Advanced IP Multicast Concepts

BRKIPM-3018

Steve Simlo

Cisco Networkers 2007

HOUSEKEEPING

- We value your feedback, don't forget to complete your online session evaluations after each session and complete the Overall Conference Evaluation which will be available online from Friday.
- Visit the World of Solutions on Level -01!
- Please remember this is a 'No Smoking' venue!
- Please switch off your mobile phones!
- Please remember to wear your badge at all times including the Party!
- Do you have a question? Feel free to ask them during the Q&A section or write your question on the Question form given to you and hand it to the Room Monitor when you see them holding up the Q&A sign.

Now is the time to leave ☺ if you do not want to hear about:

- Advanced multicast topics associated with inter-domain
- The fundamentals of MBGP and MSDP
- Source-specific multicast and how this mode solves many problems associated with traditional inter-domain multicast
- Building multicast VPNs in an MPLS VPN environment using multicast domains and multipoint LSP solutions
- A brief overview of IPv6 multicast
- Multicast using multi-topology routing (MTR)
- Multicast for a ‘Triple play’ deployment

Agenda

- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

Agenda

- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

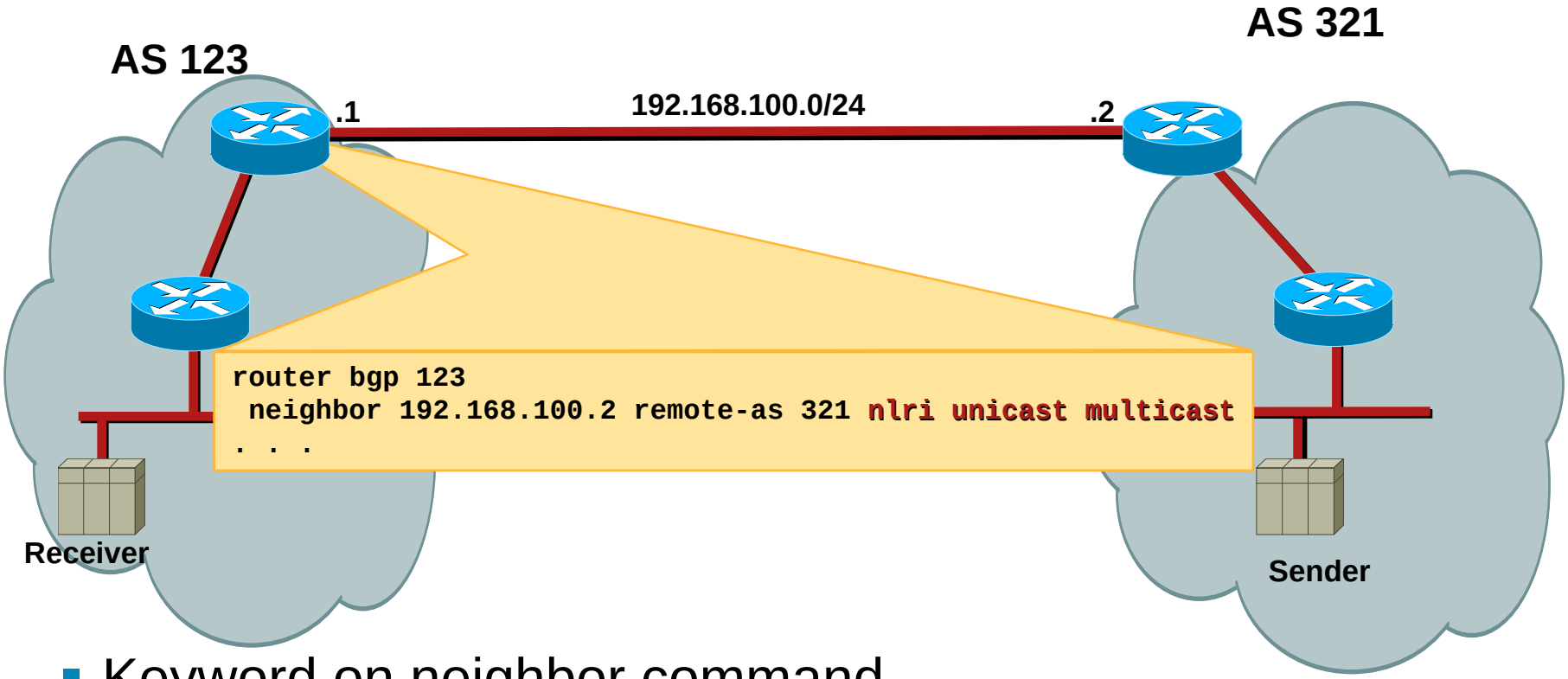
MBGP Overview

- 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
  - No multicast state propagated: still need PIM
  - Path selection and validation rules: AS-Path, LocalPref, MED
- 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



- Keyword on neighbor command

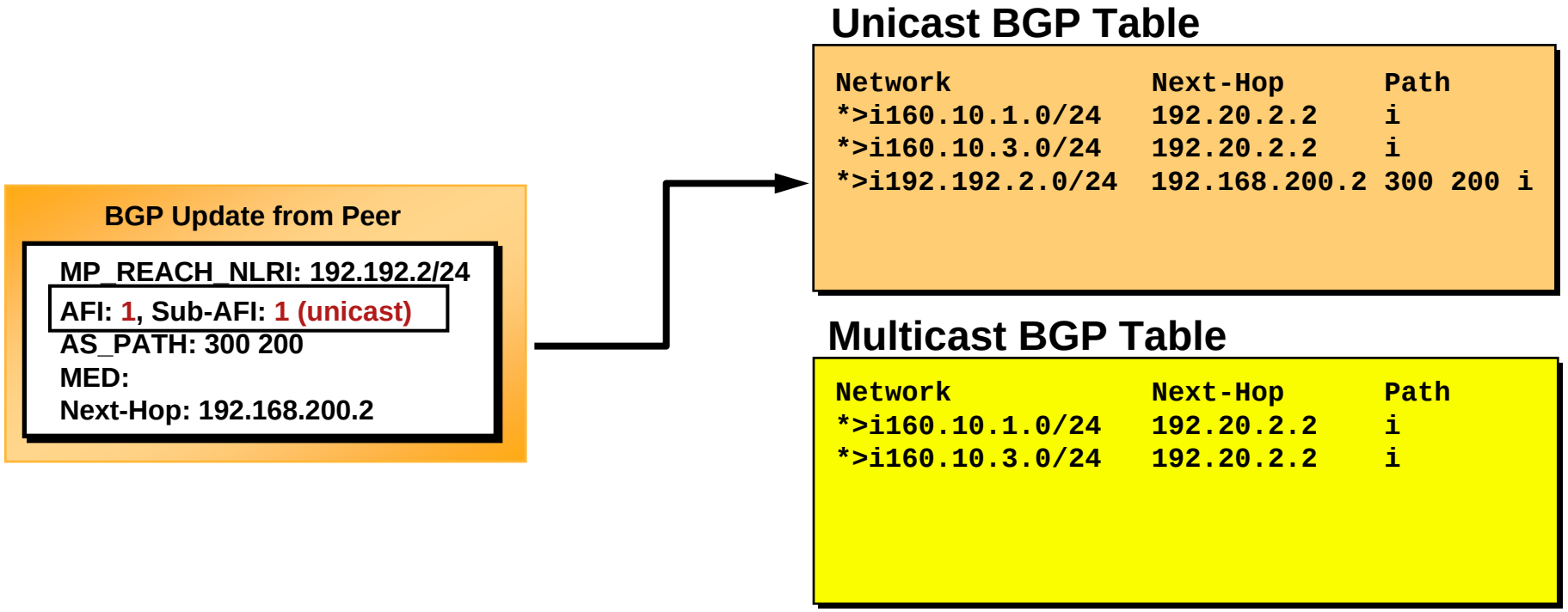
  `neighbor <foo> remote-as <asn> nlri multicast unicast`

  Configures router to negotiate either or both NLRI

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

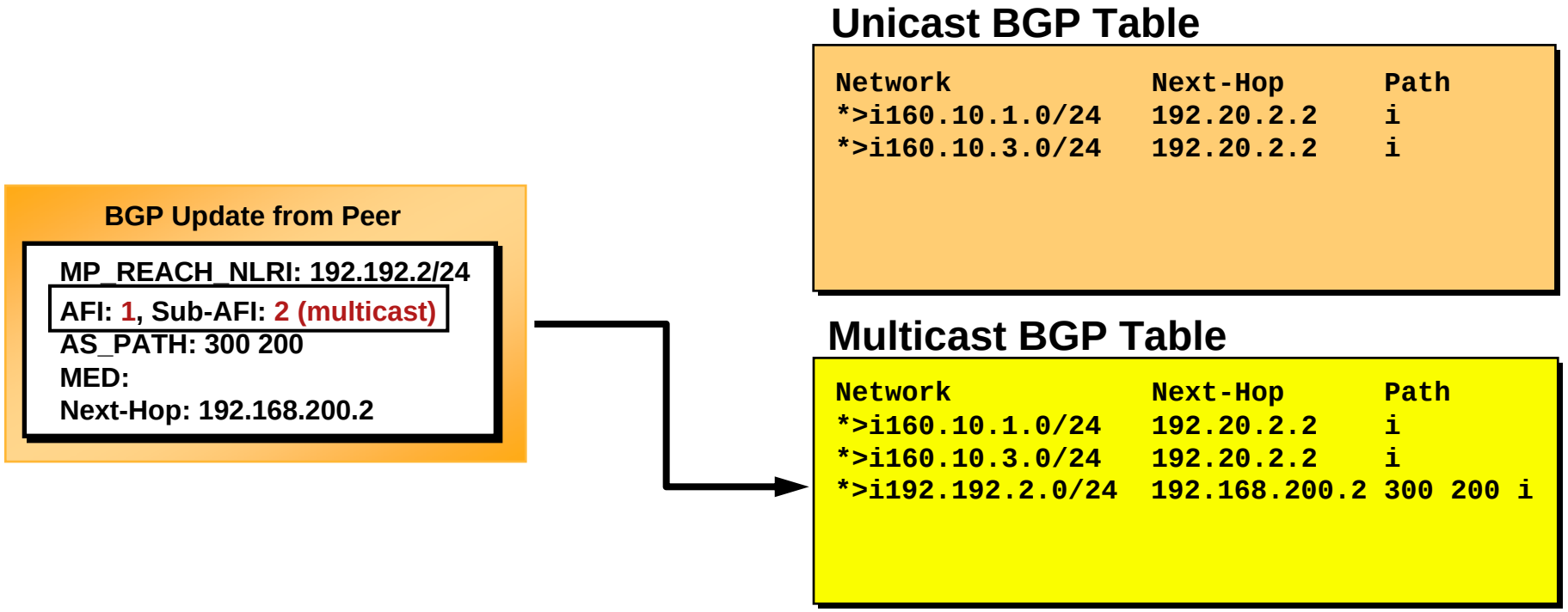
  If there is no match, notification is sent and peering doesn't come up

MBGP—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)**

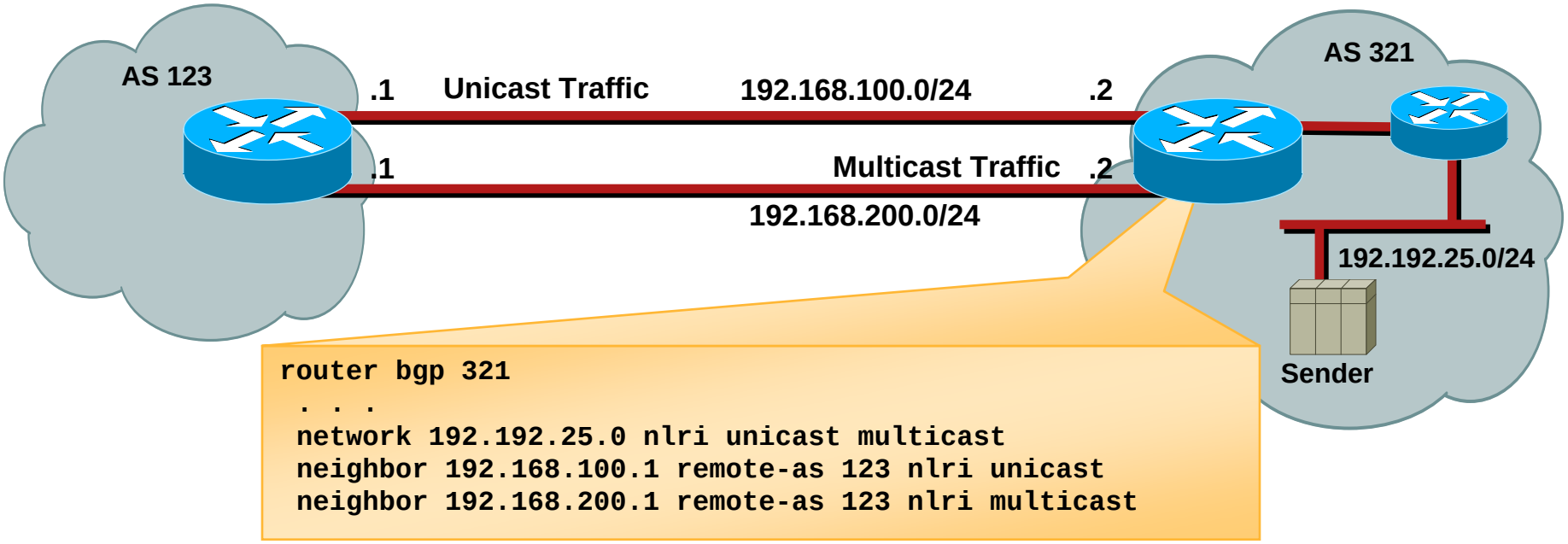
MBGP—NLRI Information



- 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

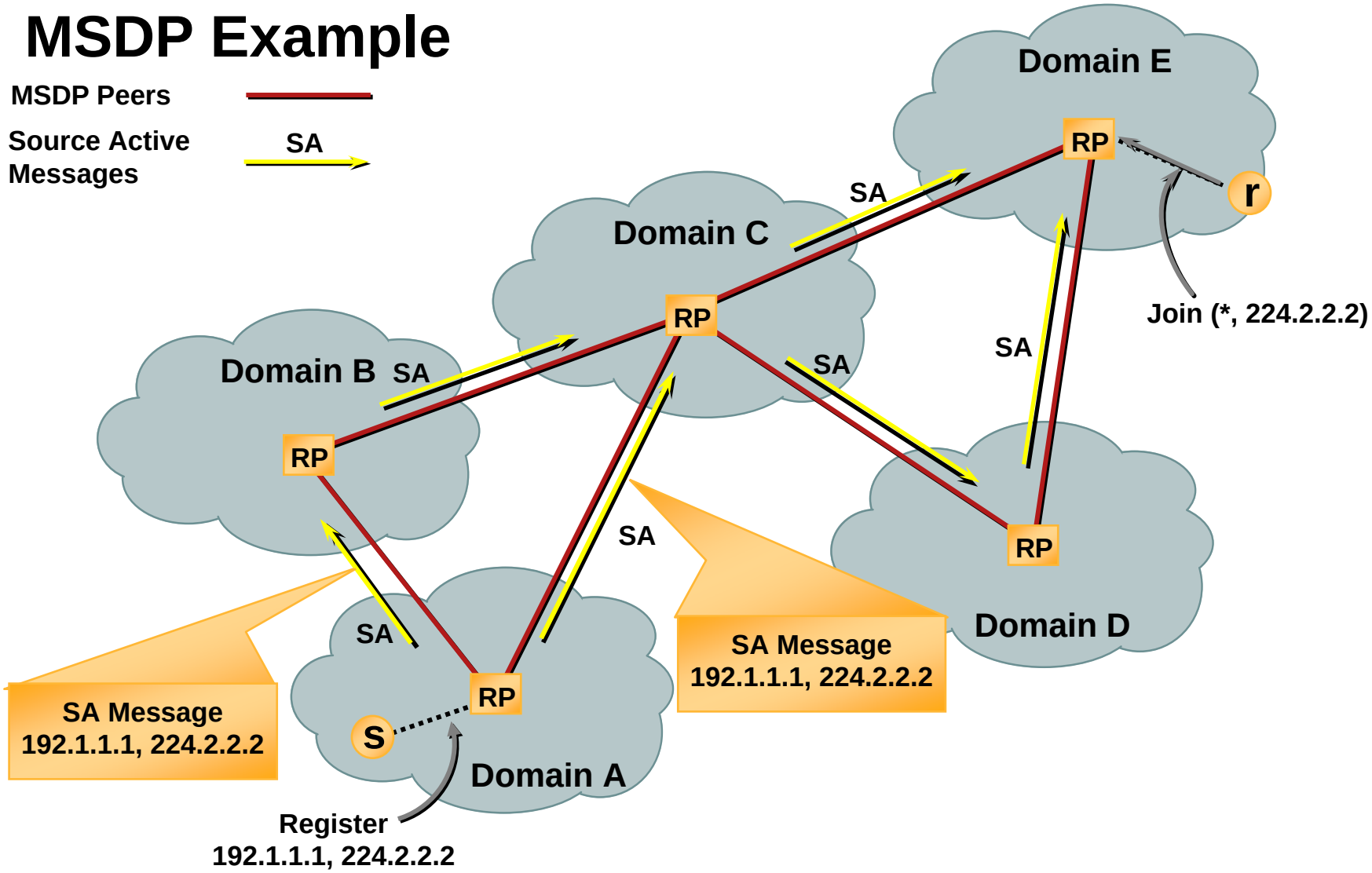
```
router bgp 5
  no bgp default ipv4-unicast
  neighbor 171.69.214.38 remote-as 2
  neighbor 171.69.214.50 remote-as 2
!
  address-family ipv4 unicast
  neighbor 171.69.214.38 activate
  network 171.69.214.0 mask 255.255.255.0
  exit-address-family
!
  address-family ipv4 multicast
  neighbor 171.69.214.50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

Agenda

- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

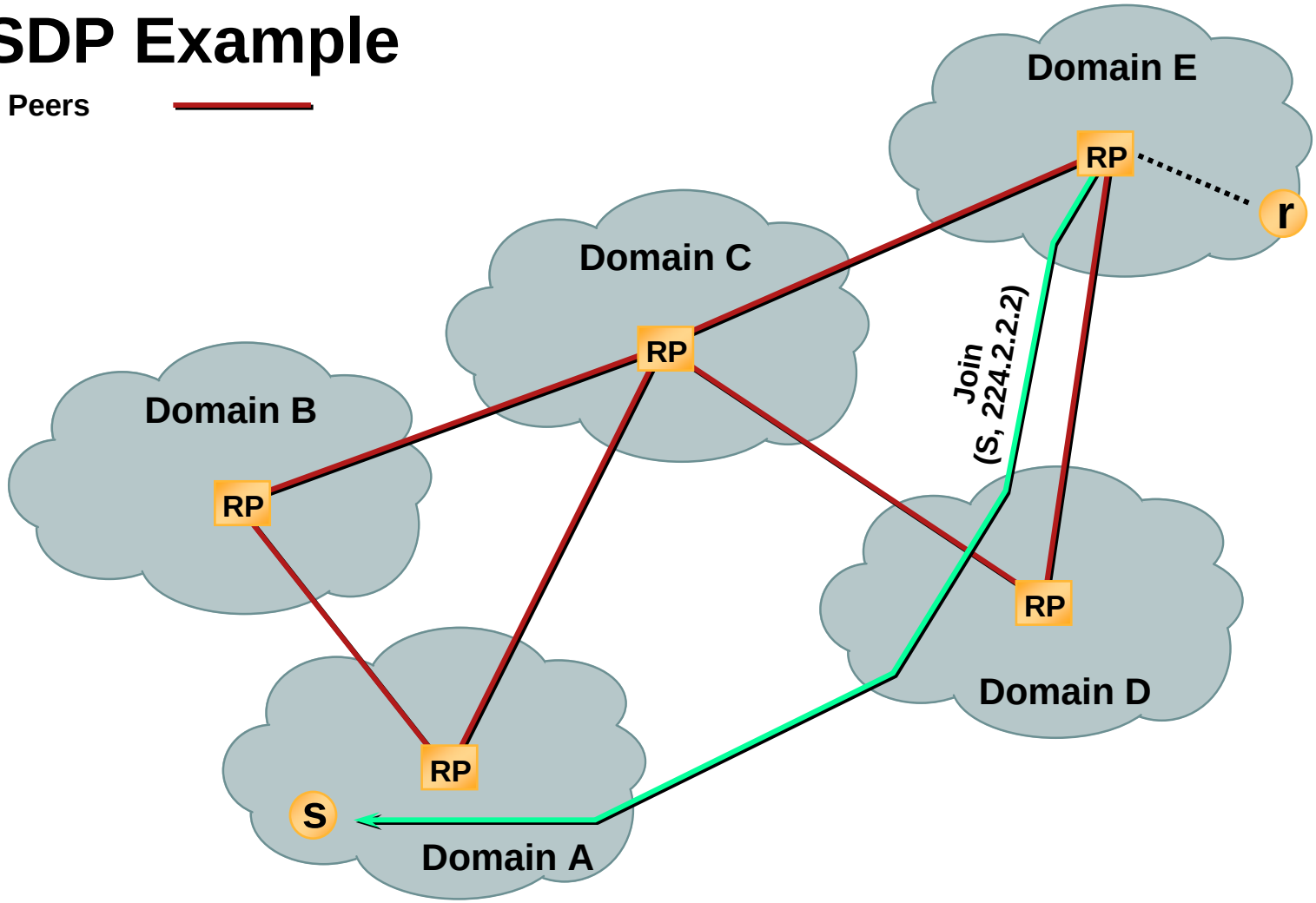
MSDP Overview



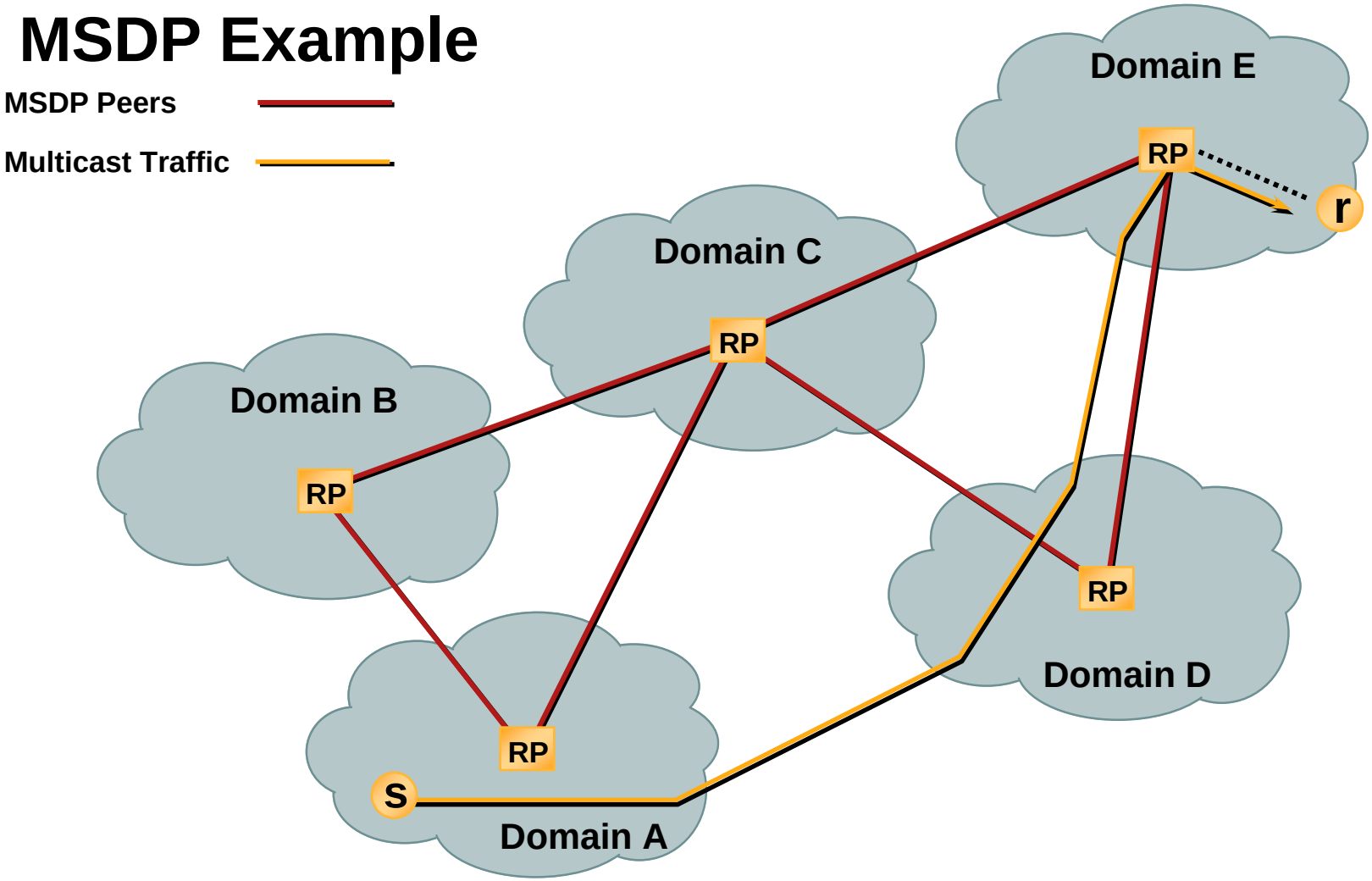
MSDP Overview

MSDP Example

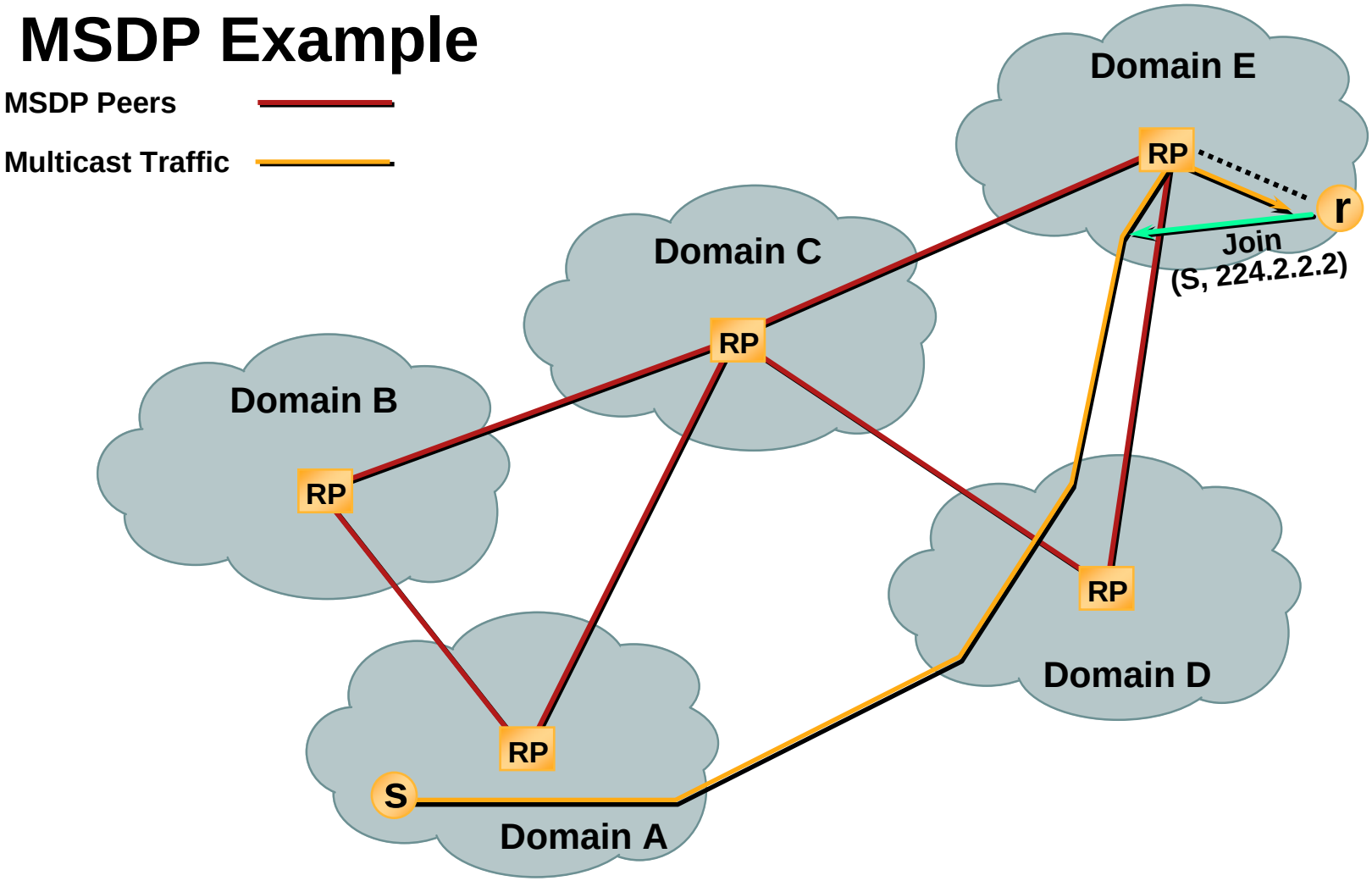
MSDP Peers



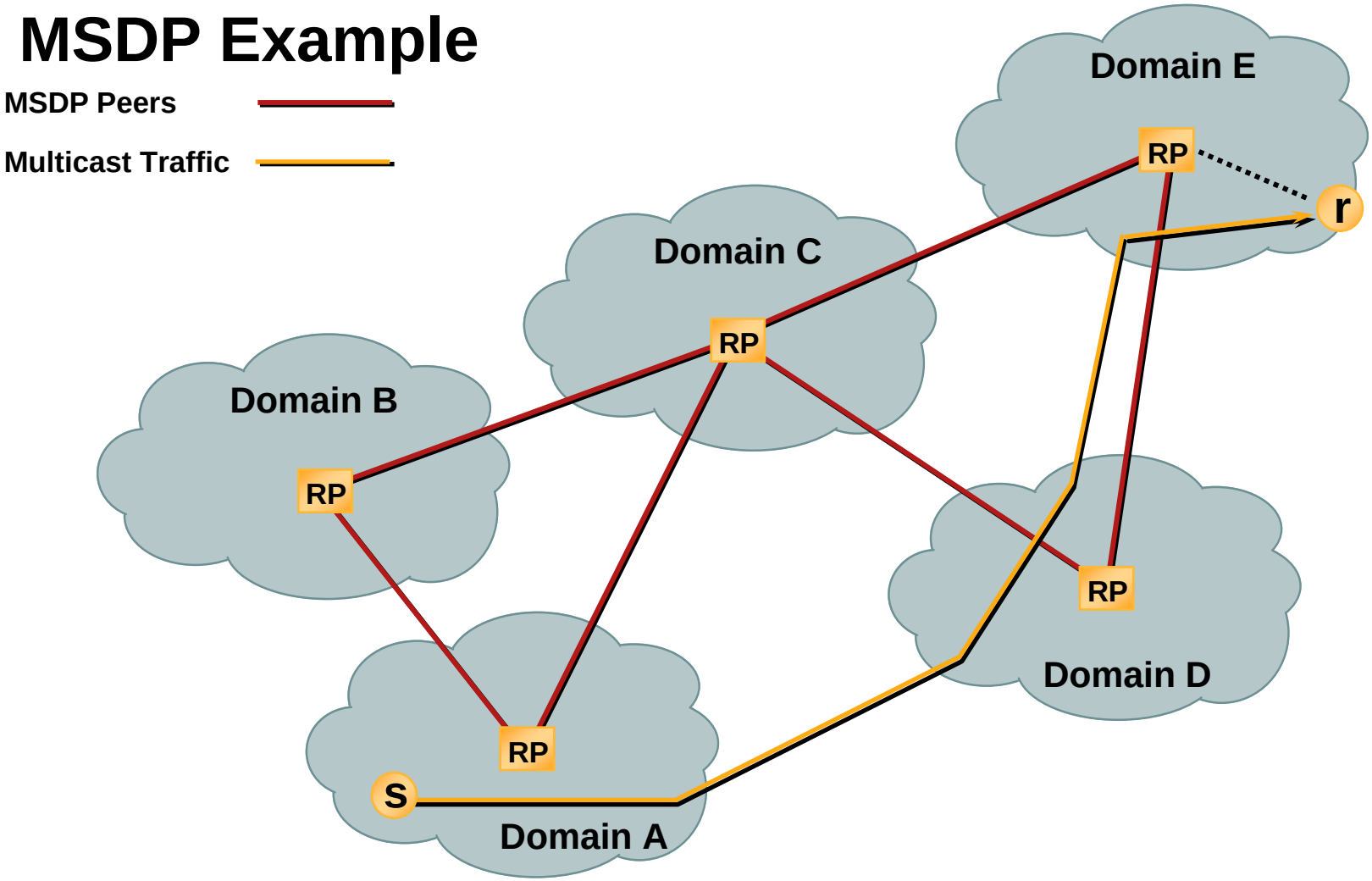
MSDP Overview



MSDP Overview



MSDP Overview



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

MSDP commands

- Filtering

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

- New IOS command

  ip msdp rpf rfc3618

  MSDP SA RPF check using IGP

  Accept SA's from BGP NEXT HOP or from closest peer along best path to originating RP

  show ip msdp rpf (12.0(27)S)

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

Agenda

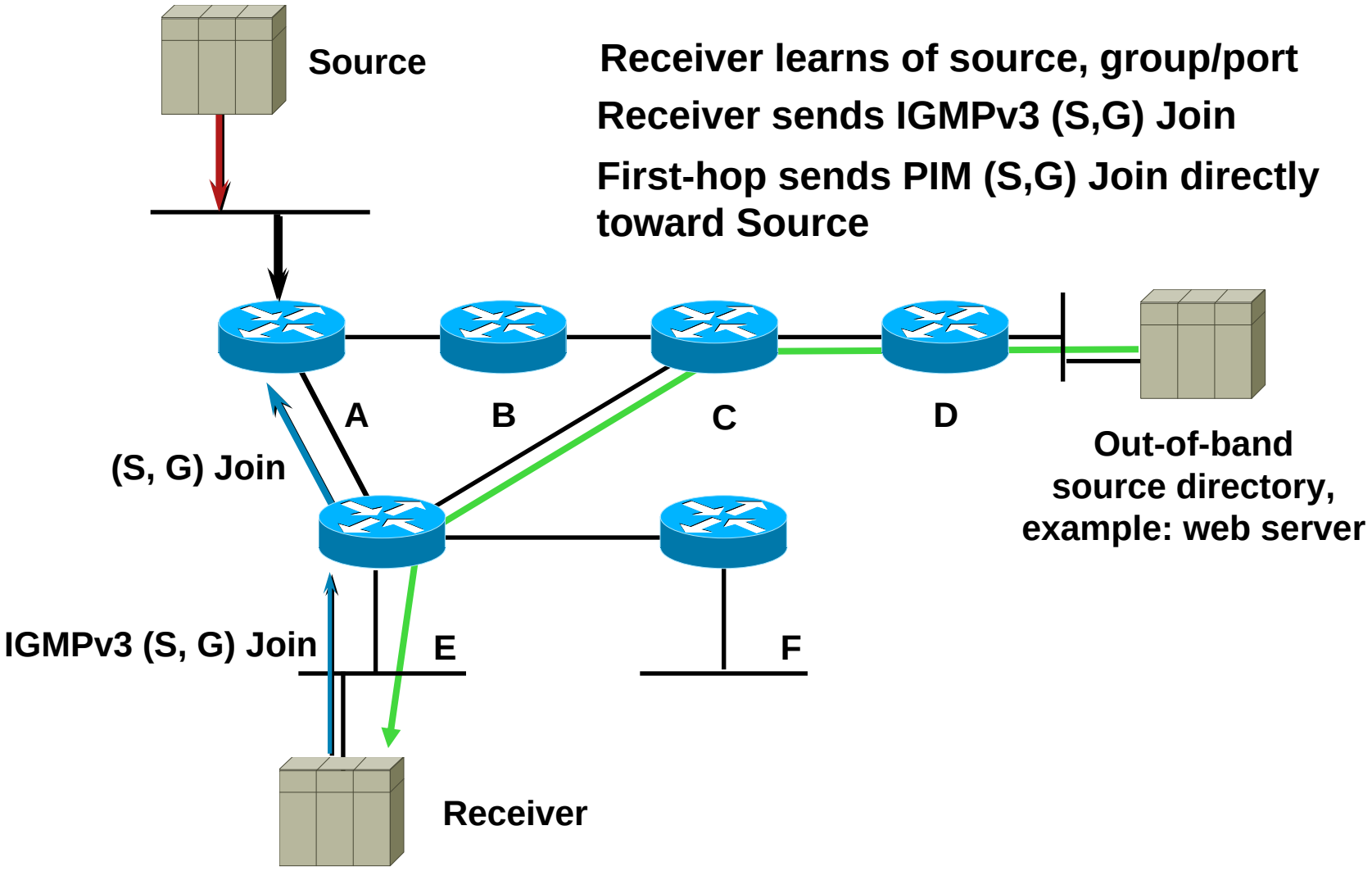
- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

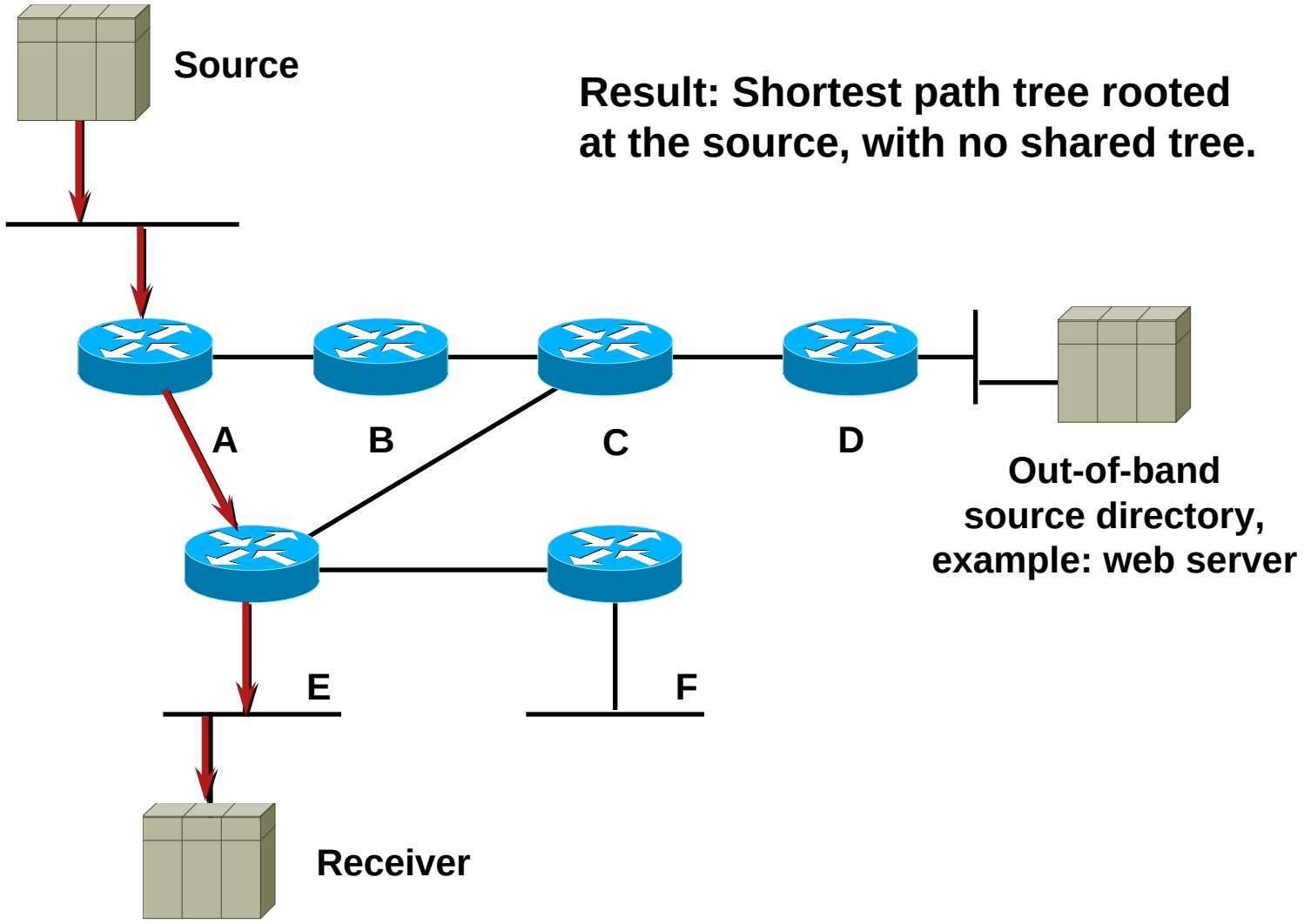
Source Specific Multicast (SSM)

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

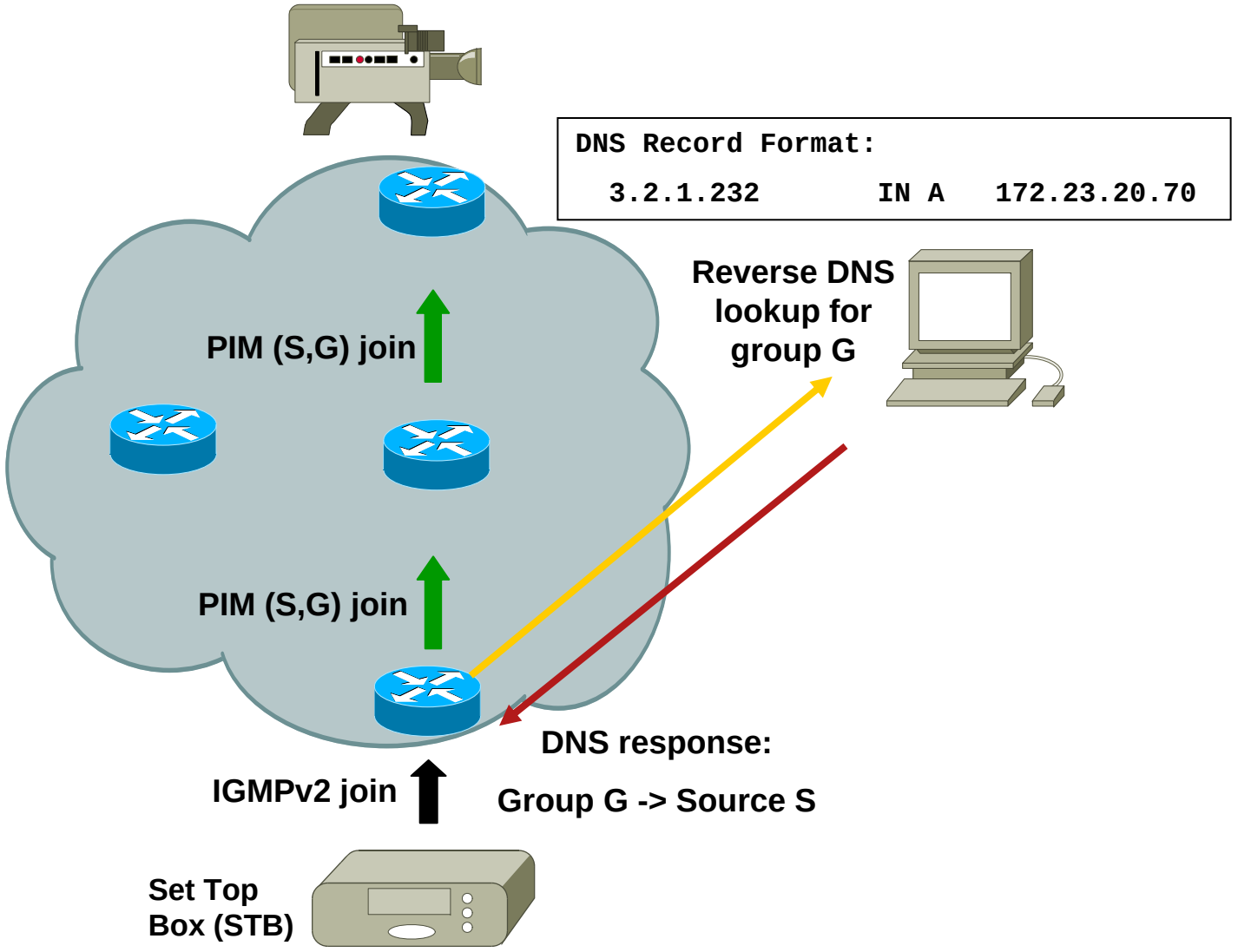
  Available starting in IOS versions

    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
- 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



SSM Mapping - 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
```

Agenda

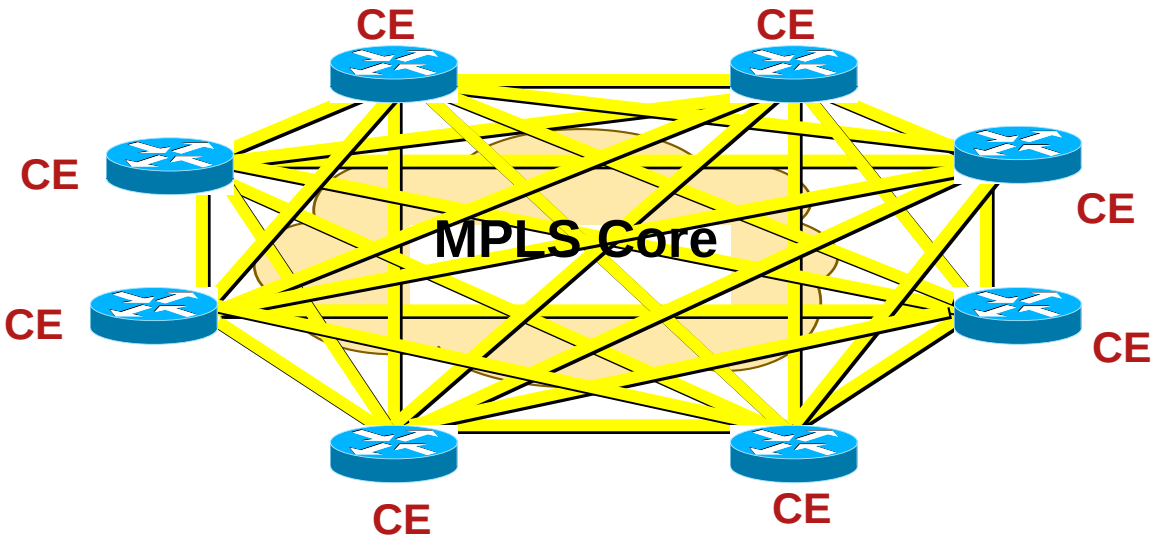
- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

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



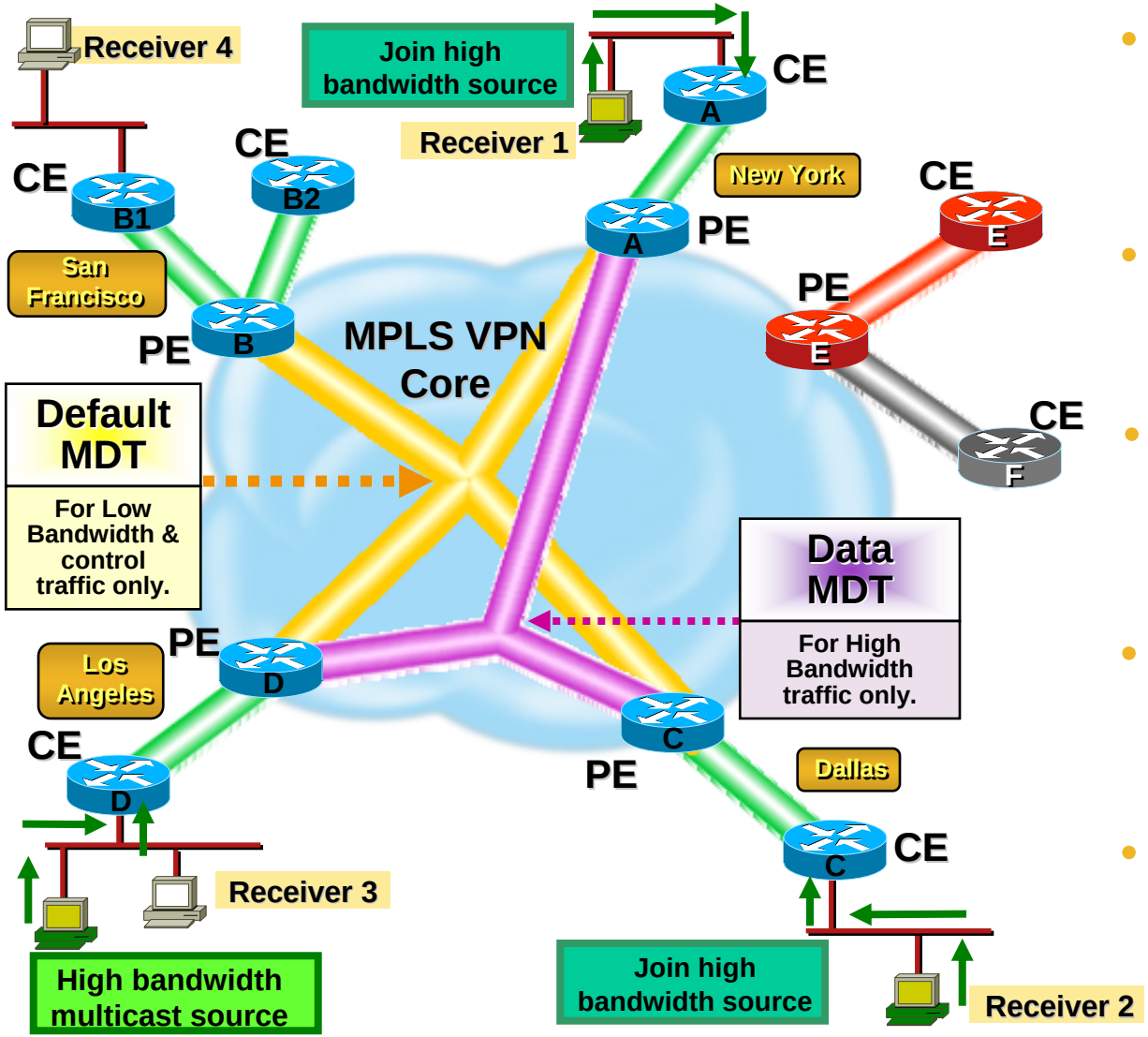
Multicast VPN – Requirements

- 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-ietf-l3vpn-2547bis-mcast-03.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.

mVPN : Concept & Fundamentals



- Customer CE devices joins the MPLS Core through provider's PE devices
- The MPLS Core forms a Default MDT for a given Customer
- A High Bandwidth source for that customer starts sending traffic
- Interested receivers 1 & 2 join that High Bandwidth source
- Data-MDT is formed for this High Bandwidth source

MVPN Concepts: Core facing side

- **Default-MDT:**

Emulate an ethernet between VRFs of an Intranet VPN.

- **No new protocols used! “**
- **”Ships in the night” with unicast MPLS/VPN**
  - **Relies on VPNv4 BGP reachability and VRFs from unicast. Nothing else!**
- **VRF constitute PIM routers on this LAN**
- **Uses GRE-Multicast encapsulation (today)**
  - **Leverages IPv4 Multicast across core – no label switching**
  - **All PE in VPN are sender and receiver**
  - **Each VPN uses on “Default-MDT group” in core.**
  - **Common deployment: PIM-SM (12.0S limitation)**

MVPN Concepts: Customer facing side

- **Multicast/PIM for VRFs**
  - **Each VRF runs an instance of PIM**
    - **Currently all implement in same process in IOS**
  - **Choice of PIM modes independent of Core PIM-Mode (MDT)**
    - **Customer can use: SSM, ASM (PIM-SM, Bidir-PIM, PIM-DM).**
- **Features available on PE-CE connection determined by "VRF-aware features" in PE software.**
  - **Most supported.**
  - **Example exceptions: DVMRP (retired), BGP-SAFI2 (TBD)**

MVPN Extensions - SSM for Default-MDT

- **Requires new BGP message exchanged between PE:**

  PEs indicate in BGP Default-MDT groups Gi configured

  PE subscribe to SSM channels (Sj, Gi), where Sj is another PEs BGP announcement for Gi.

  Before 12.0(29)S:

  Use extended community BGP message - no config

  Since 12.0(29)S:

  IPv4/MDT address/subaddress family reachability.

  Requires config of MDT-SAFI per neighbor

- **Requires Default-MDT groups to be in SSM range**

MVPN Extensions - Data-MDT

- **Resolves problem of wasted bandwidth to PE**
- **Dedicated Multicast tree from headend PE to set of tailend PEs**
  - Requires configuration of Data-MDT groups (ideally SSM)
- **Headend-PE determines traffic to put on Data-MDT**
  - Headend PE puts traffic onto Data-MDT dynamically based on bandwidth
  - 256 Data-MDT per Headend-VRF limit with re-use
  - Other mappings possible in future (eg: statically configured)
- **Uses UDP message to signal mapping to Data-MDT**
  - UDP messages multicast on Default-MDT
- **Can only place (S,G) traffic onto Data-MDT**

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 support

- **Resolves two RPF-issues**
- **Build (Sj,Gi) tree across P node where Sj is PE from remote AS not redistributed into local IGP**

  Problem with InterAS option B,C without IGP route redistribute between AS

  Resolved with InterAS-RPF-Vector

  Adds (RD, ASBR) to PIM (S,G) joins. P node RPFs towards ASBR, ASBR find next ASBR (or ultimate PE) by (RD,S) BGP lookup.

  Note: RPF-Vector (without InterAS) for BGP-free core of native IP multicast (eg: for IPTV services).

- **Send PIM join on Default-MDT to headend PE if PE is in other AS and BGP nexthop is rewritten**

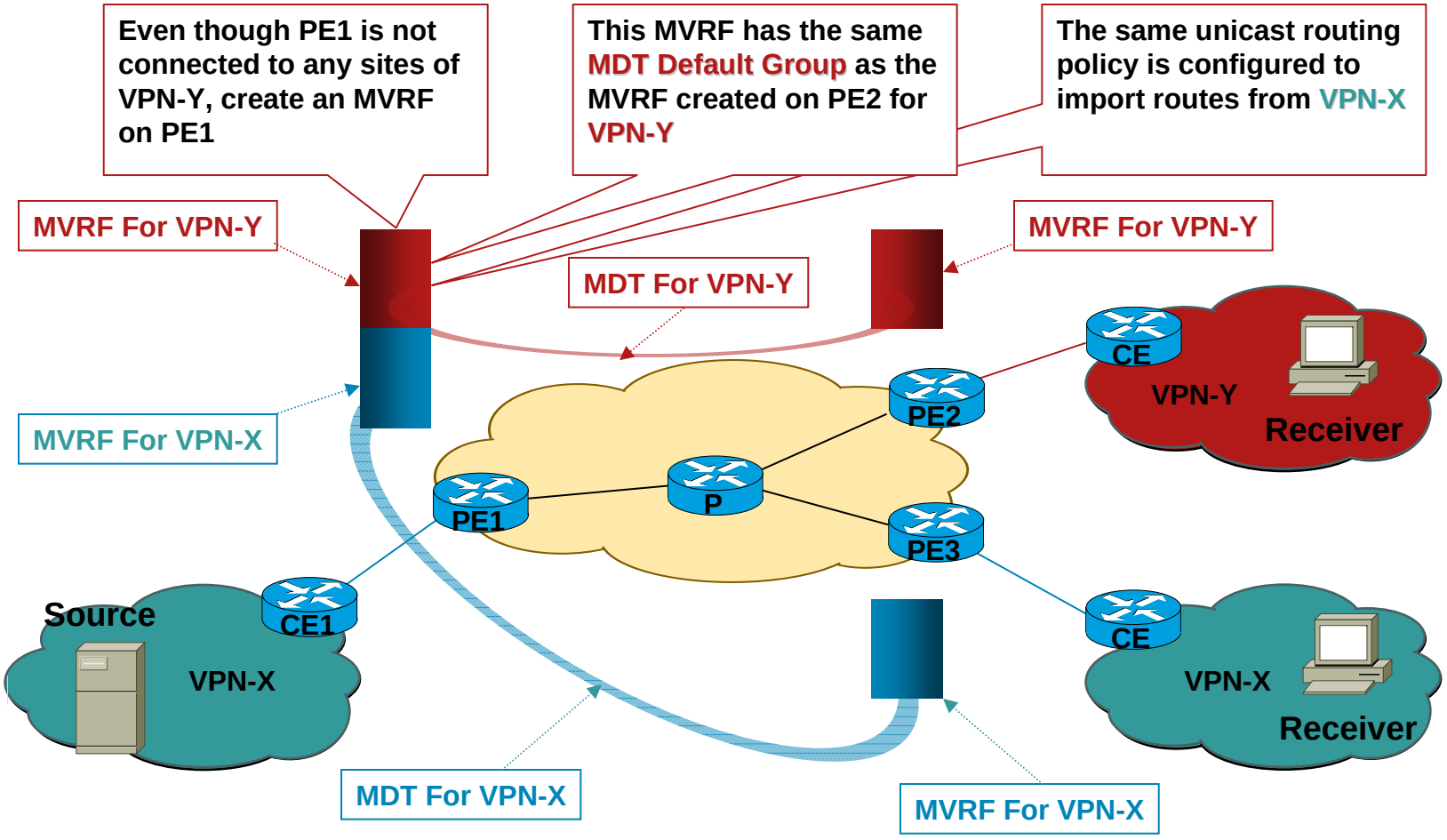
  Problem with InterAS option B with ASBR nexthop-self and option C.

  Resolved with BGP connector-attribute. Adds originator PEs (RD,S) information to VPNv4 BGP reachability of source prefix.

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
- 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



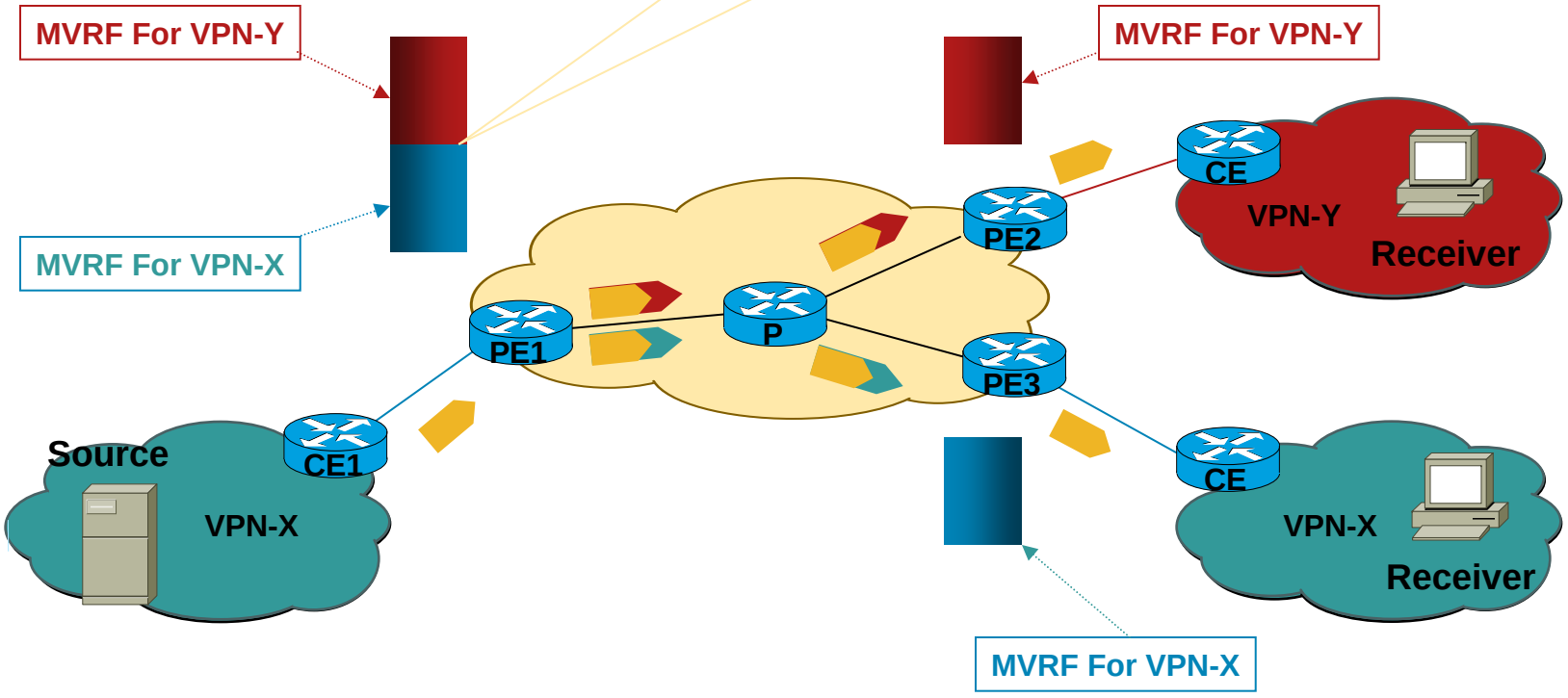
Extranet MVPN Configuration Option #1

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



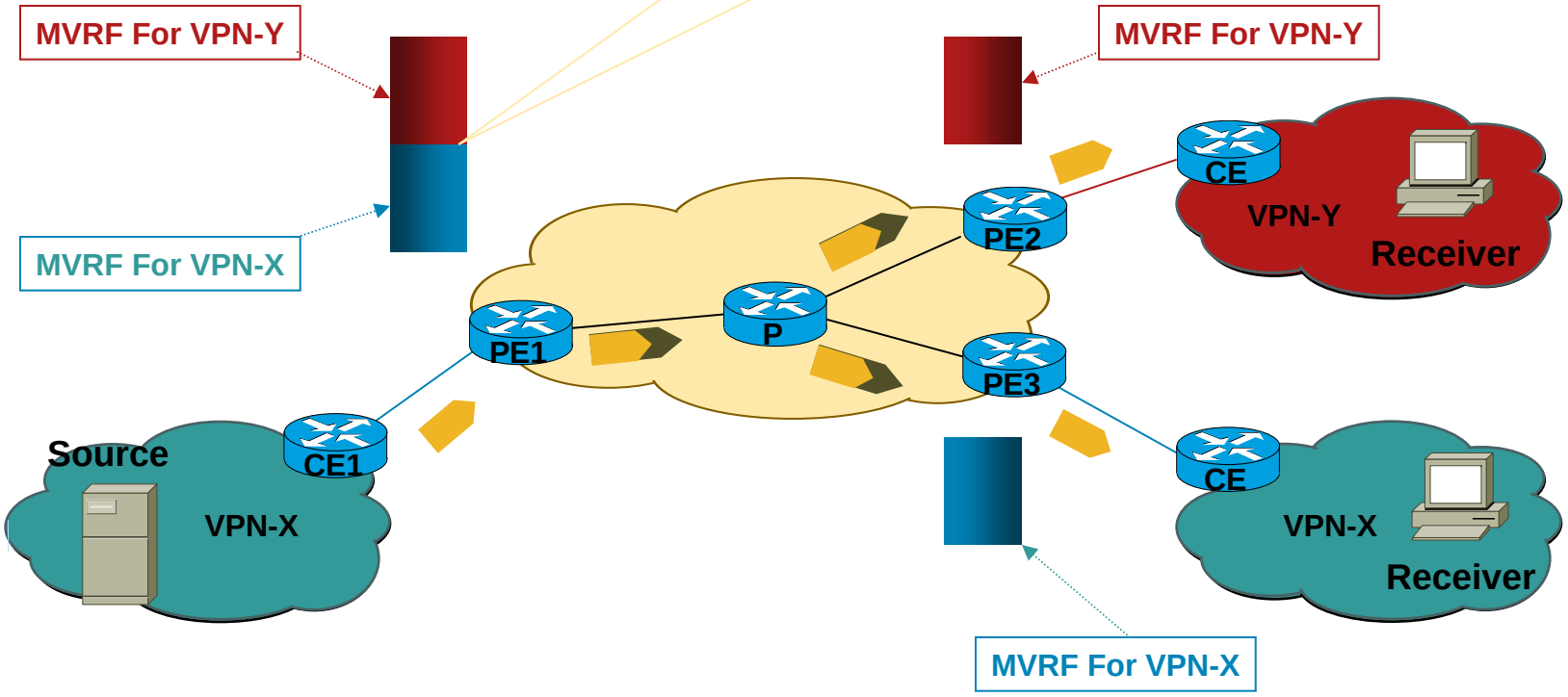
Extranet MVPN Configuration Option #1

Using a Common MDT Data Group

PE1 can optionally choose to use to same MDT Data Group to encapsulate packets

The result is that packets are only replicated once in the core independent of the number of different receiver MVRFs

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

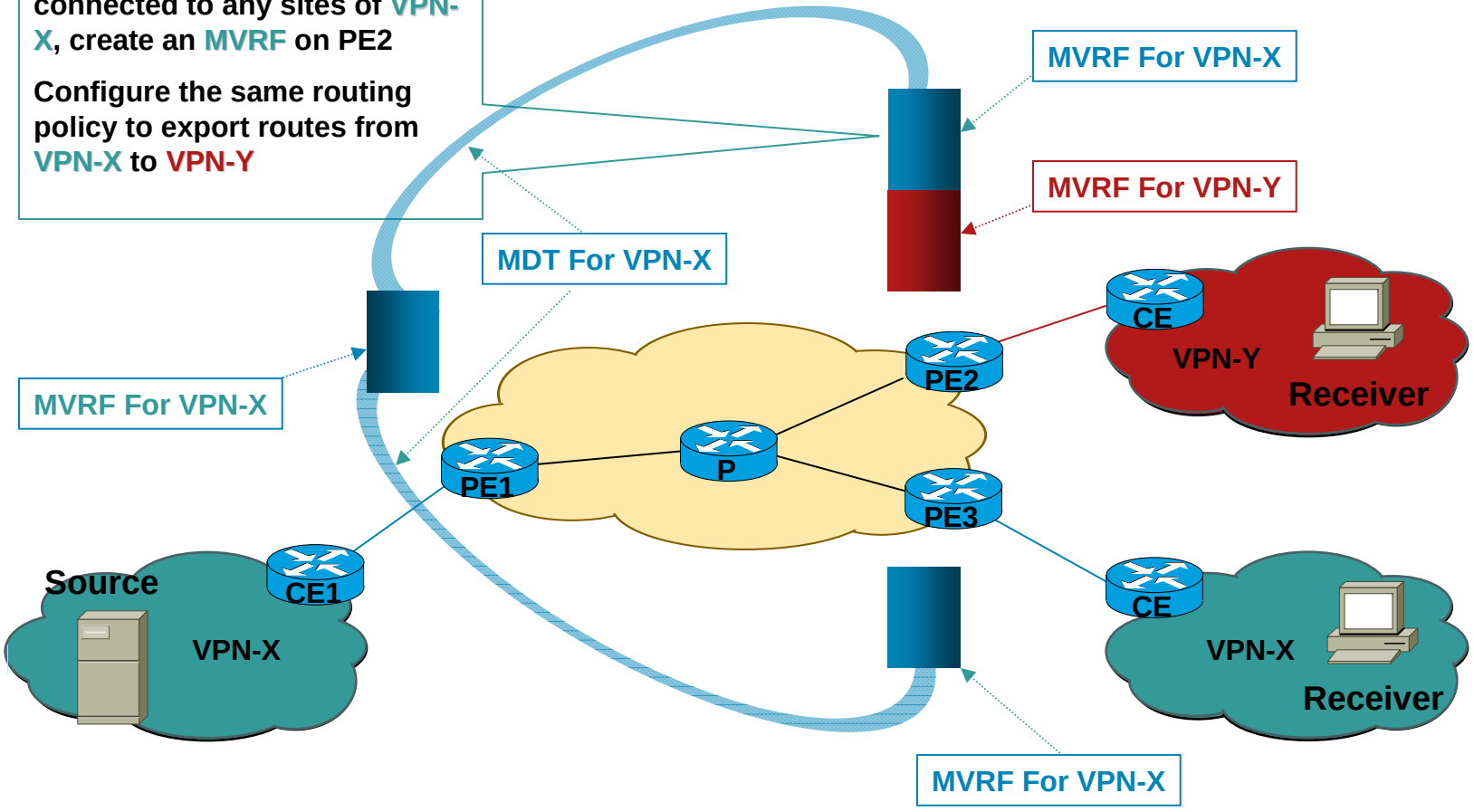


Extranet MVPN Configuration Option #2

Configuration

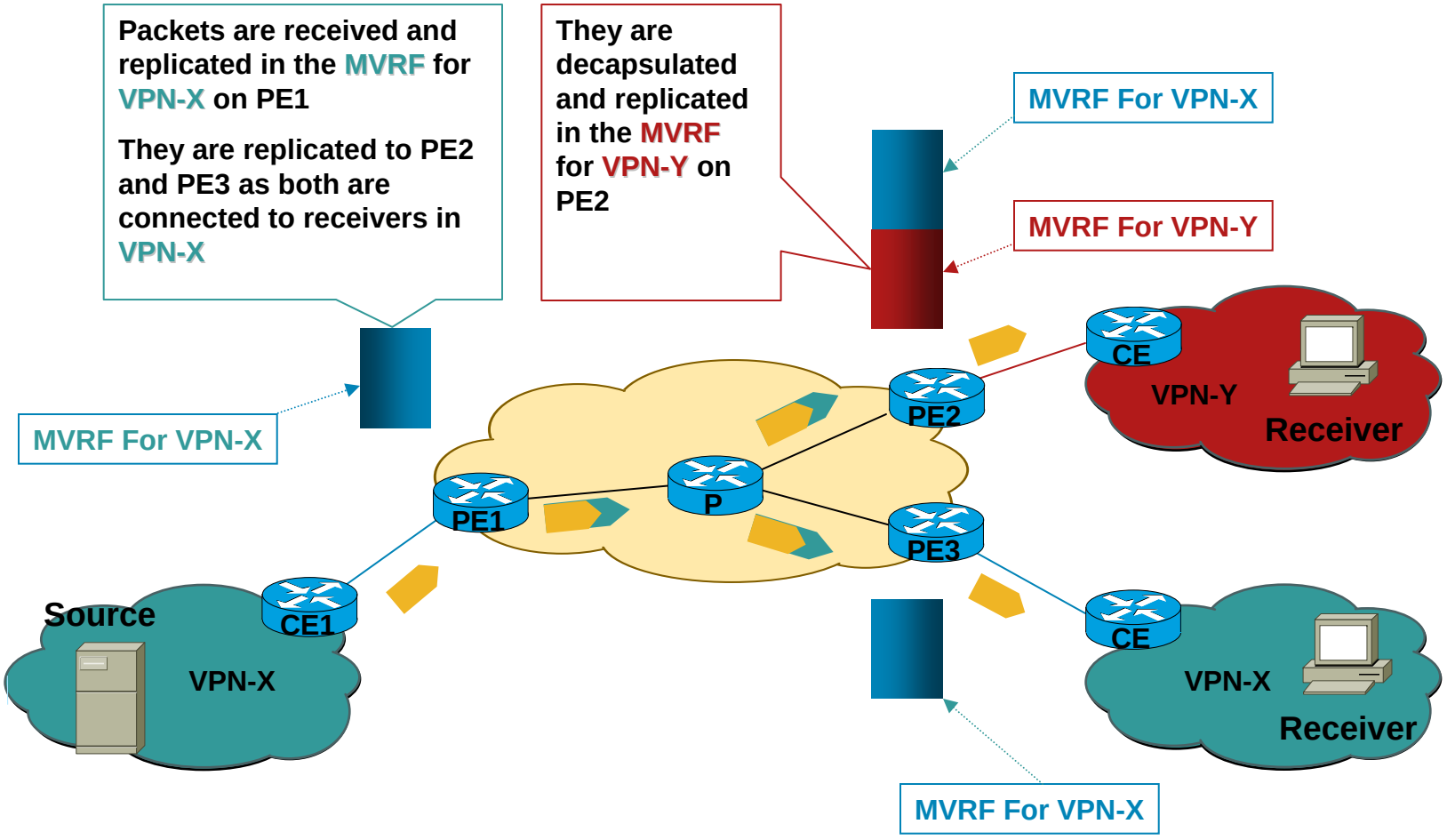
Even though PE2 is not connected to any sites of VPN-X, create an MVRF on PE2

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



Extranet MVPN Configuration Option #2

Packet Flow



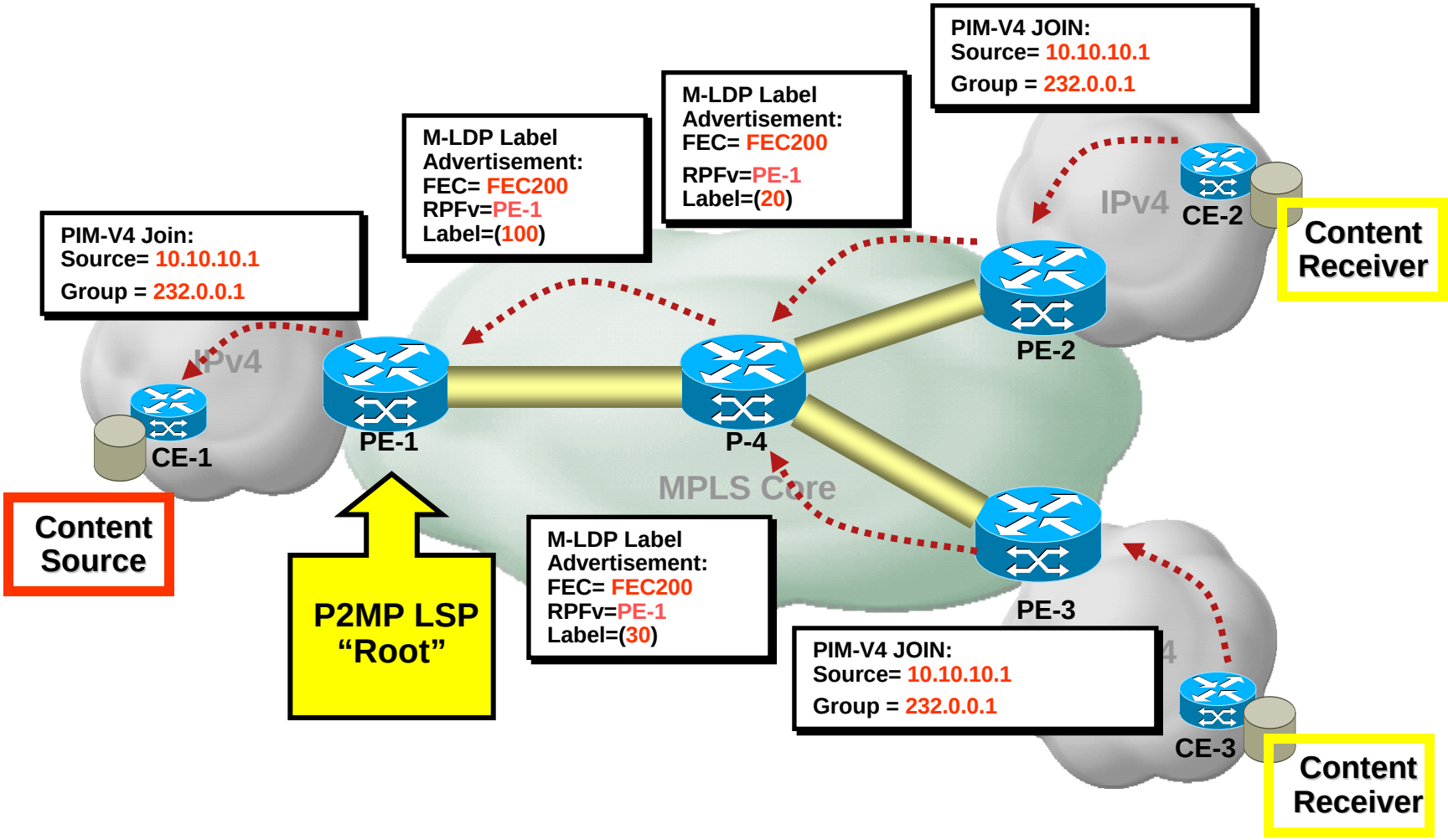
LSM framework

- **Four Multicast services:**
  - **Label Based IPv4 Multicast VPN services**
  - **Point to Multi Point IPv4 RSVP-TE Multicast services**
  - **Label Based IPv6 Multicast VPN services**
  - **Point to Multi Point IPv6 RSVP-TE Multicast services**
- **Three new protocols:**
  - **Multcast extension to LDP (mLDP) protocol**
  - **Multicast extension to RSVP-TE protocol**
  - **Multicast overlay Signaling Protocol for VPN and state aggregation**

Cisco Status – LSM

LSM Protocols	Distinct properties
MLDP **draft-ietf-mpls-ldp-p2mp-02.txt**	**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**	**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.***

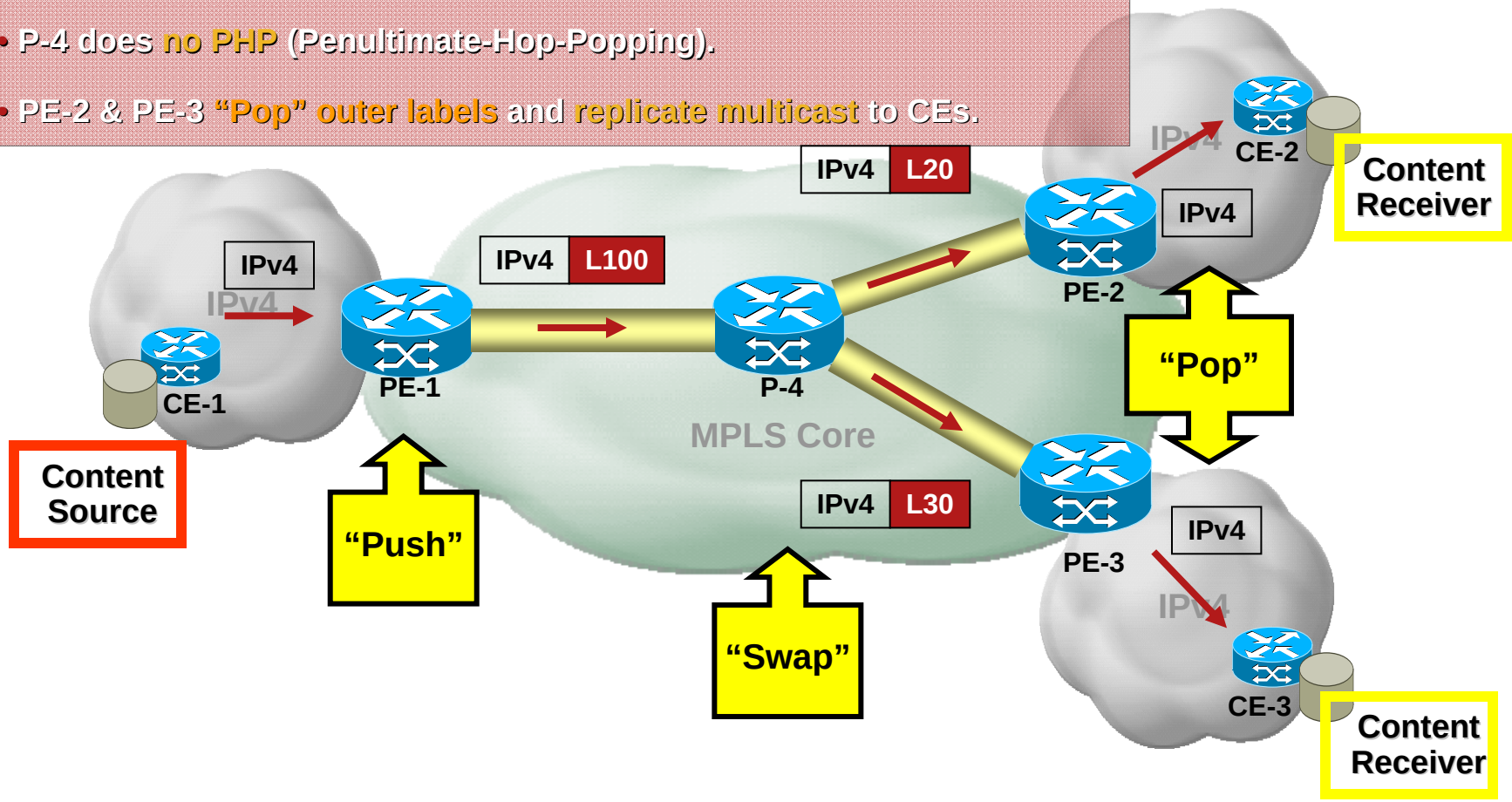
MLDP : Transiting PIM SSM (IPv4 non-VPN)



mLDP : Transiting PIM SSM (IPv4 non-VPN)

"Life of a Packet"

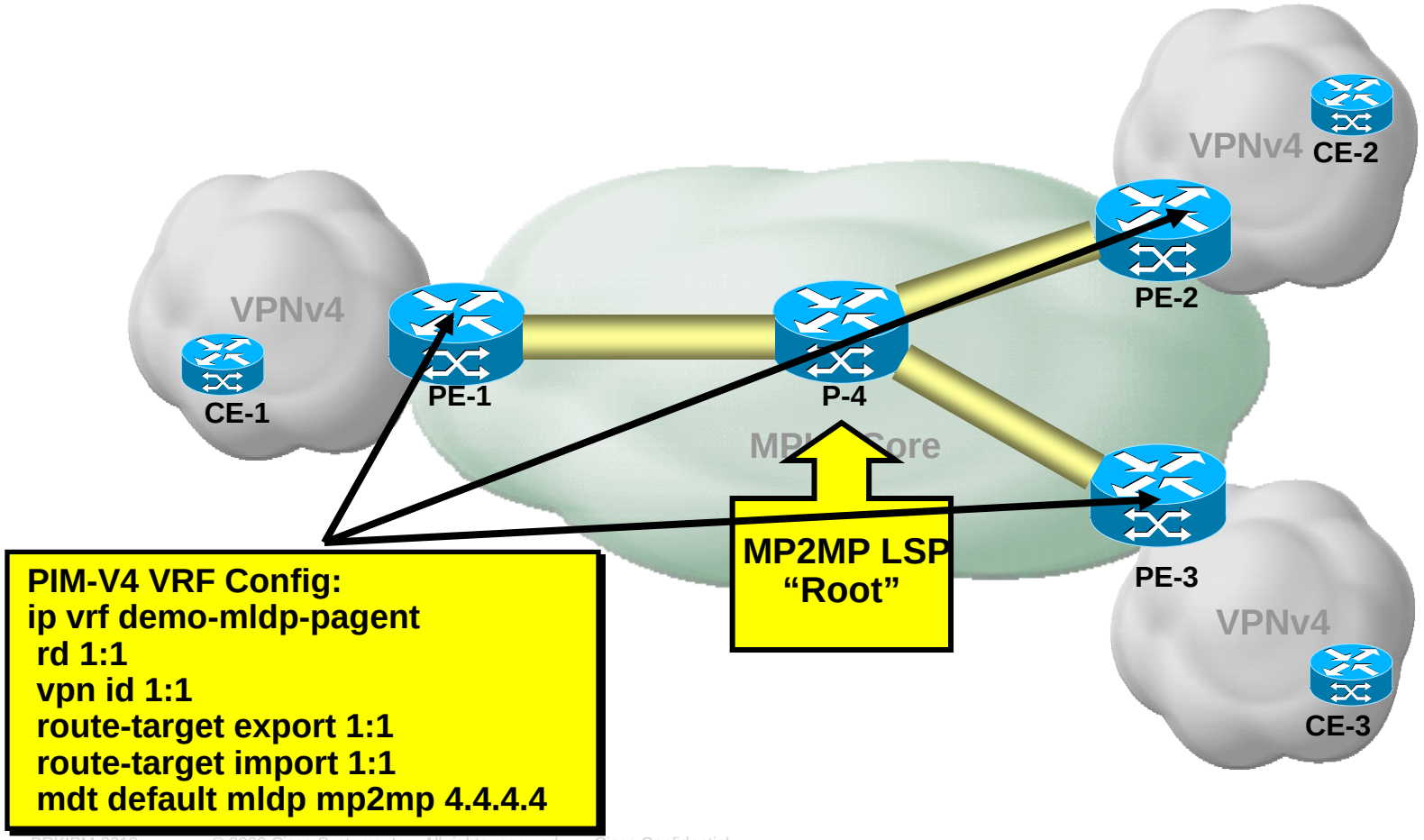
- PE-1 (Root) receives IPv4 multicast and "Pushes" label towards P-4.
- P-4 "Swaps" and replicates labels to PE-2 & PE-3.
- P-4 does no PHP (Penultimate-Hop-Popping).
- PE-2 & PE-3 "Pop" outer labels and replicate multicast to CEs.



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

MP2MP Tree Setup Summary

- All PE's configured for same VRF derive FEC from configured MDT type, VPN id and Root ID



MLDP : MPLS mapping

```
MLDP-SEVT-PE1#sh mpls mldp db
 System ID           : 97000000
 FEC tree type       : MP2MP
 FEC length          : 24 bytes
 FEC value           :  00000004 00000006 00000100 00000100 00000000 04040404
 FEC Root            : 4.4.4.4  (we are leaf)
 Root metric         : 20
 Root distance       : 115
 Opaque decoded      : VPN ID: 1:1 MDT Nr: 0
 Next Hop interface : Ethernet1/0
 Next Hop address    : 11.11.11.4
 Upstream peer ID    : 4.4.4.4:0
 Local down label    : 21
 Upstream label      : 20
 Root Node Redundancy : enabled : primary root - 4.4.4.4  Replication client(s):
 MDT              interface: Lspvif0 (vrf demo-mldp-pagent)
```

MLDP : Mapping to the FEC

```
PIM-V4 VRF Config:
ip vrf demo-mldp-pagent
 rd 1:1
 vpn id 1:1
 route-target export 1:1
 route-target import 1:1
 mdt default mldp mp2mp
4.4.4.4
```

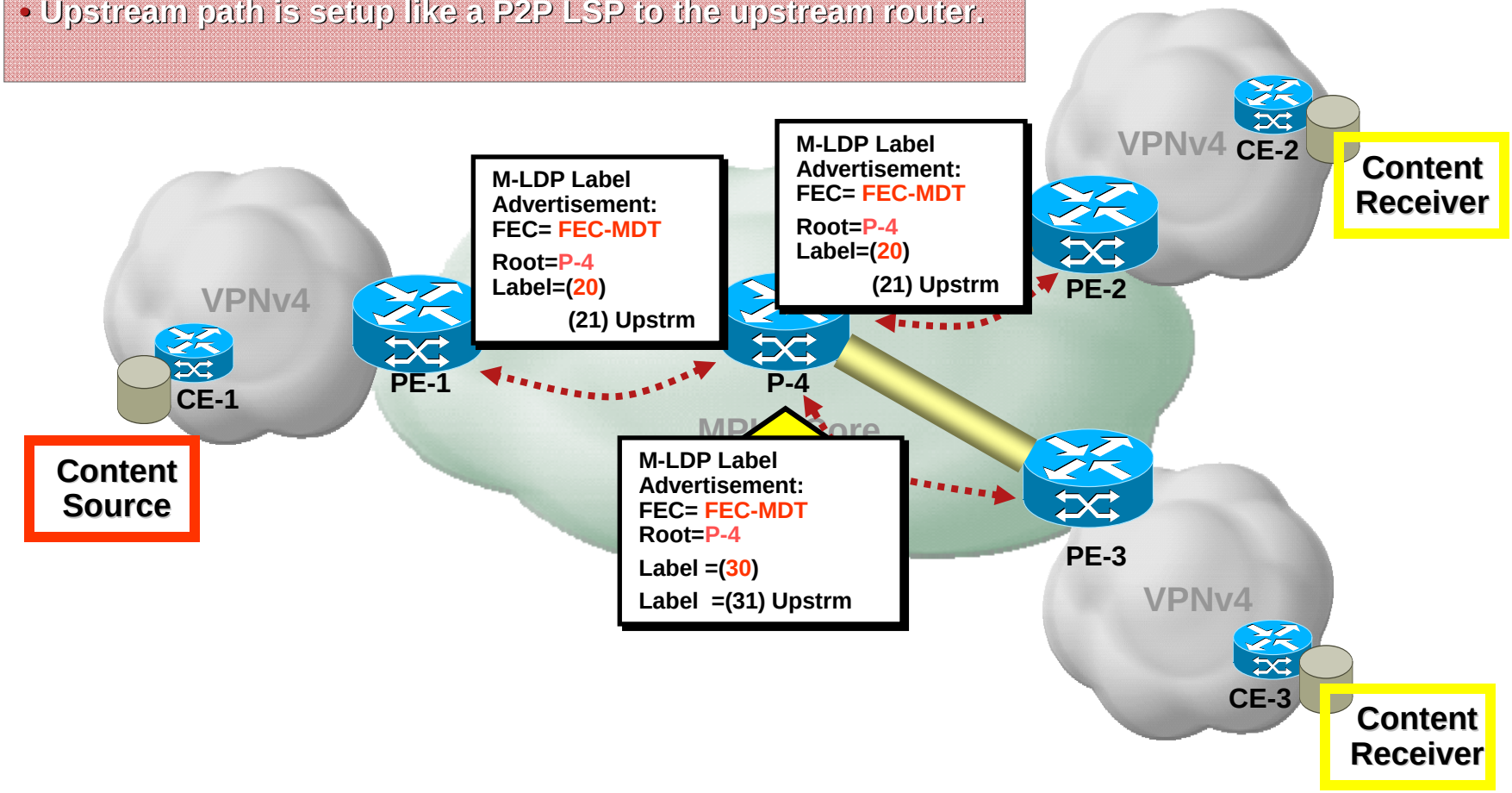
```
FEC value            :  00000004 00000006 00000100 00000100 00000000 04040404
```

- FEC Type MPMP (00000004 00000006)
- VPN ID 1.1 (00000100 00000100)
- FEC Root 4.4.4.4 (04040404)

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

MP2MP Tree Setup Summary

- Downstream path is setup like a normal P2MP LSP.
- Upstream path is setup like a P2P LSP to the upstream router.



P2MP RSVP TE

- Extend RSVP-TE to establish P2MP-LSPs

  Focus on TE requirements for relatively static P2MP-LSP topologies

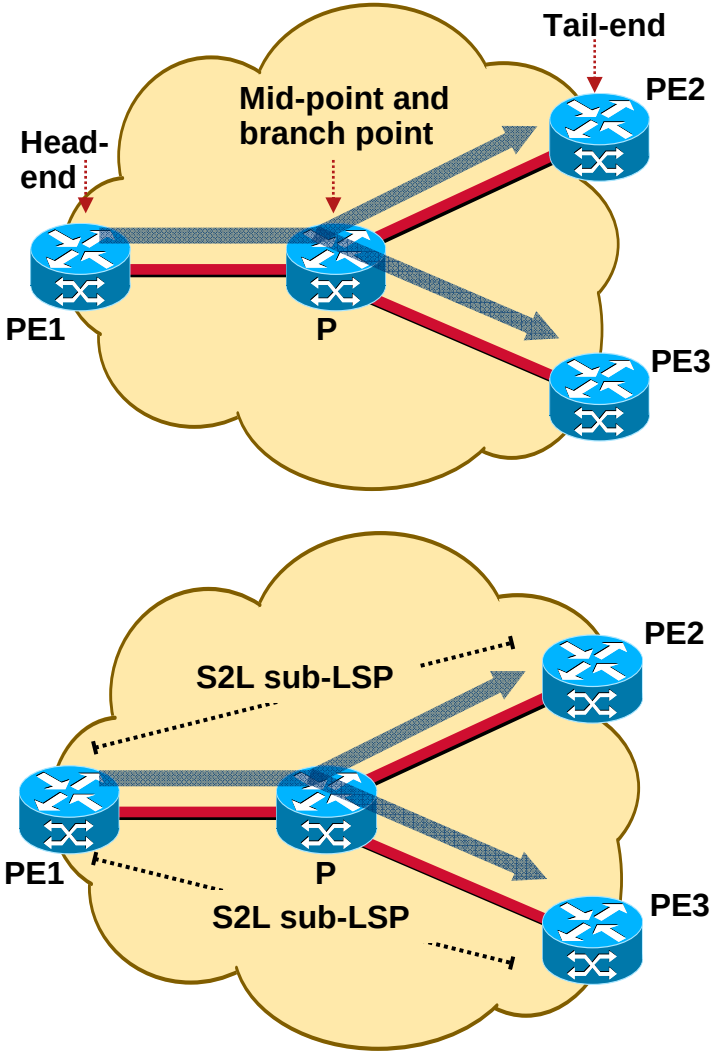
- IETF proposals are converging on RSVP-TE draft
- RSVP P2MP draft:

  draft-ietf-mpls-rsvp-te-p2mp-01.txt

- Requirements drafts

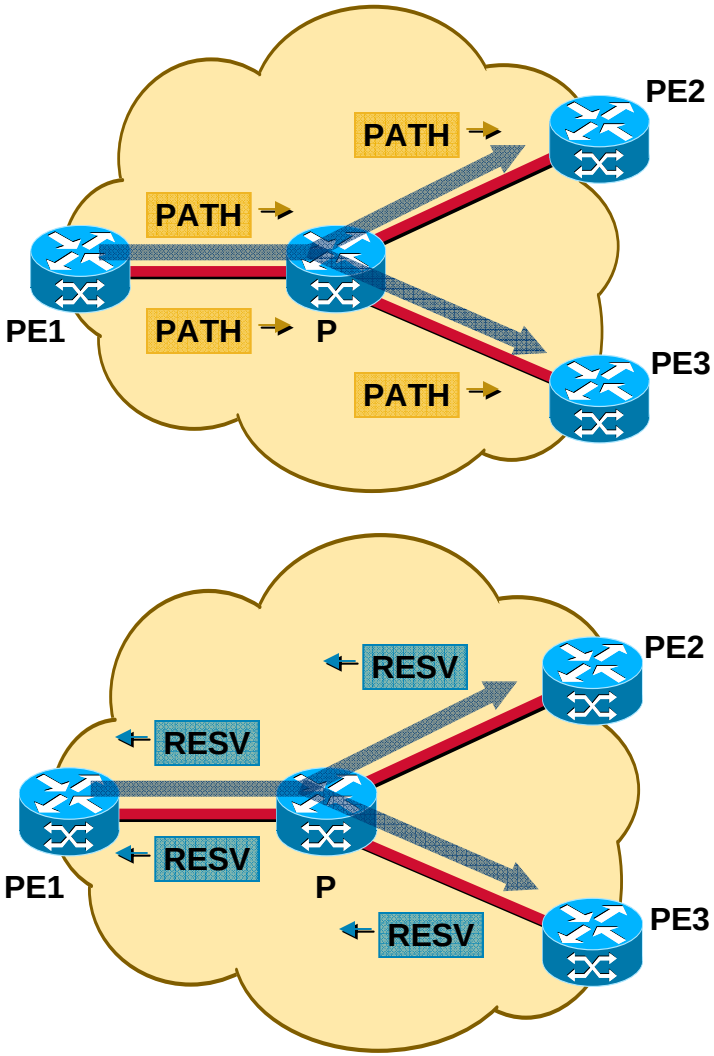
  draft-ietf-l3vpn-ppvpn-mcast-reqts-00.txt

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

Agenda

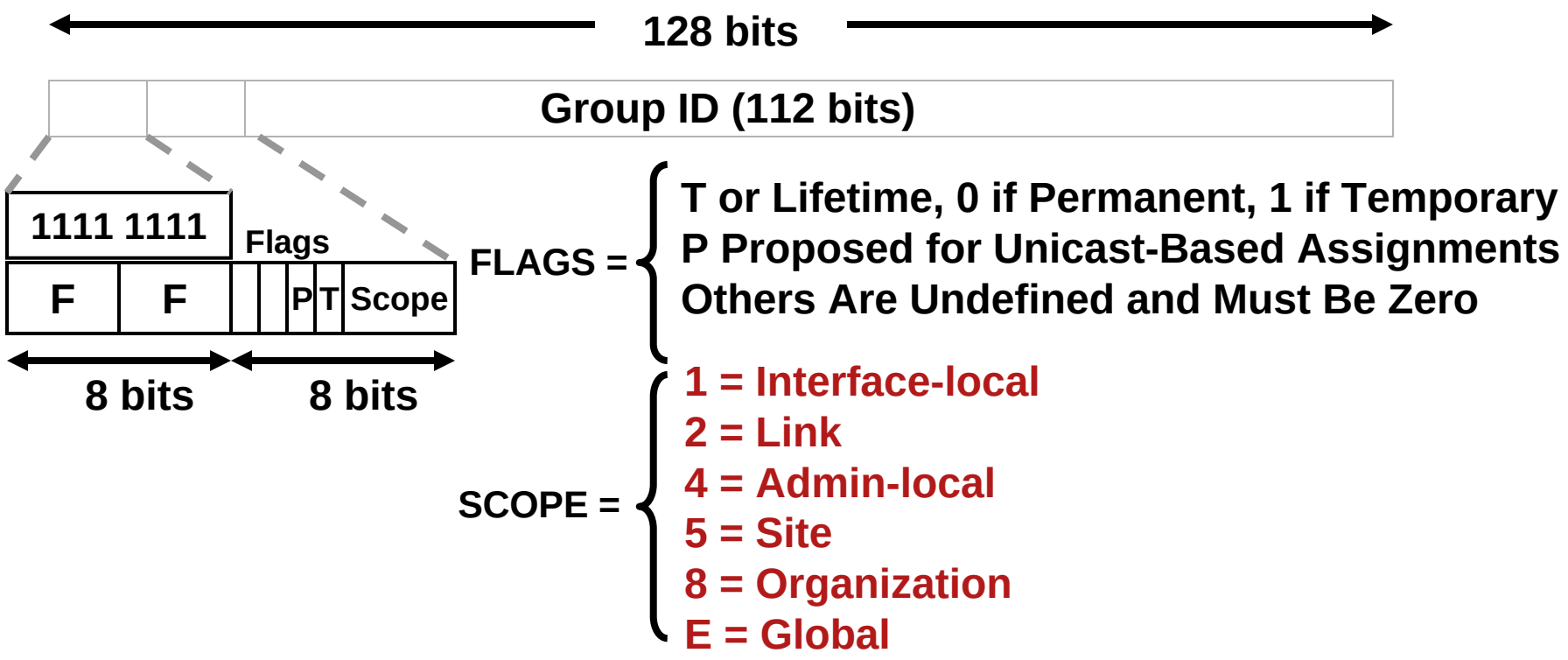
- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

IPv4 and IPv6 Multicast Comparison

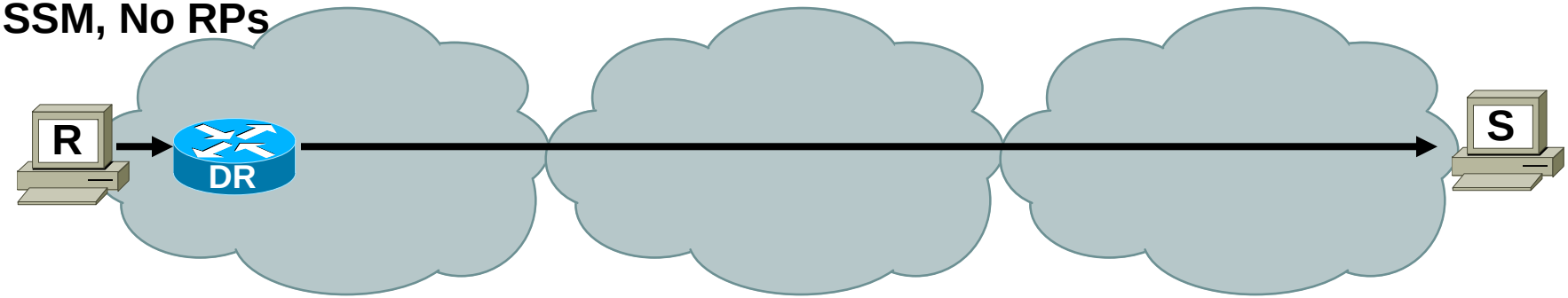
Service	IPv4 Solution	IPv6 Solution
Addressing Range	32-bit, Class D	128-bit (112-bit Group)
Routing	Protocol Independent, All IGPs and MBGP	Protocol Independent, All IGPs and MBGP with v6 mcast SAFI
Forwarding	PIM-DM, PIM-SM, PIM-SSM, PIM-bidir	PIM-SM, PIM-SSM, PIM-bidir
Group Management	IGMPv1, v2, 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)



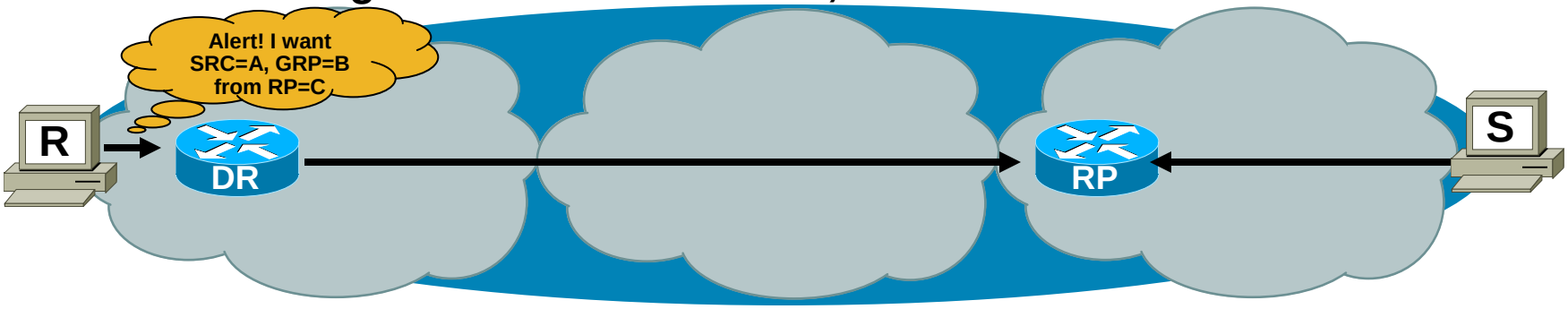
Multicast Deployment Options
With and Without Rendezvous Points (RP)

SSM, No RPs



ASM Single RP – Static definitions

ASM Across Single Shared PIM Domain, One RP – Embedded-RP

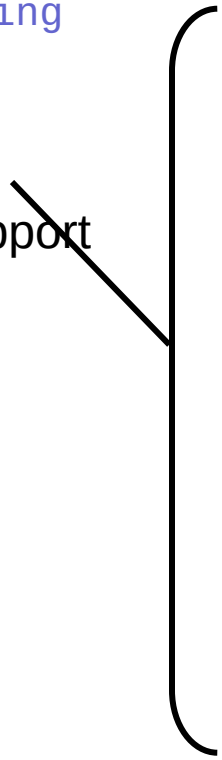


Source Specific Multicast (SSM)

- NO configuration required other than enabling

  `ipv6 multicast-routing`

- SSM group ranges are automatically defined
- Very 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
```

SSM-Mapping

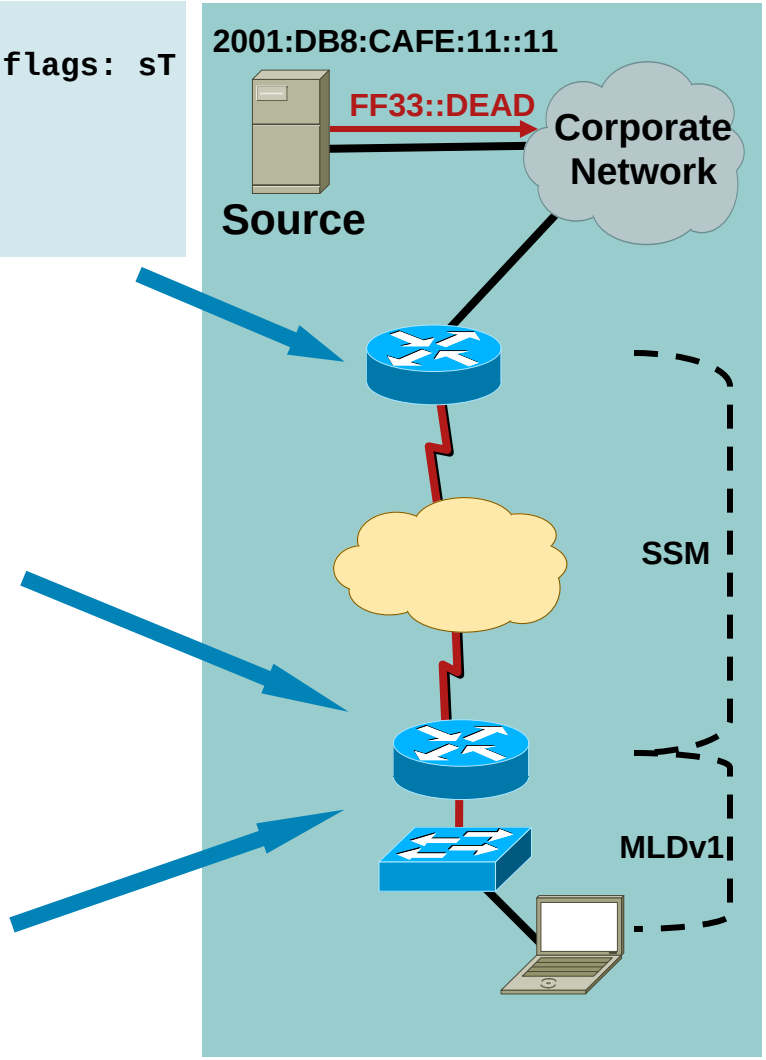
```
core-1#show ipv6 mroute | begin 2001:DB8:CAFE:11::11
(2001:DB8:CAFE:11::11, FF33::DEAD), 00:01:20/00:03:06, flags: sT
  Incoming interface: GigabitEthernet3/3
  RPF nbr: FE80::20E:39FF:FEAD:9B00
  Immediate Outgoing interface list:
    GigabitEthernet5/1, Forward, 00:01:20/00:03:06
```

Static Mapping:

```
ipv6 multicast-routing
!
ipv6 mld ssm-map enable
ipv6 mld ssm-map static MAP 2001:DB8:CAFE:11::11
no ipv6 mld ssm-map query dns
!
ipv6 access-list MAP
 permit ipv6 any host FF33::DEAD
```

DNS Mapping (the default):

```
ipv6 multicast-routing
!
ipv6 mld ssm-map enable
!
ip domain multicast ssm-map.cisco.com
ip name-server 10.1.1.1
```

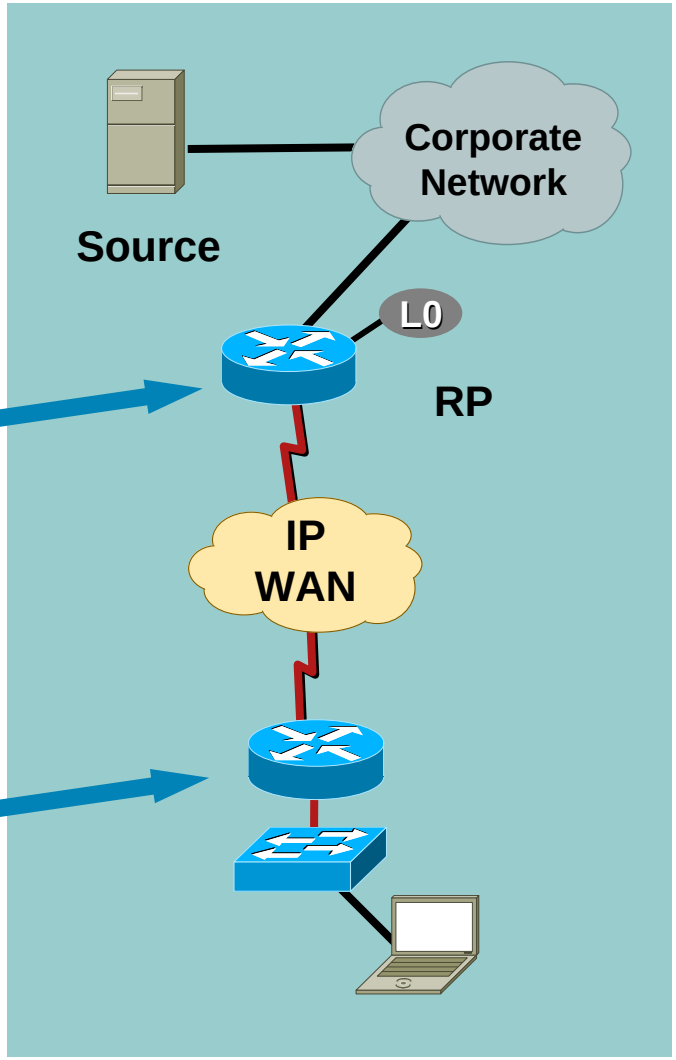


IPv6 Multicast Static RP

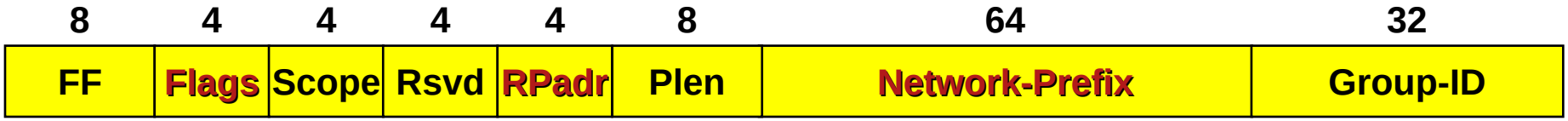
- **Easier than before as PIM is auto-enabled on every interface**

```
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
```



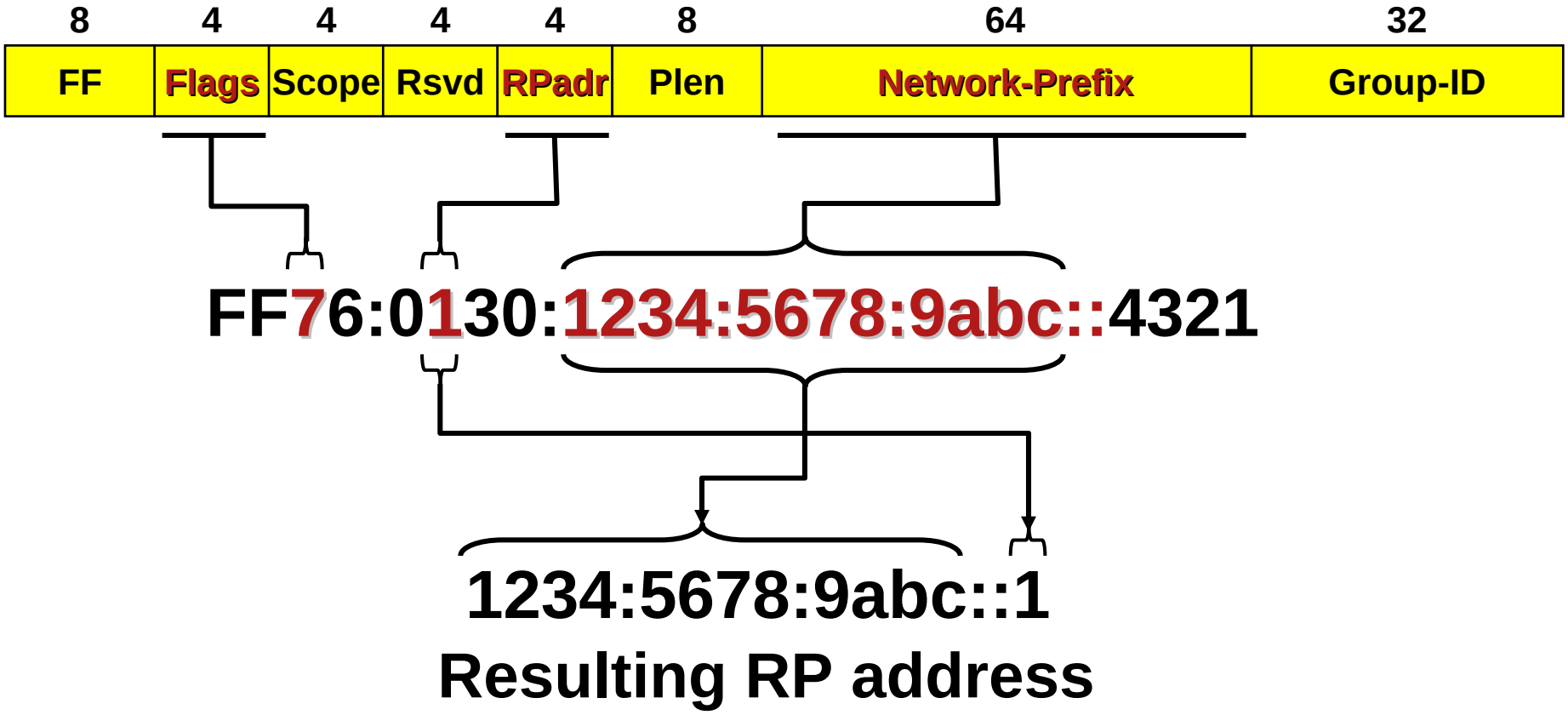
Embedded RP Addressing (rfc3956)



- **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**

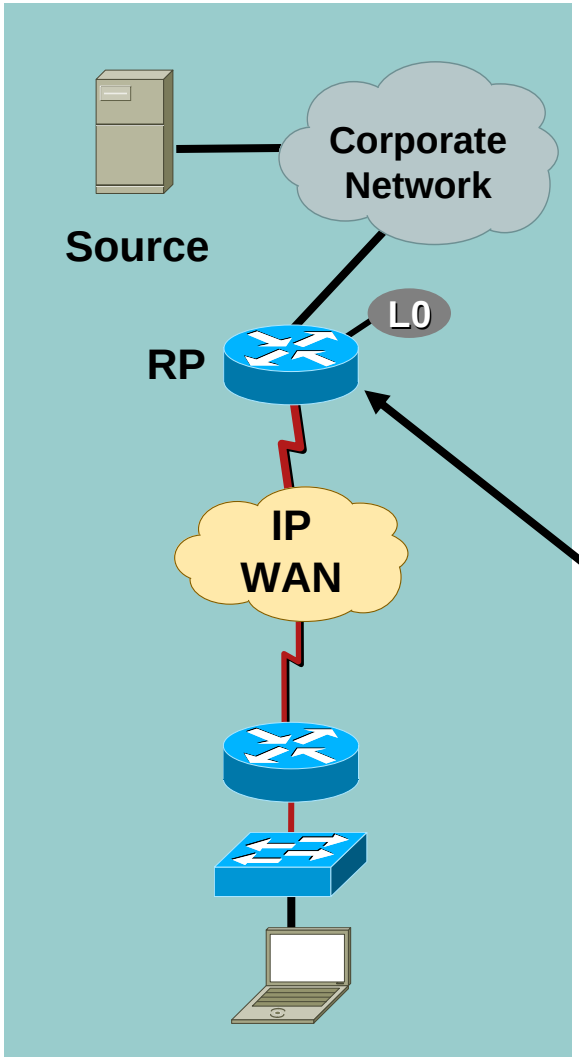
Embedded RP Addressing – Example

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
```

IPv6mc Deployment SSM / ASM

- **SSM**
  - **Use SSM whenever MLDv2 / MLD mapping is available**
  - **Intradomain and/or Interdomain deployment**
- **ASM**
  - **Bidir-PIM and/or PIM-SM**
  - **Intradomain and/or Interdomain deployment**
  - **RP-announcement: static-config, BSR, embedded-RP**
  - **RP-redundancy: BSR, anycast/prioritycast-RP**
- **PIM-Anycast-RP – not supported today: please contact your AM / SE and submit feature requests**

IPv6mc Deployment – Inter / Intra Domain

- Interdomain ASM

  - Use Embedded-RP with PIM-SM

- Intradomain ASM

  Bidir-PIM

  - Best scalable solution, no data-triggered events, arbitrarily many sources

  -use static-RP or BSR for announce and prioritycast-phantom-rp for redundancy

  PIM SM

  - only required when routers not supporting Bidir-PIM are used

  -use BSR for RP announce and redundancy

IPv6 Multicast Applications

- **Microsoft Windows Media Server/Player (9 and 10)**

  http://www.microsoft.com/windows/windowsmedia/default.aspx

- **VideoLAN—www.videolan.org**
- **DVTS (Digital Video Transport System)**

  http://www.sfc.wide.ad.jp/DVTS/

  http://www.dvts.jp/en/dvts.html

- **Internet radio stations over IPv6**

  http://www.ipv6.ecs.soton.ac.uk/virginradio/

  Supported on iTunes 4.5, Windows Media Player, XMMS 1.2.8, etc…

- **Many more applications…Google is your friend :-)**

Agenda

- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- Multicast Triple Play

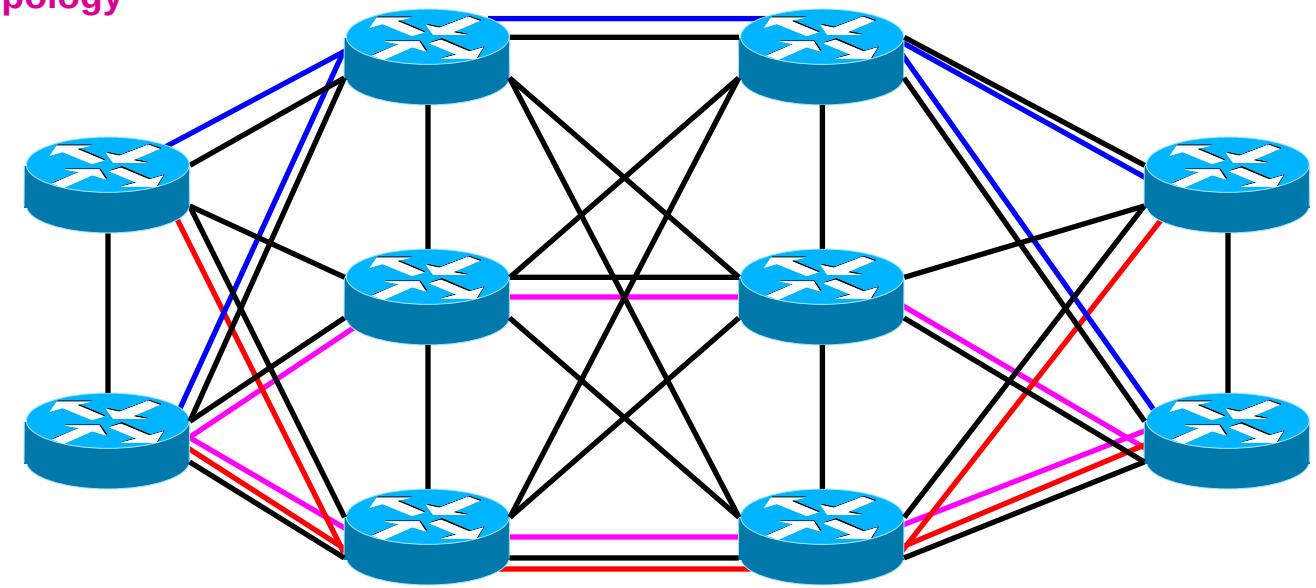
Conceptual View of MTR

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

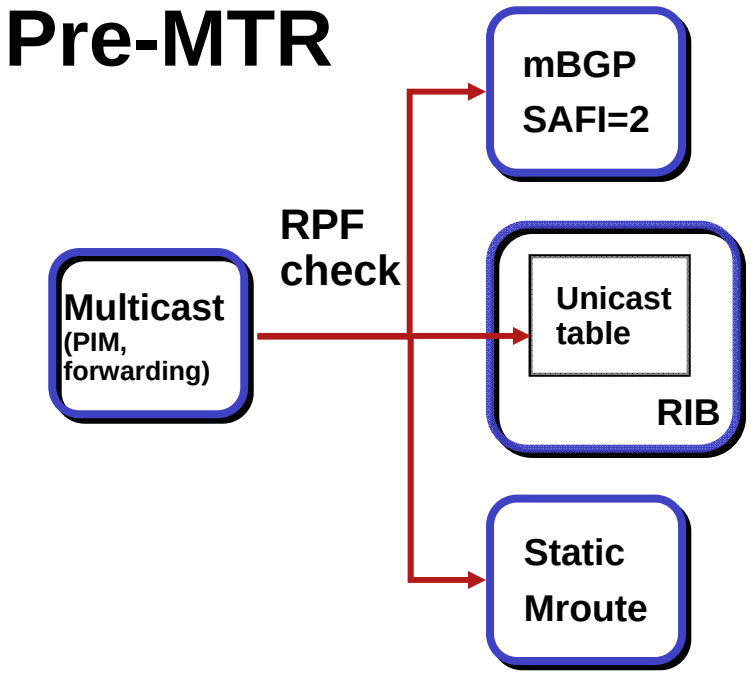


- Define the class-specific topology across a contiguous section of the network
- Individual links can belong to multiple topologies

Incongruent IPv4 Unicast and Multicast Topologies

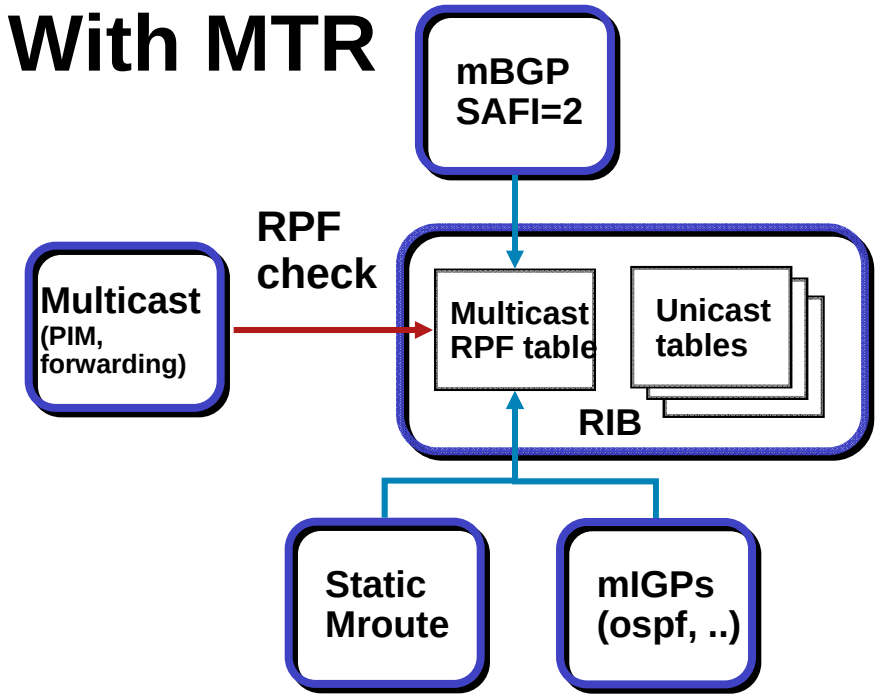
- **MTR allows incongruent unicast and multicast topologies**
  - **Metrics can be different for each on the same link**
- **Restrict traffic**
  - **Restrict multicast only to designated areas of the network**
- **No reference to unicast for multicast RPF**
  - **RPF checks based on multicast specific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

MTR Multicast Changes



- **No single database for Multicast RPF**
  - **Multicast consults multiple sources of "unicast" routes**
  - **Selects result according to custom preference rule**
- **Standard unicast routes (uRIB) *always* considered.**
  - **No way to exclude them completely**

MTR Multicast Changes

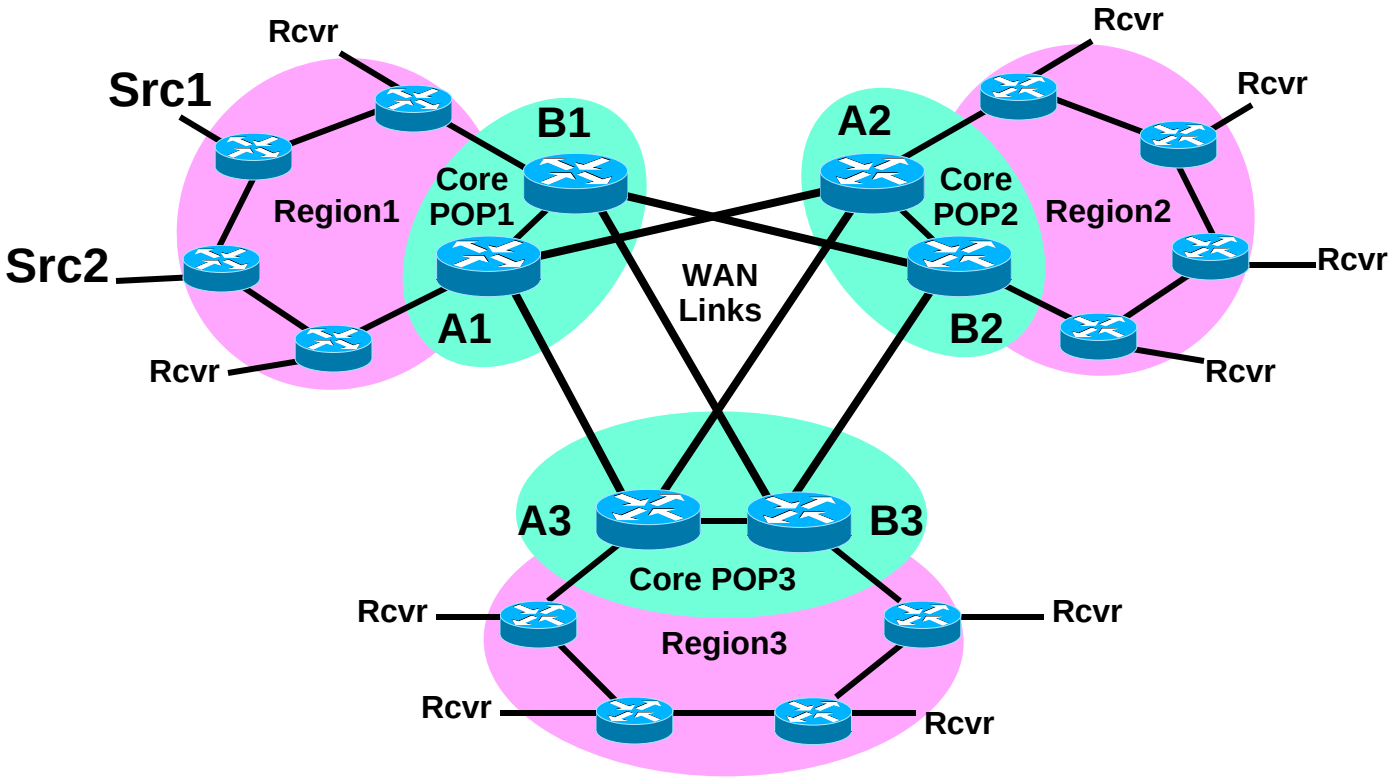


- MTR manages a routing table specifically for Multicast RPF.
  - Multicast uses that table as the sole source of RPF routes.
- Highly flexible.
  - Any protocol may be configured to contribute to the mRPF table.
- Standard configuration options are available.
  - Including redistribution among protocols.

MTR / MiMTR Overview

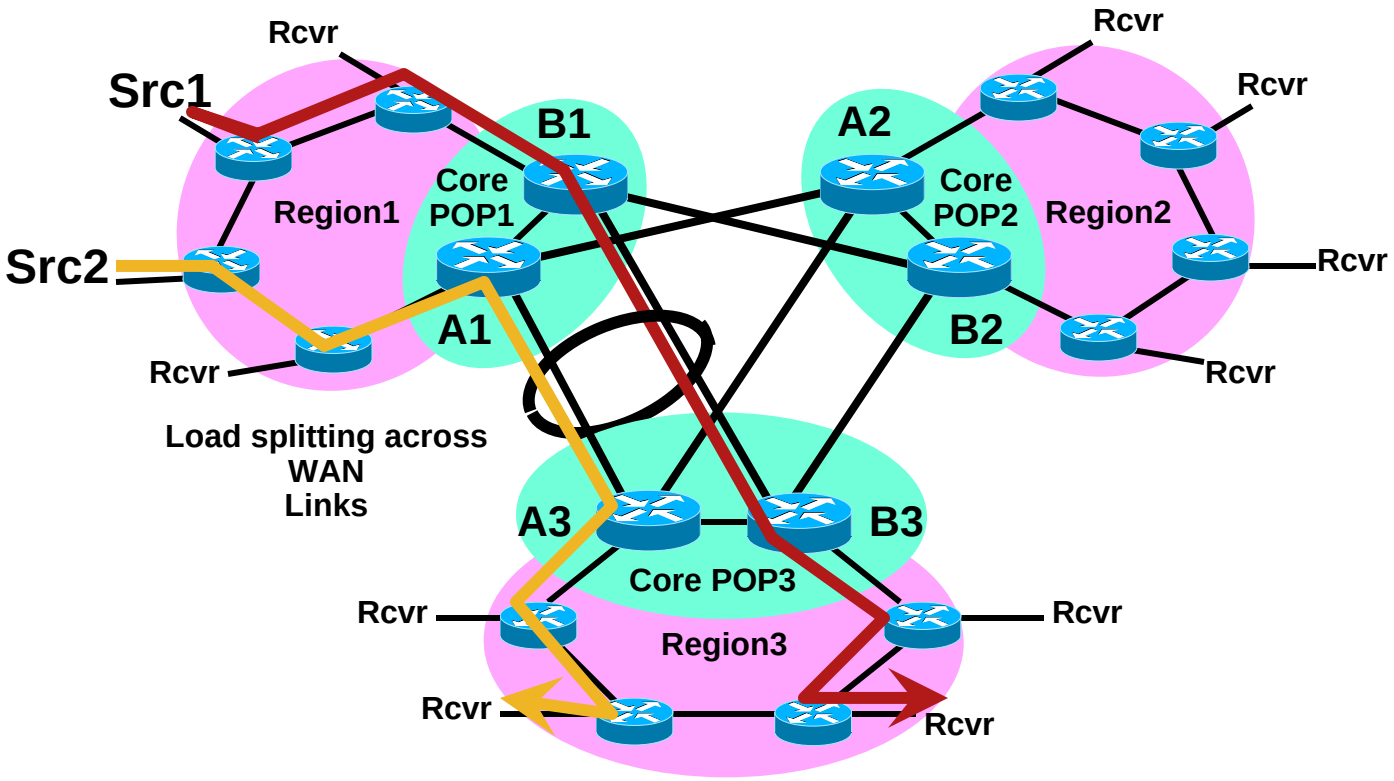
- **Functionality:**
  - **Ability to run parallel multiple instances of IGP. Each instance commits routes into different RIB/(FIB). Each "class" of traffic is assigned to be routed against particular RIB/FIB.**
- **MTR: single instance of an IGP runs multiple topologies**
  MiMTR: run multiple instances of an IGP in parallel
  - **Same result for unicast/multicast traffic**
- **Multicast: A RIB only used for multicast (not unicast) does not require a FIB – Multicast RPF only requires RIB. Called MuRIB.**

MTR/MiMTR application - *Cost optimization*



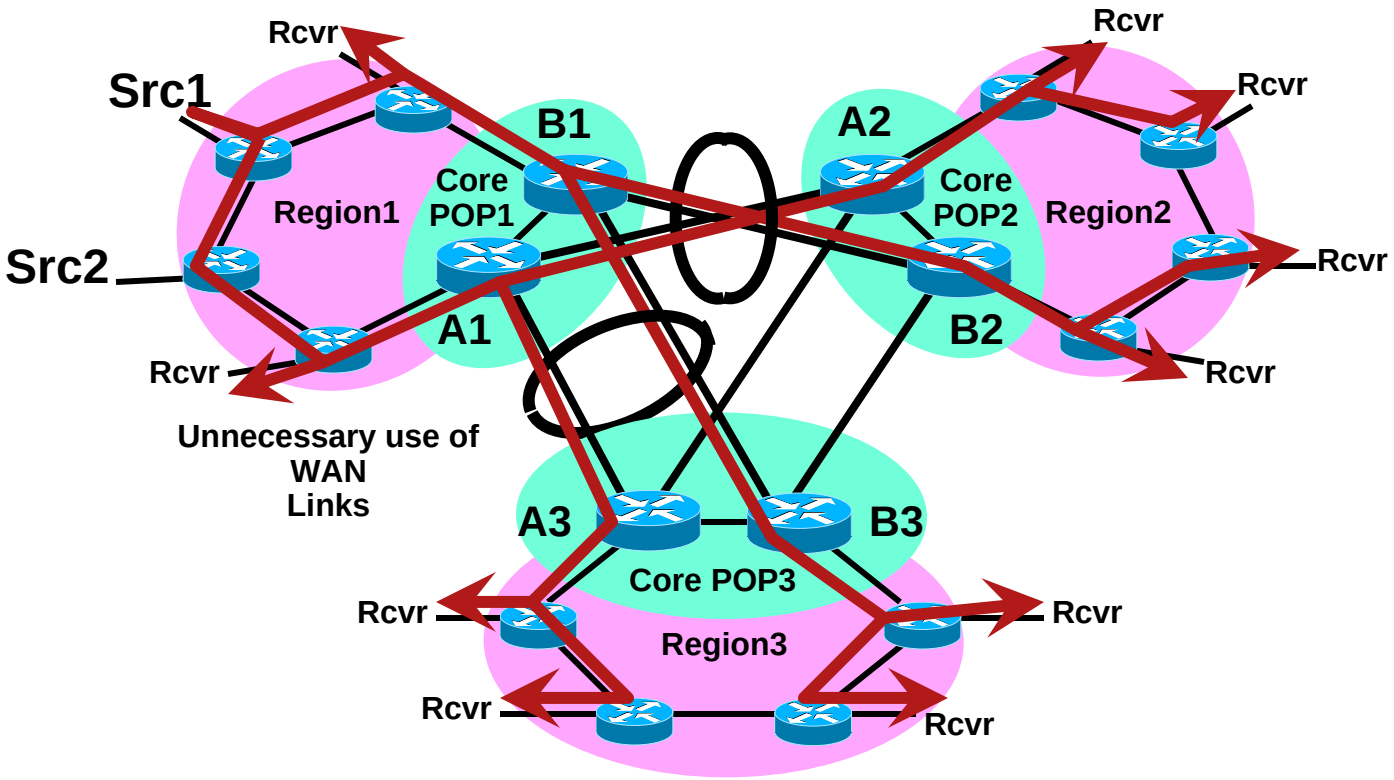
- Consider simplified example core/distribution network toplogy
- Core pops have redundant core routers, connectivity via (10Gbps) WAN links, redundant. Simple setup: A/B core routers, A/B links
- Regions use ring(s) for redundant connectivity

MTR/MiMTR application - *Cost optimization*



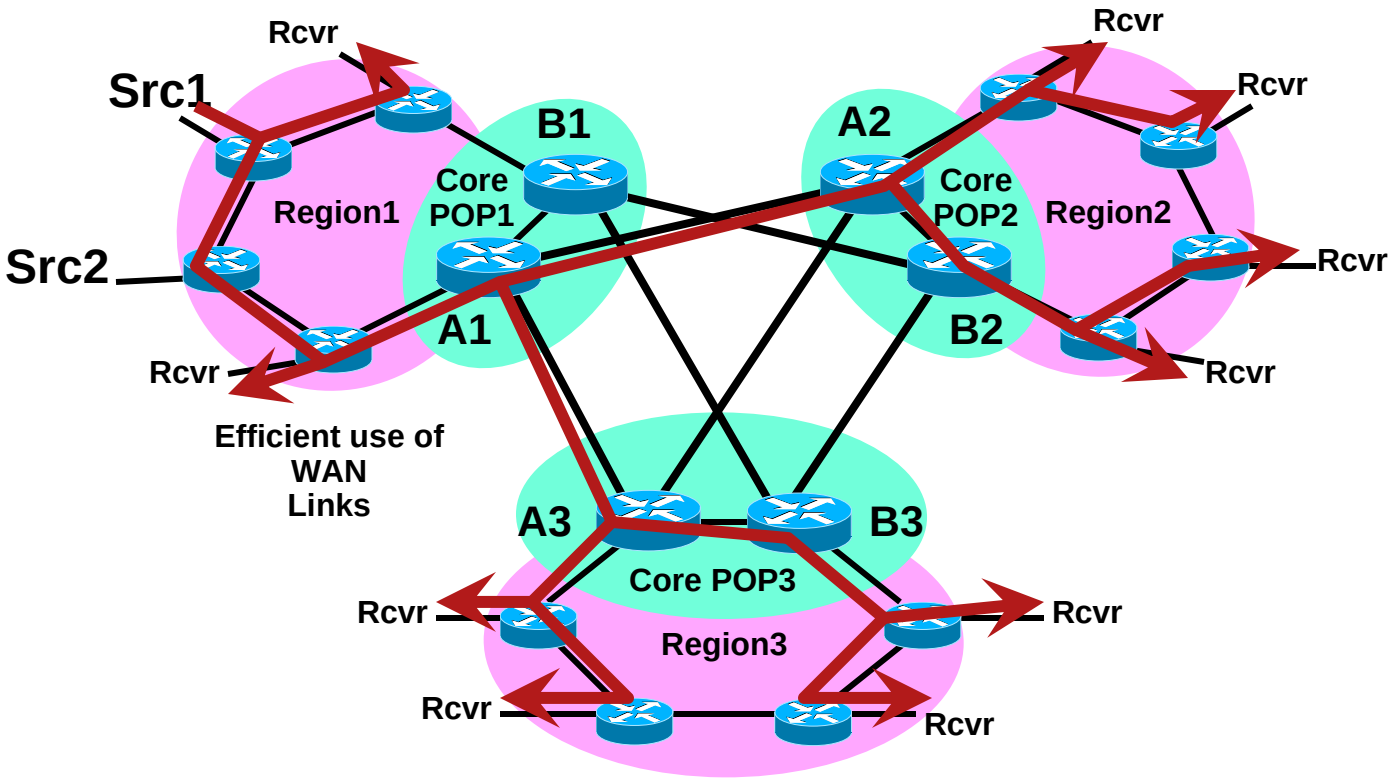
- **IGP metric are set to achieve good load distribution across redundant core.**
  - **Manual IGP metric setting and/or tools (Cariden)**
  - **Assume in the idealized topology cost of 1 on all links.**
- **Result: Unicast traffic is load split across redundant core links**

MTR/MiMTR application - *Cost optimization*



- **The same metric good for unicast load splitting causes multicast traffic to go unnecessarily across both A and B WAN links.**
  - **10 Gbps WAN links, 1..2 Gbs multicast => 10..20% WAN waste (cost factor)**
- **Cannot resolve problem without multicast specific topology**

MTR/MiMTR application - *Cost optimization*



- **Simple to minimize tree costs with a multicast specific topology**
  - **Manual or tool based**
  - **Example toplogy: make B links very expensive for multicast (cost 100), so they are only us as last resort (no A connectivity)**

Multicast MTR CLI example

```
ip multicast-routing
    ip multicast rpf multitopology
 interface TenGig0/0
 ip pim sparse-mode
! no special config here, just use  default IGP metrics.
 interface TenGig0/1 ! secondary link to other core-pop
 ip pim sparse-mode ! enables in multicast topology
 topology ipv4 unicast base
 cost 10000  ! expensive, use only when primary fails
 ! only expensive for multicast. In unicast we want load-
splitting across both links !
   router ospf 1
     network ... ! normal config
     ! Create OSPF for multicast base topology
     address-family ipv4 multicast
       topology base
     exit-address-family
```

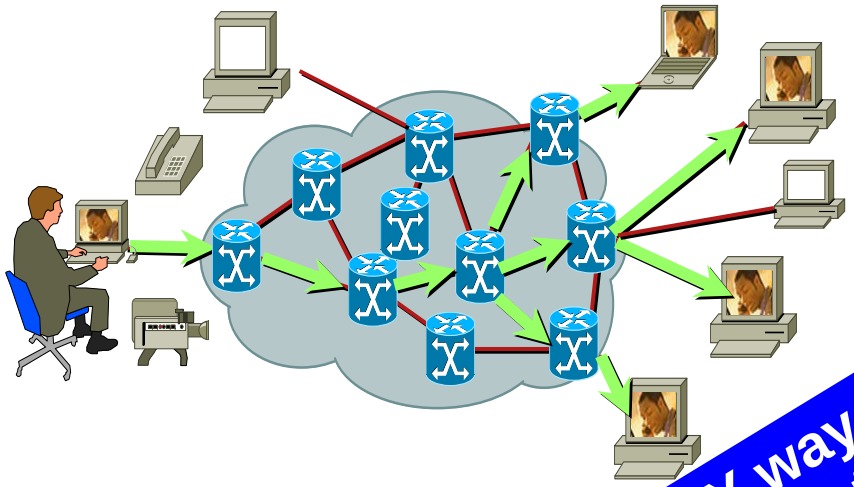
Agenda

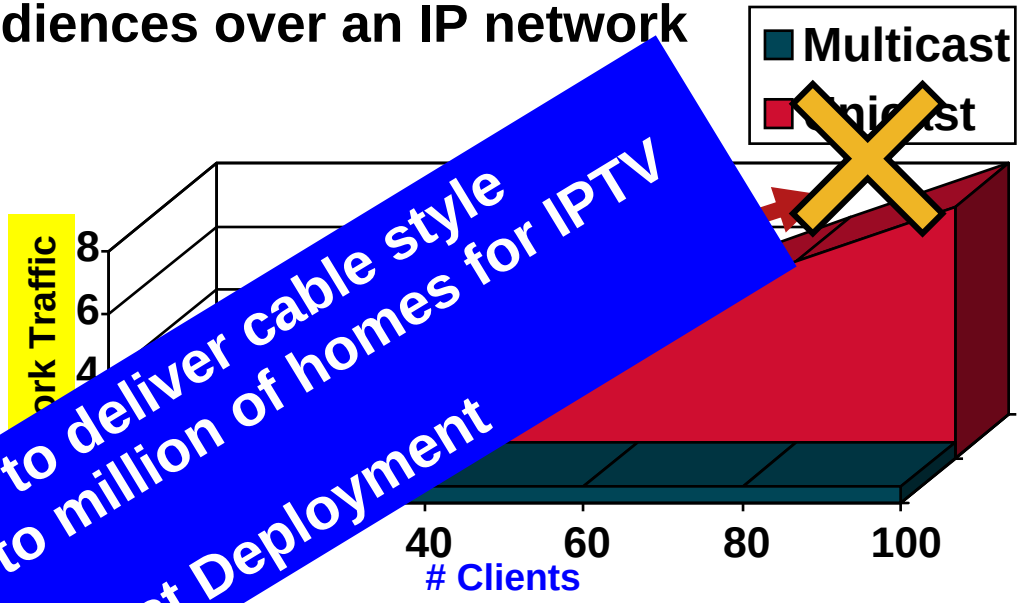
- Multiprotocol BGP (MBGP)
- Multicast Source Discovery Protocol (MSDP)
- Source Specific Multicast (SSM)
- Multicast VPN (MVPN)

- Multicast IPv6
- Multi Topology Routing (MTR)
- **Multicast Triple Play**

Relevance of Multicast for IPTV delivery

Distribute information to large audiences over an IP network





Multicast is the ONLY way to deliver cable style broadcast video content to million of homes for IPTV

IPTV Deployment == Multicast Deployment

Multicast

1. Efficiently Cont
2. Reduces
3. Eliminate ncy
4. Makes Multi pplications possible

Multicast Benefits

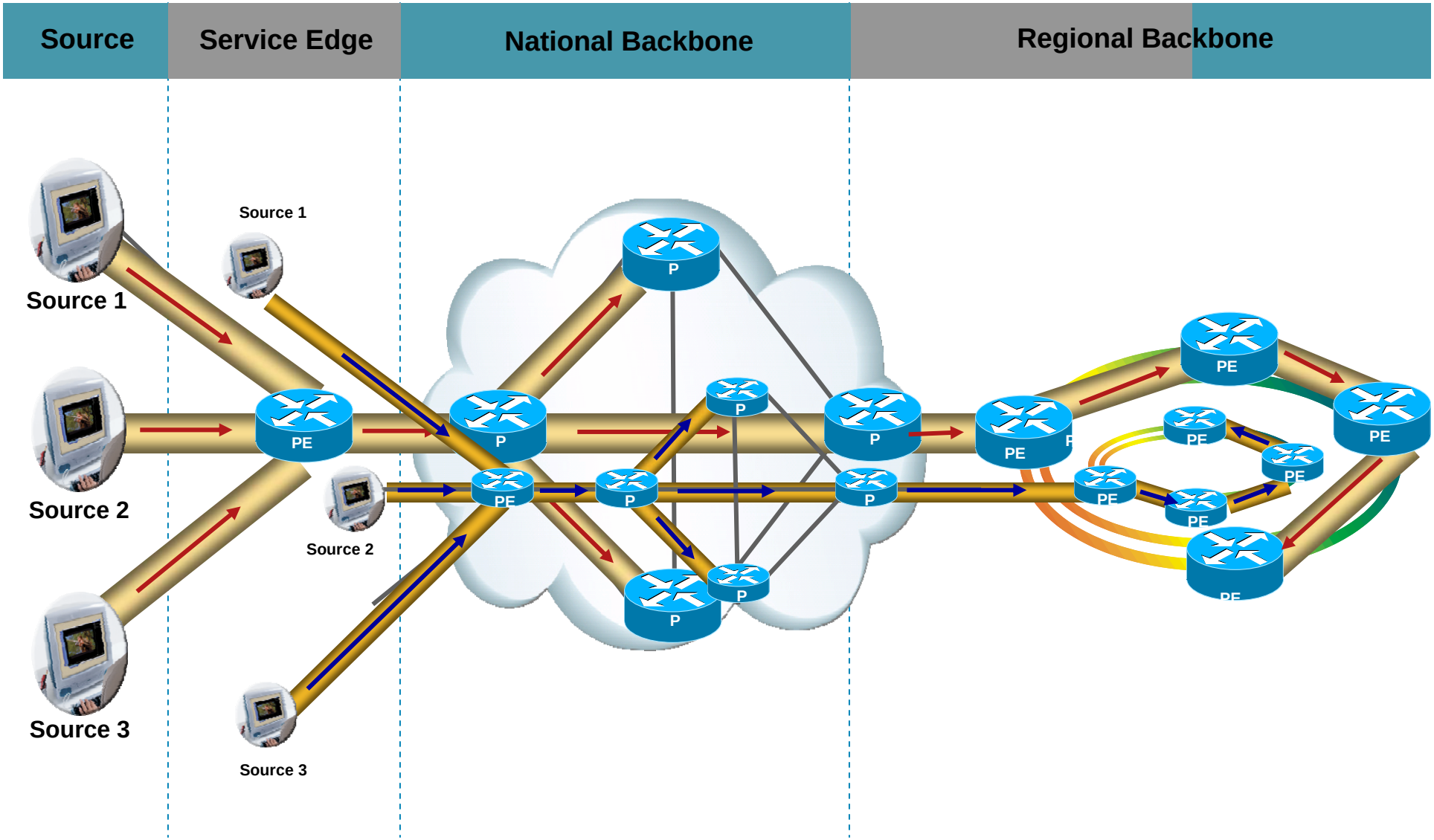
- Increase Productivity & Save Cost
- Generate New Revenue Stream

Video 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

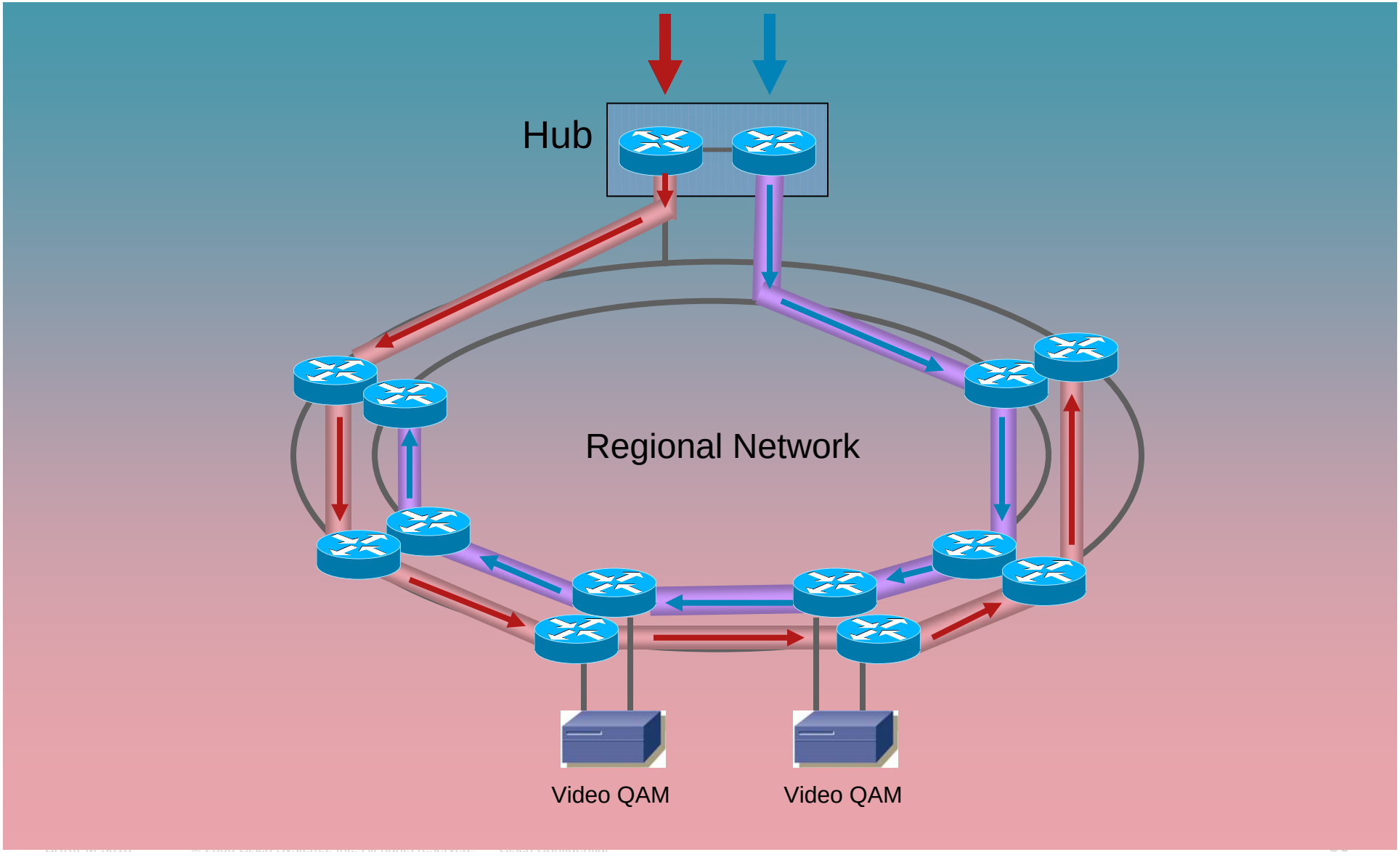
End to End Recovery Models

Hot-Hot Video Delivery Model



End to End Recovery Models

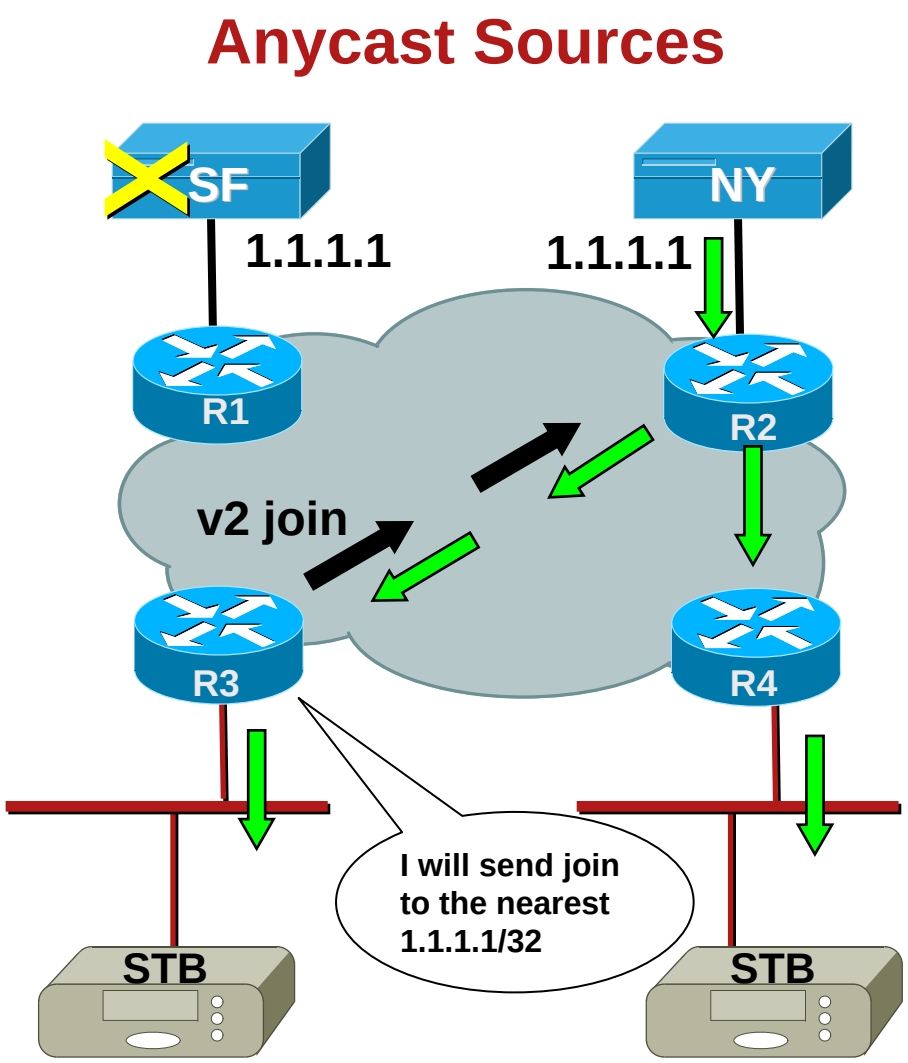
Video Hot-Hot Delivery Model Zoom in on Regional Network



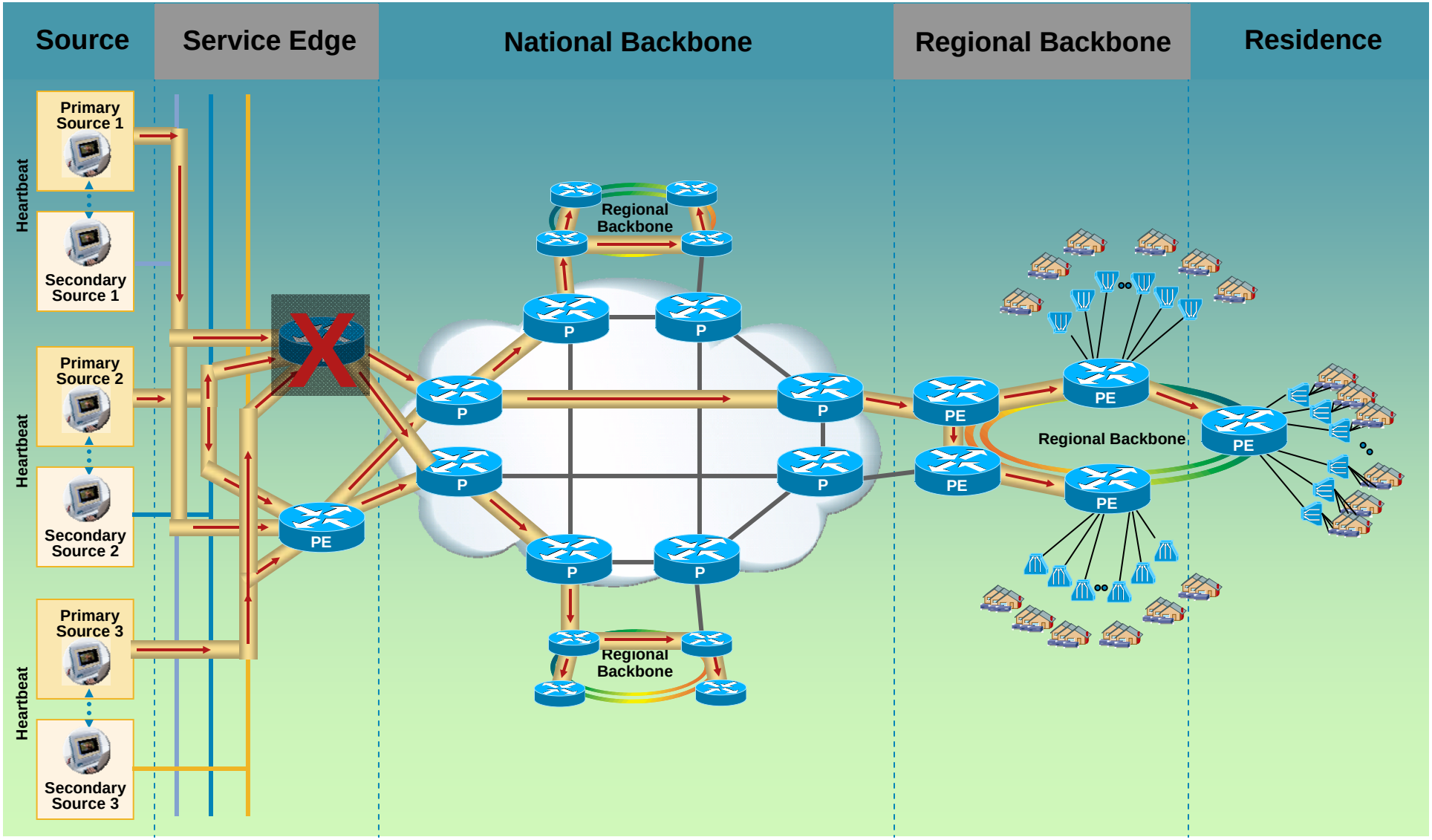
Multicast Source Redundancy Using Anycast Sources

How is source redundancy achieved in the network?

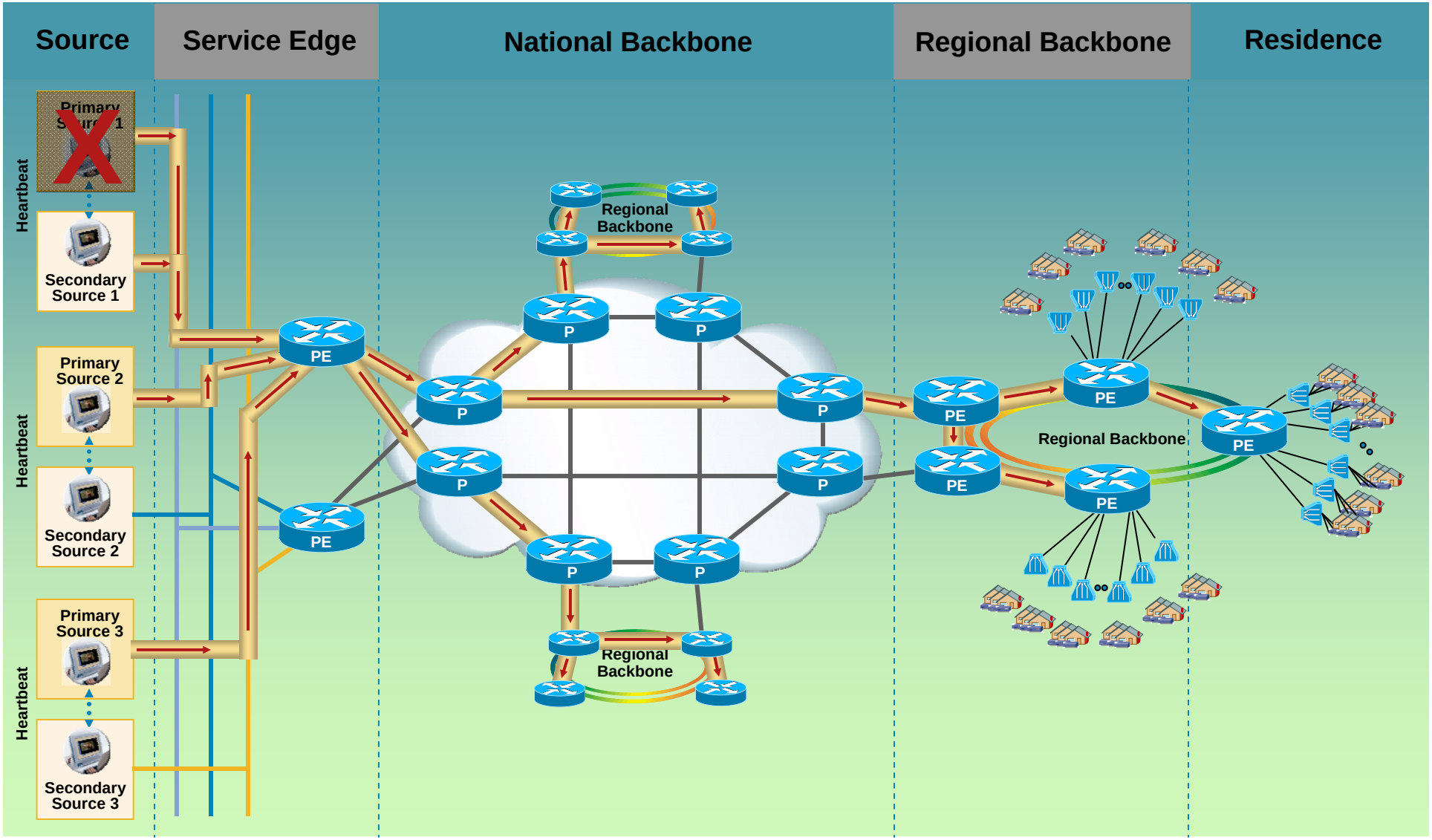
- Enable SSM on all routers
- Have R1 and R2 advertise same prefix for each source segment.
- R3 and R4 follow best path towards source based on IGP metrics.
- Let's say R3's best path to SF is through R1. The source in SF now suddenly fails.
- R3's IGP will reconverge and trigger SSM joins towards R2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



Native IP Multicast Video Triple Play
Redundancy : Source Router Failure



Native IP Multicast Video Triple Play
Redundancy : Video Source 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:

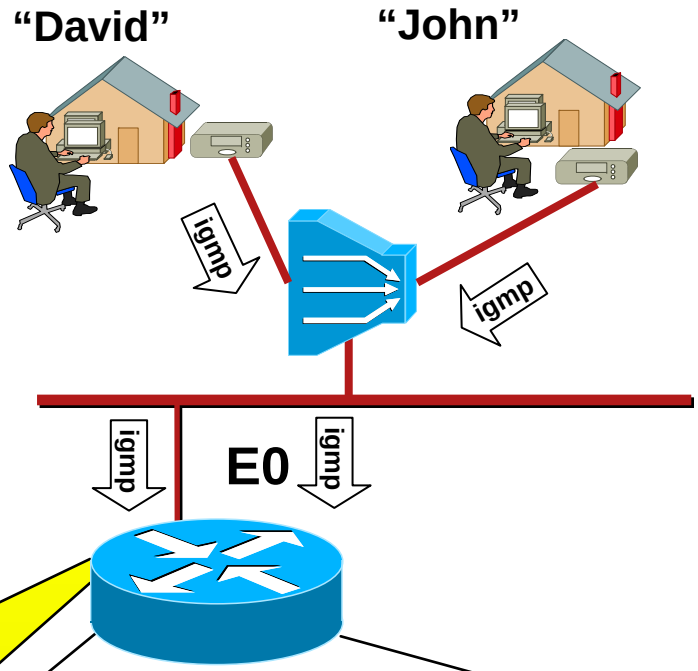
Reduce the leave latency during a channel change by extending the IGMPv3 protocol.

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)!



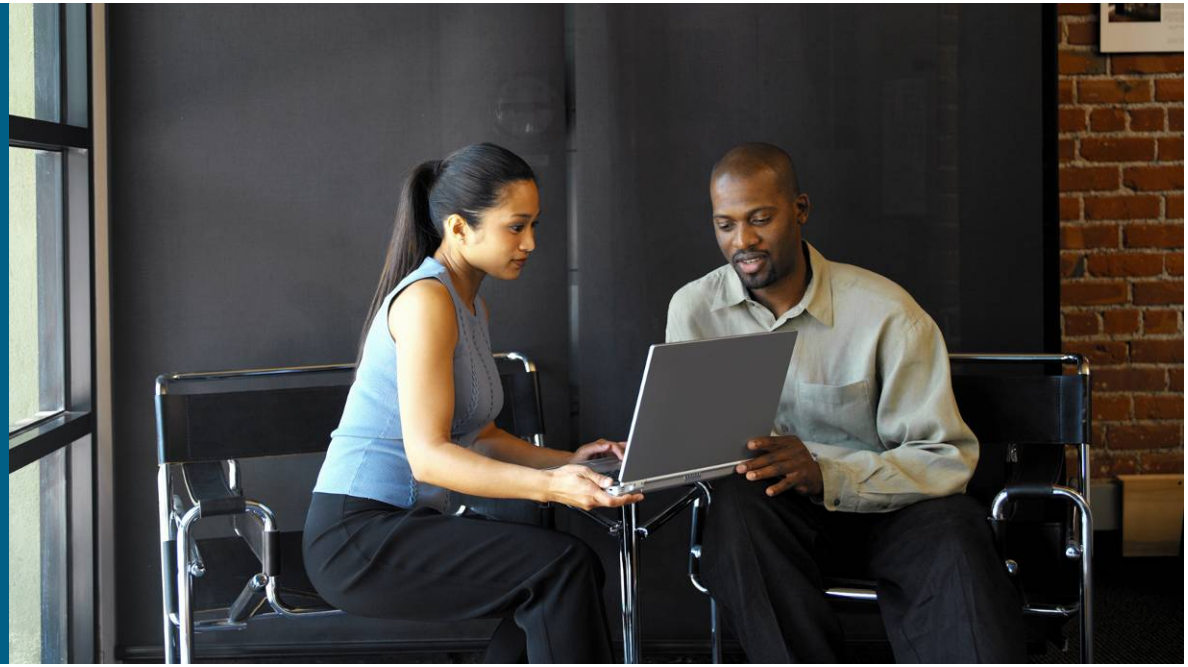
Configuration:

```
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"

Questions ?

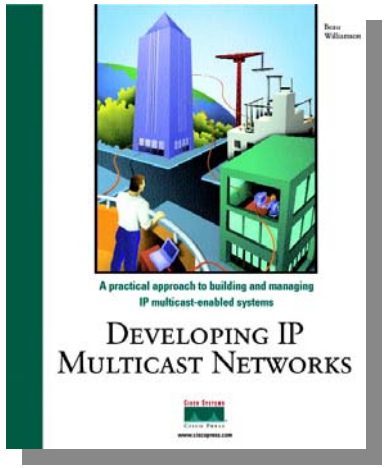


Summary of session

- Advanced multicast topics associated with inter-domain
- The fundamentals of MBGP and MSDP
- Source-specific multicast and how this mode solves many problems associated with traditional inter-domain multicast
- Building multicast VPNs in an MPLS VPN environment using multicast domains and multipoint LSP solutions
- A brief overview of IPv6 multicast
- Multicast using multi-topology routing (MTR)#
- Multicast for a ‘Triple play’ deployment

Multicast Sessions / more information

- **Multicast Techtorial**
- **Management and Security Breakouts**
- **Multicast BoF**
- **MTE: Beau Williamson, Toerless Eckert, Andy Kessler, Dino Farinacci, Steve Simlo**
- **Multicast CCO Page: www.cisco.com/go/multicast**
- **Questions: cs-ipmulticast@cisco.com**
- **Customer Support Mailing List: tac@cisco.com**
- **RTFB**



Meet the Experts
IP and MPLS Infrastructure Evolution

- Andy Kessler
  Technical Leader
- Beau Williamson
  Consulting Engineer
- Benoit Lourdelet
  IP services Product manager
- Bertrand Duvivier
  Consulting Systems Engineer
- Bruce Davie
  Cisco Fellow
- Bruce Pinsky
  Distinguished Support Engineer






Meet the Experts
IP and MPLS Infrastructure Evolution

- Gunter Van de Velde
  Technical Leader
- John Evans
  Distinguished Systems Engineer
- Oliver Boehmer
  Network Consulting Engineer
- Patrice Bellagamba
  Consulting Engineer
- Shannon McFarland
  Technical Leader





Meet the Experts
IP and MPLS Infrastructure Evolution

- Andres Gasson
  Consulting Systems Engineer

- Steve Simlo
  Consulting Engineer

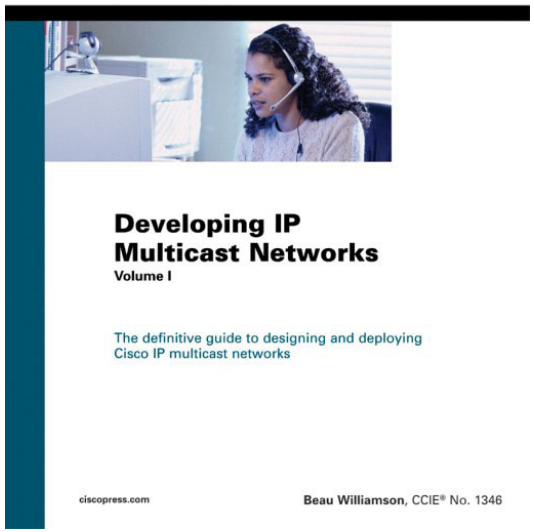
- Toerless Eckert
  Technical Leader

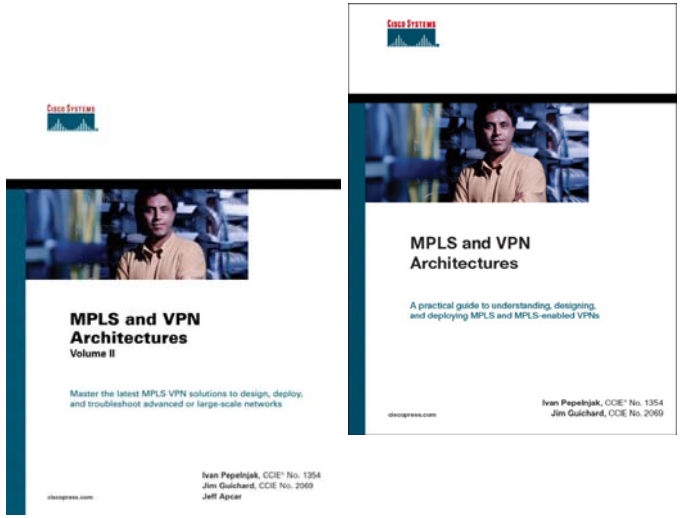
- Dino Farinacci
  Cisco Fellow & Senior Software Engineer

Recommended Reading

BRKIPM -3018

- MPLS and Next-Generation Networks
- MPLS and VPN Architectures, Volume II
- Developing IP Multicast Networks, Volume I





Available in the Cisco Company Store

Please complete your Online Session Evaluation!

- If you are uncertain what to put then 5 is always a good default !