Advanced IP Multicast

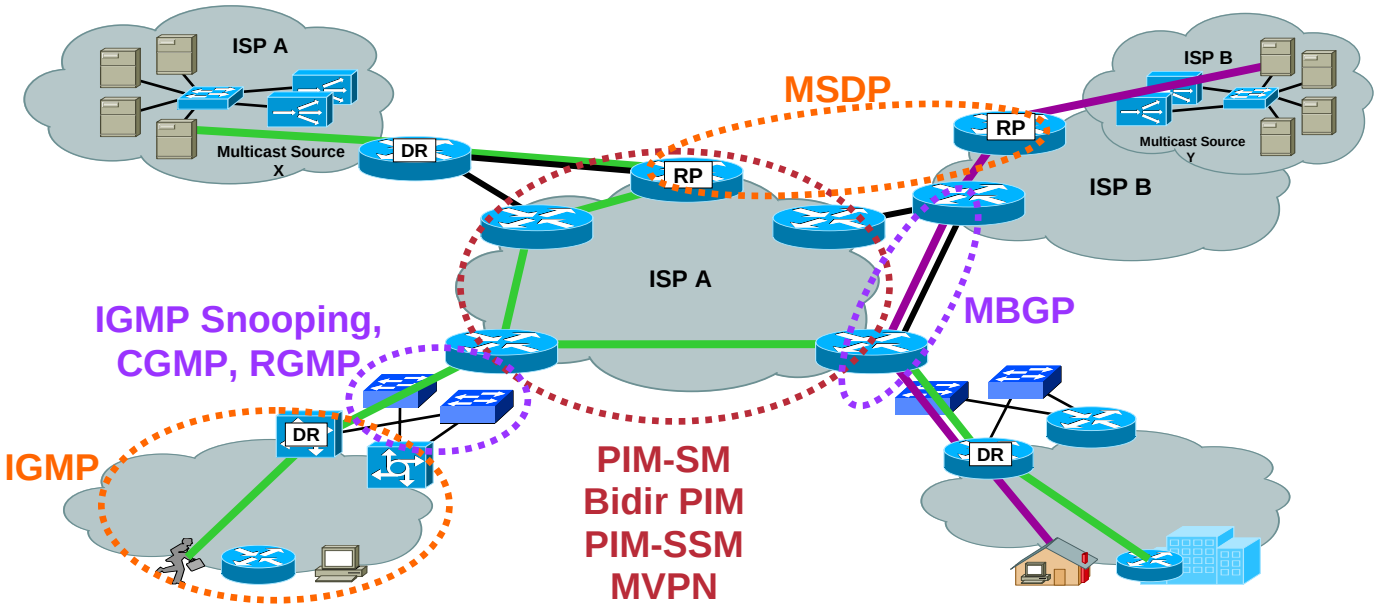
Session RST-4051

Agenda

- **MBGP (routing)**
- **MSDP (source discovery)**
- **MBGP/MSDP Examples**
- **SSM (Source Specific Multicast)**
- **MVPN (Multicast VPN)**
- **IPv6**
- **Security**

Multicast Components

Cisco End-to-End Architecture



Campus Multicast

- End Stations (hosts-to-routers):
  - IGMP
- Switches (Layer 2 Optimization):
  - CGMP, IGMP Snooping or RGMP
- Routers (Multicast Forwarding Protocol):
  - PIM Sparse Mode or Bidirectional PIM

Interdomain Multicast

- Multicast routing across domains
  - MBGP
- Multicast Source Discovery
  - MSDP with PIM-SM
- Source Specific Multicast
  - PIM-SSM

MBGP Overview

- MBGP: Multiprotocol BGP
  - Defined in RFC 2283 (extensions to BGP)
  - Can carry different types of routes
    - IPv4 Unicast IPv6 Unicast
    - IPv4 Multicast IPv6 Multicast
  - May be carried in same BGP session
  - Does not propagate multicast state info
    - Still need PIM to build Distribution Trees
  - Same path selection and validation rules
    - AS-Path, LocalPref, MED, …

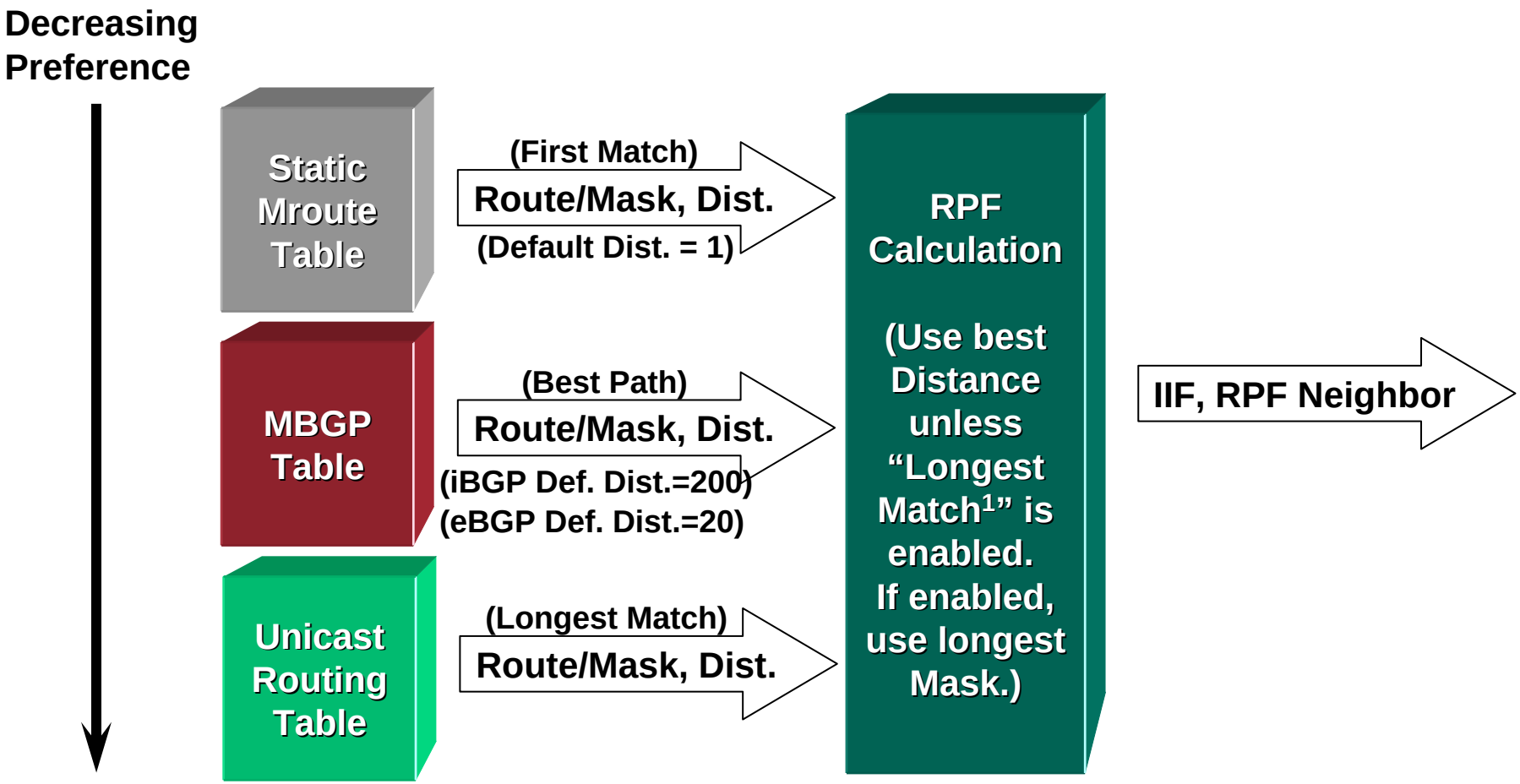
MBGP Overview

- Separate BGP tables maintained
  - Unicast BGP Table (U-Table)
  - Multicast BGP Table (M-Table)
  - BGP ‘nlri’ keyword specifies which BGP Table
  - Allows different unicast/multicast topologies or policies
- Unicast BGP Table (U-Table)
  - Contains unicast prefixes for unicast forwarding
  - Populated with BGP unicast NLRI
- Multicast BGP Table (M-Table)
  - Contains unicast prefixes for RPF checking
  - Populated with BGP multicast NLRI

MBGP Update Message

- Address Family Information (AFI)
  - Identifies Address Type (see RFC1700)
    - AFI = 1 (IPv4)
    - AFI = 2 (IPv6)
- Sub-Address Family Information (Sub-AFI)
  - Sub category for AFI Field
  - Address Family Information (AFI) = 1 (IPv4)
    - Sub-AFI = 1 (NLRI is used for unicast)
    - Sub-AFI = 2 (NLRI is used for multicast RPF check)

PIM RPF Calculation Details



Global Command: `ip multicast longest-match`

MBGP—Capability Negotiation

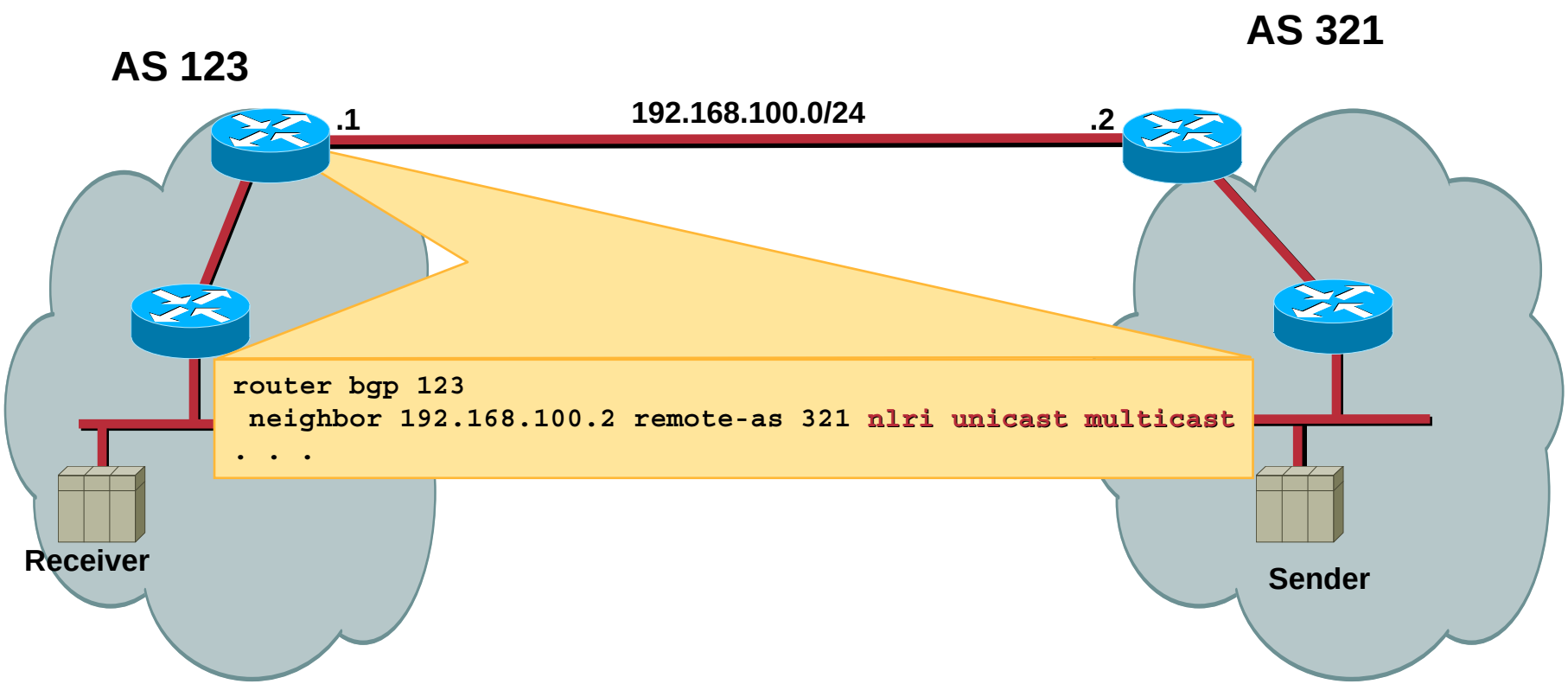
Cisco.com

- **Keyword on neighbor command**

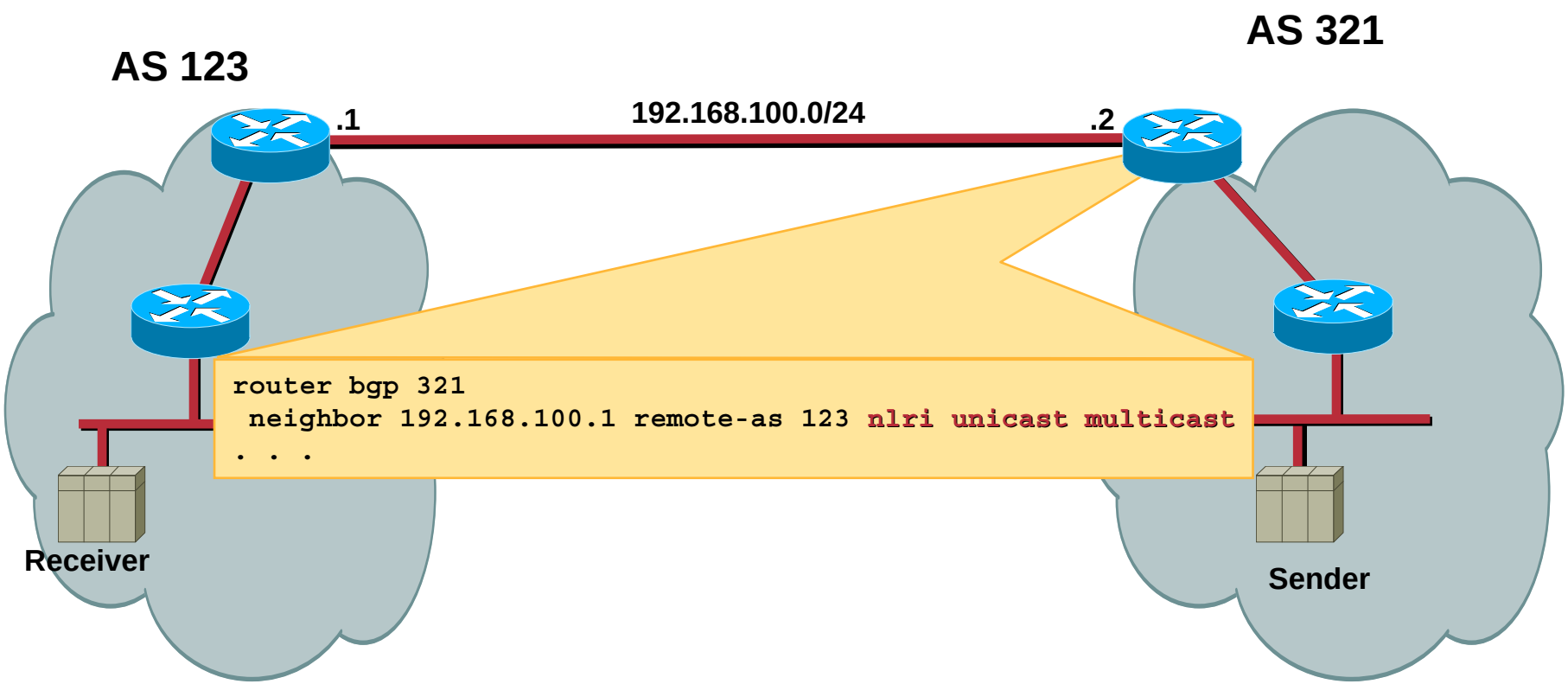
```
neighbor <foo> remote-as <asn> nlri multicast unicast
```

  - **Configures router to negotiate either or both NLRI**
  - **If neighbor configures both or subset, common NRLI is used in both directions**
  - **If there is no match, notification is sent and peering doesn't come up**

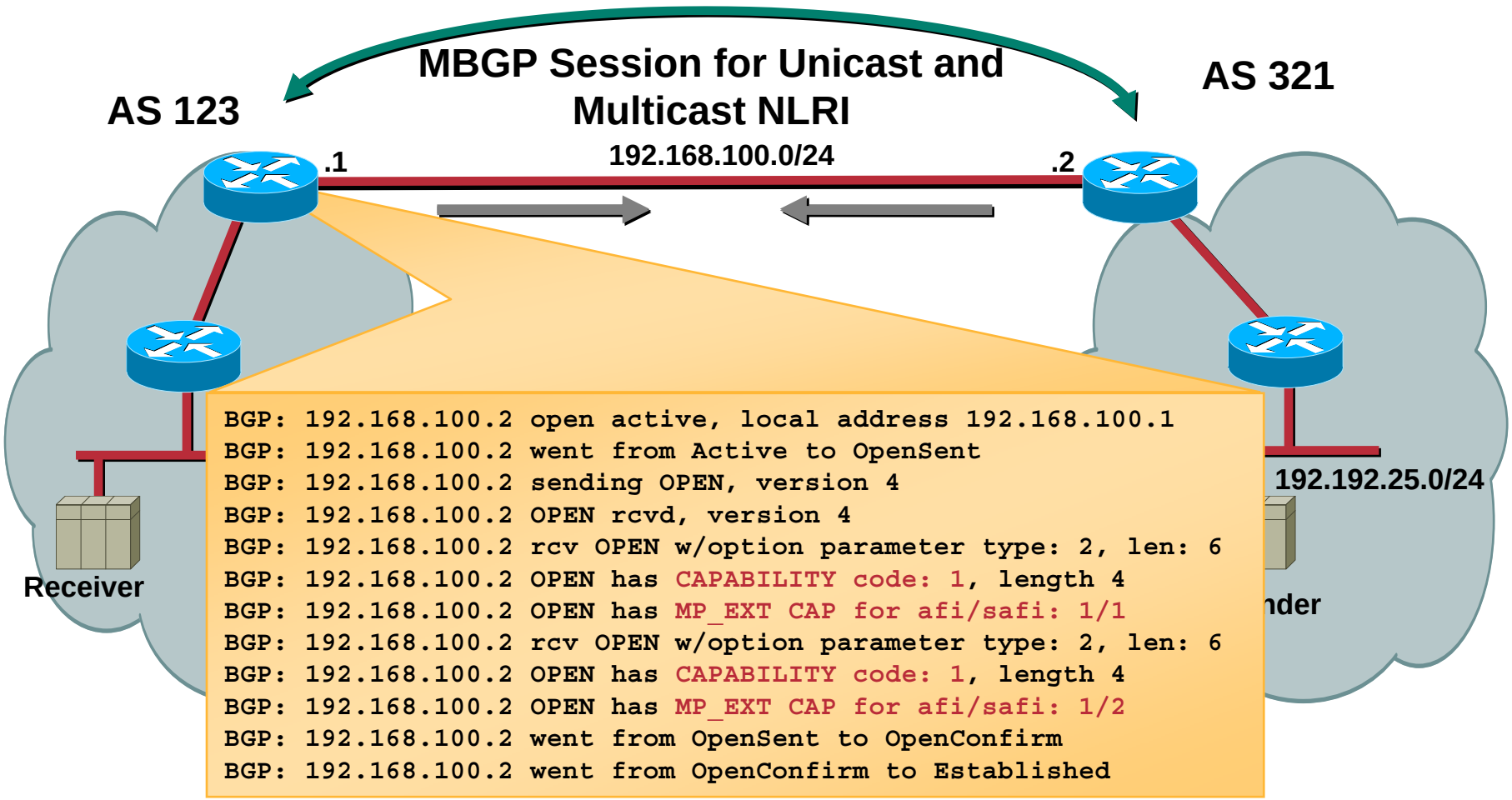
MBGP — Capability Negotiation



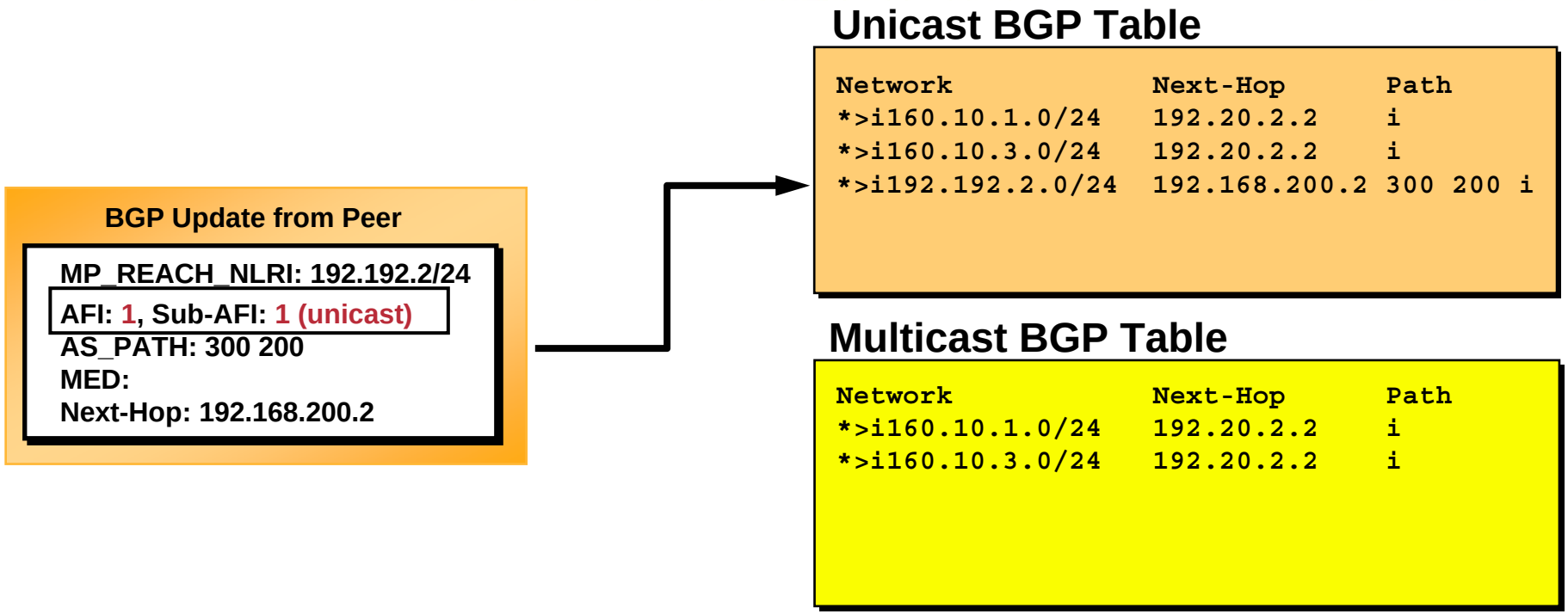
MBGP — Capability Negotiation



MBGP — Capability Negotiation



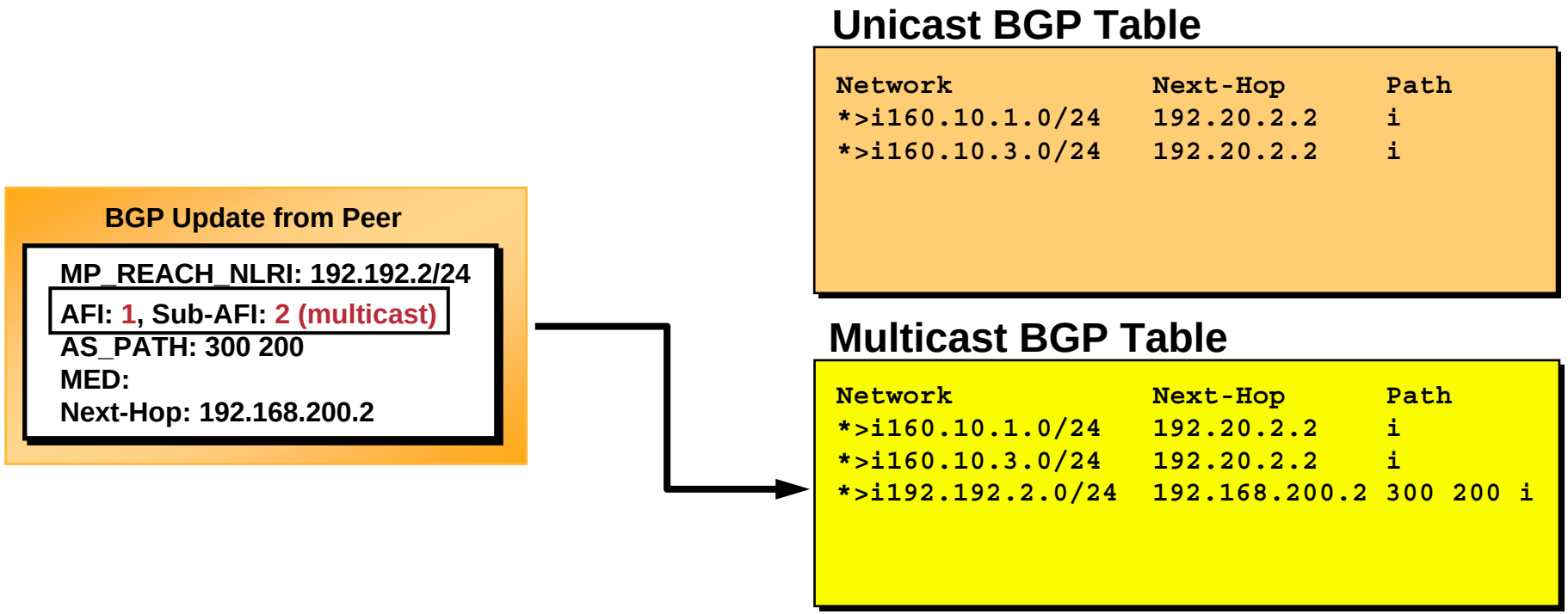
MBGP—NLRI Information



- **Storage of arriving NLRI information depends on AFI/SAFI fields in the Update message**
  - **Unicast BGP Table only (AFI=1/SAFI=1 or old style NLRI)**

MBGP—NLRI Information

Cisco.com

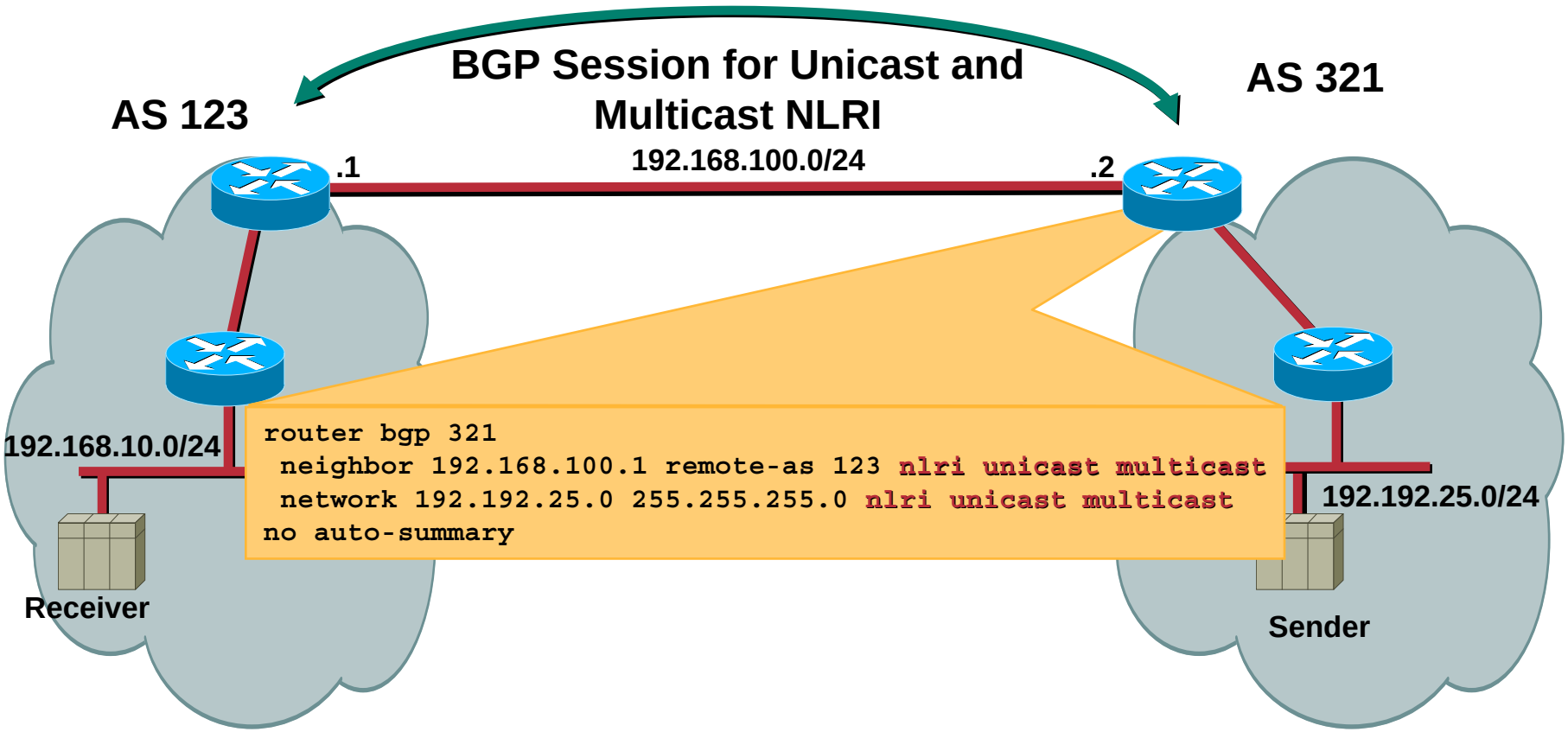


- Storage of arriving NLRI information depends on AFI/SAFI fields in the Update message
  - Unicast BGP Table only (AFI=1/SAFI=1 or old style NLRI)
  - Multicast BGP Table only (AFI=1/SAFI=2)

MBGP—NLRI Information

Cisco.com

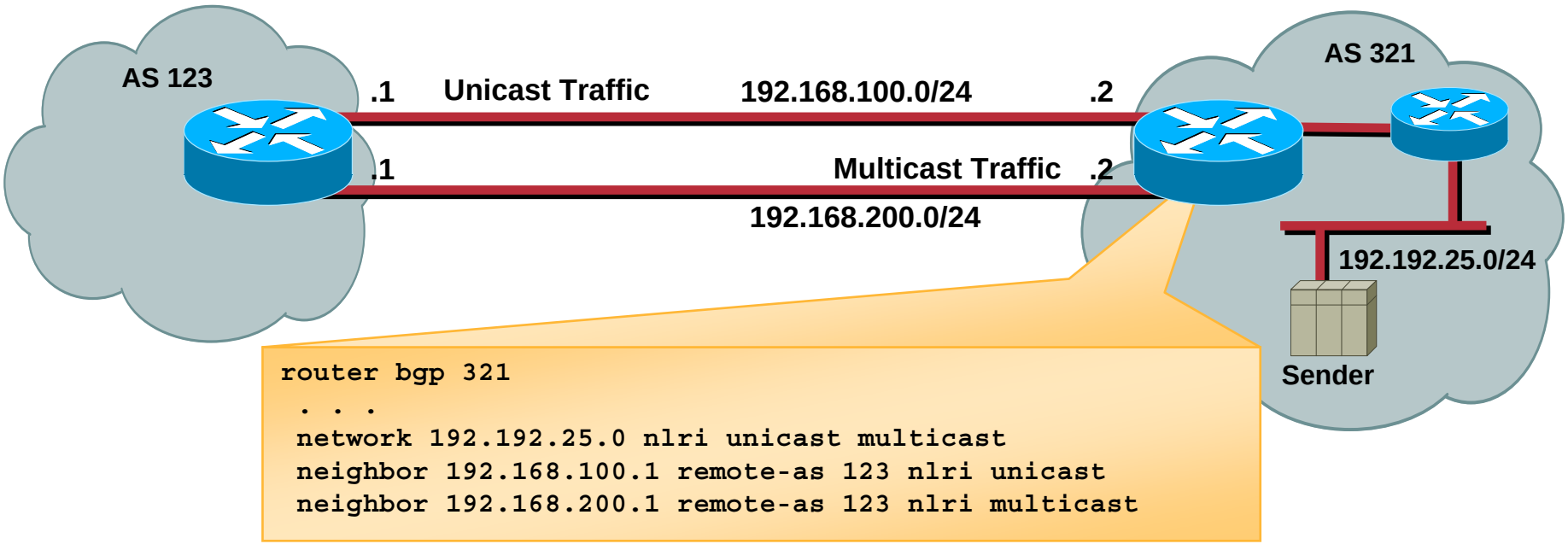
Congruent Topologies



MBGP—NLRI Information

Cisco.com

Incongruent Topologies



MBGP Syntax Change

NLRI Syntax

```
router bgp 5
  network 171.69.214.0 mask 255.255.255.0 nlri unicast multicast
  neighbor 171.69.214.38 remote-as 2 nlri unicast
  neighbor 171.69.214.50 remote-as 2 nlri multicast
```

Address-Family Syntax

```
router bgp 5
  no bgp default ipv4-unicast
  neighbor 171.69.214.38 remote-as 2
  neighbor 171.69.214.50 remote-as 2
!
  address-family ipv4 unicast
  neighbor 171.69.214.38 activate
  network 171.69.214.0 mask 255.255.255.0
  exit-address-family
!
  address-family ipv4 multicast
  neighbor 171.69.214.50 activate
  network 171.69.214.0 mask 255.255.255.0
  exit-address-family
```

MBGP—Summary

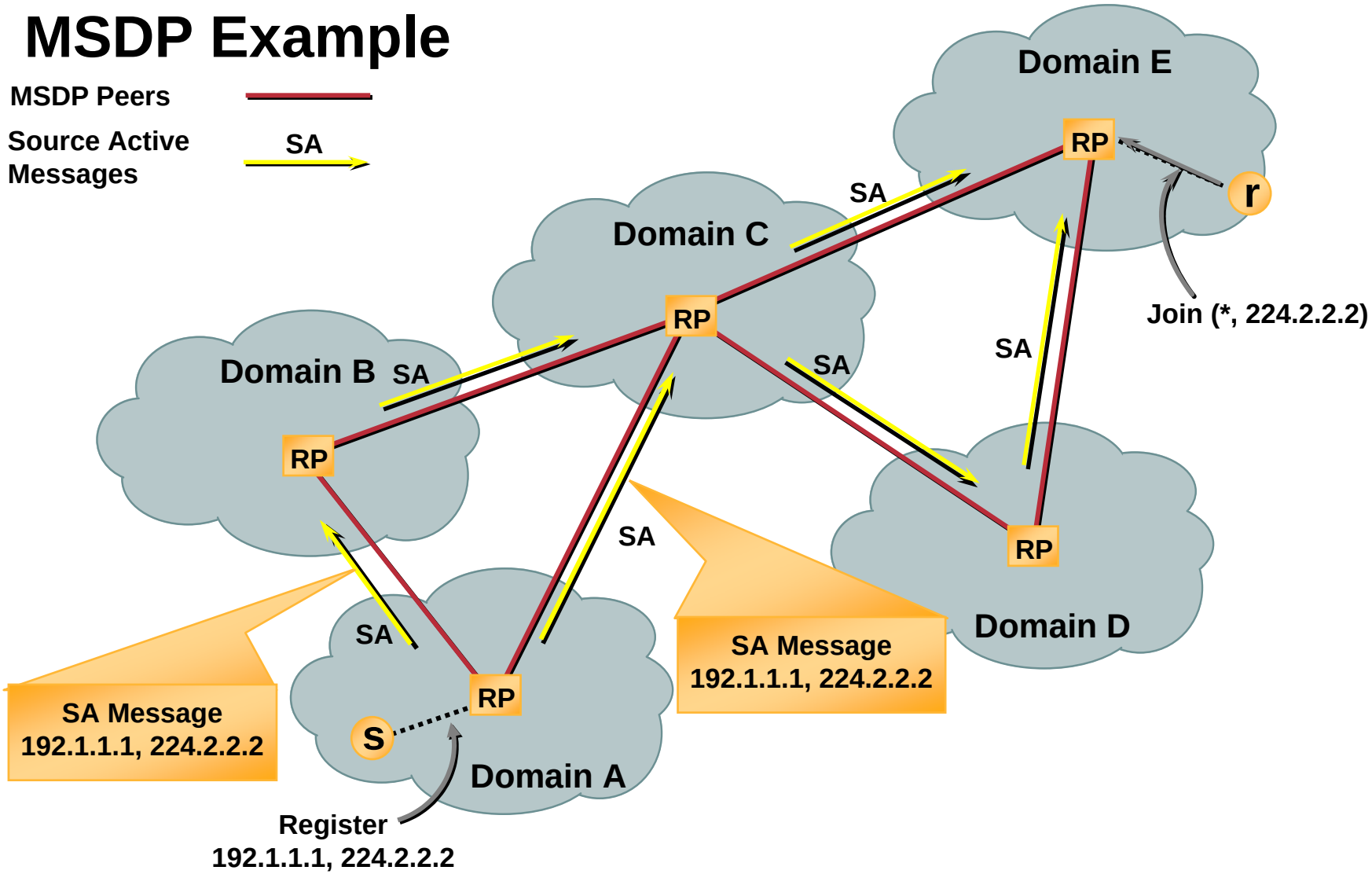
- Solves part of inter-domain problem
  - Can exchange multicast routing information
  - Uses standard BGP configuration knobs
  - Permits separate unicast and multicast topologies if desired
- Still must use PIM to:
  - Build distribution trees
  - Actually forward multicast traffic
  - PIM-SM recommended

Agenda

- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

MSDP Overview

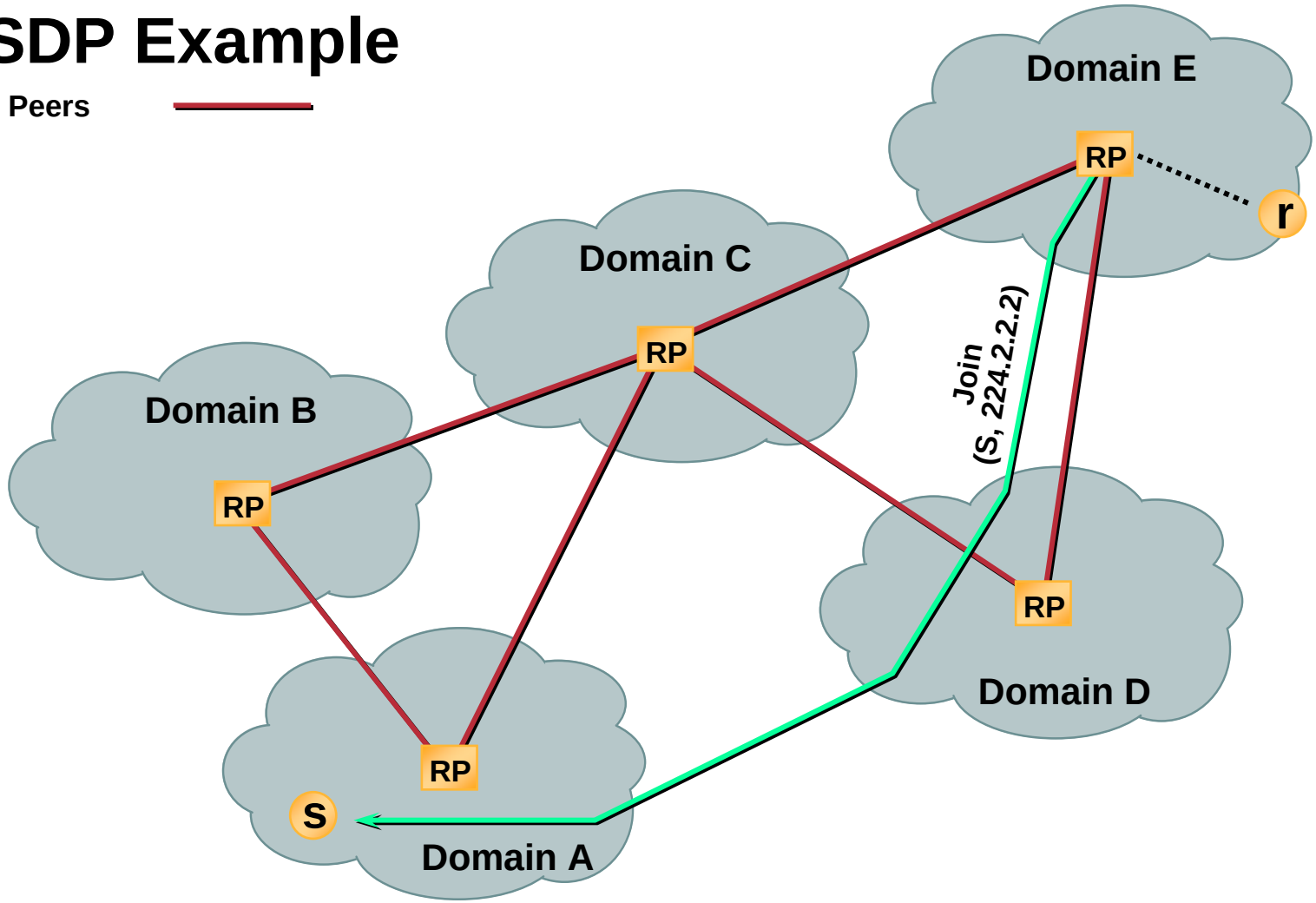
MSDP Example



MSDP Overview

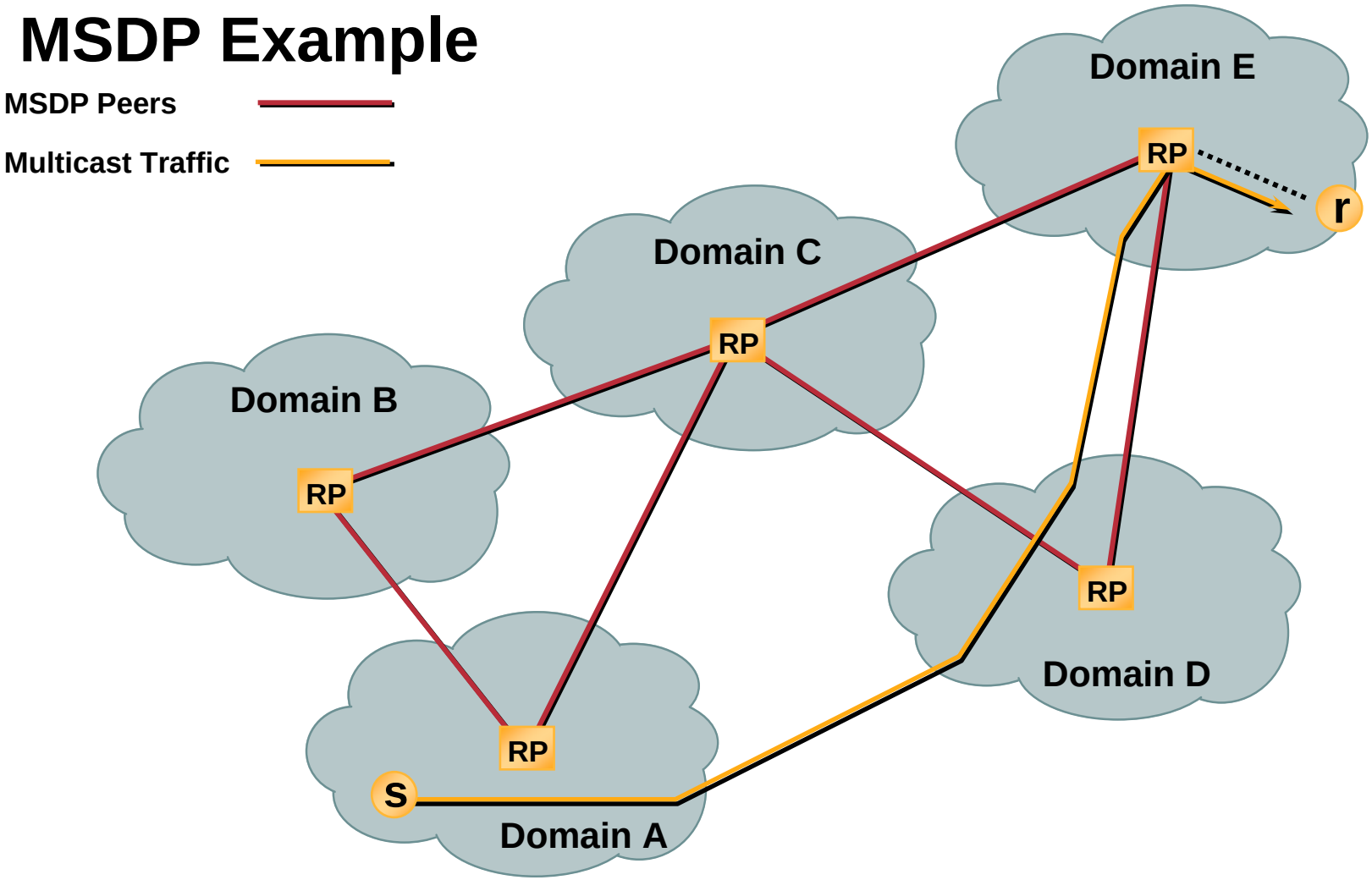
MSDP Example

MSDP Peers



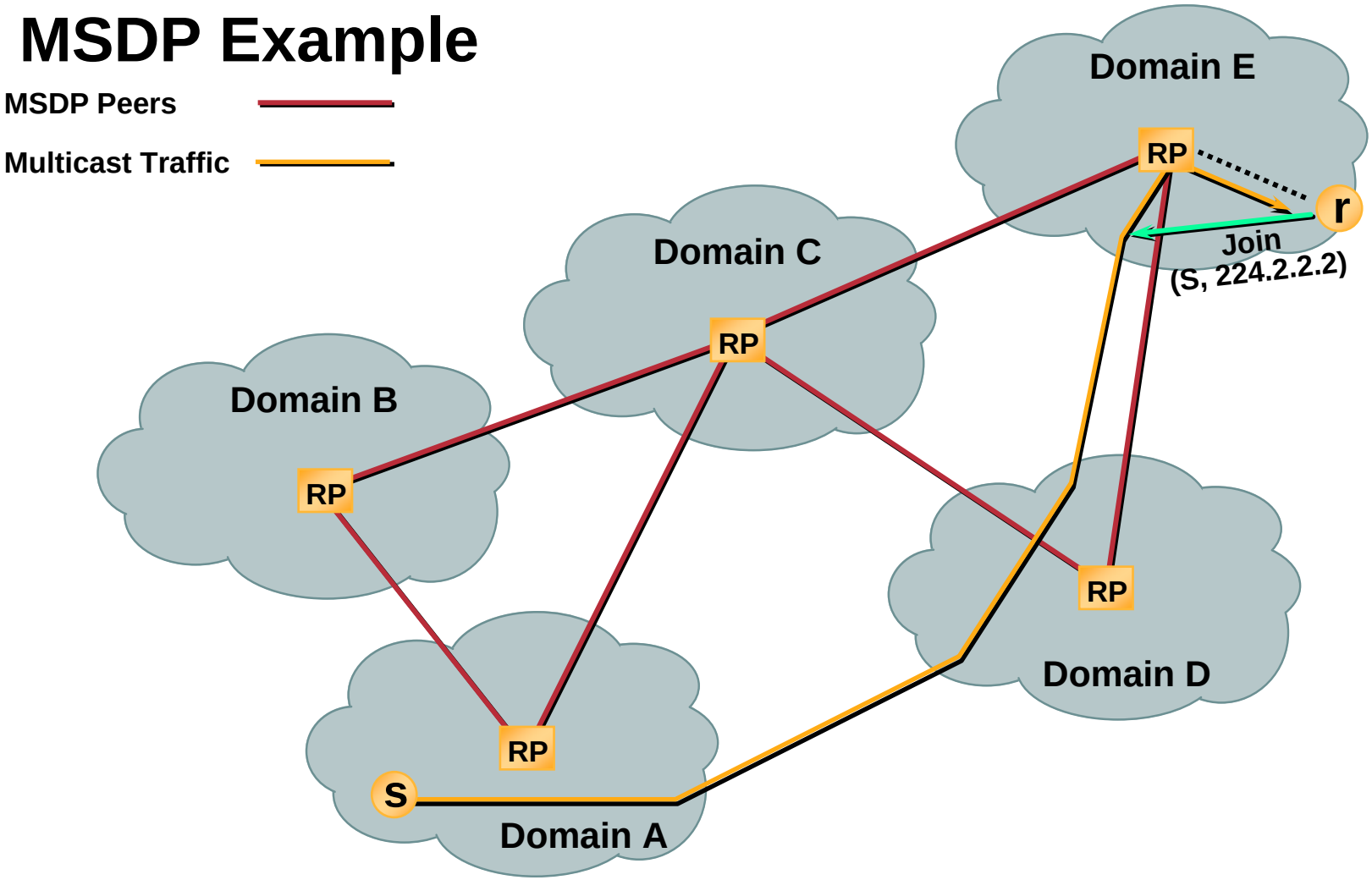
MSDP Overview

MSDP Example



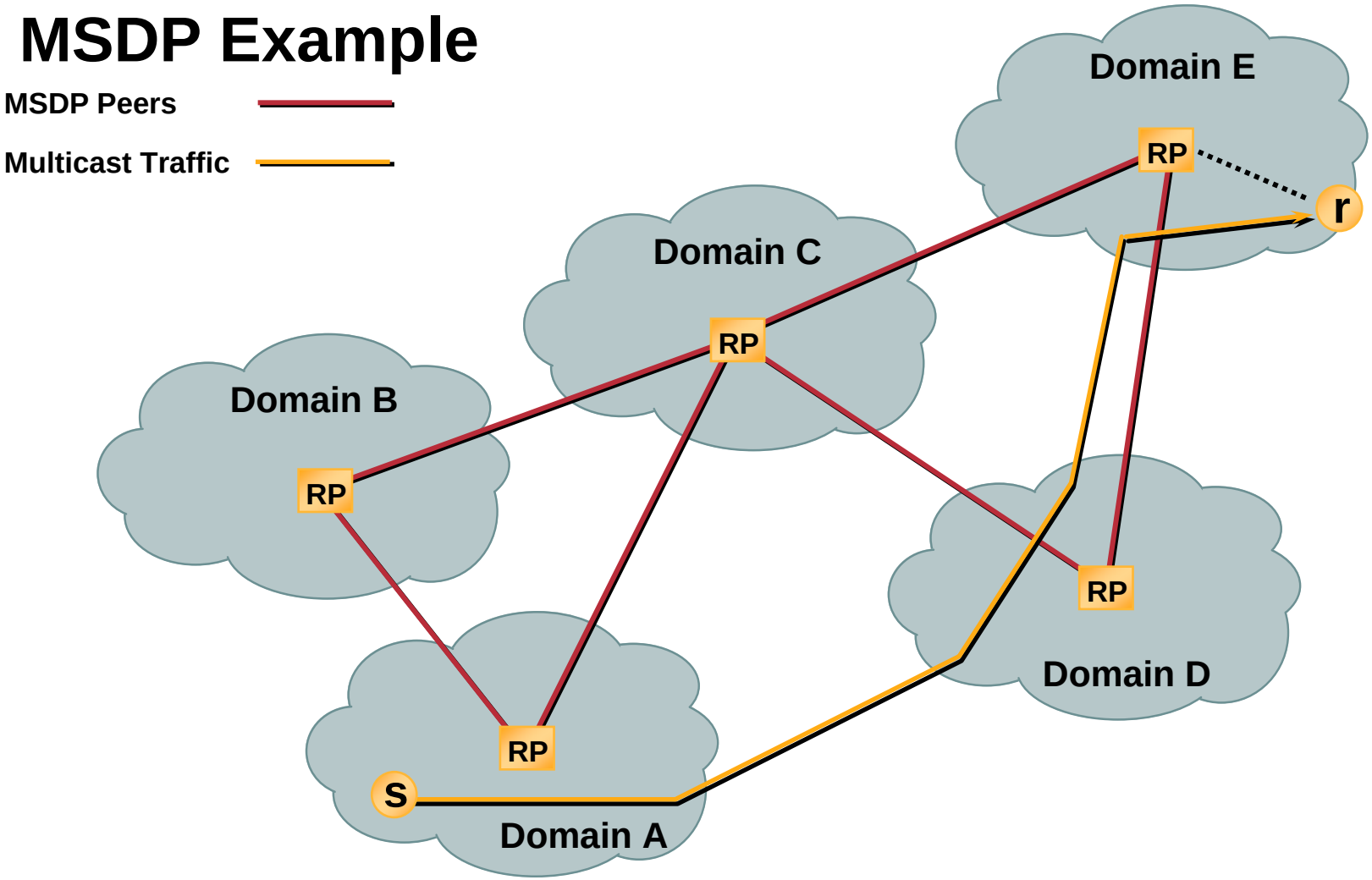
MSDP Overview

MSDP Example



MSDP Overview

MSDP Example



MSDP SA Messages

- MSDP Source Active (SA) Messages
  - Used to advertise active Sources in a domain
  - Carry 1st multicast packet from source
    - Hack for Bursty Sources (ala SDR)
  - SA Message Contents:
    - IP Address of Originator (RP address)
    - Number of (S, G)'s pairs being advertised
    - List of active (S, G)'s in the domain
    - Encapsulated Multicast packet

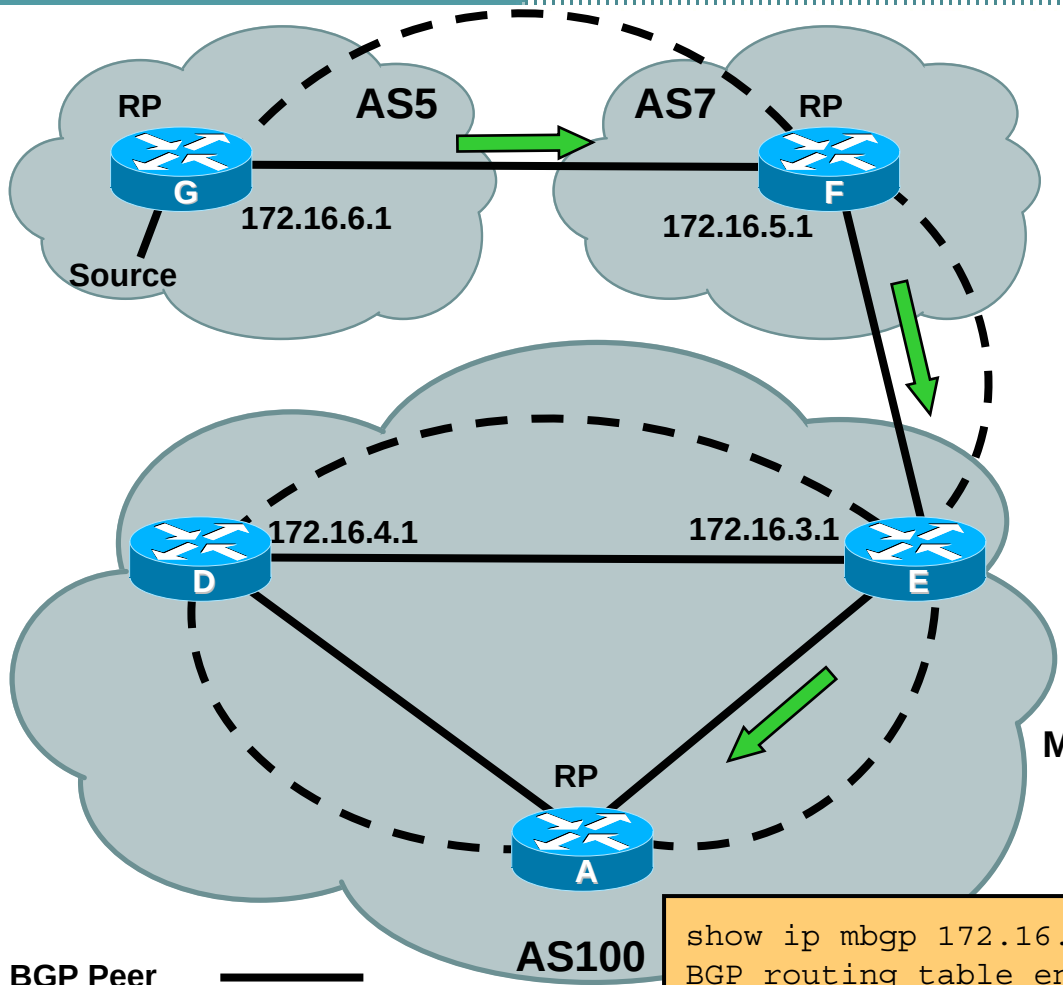
Receiving SA Messages

- RPF Check Rules depend on peering
  - Rule 1: Sending MSDP peer = i(m)BGP peer
  - Rule 2: Sending MSDP peer = e(m)BGP peer
  - Rule 3: Sending MSDP peer != (m)BGP peer
- Exceptions:
  - RPF check is skipped when:
    - Sending MSDP peer = Originating RP
    - Sending MSDP peer = Mesh-Group peer
    - Sending MSDP peer = only MSDP peer
      - (i.e. the ‘default-peer’ or the only ‘msdp-peer’ configured.)

RPF Check Rule 1

- When MSDP peer = i(m)BGP peer
  - Find “Best Path” to RP in BGP Tables
    - Search MRIB first then URIB
    - If no path to Originating RP found, RPF Fails
  - Note “BGP peer” that advertised path
    - (i.e. IP Address of BGP peer that sent us this path)
    - Warning:
      - This is not the same as the Next-hop of the path!!!
      - i(m)BGP peers normally do not set Next-hop = Self.
      - This is also not necessarily the same as the Router-ID!
  - Rule 1 Test Condition:
    - MSDP Peer address = BGP peer address?
      - If Yes, RPF Succeeds

Rule1: MSDP peer = i(m)BGP peer



i(m)BGP peer address = 172.16.3.1
(advertising best-path to RP)

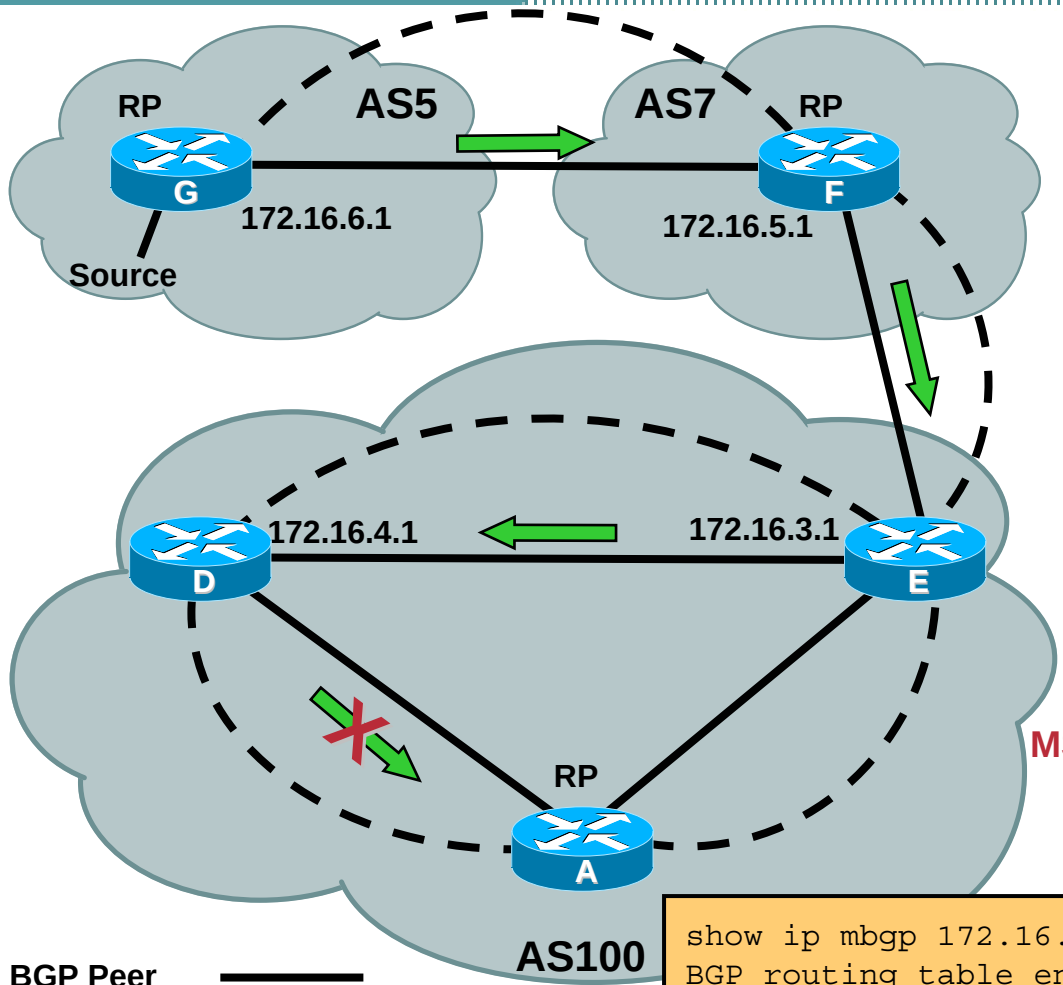
MSDP Peer address = 172.16.3.1

MSDP Peer address = i(m)BGP Peer address

SA RPF Check Succeeds

```
show ip mbgp 172.16.6.1
BGP routing table entry for 172.16.6.0/24, version 8745118
Paths: (1 available, best #1)
7 5, (received & used)
    172.16.5.1 (metric 68096) from 172.16.3.1 (172.16.3.1)
```

Rule1: MSDP peer = i(m)BGP peer



i(m)BGP Peer address = 172.16.3.1
(advertising best-path to RP)

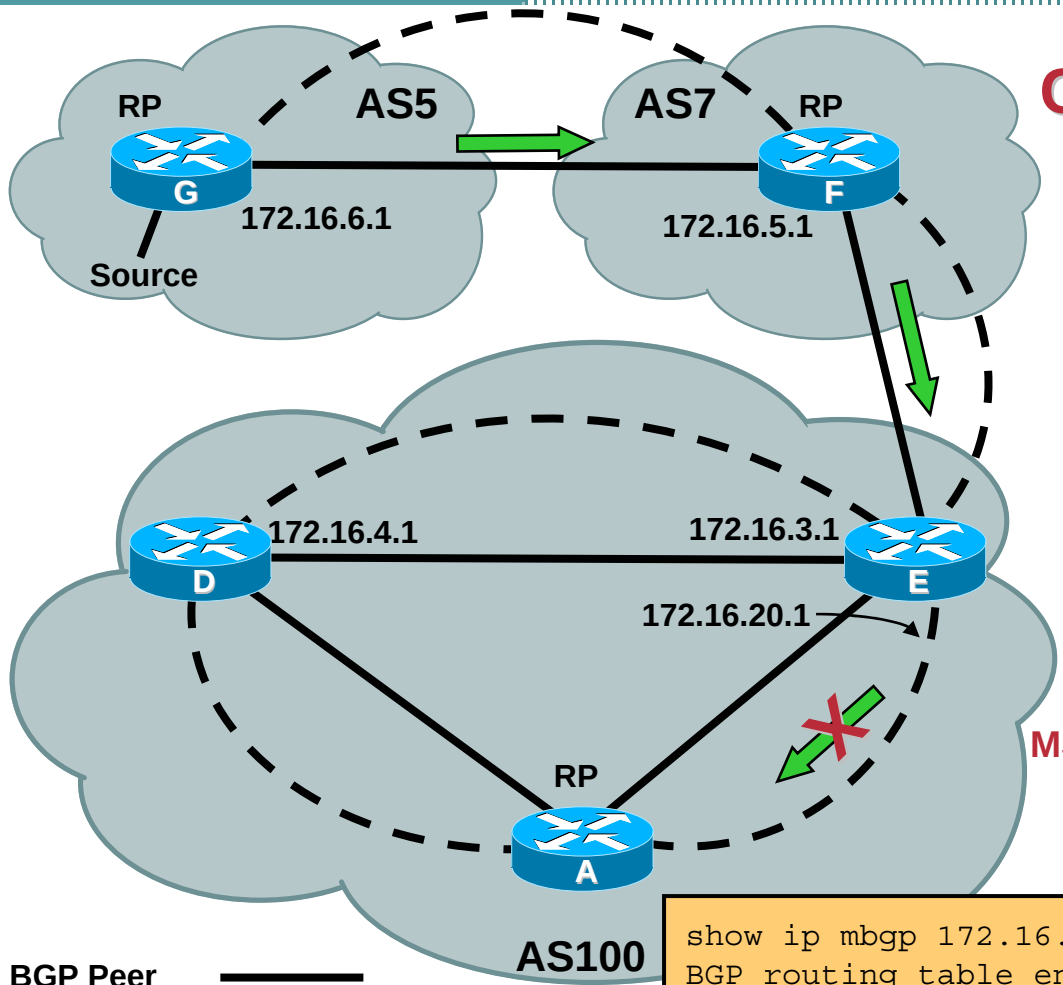
MSDP Peer address = 172.16.4.1

MSDP Peer address != i(m)BGP Peer address

SA RPF Check Fails

```
show ip mbgp 172.16.6.1
BGP routing table entry for 172.16.6.0/24, version 8745118
Paths: (1 available, best #1)
7 5, (received & used)
    172.16.5.1 (metric 68096) from 172.16.3.1 (172.16.3.1)
```

Rule1: MSDP peer = i(m)BGP peer



Common Mistake #1:

Failure to use same addresses for MSDP peers as i(m)BGP peers!

i(m)BGP Peer address = 172.16.3.1
(advertising best-path to RP)

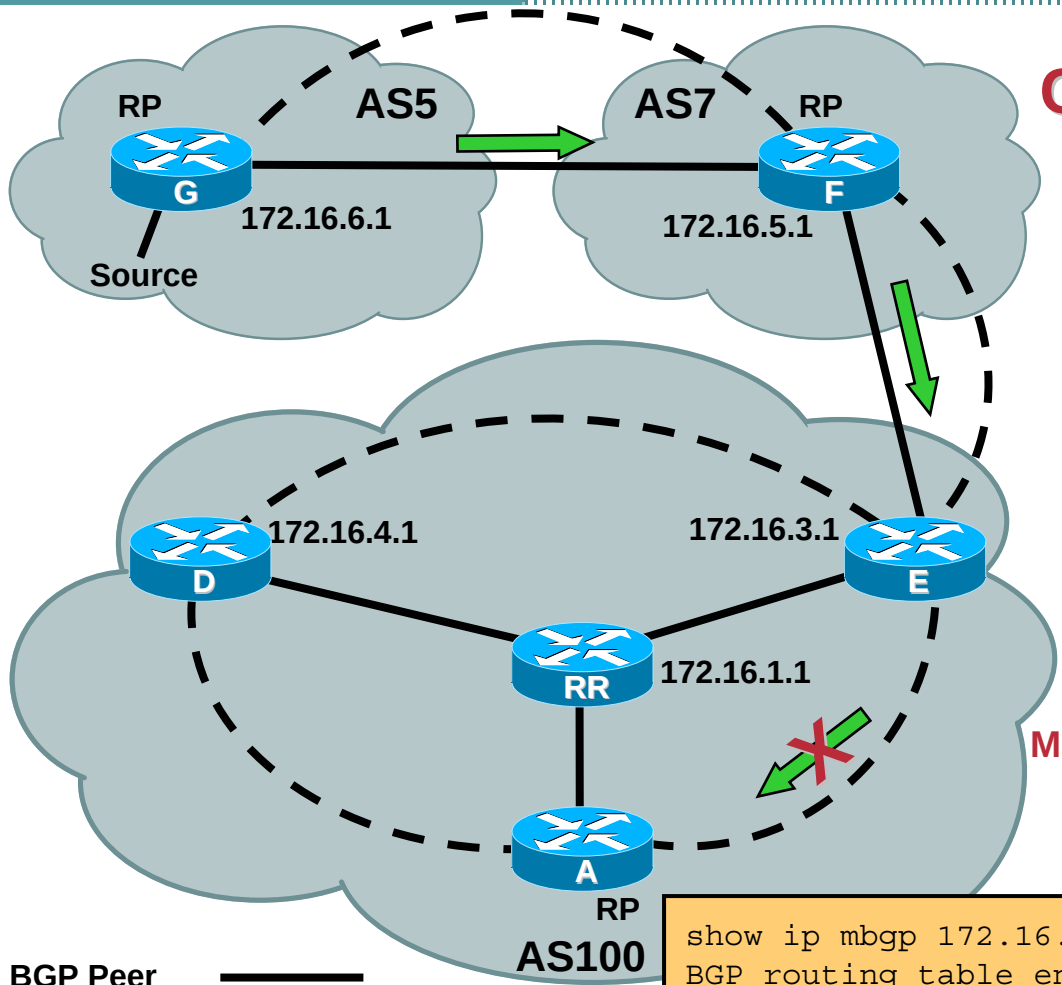
MSDP Peer address = 172.16.20.1

MSDP Peer address != i(m)BGP Peer address

SA RPF Check Fails

```
show ip mbgp 172.16.6.1
BGP routing table entry for 172.16.6.0/24, version 8745118
Paths: (1 available, best #1)
7 5, (received & used)
    172.16.5.1 (metric 68096) from 172.16.3.1 (172.16.3.1)
```

Rule1: MSDP peer = i(m)BGP peer



Common Mistake #2:

Failure to follow i(m)BGP topology!
Can happen when RR's are used.

i(m)BGP Peer address = 172.16.1.1
(advertising best-path to RP)

MSDP Peer address = 172.16.3.1

MSDP Peer address != i(m)BGP Peer address

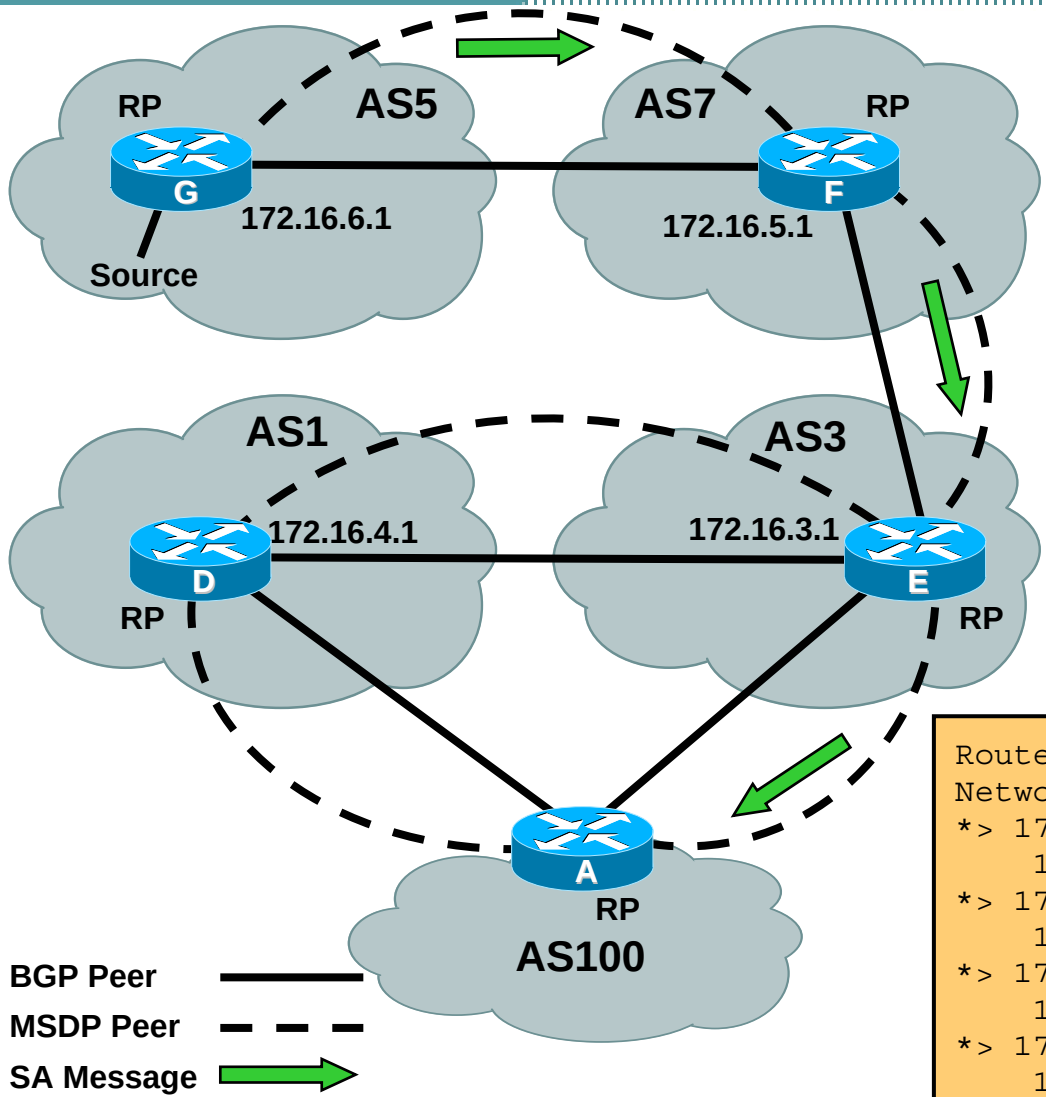
SA RPF Check Fails

```
show ip mbgp 172.16.6.1
BGP routing table entry for 172.16.6.0/24, version 8745118
Paths: (1 available, best #1)
7 5, (received & used)
    172.16.5.1 (metric 68096) from 172.16.1.1 (172.16.1.1)
```

RPF Check Rule 2

- When MSDP peer = e(m)BGP peer
  - Find (m)BGP “Best Path” to RP
    - Search MRIB first then URIB
      - If no path to Originating RP found, RPF Fails
  - Rule 2 Test Condition:
    - First AS in path to the RP = MSDP peer?
      - If Yes, RPF Succeeds

Rule2: MSDP peer = e(m)BGP peer



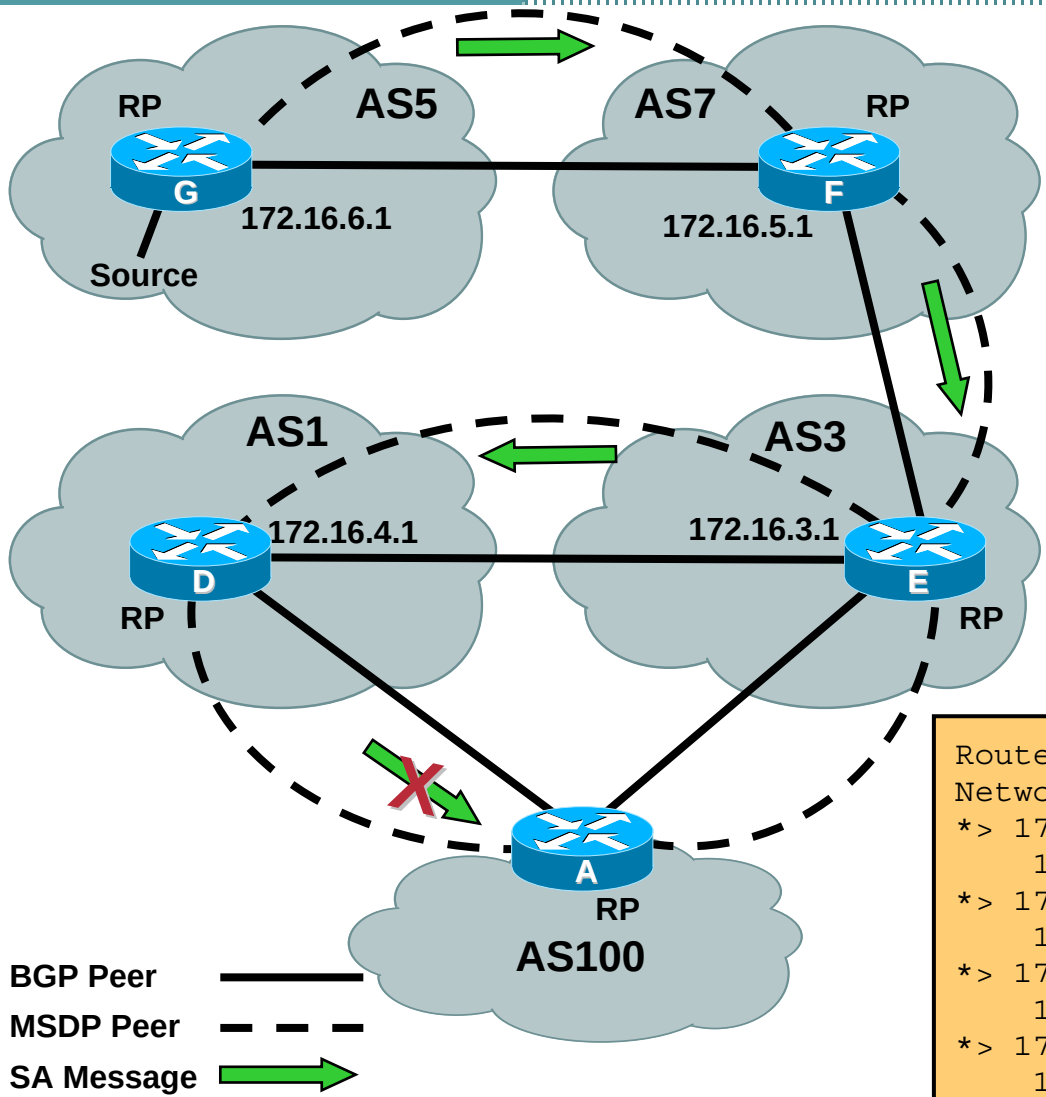
First-AS in best-path to RP = 3
AS of MSDP Peer = 3

First-AS in best-path to RP = AS of e(m)BGP Peer

SA RPF Check Succeeds

```
Router A's BGP Table
Network              Next Hop        Path
*> 172.16.3.0/24     172.16.3.1      3 i
    172.16.3.0/24     172.16.4.1     1 3 i
*> 172.16.4.0/24     172.16.4.1      1 i
    172.16.4.0/24     172.16.3.1     3 1 i
*> 172.16.5.0/24     172.16.4.1      3 7 i
    172.16.5.0/24     172.16.3.1     1 3 7 i
*> 172.16.6.0/24     172.16.3.1      3 7 5 i
    172.16.6.0/24     172.16.4.1     1 3 7 5 i
```

Rule2: MSDP peer = e(m)BGP peer



First-AS in best-path to RP = 3
AS of MSDP Peer = 1

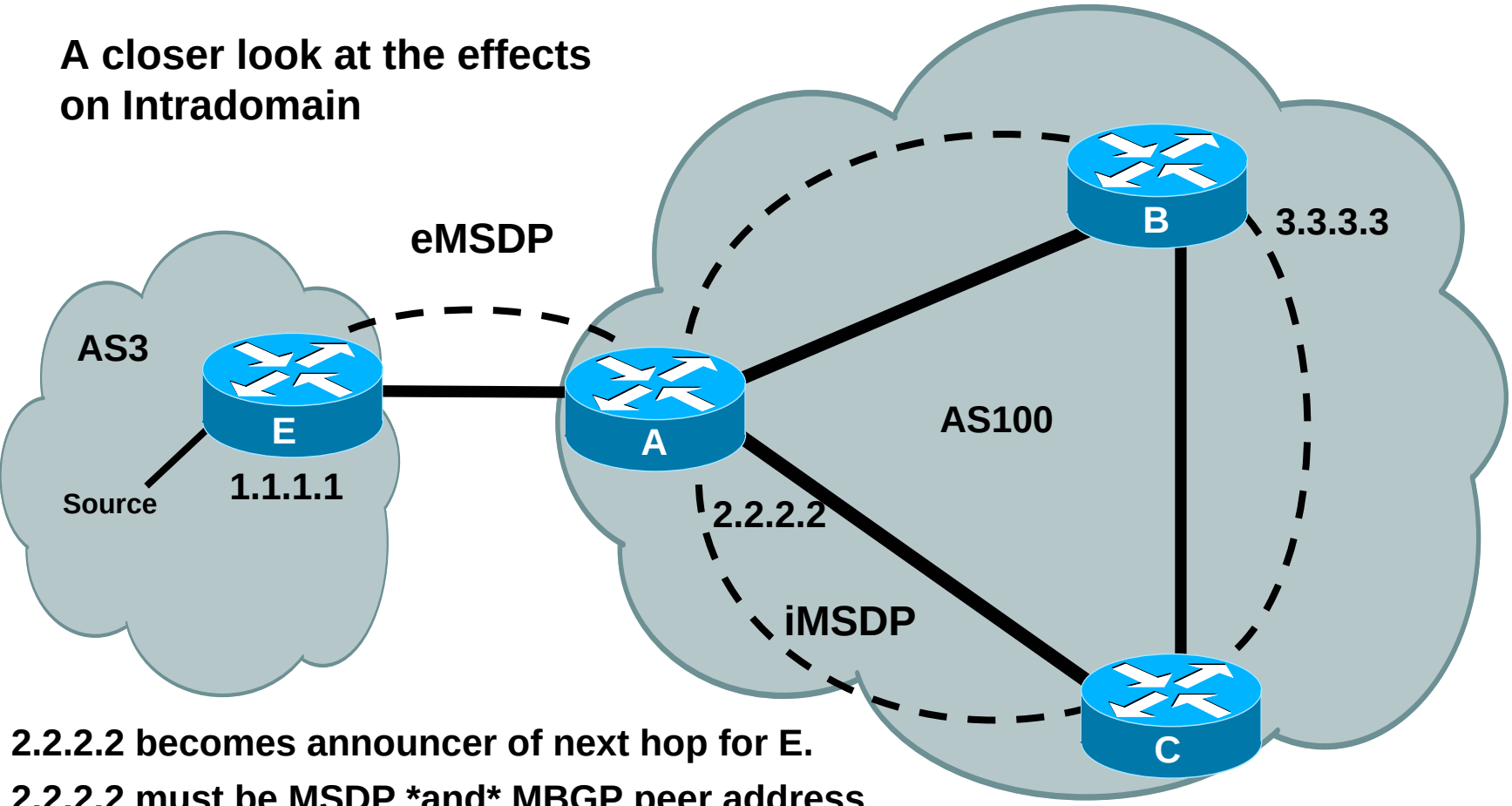
First-AS in best-path to RP != AS of e(m)BGP Peer

SA RPF Check Fails!

```
Router A's BGP Table
Network              Next Hop       Path
*> 172.16.3.0/24     172.16.3.1     3 i
    172.16.3.0/24     172.16.4.1     1 3 i
*> 172.16.4.0/24     172.16.4.1     [1] i
    172.16.4.0/24     172.16.3.1     3 1 i
*> 172.16.5.0/24     172.16.3.1     3 7 i
    172.16.5.0/24     172.16.4.1     1 3 7 i
*> 172.16.6.0/24     172.16.3.1     [3] 7 5 i
    172.16.6.0/24     172.16.4.1     1 3 7 5 i
```

Rule2: MSDP peer = e(m)BGP peer

A closer look at the effects on Intradomain



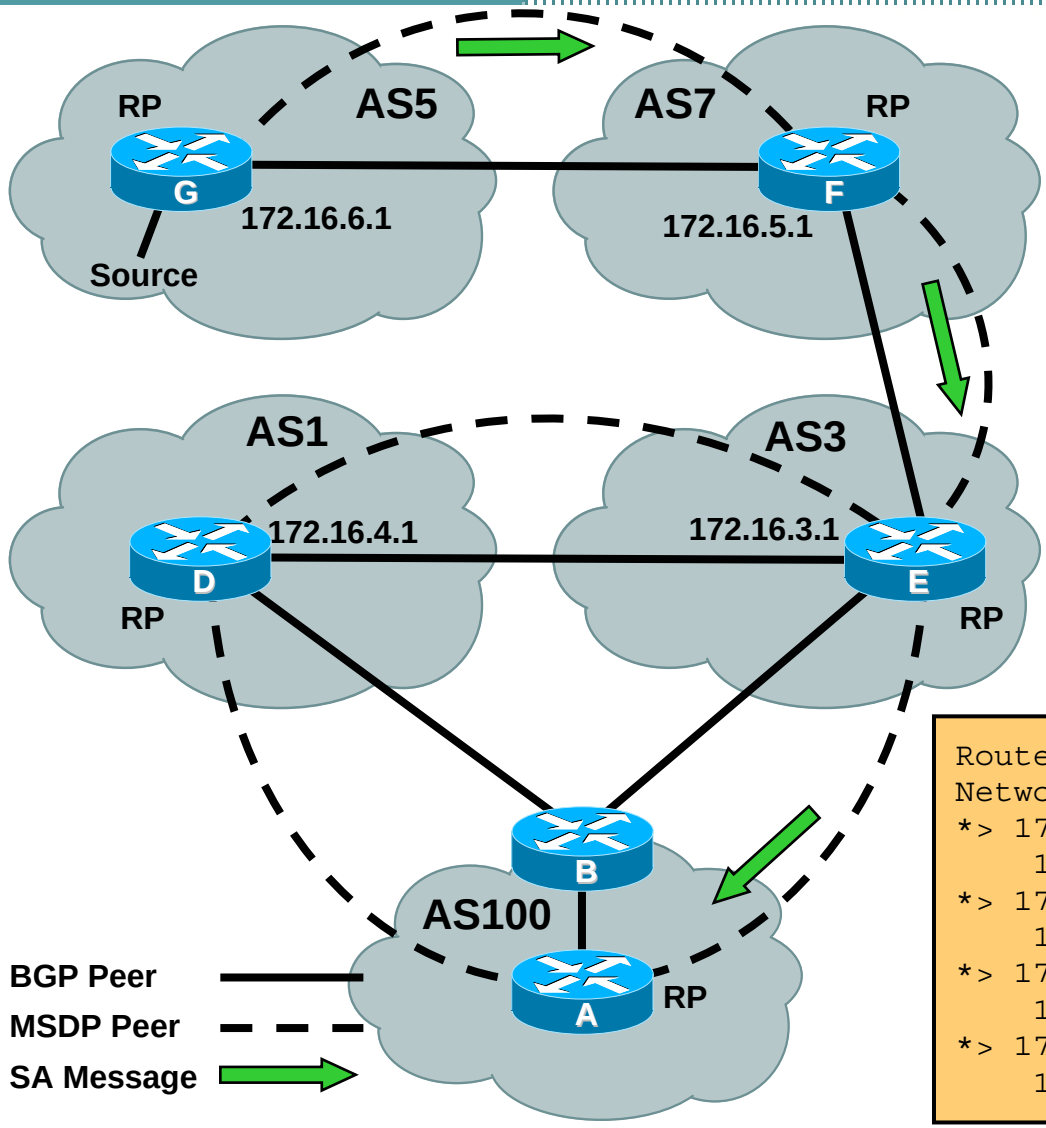
2.2.2.2 becomes announcer of next hop for E.

2.2.2.2 must be MSDP *and* MBGP peer address for B and C. Providers may want more flexibility by eMSDP peering further into their network.

RPF Check Rule 3

- When MSDP peer != (m)BGP peer
  - Find (m)BGP “Best Path” to RP
    - Search MRIB first then URIB
      - If no path to Originating RP found, RPF Fails
  - Find (m)BGP “Best Path” to MSDP peer
    - Search MRIB first then URIB
      - If no path to sending MSDP Peer found, RPF Fails
  - Note AS of sending MSDP Peer
    - Origin AS (last AS) in AS-PATH to MSDP Peer
  - Rule 3 Test Condition:
    - First AS in path to RP = Sending MSDP Peer AS ?
      - If Yes, RPF Succeeds

Rule3: MSDP peer != BGP peer



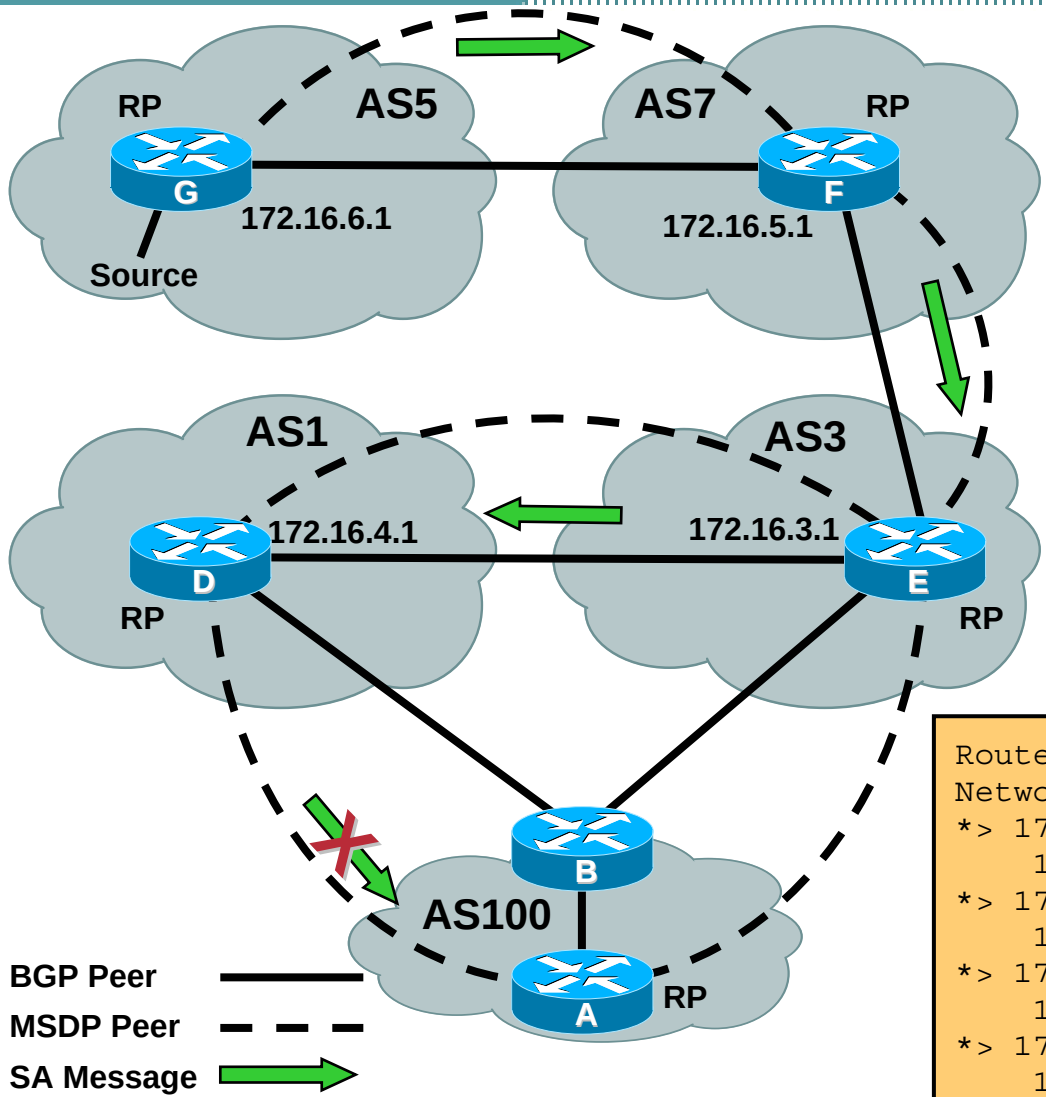
First-AS in best-path to RP = 3
AS of MSDP Peer = 3

First-AS in best-path to RP = AS of MSDP Peer

SA RPF Check Succeeds

```
Router A's BGP Table
Network              Next Hop        Path
*> 172.16.3.0/24     172.16.3.1      3 i
   172.16.3.0/24     172.16.4.1      1 3 i
*> 172.16.4.0/24     172.16.4.1      1 i
   172.16.4.0/24     172.16.3.1      3 1 i
*> 172.16.5.0/24     172.16.4.1      3 7 i
   172.16.5.0/24     172.16.3.1      1 3 7 i
*> 172.16.6.0/24     172.16.3.1      3 7 5 i
   172.16.6.0/24     172.16.4.1      1 3 7 5 i
```

Rule3: MSDP peer != BGP peer



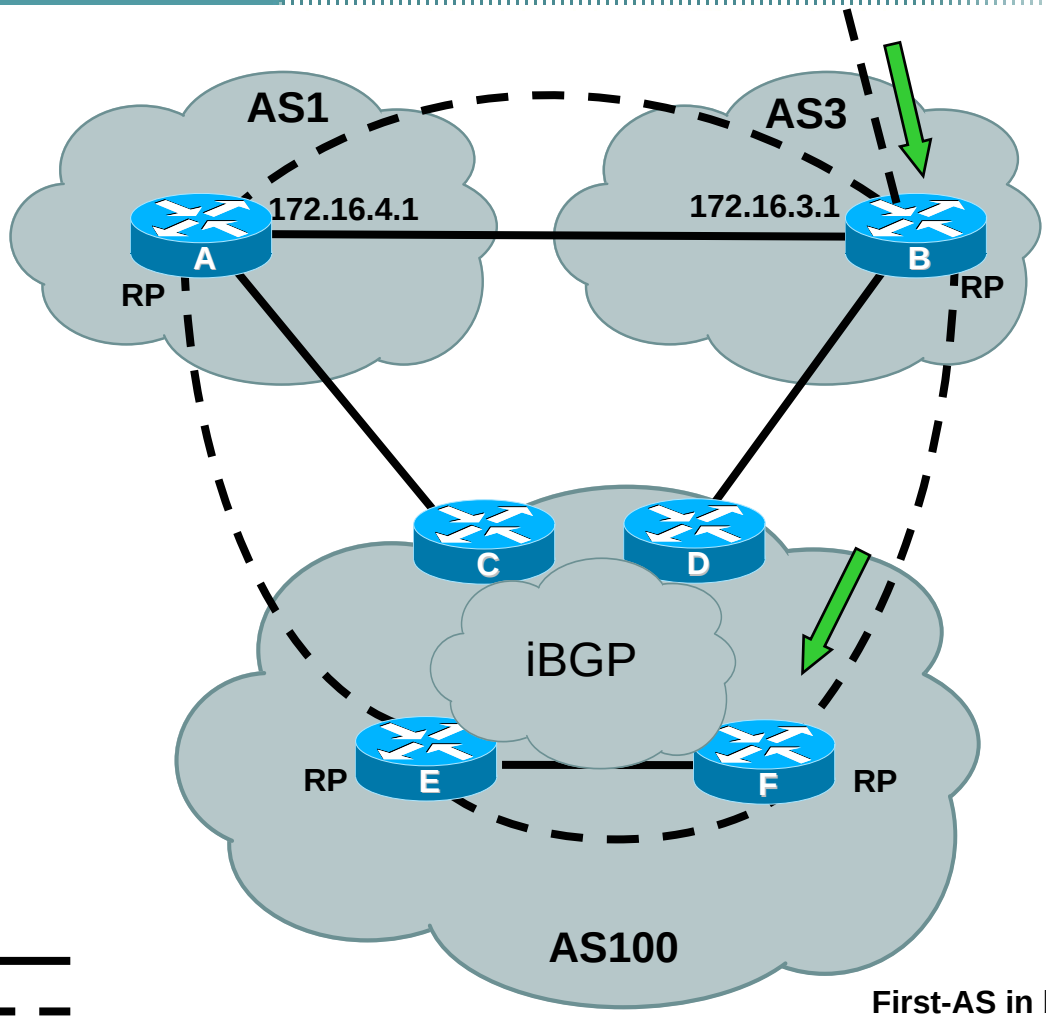
First-AS in best-path to RP = 3
AS of MSDP Peer = 1

First-AS in best-path to RP != AS of MSDP Peer

SA RPF Check Fails

```
Router A's BGP Table
Network              Next Hop       Path
*> 172.16.3.0/24     172.16.3.1     3 i
    172.16.3.0/24     172.16.4.1     1 3 i
*> 172.16.4.0/24     172.16.4.1     1 i
    172.16.4.0/24     172.16.3.1     3 1 i
*> 172.16.5.0/24     172.16.4.1     3 7 i
    172.16.5.0/24     172.16.3.1     1 3 7 i
*> 172.16.6.0/24     172.16.3.1     3 7 5 i
    172.16.6.0/24     172.16.4.1     1 3 7 5 i
```

Rule3: MSDP peer != BGP peer



More flexibility with MSDP peer placement

BGP Peer

MSDP Peer

SA Message

First-AS in best-path to RP = 3

AS of MSDP Peer = 3

First-AS in best-path to RP = AS of MSDP Peer

SA RPF Check Succeeds

MSDP Configuration

- **Configure peers**

```
ip msdp peer <ip-address> [connect-source <i/f>]
```

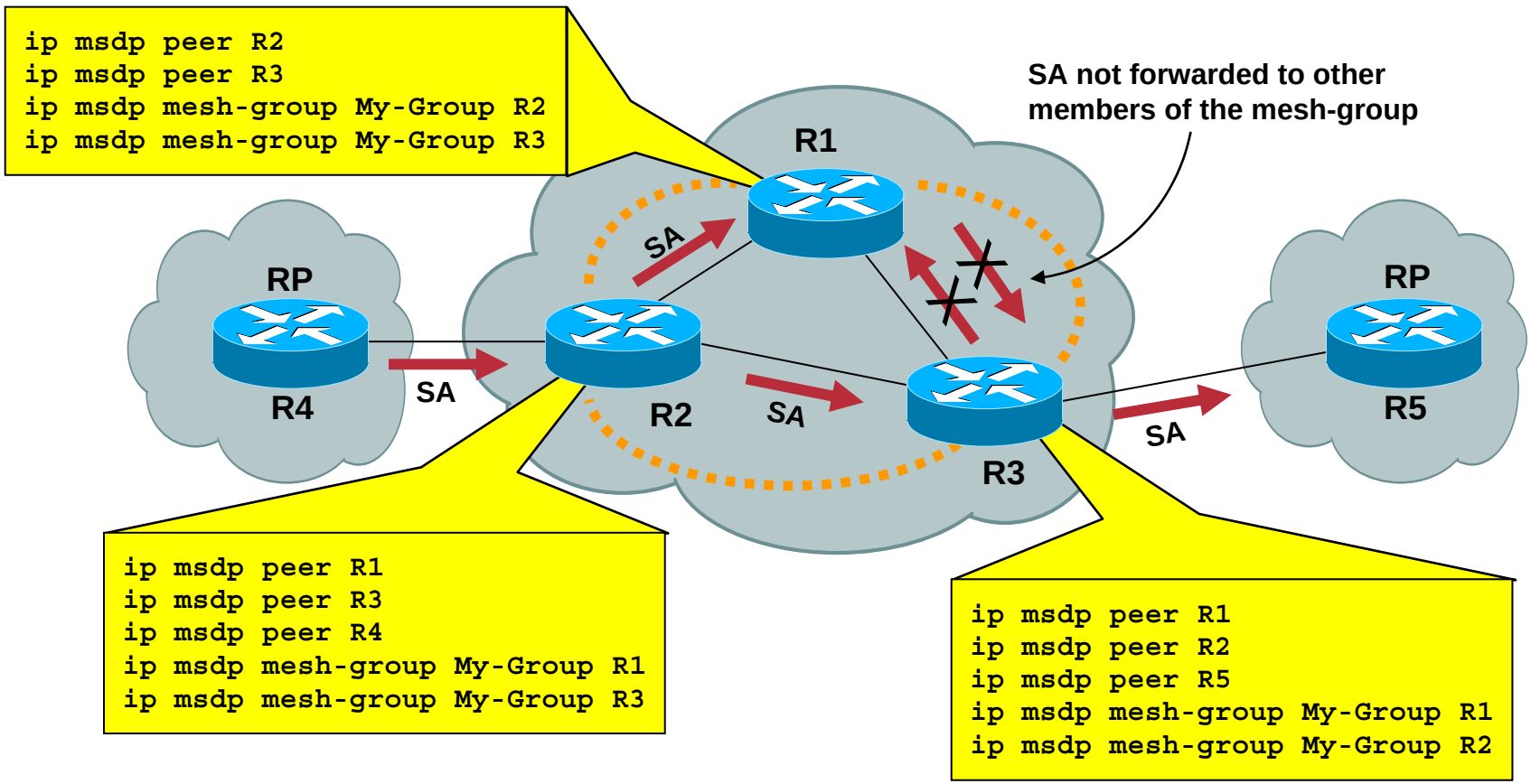
- **Configure default peer**

```
ip msdp default-peer <ip-address> [prefix-list acl]
```

- **Mesh groups**

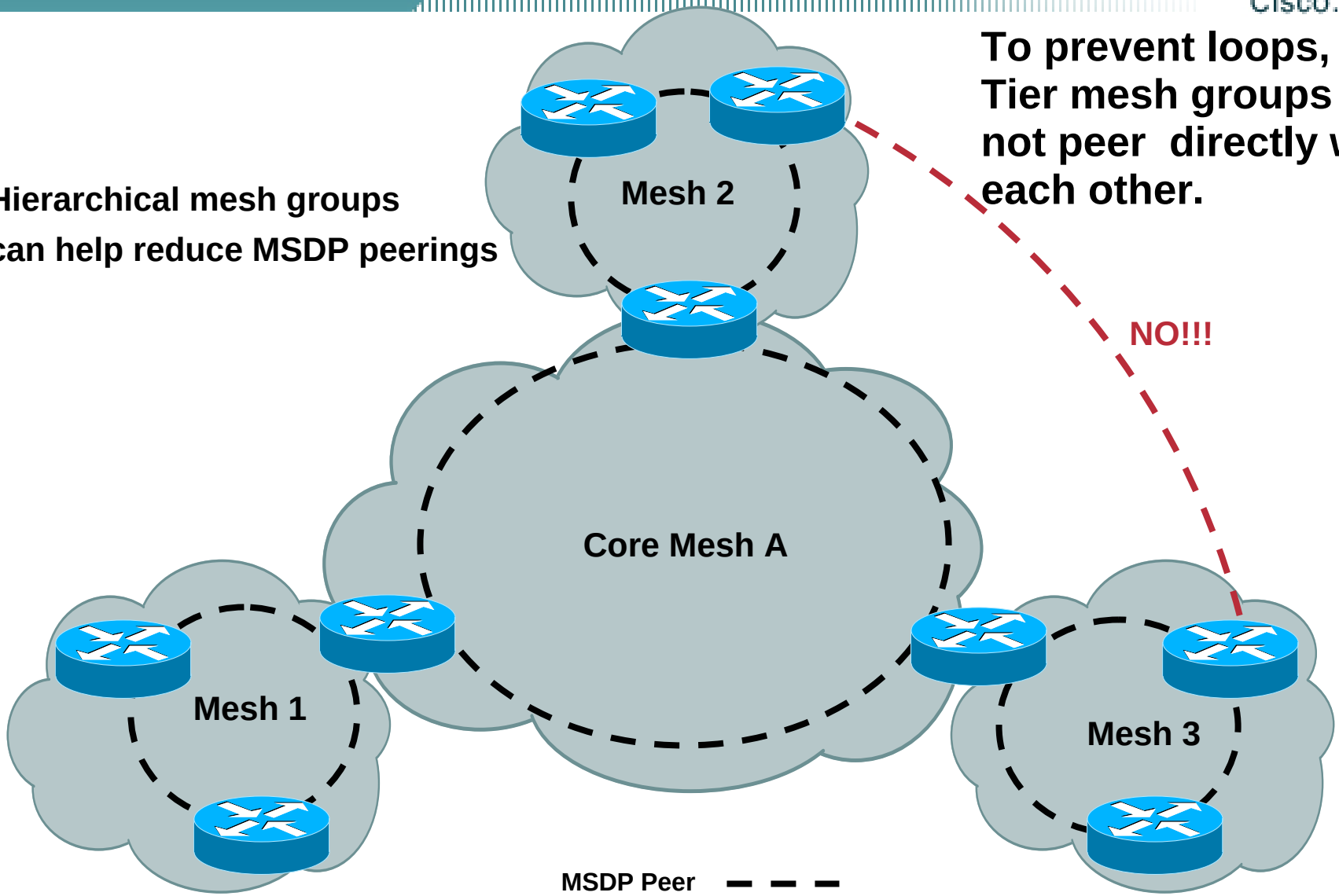
```
ip msdp mesh-group <name> <ip-address>
```

MSDP Mesh-Group Example



Intradomain MSDP peering scenarios

Hierarchical mesh groups can help reduce MSDP peerings



To prevent loops, 2nd Tier mesh groups must not peer directly with each other.

MSDP Configuration

- **Latest draft:**

  draft-ietf-msdp-spec-16.txt

- **Filtering**
  - **Can filter SA in/out, groups, with acls or route-maps**
- **For configuration commands see:**
  - **ftp://ftpeng.cisco.com/ipmulticast/Multicast-Commands**
- **For MSDP BCP (Best Current Practice) Draft:**
  - **draft-ietf-mboned-msdp-deploy-02.txt**

MSDP Enhancements

- **New IOS command**

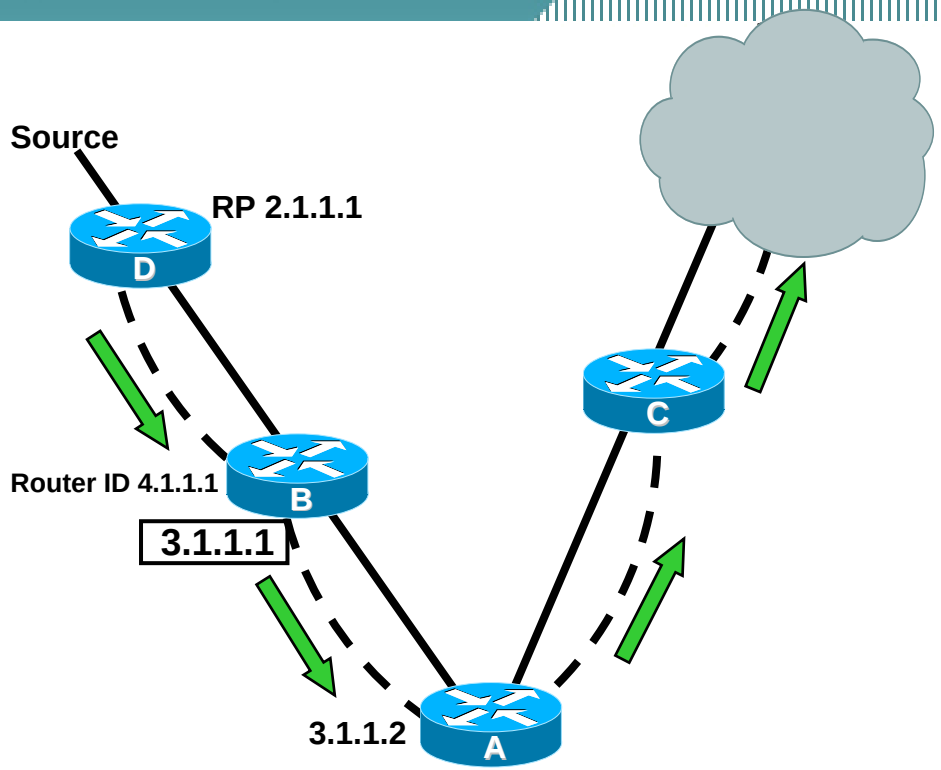
```
ip msdp new-rpf-rules
```

  - **MSDP SA RPF check using IGP**
  - **Accept SA's from BGP NEXT HOP**
  - **Accept SA's from closest peer along the best path to the originating RP**
  - **"show ip msdp rpf"**

MSDP RPF check using IGP

- **When MSDP peer = IGP peer (No BGP)**
  - **Find best IGP route to RP**
    - **Search URIB**
      - **If route to Originating RP found and:**
        - **If IGP next hop (or advertiser) address for RP is the MSDP peer and in UP state, then that is the RPF peer.**
        - **If route not found: Fall through to the next rule.**

IGP Rule: MSDP peer = IGP peer (Next hop)



MSDP Peer = 3.1.1.1

IGP next hop to originating RP = 3.1.1.1

IGP next hop to originating RP = MSDP peer

SA RPF Check Succeeds

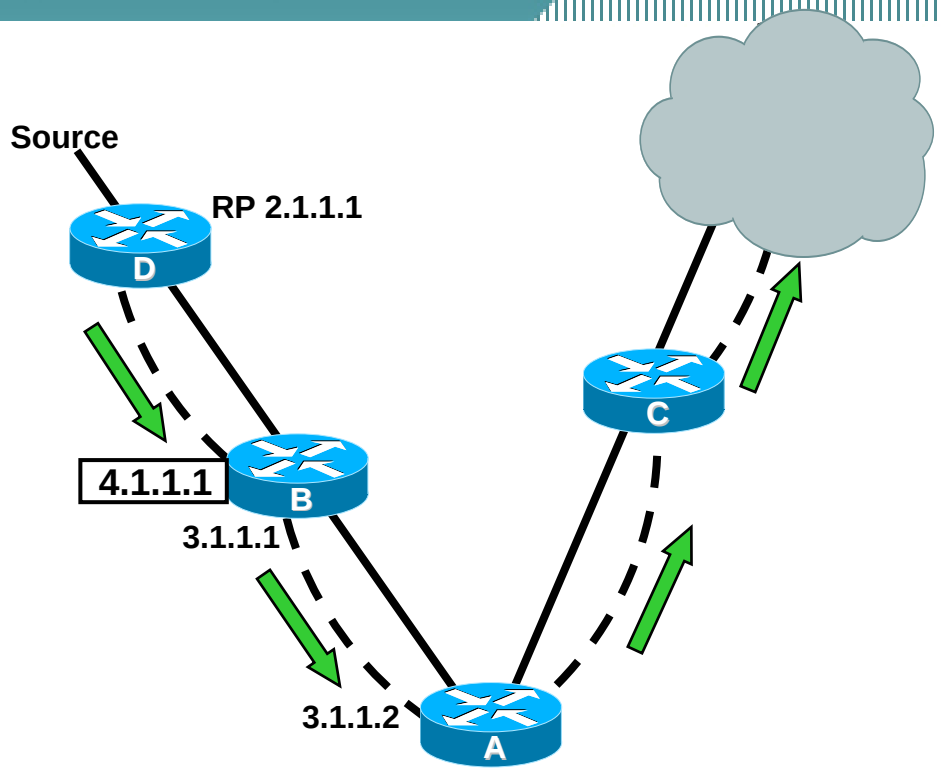
```
RouterA#show ip route 2.1.1.1
Routing entry for 2.1.1.0/24
  Known via "ospf 1", distance 110, metric 20, type intra area
  Last update from 3.1.1.1 on Ethernet2, 00:35:10 ago
  Routing Descriptor Blocks:
  * 3.1.1.1, from 4.1.1.1, 00:35:10 ago, via Ethernet2
      Route metric is 20, traffic share count is 1
```

OSPF neighbor

MSDP Peer

SA Message

IGP Rule: MSDP peer = IGP peer (Advertiser)



MSDP Peer = 4.1.1.1

IGP next hop to originating RP = ~~3.1.1.1~~

IGP advertiser to originating RP = 4.1.1.1

IGP advertiser to originating RP = MSDP peer

SA RPF Check Succeeds

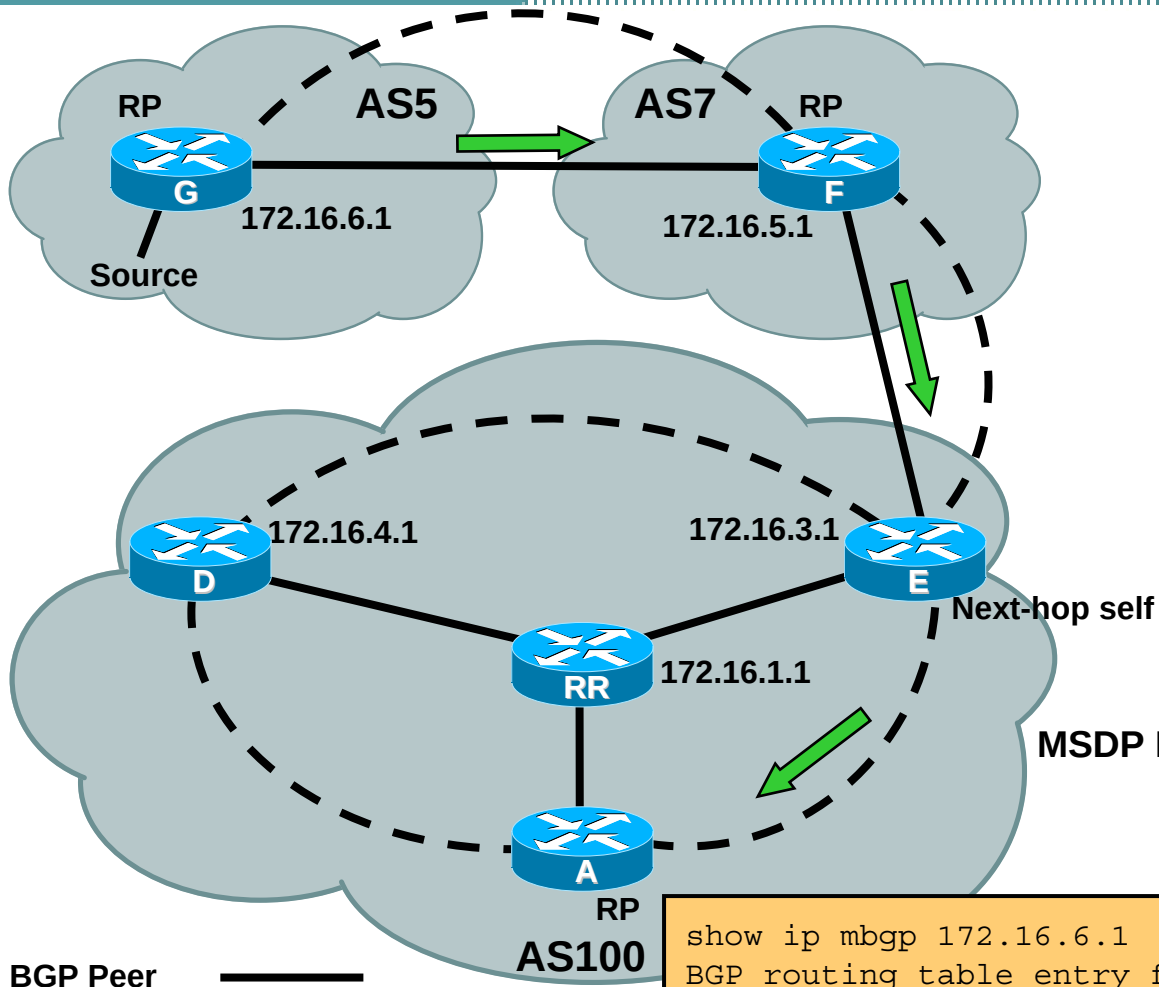
OSPF neighbor

MSDP Peer

SA Message

```
RouterA#show ip route 2.1.1.1
Routing entry for 2.1.1.0/24
  Known via "ospf 1", distance 110, metric 20, type intra area
  Last update from 3.1.1.1 on Ethernet2, 00:35:10 ago
  Routing Descriptor Blocks:
  * 3.1.1.1, from 4.1.1.1, 00:35:10 ago, via Ethernet2
      Route metric is 20, traffic share count is 1
```

SA's accepted from Next Hop



i(m)BGP Peer address = 172.16.1.1
(Advertiser of next hop)

MSDP Peer address = 172.16.3.1

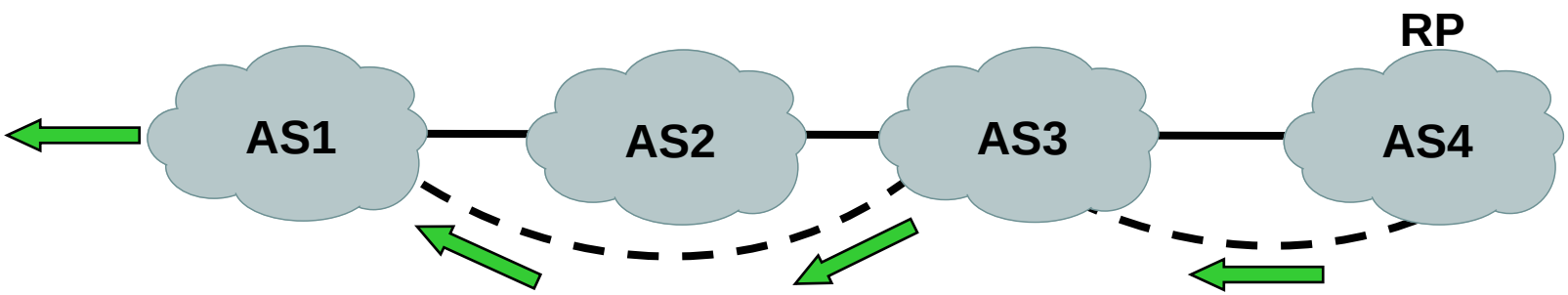
But, BGP next hop = 172.16.3.1

MSDP Peer address = BGP next hop address

SA RPF Check Succeeds

```
show ip mbgp 172.16.6.1
BGP routing table entry for 172.16.6.0/24, version 8745118
Paths: (1 available, best #1)
7 5, (received & used)
    172.16.3.1 (metric 68096) from 172.16.1.1 (172.16.1.1)
```

Accept SA along RPF path



Existing Rule: If first AS in best path to the RP != MSDP peer

RPF Fails

New code: Choose peer in CLOSEST AS along best AS path to the RP.

Loosens rule a bit.

RPF Succeeds.

BGP Peer

MSDP Peer

SA Message

New MSDP RPF command


```
Router-A# show ip msdp rpf 2.1.1.1
RPF peer information for Router-B (2.1.1.1)
  RPF peer: Router-C (3.1.1.1)
  RPF route/mask: 2.1.1.0/24
  RPF rule: Peer is IGP next hop of best route
  RPF type: unicast (ospf 1)
```

Agenda

- MBGP (routing)
- MSDP (source discovery)
- **MBGP/MSDP Examples**
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

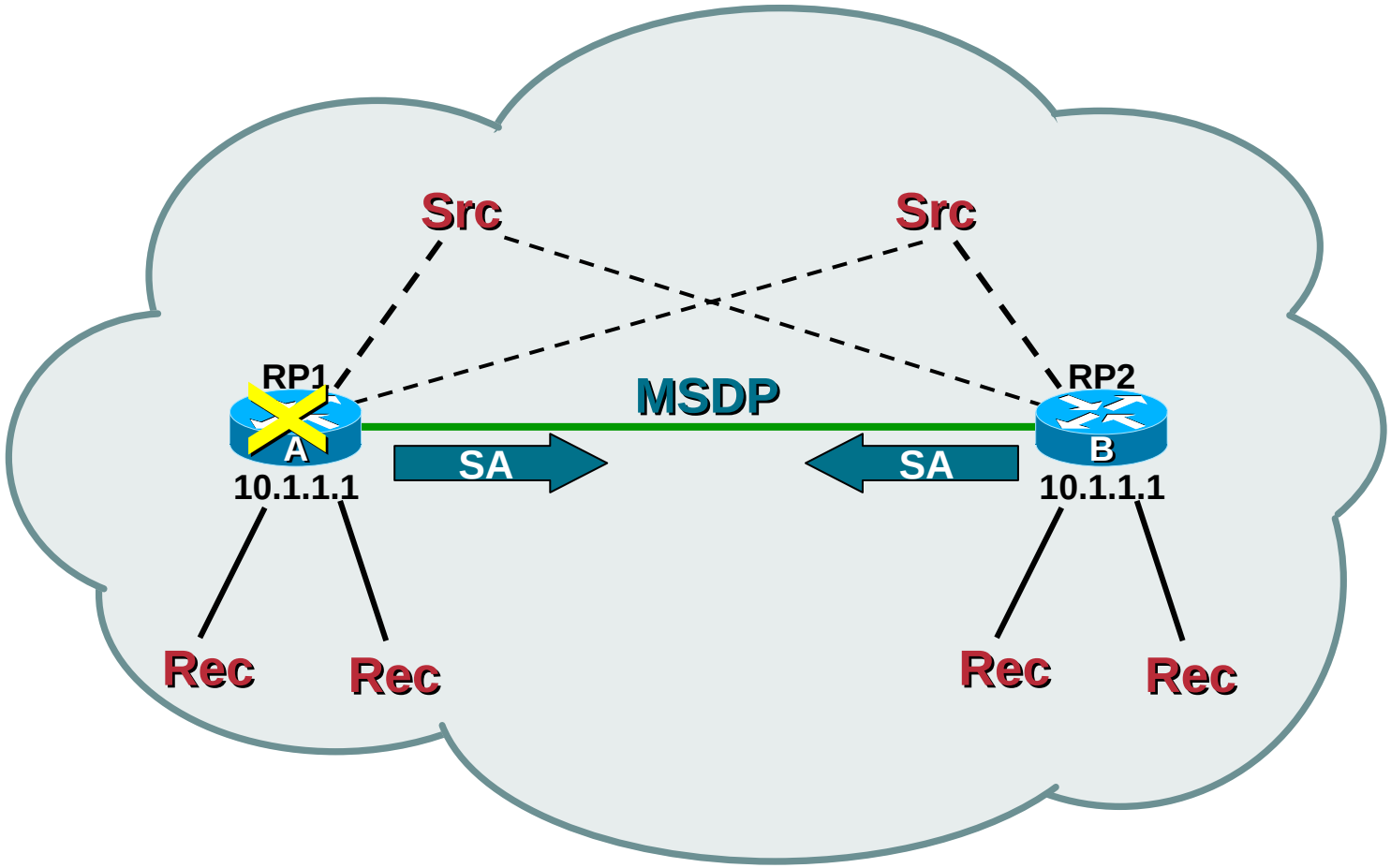
MSDP Application—Anycast RP

- RFC 3446 Anycast RP mechanism using PIM and MSDP
- Within a domain, deploy more than one RP for the same group range
- Give each RP the same IP address assignment
- Sources and receivers use closest RP
- Used intra-domain to provide redundancy and RP load sharing

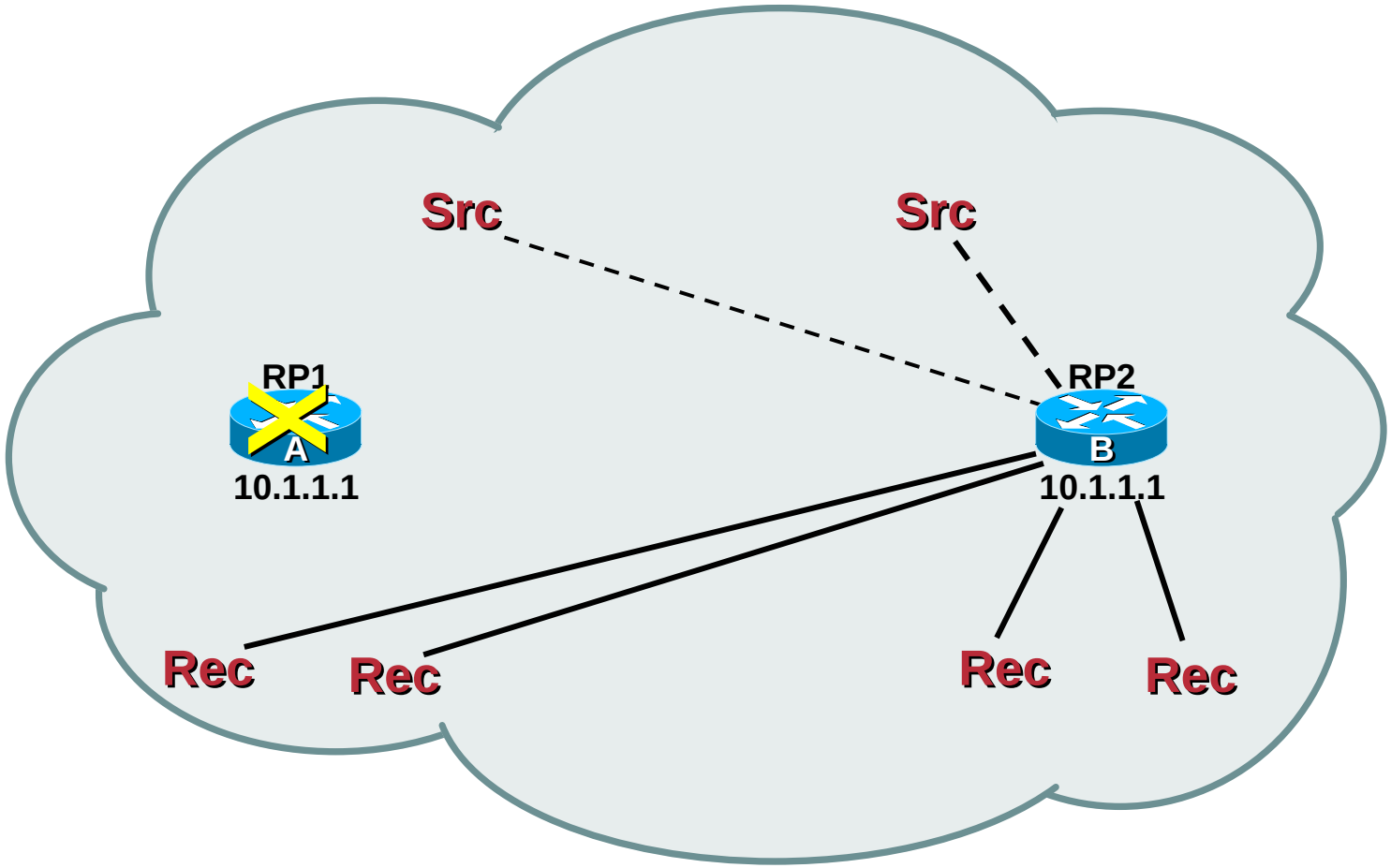
MSDP Application—Anycast RP

- Sources from one RP are made known to other RPs using MSDP
- When an RP goes down, sources and receivers are taken to new RP via unicast routing
  - Fast convergence

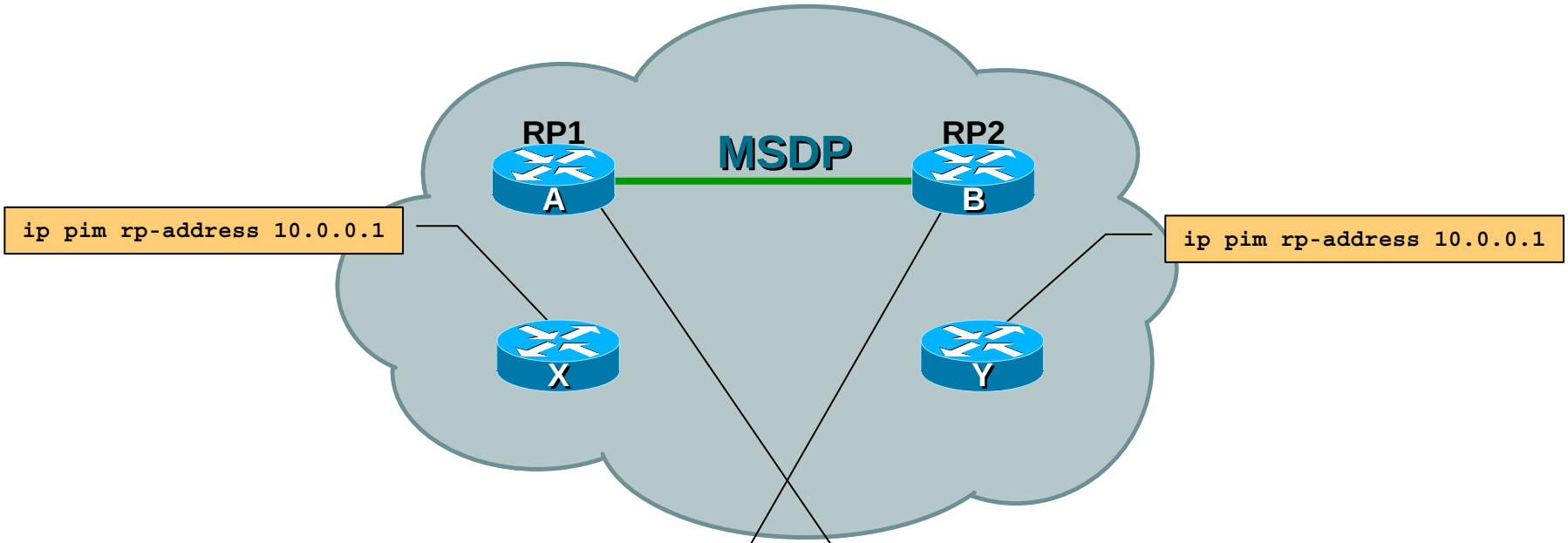
Anycast RP—Overview



Anycast RP—Overview



Anycast RP Configuration




```
Interface loopback 0
   ip address 10.0.0.2 255.255.255.255

Interface loopback 1
 ip address 10.0.0.1 255.255.255.255
!
ip msdp peer 10.0.0.3 connect-source loopback 0
ip msdp originator-id loopback 0
```

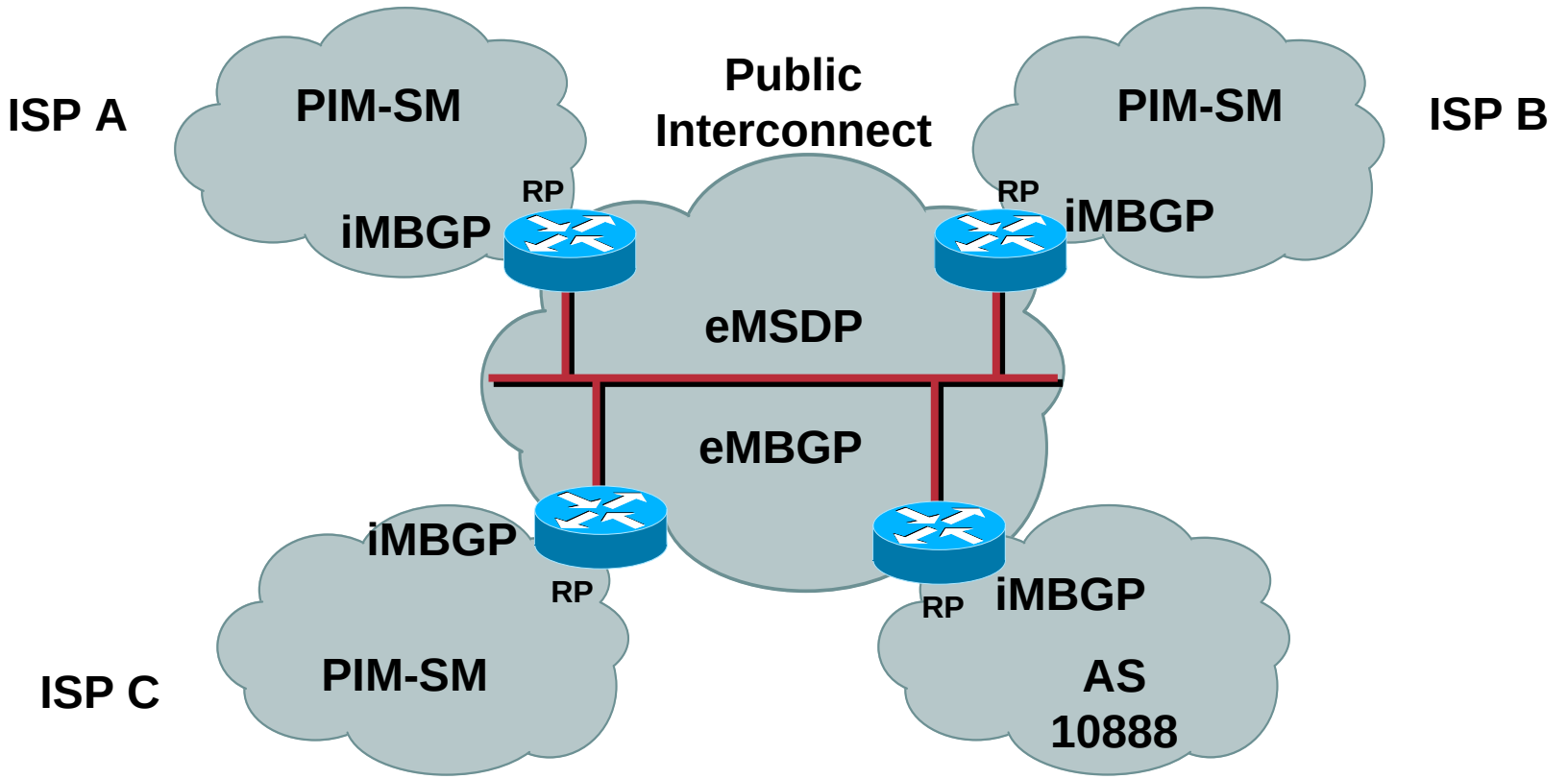
```
Interface loopback 0
 ip address 10.0.0.3 255.255.255.255

Interface loopback 1
 ip address 10.0.0.1 255.255.255.255
!
ip msdp peer 10.0.0.2 connect-source loopback 0
ip msdp originator-id loopback 0
```

ISP Requirements at the MIX

- Current solution: MBGP + PIM-SM + MSDP
  - Environment
    - ISPs run iMBGP and PIM-SM (internally)
    - ISPs multicast peer at a public interconnect
  - Deployment
    - Border routers run eMBGP
    - The interfaces on interconnect run PIM-SM
    - RPs' MSDP peering is fully meshed
    - All peers set a common distance for eMBGP

ISP Requirements at the MIX

Peering Solution: MBGP + PIM-SM +MSDP



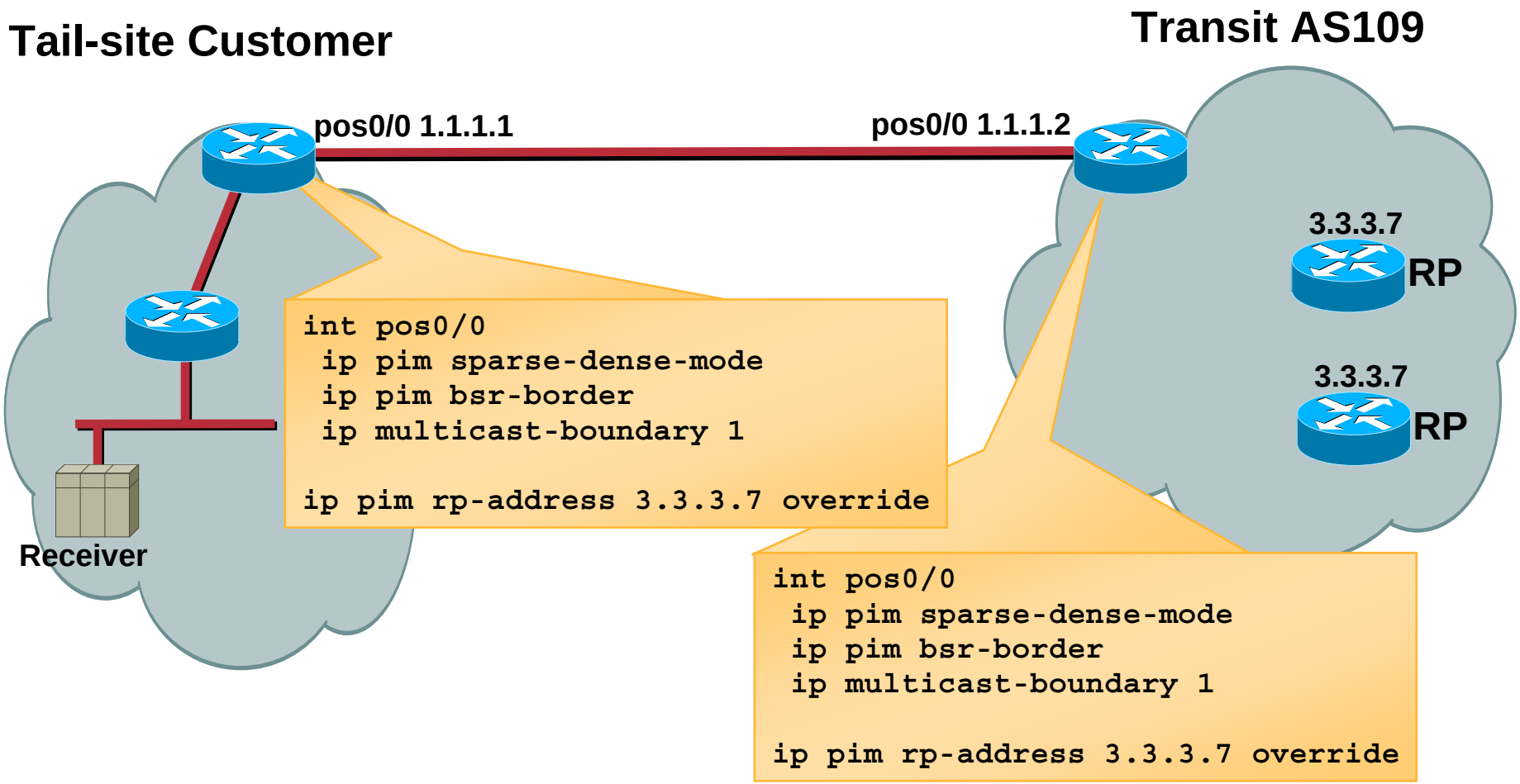
Recommended MSDP SA Filter

Cisco.com

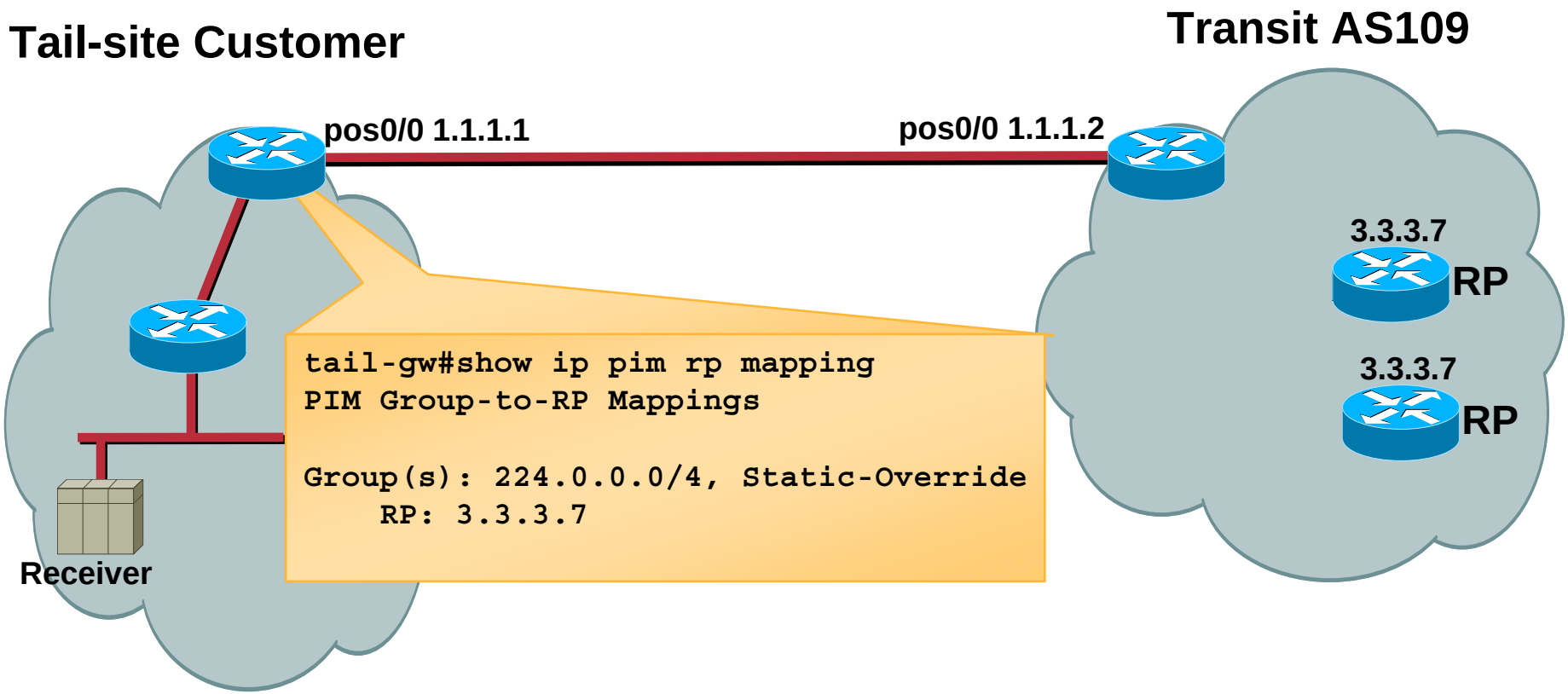
ftp://ftpeng.cisco.com/ipmulticast/config-notes/msdp-sa-filter.txt
http://www.ietf.org/internet-drafts/draft-nickless-ipv4-mcast-unusable-01.txt

```
! domain-local applications
access-list 111 deny   ip any host 224.0.2.2     !
access-list 111 deny   ip any host 224.0.1.3     ! Rwhod
access-list 111 deny   ip any host 224.0.1.24    ! Microsoft-ds
access-list 111 deny   ip any host 224.0.1.22    ! SVRLOC
access-list 111 deny   ip any host 224.0.1.2     ! SGI-Dogfight
access-list 111 deny   ip any host 224.0.1.35    ! SVRLOC-DA
access-list 111 deny   ip any host 224.0.1.60    ! hp-device-disc
!-- auto-rp groups
access-list 111 deny   ip any host 224.0.1.39
access-list 111 deny   ip any host 224.0.1.40
!-- scoped groups
access-list 111 deny   ip any 239.0.0.0 0.255.255.255
!-- loopback, private addresses (RFC 1918)
access-list 111 deny   ip 10.0.0.0 0.255.255.255 any
access-list 111 deny   ip 127.0.0.0 0.255.255.255 any
access-list 111 deny   ip 172.16.0.0 0.15.255.255 any
access-list 111 deny   ip 192.168.0.0 0.0.255.255 any
access-list 111 permit ip any any
!-- Default SSM-range. Do not do MSDP in this range
access-list 111 deny   ip any 232.0.0.0 0.255.255.255
access-list 111 permit ip any any
```

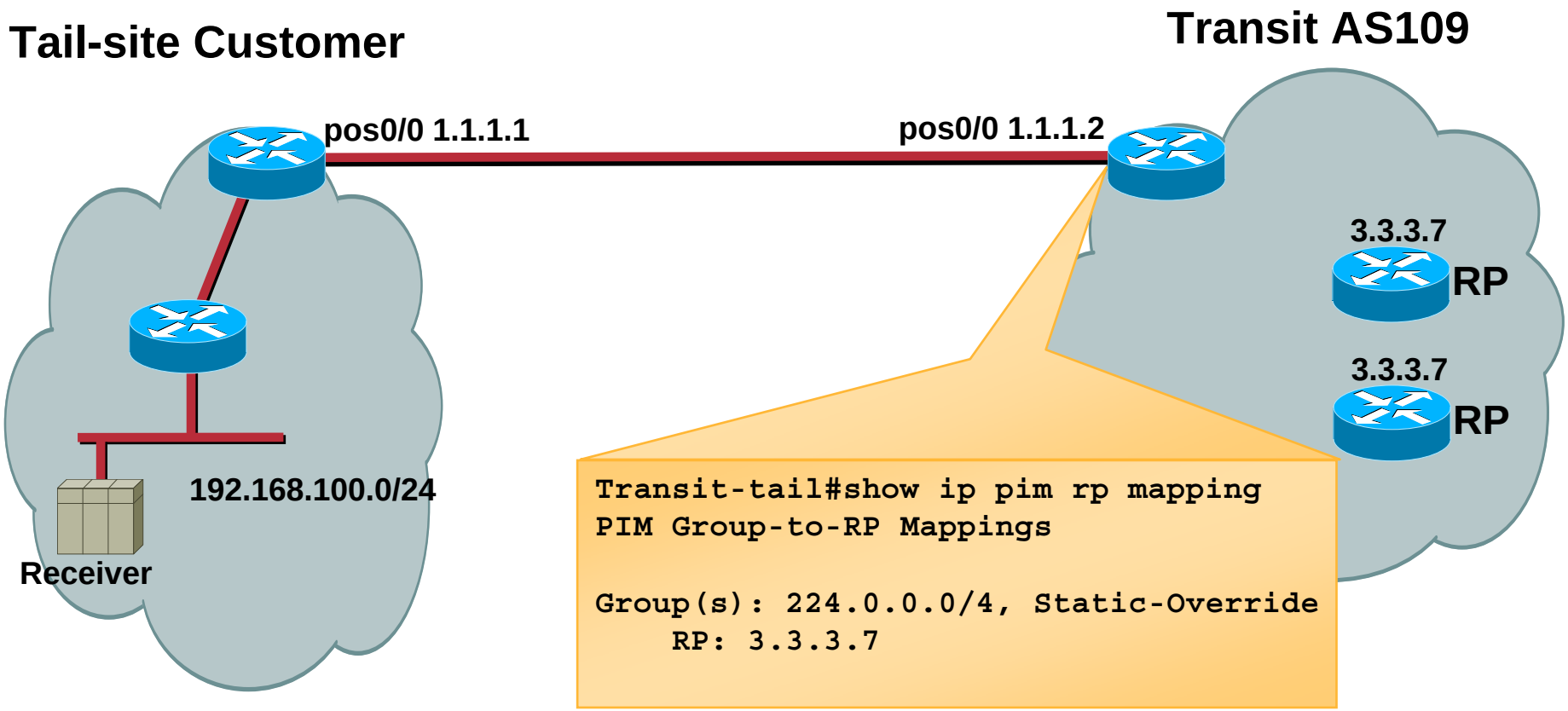
Single-Homed, ISP RP, Non-MBGP



Single-Homed, ISP RP, Non-MBGP

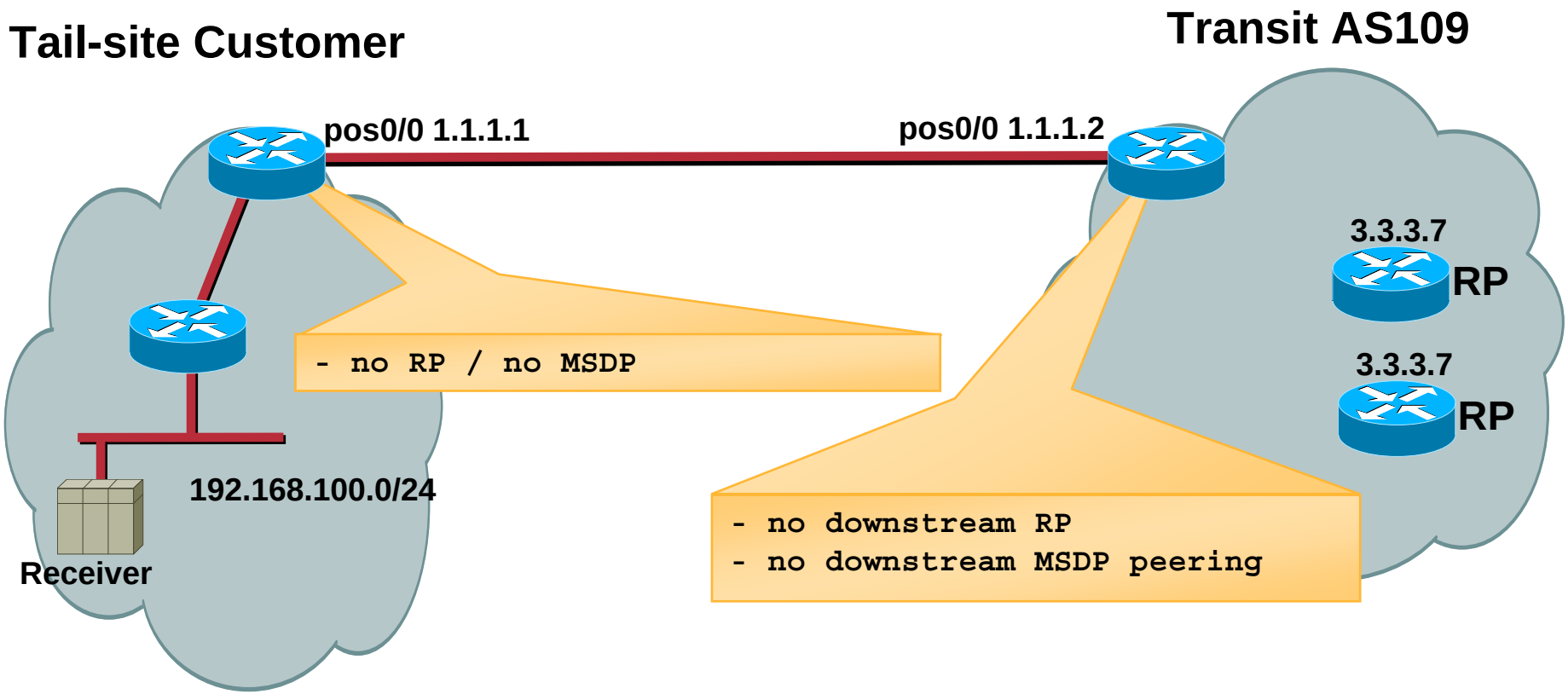


Single-Homed, ISP RP, Non-MBGP



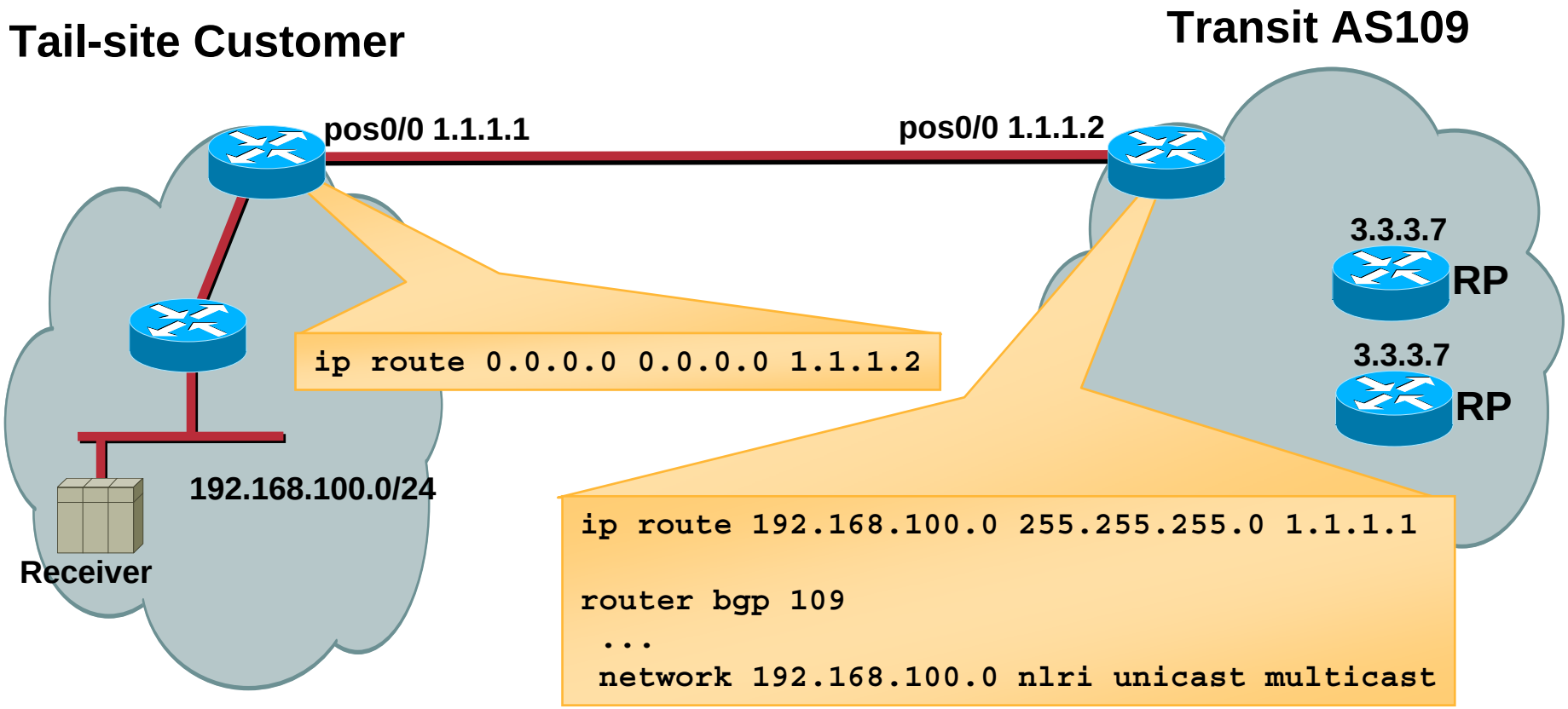
Single-Homed, ISP RP, Non-MBGP

MSDP RPF Check



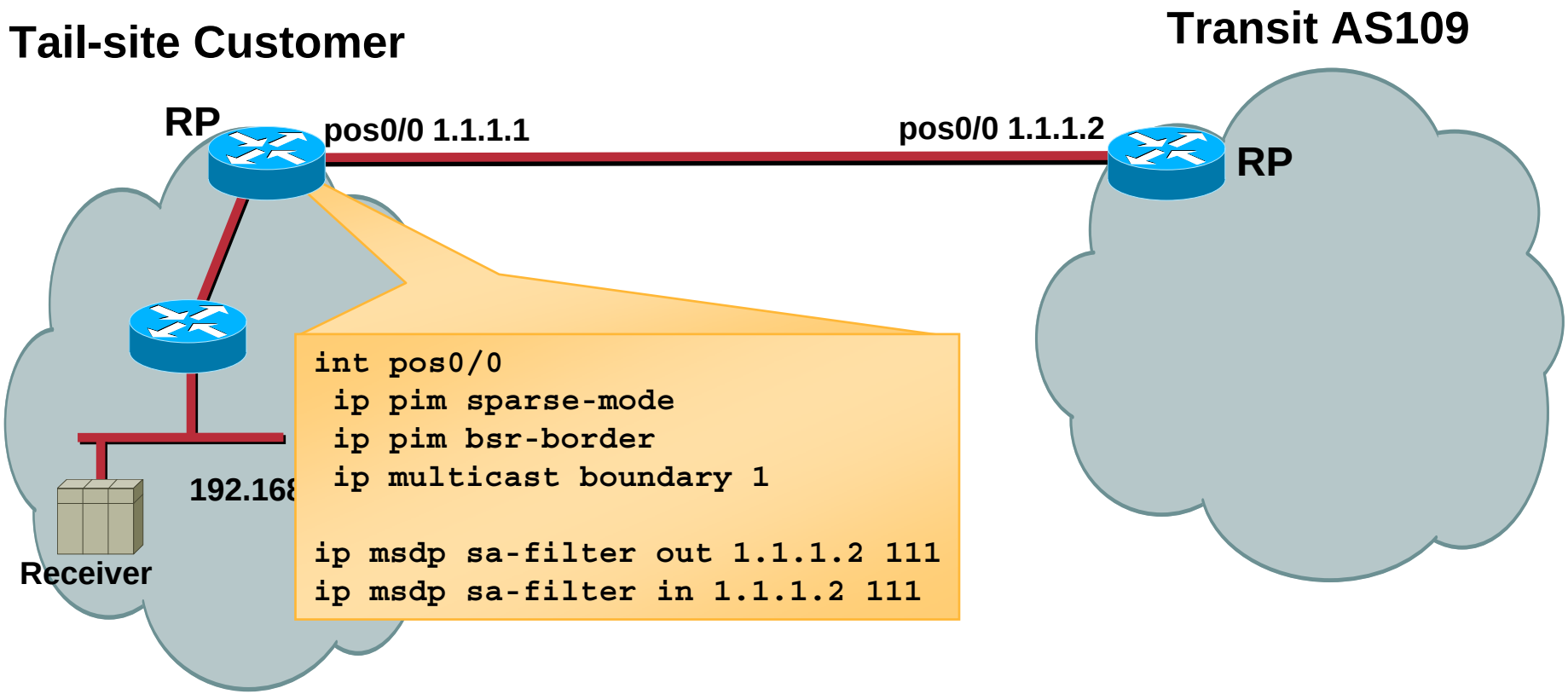
Single-Homed, ISP RP, Non-MBGP

Multicast RPF Check



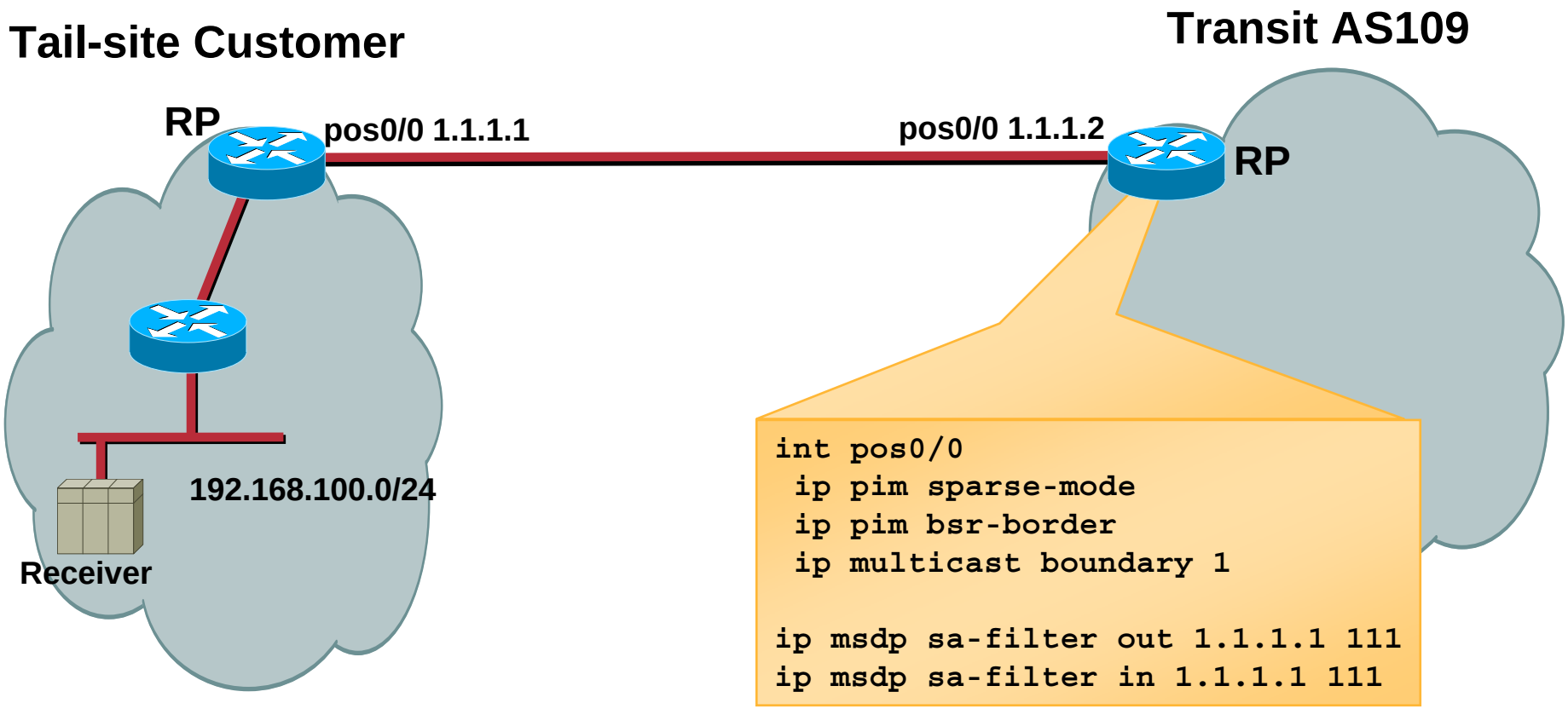
Single-Homed, Customer RP, Non-MBGP

Cisco.com



Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, Non-MBGP

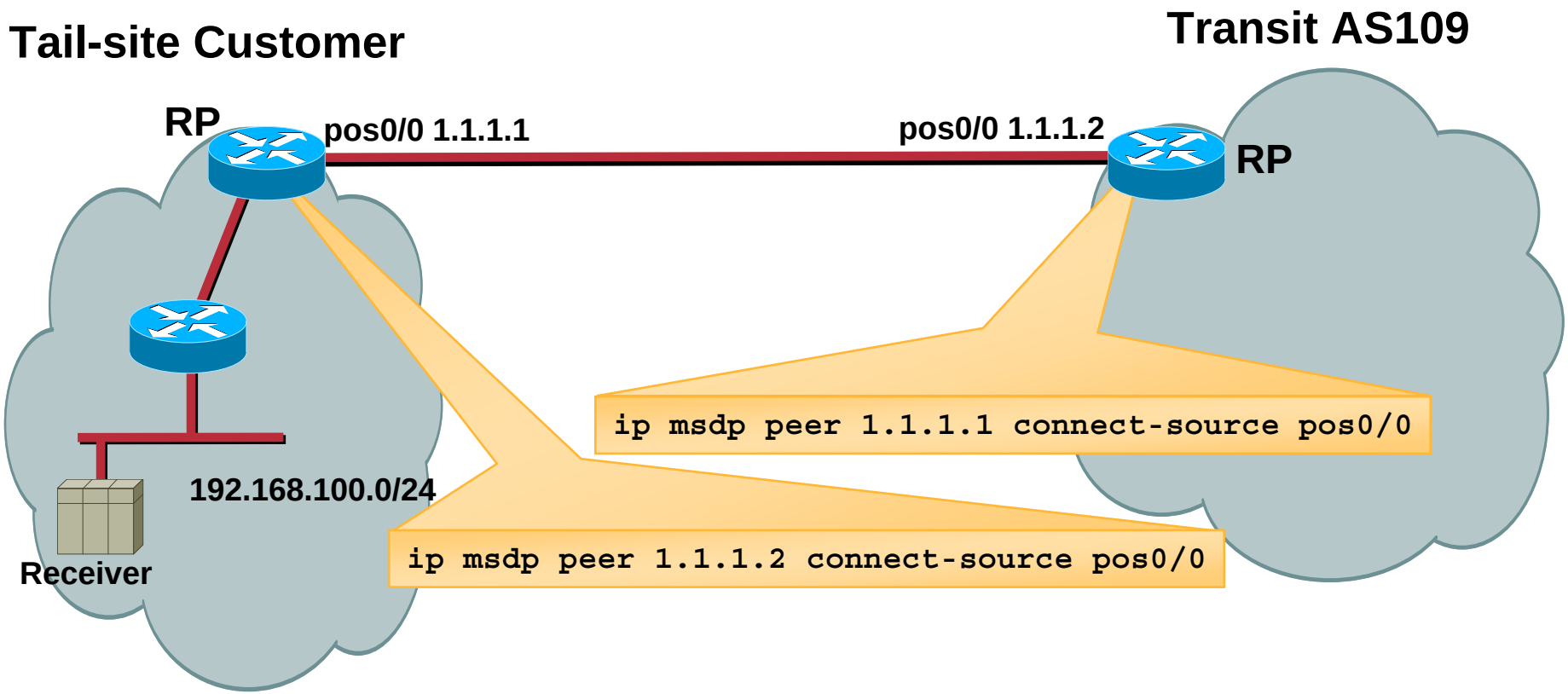


Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, Non-MBGP

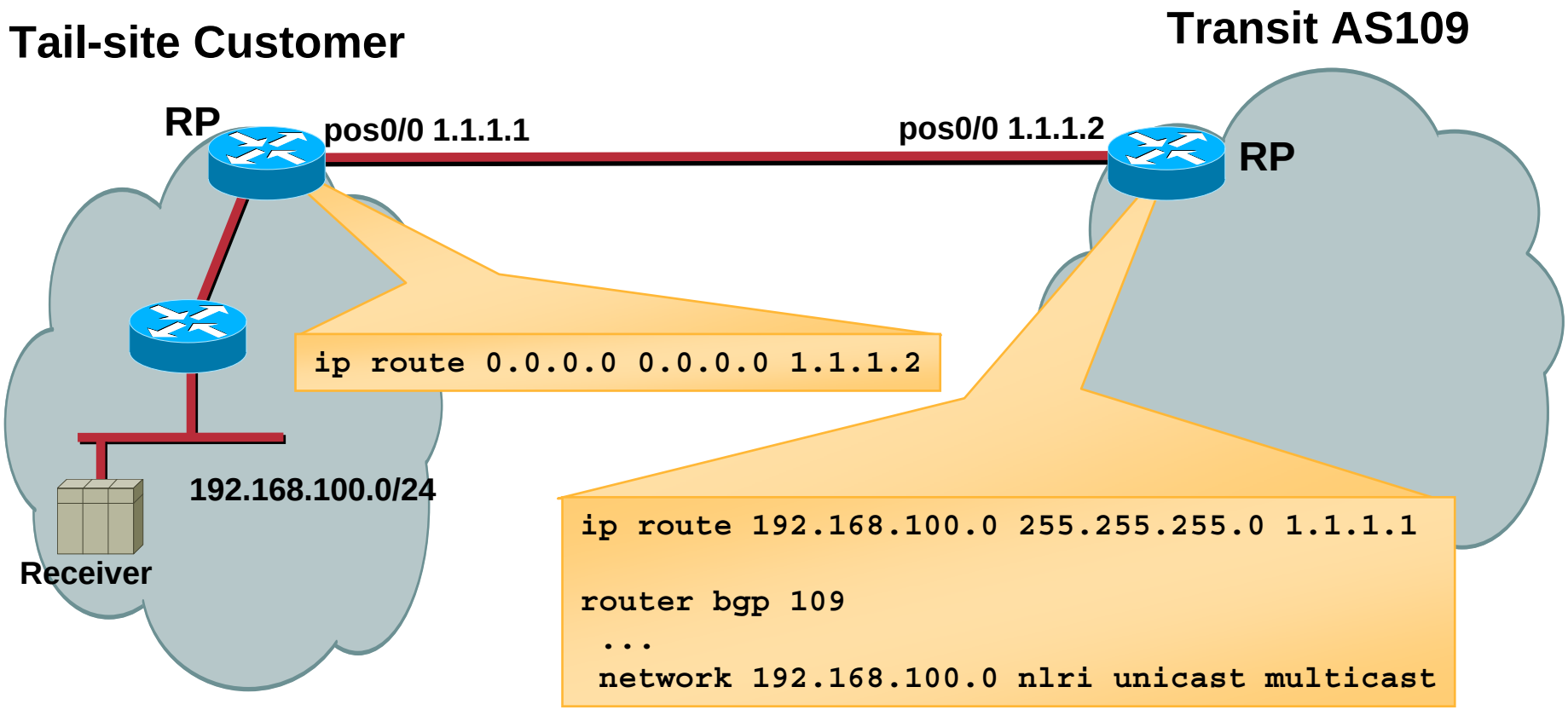
Cisco.com

MSDP RPF Check

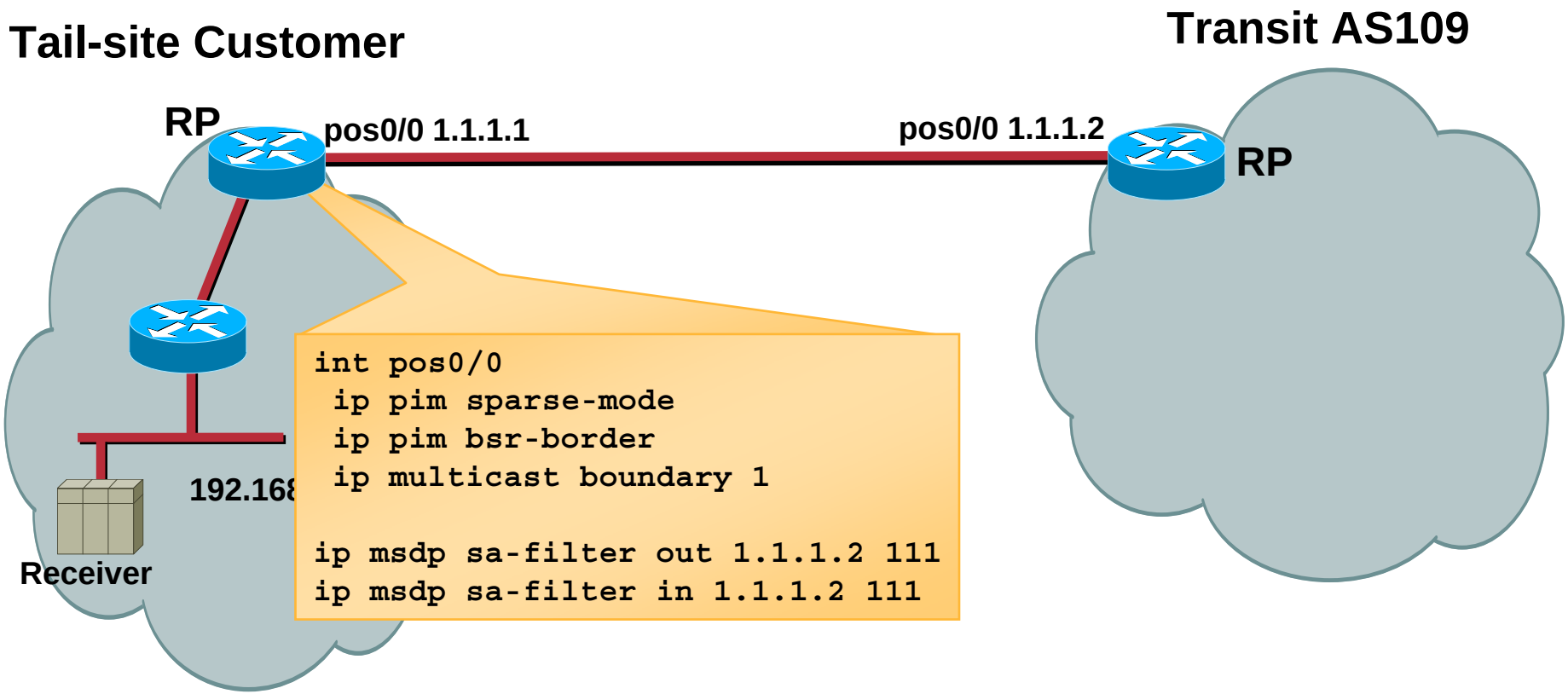


Single-Homed, Customer RP, Non-MBGP

Multicast RPF Check

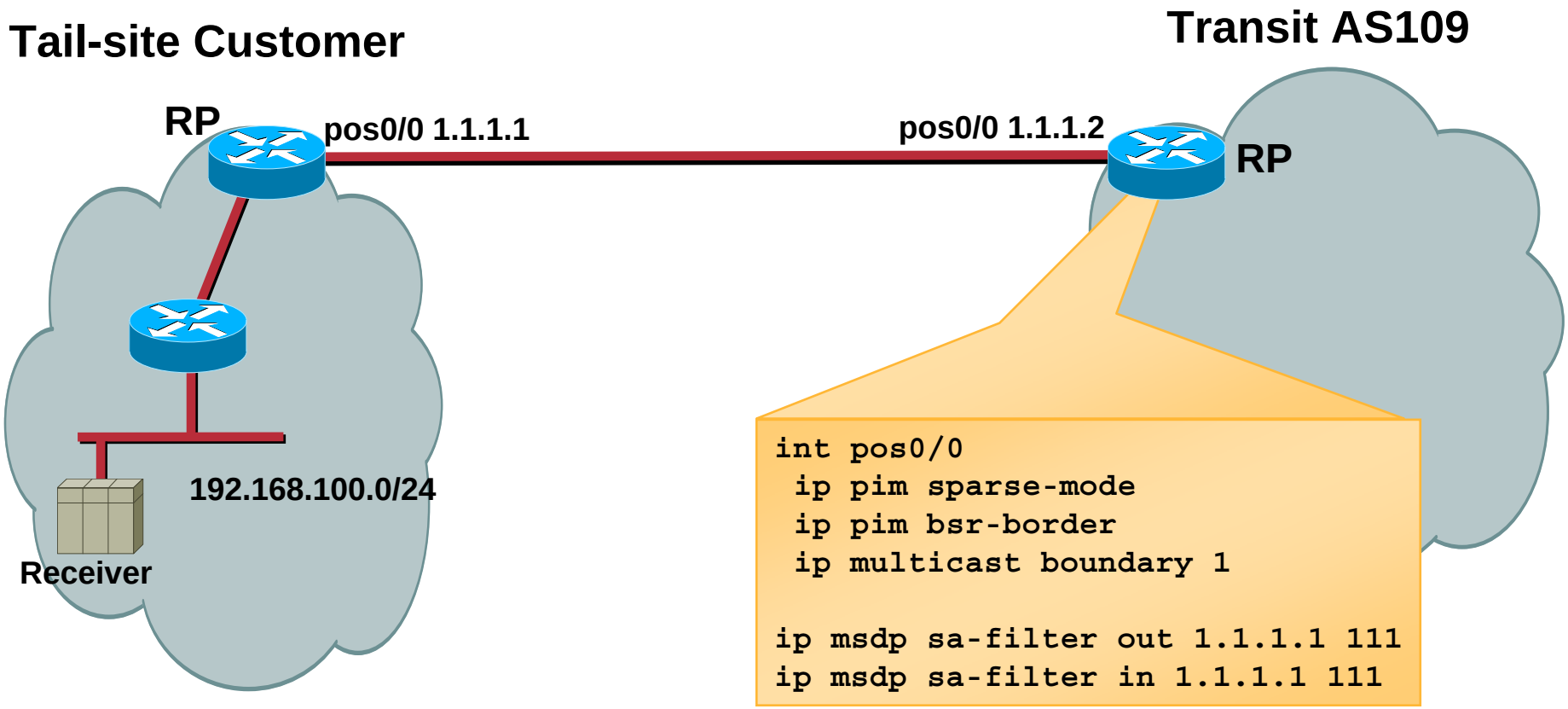


Single-Homed, Customer RP, MBGP



Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, MBGP

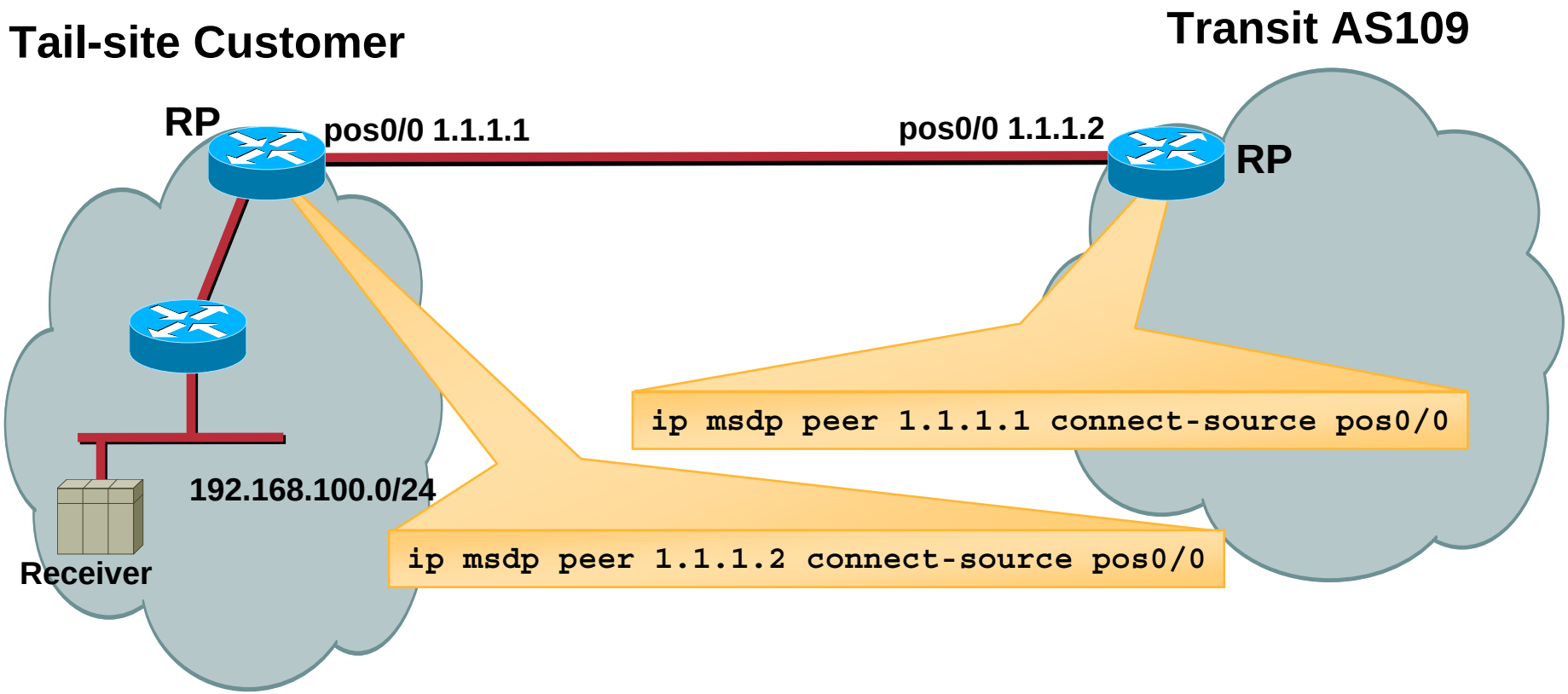


Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, MBGP

Cisco.com

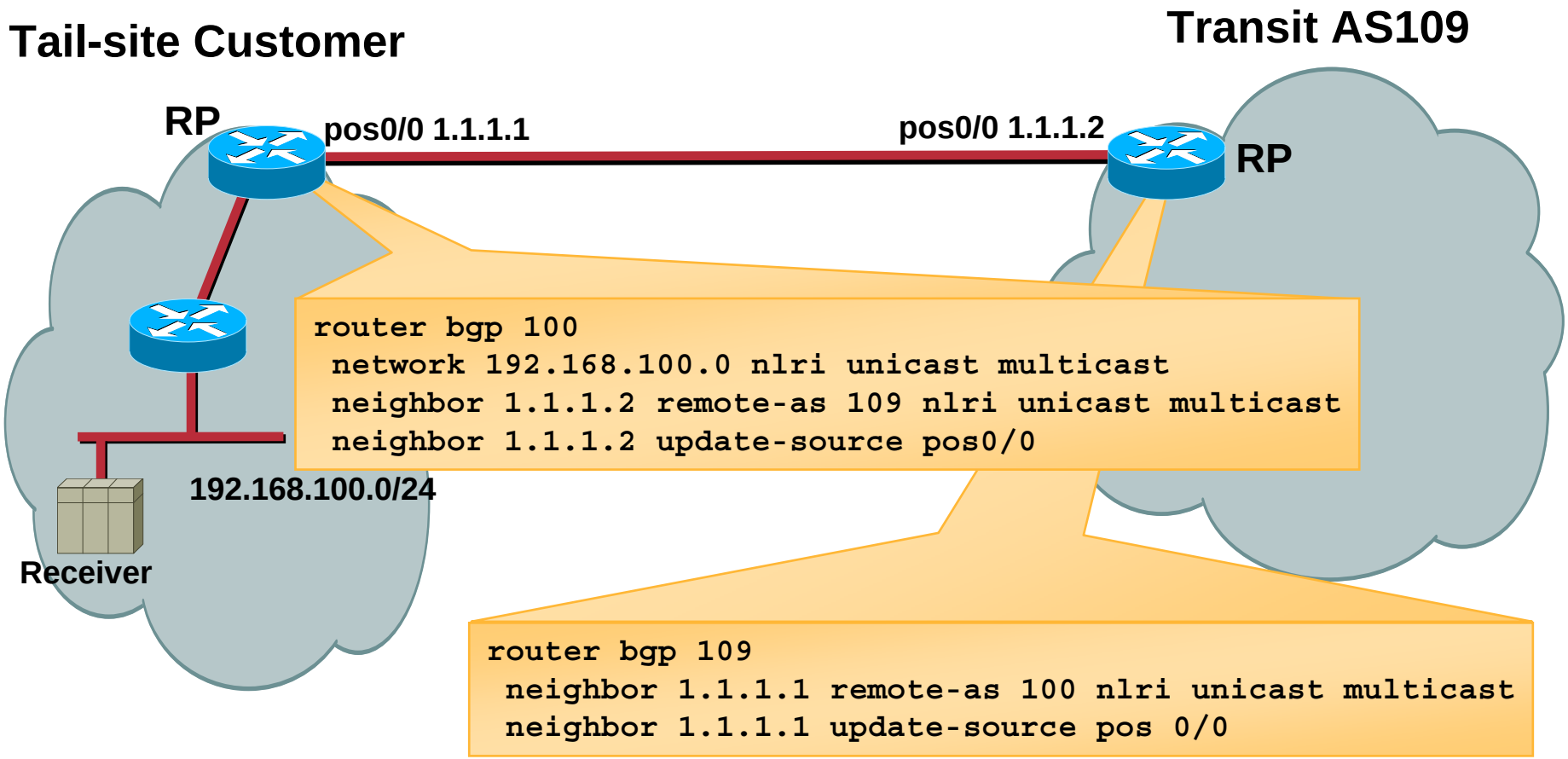
MSDP RPF Check



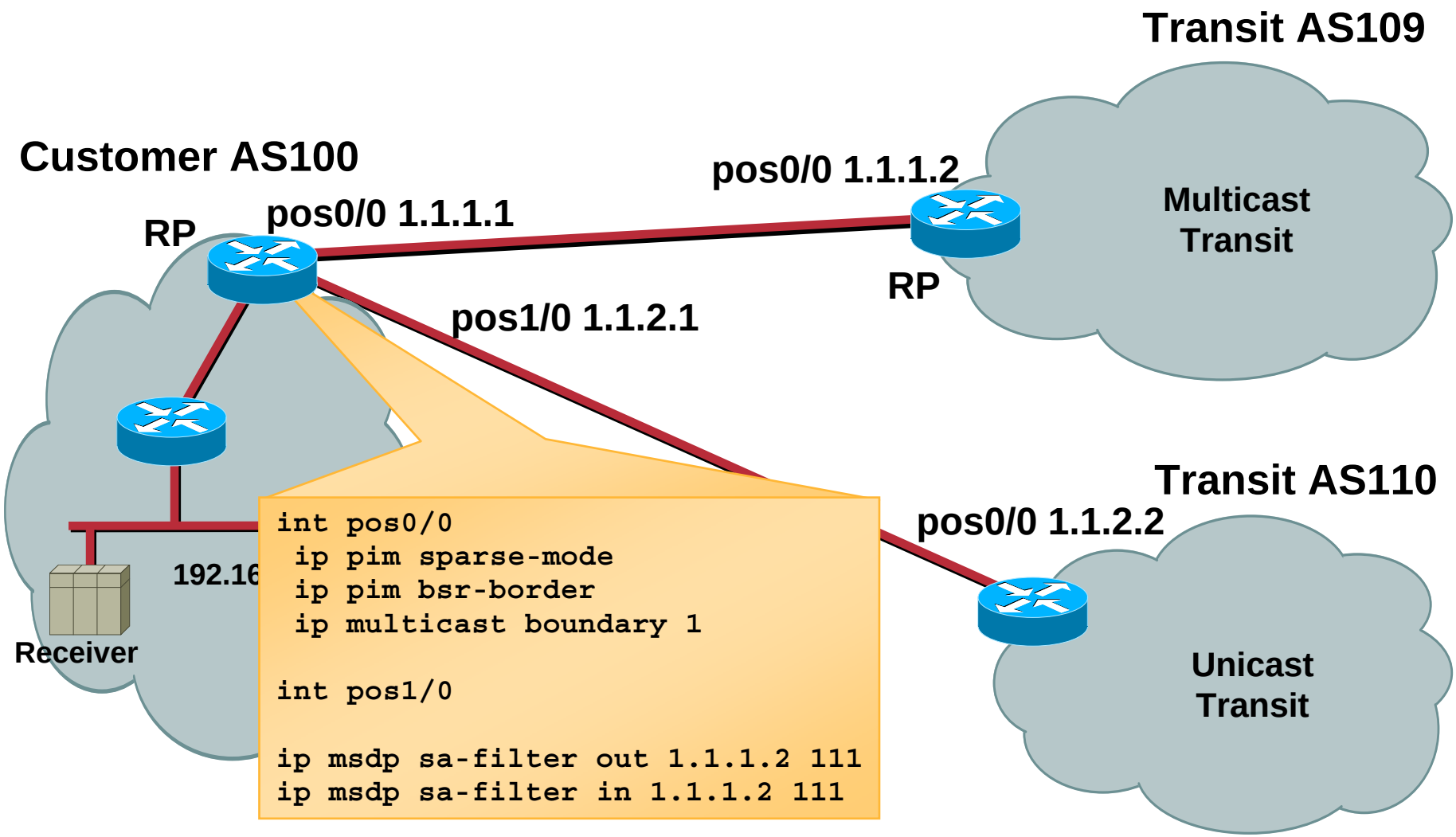
Single-Homed, Customer RP, MBGP

Cisco.com

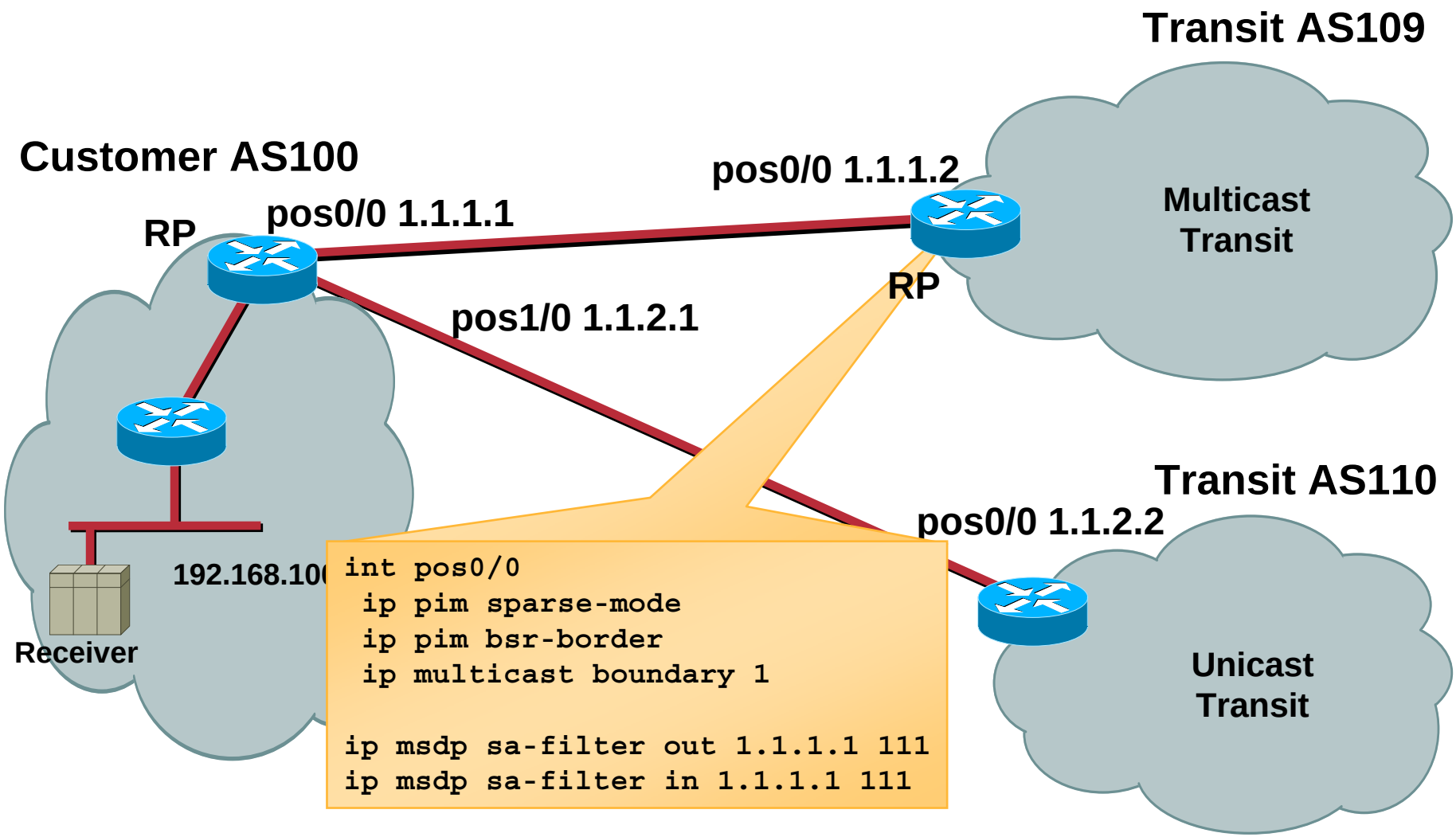
Multicast RPF Check



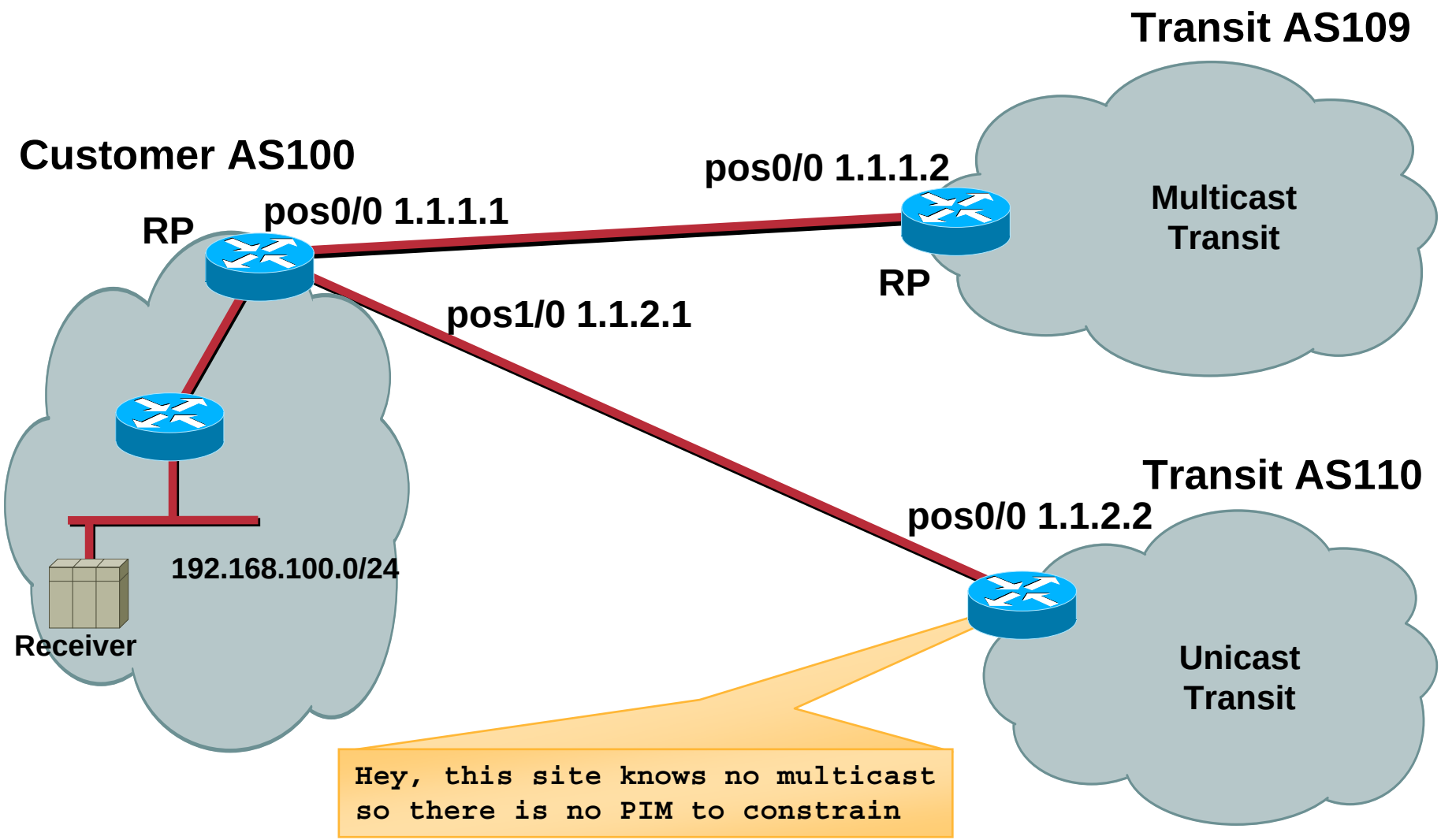
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

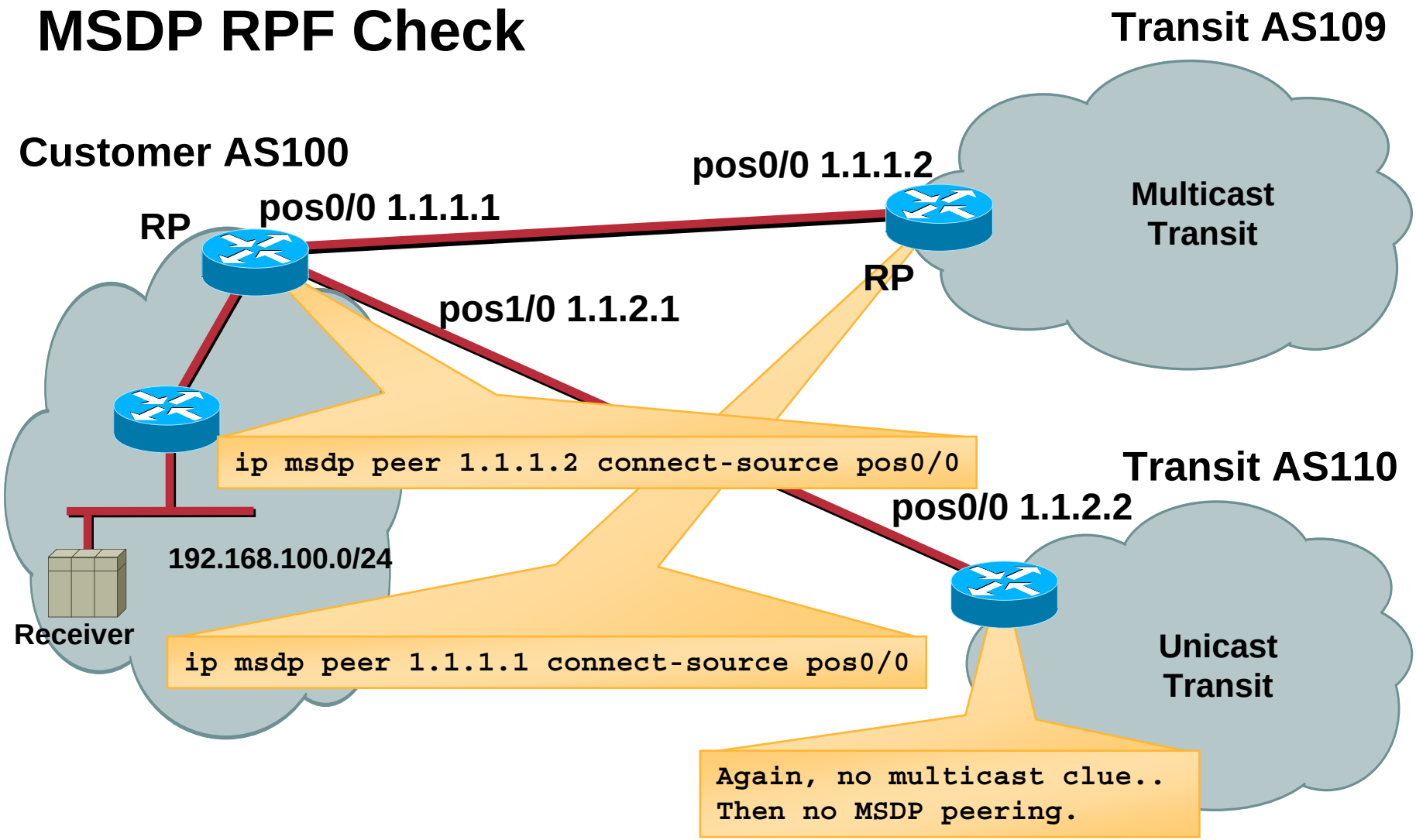


Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



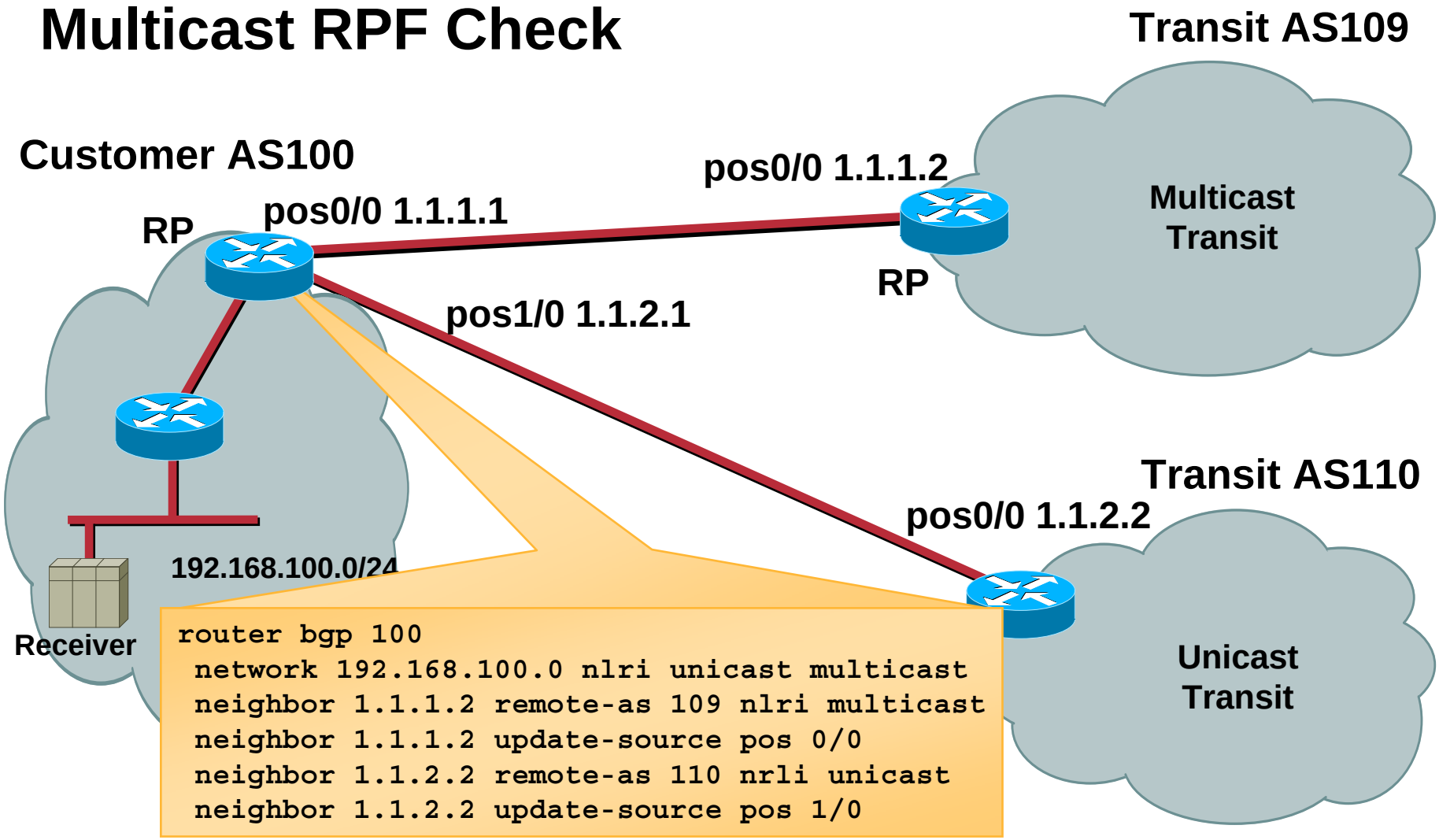
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

MSDP RPF Check



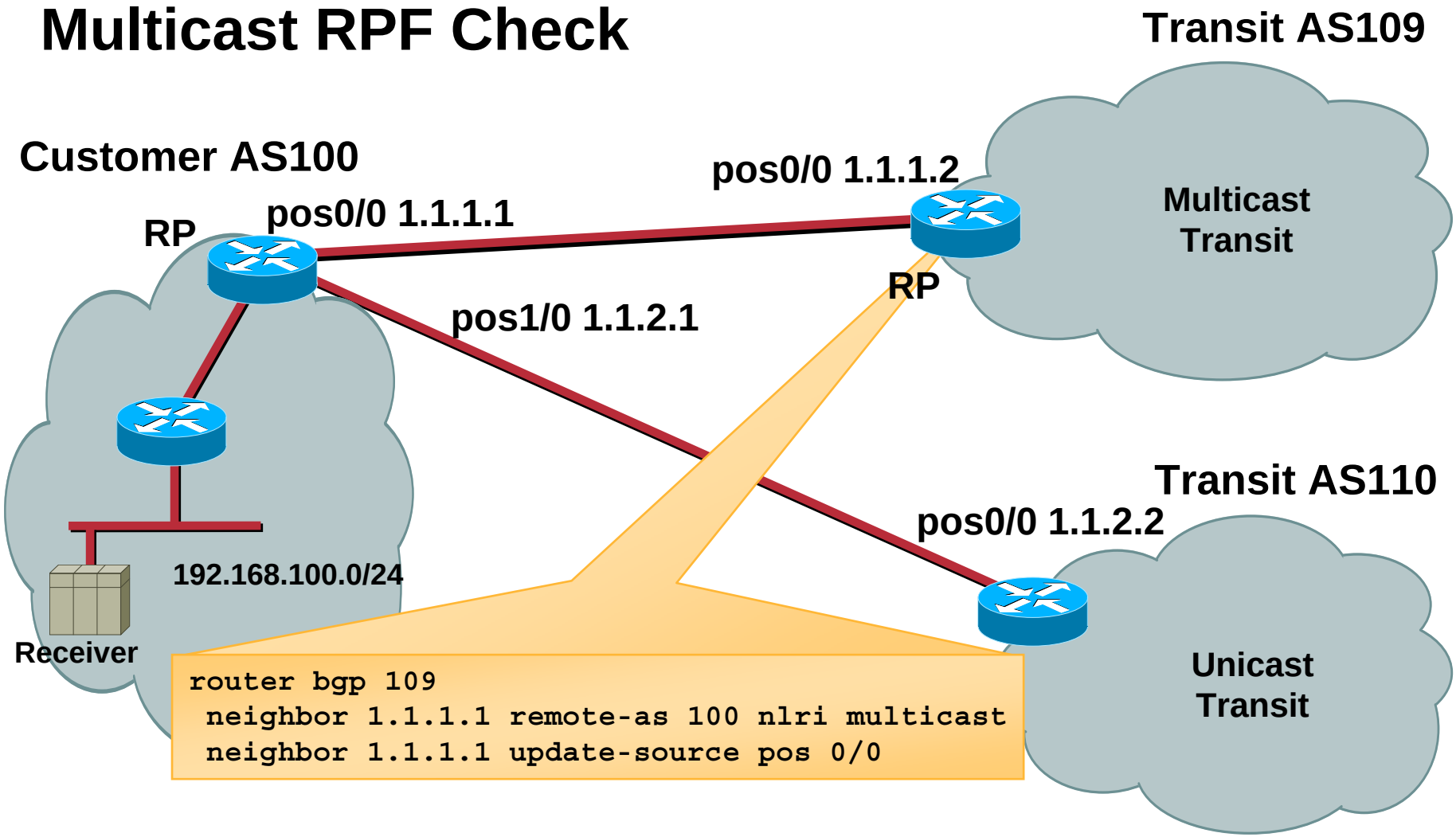
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Multicast RPF Check



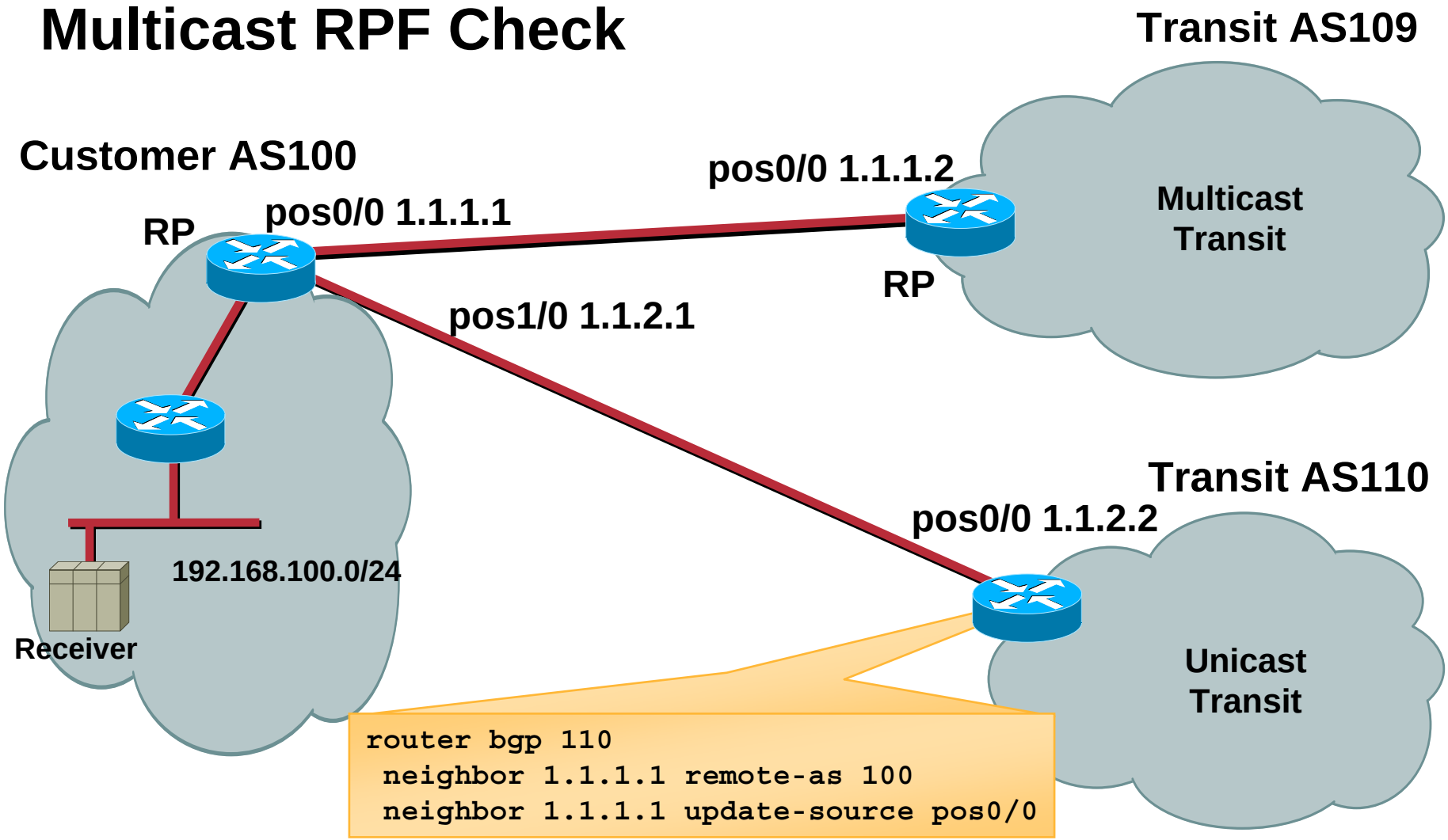
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Multicast RPF Check

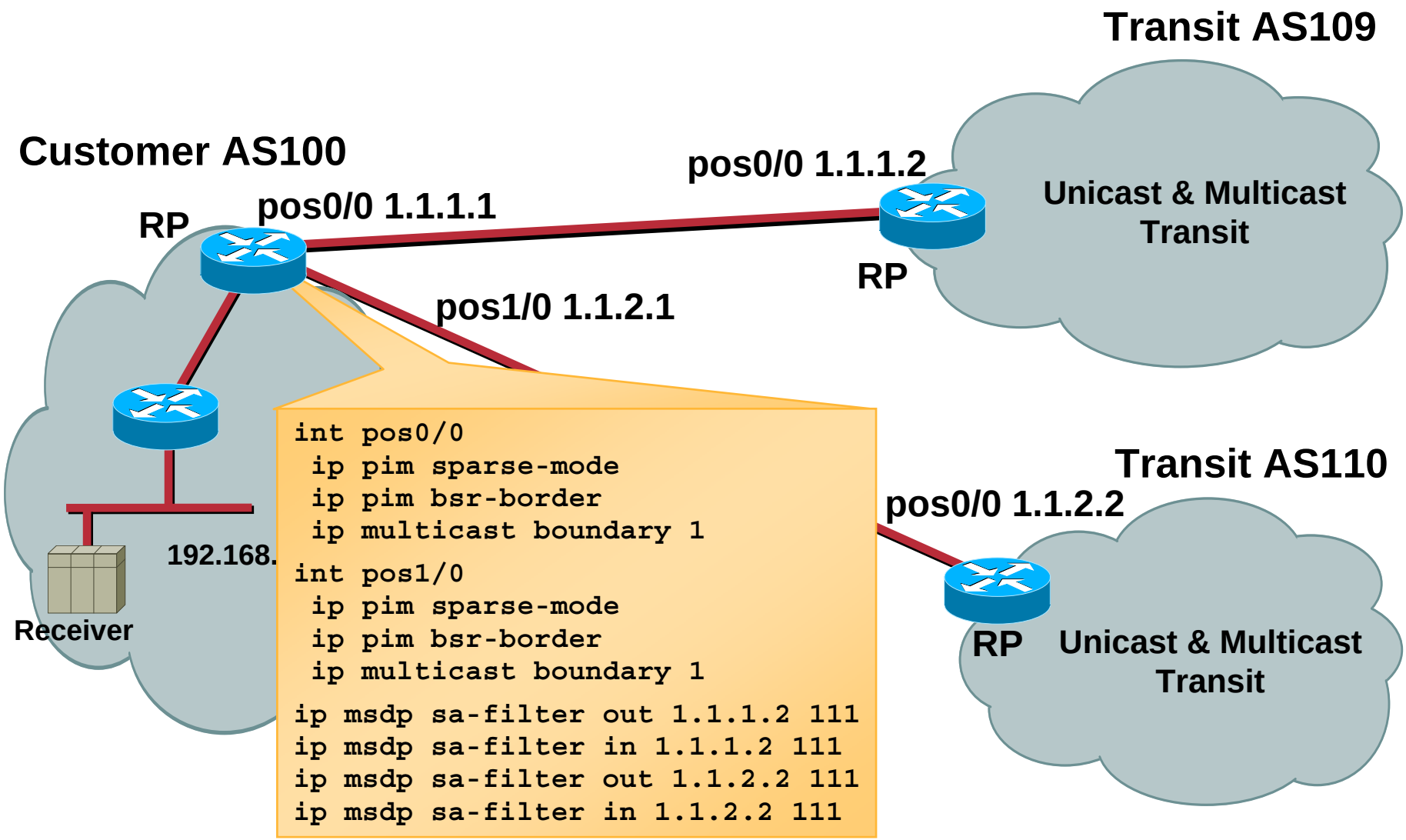


Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

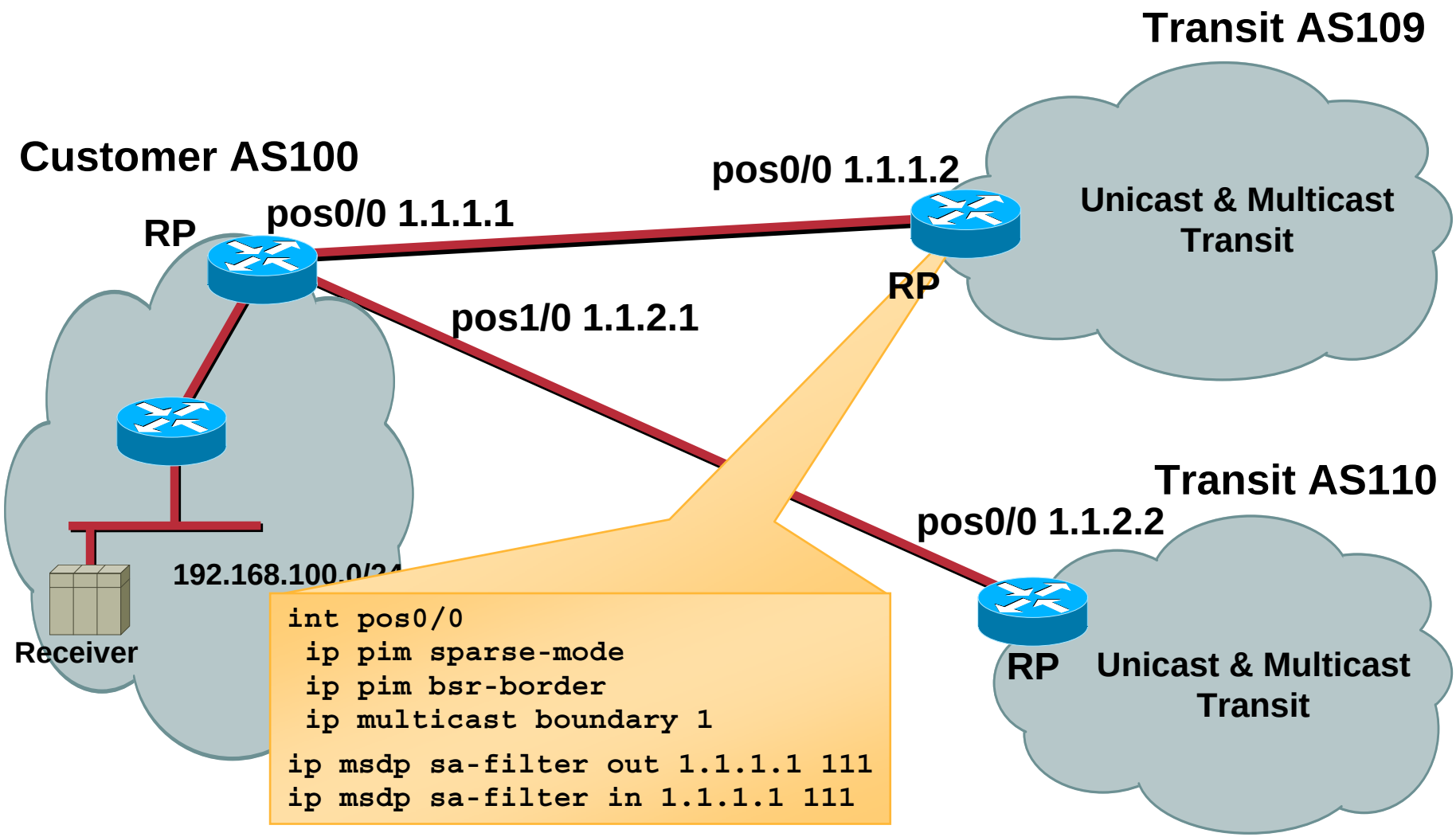
Multicast RPF Check



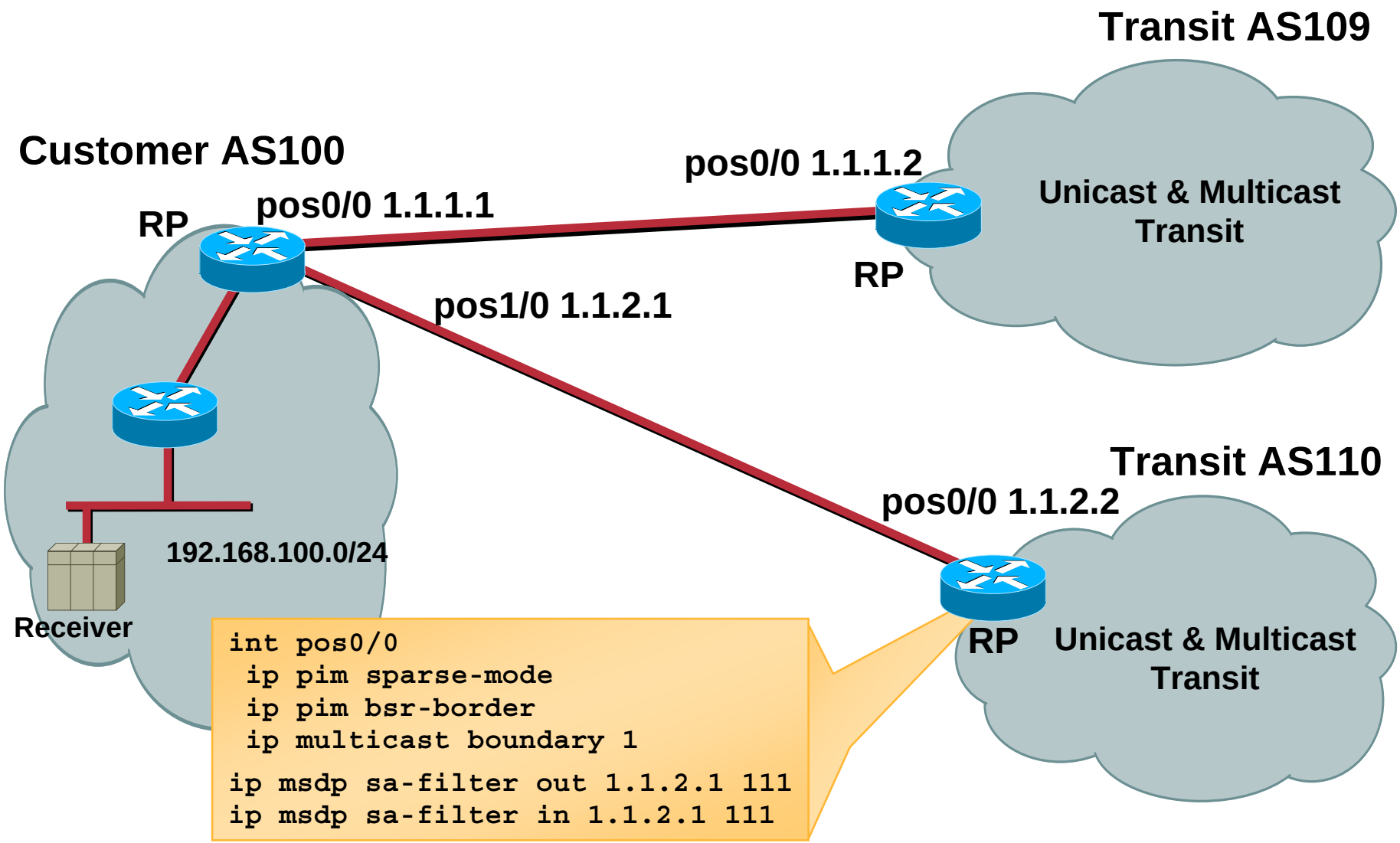
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

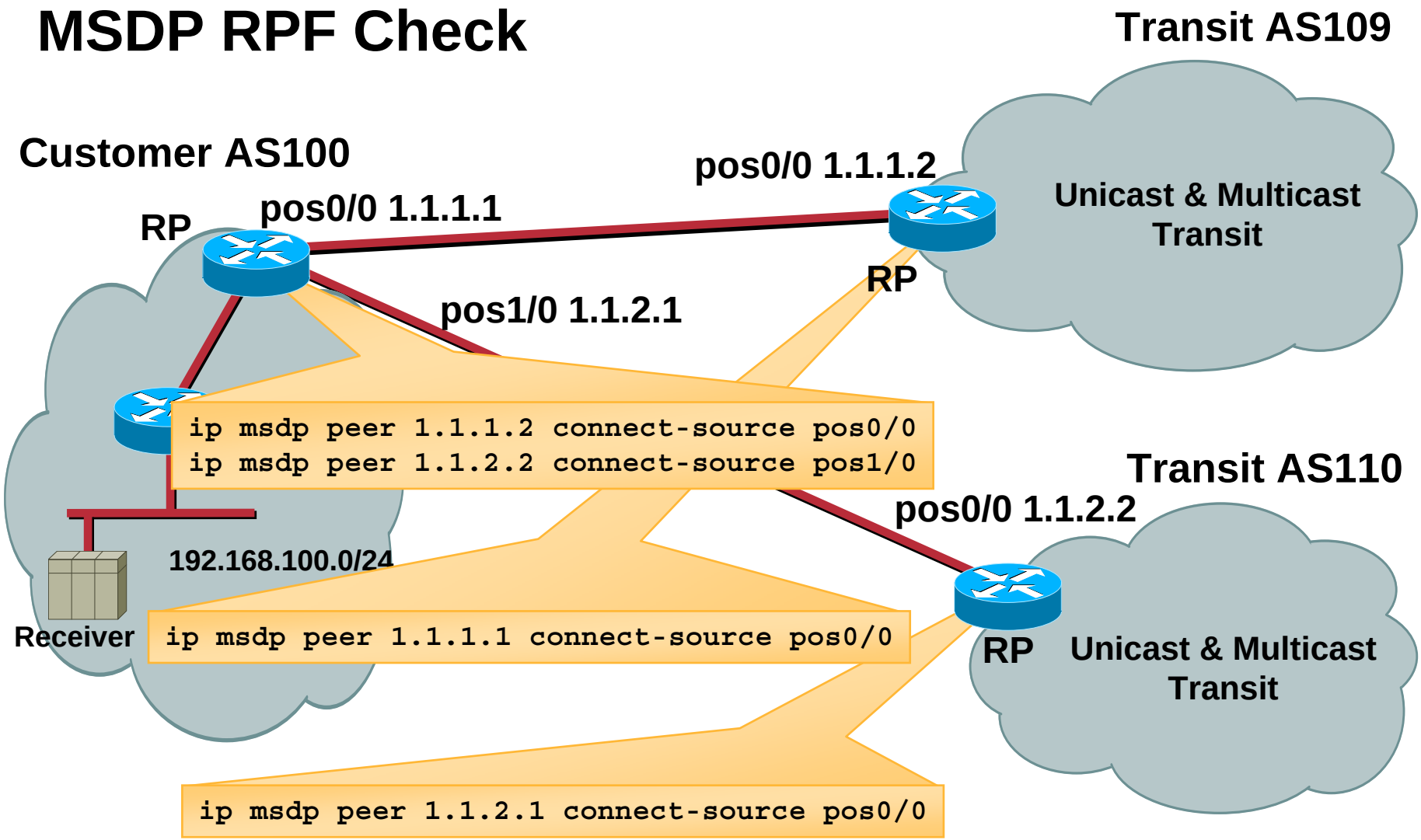


Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



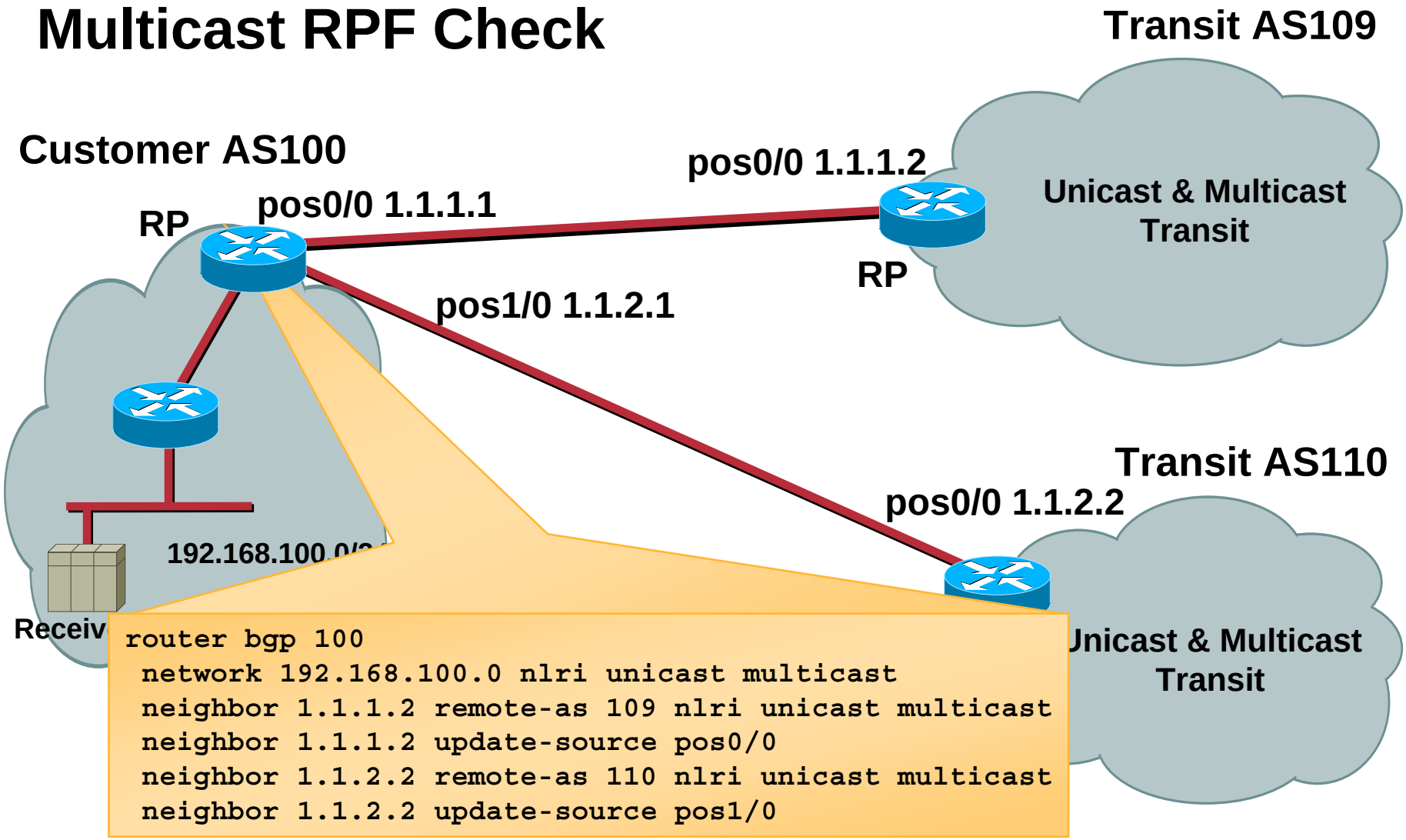
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

MSDP RPF Check



Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

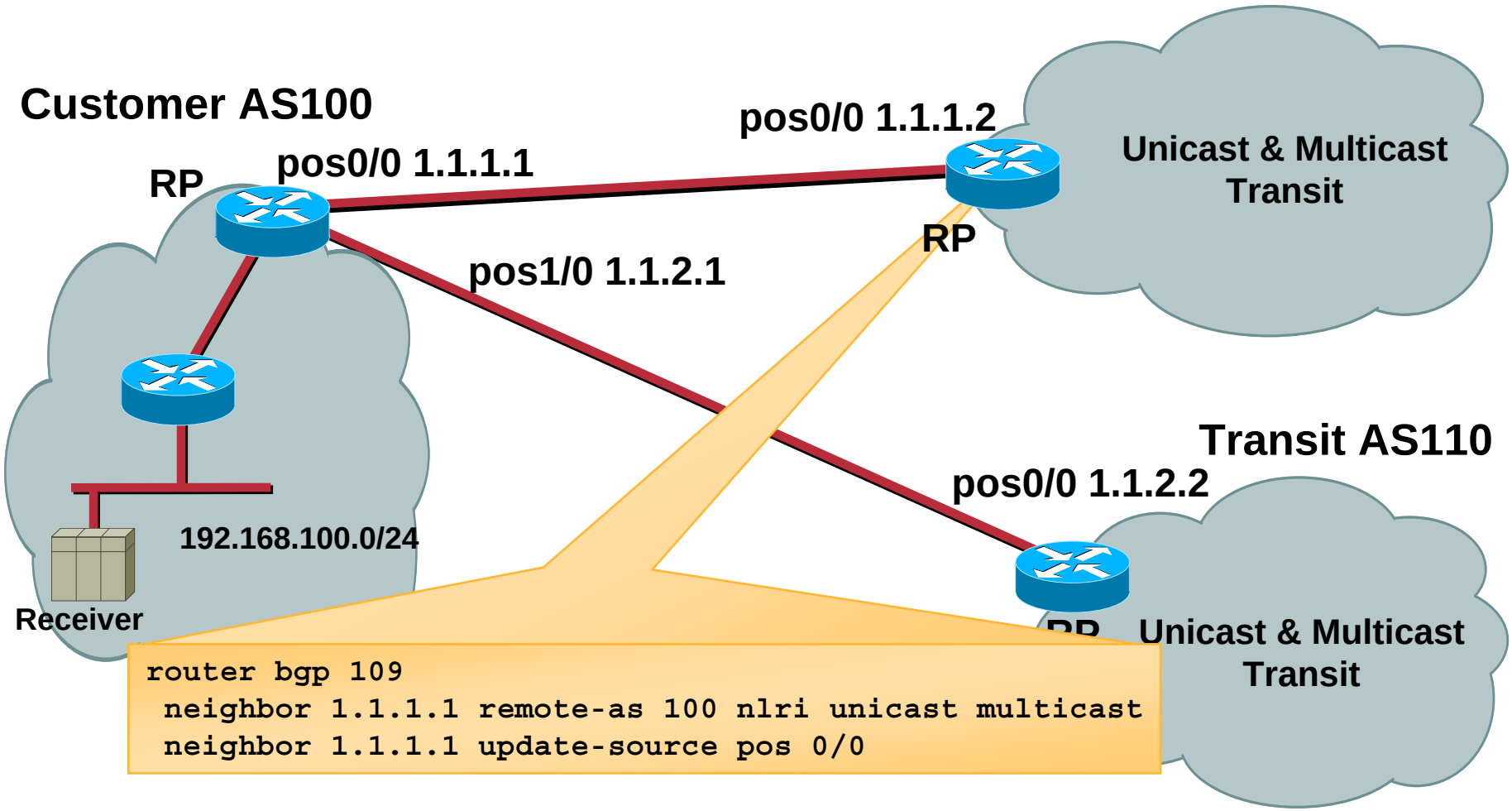
Multicast RPF Check




```
router bgp 100
 network 192.168.100.0 nlri unicast multicast
 neighbor 1.1.1.2 remote-as 109 nlri unicast multicast
 neighbor 1.1.1.2 update-source pos0/0
 neighbor 1.1.2.2 remote-as 110 nlri unicast multicast
 neighbor 1.1.2.2 update-source pos1/0
```

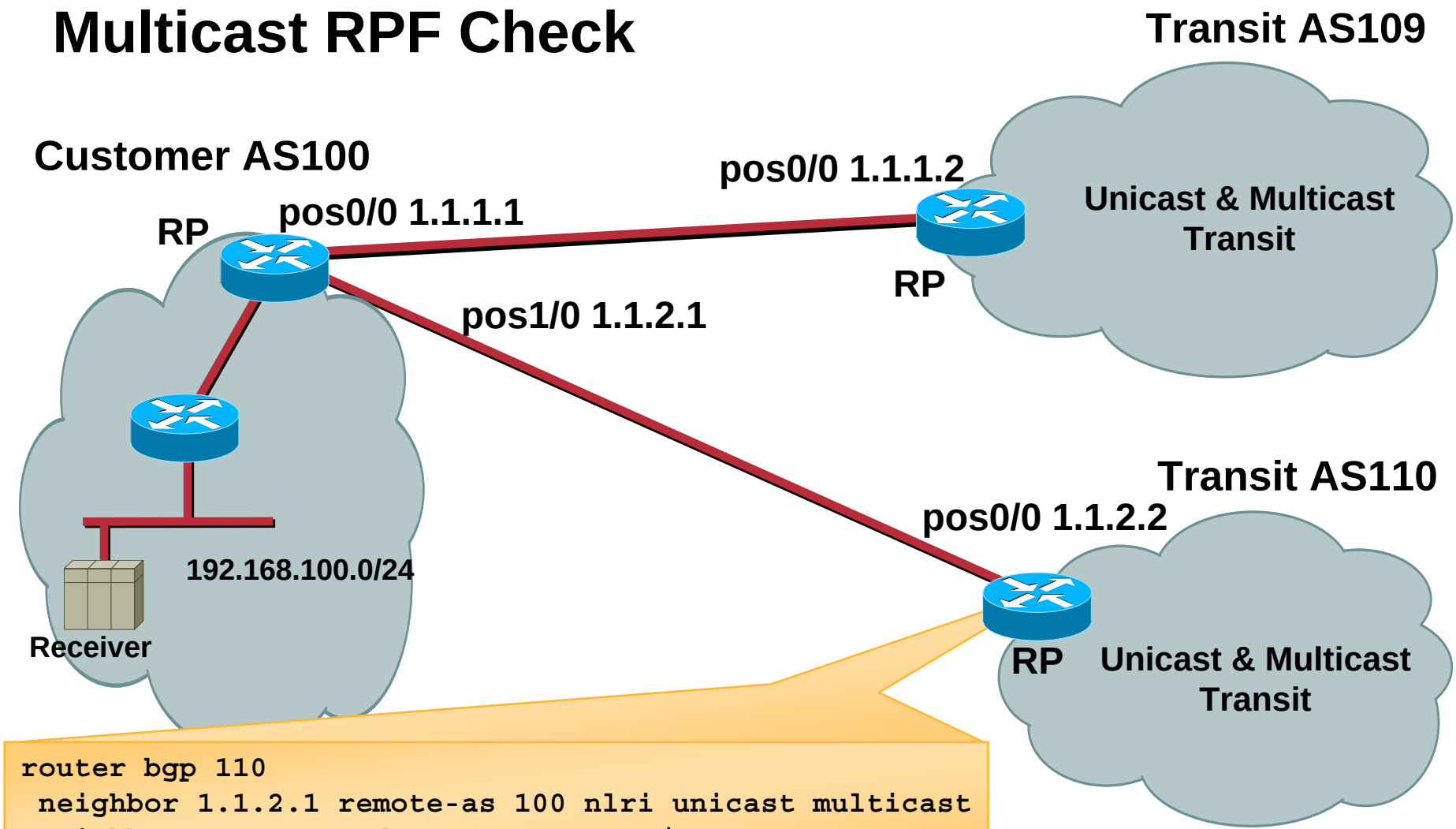
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

Multicast RPF Check



Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

Multicast RPF Check




```
router bgp 110
 neighbor 1.1.2.1 remote-as 100 nlri unicast multicast
 neighbor 1.1.2.1 update-source pos0/0
```

GLOP—Static Allocation of 233/8

- Temporary allocation of 233/8
  - RFC 2770
- Statically assigned by mapping AS number into middle octets
    - http://gigapop.uoregon.edu/glop/index.html
- Provides each AS with /24 addresses to use while waiting another solution

GLOP—Static Allocation of 233/8

- Company A owns AS 5662. How do we use GLOP to come up with a multicast address range?
- 5662 written in binary is: 0001011000011110
- Map the high order octet to the second octet of the address, and the low order octet to the third octet:
- 0 0 0 1 0 1 1 0 | 0 0 0 1 1 1 1 0
- 0 0 0 1 0 1 1 0 = 22 0 0 0 1 1 1 1 0 = 30
- AS 5662 gets 233.22.30.0/24 for multicast use over the internet

GLOP—Static Allocation of 233/8

- The hexadecimal value of 5662 is 161E. 16 hex equals 22 decimal and 1E hex equals 30 decimal. We get 233.22.30.0/24.
- The lazy (smart?) way to calculate your GLOP address space is by entering it here and it will calculate if for you:

  http://www.ogig.net/glop/

Agenda

- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

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.

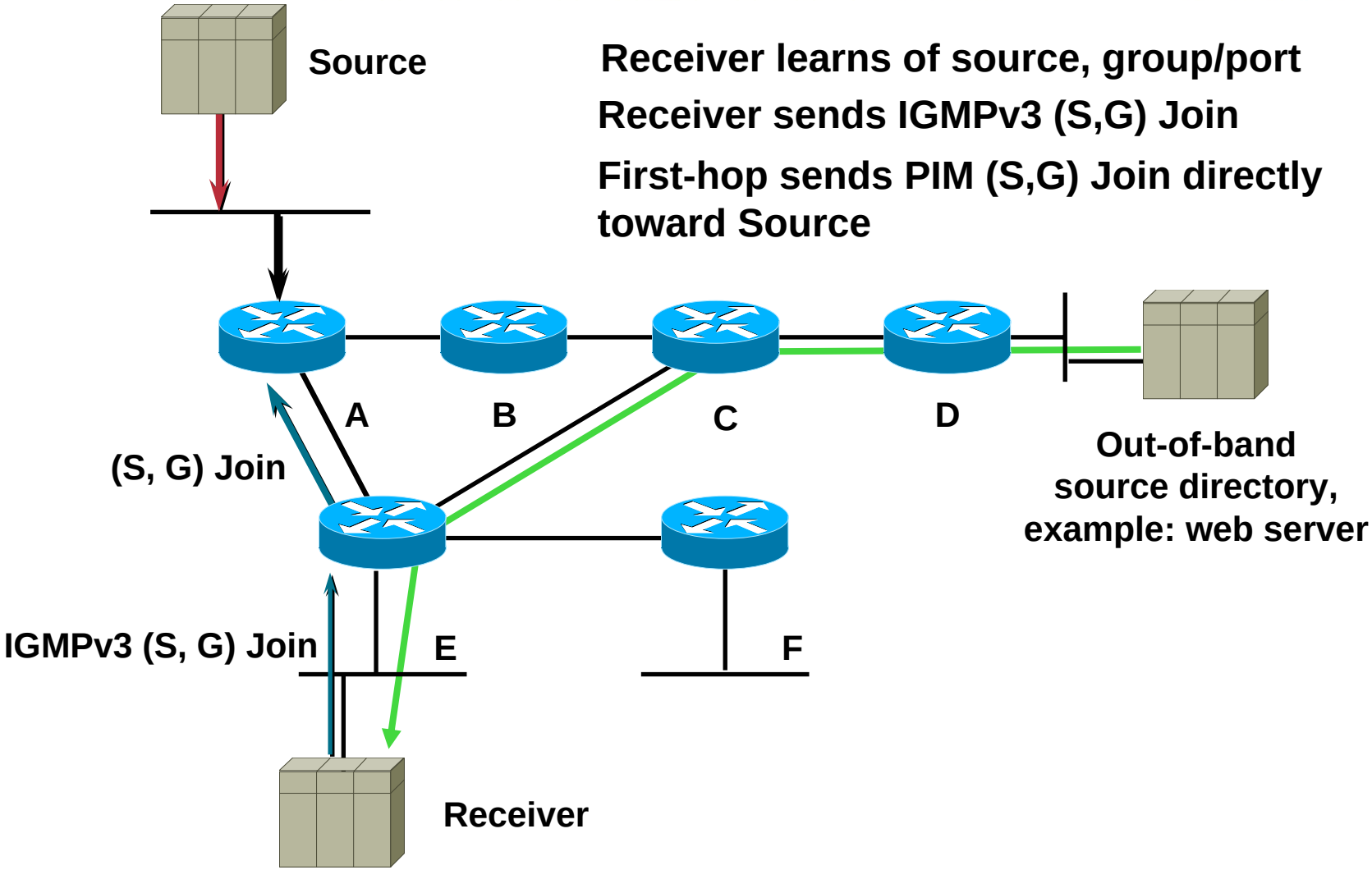
SSM Overview

- 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.

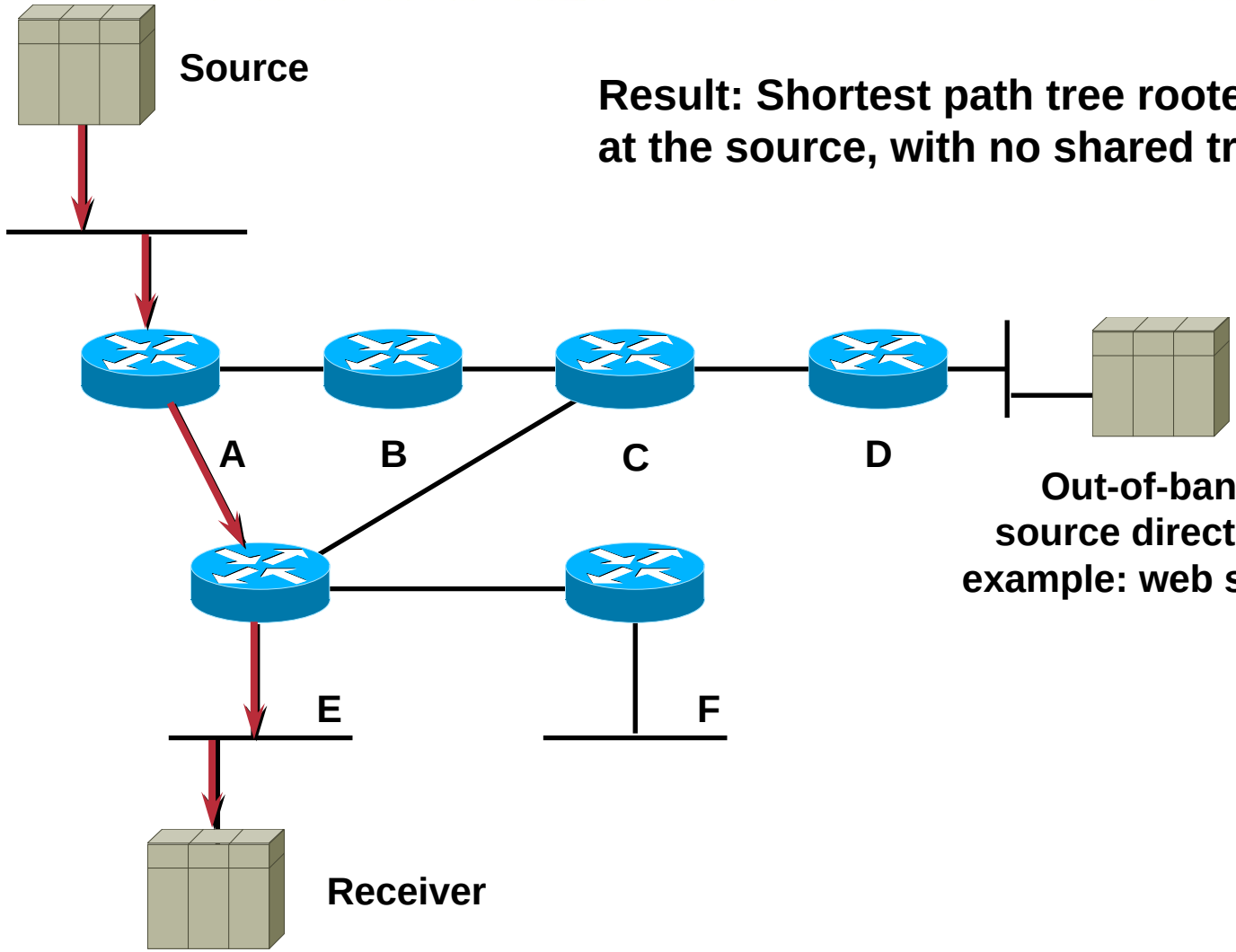
Host-Router Signaling: IGMP

- How hosts tell routers about group membership
- Routers solicit group membership from directly connected hosts
- RFC 1112 specifies version 1 of IGMP
  - Supported on Windows 95
- RFC 2236 specifies version 2 of IGMP
  - Supported on latest service pack for Windows and most UNIX systems
- RFC 3376 specifies version 3 of IGMP
  - Windows XP, FreeBSD, Linux have it.
  - ftp://ftpeng.cisco.com/ipmulticast/ssm/index.html#Stacks

PIM Source Specific Mode



PIM Source Specific Mode



Result: Shortest path tree rooted at the source, with no shared tree.

Out-of-band source directory, example: web server

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 starting in IOS versions
    - 12.1(5)T, 12.2, 12.0(15)S, 12.1(8)E

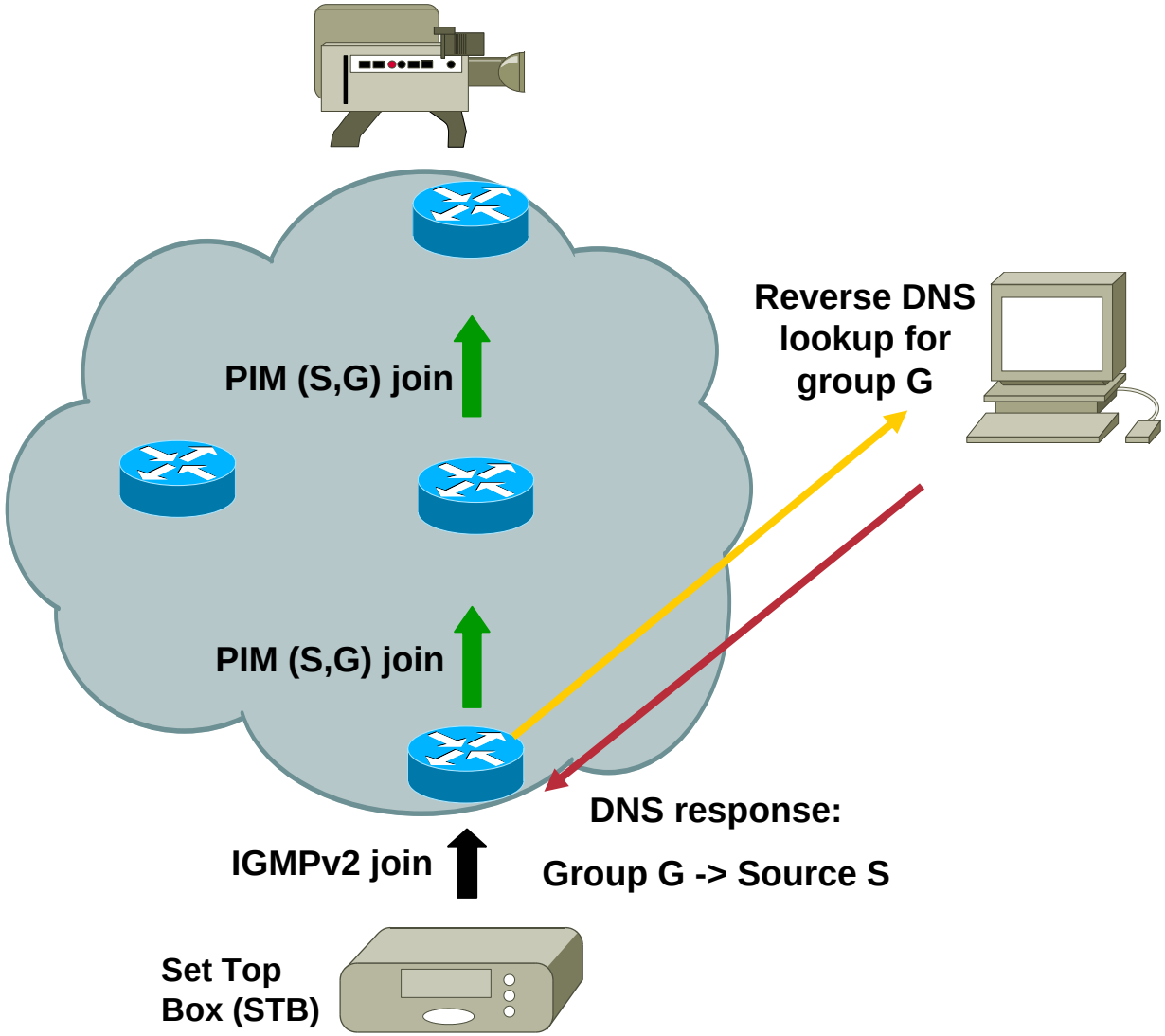
SSM Mapping

- Customers want to deploy SSM
- Hosts in network don’t support IGMPv3
- Host OS is outside of network operators control
- Network operators don’t control content
  - No knowledge about S,G mapping

SSM Mapping

- Bring Source to Group mapping from host to router
- Use an external or internal database for Source to Group mapping
  - Allows content providers to provide the mapping
  - Independent from network operators
  - Database is chosen to be static or DNS
- Allows only for one source per Group

Example



Configuration

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
```

Where Is SSM?

- **Framework**
  - **draft-holbrook-idmr-igmpv3-ssm-04.txt**
  - **draft-ietf-ssm-arch-02.txt**
- **BCP proposal**
  - **draft-ietf-mboned-ssm232-04.txt**
  - **draft-ietf-ssm-overview-04.txt**
- **Supported in:**

  IOS 12.X

  Windows XP, FreeBSD, Linux

  ftp://ftpeng.cisco.com/ipmulticast/ssm/index.html#Stacks

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.

Agenda

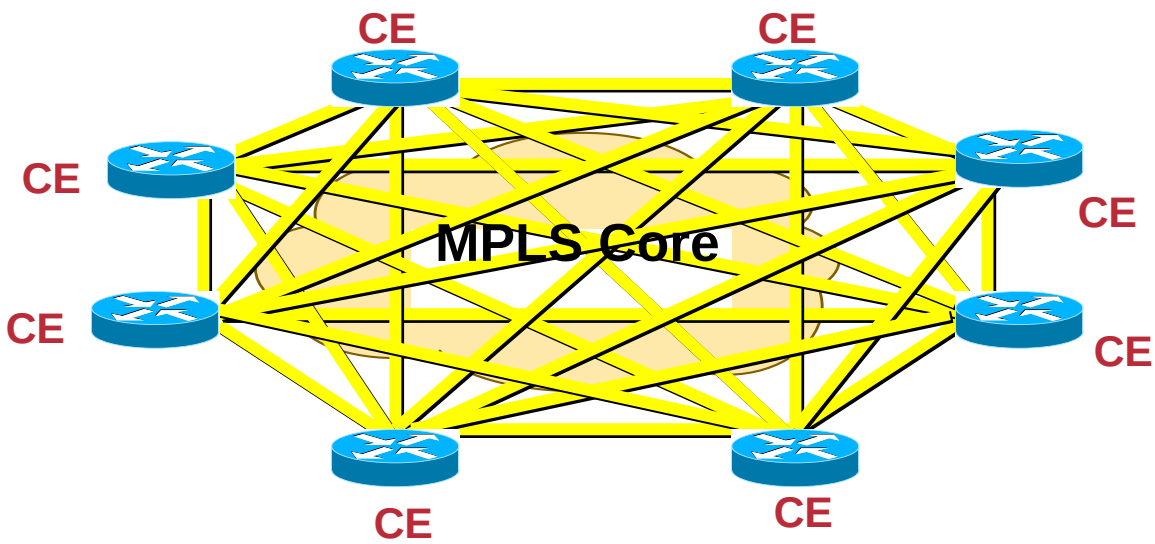
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

Why Multicast VPNs

- Until now only unicast has been supported in MPLS/BGP VPN
- VPN customers need multicast connectivity
  - Applications that require multicast
  - Internet multicast connectivity
- Service Providers want to offer additional services
  - e.g. Video streaming to its VPN customers

Multicast VPN – Challenges

- Workaround has been point-to-point GRE tunnels from CE to CE
- Not scalable with many CE routers
  - Traffic overhead
  - Administration overhead



Multicast VPN – Requirements

- Service provider may have a preferred PIM operating mode in the core.
- VPN customer may have a preferred PIM operating mode in his/her network.
- PIM mode used in the core and VPN should be independent.
- Implementation must support any PIM operating mode in customer and provider networks.
  - PIM Bidirectional (PIM-BIDIR)
  - PIM Source Specific Multicast (PIM-SSM)
  - PIM Sparse-Mode (PIM-SM)

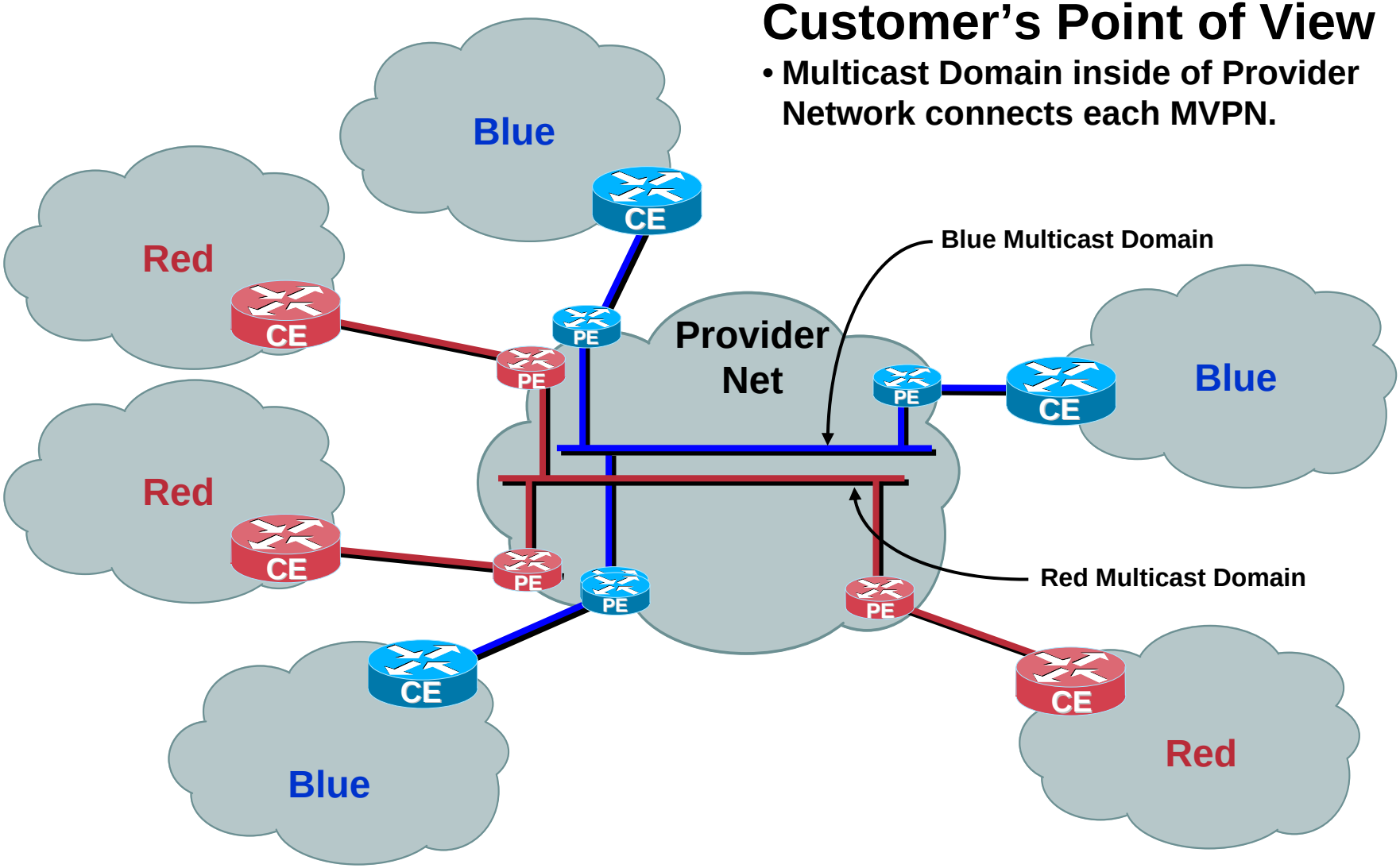
Cisco's Implementation

- **Based on Multicast Domains in draft-rosen-vpn-mcast-02.txt**
  - **Provider builds independent multicast network in the core.**
  - **All arriving customer multicast traffic is encapsulated and multicast across Provider Network.**
  - **A separate multicast group is used inside of Provider Network for each customer VPN.**
    - **Provider's multicast address space is independent of all customer address space.**
    - **Avoids VPN overlap of customers' multicast addresses.**
- **MVPN in 12.2(13)T and 12.0(23)S on 3600, 7200 and 7500. 10k in 12.0(25)S. 12K in 12.0(26)S Q2 cy03. 7600 in 12.2S Q2 cy03.**

Multicast VPN – Overview

Customer's Point of View

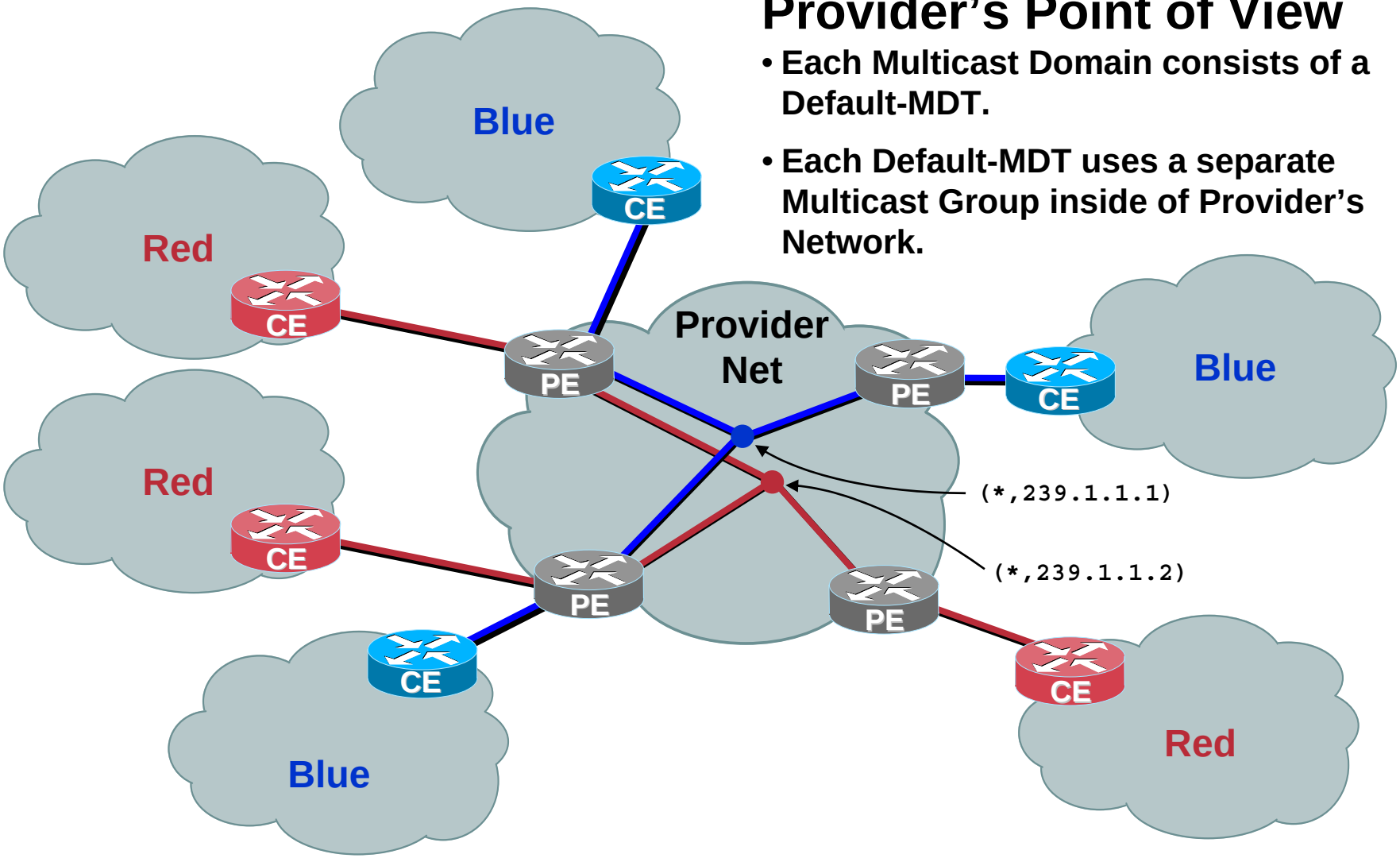
- Multicast Domain inside of Provider Network connects each MVPN.



Multicast VPN – Overview

Provider's Point of View

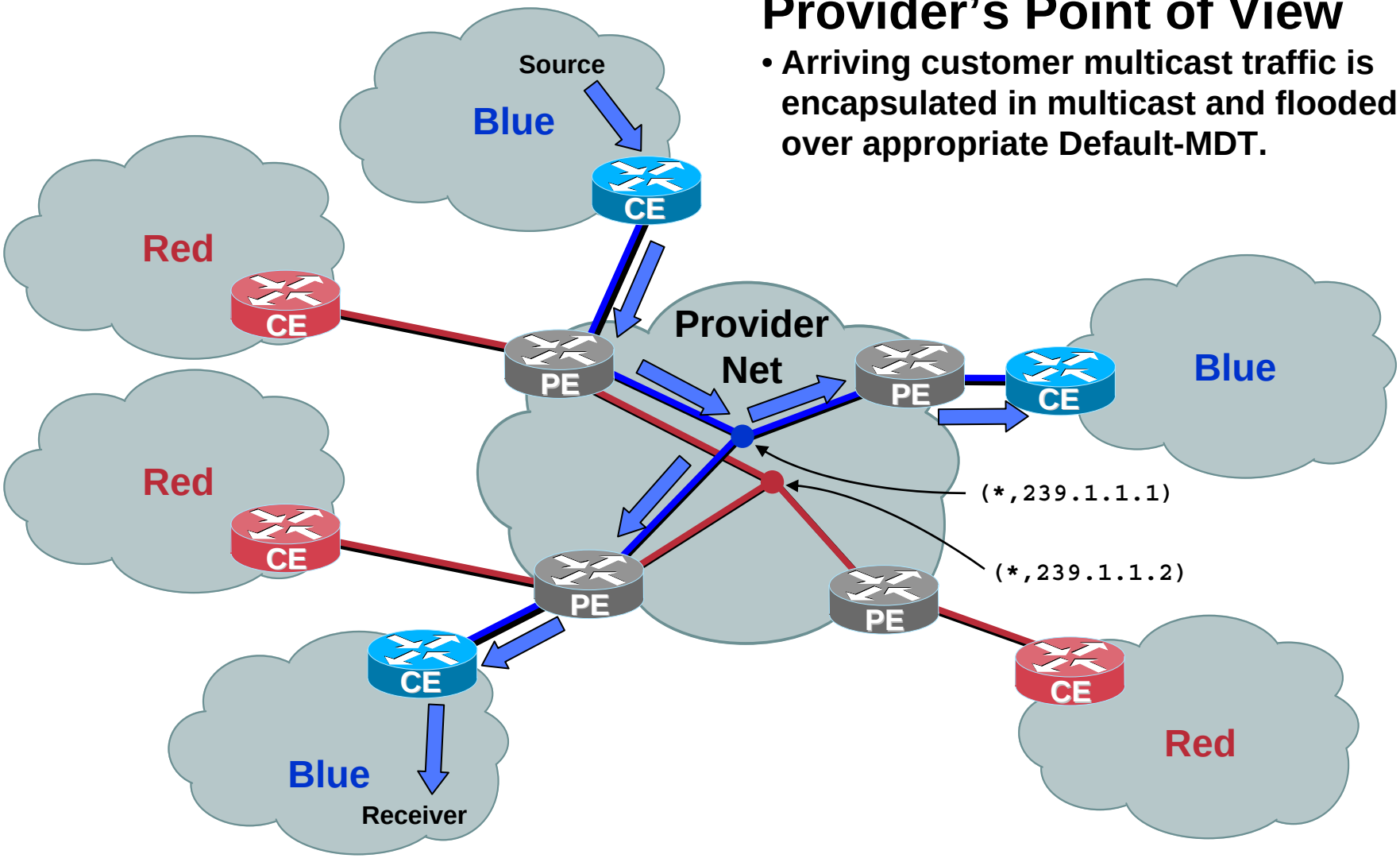
- Each Multicast Domain consists of a Default-MDT.
- Each Default-MDT uses a separate Multicast Group inside of Provider's Network.



Multicast VPN – Overview

Provider's Point of View

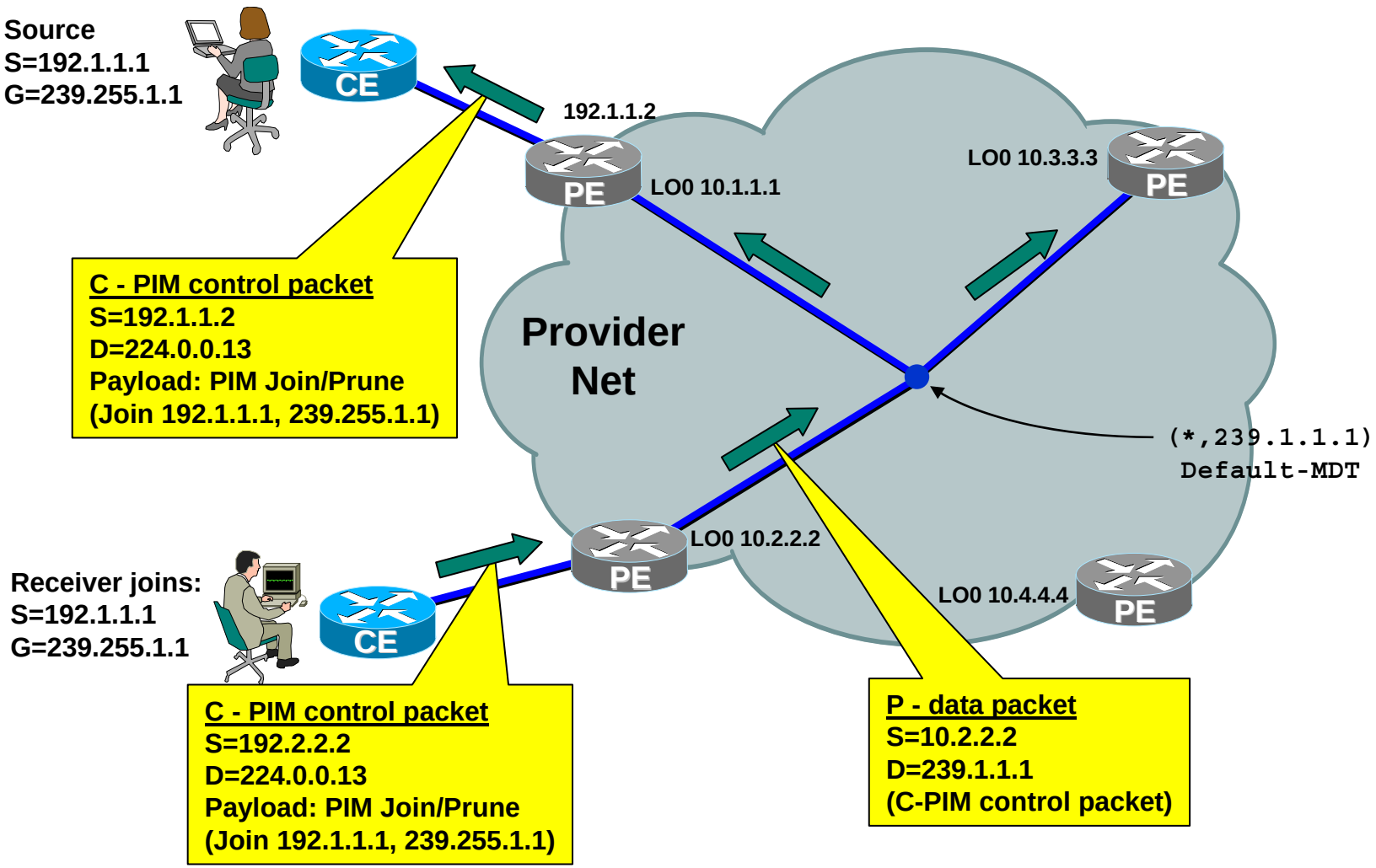
- Arriving customer multicast traffic is encapsulated in multicast and flooded over appropriate Default-MDT.



Default MDT – A Closer Look

Cisco.com

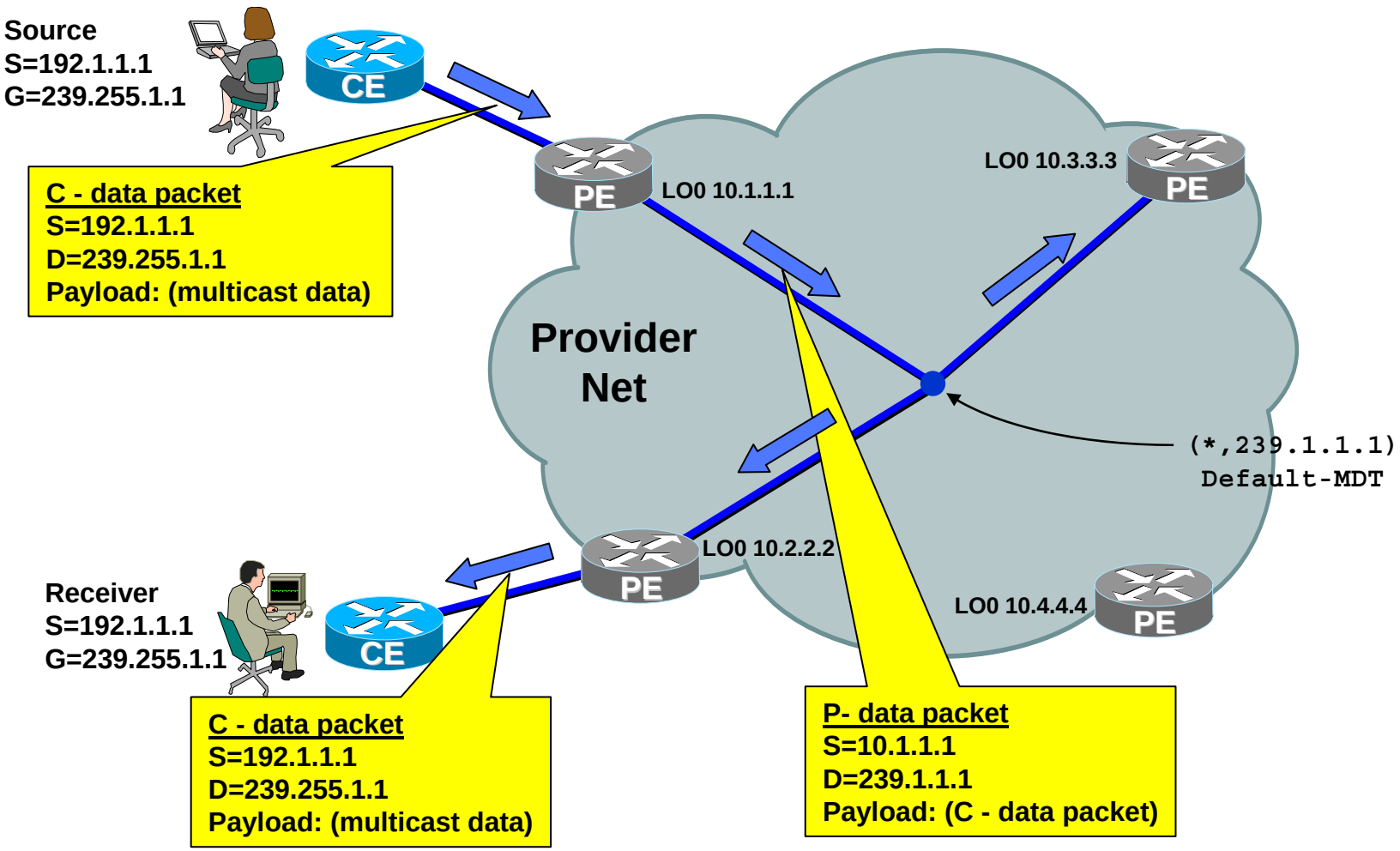
PIM Control Traffic Flow



Default MDT – A Closer Look

Cisco.com

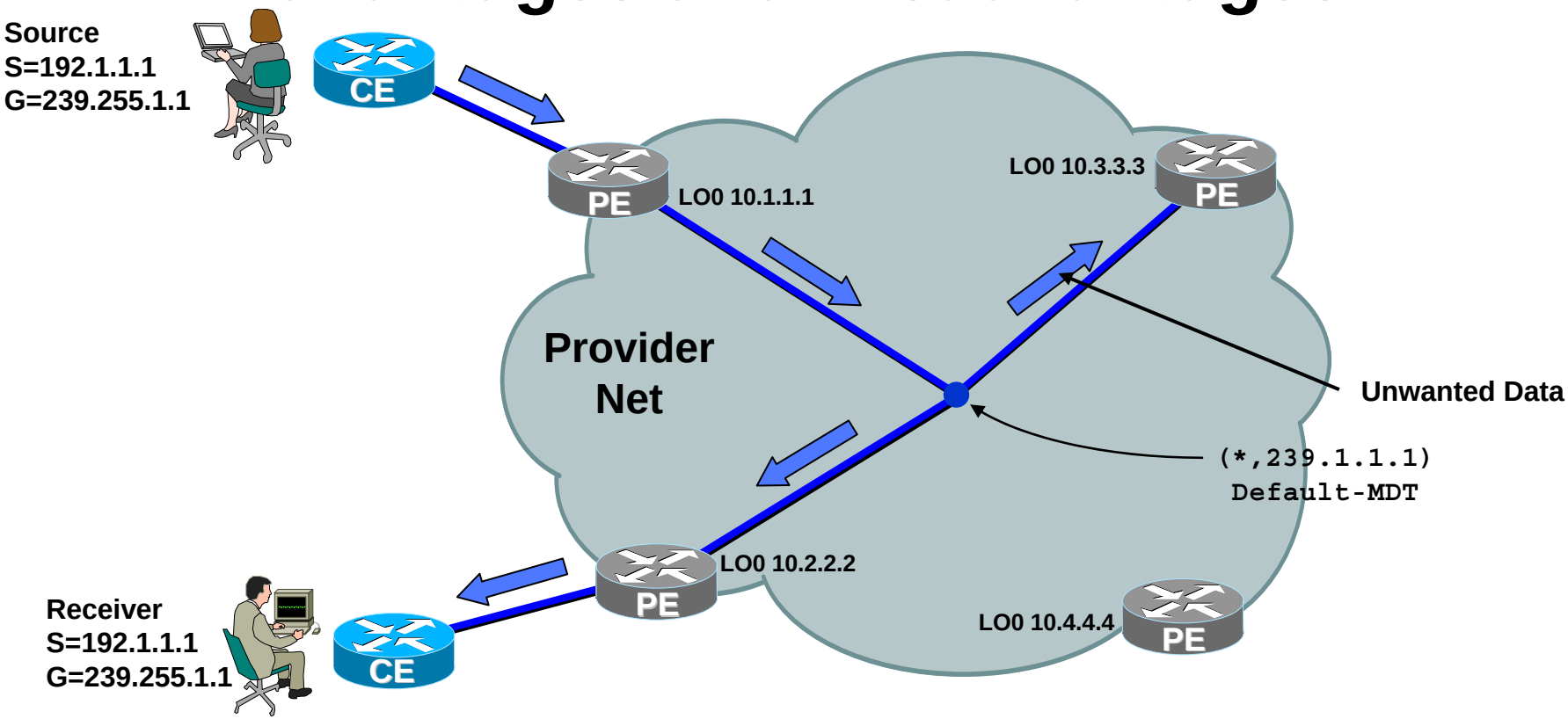
Multicast Data Traffic Flow



Default MDT – A Closer Look

Cisco.com

Advantages and Disadvantages

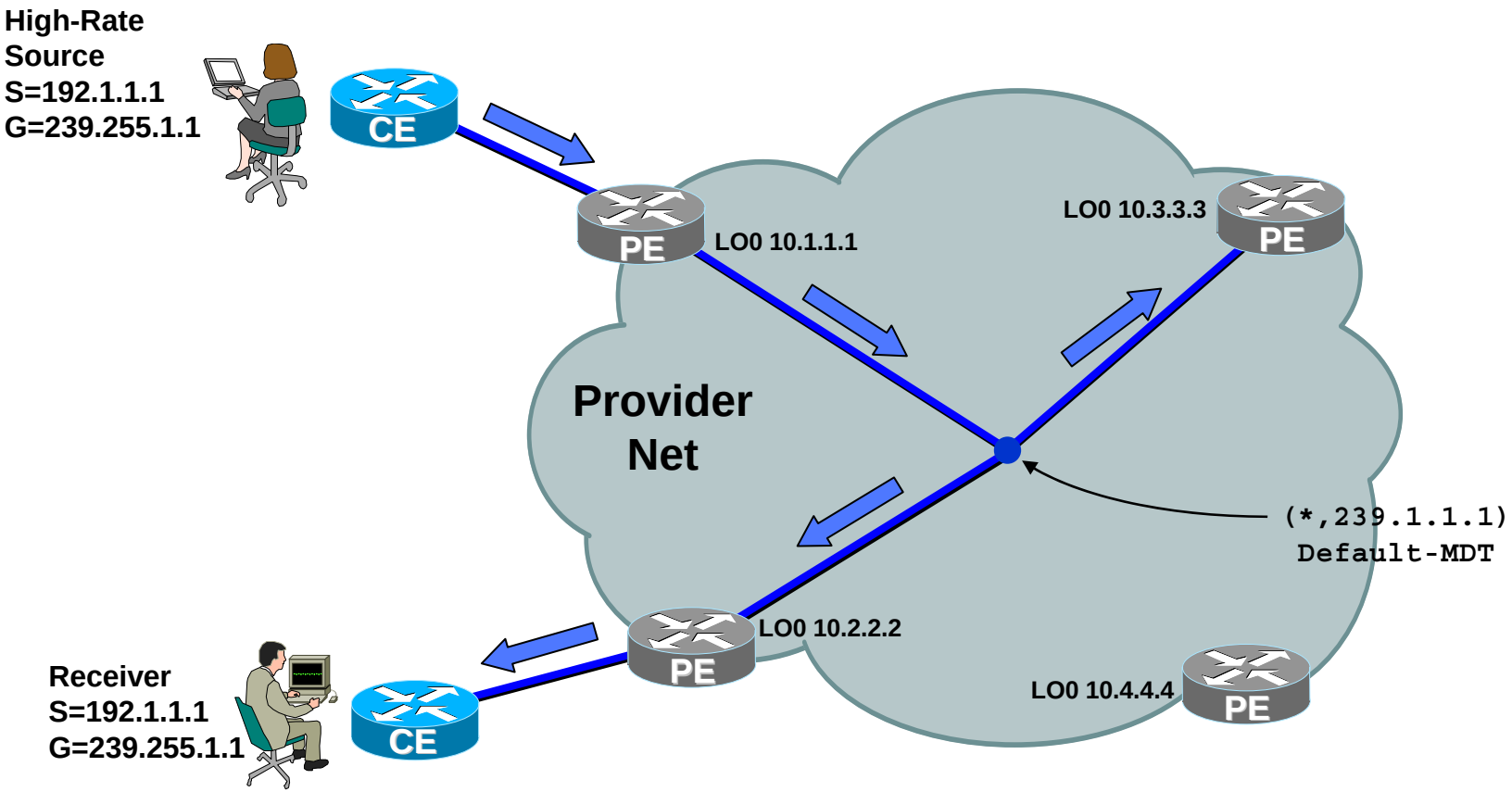


Advantage : Reduces multicast state in the P routers in the core.

Disadvantage : Can result in wasted bandwidth.

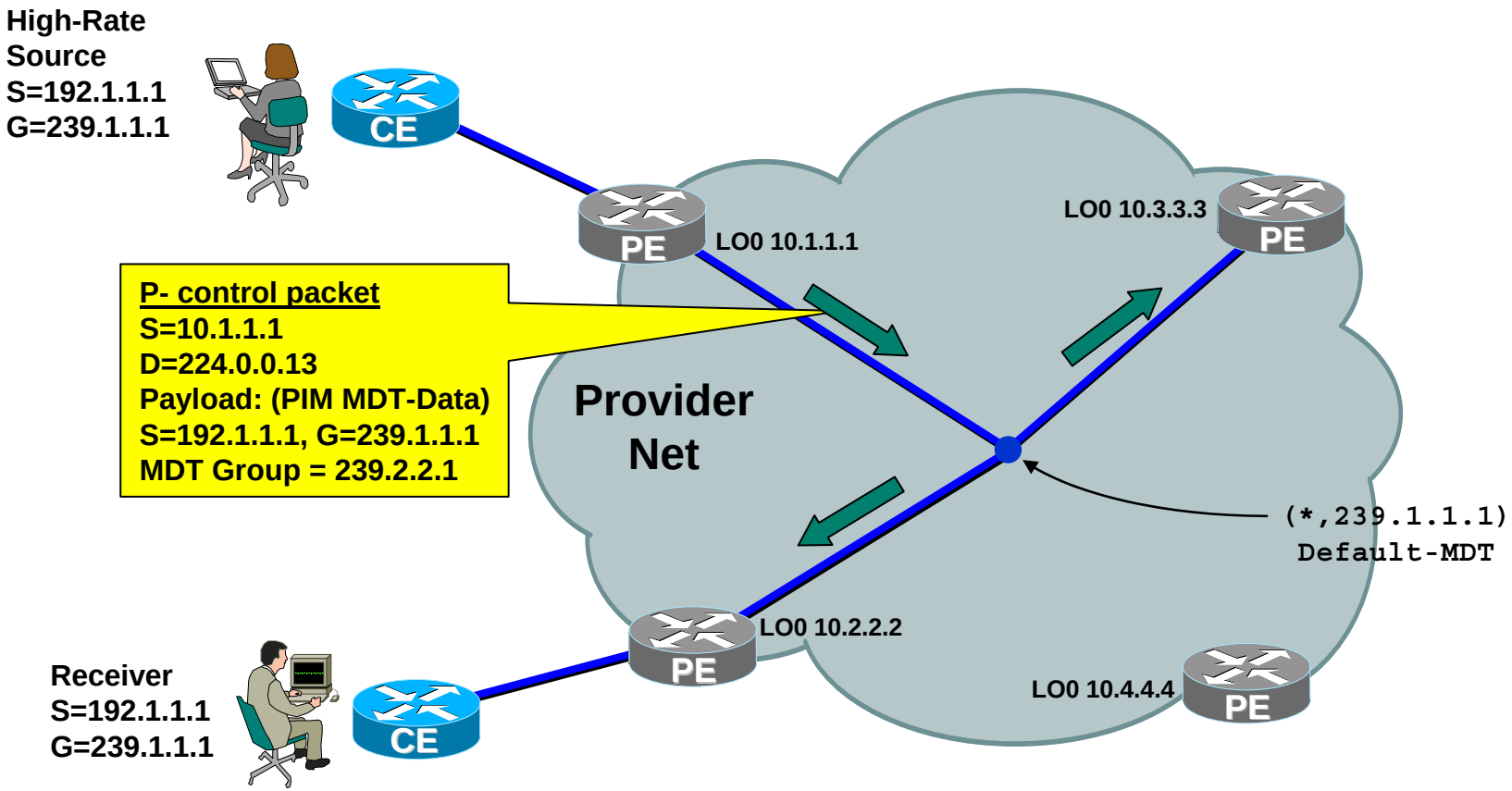
Solution : Use separate Data-MDTs for high rate sources.

Data MDTs – Concepts



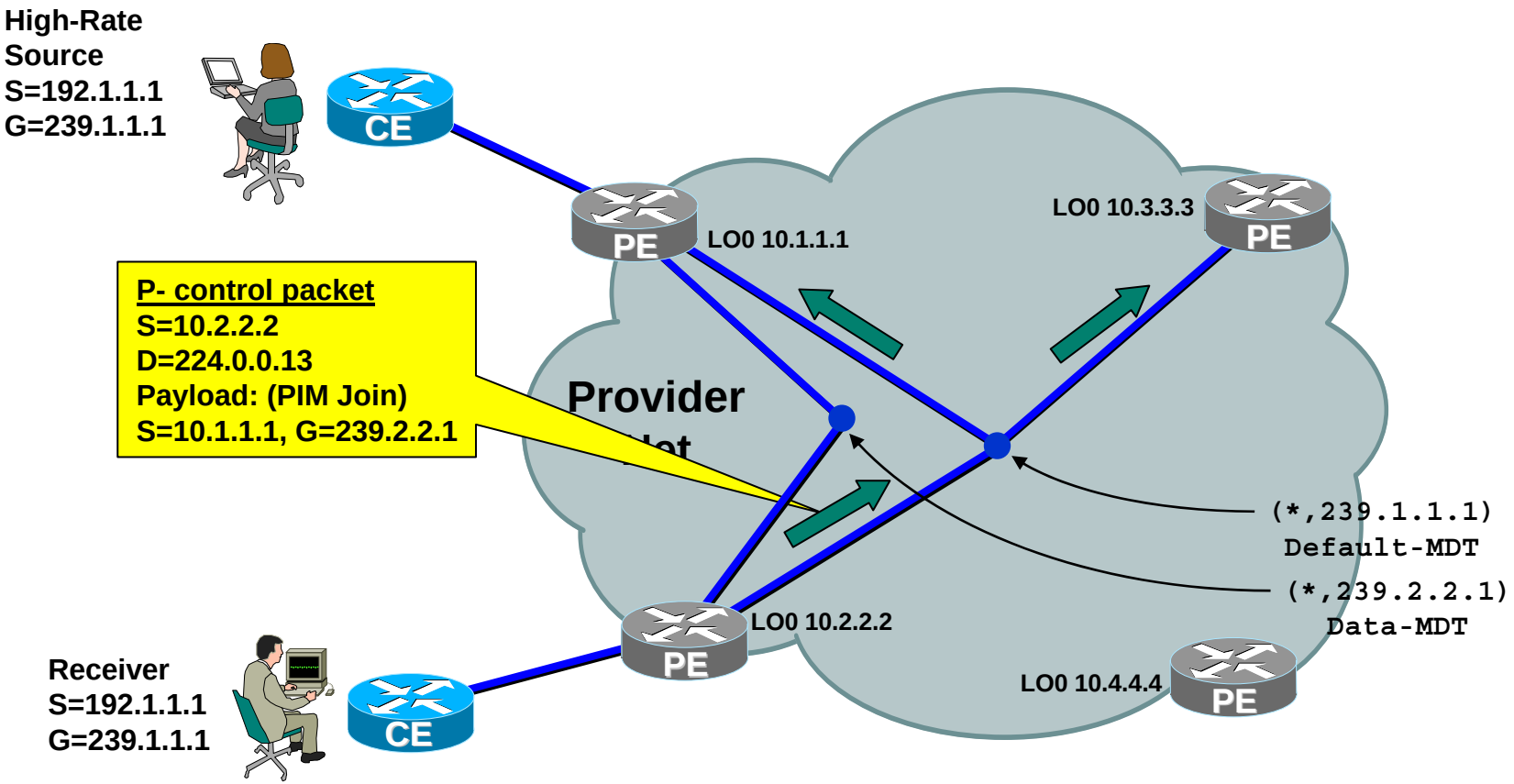
- Traffic exceeds Data-MDT threshold configured on PE router.

Data MDTs – Concepts



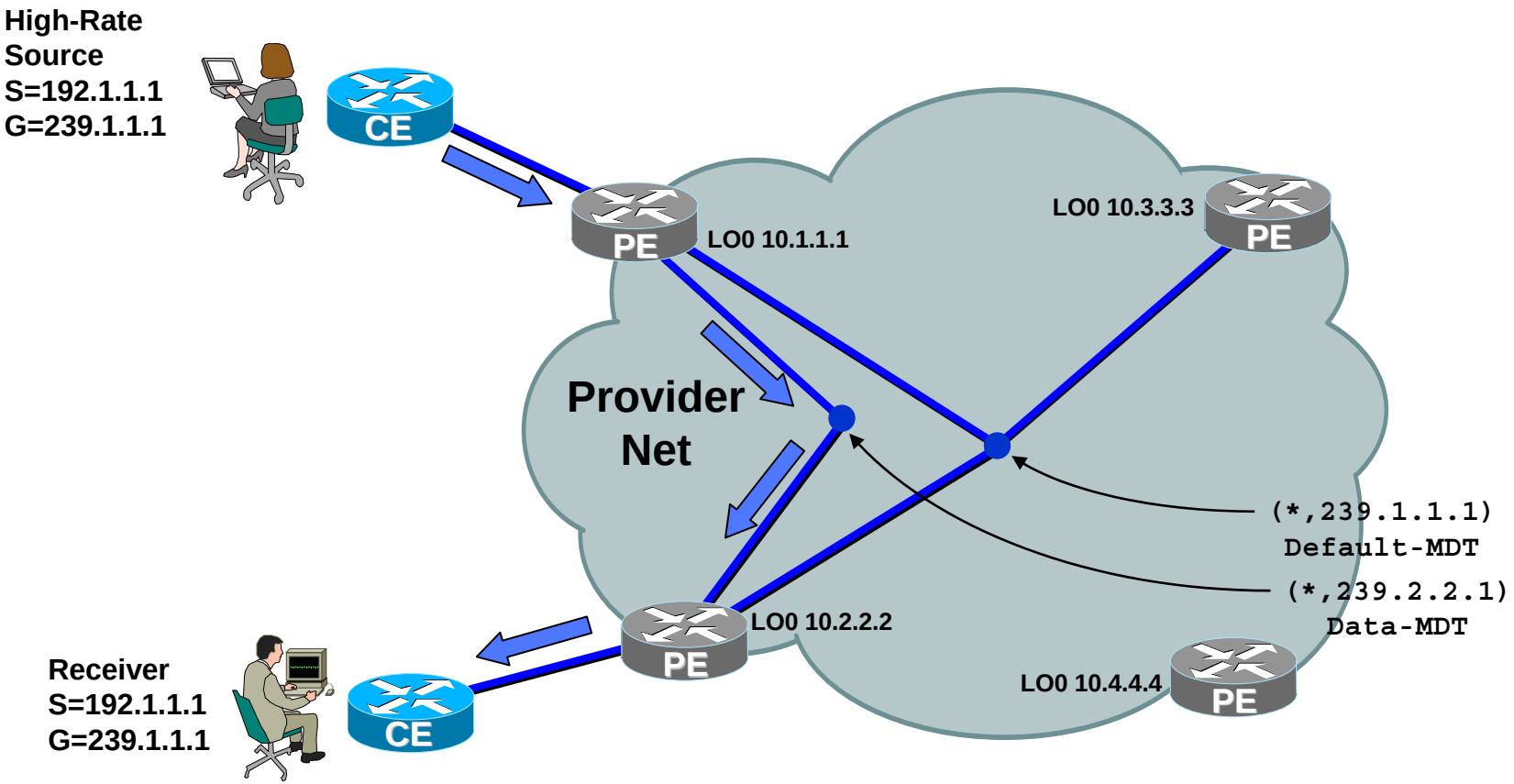
- PE router signals switch to Data-MDT using new group, 239.2.2.1

Data MDTs – Concepts



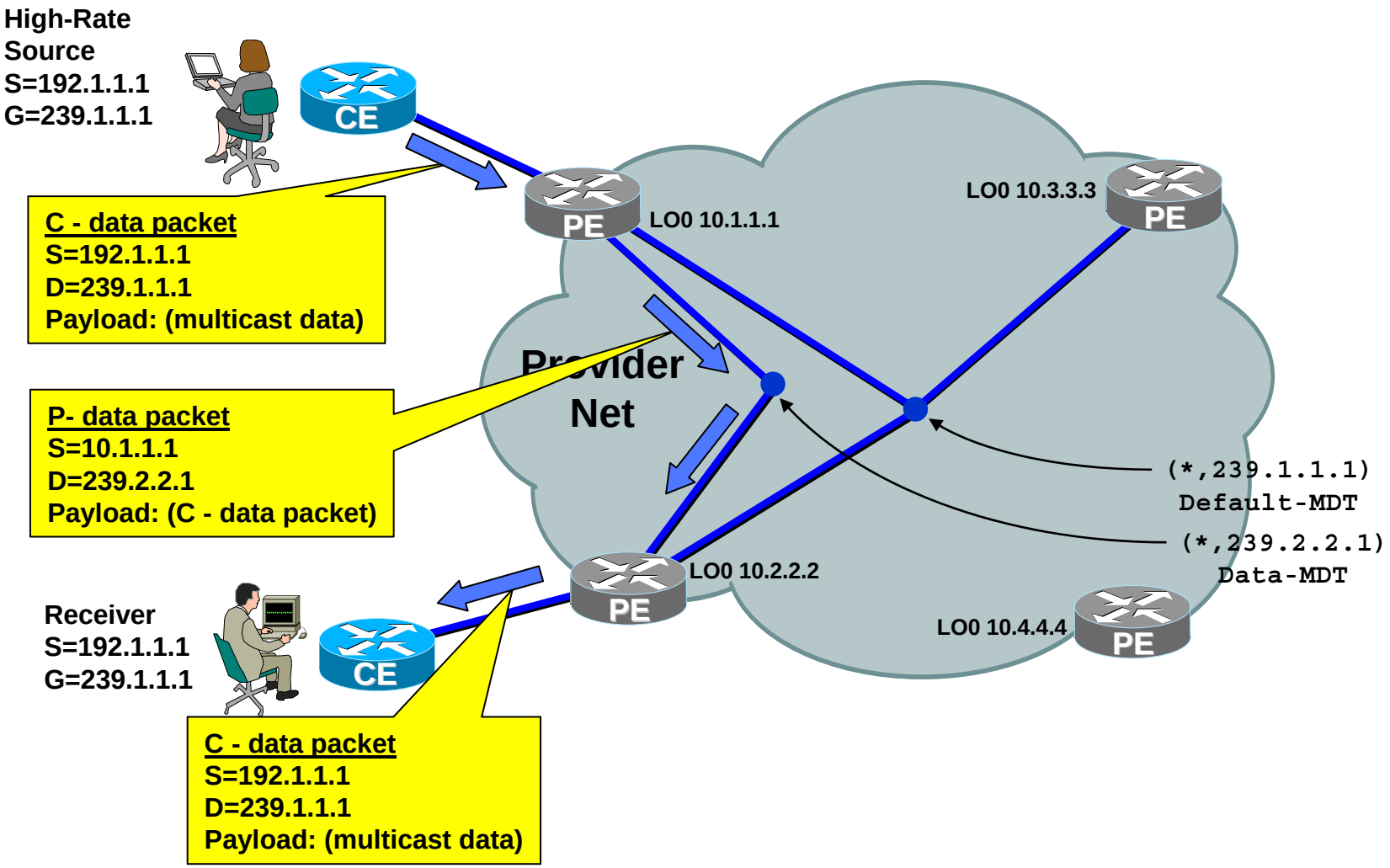
- PE routers with receivers sends Join to group 239.2.2.1.
- Data-MDT is built using group 239.2.2.1.

Data MDTs – Concepts



- High-rate data begins flowing via Data-MDT.
- Data only goes to PE routers that have receivers.

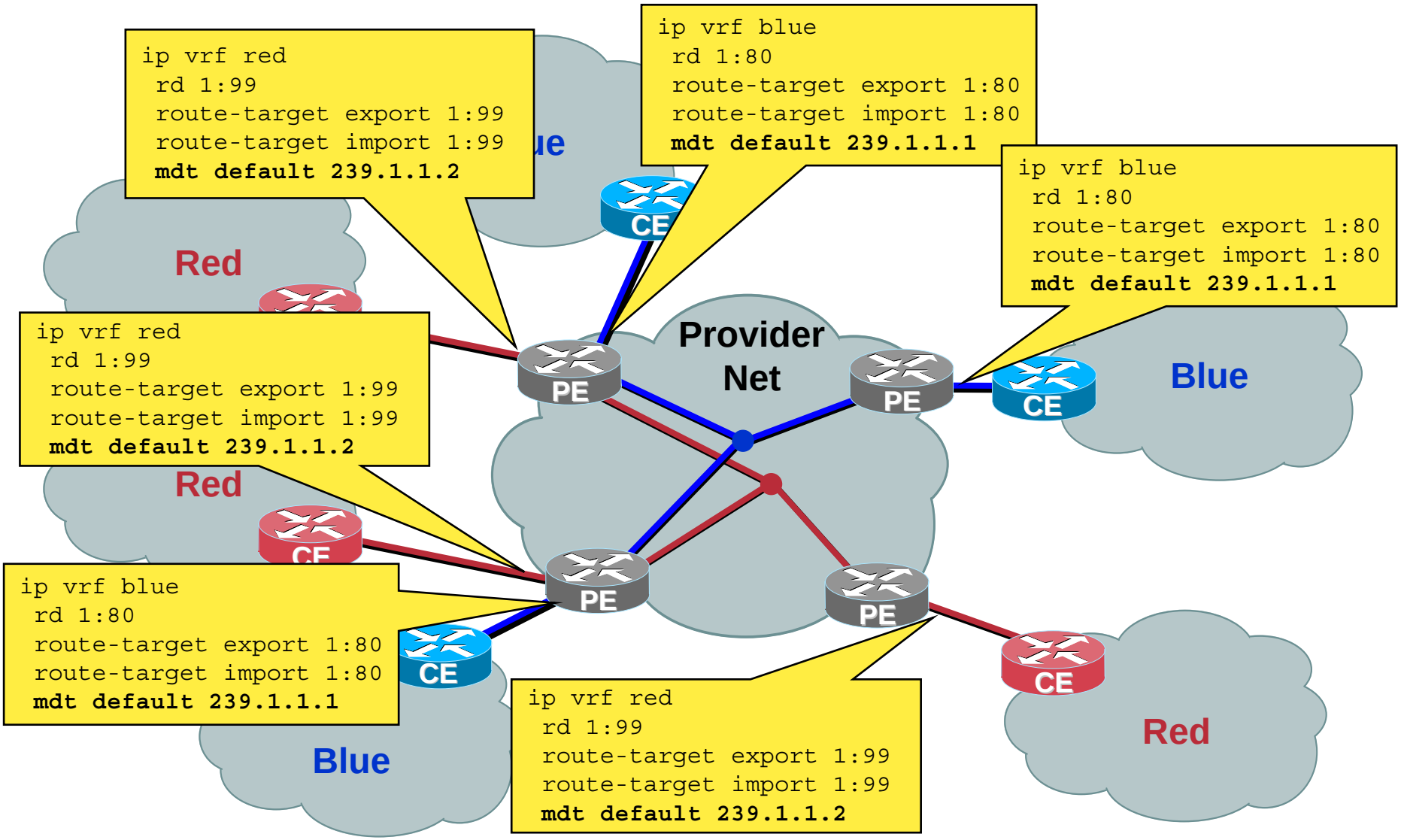
Data MDTs – Concepts



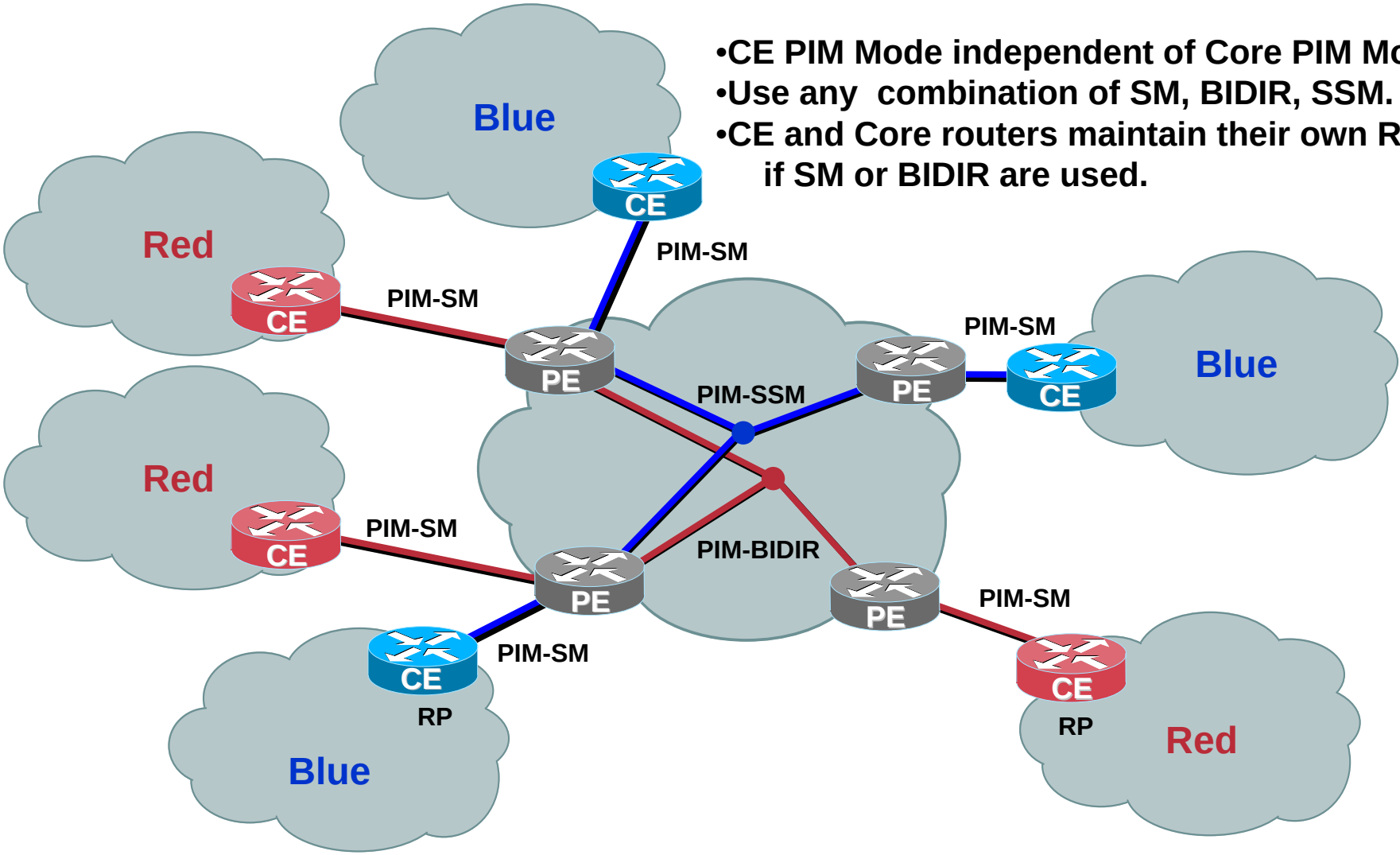
Multicast VPN Addressing

- Use Administratively Scoped Addresses
  - 239.0.0.0 – 239.255.255.255
  - Private address space
    - Similar to RFC1918 unicast addresses
    - Not used for global Internet traffic
    - Used to limit “scope” of multicast traffic
    - Same addresses may be in use at different locations for different multicast sessions

Default-MDT Group Address Example



PIM for MVPN Summary



- CE PIM Mode independent of Core PIM Mode.
- Use any combination of SM, BIDIR, SSM.
- CE and Core routers maintain their own RPs if SM or BIDIR are used.

Agenda

- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

IPv6 Multicast Addressing

- **Multicast Addresses (RFC 2373)**
  - **ff::/8 is the ipv6 equivalent of 224/4**
  - **ff02::1 is the ipv6 equivalent of the link local address 224.0.0.1**
  - **SSM address range ff3X::/32, where X represents the scope bits.**

IPv6 Multicast – O.S. & Application Support

Cisco.com

- Stacks
  - KAME host stack
  - Microsoft Windows XP
  - Mac OS 10.2
  - Linux
  - HP OpenVMS and True64
- Applications
  - Microsoft Media Player & Server 9
  - Many emerging commercial applications
    - DVTS, Videolan, etc
  - Standard MBONE Tools (vic, rat, ...) support IPv6

IPv4 versus IPv6 Multicast

IP Service	IPv4 Solution	IPv6 Solution
Address Range	32-bit, class D	128-bit
Routing	Protocol Independent All IGPs,and BGP4+	Protocol Independent All IGPs,and BGP4+ with v6 mcast SAFI
Forwarding	PIM-DM, PIM-SM, PIM-SSM, PIM-bidir	PIM-SM, PIM-SSM, PIM-bidir
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

IP Routing for Multicast

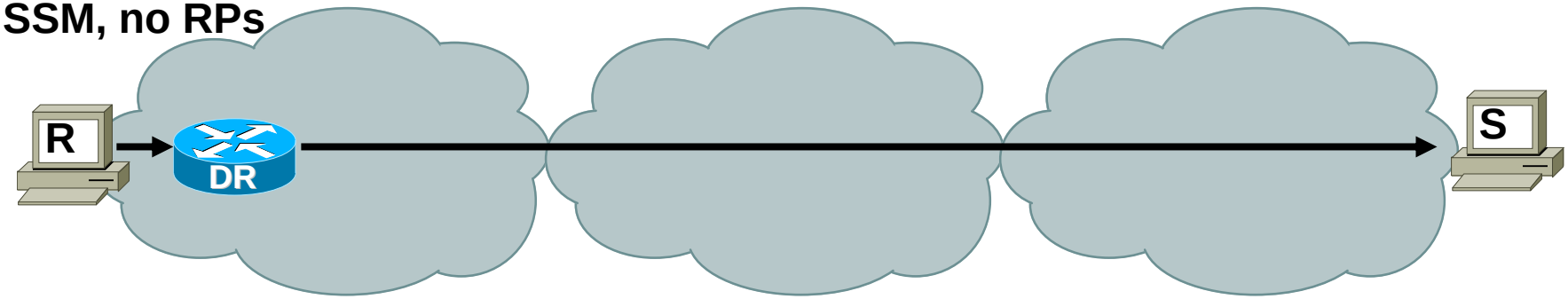
- RPF based on reachability to v6 source same as with v4 multicast
- RPF still protocol independent:
  - Static routes, mroutes
  - Unicast RIB: BGP, ISIS, OSPF, EIGRP, RIP, etc
  - Multi-protocol BGP (mBGP)
    - Support for v6 mcast sub-address family
    - Provide translate function for non-supporting peers

Domain Control

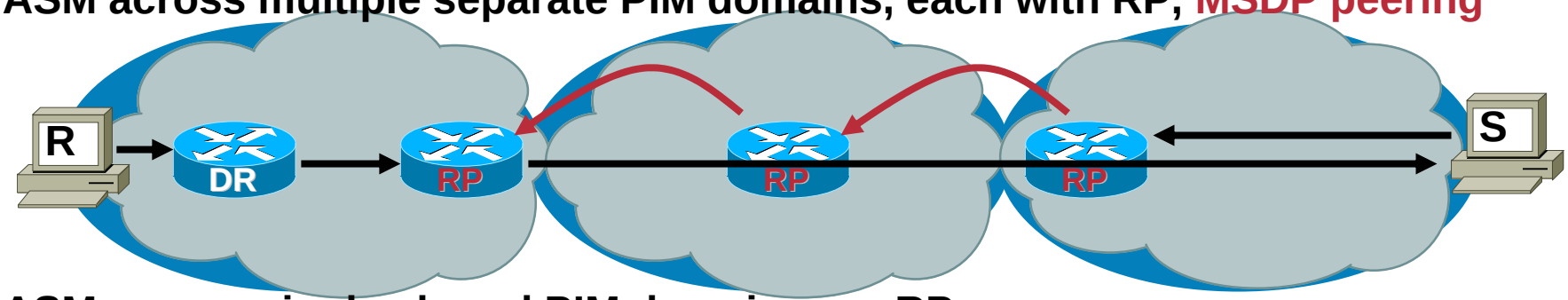
- **Definitions:**
  - **A PIM domain is topology served by common RP for all sources and receivers of same group.**
  - **A routing domain is consistent with AS.**
- **Its necessary to constrain the PIM messages, rp-mappings, and data for groups within the PIM domain:**
  - **In IPv4 we used multicast boundary/ BSR border**
  - **In IPv6 we use scopes and zones**

Interdomain v6 Multicast Options

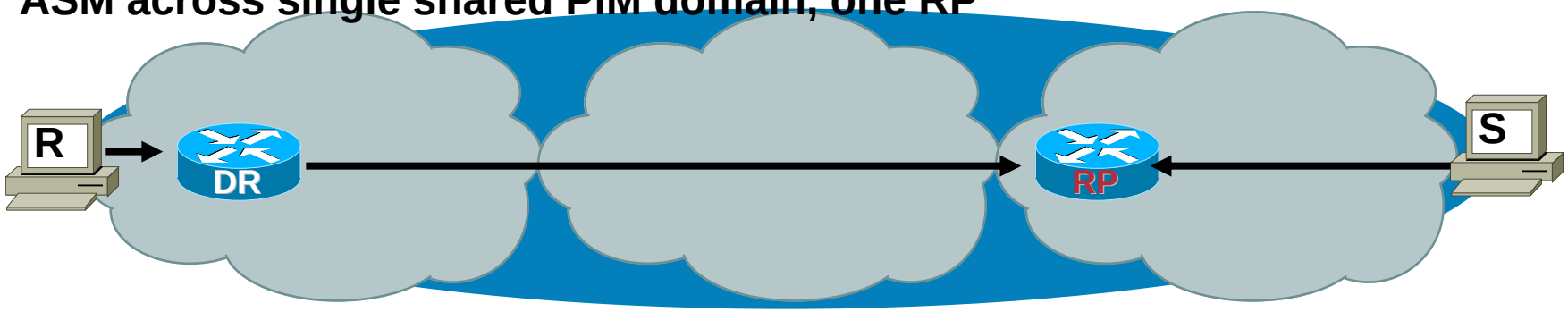
SSM, no RPs



ASM across multiple separate PIM domains, each with RP, MSDP peering



ASM across single shared PIM domain, one RP



Multicast Phase I

12.3(1)T, 12.2S RLS3, 12.0(26)S

Cisco.com

- PIM
  - Source Specific Multicast (PIM-SSM)
  - Sparse-mode (PIM-SM)
    - Full support for DR functionality (registers, etc)
    - Static RP assignment with multiple RP mapping
- Scoping support (replaces v4 boundary function)
- MLDv1 and v2
  - Support INCLUDE and EXCLUDE mode reports in MLDv2
  - Full MLDv1/v2 compatibility
  - Explicit tracking in v2 mode.
- v6-in-v4 tunneling

Multicast Phase II, active

- PIM
  - Support for embedded RP mapping
  - Support for “upstream-address-detection” PIM Hello Option
- Multicast specific Routing – mBGP
  - ‘translate update’ for seamless migration into existing BGP peerings
    - (also in 12.0(26)S for initial GSR release)
  - static mroutes
    - (also in 12.0(26)S for initial GSR release)

Multicast Phase II, active

- BSR Forwarding support for BSR messages
  - (also in 12.0(26)S for initial GSR release)
- Security and access-control
  - MLD access-groups for receiver control
  - Register filters for source control
- Distributed Fast Switching
  - (also in 12.0(26)S for initial GSR release)
- v6-in-v6 tunneling

IOS CLI: Configure like v4, but easier

Group mode determines how to forward, compared to interface mode in v4.

By default all interfaces are PIM enabled unless explicitly disabled.

Config for PIM-SSM:

```
!
ipv6 multicast-routing
!
```

Config for PIM-SM:

```
!
ipv6 multicast-routing
ipv6 pim rp-address <v6_address>
!
```

Config for PIM-bidir:

```
!
ipv6 multicast-routing
ipv6 pim rp-address <v6-address> bidir
!
```

Disable PIM on an interface

```
!
interface ethernet 0
 no ipv6 pim
!
```

Conclusion

- Cisco IOS IPv6 Multicast in initial deployment now
- Multicast Applications can be developed and tested over an infrastructure running Cisco IOS IPv6 Multicast
- IPv6 Multicast is an IPv6 service fully integrated with other Cisco IPv6 solutions

v6 Mcast Questions?:

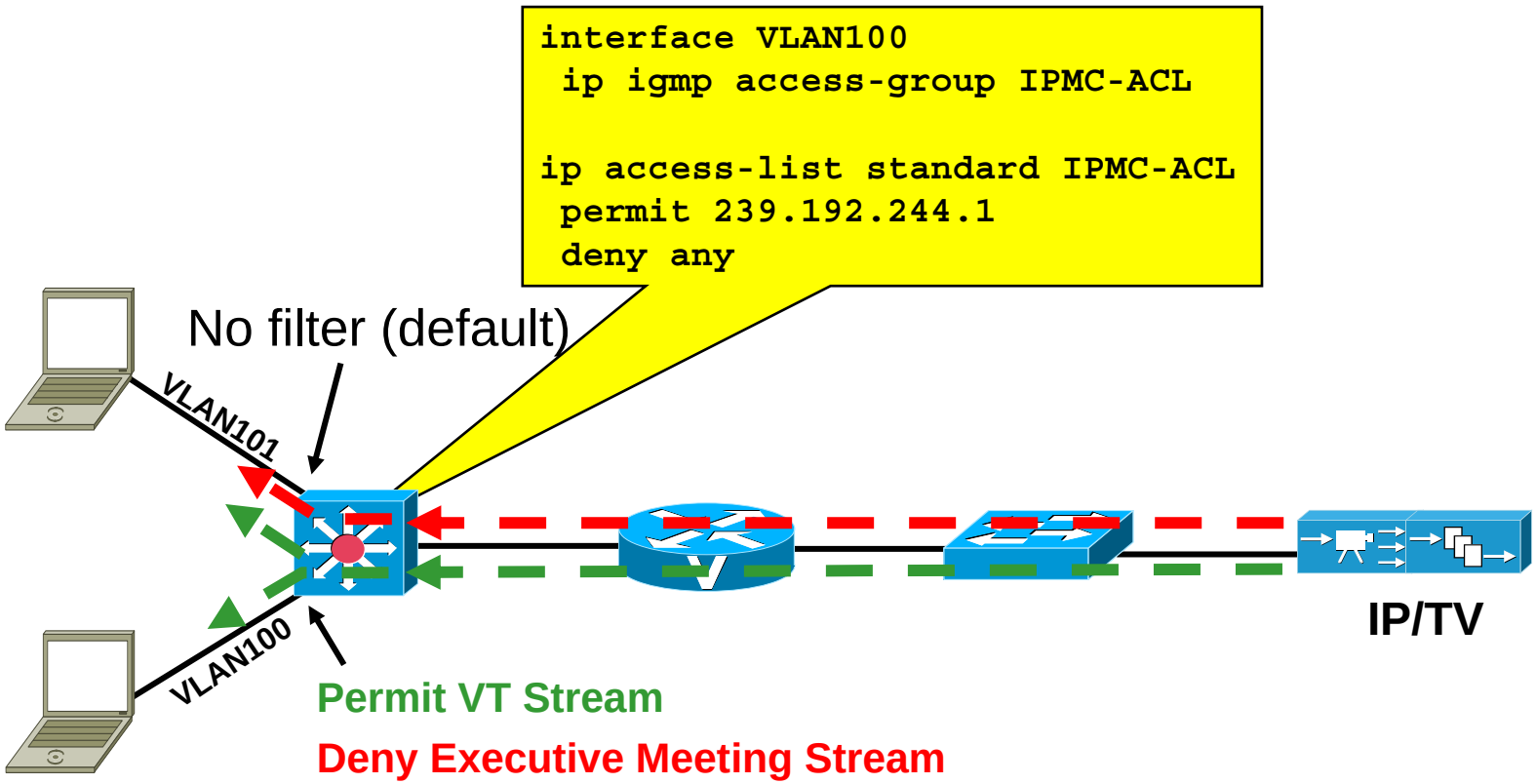
mcast-v6-support@cisco.com

Agenda

- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)
- IPv6
- Security

Controlling Receivers

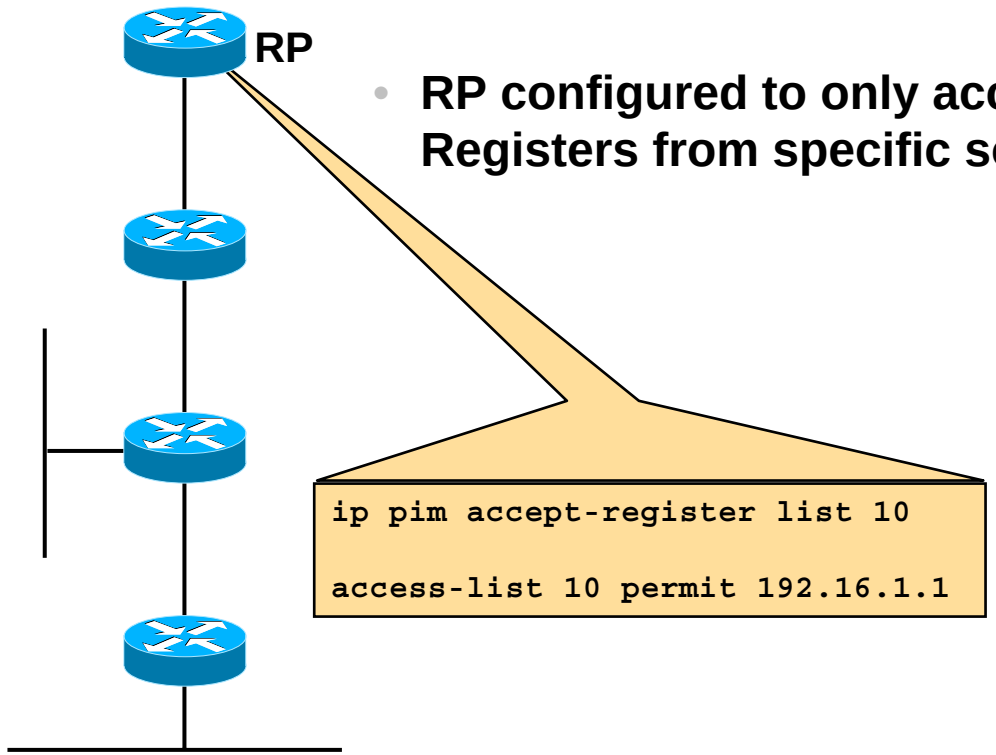
IGMP Access-Group Approach



Controlling Source Registration

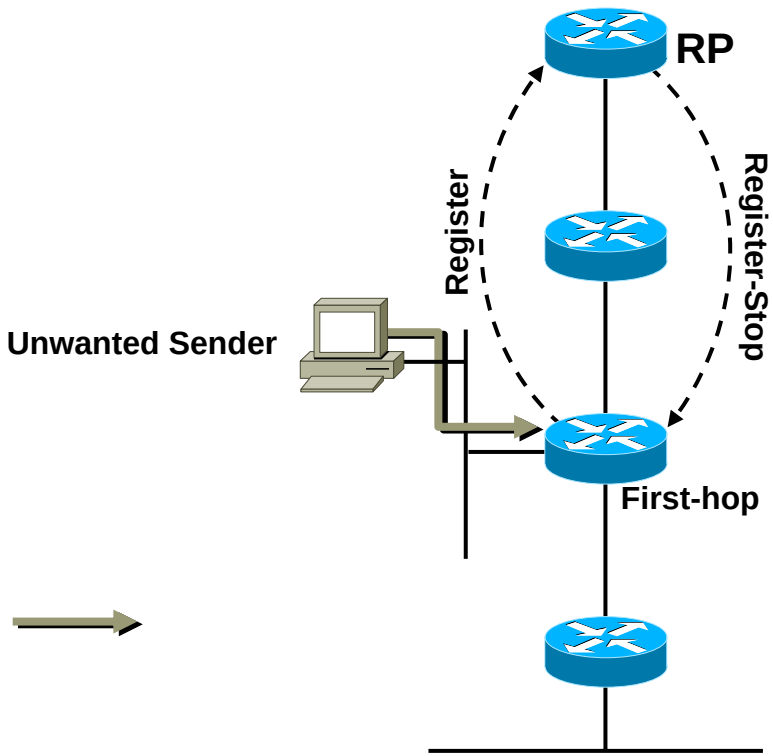
- Global command
  - ip pim accept-register [list <acl>] | [route-map <map>]
  - Used on RP to filter incoming Register messages
  - Filter on Source address alone (Simple ACL)
  - Filter on (S, G) pair (Extended ACL)
  - May use route-map to specify what to filter
    - Filter by AS-PATH if (m)BGP is in use.
- Helps prevents unwanted sources from sending
  - First hop router blocks traffic from reaching net

Controlling Source Registration



- RP configured to only accept Registers from specific source.

Controlling Source Registration



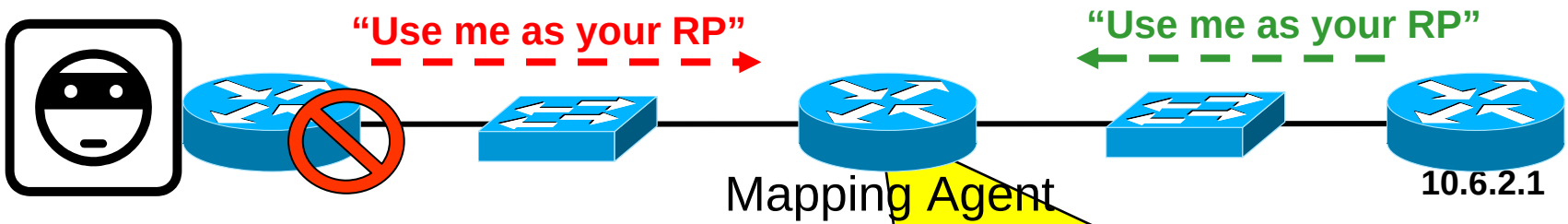
- Unwanted source traffic hits first-hop router.
- First-hop router creates (S,G) state and sends Register.
- RP rejects Register, sends back a Register-Stop.

Preventing RP-Spoofing DoS Attacks

- **Global command**
  - **ip pim rp-announce-filter rp-list <acl> [group-list <acl>]**
    - **rp-list <acl>**
      - **Specifies from which routers C-RP Announcements are accepted.**
    - **group-list <acl>**
      - **Specifies which groups in the C-RP Announcement are accepted.**
      - **If not specified, defaults to deny all groups**
- **Use on Mapping Agents to filter out bogus C-RP's**
  - **Some protection from RP-Spoofing denial-of-service attacks**
  - **Multiple commands may be configured as needed**

Preventing RP-Spoofing DoS Attacks

Use ip pim rp-announce-filter on RP




```
ip pim rp-announce-filter rp-list 11 group-list 12
access-list 11 permit 10.6.2.1                    !IP address of Permitted RP
access-list 12 permit 239.192.240.0 0.0.3.255     !Permit MoH
access-list 12 permit 239.192.244.0 0.0.3.255     !Permit Low Stream
access-list 12 permit 239.192.248.0 0.0.3.255     !Permit Medium Stream
access-list 12 permit 239.255.0.0 0.0.255.255     !Permit High Stream
access-list 12 deny 239.0.0.0 0.255.255.255       !Deny remaining Admin. Scoped range
access-list 12 permit 224.0.0.0 15.255.255.255    !Permit Link Local/Reserved Addr.
```

- Using this configuration allows the router to accept RP announcements from only the RP in 'access-list 11' for group ranges described in 'access-list 12'

Other Commands To Protect

- **ip multicast route-limit <routes>**

  Use this command to limit the impact of Denial of Service attacks based on creating useless IP multicast routing state.

- **ip pim rp-address <ip-address> [<group-access-list>] [override]**

  A single RP will be used only for certain defined groups.

- **ip pim accept-rp {<address> | auto-rp} [<acl>]**

  To configure a router to accept Joins or Prunes destined for a specified RP and for a specific list of groups.

- **ip msdp sa-filter in|out <ip-address-or-name> [list <acl>]**

  Filters incoming/outgoing SA messages to/from a peer.

- **ip msdp sa-limit <peer-address-or-name> <limit>**

  Introduced as a mean of protection against (distributed) denial of service attacks. Limits the overall number of SA messages the router will accept from a peer.

Other Commands To Protect

- **ip pim neighbor-filter <acl>**

  Used to administratively deny a misconfigured PIM neighbor from participating in PIM

- **ip pim bsr-border**

  Bootstrap messages will not be able to pass through this border in either direction.

- **ip multicast boundary <acl>**

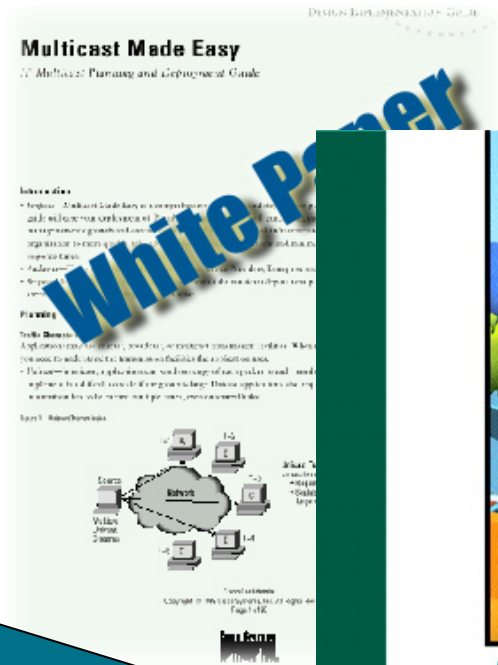
  No multicast data packets, defined in acl, will be allowed to flow across the boundary from either direction. For example, to configure a boundary for all administratively scoped addresses, do:

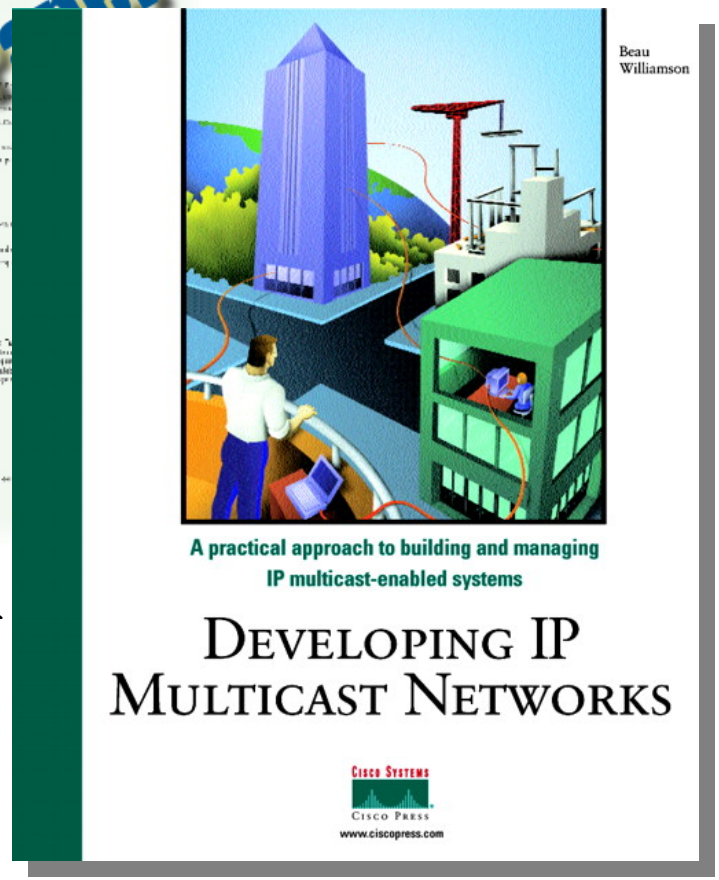
```
access-list 1 deny 239.0.0.0 0.255.255.255
access-list 1 permit 224.0.0.0 15.255.255.255
interface ethernet 0
ip multicast boundary 1
```

More Information

- White Papers
- Web and Mailers
- Cisco Press

RTAB





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

RTAB = "Read the Amazing Book"

Advanced IP Multicast

Session RST-4051

Please Complete Your Evaluation Form

Session RST-4051