Advanced IP Multicast

RST-3261

Agenda

- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)
- Multicast IPv6
- Multi Topology Routing (MTR)
- Triple Play and Multicast

Multiprotocol BGP (MBGP)

MBGP Overview

- MBGP: Multiprotocol BGP
  - Defined in RFC 2283 (extensions to BGP)
  - Can carry different types of routes
    - IPv4/v6 Unicast/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, …

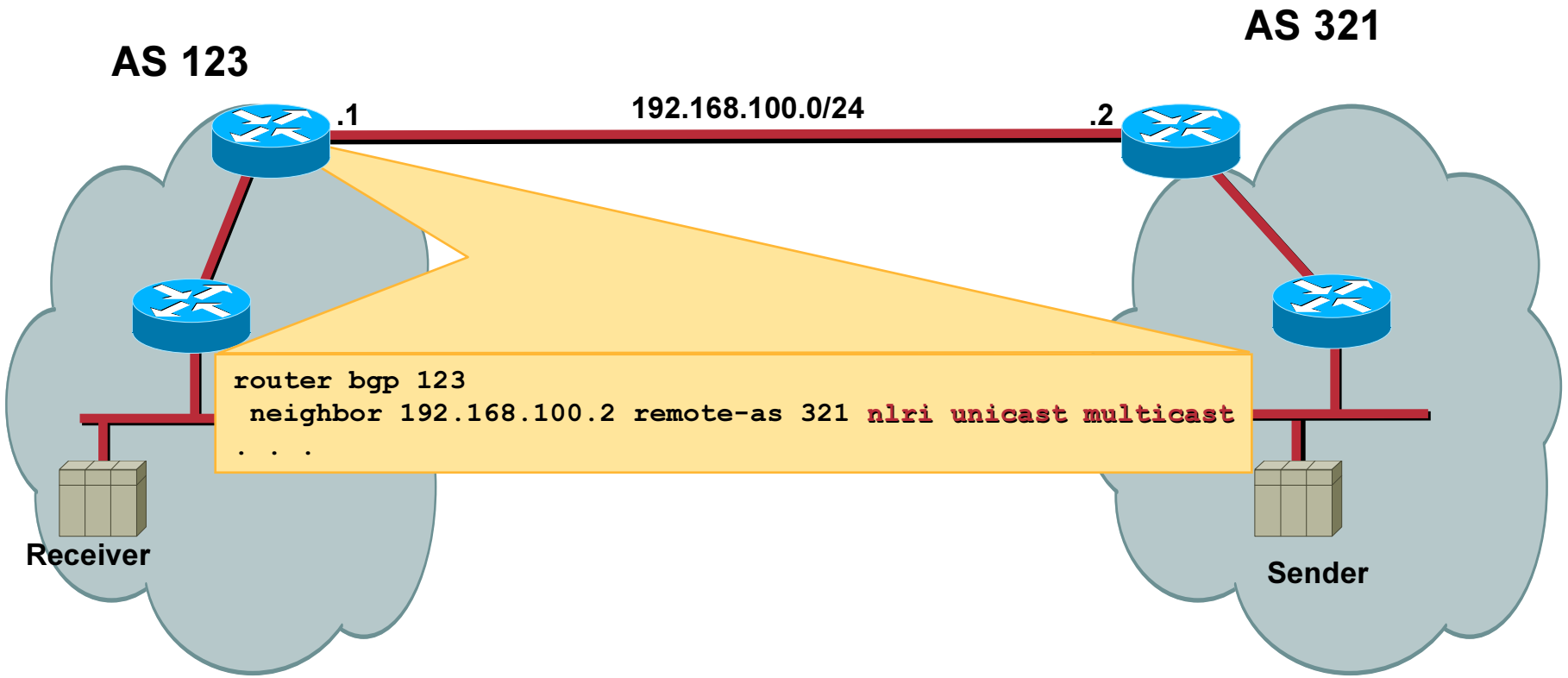
MBGP Overview

- **Separate BGP tables maintained**
  - **Unicast BGP Table (U-Table)**
  - **Multicast BGP Table (M-Table)**
  - **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**

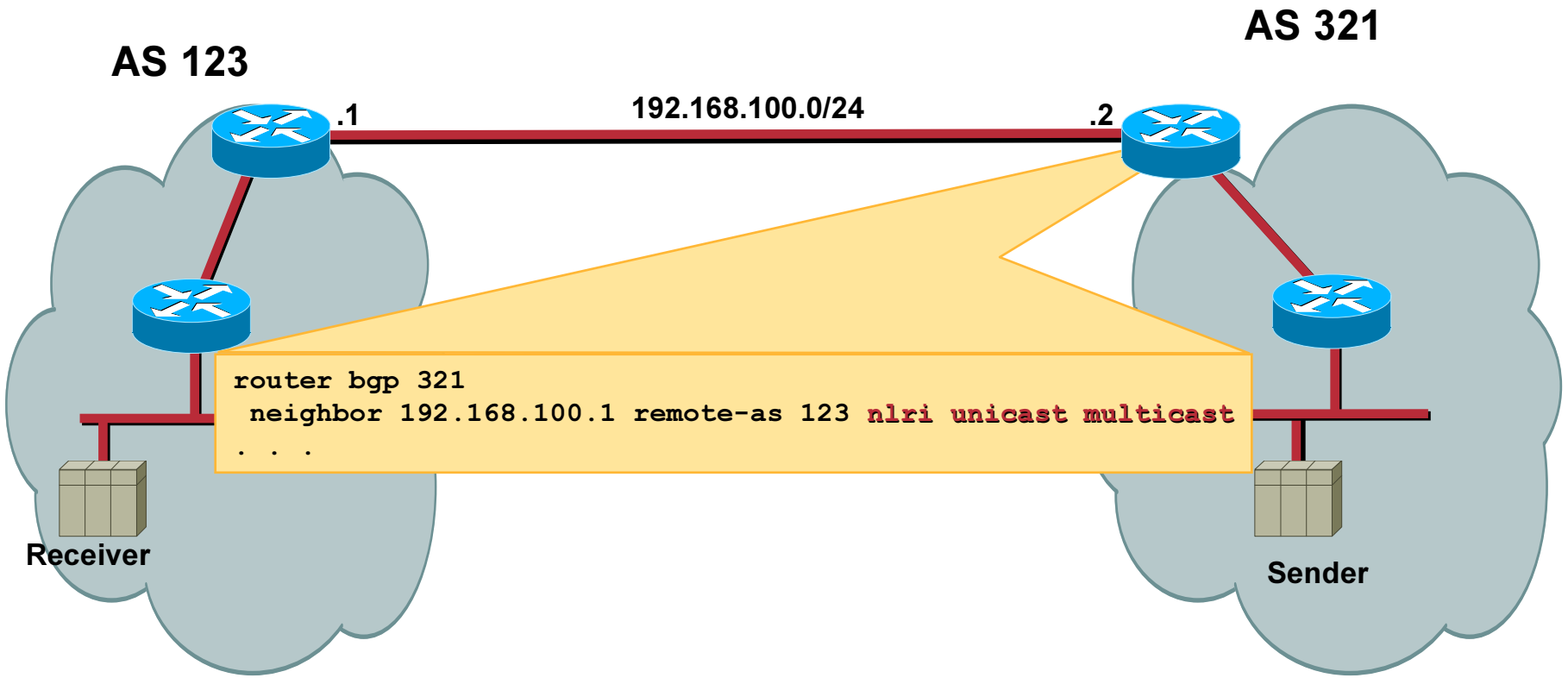
MBGP Update Message

- 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)

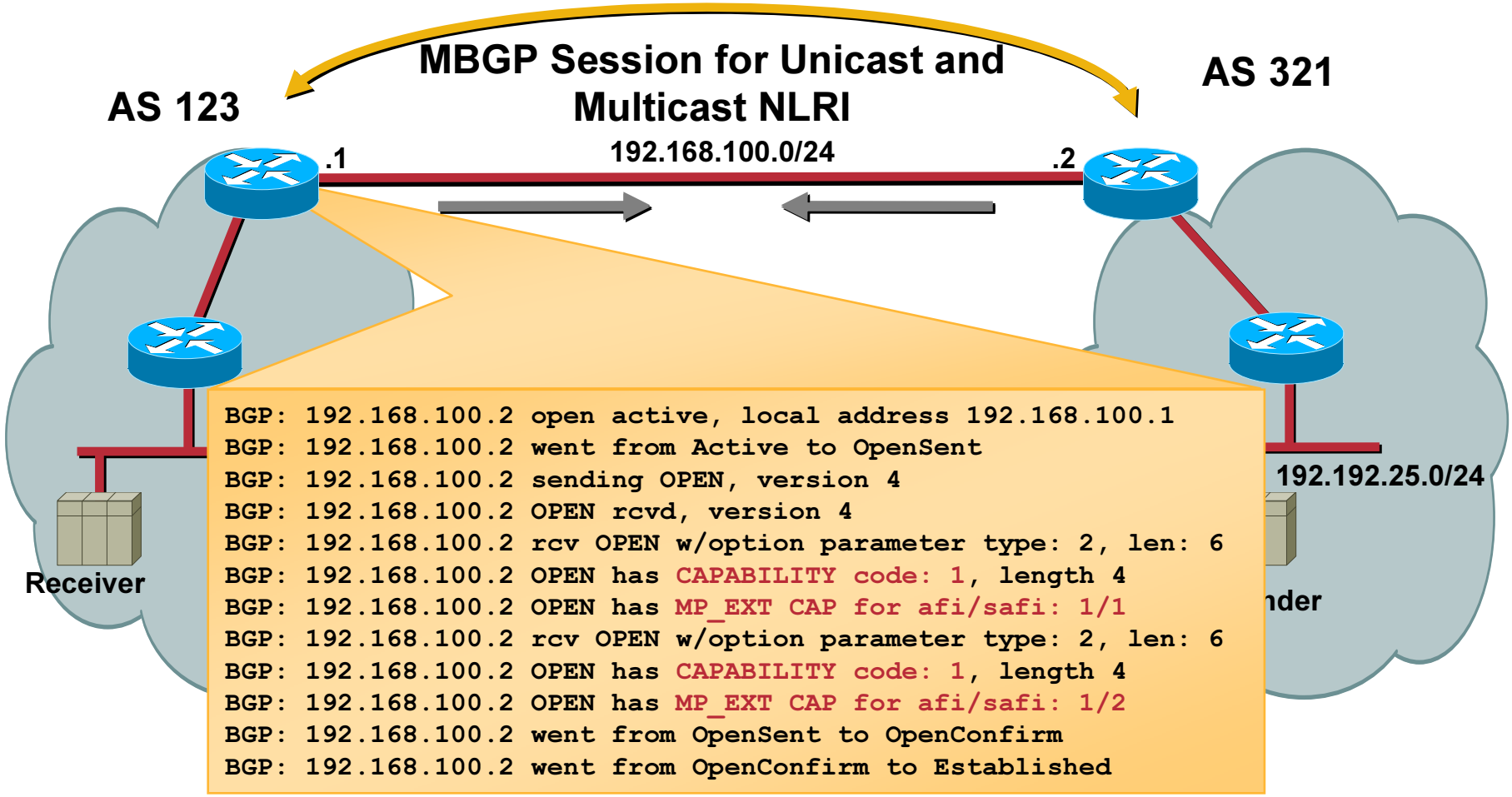
MBGP — Capability Negotiation



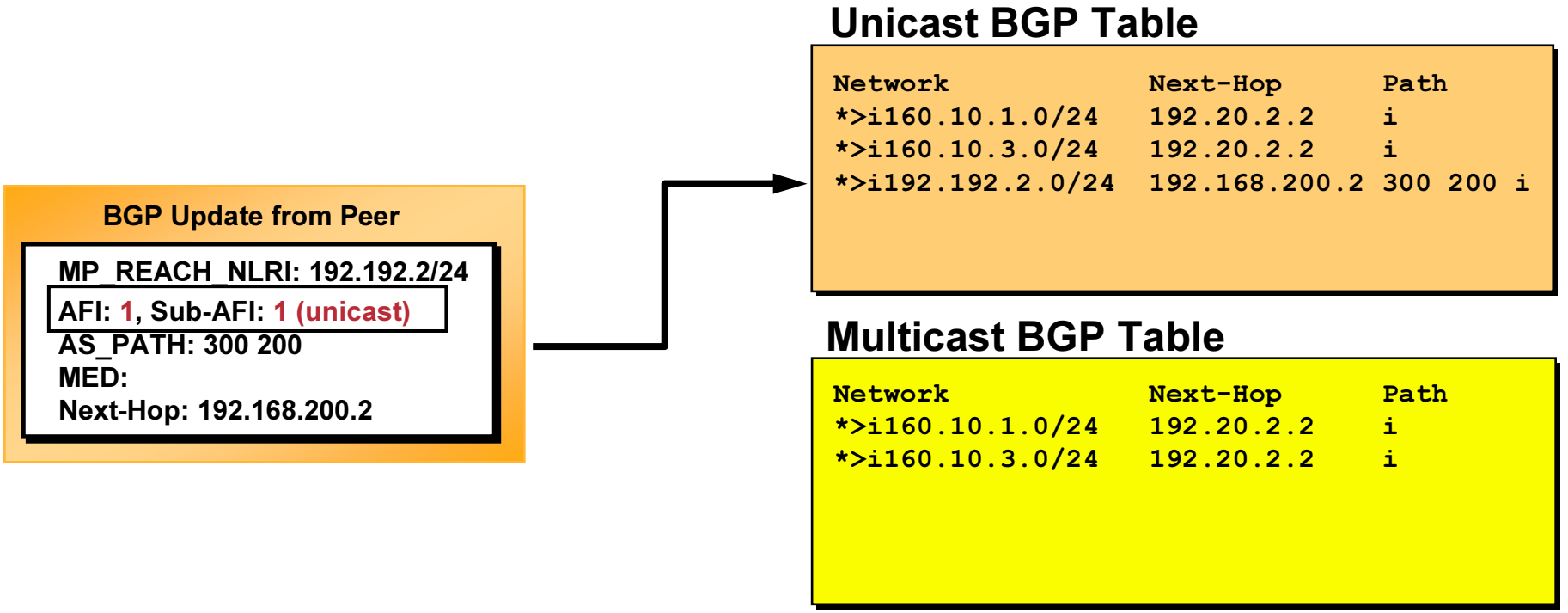
MBGP — Capability Negotiation



MBGP — Capability Negotiation

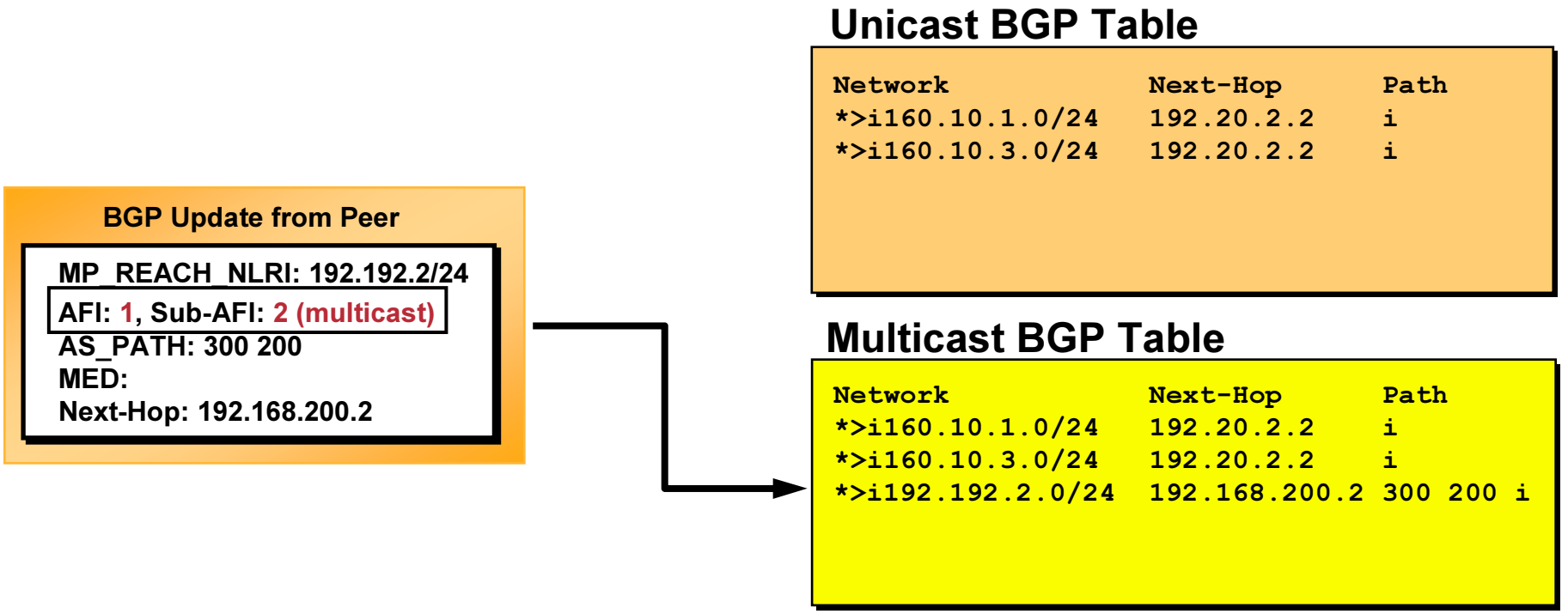


MBGP—NLRI Information



- 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)

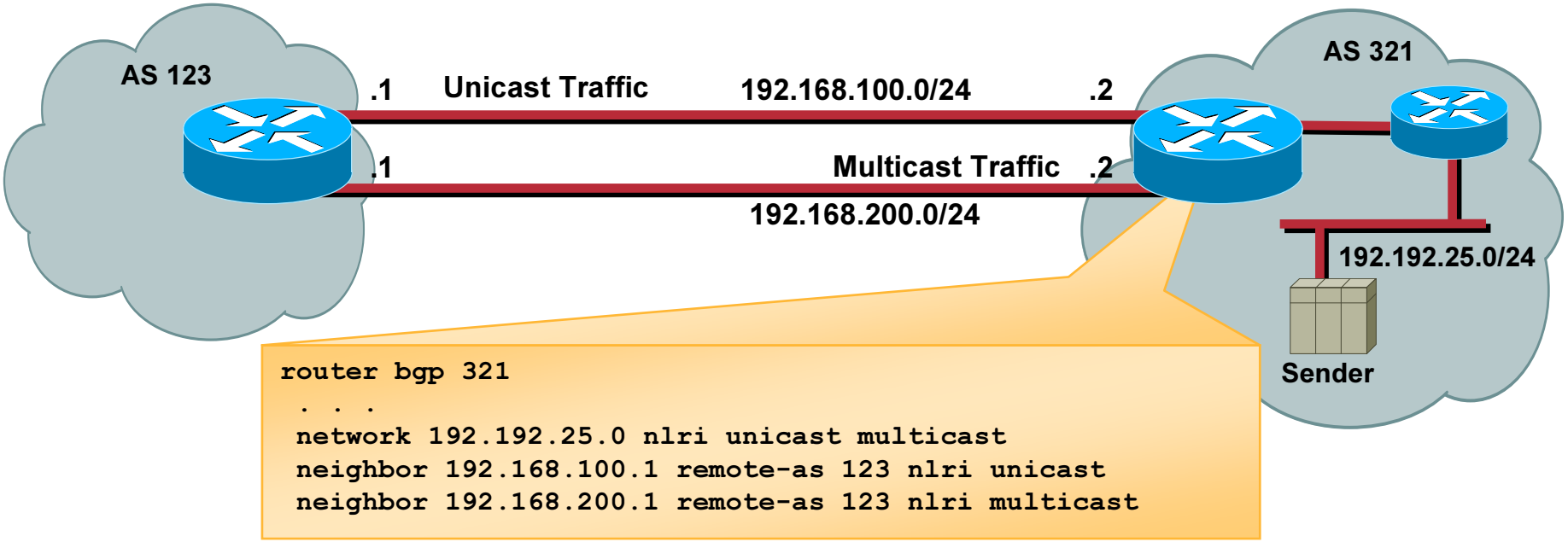
MBGP—NLRI Information



- 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

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

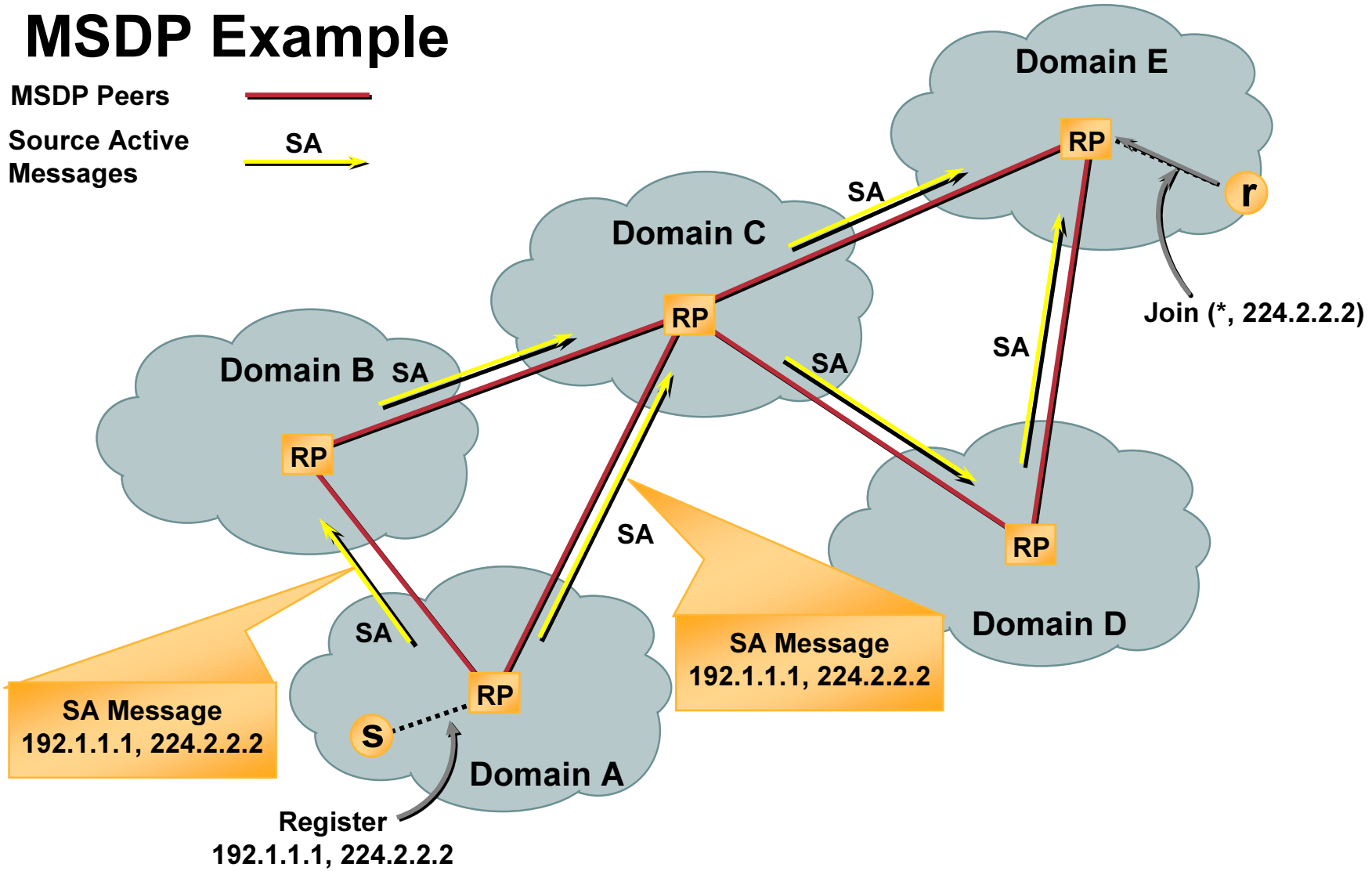
```
router bgp 5
  no bgp default ipv4-unicast
  neighbor 171.69.214.38 remote-as 2
  neighbor 171.69.214.50 remote-as 2
!
  address-family ipv4 unicast
  neighbor 171.69.214.38 activate
  network 171.69.214.0 mask 255.255.255.0
  exit-address-family
!
  address-family ipv4 multicast
  neighbor 171.69.214.50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

Multicast Source Discovery Protocol (MSDP)

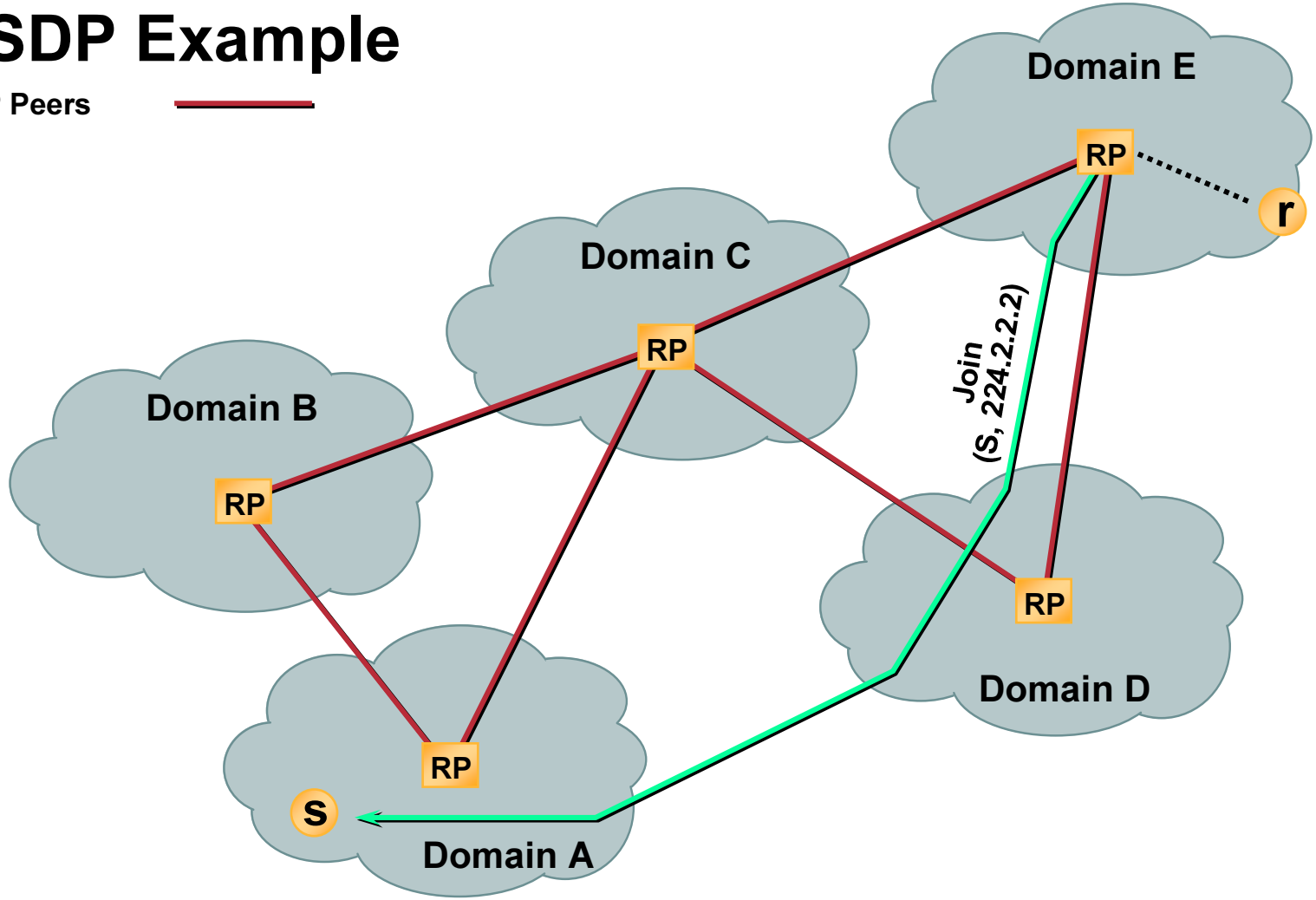
MSDP Overview



MSDP Overview

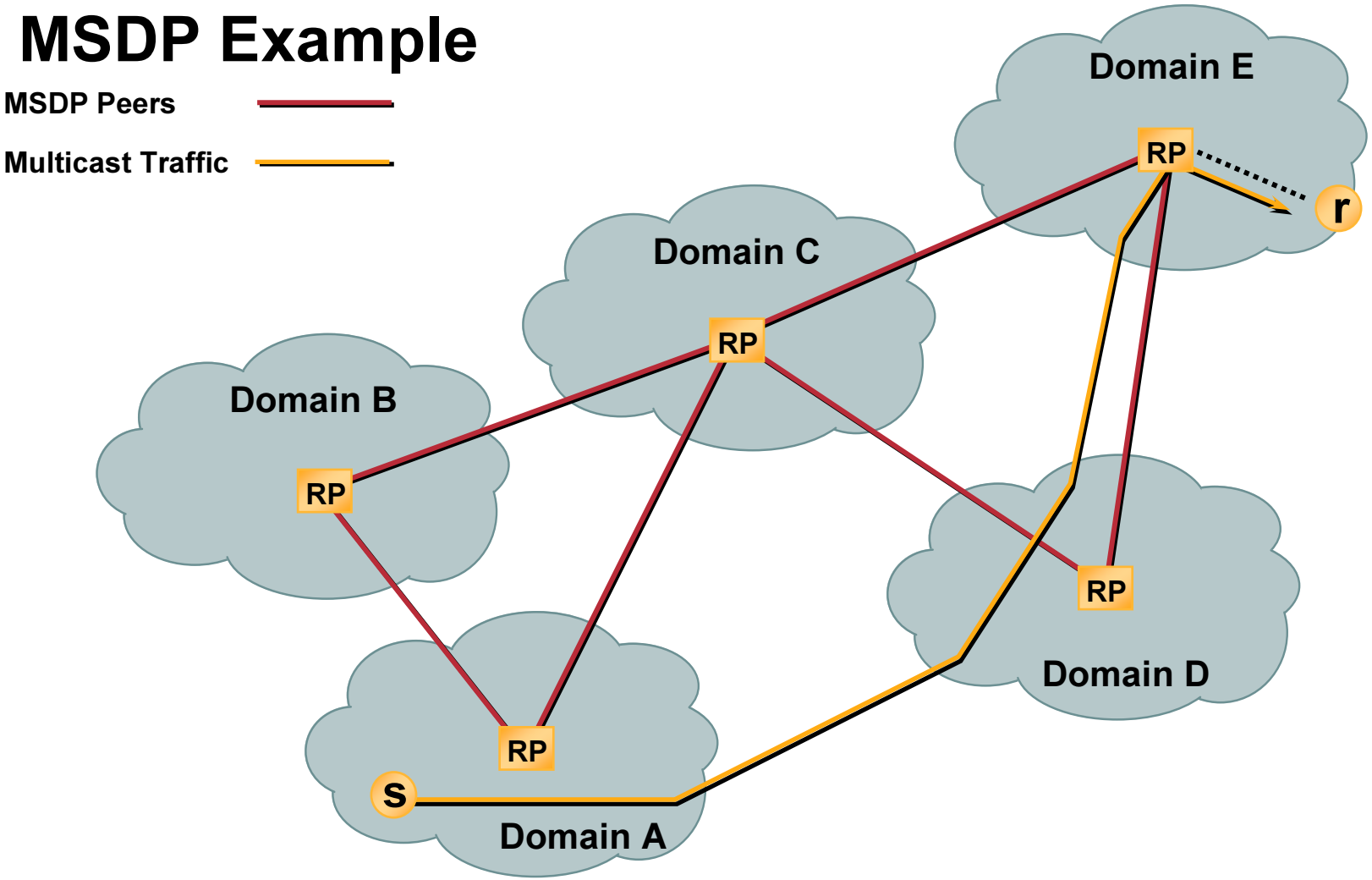
MSDP Example

MSDP Peers



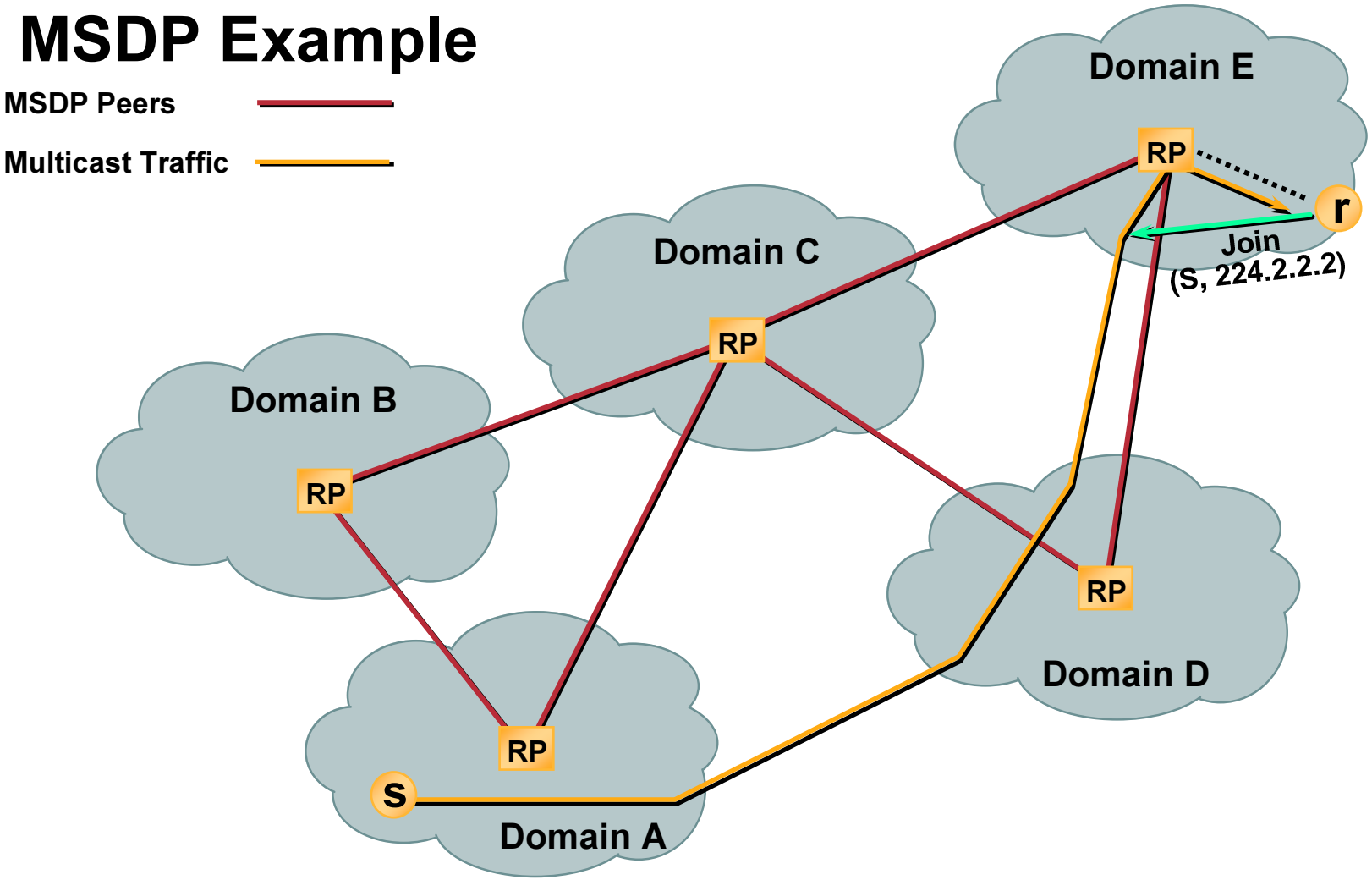
MSDP Overview

MSDP Example



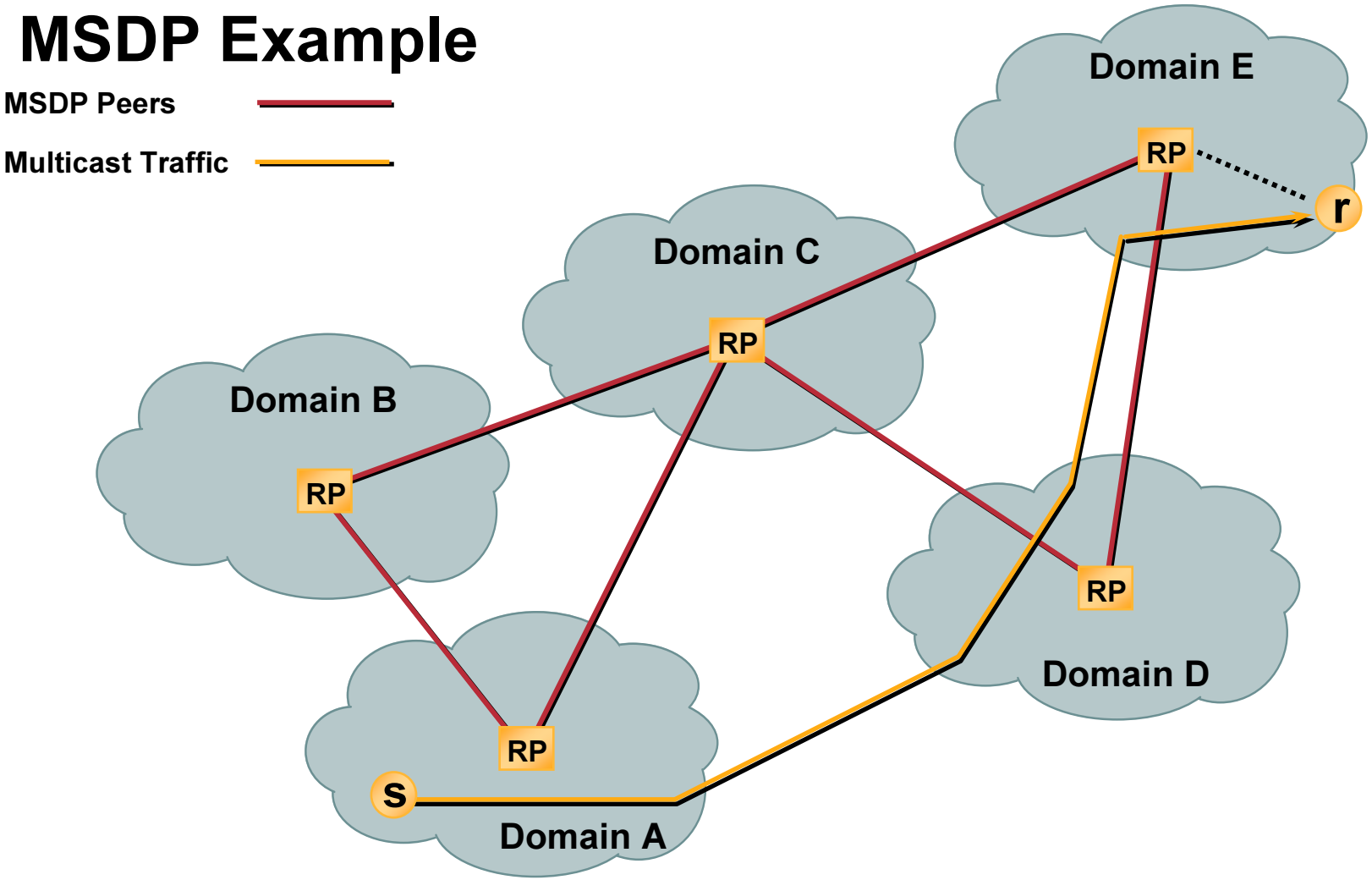
MSDP Overview

MSDP Example



MSDP Overview

MSDP Example



MSDP SA Messages

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

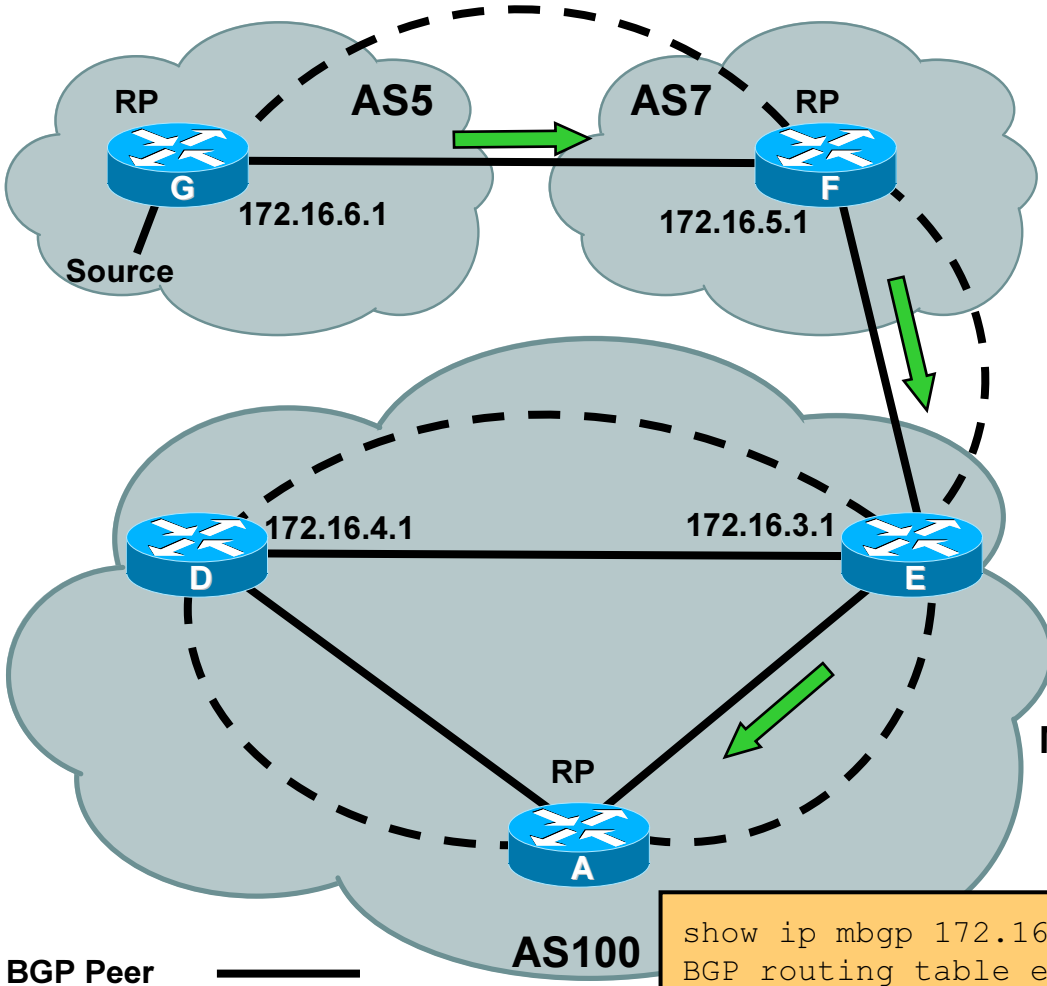
Receiving SA Messages

- RPF Check Rules depend on peering
  - Rule 1: Sending MSDP peer = i(m)BGP peer
  - Rule 2: Sending MSDP peer = e(m)BGP peer
- 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)
    - This is not the same as the Next-hop of the path!!
  - 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)

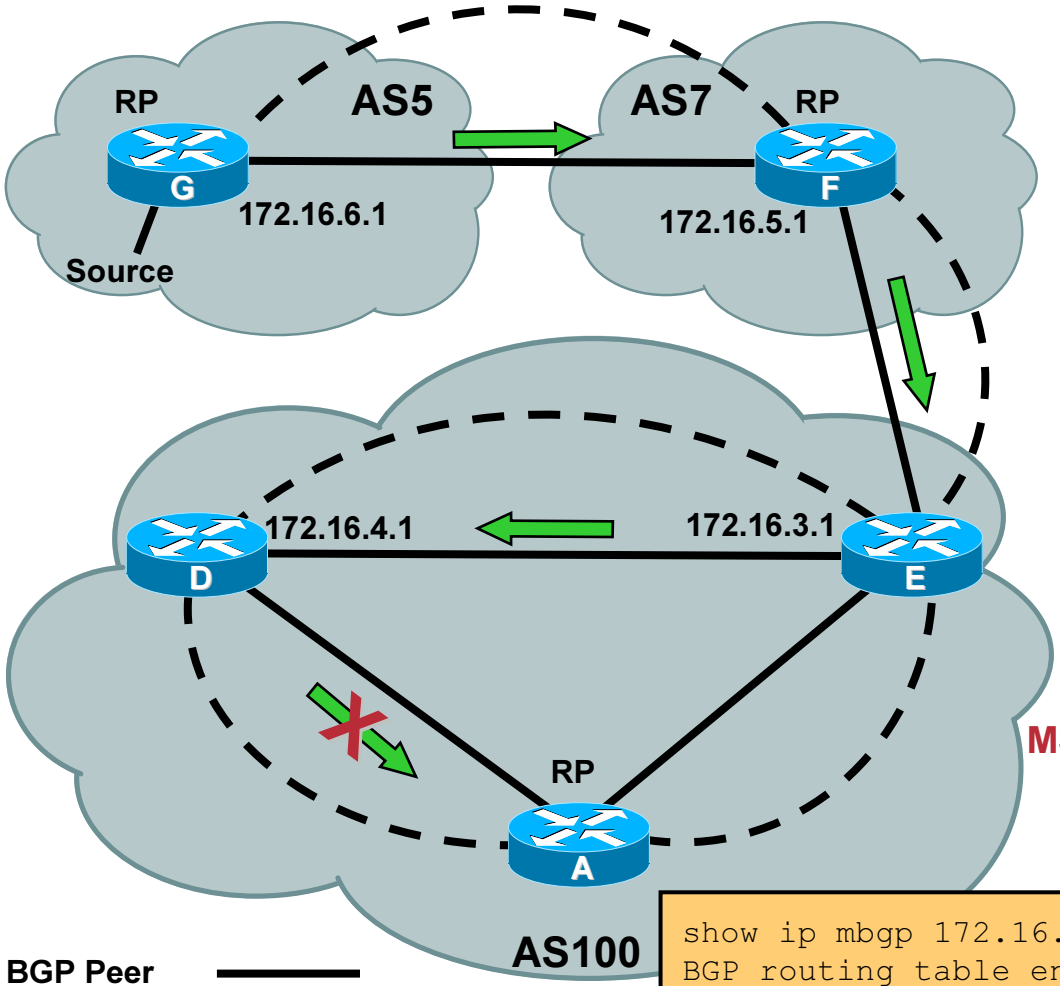
MSDP Peer address = 172.16.3.1

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)

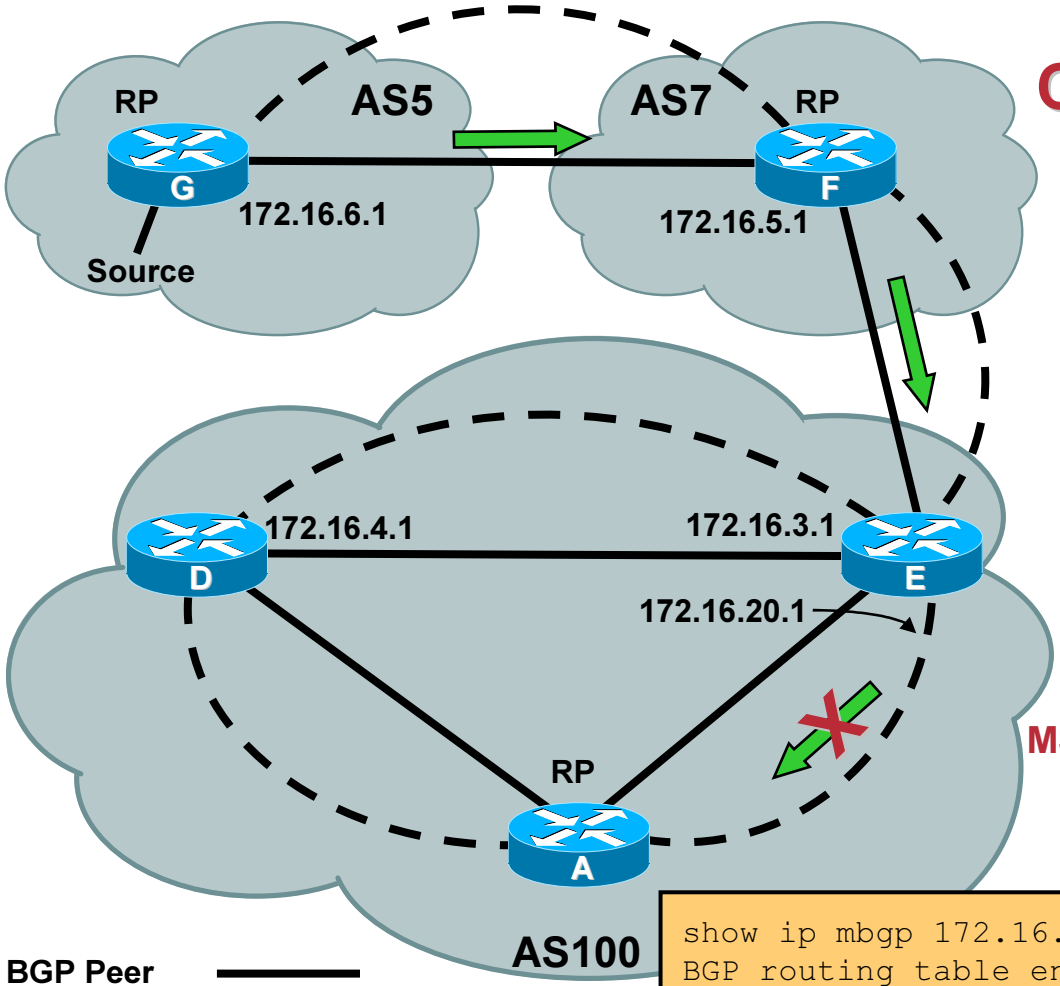
MSDP Peer address = 172.16.4.1

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)

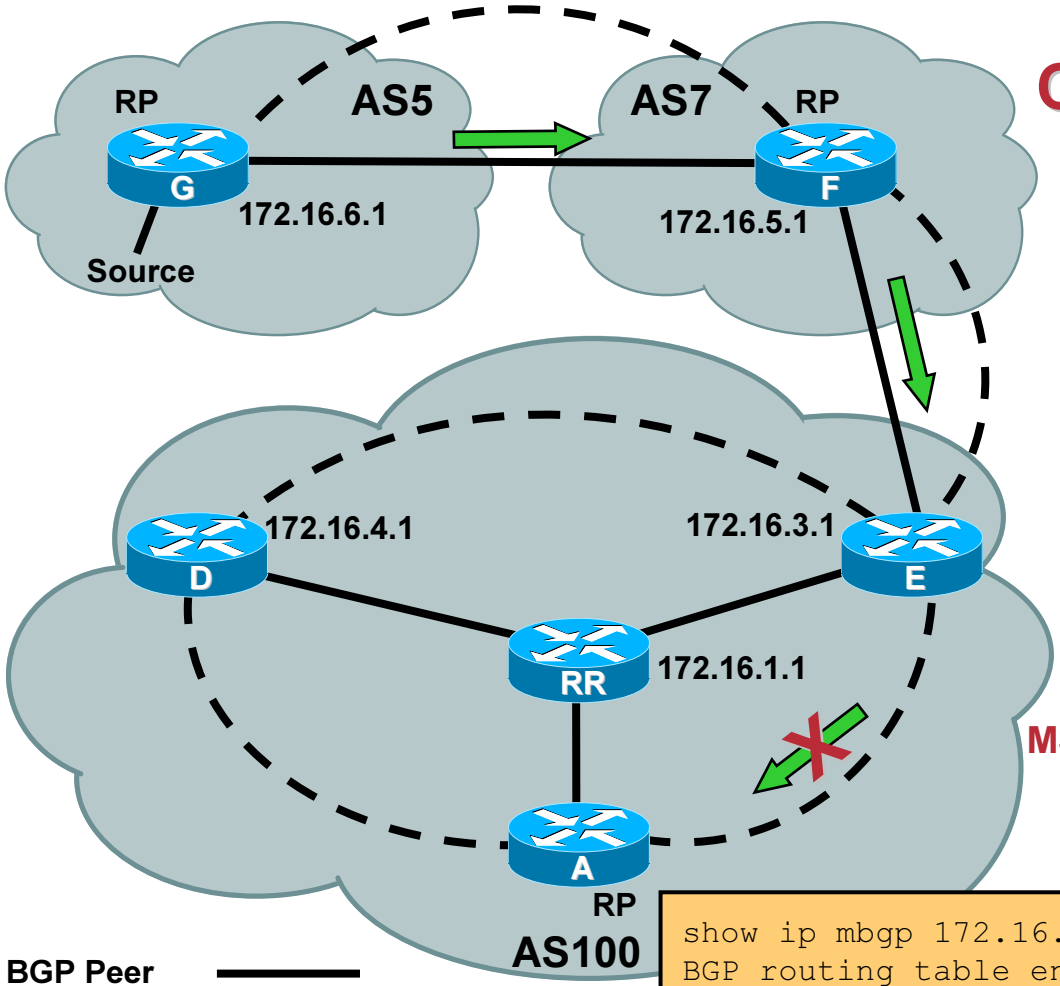
MSDP Peer address = 172.16.20.1

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

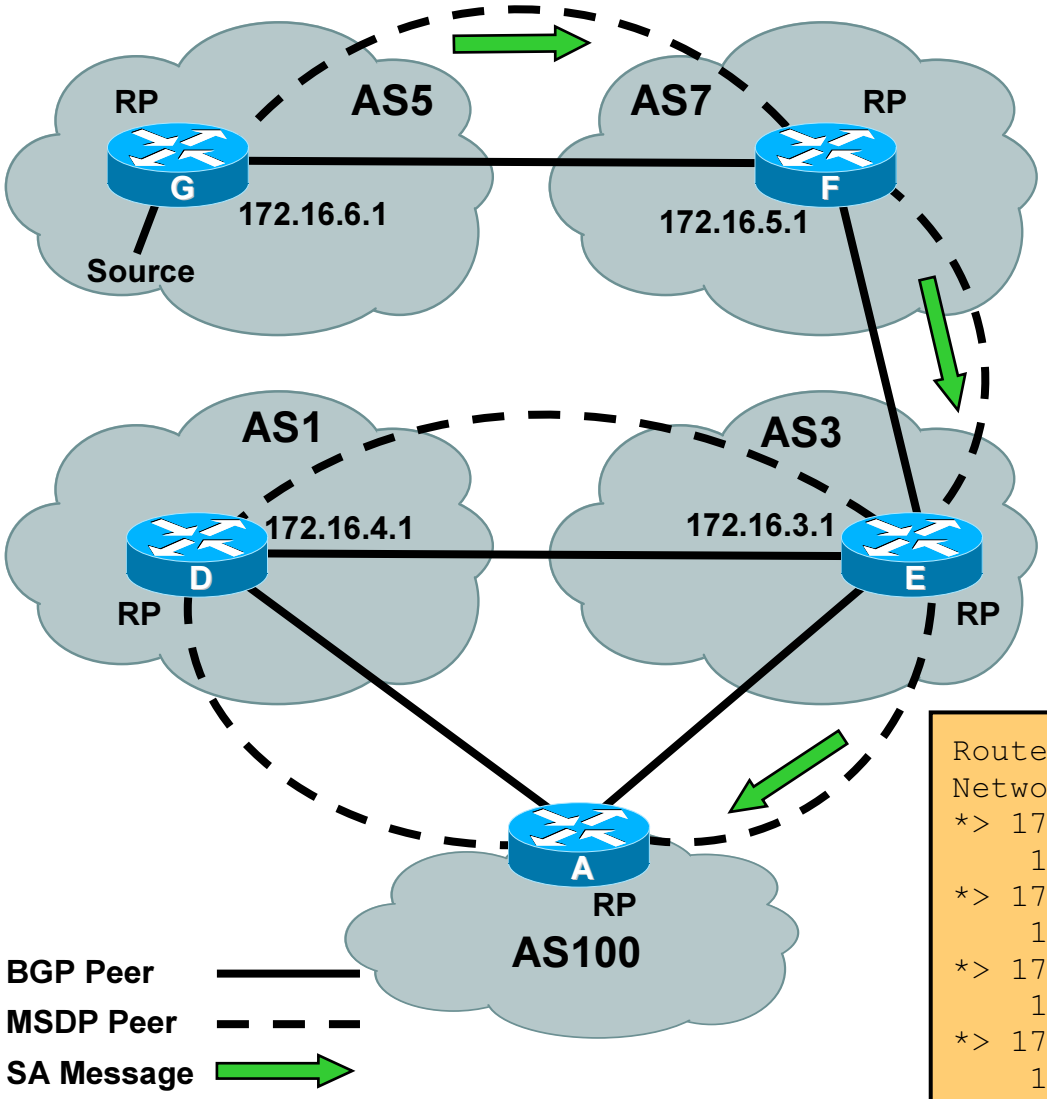
SA RPF Check Fails

```
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



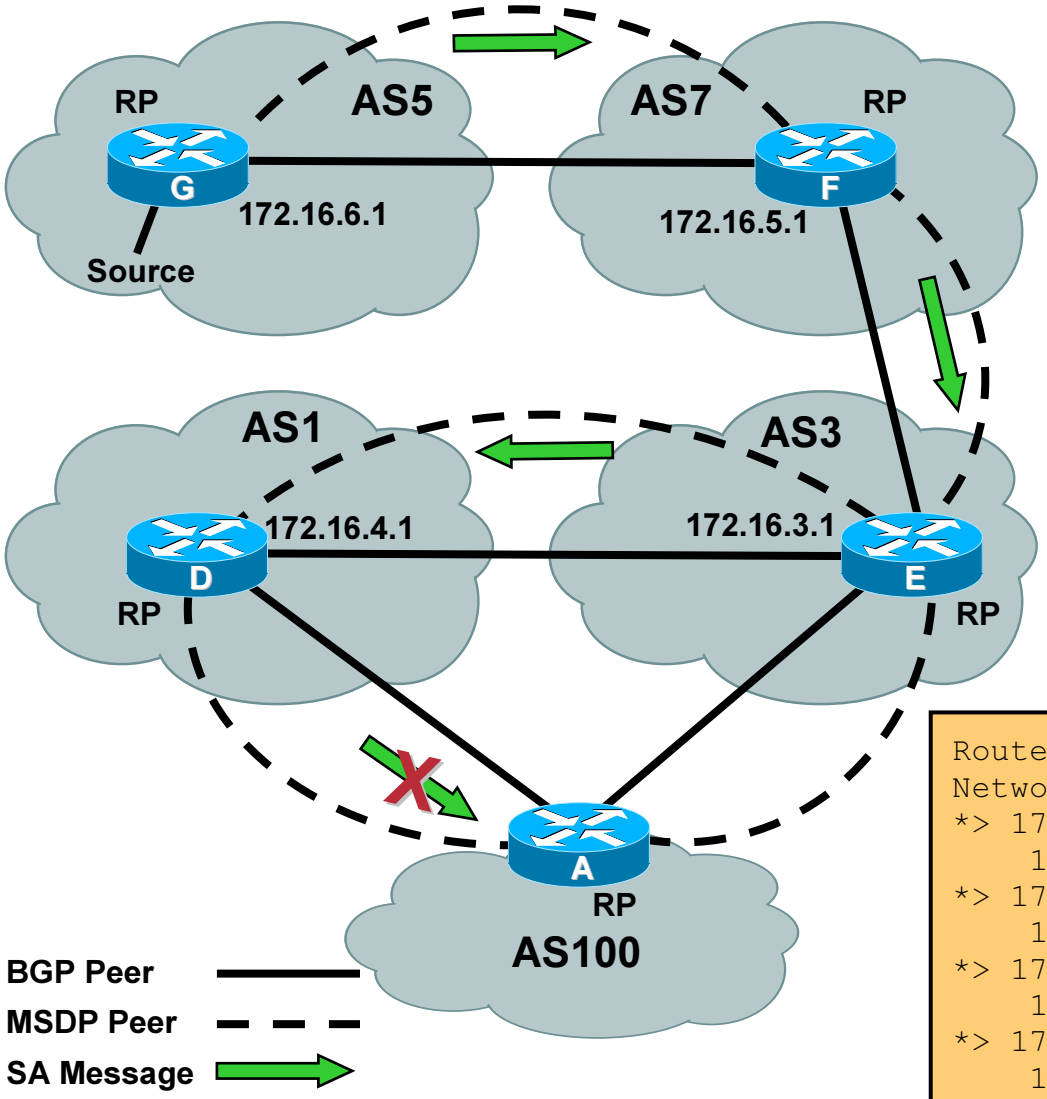
First-AS in best-path to RP = 3
AS of MSDP Peer = 3

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



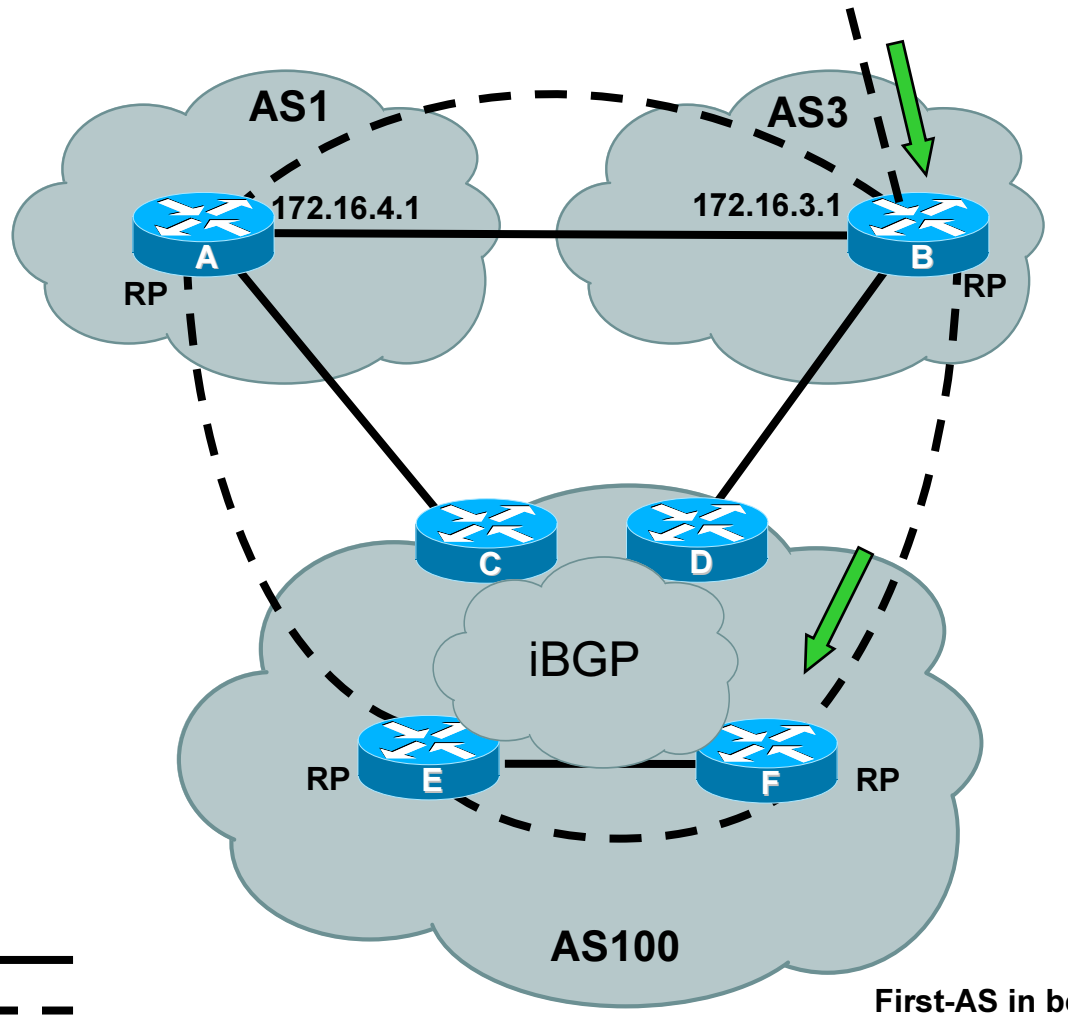
First-AS in best-path to RP = 3
AS of MSDP Peer = 1

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
```

Common MSDP Deployment



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

- RFC 3618
- Filtering
  - Can filter SA in/out, groups, (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-06.txt

MSDP Enhancements

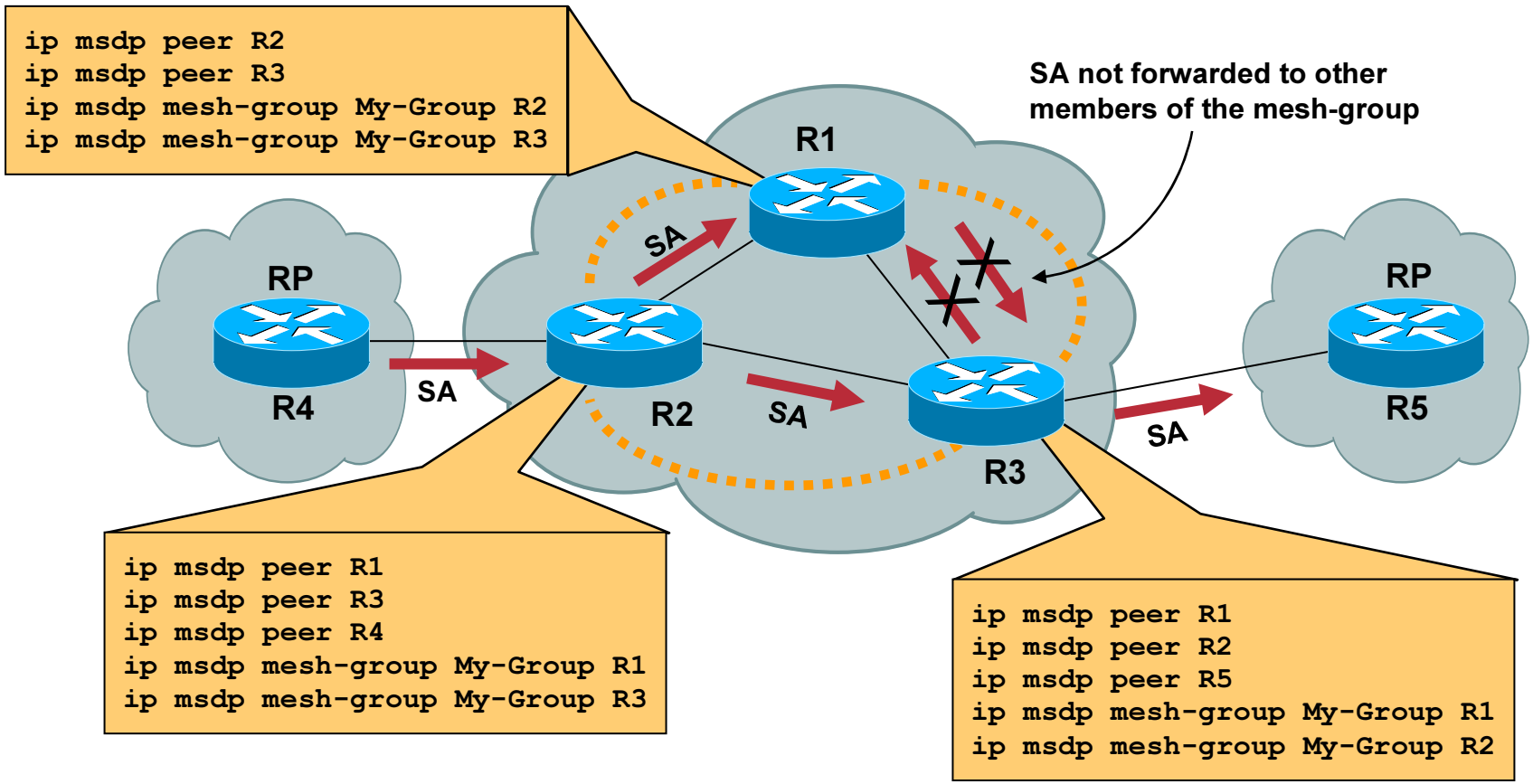
- New IOS command
  - ip msdp rpf rfc3618
  - MSDP SA RPF check using IGP
  - Accept SA's from BGP NEXT HOP
  - Accept SA's from closest peer along the best path to the originating RP
    - show ip msdp rpf
  - 12.0(27)S

New MSDP RPF command

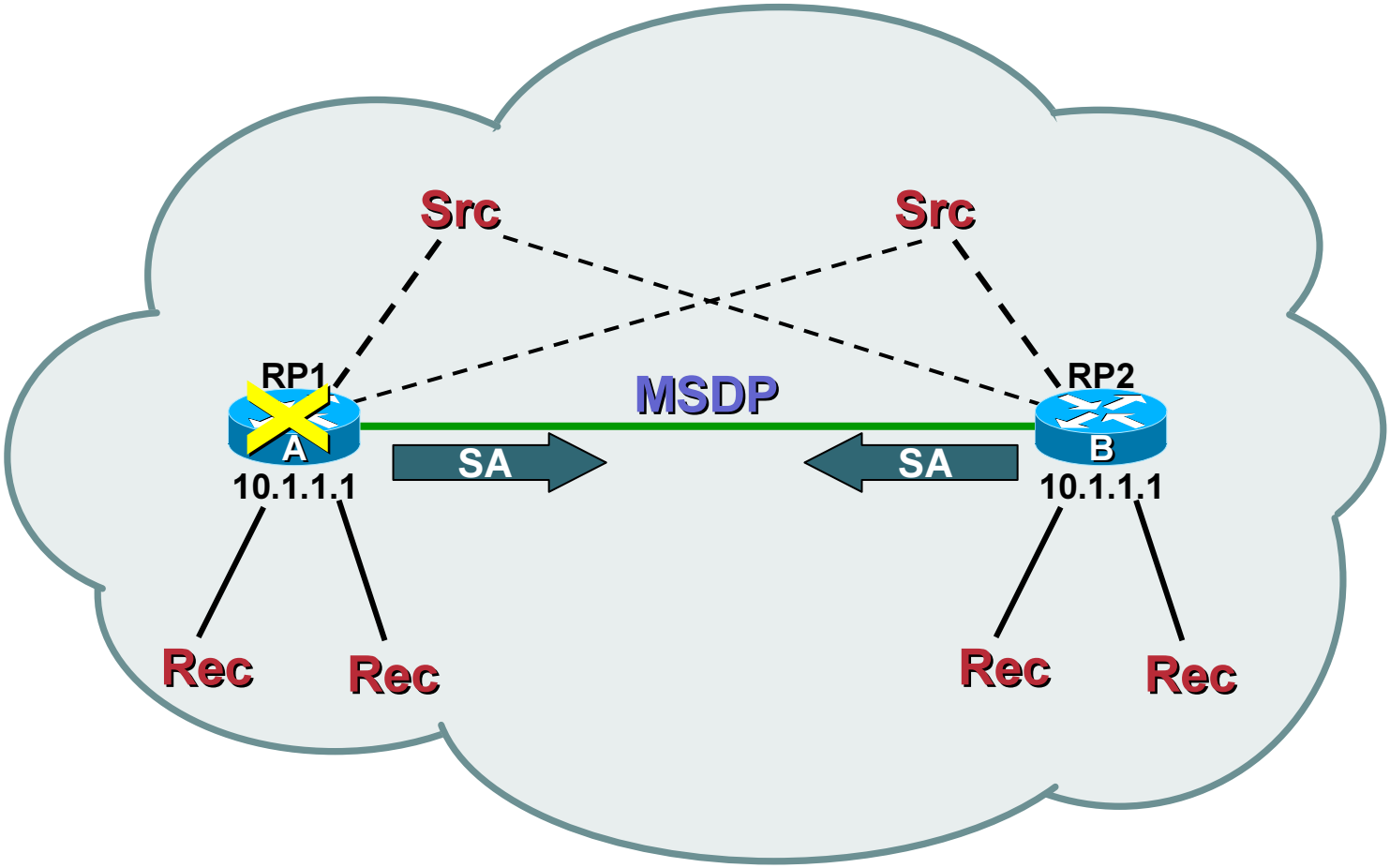
```
Router-A# show ip msdp rpf 2.1.1.1
RPF peer information for Router-B (2.1.1.1)
  RPF peer: Router-C (3.1.1.1)
  RPF route/mask: 2.1.1.0/24
  RPF rule: Peer is IGP next hop of best route
  RPF type: unicast (ospf 1)
```

MBGP/MSDP Examples

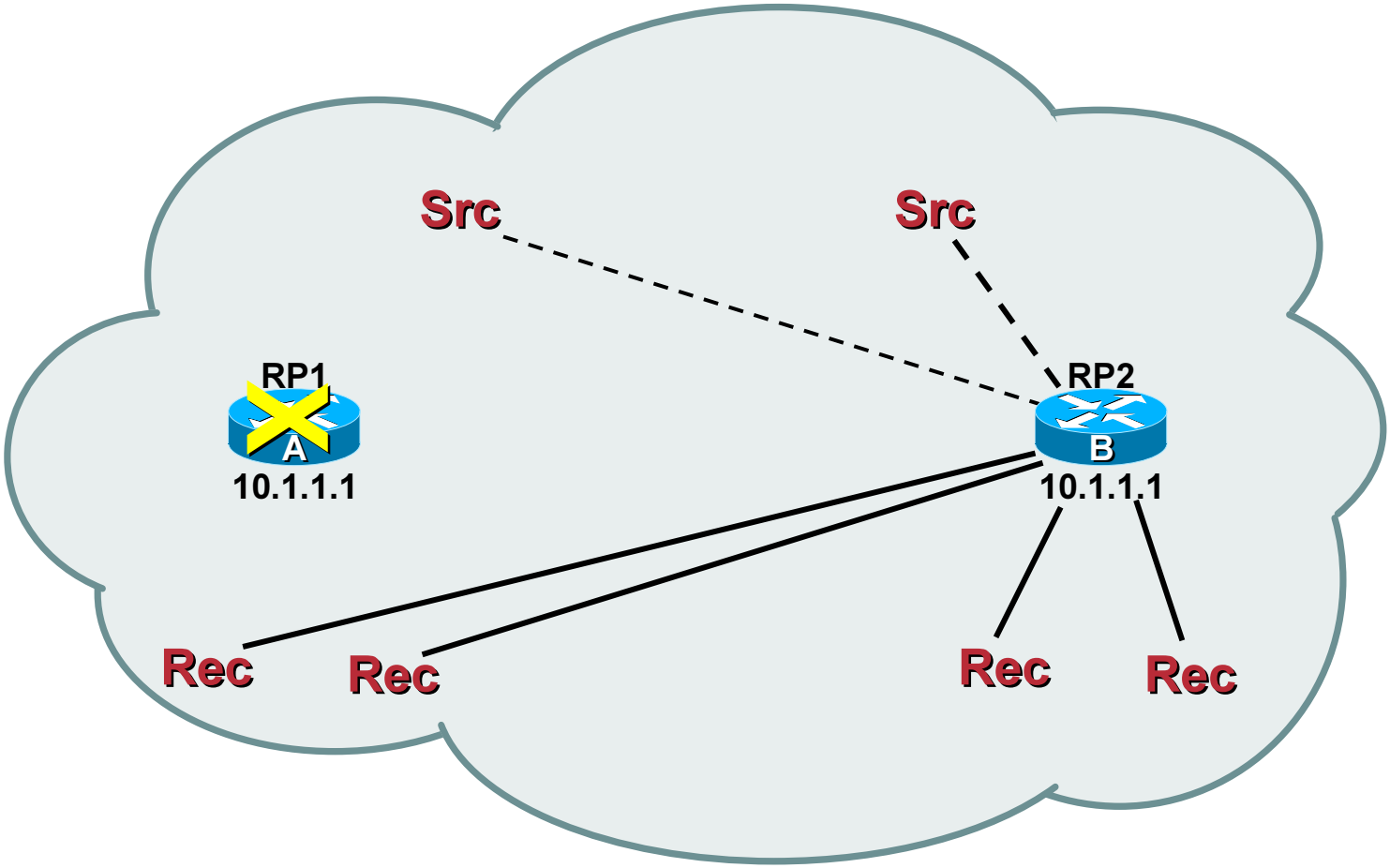
MSDP Mesh-Group Example



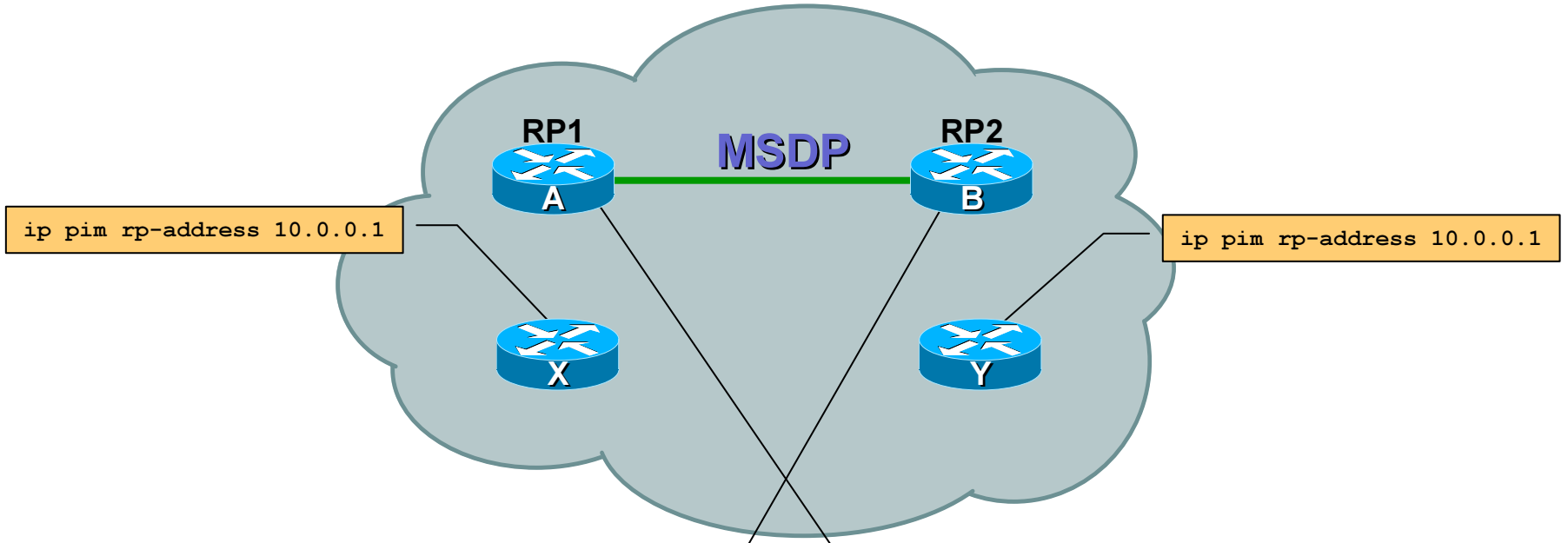
Anycast RP—Overview



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
```

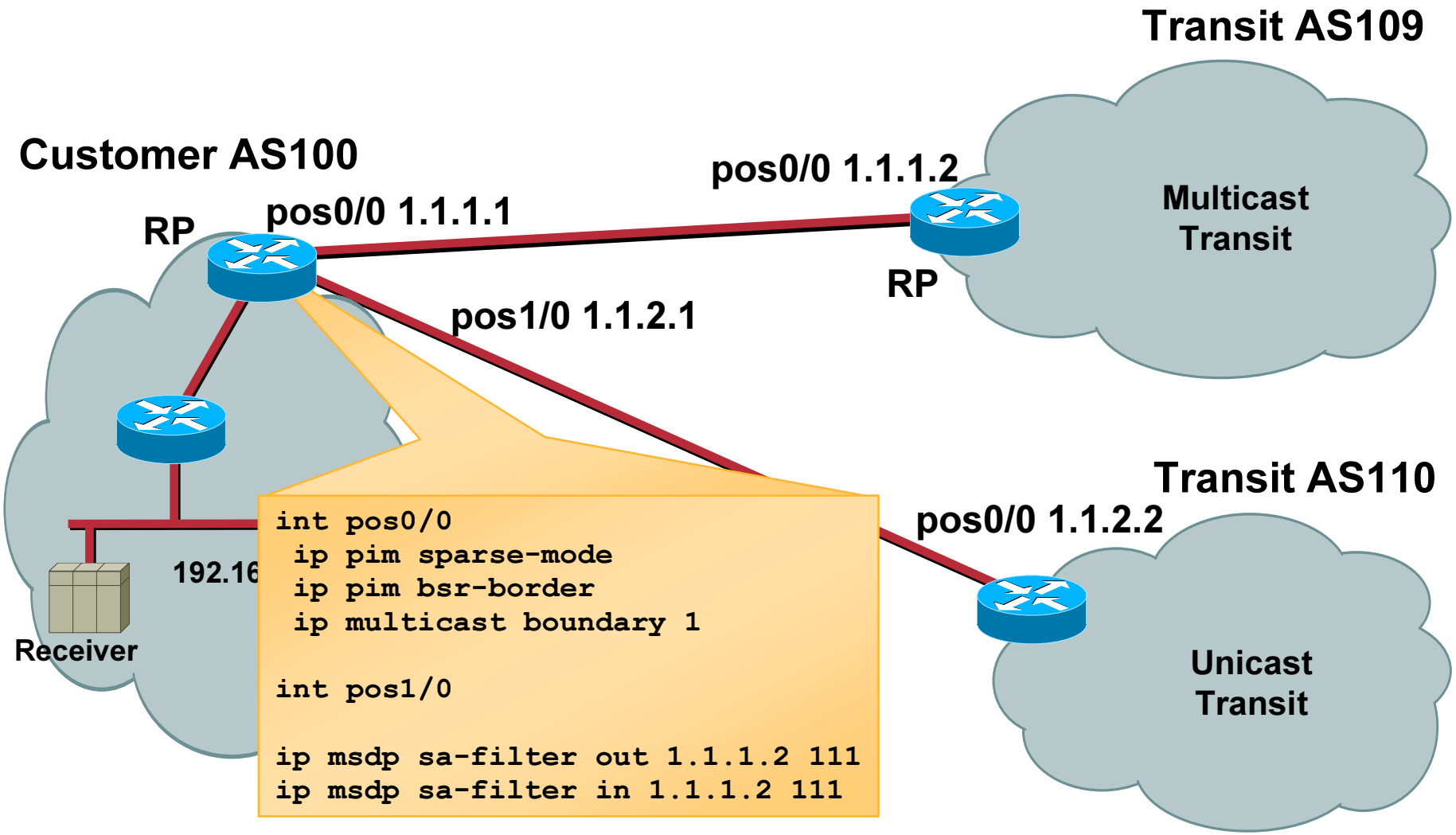
```
Interface loopback 0
 ip address 10.0.0.3 255.255.255.255

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
```

MSDP SA Filtering

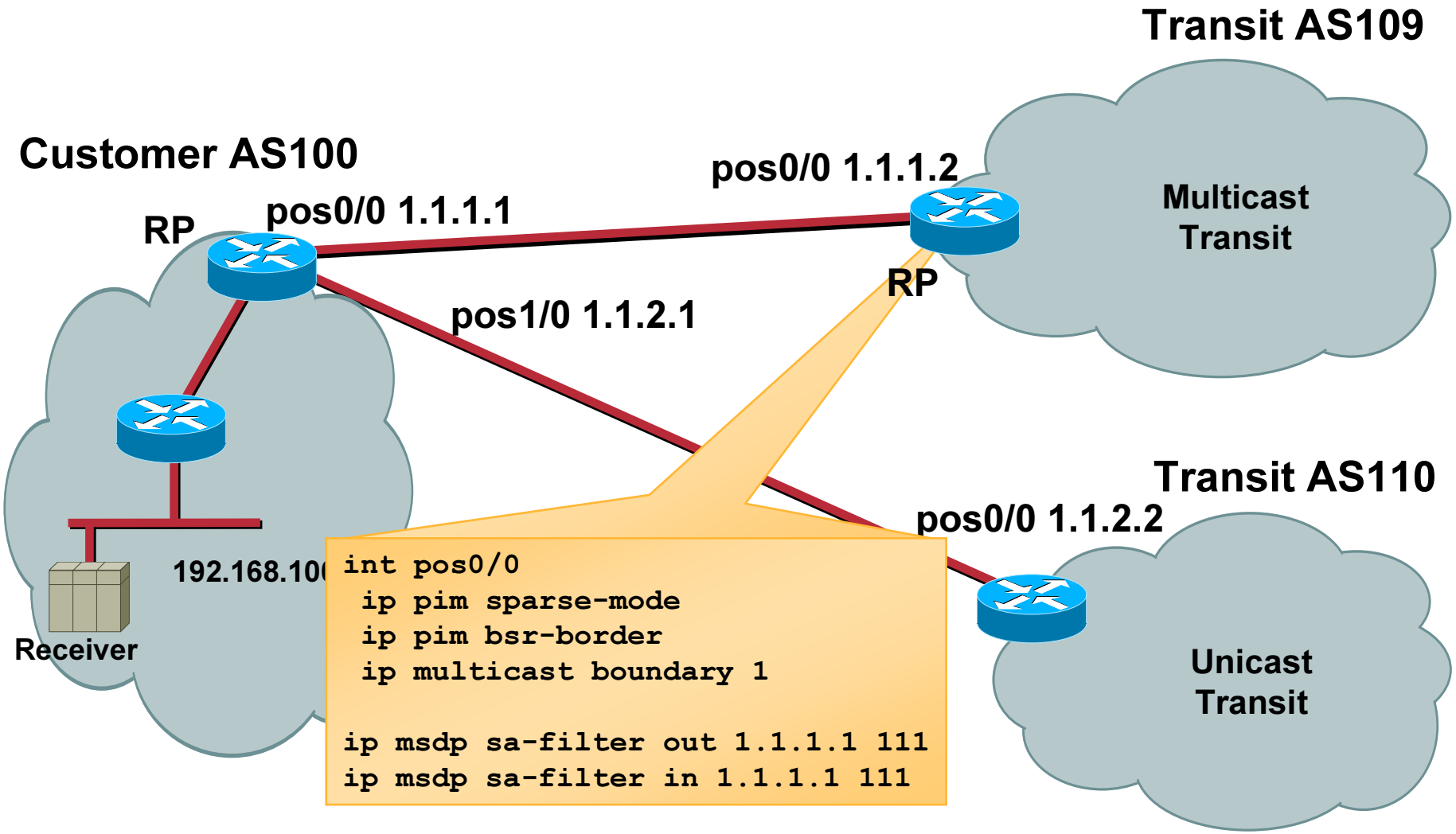
ftp://ftpeng.cisco.com/ipmulticast/config-notes/msdp-sa-filter.txt

```
! 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
```

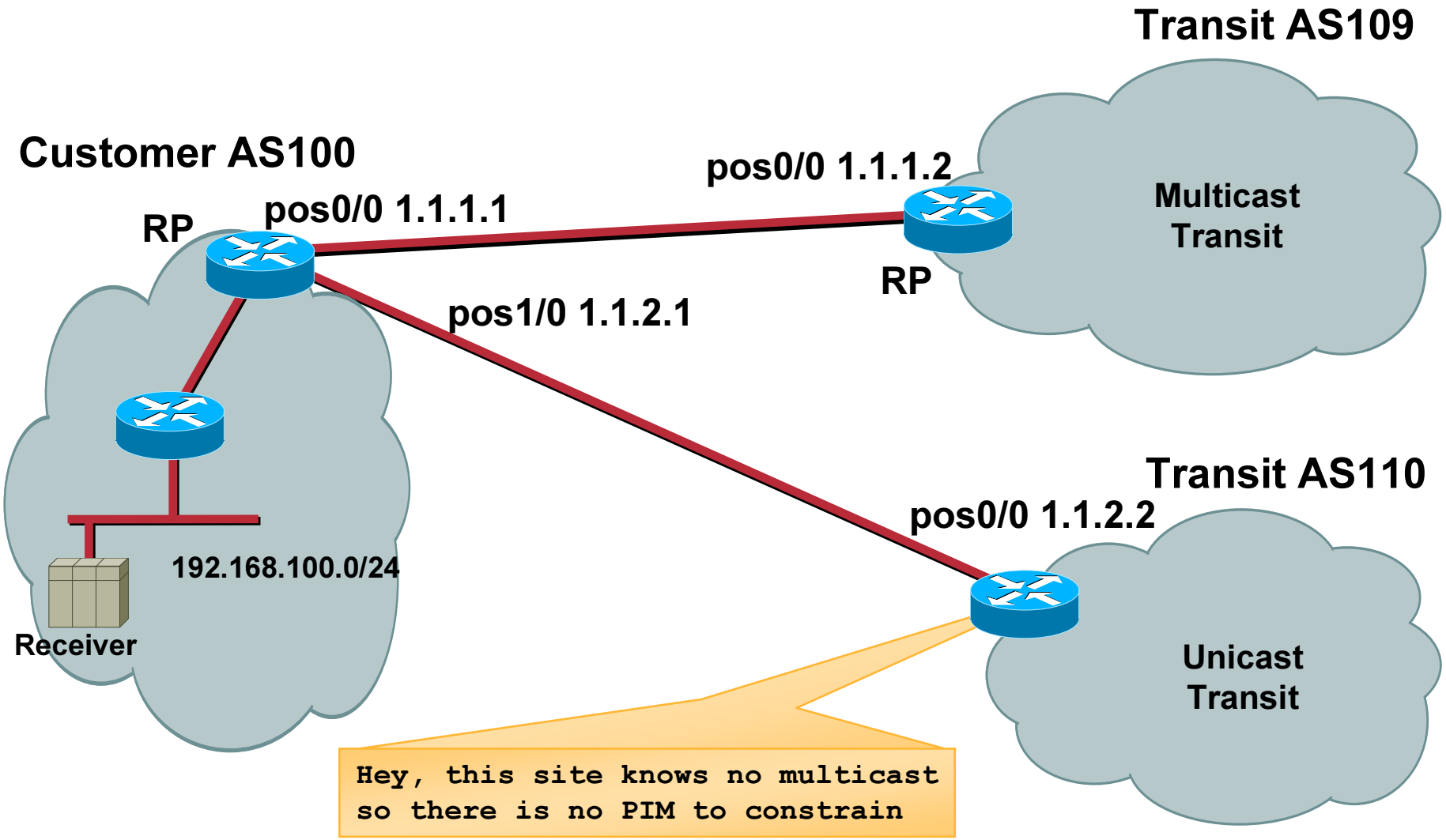
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



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

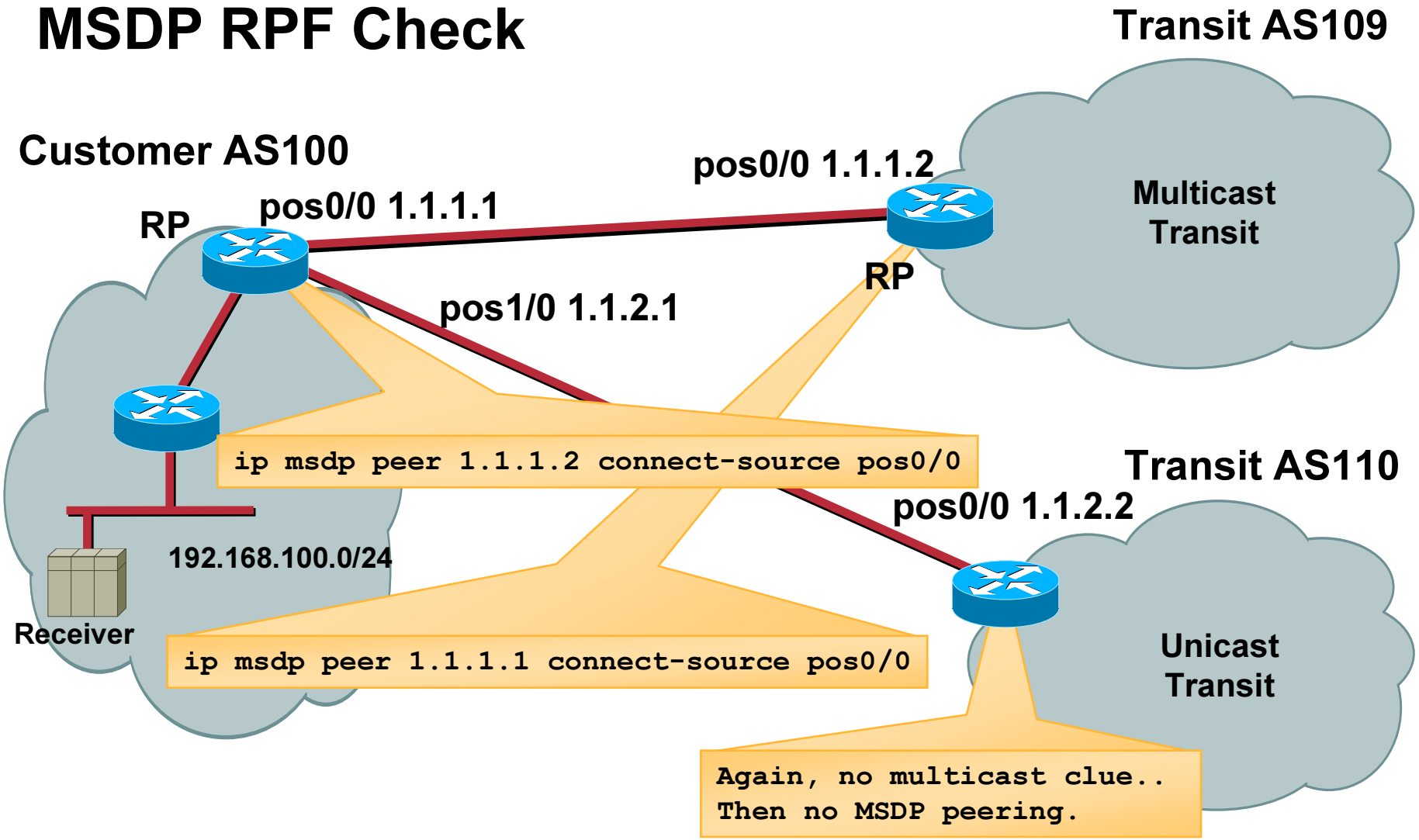


Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



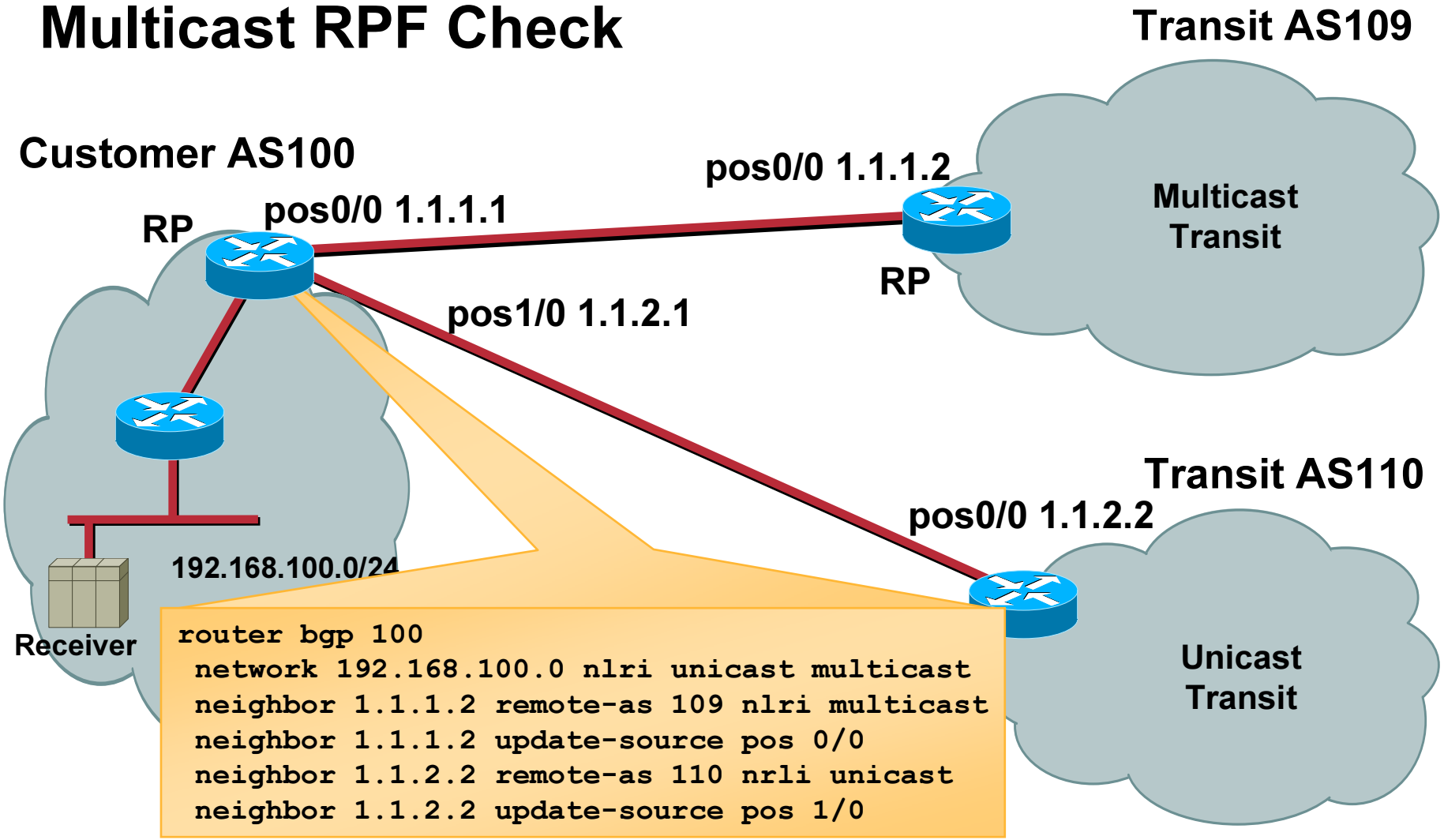
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

MSDP RPF Check



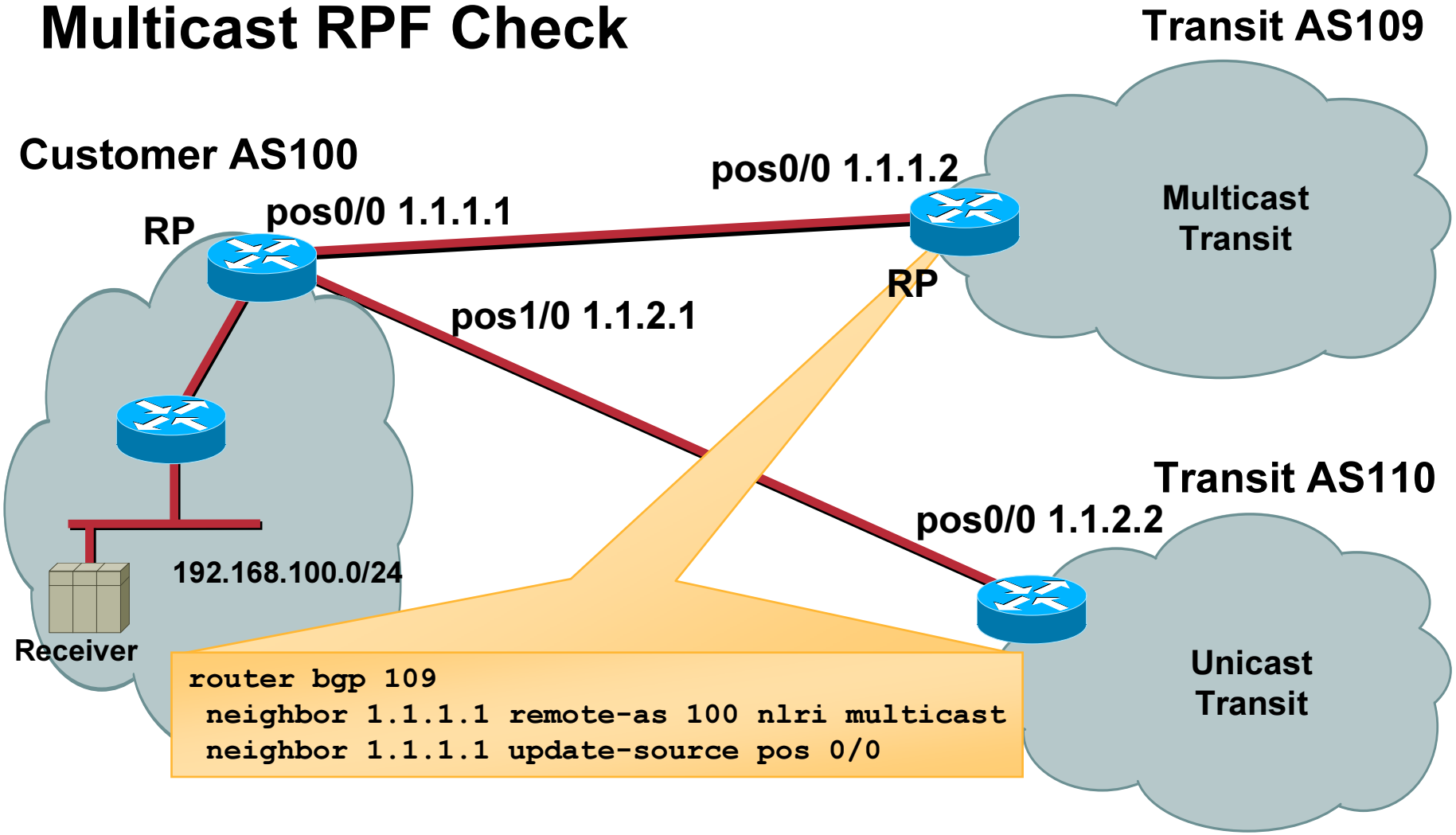
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Multicast RPF Check



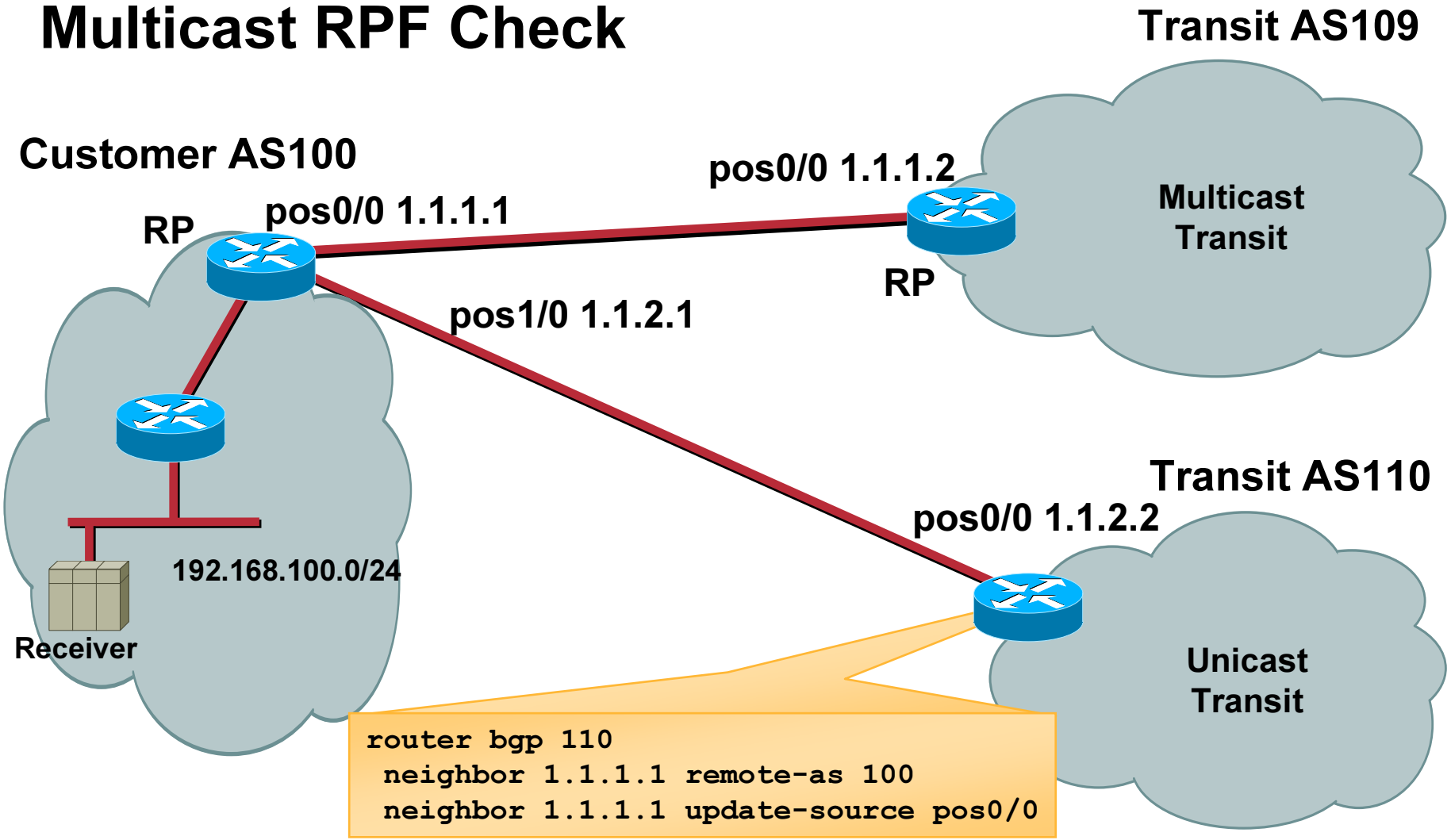
Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast

Multicast RPF Check



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

Multicast RPF Check



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

Source Specific Multicast (SSM)

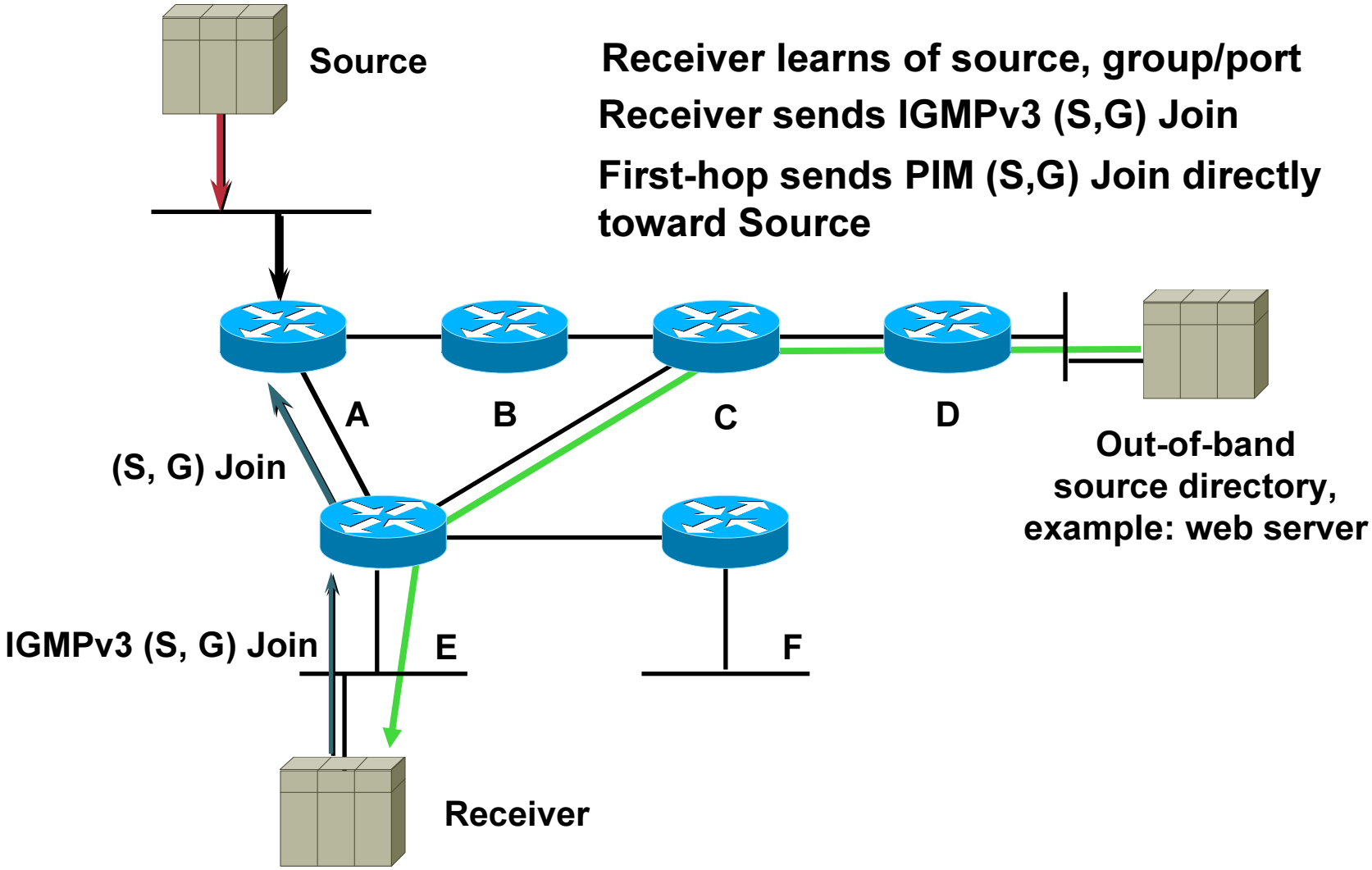
Source Specific Multicast (SSM)

- 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.

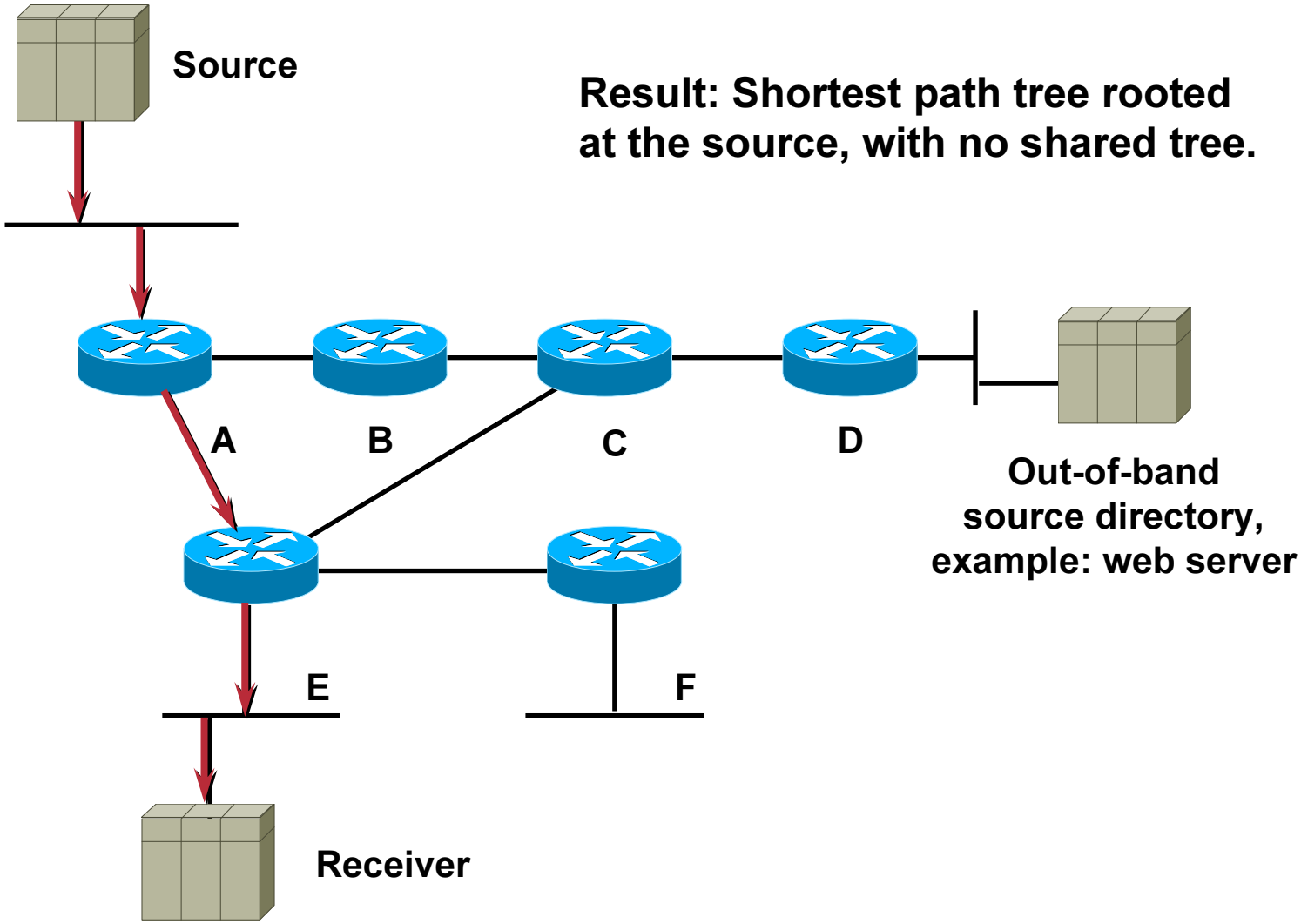
SSM Overview

- 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.

PIM Source Specific Mode



PIM Source Specific Mode



SSM Configuration

- **Global command**
  - **ip pim ssm {default | <acl>}**
  - **Defines SSM address range**
    - **Default range = 232.0.0.0/8**
  - **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**

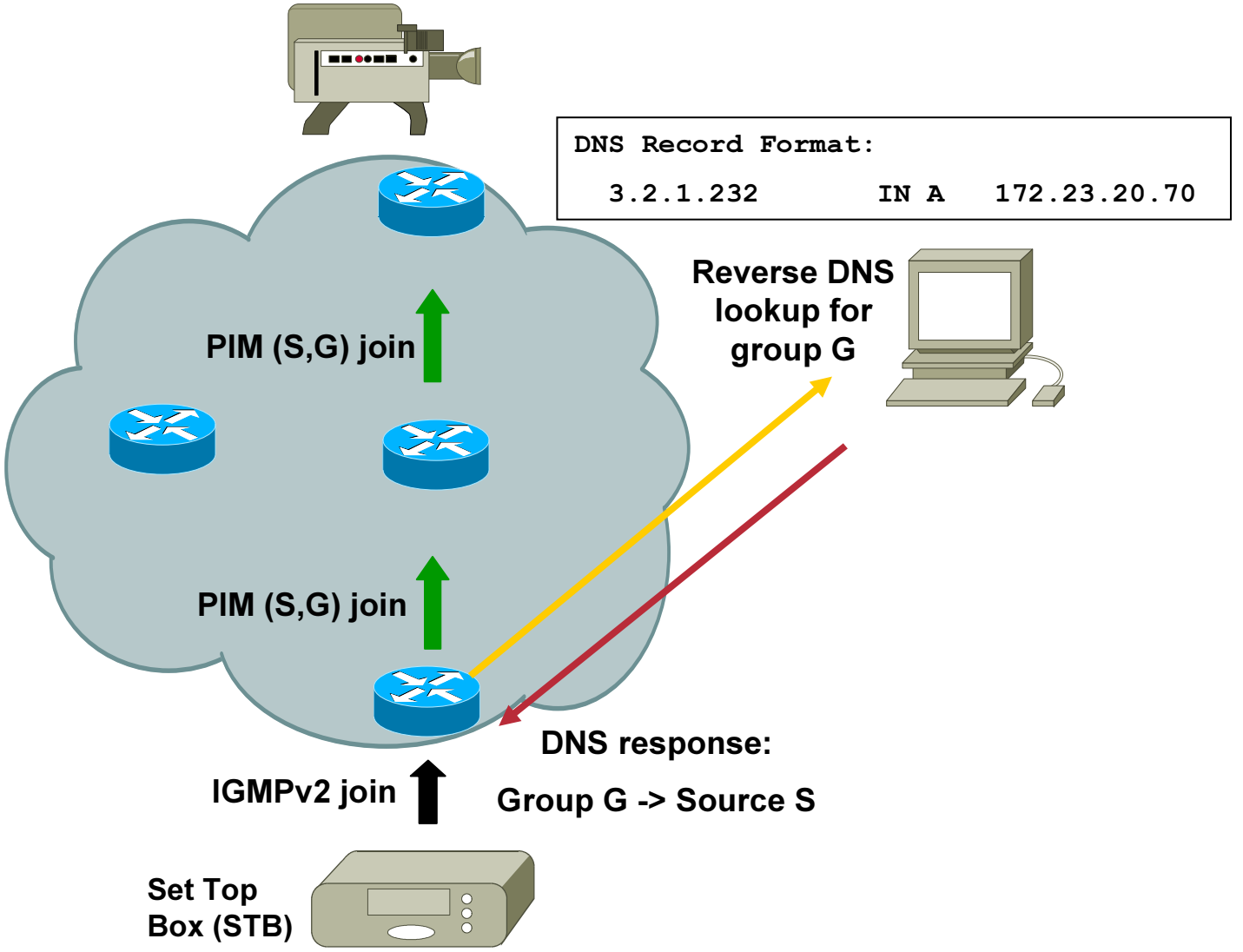
SSM Mapping

- 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

SSM Mapping – DNS 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
```

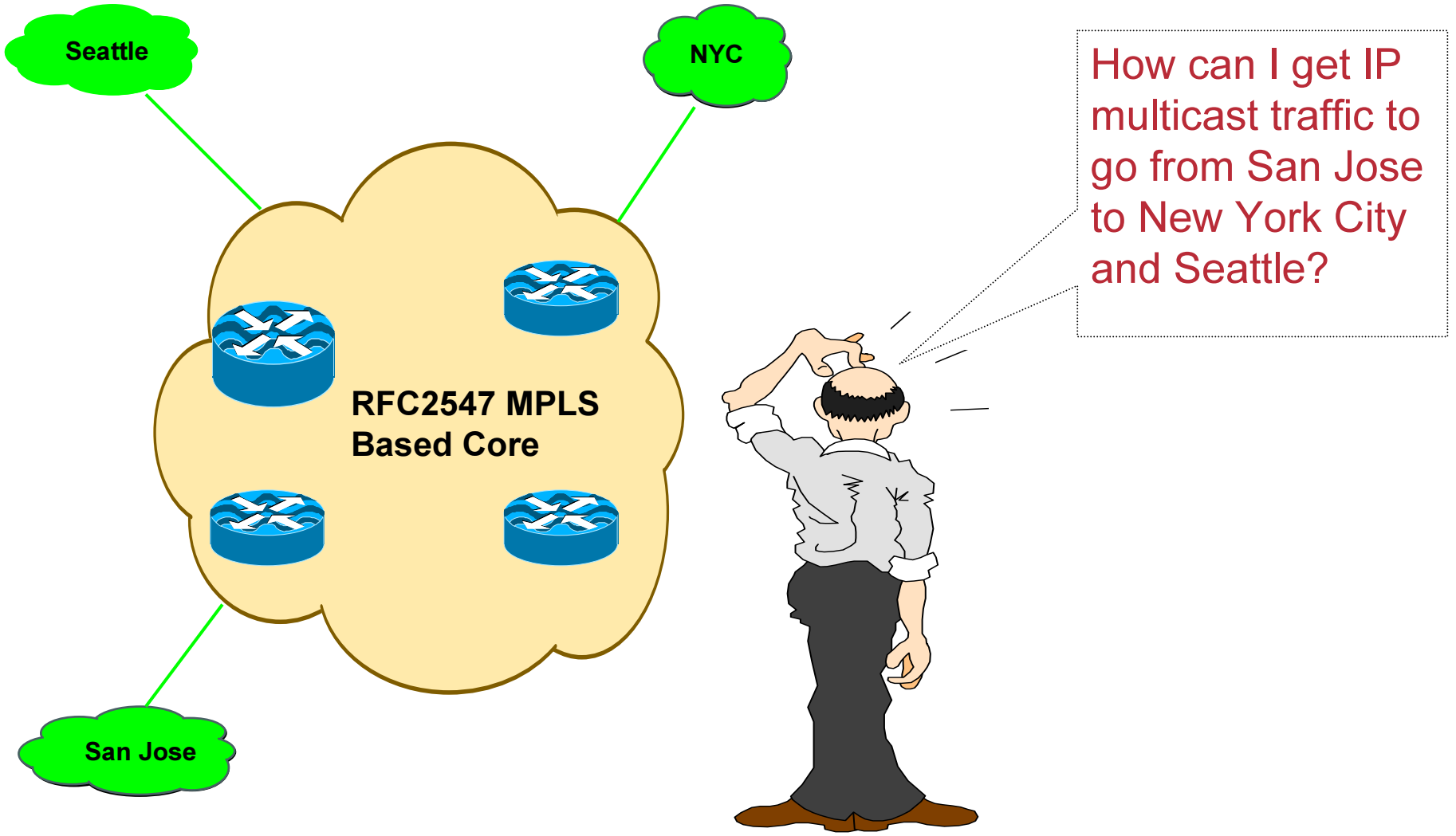
```
DNS Record Format:    3.2.1.232          IN A    172.23.20.70
```

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.

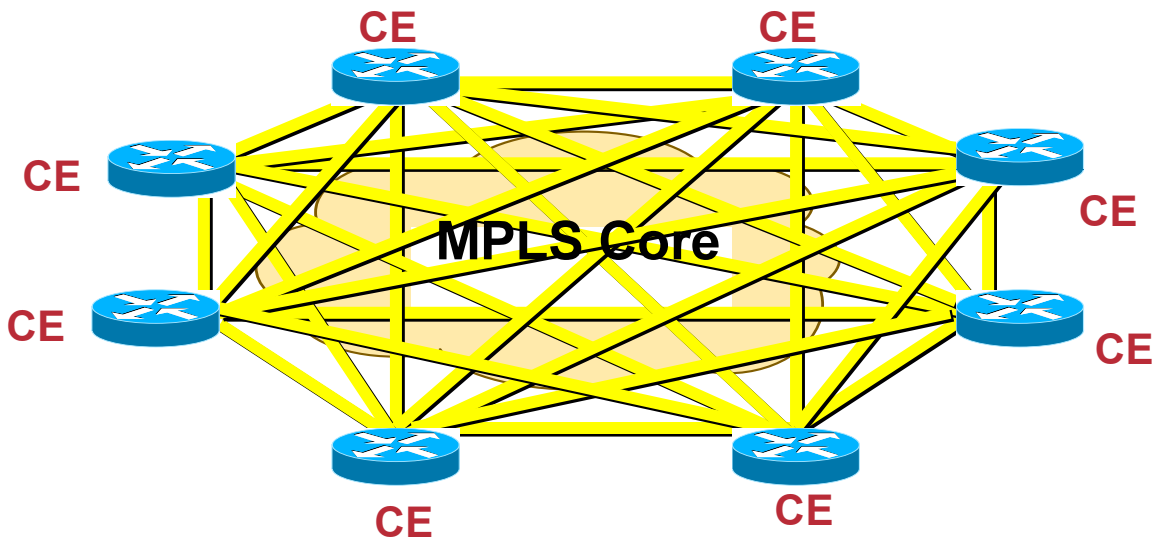
Multicast VPN (MVPN)

Deploying MPLS Based L3 VPNs and…



Multicast VPN – Challenges

- Multicast not supported with MPLS
- 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)

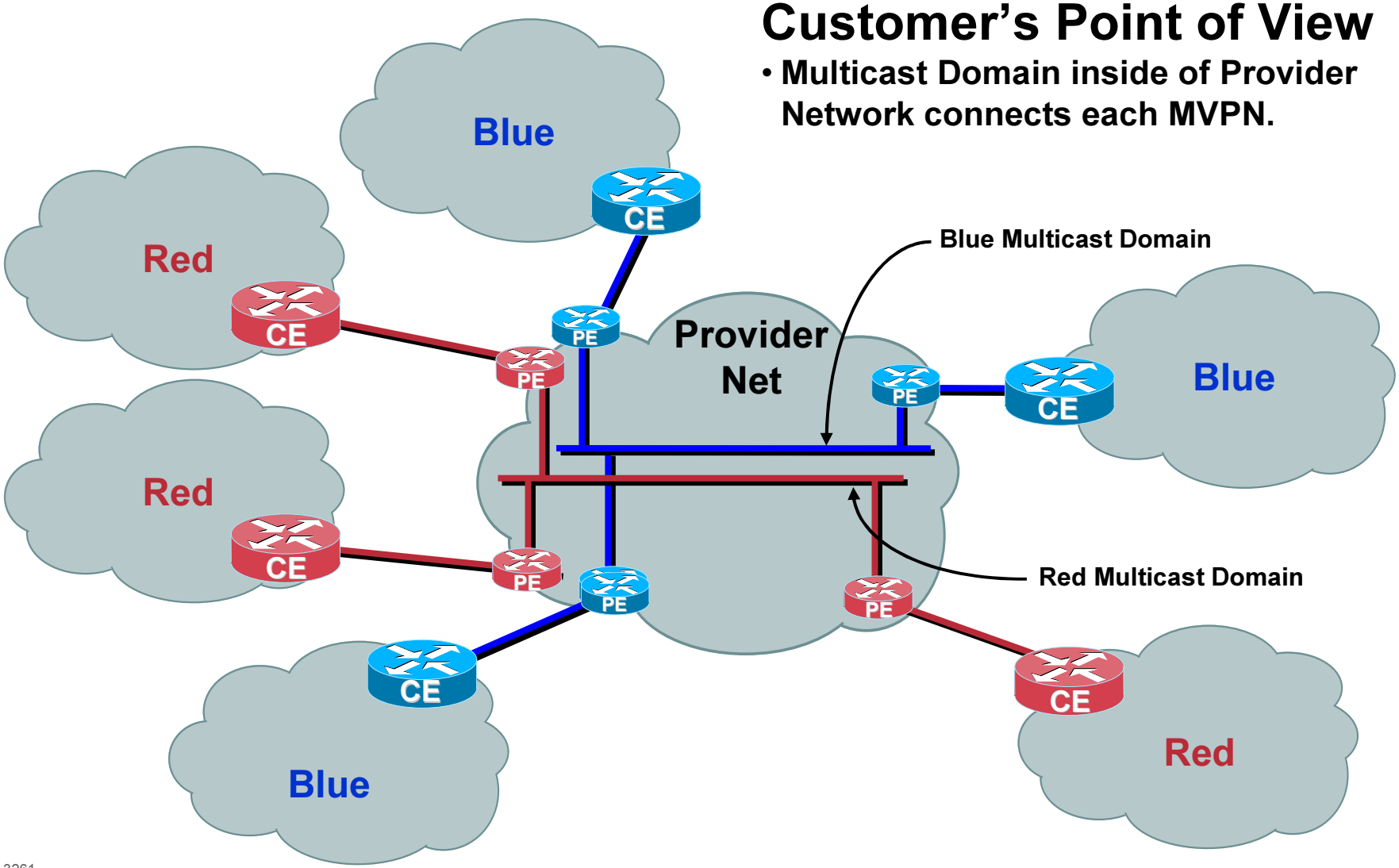
Cisco's Implementation

- **Based on Multicast Domains in draft-rosen-vpn-mcast-08.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. 7600 in 12.2S.**

Multicast VPN – Overview

Customer's Point of View

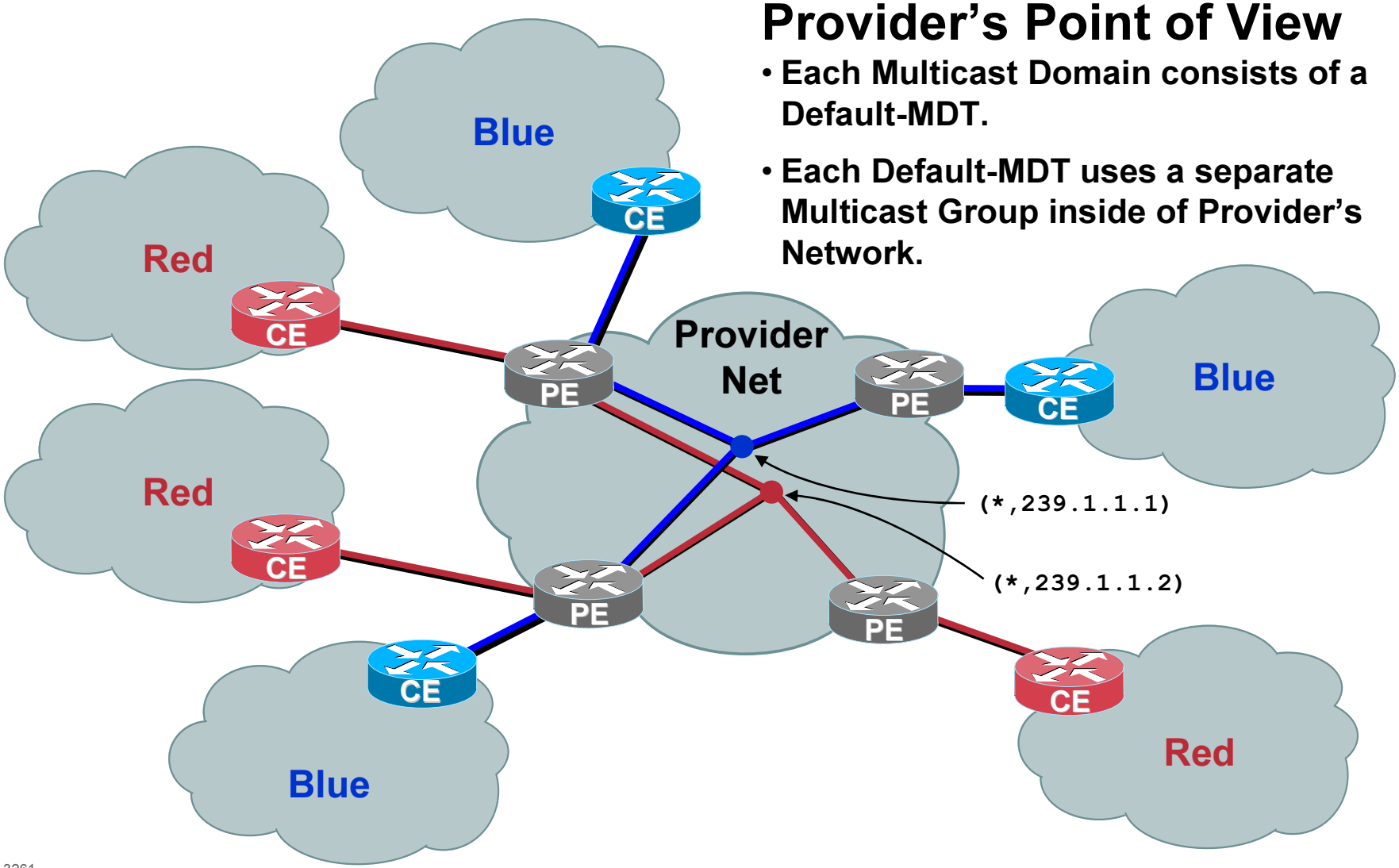
- Multicast Domain inside of Provider Network connects each MVPN.



Multicast VPN – Overview

Provider's Point of View

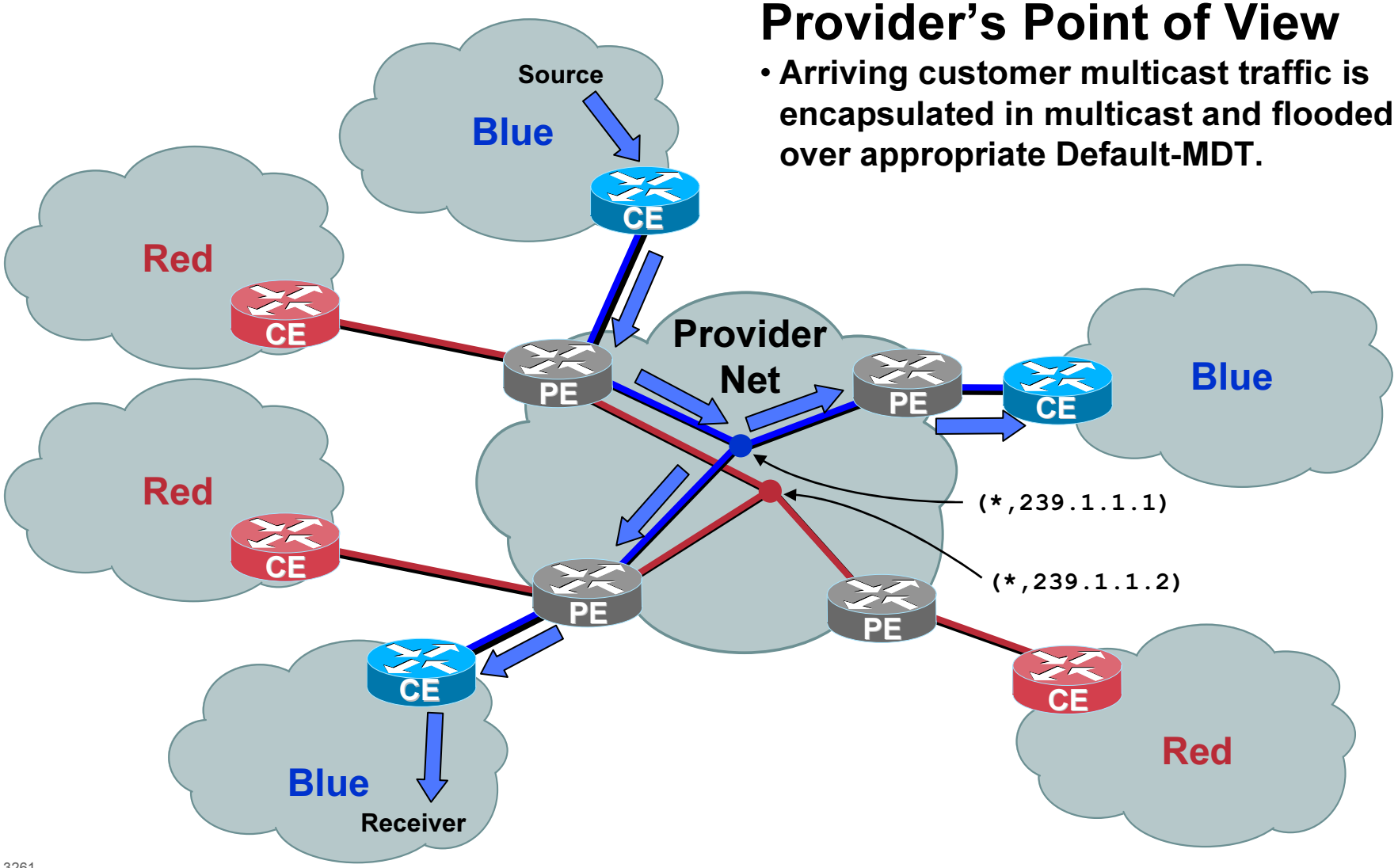
- Each Multicast Domain consists of a Default-MDT.
- Each Default-MDT uses a separate Multicast Group inside of Provider's Network.



Multicast VPN – Overview

Provider's Point of View

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



Multicast Distribution Tree (MDT)

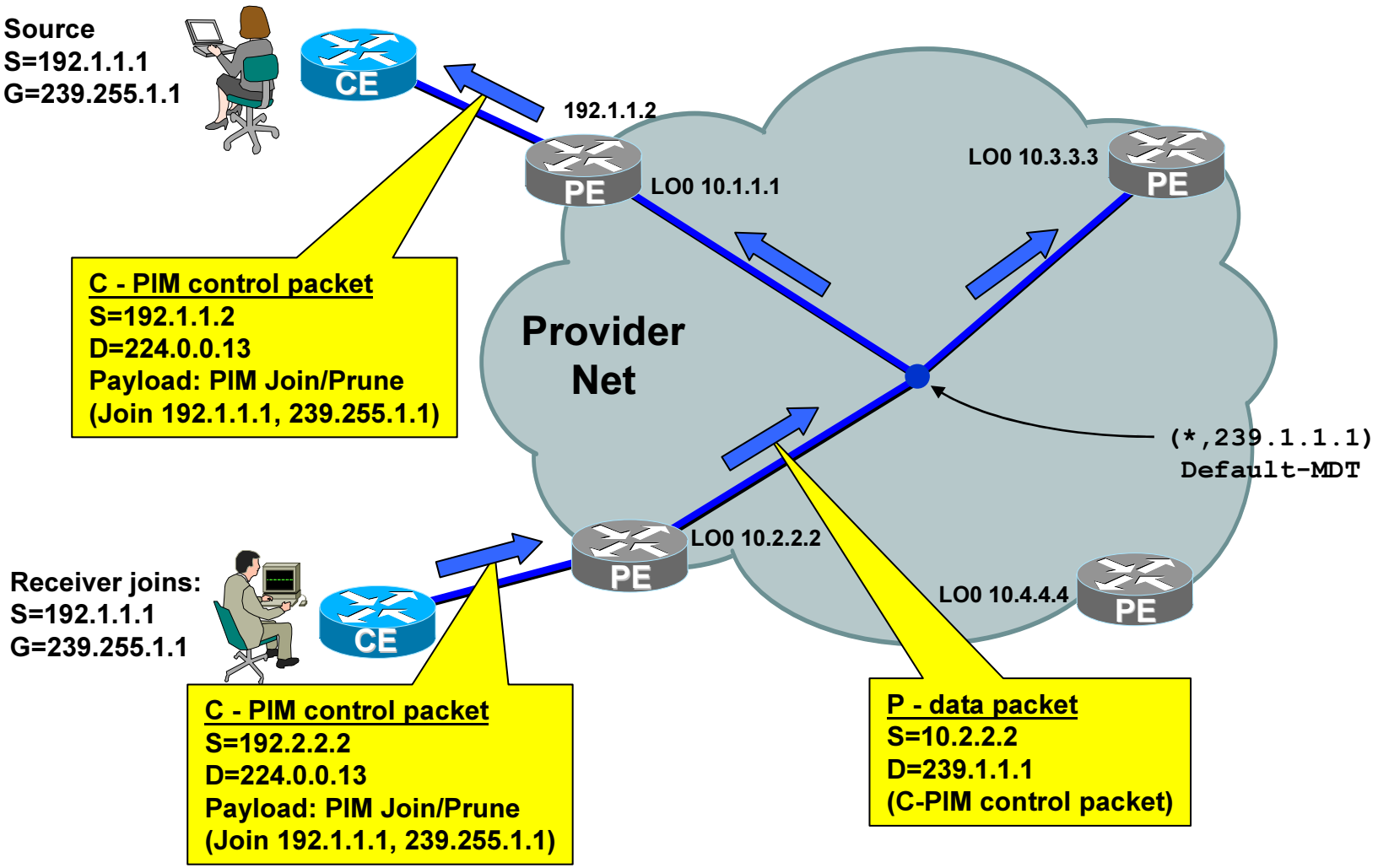
- MDTs built in service provider network
- MDTs built for each MDT group
- The number of MDTs depends on PIM modes of MDT groups
  - MDT group ranges administered by service provider

Two Types Of MDT Groups

- **Default MDT Groups**
  - **Configured for every MVRF if MPLS or IP core network present**
  - **Used for PIM control traffic, low bandwidth sources, and flooding of Dense-mode traffic**
- **Data MDT Groups**
  - **Optionally configured**
  - **Used for high bandwidth sources to reduce replication to uninterested PEs**

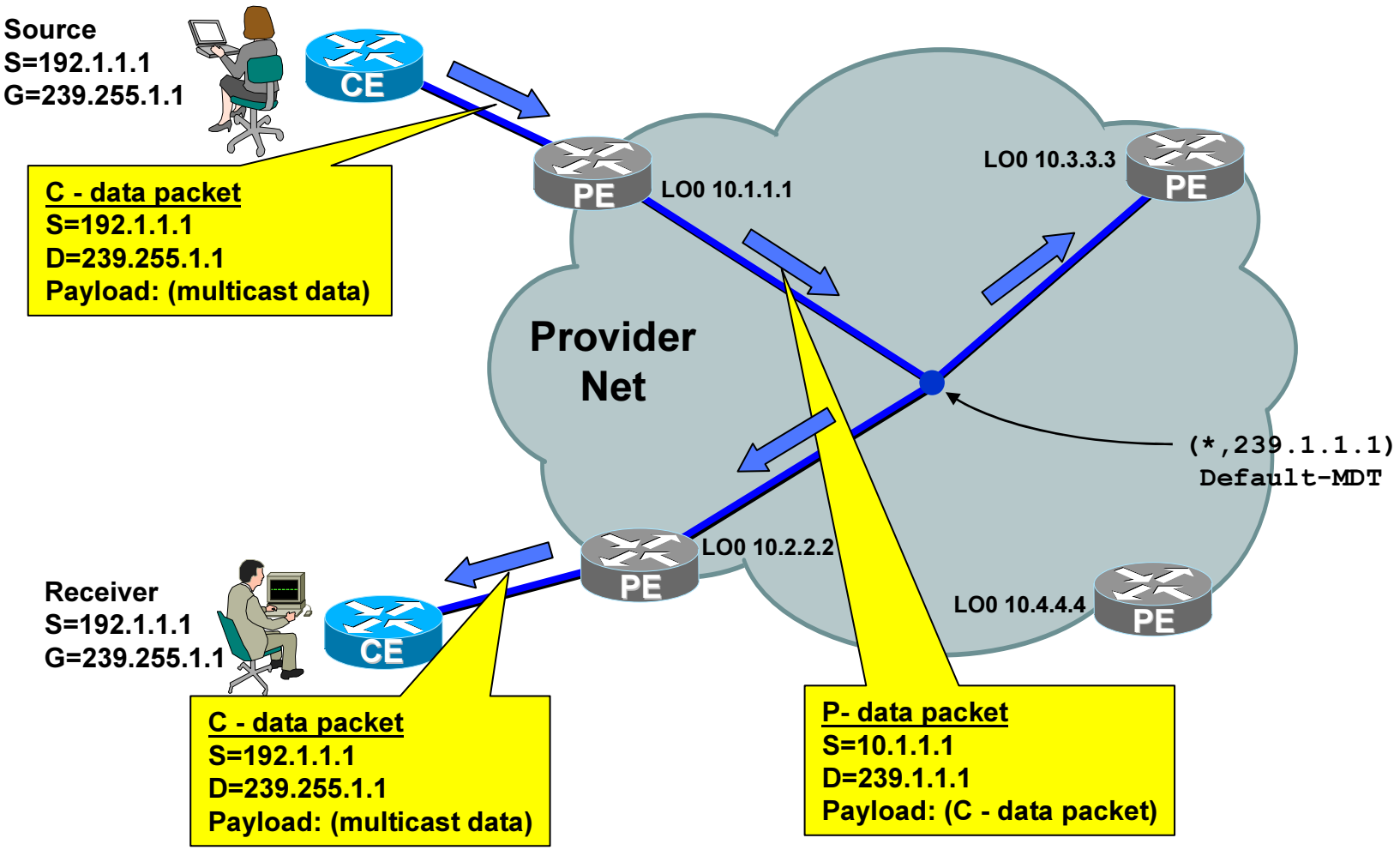
Default MDT – A Closer Look

PIM Control Traffic Flow



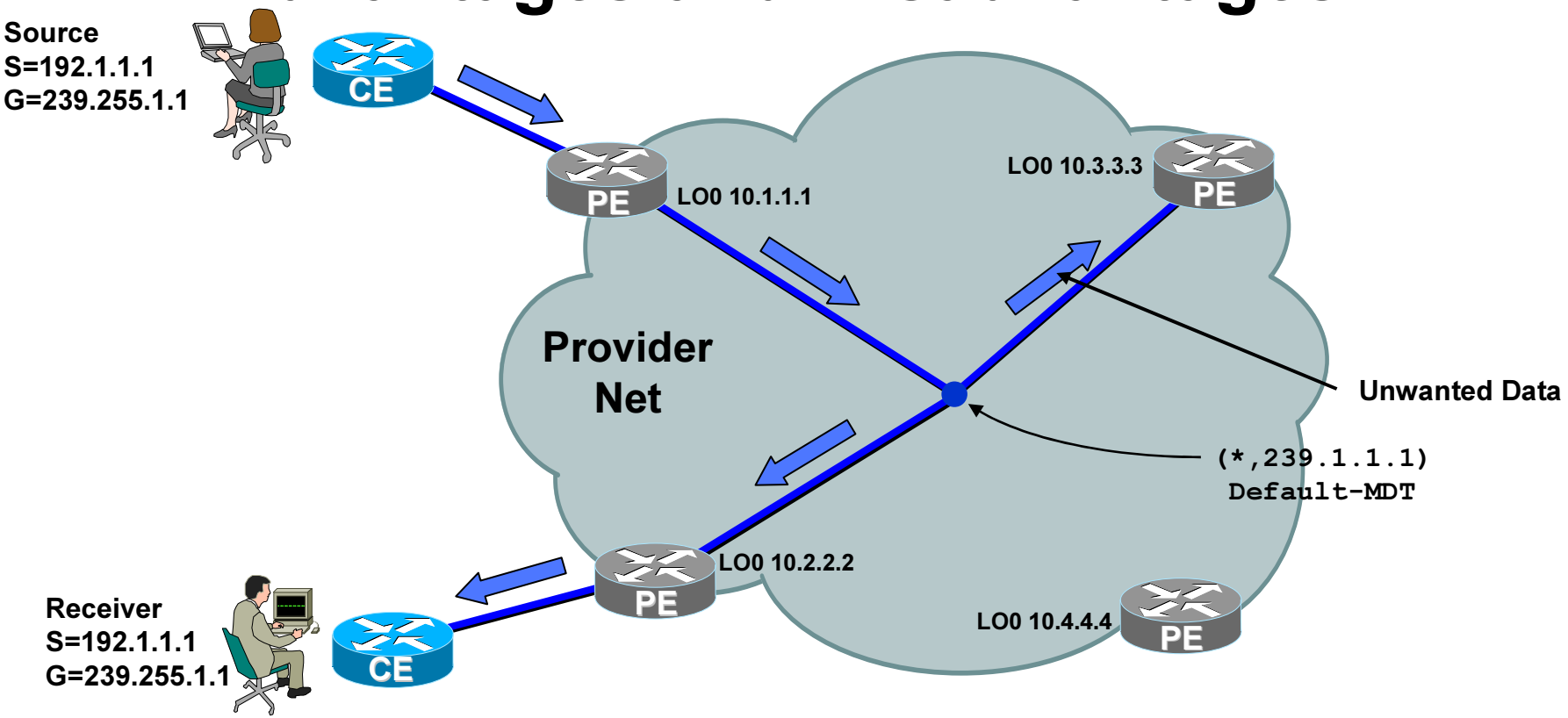
Default MDT – A Closer Look

Multicast Data Traffic Flow



Default MDT – A Closer Look

Advantages and Disadvantages

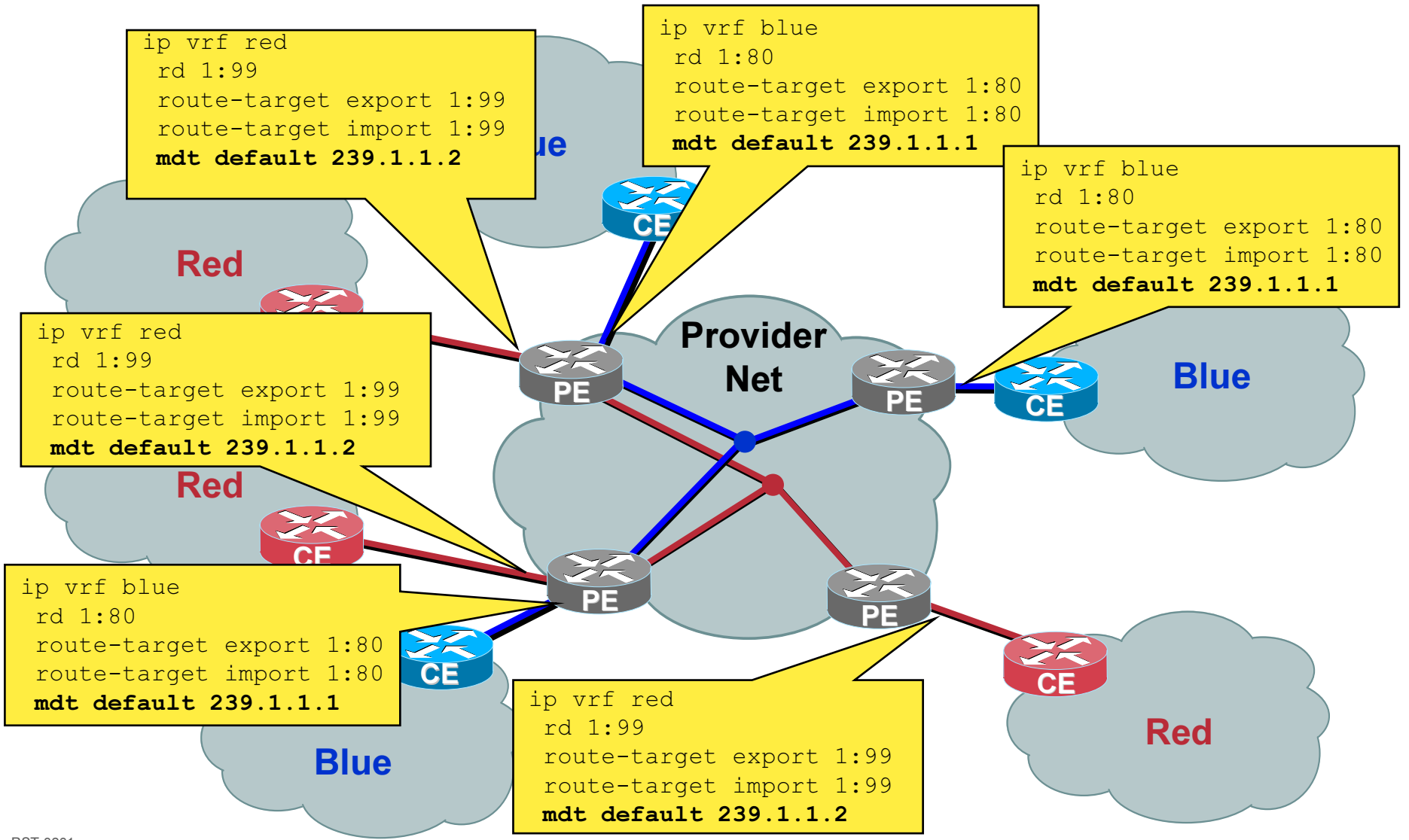


Advantage : Reduces multicast state in the P routers in the core.

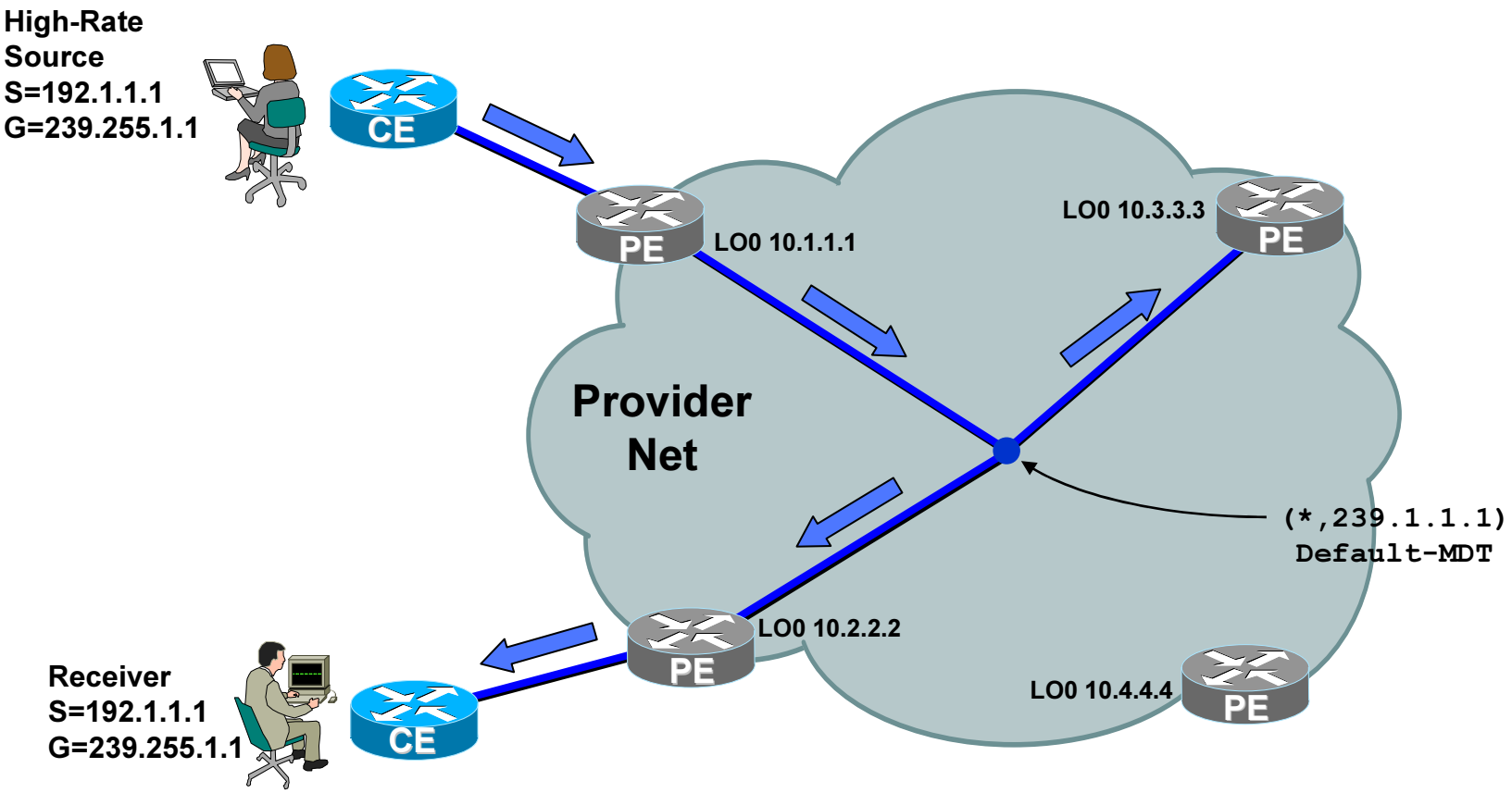
Disadvantage : Can result in wasted bandwidth.

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

Default-MDT Group Address Example

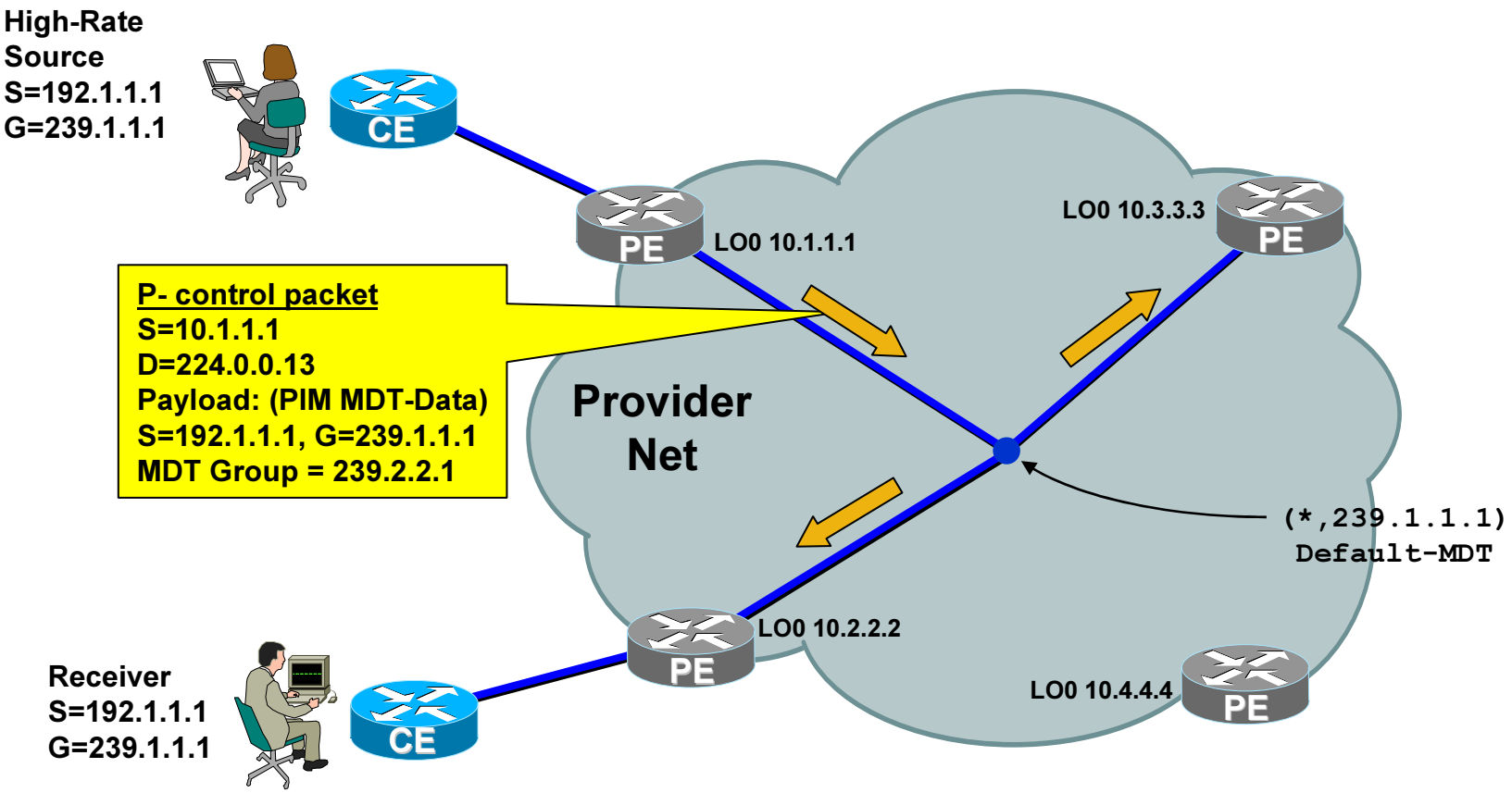


Data MDTs – Concepts



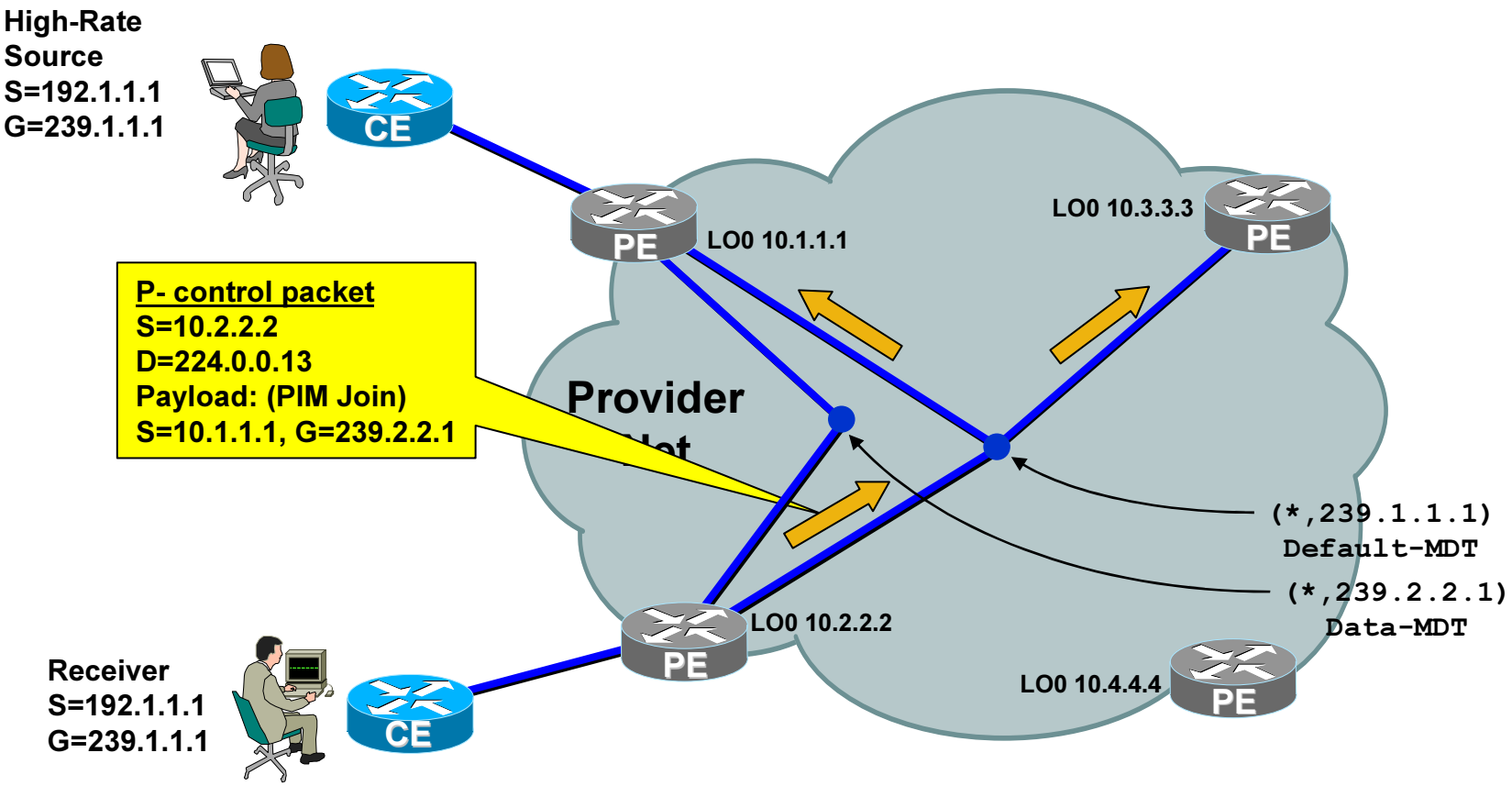
- **Traffic exceeds Data-MDT threshold configured on PE router.**

Data MDTs – Concepts



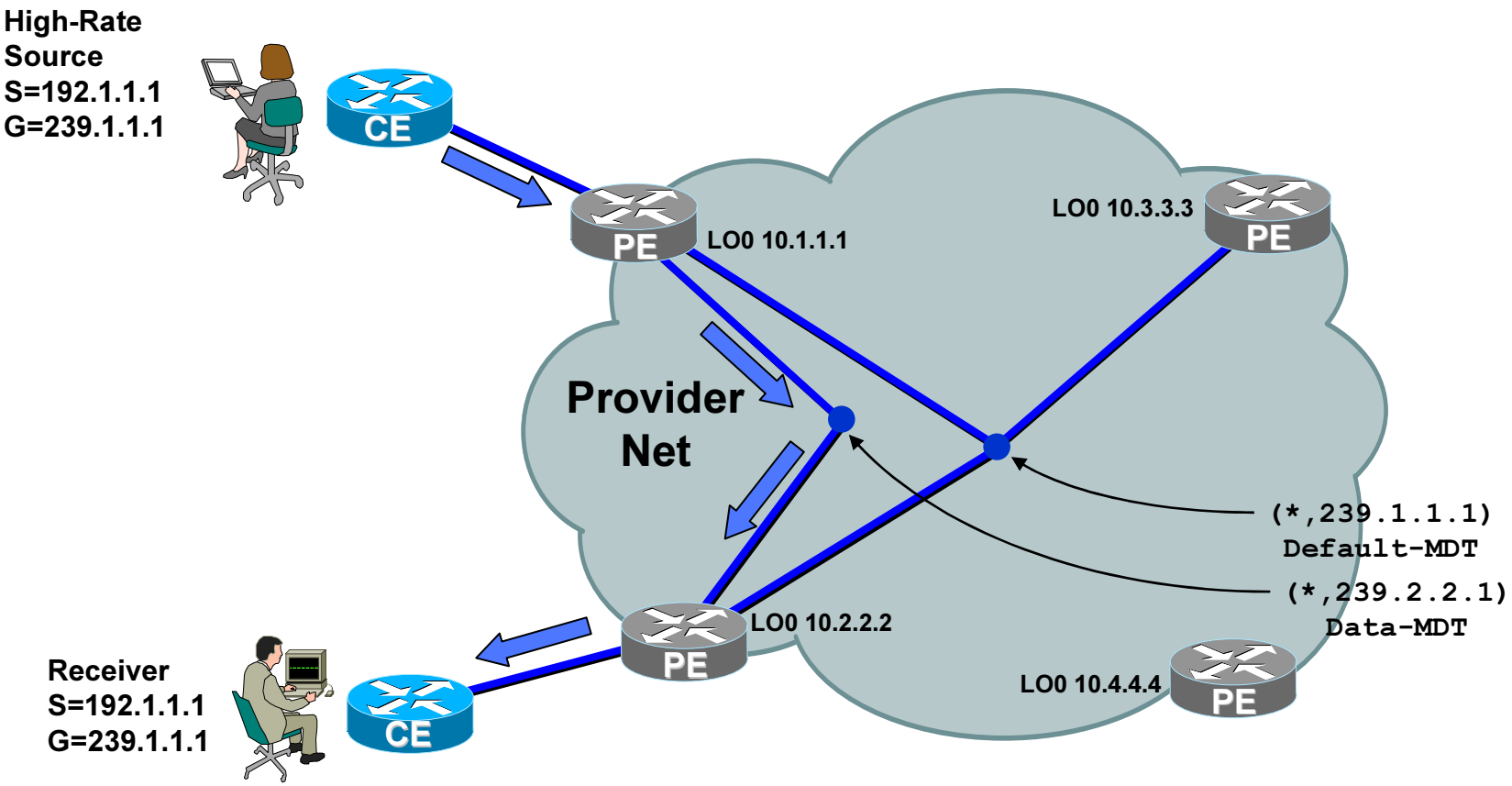
- PE router signals switch to Data-MDT using new group, 239.2.2.1

Data MDTs – Concepts



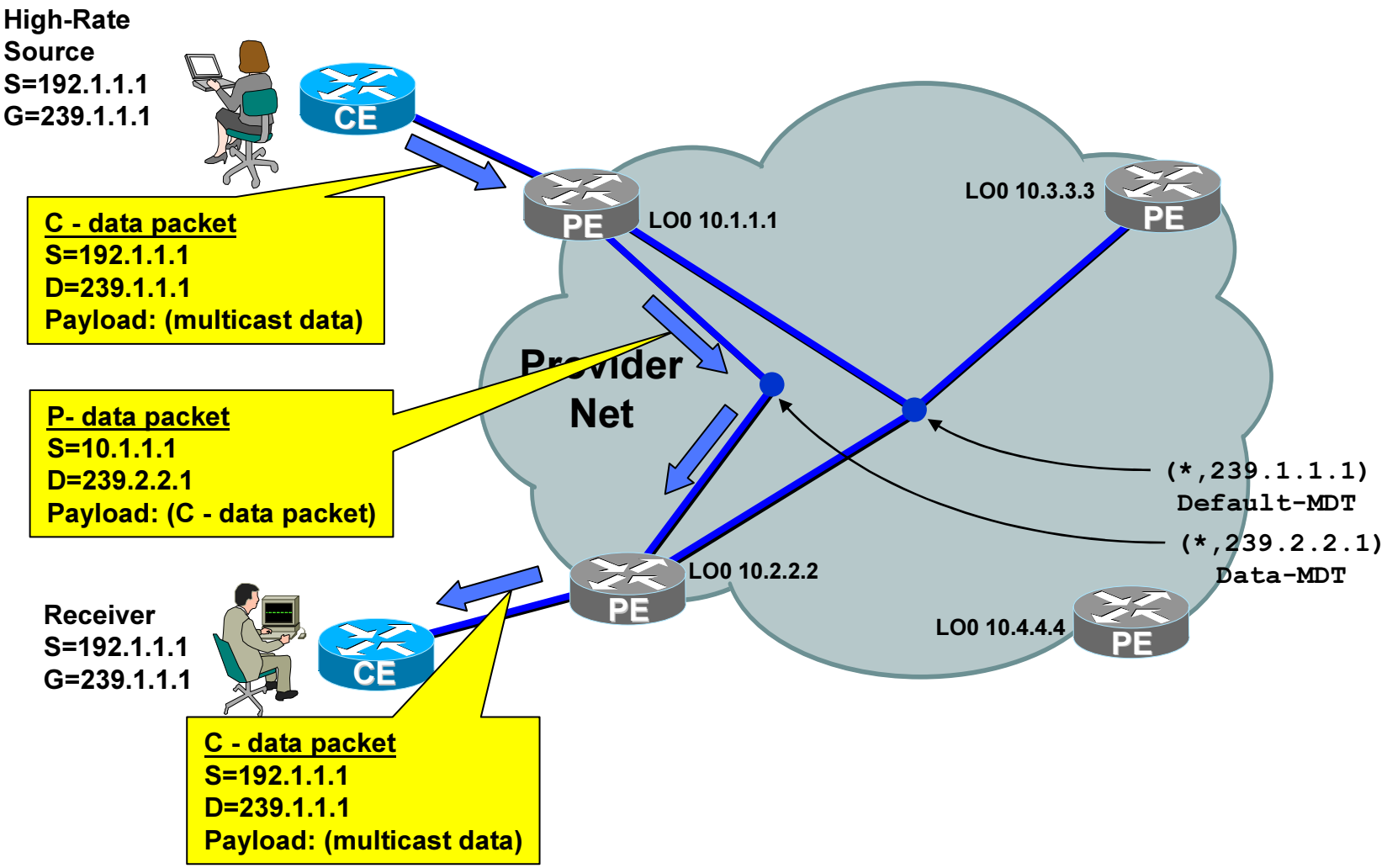
- PE routers with receivers sends Join to group 239.2.2.1.
- Data-MDT is built using group 239.2.2.1.

Data MDTs – Concepts

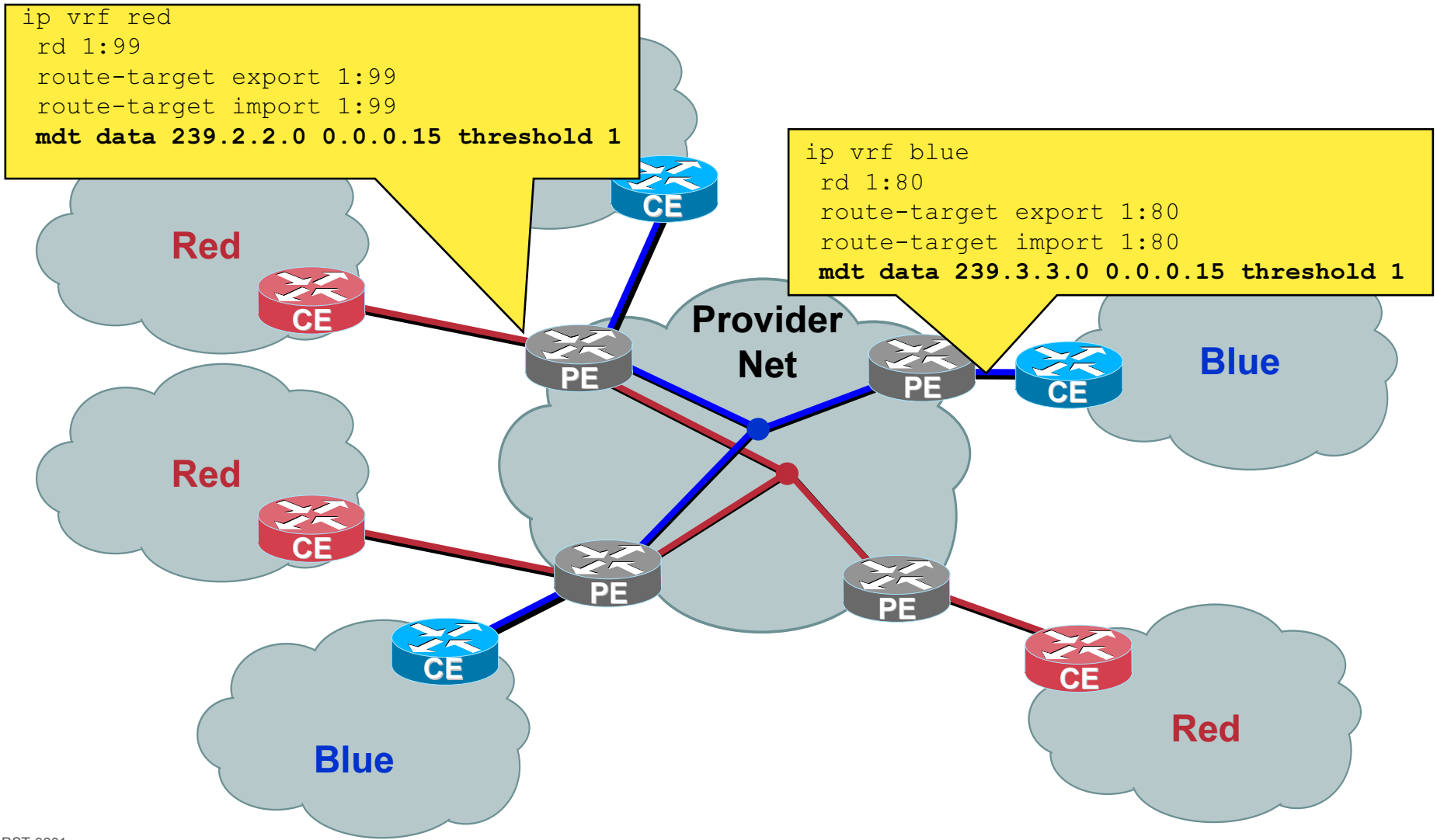


- High-rate data begins flowing via Data-MDT.
- Data only goes to PE routers that have receivers.

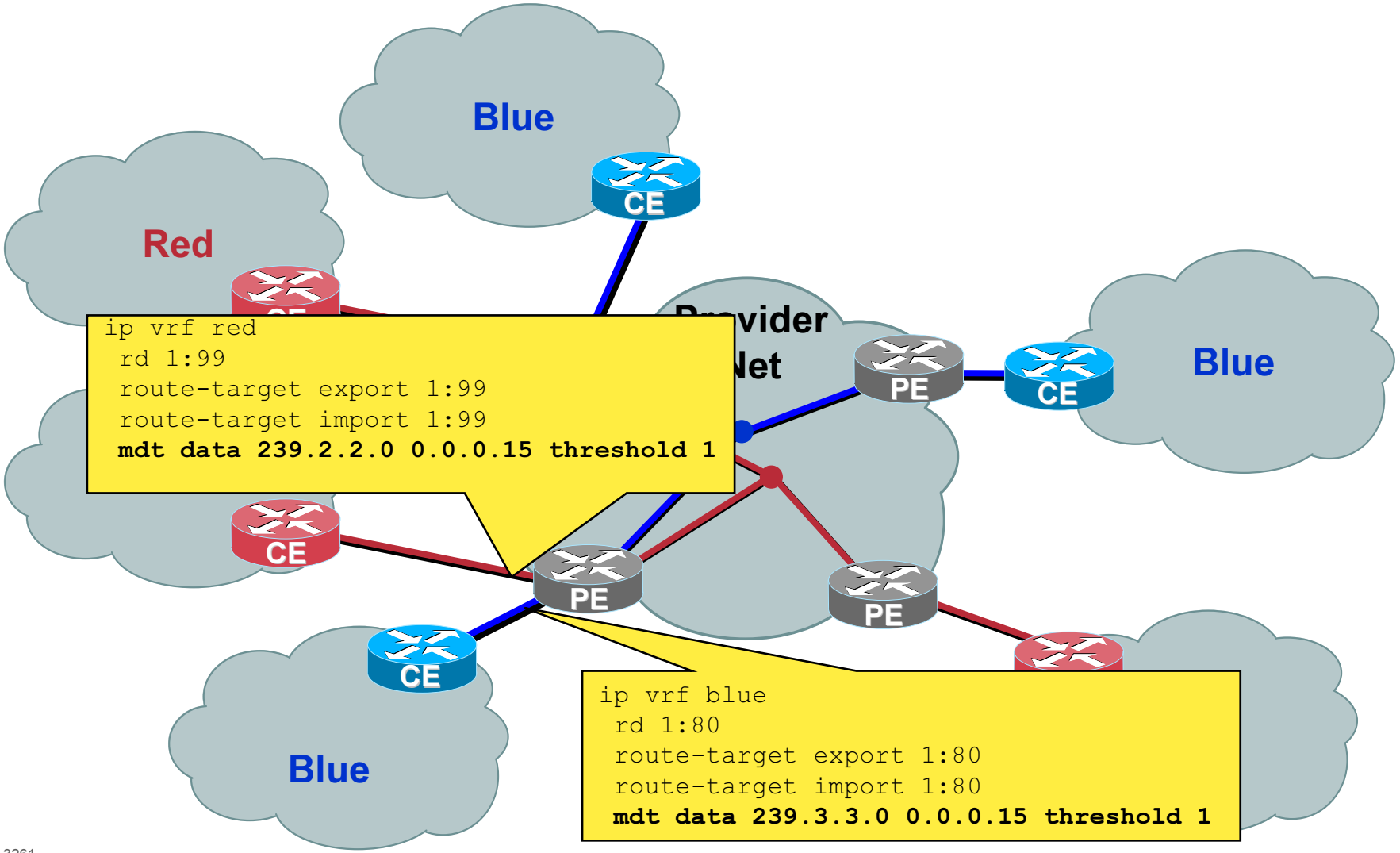
Data MDTs – Concepts



Data-MDT Group Address Example



Data-MDT Group Address Example



SSM for MDT

- Permits PE to directly join source tree rooted at another PE for MDT
- Cisco's recommended PIM mode for ALL MDTs
- No Rendezvous Point needed in service provider network
  - Reduce forwarding delay
  - Avoid management overhead to administer group/RP mapping and redundant RPs for reliability
  - Eliminate potential point of failure

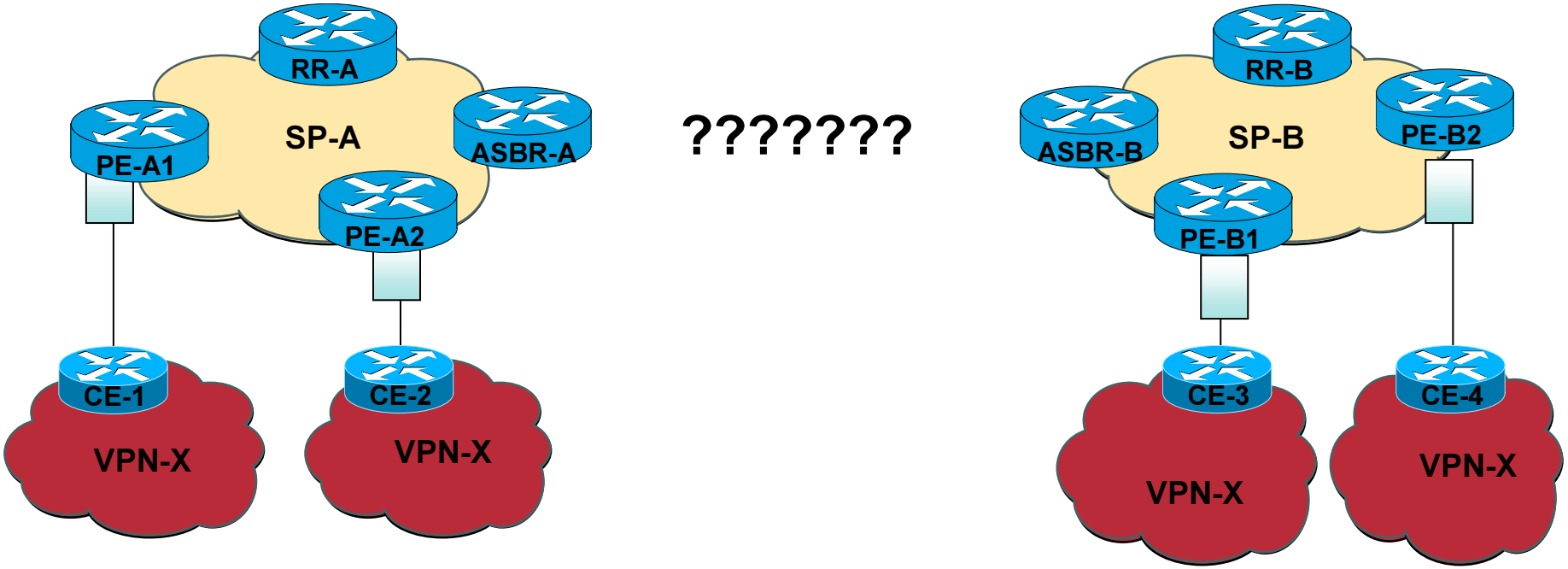
SSM for MDT - Default Group Source ID

- SSM requires PE to join an (S, G) not (*, G)
  - G already known - configured as MDT Default Group
  - PE does not directly know S, (BGP loopback identities of other PEs in same MD)
- Use MBGP to distribute the information
  - pre 12.0(29)S IOS use extended community attributes
  - Newer IOS use a new BGP address family
  - SAFI capability negotiated by BGP peers

Inter-AS MPLS/VPN - Requirement

Provide connectivity for all VPN-X sites

May be for separate providers or one provider operating National and International Backbones as separate AS's



Inter-AS MPLS/VPN Options - rfc2547bis

Three options for unicast listed in draft-ietf-l3vpn-rfc2547bis

1. **Back-to-back ASBR-PEs**
2. **ASBRs exchanging VPNv4 routes**
3. **VPNv4 routes via multi-hop MP-eBGP**

All three options are in deployment and must be supported for multicast VPN

Inter-AS MVPN - Solution components

- **BGP Connector attribute**
  - **draft-nalawade-idr-mdt-safi-03**
  - **Preserves identity of PE router originating VPNv4 prefix**
- **MBGP MDT SAFI**
  - **draft-nalawade-idr-mdt-safi-03**
  - **Help ASBR RPF to PE in remote AS**
  - **Help ASBR and receiver PE insert RPF Vector to build MDT to PE in remote AS**
- **RPF Vector**
  - **draft-ietf-pim-rpf-vector-02**
  - **Allow P routers to build MDT to PE in remote AS**

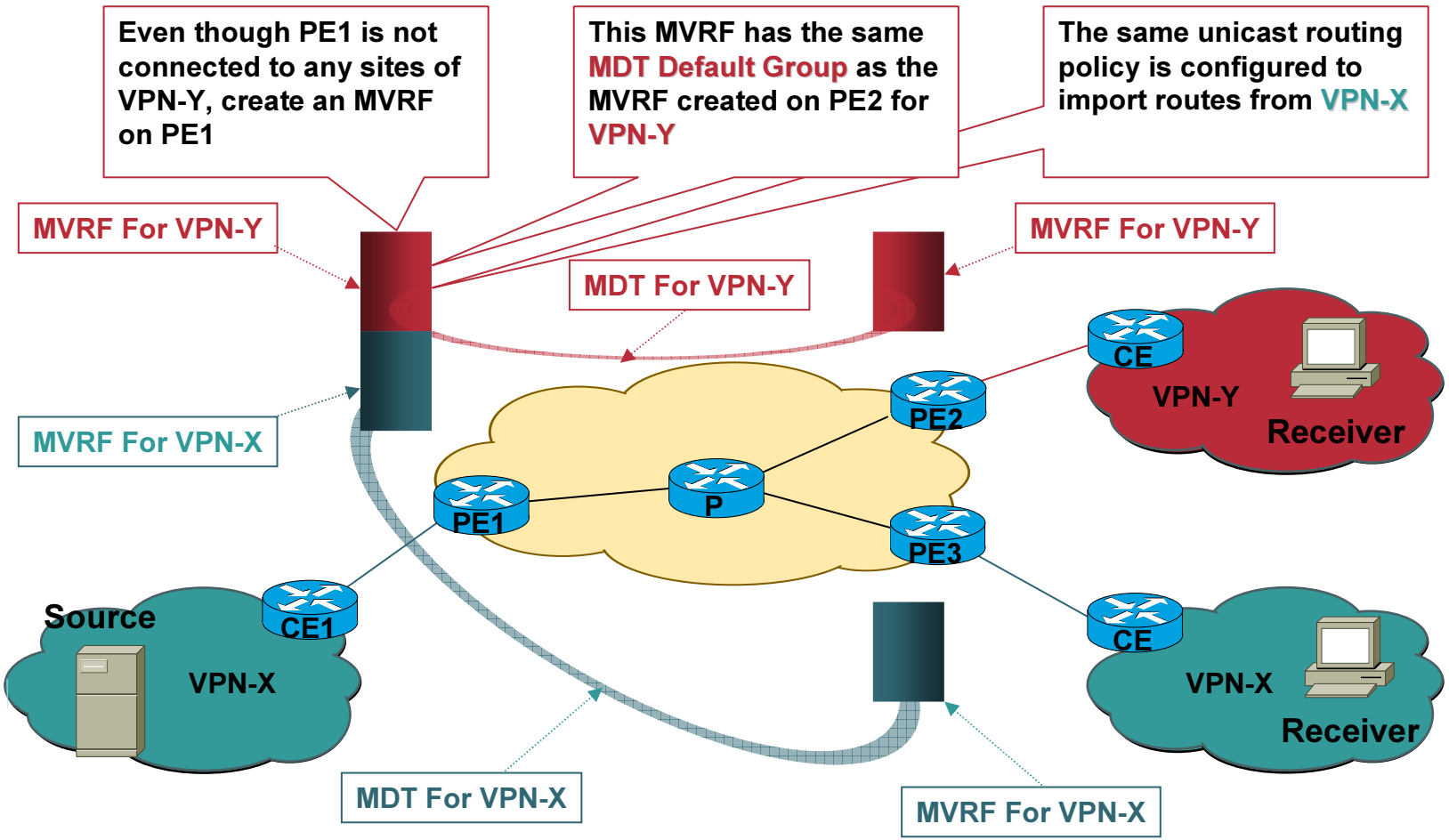
Extranet MVPN

- Allow multicast content originated from within one site to be distributed to other sites, possibly belonging to different VPNs
  - SP content provisioned to multiple vpns
- Require no new protocols
- Depend only on unicast routing policies to perform RPF
  - In case multicast and unicast topologies are not congruent, additional configuration is necessary

Extranet MVPN Configuration Options

- **Configuration Option 1:**
  - **On PE router connected to the source:**
    - **For each MVPN that wishes to receive the content**
    - **Configure an additional MVRF which has the same Default MDT Group (if the MVRF is not present).**
- **Configuration Option 2:**
  - **On PE router(s) connected to the receivers:**
    - **Configure an additional MVRF which has the same Default MDT Group as the one connected to the multicast source (if the MVRF is not present).**

Extranet MVPN Configuration Option #1



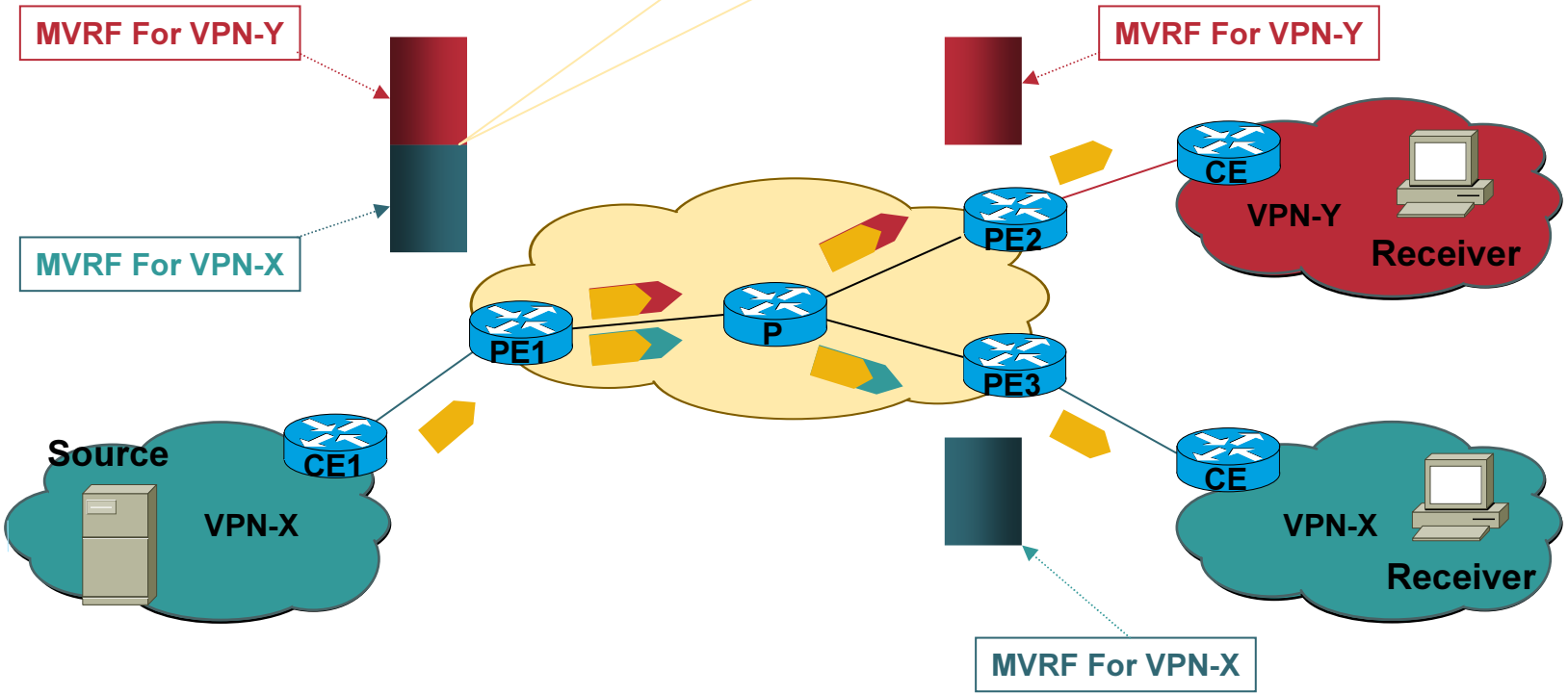
Extranet MVPN Configuration Option #1

Packet Flow

Packets received in MVRF for VPN-X from the source

Independently replicated and encapsulated in the MVRF for VPN-X and VPN-Y

PE2 and PE3 decapsulate and forward the packet to the respective MVRFs

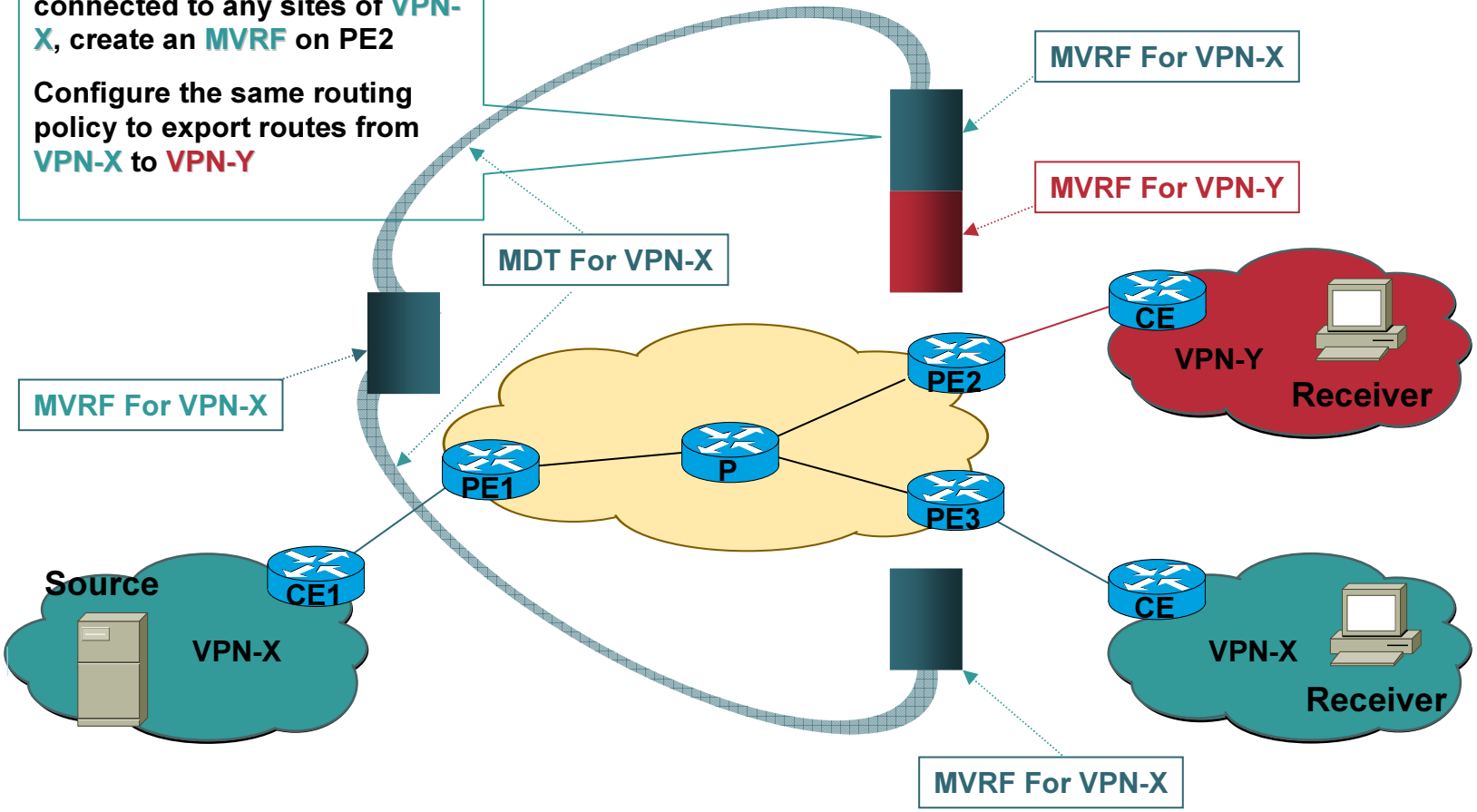


Extranet MVPN Configuration Option #2

Configuration

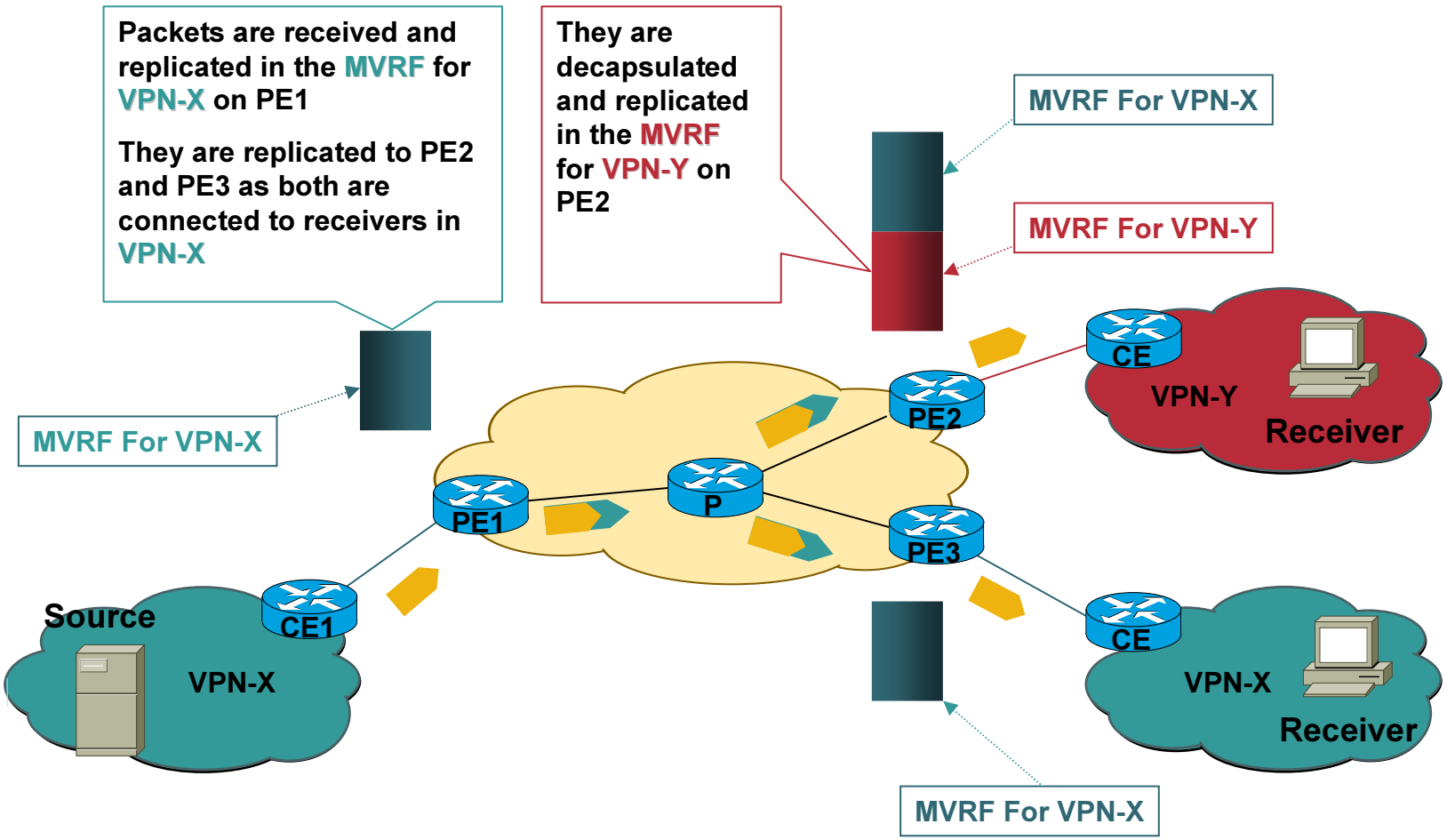
Even though PE2 is not connected to any sites of VPN-X, create an MVRF on PE2

Configure the same routing policy to export routes from VPN-X to VPN-Y



Extranet MVPN Configuration Option #2

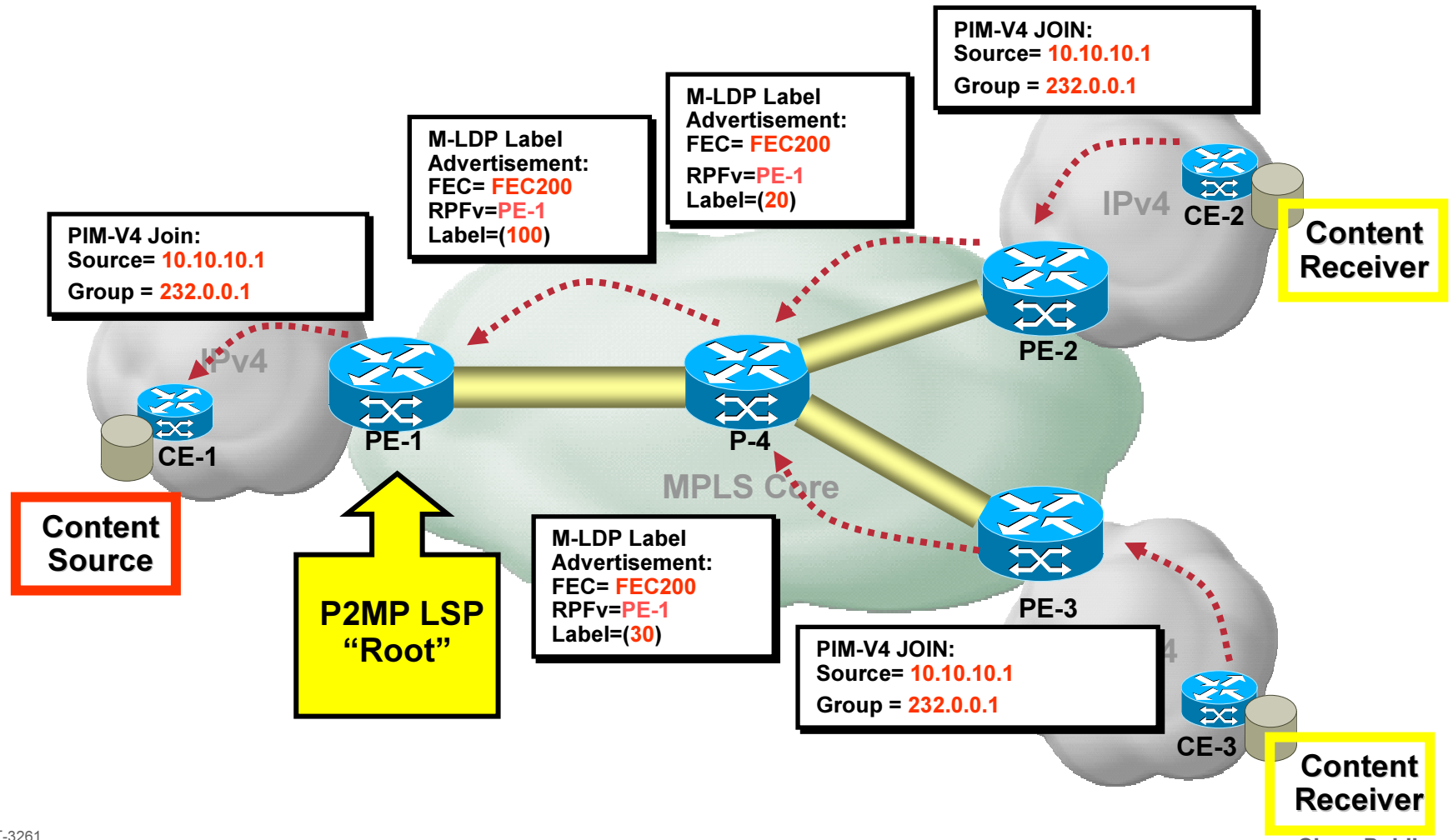
Packet Flow



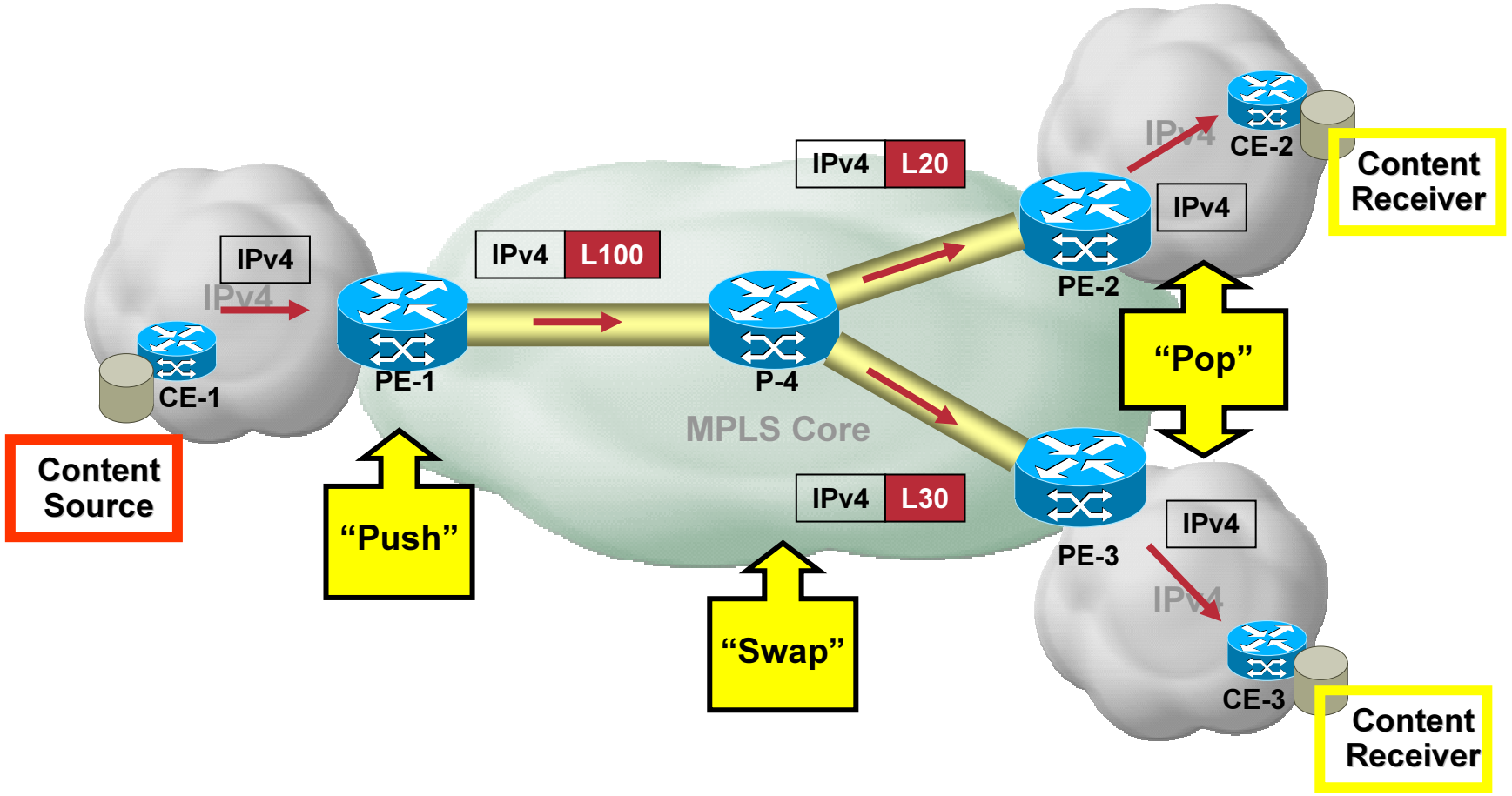
Cisco MVPN Strategy

- Customers require multiple forwarding options for transit services.
- Build upon successful MVPN model.
- Scalable modular architecture for multicast transport services
  - MVPN GRE is first deployable option
  - MVPN LSM is a new option
    - mLDP
    - P2MP RSVP-TE
  - Same operations model for IP or MPLS for ease of transition between options. May use multiple options in parallel (depending on service)
  - Focus on (necessary) migration options

MLDP : Transiting SSM (IPv4 non-VPN)



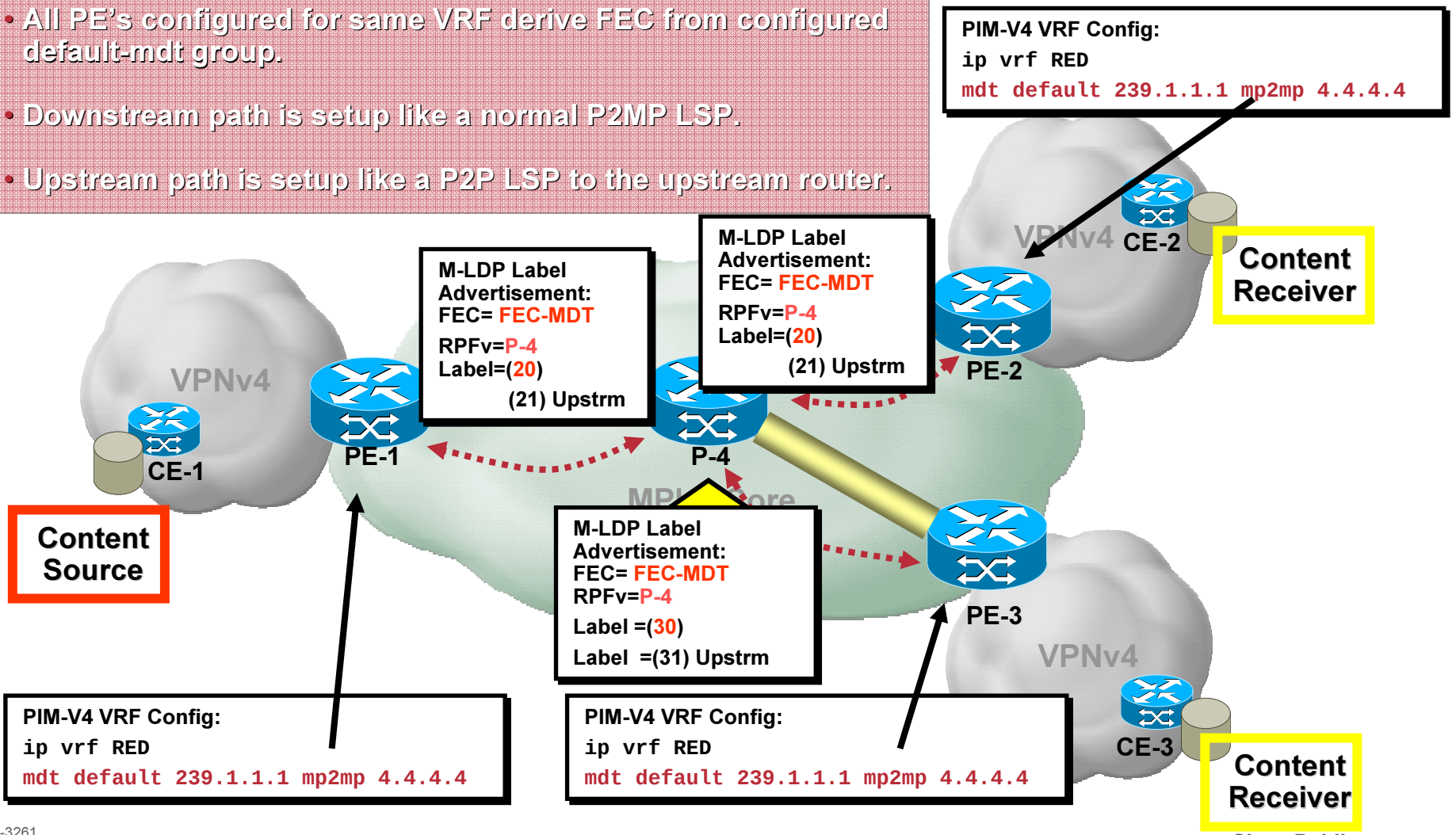
mLDP : Transiting SSM (IPv4 non-VPN)



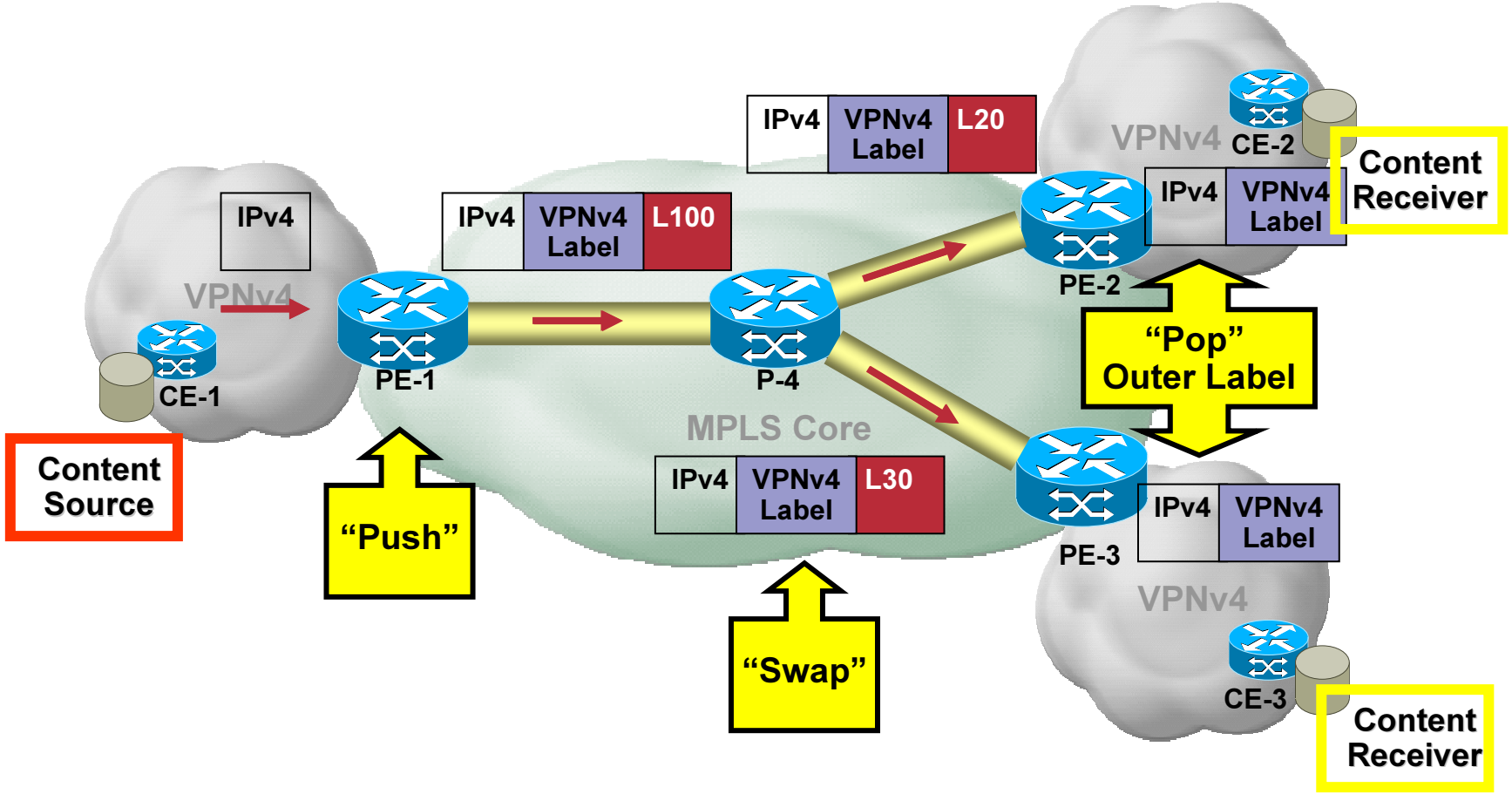
Multicast LDP based Multicast VPN *(Default-MDT)*

MP2MP Tree Setup Summary

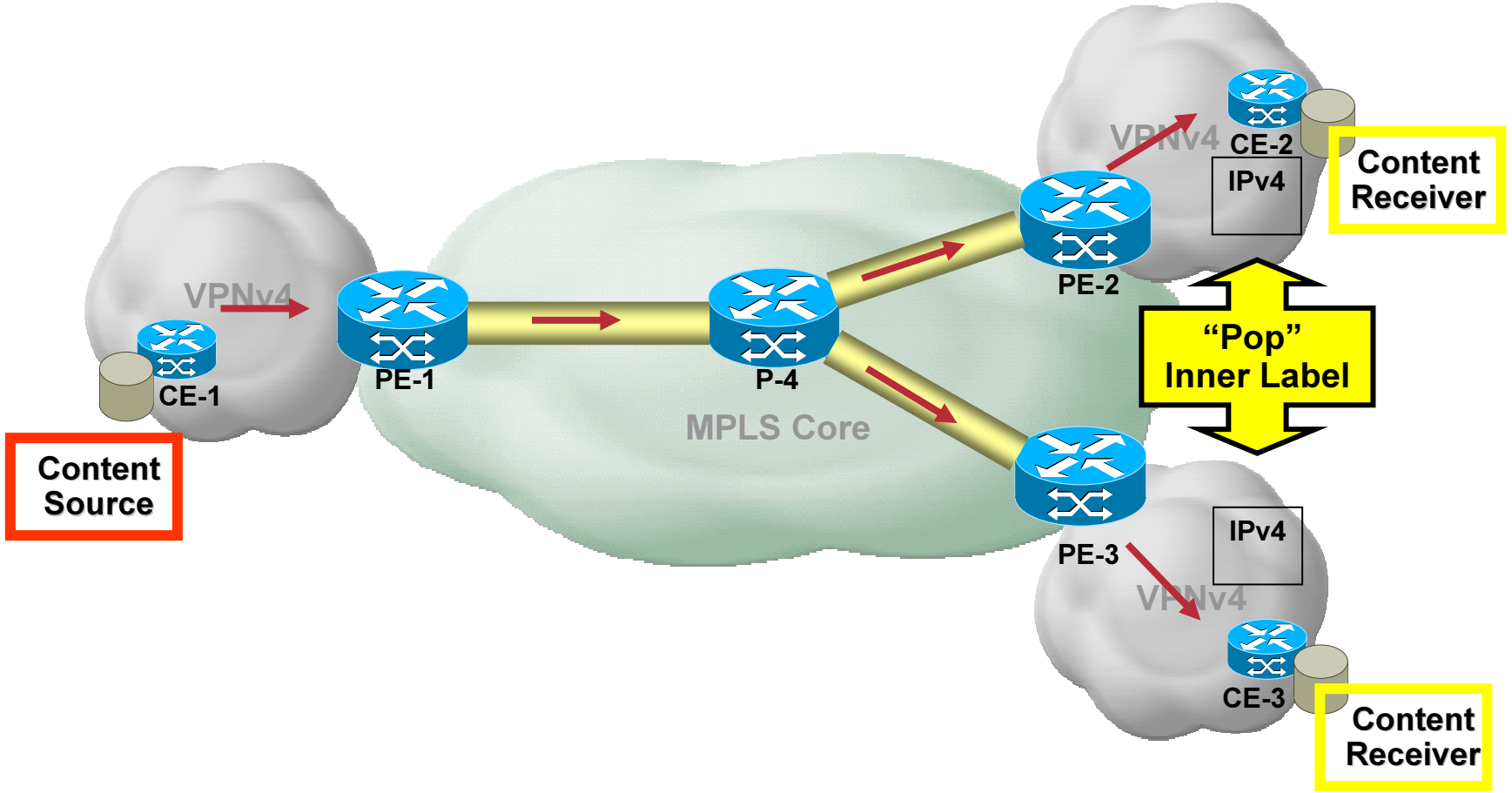
- All PE's configured for same VRF derive FEC from configured default-mdt group.
- Downstream path is setup like a normal P2MP LSP.
- Upstream path is setup like a P2P LSP to the upstream router.



Multicast LDP based Multicast VPN *(Default-MDT)*



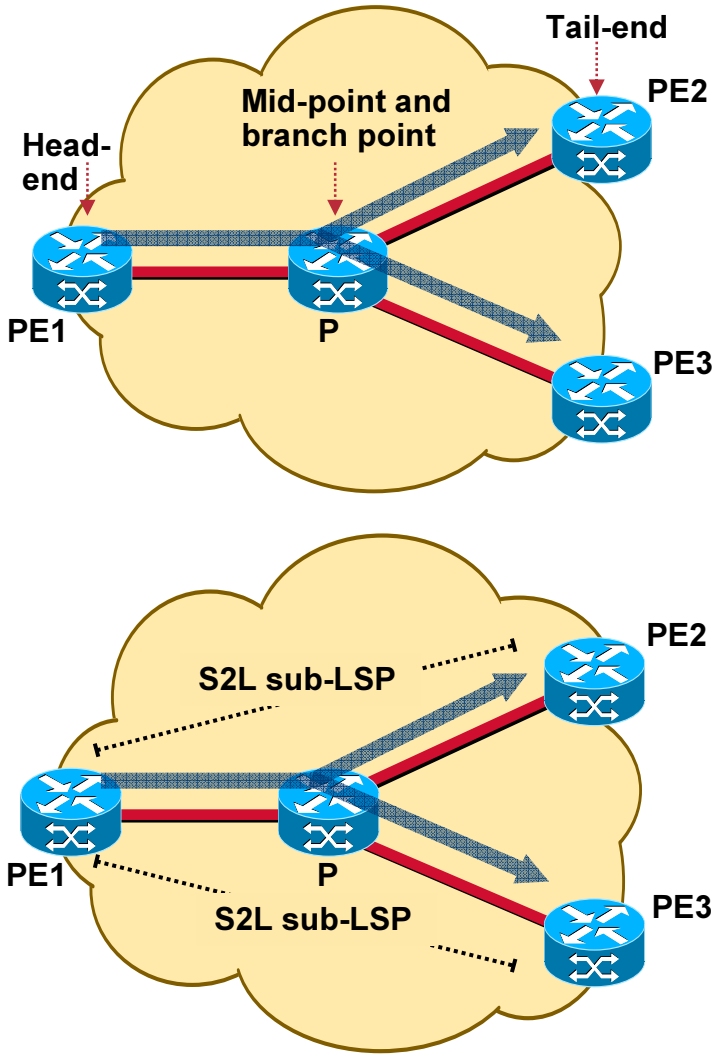
Multicast LDP based Multicast VPN *(Default-MDT)*



P2MP RSVP TE

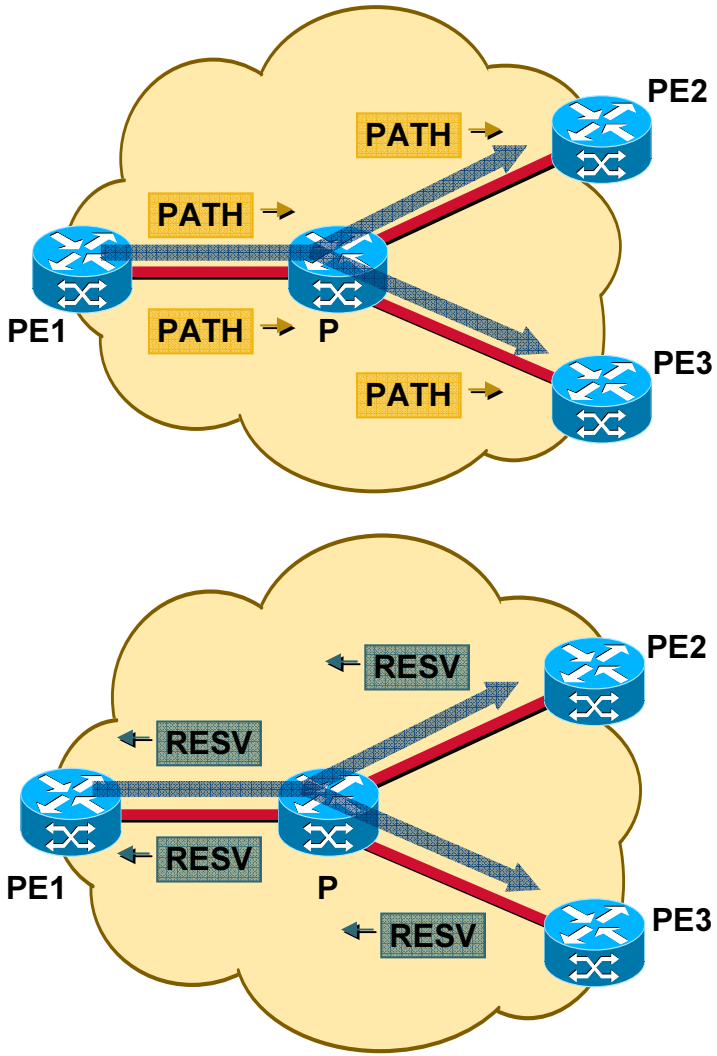
- Extend RSVP-TE to establish P2MP-LSPs
  - Focus on TE requirements for relatively static P2MP-LSP topologies
- IETF proposals are converging on RSVP-TE draft
- RSVP P2MP draft:
  - draft-ietf-mpls-rsvp-te-p2mp-05.txt
- Requirements drafts
  - draft-ietf-l3vpn-ppvpn-mcast-reqts-06.txt

Terminology



- **Head-end/source**: Node where LSP signaling is initiated
- **Mid-point**: Transit node where LSP signaling is processed (not a head-end, not a tail-end)
- **Tail-end/leaf/destination**: node where LSP signaling is terminated
- **Branch point**: node where packet replication is performed
- **Source-to-leaf (S2L) sub-LSP**: P2MP TE LSP segment that runs from source to one leaf

P2MP TE LSP Setup



- P2MP TE is defined as a collection of S2L sub-LSPs
- Each sub-LSP signaled independently
- Label replication state built during label distribution when two or more sub-LSPs diverge
- Sub-LSPs on same path receive the same label during label distribution

Cisco Status – LSM

LSM Protocols	Distinct properties
MLDP **draft-ietf-mpls-ldp-p2mp-00**	**Dynamic Tree Building suitable for broad set Multicast Applications** **FRR as optional capability** **Receiver driven dynamic tree building approach**
P2MP RSVP-TE **draft-ietf-mpls-rsvp-te-p2mp-04**	**Deterministic bandwidth guarantees over entire tree *(calculation overhead limits this to static tree scenarios)*** **Head end defined trees** **FRR inherent in tree set-up** **Useful for Small but significant subset of Multicast Application: Broadcast TV *where bandwidth restrictions exist.***

Summary

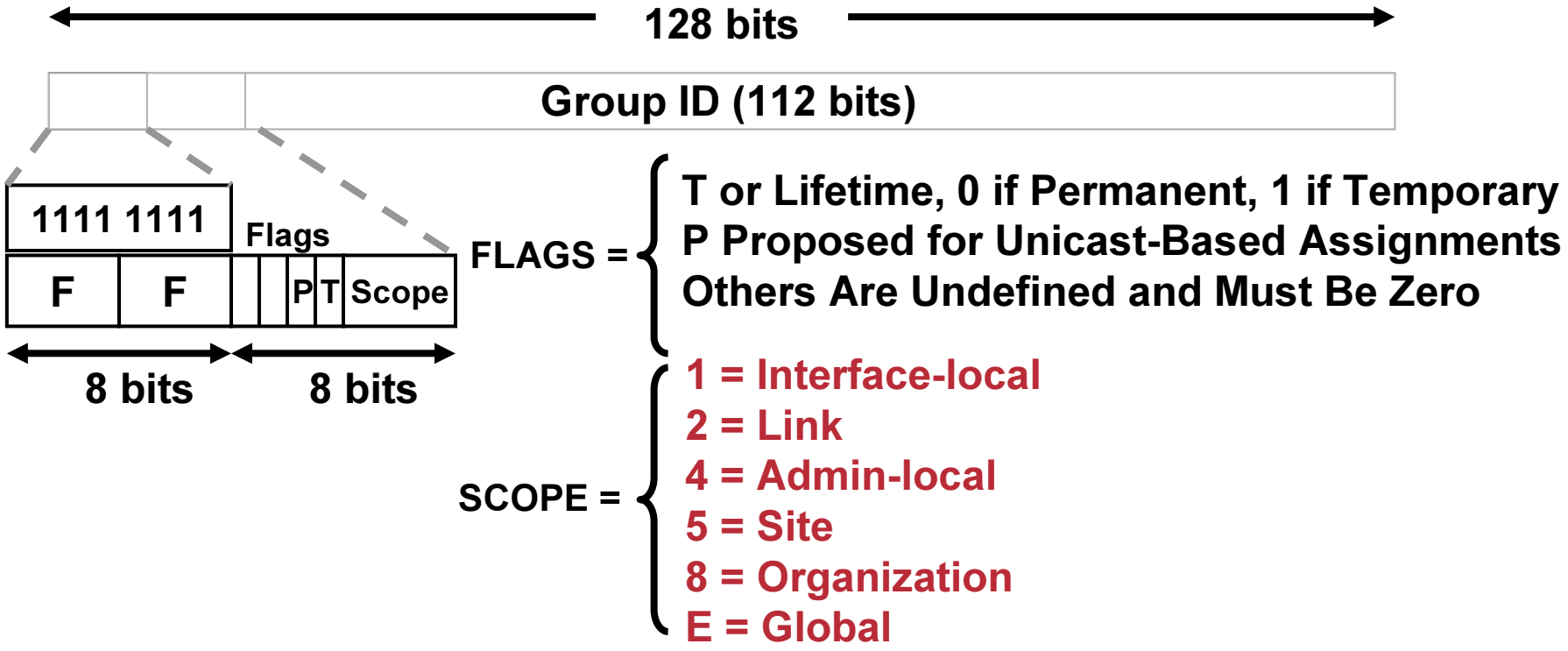
- **Native services**
  - **Continue to enhance v4 and v6 services**
- **Encapsulated services**
  - **mVPN GRE solutions**
    - **Continue to leverage leadership**
    - **Maintain feature richness**
  - **LSM solutions**
    - **Develop P2MP RSVP-TE**
    - **Develop MLDP**
    - **Incorporate into mVPN framework**

IPv6 Multicast

IPv4 and IPv6 Multicast Comparison

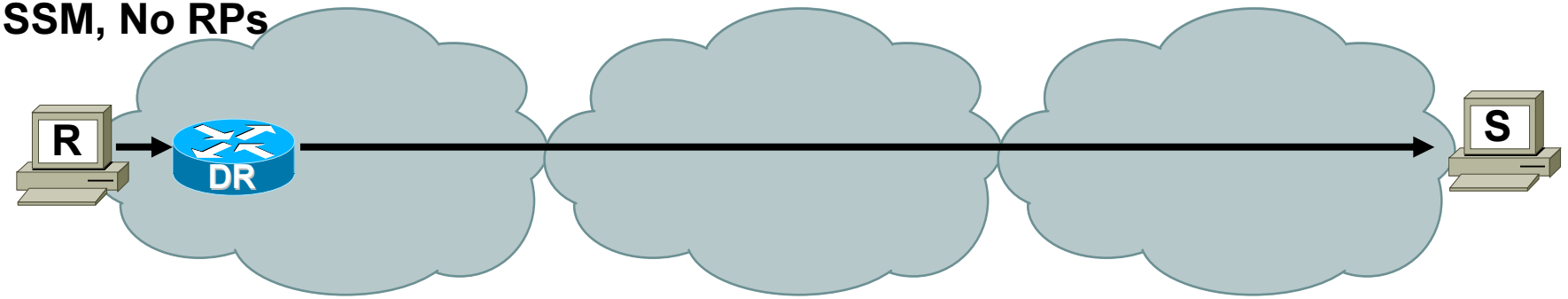
Service	IPv4 Solution	IPv6 Solution
Addressing Range	32-bit, Class D	128-bit (112-bit Group)
Routing	Protocol Independent, All IGPs and MBGP	Protocol Independent, All IGPs and MBGP with v6 mcast SAFI
Forwarding	PIM-DM, PIM-SM, PIM-SSM, PIM-bidir	PIM-SM, PIM-SSM, PIM-bidir
Group Management	IGMPv1, v2,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

IPv6 Multicast Addresses (RFC 3513)

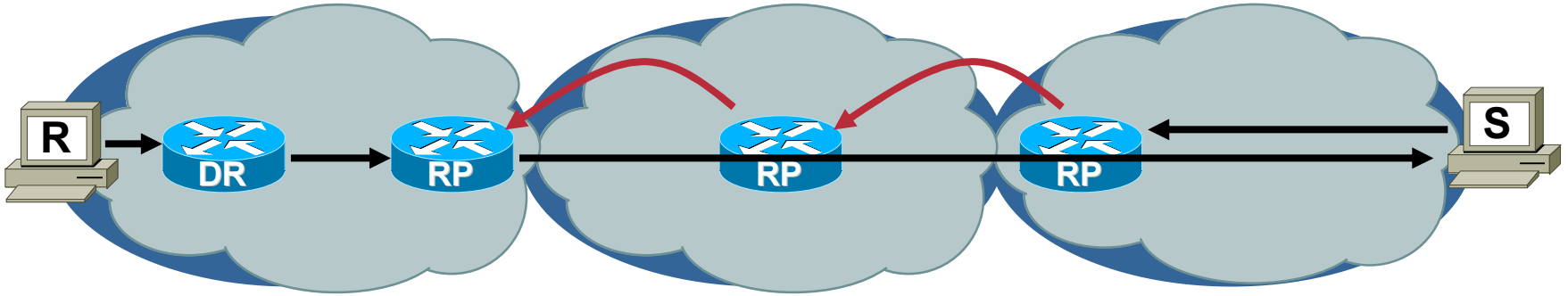


Multicast Interdomain Options
With and Without Rendezvous Points (RP)

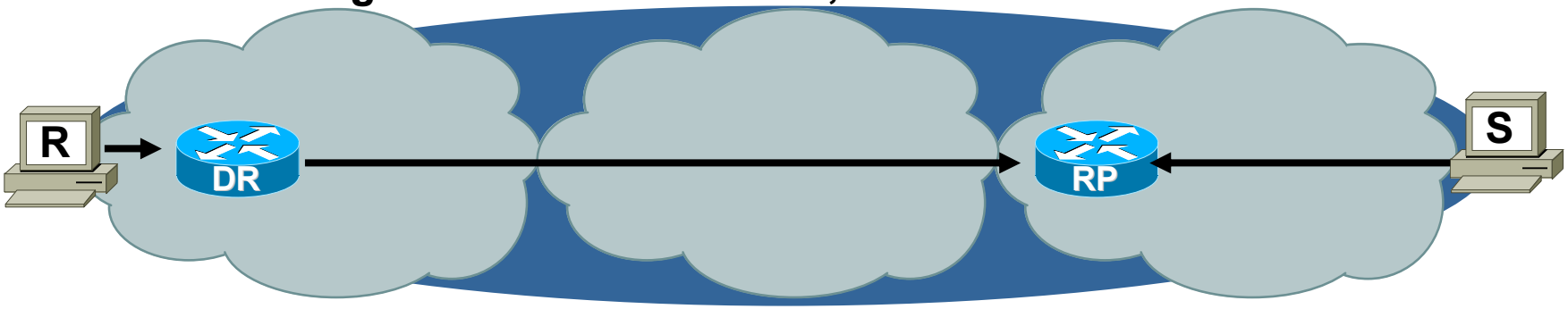
SSM, No RPs



ASM Across Multiple Separate PIM Domains, Each with RP, MSDP Peering



ASM Across Single Shared PIM Domain, One RP



Source Specific Multicast (SSM)

- **NO** configuration required other than enabling
  - `ipv6 multicast-routing`
- SSM group ranges are automatically defined
- Few applications support MLDv2…yet

```
router#show ipv6 pim range-list
config SSM Exp: never Learnt from : ::
  FF33::/32 Up: 1d00h
  FF34::/32 Up: 1d00h
  FF35::/32 Up: 1d00h
  FF36::/32 Up: 1d00h
  FF37::/32 Up: 1d00h
  FF38::/32 Up: 1d00h
  FF39::/32 Up: 1d00h
  FF3A::/32 Up: 1d00h
  FF3B::/32 Up: 1d00h
  FF3C::/32 Up: 1d00h
  FF3D::/32 Up: 1d00h
  FF3E::/32 Up: 1d00h
  FF3F::/32 Up: 1d00h
```

Rendezvous Point (RP) Deployment Types

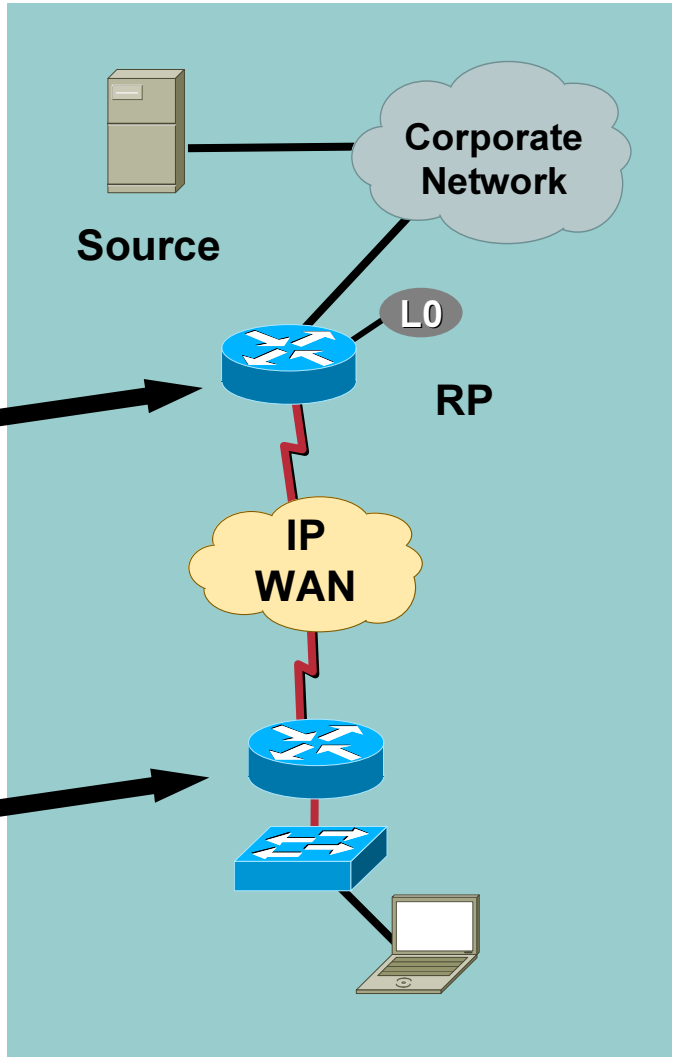
- Static RP
  - For PIM-SM and Bidir-PIM
  - Provides group-to-RP mapping, no RP redundancy
- Boot Strap Router (BSR)
  - Provides group-to-RP mapping AND RP redundancy
- Embedded-RP
  - Easy to deploy
  - Group-to-RP mapping only, no RP redundancy
  - PIM-SM only (today), no Bidir-PIM
- RP redundancy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

IPv6 Multicast Static RP

- **Easier than before as PIM is auto-enabled on every interface**

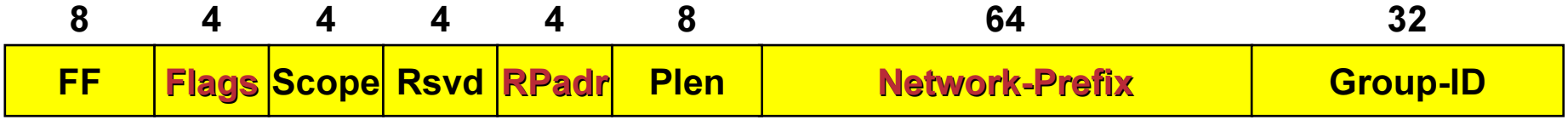
```
ipv6 multicast-routing
!
interface Loopback0
 description IPV6 IPmc RP
 no ip address
 ipv6 address 2001:DB8:C003:110A::1/64
!
ipv6 pim rp-address 2001:DB8:C003:110A::1/64
```

```
ipv6 multicast-routing
!
ipv6 pim rp-address 2001:DB8:C003:110A::1/64
```



Embedded RP Addressing

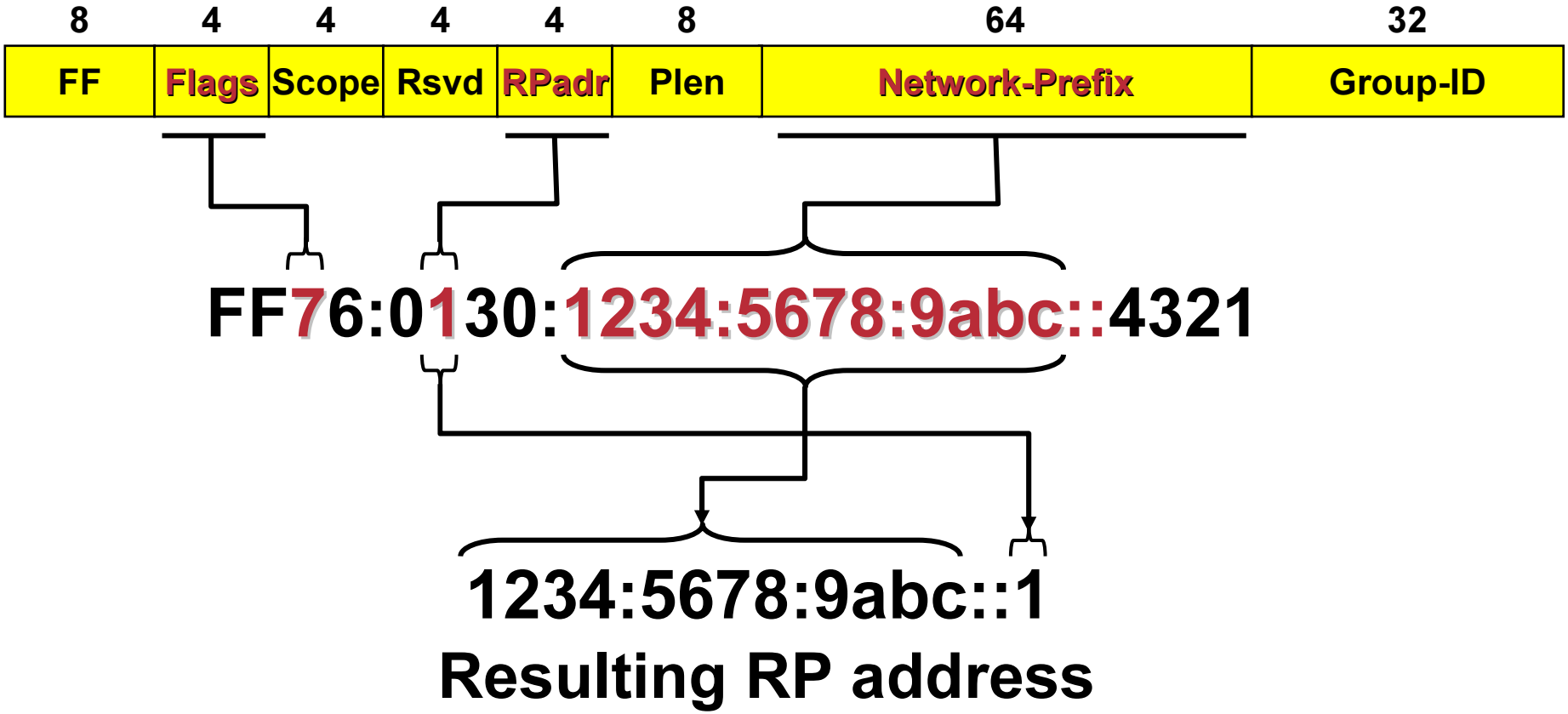
draft-savola-mboned-mcast-rpaddr-03.txt



- Proposed new multicast address type
  - Uses Unicast-Based Multicast addresses (RFC 3306)
- RP Address is embedded in multicast address.
- Flag bits = 0RPT
  - R = 1, P = 1, T = 1 => Embedded RP Address
- Network-Prefix::RPadr = RP address

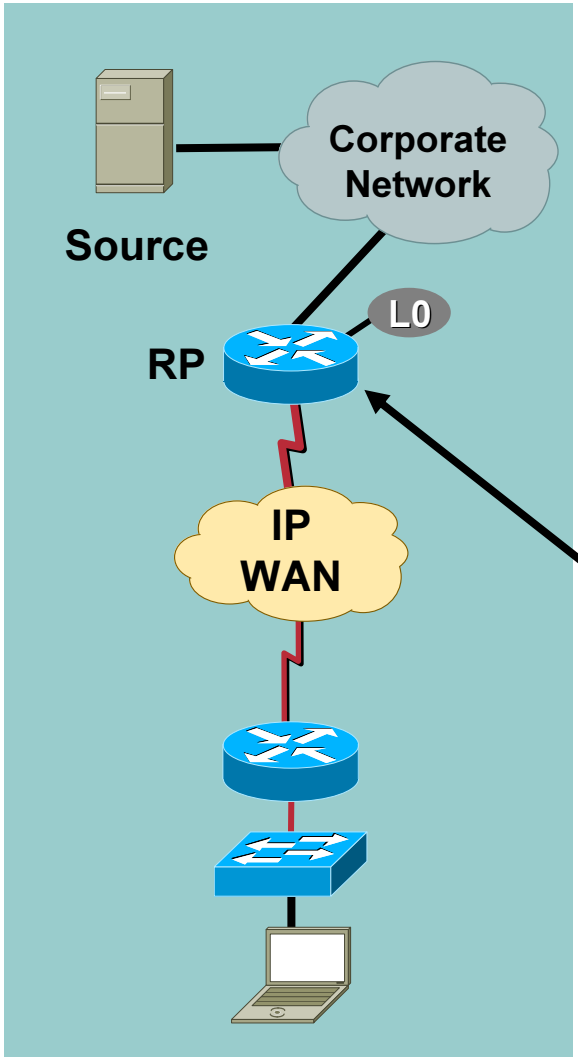
Embedded RP Addressing – Example

Multicast Address with Embedded RP address



Resulting RP address

Embedded-RP Configuration Example



- RP to be used as an Embedded-RP needs to be configured with address/ group range
- All other non-RP routers require no special configuration

```
ipv6 pim rp-address 2001:DB8:C003:111D::1 ERP
!
ipv6 access-list ERP
 permit ipv6 any FF7E:140:2001:DB8:C003:111D::/96
```

Pick Your Flavor – One size does NOT fit all

- **PIM-SSM:**
  - **One/Few-to-Many applications.**
    - **Requires MLDv2 and the app to support SSM operation.**
- **Embedded-RP:**
  - **Simple to deploy.**
  - **New and does not currently provide for RP redundancy.**
- **PIM-BSR:**
  - **Provides for easier RP deployment than static RP.**
  - **Provides for RP redundancy (albeit slow), but is a bit more complicated.**
- **Cisco is working on scalable and highly-available RP deployment methods.**

Multi Topology Routing (MTR)

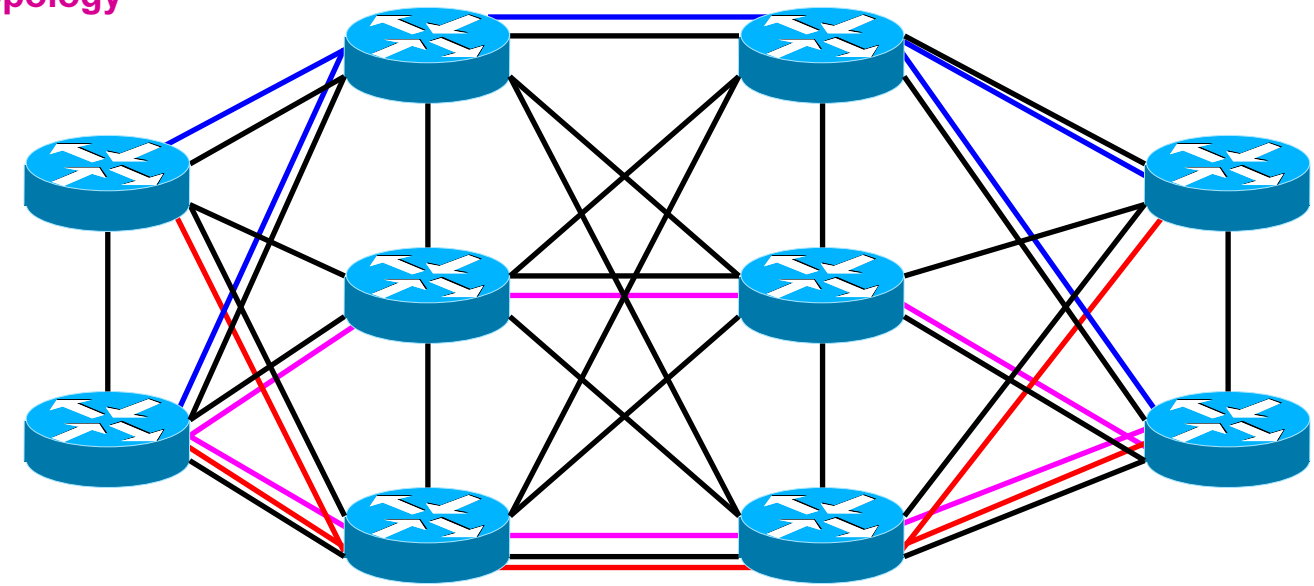
Conceptual View of MTR

- Creation of multiple topologies
  - Logical path that traffic will take across the given network
  - Each topology will route/forward a subset of the traffic as defined by the classification criteria
- Mapping of traffic to a topology
  - Determine which traffic (based on a classification criteria) is subject to topology specific forwarding
- MTR vs. QoS
  - QoS provides per-hop service differentiation within a single path
  - MTR provides PATH-BASED service differentiation within a single domain

Multi-Topology Routing
Defining Topologies

Base Topology
Voice Topology
Multicast Topology
Video Topology

Start with a Base Topology
Includes all routers and all links



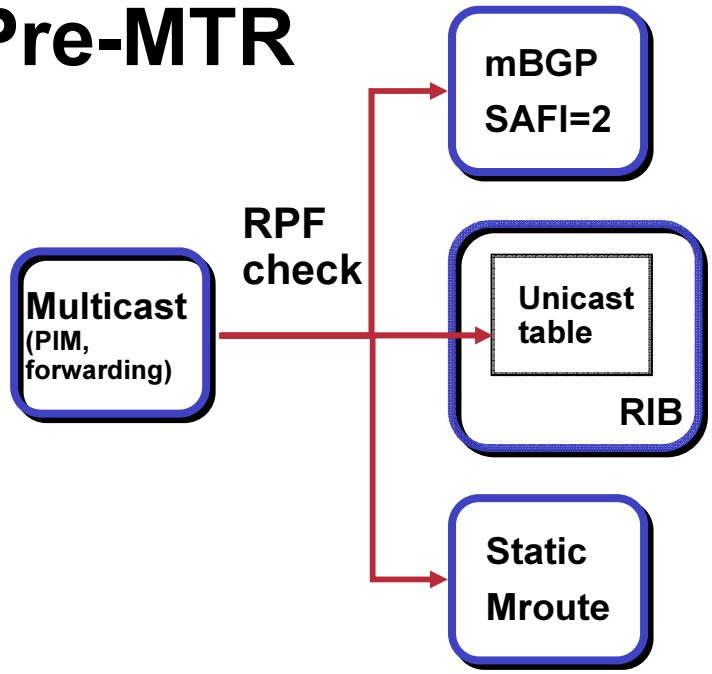
- Define the class-specific topology across a contiguous section of the network
- Individual links can belong to multiple topologies

Incongruent IPv4 Unicast and Multicast Topologies

- MTR allows incongruent unicast and multicast topologies
  - Metrics can be different for each on the same link
- Restrict traffic
  - Restrict multicast only to designated areas of the network
- No reference to unicast for multicast RPF
  - RPF checks based on multicast specific table
- Multicast specific protocols such as PIM are not topology specific

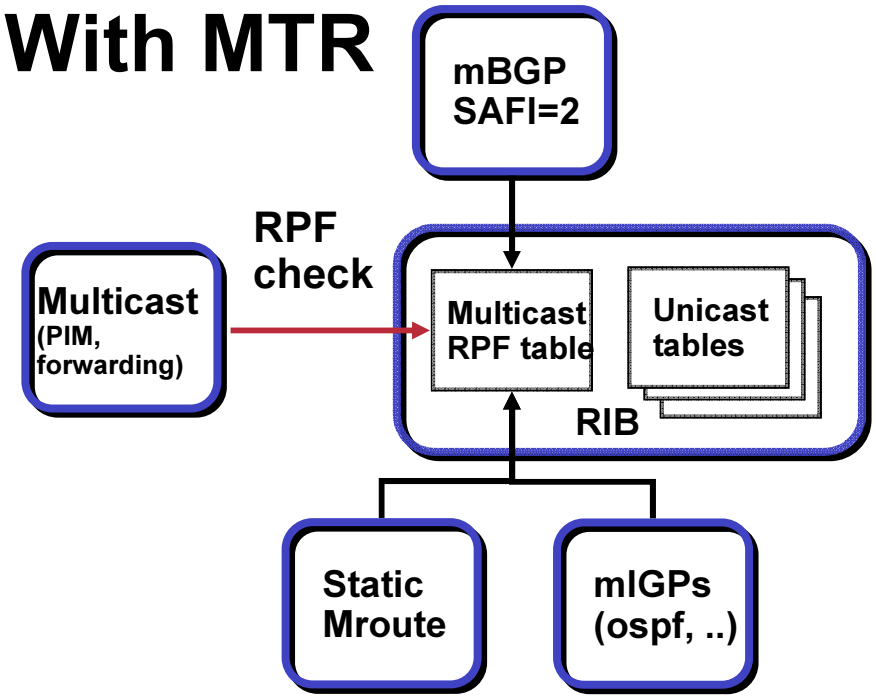
MTR Multicast Changes

Pre-MTR



- There is no single database for Multicast RPF.
  - Multicast consults multiple sources of “unicast” routes.
  - It selects the results according to a custom preference rule.
- Standard unicast routes (uRIB) are *always* considered.
  - There is no way to exclude them completely.

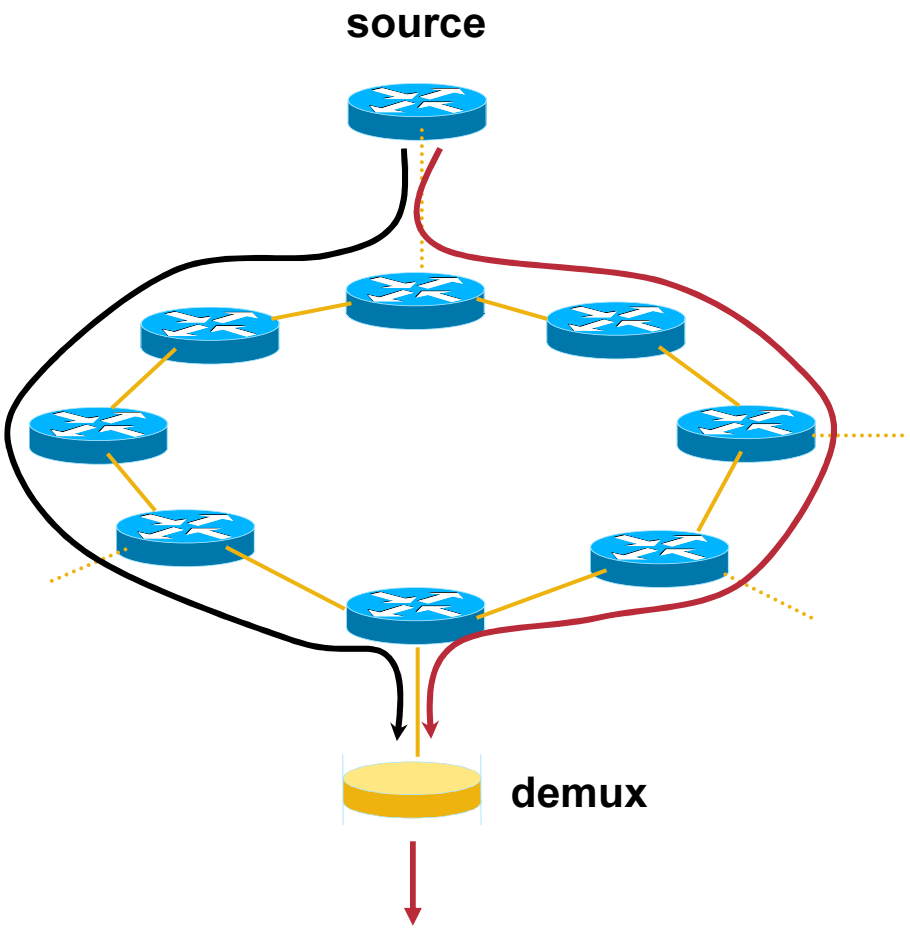
MTR Multicast Changes



- MTR manages a routing table specifically for Multicast RPF.
  - Multicast uses that table as the sole source of RPF routes.
- Highly flexible.
  - Any protocol may be configured to contribute to the mRPF table.
- Standard configuration options are available.
  - Including redistribution among protocols.

Multicast MTR Broadband Application

- Brute force fast convergence: receive twice from two different paths and re-source one of the streams
- When one of the paths fails re-source the other stream
- Potentially need to be able to receive from the same source using 2 completely different paths --> use 2 different RPF topologies
- May need to stop routing convergence of the broken topology trying to use the paths of the unaffected topology



Multicast MTR CLI example

- **Enable Multicast MTR mode**
  - ip multicast multitopology-rpf
  - The multicast RPF RIB table (multicast base topology) is empty until a protocol is enabled for multicast..
- **Enable OSPF on the multicast base topology**

```
router ospf 1
  network 10.1.0.0 255.255.0.0
  address-family ipv4 multicast
```

- **Configure ospf multicast cost on an interface**

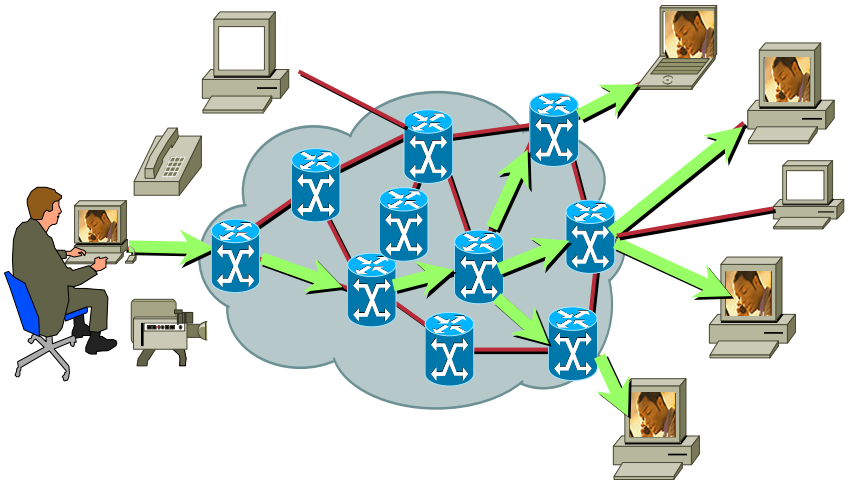
```
interface ethenet1/0
  ip address 10.1.1.1 255.255.255.0
  ip ospf multicast cost 20
```

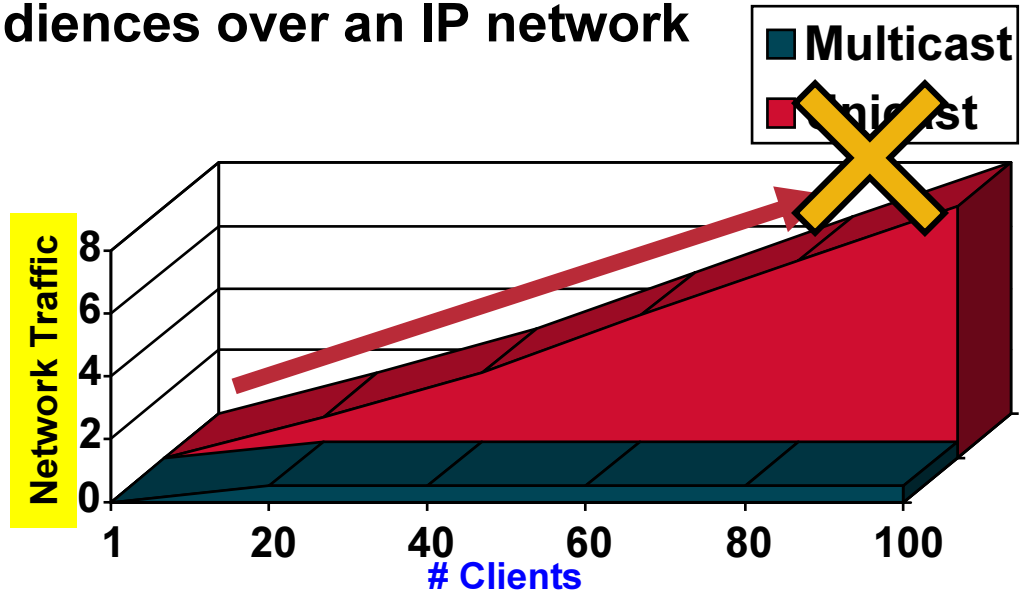
- **Display the content of the multicast base topology**
  - show ip route <usual options> multicast

Triple Play and Multicast

Relevance of Multicast for IPTV delivery

Distribute information to large audiences over an IP network





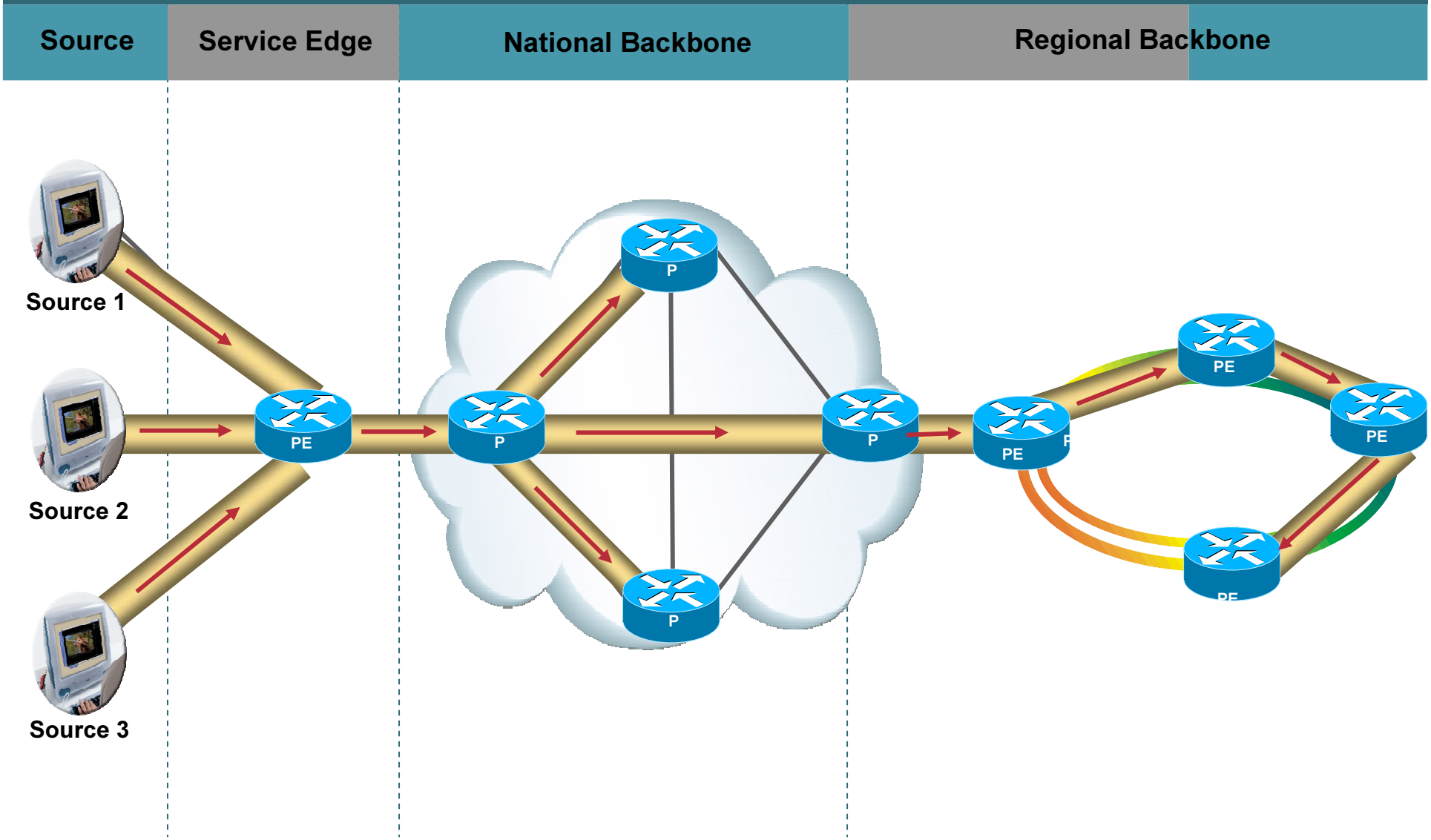
Multicast

1. Efficiently Controls network traffic
2. Reduces server and CPU loads
3. Eliminates traffic redundancy
4. Makes Multipoint applications possible

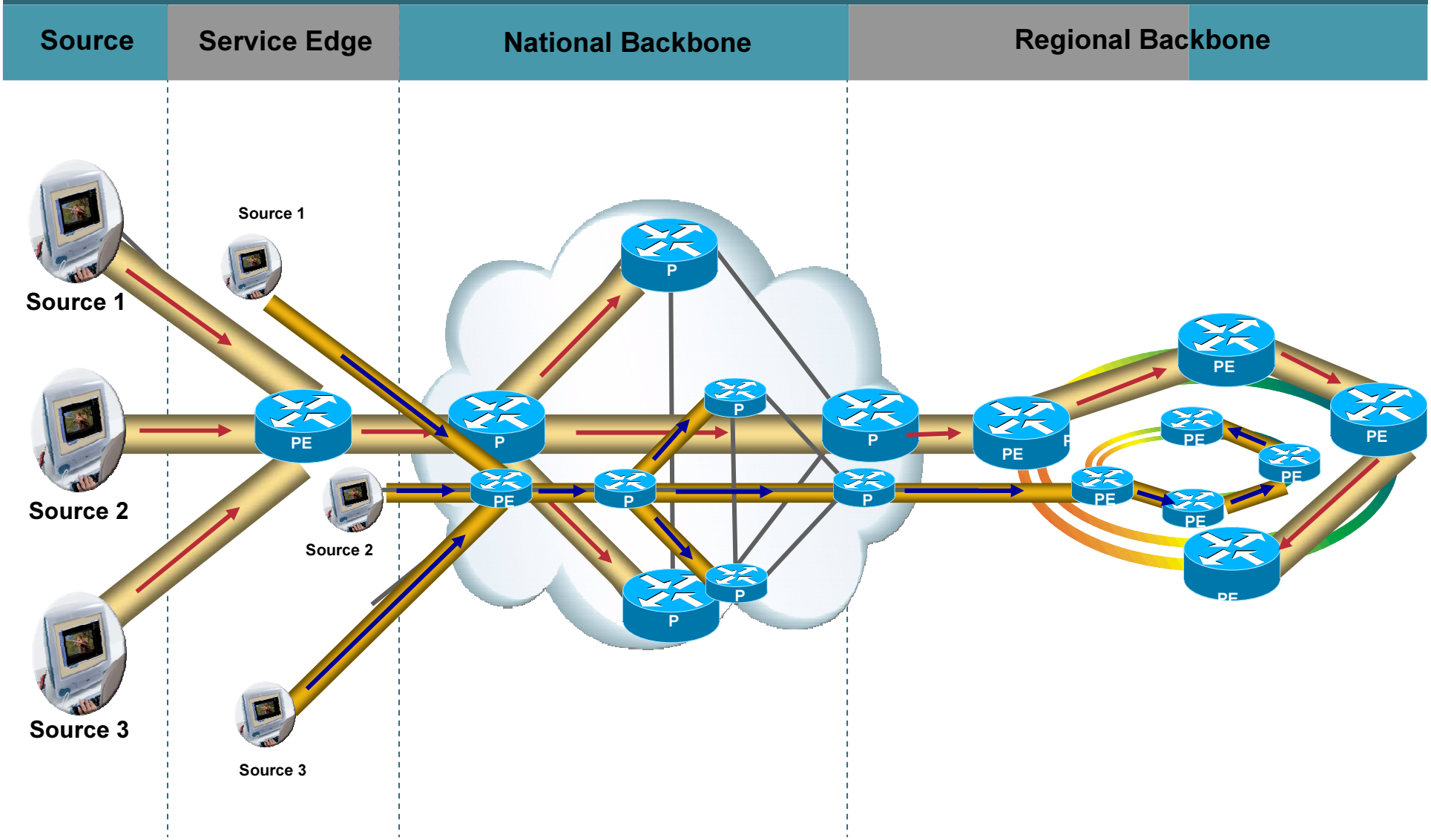
Multicast Benefits

- Increase Productivity & Save Cost
- Generate New Revenue Stream

End to End Recovery Models
Hot-Hot Video Delivery Model

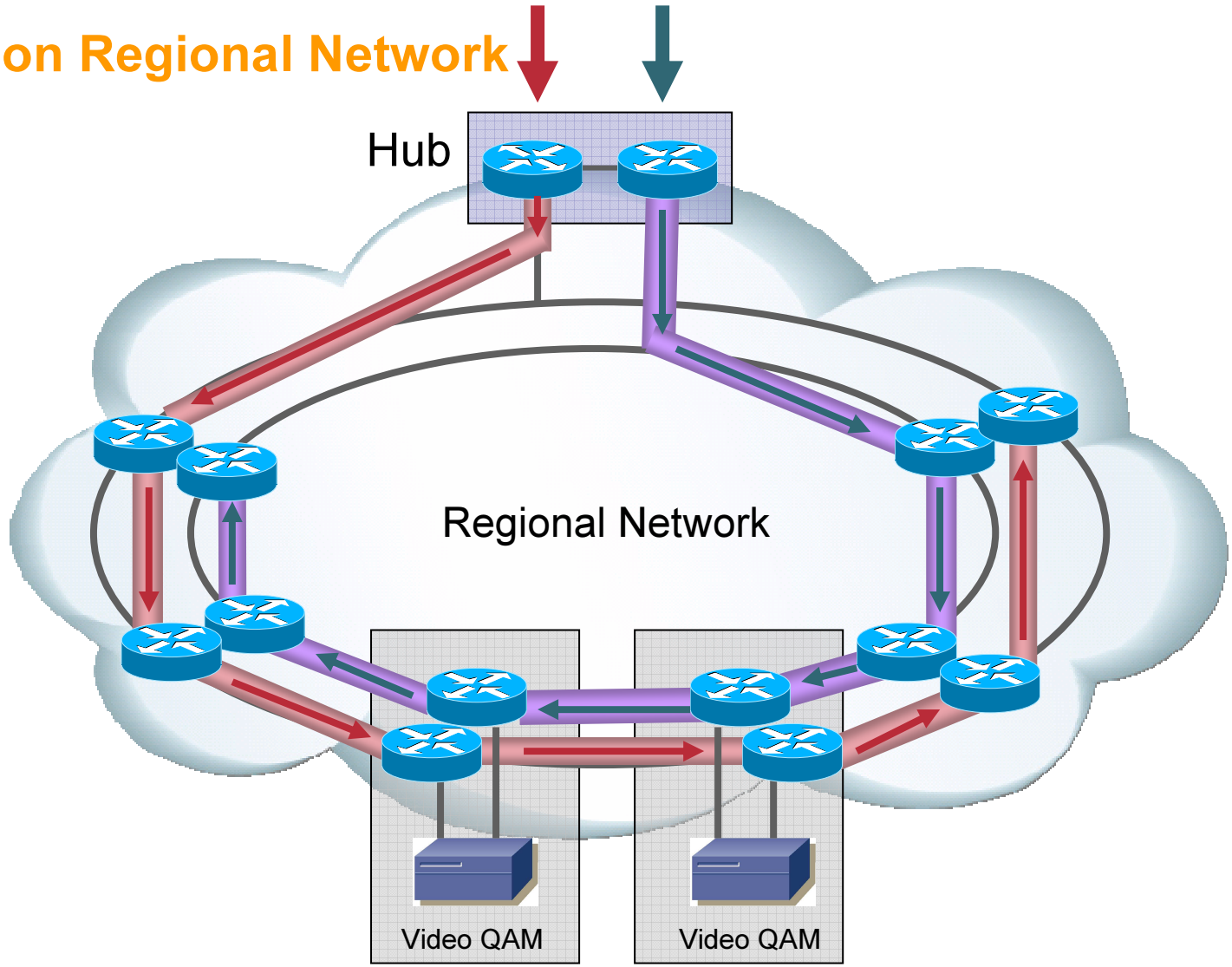


End to End Recovery Models
Hot-Hot Video Delivery Model

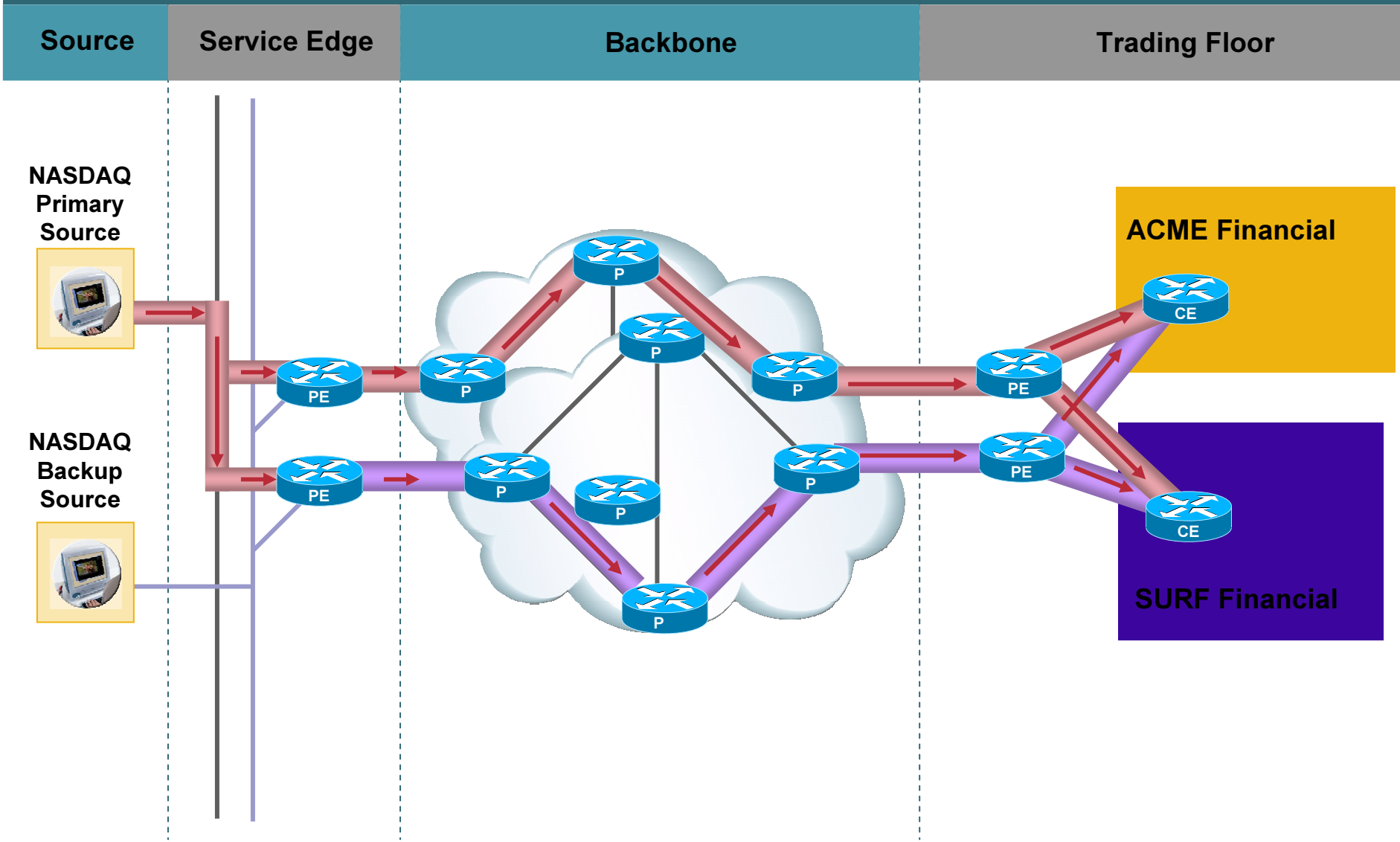


End to End Recovery Models
Hot-Hot Video Delivery Model

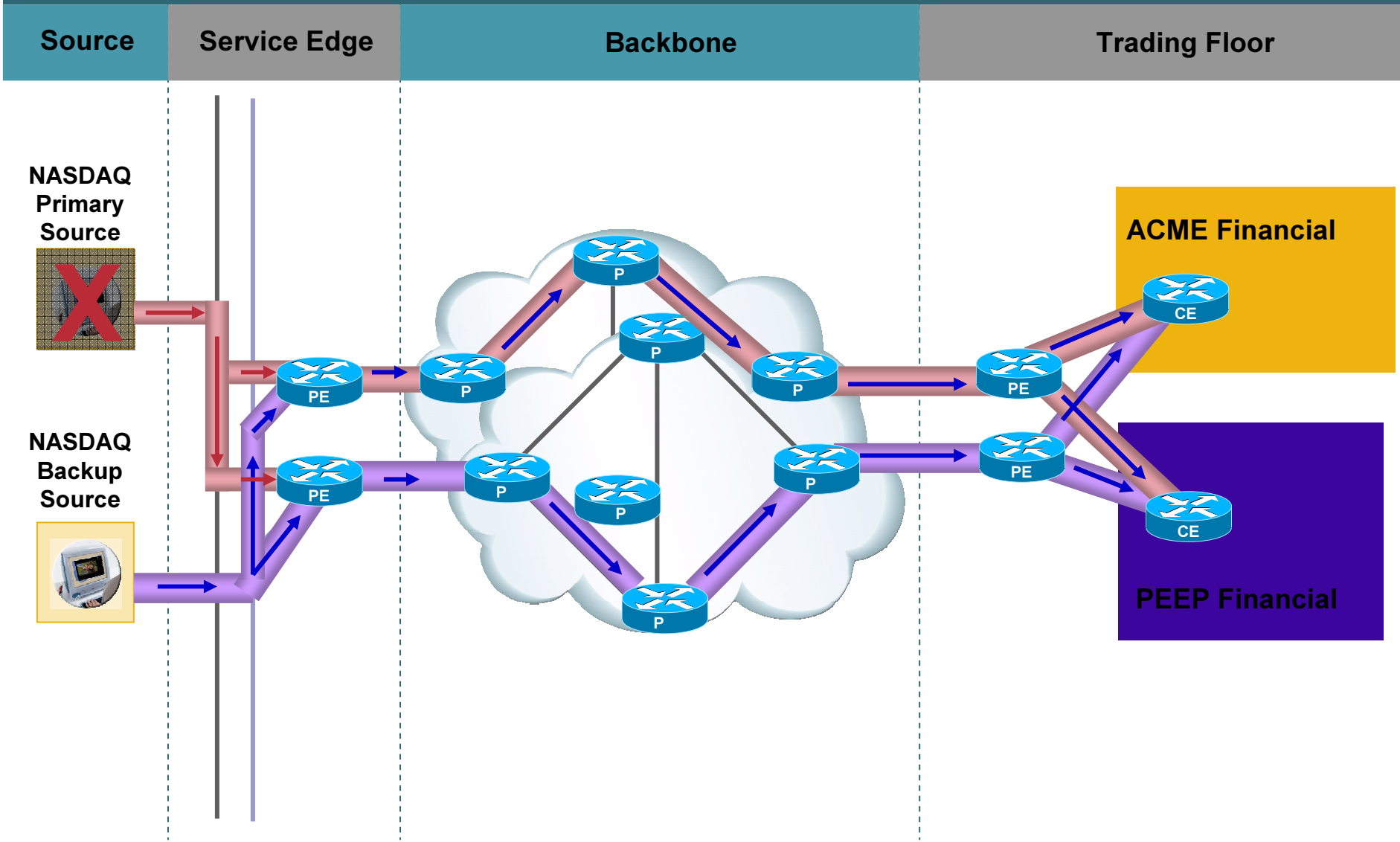
Zoom in on Regional Network



End to End Recovery Models
Financials Industry Model



End to End Recovery Models
Financials Industry Model



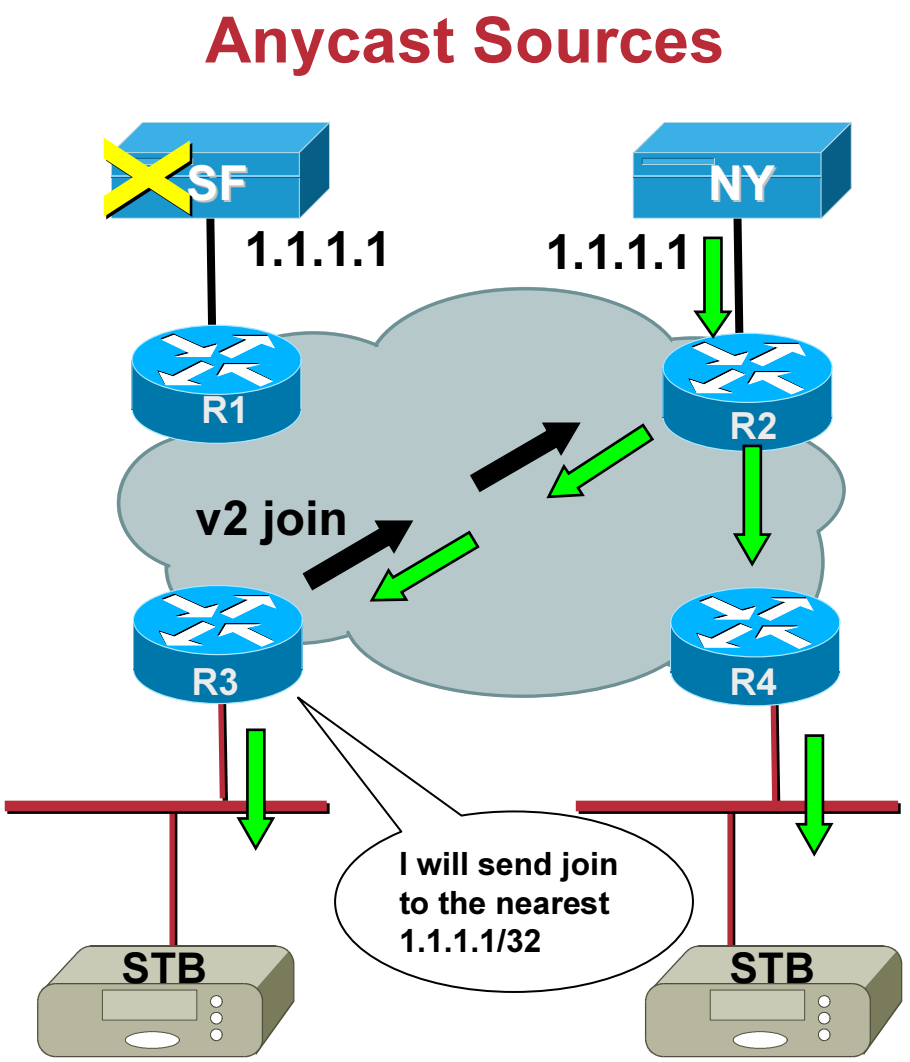
Source Redundancy : Two Approaches

Primary-Backup	Hot-Hot
Two sources, One is active and src'ing content, Second is in standby mode (not src'ing content) Heartbeat mechanism used to communicate with each other	Two sources, *both* are active and src'ing multicast into the network No Protocol between the two sources
Only one copy is on the network at any instant Single Multicast tree is built per the unicast routing table	Two copies of the multicast packets will be in the network at any instant Two Multicast tree on almost redundant Infrastructure
Uses required bandwidth	Uses 2x network bandwidth
Receiver's functionality simpler: Aware of only one src, fail-over logic handled between sources.	Receiver is smarter: Is aware/configured with two feeds (s1,g1), (s2,g2) / (*,g1), (*,g2) Joins both and receives both feeds
This approach requires the network to have fast IGP and PIM convergence	This approach does not require fast IGP and PIM convergence

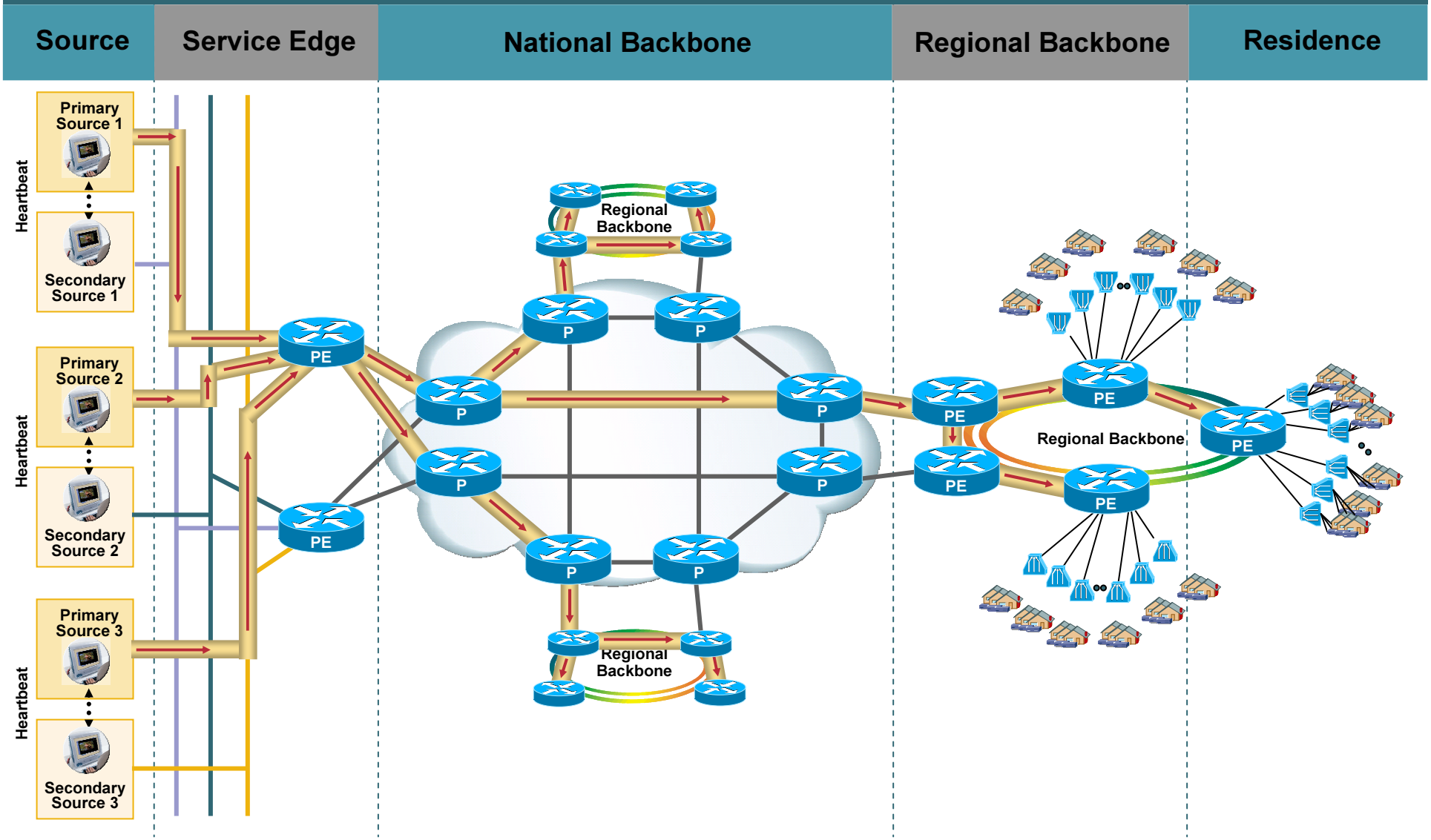
Multicast Source Redundancy Using Anycast Sources

How is source redundancy achieved in the network?

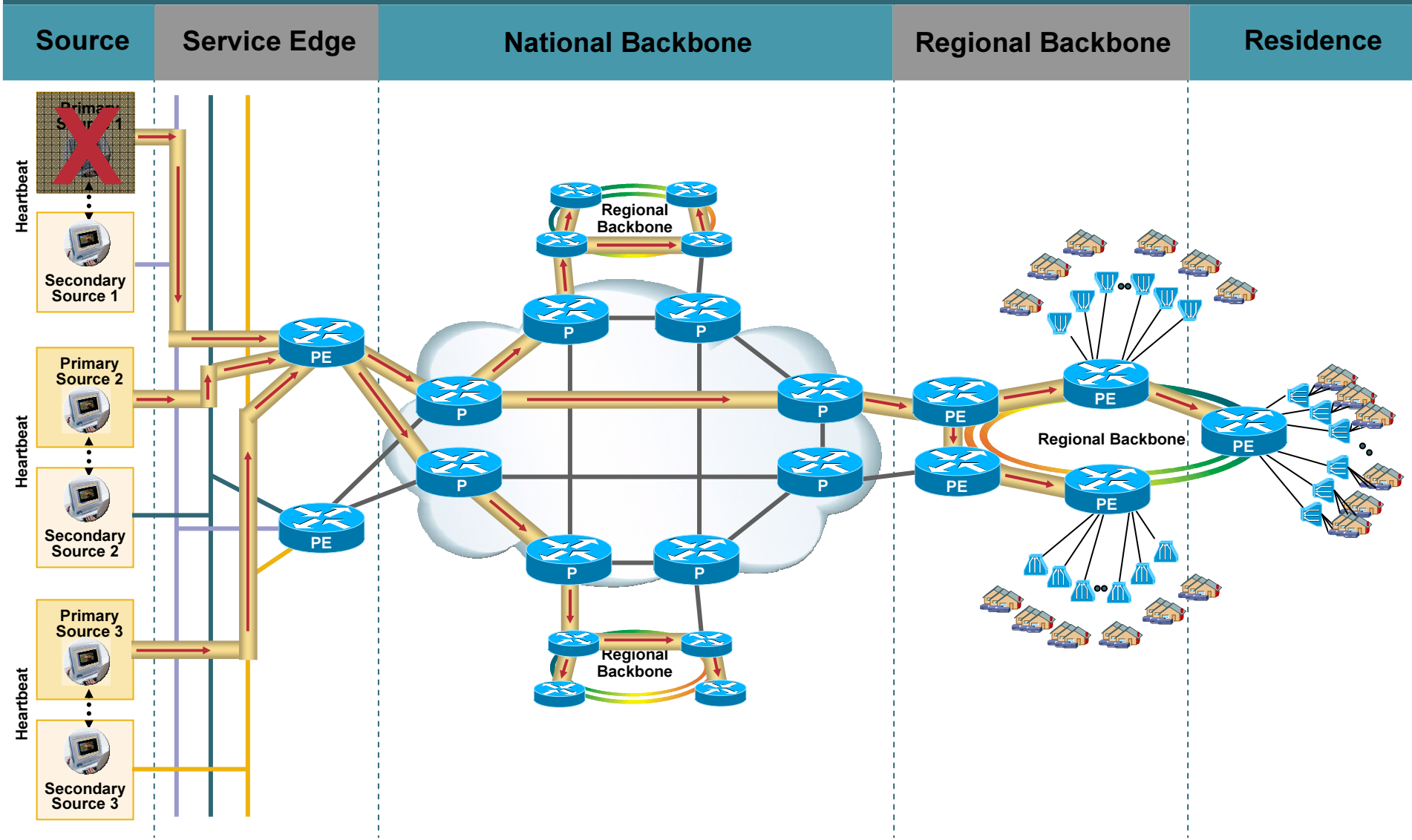
- Enable SSM on all routers
- Have R1 and R2 advertise same prefix for each source segment.
- R3 and R4 follow best path towards source based on IGP metrics.
- Let's say R3's best path to SF is through R1. The source in SF now suddenly fails.
- R3's IGP will reconverge and trigger SSM joins towards R2 in NY.



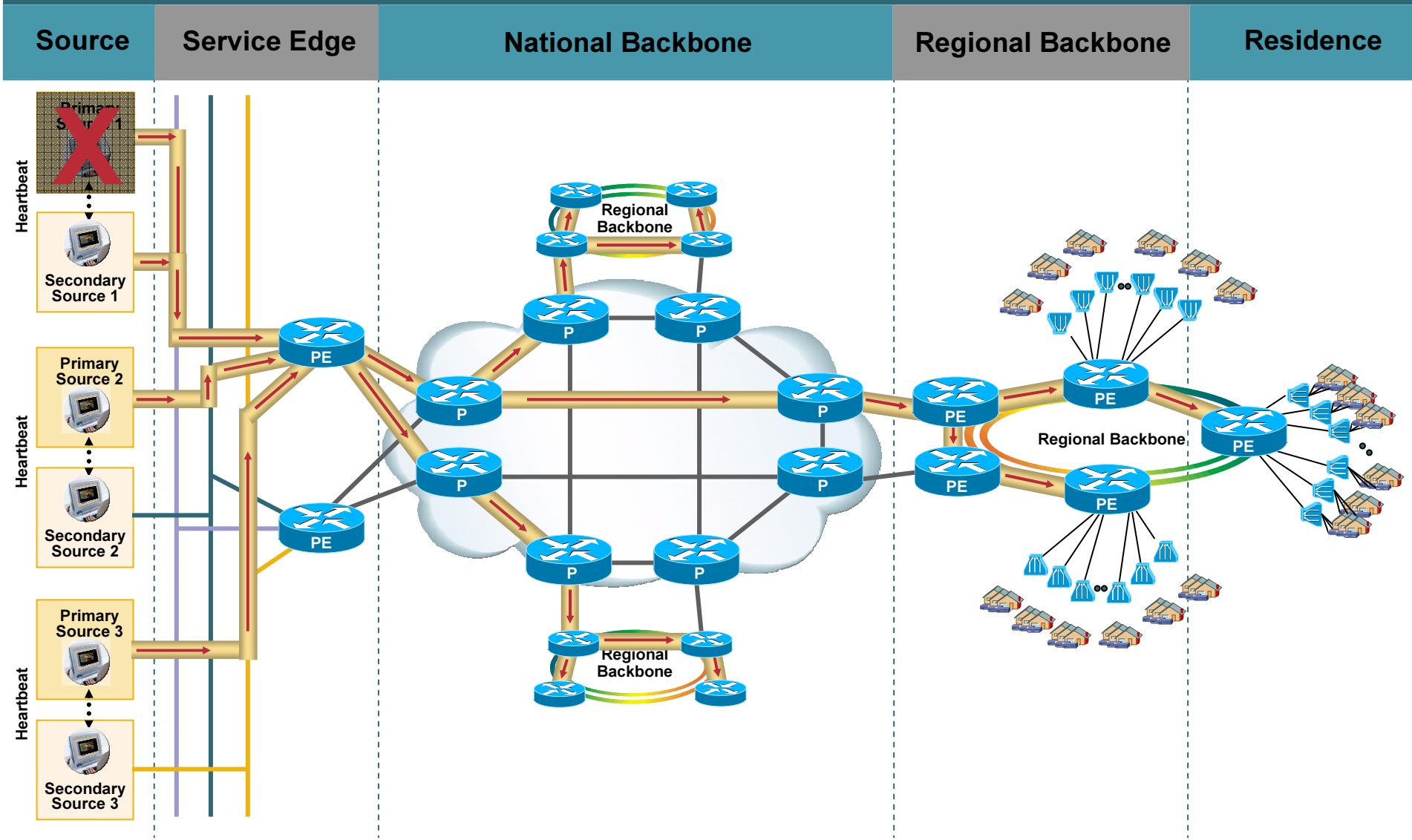
Native IP Multicast Video Triple Play
Redundancy : Video Source Failure



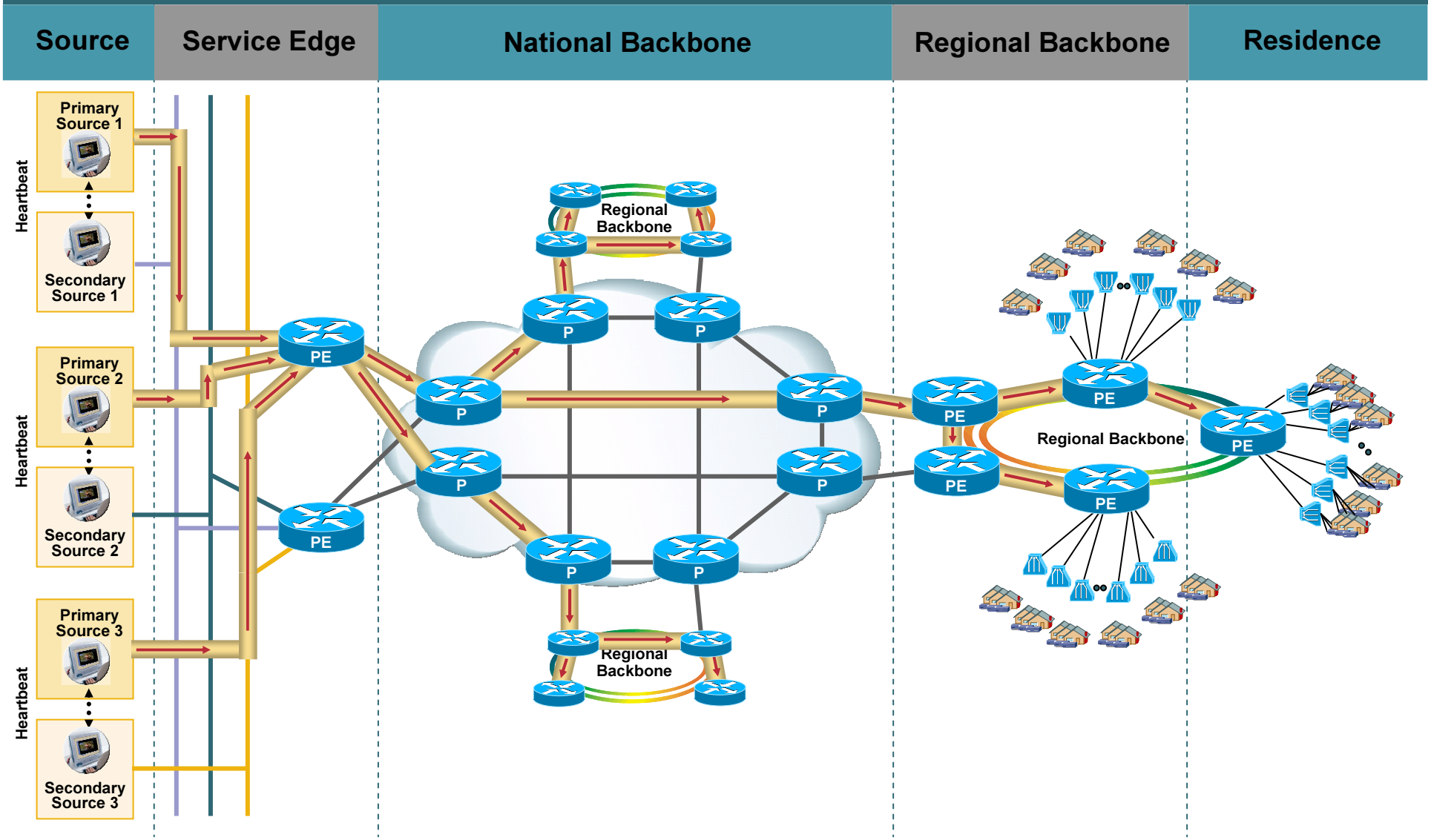
Native IP Multicast Video Triple Play
Redundancy : Video Source Failure



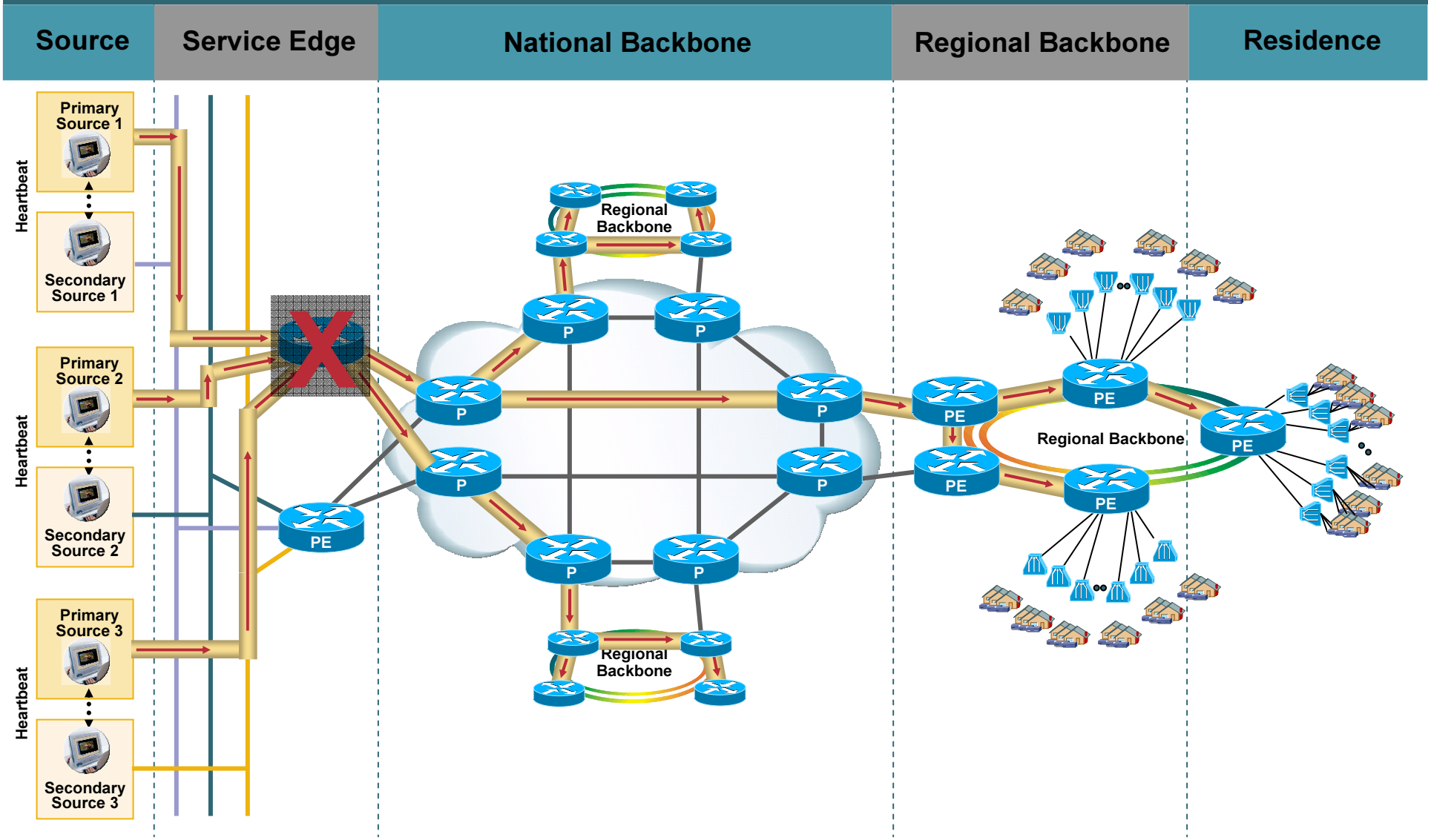
Native IP Multicast Video Triple Play
Redundancy : Video Source Failure



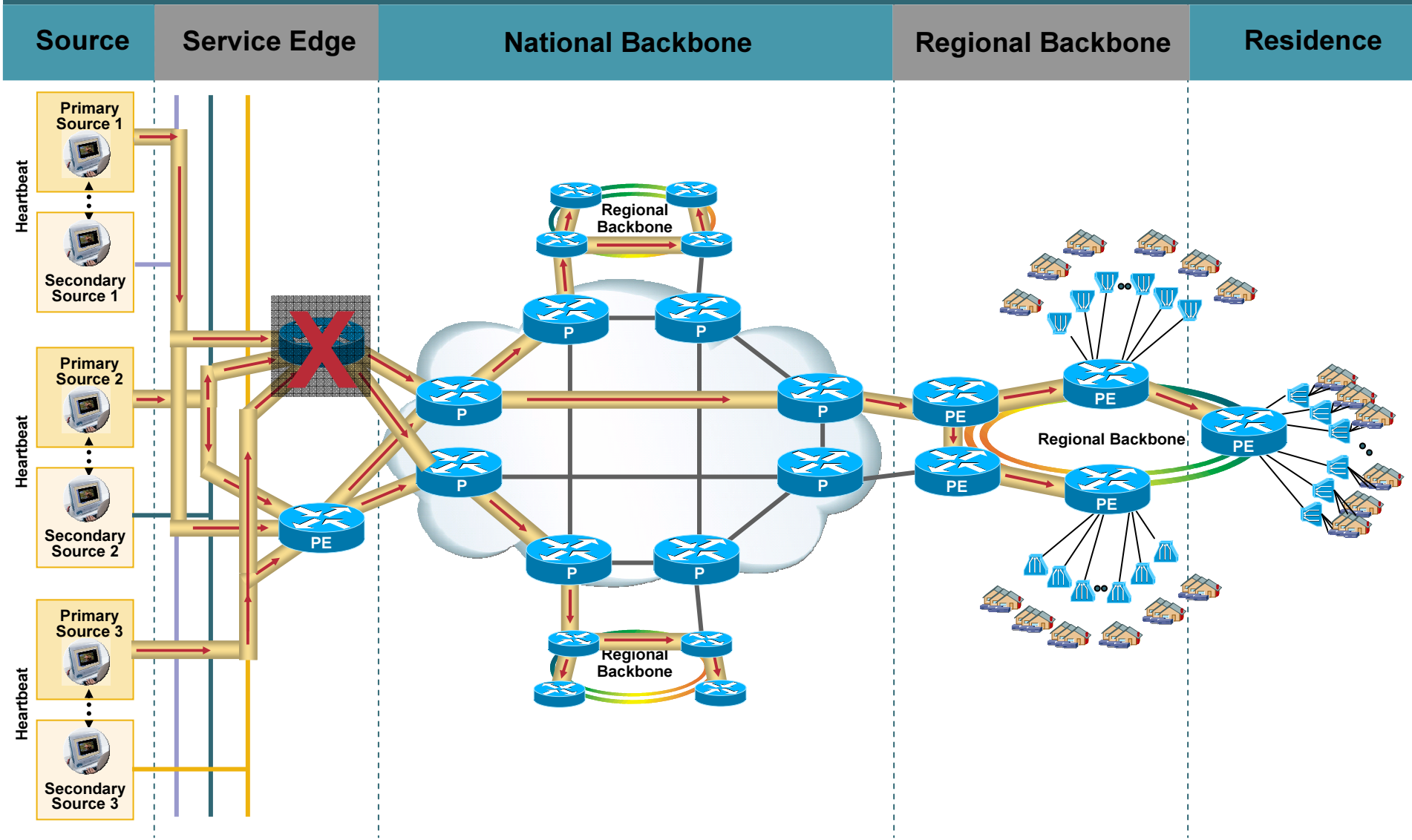
Native IP Multicast Video Triple Play
Redundancy : Source Router Failure



Native IP Multicast Video Triple Play
Redundancy : Source Router Failure



Native IP Multicast Video Triple Play
Redundancy : Source Router Failure



Fast Join/Leave for Faster Channel Change

Problem Description:

In networks where bandwidth is constrained between multicast routers and hosts (like in xDSL deployments), fast channel changes can easily lead to ***bandwidth oversubscription***, resulting in a temporary degradation of traffic flow for all users.

Solution:

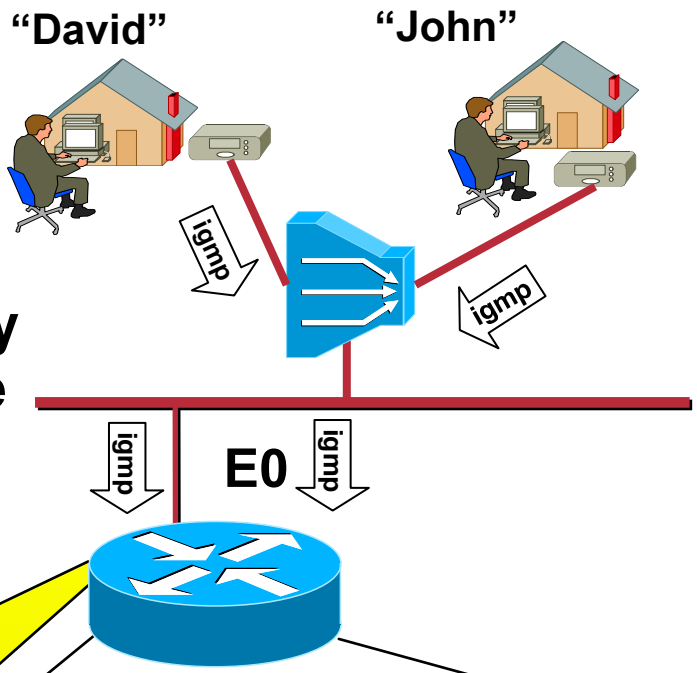
Reduce the leave latency during a channel change by extending the IGMPv3 protocol.

Benefits:

- Faster channel changing without BW oversubscription
- Improved diagnostics capabilities

Multicast Fast Join/Leave for Faster Channel Change

- Relies on IGMPv3
- Router tracks both User and Channel(s) being watched
- When user leaves channel no one else is watching, router immediately prunes the channel off the interface compared to IGMPv2 (up to 3 seconds) and IGMPv1 (up to 180 seconds)!



Configuration:

```
interface Ethernet 0
  ip pim sparse-mode
    ip igmp version 3
    ip igmp explicit-tracking
```

First introduced in 12.0(29)S

Int	Channel	User
~~E0~~	~~10.0.0.1, 239.1.1.1~~	~~"David"~~
E0	10.0.0.1, 239.2.2.2	"John"
E0	10.0.0.1, 239.3.3.3	"David"

Multicast in the Fast Path

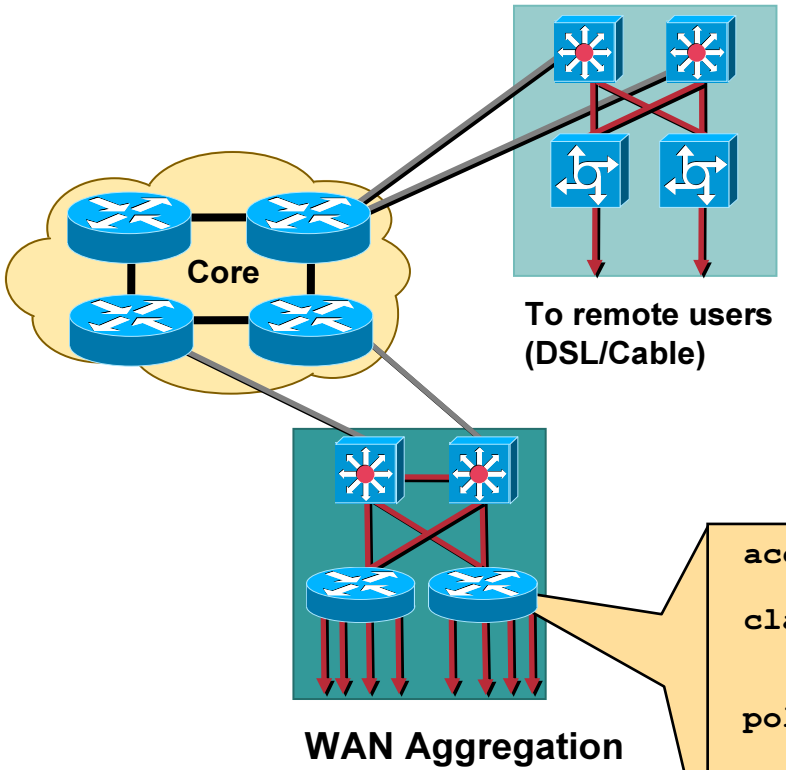
- CEF does not support Multicast traffic
- Multicast forwards traffic in the “Fast Path”
- Legacy QoS Features that were designed to run in the fast path will work with IPmc
  - Priority Queueing, Custom Queueing, WFQ, RSVP, IP RTP Priority
- PBR also supports IPmc and can be used for Marking
  - Supported on distributed platforms 12.0(13)S, 12.1(4)
- These Legacy QoS features will work but MQC features should be used where supported

Multicast with MQC

Modular QoS CLI structure was introduced in 12.0(5)T

- **MQC Features are supported in CEF/dCEF path - would not work with IPmc**
- **Changes were made so that MQC features would work with IPmc - 12.1(5)T**
  - **Classification, Policing, Queueing and Shaping are supported**

IP/TV – SSM over WAN



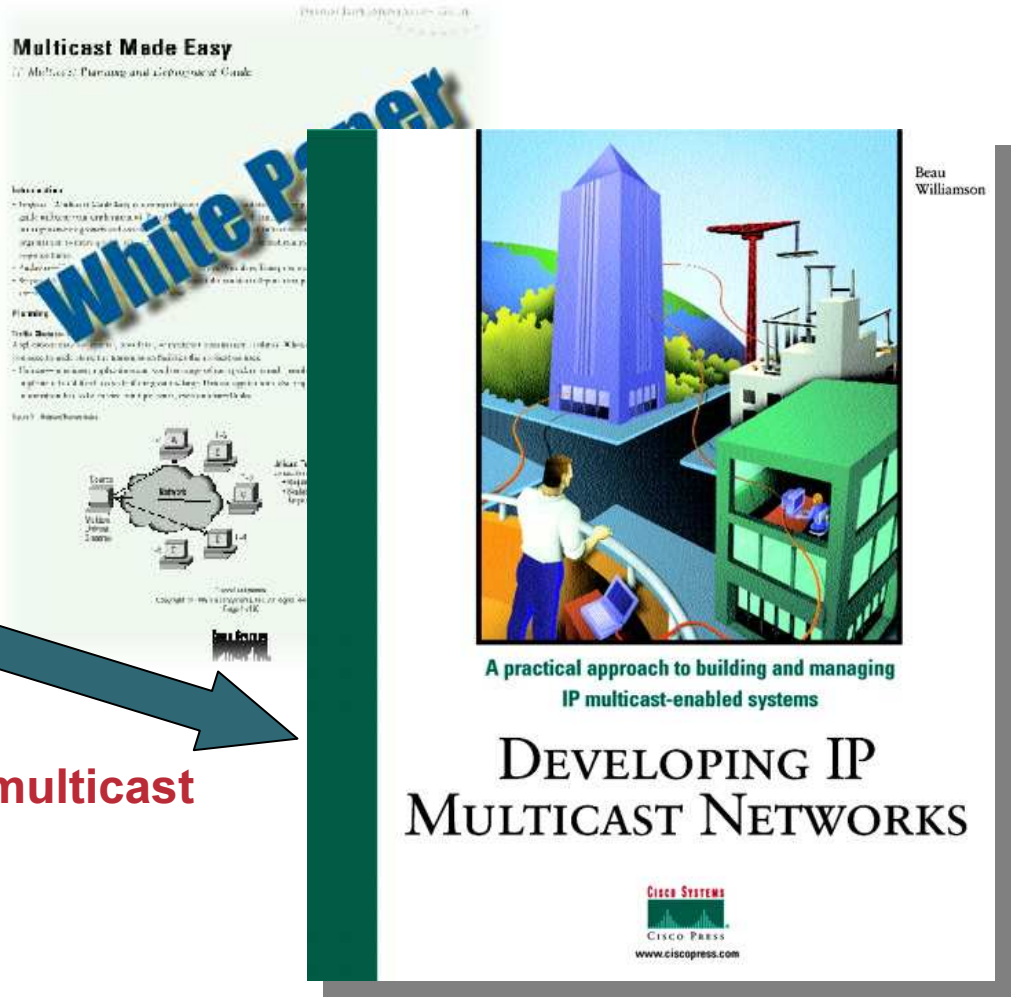
- QOS protection necessary for remote users and WAN aggregation
- Use bandwidth scoping for SSM groups with different rates (239.232.QOS.x)
- Use bandwidth scoping to deny high rate streams
- Use MQC to allocate bandwidth to IP/TV:

```
access-list 101 permit ip any 239.232.224.0 0.0.31.255

class-map match-all iptv-qos-lowbandwidth
    match access-group 101

policy-map IPTV-over-T1
    class iptv-qos-lowbandwidth
       bandwidth 512
    class default
       fair-queue
```

More Information

- White Papers
- Web and Mailers
- Cisco Press

RTIB

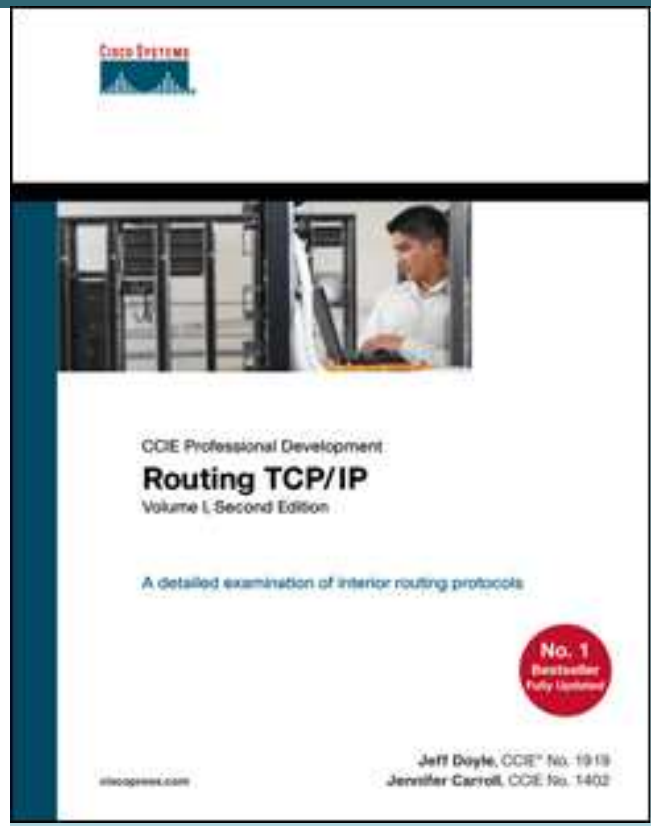


CCO Multicast page:
http://www.cisco.com/go/ipmulticast
Questions:
cs-ipmulticast@cisco.com
Customer Support Mailing List:
tac@cisco.com

RTIB = "Read the Incredible Book"

Recommended Reading

- Continue your Cisco Networkers learning experience with further reading from Cisco Press
- Check the Recommended Reading flyer for suggested books



Available Onsite at the Cisco Company Store

Complete Your Online Session Evaluation

- Win fabulous prizes; Give us your feedback
- Receive ten Passport Points for each session evaluation you complete
- Go to the Internet stations located throughout the Convention Center to complete your session evaluation
- Drawings will be held in the World of Solutions
  - Tuesday, June 20 at 12:15 p.m.
  - Wednesday, June 21 at 12:15 p.m.
  - Thursday, June 22 at 12:15 p.m. and 2:00 p.m.