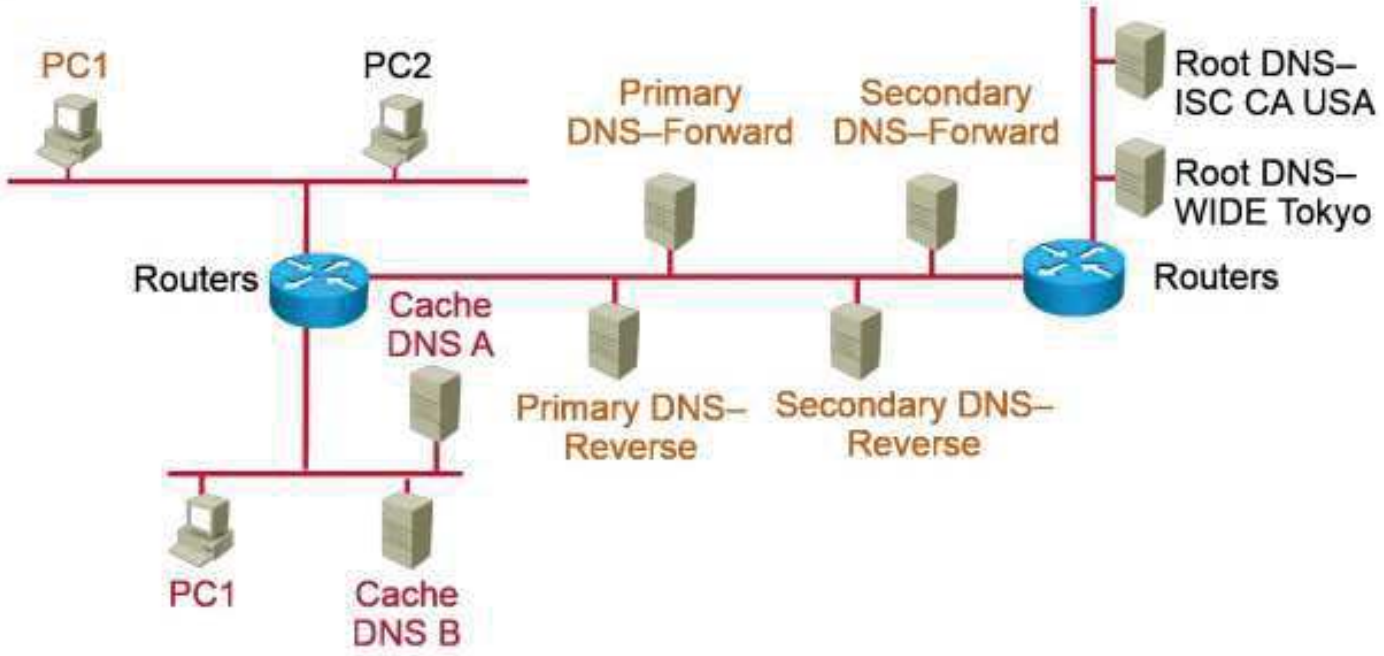
DNS tree structure characteristics:

- IPv6 needs an updated version of a DNS server and client resolver.
- DNS tree structure is identical to IPv4:
  - Root DNS server.
  - Top-level domain DNS server.
  - Authoritative DNS server for each particular domain.
- From the operational perspective, there are:
  - Authoritative DNS servers.
  - Caching DNS servers.
- The majority of DNS root servers are accessible using IPv6, many since 2008:
  - Enabled, end-to-end IPv6 communication without using IPv4 for communication with the root DNS server.
  - Removed the need for dual stack (from DNS perspective).

DNS IPv6 Support (Cont.)

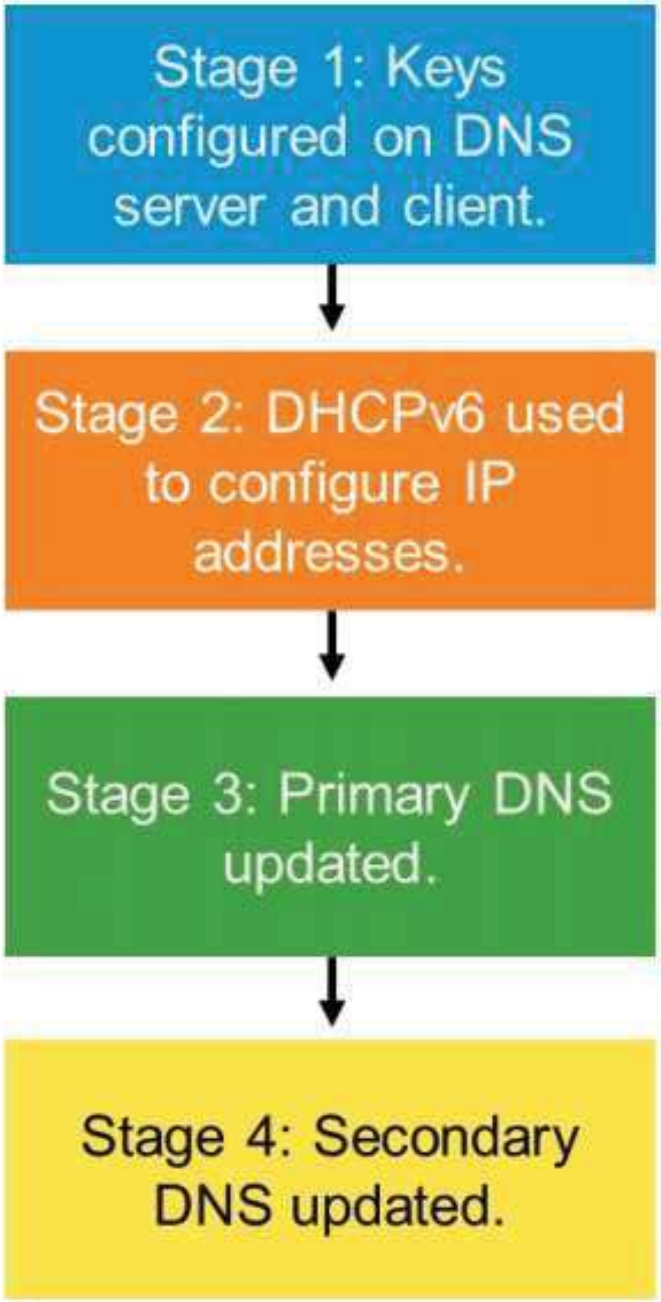
DNS tree structure components characteristics:

- Authoritative primary and secondary DNS servers support both IPv6 and IPv4 records:
  - Forward and reverse zones are not often on the same system.
  - Reverse zones are often maintained by an ISP.
- Caching DNS is typically provided by ISPs (for home or small business) or by large enterprises for in-house clients.



Dynamic DNS

- Dynamic DNS allows IPv6 clients to update resource records in their authoritative DNS server.
- Updates should be authenticated to prevent domain hijackings, man-in-the-middle attacks, and so on.
- There are two types of DDNS implementations:
  - Client-based implementations for endpoints, routers, and so on.
  - DHCPv6-based implementations.



DHCPv6 Operations

DHCPv6 operates the same as IPv4, with these exceptions:

- The client first detects the presence of routers on the link.
- If found, the client examines router advertisements to determine if DHCP can be used.
- If no router is found, or if DHCP can be used, the client:
  - Sends a DHCP solicit message to the all-DHCP-agents multicast address.
  - Uses the link-local address as the source address.

DHCPv6 Operations (Cont.)

DHCPv6 operates using the following multicast adresses:

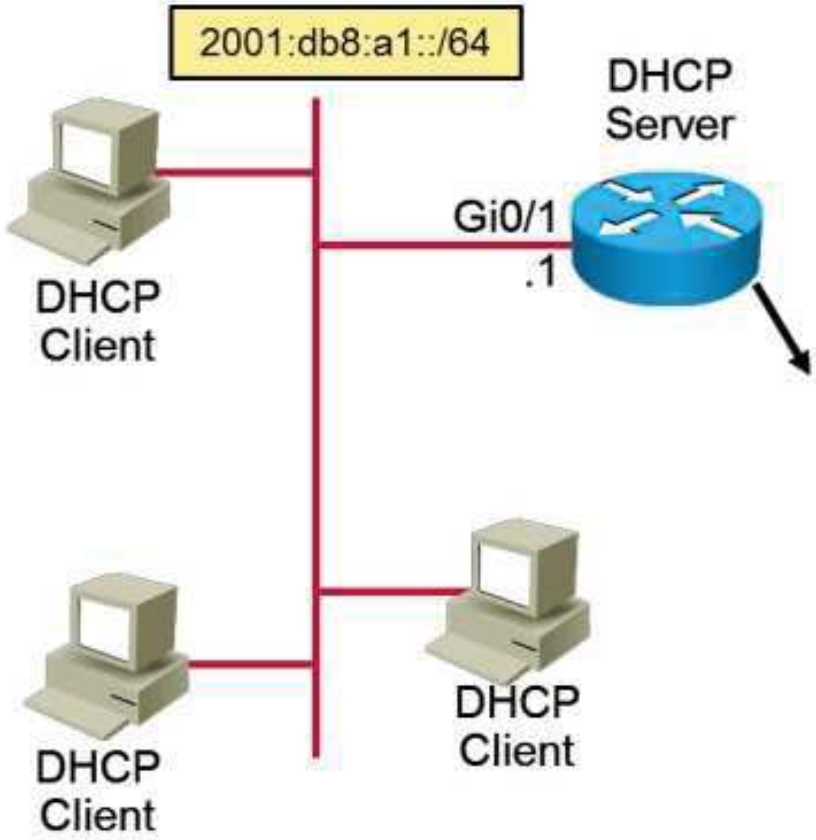
IPv6 Multicast Address	Description
FF02::1:2	All-DHCP-agents (servers or relays), link-local scope
FF05::1:3	All-DHCP-servers, site-local scope
FF05::1:4	All-DHCP-relays, site-local scope

DHCPv6 Server Router Configuration

DHCPv6 server router configuration characteristics:

- A router can act as a DHCP server.
- Operation is similar to IPv4 DHCP:
  - Clients get an address assigned.
  - Servers keep track of all bindings.
  - A bindings database can be uploaded to a remote server.
- Configuration options include:
  - DHCP pool name
  - Prefix information
  - Addresses for particular clients
  - List of DNS servers
  - Domain name

DHCPv6 Server Router Configuration (Cont.)




```
ipv6 dhcp pool Pool1
 address prefix 2001:db8:a1::/64
 dns-server 2001:db8:c1::53
 dns-server 2001:db8:c2::53
 domain-name example.org
!
interface GigabitEthernet 0/1
 ipv6 dhcp server Pool1
 ipv6 address 2001:db8:a1::1/64
 ipv6 nd managed-config-flag
 ipv6 nd other-config-flag
```

Configure the IPv6 DHCP pool.

Apply the address pool to the interface.

DHCPv6 Lite Operation (Stateless DHCPv6)

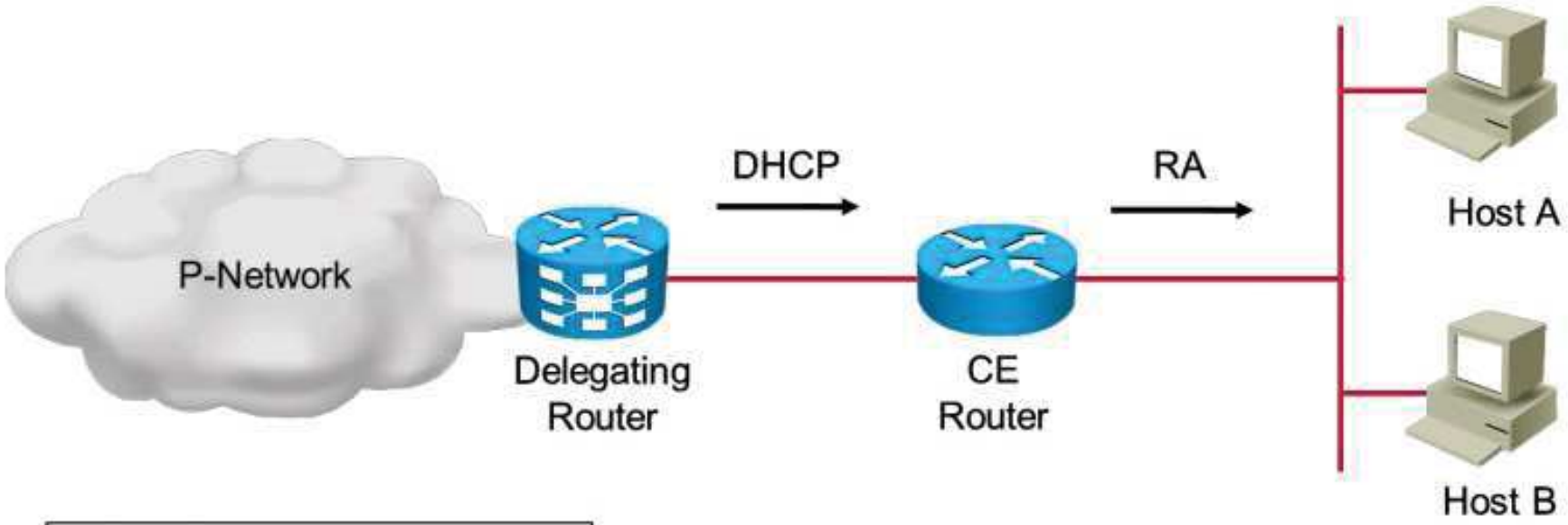
DHCPv6 Lite operation:

- DHCPv6 Lite is used for providing additional information:
  - DNS servers
  - SIP servers
  - Domain name
- DHCPv6 Lite does not perform address assignment
- Nodes need to acquire addresses through other means
- DHCPv6 Lite is configured similarly to DHCPv6:
  - Without address pool
  - Without **managed-config-flag** on an interface

DHCPv6 Prefix Delegation

DHCPv6 prefix delegation characteristics:

- A service provider allocates a block of addresses for delegation to customers.
- The customer receives a prefix (for example, /56).
- The customer assigns /64 prefixes to LAN interfaces.

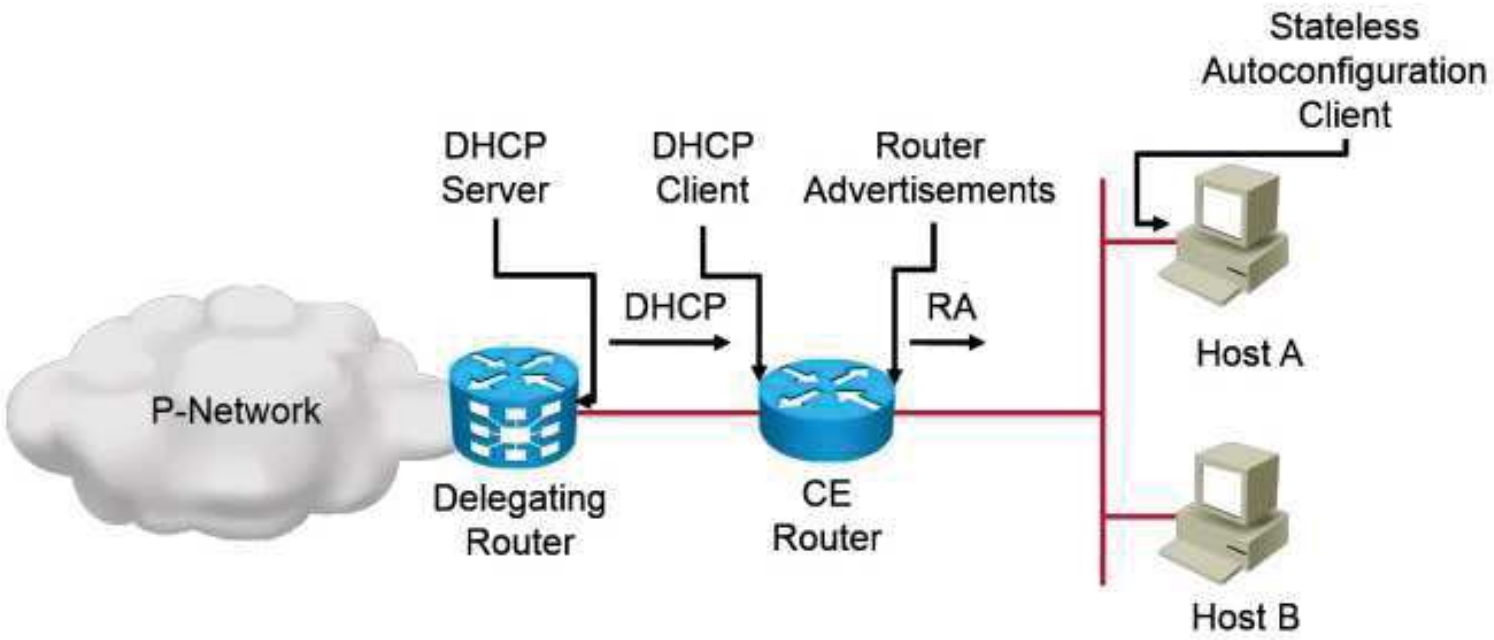


RA = Route advertisement

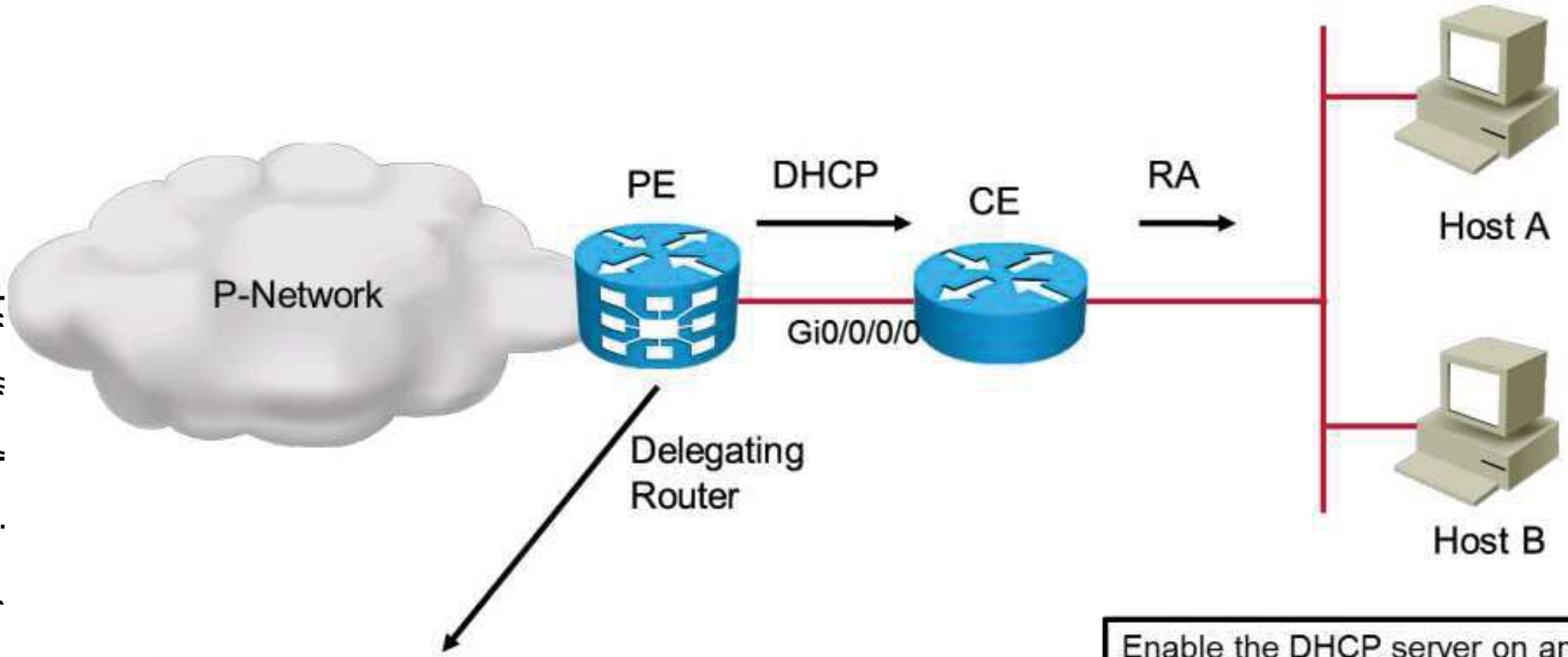
DHCPv6 Prefix Delegation (Cont.)

Interface configuration:

- PE as the delegating DHCP server.
- CE as the DHCP client and IPv6 router.



DHCPv6 Prefix Delegation (Cont.)

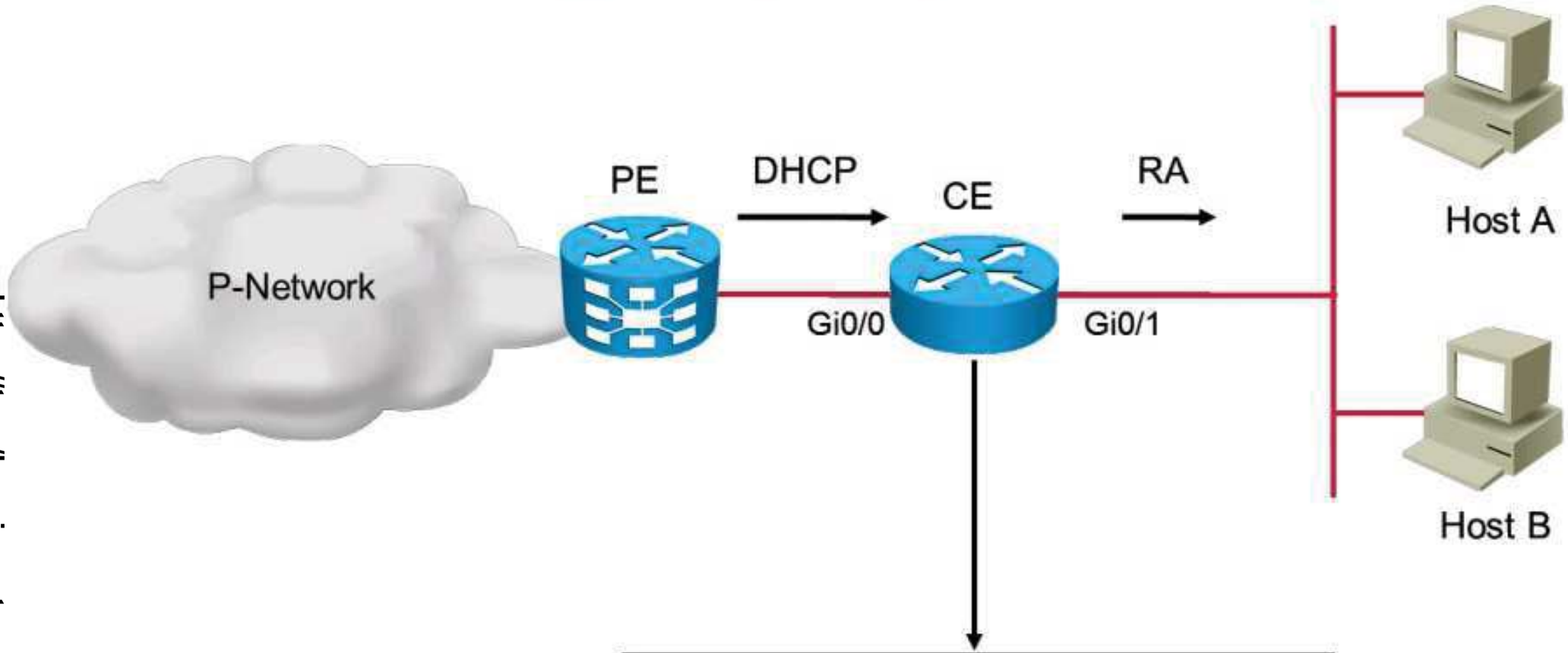



```
dhcp ipv6
 interface GigabitEthernet0/0/0/0 server
  pd 2001:db8:1::/64
!
interface GigabitEthernet0/0/0/0
 ipv6 address 2001:db8:a::1/64
```

Enable the DHCP server on an interface.

Configure the delegated prefix.

DHCPv6 Prefix Delegation (Cont.)




```
interface GigabitEthernet0/0
 ipv6 address 2001:db8:a::2/64
 ipv6 dhcp client pd PREFIX
!
interface GigabitEthernet0/1
 ipv6 address PREFIX ::1/64
```

DHCPv6 Verification

DHCPv6 verification commands:

- Shows all DHCPv6 pools on a router.

```
Router# show dhcp ipv6 pool
```

- Shows the state of all current clients of the DHCP server.

```
Router# show dhcp ipv6 binding
```

DHCPv6 Verification (Cont.)

- Displays whether an interface is in client mode or in server mode.

```
Router# show dhcp ipv6 interface
```

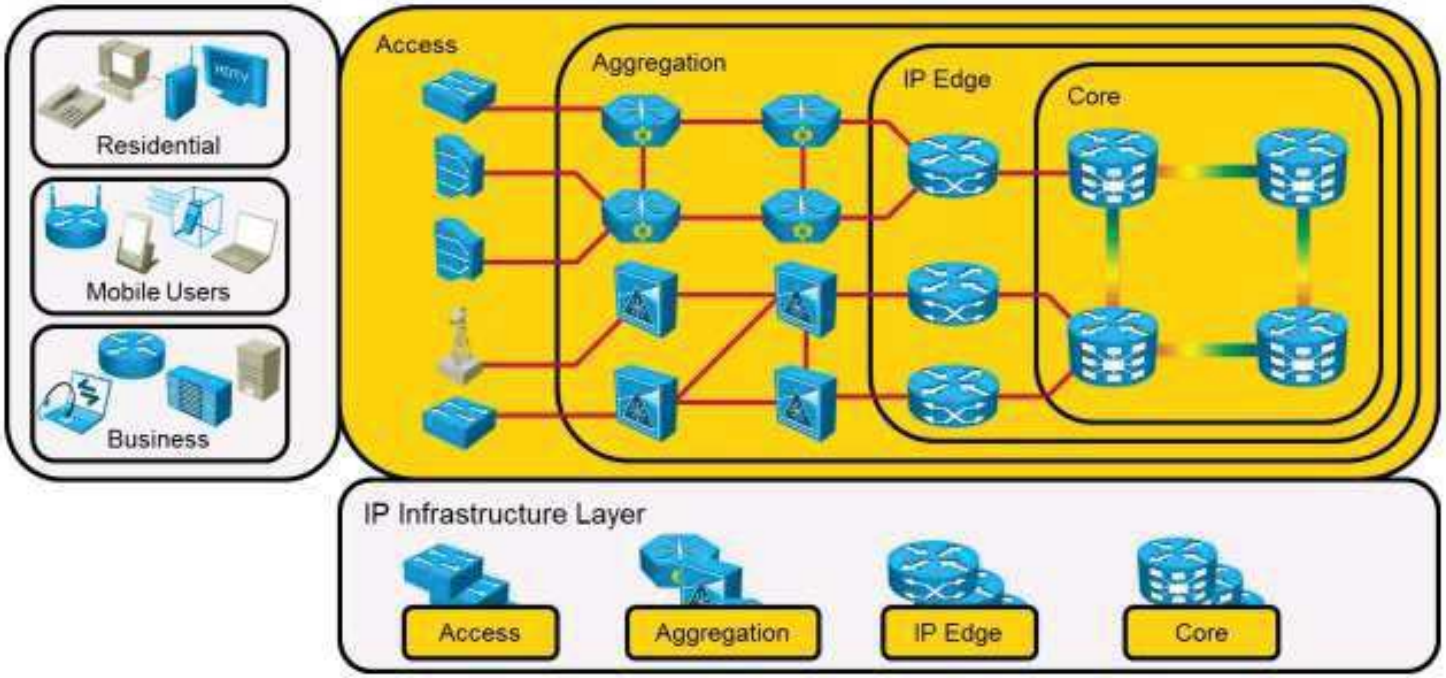
- Used when debugging either DHCP server or DHCP client functionality on a router.

```
Router# debug dhcp ipv6 detail
```

QoS in the Cisco IP NGN Infrastructure Layer

The QoS is used:

- On the IP infrastructure layer of the Cisco IP NGN.
- End-to-end QoS must be implemented to satisfy requirements for the most demanding services.



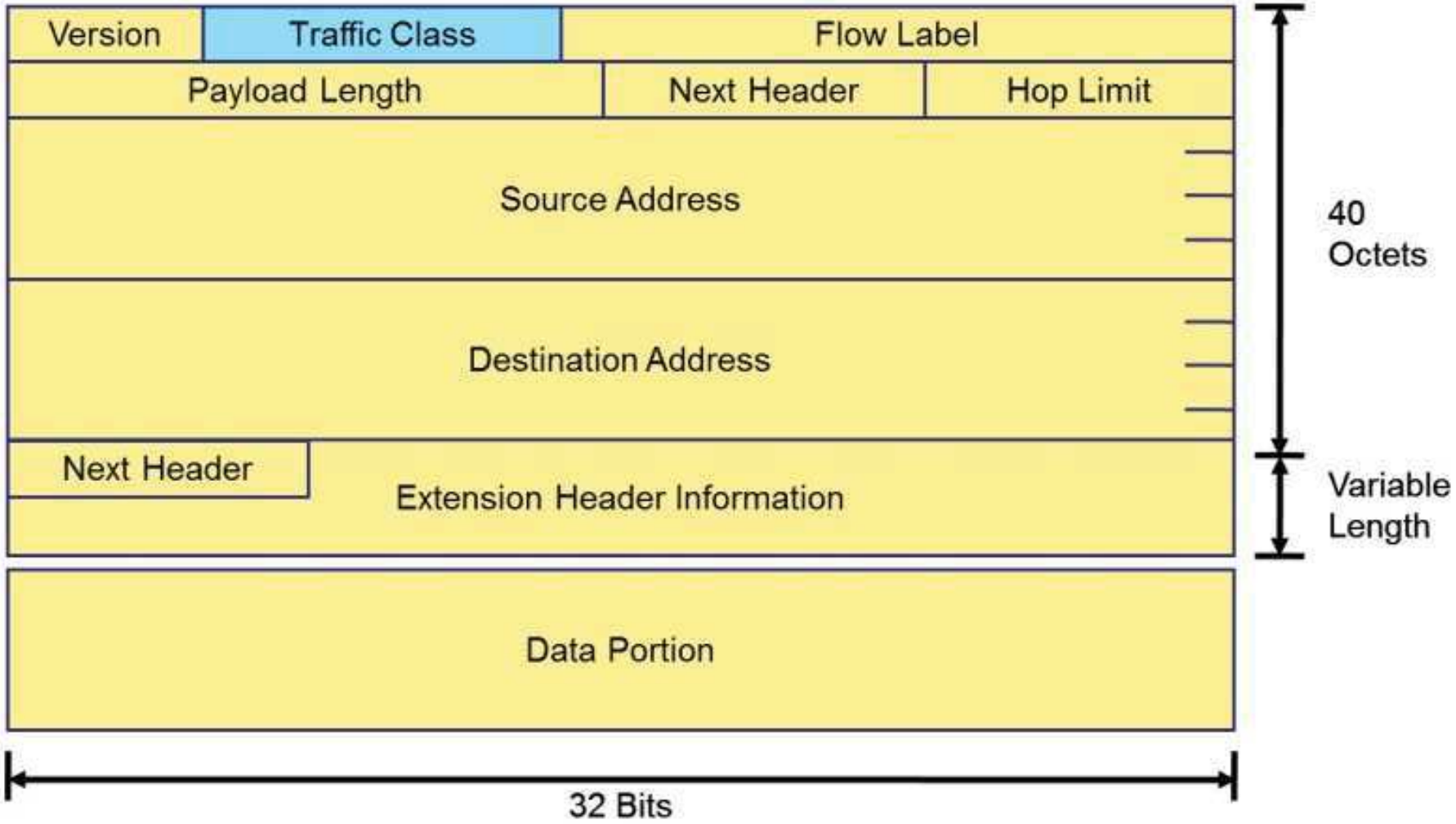
IPv6 Header Fields Used for QoS

IPv6 header fields used for QoS characteristics:

- IPv6 was designed to support QoS natively.
- Two fields in an IPv6 header enable awareness of QoS:
  - Traffic Class
  - Flow Label
- Additionally, IPv6 can be extended via extension headers to possibly support entirely new QoS mechanisms.
- QoS processing must be defined on network devices, using IntServ or DiffServ modes of operation.

IPv6 Traffic Class Field

- The Traffic Class field is the same as the IPv4 ToS field.



IPv6 Traffic Class Field (Cont.)

IPv6 traffic class field characteristics:

- The 8-bit field is identical to the IPv4 ToS field.
- A total of 6 bits are used for DSCP.
- The remaining two bits are used for ECN.
- The Traffic Class field is mutable between the source and destination nodes (which may be changed).
- The Traffic Class field is used to preserve packet QoS information end to end and also when the packet crosses Layer 2 domains.
- Traffic Class or Flow Label field change does not affect IPsec integrity and security because these are mutable fields.

IPv6 Flow Label Field

- Flow Label is a new IPv6 header field.

IPv6 Flow Label Field (Cont.)

IPv6 flow label field characteristics:

- A new field is used to label packet flows.
- A flow can be used to request nondefault QoS.
- The Flow Label field is immutable between the source and destination nodes (which may not be changed, unlike the Traffic Class field).
- There are no existing implementations or standards defining the Flow Label field for QoS; they could be used to mark media streams.

IPv6 Flow Label Field (Cont.)

IPv6 flow label field characteristics continuation:

- A Flow Label field can be used if the encryption protocol “hides” the Layer 4 port number, which would be the base for traffic classification.
- The transport layer information can be located at a variable offset due to the presence of option headers.
- A flow label can be used to classify this traffic and to ensure QoS, based on the information in the first header.

IPv6 QoS Configuration

IPv6 QoS configuration characteristics:

- IPv6 QoS configuration is nearly identical to the IPv4 model
- On Cisco IOS, IOS XE, and IOS XR Software, the following is supported:
  - MQC
  - Class maps, policy maps, and service policy constructs
  - Support for most QoS features for managing IPv6 traffic
- QoS features supported for IPv6:
  - Packet classification
  - Queuing
  - Traffic shaping
  - Traffic policing
  - WRED
  - Class-based packet marking
  - NBAR

IPv6 QoS Configuration (Cont.)

```
class- map VOIP
 match dscp ipv6 ef
!
policy-map QOS
 class VOIP
  priority
  police rate 10 mbps
   conform-action transmit
   exceed-action drop
 !
 class class-default
  police rate 100 mbps
   conform-action transmit
   exceed-action drop
!
interface GigabitEthernet0/0/0/1
 service-policy output QOS
```

```
class-map VOIP
 match dscp ef
!
policy-map QOS
 class VOIP
  priority
  police rate 10000
   conform-action transmit
   exceed-action drop
!
 class class-default
  police rate 100000
   conform-action transmit
   exceed-action drop
!
interface GigabitEthernet0/0/1
 service-policy output QOS
```

Omit the **ip** keyword to match IPv4 and IPv6.

Cisco IOS Software Features

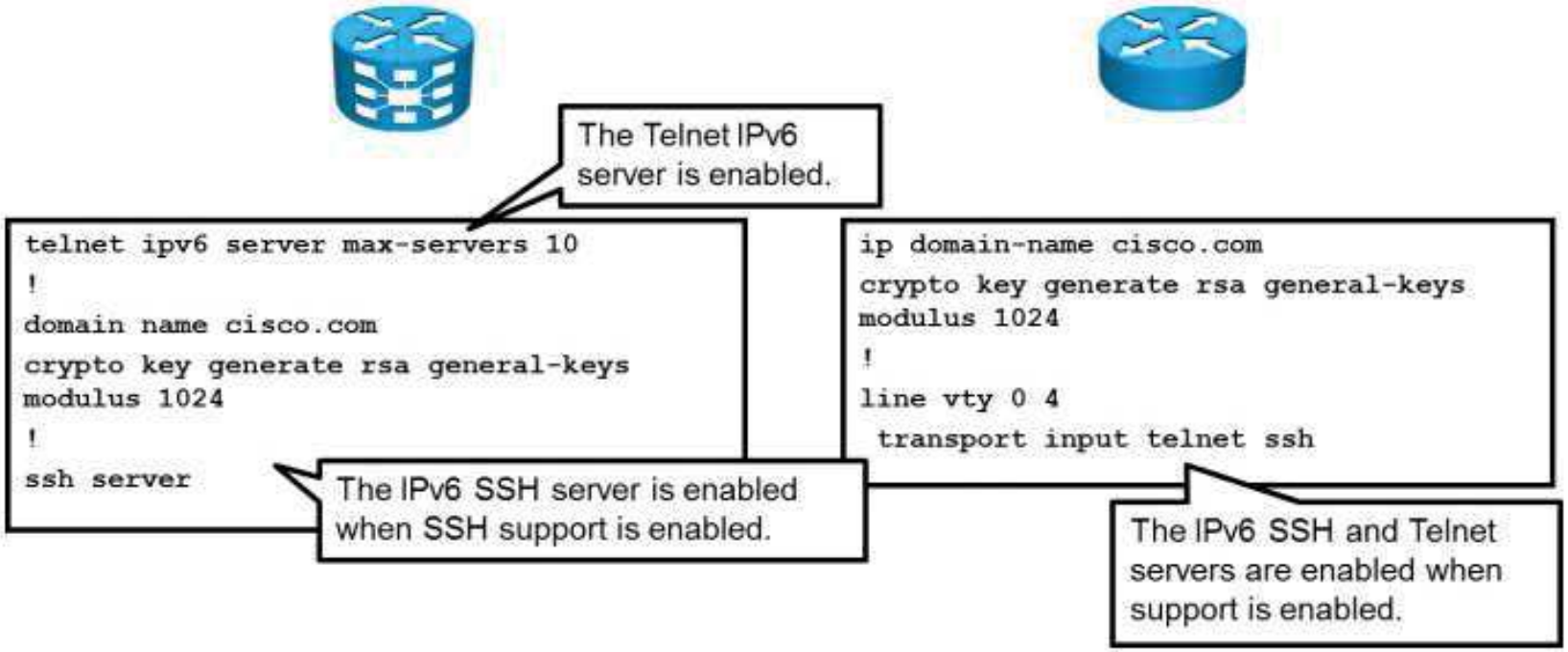
Cisco IOS software features characteristics:

- A router running Cisco IOS Software can act as a client or a server for many services.
  - Routing protocols
  - Network services
  - Management access
- To fully support IPv6, all of these services must be IPv6-capable.
- Configuration may differ slightly compared to IPv4.

Cisco IOS IPv6 Telnet and SSH Server and Client Support

Following are supported on the Cisco IOS:

- IPv6 Telnet client and server.
- IPv6 SSH client and server.



Cisco IOS IPv6 Telnet and SSH Server and Client Support (Cont.)

Cisco IOS IPv6 telnet and SSH server and client support example:

- To connect to an IPv6-enabled Telnet or SSH server, specify the IPV6 address or a hostname.

```
RP/0/RSP0/CPU0:P# telnet 2001:db8:1:1001::f
```

- To connect to an IPv6-enabled SSH server, specify the IPV6 address or a hostname.

```
RP/0/RSP0/CPU0:P# ssh 2001:db8:1:1001::f username student
```

Cisco IOS IPv6 Tools

Cisco IOS IPv6 tools characteristics:

- These IPv6 applications are available in Cisco IOS Software for network diagnostics:
  - Traceroute
  - Ping
- The following protocols are available for data transfer and remote management, and they support IPv6:
  - TFTP
  - HTTP
  - NTP version 4
  - Syslog
  - SNMP
- Tcl scripting can be used for automating complex tasks.

Cisco Discovery Protocol Support for IPv6

Cisco Discovery Protocol characteristics:

- Cisco Discovery Protocol:
  - Used to discover protocol addresses and platform information of neighboring devices.
  - Runs on all Cisco devices (that is, routers, switches, and so on).
- Cisco Discovery Protocol IPv6 support:
  - Adds IPv6 address and address type information.

Cisco Discovery Protocol Support for IPv6 (Cont.)

- Displays detailed information about neighbors.

```
RP/0/RSP0/CPU0:P# show cdp neighbors detail
<… output omitted …>
-------------------------
Device ID: PE8.lab.com
SysName :
Entry address(es):
IPv4 address: 192.168.82.80
IPv6 address: 2001:db8:192:168:82::80
IPv6 address: fe80::207:7dff:fe33:2783
Platform: cisco ASR1001, Capabilities: Router IGMP
Interface: GigabitEthernet0/0/0/9
Port ID (outgoing port): GigabitEthernet0/0/3
Holdtime : 176 sec
```

Cisco Express Forwarding IPv6

Cisco Express Forwarding characteristics:

- Similarities between Cisco Express Forwarding, v6, and Cisco Express Forwarding, v4:
  - Rapid packet forwarding on interfaces.
  - Behavior of commands.
- Cisco Express Forwarding is mandatory on Cisco IOS XR Software routers and cannot be disabled.
- On Cisco IOS Software, Cisco Express Forwarding, v6, uses a subset of the Cisco Express Forwarding, v4, commands:

```
router(config)# ipv6 cef
```

- To remove Cisco Express Forwarding, v6, use this command:

```
router(config)# no ipv6 cef
```

IP Service Level Agreement (SLA) for IPv6

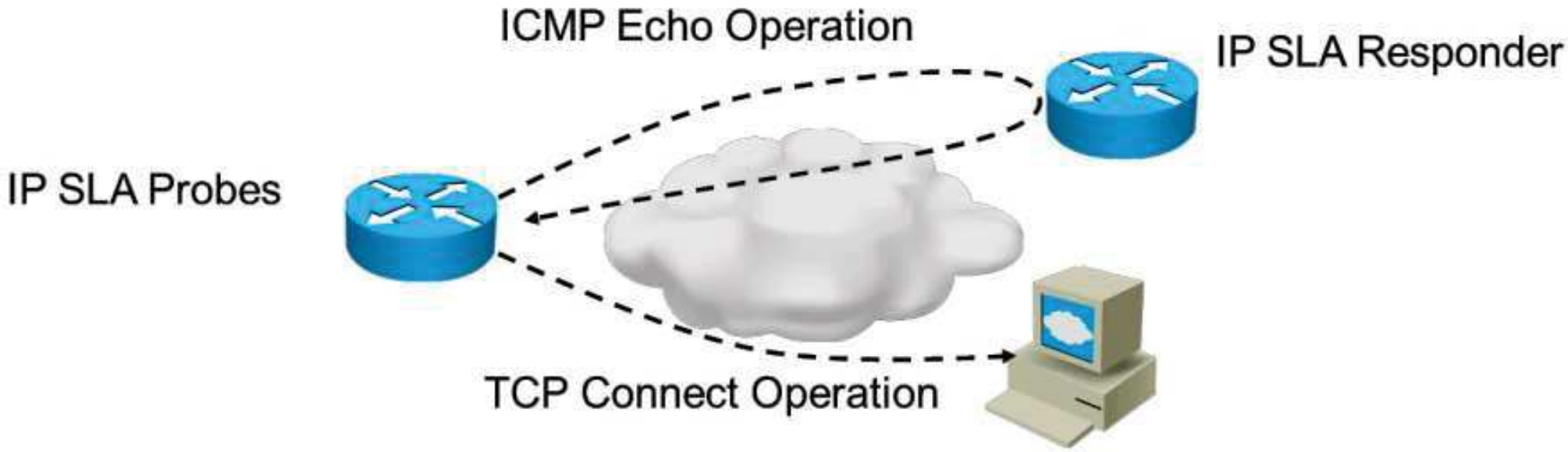
IP Service Level Agreements characteristics:

- The IP SLA software can be used as a performance tracking tool.
- Active monitoring of network infrastructure:
  - Monitors connectivity and throughput.
  - Monitors availability of network services (that is, web and so on).
- Monitoring capabilities include:
  - Monitoring network delay and packet loss.
  - Monitoring network latency and jitter.
  - Checking conformity with service provider service level agreements.
- IP SLAs are not available for IPv6 on Cisco IOS XR Software.
- IP SLAs are available for both IPv4 and IPv6 (currently not all probes are available for IPv6) on Cisco IOS and IOS XE Software.

IP Service Level Agreement (SLA) for IPv6 (Cont.)

IP SLA key components:

- The IP SLA architecture consists of an IP SLA source and an IP SLA target.
- A probe is configured on the source, checking connectivity from the source to the target.
- The target device can be a router with an IP SLA responder or an IPv6 endpoint.



Configuring IP SLA

IP SLA configuration examples:

- Creates an IP SLA probe with a number and enters IP SLA configuration mode.

```
Router(config)# ip sla number
```

- Available probes for IPv6.

```
Router(config-ip-sla)# udp-jitter
Router(config-ip-sla)# udp-echo
Router(config-ip-sla)# icmp-echo
Router(config-ip-sla)# tcp-connect
```

Configuring IP SLA (Cont.)

- Schedules an IP SLA probe on the router.

```
Router(config)# ip sla schedule number [life {forever | seconds}] [starttime
{hh:mm[:ss] [month day | day month] | pending | now | after hh:mm:ss}] [ageout
seconds] [recurring]]
```

Configuring IP SLA (Cont.)

Available IPv6-related options:

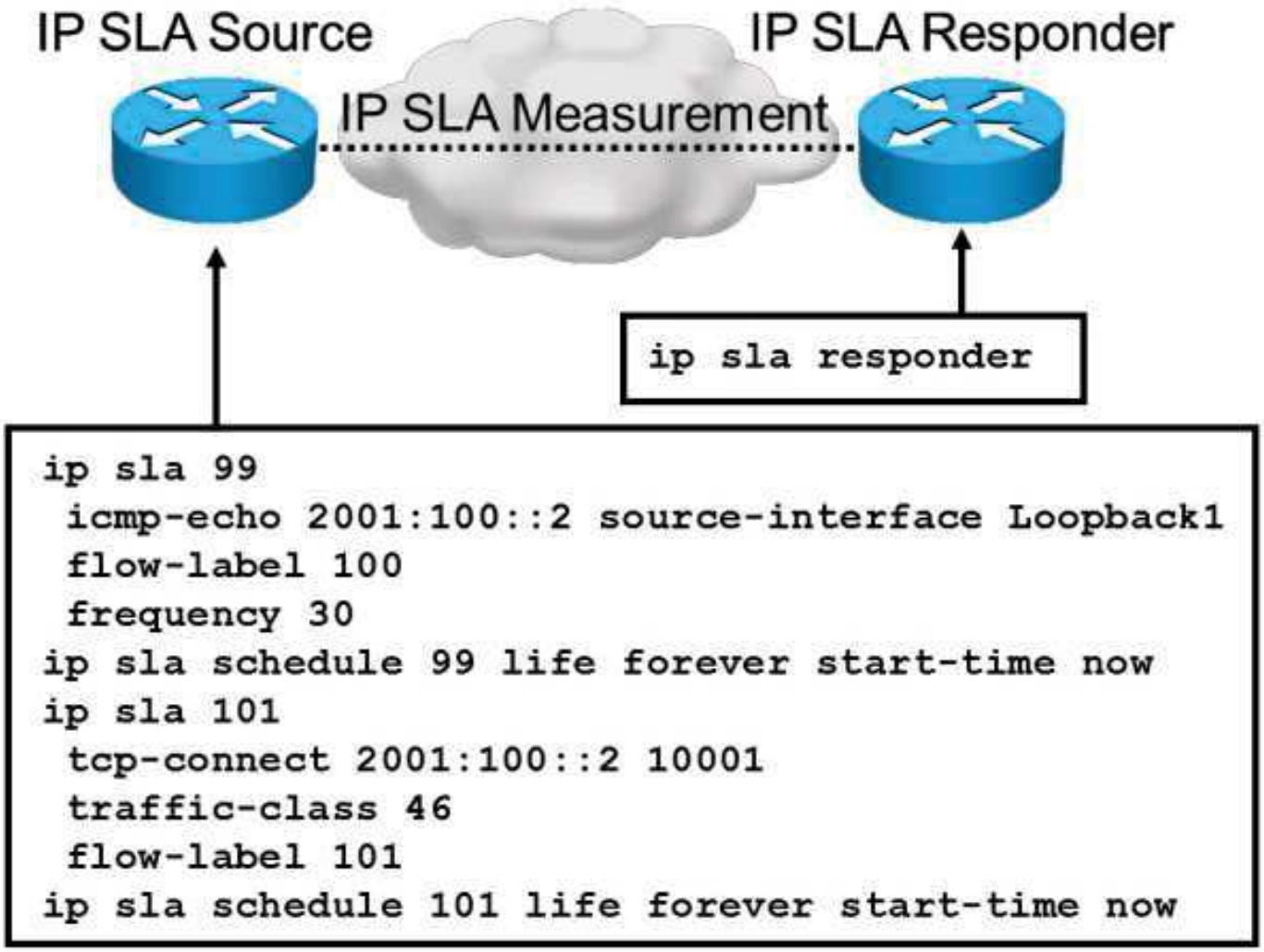
- Flow Label Identifier:

```
router(config-ip-sla-echo)# flow-label
```

- Traffic Class:

```
router(config-ip-sla-echo)# traffic-class
```

Configuring IP SLA (Cont.)



- An IP SLA probe is configured on the SLA source device.
- The responder on the IP SLA responding device is enabled.

Summary

- Multicast is used in the infrastructure layer of the Cisco IP NGN.
- IPv6 multicast addresses start with “FF”.
- IPv6 multicast address scope defines the reach of the multicast address group.
- Solicited node multicast address is used when doing neighbor discovery.
- Global scope represents the entire Internet.
- Multicast in IPv6 is conceptually similar to multicast in IPv4. The biggest difference is in domain control.
- PIM is protocol independent: MRIB is built on underlying unicast routing table.
- The RP address can be embedded into a multicast group address.

Summary (Cont.)

- To enable multicast routing use multicast-routing IOS XR command or ipv6 multicast-routing IOS/IOS XE command.
- MLD handles join and leave processes on the access segment between the listener and the first multicast router.
- MLD v1 messages are Query, Report and Done.
- MLDv1 Query: “Which multicast addresses have listeners on the link?”
- MLDv1 Report: Nodes reply with the list of multicast groups they are receiving.
- MLDv1 Done: A done message is a node leaving a group.
- Multicast-address-specific queries have their Maximum Response Delay set to Last Listener Query Interval.
- MLDv2 defines new messages.

Summary (Cont.)

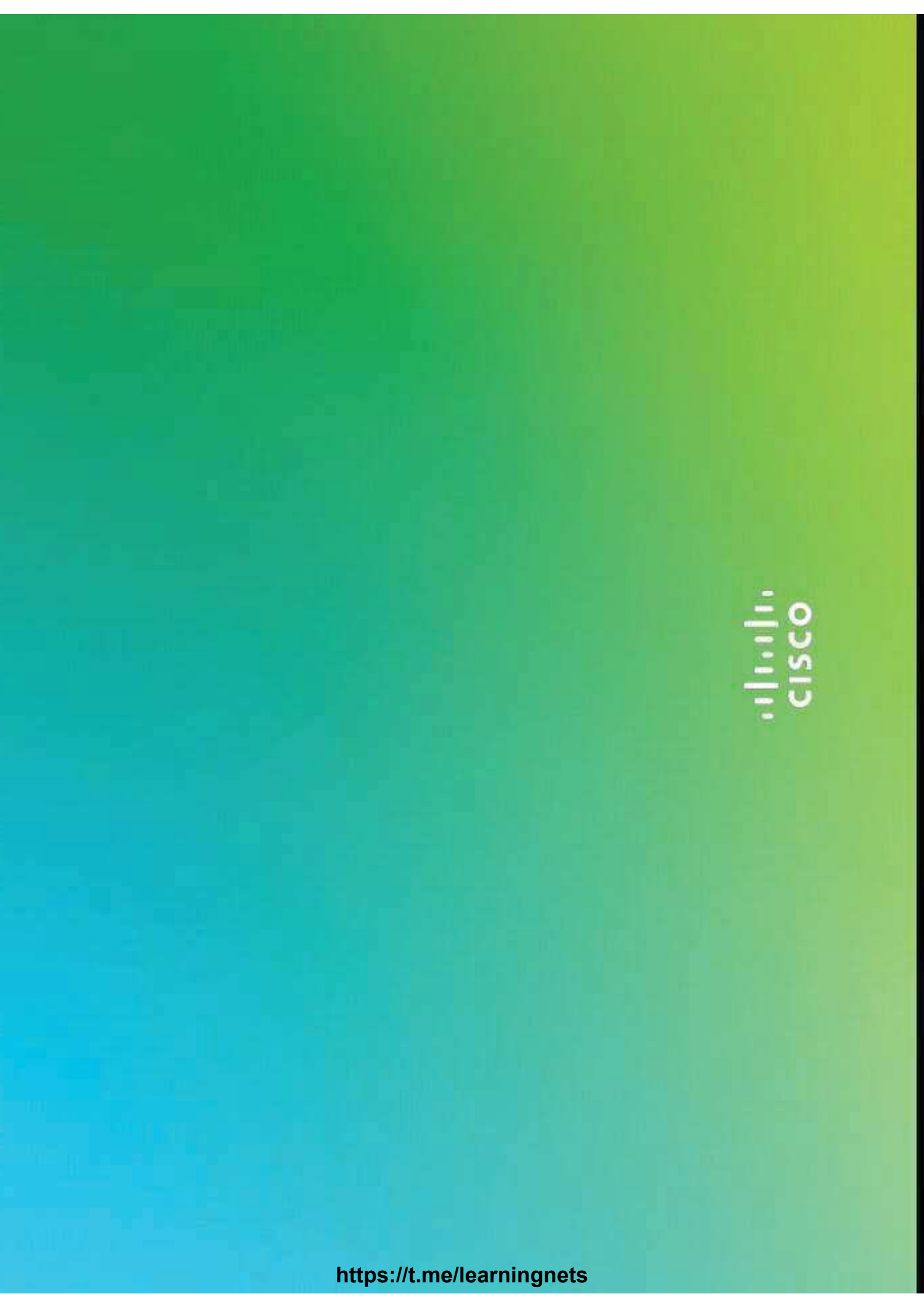
- Access groups enable access control for multicast group receivers on a MLD router.
- MLD join-group and static-group allow routers to join multicast groups.
- MLD is used with IPv6 multicast traffic to handle join and leave processes on the access segment between the listener and the first multicast router.
- Snooping of MLD packets is used to determine and store the information that switch ports receive data for multicast groups.
- DNS and DHCPv6 are used in the services layer of the Cisco IP NGN.
- DNS replaces A record with AAAA record for IPv6 addresses.
- Dynamic DNS works on the same principle as in IPv4.
- Before using DHCPv6, the client will check for the presence of any routers.

Summary (Cont.)

- DHCPv6 cannot send gateway information to the client.
- DHCPv6-Lite is used to send DNS information to the client, without configuring the client address.
- DHCPv6 can be used to delegate prefixes to client routers instead of single addresses.
- DHCPv6 uses similar commands as IPv4 DHCP.
- QoS is used on the IP infrastructure layer of the Cisco IP NGN.
- IPv6 keeps ToS fields under new name Traffic class. It has an additional flow label field.
- Traffic class field is just renamed ToS field.
- Flow label field is new addition to IPv6 and is used to distinguish flows at layer 3.

Summary (Cont.)

- QoS can be configured to be protocol agnostic.
- To simplify transition to IPv6, systems should be dual stacked to support IPv6 in conjunction with IPv4.
- Cisco IOS SSH and telnet support IPv6 seamlessly.
- Most other tools in Cisco IOS also support IPv6.
- Cisco Discovery Protocol supports IPv6.
- To take the full advantage of IPv6 forwarding you must enable Cisco Express Forwarding for IPv6.
- IP SLA supports IPv6, but not all features may be available.
- The commands for the currently available probes of IP SLA that can be used for IPv6 are **udp-jitter, udp-echo, icmp-echo**, and **tcp-connect**.



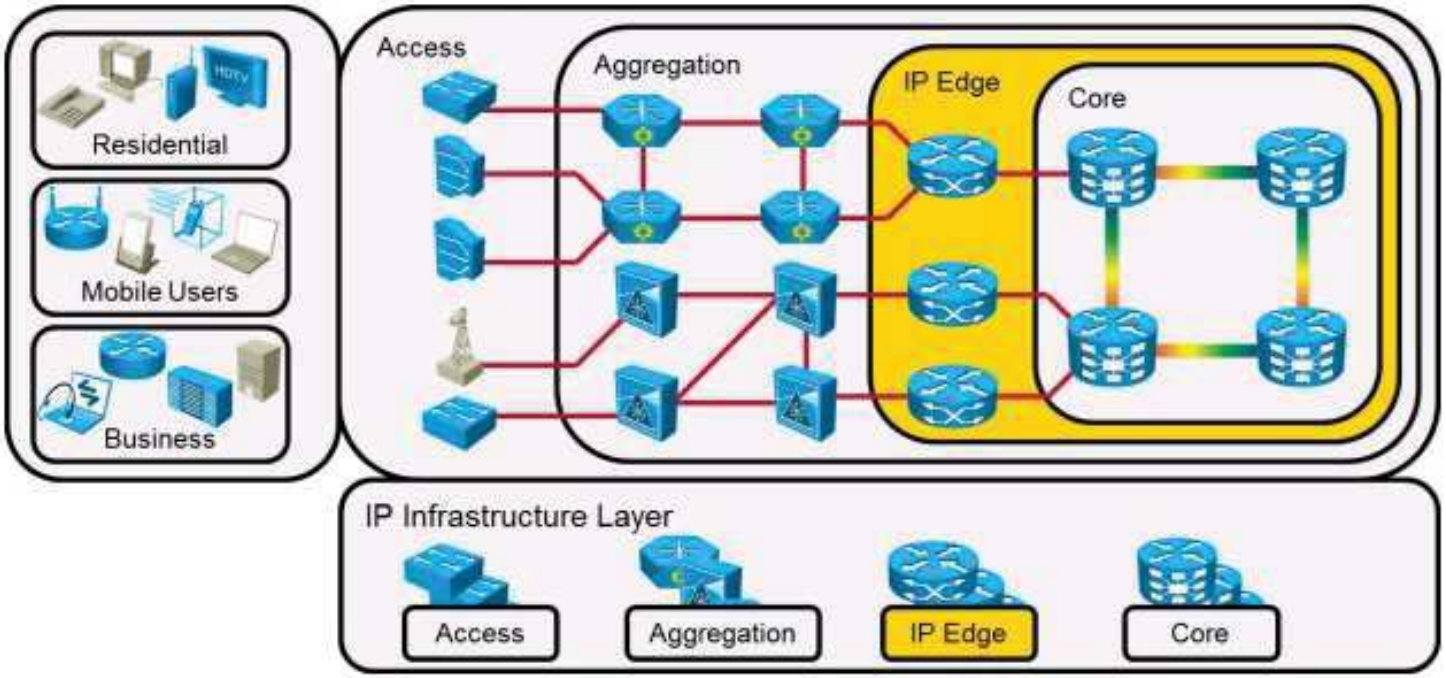
Defining IPv6 Transition Mechanisms

Service Provider IPv6 Transition Implementations

Dual Stack, CGN, and NAT64

Dual Stack, CGN, and NAT64 are used in the Cisco IP NGN infrastructure layer:

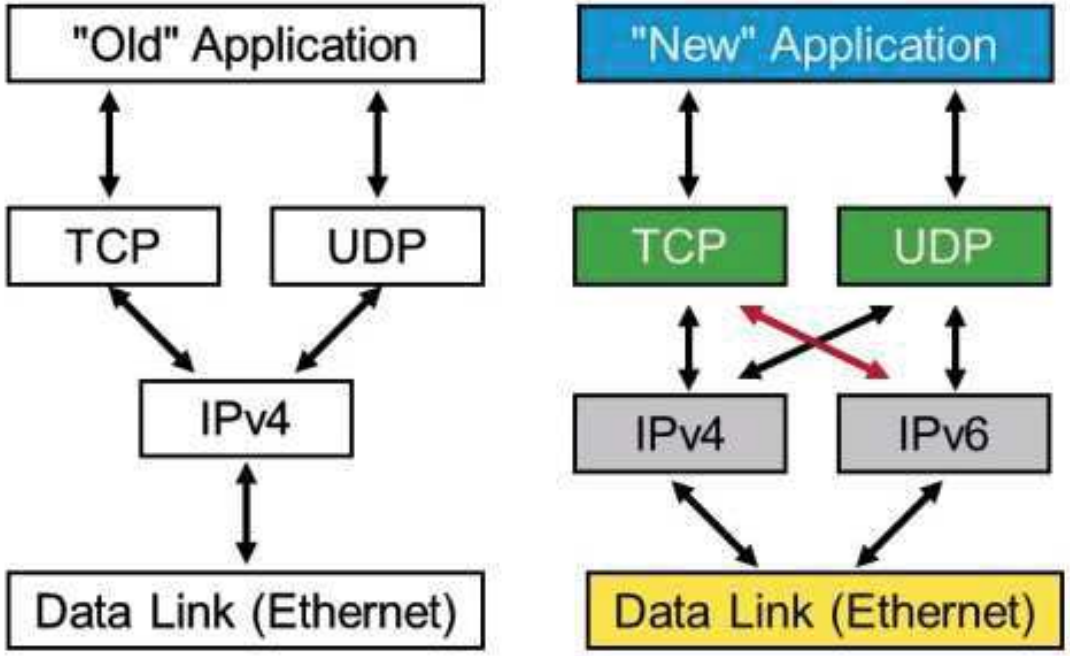
- Dual stack is used on the IP infrastructure layer of the Cisco IP NGN.
- Dual stack is implemented on all devices that require IPv4 and IPv6 connectivity.
- CGN and NAT64 are implemented on IP edge devices.



Dual-Stack Operations Overview

Overview:

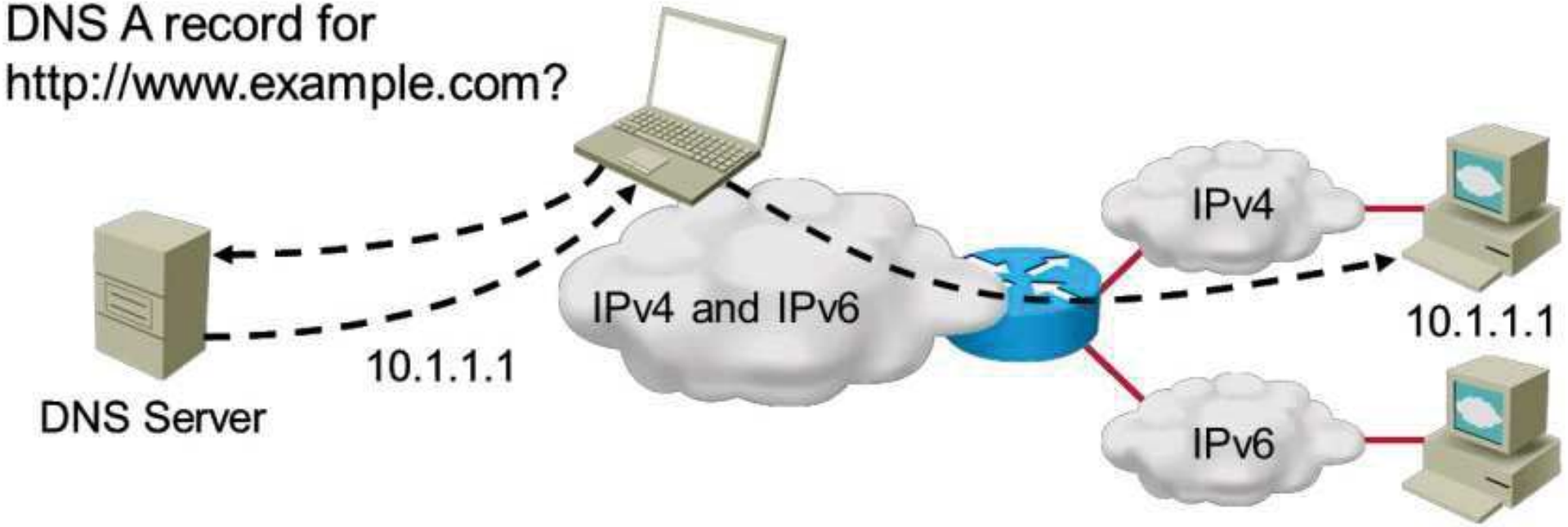
- If a host is required to communicate with both IPv4 and IPv6 natively, dual stack is required.
- Both IPv4 and IPv6 stacks are concurrently enabled.
- Applications can talk to both stacks.
- IP version choice is based on name lookup and application preference; the IPv6 path is preferred.



Dual-Stack Operations Overview (Cont.)

A dual-stack operation is an application that is not aware of IPv6, or if the destination is only IPv4-enabled:

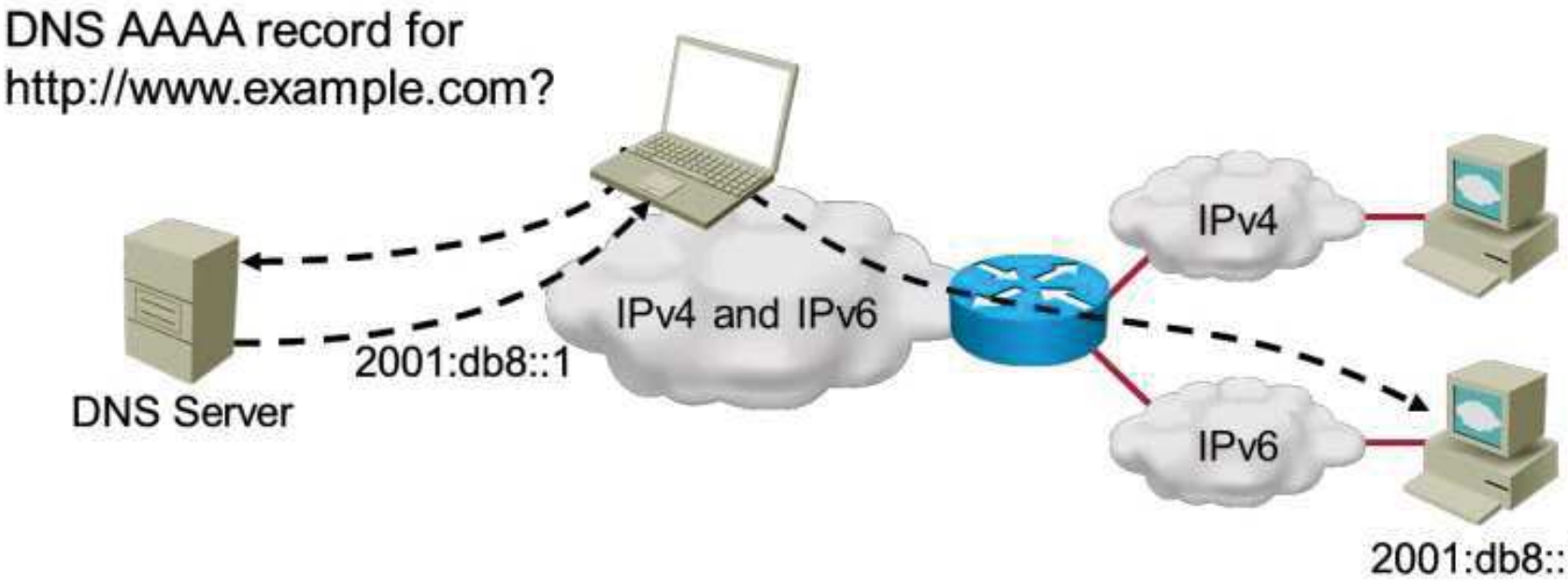
- It asks the DNS for an IPv4 address.
- It connects to the IPv4 address.
- Legacy applications exist; new applications should be designed as protocol-independent.



Dual-Stack Operations Overview (Cont.)

An application or destination that is only IPv6-enabled has to use IPv6 because it is the only stack:

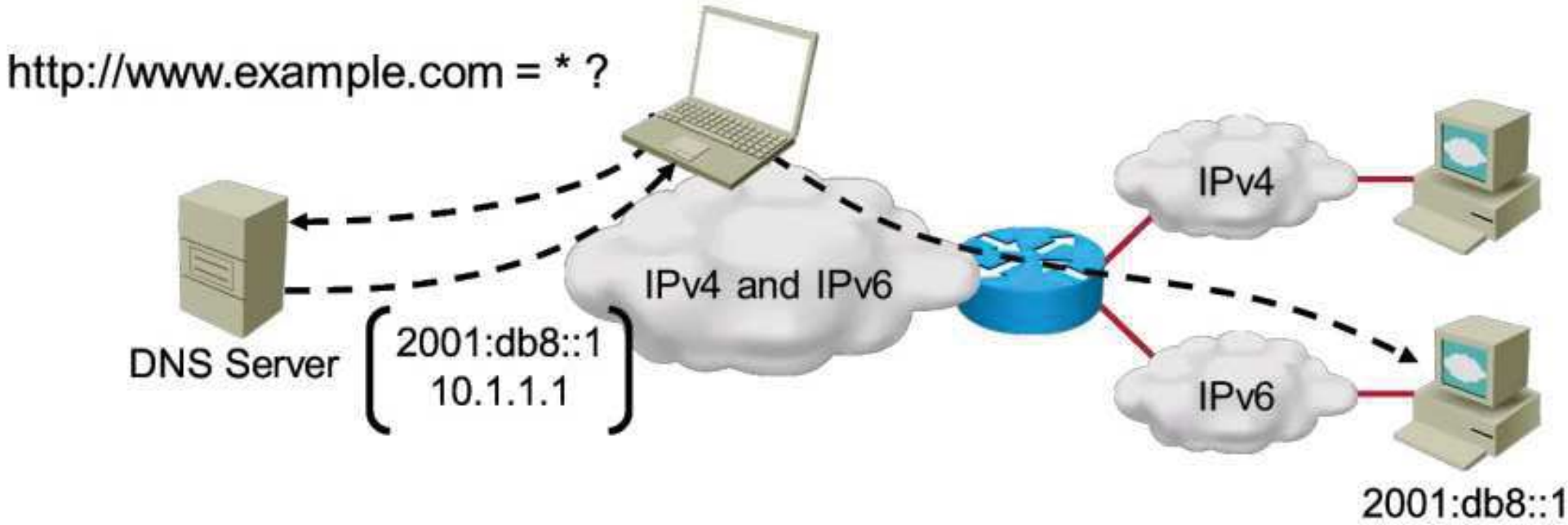
- It asks the DNS for an IPv6 address.
- It connects to the IPv6 address.



Dual-Stack Operations Overview (Cont.)

If an application and destination support both IPv4 and IPv6, then the following occurs:

- It chooses one address.
- It connects, for example, to the IPv6 address.
- Some content providers require you to register to receive AAAA DNS responses.



Dual-Stack Considerations

Dual-stack considerations follow:

- Advantages:
  - Relatively simple approach.
  - Can be less costly in the short term.
  - Continues to leverage existing IPv4 infrastructure.
  - Allows indefinite coexistence between IPv4 and IPv6.
- Issues:
  - Can be more costly over the long term.
  - Increases network complexity.

Dual Stack with Carrier-Grade NAT

Dual stack with carrier-grade NAT characteristics follow:

- Carrier-grade NAT can be used to alleviate public IPv4 address exhaustion on dual-stacked hosts.
- Host addressing:
  - Public IPv6 address.
  - Private IPv4 address.
- IPv6 content is accessed natively over IPv6.
- IPv4 content is accessed natively over IPv4 and is also called large-scale NAT.

NAT444

NAT444 characteristics:

- Translating IPv4 addresses into IPv4 addresses is called NAT44.
- The customer also can use its own private address space and translate this address space to be assigned a private IP address.
- This solution is called NAT444 because of double IPv4-to-IPv4 translation.

Carrier-Grade NAT on Cisco Routers

- The platform for CGN should be capable of providing the following:
  - An order of millions of translations.
  - 10-Gbps, full-duplex, bandwidth throughput.
- Cisco Carrier-Grade Services Engine PLIM is a high-performance, multi-CPU module that performs carrier-grade NAT:
  - Available for Cisco CRS-1 and CRS-3 platforms.
  - Part of the Cisco Carrier-Grade IPv6 solution.

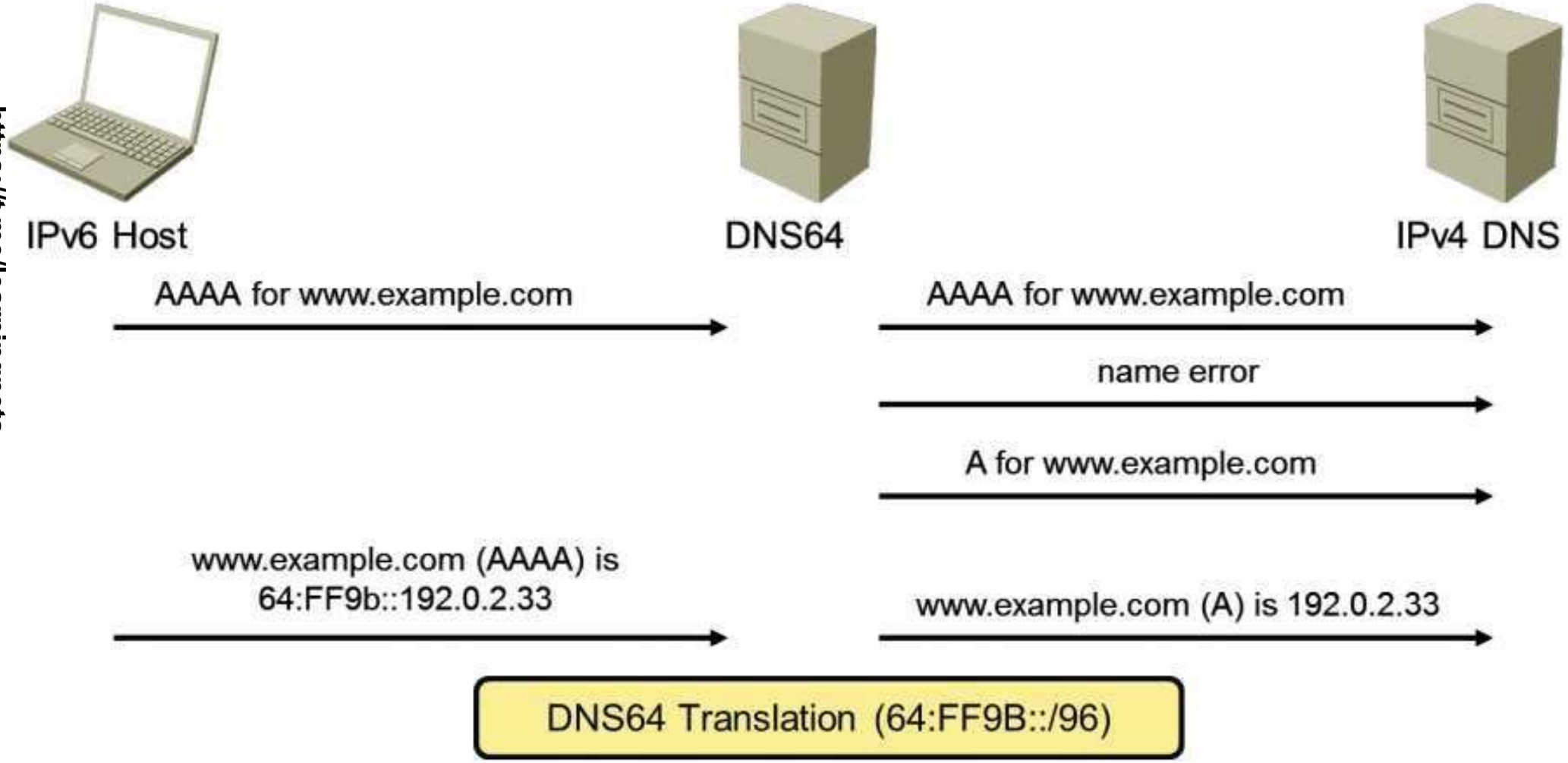
NAT64

NAT64 characteristics:

- Instead of CGN, another solution is to assign only IPv6 addresses to customers and to translate IPv6 packets to IPv4 packets on the edge of the P-network.
- The solution is called NAT64 and works with DNS64.
- NAT64 comes in two flavors:
  - Stateless NAT64
  - Stateful NAT64
- NAT64 is supported on the Cisco ASR 1000 and Cisco CRS with Cisco CGSE module.

DNS64

- DNS64 synthesizes an AAAA record based on the received A record and a well-known or service provider-assigned translation prefix.



Stateless NAT64

Stateless NAT64 characteristics follow:

- Source and destination IPv4 addresses are embedded in the IPv6 addresses.
- Specific IPv6 addresses have to be assigned to IPv6 hosts:
  - Combination of prefix and translatable IPv4 address.
- One IPv6 address is translated into one IPv4 address, so no IPv4 address conservation is achieved.

SRC IPv6	DST IPv6		SRC IPv4	DST IPv4
2001:db8::192.0.2.33	2001:db8::198.51.100.2	→	192.0.2.33	198.51.100.2

SRC IPv6	DST IPv6		SRC IPv4	DST IPv4
2001:db8::198.51.100.2	2001:db8::192.0.2.33	←	198.51.100.2	192.0.2.33

Stateful NAT64

Stateful NAT64 characteristics follow:

- The destination IPv4 address is embedded in the IPv6 address.
- The source IPv6 address is translated into one of the IPv4 addresses assigned to the NAT64 pool. IPv4 addresses can be overloaded using PAT.
- Many IPv6 addresses are translated into one IPv4 address, so IPv4 address conservation is achieved.
- Any IPv6 address can be assigned to IPv6 hosts.

SRC IPv6	DST IPv6	SRC	DST
2001:db8::2	2001:db8::198.51.100.2	1025	80

→

SRC IPv4	DST IPv4	SRC	DST
192.0.2.33	198.51.100.2	1026	80

SRC IPv6	DST IPv6	SRC	DST
2001:db8::198.51.100.2	2001:db8::2	80	1025

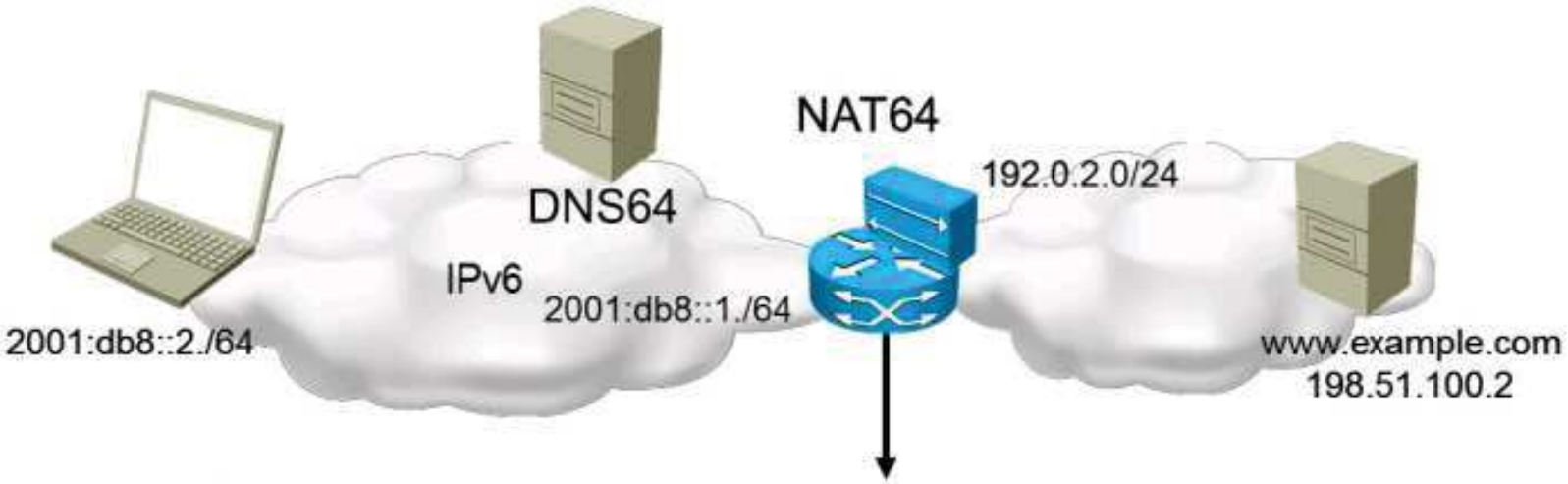
←

SRC IPv4	DST IPv4	SRC	DST
198.51.100.2	192.0.2.33	80	1026

Stateless versus Stateful NAT64 Comparison

Stateless NAT64	Stateful NAT64
1:1 translation	1:N translation
No conservation of IPv4 address	Conserves IPv4 address
Assures end-to-end address transparency and scalability	Uses address overloading, so it lacks an end-to-end address transparency
No state or bindings created on the translation	State or bindings are created on every unique translation
Requires IPv4-translatable IPv6 address assignment (mandatory requirement)	No requirement on the nature of IPv6 address assignment
Requires either manual or DHCPv6-based address assignment for IPv6 hosts	Free to choose any mode of IPv6 address assignment (DHCPv6 and stateless autoconfiguration)

Stateful NAT64 Configuration on ASR 1000




```
interface GigabitEthernet0/0/0
 ipv6 address 2001:db8::1/64
 nat64 enable
!
interface GigabitEthernet0/0/1
 ip address 192.0.2.1 255.255.255.0
 nat64 enable
!
ipv6 access-list LIST
 permit ipv6 2001:db8::/64 any
!
nat64 prefix stateful 2001:db8::/96
nat64 v4 pool POOL 192.0.2.2 192.0.2.254
nat64 v6v4 list LIST pool POOL overload
```

Enable NAT64 on the IPv6-facing interface.

Enable NAT64 on the IPv4-facing interface.

Create an ACL to specify which IPv6 hosts can translate.

Specify NAT64 prefix.

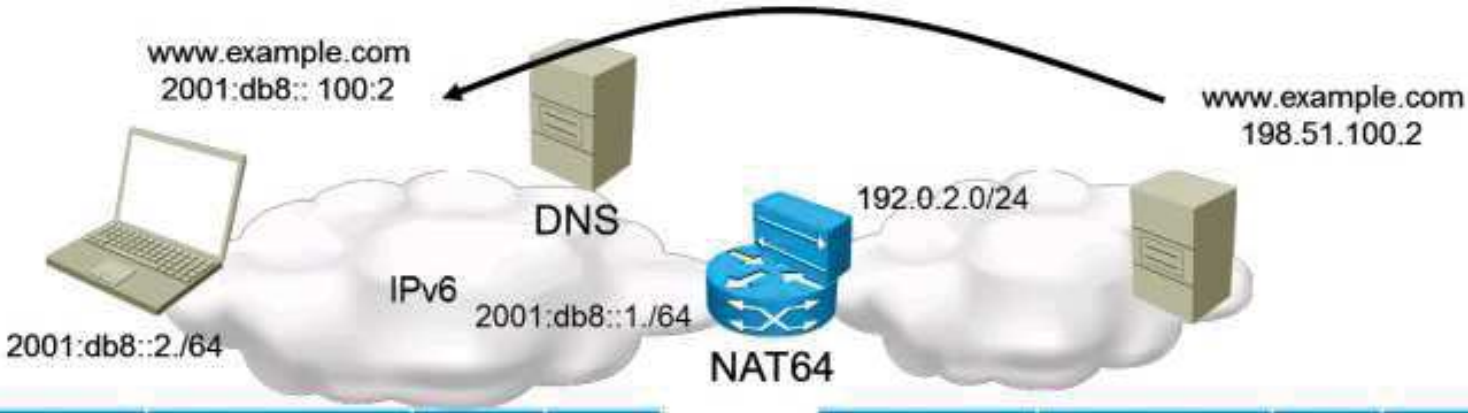
Specify IPv4 pool.

Enable NAT64 with PAT.

Static, Stateful NAT64 Configuration on ASR 1000

Static, stateful NAT64 characteristics follow:

- NAT64 is used to provide services to the IPv6 world from existing IPv4 services.
- Instead of algorithmic translation for IPv4 servers, static translations can be implemented.
- NAT64 does not require DNS64 but does require AAAA records on DNS instead.



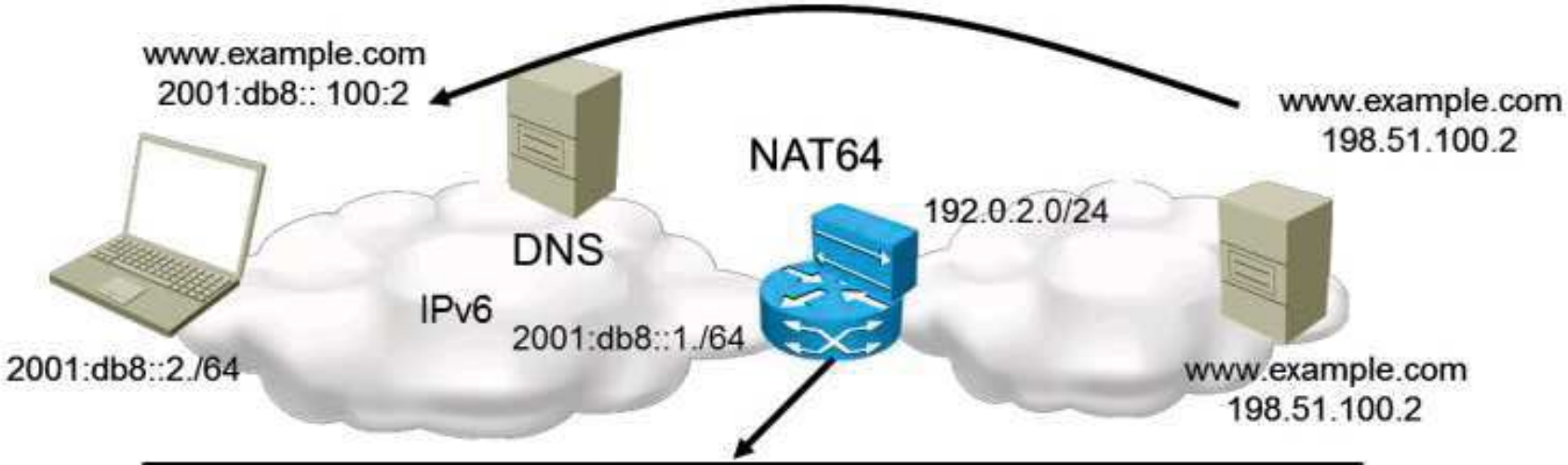
SRC IPv6	DST IPv6	SRC	DST
2001:db8::2	2001:db8::100.2	1025	80

SRC IPv4	DST IPv4	SRC	DST
192.0.2.33	198.51.100.2	1026	80

SRC IPv6	DST IPv6	SRC	DST
2001:db8::100.2	2001:db8::2	80	1025

SRC IPv4	DST IPv4	SRC	DST
198.51.100.2	192.0.2.33	80	1026

Static Stateful NAT64 Configuration on ASR 1000 (Cont.)




```
interface GigabitEthernet0/0/0
 ipv6 address 2001:db8::1/64
 nat64 enable
!
interface GigabitEthernet0/0/1
 ip address 192.0.2.1 255.255.255.0
 nat64 enable
!
ipv6 access-list LIST
 permit ipv6 2001:db8::/64 any
!
nat64 prefix stateful 2001:db8::/96
nat64 v4 pool POOL 192.0.2.2 192.0.2.254
nat64 v4v6 static 198.51.100.2 2001:db8::100.2
nat64 v6v4 list LIST pool POOL overload
```

Enable static NAT64.

Static, Stateful NAT64 Configuration on ASR 1000 (Cont.)

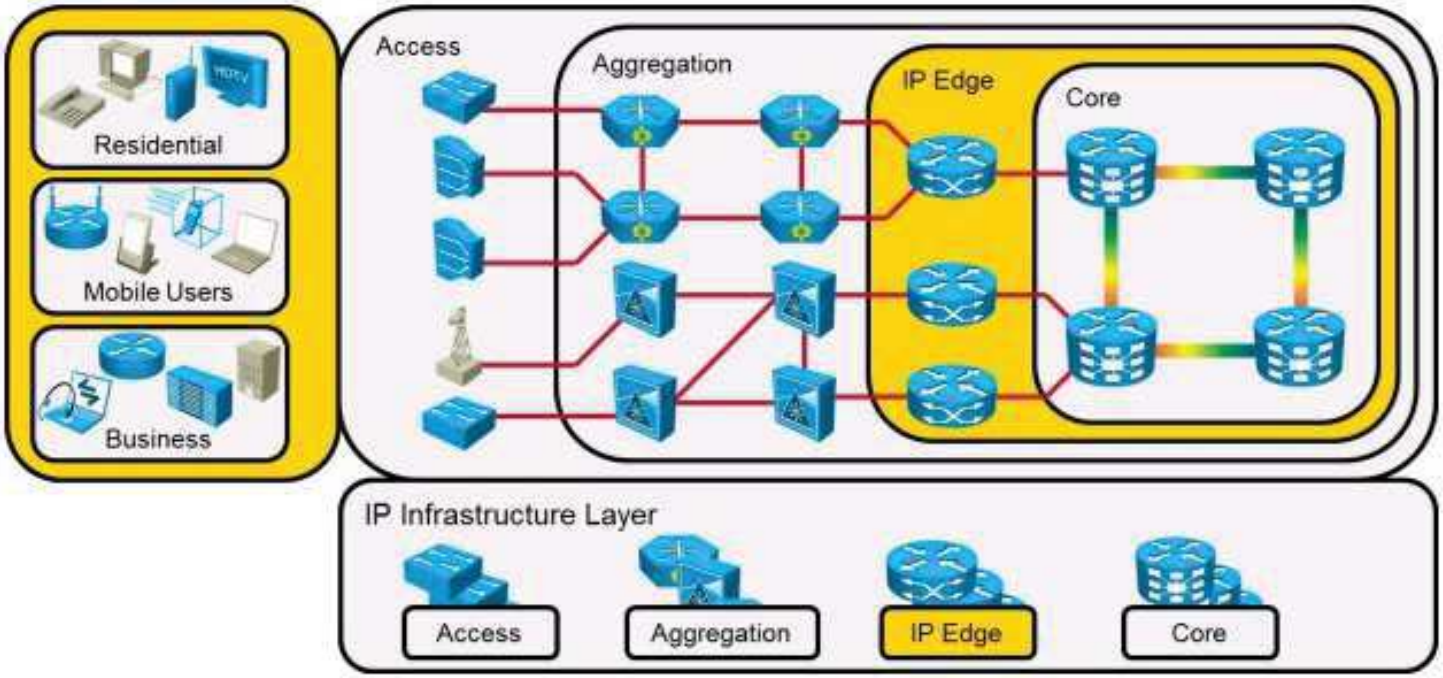
- Displays NAT64 translations:

```
NAT64# show nat64 translations
Proto  Original IPv4       Translated IPv4
       Translated IPv6     Original IPv6
------------------------------------------------------------------------
---    198.51.100.2        2001:db8::100.2
       ---
icmp   198.51.100.2 :1     [2001:db8::100.2]:6520
       192.0.2.2 :1        [2001:db8::2]:6520
Total number of translations: 2
```

IPv6 Tunneling Mechanisms

IPv6 tunneling mechanisms in the Cisco IP NGN infrastructure layer follow:

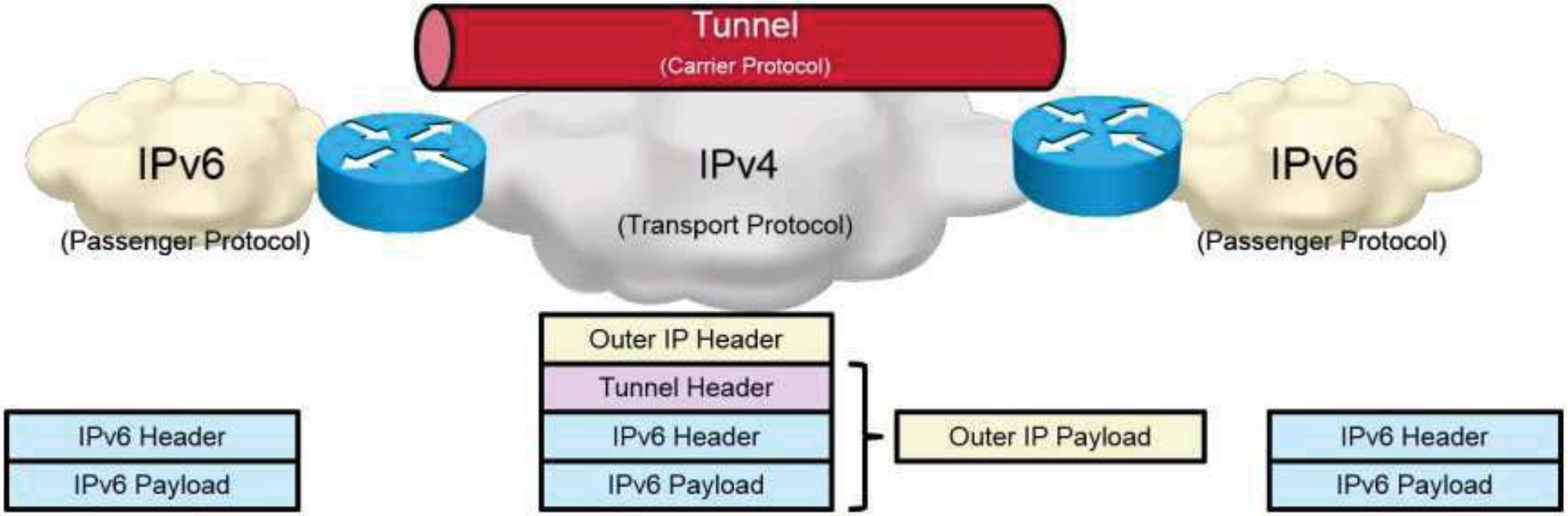
- IPv6 tunneling is used on the IP infrastructure layer of the Cisco IP NGN.
- IPv6 tunnel endpoints are implemented on IP edge devices or customer edge devices.



IPv6 Tunneling Mechanisms (Cont.)

IPv6 tunneling mechanism characteristics follow:

- Tunneling is used to transport one network protocol over another by encapsulating packets.
- During the transition, not all devices could be configured with dual stack:
  - Not all devices are under common administration.
  - There is almost twice as much of an administrative burden as with single-stack networks.



IPv6 Tunneling Mechanisms (Cont.)

Many techniques are available for establishing a tunnel:

- **Manually configured:** IPv6-in-IPv4 and GRE
- **Semiautomatic:** Tunnel broker
- **Automatic:** 6to4 tunneling and 6RD

Manually Configured Tunnels

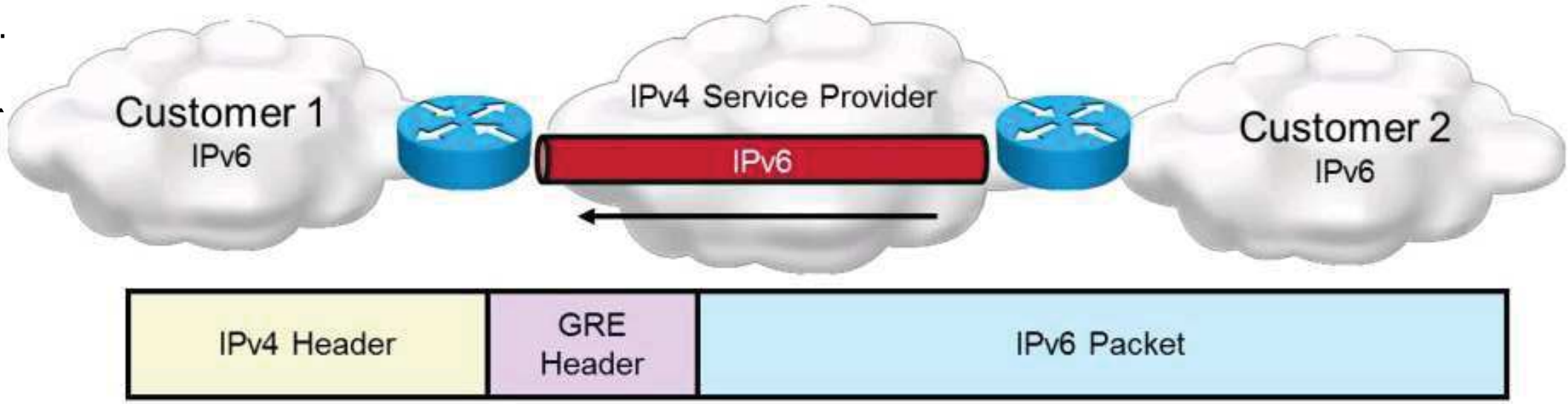
Configured tunnels connect IPv4 and IPv6 dual-stack hosts or networks to larger IPv6 networks:

- Local network administrators arrange for a tunnel between IPv6 networks across IPv4-only networks.
- Configured tunnels are simple to deploy.
- Configured tunnels allow the transport of IPv6 packets over an IPv4 network.
- Configured tunnels are available on most platforms.
- Configured tunnels do not scale well.

GRE Tunnels

GRE tunnel characteristics follow:

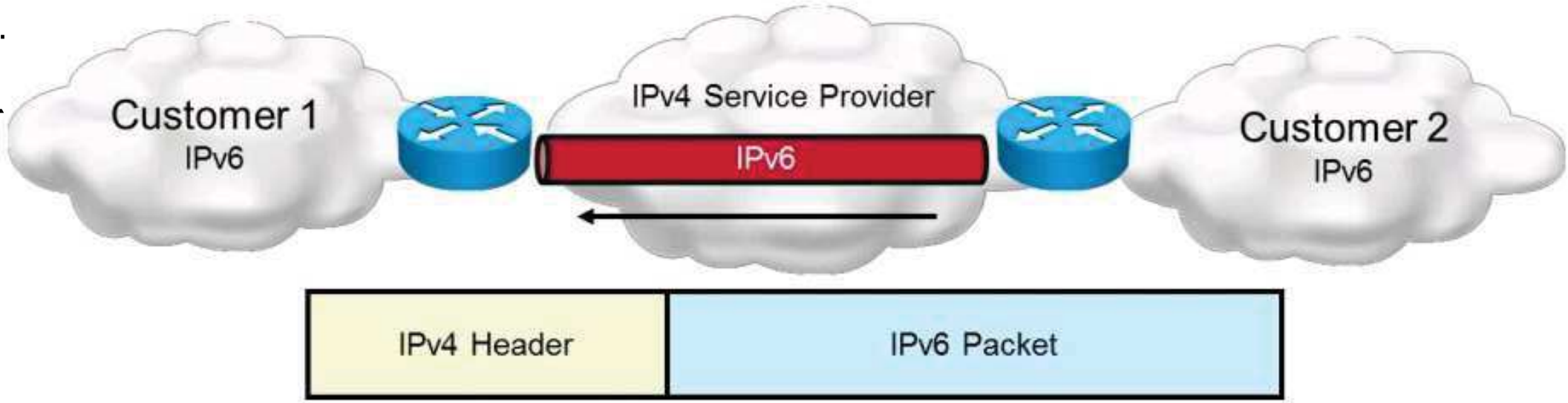
- IPv6 traffic is sent over IPv4 over explicitly configured GRE tunnels.
- GRE tunnels are an arbitrary Layer 3 protocol in IP.
- GRE uses protocol number 47 in the IPv4 header.
- The tunnel interface can be dual stack.
- The overhead is the IPv4 header + the GRE header.



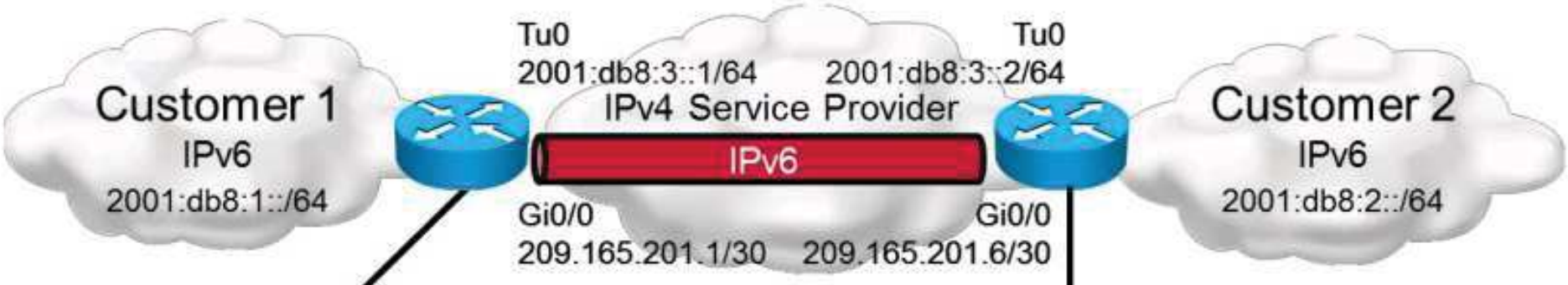
6in4 Tunnels

6in4 tunnel characteristics follow:

- IPv6 traffic is sent over IPv4 over explicitly configured tunnels.
- 6in4 tunneling uses protocol number 41 in the IPv4 header.
- The tunnel interface is IPv6-stack only.
- The only overhead is the IPv4 header.



6in4 Tunnel Configuration



Set IPv6 address.

```
interface Tunnel0
 ipv6 address 2001:db8:3::1/64
 tunnel source GigabitEthernet0/0
 tunnel destination 209.165.201.6
 tunnel mode ipv6ip
!
ipv6 route 2001:db8:2::/64 Tunnel0
2001:db8:3::2
```

Set tunnel mode.

Configure static route for IPv6 traffic.

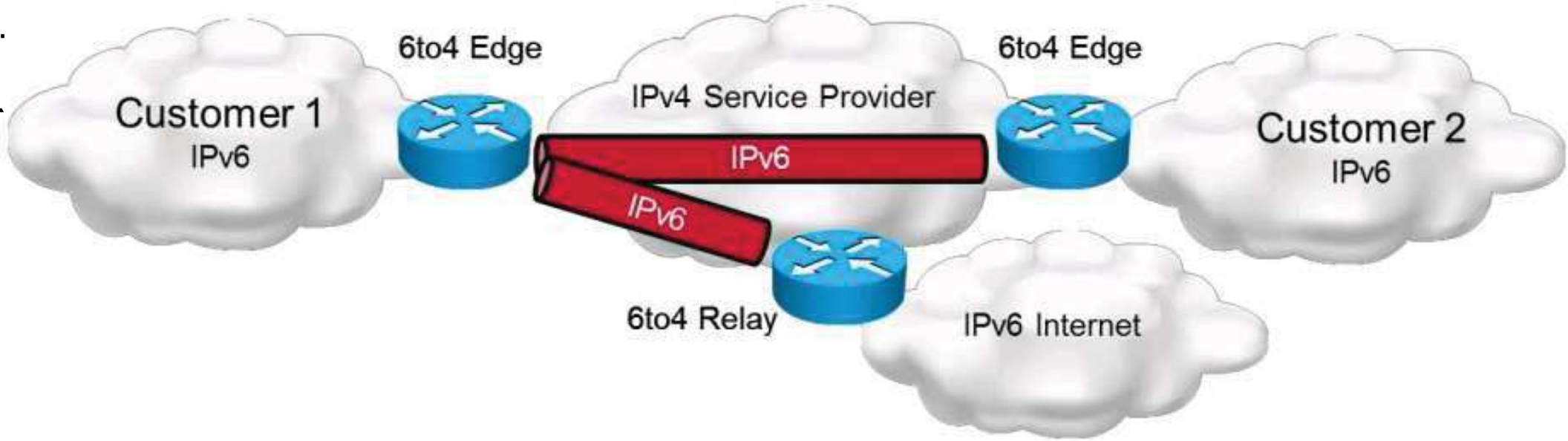
Set tunnel source and destination.

```
interface Tunnel0
 ipv6 address 2001:db8:3::2/64
 tunnel source GigabitEthernet0/0
 tunnel destination 209.165.201.1
 tunnel mode ipv6ip
!
ipv6 route 2001:db8:1::/64 Tunnel0
2001:db8:3::1
```

6to4 Automatic Tunnels

6to4 automatic tunnel characteristics follow:

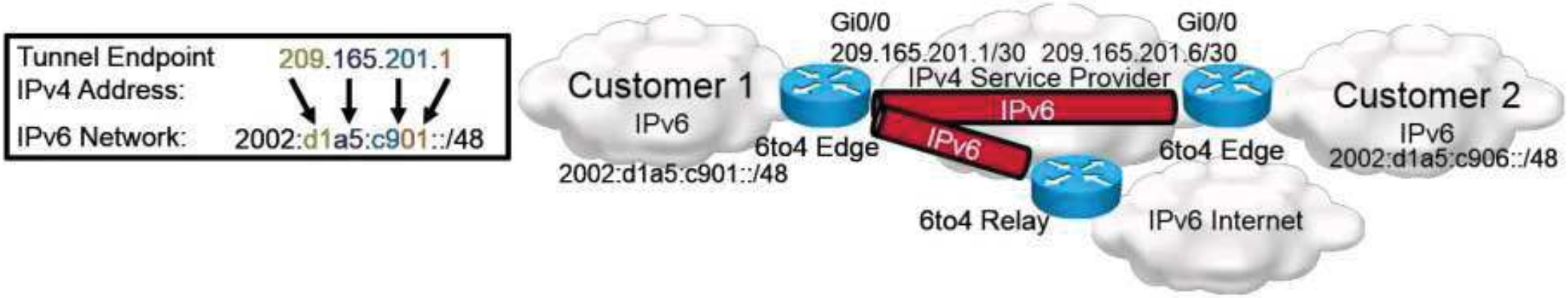
- 6to4 is an automatic tunneling method.
- 6to4 uses 6in4 encapsulation to tunnel IPv6 in IPv4.
- To access native IPv6 networks, relay routers have to be established.
- The relay router should be available on a reserved IPv4 address—192.88.99.1.



6to4 Automatic Tunnels (Cont.)

6to4 addressing characteristics follow:

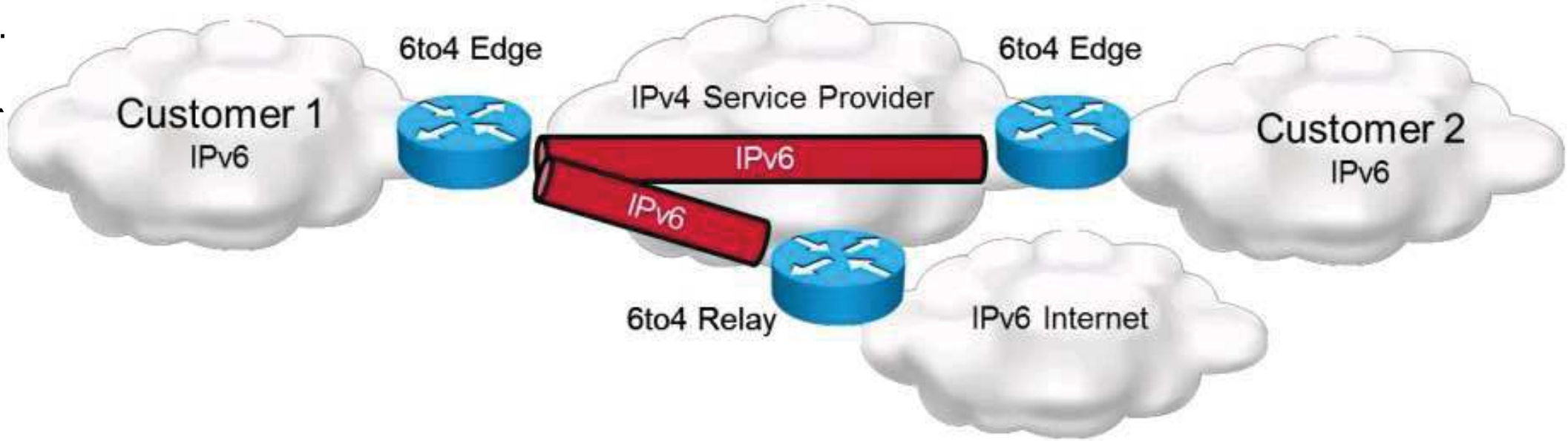
- IPv6 networks have to use specially assigned prefixes—2002::/16.
- The well-known prefix is concatenated with the IPv4 address that is assigned to the customer. The resulting IPv6 network is assigned to the customer.
- Traffic between IPv6 islands: The tunnel destination is determined from the IPv6 destination prefix.
- Traffic to the IPv6 Internet: The 6to4 relay router IPv4 address is known.



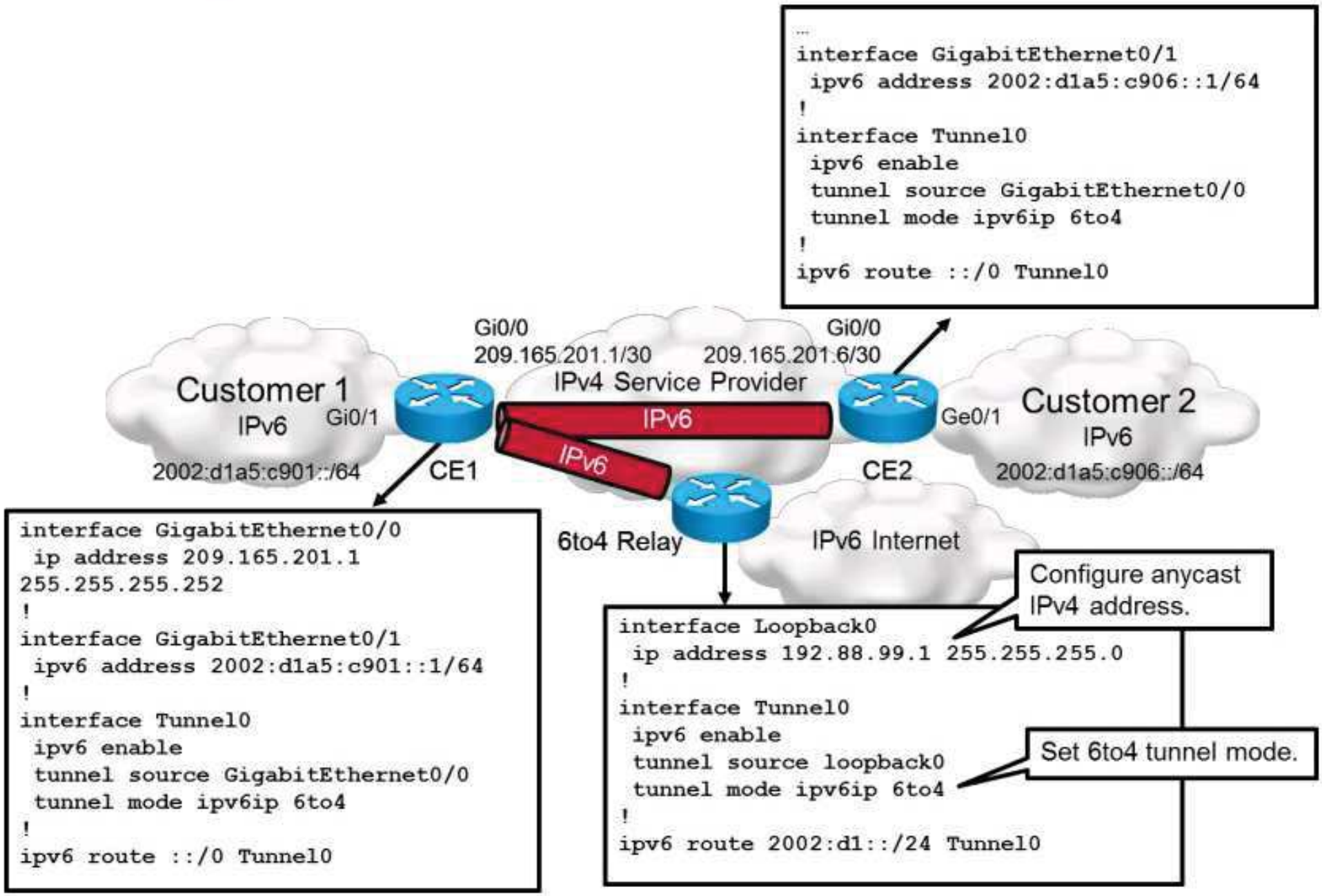
6to4 Considerations

These are 6to4 considerations:

- 6to4 should scale well and is easy to configure.
- 6to4 is obsolete because of the following:
  - Predefined IPv6 addressing: There are problems with readdressing when migrating to native IPv6.
  - Nobody guarantees the existence of 6to4 relay routers.



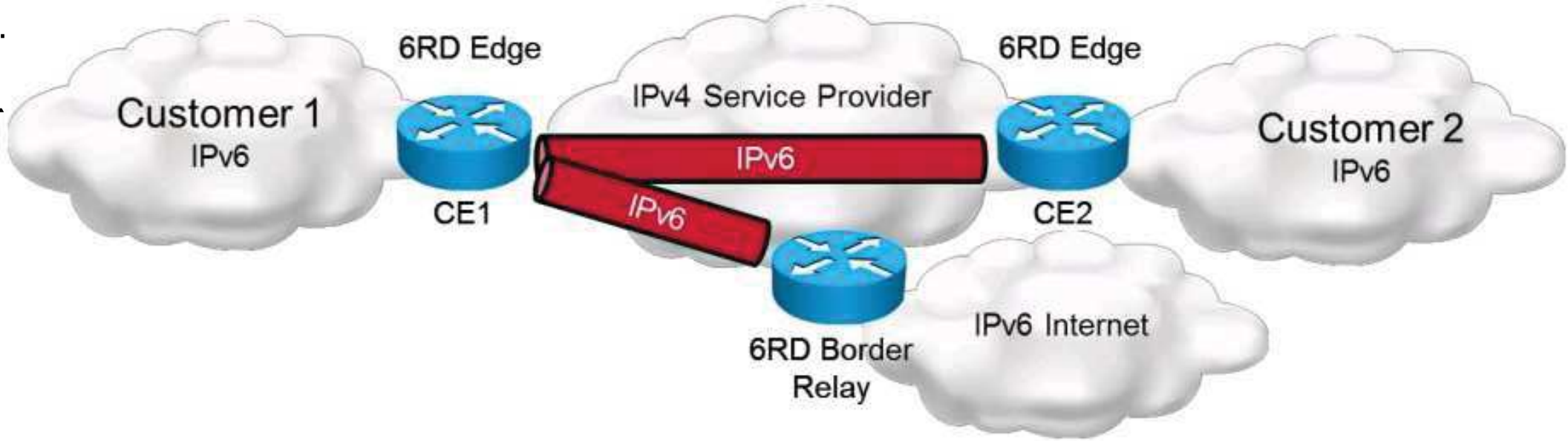
6to4 Configuration



6RD Automatic Tunnels

6RD automatic tunnels are similar to 6to4 tunneling:

- 6RD uses 6in4 encapsulation to tunnel IPv6 in IPv4.
- Customers use the service provider-assigned prefix.
- The 6RD border relay router is under service provider control, and the service provider is responsible to route traffic from customers to native IPv6 addresses.



6RD Automatic Tunnels (Cont.)

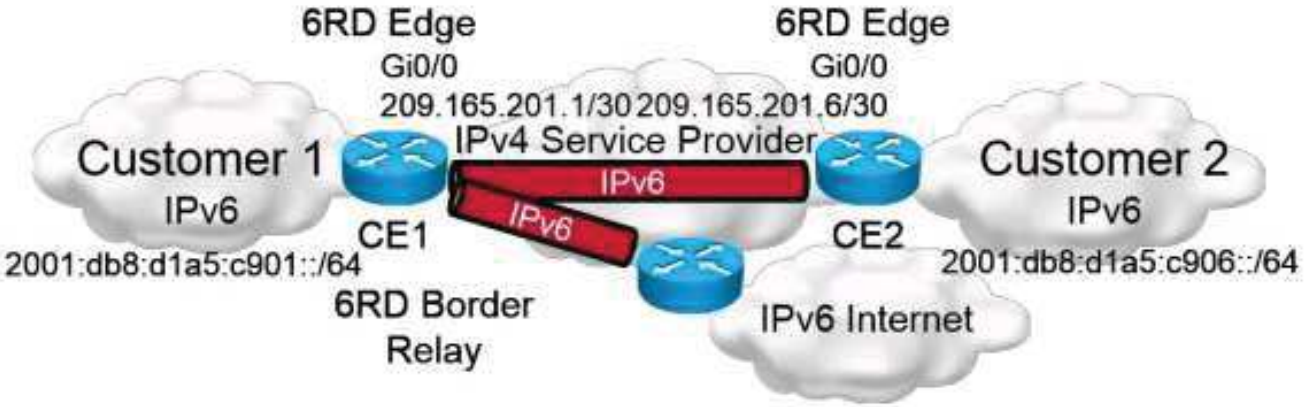
6RD addressing characteristics follow:

- The service provider selects a globally routable 6RD prefix from its IPv6 address space.
- The 6RD prefix is concatenated with the IPv4 address that is assigned to the customer. The resulting IPv6 network is assigned to the customer.
- Traffic between customers: The destination address falls within the 6RD prefix, and the tunnel destination is determined from the IPv6 prefix.
- Traffic to the IPv6 Internet: The destination address does not fall within the 6RD prefix, and traffic is sent to a preconfigured 6RD border relay.

6RD Prefix: 2001:db8::/32

6RD Edge IPv4 Address: 209.165.201.1

IPv6 Network: 2001:db8:d1a5:c901::/64



6RD Automatic Tunnels (Cont.)

A 6RD addressing example follows:

- 6RD CE routers share the first octet of the IPv4 address.
- The service provider uses 2001:db8:aa::/40 as the 6RD prefix.
- The service provider would like to assign /64 networks to customers.

6RD Prefix: 2001:db8::aa/40

6RD Edge IPv4 Address: 10.1.2.3

IPv6 Network: 2001:db8:aa01:0203::/64

6RD prefix: 2001:db8::aa/40

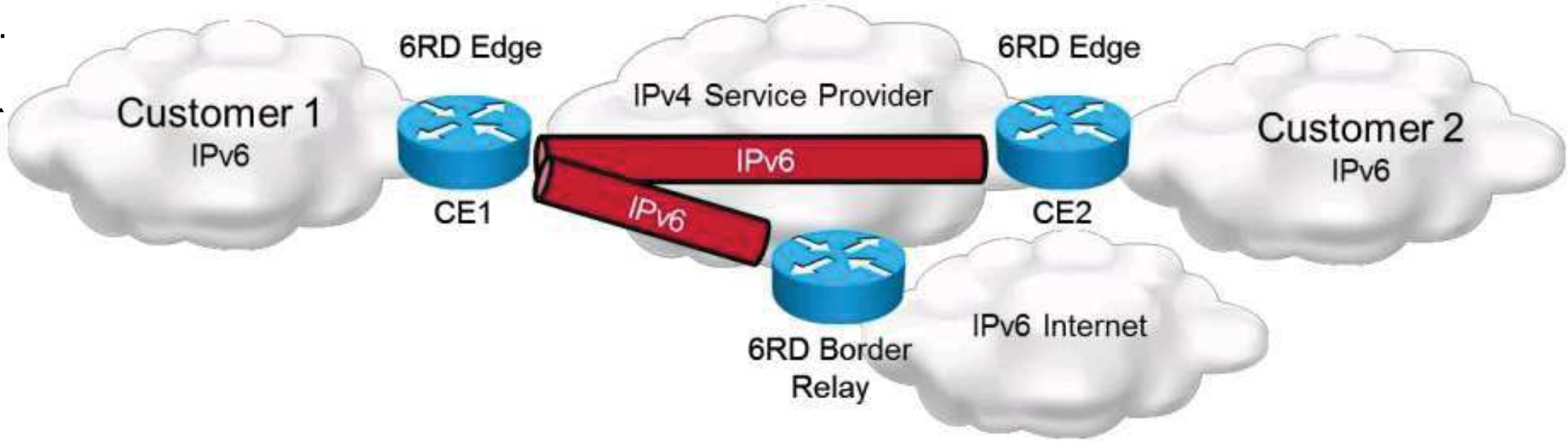
6RD Edge IPv4 address: 10.2.3.4

IPv6 Network: 2001:db8:aa02:0304::/64

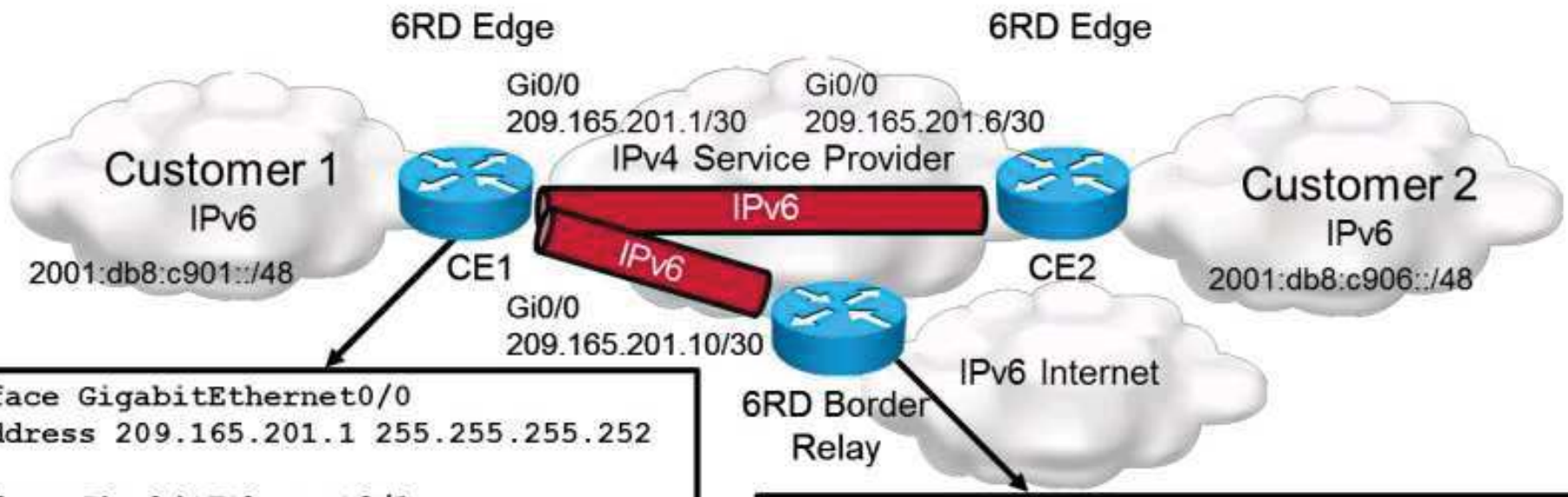
6RD Considerations

6RD considerations follow:

- 6RD allows service providers to instantly offer IPv6 services without migrating the core network.
- CE routers should be under the service provider administration.
- 6RD is supported on the Cisco ISR Series and Cisco ASR 1000 Series routers.



6RD Configuration




```
interface GigabitEthernet0/0
 ip address 209.165.201.1 255.255.255.252
!
interface GigabitEthernet0/1
 ipv6 address 2001:db8:d1a5:c901::1/64
!
interface Tunnel0
 ipv6 enable
 tunnel source GigabitEthernet0/0
 tunnel mode ipv6ip 6rd
 tunnel 6rd ipv4 prefix-len 16
 tunnel 6rd prefix 2001:db8::/32
 tunnel 6rd br 209.165.201.10
!
ipv6 route 2001:db8::/32 Tunnel0
ipv6 route ::/0 2001:db8:c90a::
```

Set the common prefix.

Set the 6RD border relay.

```
interface GigabitEthernet0/0
 ip address 209.165.201.10 255.255.255.252
!
interface Tunnel0
 ipv6 enable
 tunnel source GigabitEthernet0/0
 tunnel mode ipv6ip 6rd
 tunnel 6rd ipv4 prefix-len 16
 tunnel 6rd prefix 2001:db8::/32
!
ipv6 route 2001:db8::/32 Tunnel0
```

Set the 6RD tunnel mode.

Set the 6RD prefix.

Summary

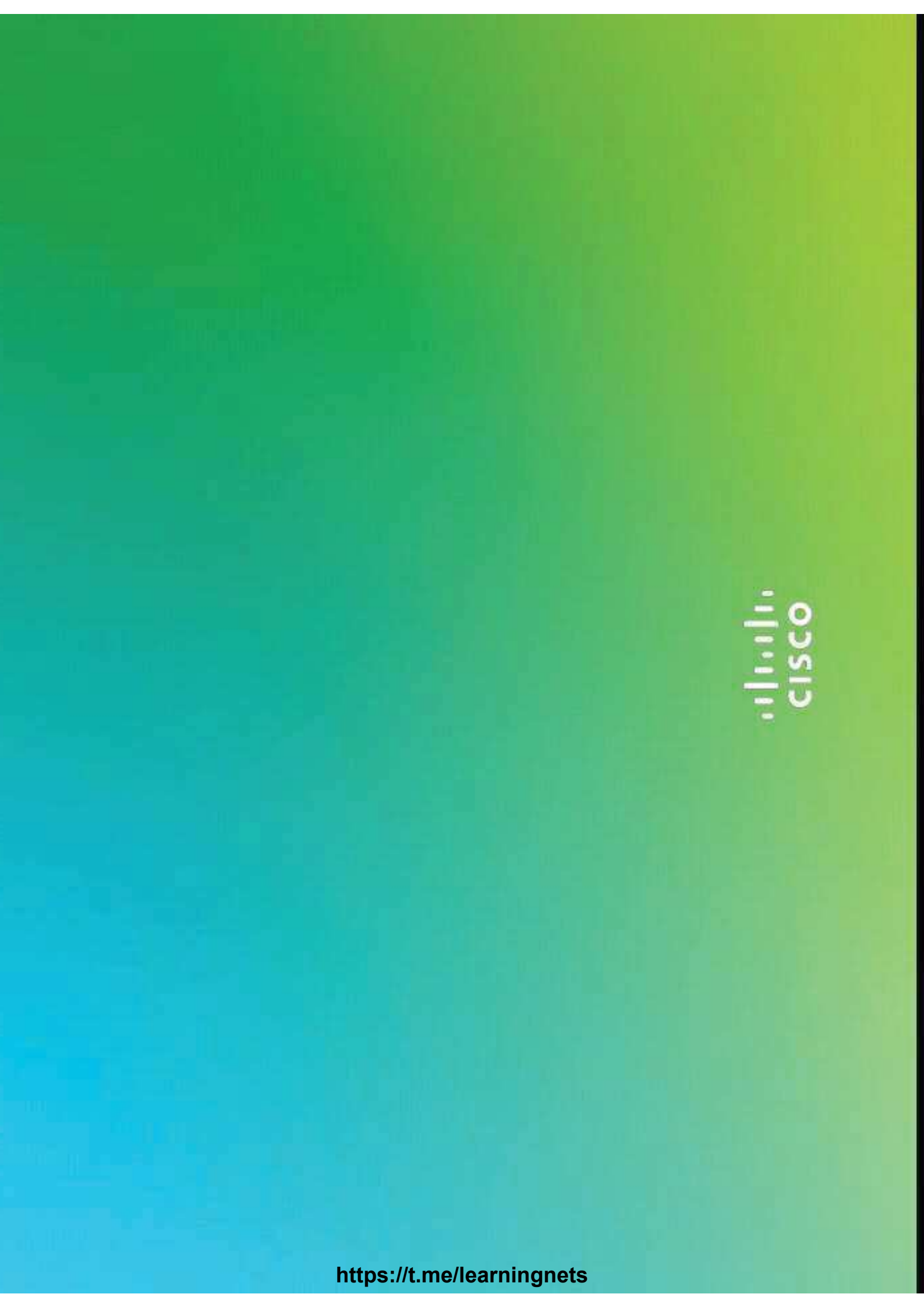
- Cisco IOS supports several ways of managing IPv4 address shortage; three of these are dual stack, CGN, and NAT64.
- Dual stack allows for coexistence of IPv6 and IPv4.
- Dual stack is less costly in the short term but more costly over the long term.
- To allow private IPv4 addresses to reach the Internet, you must implement CGN as well as dual stack.
- NAT444 means translating the IPv4 source address twice.
- Cisco supports carrier-grade NAT for deployment of the scalable NAT444 scenario.
- NAT64 means translating IPv6 addresses into IPv4 addresses and vice versa.
- NAT64 works together with (out of band) DNS64.

Summary (Cont.)

- Stateless NAT64 can be used for static one-to-one translations.
- Stateful NAT64 is useful in general deployment.
- Stateless NAT64 is only useful in specific scenarios; for general purpose, use stateful NAT64.
- NAT64 is configured very similarly to regular NAT.
- For static NAT64, you must configure static translation in addition to NAT64.
- IPv6 tunneling is a convenient and simple way to send IPv6 across IPv4-only networks.
- IPv6 tunneling can be implemented in a manual or automatic way.
- GRE by definition allows tunneling of arbitrary Layer 3 protocols, and IPv6 is no exception.

Summary (Cont.)

- 6in4 is simpler method of tunneling but allows only IPv6 traffic in the tunnel.
- To set up 6in4 tunneling, you need to specify ipv6ip tunnel mode.
- 6to4 is an old method of building IPv6 tunnels that are specially suitable for scenarios where you do not have PA or PI address space.
- 6to4 tunneling should scale well and should be easy to configure due to automatic establishment of IPv6 tunnels.
- To create a 6to4 tunnel, the tunnel mode has to be set using the **ipv6ip 6to4** command.
- 6RD tunnels are an evolution of 6to4 tunneling that is specifically suited for rapid deployment within ISP networks.
- 6RD allows service providers to instantly offer IPv6 services without migrating the core network.
- To create a 6RD tunnel, the tunnel mode has to be set using the **ipv6ip 6rd** command.



Deploying IPv6 in the Service Provider Network

Service Provider IPv6 Transition Implementations

IPv6 Service Provider Deployment

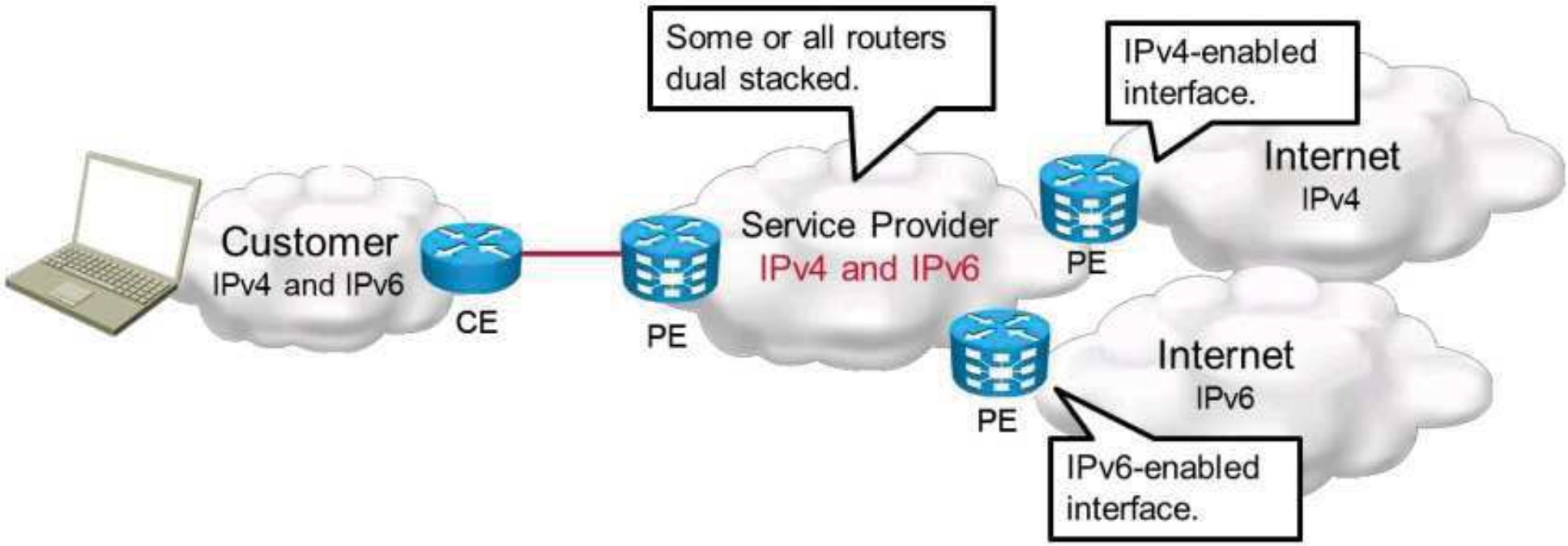
IPv6 deployment options follow:

- Existing service provider core network is IPv4-only.
- How do you offer IPv6 connectivity to customers?
  - IPv6 is in native IPv4 environments:
    - Dual stack
    - Native IPv6 over dedicated Layer 2 infrastructure
    - Tunneling of IPv6 in IPv4
  - IPv6 is in MPLS environments:
    - Cisco IPv6 Provider Edge Router over MPLS (Cisco 6PE)
    - Cisco IPv6 VPN Provider Edge Router over MPLS (Cisco 6VPE)

Dual-Stack Option

Dual-stack characteristics follow:

- CE, PE, and P routers are all capable of IPv4 and IPv6 support.
- Two IGPs are needed to support IPv4 and IPv6.
- Dual stack offers native IPv6 multicast support.



Dual Stack Pros and Cons

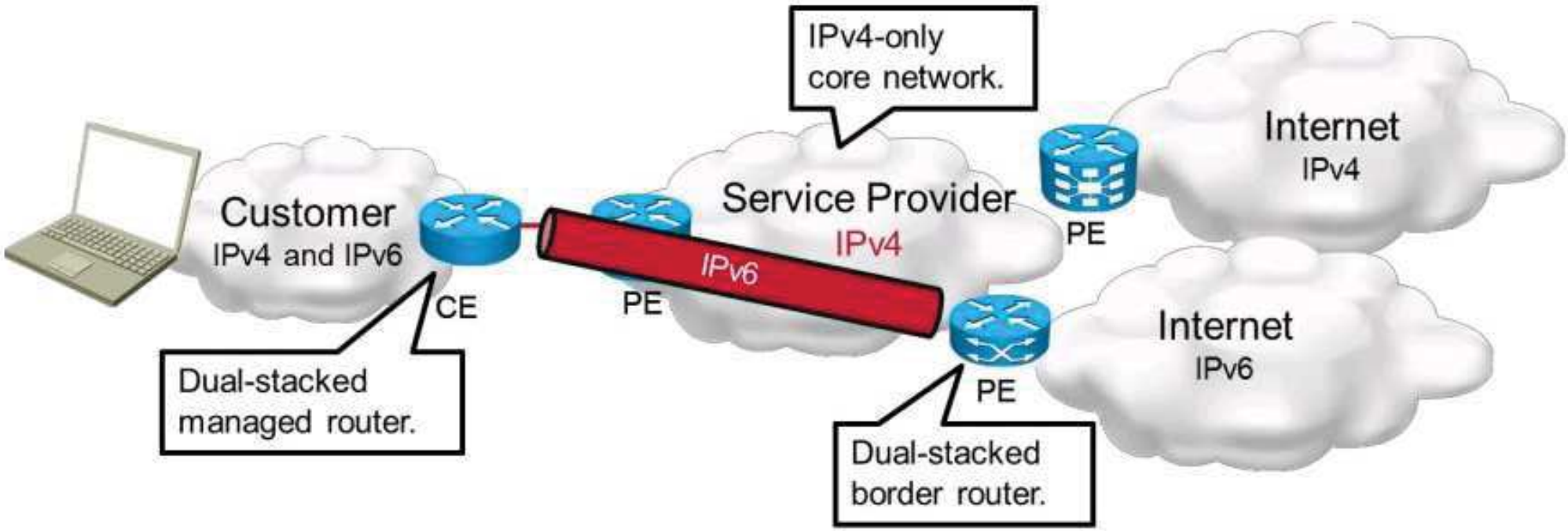
Dual-stack considerations follow:

- **Advantages:**
  - Only those systems that are required to facilitate IPv6 connectivity need to be dual stacked.
  - IPv4-only applications should continue to function without alteration.
- **Drawbacks:**
  - There is increased management of DNS, routing protocols, and address management.
  - The IPv6 software feature is compatible with existing IPv4 features.
  - The approach for larger deployments is costly.

Tunneling of IPv6 in IPv4

Tunneling of IPv6 in IPv4 characteristics follow:

- Dual stack at service provider or customer edge, IPv4 in core.
- Tunneling options:
  - Manual 6in4 tunnels, GRE, 6to4, 6RD, and so on.
- Requires managed CE routers if tunnels are established from CE devices.



Tunneling of IPv6 in IPv4 Pros and Cons

Tunneling of IPv6 in IPv4 considerations follow:

- **Advantages:**
  - This targeted IPv6 deployment method enables the delivery of IPv6 connectivity without significant changes to the network.
  - Service providers can gauge demand and observe traffic volumes before making substantial capital expenditures in IPv6 deployment.
- **Drawbacks:**
  - Manual tunnels do not scale well.
  - As the volume of tunnels scales beyond dozens into hundreds and thousands, configuration and management requirements become ungainly.
  - Additional overhead and processing exist due to tunneling and possible MTU problems.

Key Service Provider Strategies

Service providers may roll out IPv6 in phases:

- At the customer edge, allowing service provider investments in IPv6 services near paying customers and without a large investment in the core network.
- In the core, taking advantage of the economics of running a single-architecture, end-to-end in the network—in the core, access, and distribution layers.

IPv6 Services

Most service providers offer services other than simple connectivity, including the following:

- **QoS:** Provided within the core of the ISP networks and provides customers with the ability to receive a tailored response from the network for critical applications.
- **Multicasting:** Can be provided by ISPs in order to more efficiently carry rich media services such as streaming audio or IPTV.
- DNS **and** DHCP**:** End users need DNS for Internet access.
- **Multicasting:** Can be provided by ISPs in order to more efficiently carry rich media services such as streaming audio or IPTV.

IPv6 Address Allocation

Address allocation policies follow:

- **Permanent /48:** for larger sites (up to 65,536 subnets)
- **Permanent /56:** for smaller sites with only a few subnets (up to 256 subnets)
- **Permanent /64:** when only one subnet will be used
- **Short-lived /64:** for client-oriented customers that do not need stable addresses and do not run IPv6-based services
- **Permanent /128:** single, stable address for an individual device
- **Short-lived /128:** single, variable address for individual device

IPv6 Address Selection Guidelines

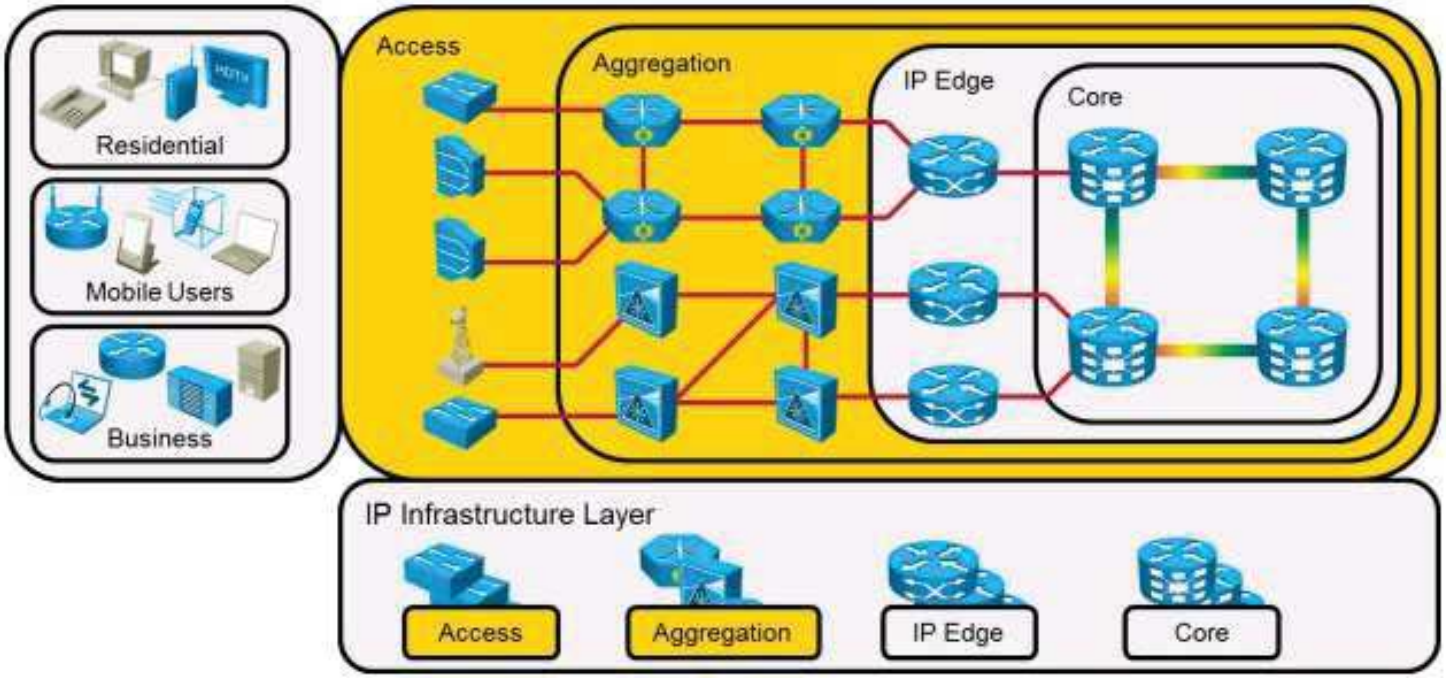
IPv6 address selection guidelines follow:

- Enterprises typically have multiple networks and desire stable addressing. For these customers, a permanent /48 or permanent /56 is used, depending on size.
- Subscribers with only one subnet will need only one permanent /64.
- Any customer that is running servers, or running peer-to-peer applications, will want a consistent prefix and therefore a permanent assignment.
- In general, users that are not running servers themselves do not need stable addresses and may desire the anonymity that comes from a dynamic /64 assignment.

IPv6 Broadband Access Services

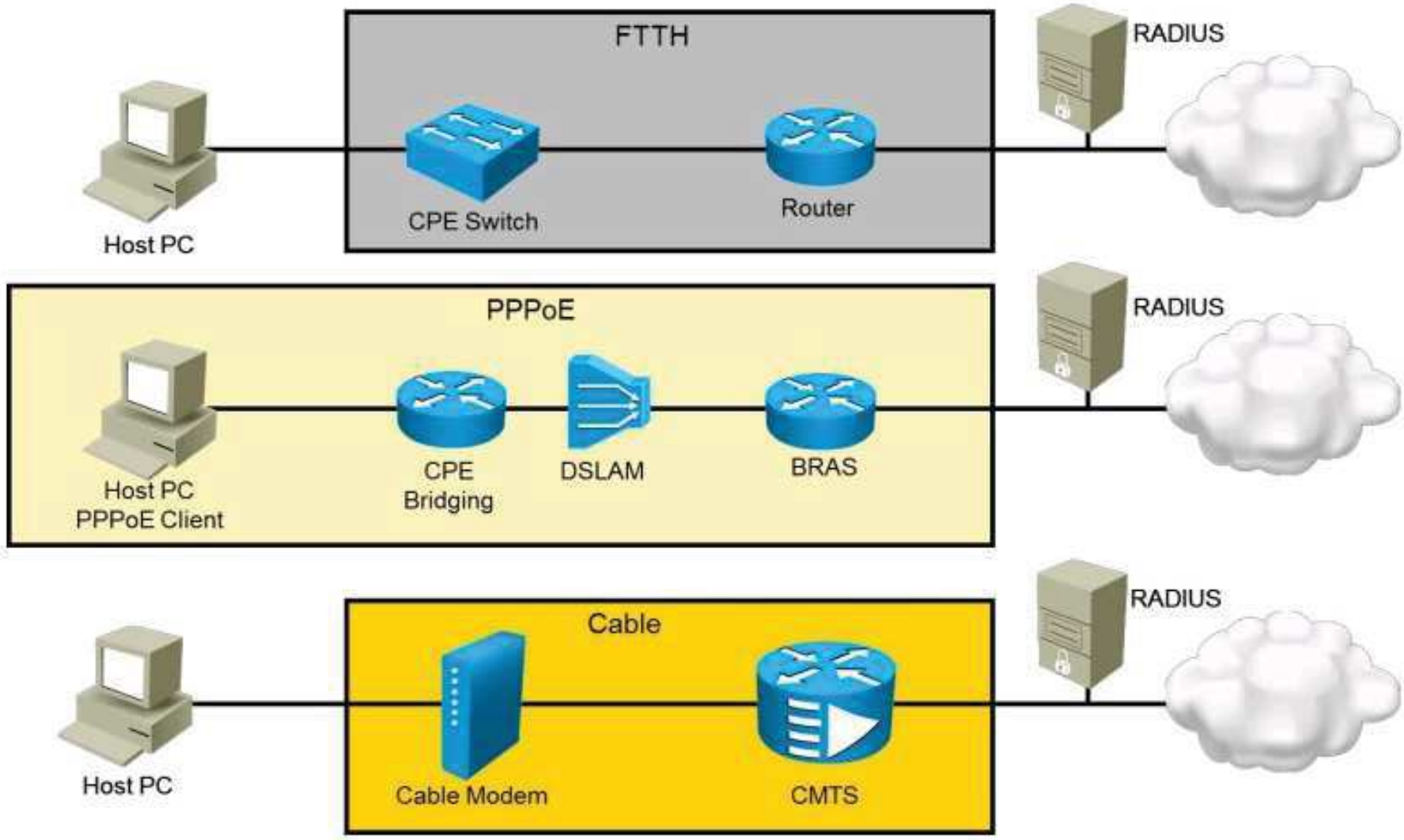
IPv6 broadband access services are in the Cisco IP NGN infrastructure layer:

- Broadband access services are part of the IP infrastructure layer of the Cisco IP NGN.
- Broadband access services are implemented on access and aggregation devices.

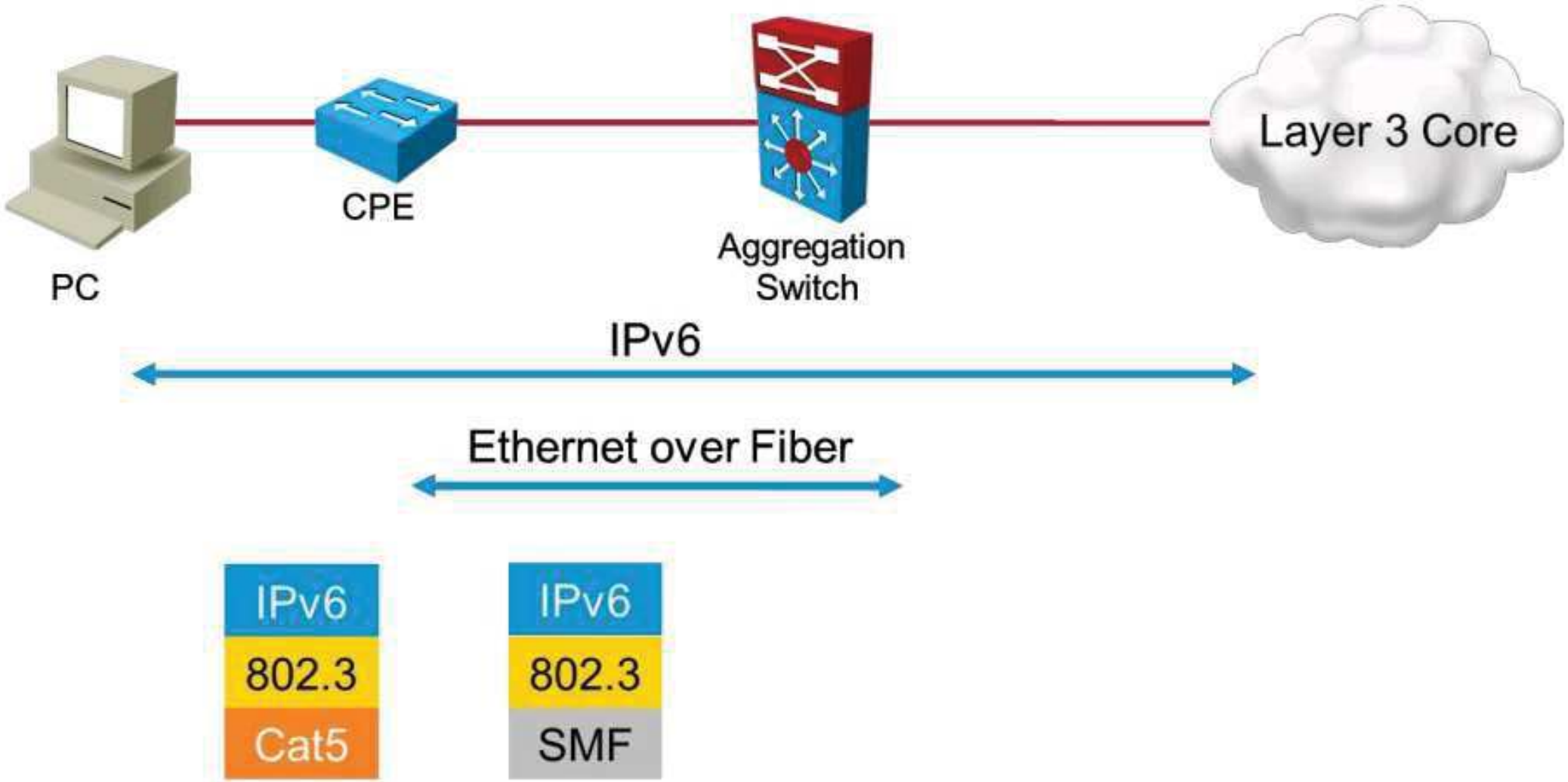


IPv6 Broadband Access Services (Cont.)

Broadband Access Last Mile Support:

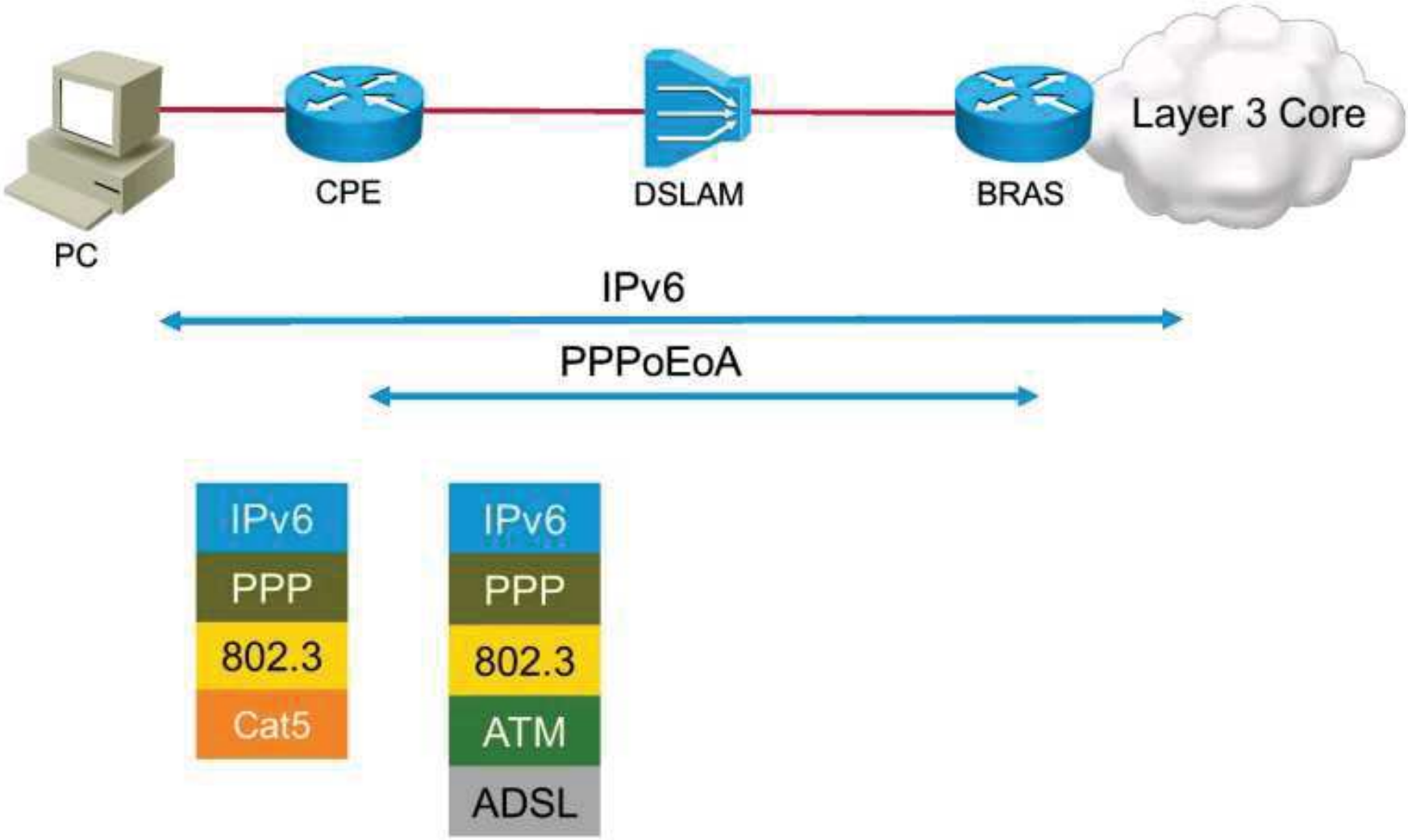


FTTH Access Architecture



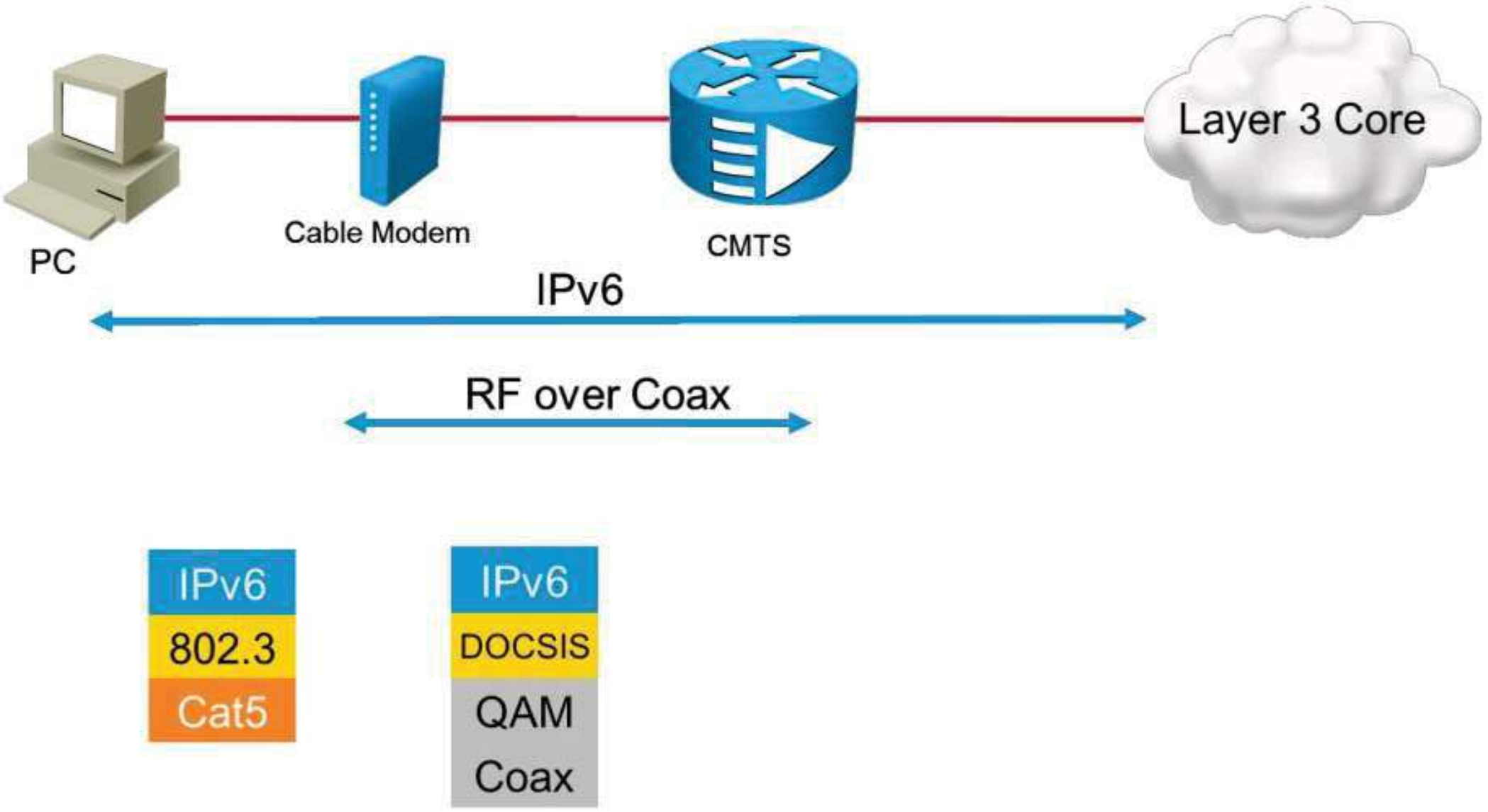
SMF = single-mode fiber

DSL Access Architecture



ADSL = Asymmetric DSL

Cable Access Architecture

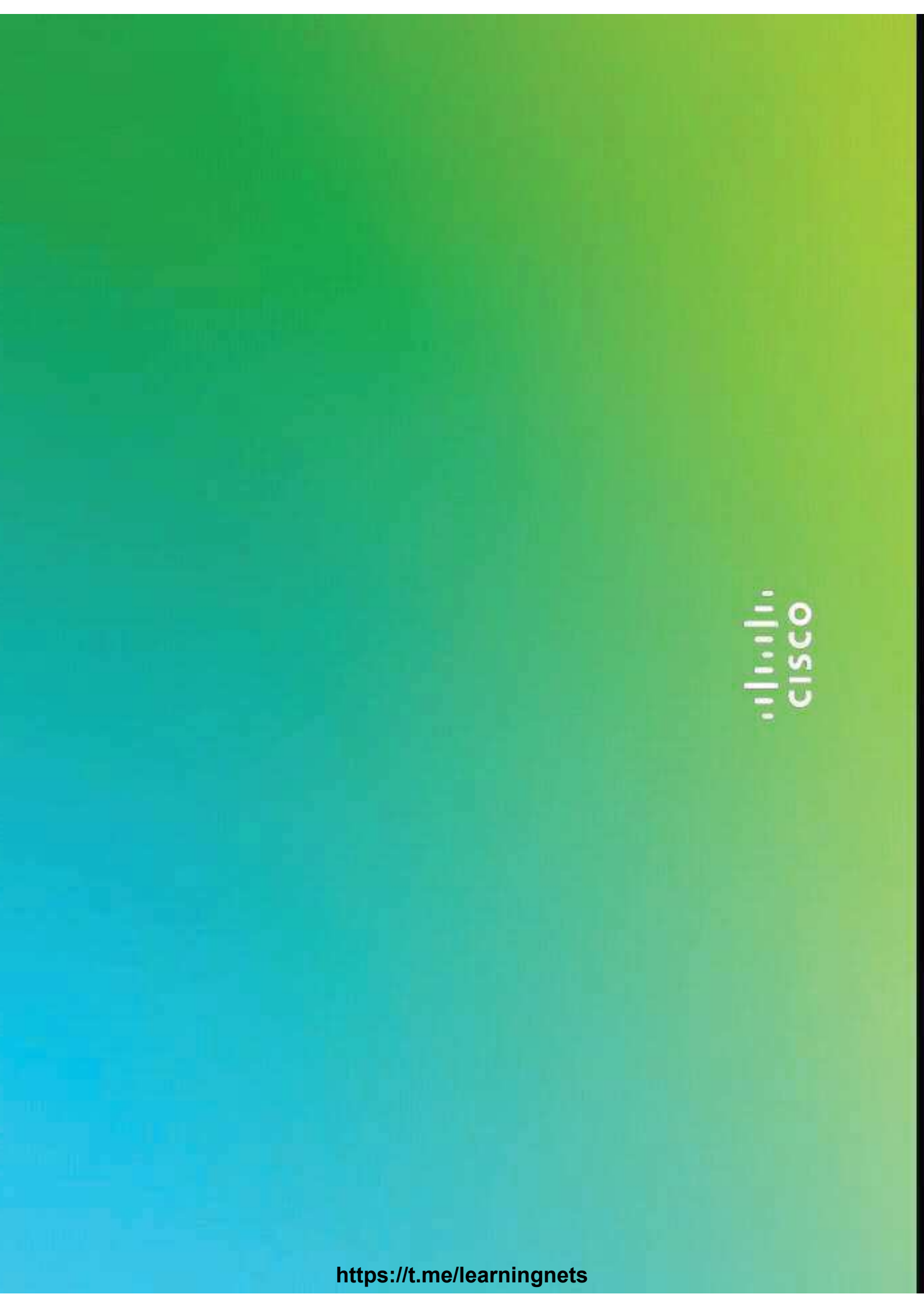


Summary

- Service providers that only have IPv4 networks can use workarounds to still provide IPv6 to end users while maintaining the IPv4-only core network.
- Implementing IPv6 side by side with IPv4 will not disrupt IPv4, but it may increase network core complexity.
- There are many advantages when deploying dual stack, but the drawbacks should be considered.
- Implementing IPv6 tunneling will increase complexity only at the edge and may also affect scalability depending on the type of tunneling.
- Tunneling of IPv6 in IPv4 enables the delivery of IPv6 connectivity without significant changes to the network, but additional overhead and processing exist.
- Service providers will want to invest step by step into new technologies at the edge and avoid large investments into the core by keeping changes away from it.

Summary (Cont.)

- Network services that may be affected by the introduction of IPv6 are QoS, multicast, DNS, DHCP, and managed services.
- Assignment of IPv6 addresses is different than assignment of IPv4 addresses because IPv6 addresses are not as scarce. Policies must therefore be different to allow customers to take advantage of all IPv6 features.
- ISPs should assign address space based on the number of segments that are required and not the number of nodes that are present.
- Most Layer 2 technologies support IPv6, and you can easily deploy IPv6 over the following:
  - FTTH access architecture
  - DSL access architecture
  - Cable access architecture



Module Summary

- Several services, such as DNS, DHCP, multicast, and QoS, are available for IPv6.
- Several transition technologies, such as dual stack, tunneling, and NAT64, are available to migrate to IPv6.
- The recommended approach for transition to IPv6 is dual stack in the entire network or IPv4 in the core and dual stack at the edge and use of IPv6 tunneling.