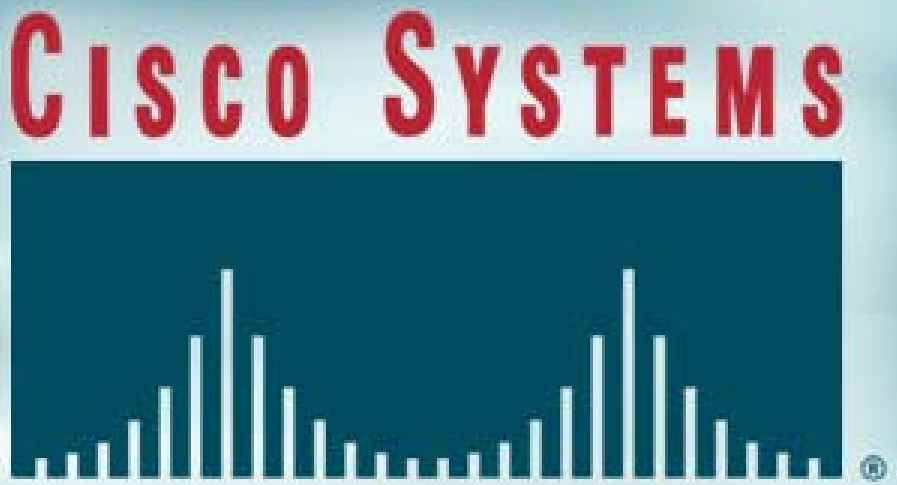
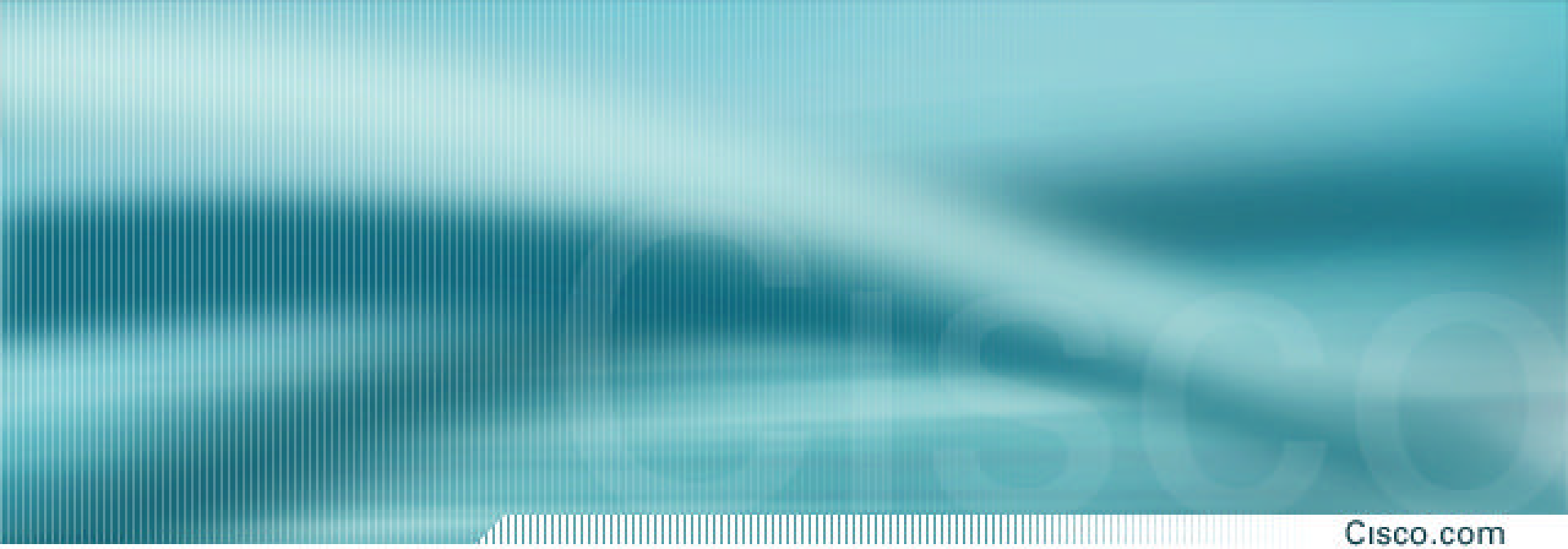

CISCO SYSTEMS



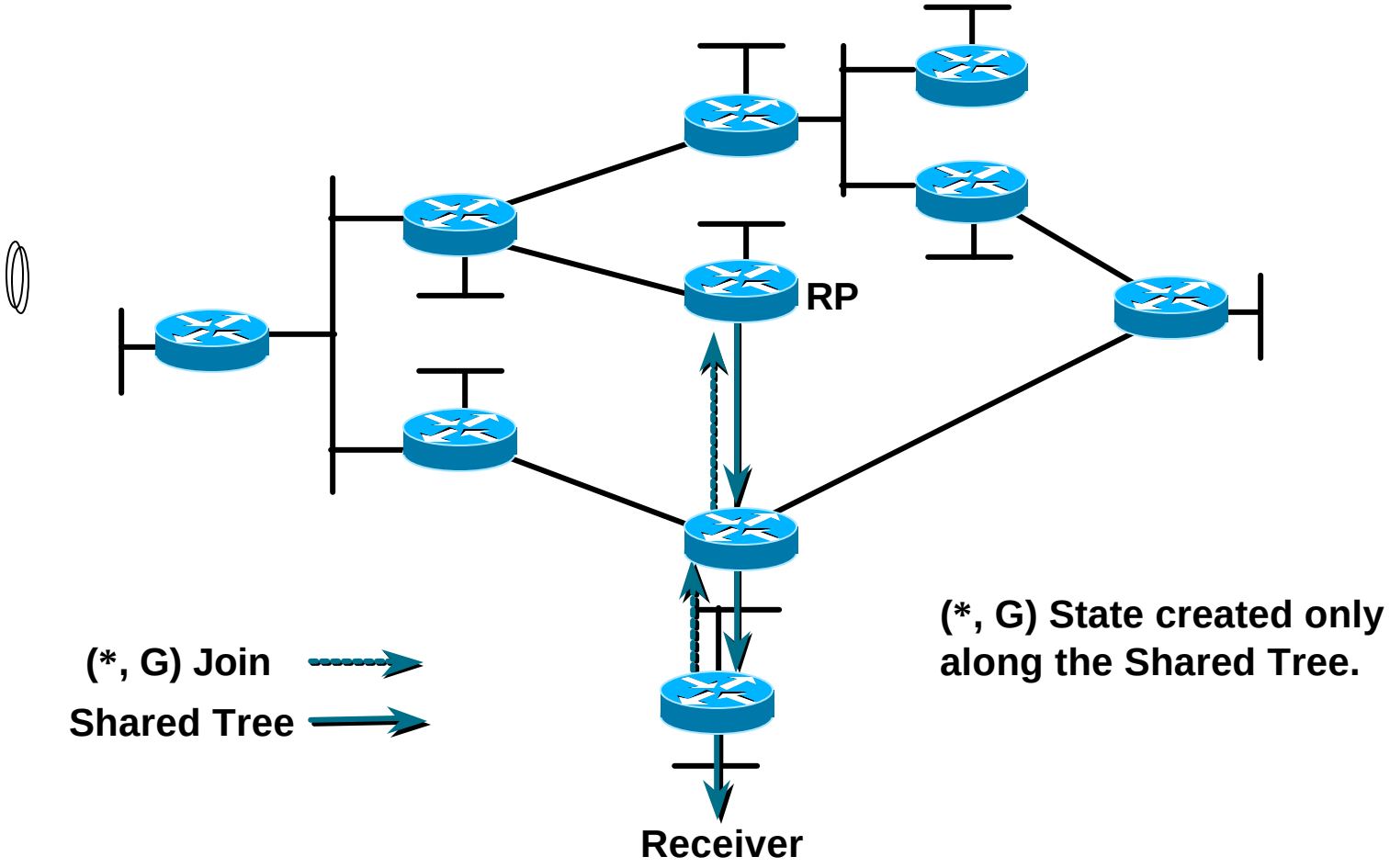
Deploying Inter-Domain IP Multicast

Session RST-261

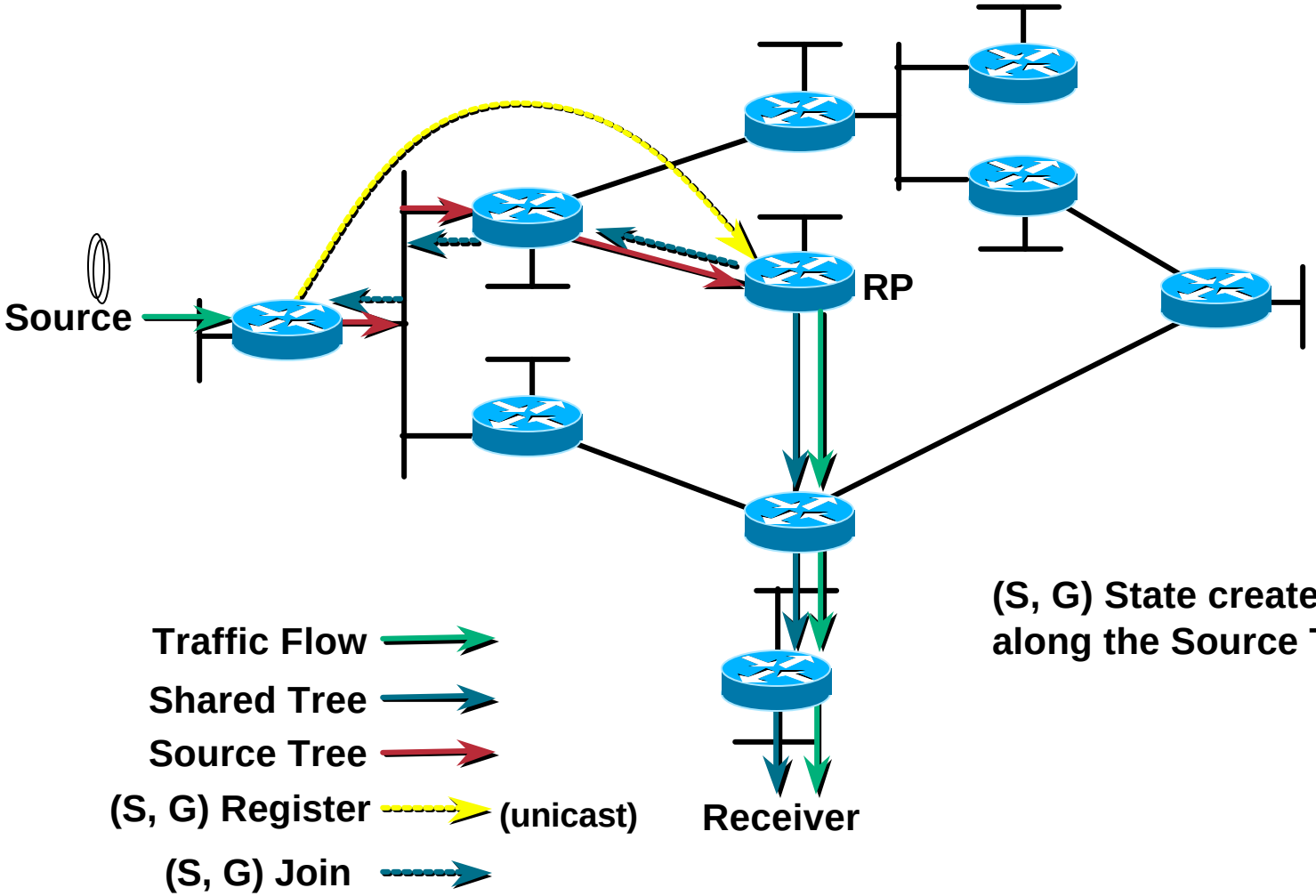
Agenda

- **PIM-SM review (forwarding)**
- **MBGP (routing)**
- **MSDP (source discovery)**
- **MBGP/MSDP Examples**
- **SSM (Source Specific Multicast)**
- **MVPN (Multicast VPN)**

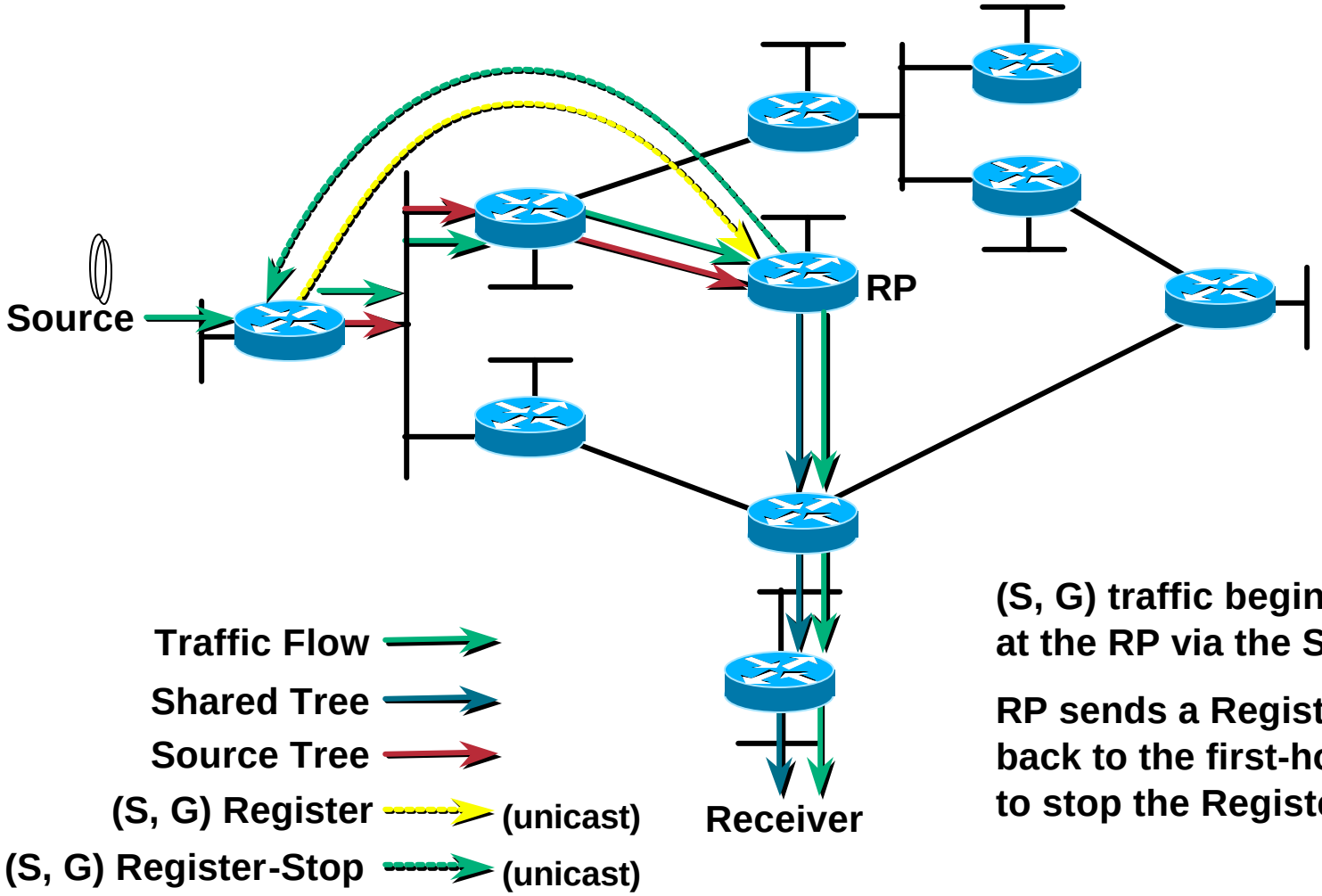
PIM-SM Shared Tree Join



PIM-SM Sender Registration



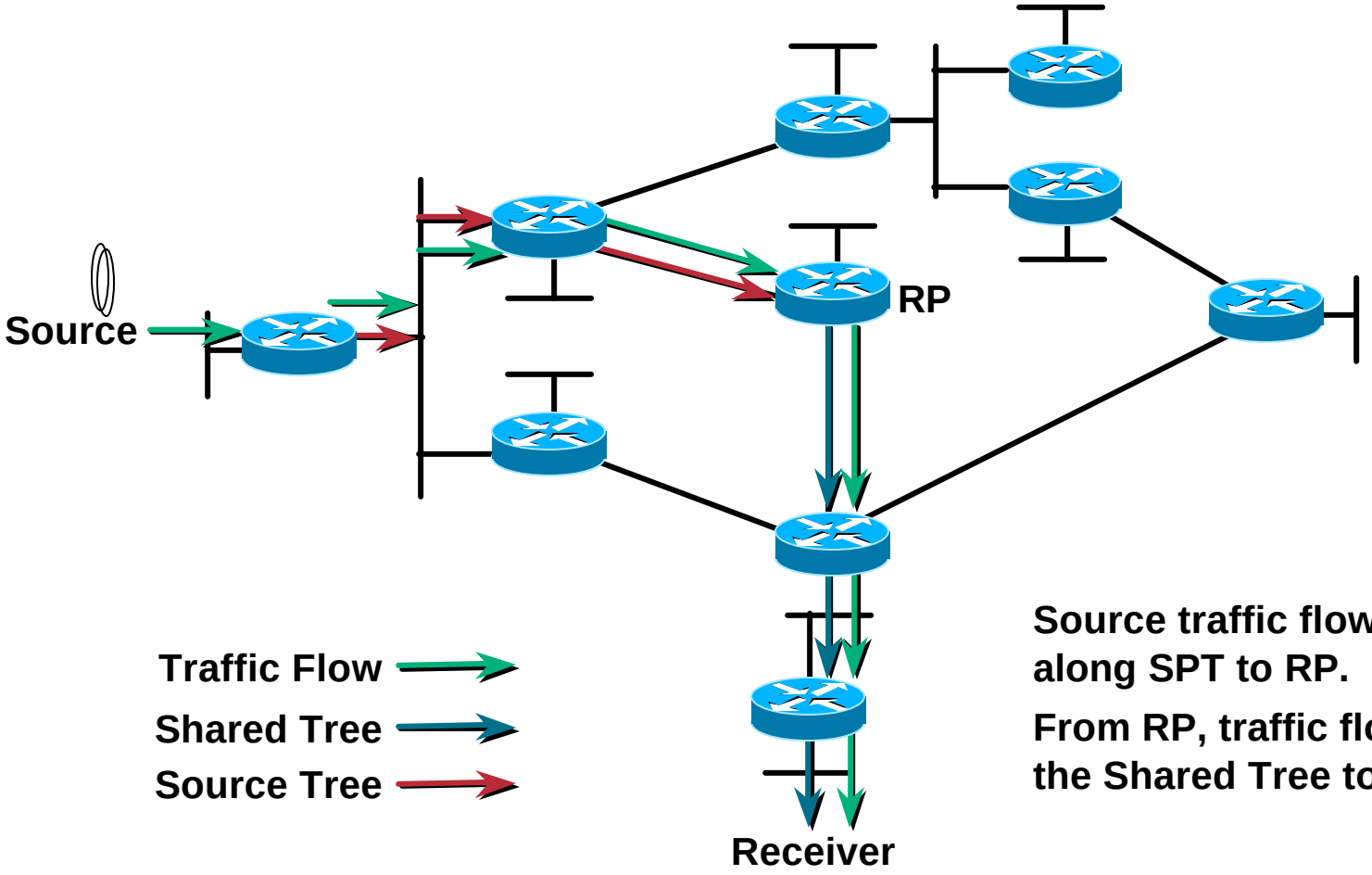
PIM-SM Sender Registration



(S, G) traffic begins arriving at the RP via the Source tree.

RP sends a Register-Stop back to the first-hop router to stop the Register process.

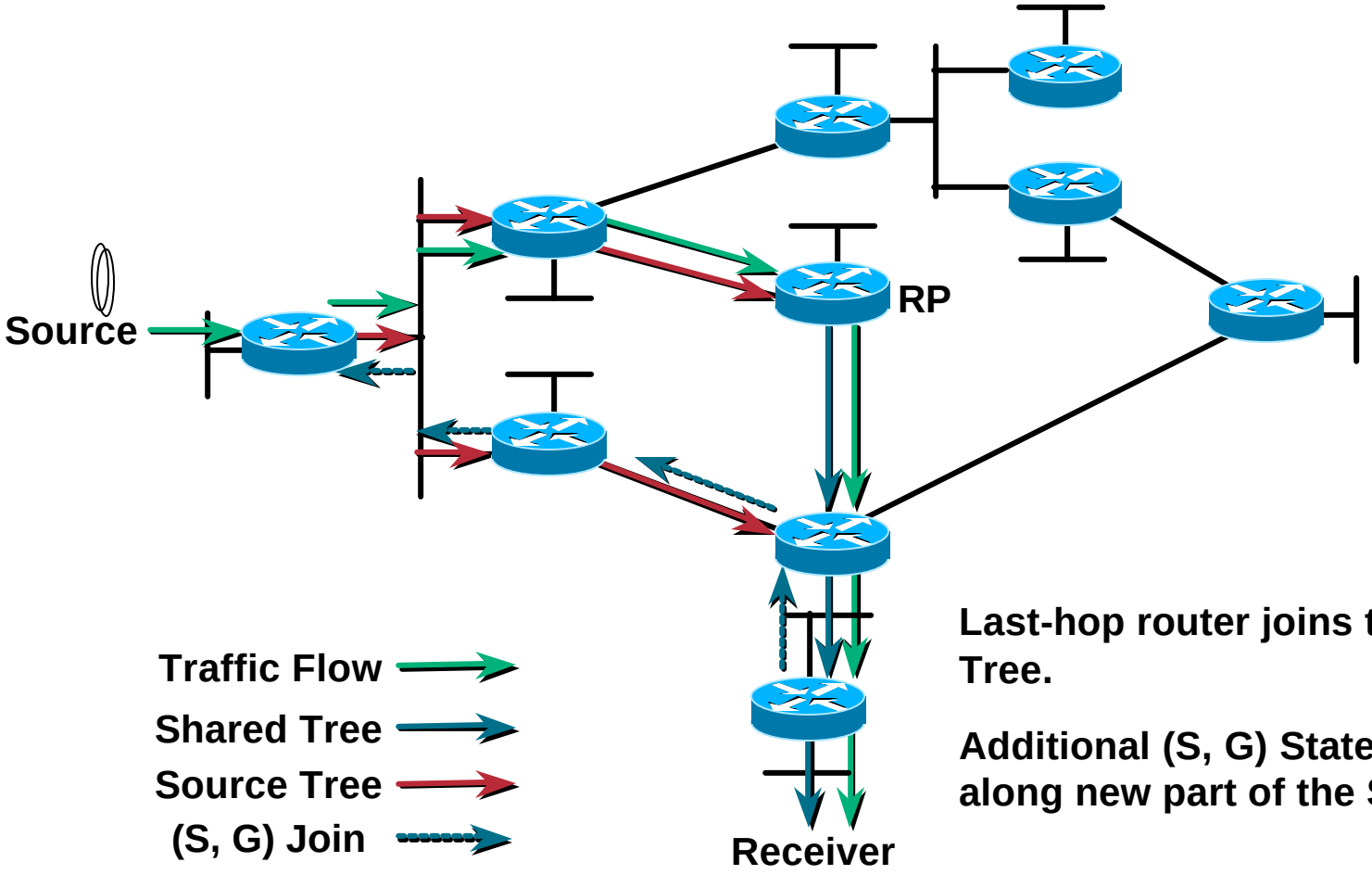
PIM-SM Sender Registration



Source traffic flows natively along SPT to RP.

From RP, traffic flows down the Shared Tree to Receivers.

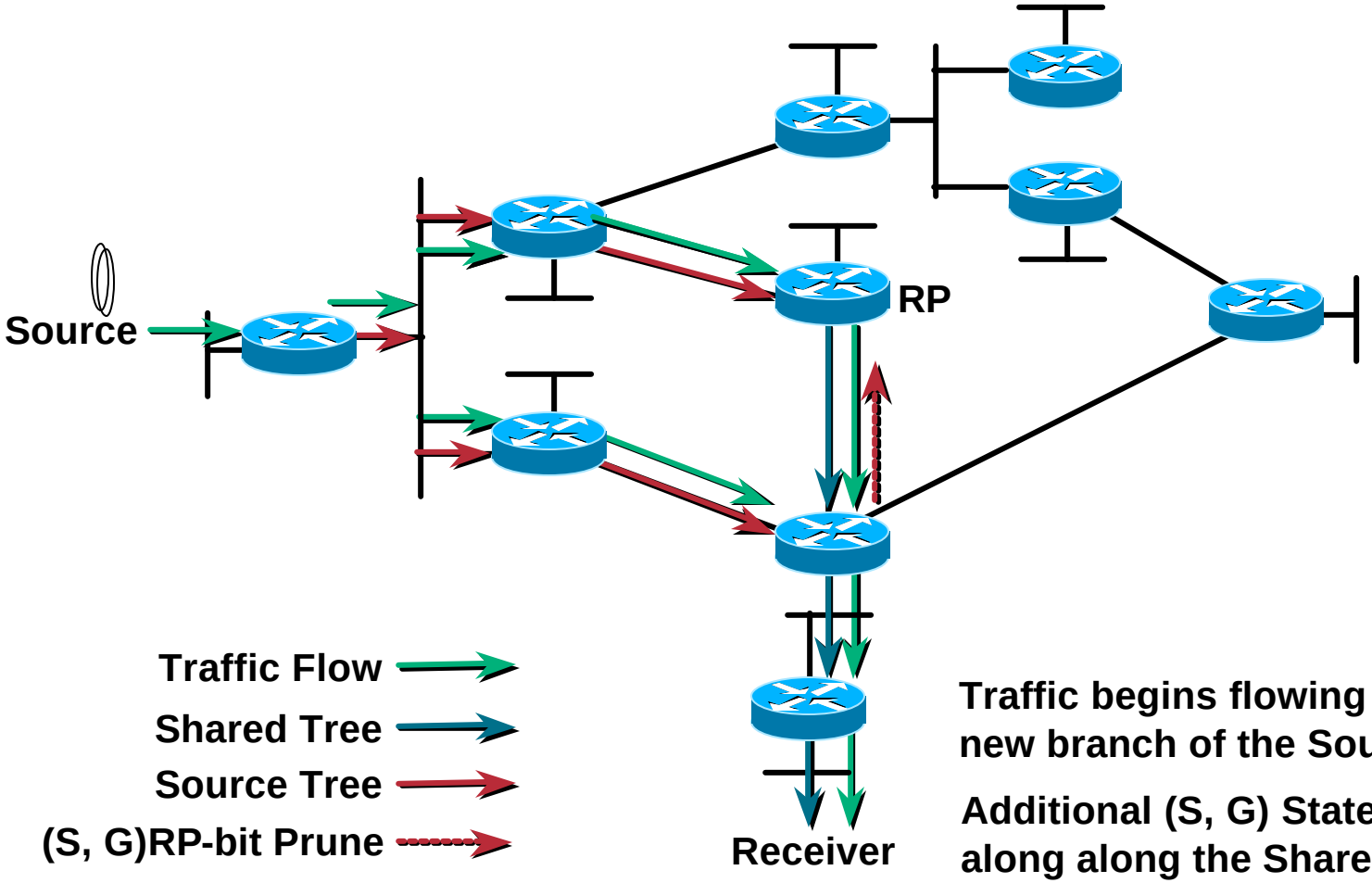
PIM-SM SPT Switchover



Last-hop router joins the Source Tree.

Additional (S, G) State is created along new part of the Source Tree.

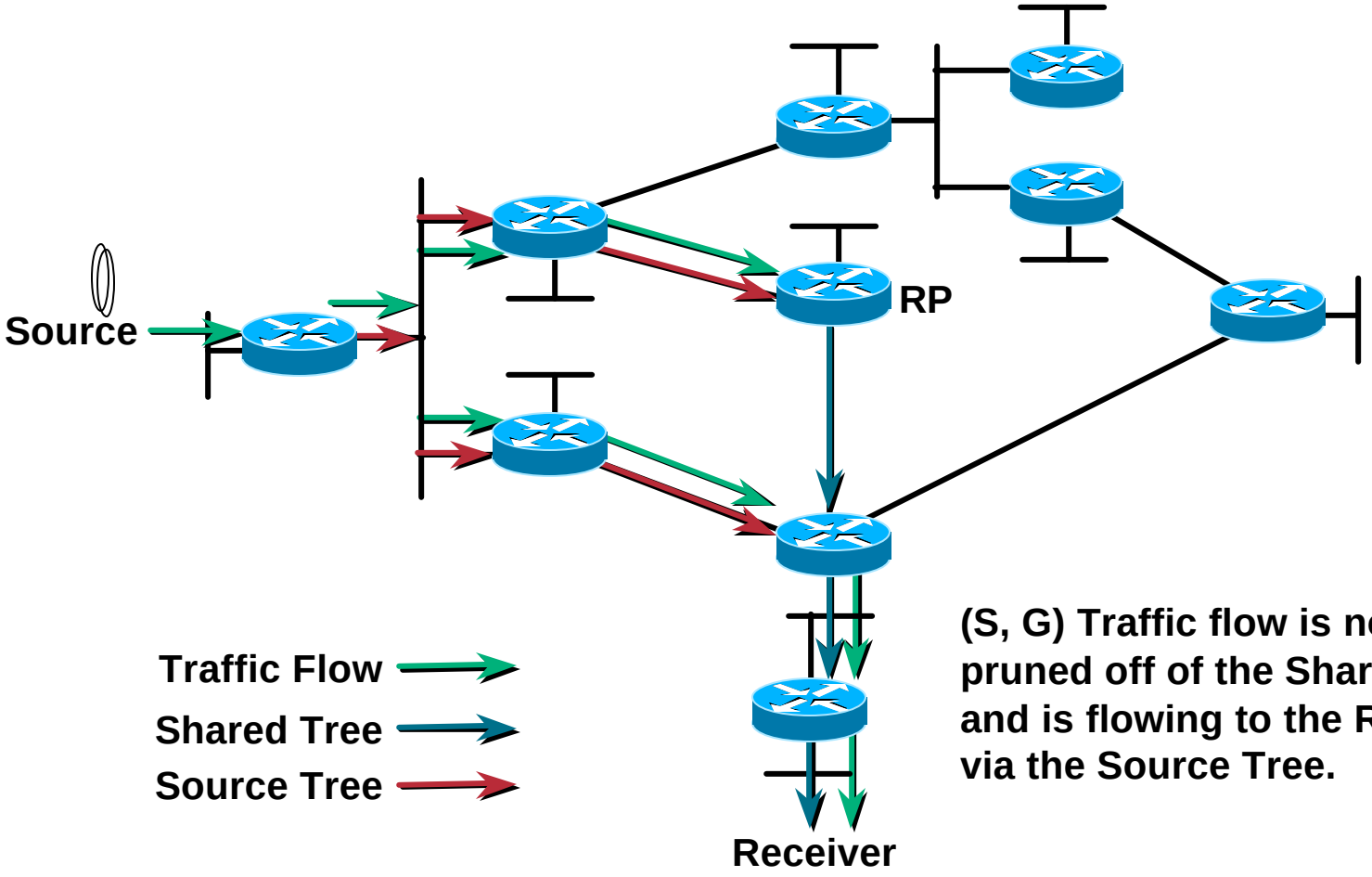
PIM-SM SPT Switchover



Traffic begins flowing down the new branch of the Source Tree.

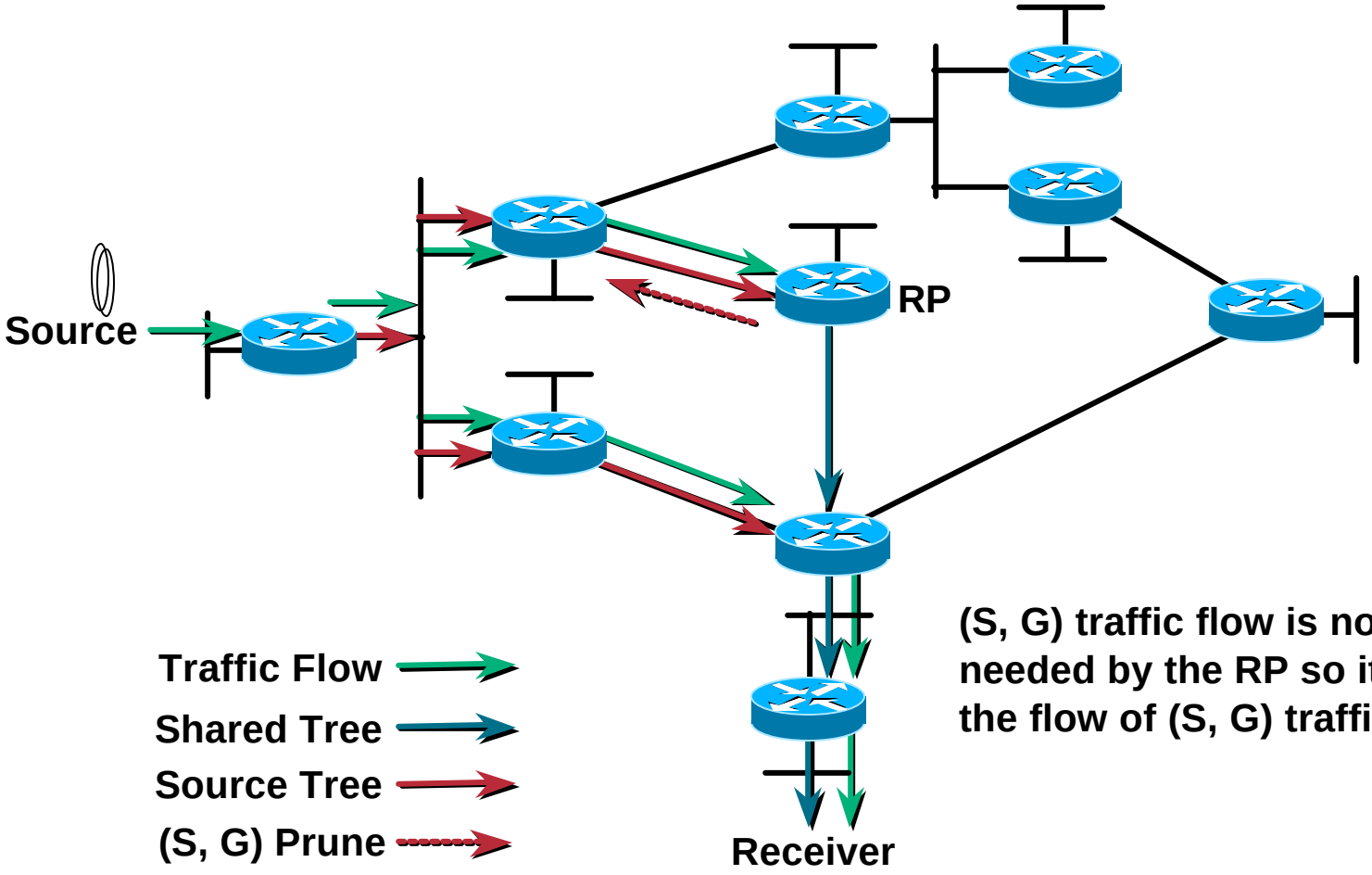
Additional (S, G) State is created along along the Shared Tree to prune off (S, G) traffic.

PIM-SM SPT Switchover



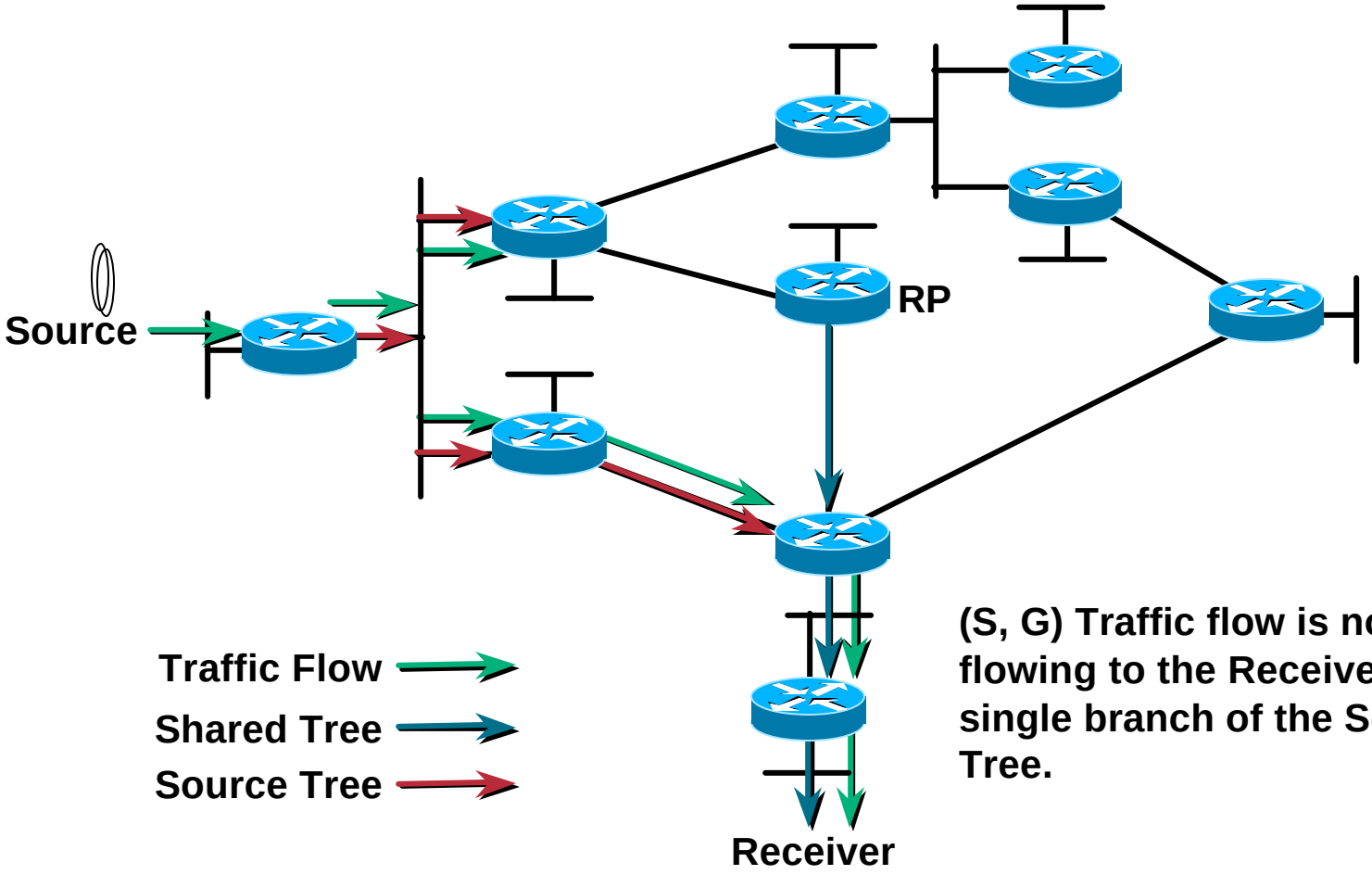
(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver via the Source Tree.

PIM-SM SPT Switchover



(S, G) traffic flow is no longer needed by the RP so it Prunes the flow of (S, G) traffic.

PIM-SM SPT Switchover



(S, G) Traffic flow is now only flowing to the Receiver via a single branch of the Source Tree.

Agenda

- PIM-SM review (forwarding)
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)

MBGP—Multiprotocol BGP

- MBGP overview
- MBGP capability negotiation
- MBGP NLRI exchange
- New MBGP syntax
- Unicast to Multicast NLRI Translation

MBGP Overview

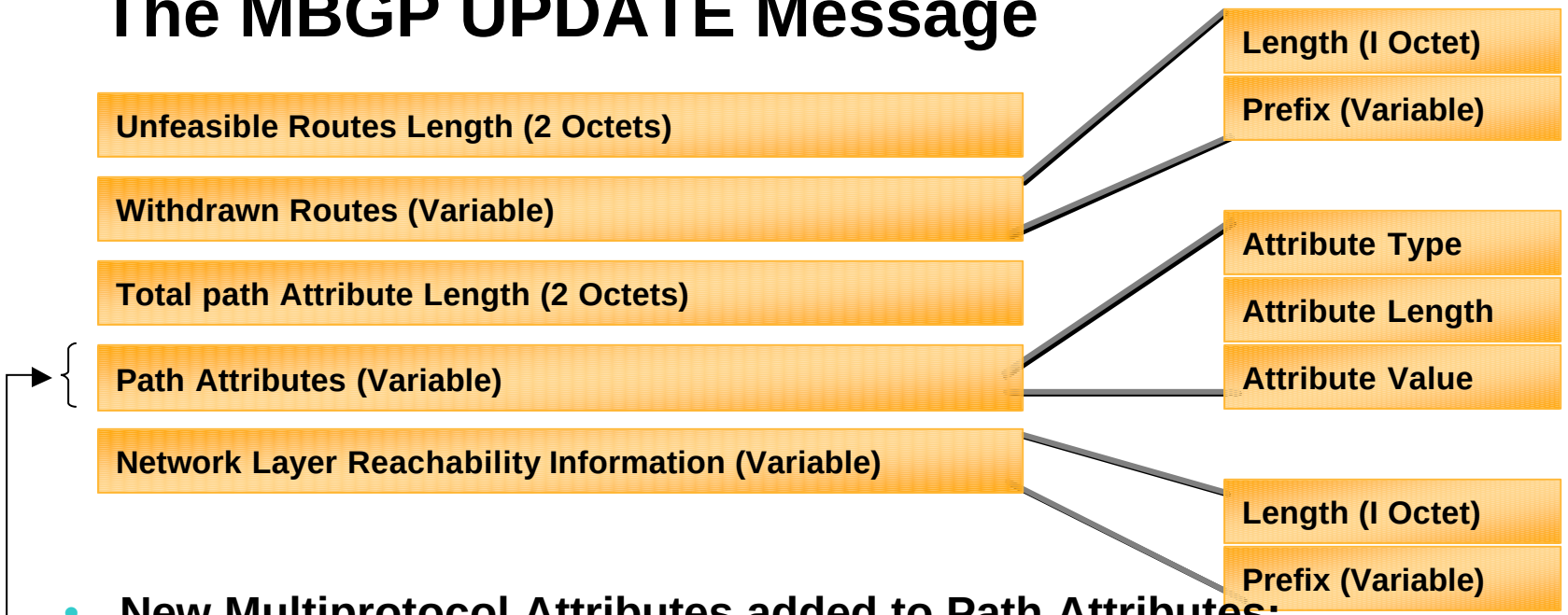
- MBGP: Multiprotocol BGP
  - Defined in RFC 2283 (extensions to BGP)
  - Can carry different types of routes
    - IPv4 Unicast
    - IPv4 Multicast
    - IPv6 Unicast
  - May be carried in same BGP session
  - Does not propagate multicast state info
    - Still need PIM to build Distribution Trees
  - 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

The MBGP UPDATE Message

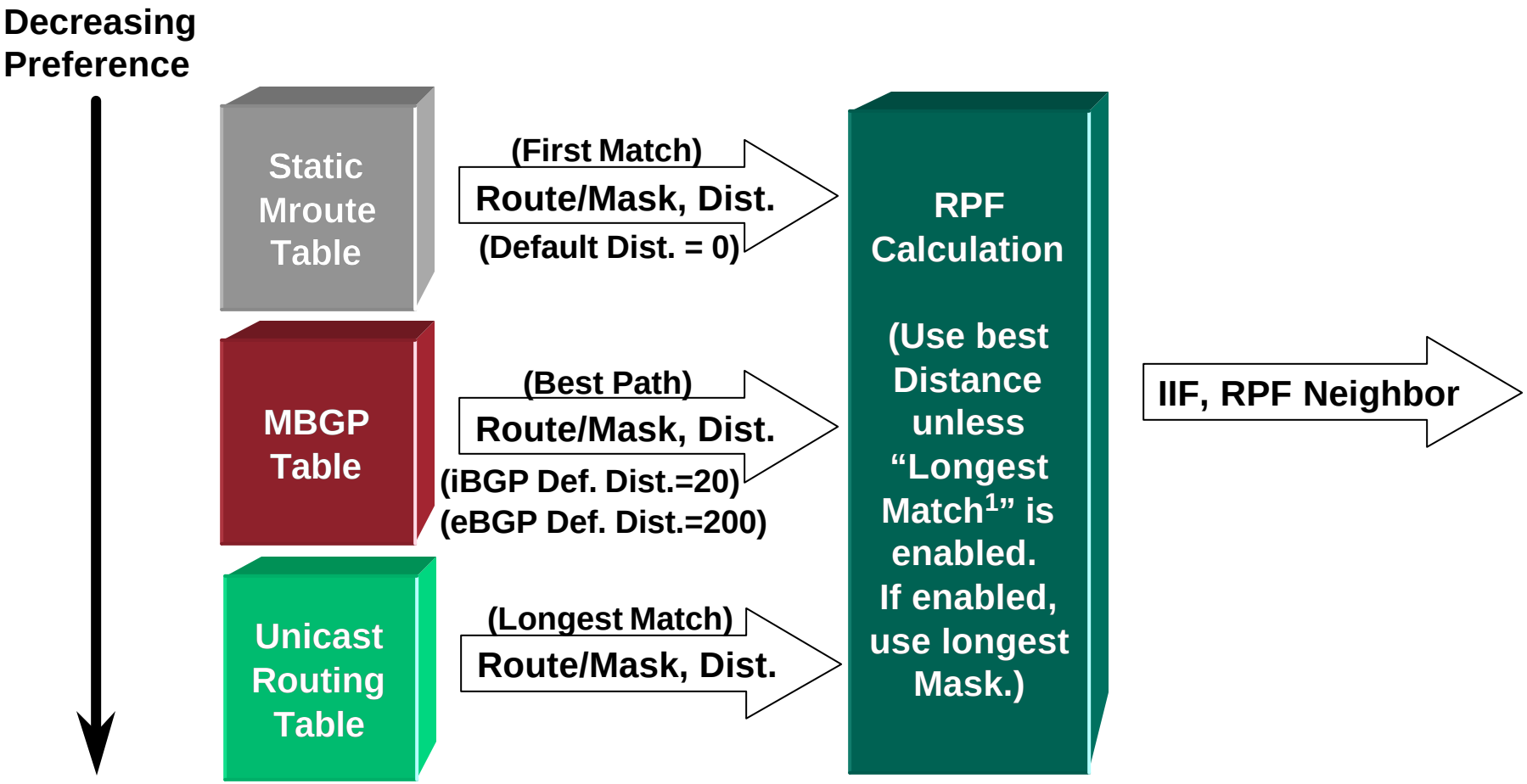


- New Multiprotocol Attributes added to Path Attributes:
  - MP_REACH_NLRI
  - MP_UNREACH_NLRI

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)**
    - **Sub-AFI = 3 (NLRI is used for both unicast and multicast RPF check)**

PIM RPF Calculation Details



Global Command: `ip multicast longest-match`

MBGP—Capability Negotiation

- BGP routers establish BGP sessions through the OPEN message
- OPEN message contains optional parameters
- BGP session is terminated if OPEN parameters are not recognized
- New parameter: CAPABILITIES
  - Multiprotocol extension
  - Multiple routes for same destination

MBGP—Capability Negotiation

Cisco.com

- **Keyword on neighbor command**

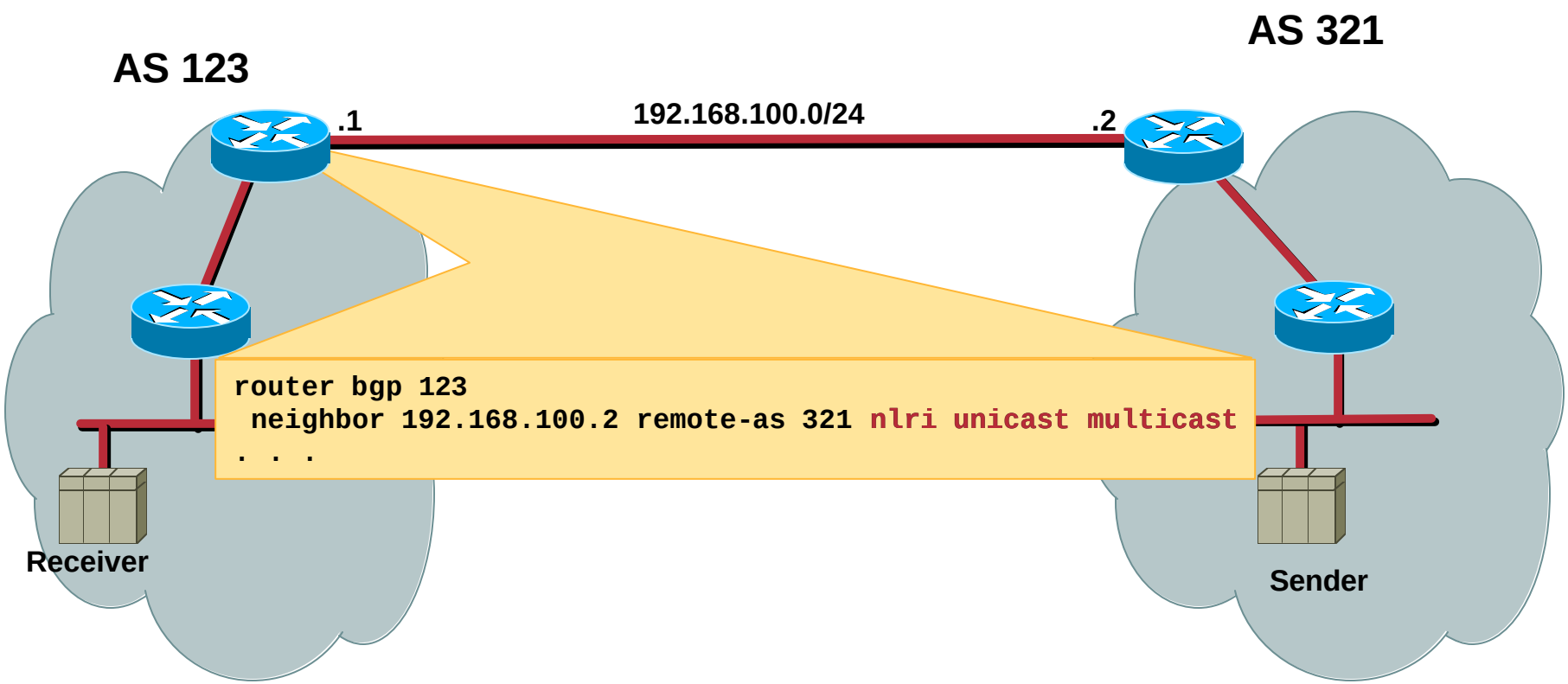
```
neighbor <foo> remote-as <asn> nlri multicast unicast
```

Configures router to negotiate either or both NLRI

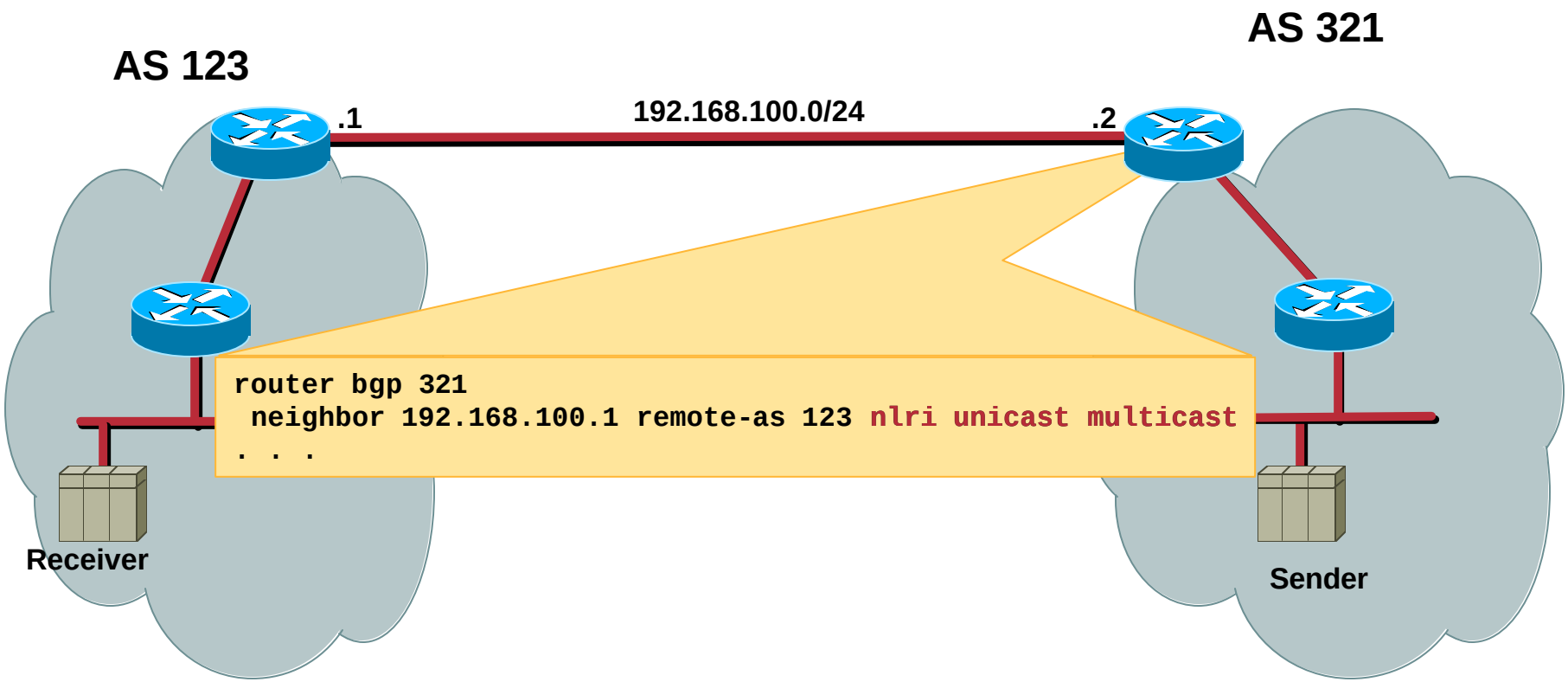
If neighbor configures both or subset, common NRLI is used in both directions

If there is no match, notification is sent and peering doesn’t come up

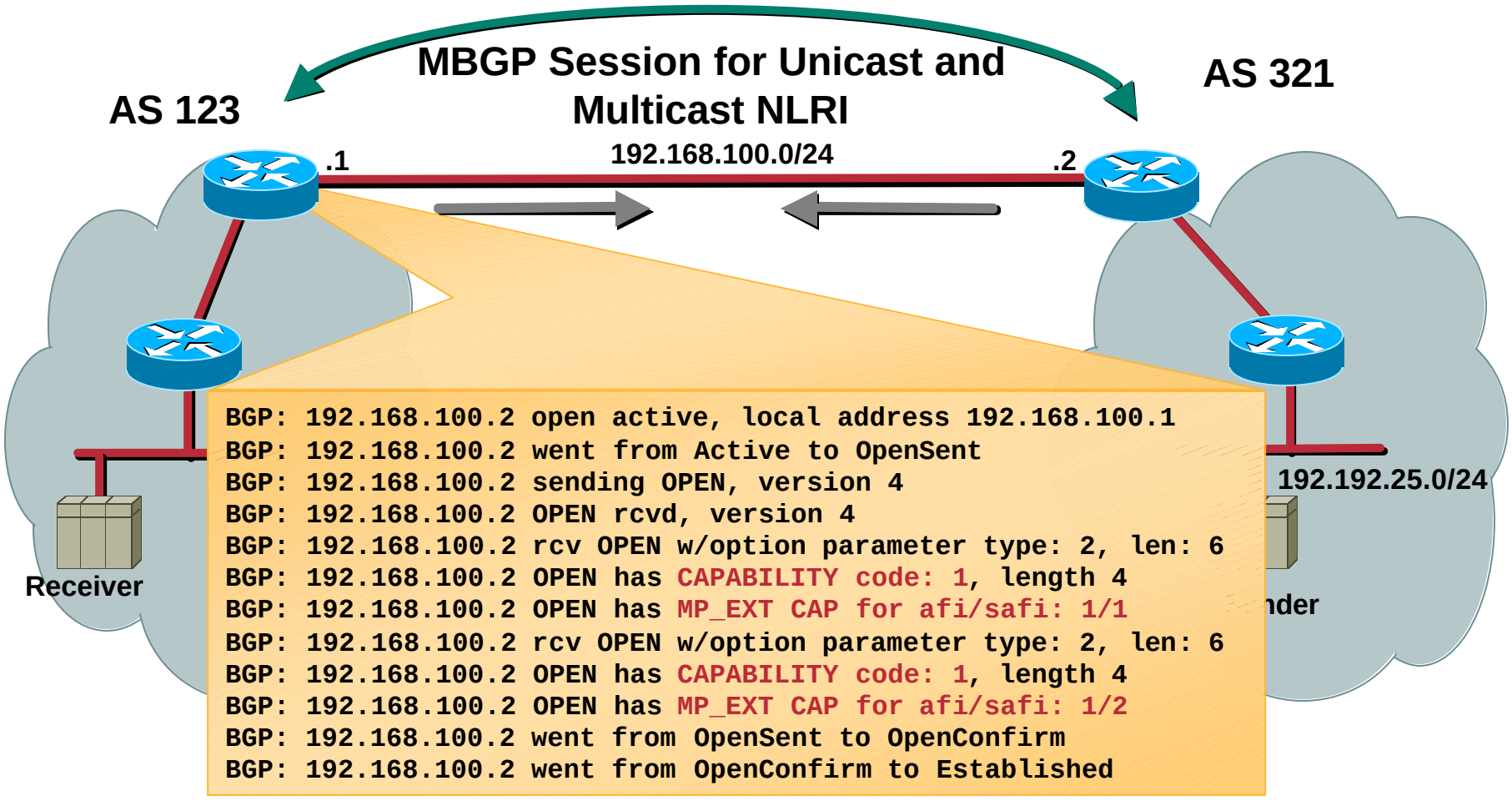
MBGP — Capability Negotiation



MBGP — Capability Negotiation

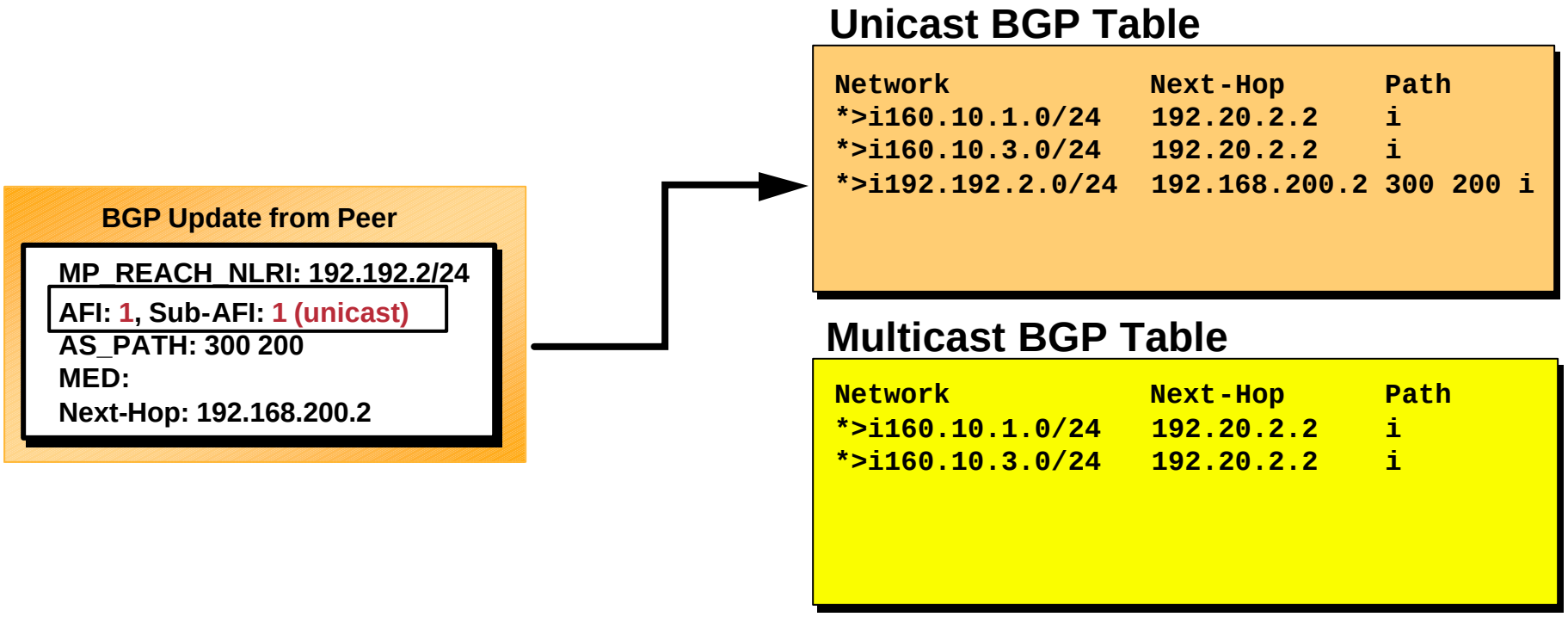


MBGP — Capability Negotiation



MBGP—NLRI Information

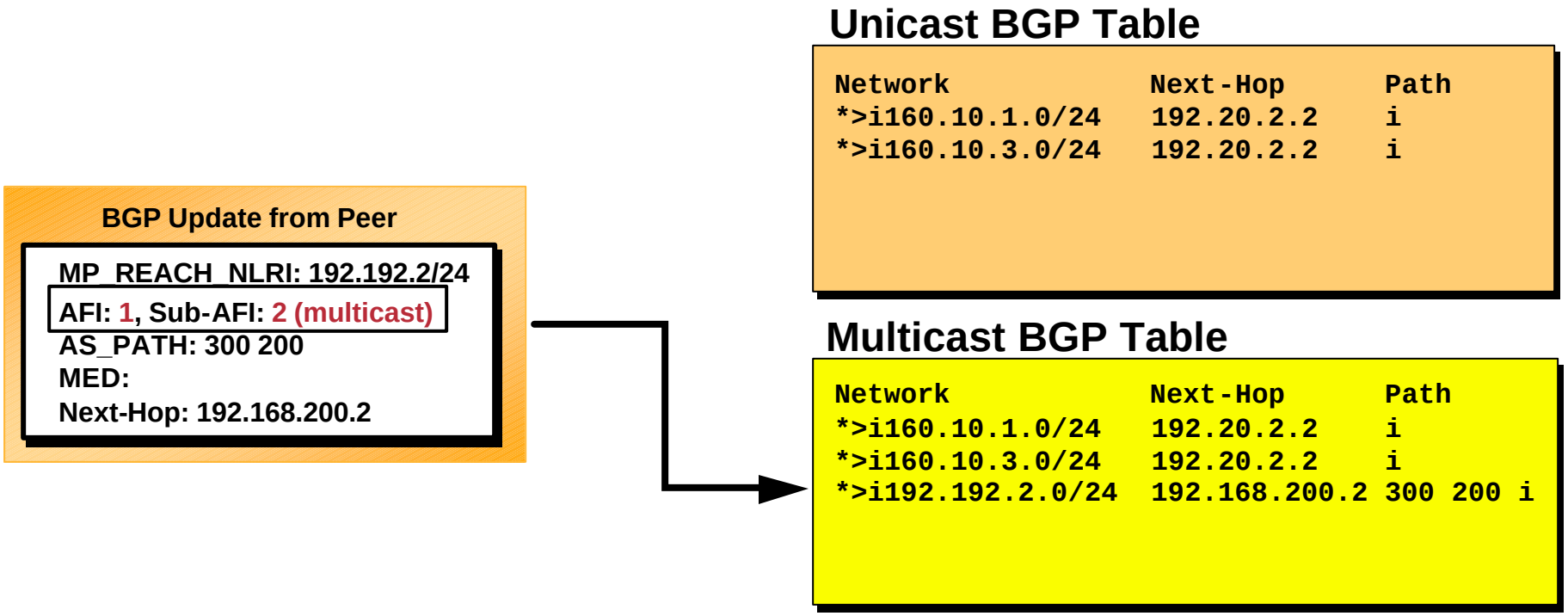
Cisco.com



- 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)

MBGP—NLRI Information

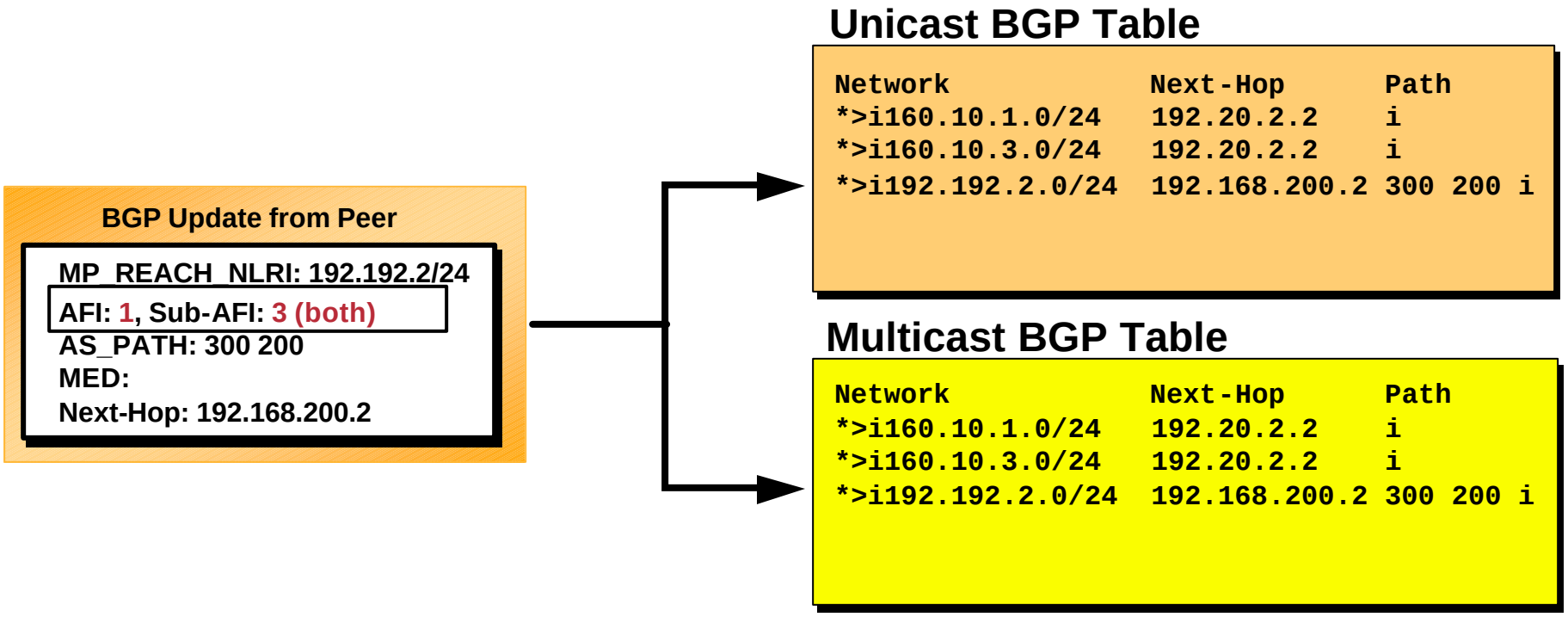
Cisco.com



- 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

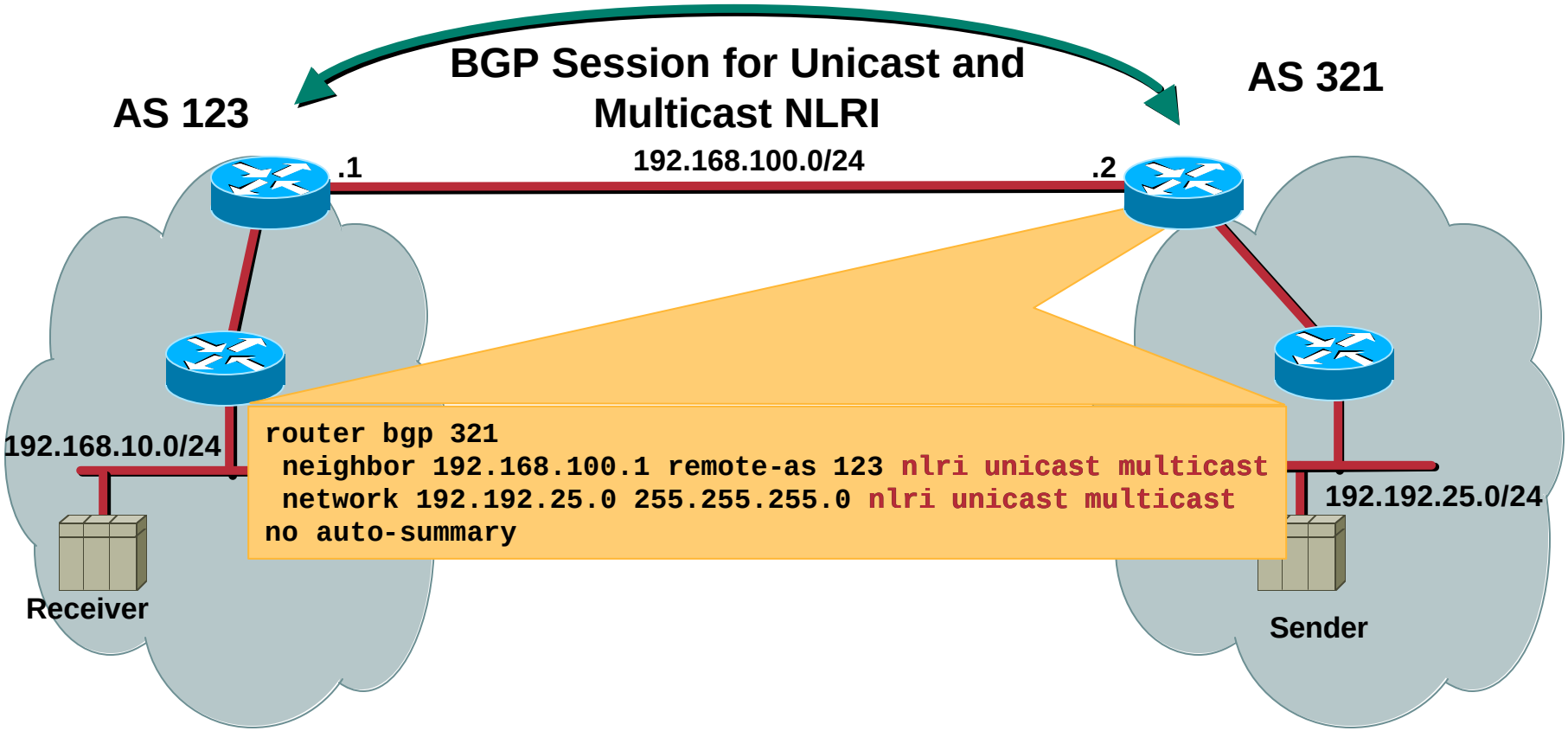


- Storage of arriving NLRI information depends on AFI/SAFI fields in the Update message
  - Unicast BGP Table only (AFI=1/SAFI=1 or old style NLRI)
  - Multicast BGP Table only (AFI=1/SAFI=2)
  - Both BGP Tables (AFI=1/SAFI=3)

MBGP—NLRI Information

Cisco.com

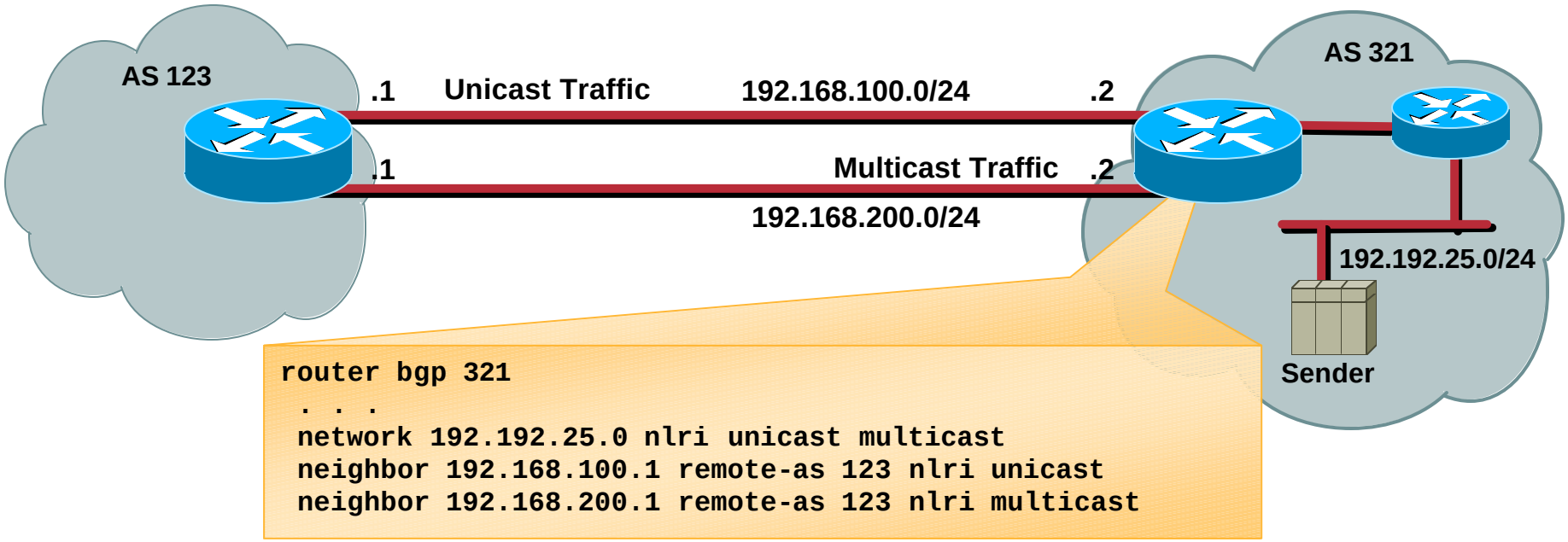
Congruent Topologies



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 Syntax Change

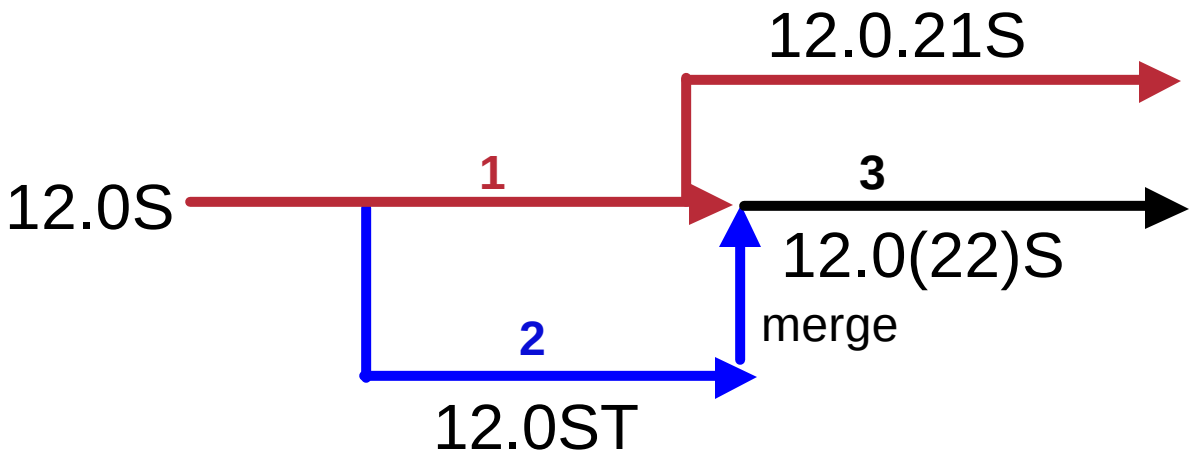
- **NLRI syntax is in 12.0S**
- **‘address –family’ syntax introduced in 12.0(7)T,12.1.**
- **12.0ST maintains both syntax structures. 12.0ST provides an easy option to move to new syntax:**

  RouterA(config)#router bgp 18

  RouterA(config-router)#bgp upgrade-cli

- **12.0S retains NLRI syntax until 12.0(22)S where both NLRI and address family config exists in Hybrid mode. Hybrid, allows NLRI and address family syntax in same configuration.**

ST to S Merge



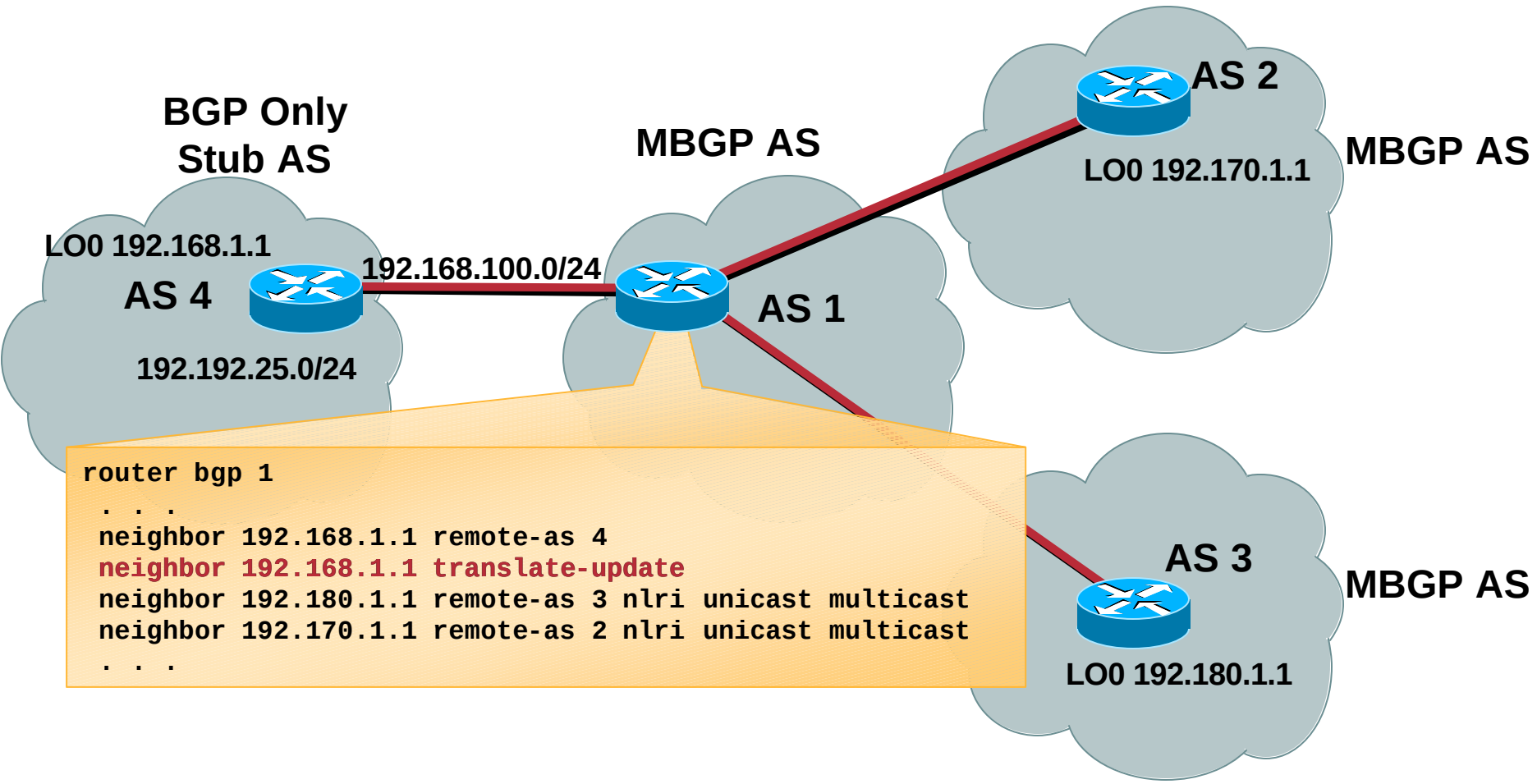
1 = traditional nlri 120s
2 = bgp_reorg with AF
3 = post merge 120s

Unicast-Multicast NLRI Translation

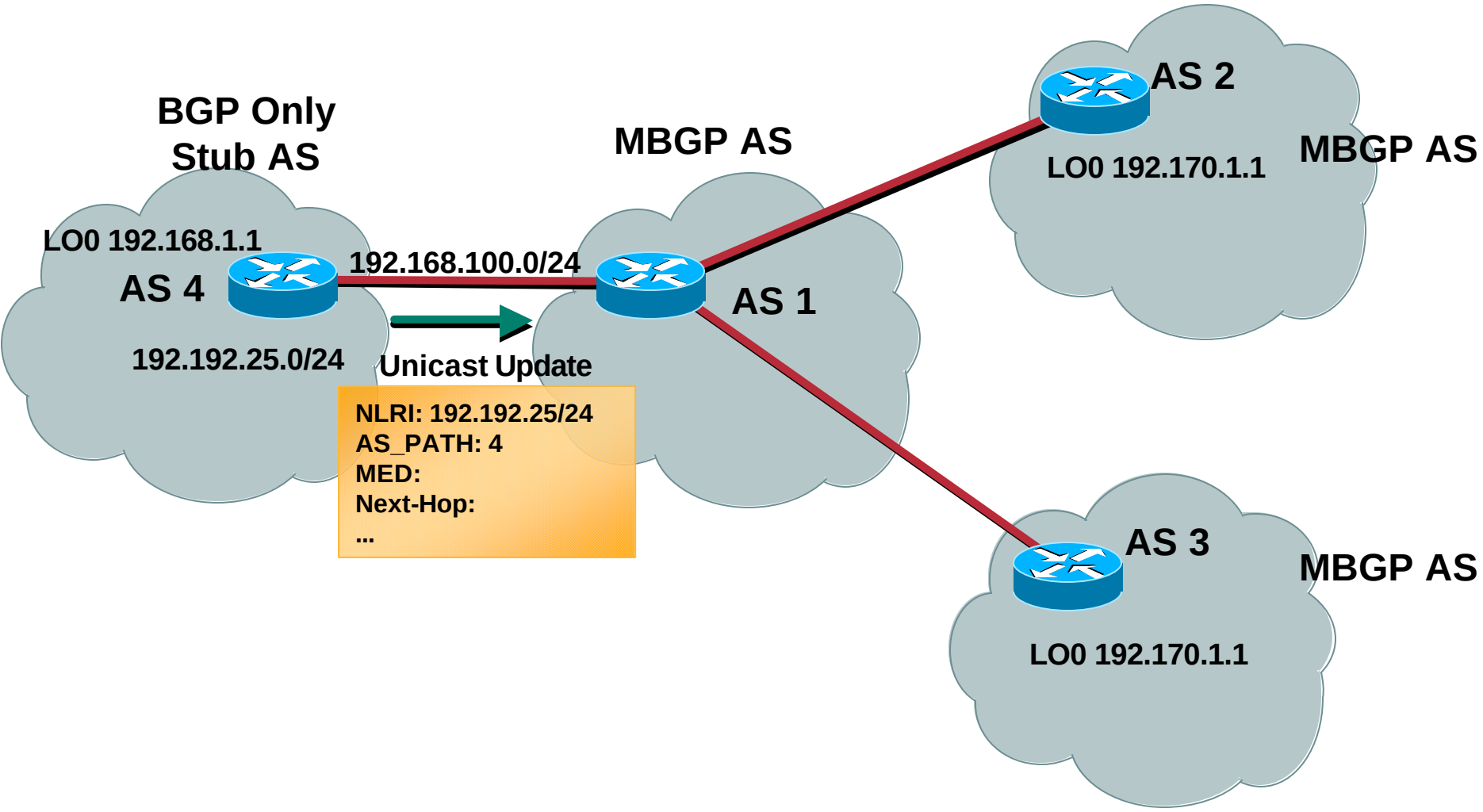
- **BGP stubs that don’t have MBGP support need to get their routes into the MBGP backbone**
- **Stub gets external routes via existing BGP peering**
- **Use translate-update command**

```
neighbor <foo> translate-update [nlri unicast multicast]
```

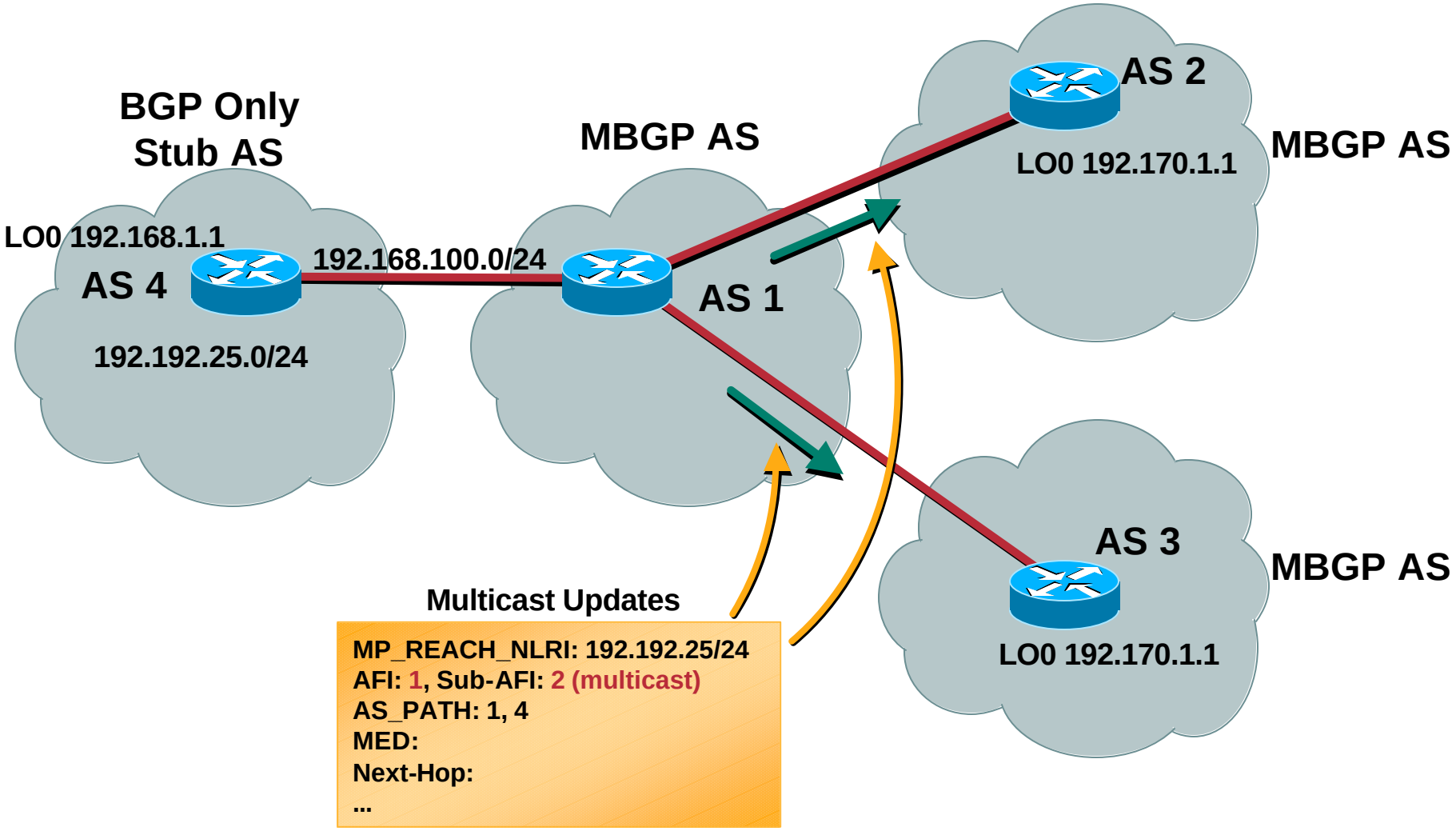
Unicast-Multicast NLRI Translation



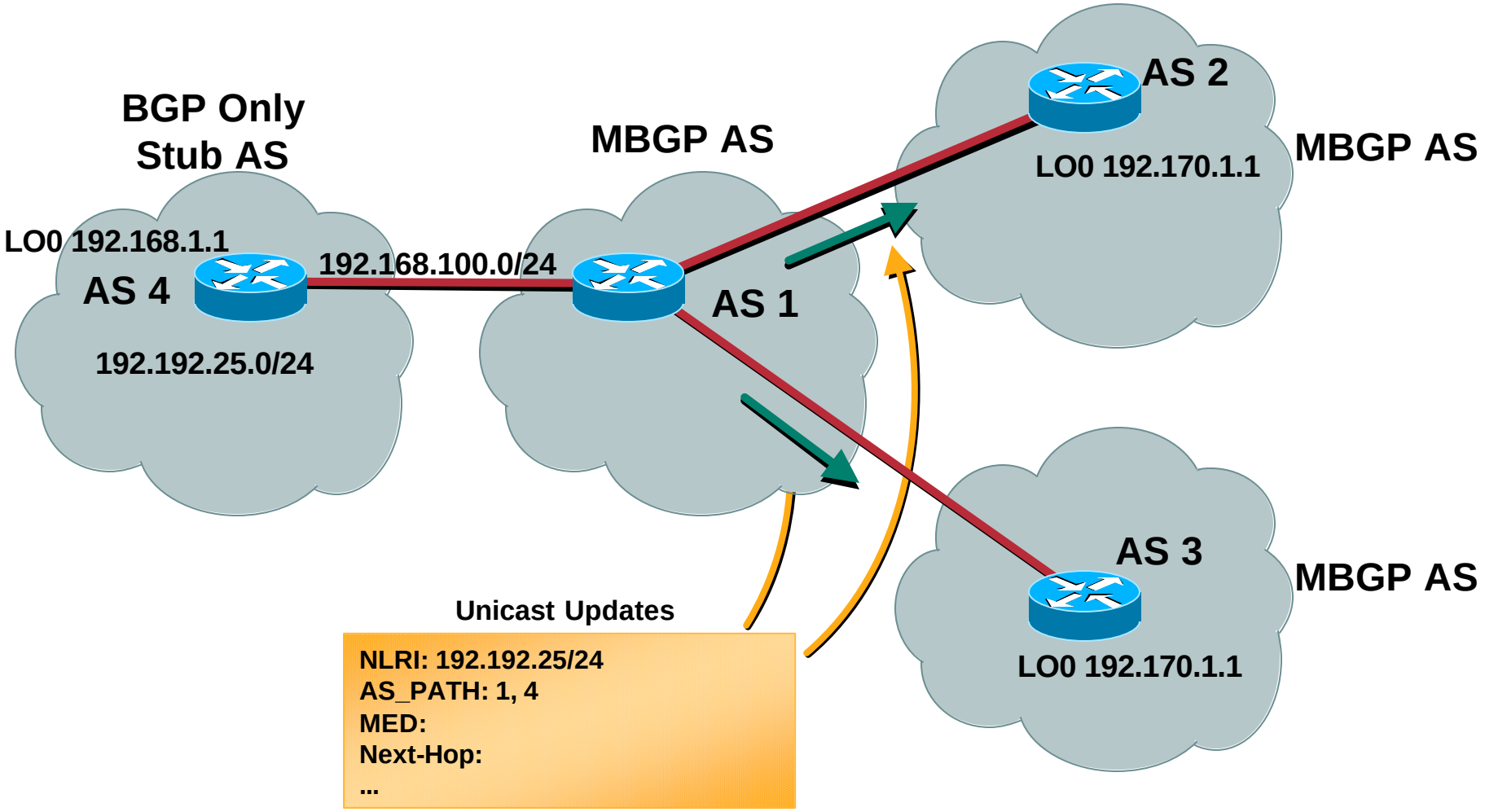
Unicast-Multicast NLRI Translation



Unicast-Multicast NLRI Translation



Unicast-Multicast NLRI Translation



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

- PIM-SM review (forwarding)
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)

MSDP Concept

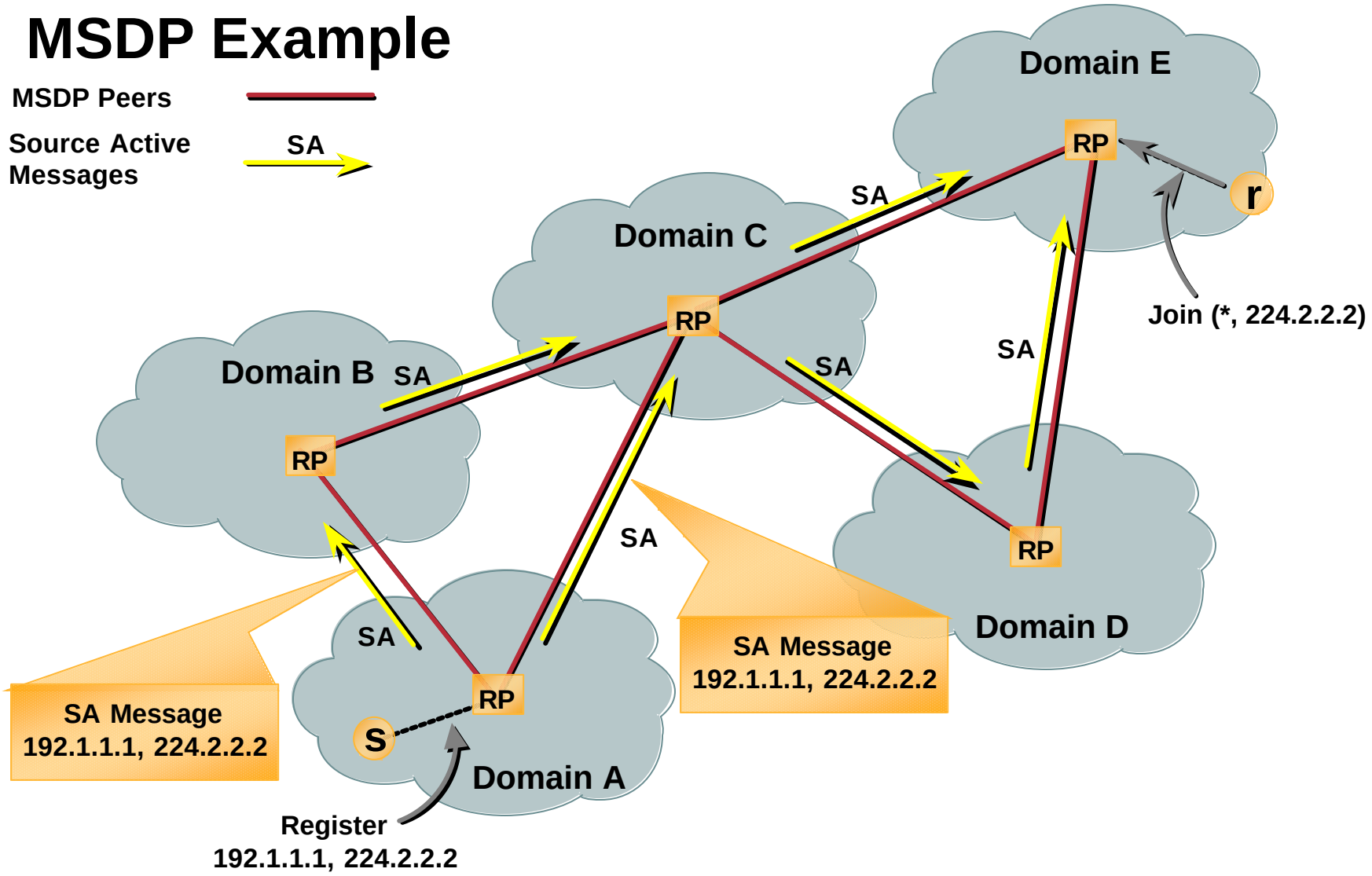
- **Simple but elegant**
  - **Utilize inter-domain source trees**
  - **Reduces problem to locating active sources**
  - **RP or receiver last-hop can join inter-domain source tree**

MSDP Concepts

- **Works with PIM-SM only**
  - **RP's knows about all sources in a domain**
    - **Sources cause a "PIM Register" to the RP**
    - **Can tell RP's in other domains of its sources**
      - **via MSDP SA (Source Active) messages**
  - **RP's know about receivers in a domain**
    - **Receivers cause a "(*, G) Join" to the RP**
    - **RP can join the source tree in the peer domain**
      - **via normal PIM (S, G) joins**

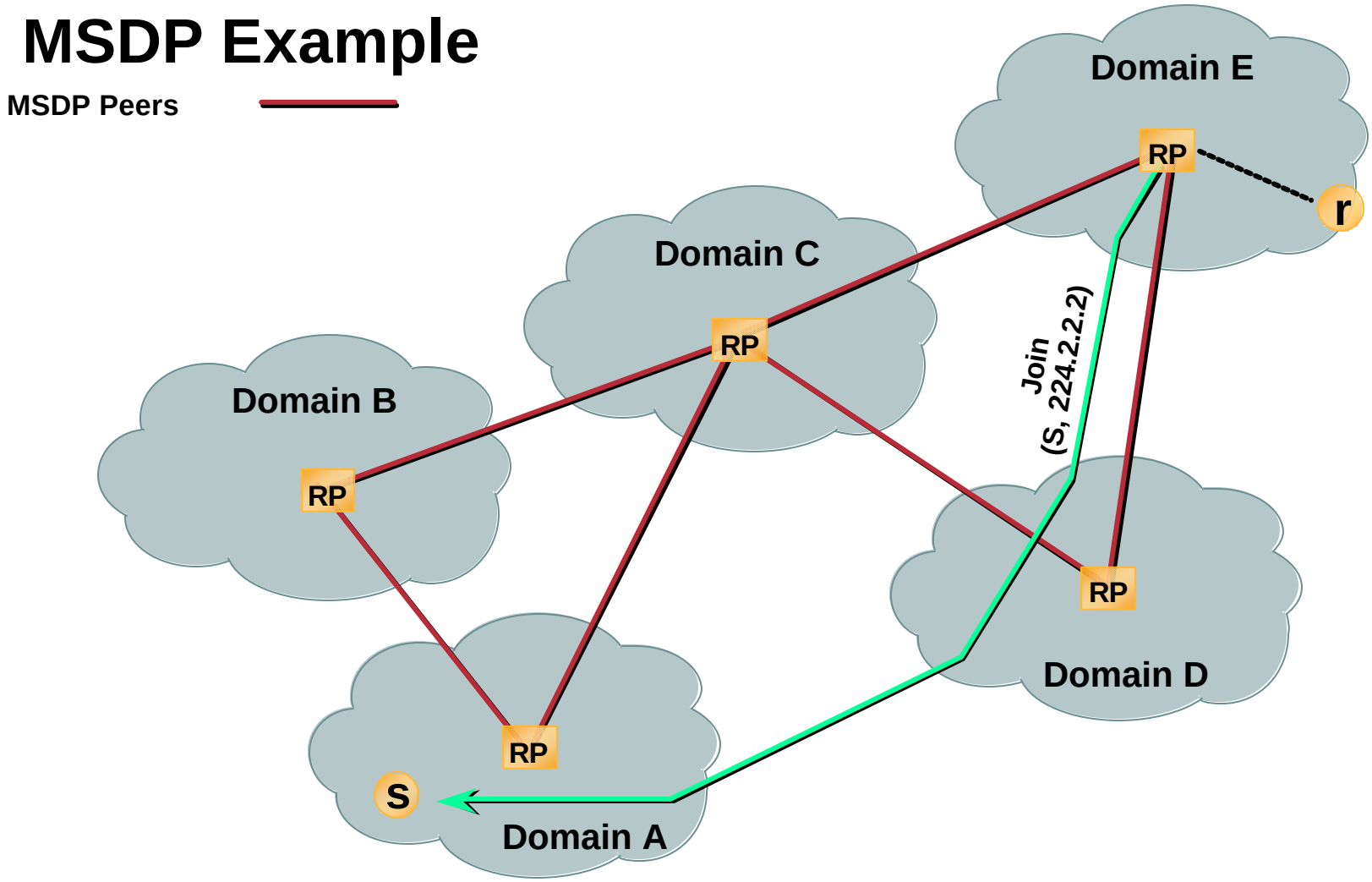
MSDP Overview

MSDP Example



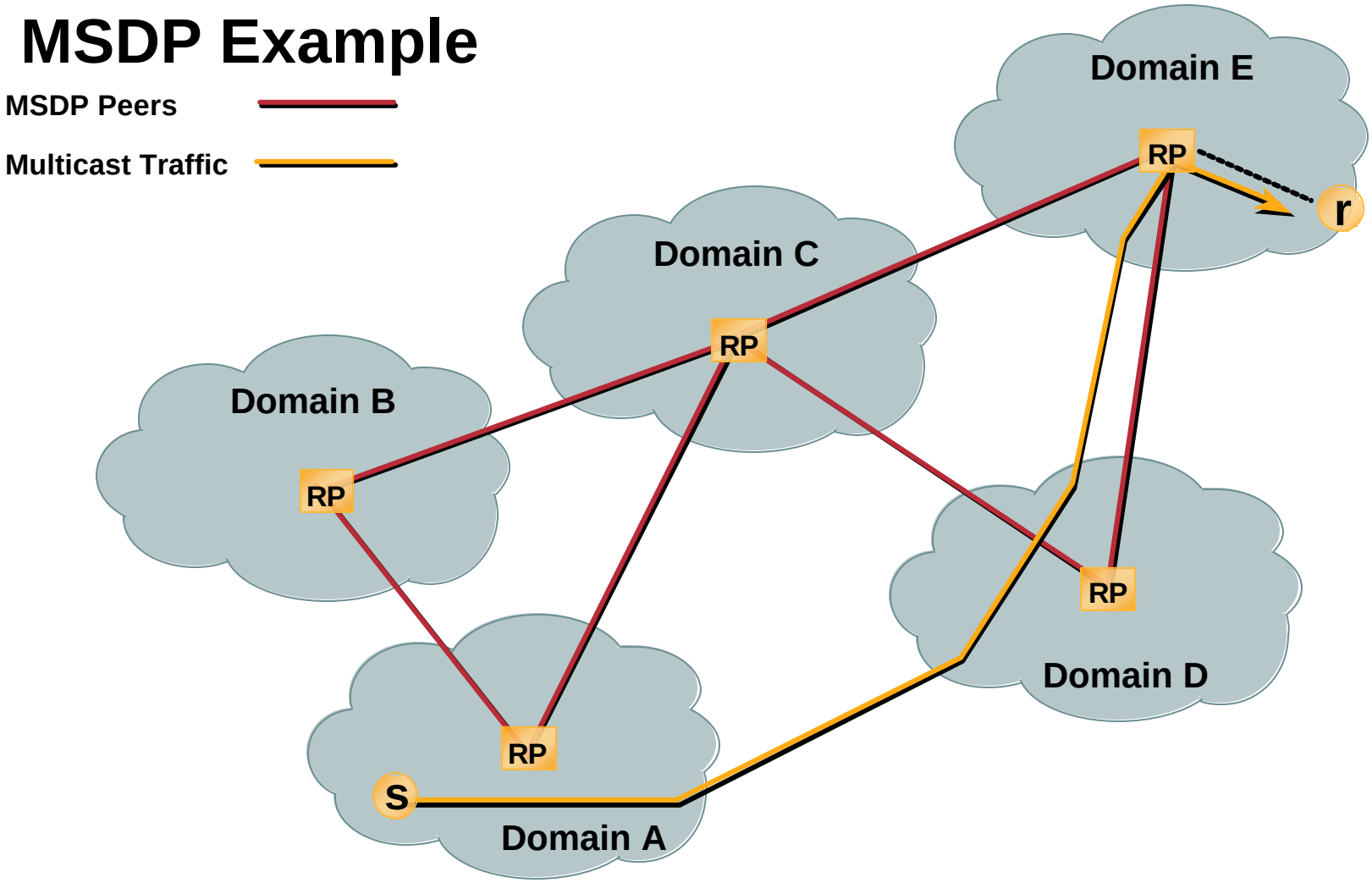
MSDP Overview

MSDP Example



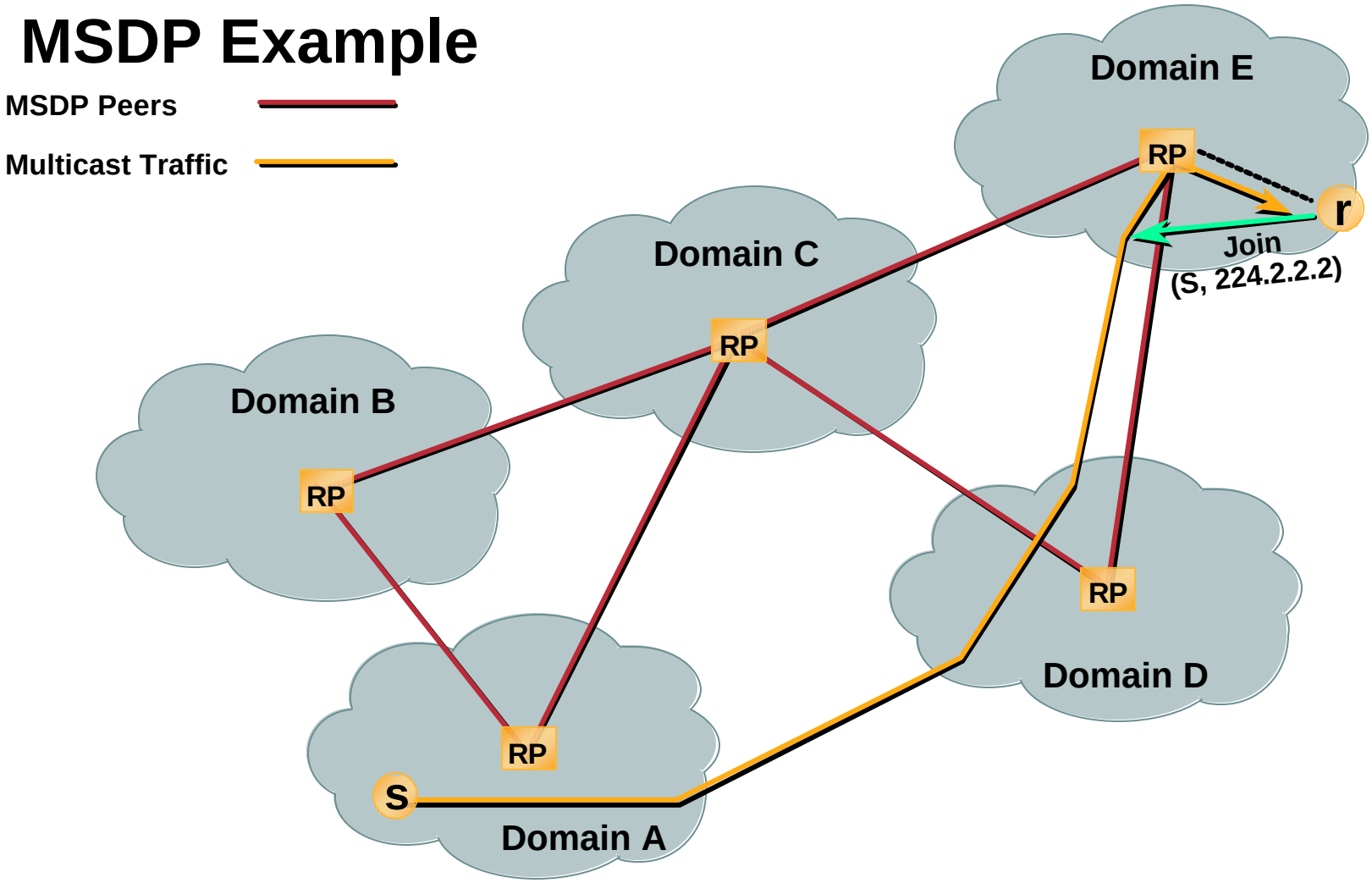
MSDP Overview

MSDP Example



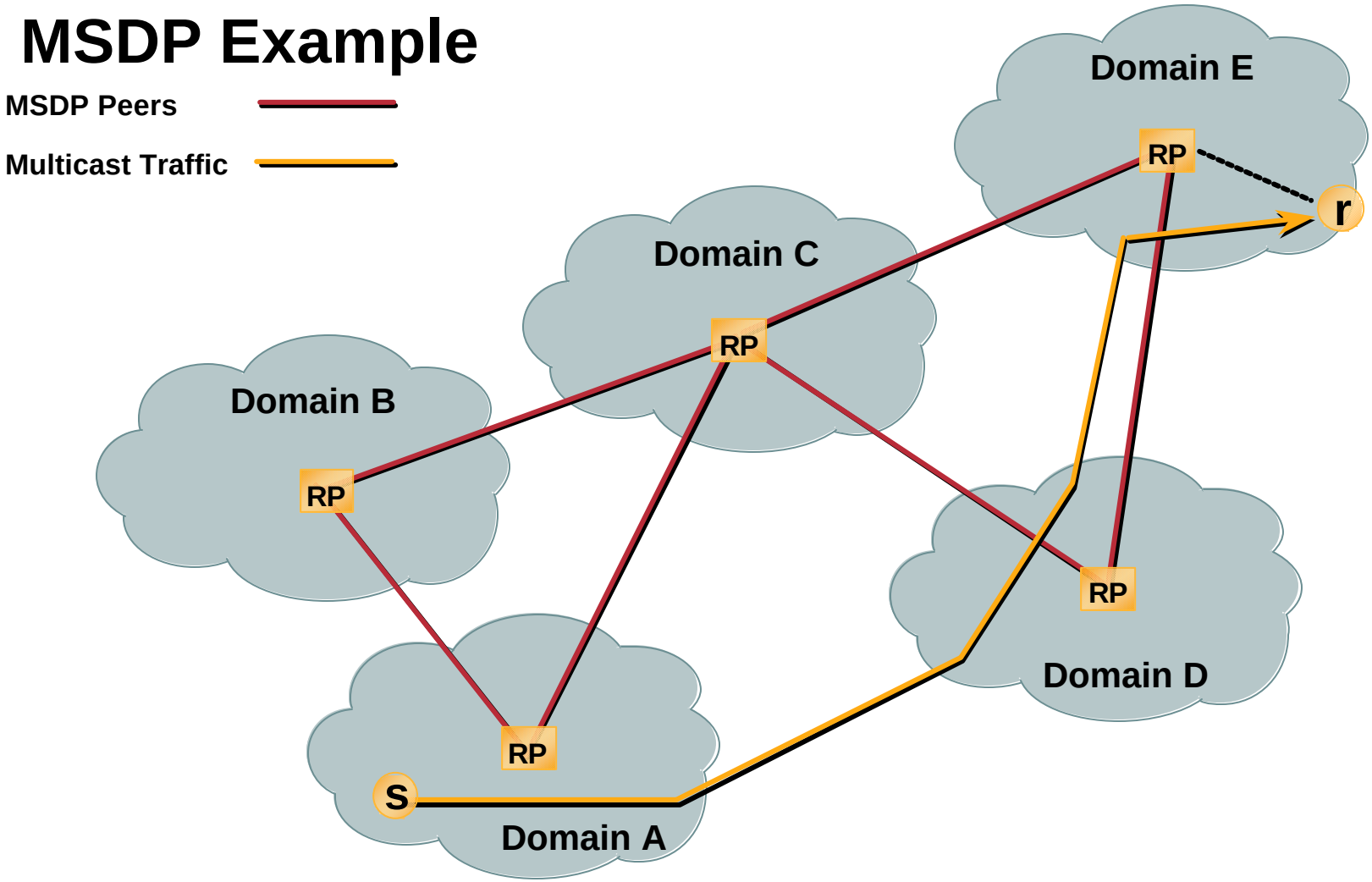
MSDP Overview

MSDP Example



MSDP Overview

MSDP Example



MSDP Design Points

- MSDP peers talk via TCP connections
- Source Active (SA) messages
  - Peer-RPF forwarded to prevent loops
    - RPF check on AS-PATH back to the peer RP
    - If successful, flood SA message to other peers
    - Stub sites can be configured to accept all SA messages (similar to static default for routing)
      - Since they have only one exit (e.g., default peer)
  - SA messages stored in local cache
    - On by default
    - Reduces join latency

MSDP Peers

- **MSDP establishes a neighbor relationship between MSDP peers**
  - **Peers connect using TCP port 639**
  - **Peers send keepalives every 60 secs (fixed)**
  - **Peer connection reset after 75 seconds if no MSDP packets or keepalives are received**
- **MSDP peers almost always run BGP**
  - **May be an MBGP peer, a BGP peer or both**

MSDP SA Messages

- **MSDP Source Active (SA) Messages**
  - **Used to advertise active Sources in a domain**
  - **Can also 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

- **SA Message RPF Check**
  - **Accept SA's via single deterministic path**
    - **Ignore all other arriving SA's**
    - **Necessary to prevent SA's from looping**
- **Problem**
  - **Need to know MSDP topology of Internet**
    - **But, MSDP does not distribute topology data!**
- **Solution**
  - **Use (m)BGP data to infer MSDP topology**
    - **Impact:**
      - **The MSDP topology should follow (m)BGP topology**
      - **An MSDP peer will *generally* also be an m(BGP) peer**

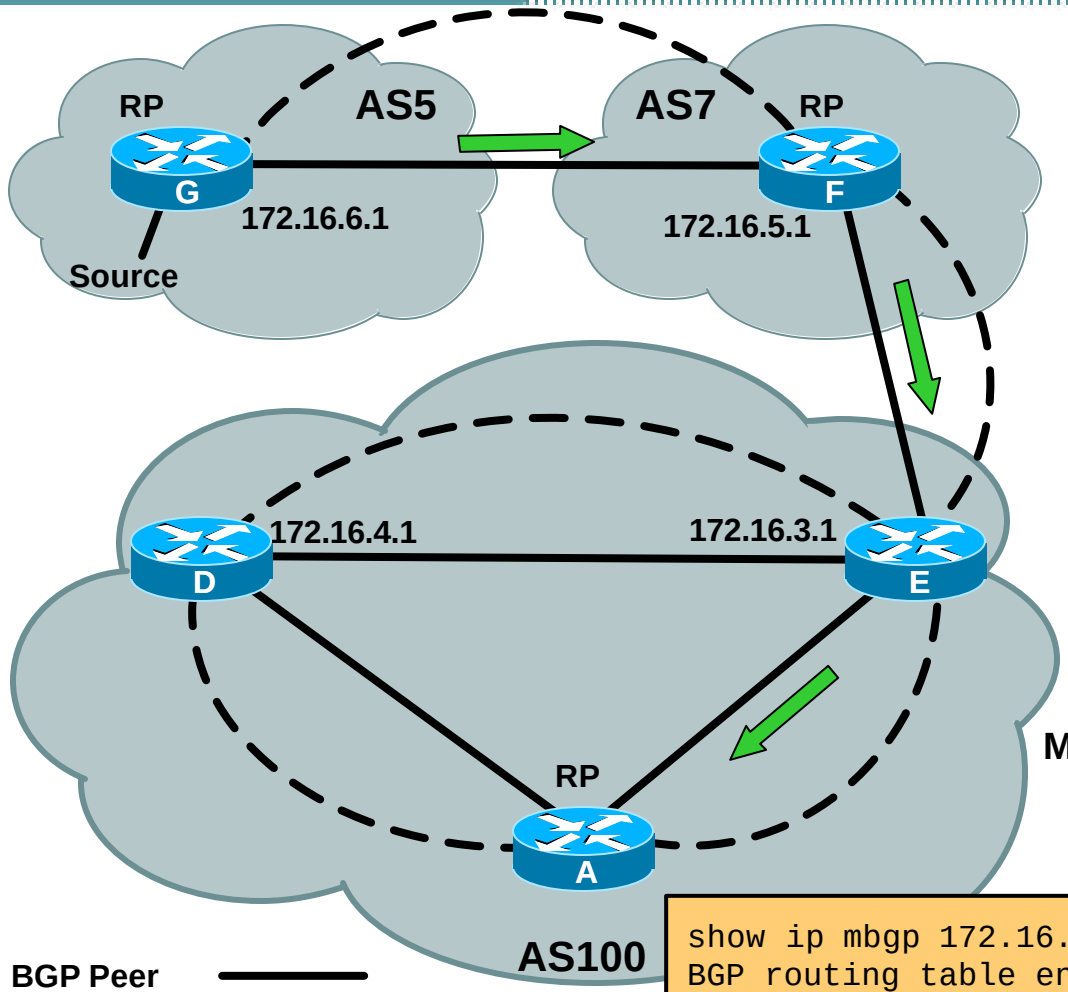
Receiving SA Messages

- **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)

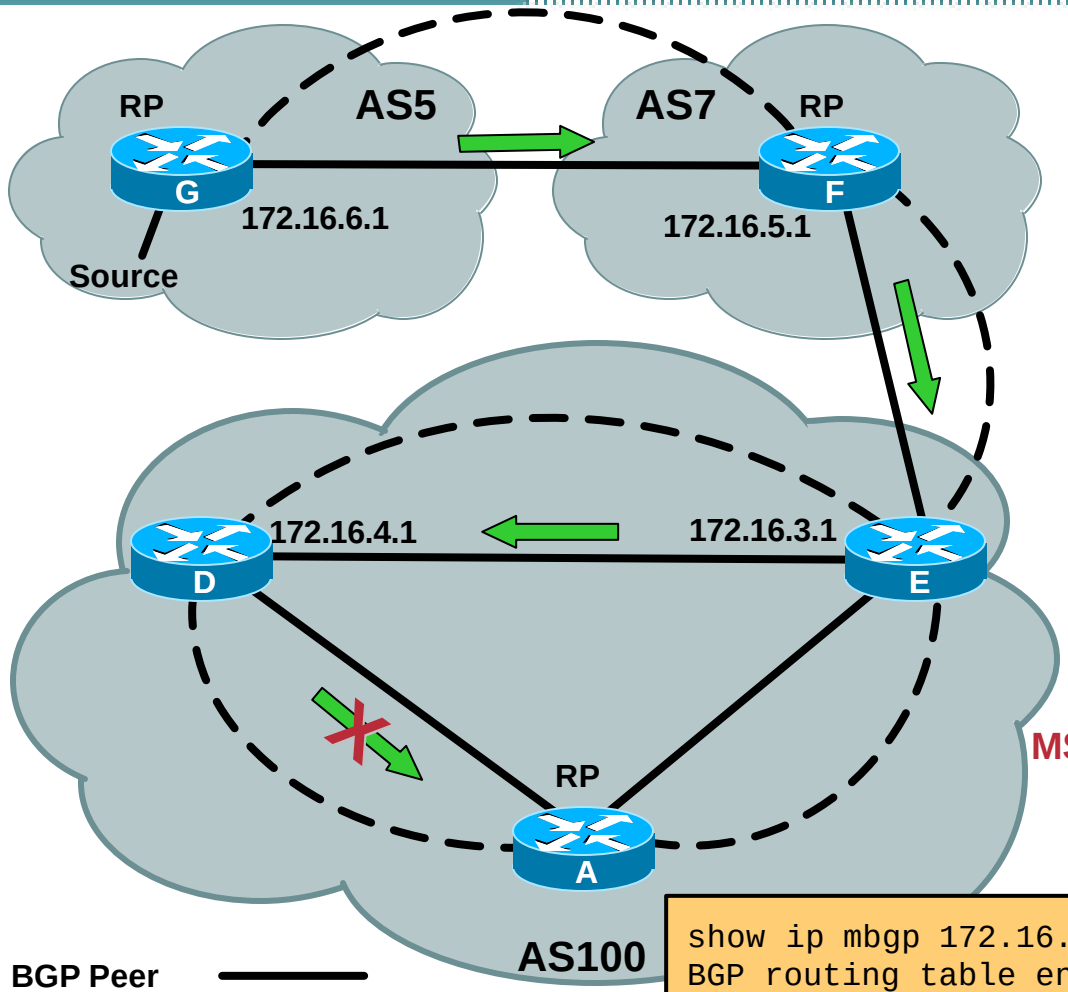
MSDP Peer address = 172.16.3.1

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)

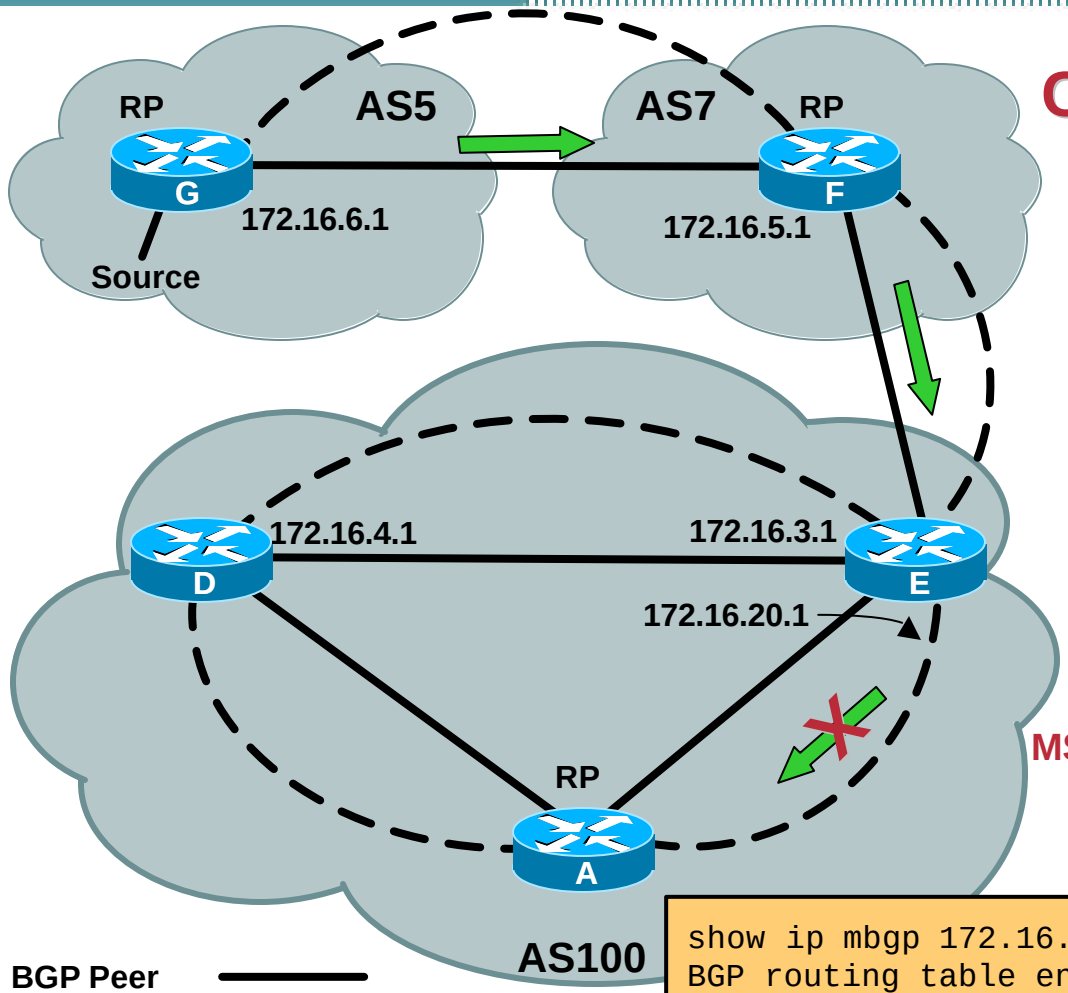
MSDP Peer address = 172.16.4.1

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)

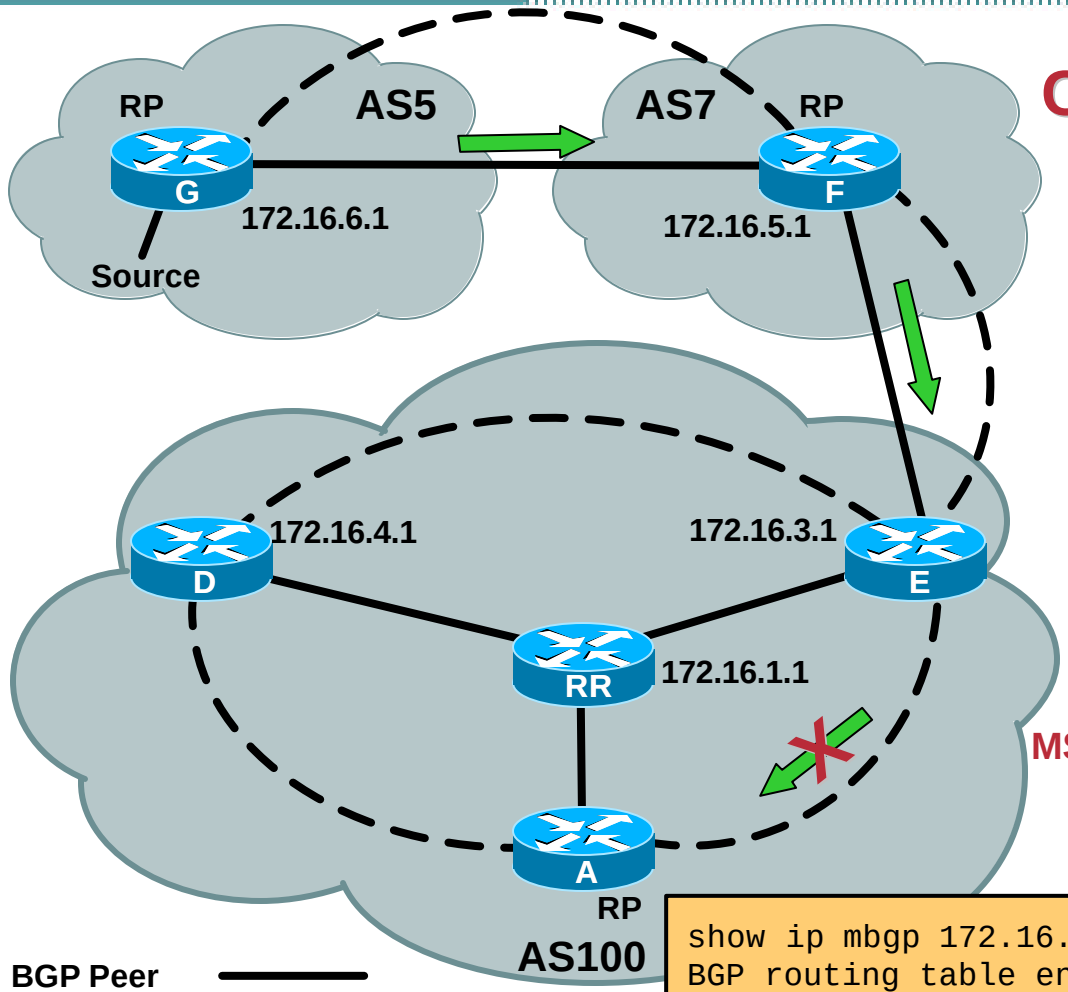
MSDP Peer address = 172.16.20.1

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

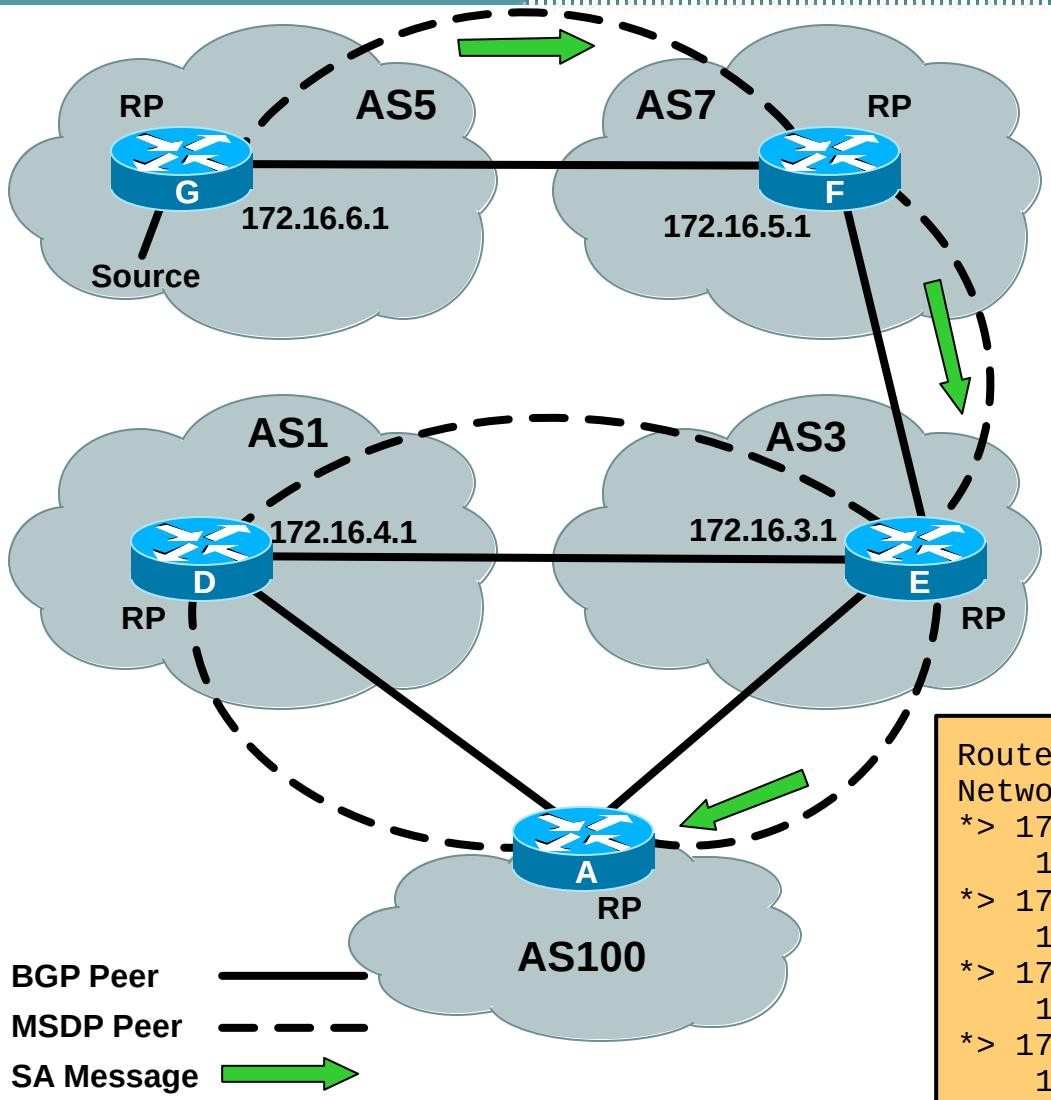
SA RPF Check Fails

```
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 = AS of e(m)BGP peer?**
      - **If Yes, RPF Succeeds**

Rule2: MSDP peer = e(m)BGP peer



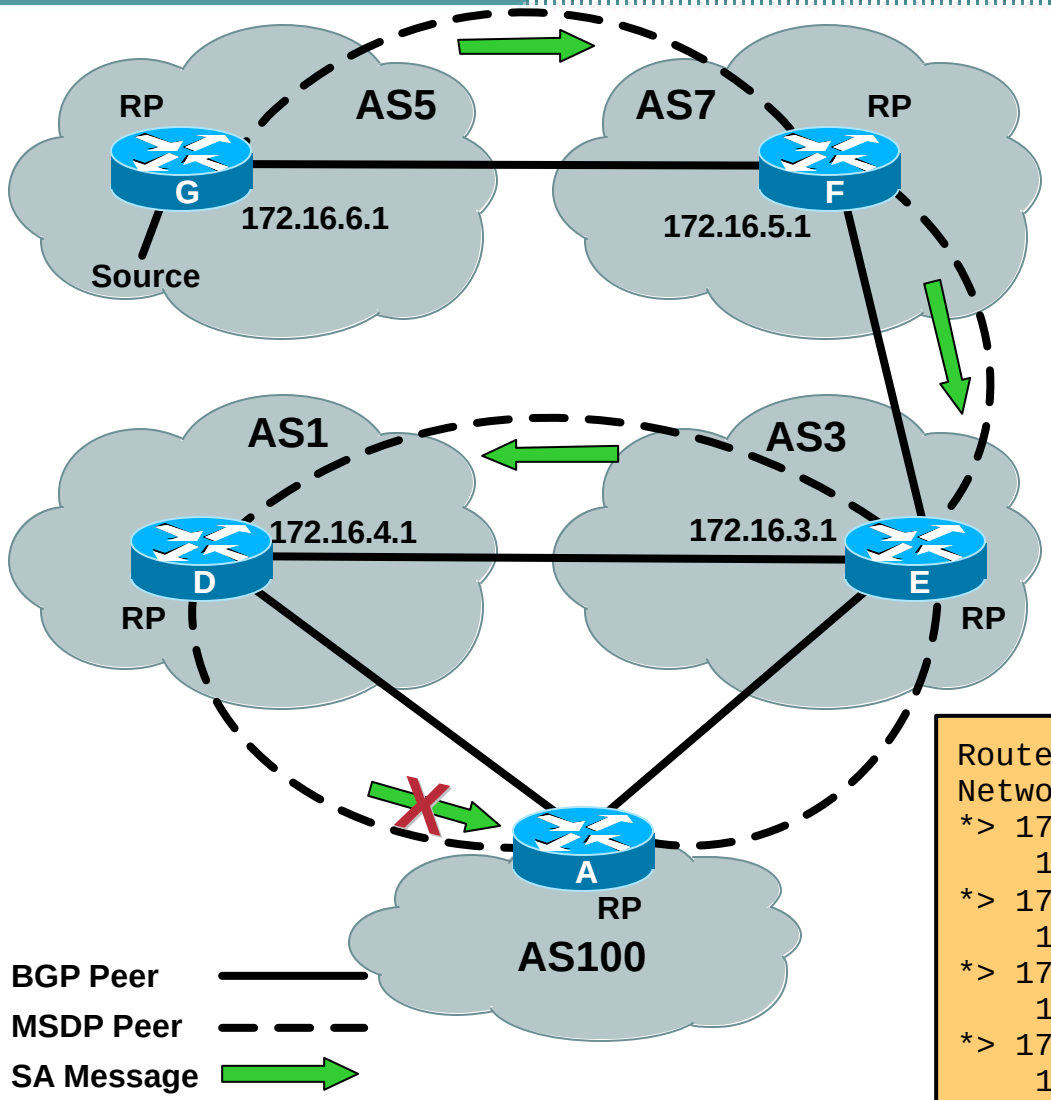
First-AS in best-path to RP = 3
AS of MSDP Peer = 3

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    172.16.4.1     [3] 7 5 i
    172.16.6.0/24    172.16.3.1     1 3 7 5 i
```

Rule2: MSDP peer = e(m)BGP peer



First-AS in best-path to RP = 3
AS of e(m)BGP Peer = 1

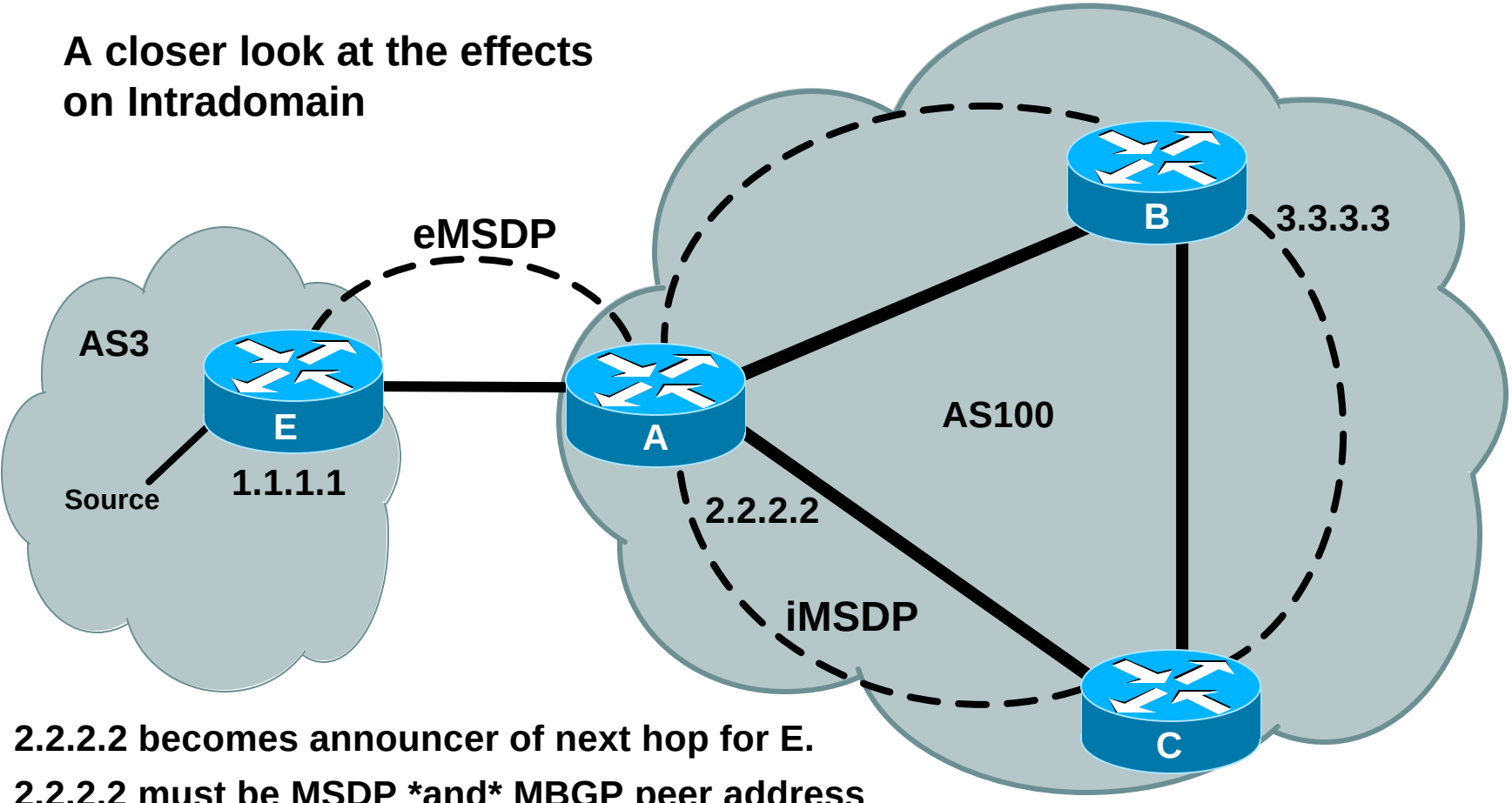
First-AS in best-path to RP != AS of e(m)BGP Peer

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



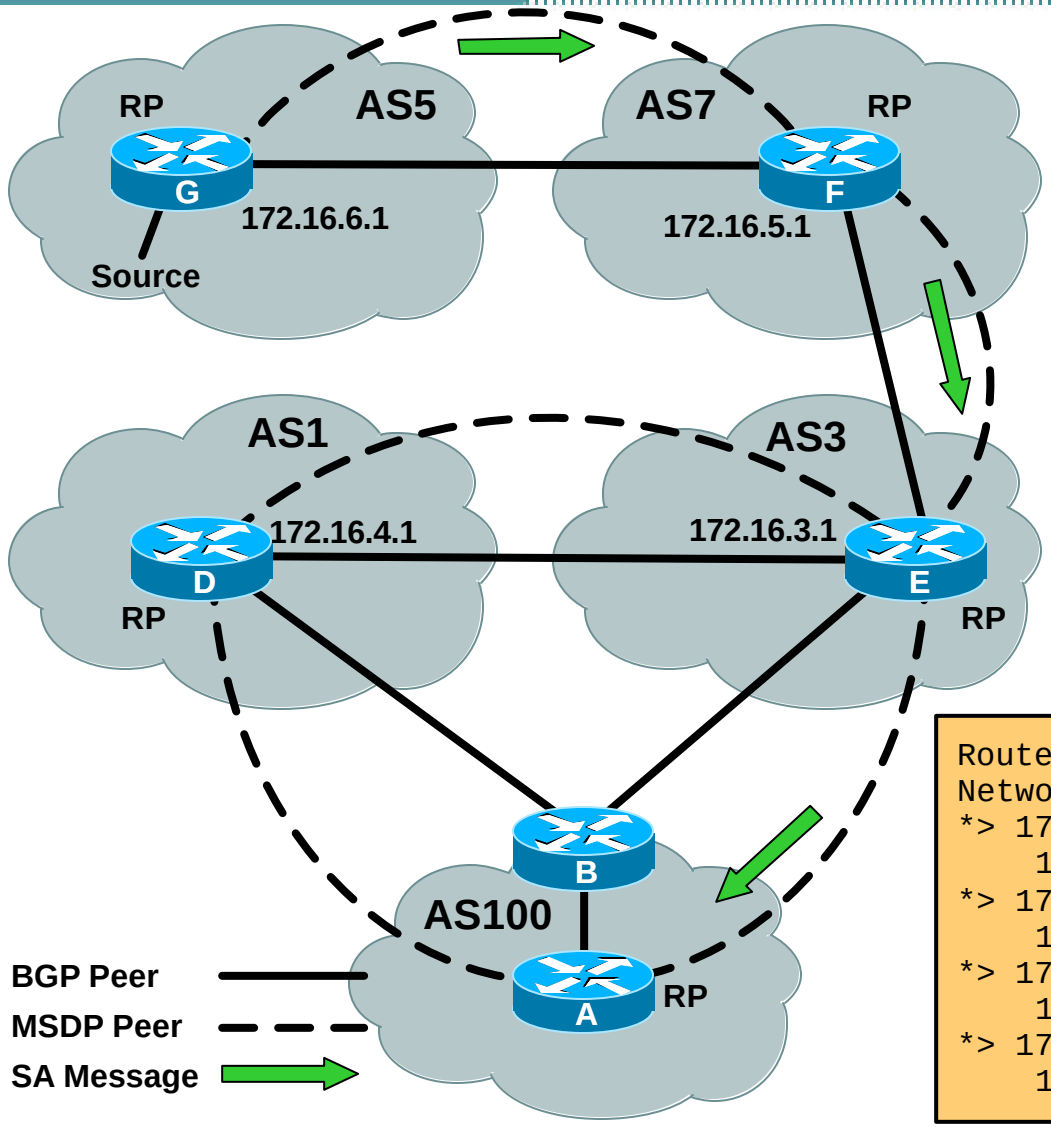
2.2.2.2 becomes announcer of next hop for E.

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



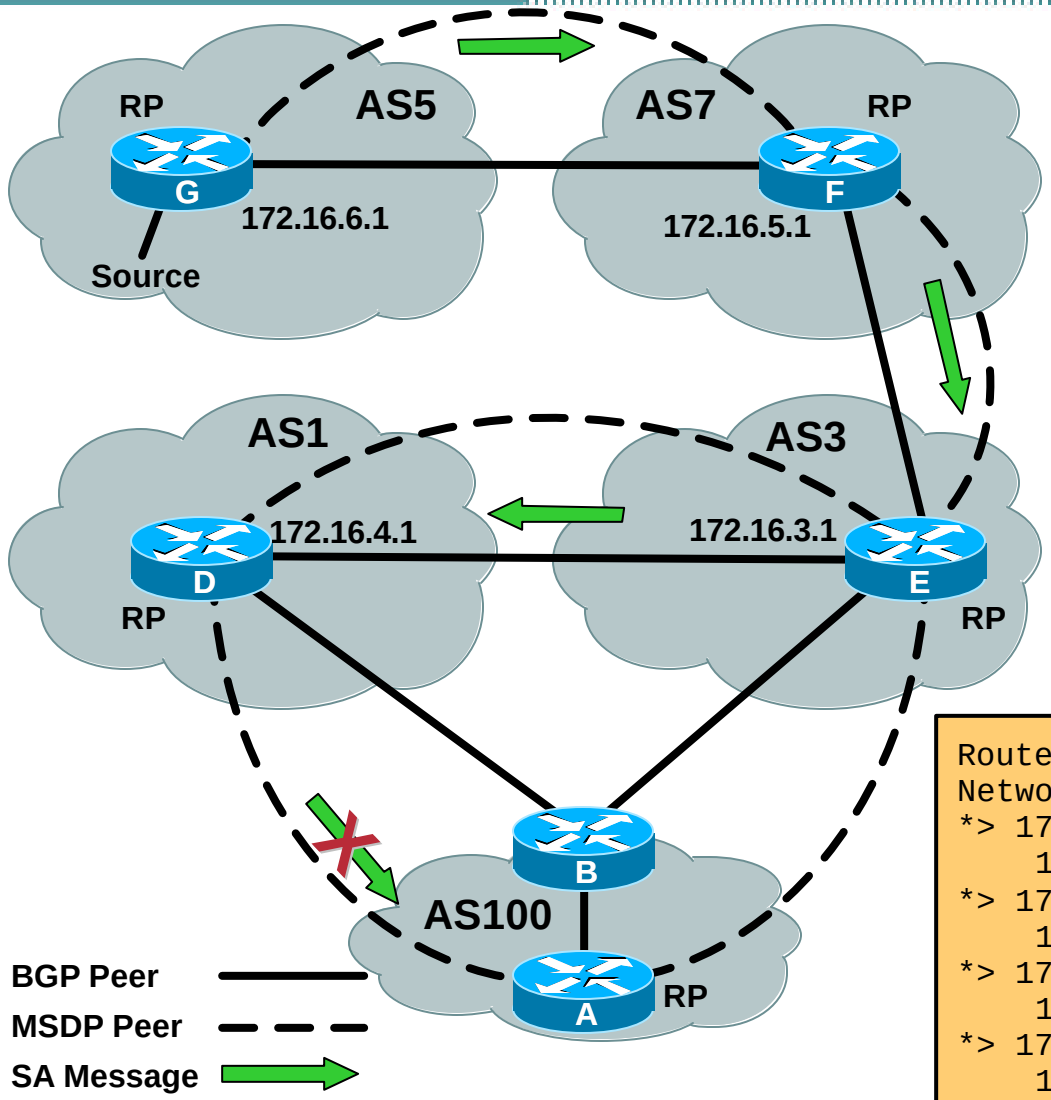
First-AS in best-path to RP = 3
AS of MSDP Peer = 3

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    172.16.4.1    [3] 7 5 i
    172.16.6.0/24    172.16.3.1    1 3 7 5 i
```

Rule3: MSDP peer != BGP peer



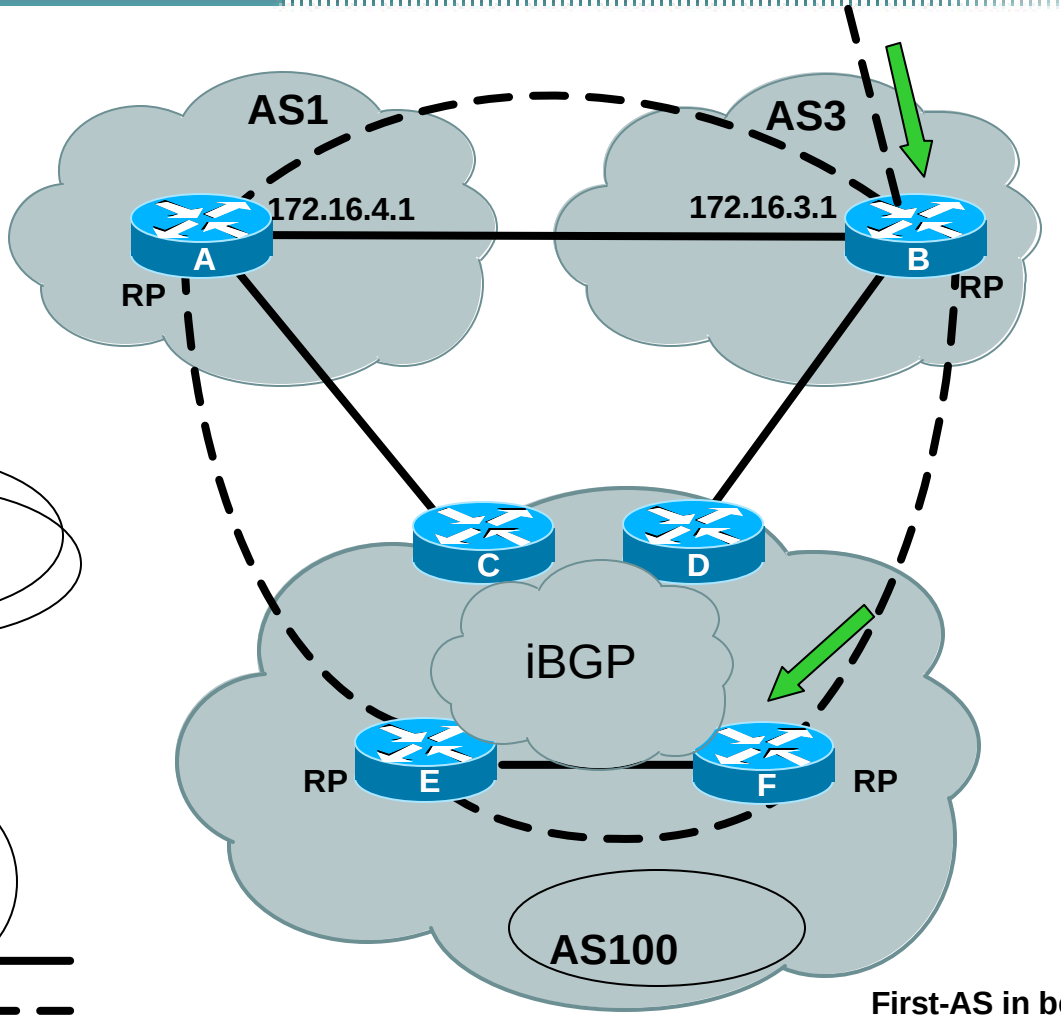
First-AS in best-path to RP = 3
AS of MSDP Peer = 1

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    172.16.4.1     [3] 7 5 i
    172.16.6.0/24    172.16.3.1     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 Configuration

- **Filtering**

  Can filter SA in/out, groups, with acls or route-maps

- **For configuration commands see:**

  ftp://ftpeng.cisco.com/ipmulticast/msdp-commands

- **For MSDP BCP (Best Current Practice) Draft:**

  draft-ietf-mboned-msdp-deploy-00.txt

Agenda

- PIM-SM review (forwarding)
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)

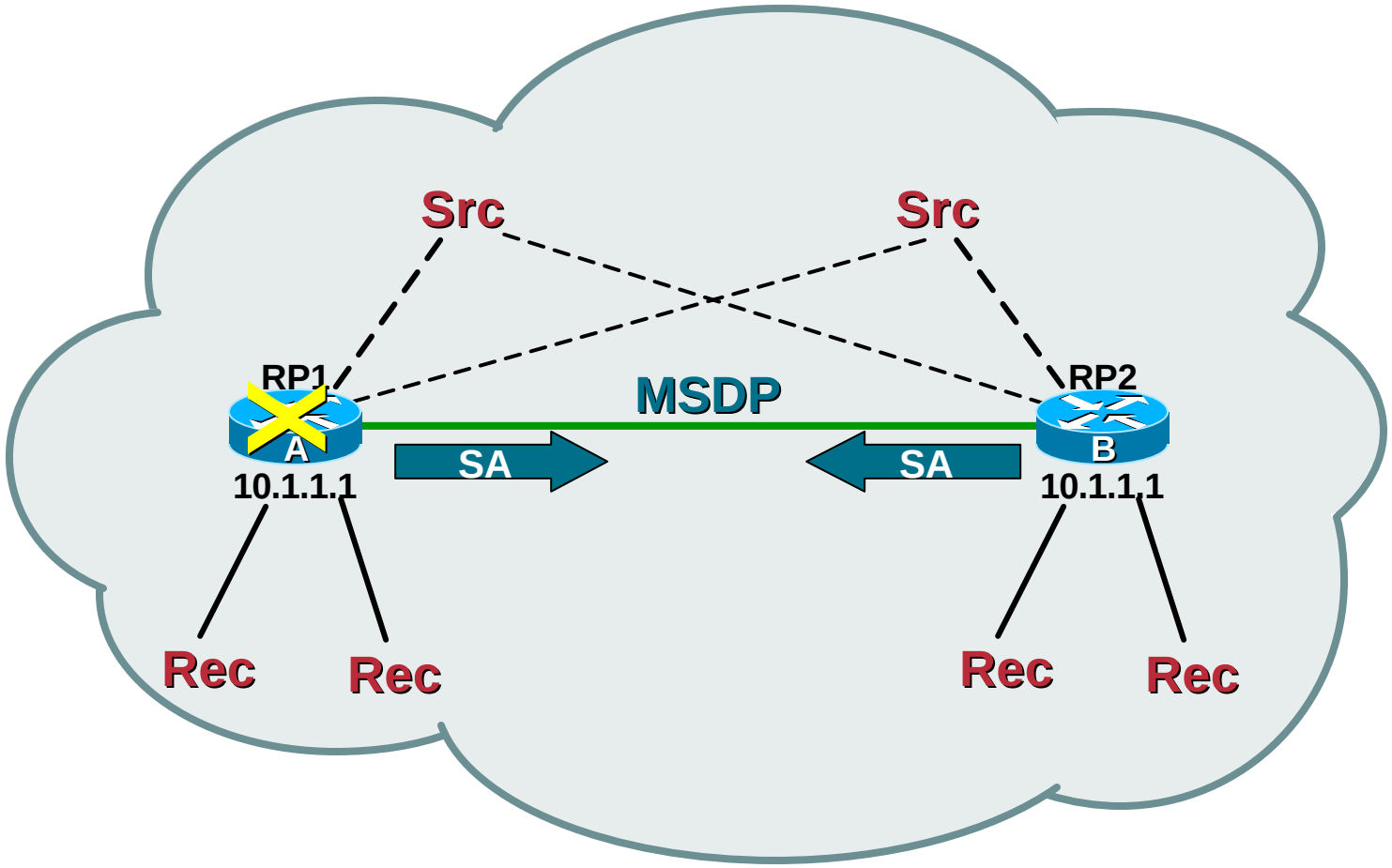
MSDP Application—Anycast RP

- draft-ietf-mboned-anycast-rp-05.txt
- 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
- May be used intra-domain (enterprise) to provide redundancy and RP load sharing

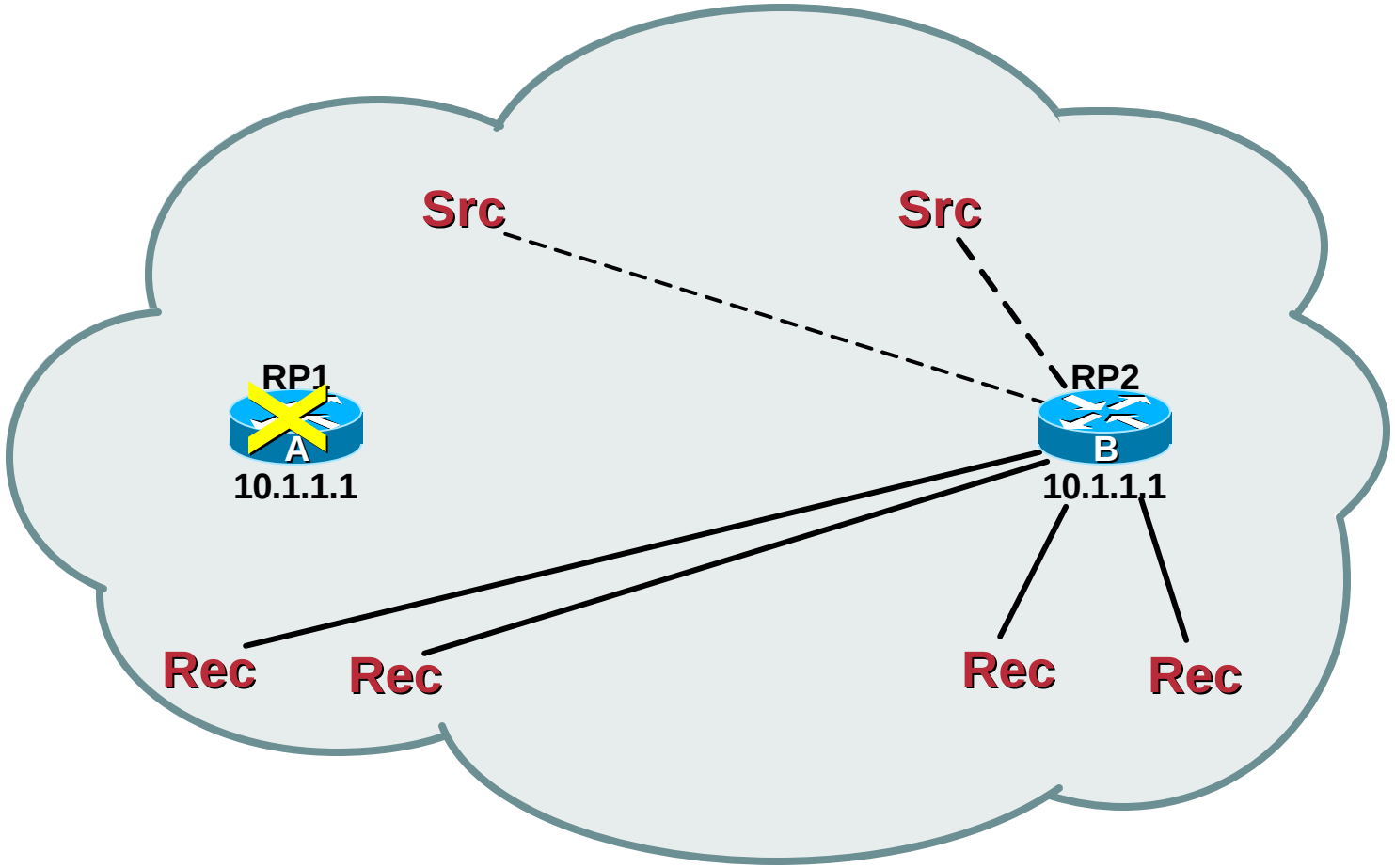
MSDP Application—Anycast RP

- 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

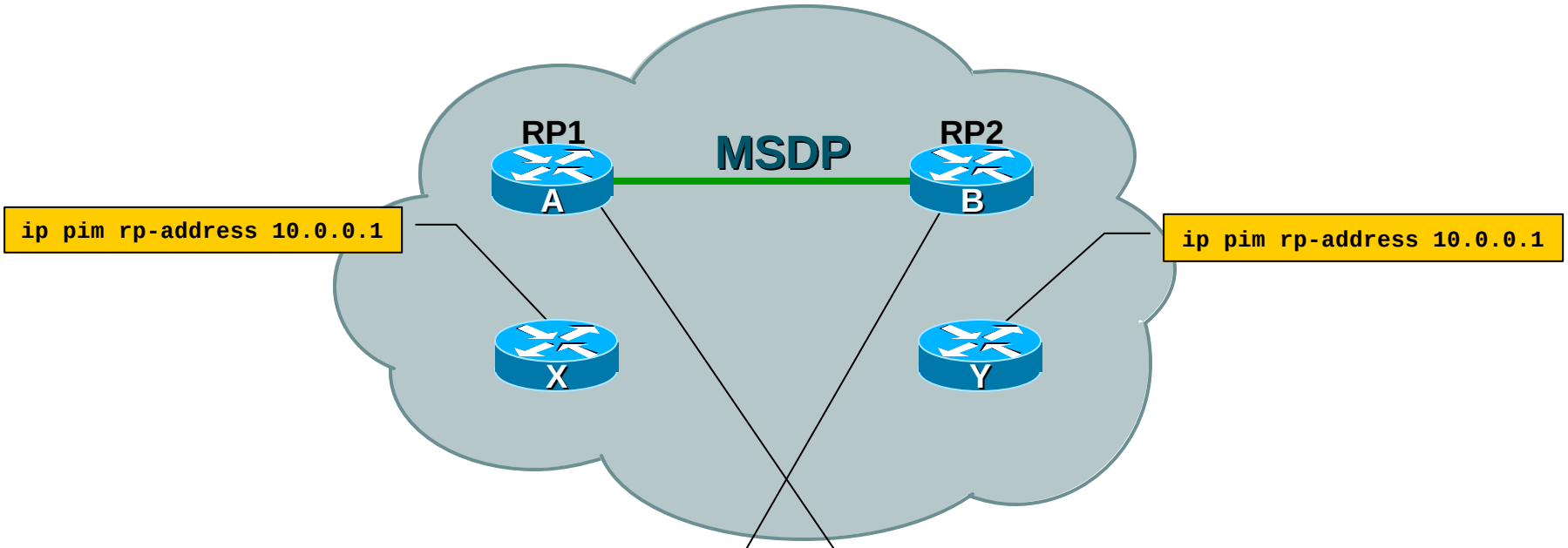
Anycast RP—Overview



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
```

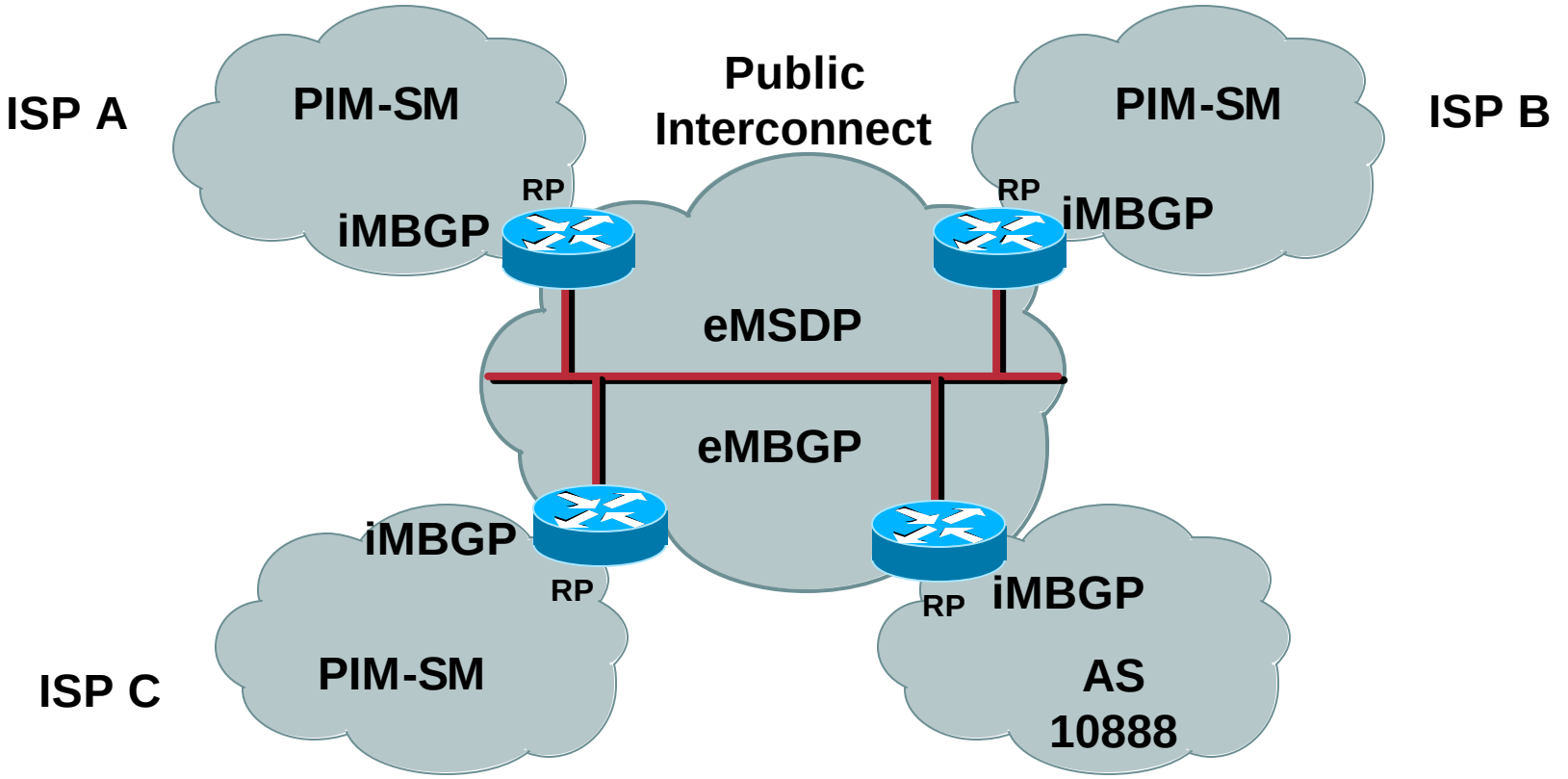
```
Interface loopback 0
 ip address 10.0.0.3 255.255.255.255

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

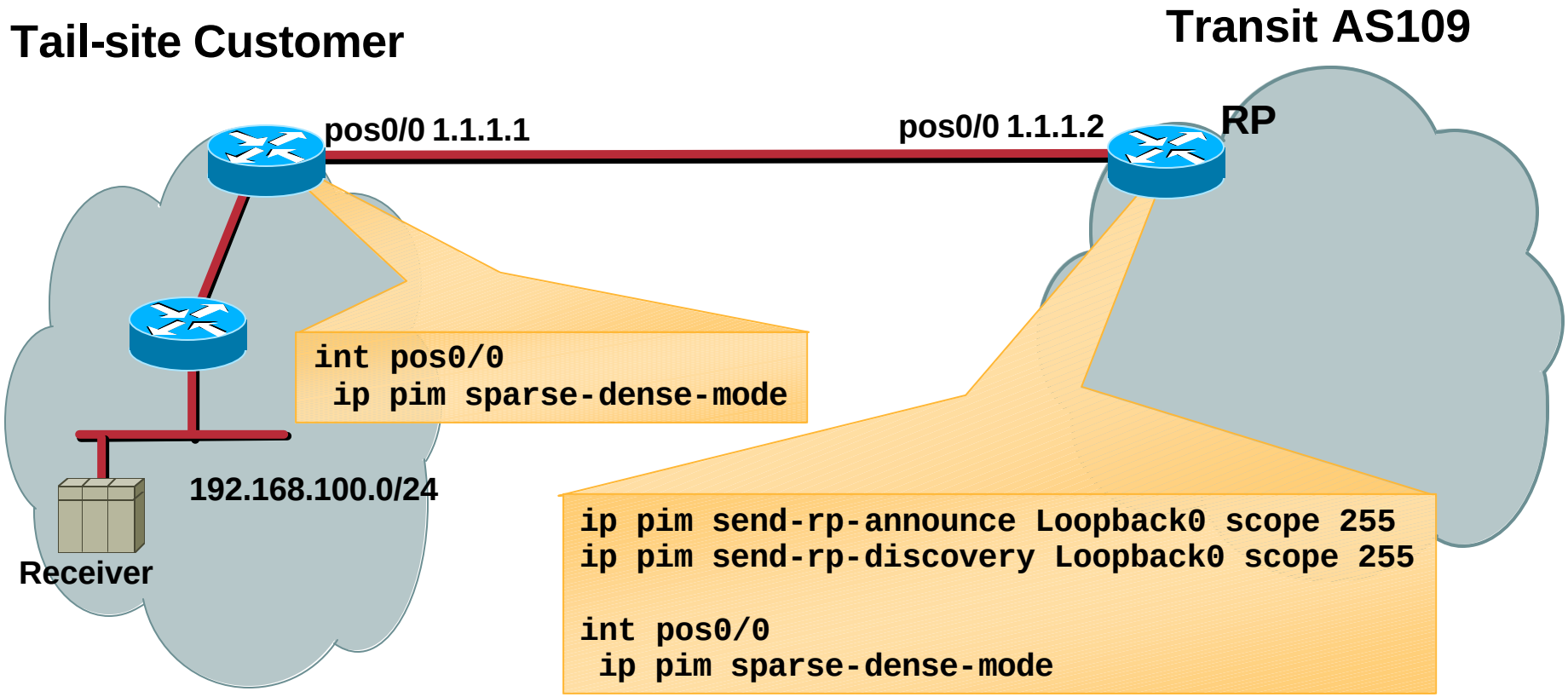
Cisco.com

ftp://ftp-eng.cisco.com/ipmulticast/msdp-sa-filter.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
```

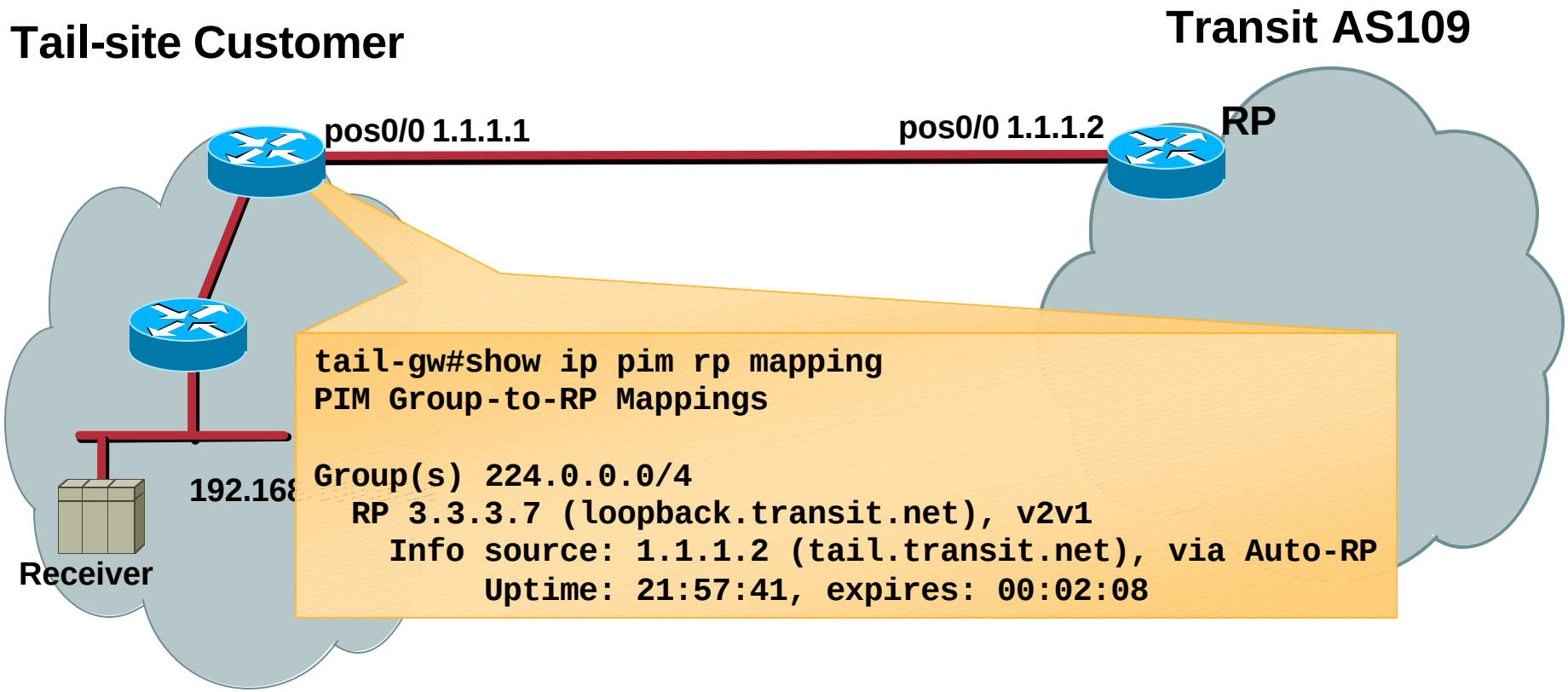
Single-Homed, ISP RP, Non-MBGP

PIM Border Constraints



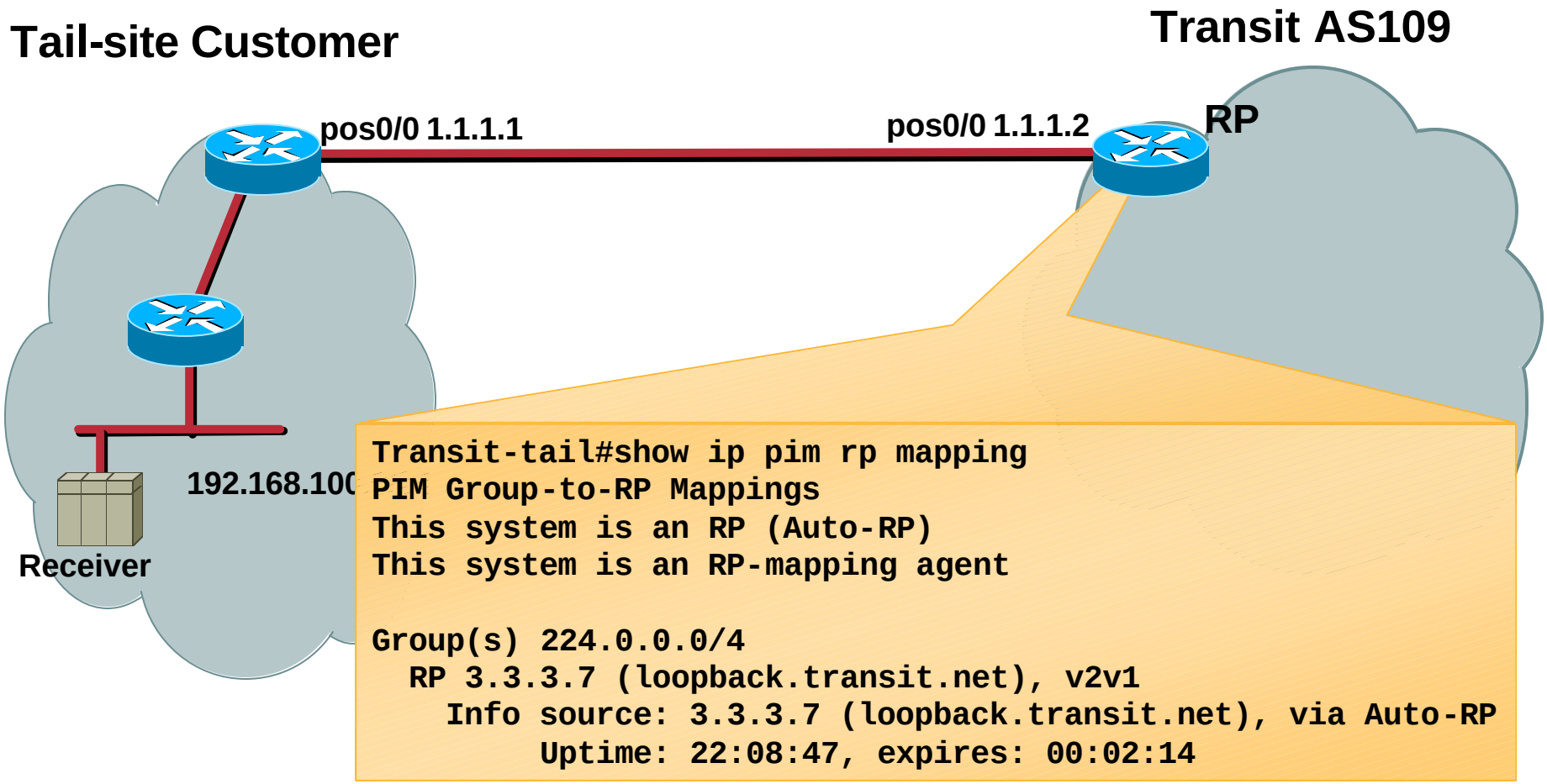
Single-Homed, ISP RP, Non-MBGP

Checking PIM Border (RP Mapping)



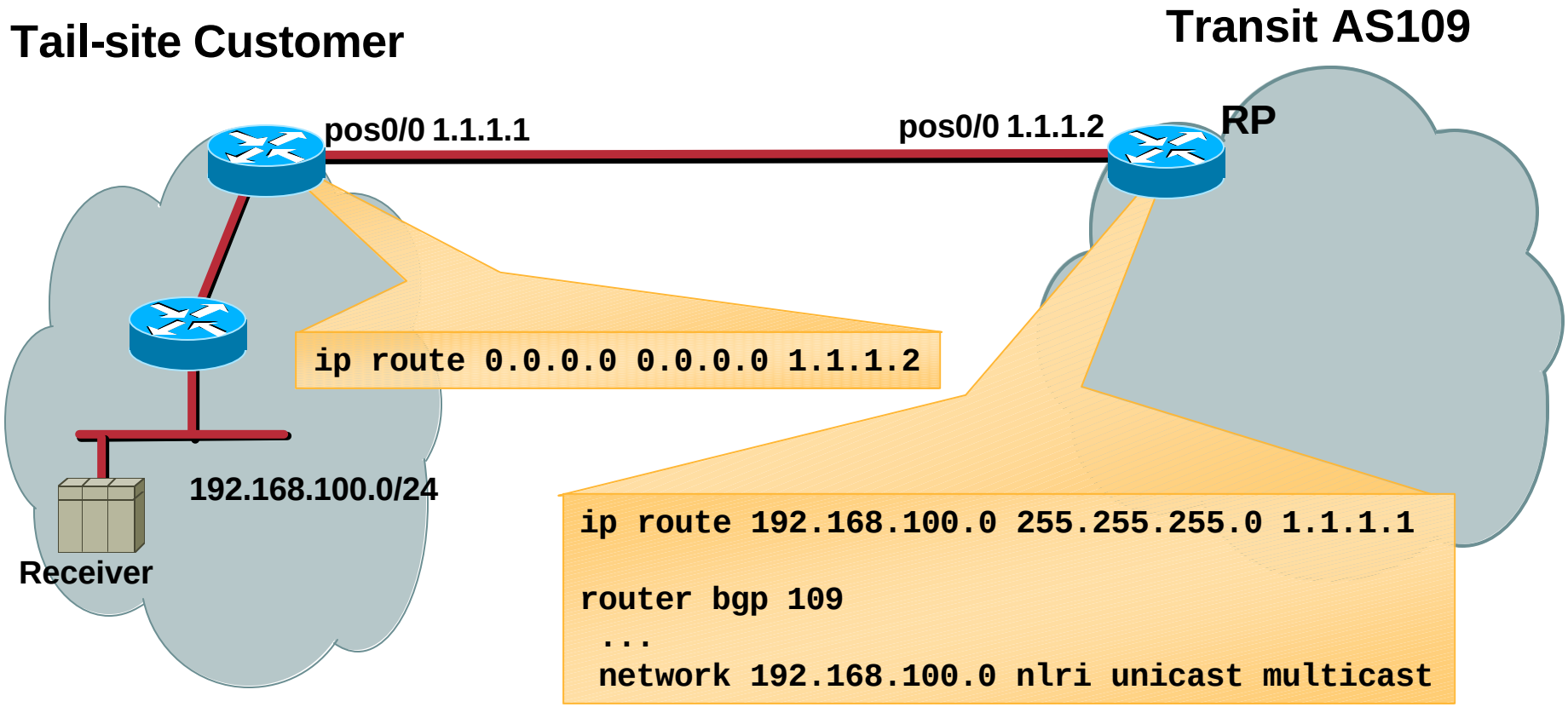
Single-Homed, ISP RP, Non-MBGP

Checking PIM Border (RP Mapping)



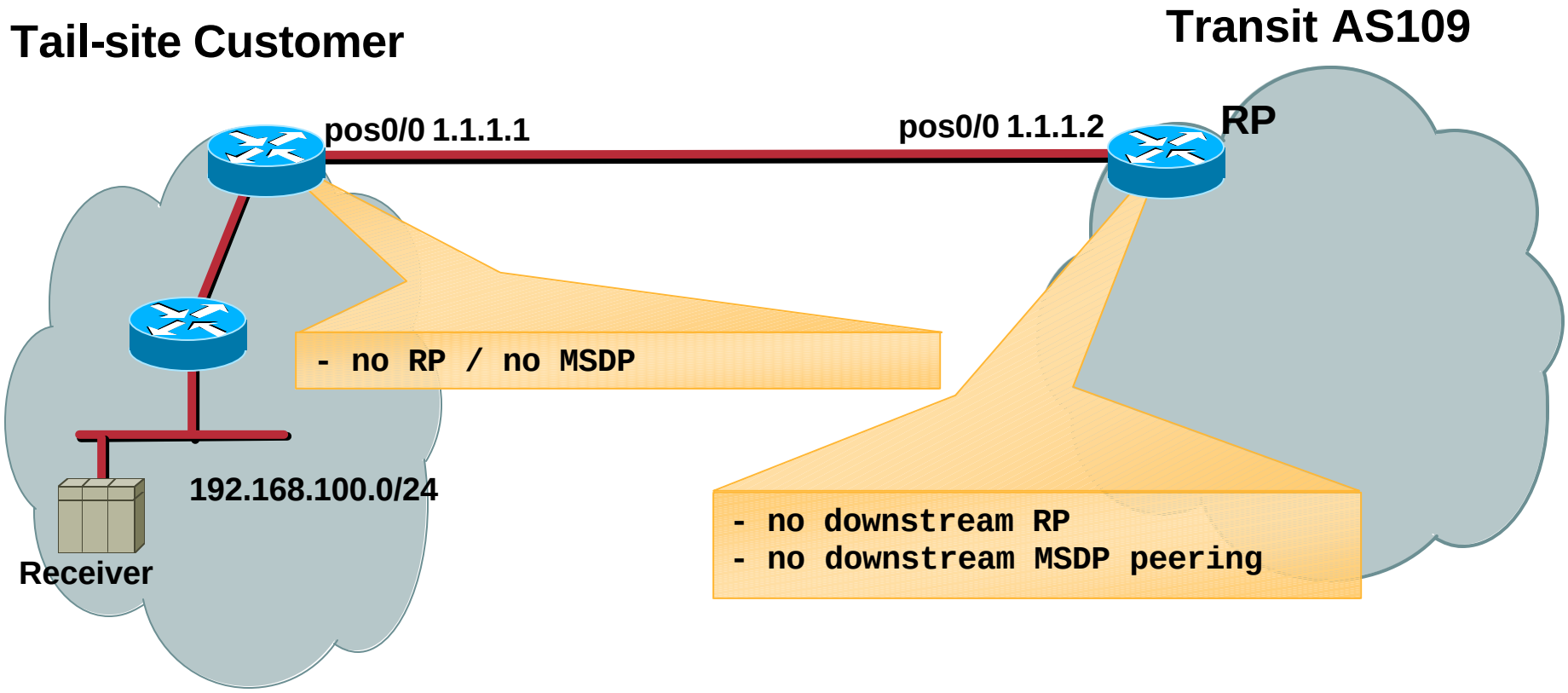
Single-Homed, ISP RP, Non-MBGP

Border RPF Check



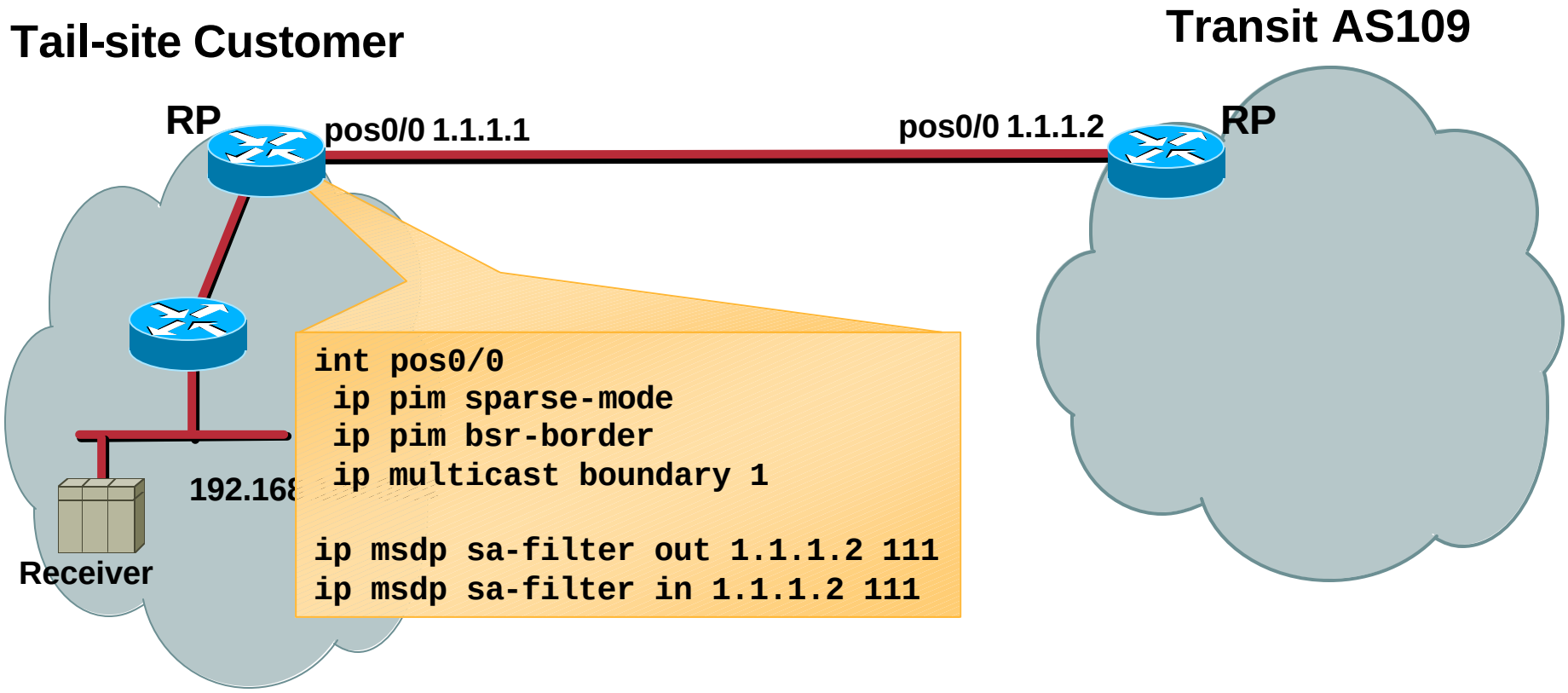
Single-Homed, ISP RP, Non-MBGP

MSDP RPF Check



Single-Homed, Customer RP, Non-MBGP

PIM Border Constraints

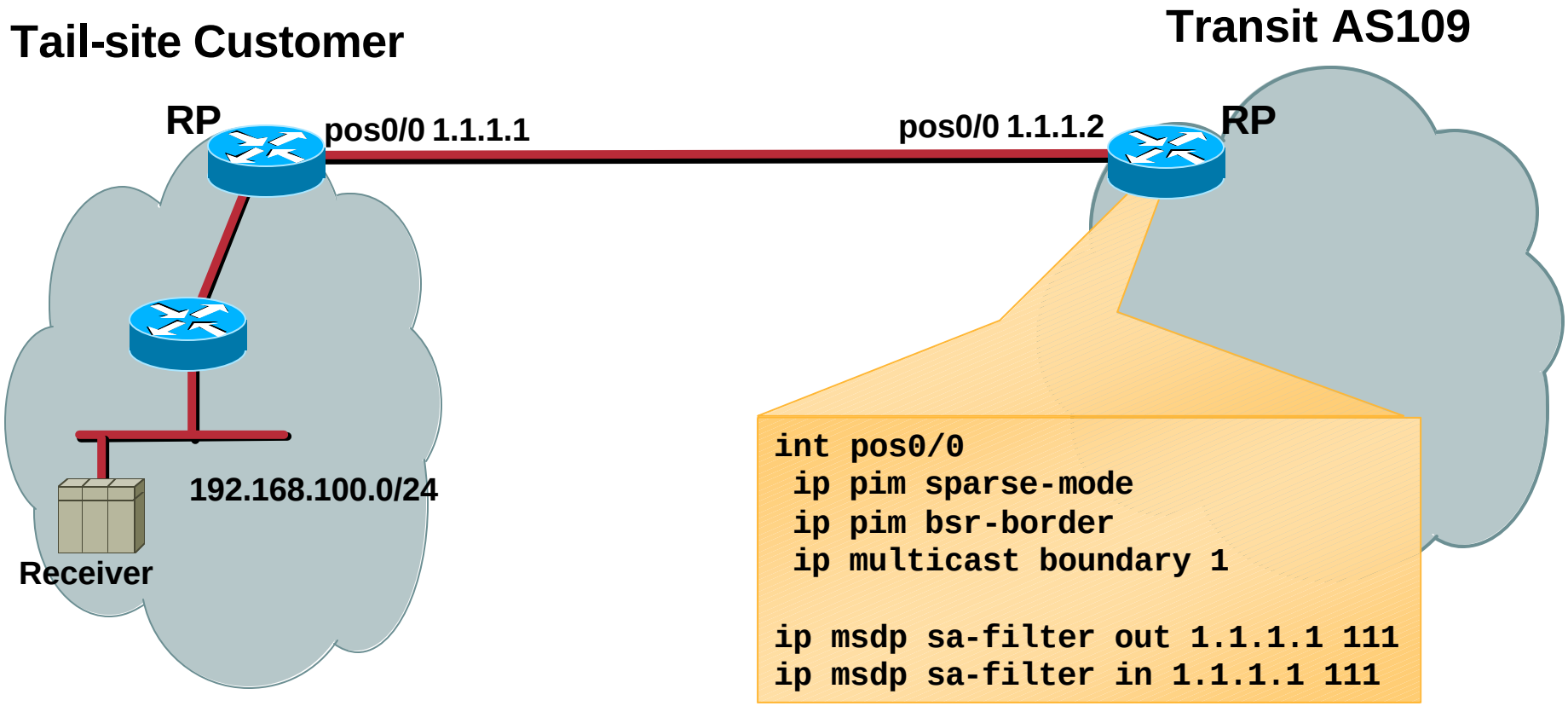


Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, Non-MBGP

Cisco.com

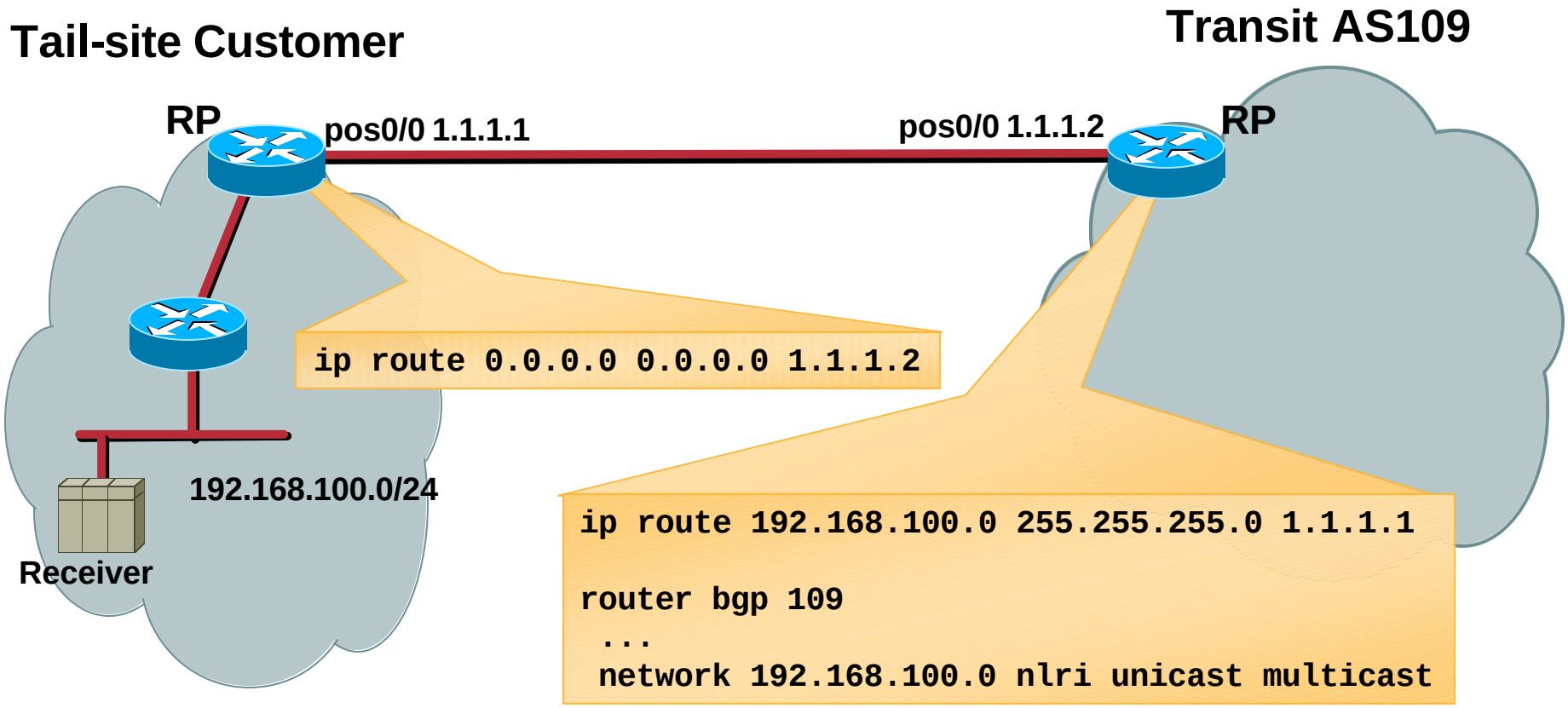
PIM Border Constraints



Note: Access-list 111 = Recommended SA Filter

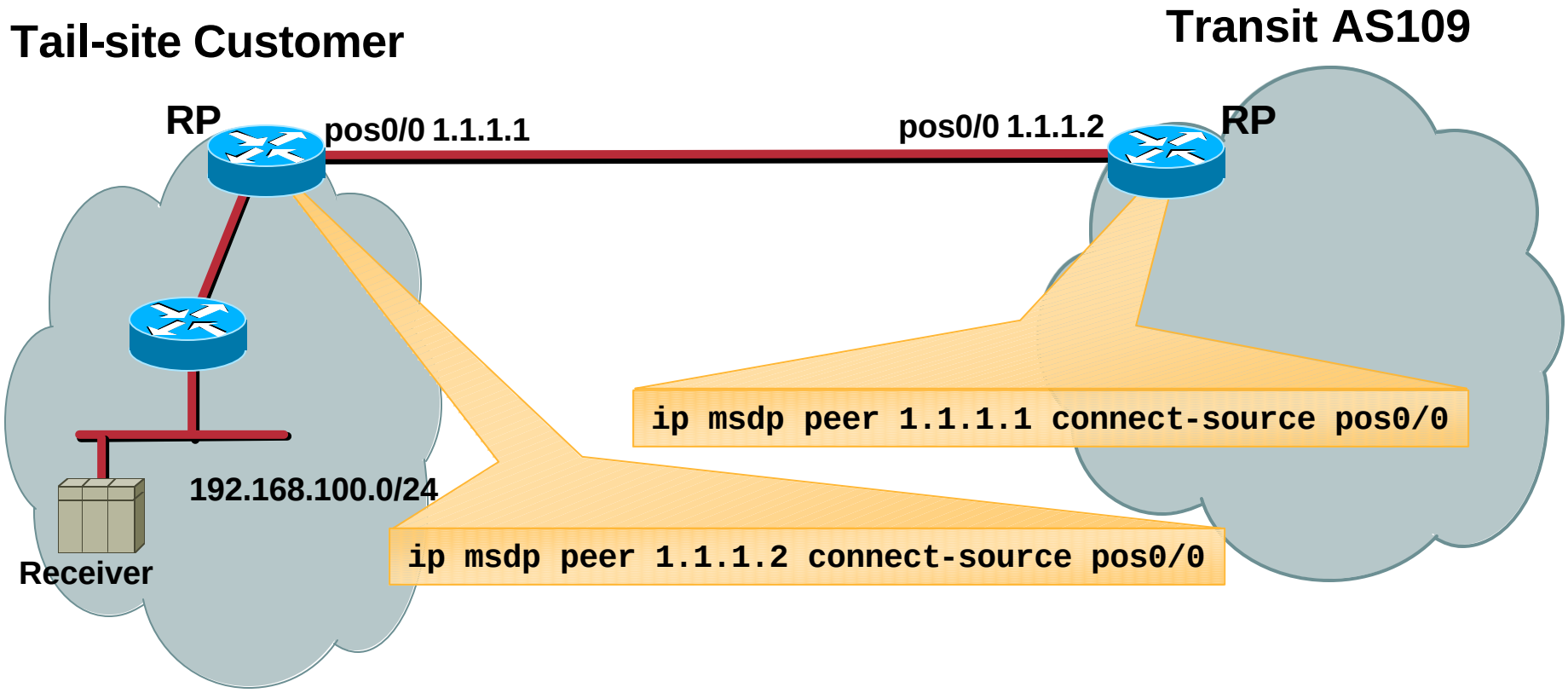
Single-Homed, Customer RP, Non-MBGP

Border RPF Check



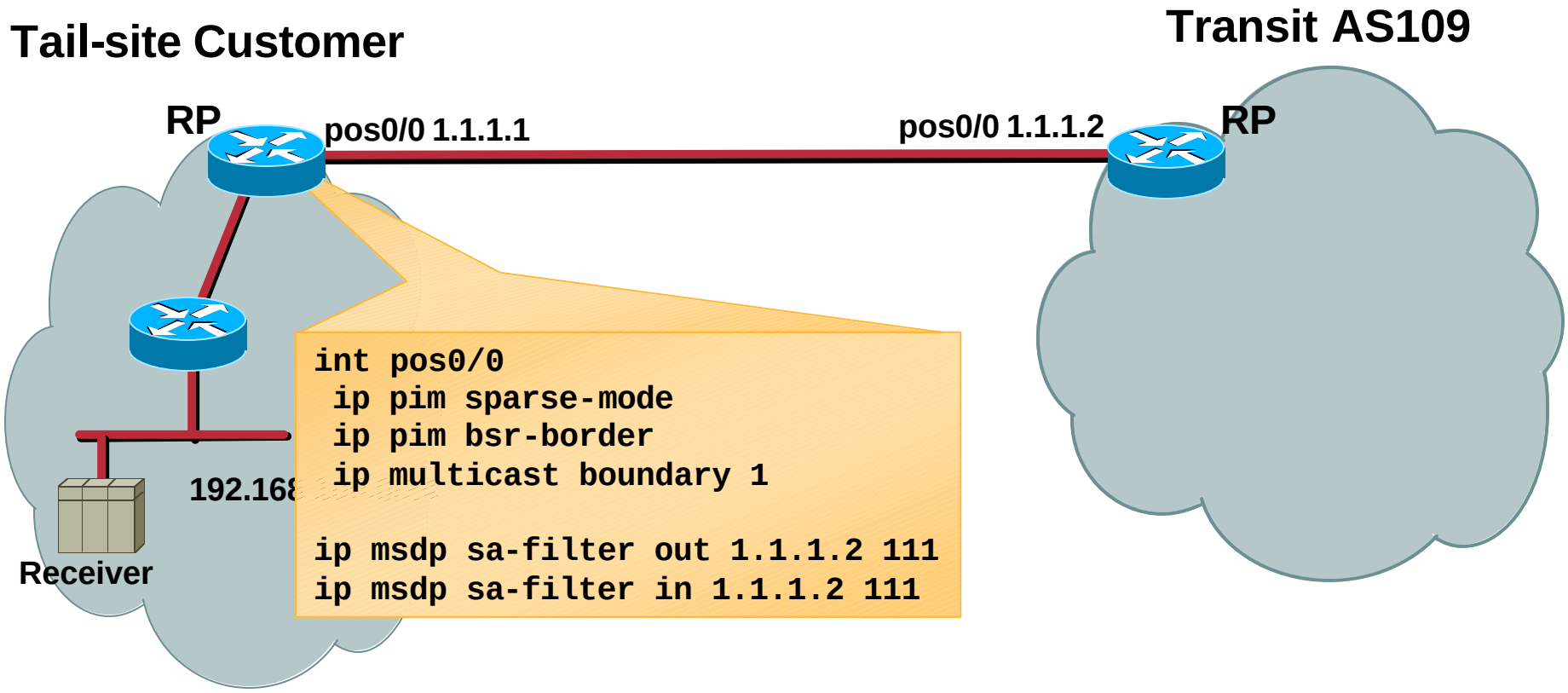
Single-Homed, Customer RP, Non-MBGP

MSDP RPF Check



Single-Homed, Customer RP, MBGP

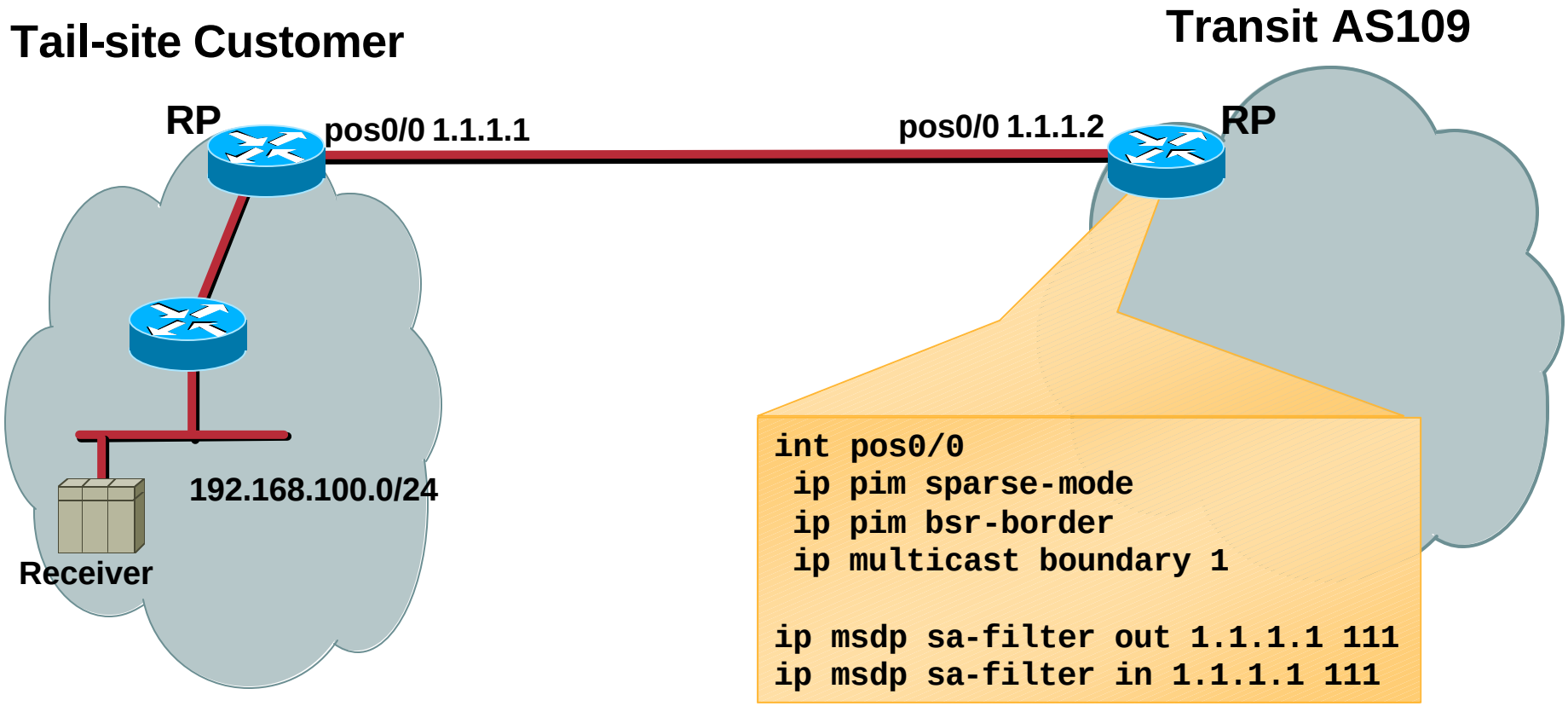
PIM Border Constraints



Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, MBGP

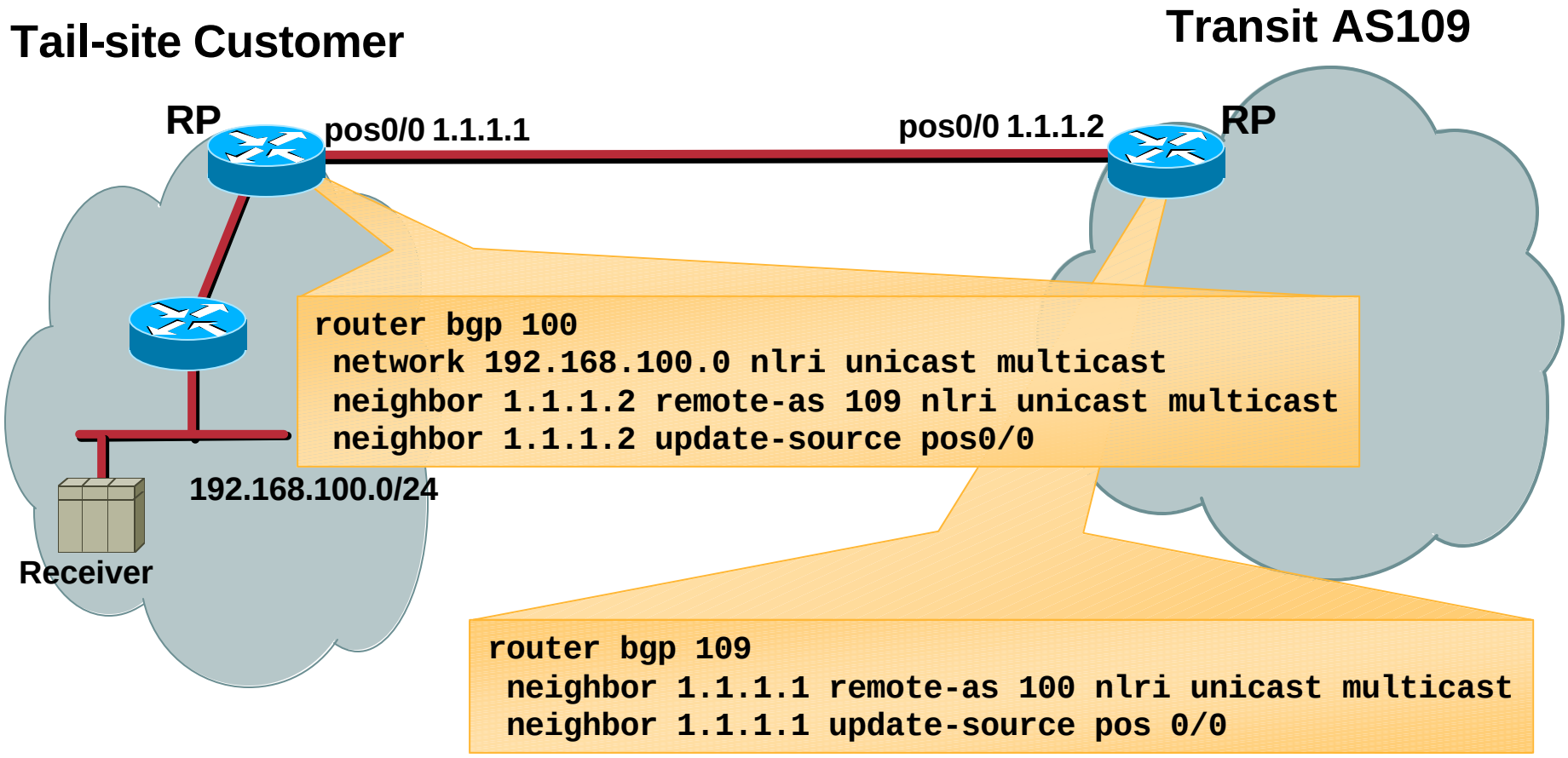
PIM Border Constraints



Note: Access-list 111 = Recommended SA Filter

Single-Homed, Customer RP, MBGP

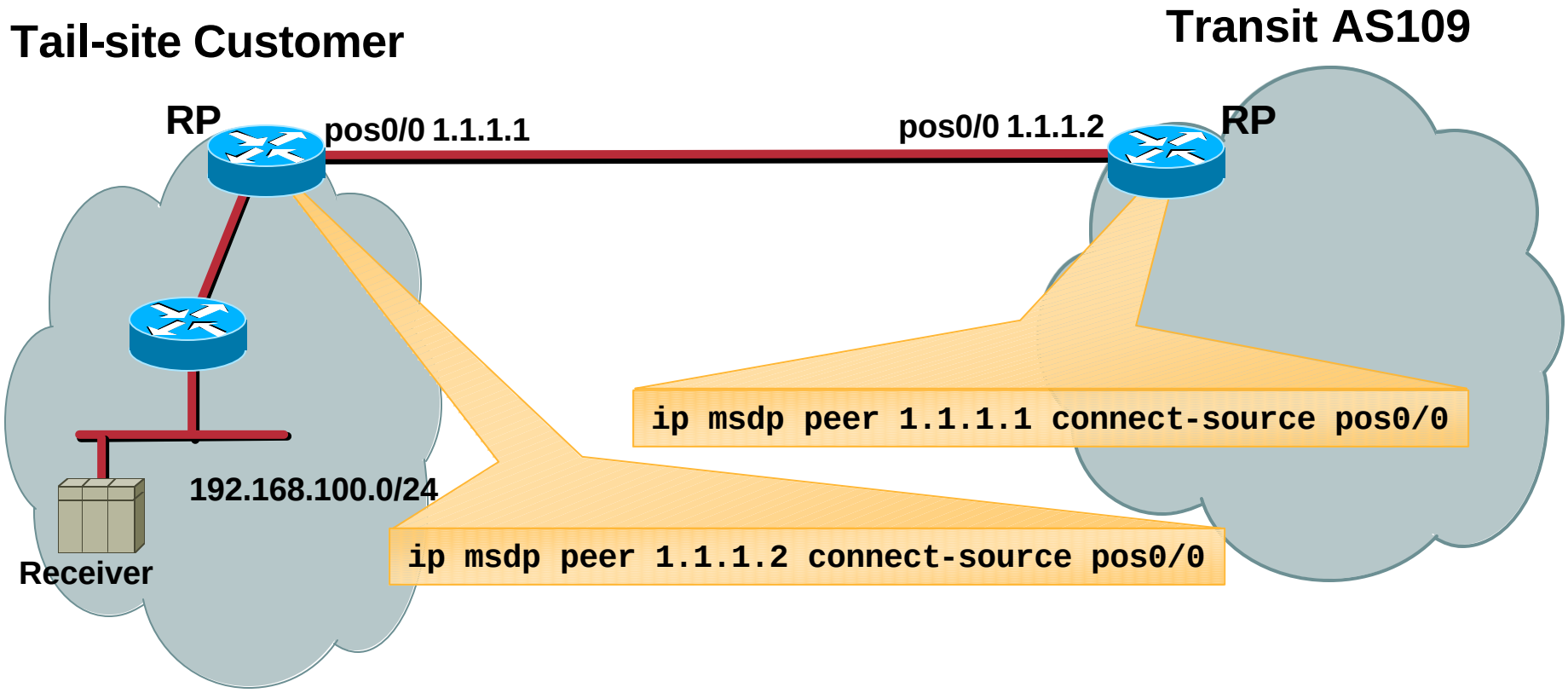
Border RPF Check



Single-Homed, Customer RP, MBGP

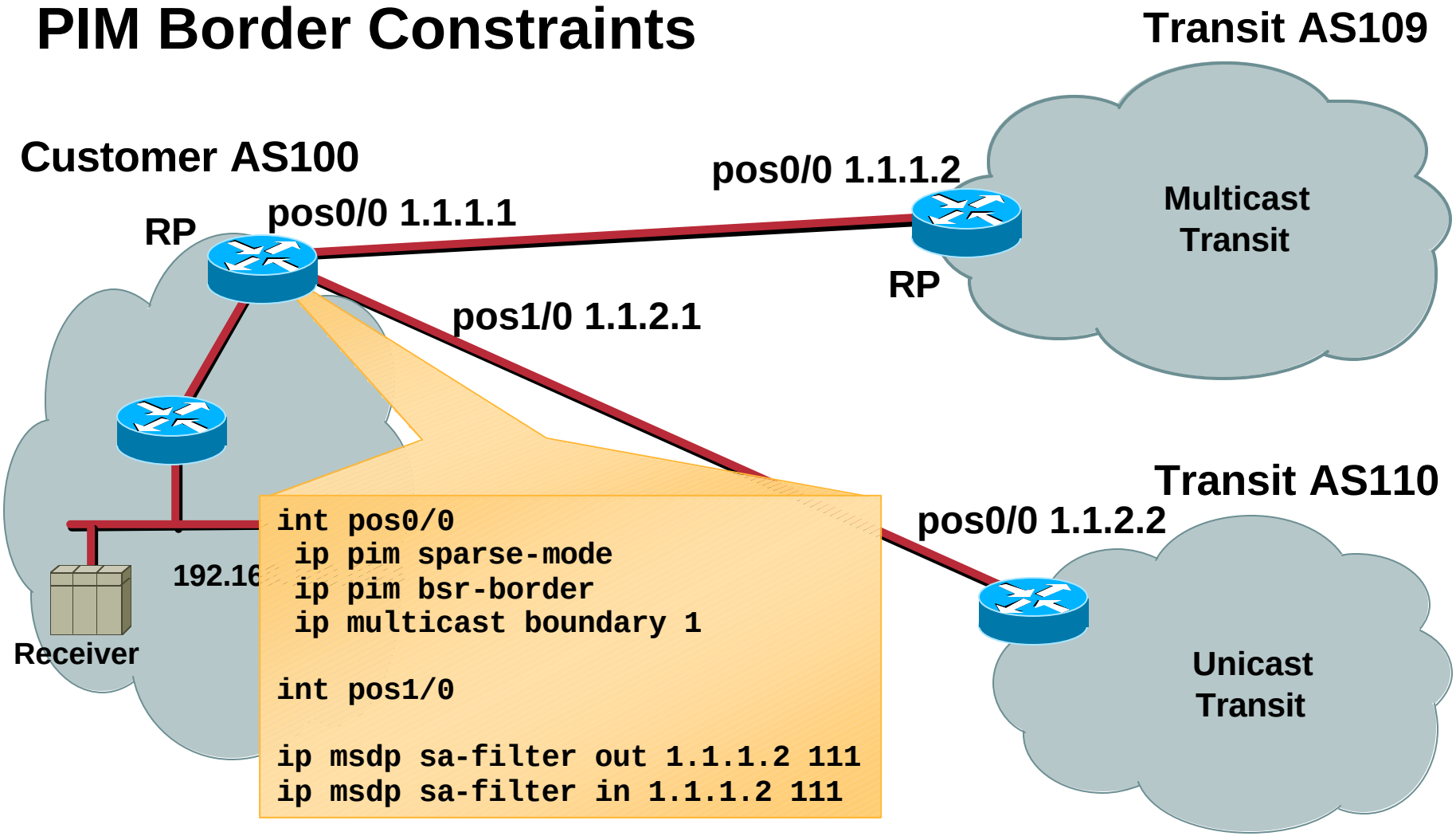
Cisco.com

MSDP RPF Check



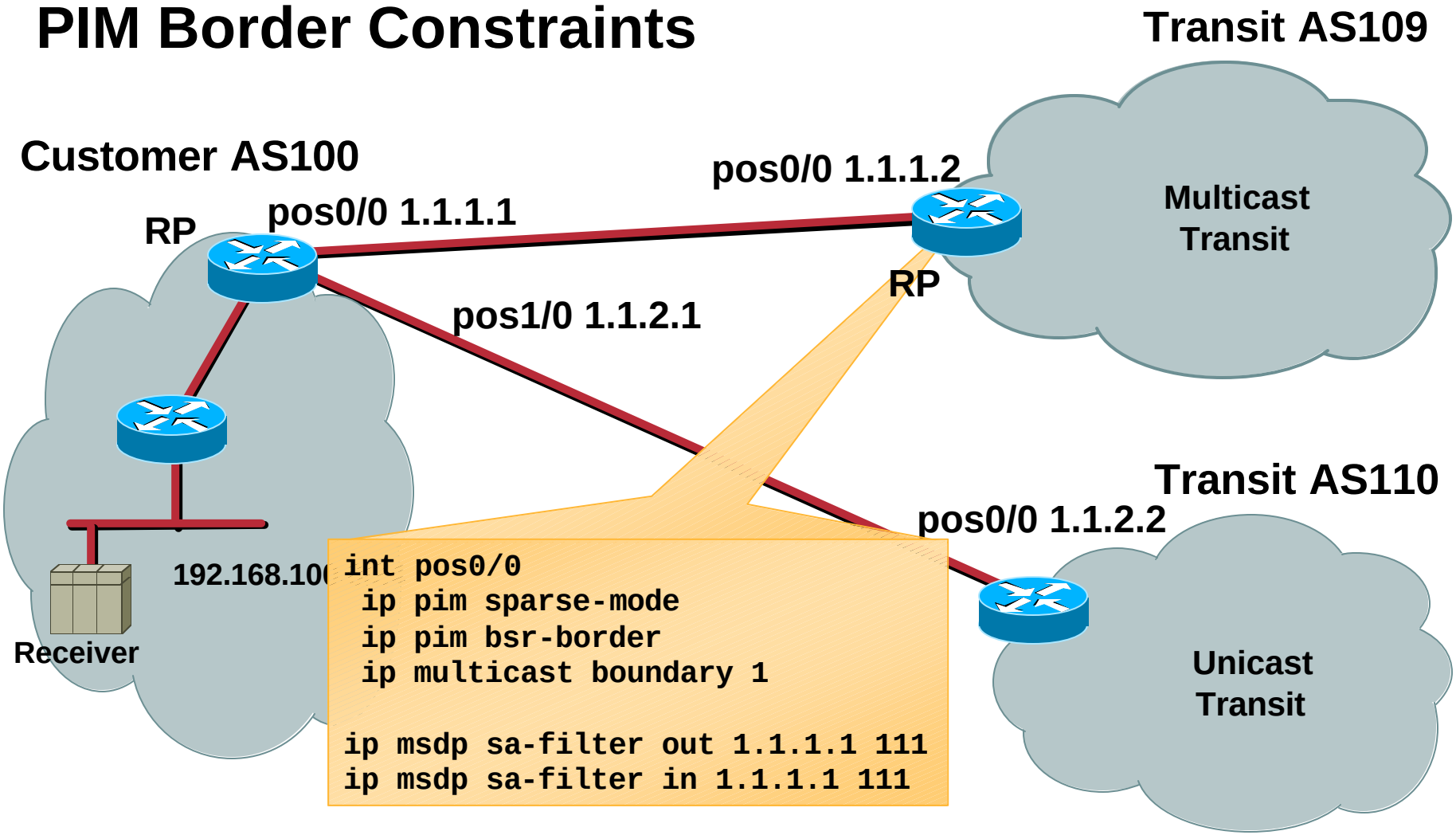
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

PIM Border Constraints



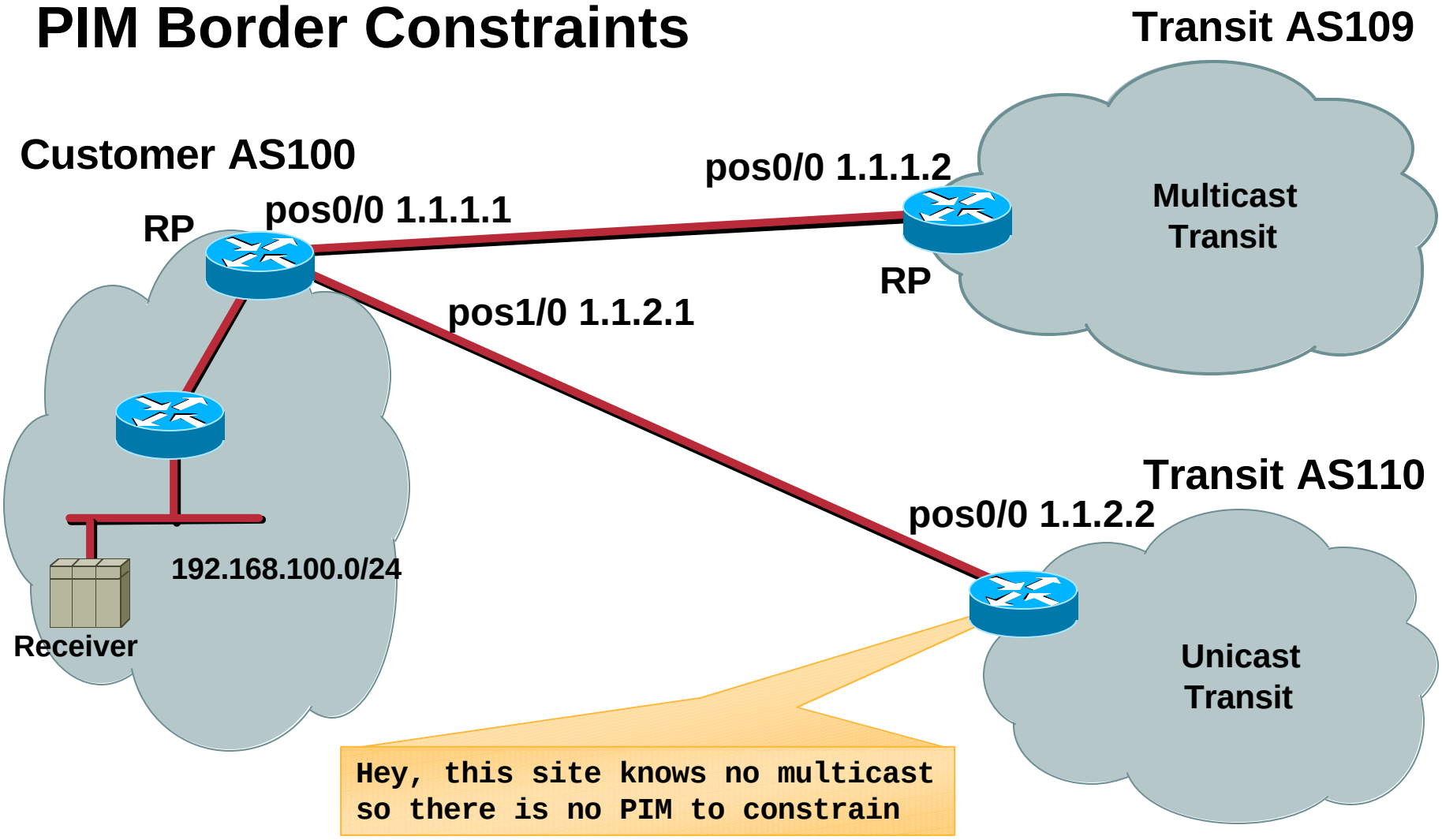
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

PIM Border Constraints



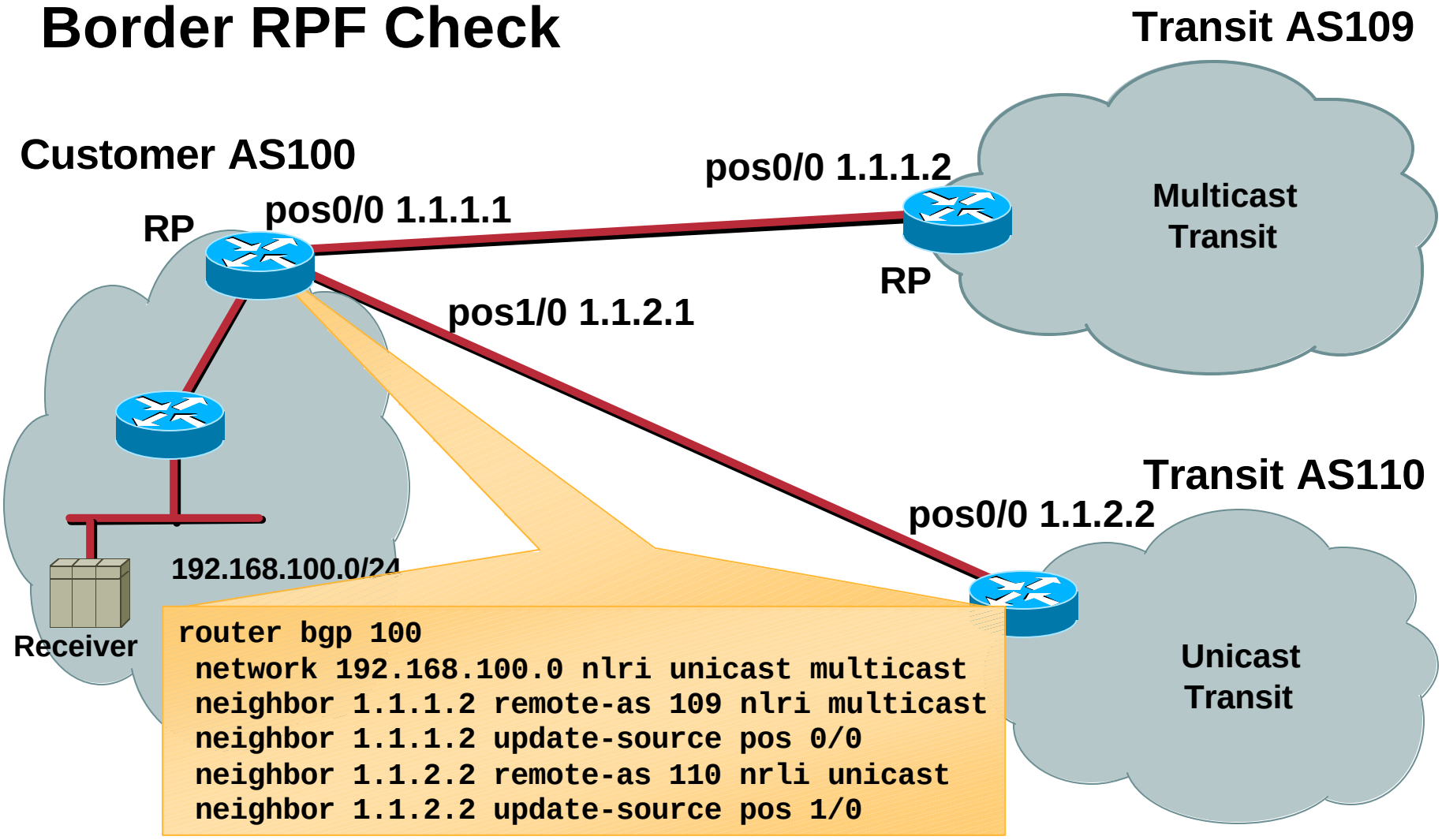
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

PIM Border Constraints



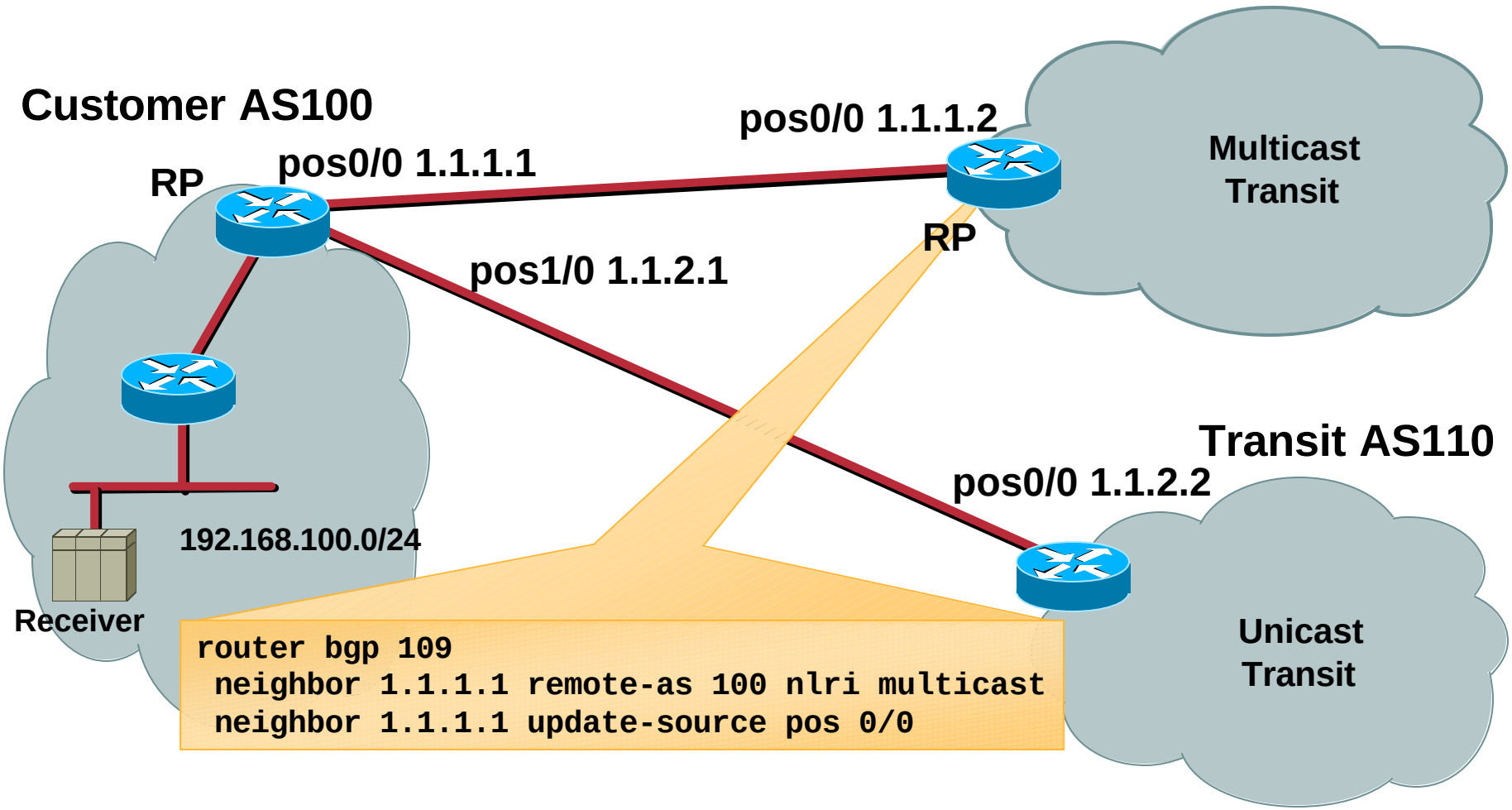
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Border RPF Check



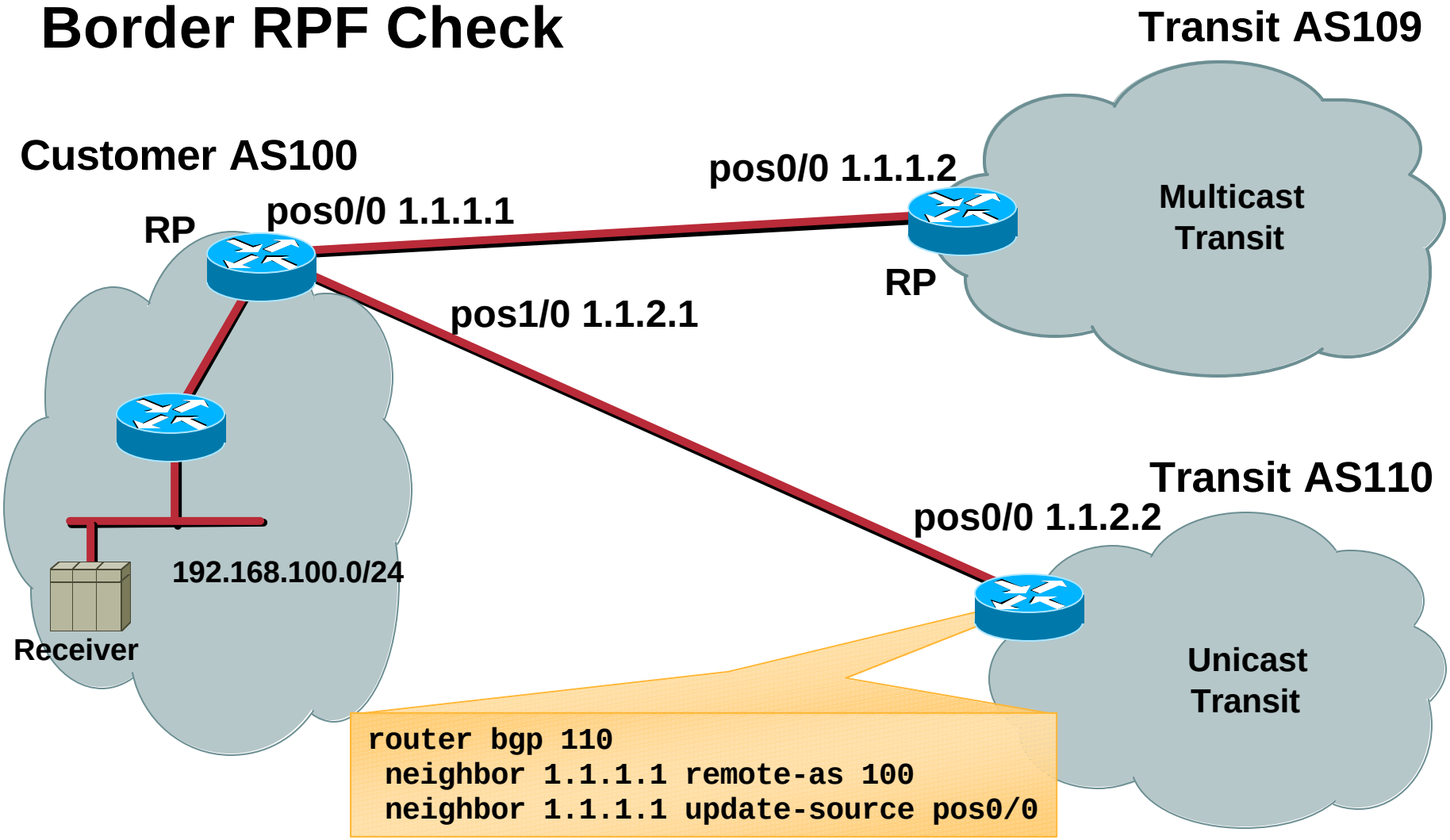
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Border RPF Check



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

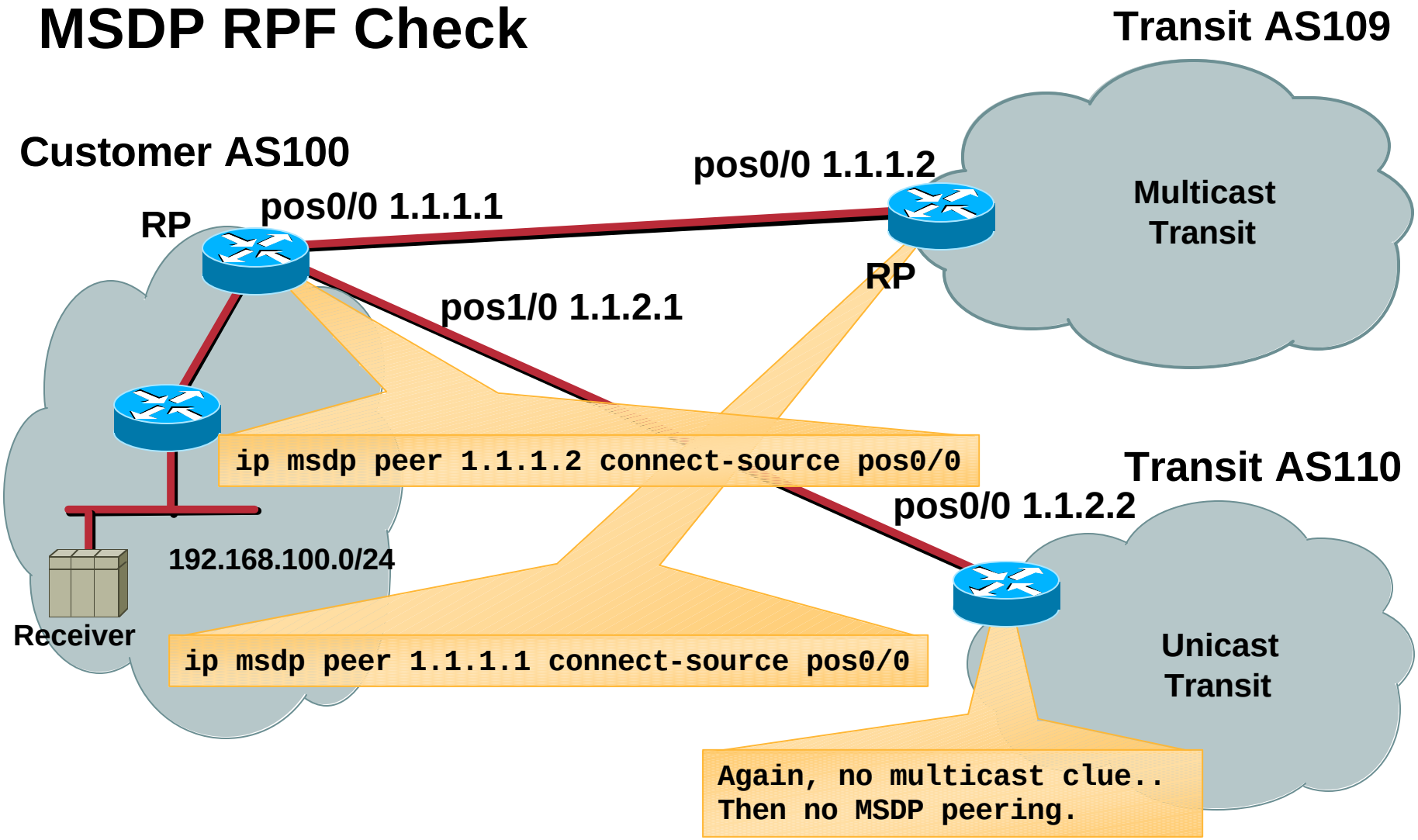
Border RPF Check



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

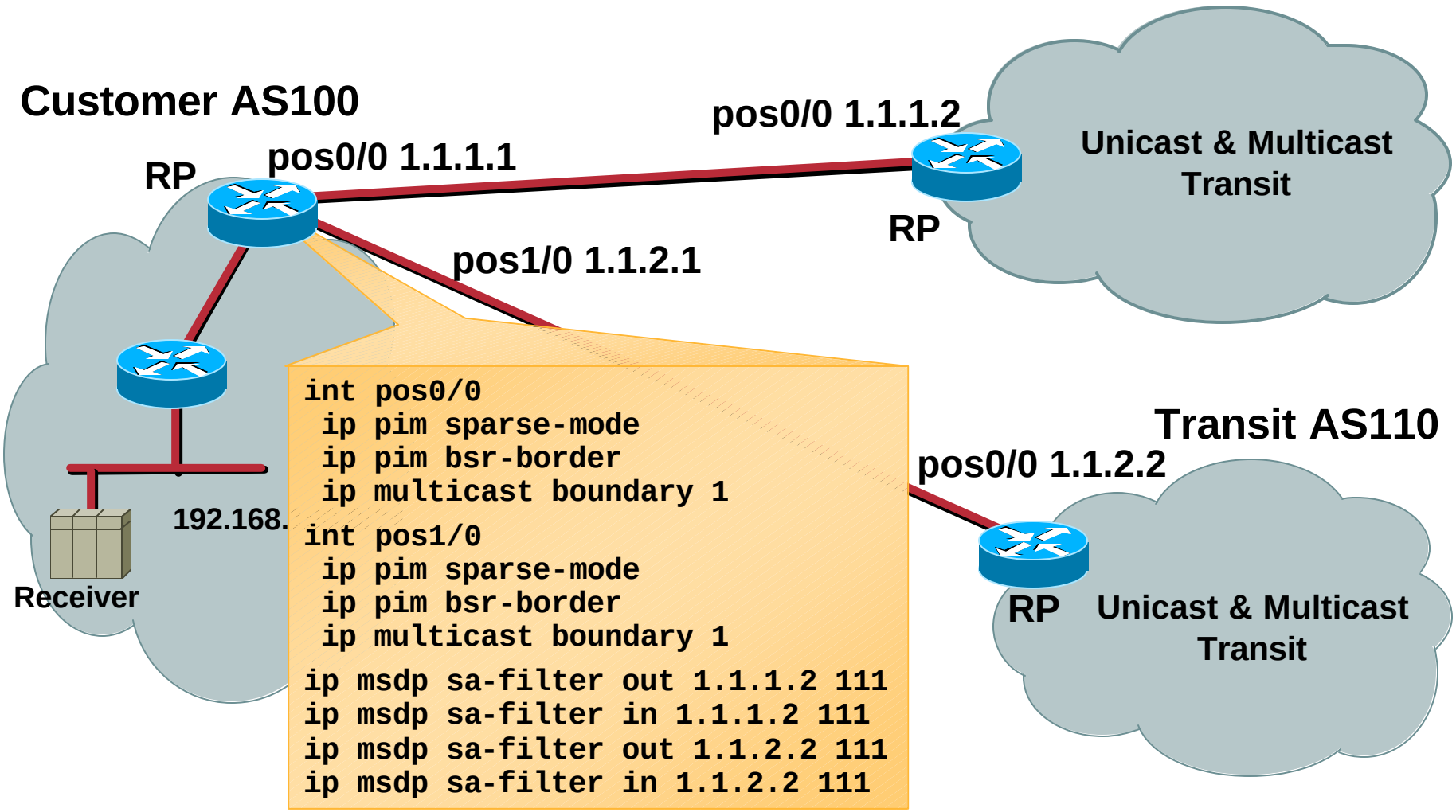
Cisco.com

MSDP RPF Check



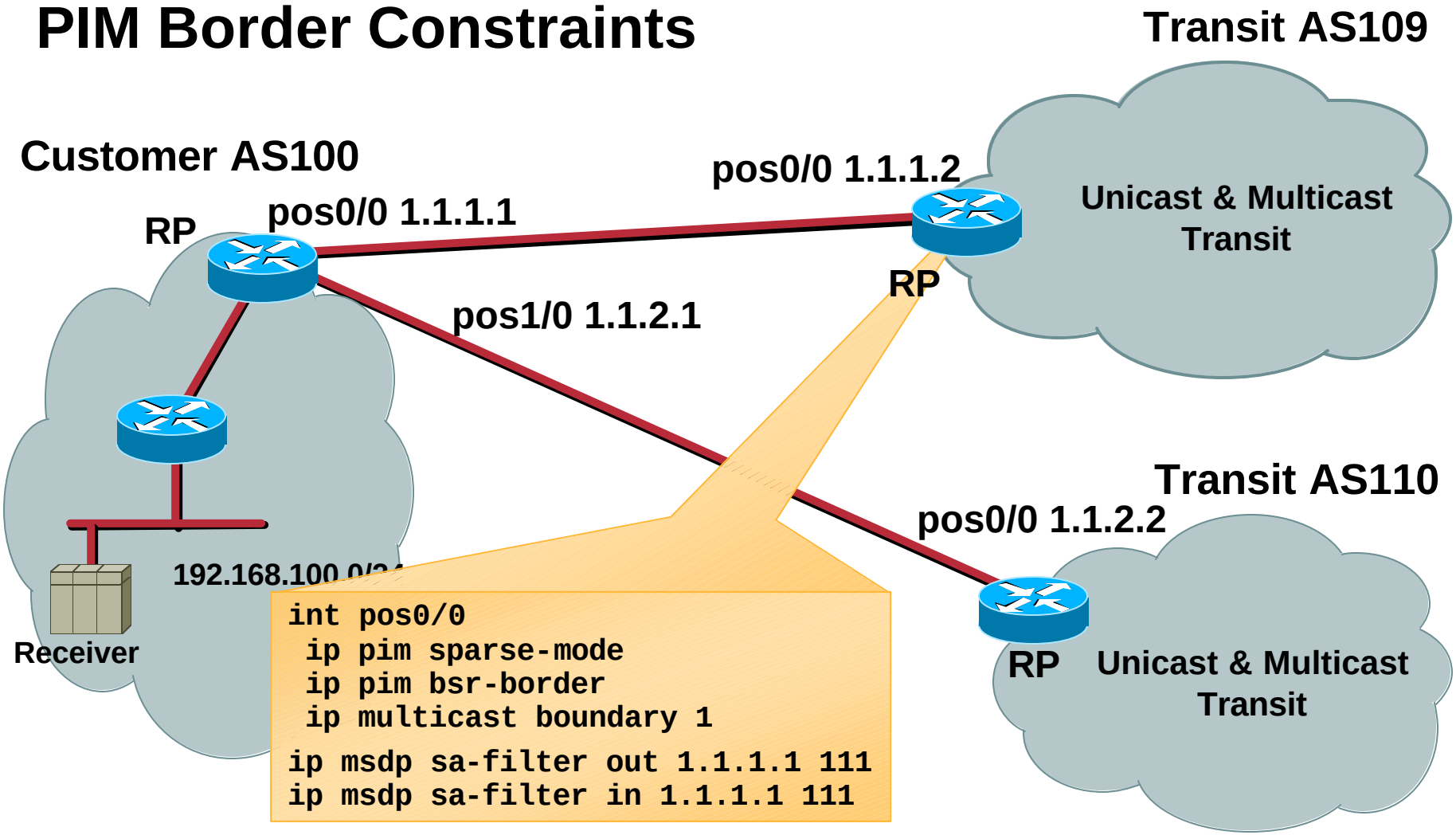
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

PIM Border Constraints



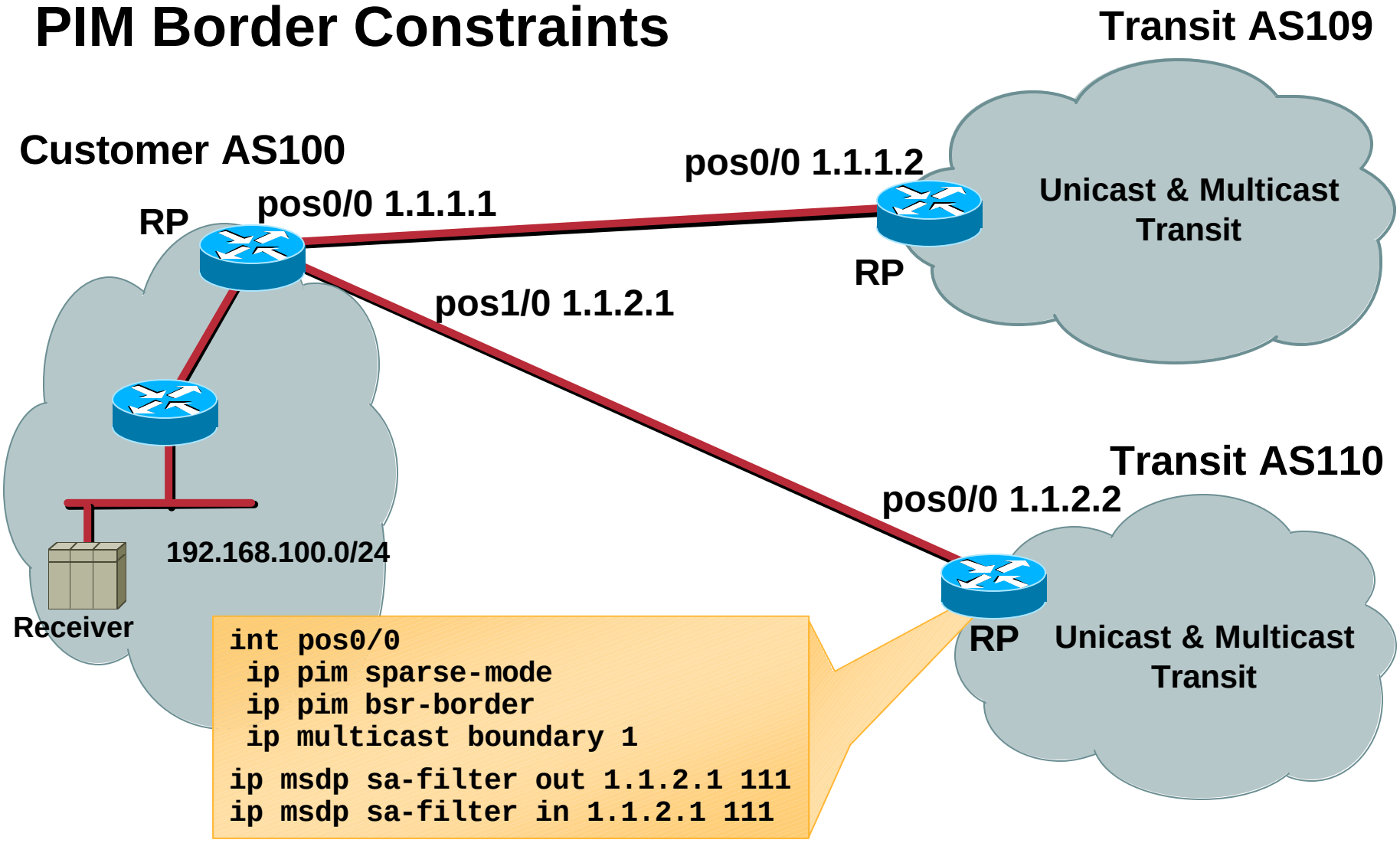
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

PIM Border Constraints



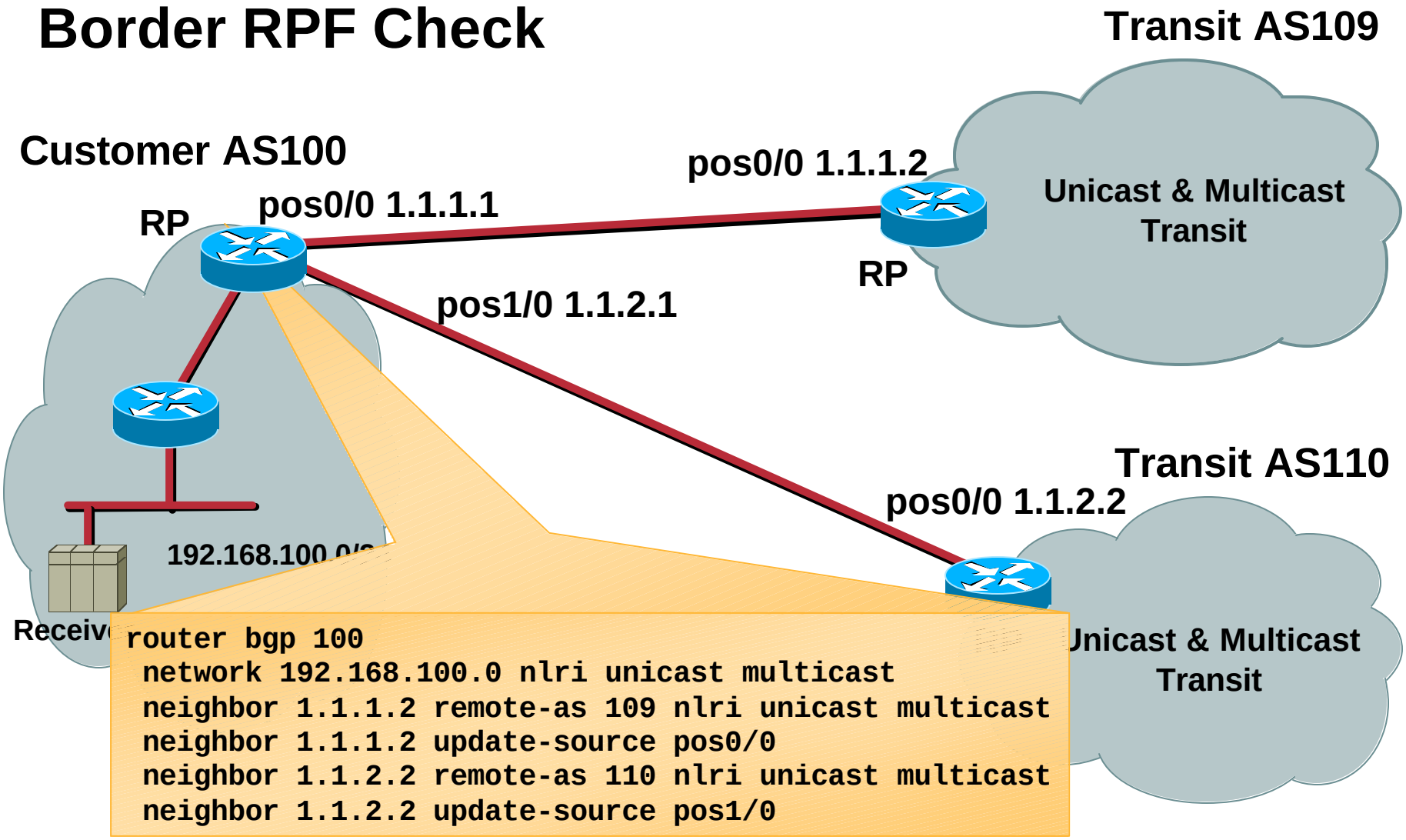
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

PIM Border Constraints



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

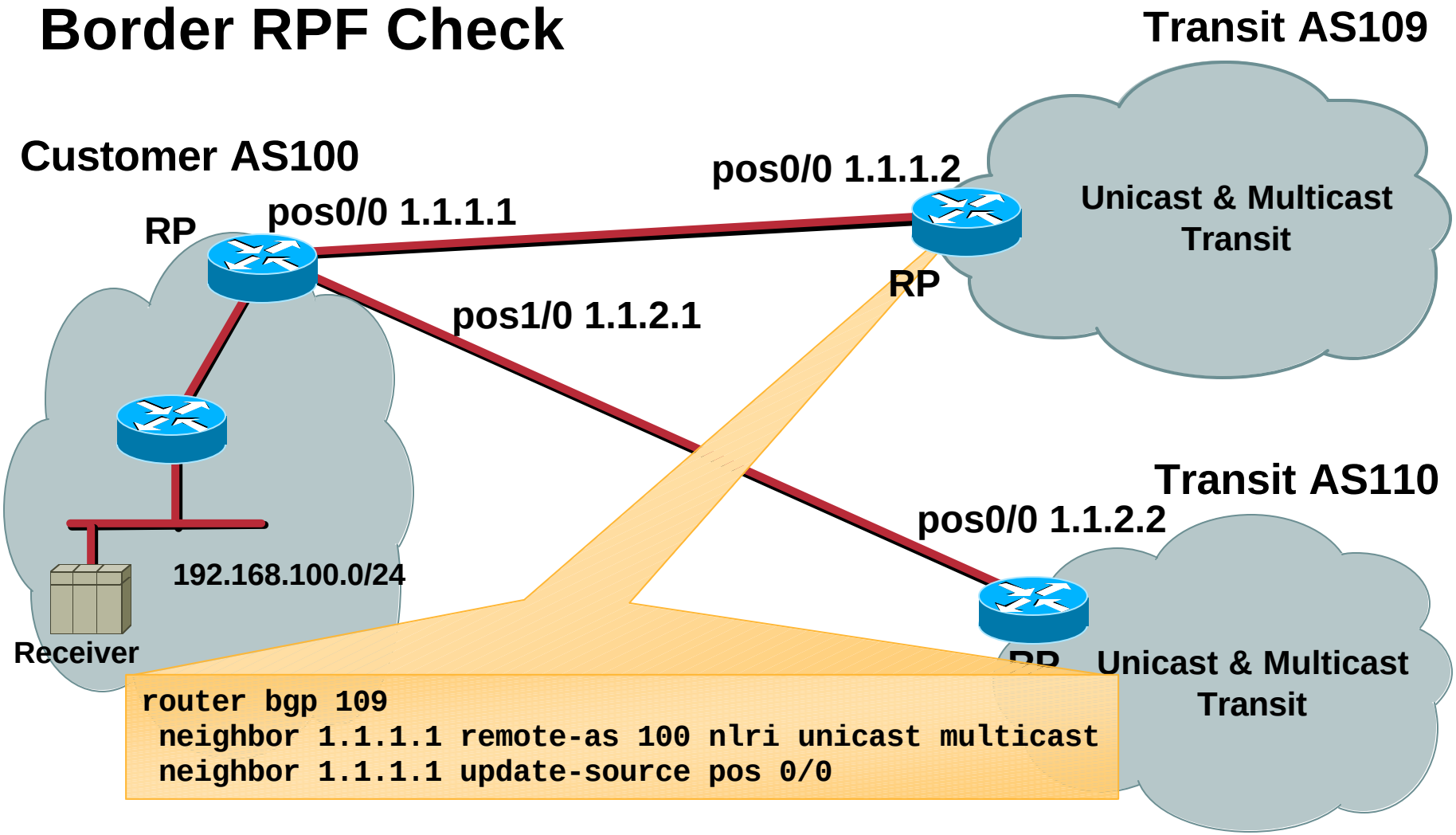
Border RPF Check




```
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
```

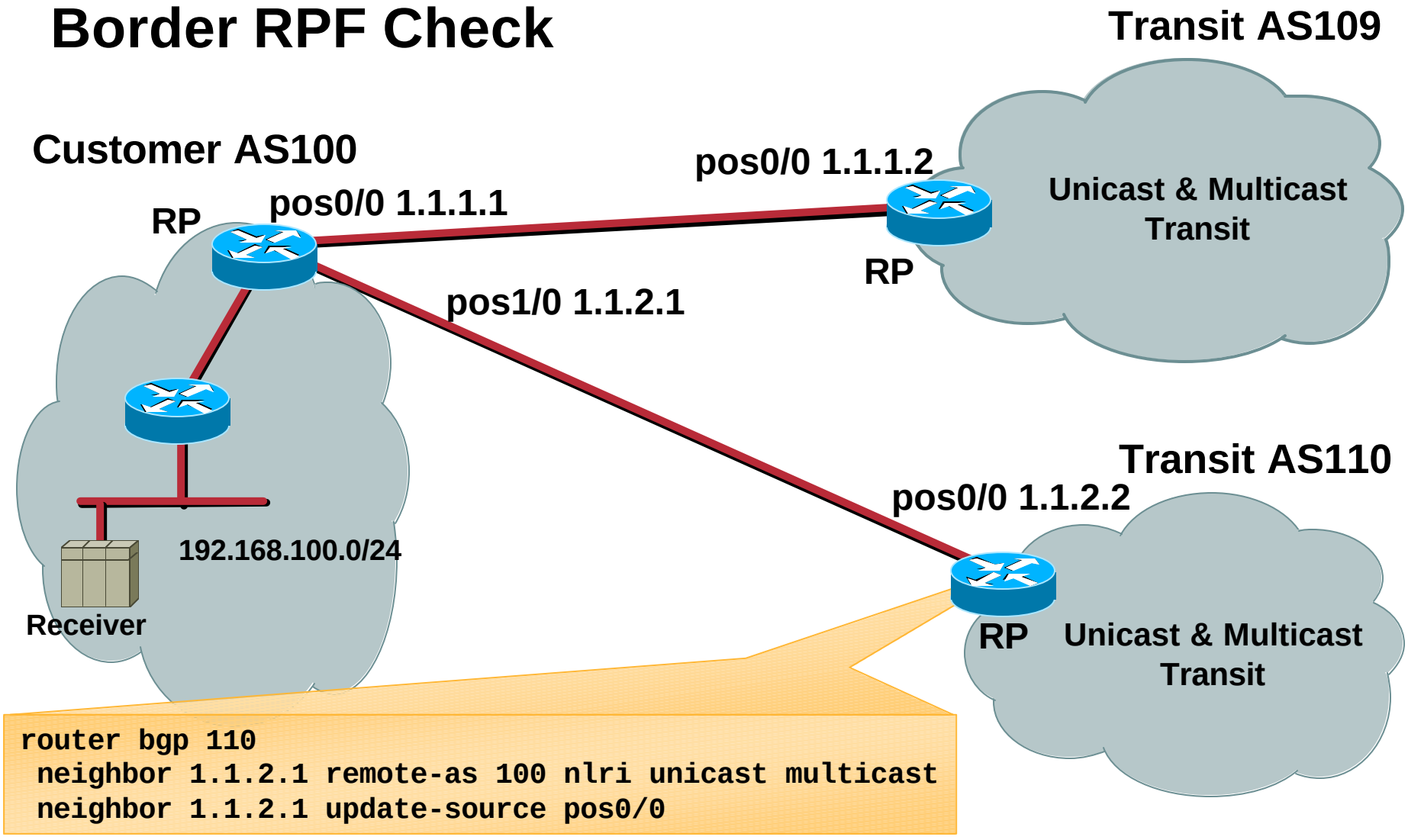
Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

Border RPF Check



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

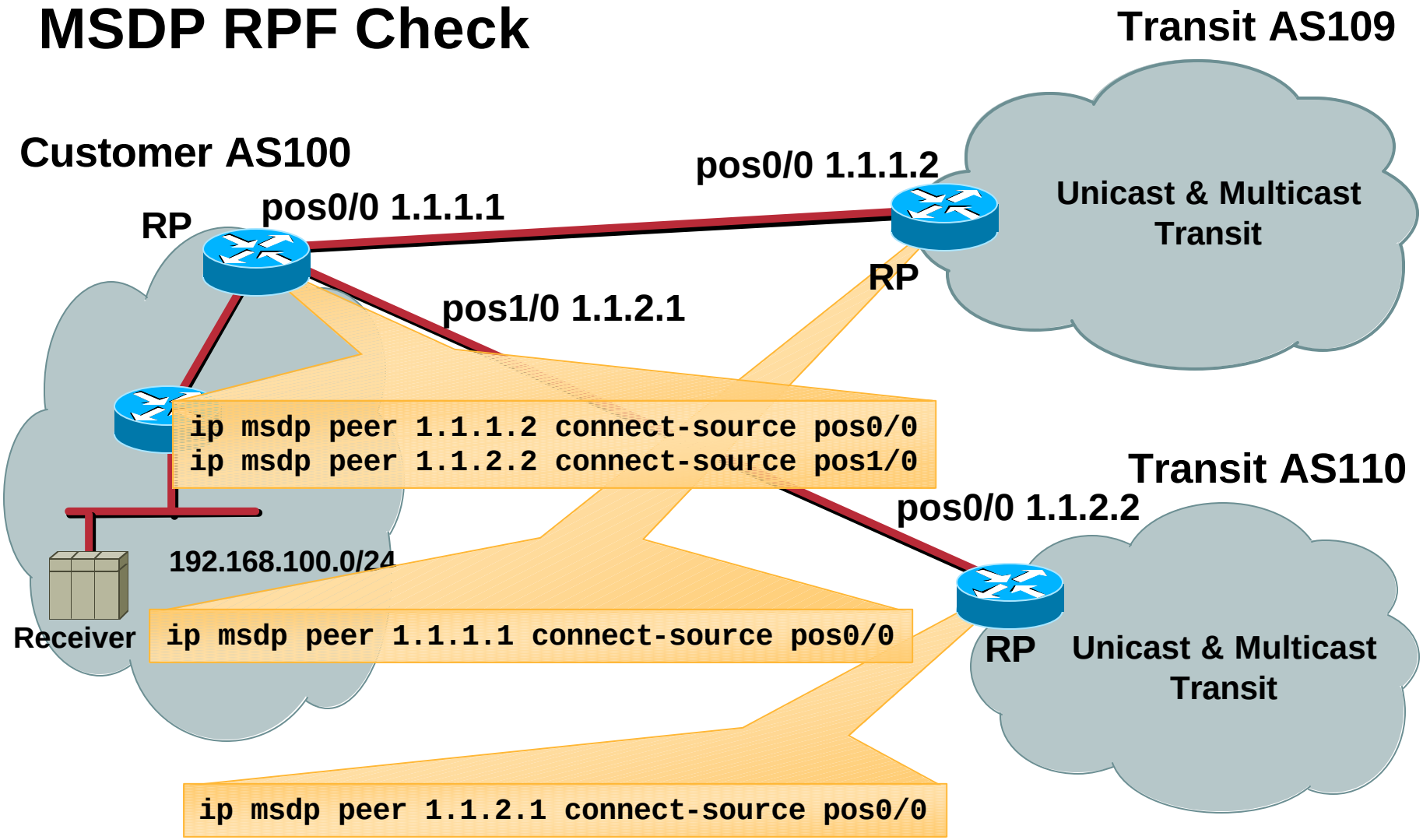
Border RPF Check



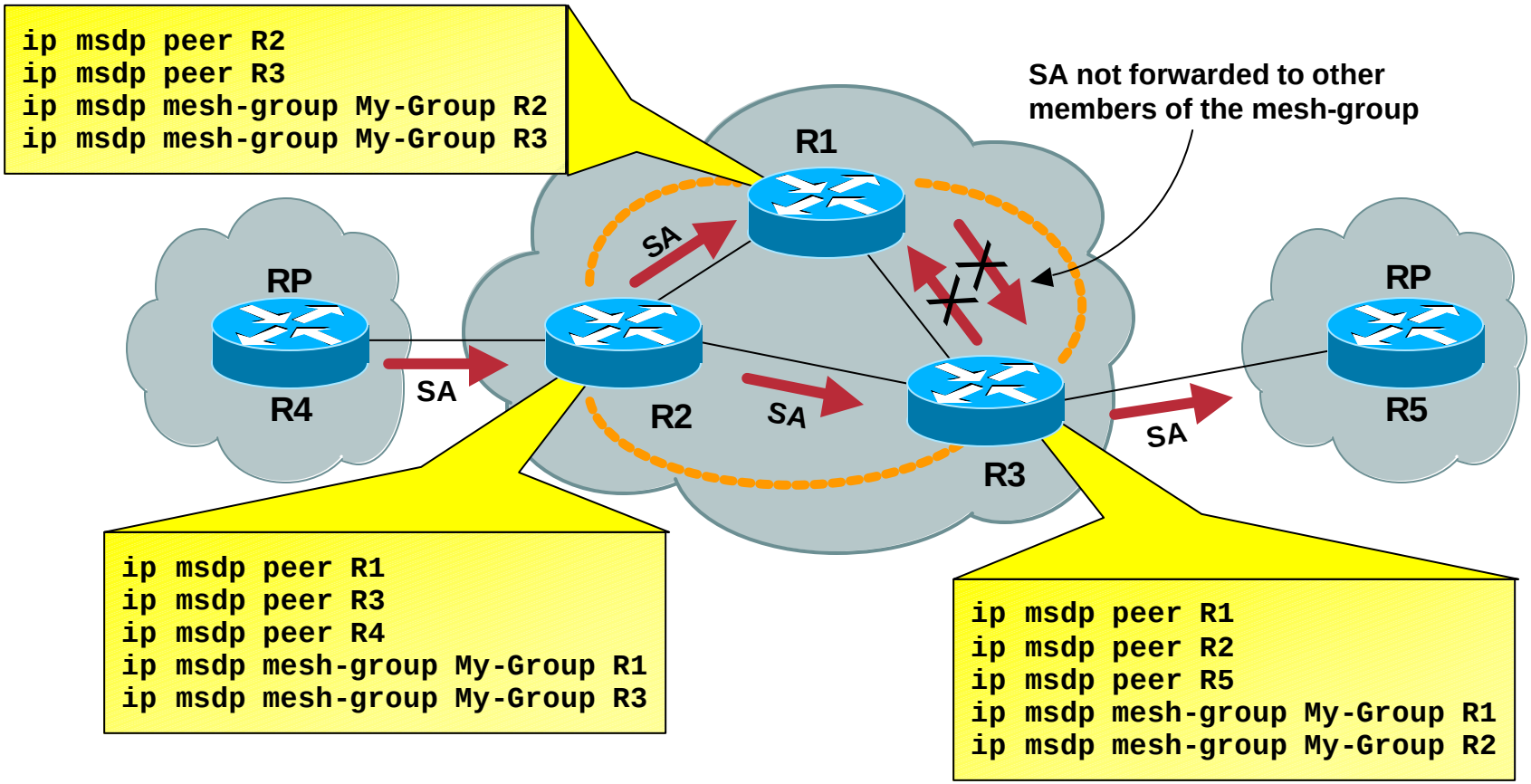

```
router bgp 110
 neighbor 1.1.2.1 remote-as 100 nlri unicast multicast
 neighbor 1.1.2.1 update-source pos0/0
```

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

MSDP RPF Check

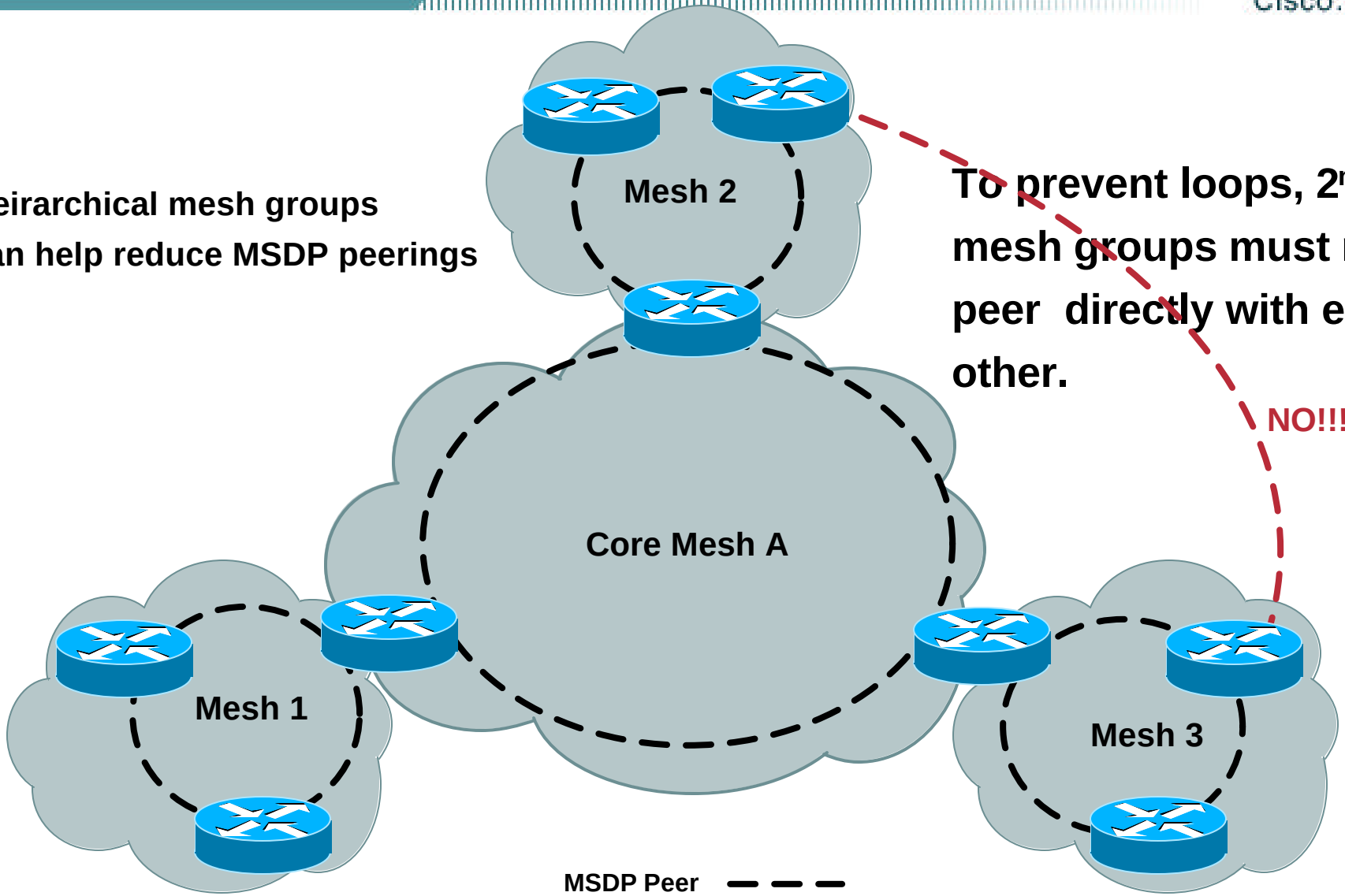


MSDP Mesh-Group Example



Intradomain MSDP peering scenarios

Heirarchical mesh groups can help reduce MSDP peerings



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

NO!!!

GLOP—Static Allocation of 233/8

- Temporary allocation of 233/8

  rfc2770

- 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

- CompanyA 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 0 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

- Another way to calculate the GLOP address is take the hexadecimal value of 5662, which is 161E. 16 hex equals 22 decimal and 1E hex equals 30 decimal. Again, 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

- PIM-SM review (forwarding)
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)

Source Specific Multicast (SSM)

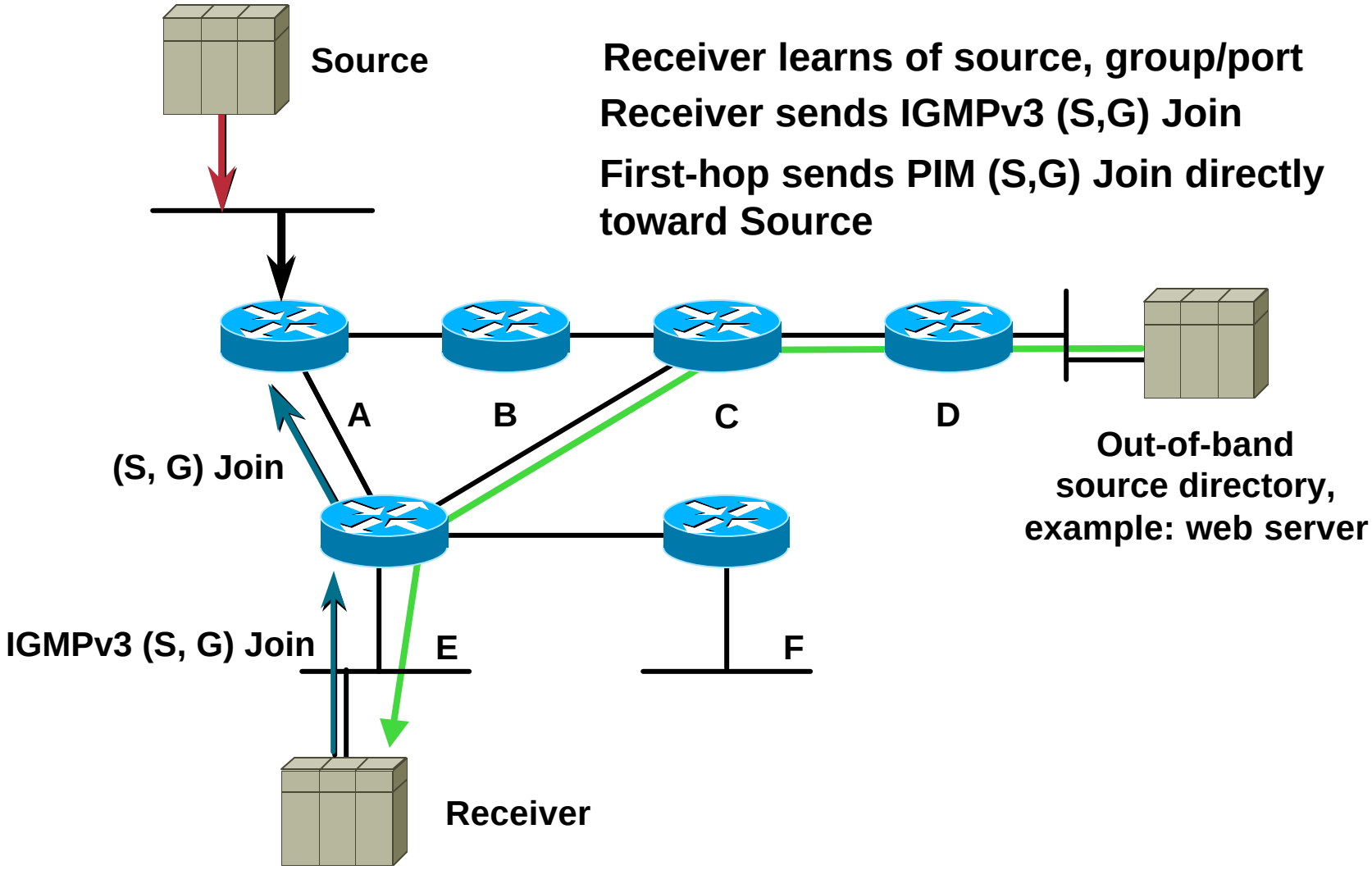
Cisco.com

- 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.

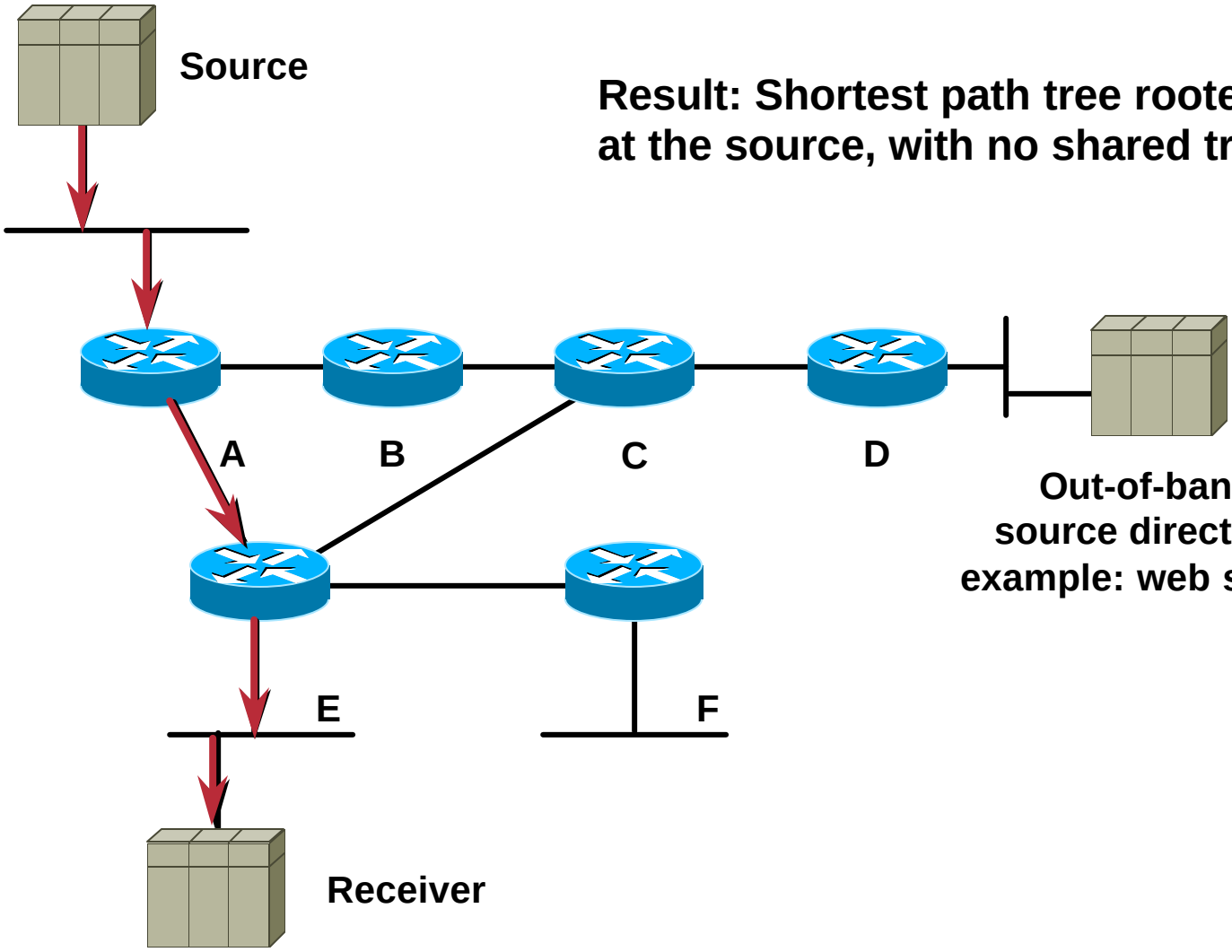
SSM Overview

- **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.**

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 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.**

Where Is SSM?

- **Framework**
  - **draft-holbrook-idmr-igmpv3-ssm-02.txt**
  - **draft-ietf-ssm-arch-00.txt**
- **BCP proposal**
  - **draft-ietf-mboned-ssm232-02.txt**
  - **draft-ietf-ssm-overview-02.txt**
- **Supported in:**
  - **IOS 12.X**
  - **Windows XP, FreeBSD, Linux**
  - **ftp://ftpeng.cisco.com/ipmulticast/ssm/index.html#Stacks**

Agenda

- PIM-SM review (forwarding)
- MBGP (routing)
- MSDP (source discovery)
- MBGP/MSDP Examples
- SSM (Source Specific Multicast)
- MVPN (Multicast VPN)

Agenda

Cisco.com

- **Why Multicast VPNs (MVPN's)**
- **Multicast VPN Terminology**
- **Multicast VPN Overview**
- **Configuring Multicast VPN's**

Why Multicast VPNs

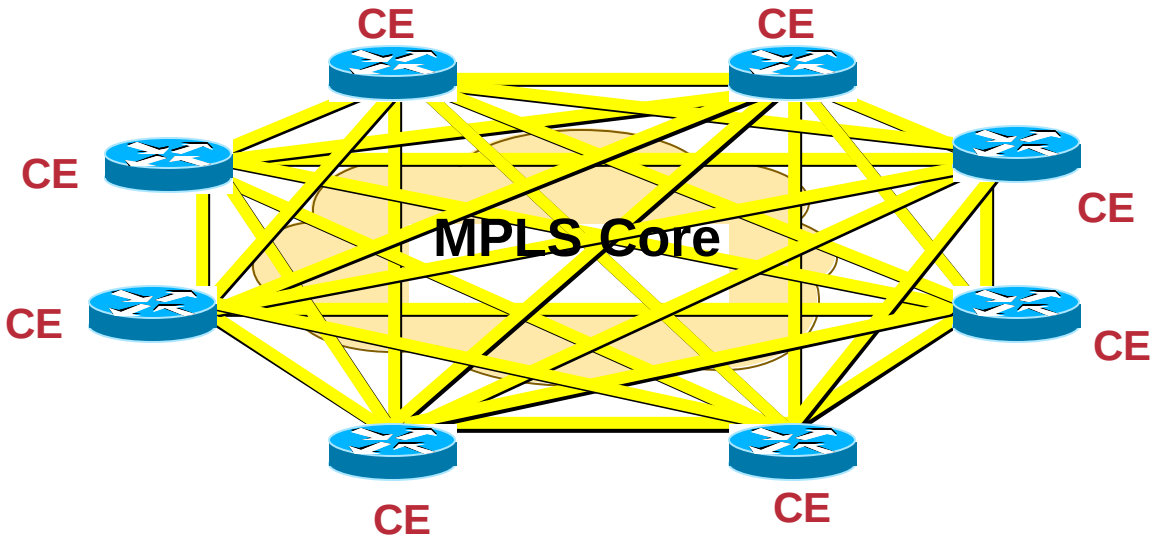
- Until now only unicast has been supported in MPLS/BGP VPN
- VPN customers need multicast connectivity
  - Applications that require multicast
  - Internet multicast connectivity
- Service Providers want to offer additional services
  - e.g. Video streaming to its VPN customers

Why Multicast VPNs

- MPLS VPN customers want to run multicast within their VPNs
- Multicast deployment is expanding
- MPLS VPNs do not support multicast today
- Multicast options in MPLS VPNs today

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.**
  - **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.**
- **Scheduled for 12.2(12)T on 7200 and 7500.**

Multicast VPN – Terminology

- **VPN: Virtual Private Network**

  Although different VPN models exist, the discussion here is for MPLS based VPNs

- **VRF: VPN Routing and Forwarding**

  per-site forwarding tables

- **MVPN: Multicast VPN**

  A VPN that supports multicast natively

- **MVRF: Multicast VRF**

  A VRF that supports unicast and multicast tables

Multicast VPN – Terminology

- **Multicast Domain (MD):**

  Set of MVRFs that can send multicast to each other.

- **MDT: Multicast Distribution Tree**

  A multicast tree, built in the core network that tunnels customer multicast traffic between sites.

- **Default-MDT:**

  Tunnels PIM control traffic and multicast data between sites.

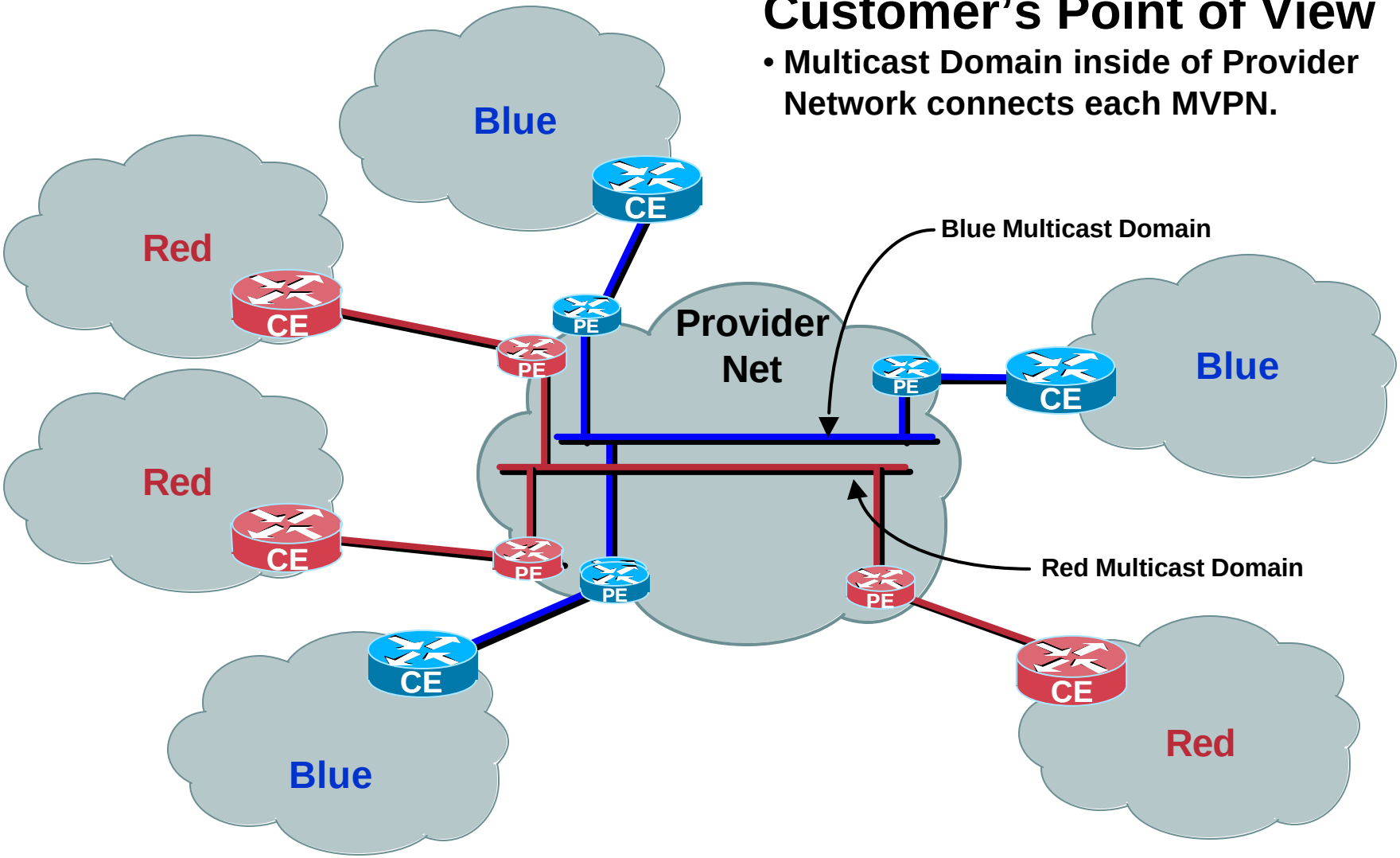
- **Data-MDT:**

  MDT group created on demand for MVPN (S,G) pairs, usually high bandwidth traffic.

Multicast VPN – Overview

Customer's Point of View

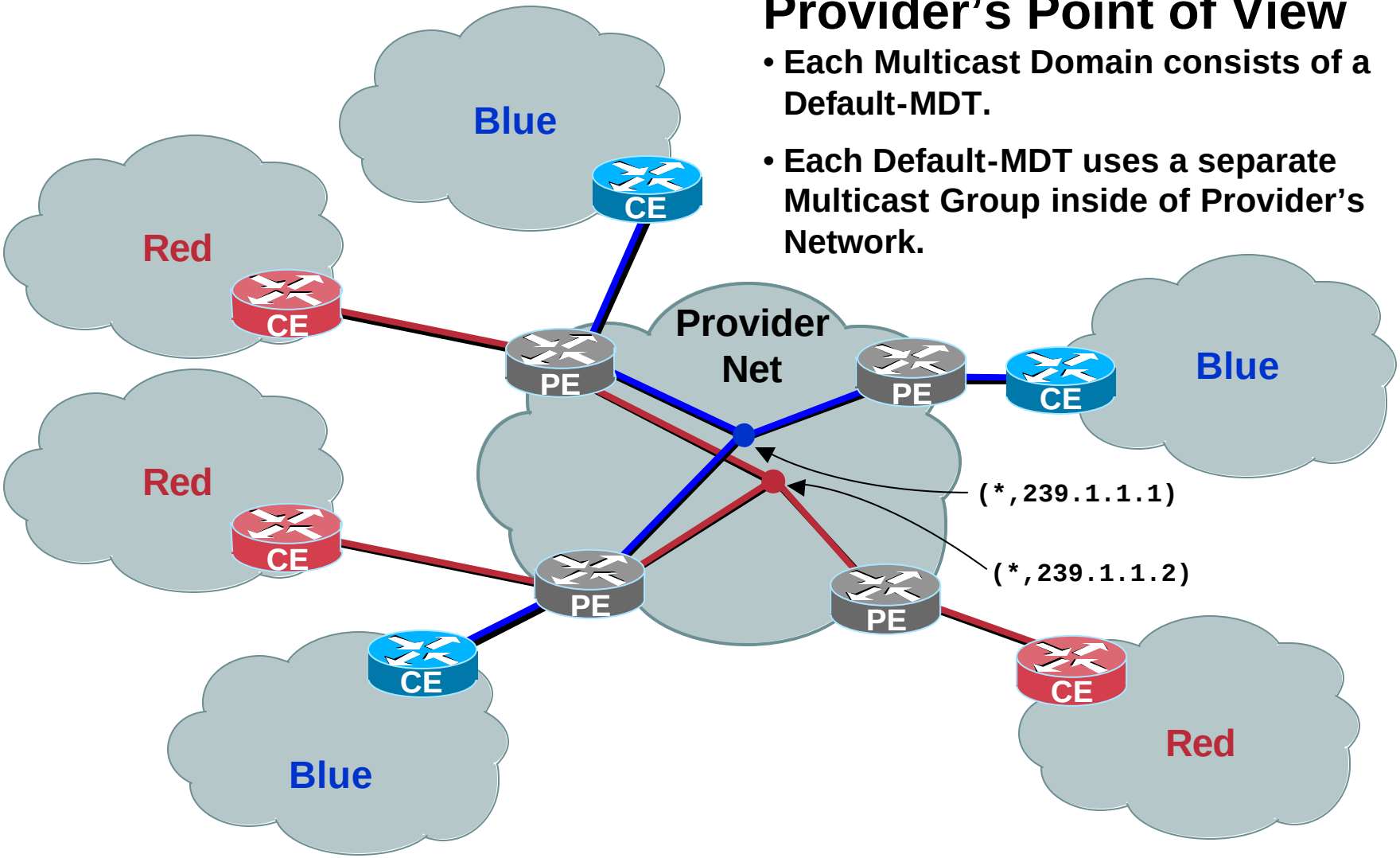
- Multicast Domain inside of Provider Network connects each MVPN.



Multicast VPN – Overview

Provider's Point of View

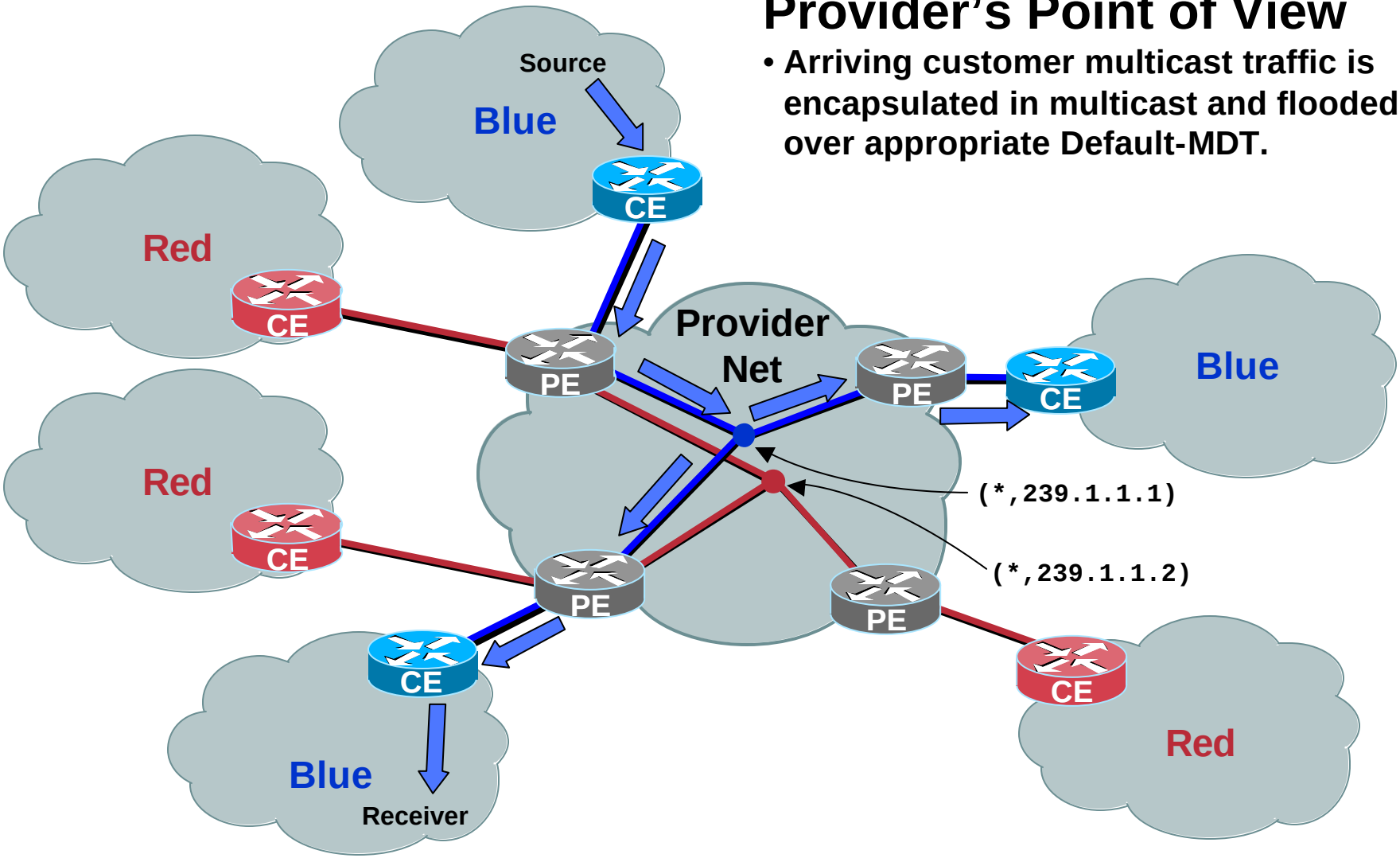
- Each Multicast Domain consists of a Default-MDT.
- Each Default-MDT uses a separate Multicast Group inside of Provider's Network.



Multicast VPN – Overview

Provider's Point of View

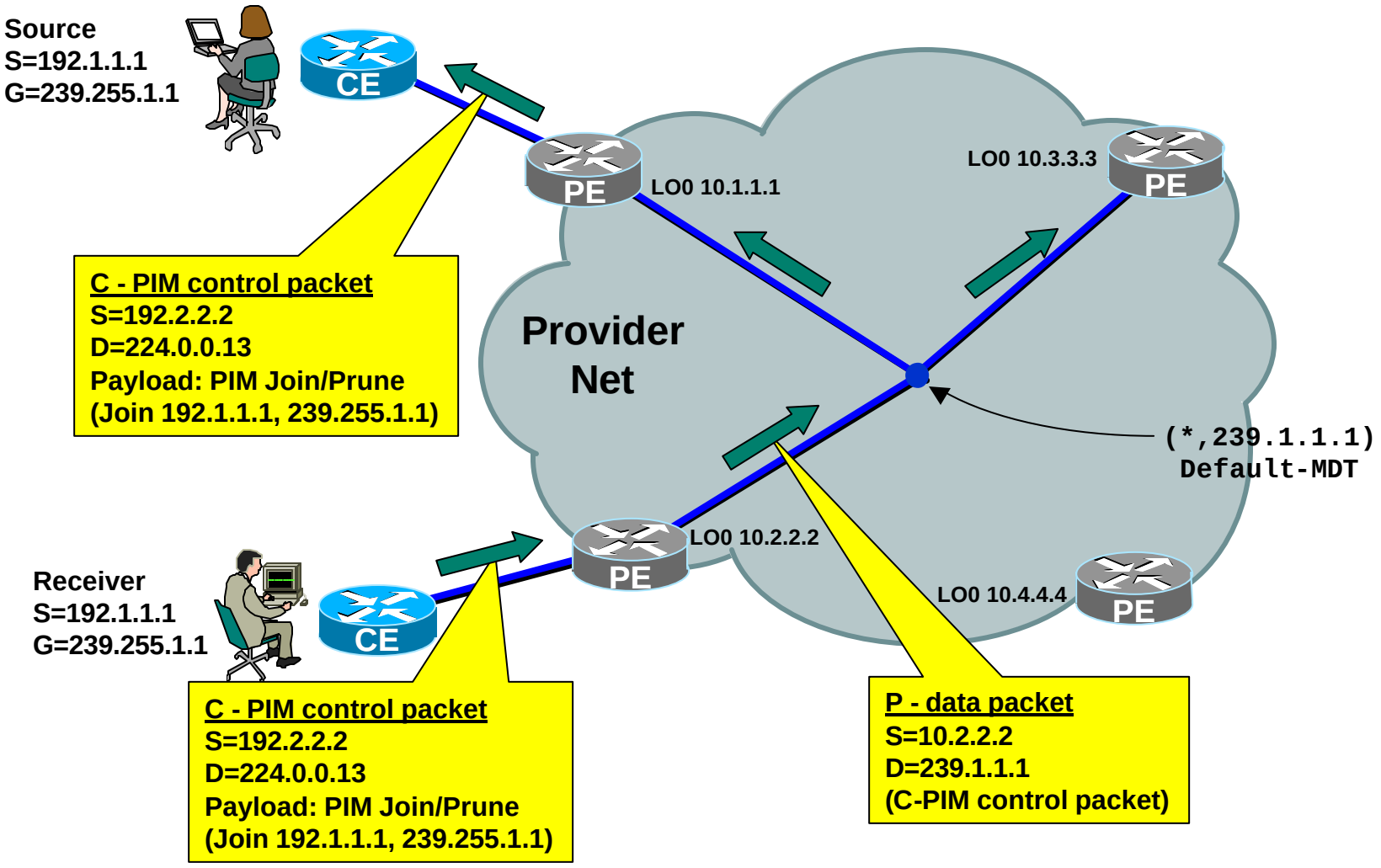
- Arriving customer multicast traffic is encapsulated in multicast and flooded over appropriate Default-MDT.



Default MDT – A Closer Look

Cisco.com

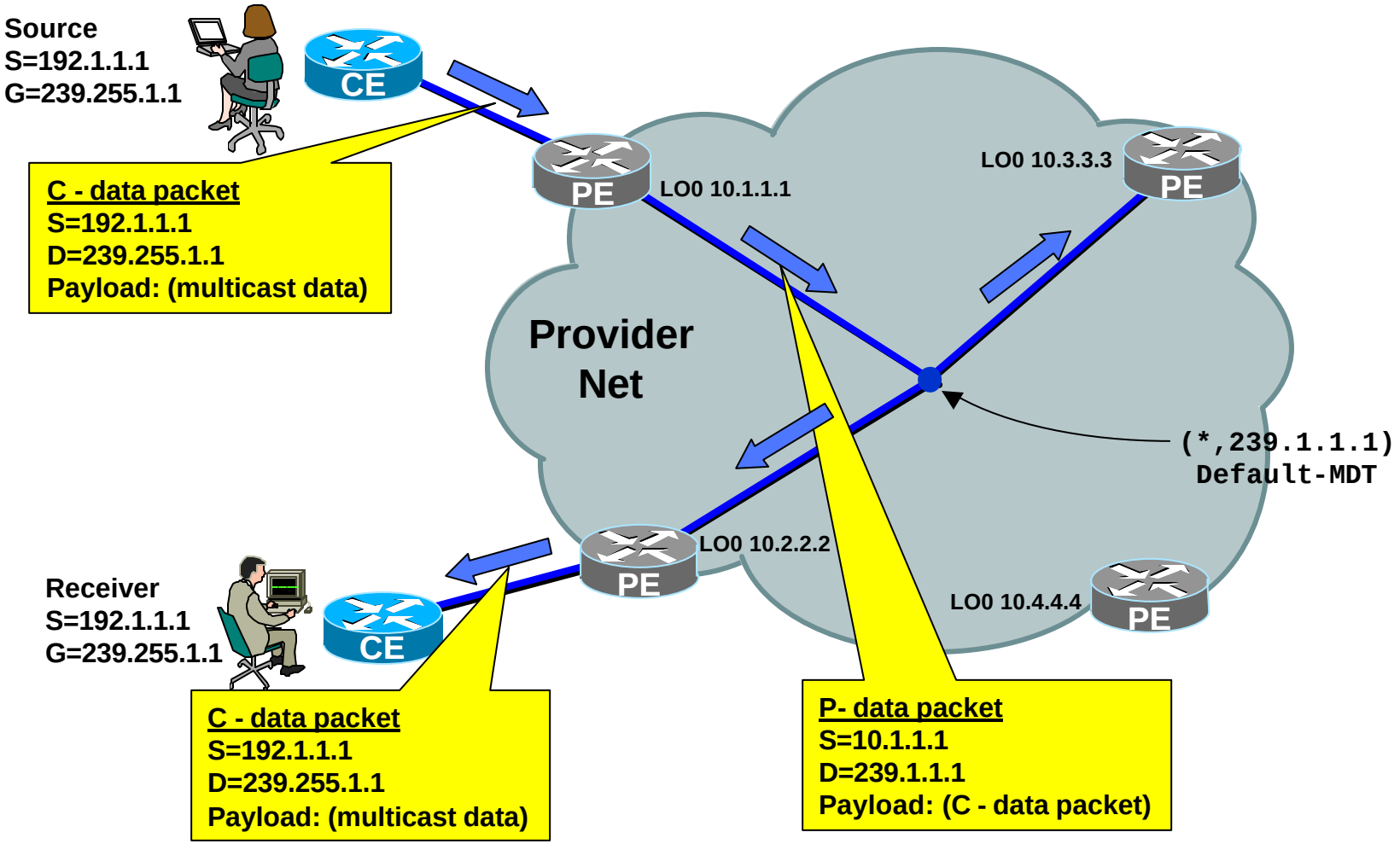
PIM Control Traffic Flow



Default MDT – A Closer Look

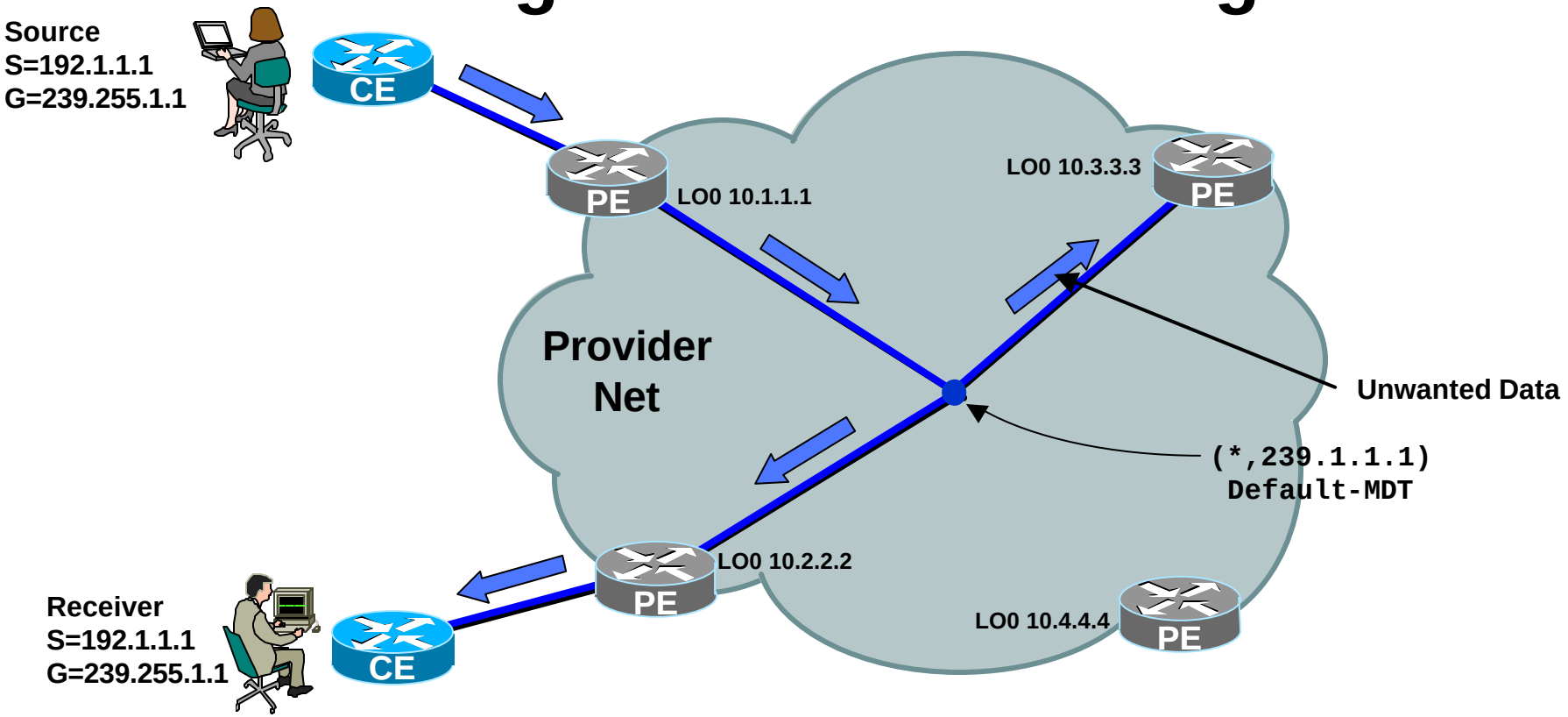
Cisco.com

Multicast Data Traffic Flow



Default MDT – A Closer Look

Advantages and Disadvantages

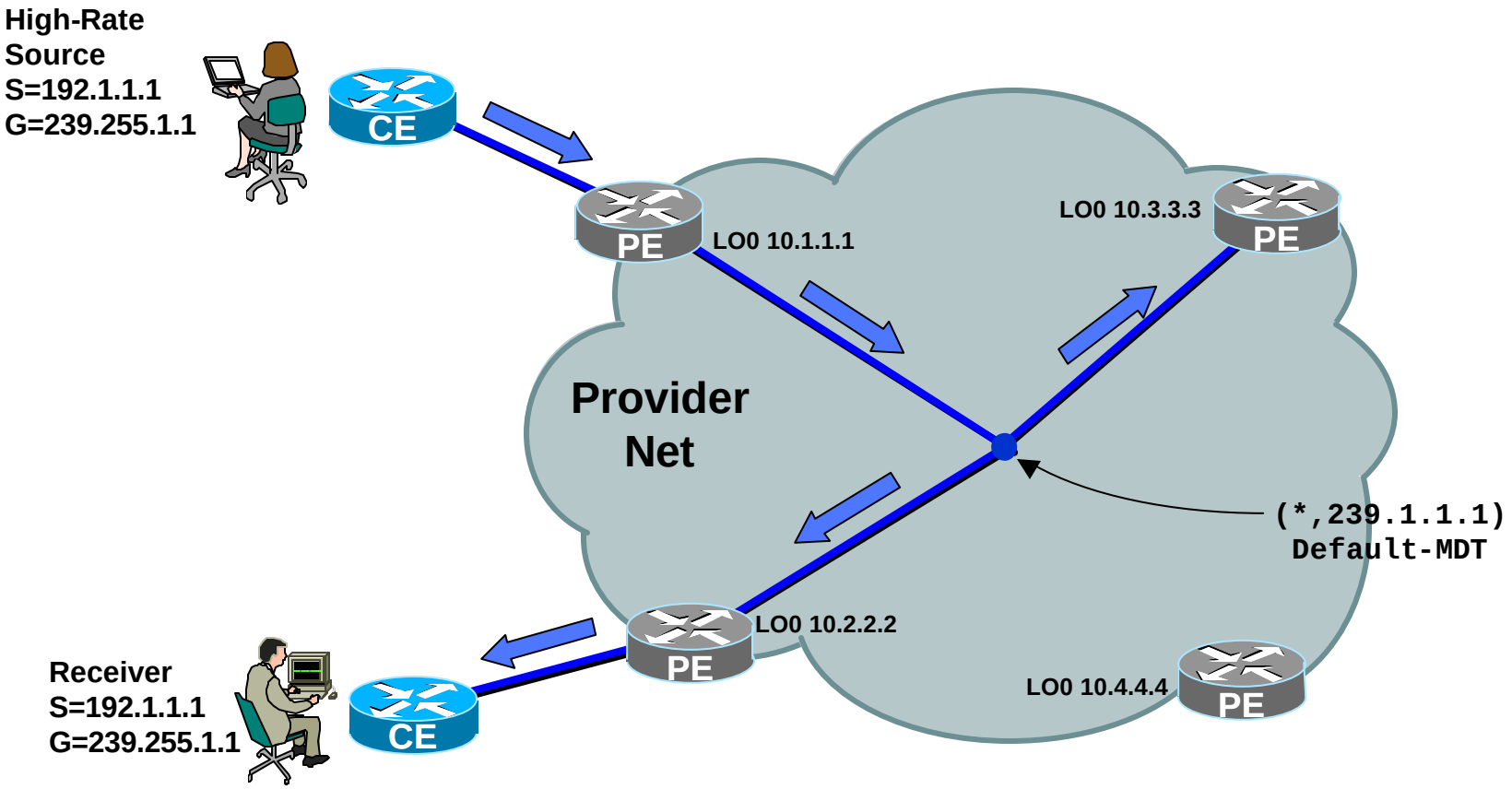


Advantage : Reduces multicast state in the P routers in the core.

Disadvantage : Can result in wasted bandwidth.

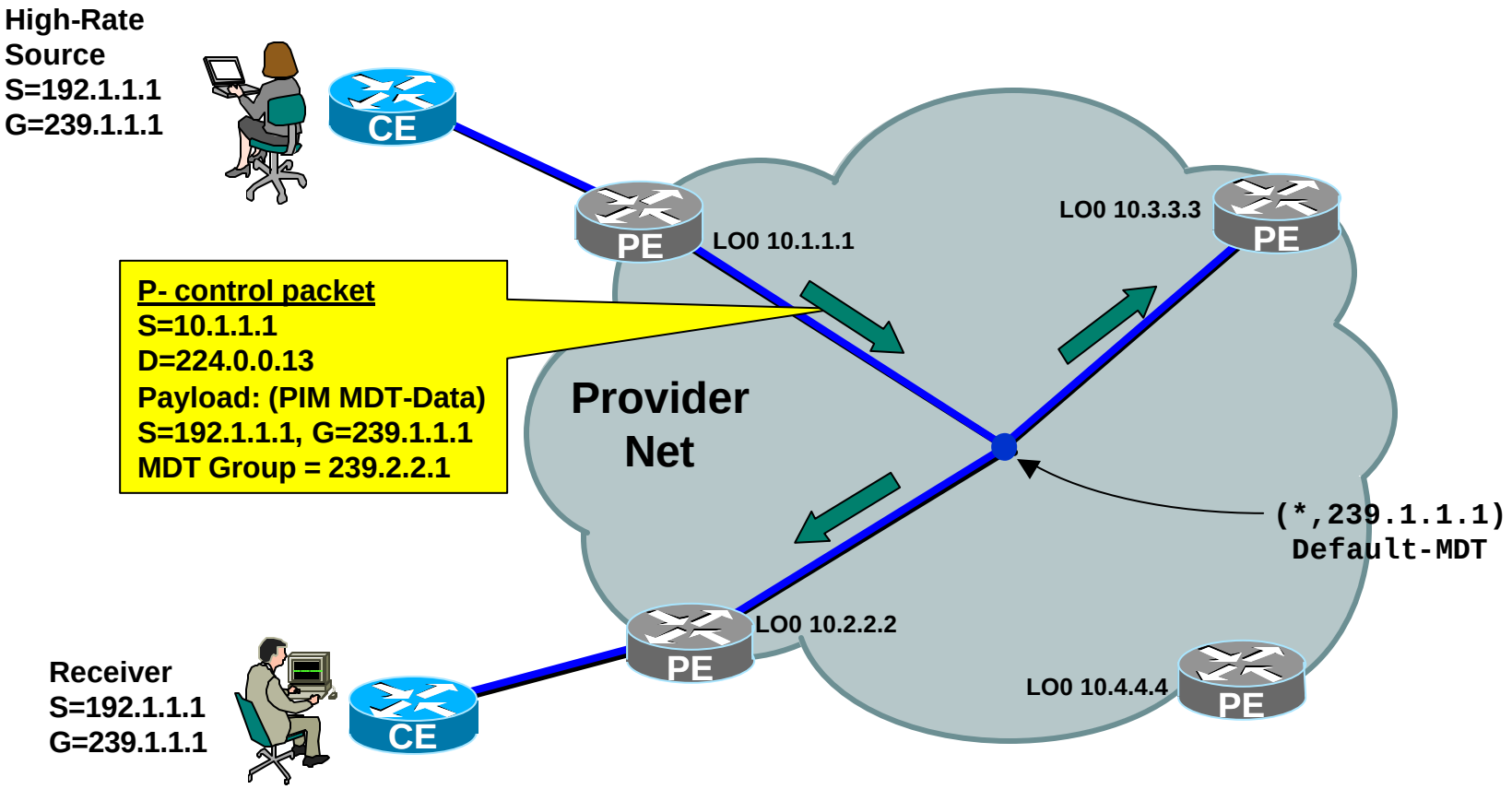
Solution : Use separate Data-MDTs for high rate sources.

Data MDTs – Concepts



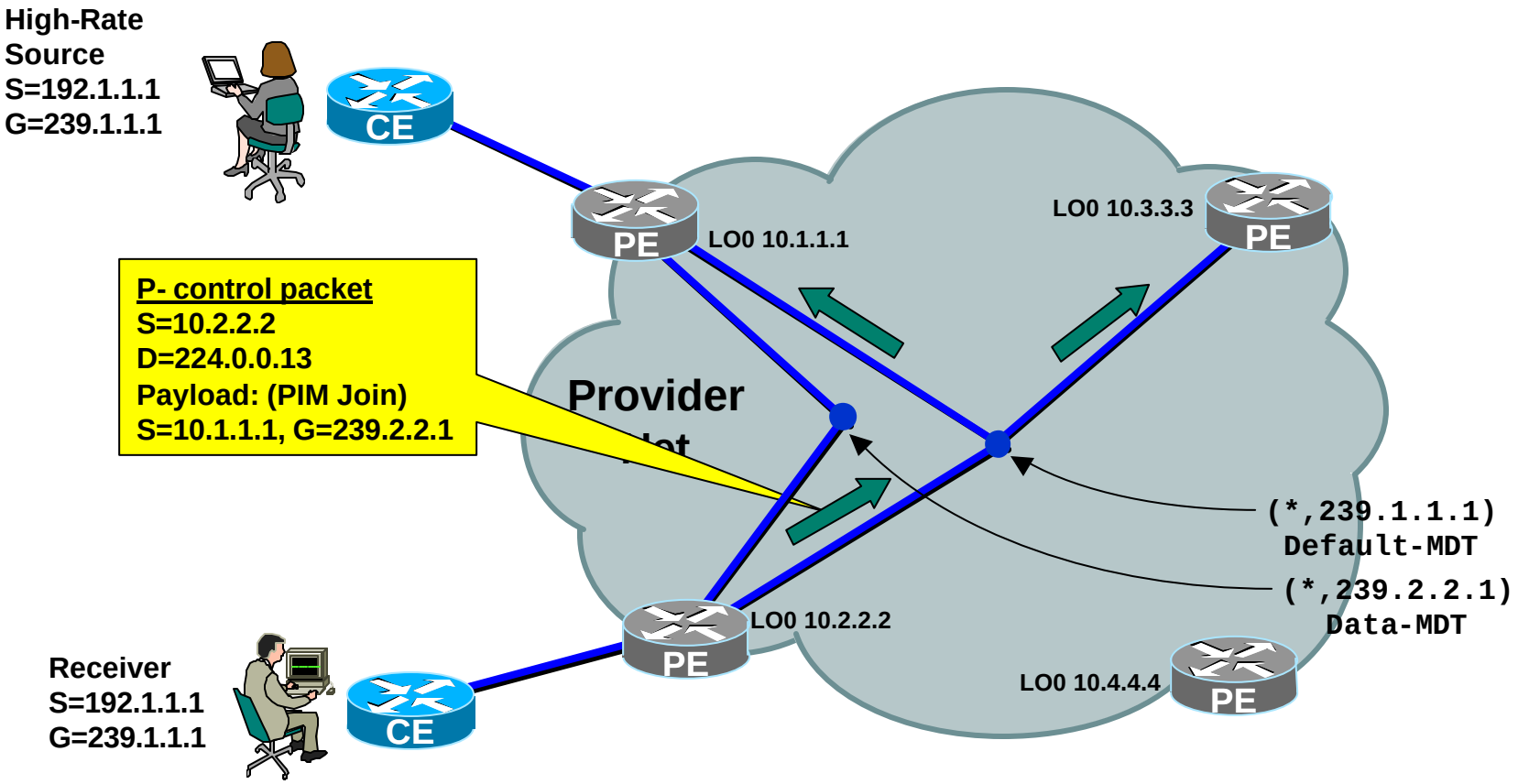
- Traffic exceeds Data-MDT threshold configured on PE router.

Data MDTs – Concepts



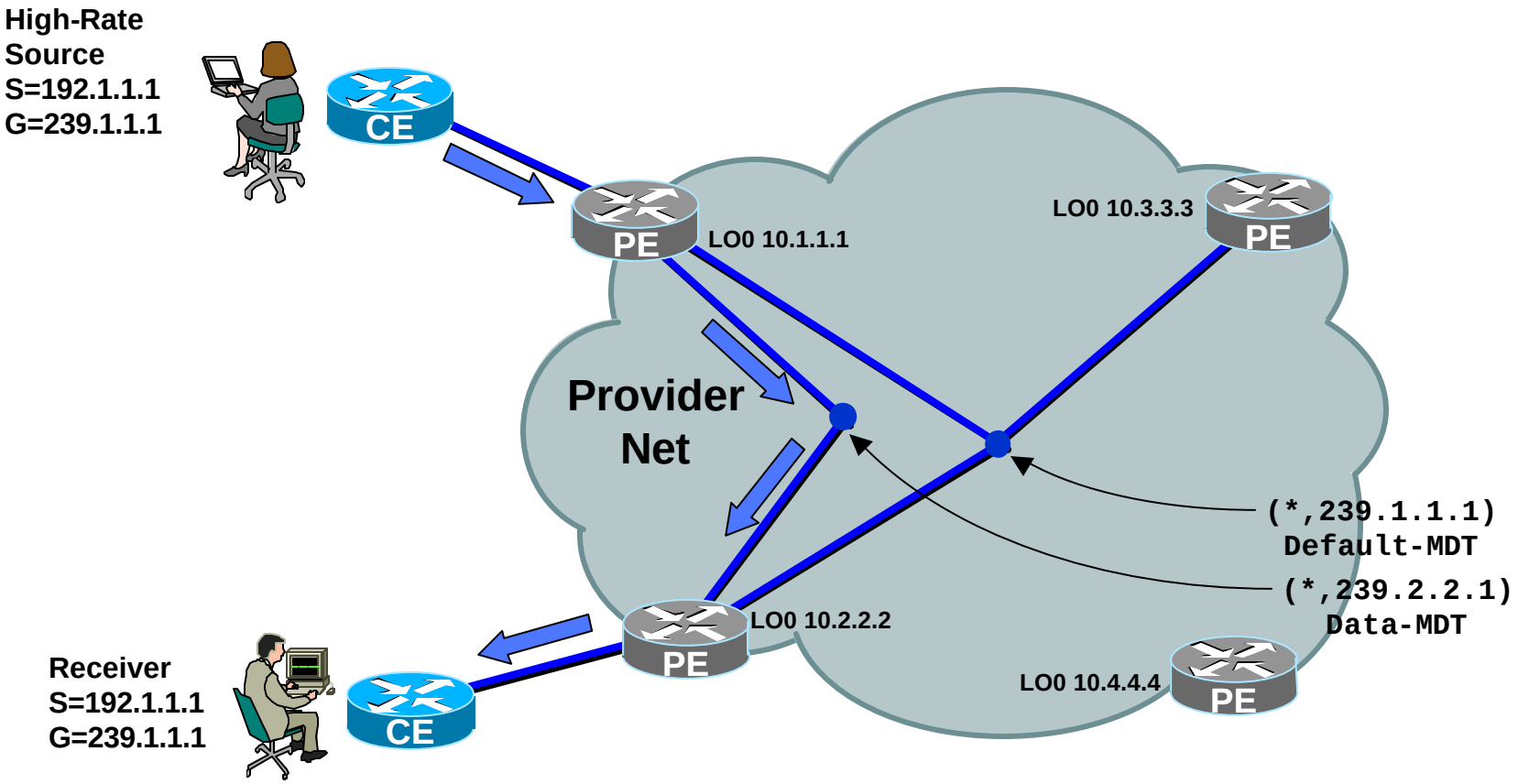
- PE router signals switch to Data-MDT using new group, 239.2.2.1

Data MDTs – Concepts



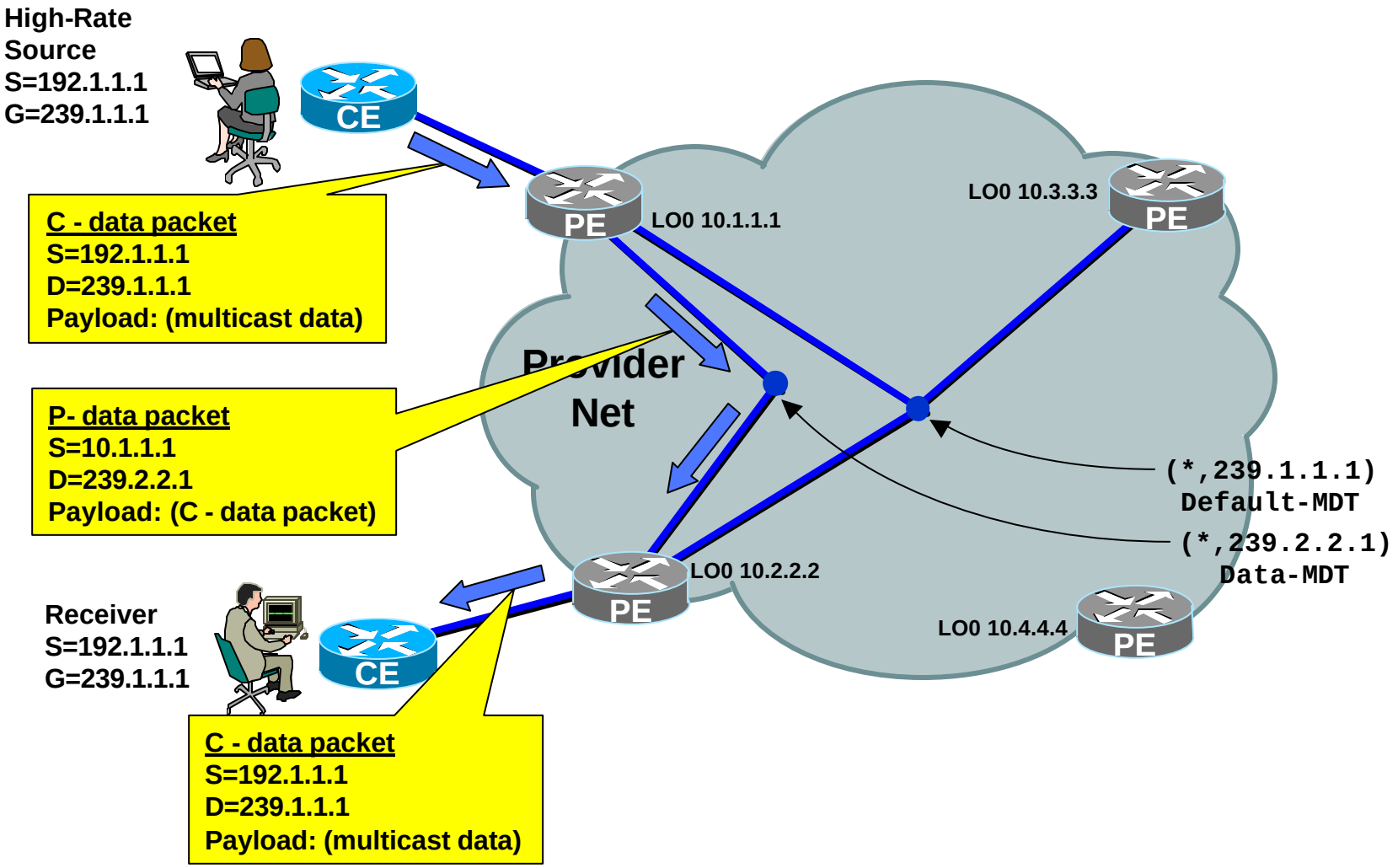
- PE routers with receivers sends Join to group 239.2.2.1.
- Data-MDT is built using group 239.2.2.1.

Data MDTs – Concepts



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

Data MDTs – Concepts



Configuration of Default-MDT

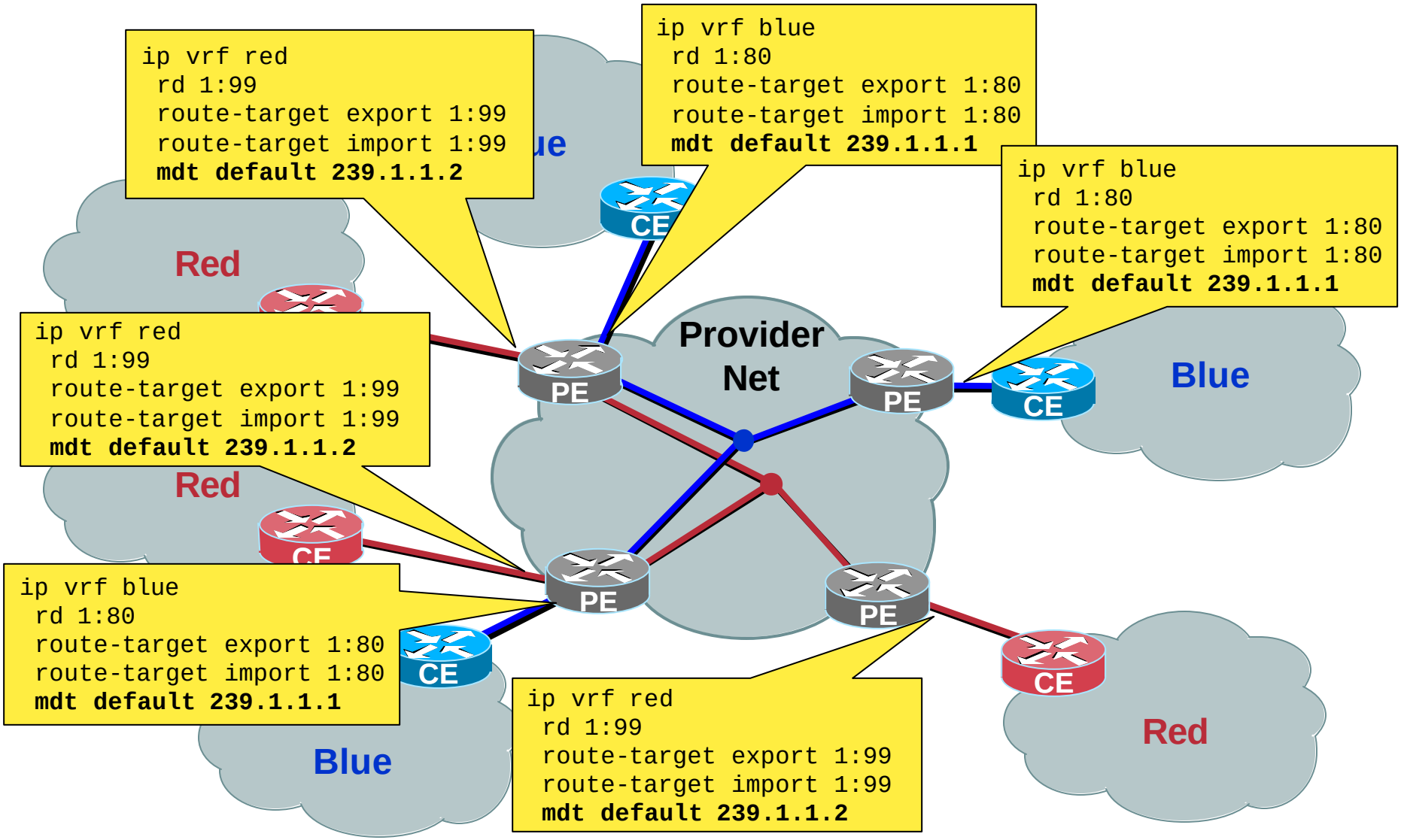
- **Default MDT Group address**

  Must be configured.

  Group address MUST be the same for all MVRFs belonging to the same MVPN

  Group address MUST be different for all MVRFs belonging to different VPNs that are configured on the same PE router

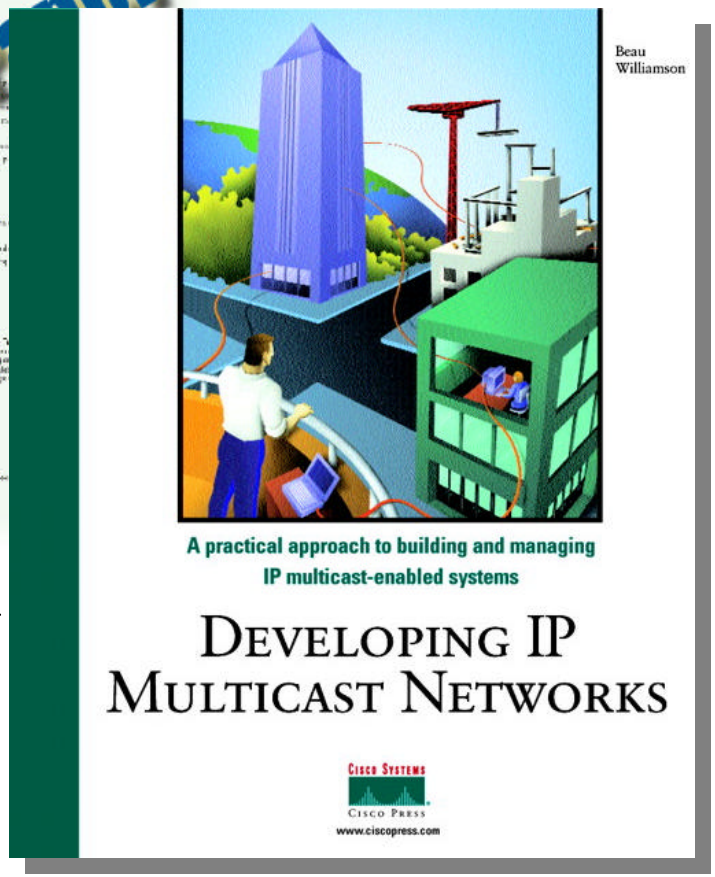
Default-MDT Group Address Example



More Information

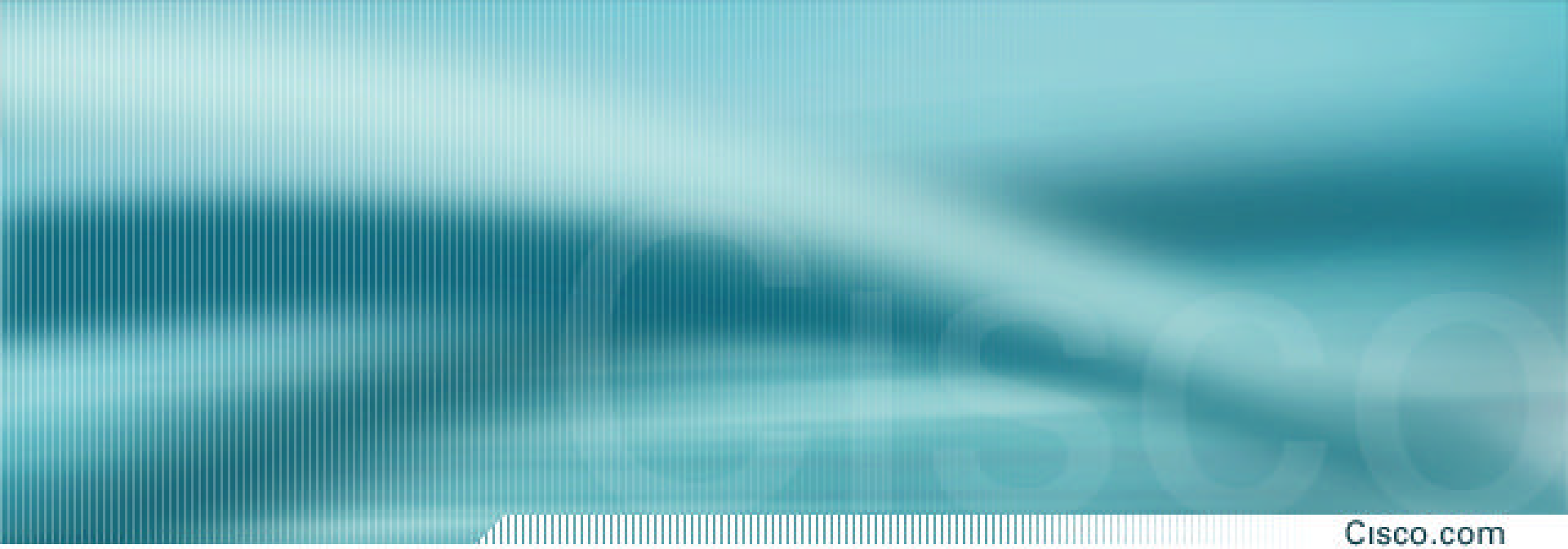
- White Papers
- Web and Mailers
- Cisco Press

RTFB



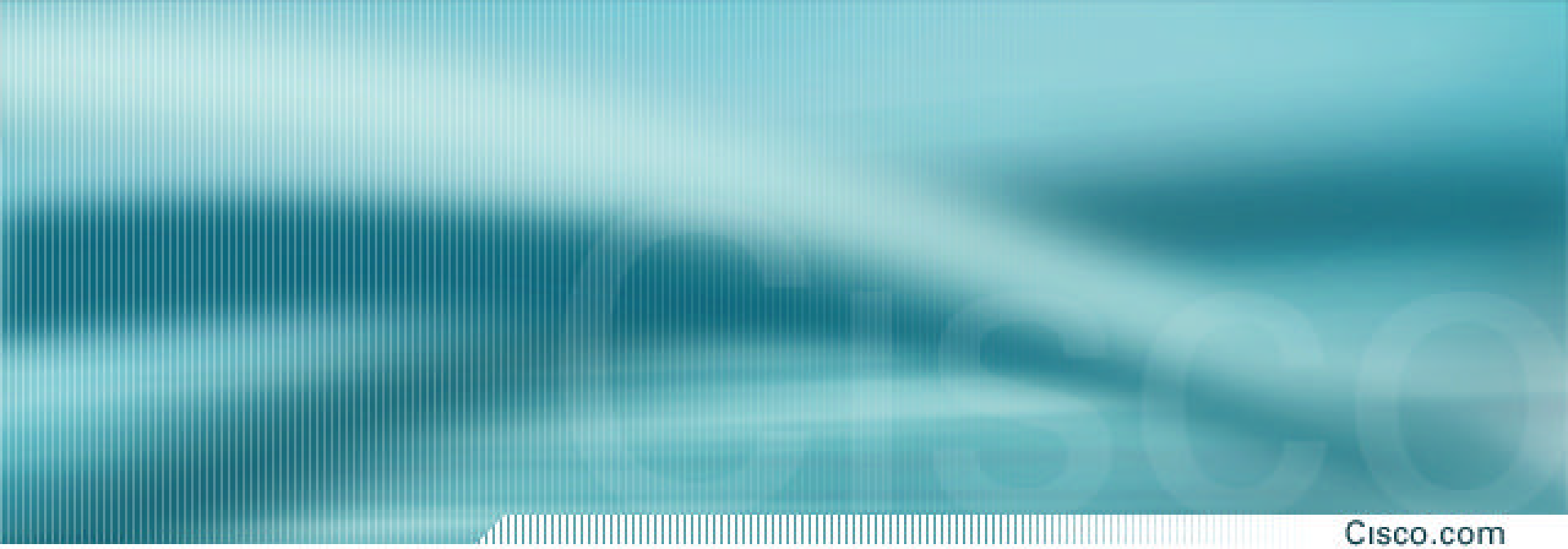
CCO Multicast page:
http://www.cisco.com/go/ipmulticast
Questions:
cs-ipmulticast@cisco.com
Customer Support Mailing List:
tac@cisco.com

RTFB = "Read the Fabulous Book"



Deploying Inter-Domain IP Multicast

Session RST-261



Please Complete Your Evaluation Form

Session RST-261