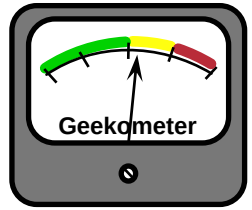
Deploying IP Multicast

Session RST-2051

Agenda



Cisco.com

- **Basic Multicast Engineering**
- **Advanced Multicast Engineering**

Basic Multicast Engineering

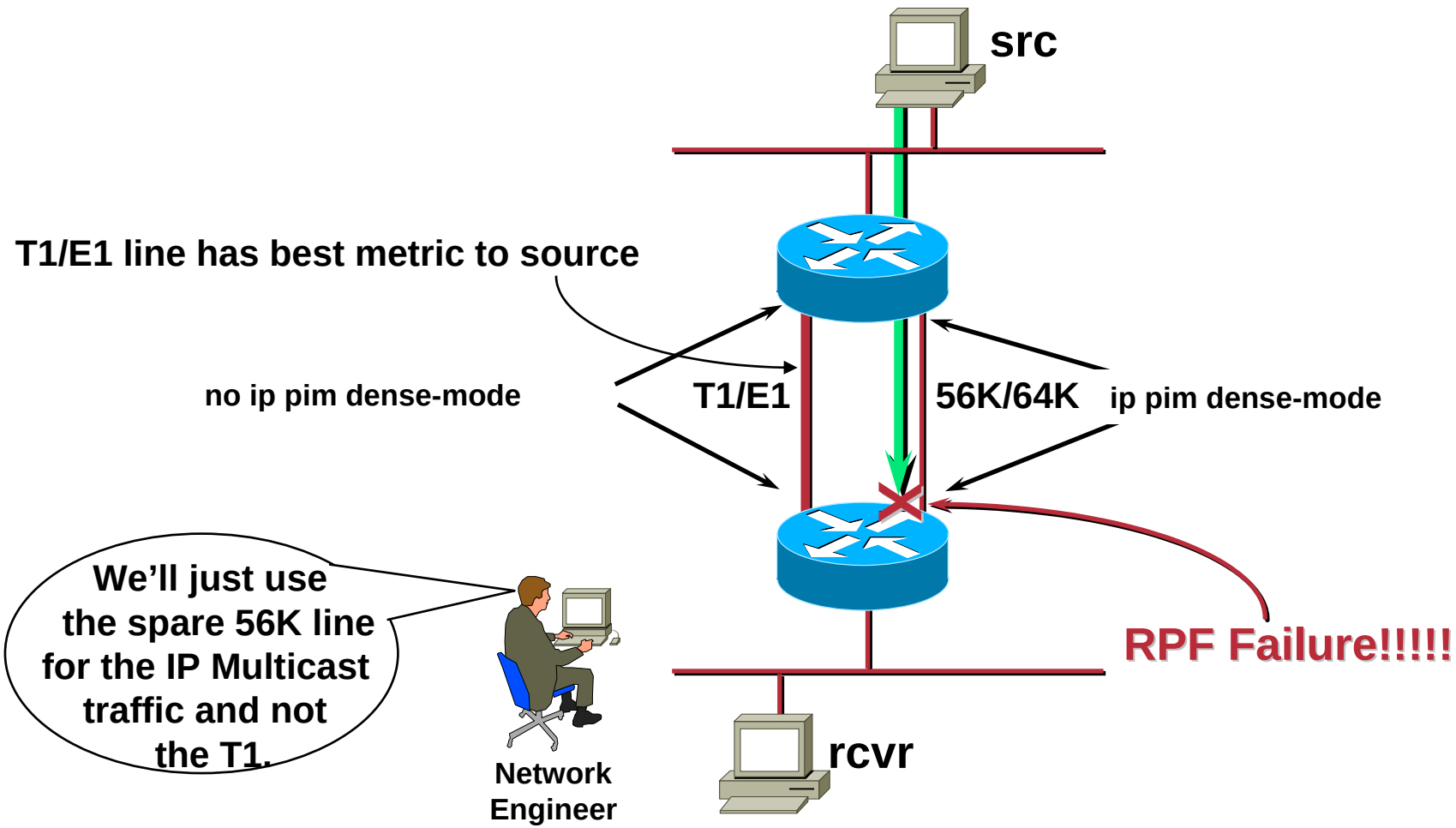
- **PIM Configuration Steps**
- **Which Mode: Sparse or Dense?**
- **RP Engineering**

PIM Configuration Steps

- Enable Multicast Routing on *every* router
- Configure *every* interface for PIM
- Configure the RP
  - Using Auto-RP or BSR
    - Configure certain routers as Candidate RP(s)
    - All other routers automatically learn elected RP
  - Anycast/Static RP addressing
    - RP address must be configured on every router
    - Note: Anycast RP requires MSDP

Configure PIM on Every Interface

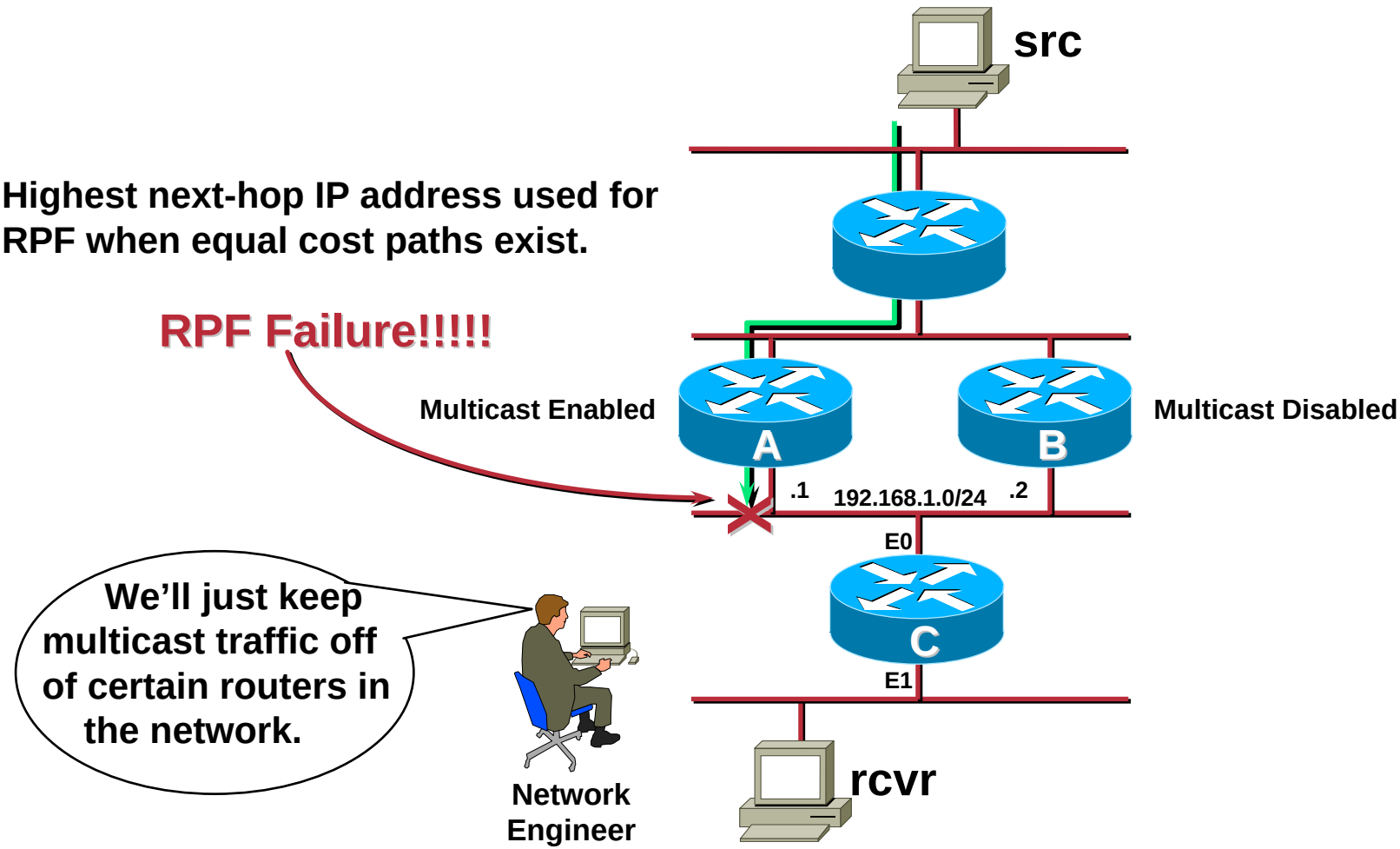
Classic Partial Multicast Cloud Mistake #1



Configure PIM on Every Router

Classic Partial Multicast Cloud Mistake #2

Highest next-hop IP address used for RPF when equal cost paths exist.



Basic Multicast Engineering

- PIM Configuration Steps
- **Which Mode: Sparse or Dense?**
- RP Engineering

Which Mode—Sparse or Dense

- Dense mode
  - Flood and Prune behavior very inefficient
    - Can cause problems in certain network topologies
  - Creates (S, G) state in EVERY router
    - Even when there are no receivers for the traffic
  - Complex Assert mechanism
  - Mixed control and data planes
    - Results in (S, G) state in every router in the network
    - Can result in non-deterministic topological behavior
      - *Read: It can black-hole traffic and/or melt down your network!*
  - Primarily usage:
    - Testing a router's performance in the lab

Which Mode—Sparse or Dense

- Sparse mode
  - Must configure a Rendezvous Point (RP)
  - Very efficient
    - Uses Explicit Join model
    - Traffic only flows to where it's needed
  - Separated control and data planes
    - Router state only created along flow paths
    - Deterministic topological behavior
  - Scales well
    - Works for both sparsely or densely populated networks

Which Mode—Sparse or Dense

CONCLUSION

“Sparse mode Good! Dense mode Bad!”

Source: ***“The Caveman’s Guide to IP Multicast”, ©2000, R. Davis***

Group Mode vs. Interface Mode

- **Group & Interface mode are independent.**
  - **Interface Mode**
    - **Determines how the *interface* operates when sending/receiving multicast traffic.**
  - **Group Mode**
    - **Determines whether the group is Sparse or Dense.**

Group Mode

- Group mode is controlled by local RP info
  - Local RP Information
    - Stored in the Group-to-RP Mapping Cache
    - May be statically configured or learned via Auto-RP or BSR
  - If RP info exists, Group = Sparse
  - If RP info does not exist, Group = Dense
  - Mode Changes are automatic.

    i.e. if RP info is lost, Group falls back to Dense.

Configuring Interface

- **Interface Mode Configuration Commands**
  - **Enables multicast forwarding on the interface.**
  - **Controls the *interface's* mode of operation.**

`ip pim dense-mode`

- **Interface mode is set to Dense mode operation.**

`ip pim sparse-mode`

- **Interface mode is set to Sparse mode operation.**

`ip pim sparse-dense-mode`

- **Interface mode is determined by the Group mode.**
  - **If Group is Dense, interface operates in Dense mode.**
  - **If Group is Sparse, interface operates in Sparse mode.**

Basic Multicast Engineering

- PIM Configuration Steps
- Which Mode: Sparse or Dense?
- RP Engineering

RP Engineering

- **RP Configuration Methods**
- General RP Recommendations
- Avoiding DM Fallback
- Using Multiple Group Ranges

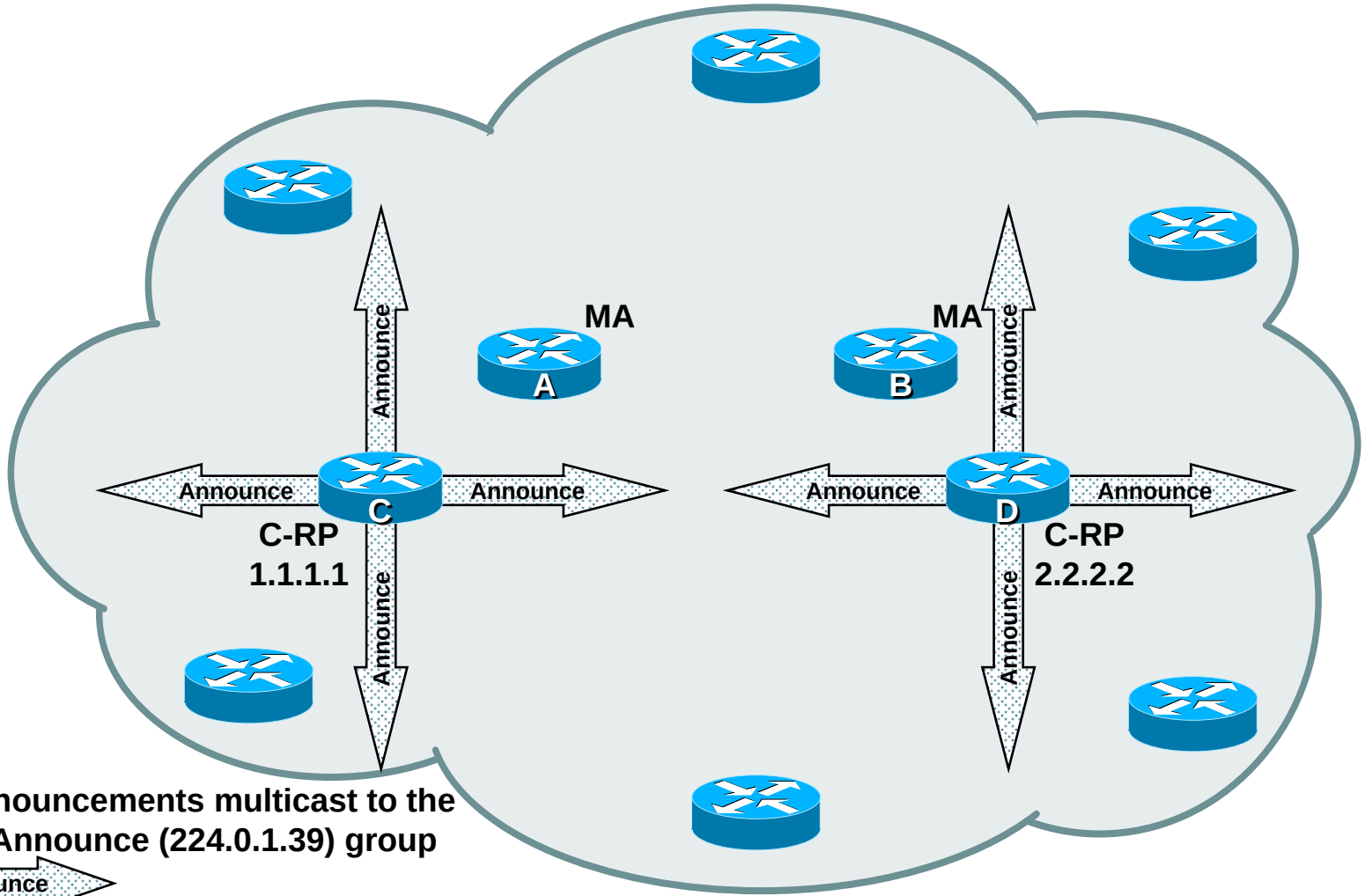
RP Configuration Methods

- **Static**
- **Auto-RP**
- **BSR**
- **Anycast-RP's**

Static RP's

- Hard-coded RP address
  - When used, must be configured on every router
  - All routers must have the same RP address
  - RP fail-over not possible
    - Exception: If Anycast RPs are used. (More on that later.)
  - Group can never fall back into Dense mode.

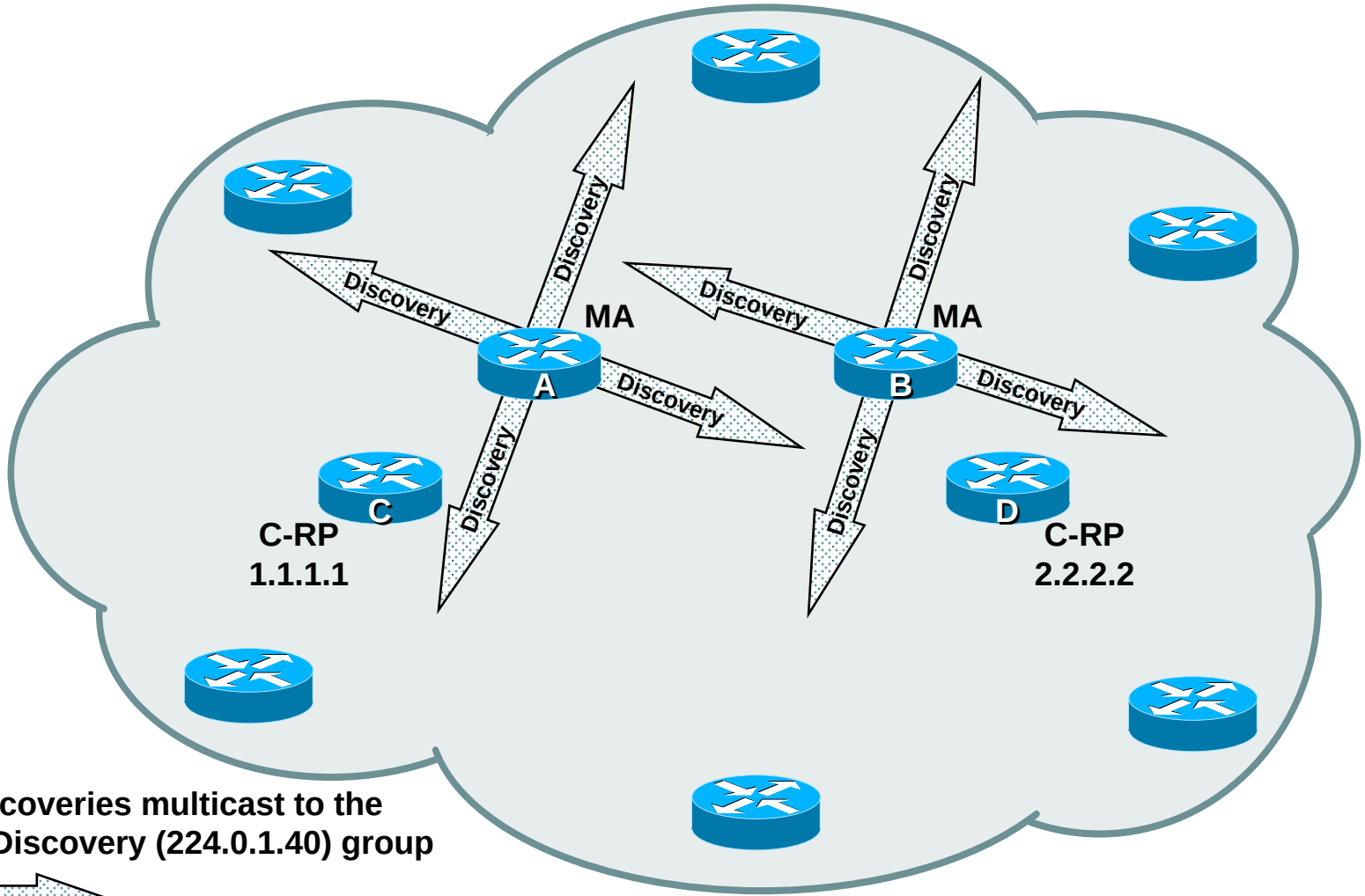
Auto-RP Overview



RP-Announcements multicast to the Cisco Announce (224.0.1.39) group

Announce

Auto-RP Overview

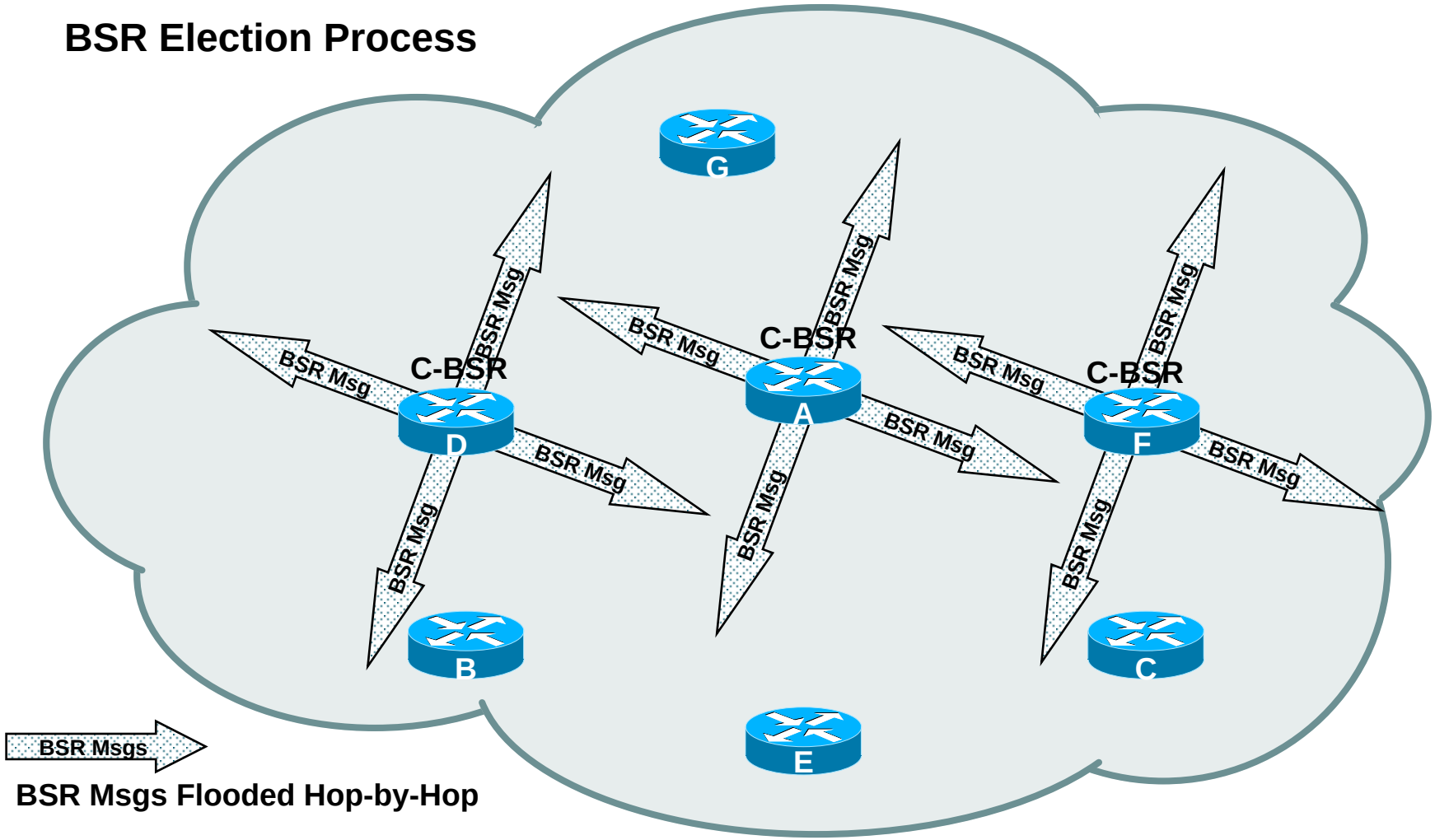


RP-Discoveries multicast to the Cisco Discovery (224.0.1.40) group

Discovery

BSR Overview

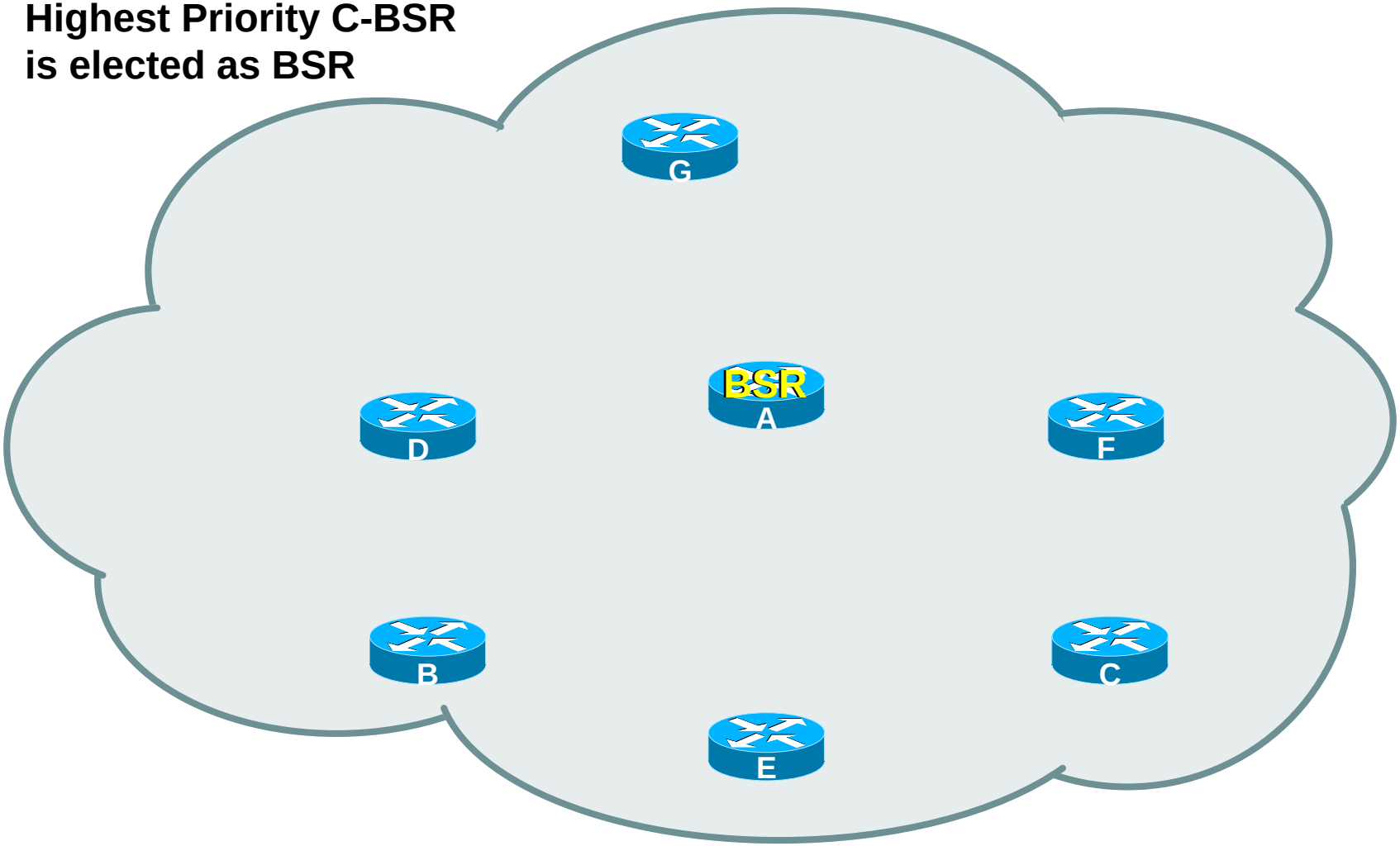
BSR Election Process



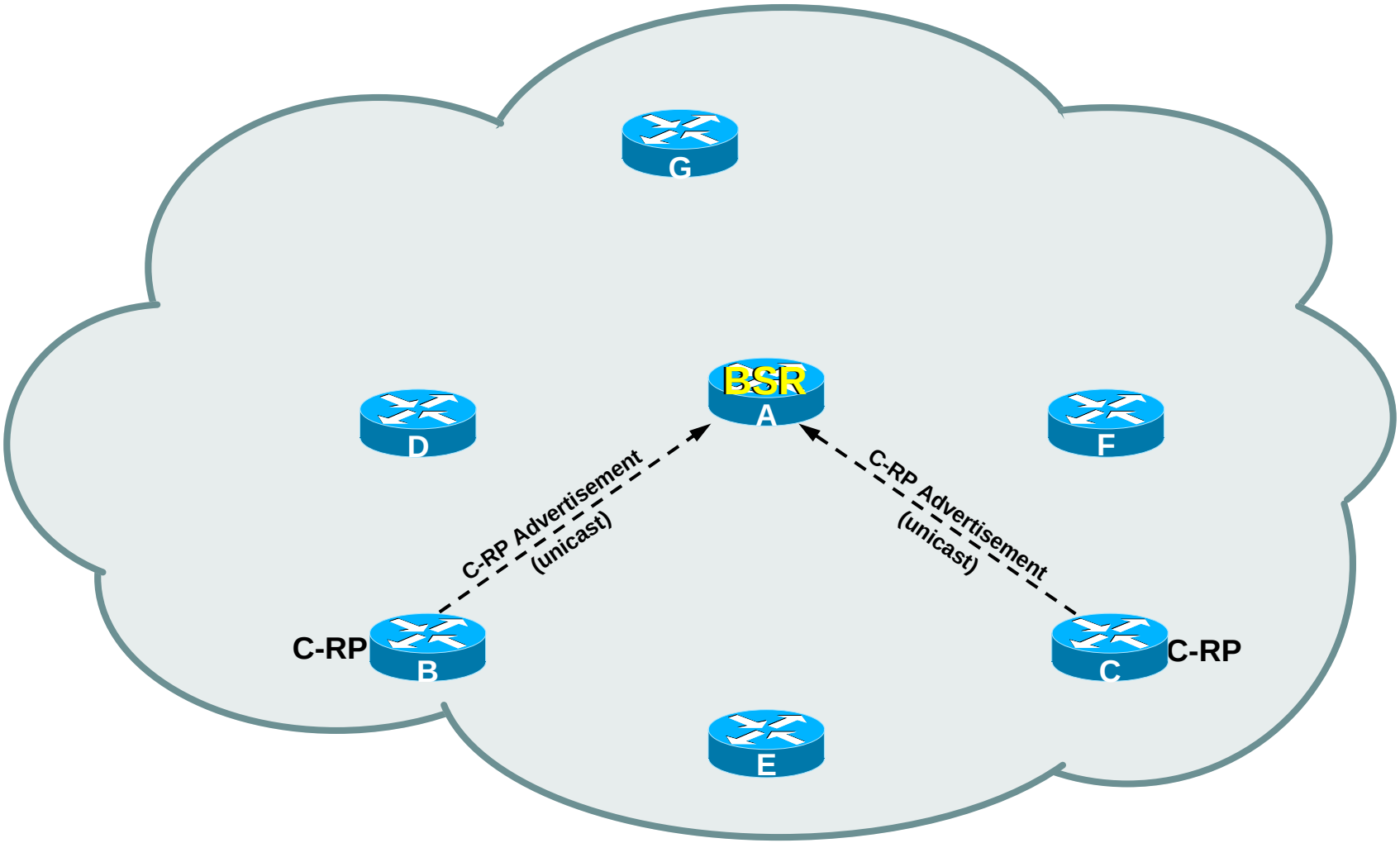
BSR Msgs Flooded Hop-by-Hop

BSR Overview

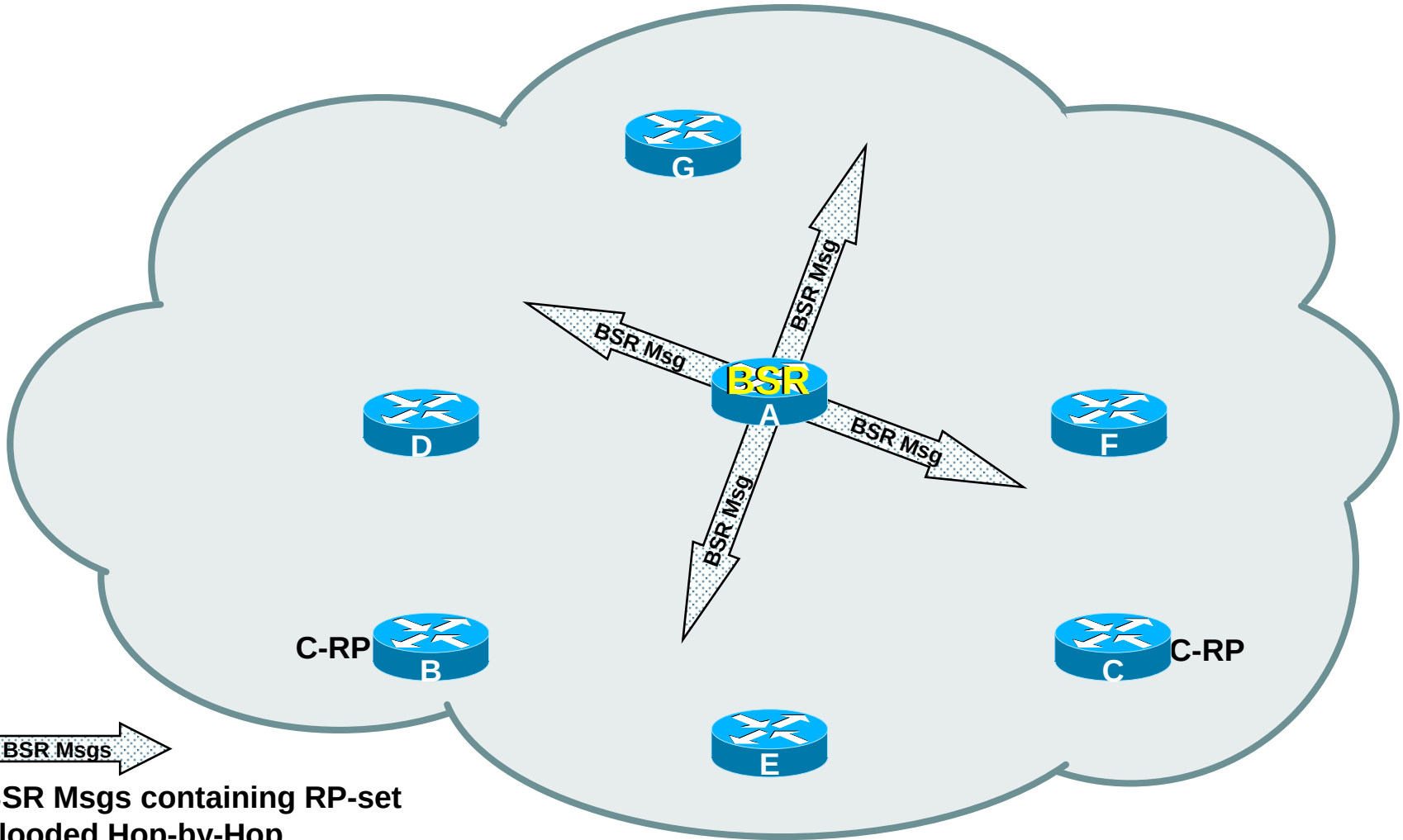
Highest Priority C-BSR is elected as BSR



BSR Overview

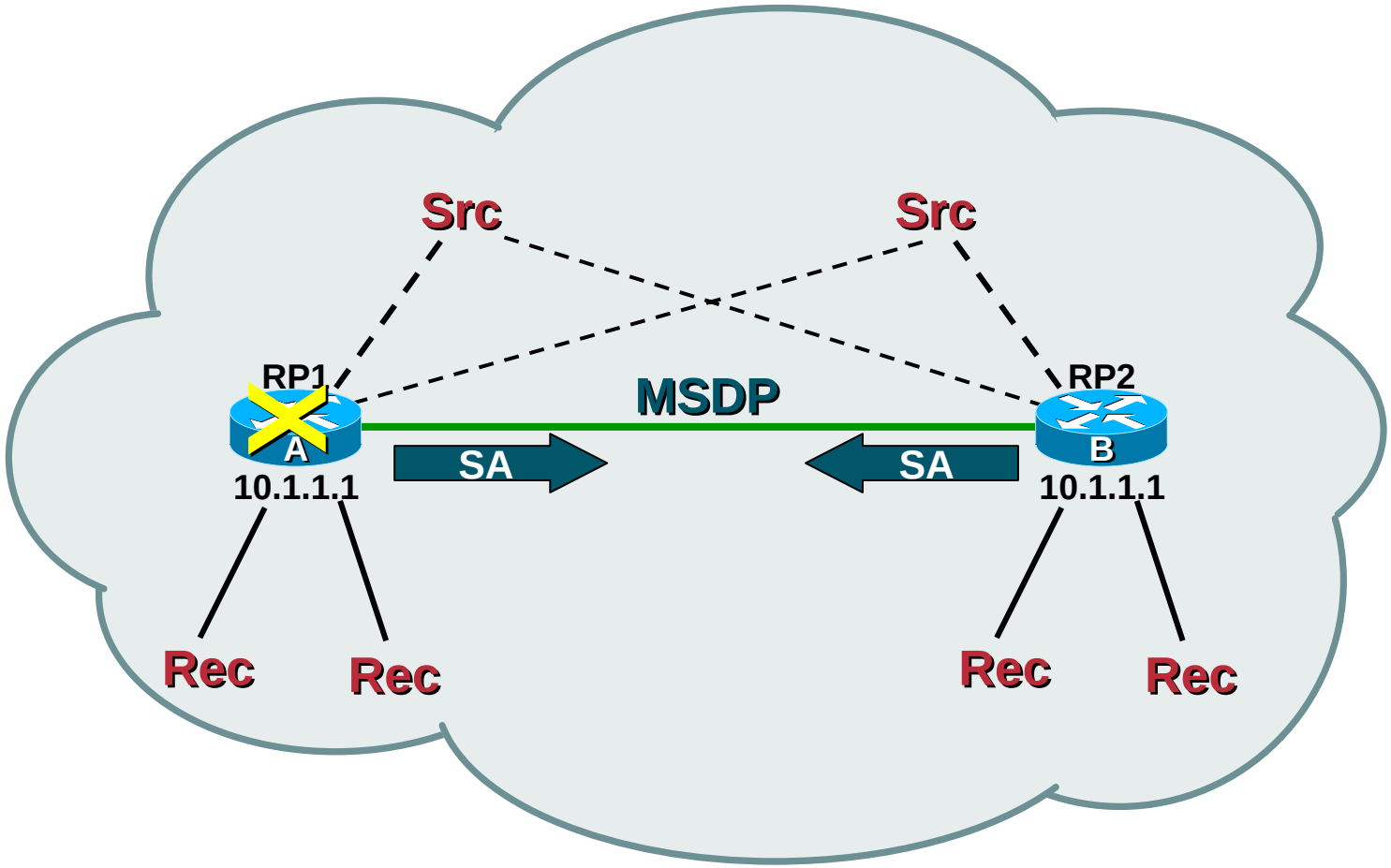


BSR Overview

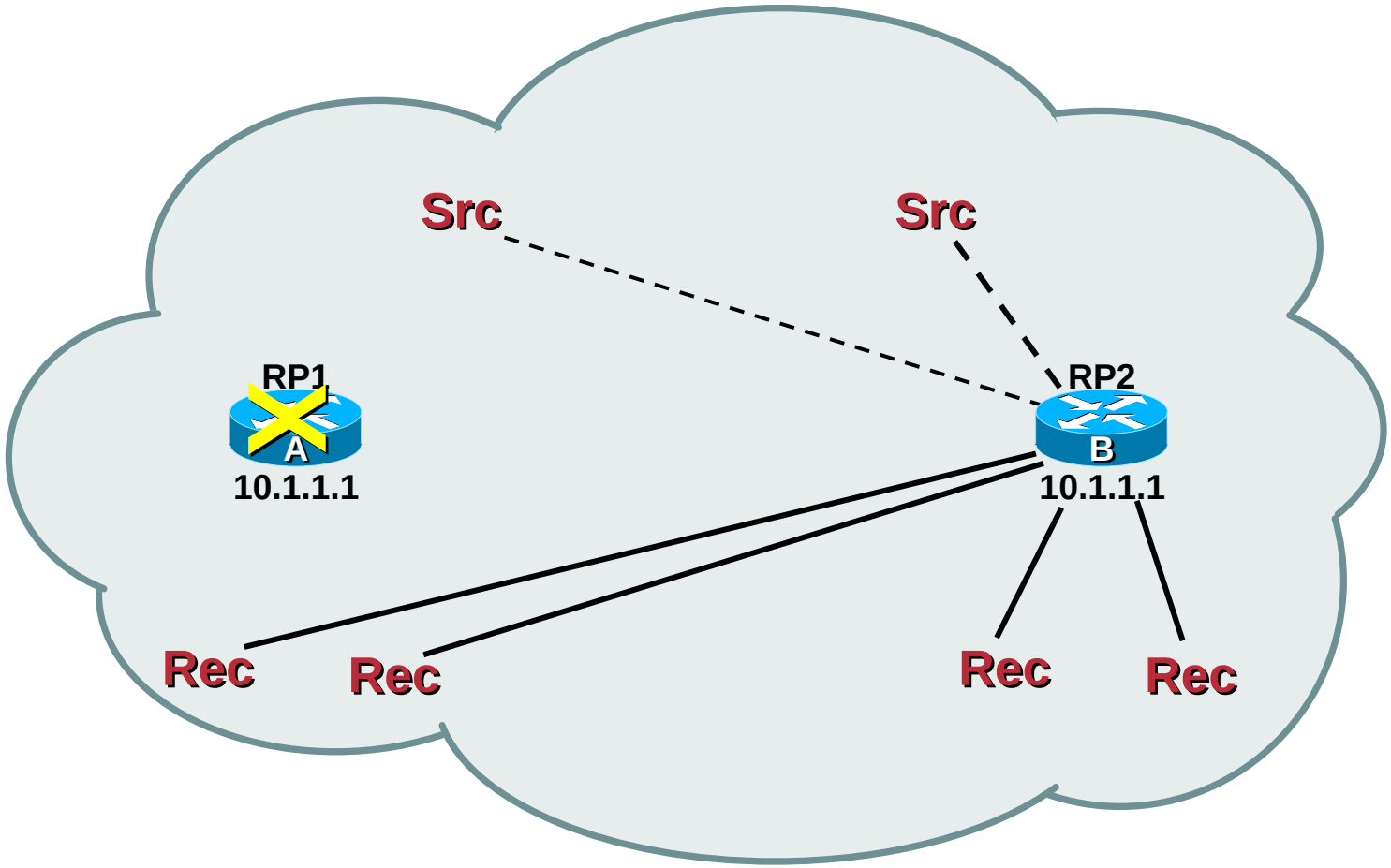


BSR Msgs containing RP-set
Flooded Hop-by-Hop

Anycast RP—Overview



Anycast RP—Overview



RP Engineering

- RP Configuration Methods
- General RP Recommendations
- Avoiding DM Fallback
- Using Multiple Group Ranges

General RP Recommendations

- Use Anycast RP's:
  - When network must connect to Internet or
  - When rapid RP failover is critical
- Pros
  - Fastest RP Convergence method
  - Required when connecting to Internet
- Cons
  - Requires more configuration
  - Requires use of MSDP between RP's

General RP Recommendations

- Use Auto-RP
  - When minimum configuration is desired and/or
  - When maximum flexibility is desired
- Pros
  - Most flexible method
  - Easiest to maintain
- Cons
  - Increased RP Failover times vs Anycast
  - Special care needed to avoid DM Fallback
    - Some methods greatly increase configuration

General RP Recommendations

- Use BSR:
  - When Static/Anycast RP's cannot be used and
  - When maximum interoperability is needed
- Pros
  - Interoperates with all Vendors
- Cons
  - Increased RP Failover times vs Anycast
  - Special care needed to avoid DM Fallback
    - Some methods greatly increase configuration
  - Not as "field-proven" as other methods

Dense Mode Fallback

- Caused by loss of local RP information.
  - Entry in Group-to-RP mapping cache times out.
- Can happen when:
  - All C-RP's fail.
  - Auto-RP/BSR mechanism fails.
    - Generally a result of network congestion.
- Group is switched over to Dense mode.
  - Dense mode state is created in the network.
  - Dense mode flooding begins if interfaces configured as `ip pim sparse-dense-mode`.

Dense Mode Fallback

Avoiding Dense Mode Fallback

To always guarantee Sparse mode operation (and avoid falling back to Dense mode), make sure that every router *always* knows of an RP for every group.

Avoiding DM Fallback – Current Workaround

- **Define an "RP-of-last-resort"**
  - **Configure as a Static RP on every router**
    - **Will only be used if all Candidate-RP's fail**
    - **Can be a dummy address or local Loopback**
      - **Recommendation: Use local Loopback on each router**
  - ***MUST use ACL to avoid breaking Auto-RP!***

```
ip pim rp-address <RP-of-last-resort> 10

access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

Avoiding DM Flooding – Future

- New IOS global command

  `ip pim autorp-listener`

- Added support for Auto-RP Environments
  - Modifies interface behavior
    - Interface always uses DM for Auto-RP groups
    - Permits use of `ip pim sparse-mode` interfaces and Auto-RP.
  - Prevents DM Flooding
    - When `ip pim sparse-mode` used on interfaces.
  - *Does not prevent DM Fallback!*
- Available soon

Avoiding DM Flooding – Future

- Deploying `ip pim autorp-listener`
  - Must be configured on every router.
  - Use RP-of-last-resort on older IOS versions until upgraded
    - Assign local Loopback as RP-of-last-resort on each router.
    - Example

```
ip pim rp-address <local_loopback> 10

access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

Avoiding DM *Fallback* – Future

- **New IOS global command**

```
no ip pim dm-fallback
```

- ***Totally prevents DM Fallback!!***
  - **No DM Flooding since all state remains in SM**
- **Default RP Address = 0.0.0.0 [nonexistent]**
  - **Used if all RP's fail.**
    - **Results in loss of Shared Tree.**
    - **All SPT's remain active.**
- **Available soon**

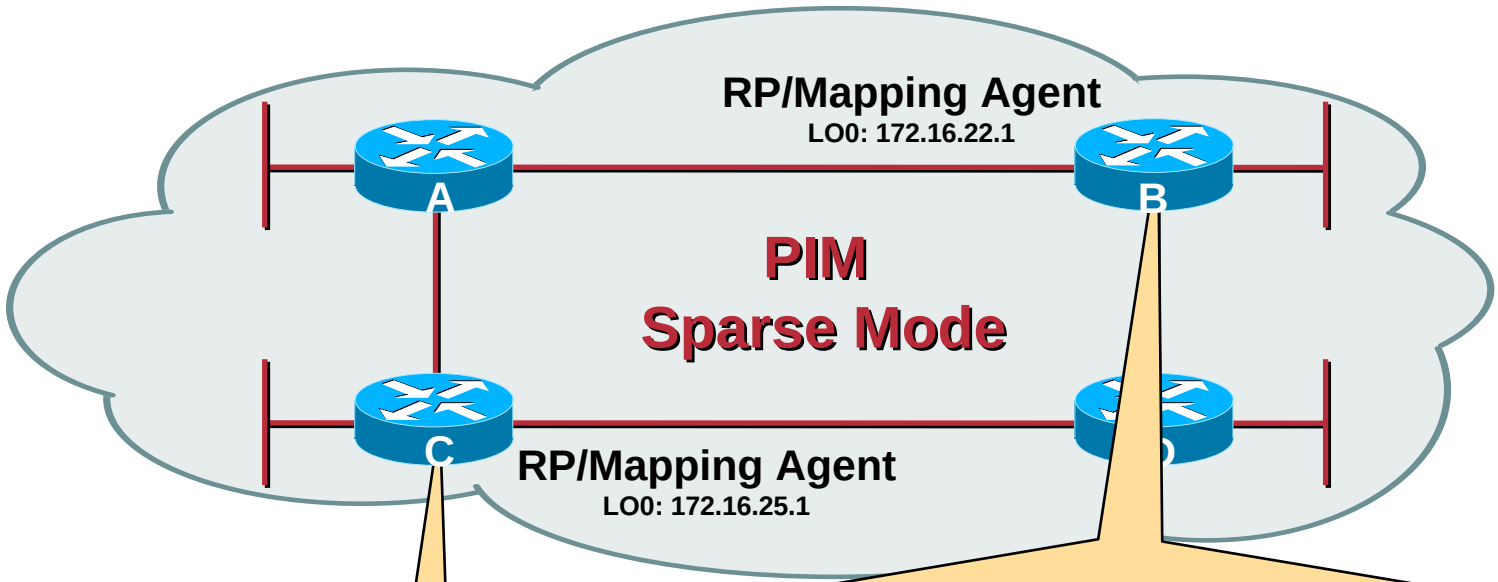
RP Engineering

- RP Configuration Methods
- General RP Recommendations
- Avoiding DM Fallback
- **Using Multiple Group Ranges**

Using Multiple Group Ranges

- **Definition:**
  - **Different RPs for different group ranges**
- **Often used to:**
  - **Directly connect an RP to group sources**
    - **Assumes Few-to-many application model**
  - **Split up RP workload over multiple RP's**
  - **Provide different Shared Tree topologies**
    - **Used with 'spt-threshold = infinity"**
- **Caveats:**
  - **Try to avoid overlapping group ranges**
    - **Can cause unexpected RP election results**

Using Multiple Group Ranges

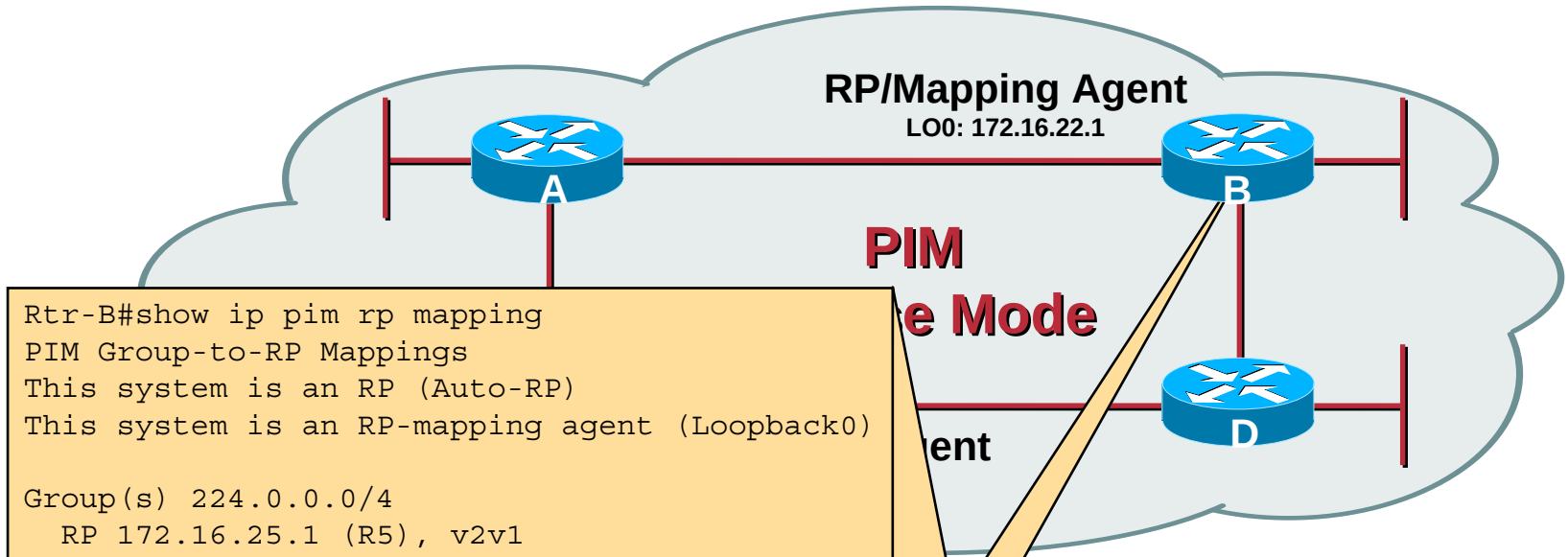



```
ip pim send-rp-announce Loopback0 scope 16 group-list 20
ip pim send-rp-discovery Loopback0 scope 16
!
access-list 20 permit 224.0.0.0 15.255.255.255
access-list 20 permit 239.255.0.0 0.0.255.255
```

```
ip pim send-rp-announce Loopback0 scope 16 group-list 20
ip pim send-rp-discovery Loopback0 scope 16
!
access-list 20 permit 224.0.0.0 15.255.255.255
access-list 20 permit 225.0.0.0 0.255.255.255
```

C-RP/Mapping Agent Configs

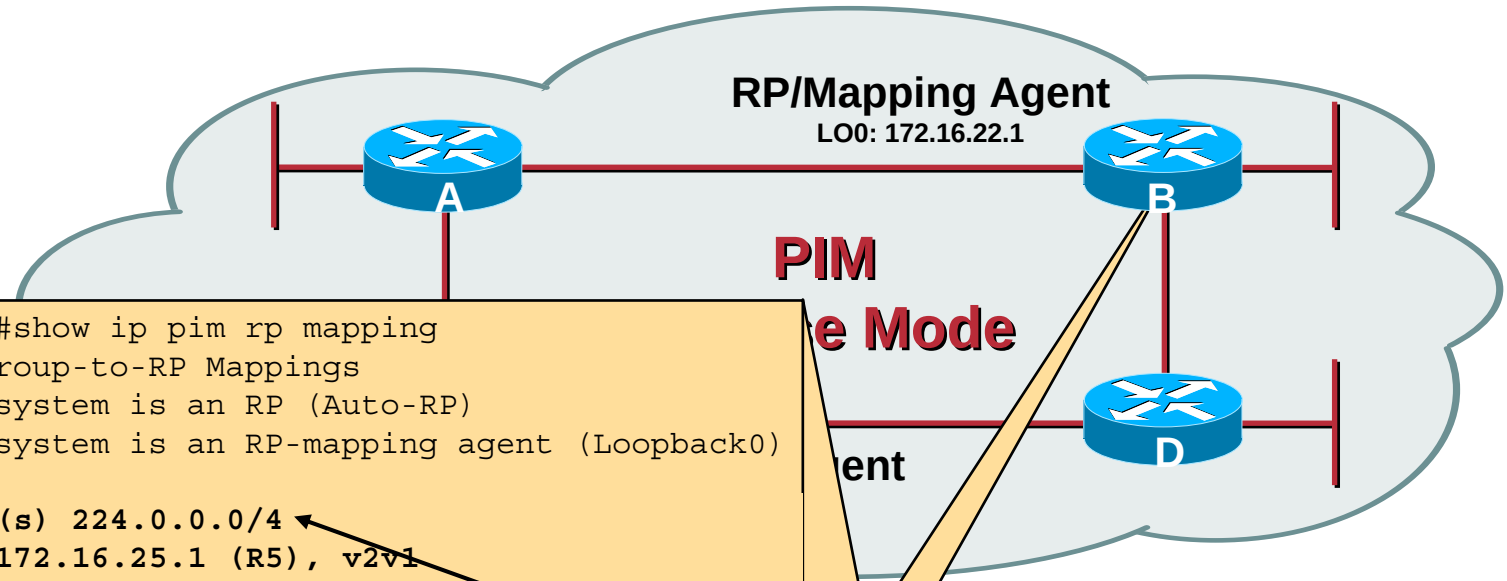
Using Multiple Group Ranges




```
Rtr-B#show ip pim rp mapping
PIM Group-to-RP Mappings
This system is an RP (Auto-RP)
This system is an RP-mapping agent (Loopback0)

Group(s) 224.0.0.0/4
  RP 172.16.25.1 (R5), v2v1
    Info source: 172.16.25.1 (R5), via Auto-RP
         Uptime: 01:05:03, expires: 00:02:27
  RP 172.16.22.1 (R2), v2v1
    Info source: 172.16.22.1 (R2), via Auto-RP
         Uptime: 01:05:11, expires: 00:02:43
Group(s) 225.0.0.0/8
  RP 172.16.25.1 (R5), v2v1
    Info source: 172.16.25.1 (R5), via Auto-RP
         Uptime: 00:17:29, expires: 00:02:29
Group(s) 239.255.0.0/16
  RP 172.16.22.1 (R2), v2v1
    Info source: 172.16.22.1 (R2), via Auto-RP
         Uptime: 00:56:18, expires: 00:02:41
```

Mapping Agent's Group-to-RP mapping cache contains all learned C-RP's ordered by Group, Mask length, and C-RP address.

Using Multiple Group Ranges

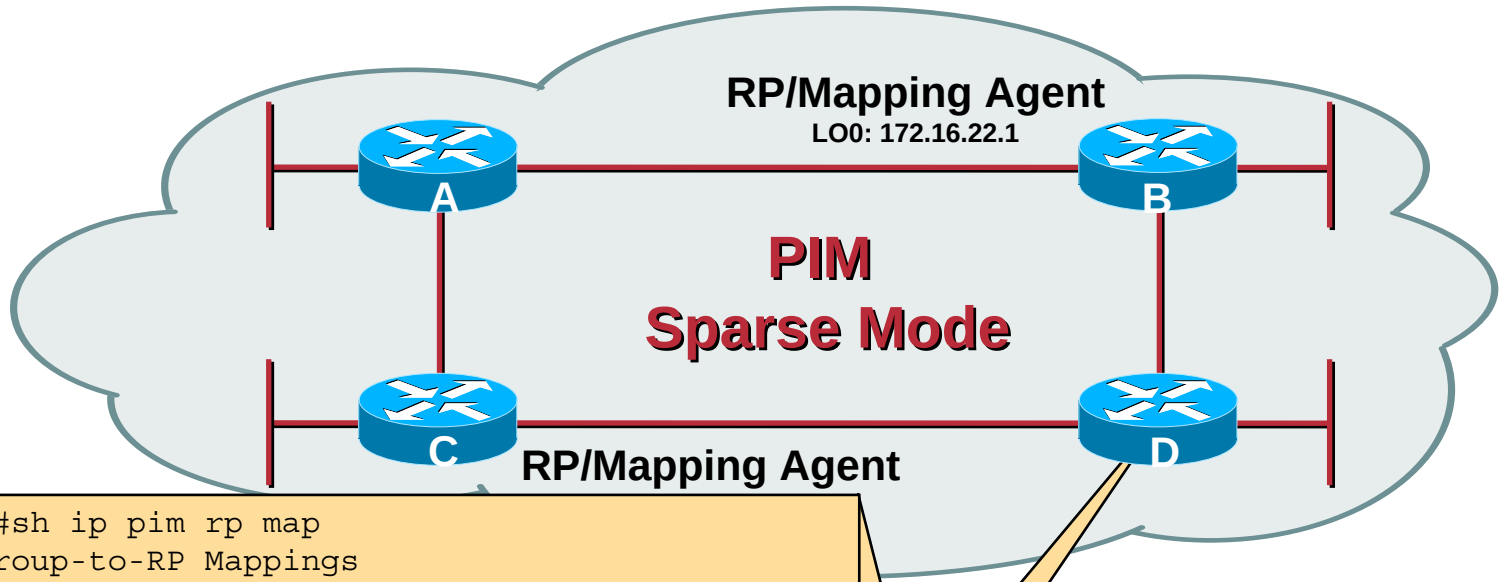



```
Rtr-B#show ip pim rp mapping
PIM Group-to-RP Mappings
This system is an RP (Auto-RP)
This system is an RP-mapping agent (Loopback0)

Group(s) 224.0.0.0/4
  RP 172.16.25.1 (R5), v2v1
    Info source: 172.16.25.1 (R5), via Auto-RP
         Uptime: 01:05:03, expires: 00:02:27
  RP 172.16.22.1 (R2), v2v1
    Info source: 172.16.22.1 (R2), via Auto-RP
         Uptime: 01:05:11, expires: 00:02:43
Group(s) 225.0.0.0/8
  RP 172.16.25.1 (R5), v2v1
    Info source: 172.16.25.1 (R5), via Auto-RP
         Uptime: 00:17:29, expires: 00:02:29
Group(s) 239.255.0.0/16
  RP 172.16.22.1 (R2), v2v1
    Info source: 172.16.22.1 (R2), via Auto-RP
         Uptime: 00:56:18, expires: 00:02:41
```

Mapping Agent advertises elected RP's in Discovery Messages

Note: Elected RP is always listed as first C-RP for a Group Range.

Using Multiple Group Ranges

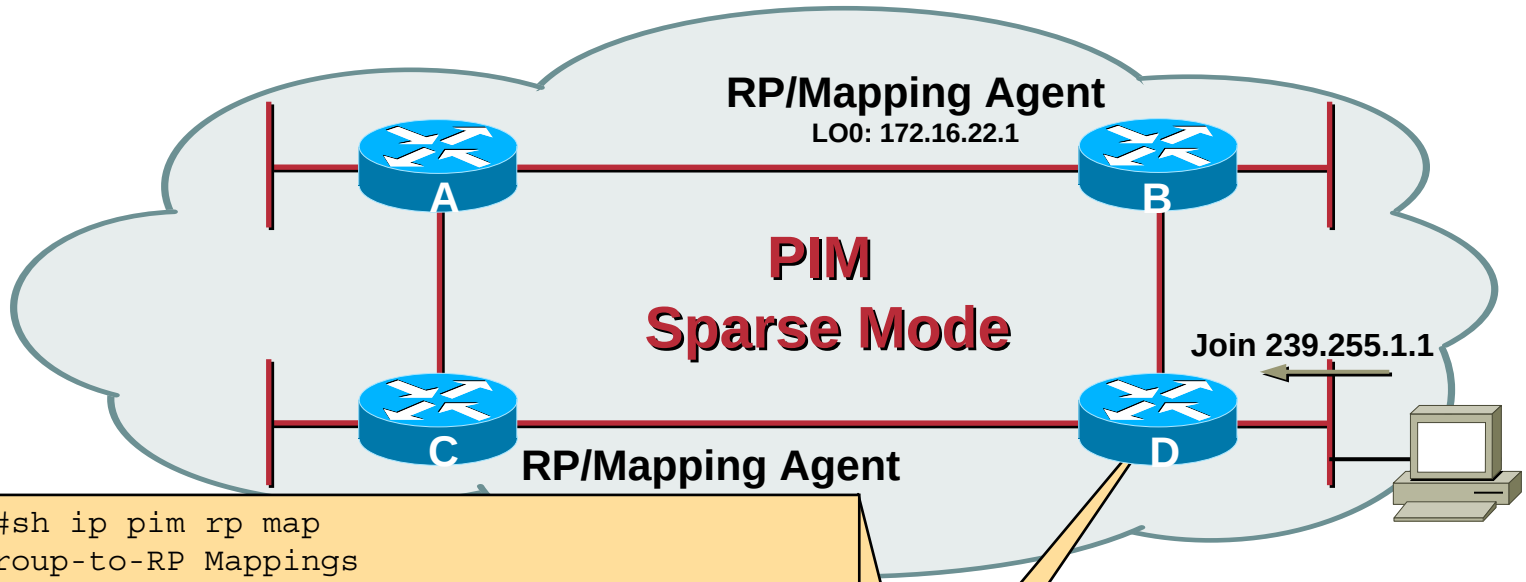



```
Rtr-D#sh ip pim rp map
PIM Group-to-RP Mappings

Group(s) 224.0.0.0/4
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 01:00:14, expires: 00:02:25
Group(s) 225.0.0.0/8
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:12:40, expires: 00:02:24
Group(s) 239.255.0.0/16
  RP 172.16.22.1 (Rtr-B), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:51:28, expires: 00:02:24
```

Resulting Group-to-RP Mapping Cache in all non-MA routers

(Notice that only elected RP's are contained the Group-to-RP mapping cache in non-MA routers.)

Using Multiple Group Ranges



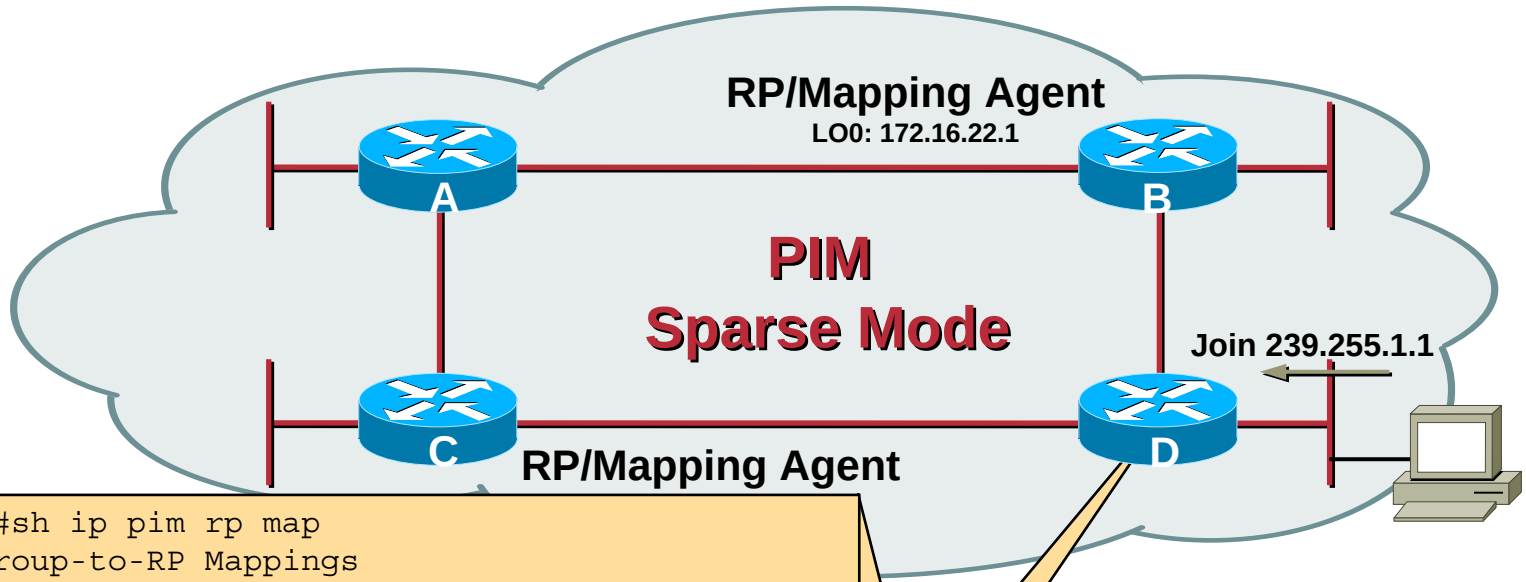

```
Rtr-D#sh ip pim rp map
PIM Group-to-RP Mappings

Group(s) 224.0.0.0/4
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 01:00:14, expires: 00:02:25
Group(s) 225.0.0.0/8
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:12:40, expires: 00:02:24
Group(s) 239.255.0.0/16
  RP 172.16.22.1 (Rtr-B), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:51:28, expires: 00:02:24
```

Which entry will the router use to determine RP address?

This one? (It has the highest RP address.)

Or this one? (It has the longest mask.)

Using Multiple Group Ranges




```
Rtr-D#sh ip pim rp map
PIM Group-to-RP Mappings

Group(s) 224.0.0.0/4
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 01:00:14, expires: 00:02:25
Group(s) 225.0.0.0/8
  RP 172.16.25.1 (Rtr-C), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:12:40, expires: 00:02:24
Group(s) 239.255.0.0/16
  RP 172.16.22.1 (Rtr-B), v2v1
    Info source: 172.16.25.1 (Rtr-C), via Auto-RP
         Uptime: 00:51:28, expires: 00:02:24
```

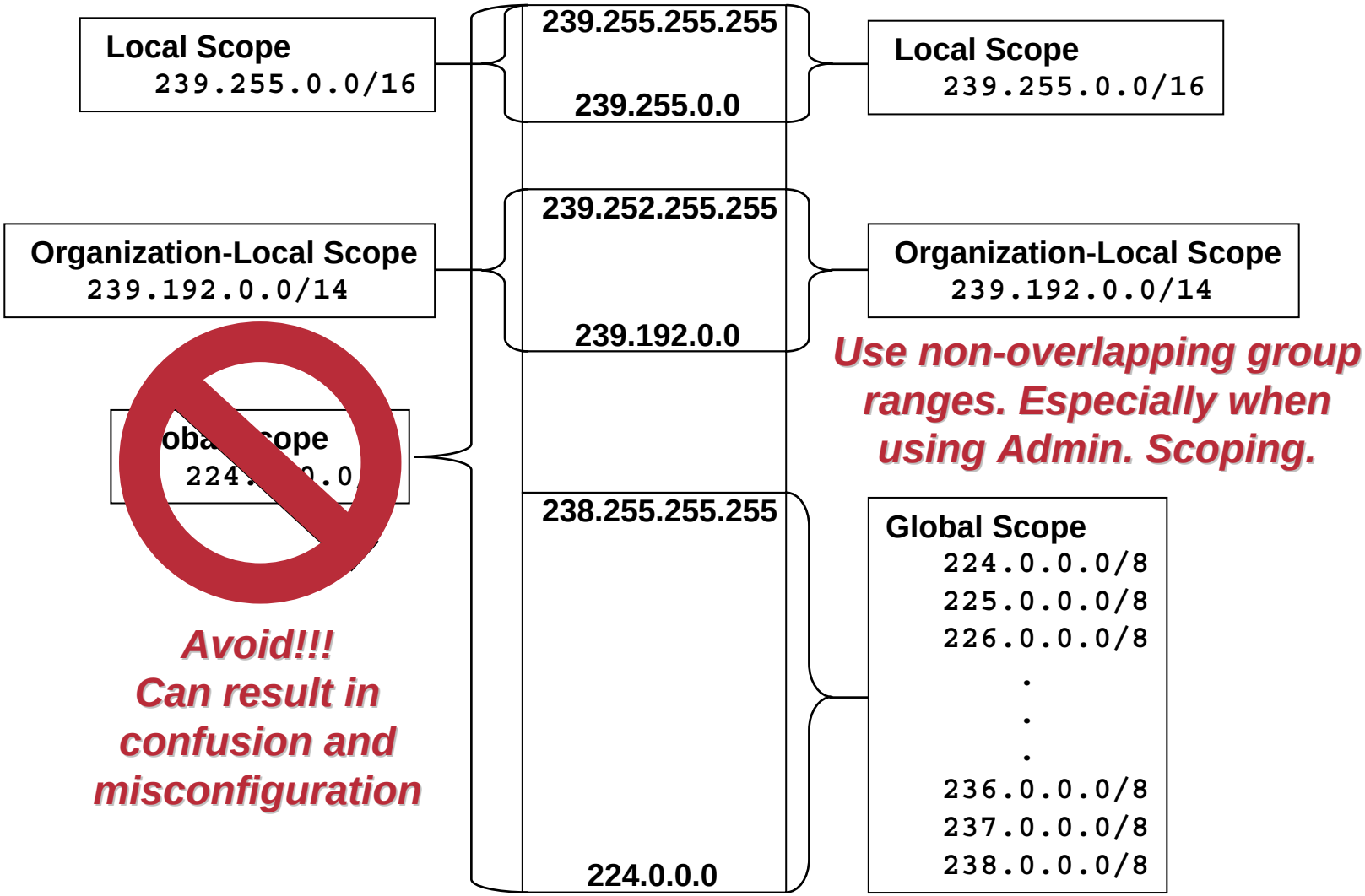
Answer:

Router uses “longest match” to find matching entry in the Group-to-RP mapping cache.

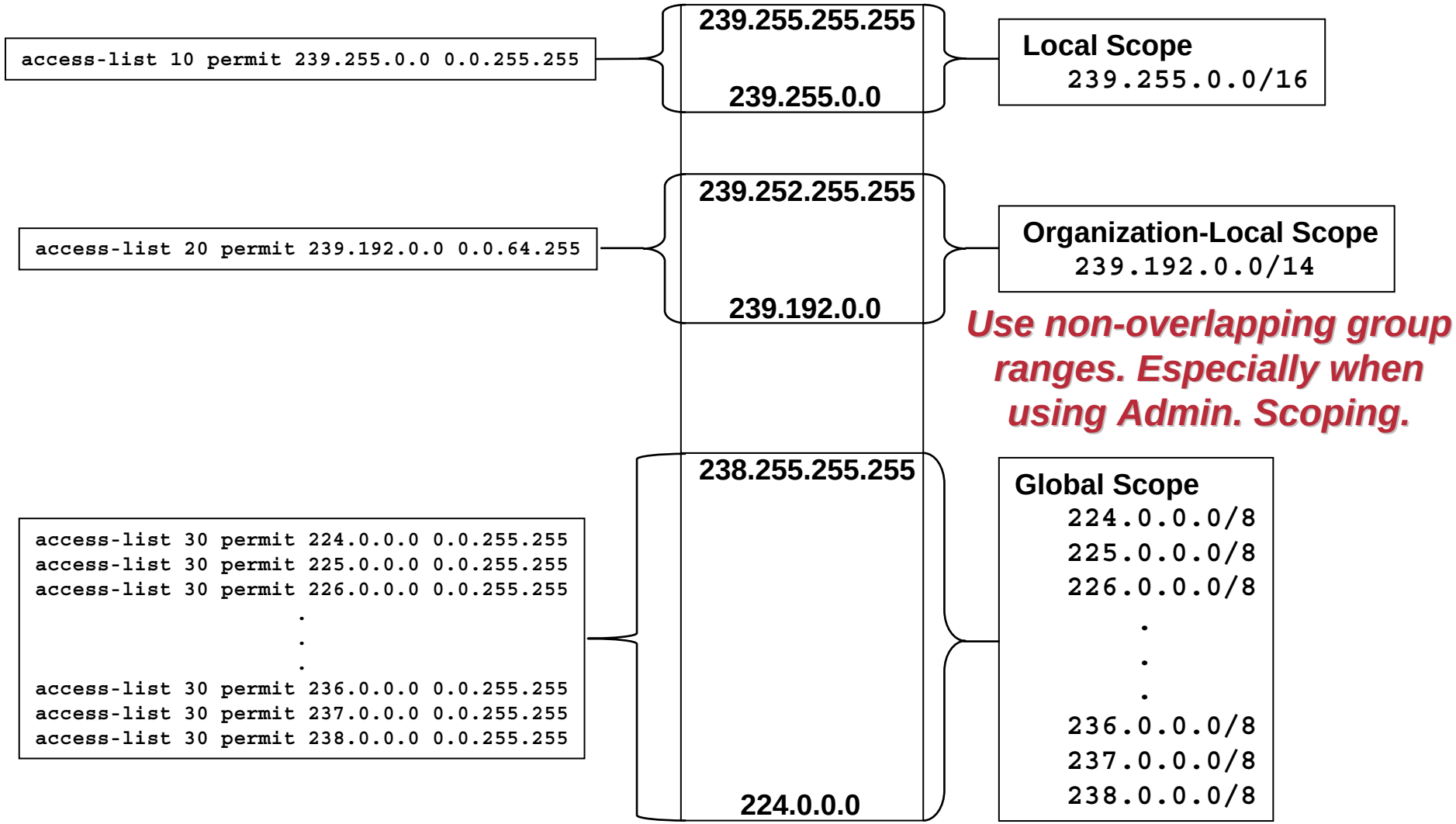
Moral:

Avoid overlapping group ranges to reduce the chances of incorrect RP selection.

Overlapping Group Ranges



Overlapping Group Ranges



Overlapping Group Ranges

- **Avoiding Overlapping Group Ranges**
  - **Can't use "deny" clause in C-RP ACL's**
    - **Implies "Dense-mode Override"**

```
ip pim send-rp-announce loopback0 scope 16 group-list 10
access-list 10 deny 239.0.0.0 0.255.255.255
access-list 10 permit 224.0.0.0 15.255.255.255
```

  - **Must only use "permit" clauses**

```
ip pim send-rp-announce loopback0 scope 16 group-list 10
access-list 10 permit 224.0.0.0 0.255.255.255
access-list 10 permit 225.0.0.0 0.255.255.255
                       .
                       .
                       .
access-list 10 permit 238.0.0.0 0.255.255.255
```

Agenda

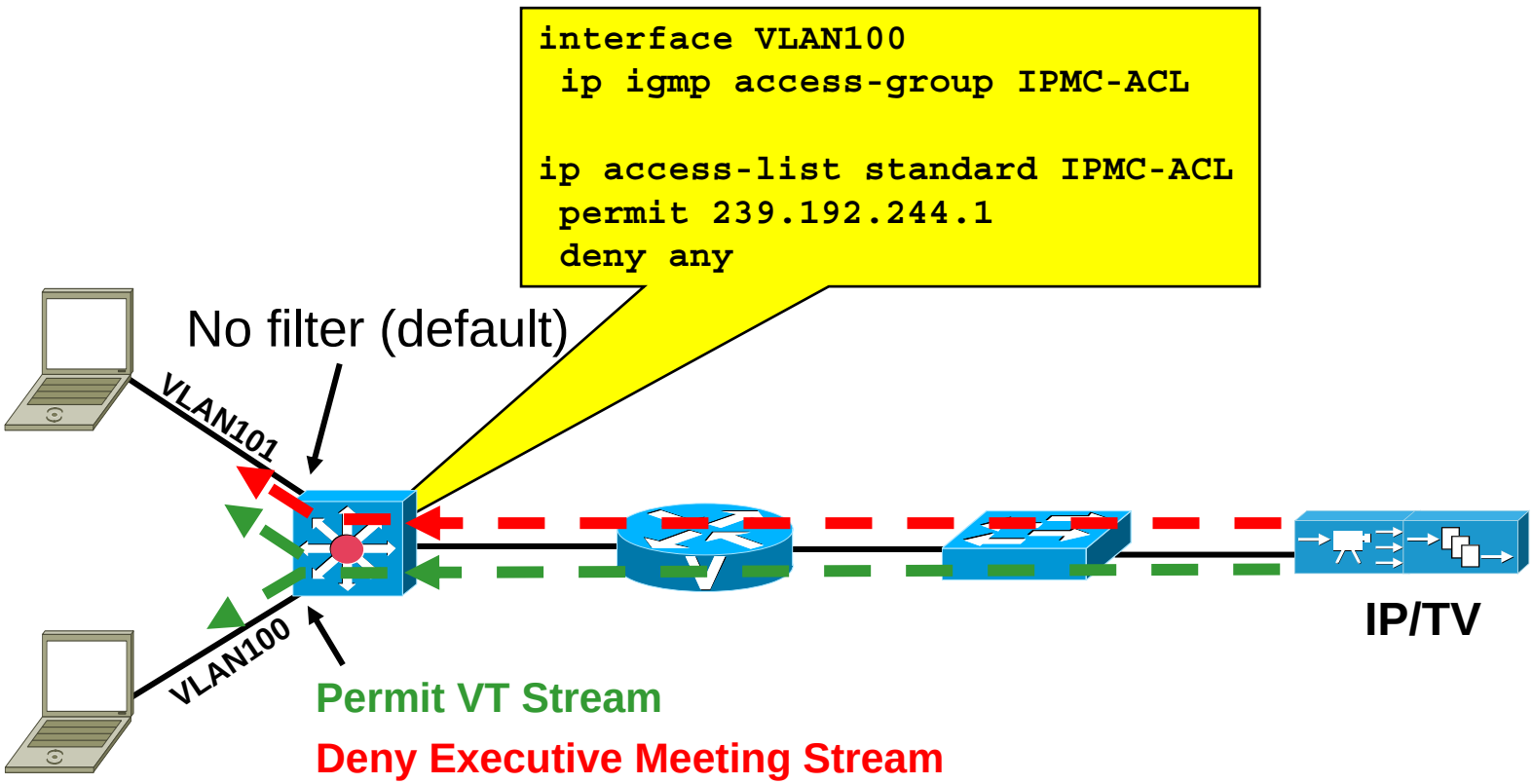
- Basic Multicast Engineering
- **Advanced Multicast Engineering**

Advanced Multicast Engineering

- **Multicast Group Control**
- **Using Admin. Scoped Zones**
- **PIM Protocol Extensions**

Controlling Receivers

IGMP Access-Group Approach



This is micro-management of IP Multicast traffic!!!

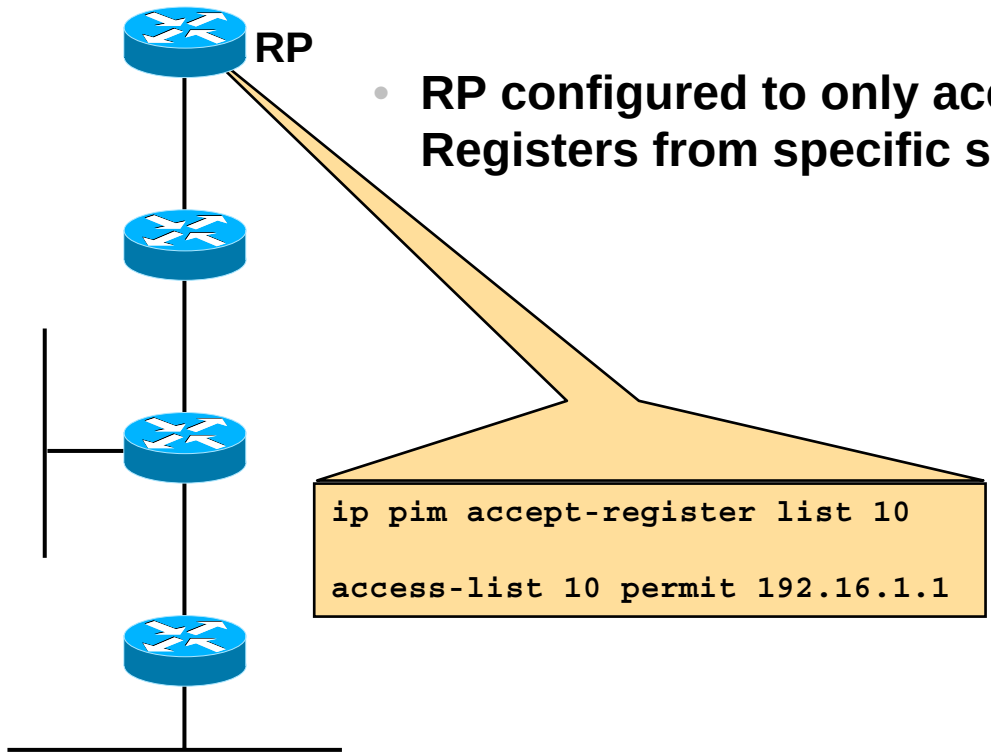
Controlling Source Registration

- **Global command**

```
ip pim accept-register [list <acl>] | [route-map <map>]
```

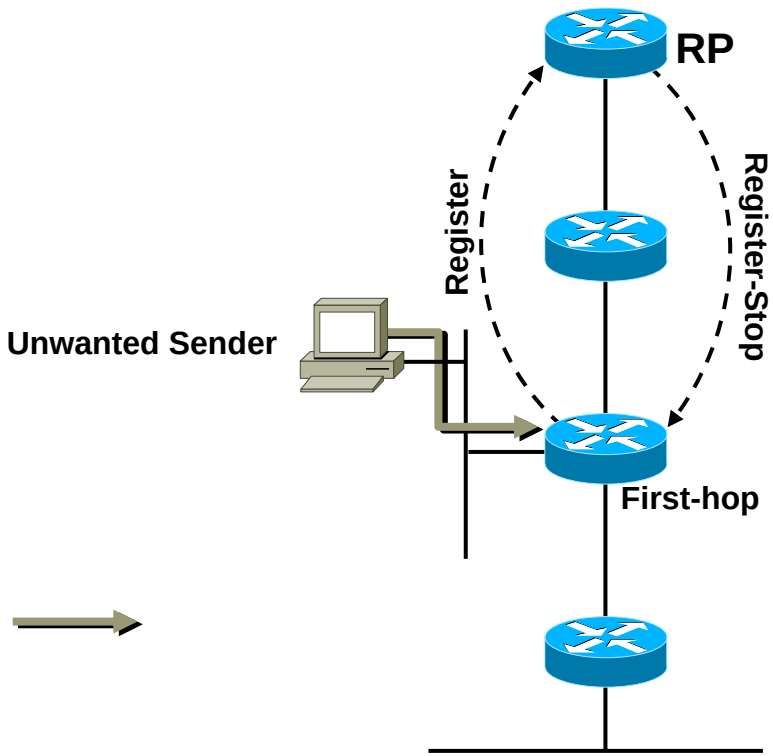
  - **Used on RP to filter incoming Register messages**
  - **Filter on Source address alone (Simple ACL)**
  - **Filter on (S, G) pair (Extended ACL)**
  - **May use route-map to specify what to filter**
    - **Filter by AS-PATH if (m)BGP is in use.**
- **Helps prevents unwanted sources from sending**
  - **First hop router blocks traffic from reaching net**
  - **Note: Traffic can still flow under certain situations**

Controlling Source Registration



- RP configured to only accept Registers from specific source.

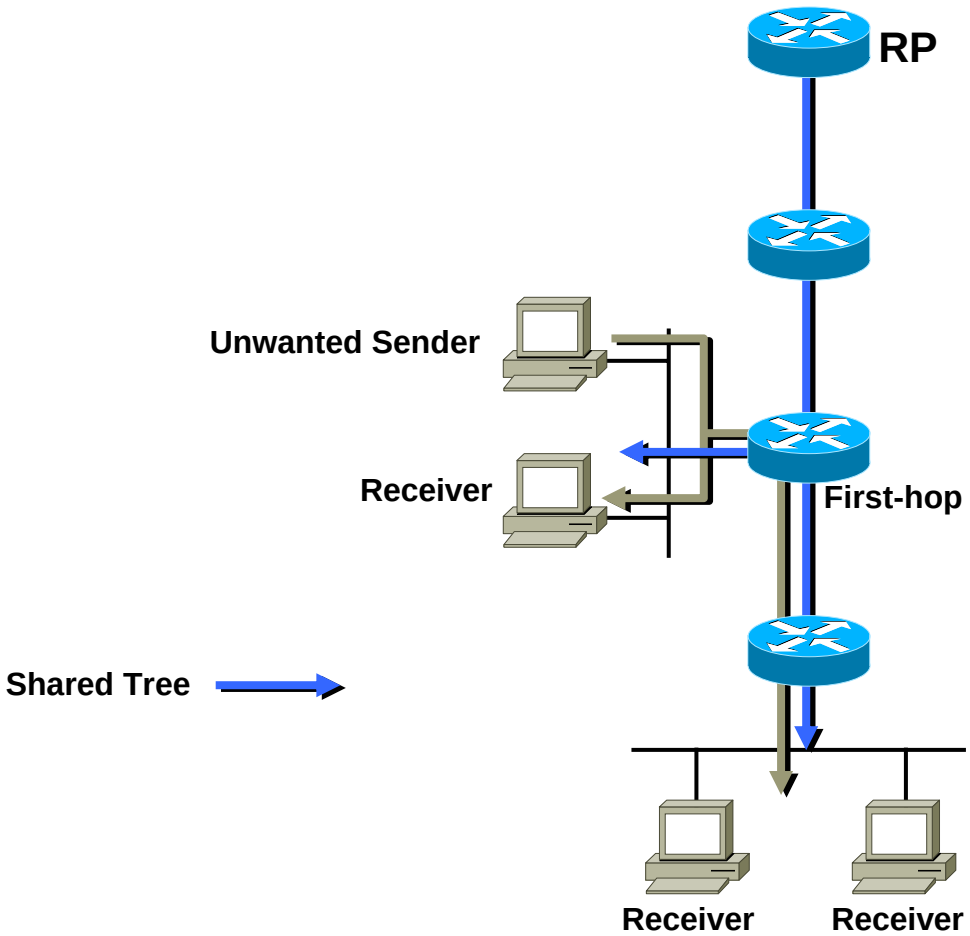
Controlling Source Registration



- Unwanted source traffic hits first-hop router.
- First-hop router creates (S,G) state and sends Register.
- RP rejects Register, sends back a Register-Stop.

Controlling Source Registration

Weaknesses in ‘accept-register’ usage.



- Traffic will flow on local subnet where source resides.
- Traffic will flow from first-hop router down any branches of the Shared Tree.
  - Results when (*,G) OIL is copied to (S,G) OIL at first-hop router.
  - Causes (S,G) traffic to flow down all interfaces in (*,G) OIL of first-hop router.
  - Fundamental limitation of PIM protocol.

Disabling Entire Group Ranges

- **Accept-Register Method**

```
ip pim accept-register group-list 10

access-list 10 deny 224.2.0.0 0.0.255.255
access-list 10 permit any
```

- **Pros**
  - **Only configured on RP(s)**
- **Cons**
  - **Shared Trees and (*,G) state still created.**
    - **Results in unwanted (*,G) PIM Control Traffic.**
  - **Source traffic can still flow.**

    (See previous section on Accept-Register)

Disabling Entire Group Ranges

- Garbage Can RP Method
  - Concept:
    - Separate RP for “disabled” groups
      - Could be non-existant router
    - Blackholes all Registers and Joins
  - Implementation:
    - Define separate RP for disabled groups
      - Use Auto-RP, BSR or Static RP definition
    - Disable RP functionality on Garbage Can RP
      - Use ‘`accept-rp`’ command on GC RP to “deny” it from serving as RP for the disabled group range.

Disabling Entire Group Ranges

- Garbage Can RP Method
  - Pros:
    - Few if any.
  - Cons:
    - Periodic Registers still sent to GC RP
    - Periodic Joins still sent to GC RP
    - Has same source issues as Accept-Register
      - Source traffic can still flow under certain conditions.
    - Adds ***significant*** complexity to network

Disabling Entire Group Ranges

- **Local Loopback RP Method**
  - **Concept:**
    - **Only Auto-RP-learned groups are authorized.**
    - **All other groups are considered *unauthorized.***
  - **Implementation:**
    - **Define local Loopback as RP for unauthorized groups on each router.**

```
ip pim rp-address <local_loopback> 10
access-list 10 permit 224.2.0.0 0.0.255.255
```

Note: The permit clause defines the unauthorized group.

Disabling Entire Group Ranges

- **Local Loopback RP Method**
  - **Operation:**
    - **Each router serves as RP for unauthorized groups.**
      - **Collapses PIM-SM domain of unauthorized groups down to the local router.**
    - **Unauthorized group traffic cannot flow beyond local router.**

Disabling Entire Group Ranges

- Local Loopback RP Method
  - Pros:
    - No PIM control traffic sent.
      - Local router is RP so no Registers/Joins are sent.
    - No additional workload on local router.
      - First-hop routers always have to create state anyway.
    - Can also serve as RP-of-last-resort
      - Solving DM Fallback problem at the same time.
  - Cons:
    - Must be configured on every router.
    - Local sources can still send to local receivers.

Disabling Entire Group Ranges

- **Recommendation**
  - **Use Local Loopback RP Method**
    - ***Effectively*** **disables unauthorized group traffic.**
    - **Can also serve as RP-of-last-resort**

```
ip pim rp-address <local_loopback> 10

access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

Disabling Entire Group Ranges – Future

- New ‘no ip pim dm-fallback’ command
  - Undefined (via Auto-RP or BSR) groups default to an RP address of 0.0.0.0.
  - Effectively disables any group unlearned groups.
- Available soon.

Advanced Multicast Engineering

- Multicast Group Control
- **Using Admin. Scoped Zones**
- PIM Protocol Extensions

Administratively-Scoped Zones

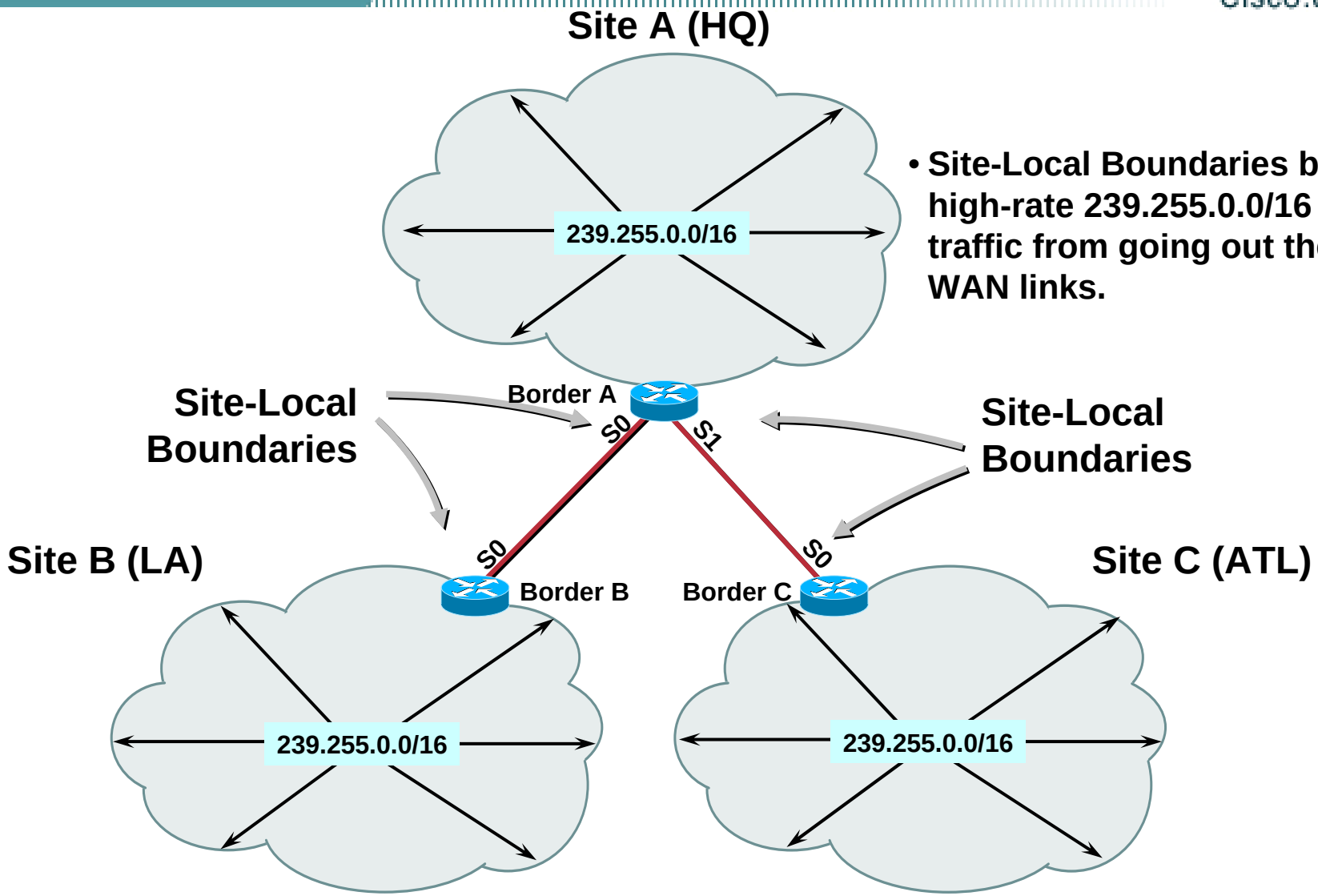
- **Used to limit:**
  - **High-BW sources to local site**
  - **Control sensitive multicast traffic**
- **Simple scoped zone example:**

  239.255.0.0/16 = (Site) Local Scope

  239.192.0.0/14 = Organization-Local Scope

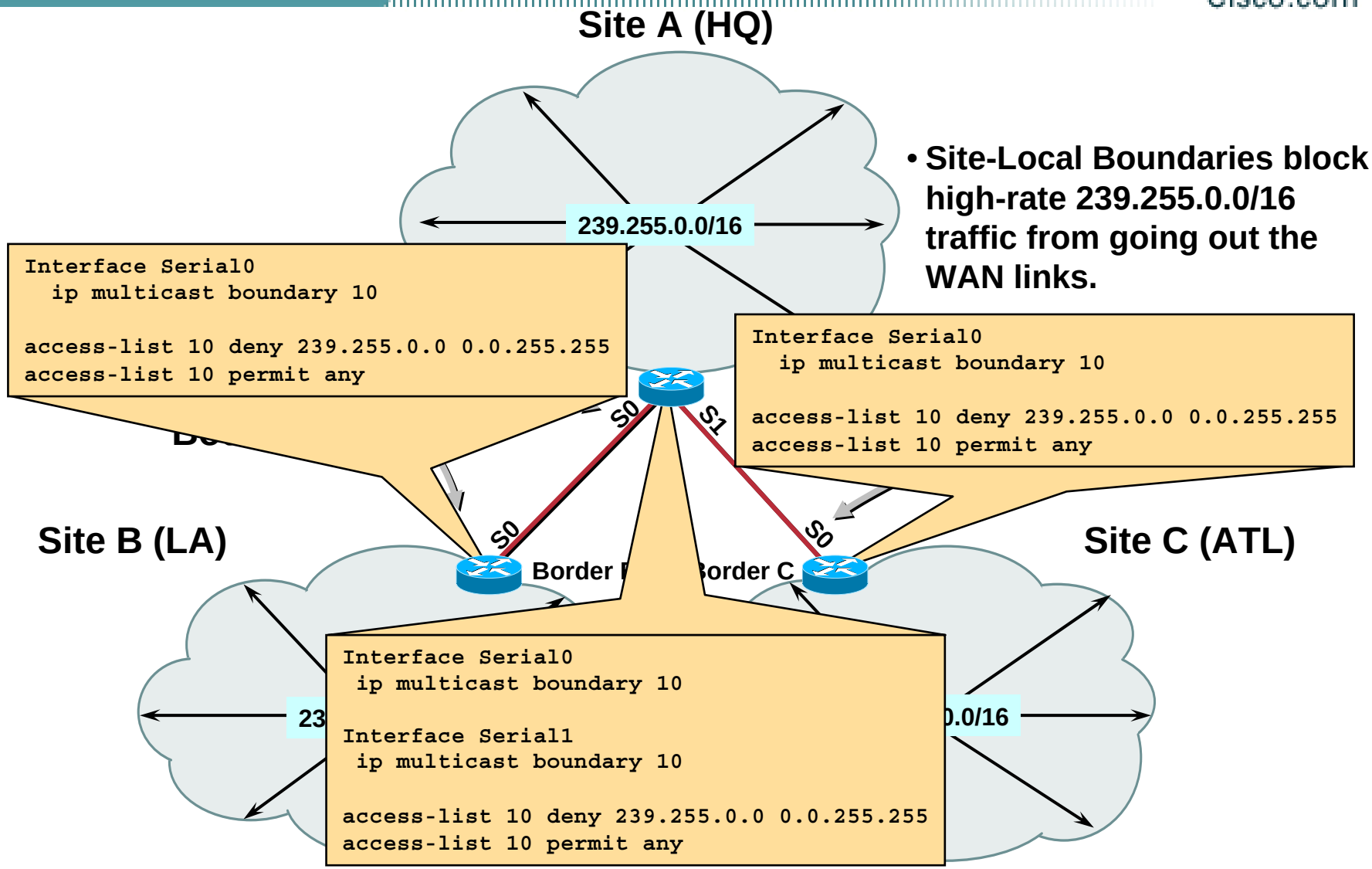
  224.1.0.0 - 238.255.255.255 = Global scope (Internet) zone
  - **High-BW sources use Site-Local scope**
  - **Low-Med. BW sources use Org.-Local scope**
  - **Internet-wide sources use Global scope**

Administratively-Scoped Zones



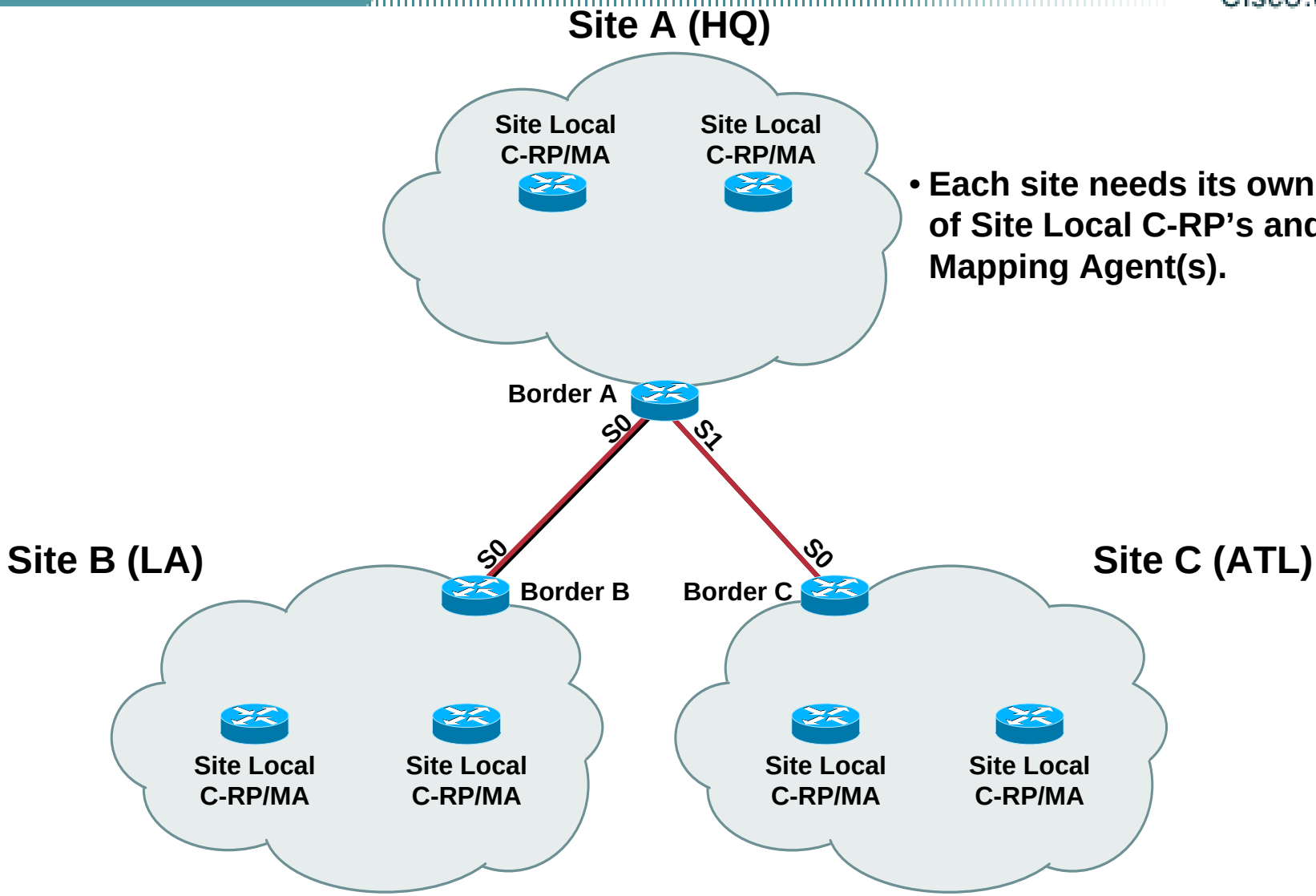
- Site-Local Boundaries block high-rate 239.255.0.0/16 traffic from going out the WAN links.

Administratively-Scoped Zones



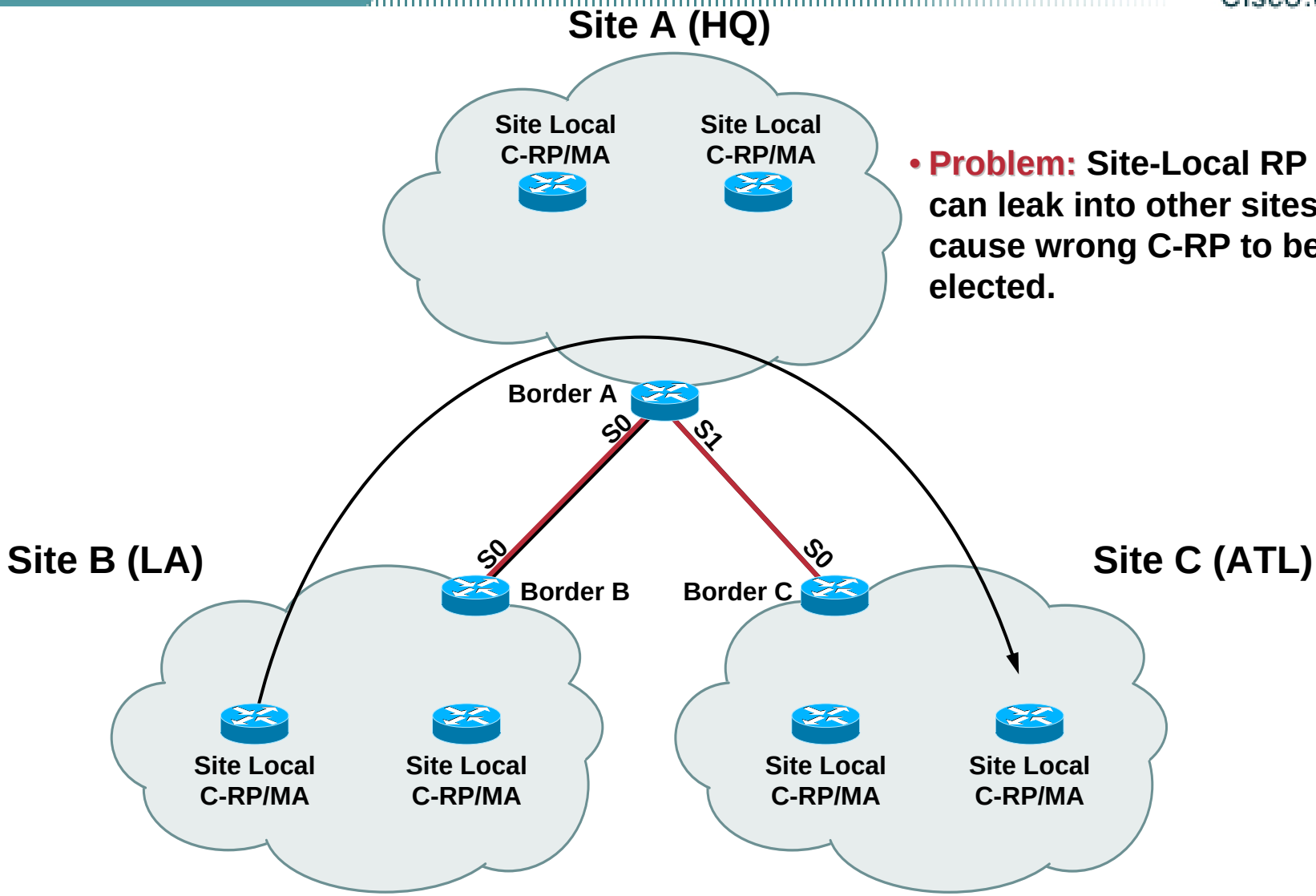
- Site-Local Boundaries block high-rate 239.255.0.0/16 traffic from going out the WAN links.

Administratively-Scoped Zones
Auto-RP Example



- Each site needs its own set of Site Local C-RP's and Mapping Agent(s).

Administratively-Scoped Zones
Auto-RP Example



- **Problem:** Site-Local RP info can leak into other sites and cause wrong C-RP to be elected.

Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

- **Multicast Boundary Command**

```
ip multicast boundary <acl> [filter-autorp]
```

  - **New ‘filter-autorp’ option**
    - **Filters contents of Auto-RP packets**
      - **Filters both Announcement and Discovery messages**
      - **C-RP entries that fail <acl> are removed from packet**
    - **Prevents C-RP information from leaking in/out of scoped zone.**
    - **Greatly simplifies Admin. Scoped Zone support in Auto-RP.**
    - **Available in 12.0(22)S, 12.2(12).**

Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

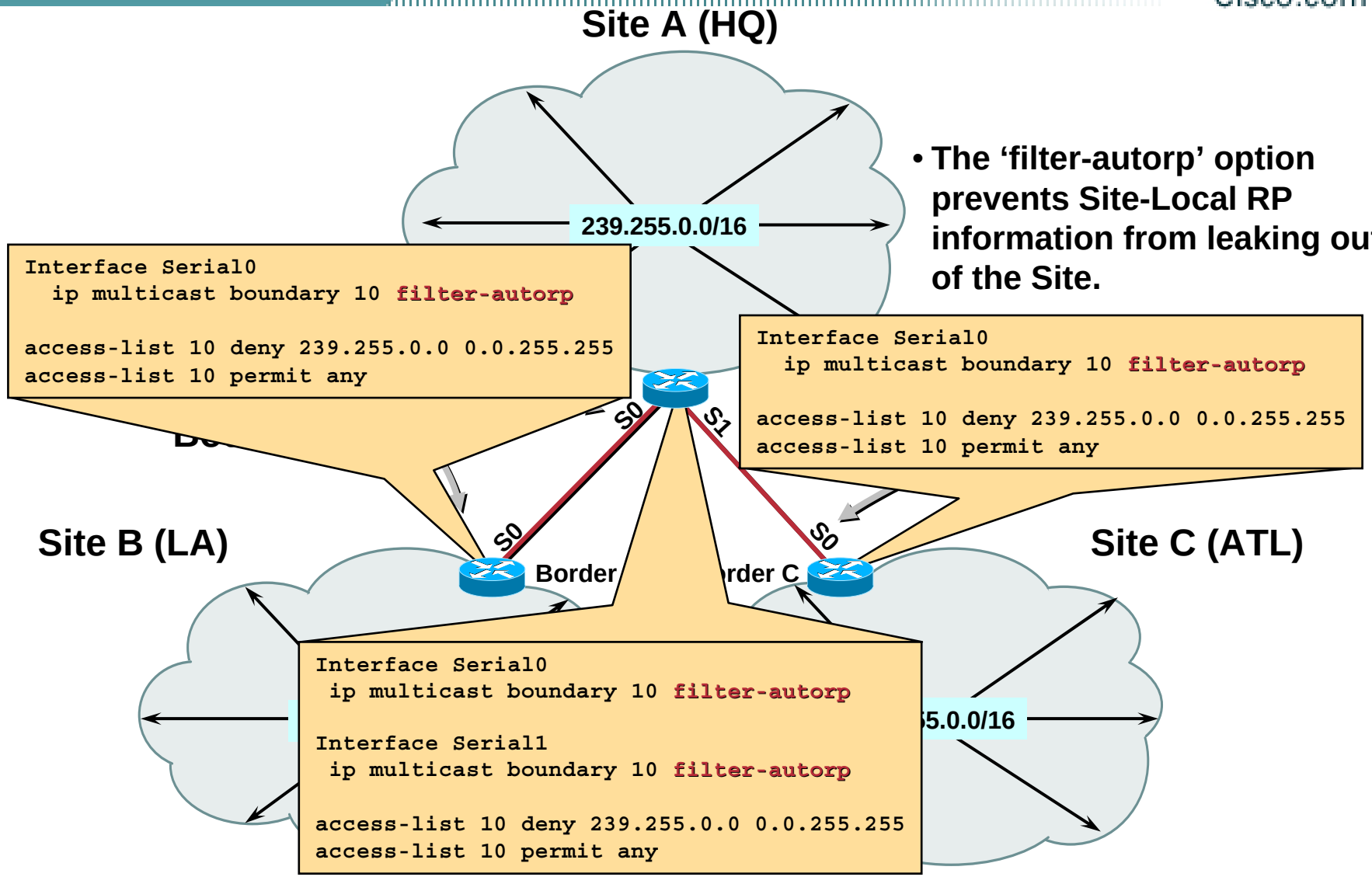
- **How ‘filter-autorp’ option works:**
  - **For each RP Entry in Auto-RP packet:**
    - **If group-range in RP-Entry *‘intersects’* any ‘denied’ group-range in the Multicast Boundary ACL, delete RP Entry from Auto-RP packet.**
  - **If resulting Auto-RP packet is non-empty, forward across multicast boundary.**

Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

- **Using Multicast Boundary ‘filter-autorp’**
  - **Avoid Auto-RP Group-Range Overlaps**
    - **Overlapping ranges can “intersect” denied ranges at multicast boundaries.**
      - **Can cause unexpected Auto-RP info filtering at multicast boundaries.**
      - **Results in loss of Auto-RP info to other parts of network.**
  - **Rule of Thumb:**
    - **Make sure Auto-RP Group-Ranges match exactly any Multicast Boundary Ranges!**

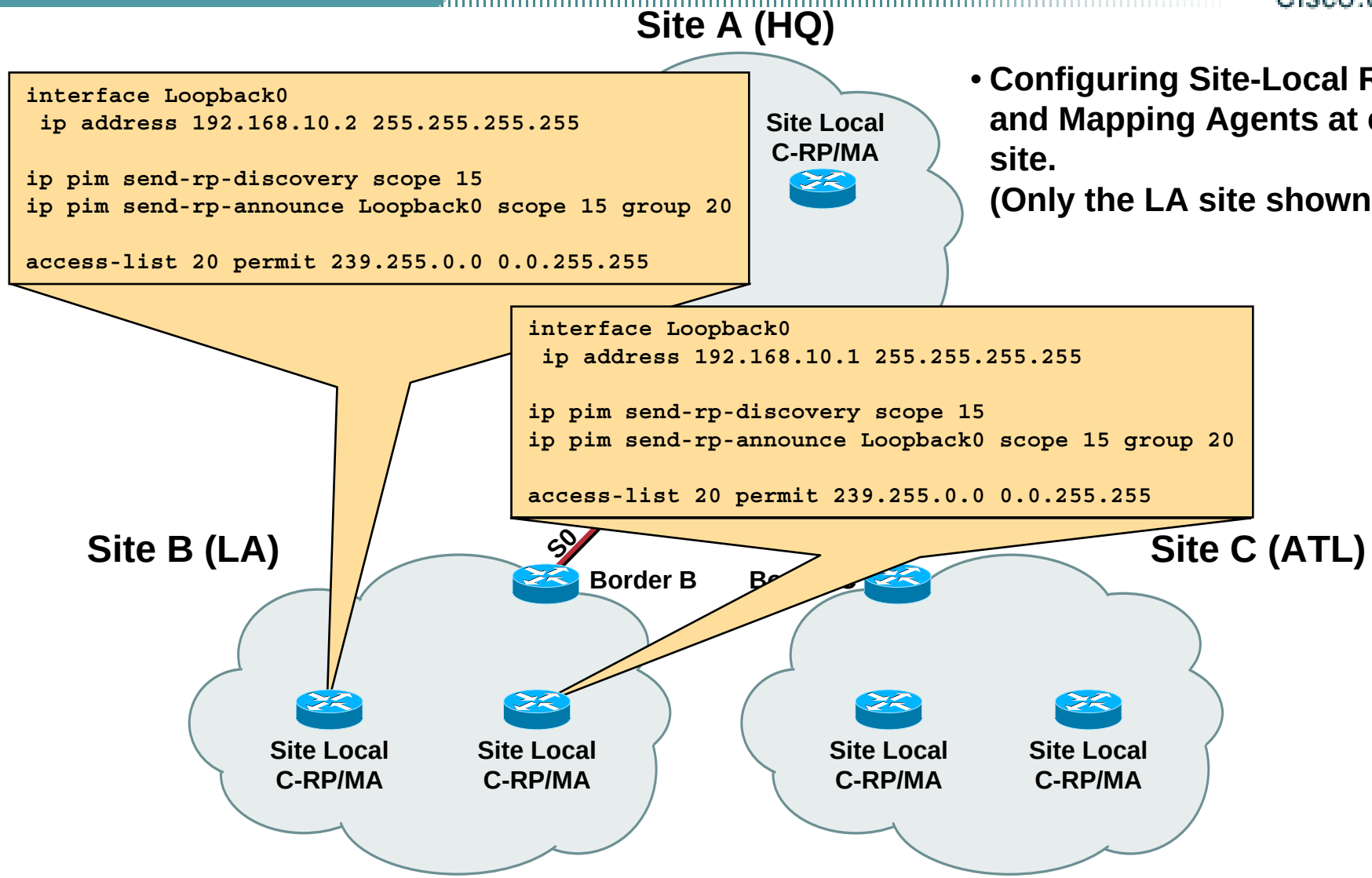
      (i.e. don’t use overlapping Auto-RP group ranges.)

Administratively-Scoped Zones
Auto-RP Example with 'filter-autorp' boundaries

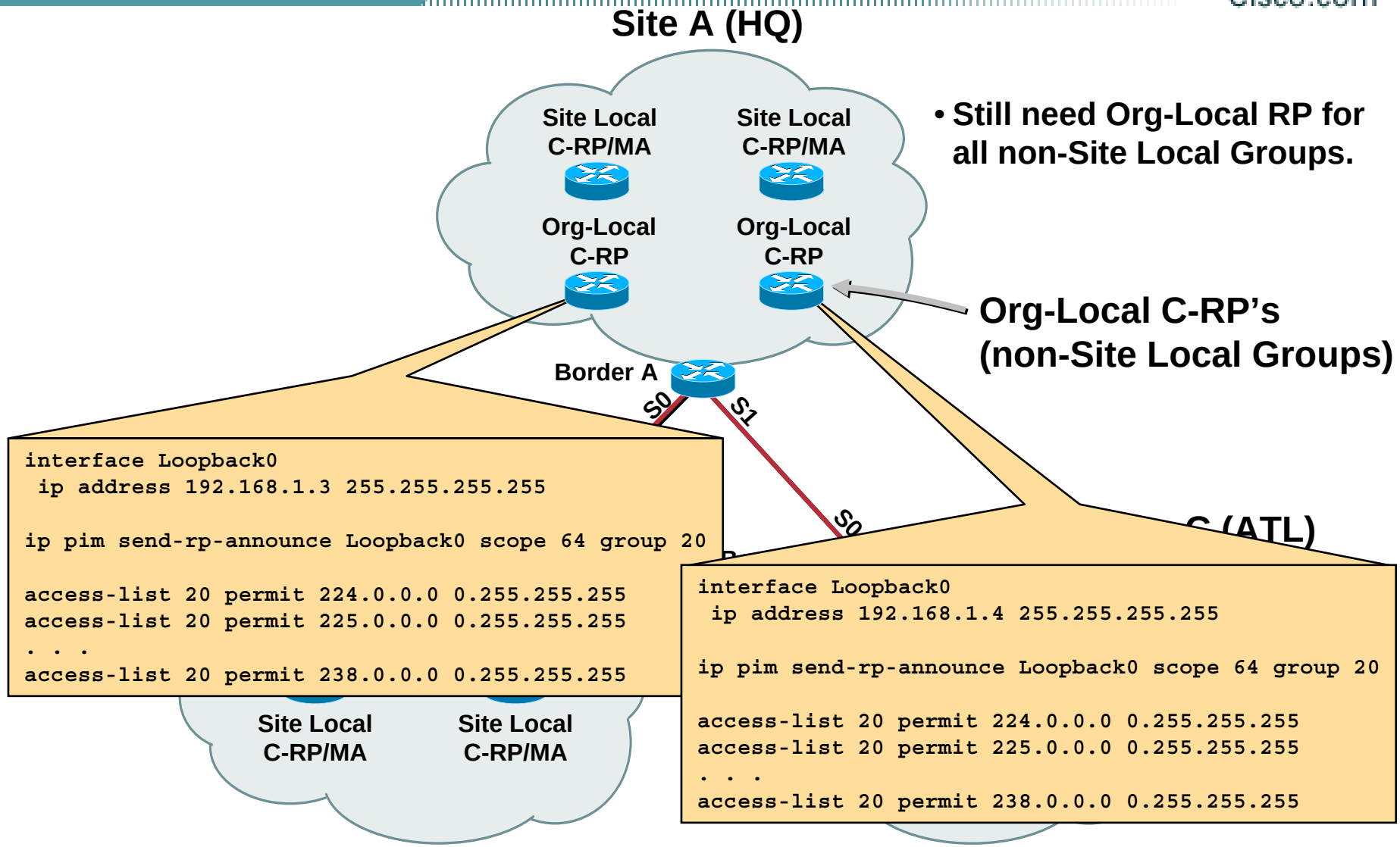


- The 'filter-autorp' option prevents Site-Local RP information from leaking out of the Site.

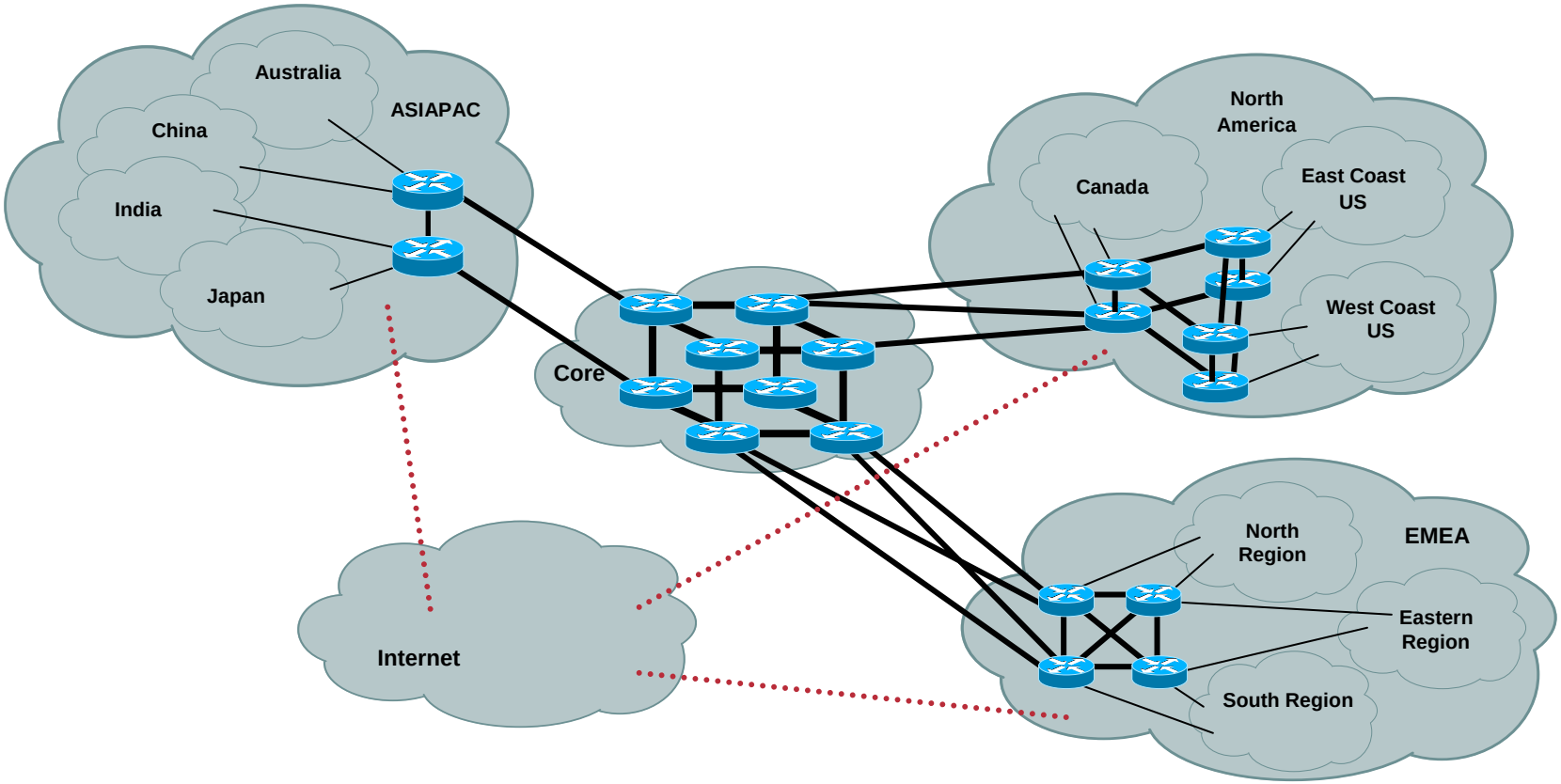
Administratively-Scoped Zones
Auto-RP Example with 'filter-autorp' boundaries



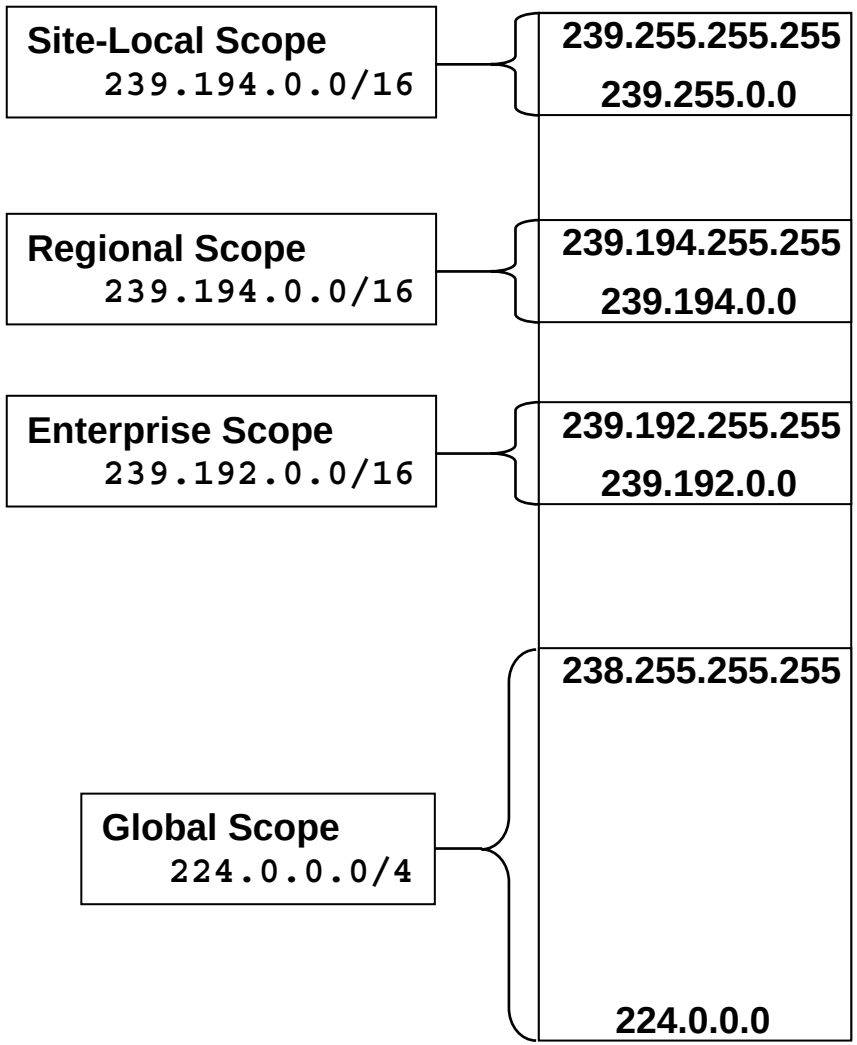
Administratively-Scoped Zones
Auto-RP Example with 'filter-autorp' boundaries



Administratively-Scoped Zones Example



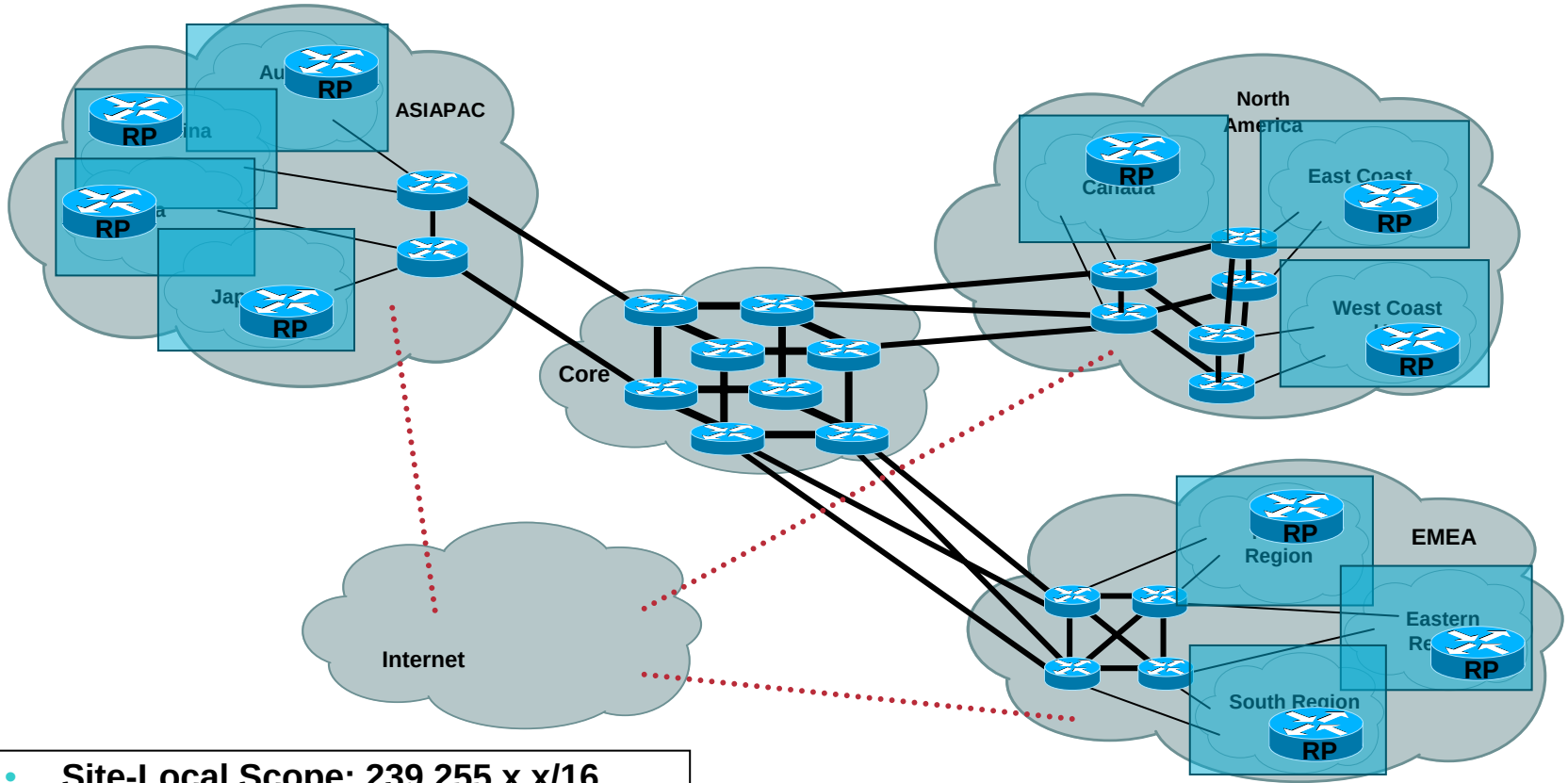
Administratively-Scoped Zones Example



Administratively-Scoped Zones Example

Cisco.com

Level1: Site Scope

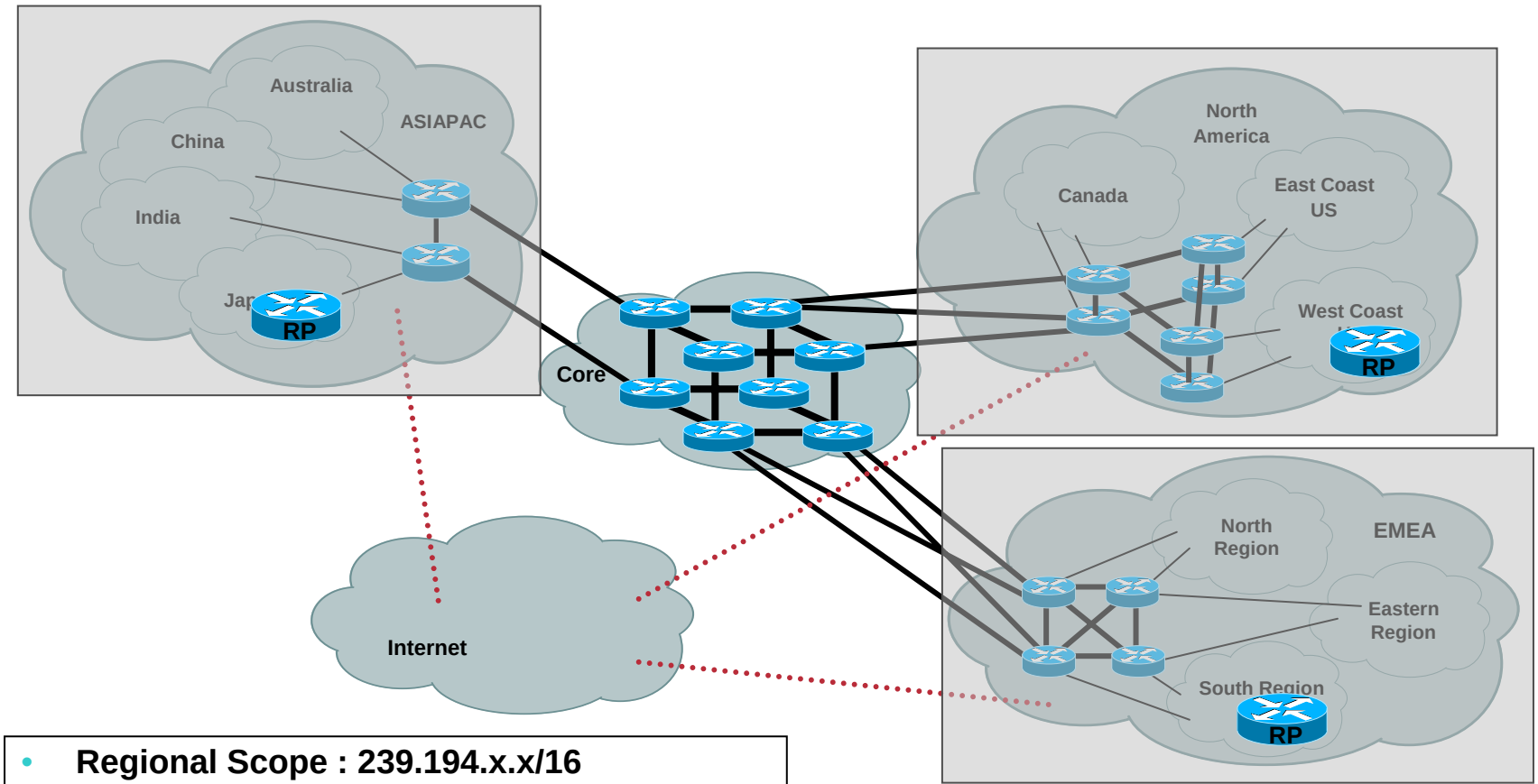


- Site-Local Scope: 239.255.x.x/16
- RP per Site
- Reusable range

Administratively-Scoped Zones Example

Cisco.com

Level2: Regional Scope

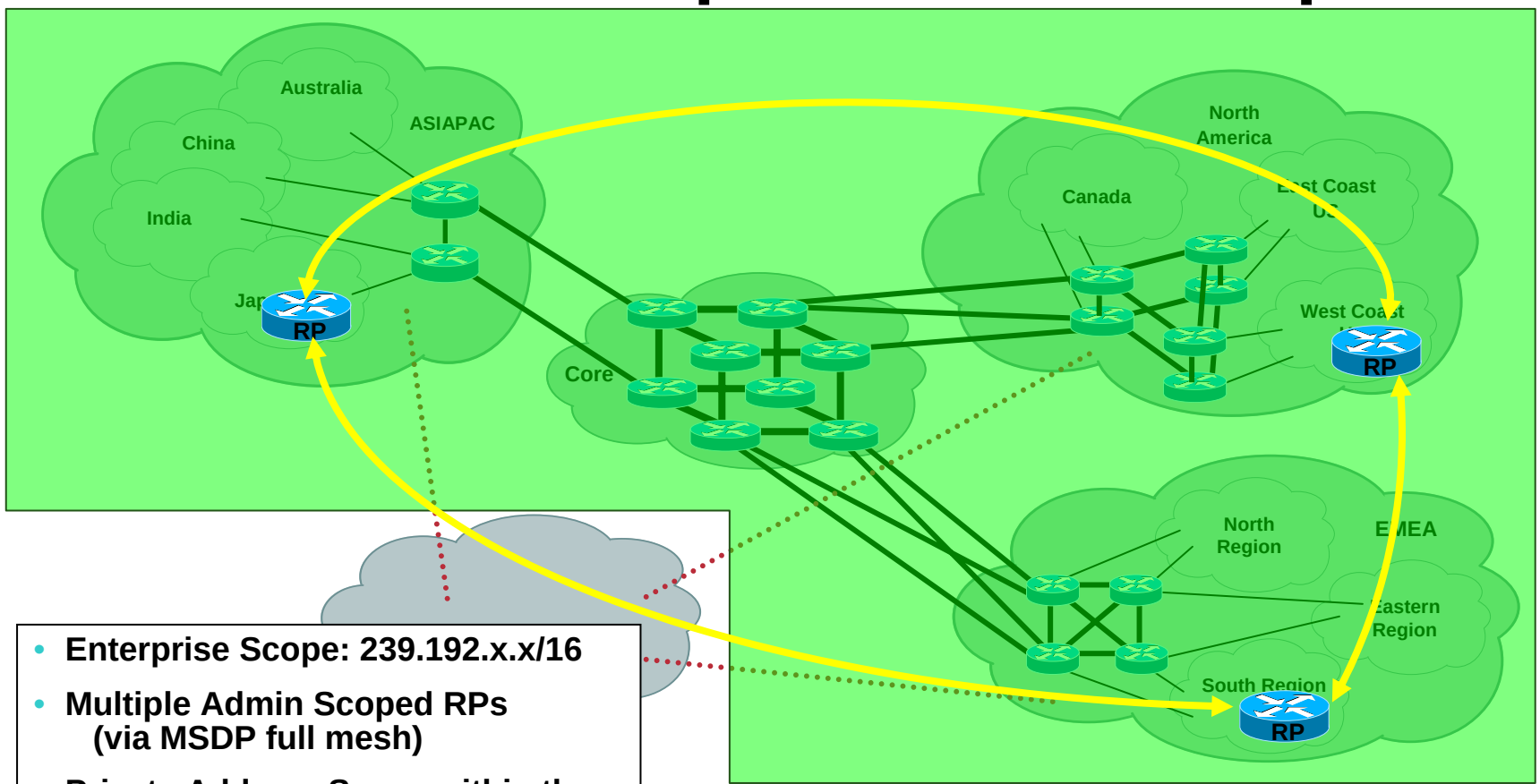


- Regional Scope : 239.194.x.x/16
- RP per Region
- Reusable Range

Administratively-Scoped Zones Example

Cisco.com

Level3: Enterprise Global Scope

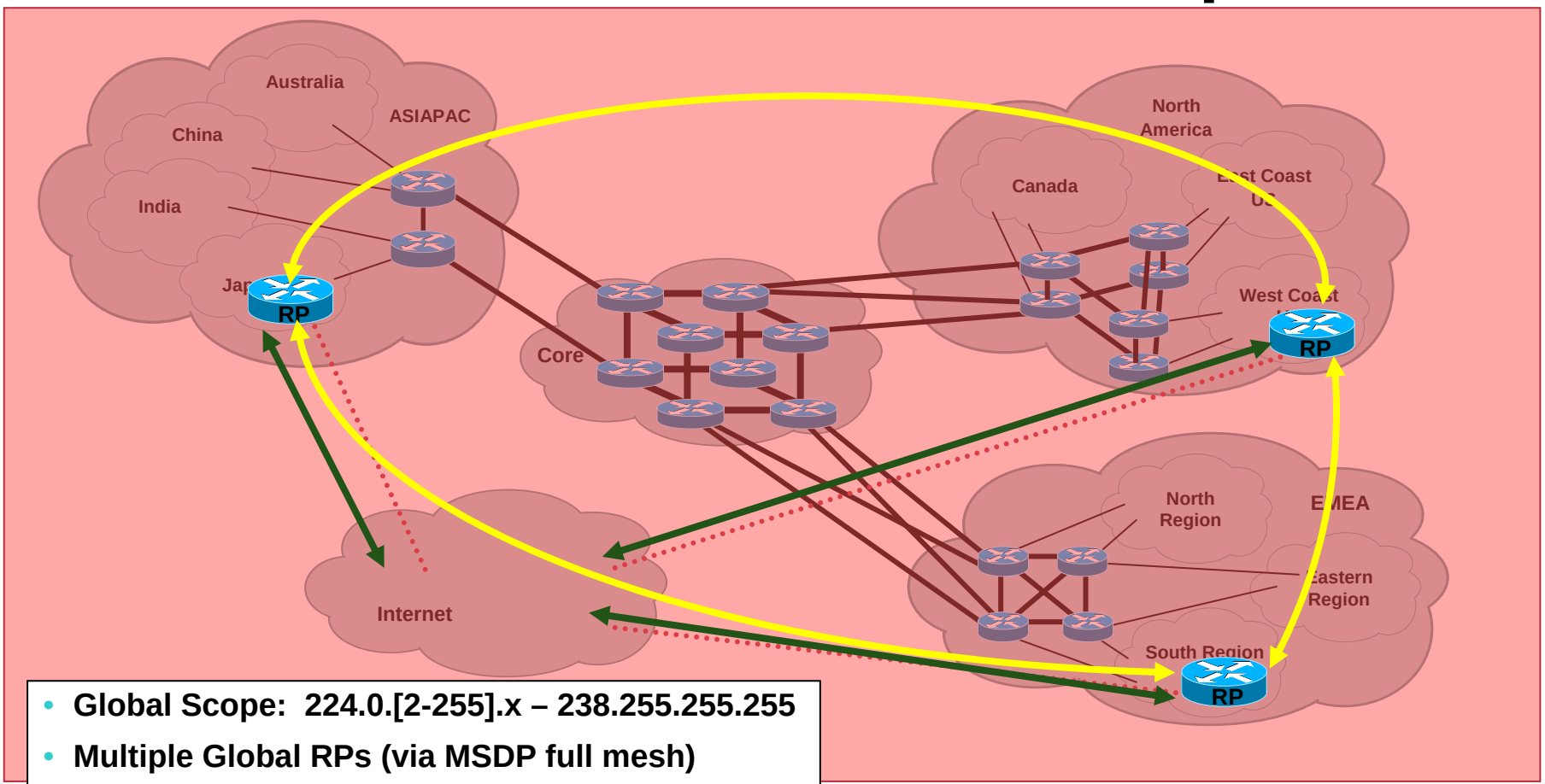


- **Enterprise Scope: 239.192.x.x/16**
- **Multiple Admin Scoped RPs (via MSDP full mesh)**
- **Private Address Space within the Enterprise, no Internet connectivity for this range**

Administratively-Scoped Zones Example

Cisco.com

Level 4: Internet Global Scope



- Global Scope: 224.0.[2-255].x – 238.255.255.255
- Multiple Global RPs (via MSDP full mesh)
- MSDP connectivity to provider's RPs

Advanced Multicast Engineering

- Multicast Group Control
- Using Admin. Scoped Zones
- PIM Protocol Extensions

PIM Protocol Extensions

- **Source Specific Multicast**
- **Bidirectional (Bidir) PIM**

Barriers to Multicast Deployment

- Global Multicast Address Allocation
  - Dynamic Address Allocation
    - No adequate dynamic address allocation methods exist
    - SDR – Doesn't scale
    - MASC – Long ways off!
  - Static Address Allocation (GLOP)
    - Based on AS number.
    - Insufficient address space for large Content Providers.
- Multicast Content "Jammers"
  - Undesirable sources on a multicast group.
    - "Capt. Midnight" sources bogus data/noise to group.
    - Can cause DoS attack by congesting low speed links.

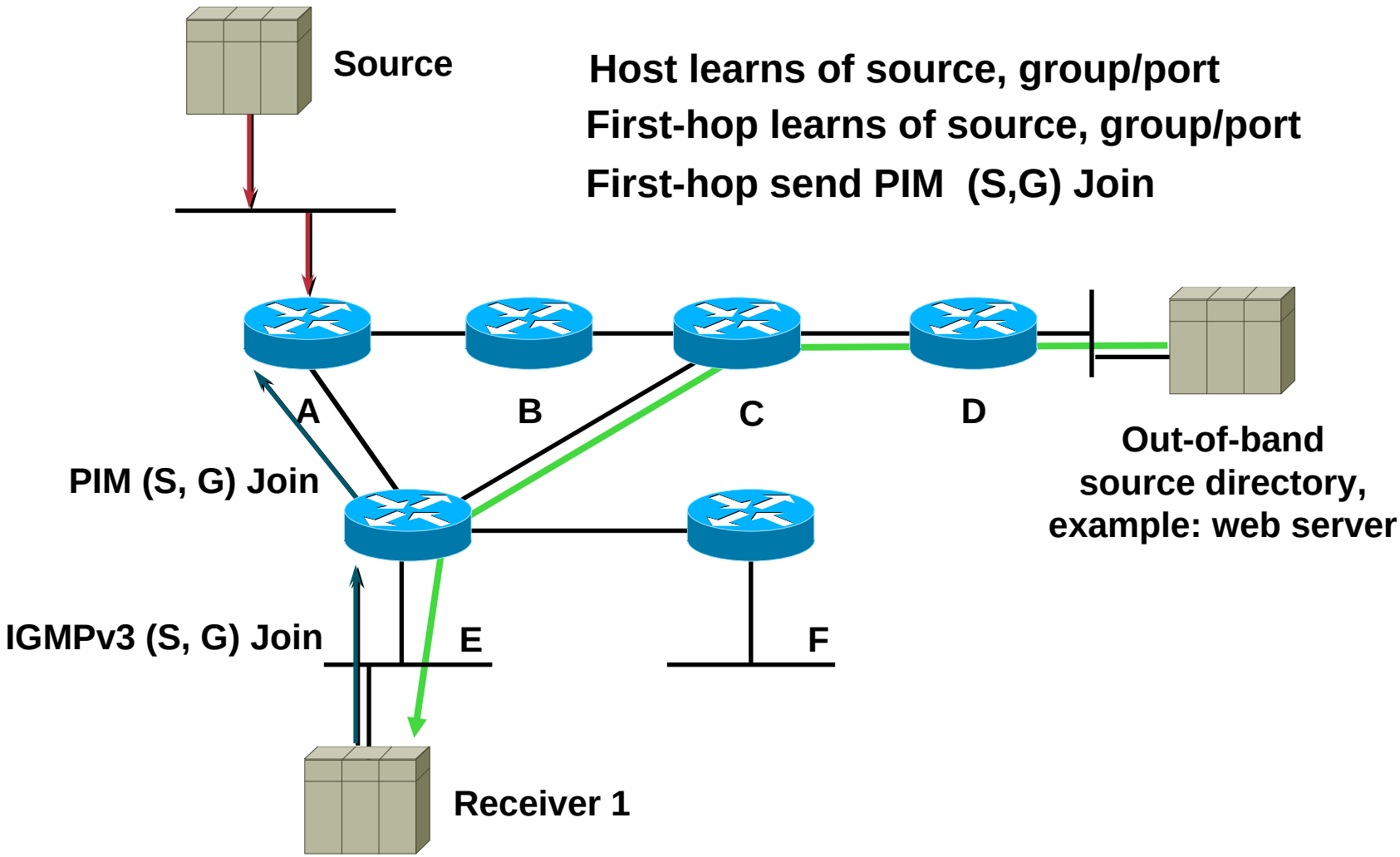
Source Specific Multicast (SSM)

- Uses Source Trees only.
- Assumes One-to-Many model.
  - Most Internet multicast fits this model.
  - IP/TV also fits this model.
- Hosts responsible for source discovery.
  - Typically via some out-of-band mechanism.
    - Web page, Content Server, etc.
  - Eliminates need for RP and Shared Trees.
  - Eliminates need for MSDP.

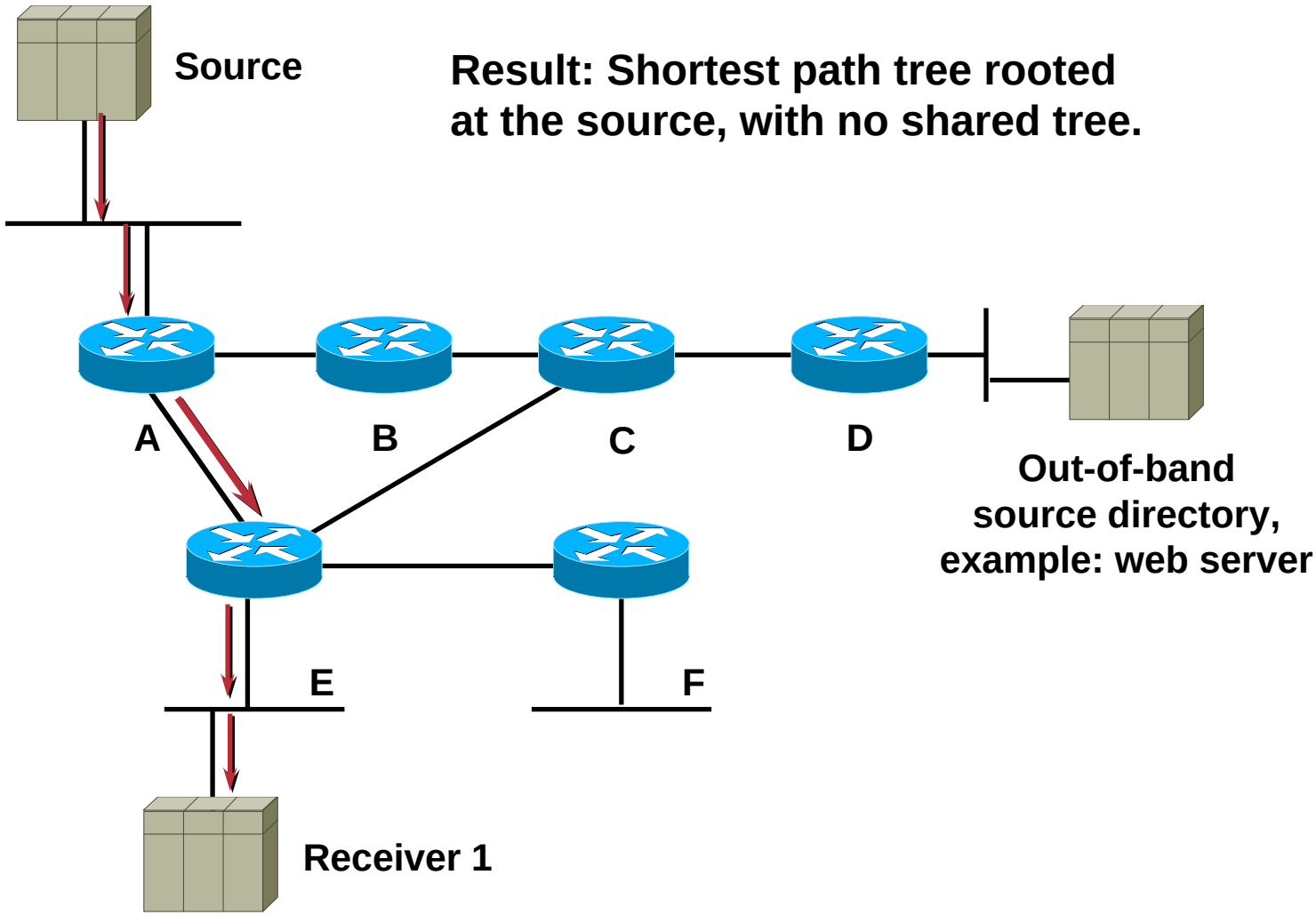
SSM Overview

- **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 each other.**

SSM Example



SSM Example



SSM Configuration

- **Global command**

```
ip pim ssm {default | <acl>}
```

  - **Defines SSM address range**
    - **Default range = 232.0.0.0/8**
    - **Use ACL for other ranges**
  - **Prevents Shared Tree Creation**
    - **(*, G) Joins never sent or processed**
    - **PIM Registers never sent or processed**
  - **Available in IOS versions**
    - **12.1(5)T, 12.2, 12.0(15)S, 12.1(8)E**

SSM Configuration of Legacy Routers

- Only Last-Hop routers ***must*** be upgraded.
  - Core may be upgraded later.
- Must insure no Shared Trees in SSM range.
  - Use '`ip pim accept-register`' at RP.
    - Prevents sources from registering in 232/8.
  - Use '`ip pim accept-rp`' on all routers.
    - Prevents (*,G) Joins from being processed for 232/8.
  - Use '`ip msdp sa-redistribute`' at RP.
    - Stops SA message origination in 232/8.
  - Use '`ip msdp sa-filter`' on MSDP peers.
    - Prevents forwarding of SA messages in 232/8.

SSM – Summary

- Uses Source Trees only.
  - Hosts are responsible for source & group discovery.
  - Hosts must signal router which (S,G) to join.
- Solves multicast address allocation problems.
  - Flows differentiated by ***both*** source and group.
  - Content providers can use same group ranges.
    - Since each (S,G) flow is unique.
- Helps prevent certain DoS attacks
  - "Bogus" source traffic:
    - Can't consume network bandwidth.
    - Not received by host application.

So where is SSM?

- **Dependant on IGMPv3**
  - **Microsoft supports IGMPv3 in Windows XP**
    - **Workarounds**
  - **IGMPv3 lite**
    - **Free available API/Library/DLL - www.talarian.com**
    - **Used by Cisco IP/TV 3.2**
  - **URL RenDezvous (URD)**
    - **Redirect from Web page with specific information intercepted by Router**
  - **Static Source Mapping**
    - **Router maps IGMPv2 Joins in SSM range to well-known sources via DNS or static configuration**

IGMP v3lite

- Source side:
  - No application changes required!
- Receiver side:
  - Application must use IGMPv3 API:
  - IGMP v3lite Library Component
    - Provides the IP SSM subset of IGMPv3 API
      - Applications must still filter out unwanted traffic.
  - IGMP v3lite Daemon Component
    - Sends special (S,G) Join to local router via UDP port 465

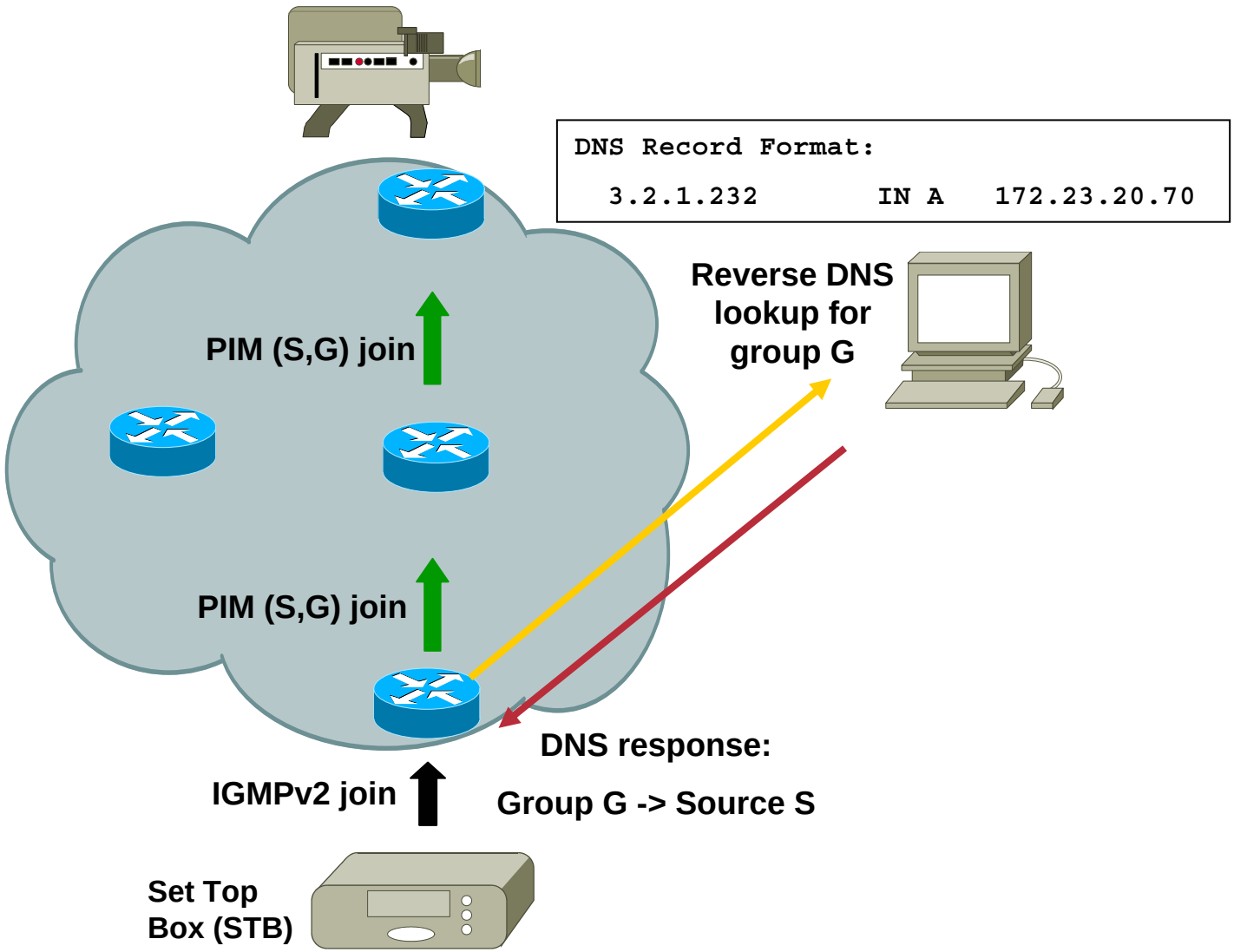
URD

- A content provider builds a web page that contains URD links.
  - List of sources willing to provide multicast content
- The user (receiver) clicks on one of the links
- Web Server sends back an HTTP redirect containing source and group info to TCP port 465
- Host sends the redirect via TCP port 465
- Local router intercepts TCP port 465 traffic
  - Uses source/group information in the redirect to identify the requested SSM flow.

SSM Mapping

- Allows only for one source per Group
- Router maps group to a single source
  - Uses either DNS or static internal database
    - DNS method allows content providers to provide the mapping
    - DNS Method independent from network operators

SSM Mapping – DNS Example



SSM Mapping Configuration

Cisco.com

Enabling SSM mapping on the router

```
ip igmp ssm-map enable
```

For static mapping:

```
ip igmp ssm-map static <acl-1> <source-1 IP address>
ip igmp ssm-map static <acl-2> <source-2 IP address>
```

For DNS mapping (existing commands):

```
ip domain-server <ip address>
ip domain-name <domain.com>
```

To disable DNS mapping

```
no ip igmp ssm-map query dns
```

```
DNS Record Format:   3.2.1.232        IN A   172.23.20.70
```

PIM Protocol Extensions

- Source Specific Multicast
- Bidirectional (Bidir) PIM

Multicast Application Categories

- One-to-Many Applications
  - Video, TV, Radio, Concerts, Stock Ticker, etc.
- Few-to-Few Applications
  - Small (<10 member) Video/Audio Conferences
- Few-to-Many Applications
  - TIBCO RV Servers (Publishing)
- Many-to-Many Applications
  - Stock Trading Floors, Gaming
- Many-to-Few Applications
  - TIBCO RV Clients (Subscriptions)

Multicast Application Categories
PIM-SM (S, G) State

- One-to-Many Applications
  - Single (S,G) entry
- Few-to-Few Applications
  - Few (<10 typical) (S,G) entries
- Few-to-Many Applications
  - Few (<10 typical) (S,G) entries
- Many-to-Many Applications
  - Unlimited (S,G) entries
- Many-to-Few Applications
  - Unlimited (S,G) entries

Multicast State Maintenance

- CPU load factors
  - Must send/receive Registers
  - Must send periodic Joins/Prunes
  - Must perform RPF recalculation every 5 seconds
    - Watch the total number of mroute table entries
    - Unicast route table size impacts RPF recalculation
- Memory load factors
  - (*, G) entry ~ 380 bytes + OIL size
  - (S, G) entry ~ 220 bytes + OIL size
  - Outgoing interface list (OIL) size
    - Each oil entry ~ 150 bytes

Many-to-Any State Problem

- Creates huge amounts of (S,G) state
  - State maintenance workloads skyrocket
    - High OIL fanouts make the problem worse
  - Router performance begins to suffer
- Using Shared-Trees only
  - Provides some (S,G) state reduction
    - Results in (S,G) state only along SPT to RP
    - Frequently still too much (S,G) state
    - Need a solution that only uses (*,G) state

Bidirectional (Bidir) PIM

- **Idea:**
  - **Use the same tree for traffic from sources towards RP and from RP to receivers**
- **Benefits:**
  - **Less state in routers**
    - **Only (*, G) state is used**
    - **Source traffic follows the Shared Tree**
      - **Flows up the Shared Tree to reach the RP.**
      - **Flows down the Shared Tree to reach all other receivers.**

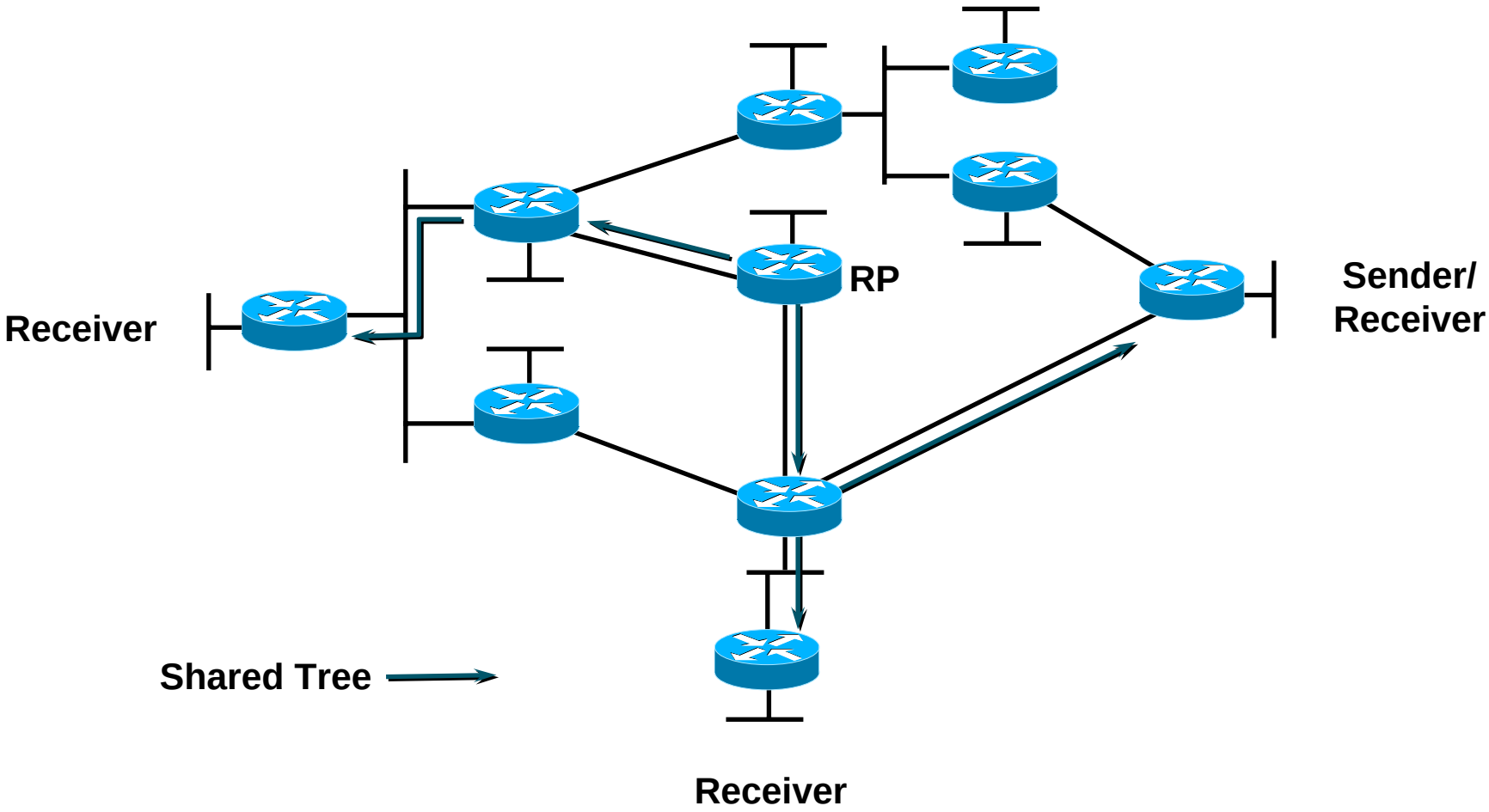
Bidirectional (Bidir) PIM

- **Bidirectional Shared-Trees**
  - **Violates current (*,G) RPF rules**
    - **Traffic often accepted on *outgoing* interfaces.**
    - **Care must be taken to avoid multicast loops**
  - **Requires a Designated Forwarder (DF)**
    - **Responsible for forwarding traffic up Shared Tree**
      - **DF's will accept data on the interfaces in their OIL.**
      - **Then send it out all other interfaces. (Including the IIF.)**

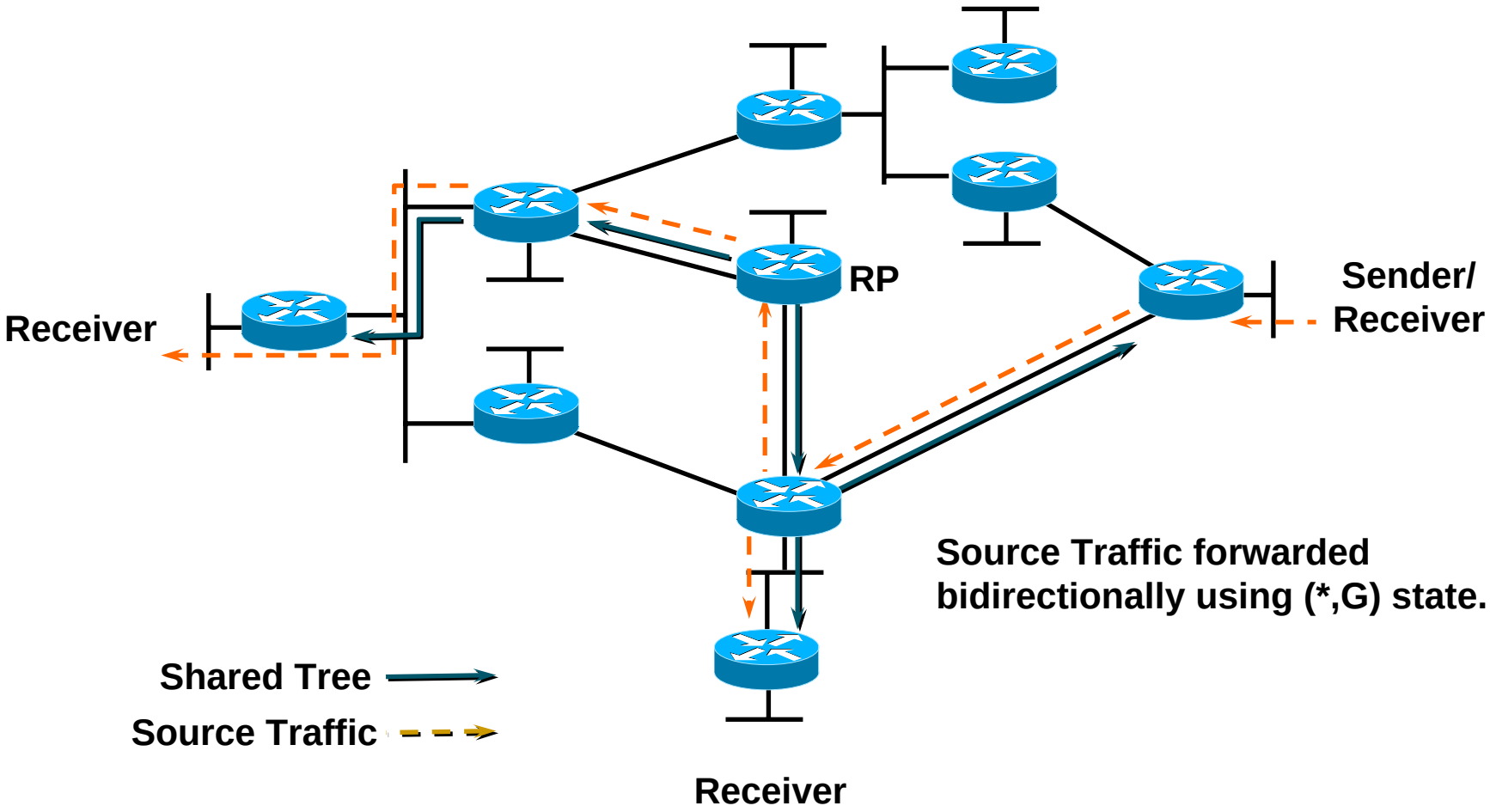
Bidirectional (Bidir) PIM

- Designated Forwarders (DF)
  - On each link the router with the best path to the RP is elected to be the DF
    - Note: Designated Routers (DR) are not used for bidir groups
  - The DF is responsible for forwarding traffic upstream towards the RP
  - No special treatment is required for local sources

Bidirectional PIM — Example



Bidirectional PIM — Example

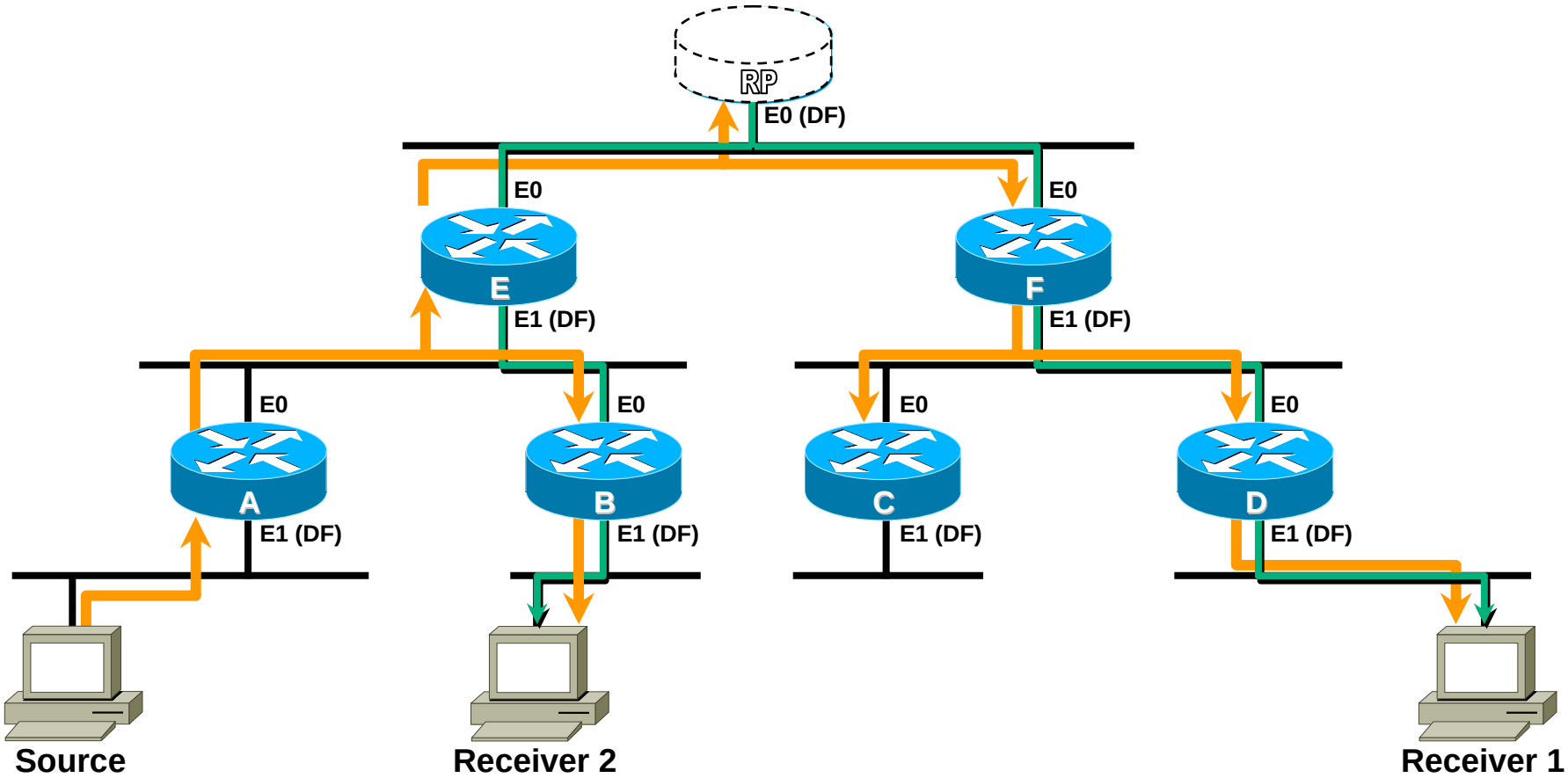


Configuring Bidir PIM
(Auto-RP Example)

- **Define Candidate RP and groups / modes it is willing to serve**

```
ip pim send-rp-announce Loopback0 scope 10 group-list 45 bidir
ip pim send-rp-announce Loopback1 scope 10 group-list 46
! Two loopbacks needed due to a nature of ACLs (permit, deny)
ip pim send-rp-discovery scope 10

access-list 45 permit 224.0.0.0 0.255.255.255
access-list 45 permit 227.0.0.0 0.255.255.255
! 224/8 and 227/8 will be PIM Bidir groups
access-list 45 deny 225.0.0.0 0.255.255.255
! 225/8 will be a PIM Dense Mode group

access-list 46 permit 226.0.0.0 0.255.255.255
! 226/8 will be a PIM Sparse Mode group
```

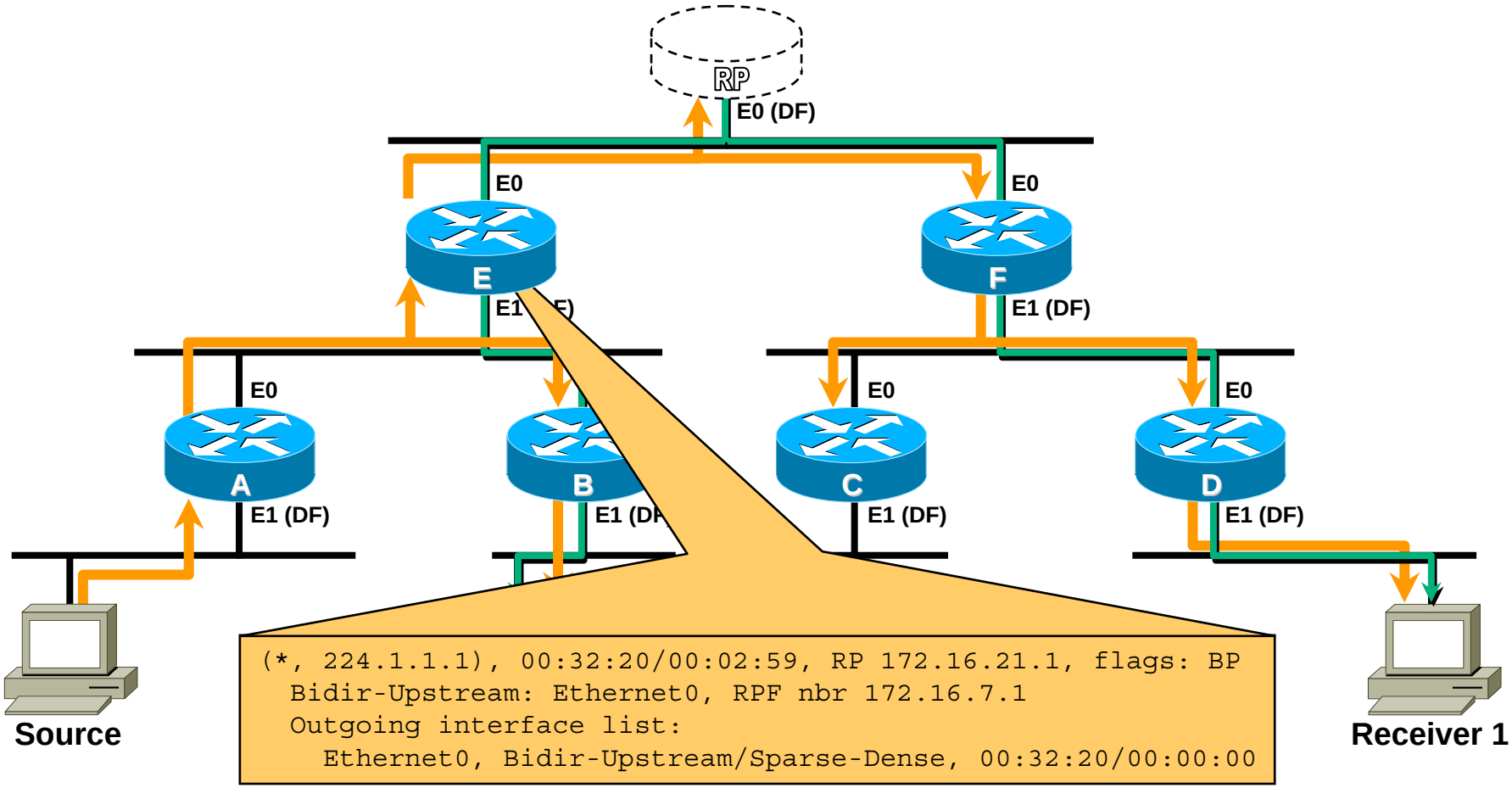
Bidir PIM – Phantom RP



Question: Does a Bidir RP even have to physically exist?

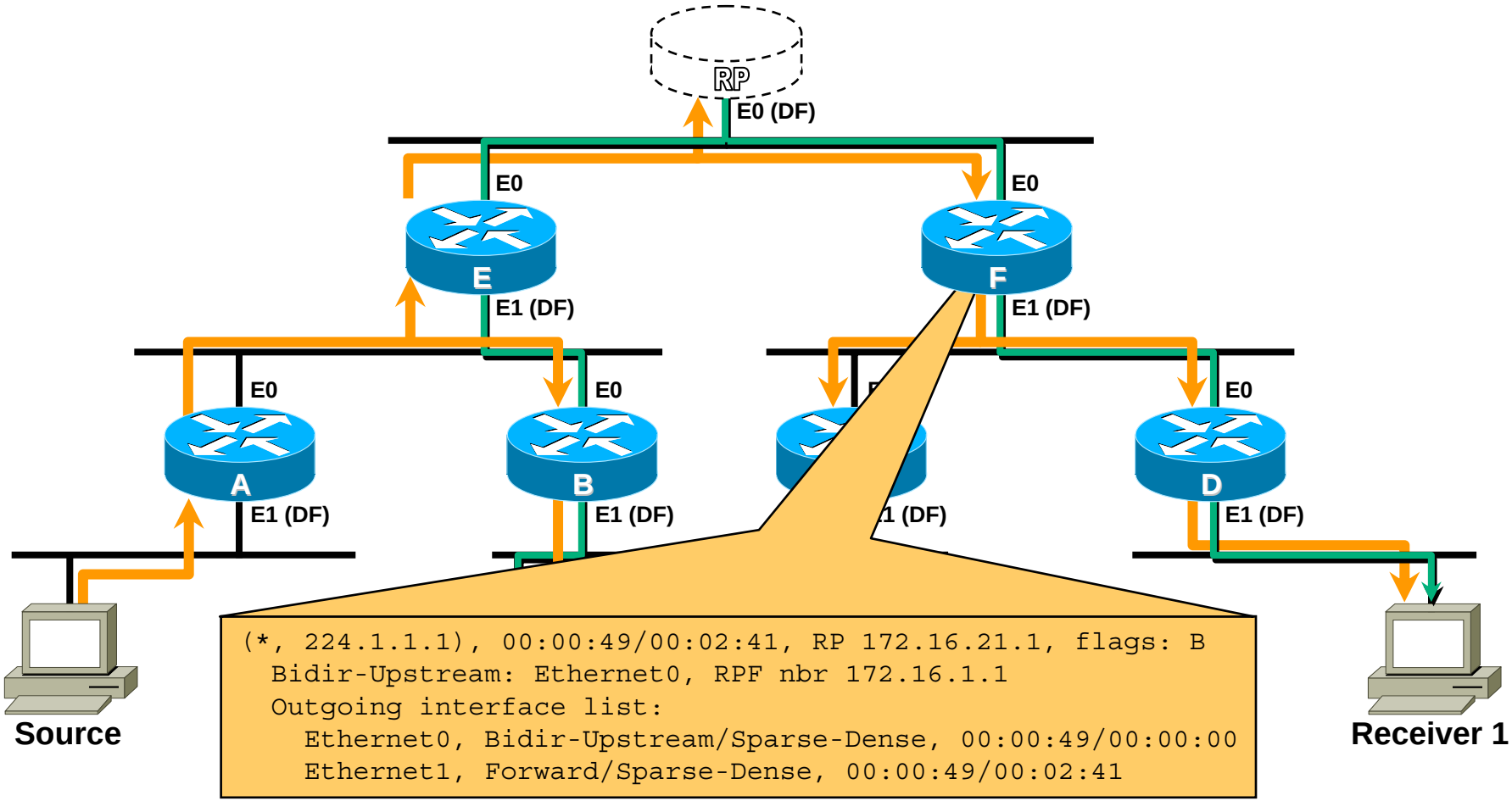
Answer: No. It can just be a phantom address.

Bidir PIM – Phantom RP



Router “E” forwards traffic onto core LAN segment.

Bidir PIM – Phantom RP

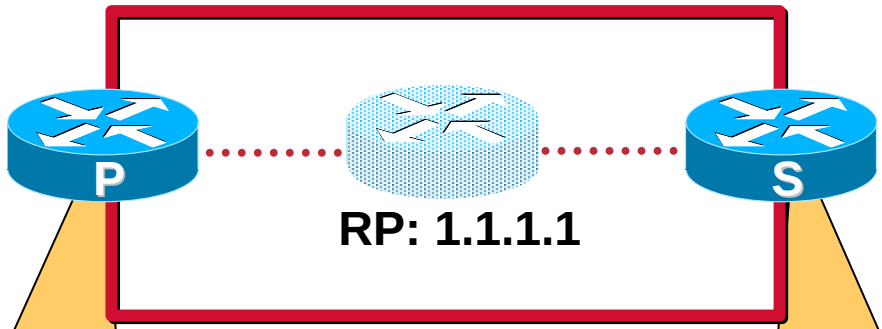


Router “F” forwards traffic on down the Shared Tree ala normal PIM-SM.

RP doesn’t even have to physically exist.

Phantom RP on Point-to-Point Core

Static Route Method

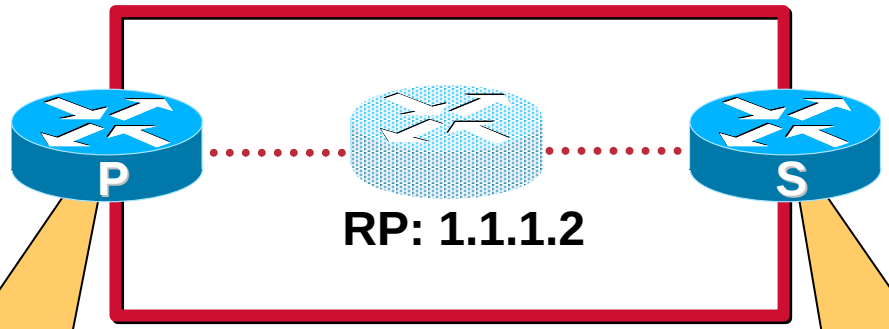



```
ip multicast-routing

interface Loopback0
 ip address 11.0.0.1 255.255.255.255
 ip pim sparse-mode

router ospf 11
 redistribute static subnets

ip route 1.1.1.1 255.255.255.255 Loopback0

ip pim bidir-enable
ip pim rp-address 1.1.1.1 bidir
```

```
ip multicast-routing

interface Loopback0
 ip address 11.0.0.2 255.255.255.255
 ip pim sparse-mode

router ospf 11
 redistribute static subnets

ip route 1.1.1.0 255.255.255.254 Loopback0

ip pim bidir-enable
ip pim rp-address 1.1.1.1 bidir
```

Phantom RP on Point-to-Point Core

Netmask Method




```
ip multicast-routing
!
interface Loopback0
 ip address 1.1.1.1 255.255.255.252
 ip pim sparse-mode
 ip ospf network point-to-point
!
router ospf 11
 network 1.1.1.0 0.0.0.3 area 0
 network 10.1.1.0 0.0.0.255 area 0
 network 10.1.2.0 0.0.0.255 area 0
!
ip pim bidir-enable
ip pim rp-address 1.1.1.1
ip pim rp-address 1.1.1.2 bidir
```

```
ip multicast-routing
!
interface Loopback0
 ip address 1.1.1.1 255.255.255.248
 ip pim sparse-mode
 ip ospf network point-to-point
!
router ospf 11
 network 1.1.1.0 0.0.0.7 area 0
 network 10.1.1.0 0.0.0.255 area 0
 network 10.1.2.0 0.0.0.255 area 0
!
ip pim bidir-enable
ip pim rp-address 1.1.1.1
ip pim rp-address 1.1.1.2 bidir
```

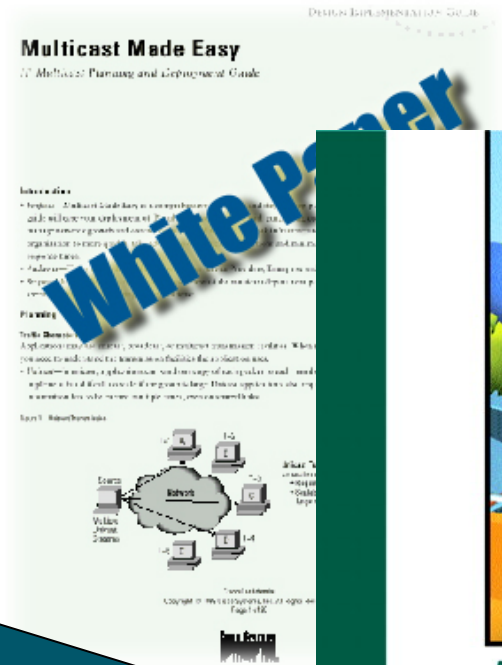
Bidir PIM—Summary

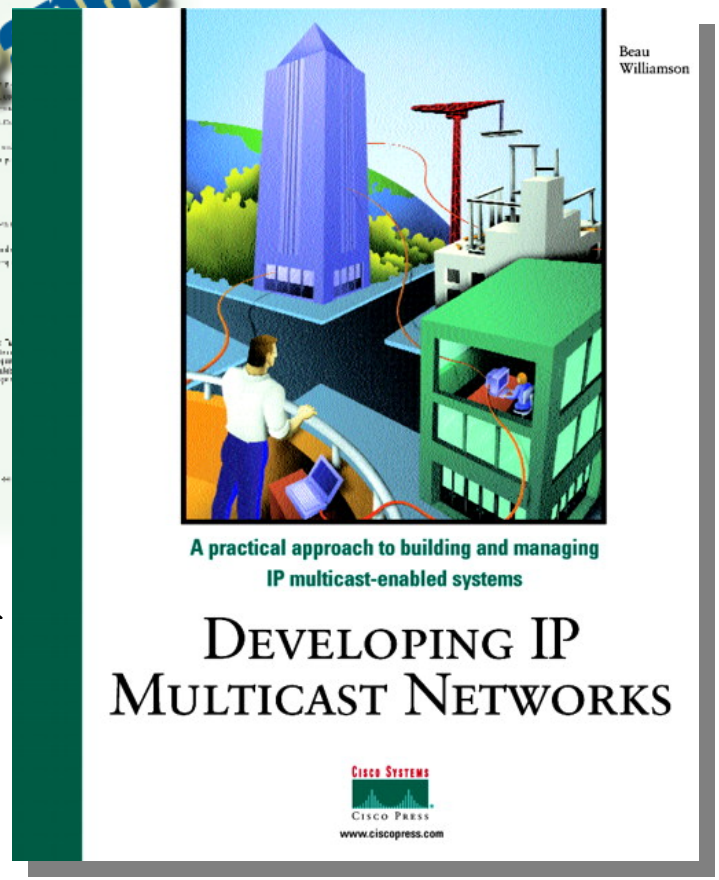
- **Drastically reduces network mroute state**
  - **Eliminates *ALL* (S,G) state in the network**
    - **SPT's between sources to RP eliminated**
    - **Source traffic flows both up and down Shared Tree**
  - **Allows Many-to-Any applications to scale**
    - **Permits virtually an unlimited number of sources**

More Information

- White Papers
- Web and Mailers
- Cisco Press

RTFB





CCO Multicast page:
http://www.cisco.com/go/ipmulticast
Questions:
cs-ipmulticast@cisco.com
Customer Support Mailing List:
tac@cisco.com

RTAB = "Read the Fine Book"

Please Complete Your Evaluation Form

Session RST-2051