You make **possible**

# mVPN Deployment Models

Luc De Ghein – Technical Leader

BRKIPM-3017

Cisco live!
June 9-13, 2019 • San Diego, CA

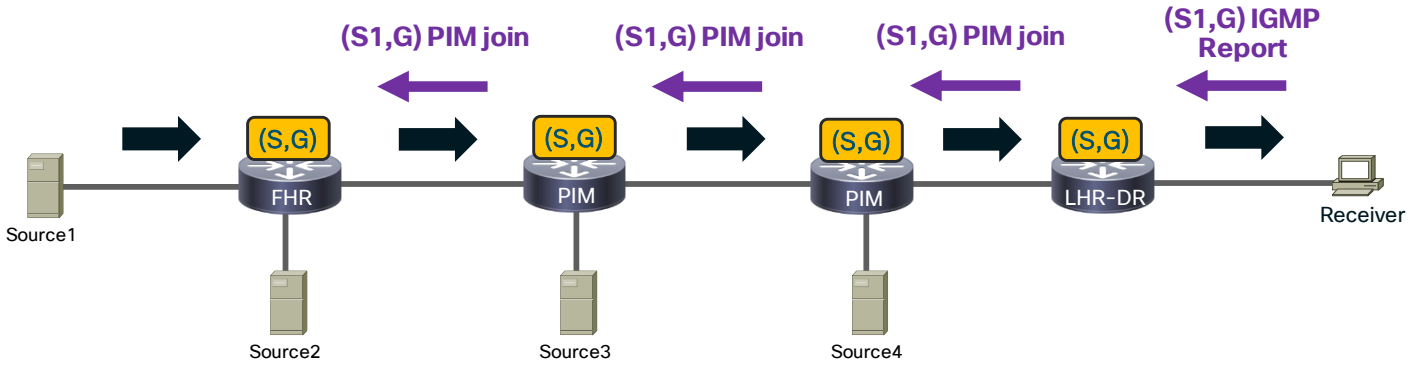
# Agenda

- Introduction
- mVPN Architecture
  - Core trees
  - Overlay Signaling
- mVPN Deployment Models
- Deployment and Examples
- Key Takeaways

# Introduction

- Purpose
  - Provide key takeaway's
  - Explain building blocks of classic mVPN and Next Generation (NG) mVPN
  - No operating system/platform specifics (including release support)
  - Little configuration
    - Configure mVPN Profiles Within Cisco IOS
    - Configure mVPN Profiles Within Cisco IOS-XR
  - Provide guidance which mVPN profile to chose
- Before we start
  - C- stands for customer
  - P- stands for Provider

# Basic Multicast Recap: SSM

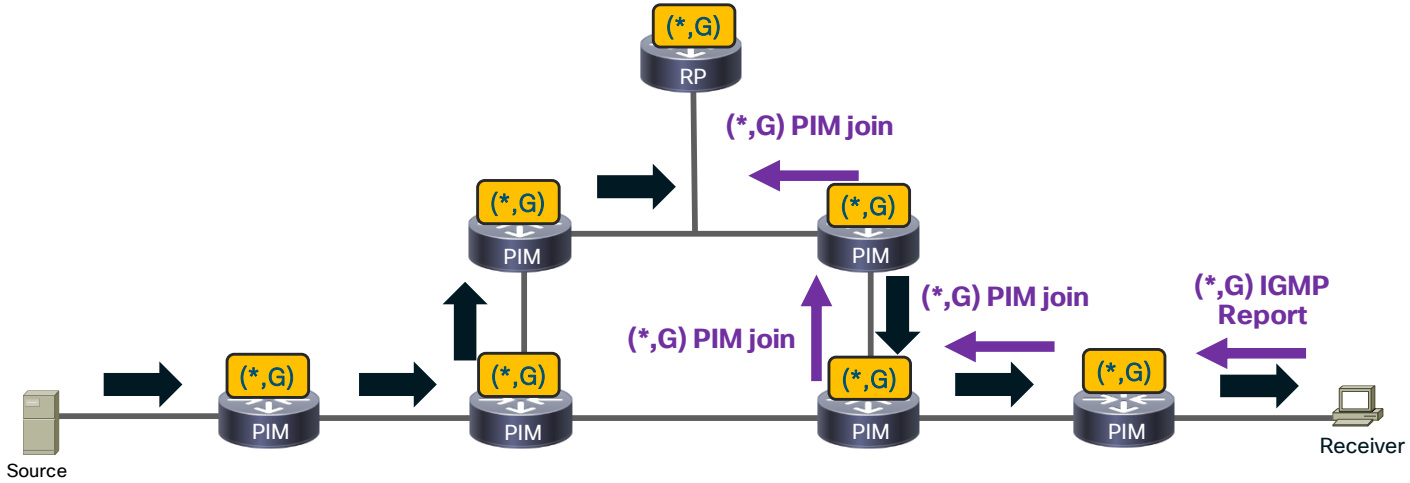



- SSM: Source Specific Multicast
- Only (S,G) state
- Optimal path
- Easy!
- Receiver needs IGMP v3 (SSM mapping can be used)

multicast traffic ➡

FHR: First Hop Router
LHR: Last Hop Router

# Basic Multicast Recap: BiDir

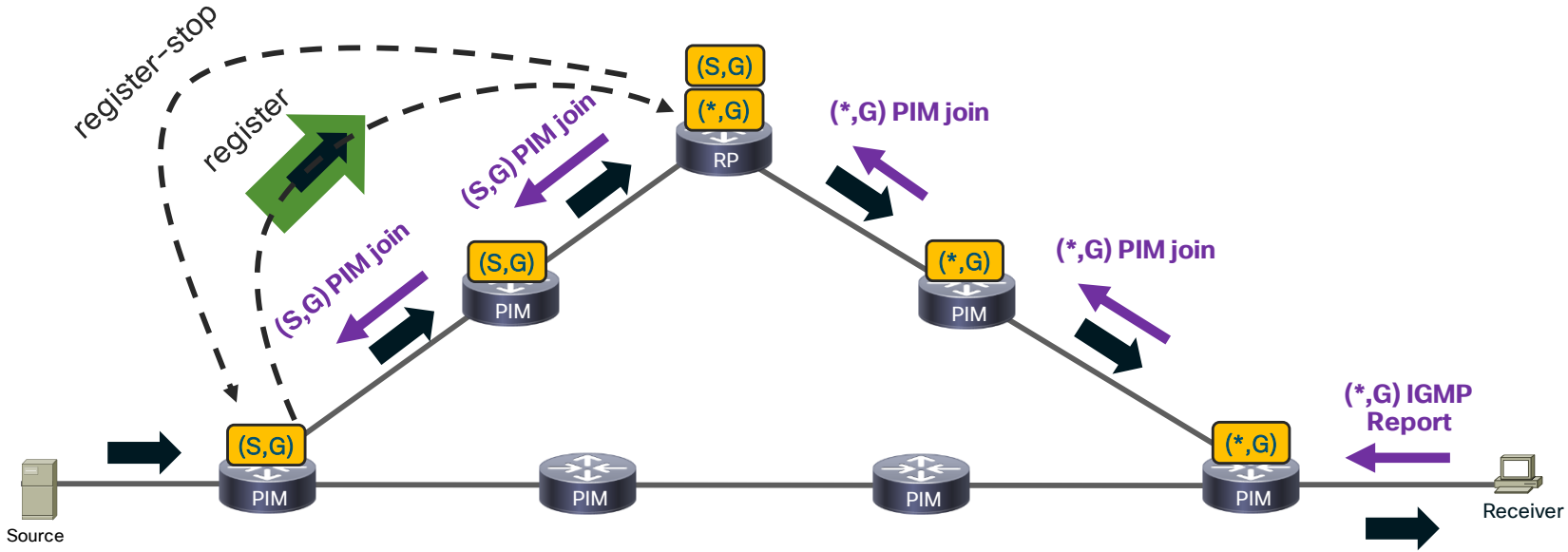



- BiDir: Bi-Directional
- Only (*,G) state
- RP needed (phantom)
- Not optimal path

- Multicast traffic going from source to the RP is moving upstream
- Multicast traffic going from the RP to receivers is moving downstream

multicast traffic

# Basic Multicast Recap: Sparse Mode

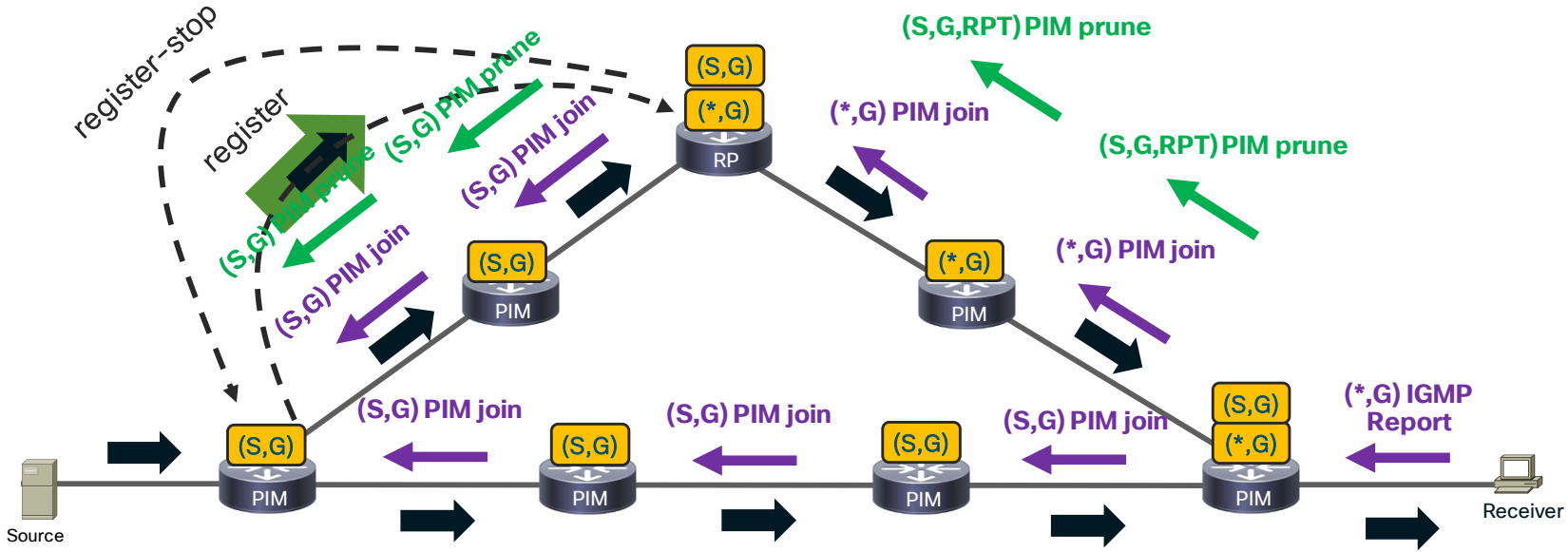



- (*,G) and (S,G) state
- RP needed

- (*,G) to (S,G) switchover
- Complex to troubleshoot

multicast traffic

# Basic Multicast Recap: Sparse Mode

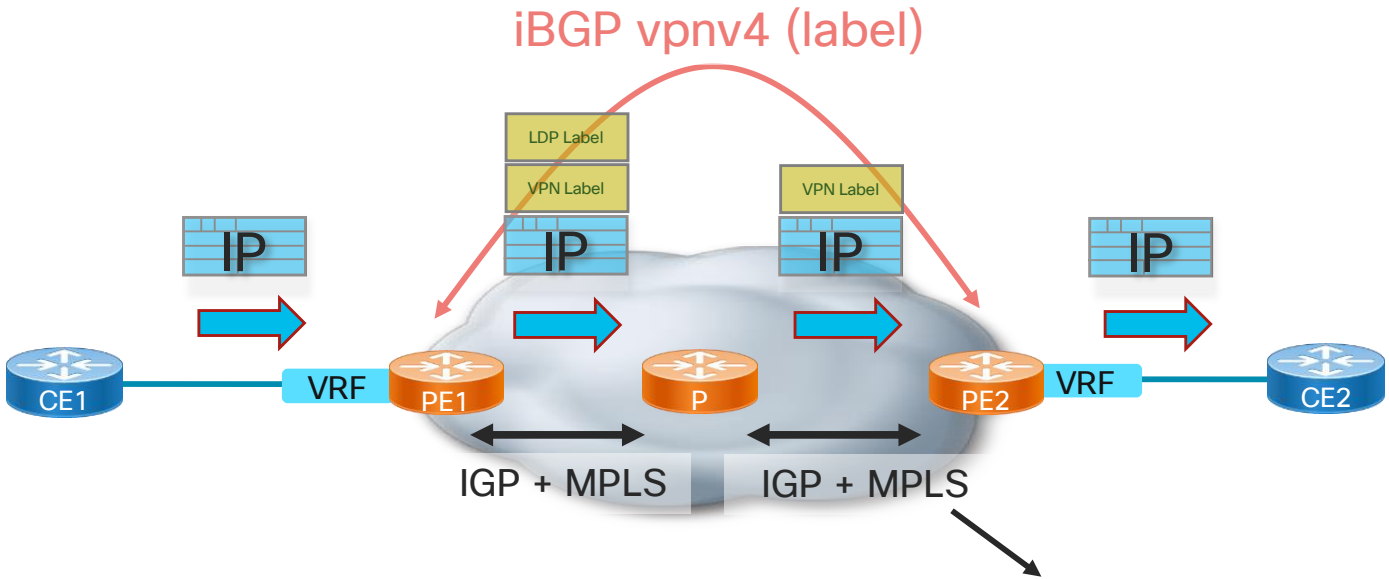



- (*,G) and (S,G) state
- RP needed

- (*,G) to (S,G) switchover
- Complex to troubleshoot

multicast traffic

# MPLS VPN Unicast Recap




- Without unicast, no multicast: RPF would fail
- Multicast forwarding in core?

MPLS =

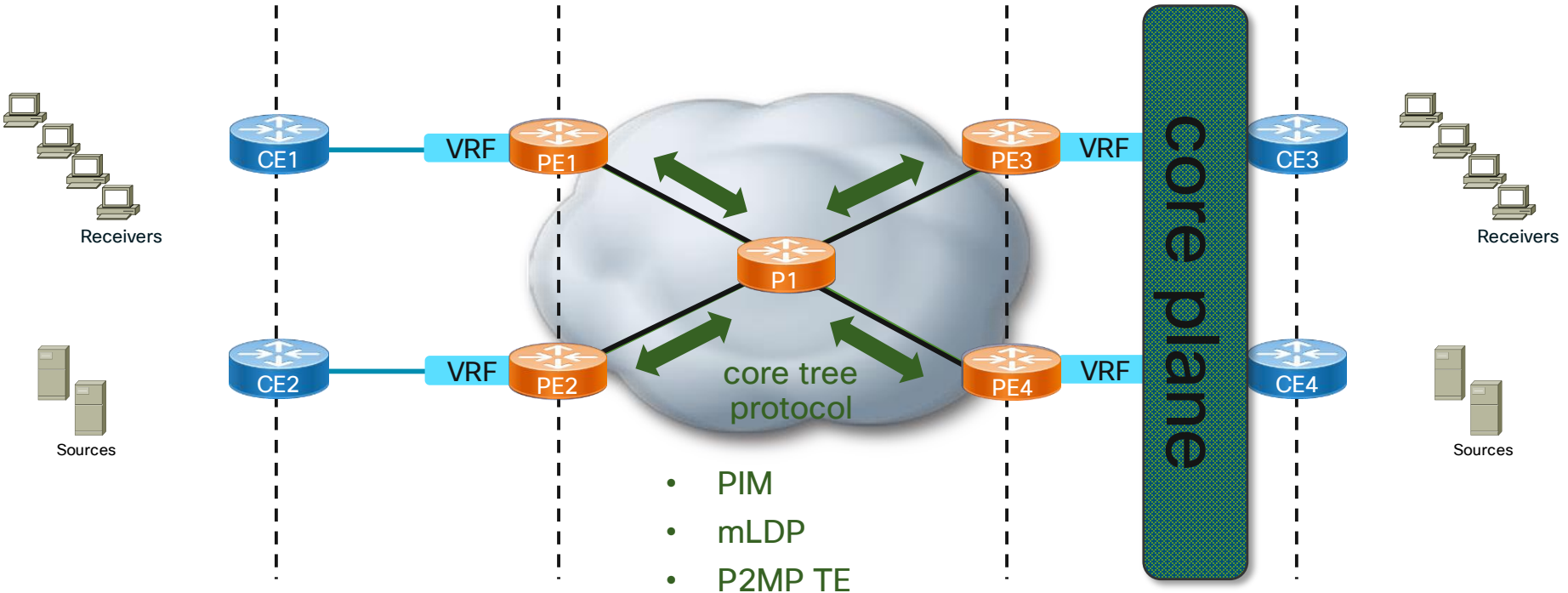
- LDP
- MPLS TE
- Segment Routing MPLS

# mVPN Architecture

You make networking **possible**

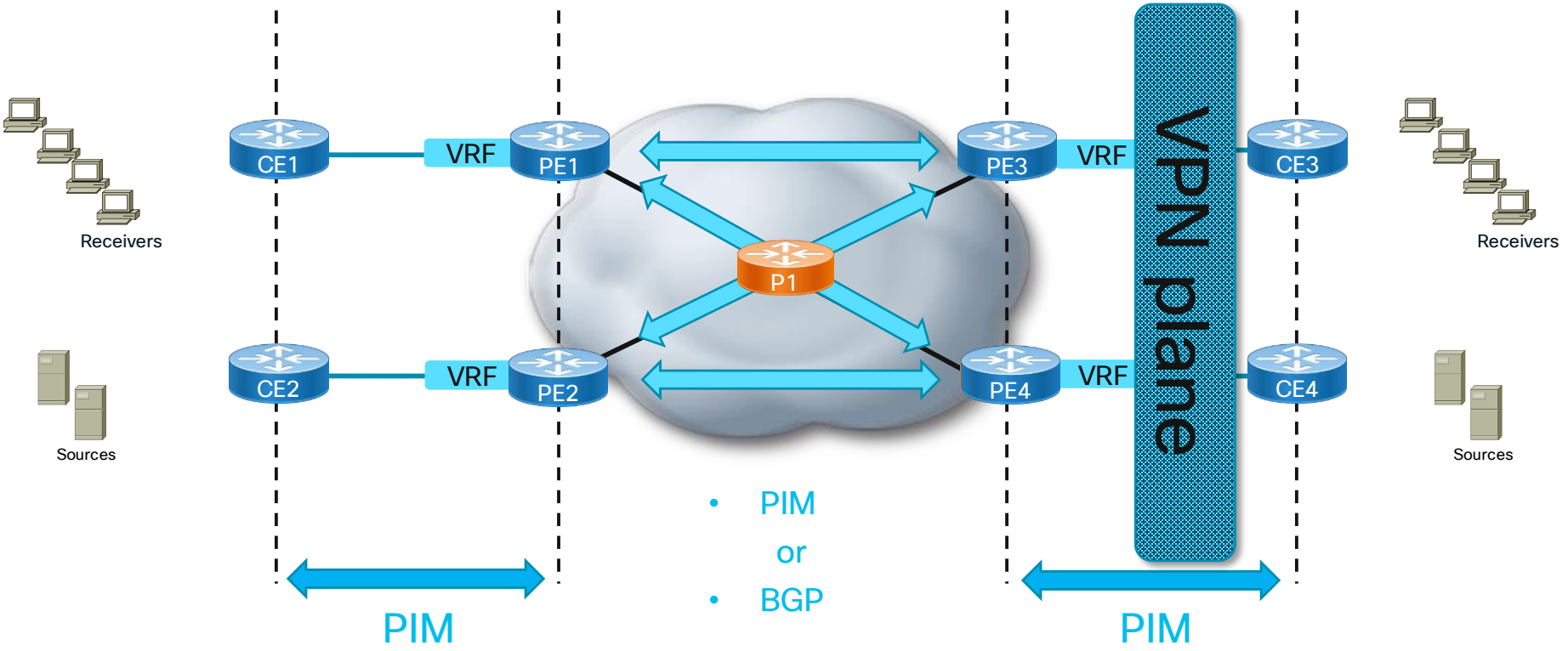
Cisco live!

# Planes – Underlay

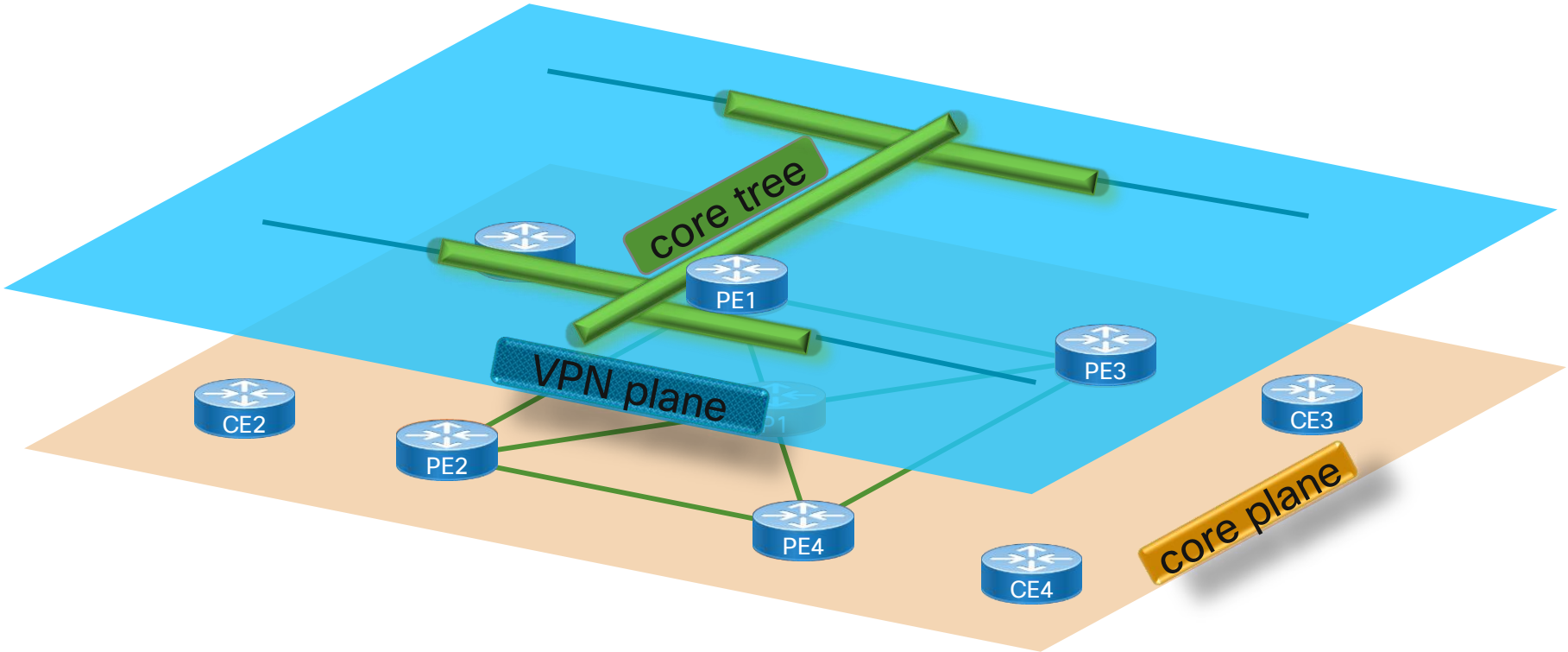



- PIM
- mLDP
- P2MP TE
- IR

# Planes – Overlay

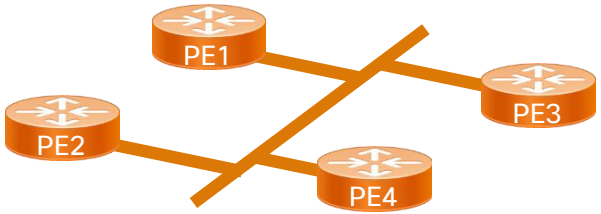


# Planes – Overlay Signaling



# Core Tree Types

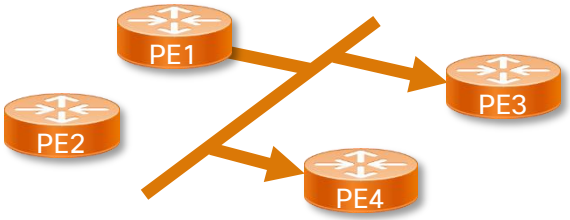
## Default MDT




- Connects all PEs
- Bidirectional
- Always present

Multi-Directional Inclusive PMSI
MI-PMSI

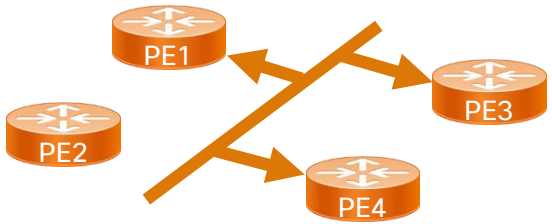
## Data MDT




- Connects subset of PEs
- Unidirectional
- On-demand

Selective PMSI
S-PMSI

## Partitioned MDT



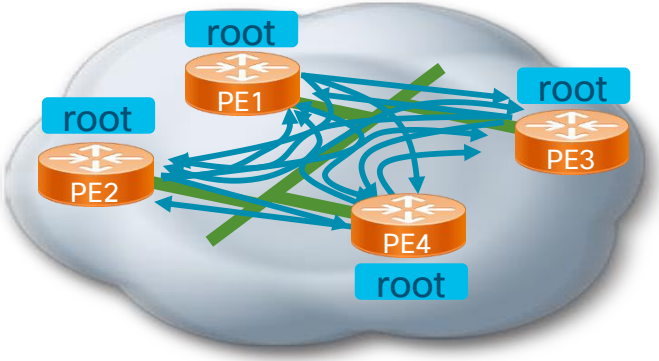

- Connects subset of PEs
- Uni- or Bidirectional
- On-demand

Multidirectional Selective PMSI
MS-PMSI

MDT = Multicast Distribution Tree
PMSI = Provider Multicast Service Interface

# Core Tree Type Construction – PIM

## Default MDT




S,G

or

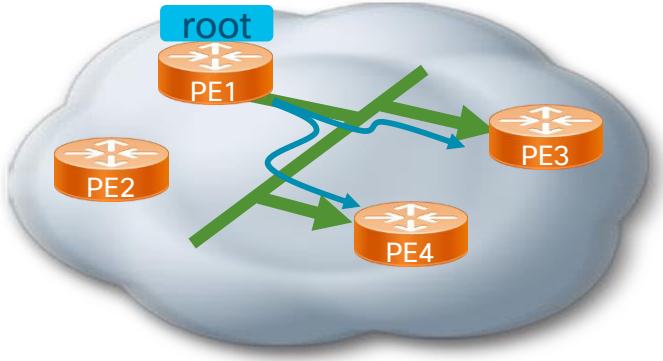
*,G

1.
2.
3.
...

1.
2.
3.
...

∑ S,G = Default MDT

## Data MDT




S,G

or

*,G

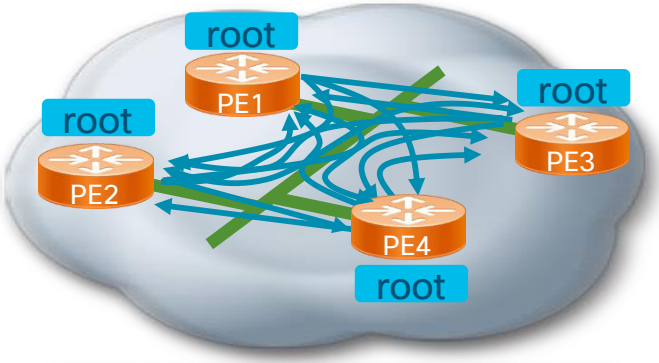
1.

1.

S,G = Data MDT

# Core Tree Type Construction – mLDP

Default MDT

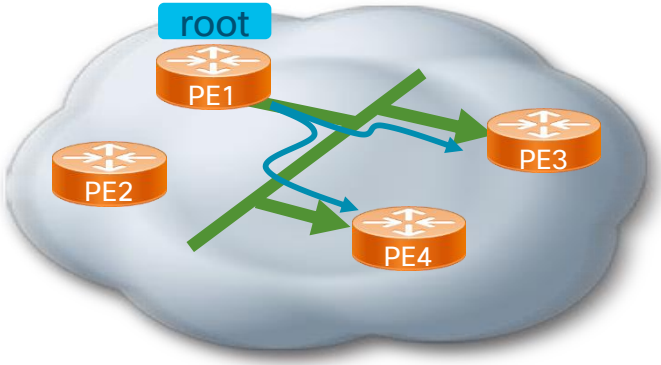



Full mesh of P2MP mLDP trees = Default MDT

or

1 MP2MP mLDP tree = Default MDT

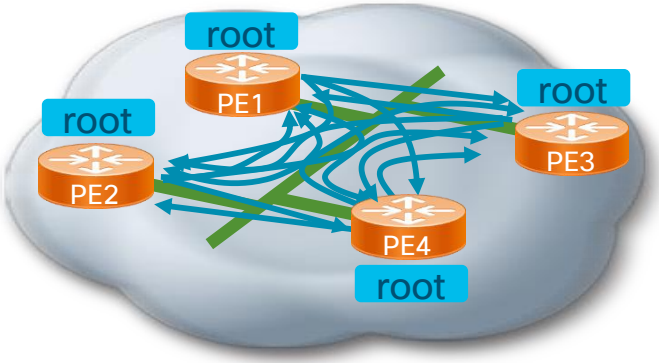
Data MDT




1 P2MP mLDP tree = Data MDT

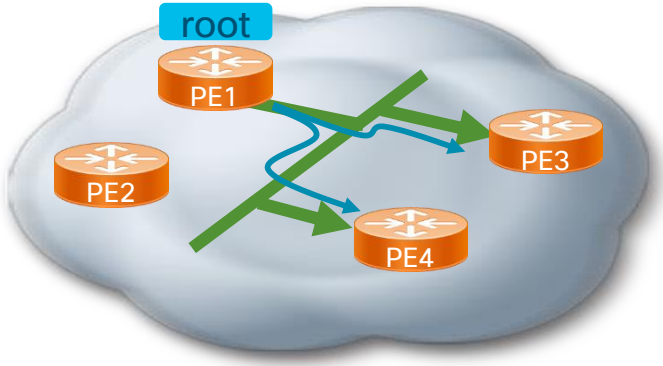
# Core Tree Type Construction – P2MP TE

Default MDT




Full mesh of P2MP TE tunnels
= Default MDT

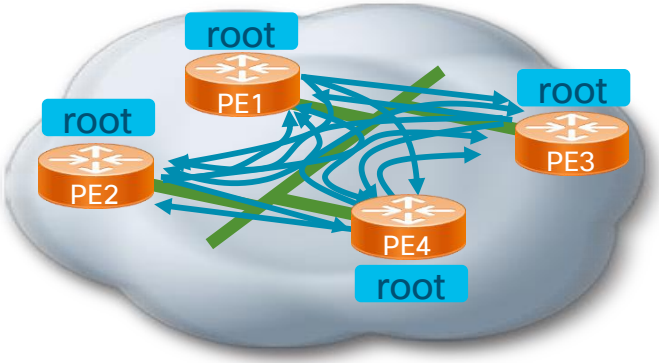
Data MDT




1 P2MP TE tunnel = Data MDT

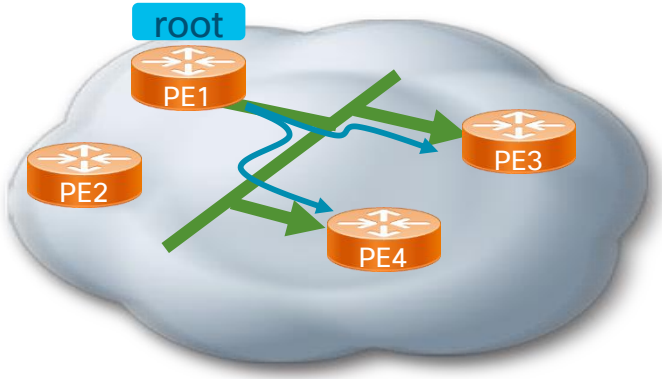
# Core Tree Type Construction – IR

Default MDT



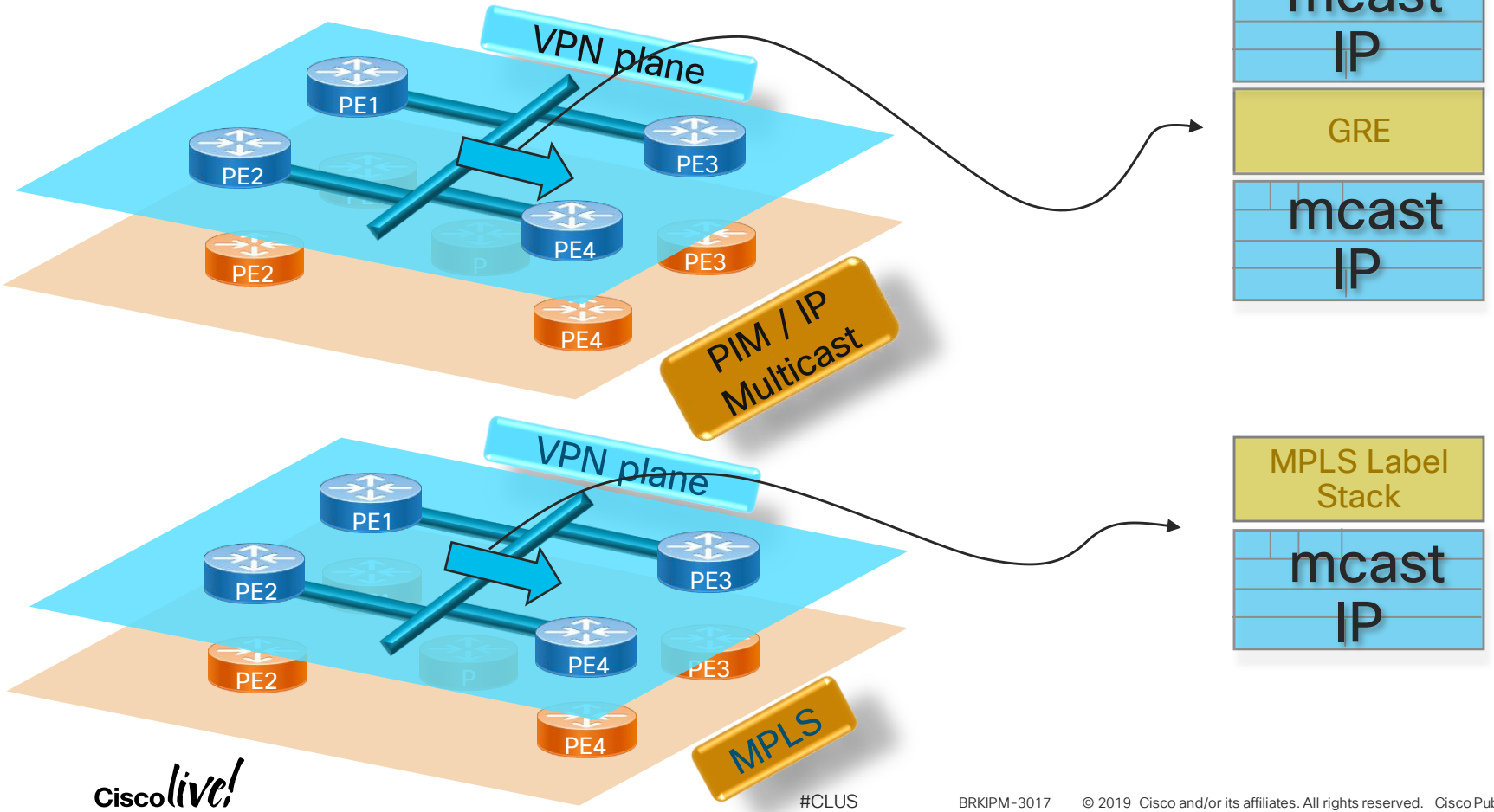

Full mesh of IR LSP sets = Default MDT

Data MDT




1 IR LSP set = Data MDT

# Encapsulation



# Core Trees

You make networking **possible**

Cisco live!

# Core Trees

- PIM
- mLDP
- P2MP TE
- Ingress Replication

You make security **possible**

# Core Tree - PIM

- PIM/IP Multicast as we know it
  - Pretty complex, but well-known
- (*,G) and (S,G) / PIM Sparse Mode, PIM SSM, PIM BiDir
- PIM for core tree
- Used for Default Model / Rosen
- Replication of multicast on the core routers

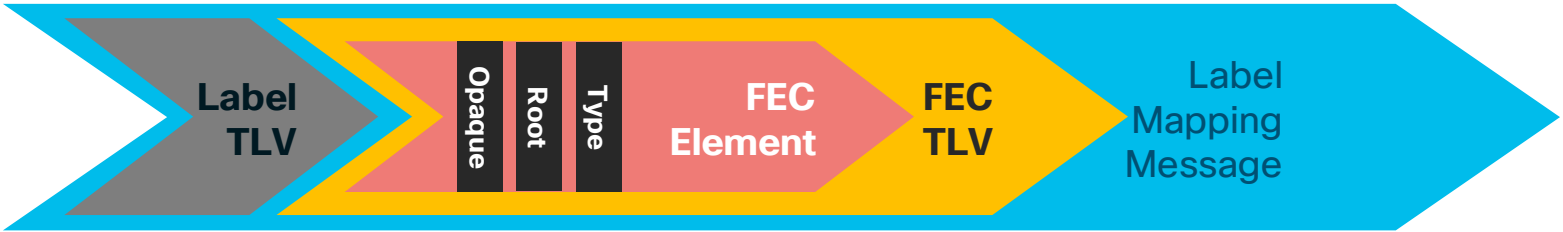
# Core Trees

- PIM
- mLDP
- P2MP TE
- Ingress Replication

You make security **possible**

# Core Tree - mLDP

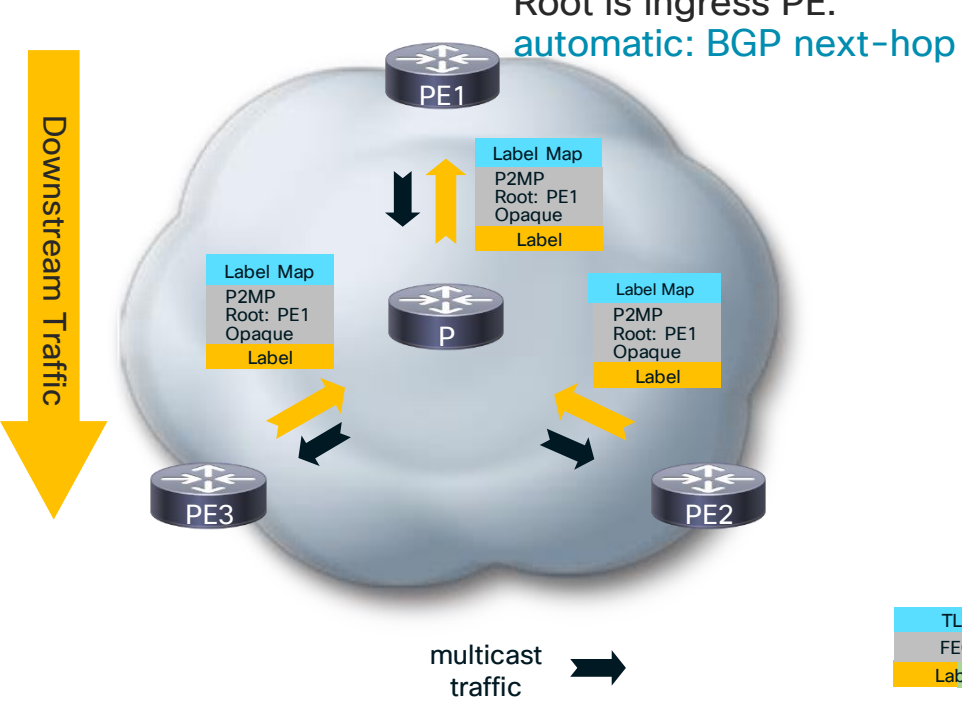
- Multipoint LDP = LDP + extensions
- P2MP tree - Receiver driven – Root learned from routing
- MP2MP tree – Configuration driven – Root configured
- Protection by MPLS TE or Loop-Free Alternate (LFA)
- No PHP – top label identifies the tree
- Replication of mcast on the core routers
- FEC elements holds: Type of tree + Root + Opaque value: (S,G), MDT number, LSP ID, …



# mLDP Signaling and Packet forwarding

## P2MP Tree

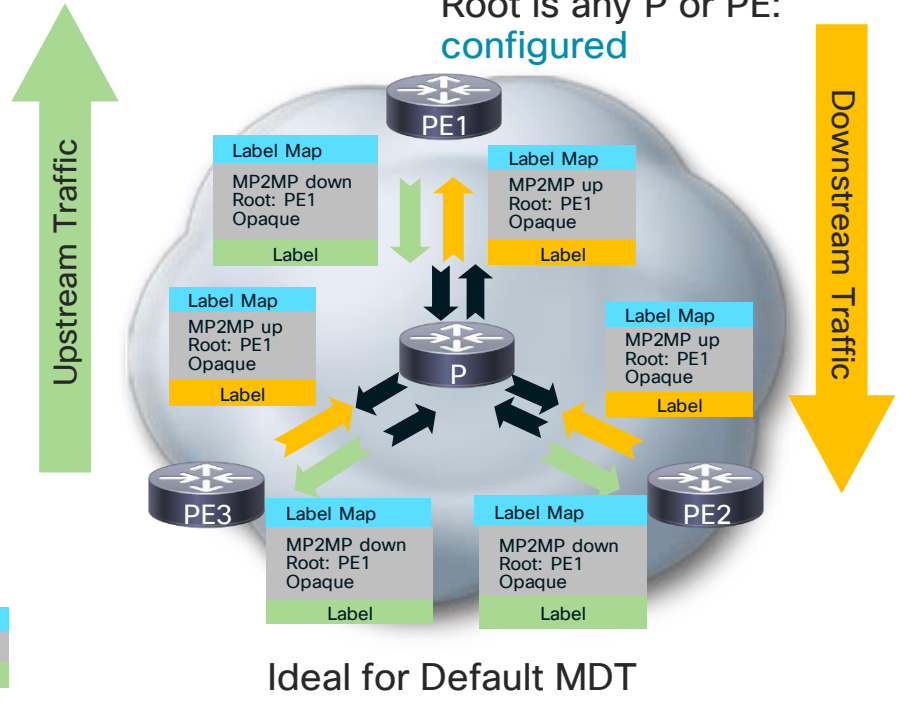
Root is ingress PE:
automatic: BGP next-hop




TLV
FEC
Label

## MP2MP Tree

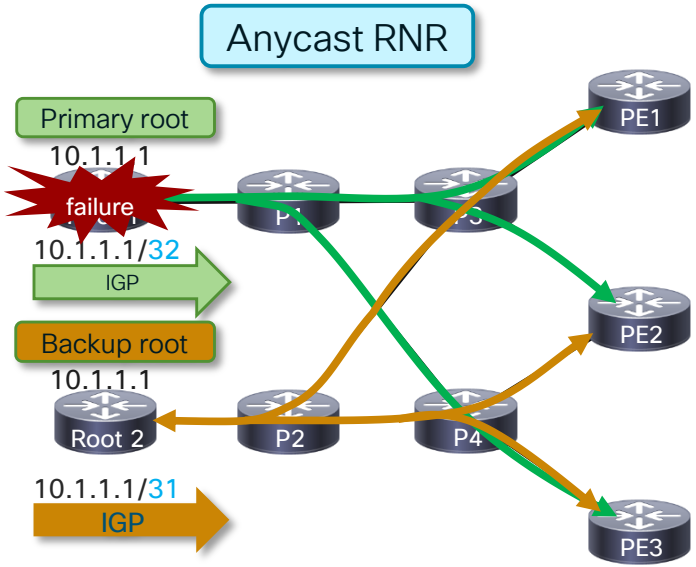
Root is any P or PE:
configured



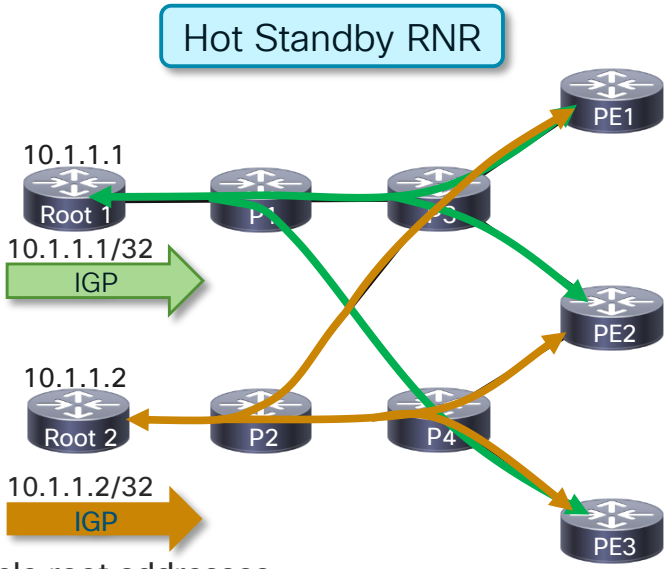

Ideal for Default MDT
Data plane is still P2MP

# MP2MP Tree and Root

- One root = single point of failure
- Multiple roots (Root Node Redundancy) → multiple trees




- One root address only
- Longest match routing
- 10.1.1.1 can be phantom




- Multiple root addresses
- Both trees active at the same time
- PE router selects preferred root and sends onto that tree only (no duplicate traffic), but receives from any tree

# mLDP MoFRR

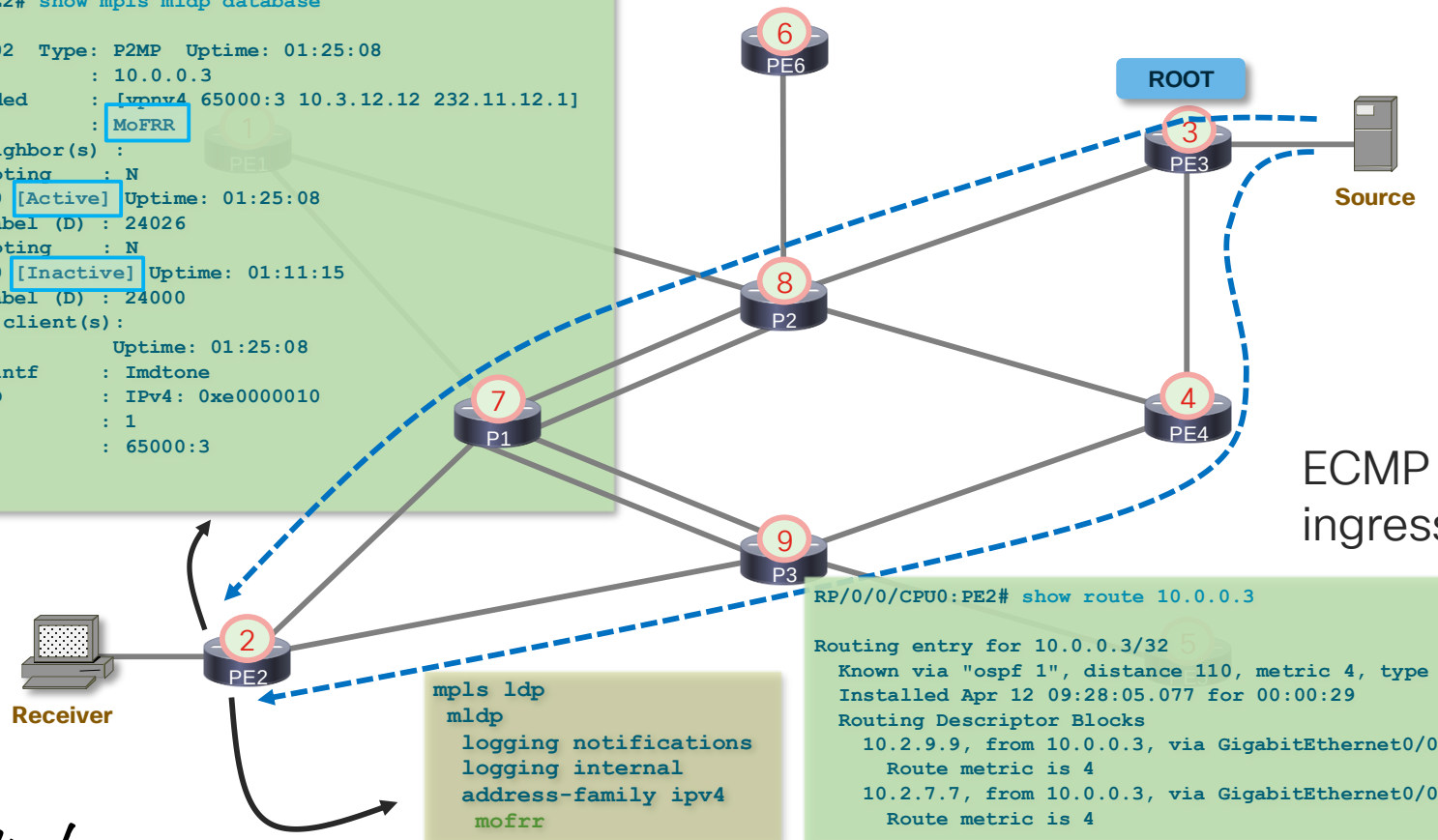
- Multicast Only Fast ReRoute
- Requirements:
  - Only inband signalling
  - Multipath: IGP
  - Only RIB-based, not Netflow-based
    - 100's msec of restoration
- Egress PE initiates two mLDP trees towards ingress PE
- Traffic is forwarded on two trees, egress PE forwards only from one tree towards receiver
- If primary path goes down, egress PE starts forwarding from backup mLDP path/tree

UCMP ( + variance) is possible

```
router ospf 1
 router-id 10.100.1.2
 ucmp variance 500
```

# mLDP MoFRR

```
RP/0/0/CPU0:PE2# show mpls mldp database
mLDP database
LSM-ID: 0x00002  Type: P2MP  Uptime: 01:25:08
  FEC Root           : 10.0.0.3
  Opaque decoded     : [vpnv4 65000:3 10.3.12.12 232.11.12.1]
  Features           : MoFRR
  Upstream neighbor(s) :
  Is CSI accepting   : N
    10.0.0.9:0 [Active] Uptime: 01:25:08
      Local Label (D) : 24026
  Is CSI accepting   : N
    10.0.0.7:0 [Inactive] Uptime: 01:11:15
      Local Label (D) : 24000
  Downstream client(s):
    PIM MDT           Uptime: 01:25:08
      Egress intf     : Imdtone
      Table ID        : IPv4: 0xe0000010
      RPF ID          : 1
      RD              : 65000:3
```




ECMP towards ingress PE

```
mpls ldp
 mldp
  logging notifications
  logging internal
  address-family ipv4
   mofrr
```

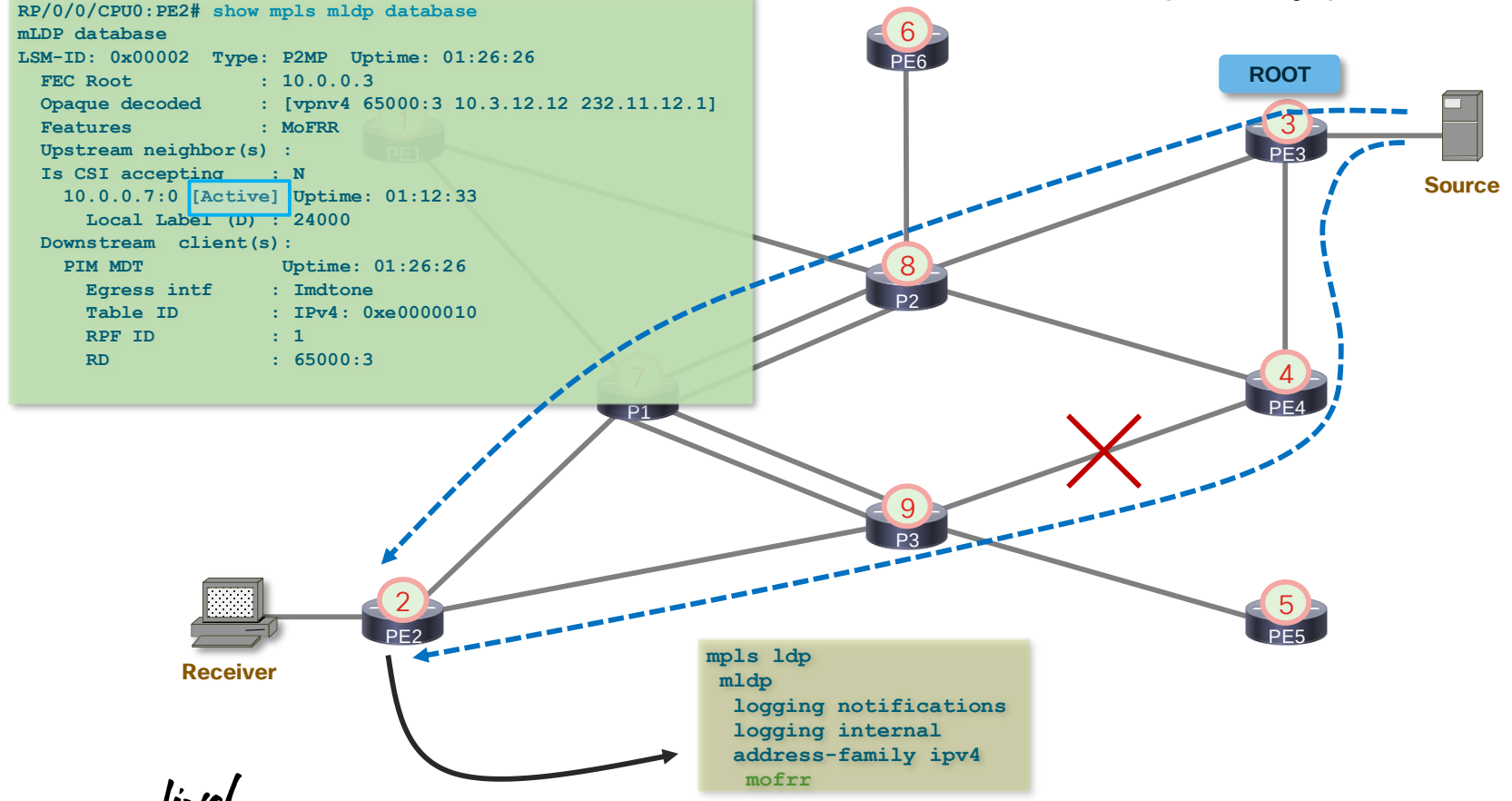
```
RP/0/0/CPU0:PE2# show route 10.0.0.3

Routing entry for 10.0.0.3/32
  Known via "ospf 1", distance 110, metric 4, type intra area
  Installed Apr 12 09:28:05.077 for 00:00:29
  Routing Descriptor Blocks
    10.2.9.9, from 10.0.0.3, via GigabitEthernet0/0/0/0
      Route metric is 4
    10.2.7.7, from 10.0.0.3, via GigabitEthernet0/0/0/1
      Route metric is 4
```

# mLDP MoFRR

primary path is down

```
RP/0/0/CPU0:PE2# show mpls mldp database
mLDP database
LSM-ID: 0x00002  Type: P2MP  Uptime: 01:26:26
  FEC Root           : 10.0.0.3
  Opaque decoded     : [vpnv4 65000:3 10.3.12.12 232.11.12.1]
  Features           : MoFRR
  Upstream neighbor(s) :
  Is CSI accepting   : N
    10.0.0.7:0 [Active] Uptime: 01:12:33
      Local Label (D) : 24000
  Downstream  client(s):
    PIM MDT            Uptime: 01:26:26
      Egress intf     : Imdtone
      Table ID        : IPv4: 0xe0000010
      RPF ID          : 1
      RD              : 65000:3
```




```
mpls ldp
 mldp
  logging notifications
  logging internal
  address-family ipv4
   mofrr
```

# mLDP-only SAC

- RFC 7473: State Advertisement Control for Non-negotiated LDP apps
- Have an LDP peer negotiate to advertise label bindings for certain MPLS apps or not by means of capability exchange at LDP session establishment
- Configure LDP to negotiate the label advertisement for IPv4, IPv6, FEC128, FEC129, and mLDP
- Request: run (m)LDP for advertisement of mLDP label bindings, but not for unicast label bindings
  - Use-case: Segment Routing network (no LDP for unicast needed)

```
RP/0/0/CPU0:PE(config-ldp)# capabilities sac ?
  fec128-disable  Disable exchanging PW FEC128 label bindings
  fec129-disable  Disable exchanging PW FEC129 label bindings
  ipv4-disable    Disable exchanging IPv4 prefix label bindings
  ipv6-disable    Disable exchanging IPv6 prefix label bindings
  mldp-only       Only exchange mLDP label bindings
  <cr>
```

# Core Trees

- PIM
- mLDP
- P2MP TE
- Ingress Replication

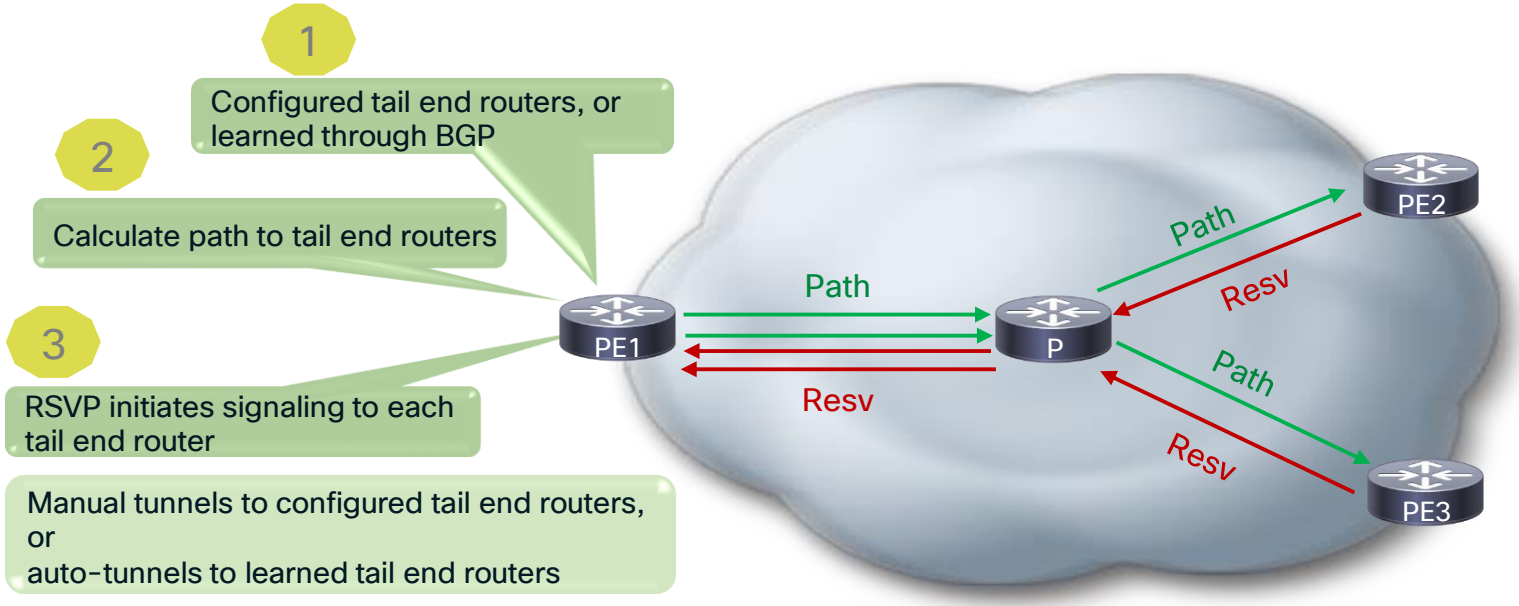
You make security **possible**

# Core Tree – P2MP TE

- Explicit (source) routing
- Bandwidth reservation
- Fast ReRoute (FRR) protection
- Uses RSVP for TE
- P2MP: extensions for RSVP-TE and IGP
- P2MP TE: looks and feels like P2P TE
- Replication of multicast on the core routers

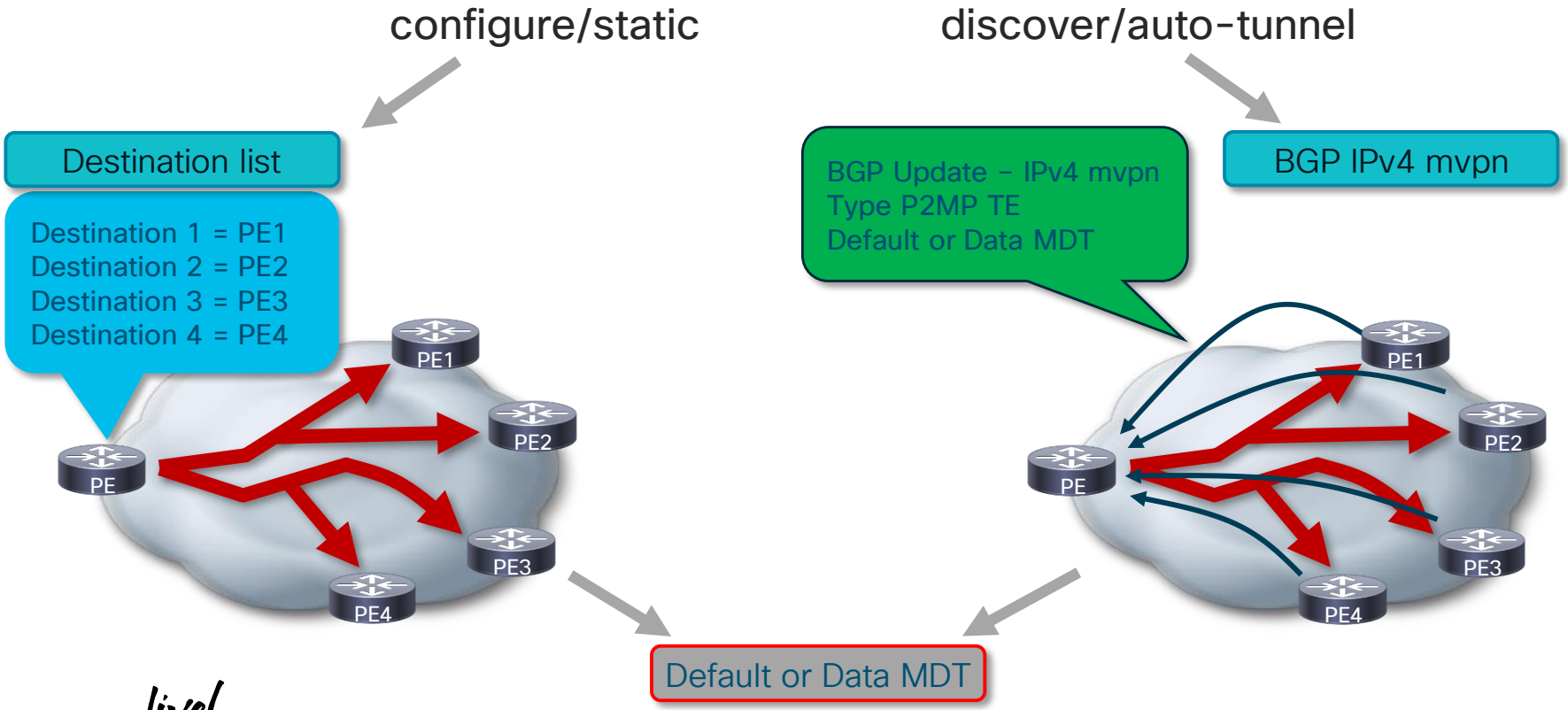
# P2MP TE

- P2MP tunnel signaled by RSVP, to multiple tail end routers



# P2MP TE - Tunnels

- Configure or discover the tail end routers of P2MP TE tunnels



# Core Trees

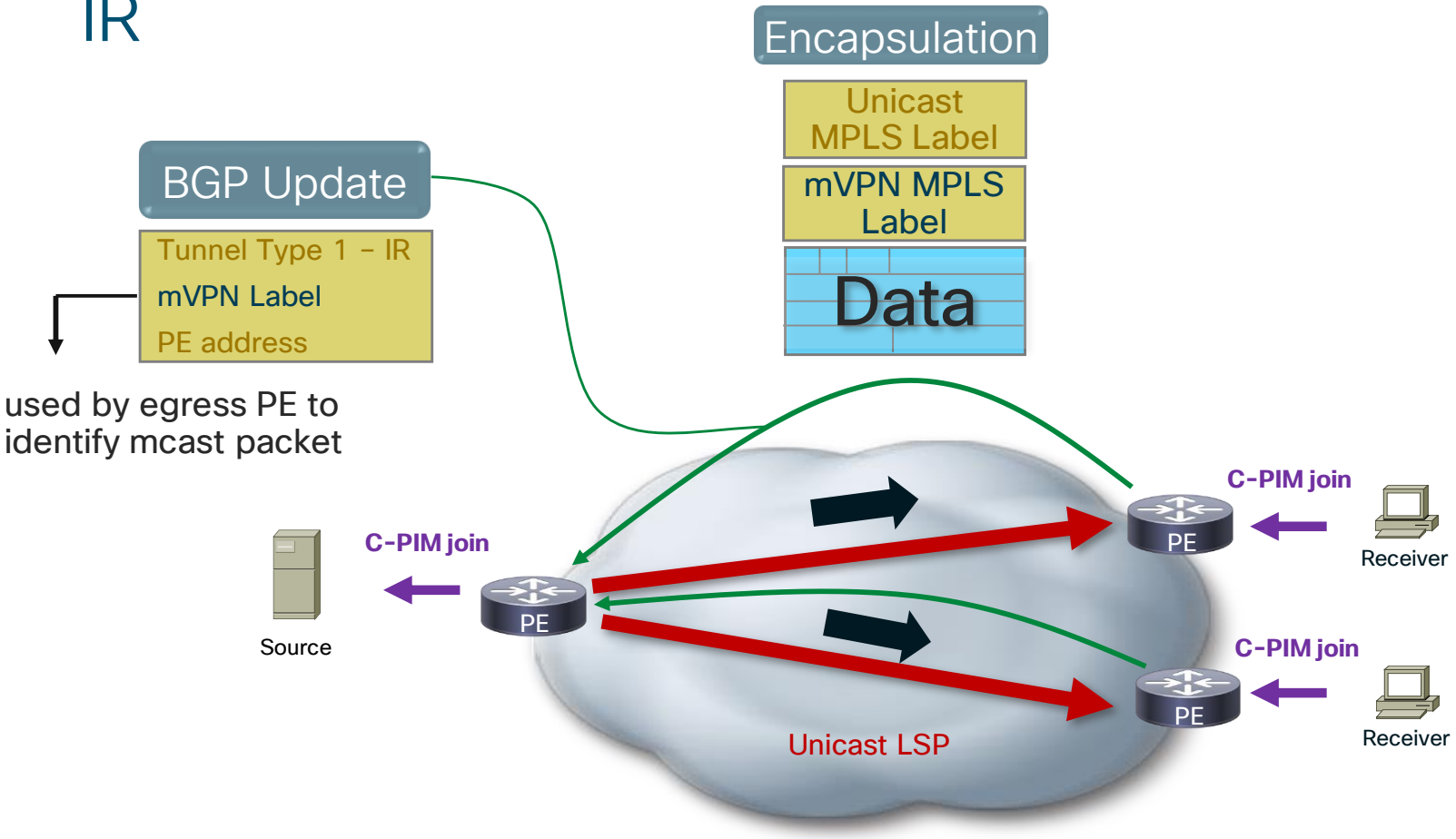
- PIM
- mLDP
- P2MP TE
- Ingress Replication

You make security **possible**

# Core Tree - IR

- No replication of multicast on the core routers
- Let's not use a P2MP tree in the core
- Let's re-use unicast MPLS Label Switched Paths (LSPs)
- Ingress replication!
- Protection by LFA/Ti-LFA
- Used when
  - Routers do not understand P2MP TE, mLDP or other inter-op issues
  - Simple solution is good enough: inter-as links
  - Amount of traffic is low – could be used in certain part of the network
- Packets have extra MPLS label to differentiate unicast vs. multicast traffic on same LSP
- BGP AD is needed to transport mVPN MPLS label

# IR



# Comparison Core Trees

| PIM | mLDP | P2MP TE |
|---|---|---|
| Mature / Well known | Enhancement to existing protocol | Enhancement to existing protocol |
| Soft state (periodic refresh) | Hard state (no periodic updates) | Soft state (periodic refresh) |
| GRE enacpsulation | MPLS encapsulation | MPLS encapsulation |
| No Fast Restoration | Fast Restoration (povided by LFA or MPLS TE) | Fast Restoration |
| No bandwidth reservation | No bandwidth reservation | Bandwidth reservation |
| P2MP trees only | P2MP and MP2MP trees | P2MP trees only |
| Inter-as and CsC | Inter-as and CsC | Inter-as, but no CsC |
| High complexity | Medium complexity | High complexity |
| Medium core state<br>C-state present in core with Data MDT | Medium core state<br>C-state present in core with Data MDT<br>All C-state in core with in-band signaling | High core state |
| Follows unicast routing | Follows unicast routing | Allows explicit or bandwidth contraint routing |
| Suitable for all mcast applications | Suitable for all mcast applications<br>Best for many-to-many | Mostly suitable for video delivery<br>Best for few-to-many |

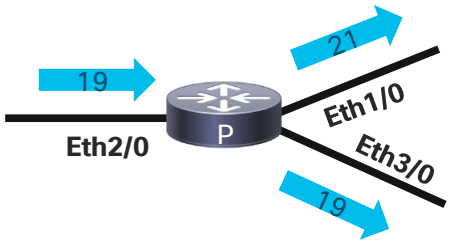
# Packet Forwarding Labeled Switched Multicast

P2MP

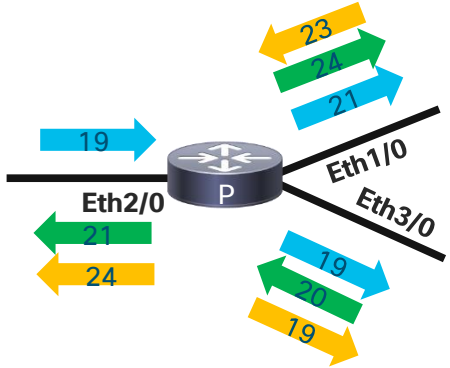
- P2MP TE
- P2MP mLDP

MP2MP

- MP2MP mLDP




Packet replication on core routers




```
P# show mpls forwarding-table
Local      Outgoing    Prefix           Outgoing    Next Hop
Label      Label       or Tunnel Id     interface

19           21        blue LSP         Et1/0       10.1.1.1
             19        blue LSP         Et3/0       10.1.3.3
```

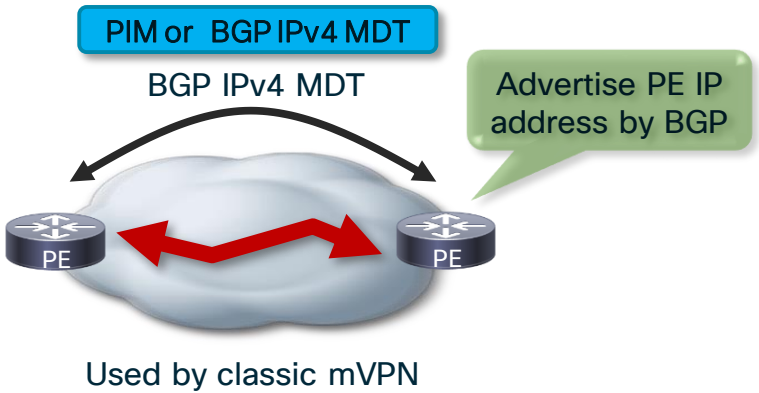
```
P# show mpls forwarding-table
Local      Outgoing    Prefix           Outgoing    Next Hop
Label      Label       or Tunnel Id     interface

19         21          blue LSP         Et1/0       10.1.2.2
           19          blue LSP         Et3/0       10.1.1.1
20         24          green LSP        Et1/0       10.1.3.3
           21          green LSP        Et2/0       10.1.2.2
23         24          orange LSP       Et2/0       10.1.3.3
           19          orange LSP       Et3/0       10.1.1.1
```

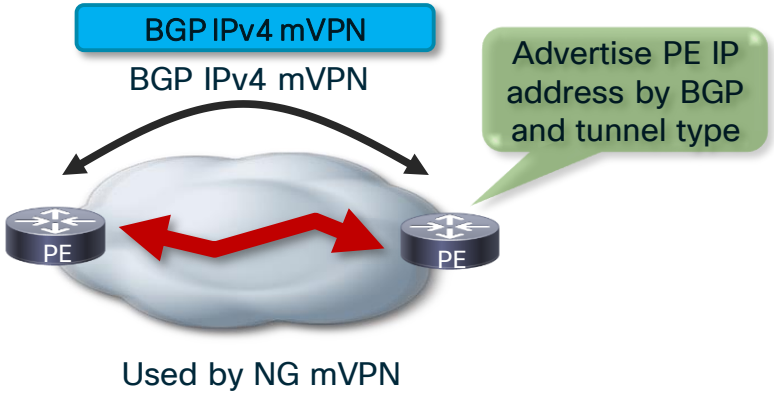
P2MP forwarding in data plane!

# Auto-Discovery

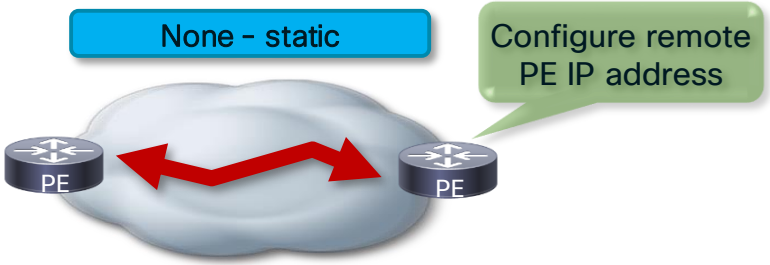
- The process of discovering all the PEs with members in a given mVPN
- In order to build the MDT




Used by classic mVPN




Used by NG mVPN




Rarely used (e.g. P2MP TE)

# FRR

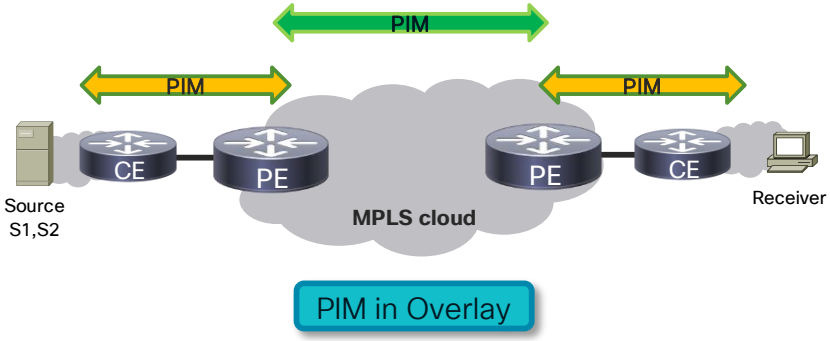
- PIM
  - No FRR
  - Rerouting as fast as IGP
    - Tune IGP for unicast
- P2MP TE
  - Inherent FRR using backup tunnels
  - Link/node protection
  - Auto-tunnels possible
- mLDP
  - mLDP with TE FRR
  - mLDP with LFA (IP FRR)

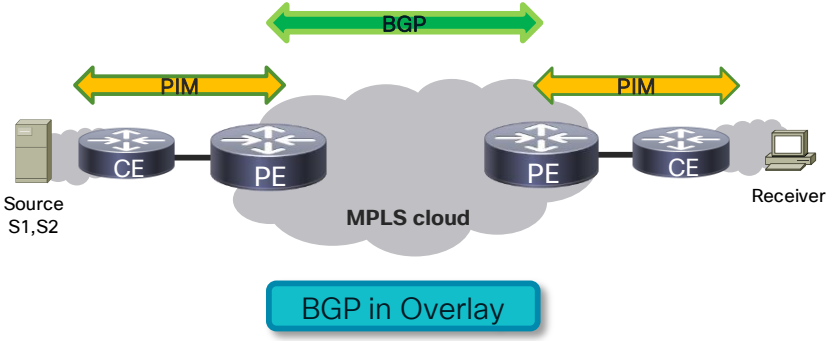
# Overlay Signaling

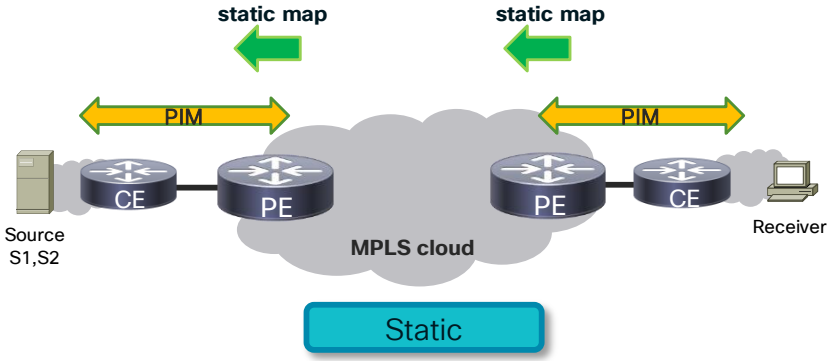
You make networking **possible**

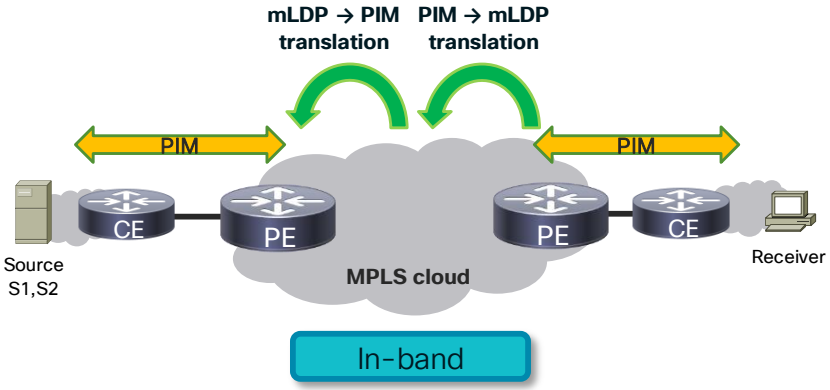
Cisco live!

# Overlay Signaling Possibilities



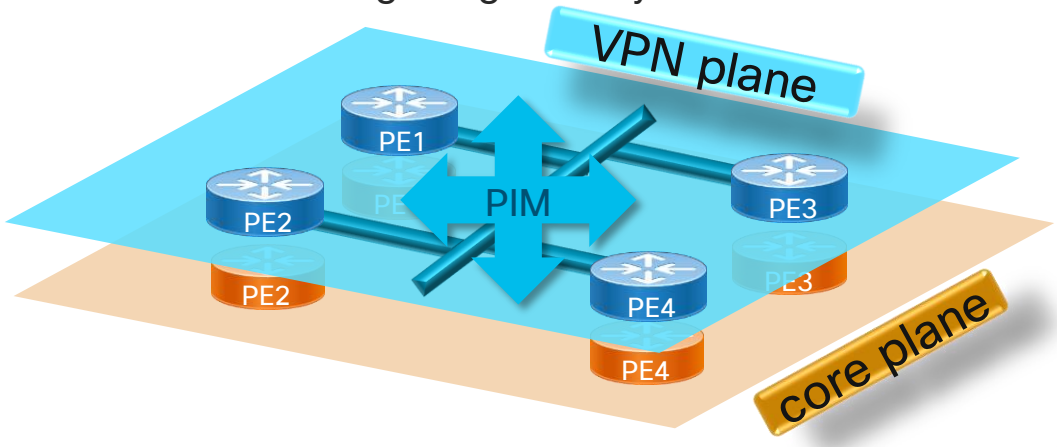



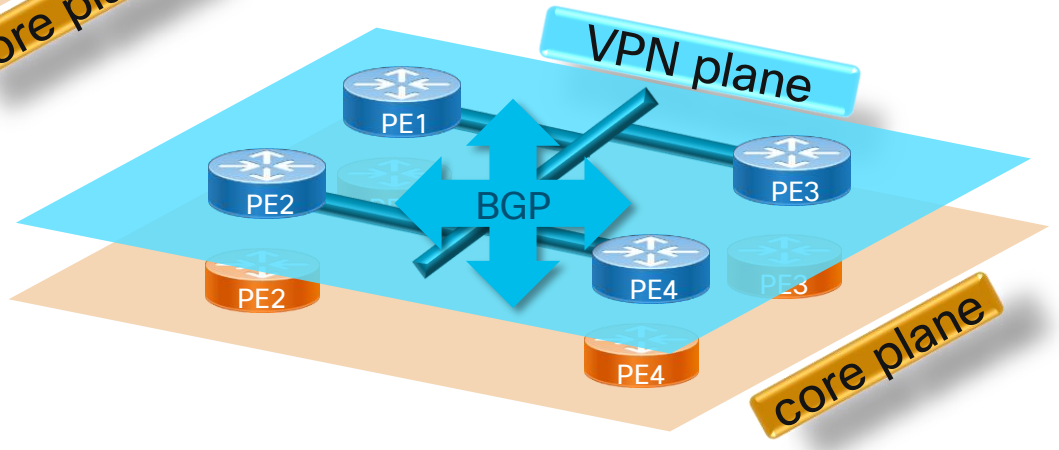
# Overlay Signaling

- PE-CE mcast signaling is always PIM




- PIM neighborship: one-hop between PEs over Default (or Partitioned) MDT
- PIM signaling: Joins/Prunes/Asserts one-hop over Default (Partitioned) MDT
- (*,G) or (S,G) over MDT

- iBGP multihop signalling between PEs
- BGP was not designed for mcast
- BGP can provide Auto-Discovery
- Signal (*,G) and (S,G)



# BGP Signaling

- BGP scales well
- BGP is not a multicast signaling protocol per design
  - Receiver to Source signaling …
  - PIM Sparse Mode works differently in BGP
    → new procedures
- New address family “IPv4 mVPN”
  1. Signal Auto-Discovery (AD)
  2. Signal multicast information
     - (*,G) or (S,G)
     - Which tunnel to use (core tree protocol and tunnel type)

```
vrf definition one
rd 1:1
!
address-family ipv4
  mdt auto-discovery mldp
  mdt default mpls mldp p2mp
  mdt overlay use-bgp
  route-target export 1:1
  route-target import 1:1

router bgp 1
 neighbor 10.100.1.7 remote-as 1
 neighbor 10.100.1.7 update-source Loopback0
!
 address-family ipv4 mvpn
  neighbor 10.100.1.7 activate
  neighbor 10.100.1.7 send-community extended
```

# A New Found Role for BGP

**Auto-Discovery**

Discovering PE endpoints automatically
- Replacing manual configuration in case of MPLS TE
- Replacing some PIM signalling, signalling Data MDT

**Multicast Signalling**

Control plane replacing PIM
- Shared tree (*,G)
- Source tree (S,G)

Joins, Prunes, Hellos

- New BGP address family
- PMSI Tunnel Attribute (PTA) information
  - Describes the core tree (PIM, mLDP, MPLS TE, IR)
- Prefix (NLRI)
  - Describes multicast state
  - Source, Group, Originator, Route Distinguisher

# BGP Address Family IPv4 mVPN

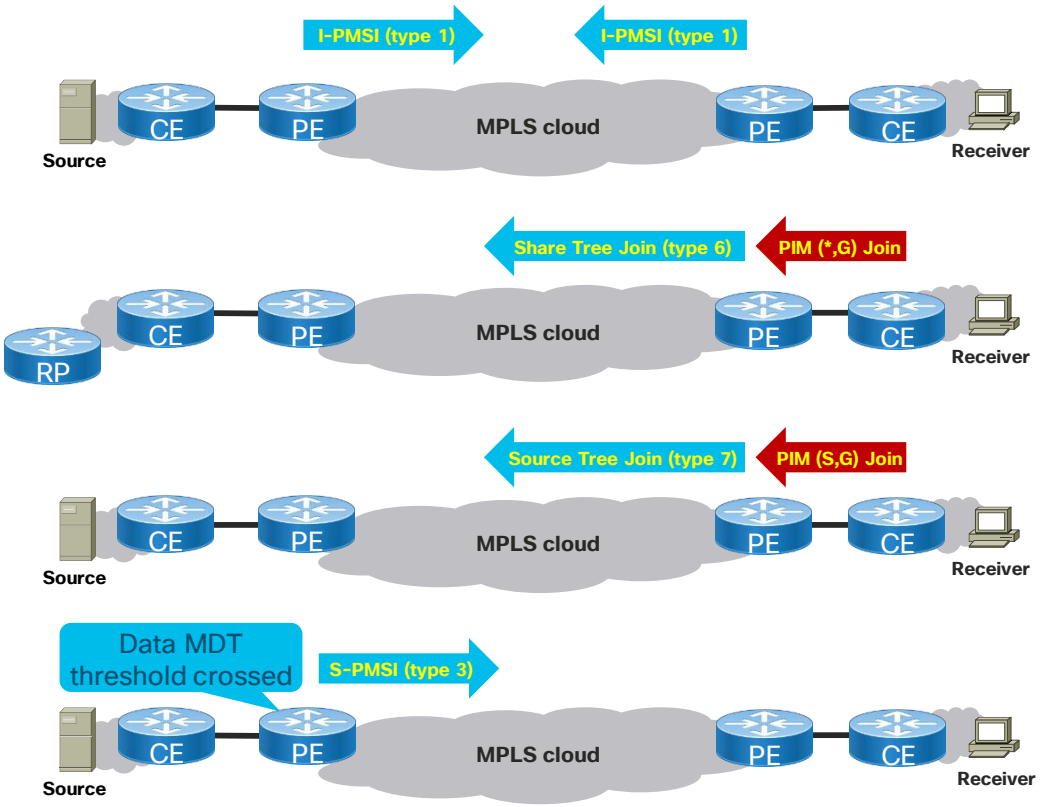

| Route Type | Meaning | Usage |
|---|---|---|
| 1 | Intra-AS I-PMSI A-D route | AD Signaling |
| 2 | Inter-AS I-PMSI A-D route | AD Signaling |
| 3 | S-PMSI A-D route | AD Signaling |
| 4 | Leaf A-D route | AD Signaling |
| 5 | Source Active A-D route | AD Signaling |
| 6 | Shared Tree Join route | C-signaling |
| 7 | Source Tree Join route | C-signaling |

Encoding can be : RD (8 octets) , MCAST source length (1 octet), MCAST source (variable) , MCAST group length (1 octet), MCAST group (variable), Originating router's IP address

| Tunnel Type | Meaning | Info encoded |
|---|---|---|
| 0 | No tunnel info present | - |
| 1 | P2MP TE tunnel | Ext tunnel ID / Tunnel ID / P2MP ID |
| 2 | mLDP P2MP | P2MP FEC Element |
| 3 | PIM SSM | Root address / P-Group |
| 4 | PIM Sparse Mode | Sender Address / P-Group |
| 5 | PIM BiDirectional | Sender Address / P-Group |
| 6 | Ingress Replication | Unicast tunnel endpoint address |
| 7 | mLDP MP2MP | MP2MP FEC Element |
| 8 | Transport Tunnel | Source PE address / local number |

# BGP Signaling

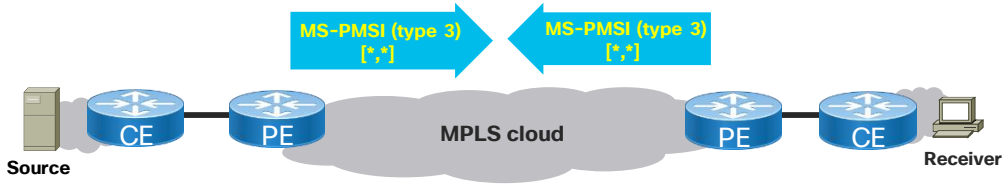
| | |
|---|---|
| Build Default MDT | config driven |
| Relay PIM Join | PIM driven |
| Relay PIM Join | PIM driven |
| Build Data MDT | data driven |



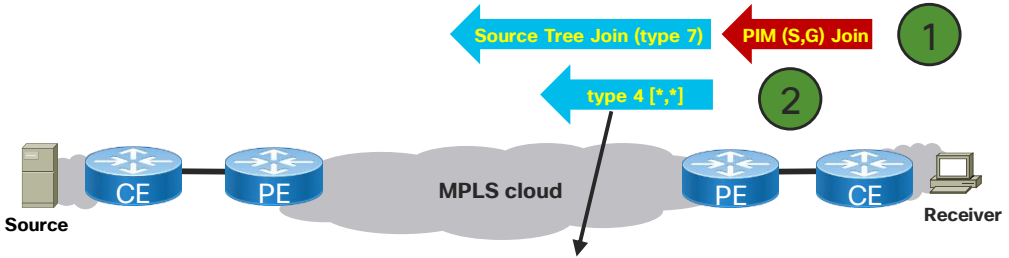
# BGP Signaling

**Type 3** — config driven — Partitioned MDT

LSP Identifier is [*,*]
PEs routers indicate they are willing to participate in the MS-PMSI (Partitioned MDT)




**Type 3** — PIM driven — Partitioned MDT — P2MP TE Auto-Tunnels




LSP Identifier is [*,*]
PE indicates it is egress PE
Head end needs to signal P2MP TE to egress PE

# BGP IPv4 mVPN Prefix Example: Route-Type 1

Route-type 1 = Intra-AS I-PMSI A-D route

```
RP/0/0/CPU0:PE1# show bgp ipv4 mvpn vrf one [1][10.100.1.7]/40
BGP routing table entry for [1][10.100.1.7]/40, Route Distinguisher: 1:5
 10.100.1.7 (metric 70) from 10.100.1.3 (10.100.1.7)
      Origin IGP, localpref 100, valid, internal, best, group-best, import-candidate, imported
      Received Path ID 0, Local Path ID 1, version 25
      Community: no-export
      Extended community: RT:100:100
      Originator: 10.100.1.7, Cluster list: 10.100.1.3
      PMSI: flags 0x00, type 7, label 0, ID 0x080001040a640107000e02000b000020000002000000000
      Source AFI: IPv4 MVPN, Source VRF: one, Source Route Distinguisher: 1:5
```

Tunnel type 7 = mLDP MP2MP

PMSI Tunnel ID / FEC Element

FEC Element Type : 8 : MP2MP-Down
AF Type : 1
Address Length : 4
Root Node Address : 10.100.1.7
MP Opaque Length : 14
MP Opaque Value Element
Opaque Type : 2 : MDT VPNID
Opaque Length : 11
OUI:VPN-Index (VPN ID) : 20:20 ... in hex
MDT ID : 0

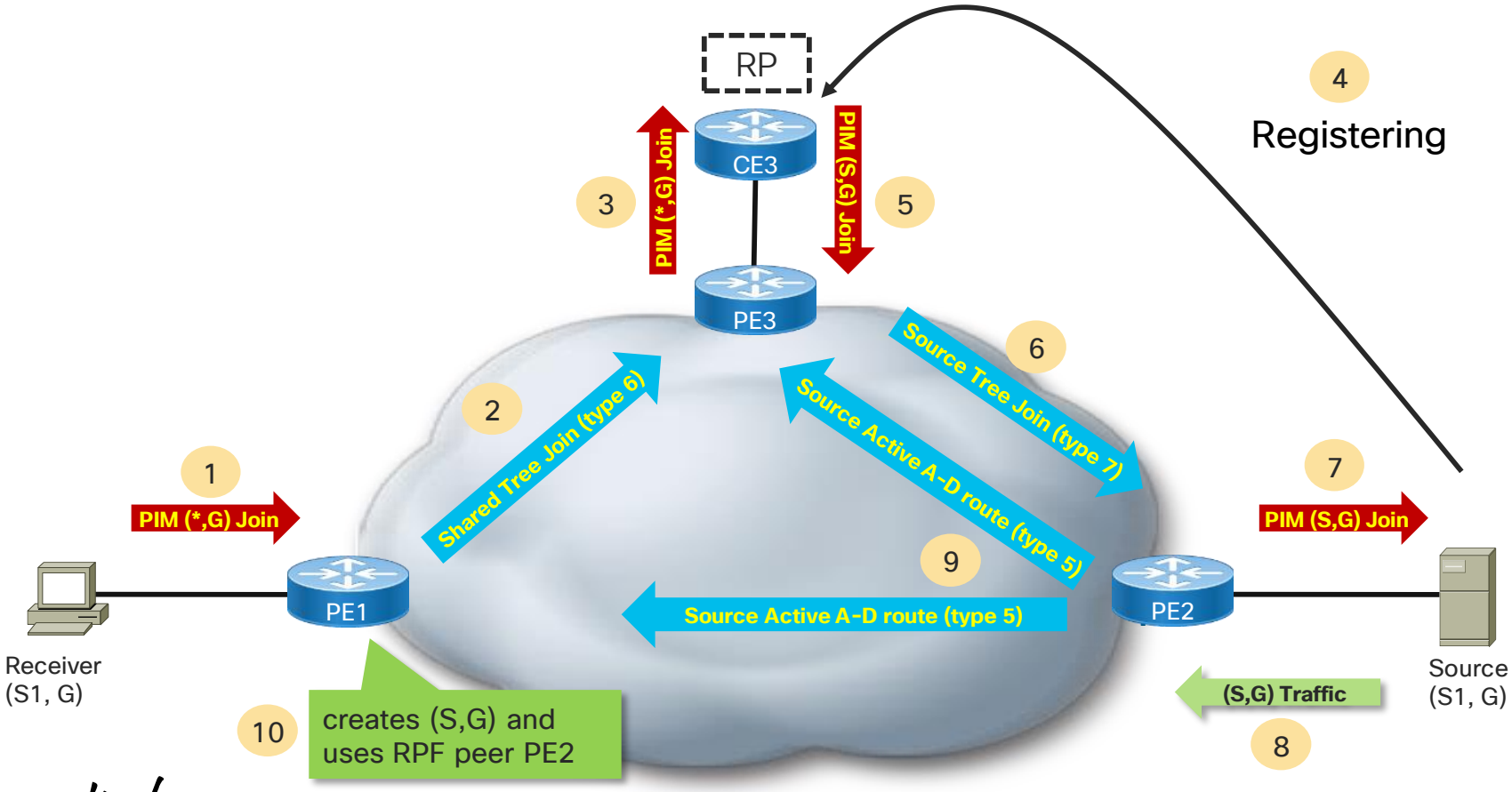
# BGP IPv4 mVPN Prefix Example: Route-Type 7

Route-type 7 = Source Tree Join route

```
RP/0/0/CPU0:PE1# show bgp ipv4 mvpn vrf one [7][1:5][1][32][10.2.1.8][32][232.1.1.1]/184
BGP routing table entry for [7][1:5][1][32][10.2.1.8][32][232.1.1.1]/184, Route Distinguisher:
1:5
 10.100.1.7 (metric 70) from 10.100.1.3 (10.100.1.7)
      Origin IGP, localpref 100, valid, internal, best, group-best, import-candidate, imported
      Received Path ID 0, Local Path ID 1, version 26
      Extended community: RT:10.100.1.1:17
      Originator: 10.100.1.7, Cluster list: 10.100.1.3
      Source AFI: IPv4 MVPN, Source VRF: one, Source Route Distinguisher: 1:5
```

No PMSI Tunnel Attribute (PTA)

# BGP Signaling for Sparse mode



# VRF Static over P2MP TE

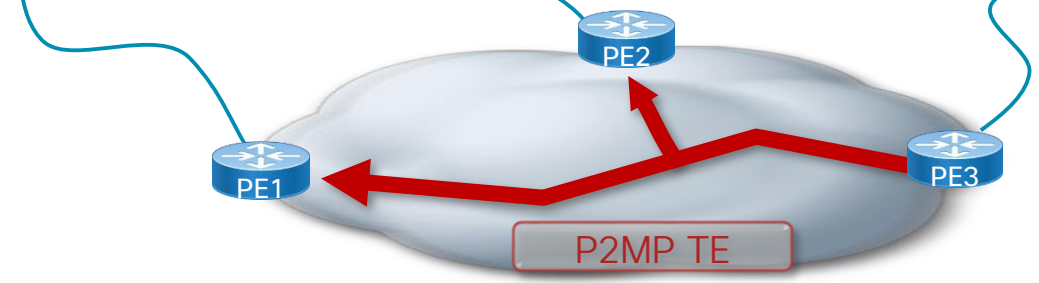
- The only model using static overlay signaling
- Allowing of aggregation of multiple C-(S,G) over one TE tunnel

head end has (S,G) mapping

```
interface Tunnel0
ip unnumbered Loopback0
ip pim passive
ip igmp static-group 232.1.1.1 source 10.2.3.10
tunnel mode mpls traffic-eng point-to-multipoint
tunnel destination list mpls traffic-eng name from-PE3
```

Tail-end has RPF statement → head end router

```
ip mroute 10.2.3.0 255.255.255.0 10.100.1.3
```

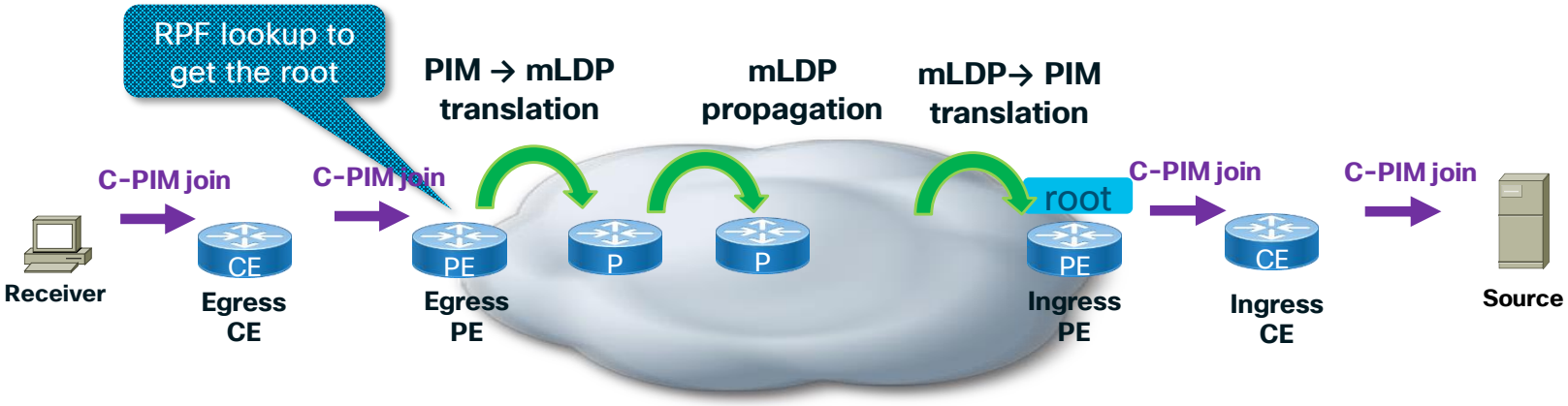


# In-band Signaling

mapping

| PIM Join | LDP Label Mapping Message |
|---|---|
| PIM Prune | LDP Label Withdraw Message |

- No overlay signaling
- Method to stitch a PIM tree to a mLDP LSP without any additional signaling
- PIM (or IGMP) can be mapped to mLDP



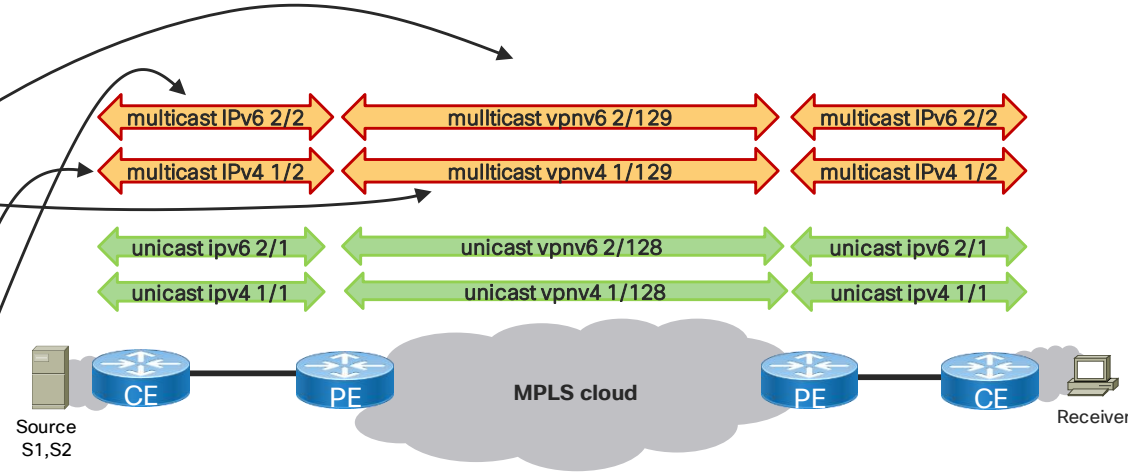
# SAFI (Subsequent Address Family Identifier) 2

- SAFI 2 in VRF and SAFI 129 across the core
  - Allows for different mcast vs unicast topologies across MPLS
  - SAFI 129 = VPN mcast SAFI
  - A PE can select an upstream mcast hop which is different than the unicast next hop (RPF is not the unicast route)

```
router bgp 1
 address-family vpnv4 unicast
 address-family vpnv6 unicast
 address-family vpnv4 multicast
 address-family vpnv6 multicast

and

address-family ipv4 unicast vrf vrf1
address-family ipv6 unicast vrf vrf1
address-family ipv4 multicast vrf vrf1
address-family ipv6 multicast vrf vrf1
```



# Comparison Customer Signaling Protocols

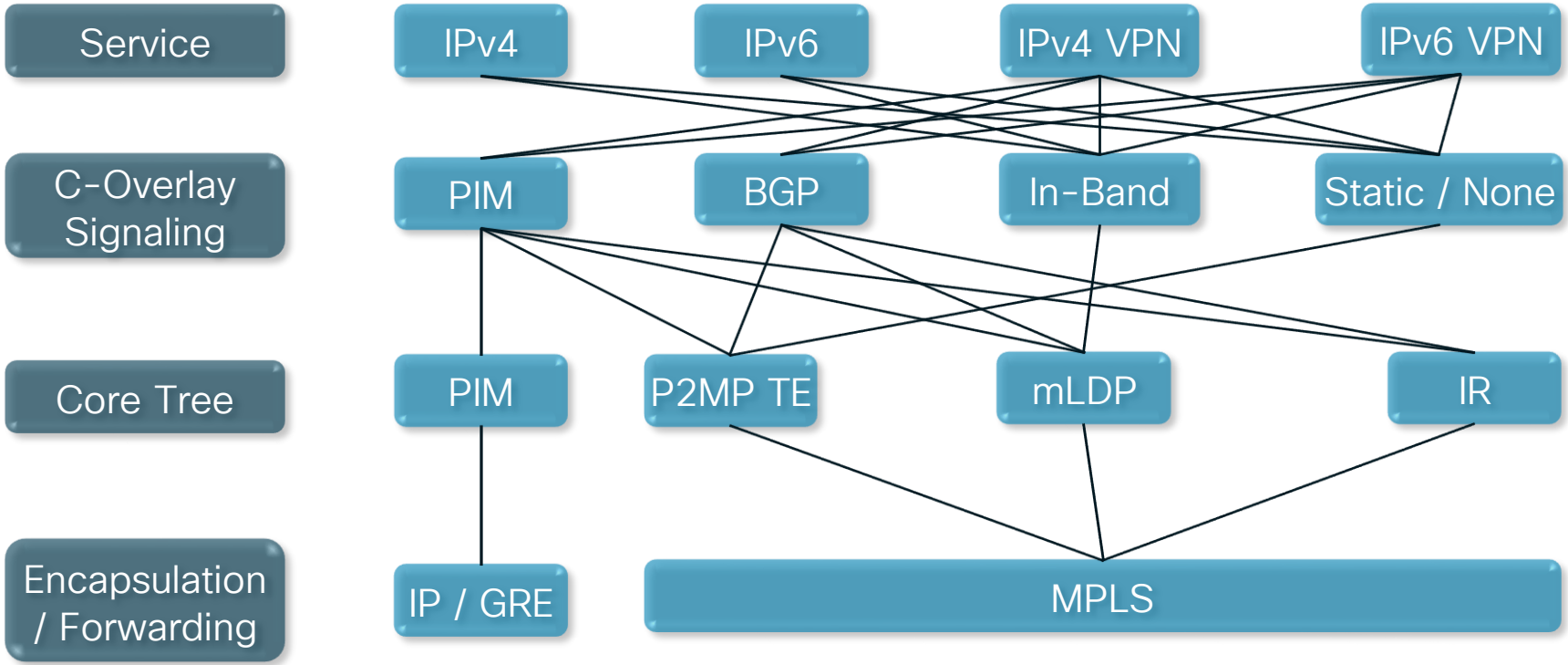
| PIM | BGP |
| --- | --- |
| Older protocol, proven, well known | New enhancement to existing protocol |
| No changes needed<br>Complex with ASM, but well known | New procedures (troubleshooting!)<br>Complex for ASM |
| Soft state (periodic refresh) | Hard state (no periodic updates) |
| Info driven to specific PE router | Info driven to all PE routers |
| PIM adjacencies to all PE routers | BGP adjacencies to all PE routers<br>but likely only to RRs |
| Medium scalability | Very high scalability |

# mVPN Deployment Models

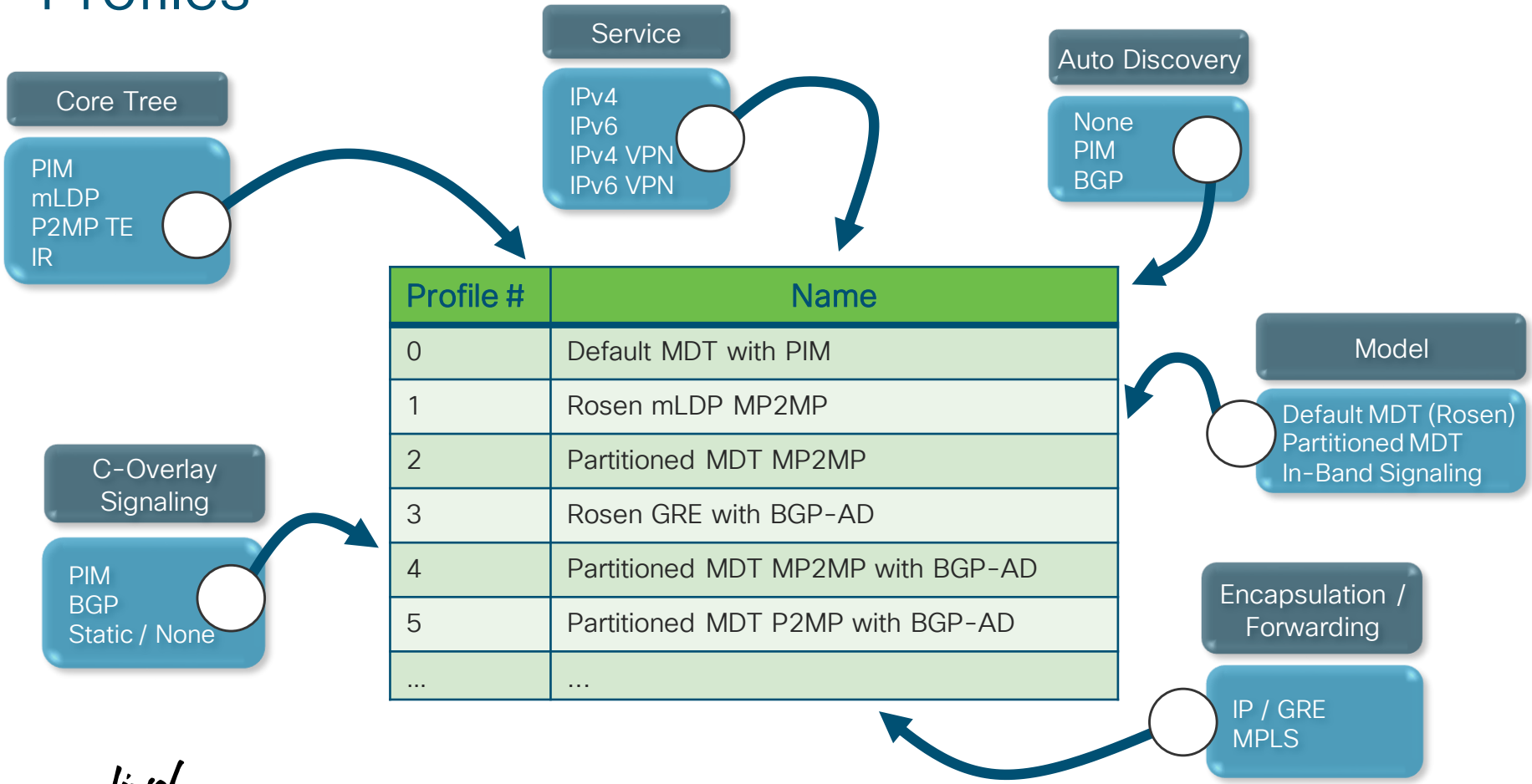
You make networking **possible**

Cisco live!

# Putting it all Together

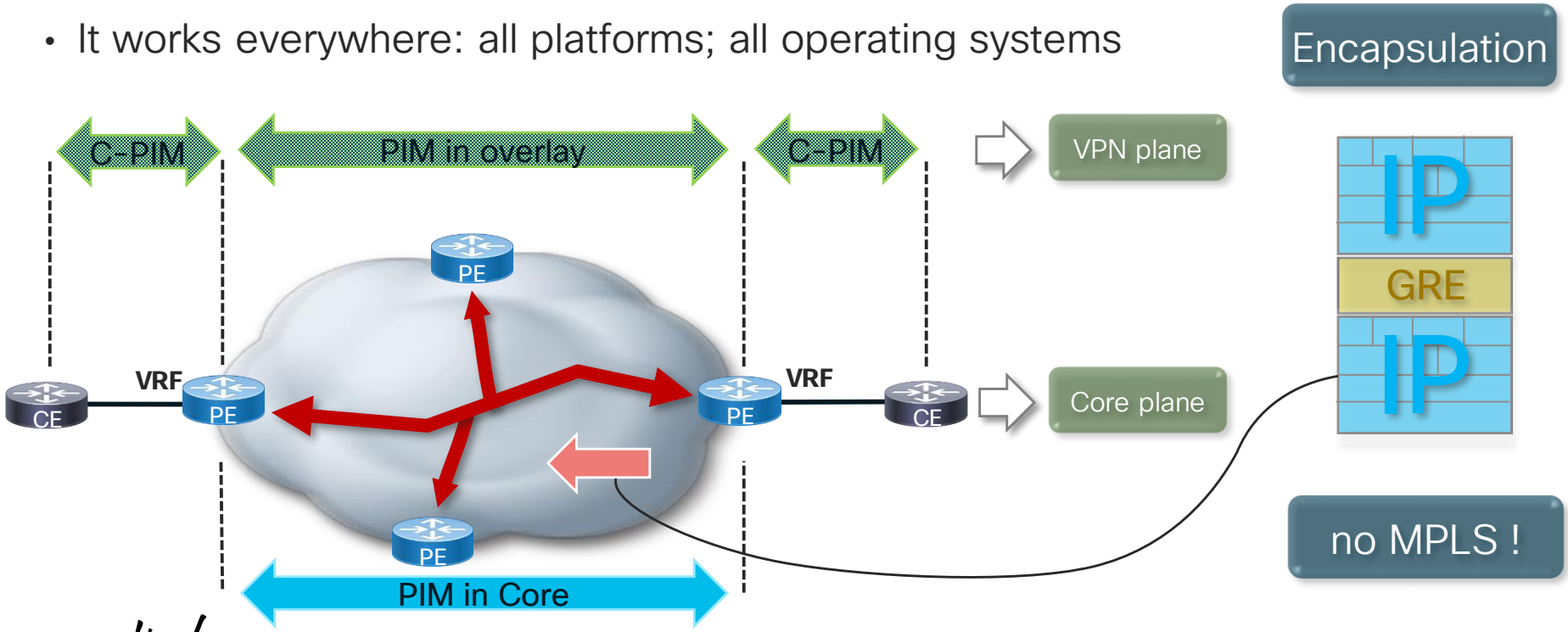


# Profiles

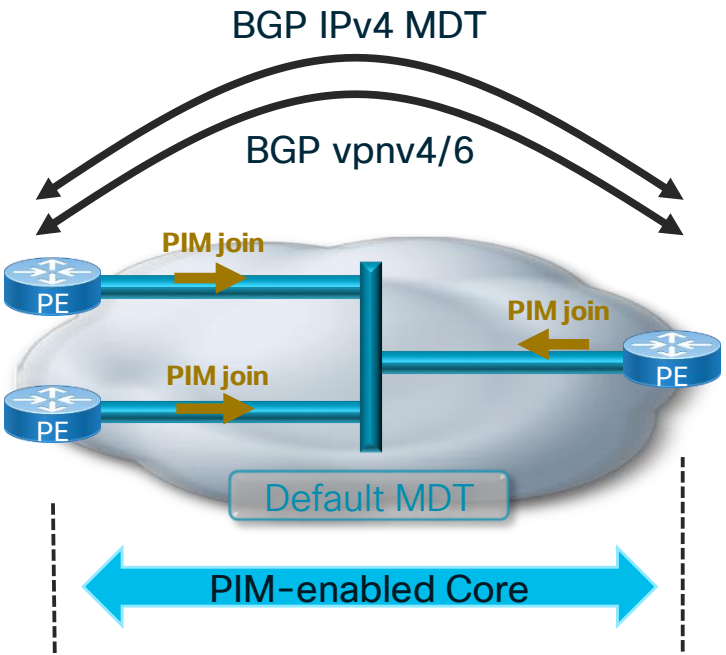


# mVPN Classic (aka Rosen / aka Default MDT)

- Since 2000
- It works everywhere: all platforms; all operating systems



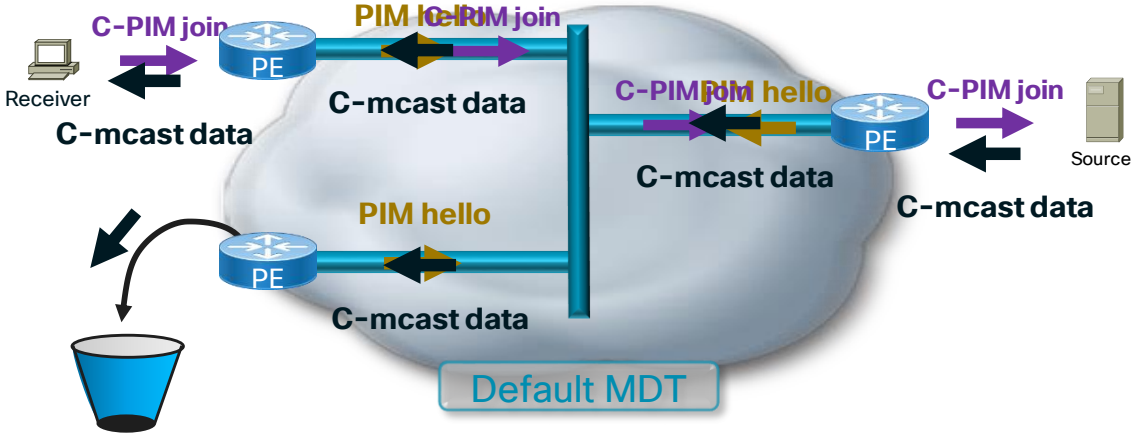
# Default MDT Recap




1. Unicast routing (RPF)
2. Build Default MDT
   - 2.1 Advertise PE as member of mVPN by BGP IPv4 MDT
   - 2.2 PE routers join Default MDT PIM tree

Default MDT connects all mVPN PE routers and carries all PIM signaling and all mcast traffic by default

# Default MDT Recap

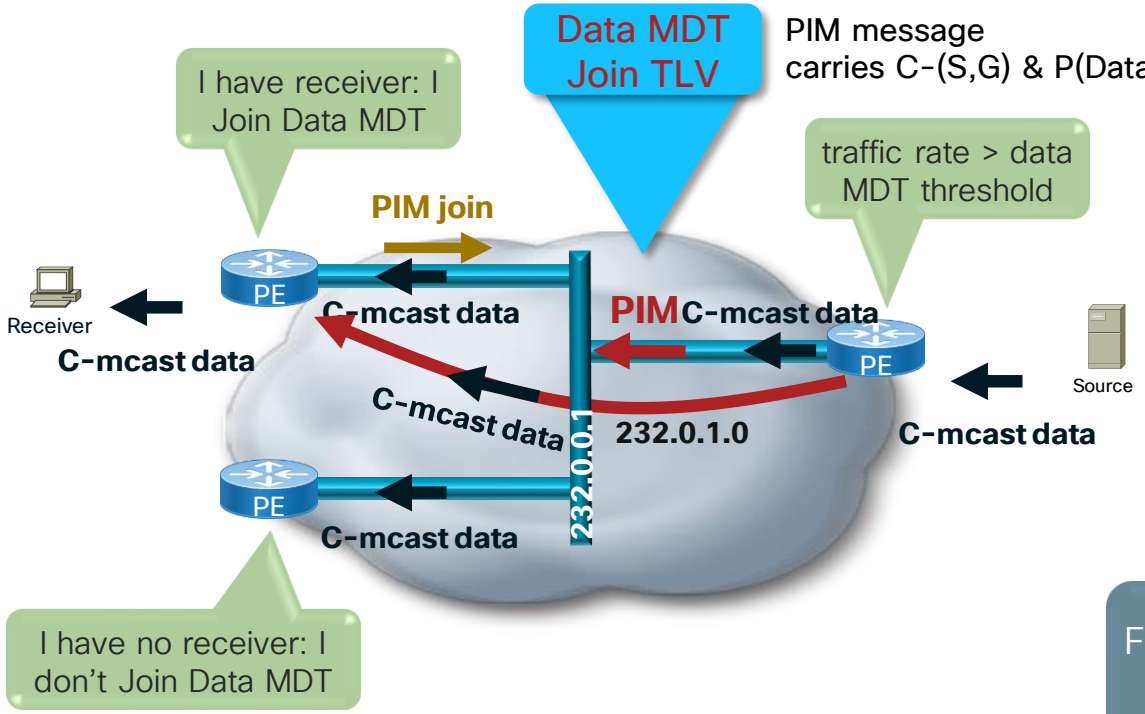



1. C-PIM neighbors across Default MDT
2. C-PIM signaling across Default MDT
3. C-multicast data

One caveat of Default MDT: drop mcast traffic on egress PE if no receiver

# Data MDT Recap

Example
Default MDT: 232.0.0.1
Data MDT: 232.0.1.0 – 232.0.1.255

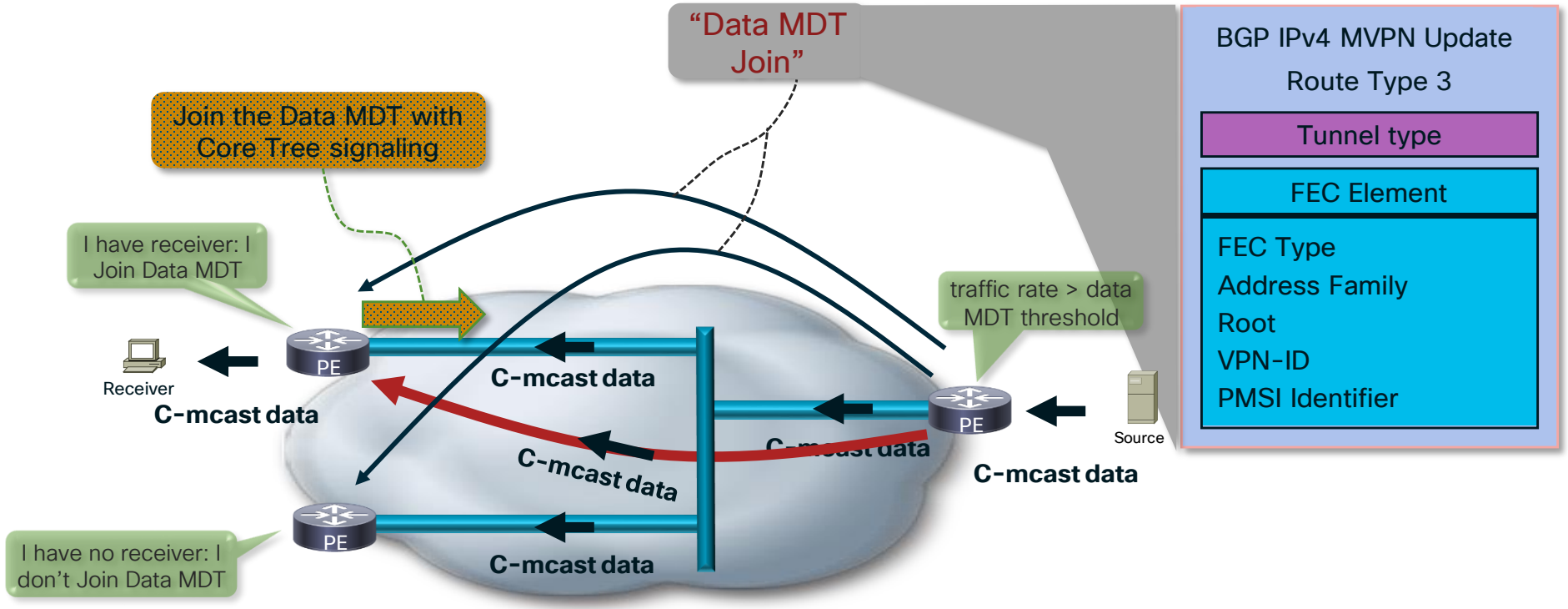



1. Signal Data MDT
2. PEs with receiver Join Data MDT
3. Ingress PE offloads traffic after 3 seconds from Default MDT onto Data MDT

For high rate C-mcast traffic: go onto Data MDT

# BGP Signaling for Data MDT

- BGP AD can also be used to discover set of PEs interested in a given C-group to enable Data MDT creation



# Why Labeled Switched Multicast (LSM)?

## Classic mVPN

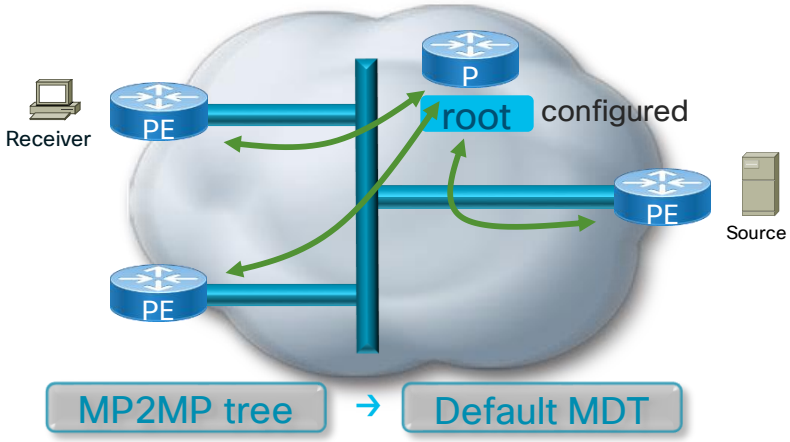
- Only GRE encapsulation
- Only PIM in core
- Default MDT / Data MDT

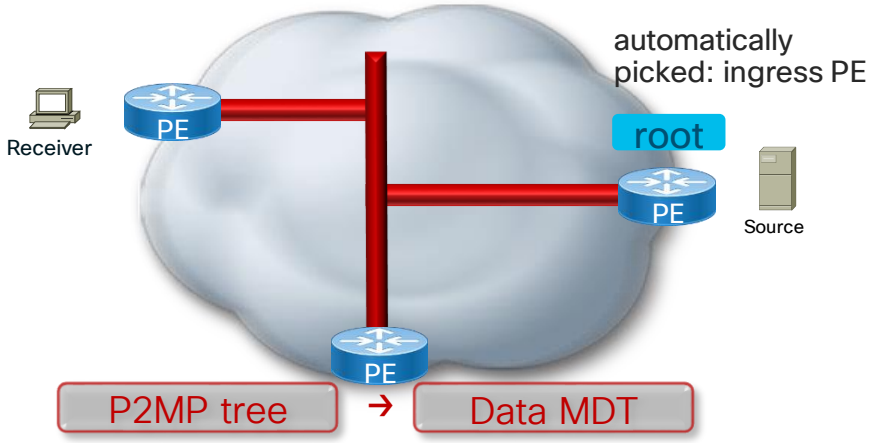
## LSM / NG mVPN

- Leverage MPLS encapsulation
  - Share control and data plane with unicast
- Leverage new core tree protocols
  - mLDP, P2MP TE
- Fast ReRoute (FRR)
- Leverage proven functionality:
  - Default & Data MDT
- More flexible designs per VPN
- Manageability: no need to track Multicast Groups per VPN/Default/Data MDT

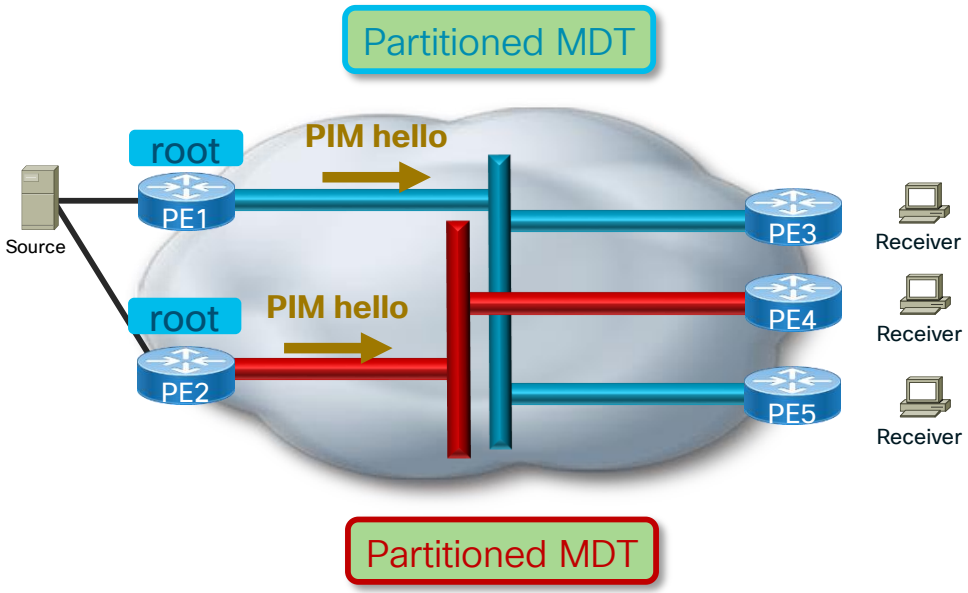
# Default MDT / Rosen - mLDP

- Re-use all of the Rosen model:
  - Default MDT
  - Data MDT
  - PIM Signaling
  - Timers
- But, use mLDP signaling and MPLS encapsulation



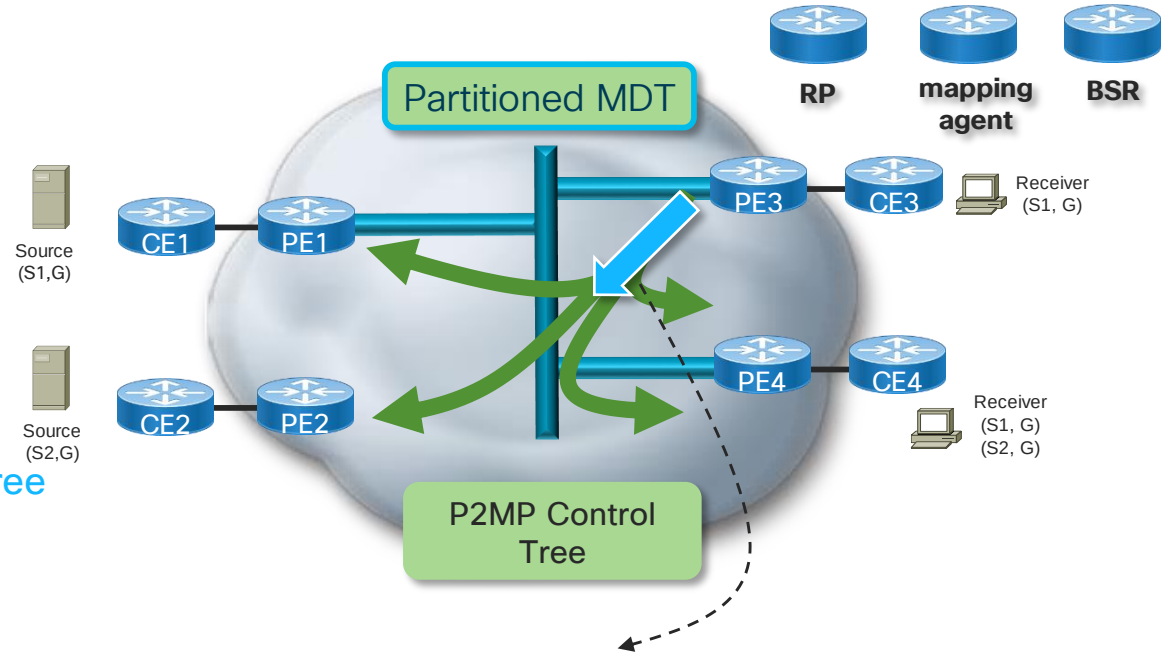
# Partitioned MDT




- Dynamic version of Default MDT model
- BGP AD is needed
- MDT built on-demand when customer traffic is present
  - Optimized for sources mostly co-located in few sites
  - Supports Anycast sources
    - It is possible to have two different trees/roots forwarding the same (S,G)
    - With Default MDT → Asserts
- Scalability: one-way PIM neighborship

# Partitioned MDT and PIM Sparse Mode (SM)

- C-PIM is SM or BiDir
- How do the PE & C-routers learn RP?

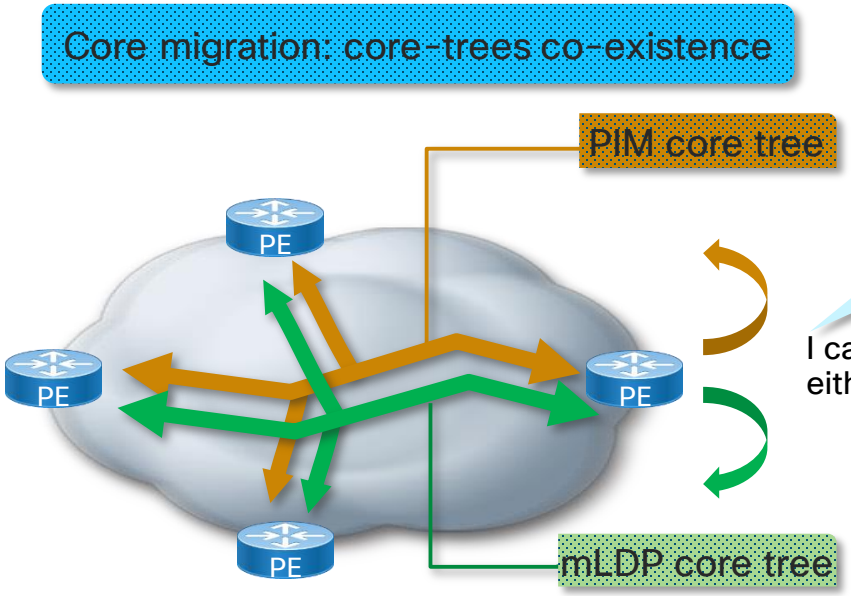



- C-PIM
- Control-tree = additional P2MP tree
- Only for RP discovery (AutoRP or BSR)
- Only for Partitioned MDT
- Automatic: when PE sees RP info
- BGP-AD

(*,224.0.0.13) for BSR announcements
(*,224.0.1.39) for AutoRP-CRP announcements (ip pim send-rp-announce)
(*,224.0.1.40) for AutoRP-MA announcements (ip pim send-rp-discovery)

# Migration Core Tree Protocol

- Classic mVPN to NG mVPN
- Likely to be PIM to mLDP, P2MP TE




On Egress PE

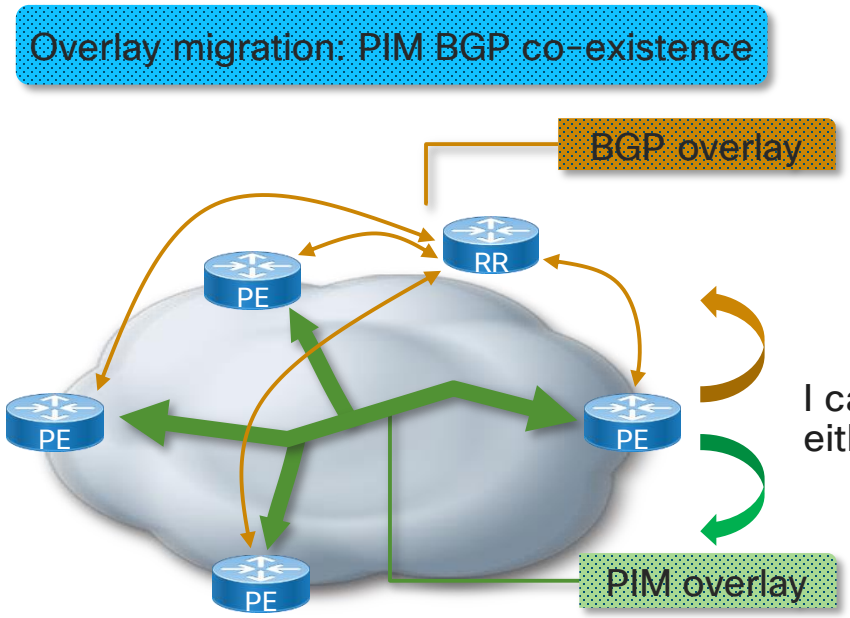
Flexible policy to do core tree selection

- Per multicast Group
- Per multicast Source
- Per Source ingress router
- ...

With RPL (Route Policy Language) in IOS-XR

I can chose either one

# Migration PIM to BGP Overlay Signaling




I can chose either one

On Egress PE

Flexible policy to do overlay selection

- Per multicast Group
- Per multicast Source
- Per Source ingress router
- …

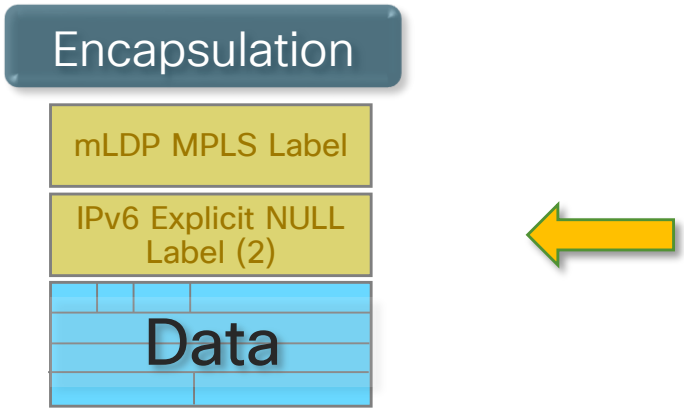
With RPL (Route Policy Language) in IOS-XR

# IPv6

- No core tree support for IPv6 (PIM, mLDP, P2MP TE, BIER)
- IPv4 core tree re-used for IPv6
  - PIM, mLDP, P2MP-TE, BIER
- Overlay signaling supports IPv6
  - PIMv6
  - BGP (IPv6 mvpn)

```
router bgp 1
…
 address-family ipv4 mvpn
  neighbor 10.100.1.4 activate
  neighbor 10.100.1.4 send-community both
!
 address-family ipv6 mvpn
  neighbor 10.100.1.4 activate
  neighbor 10.100.1.4 send-community both
```

- Note: Encapsulation of IPv6 over mLDP : explicit null label at the bottom to differentiate between IPv4 & IPv6 mcast on the same MDT



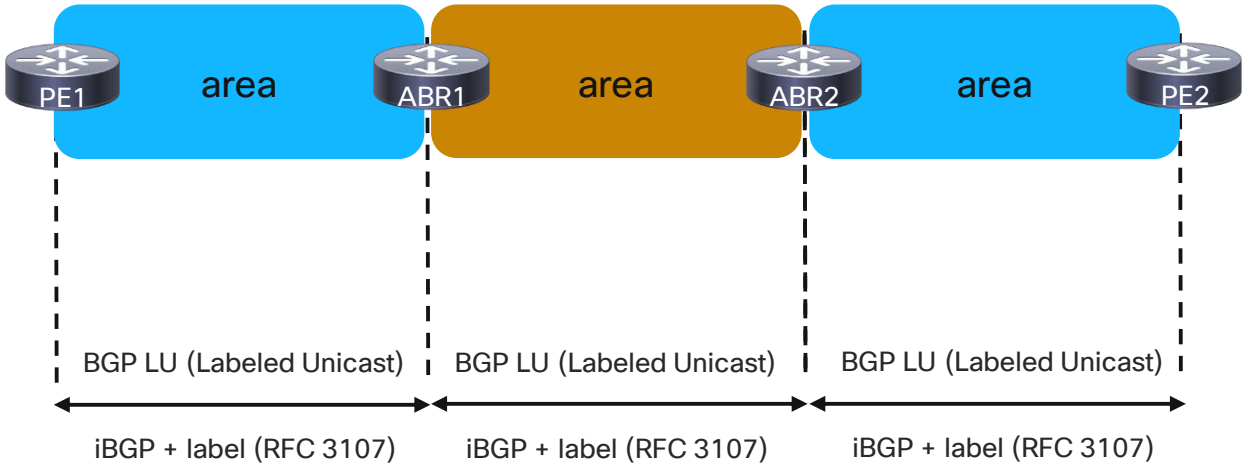
# Complex mVPN Models

- Unified/Seamless MPLS
- Segmented mVPN
- Global Table Multicast (GTM)

You make security **possible**

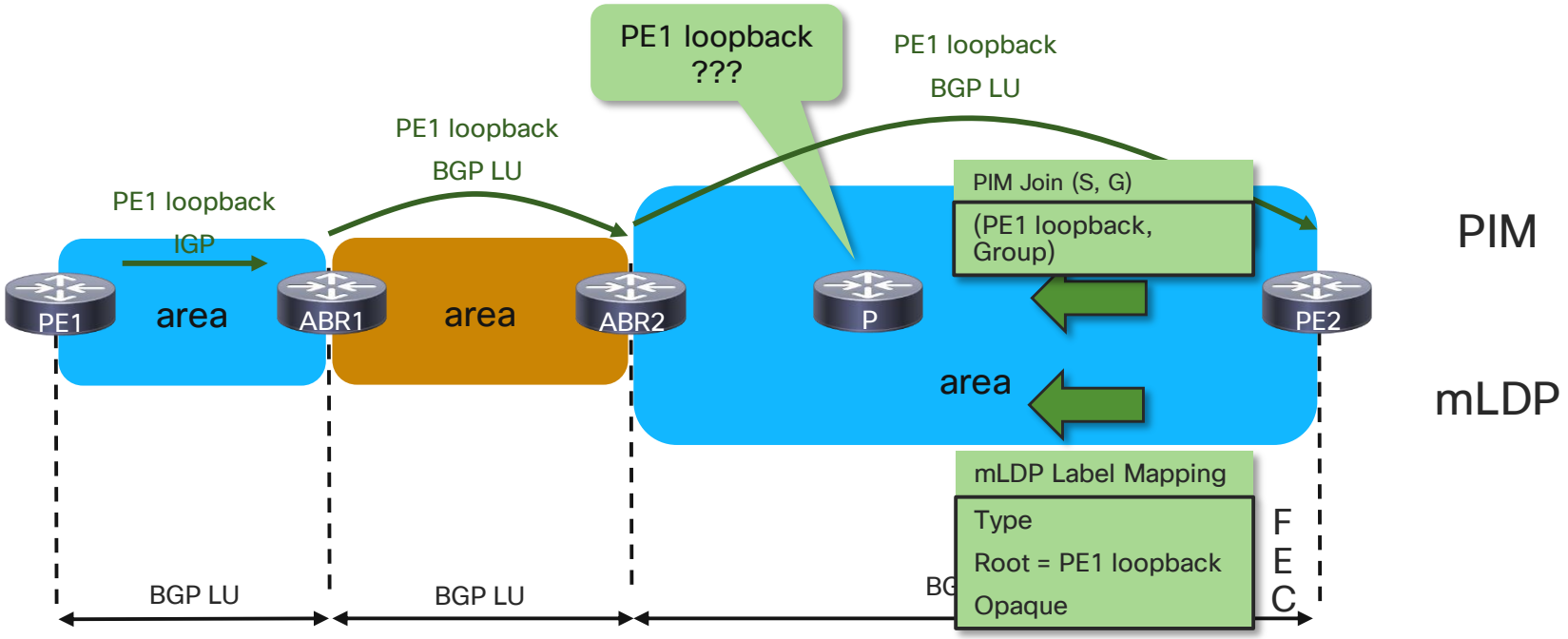
# Seamless MPLS

- Better scalability
- Filtering routes on ABR
- ABRs are BGP Route Reflectors



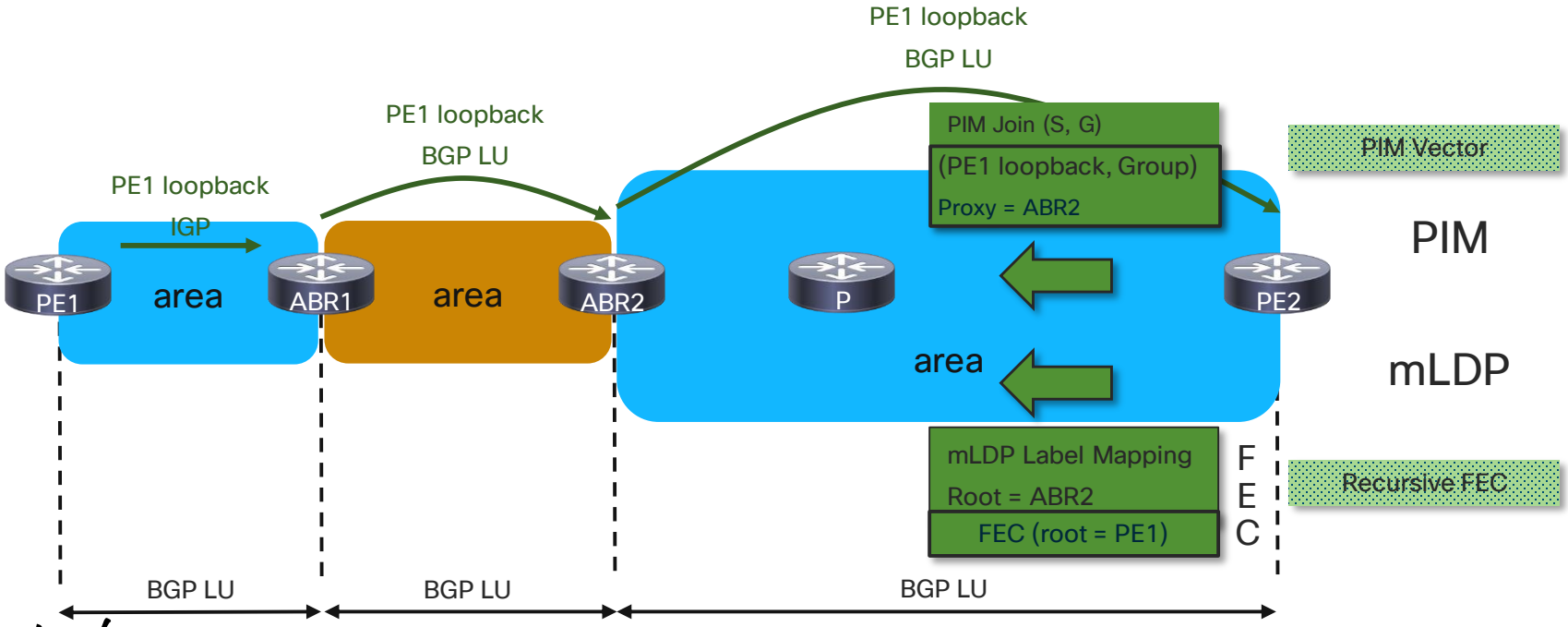
# Seamless MPLS + mVPN

- Issue: Router in one area does not know PE loopback address in other area



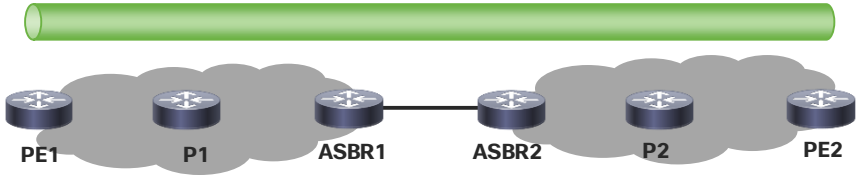
# Seamless MPLS + mVPN

- Issue: Router in one area does not know PE loopback address in other area
- Solution: PIM Vector or Recursive FEC (mLDP)

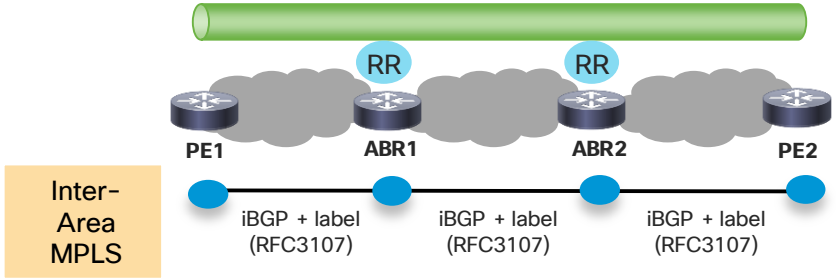


# Segmented mVPN

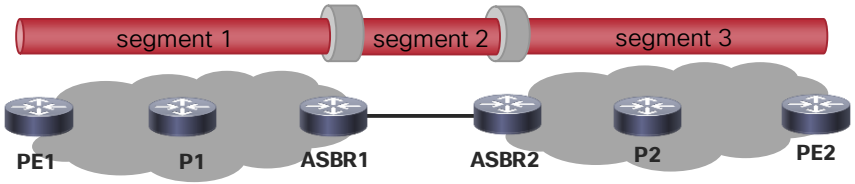
## non-segmented inter-as multicast tunnel




## non-segmented seamless multicast tunnel



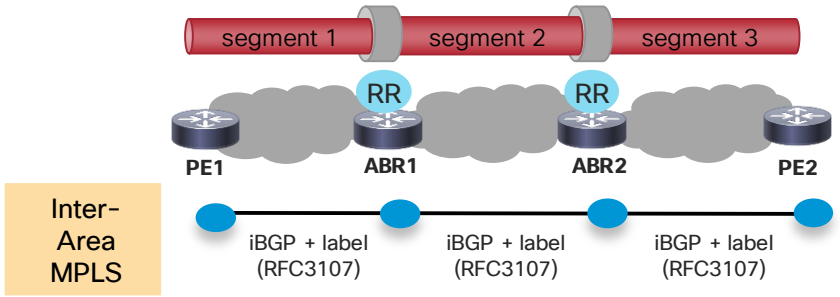

## segmented inter-as multicast tunnel




*Seamless MPLS = Unified MPLS

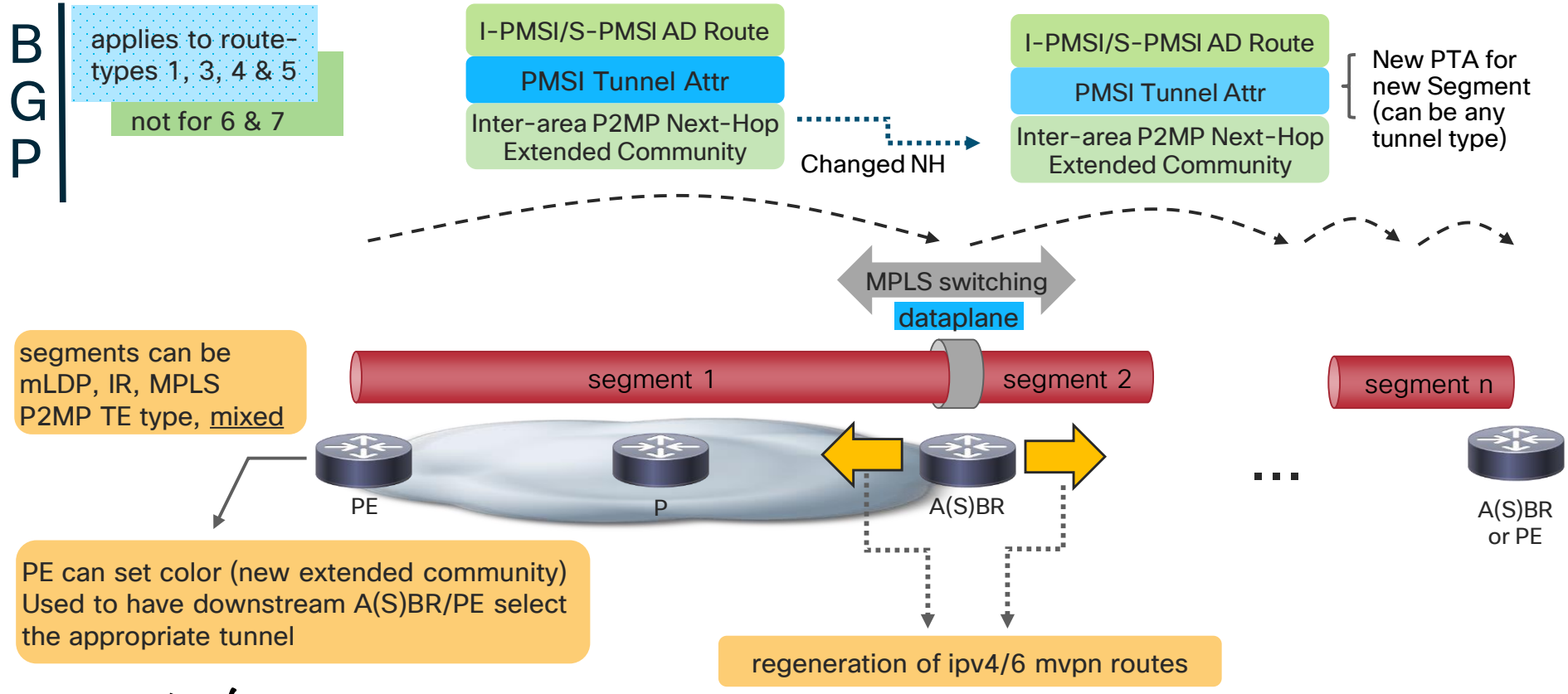
stitching

## segmented seamless multicast tunnel



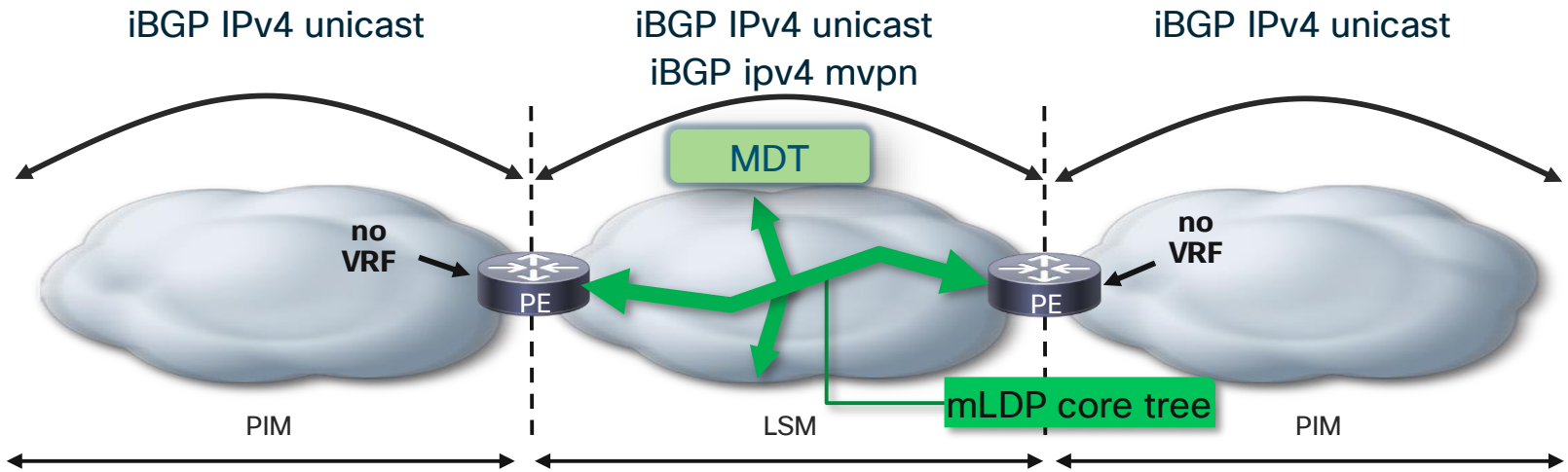

Segmented mVPN can be for Global or VPN contex

# Segmented mVPN



# Global Table Multicast (GTM)

- Have all the profiles available without VRF (so global context)
- Core tree is LSM (Labeled Switched Multicast)
- Re-use all procedures from mVPN
  - Multicast attributes (VRF Route-Import EC) will be attached to IPv4 unicast prefixes and all-zeros RD are used, to mimic vpnv4 routes
  - Unique Route-Target is used for IPv4 mvpn routes for GTM



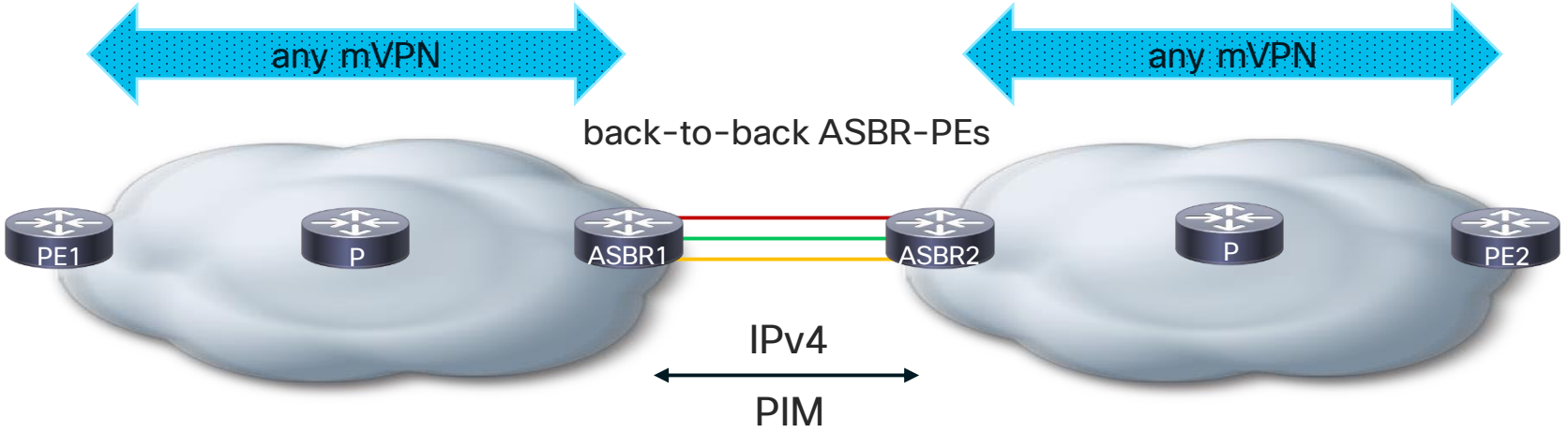
# Inter-Autonomous mVPN

You make networking **possible**

Cisco live!

# Inter-AS Option A

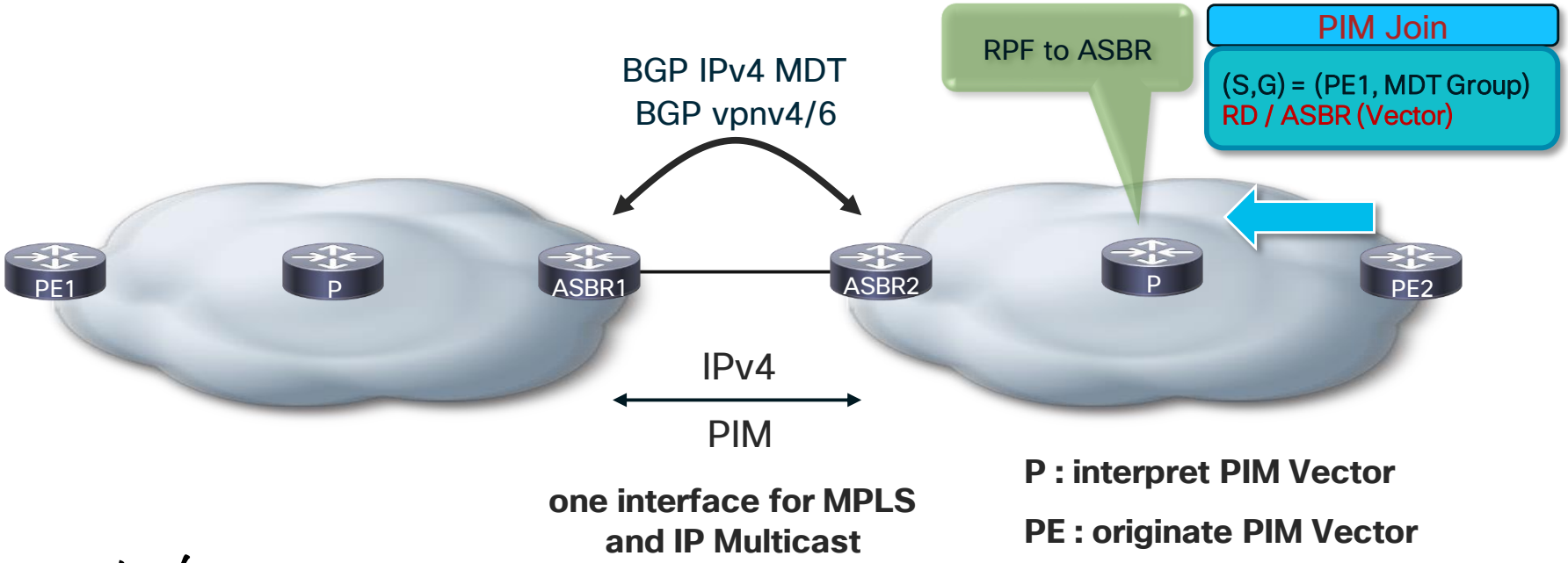
- Issue: none, each AS is completely separate
- Solution: no need for anything special
- Any AS : any core-tree protocol + any overlay protocol




**one logical interface per VRF**

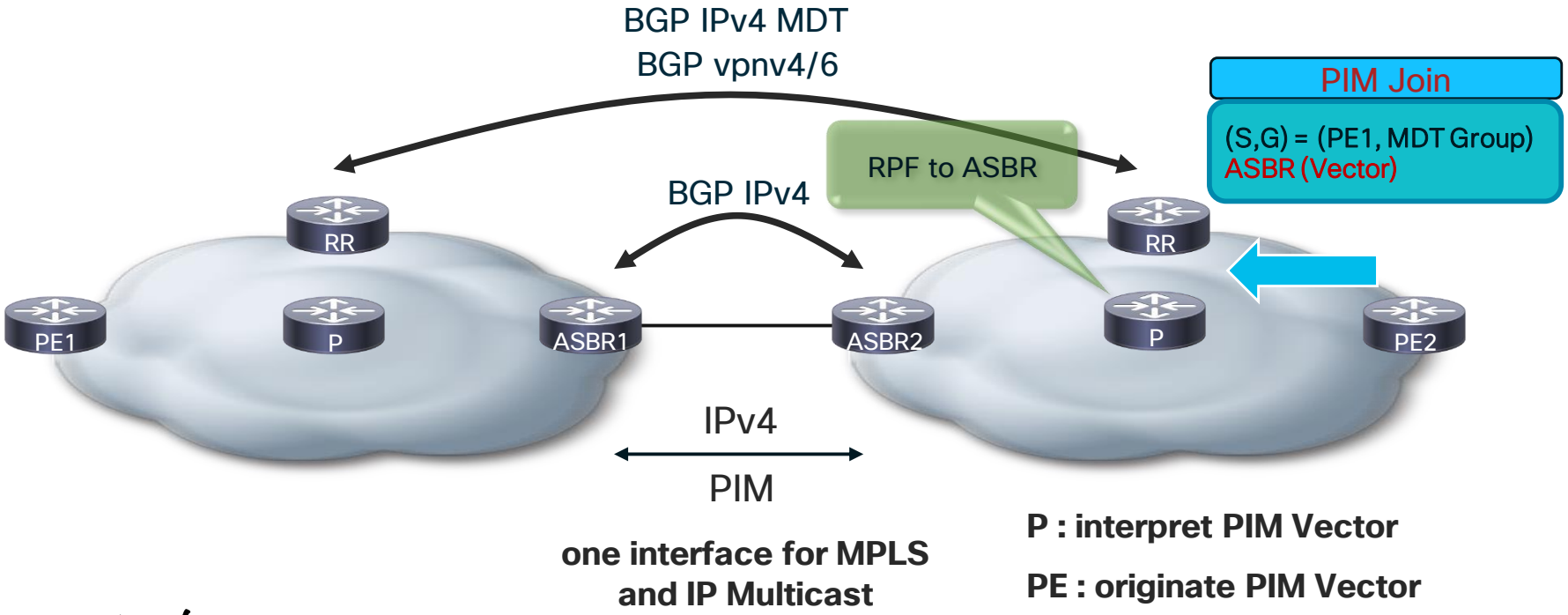
# Inter-AS Option B with PIM

- Issue: P does not know PE loopback address in other AS
- Solution: PIM Vector



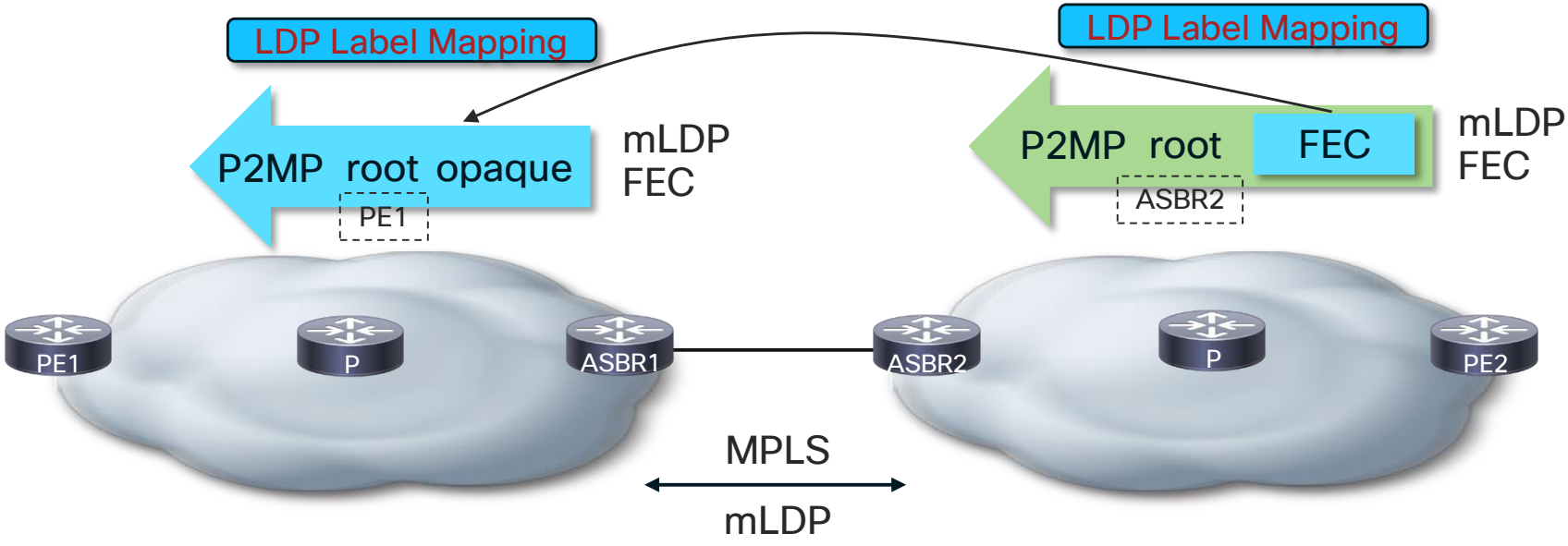
# Inter-AS Option C with PIM

- Issue: P does not know PE loopback address in other AS
- Solution: PIM Vector



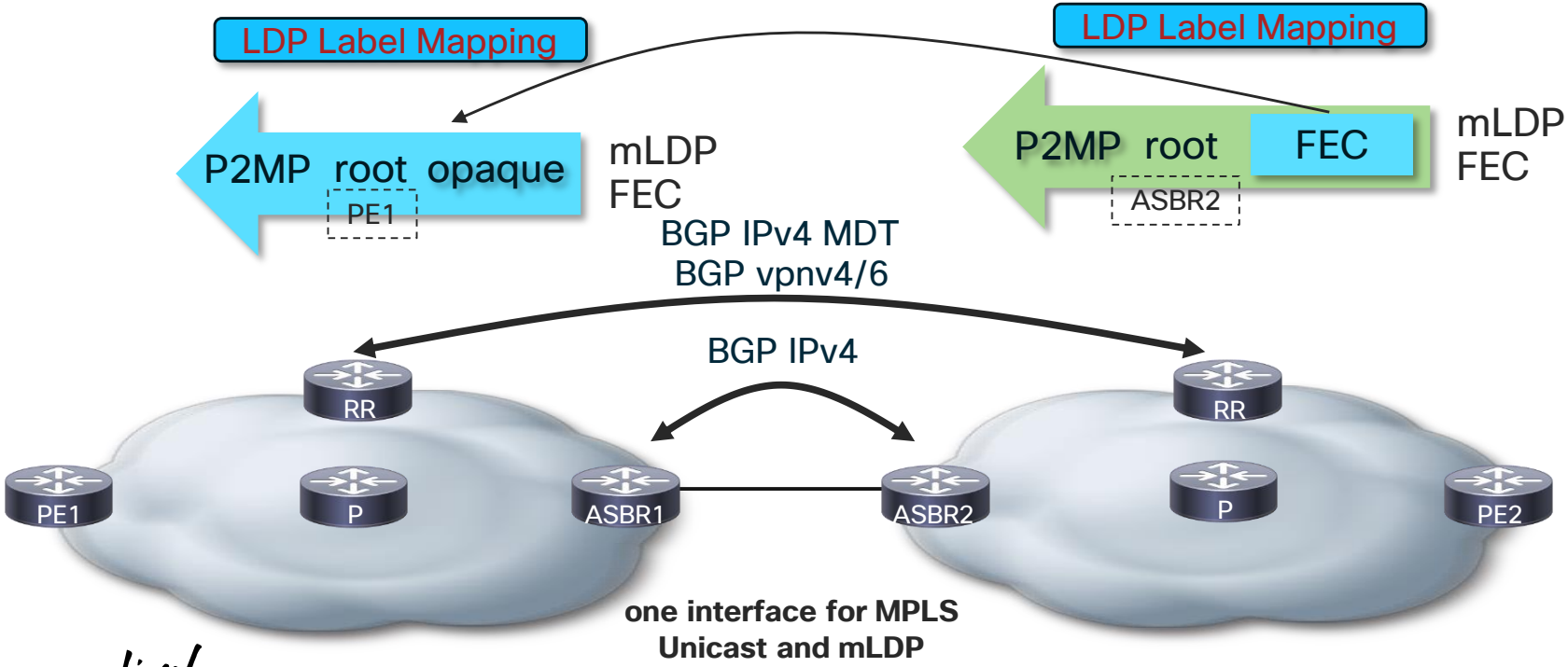
# Inter-AS Option B with mLDP

- Issue: P does not know PE loopback address in other AS
- Solution: Recursive FEC (mLDP)



# Inter-AS Option C with mLDP

- Issue: P does not know PE loopback address in other AS
- Solution: Recursive FEC (mLDP)



# Carrier's Carrier (CsC)

You make networking **possible**

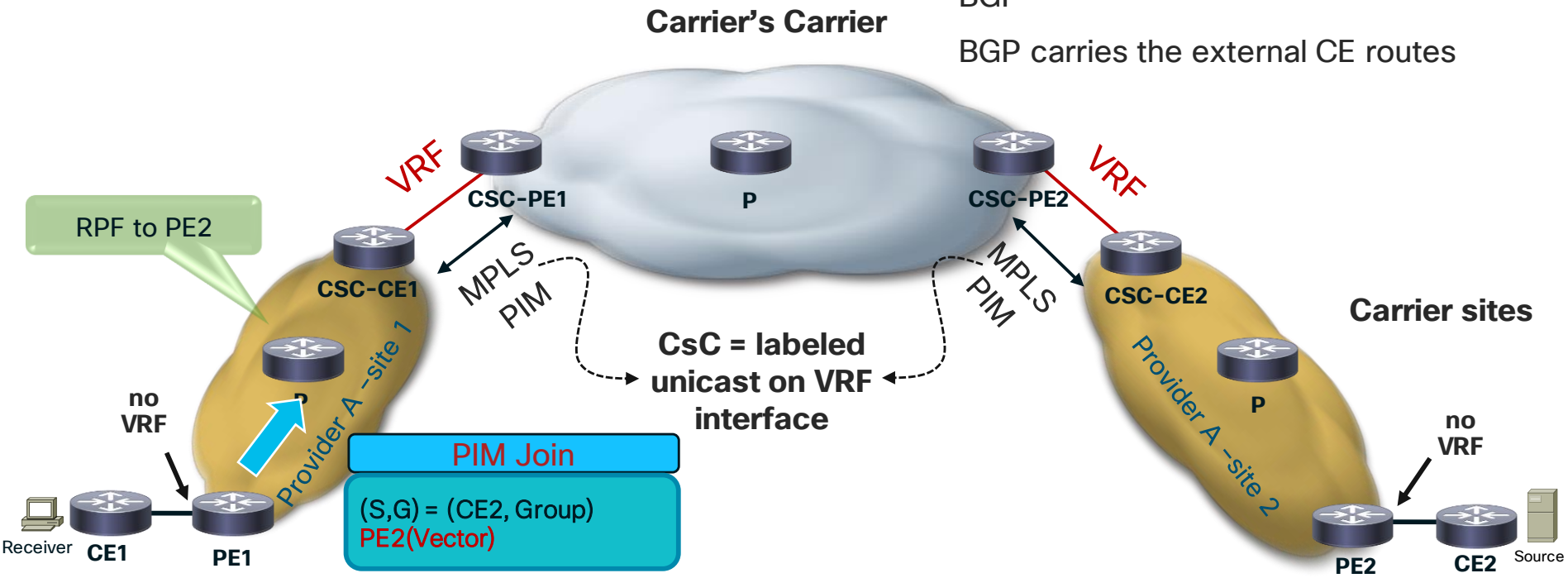
Cisco live!

# Carrier's Carrier (CsC) with PIM

PIM Vector needed (if not hierarchical CsC)

Issue : CSC-PE routers do not run BGP inside VRF and/or P routers do not run BGP

BGP carries the external CE routes

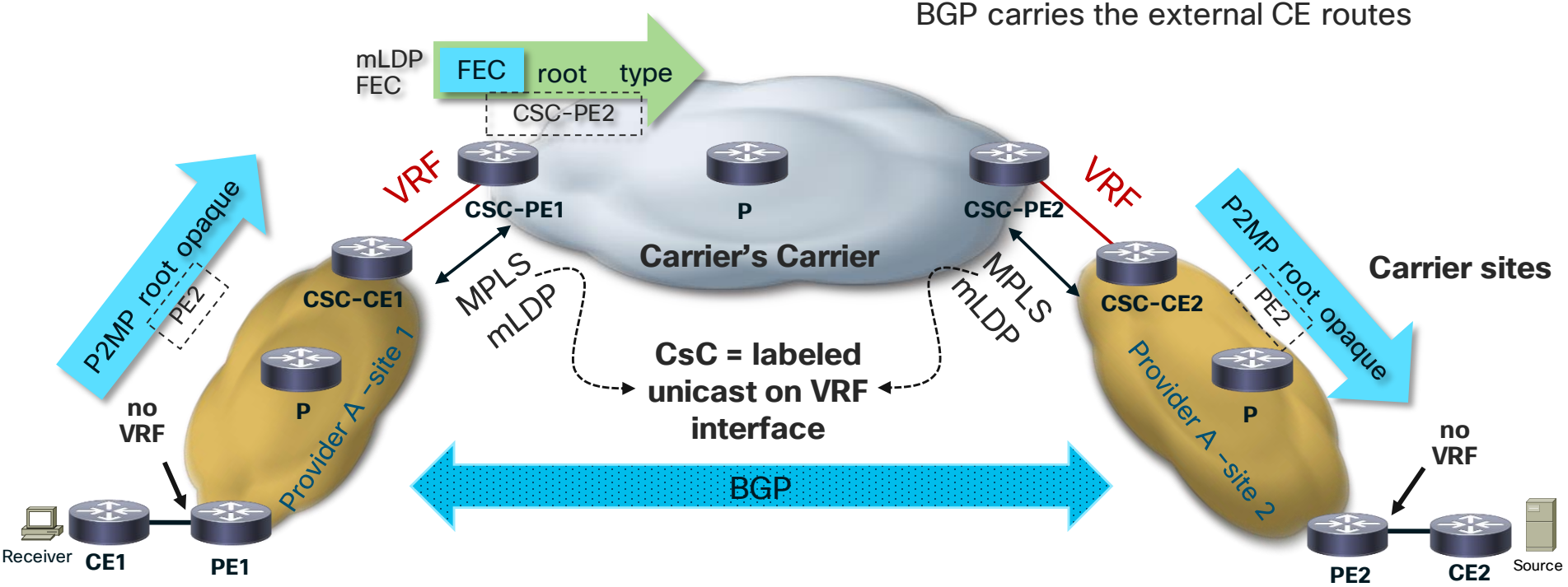


# Carrier's Carrier (CsC) with mLDP

Recursive FEC needed

Issue : CSC-PE routers do not run BGP inside VRF and/or P routers do not run BGP

BGP carries the external CE routes



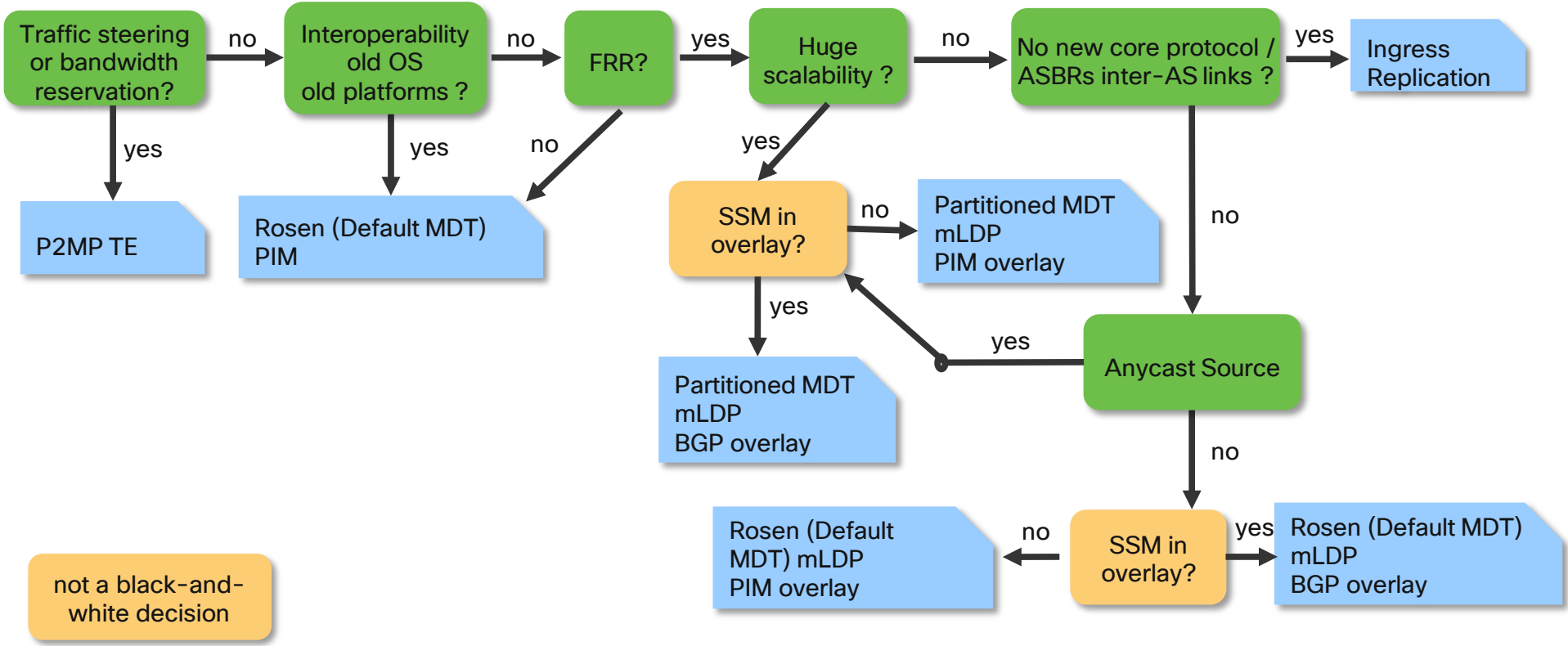
# Deployment And Examples

You make networking **possible**

Cisco live!

# Which Profile to Chose?

Nothing proprietary

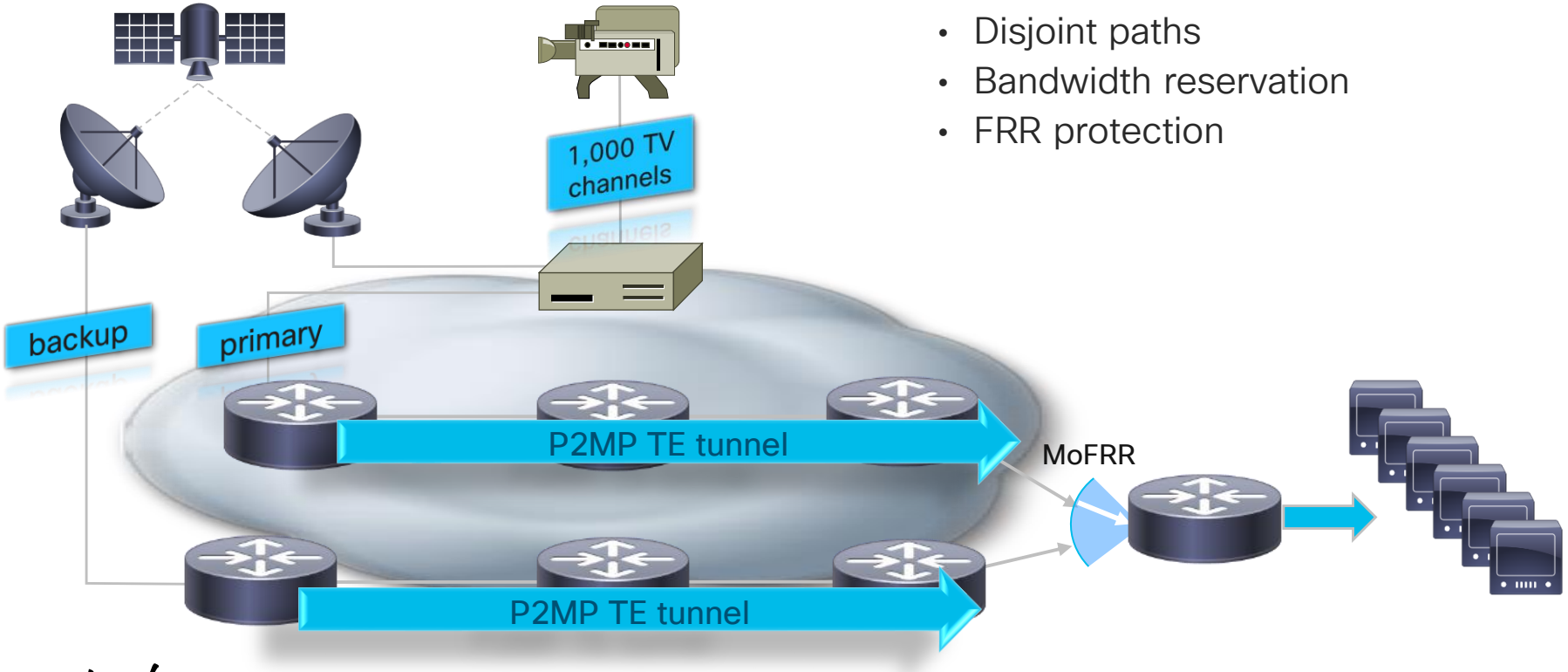


# Deployment Examples

1. TV distribution → P2MP TE
2. Railroad CCTV → mLDP + BGP
3. Migration of overlay signaling and core tree
4. Service Provider migration Rosen to mLDP, using GTM
5. Anycast Source

# Example 1: Bandwidth reservation and Protection

- TV distribution

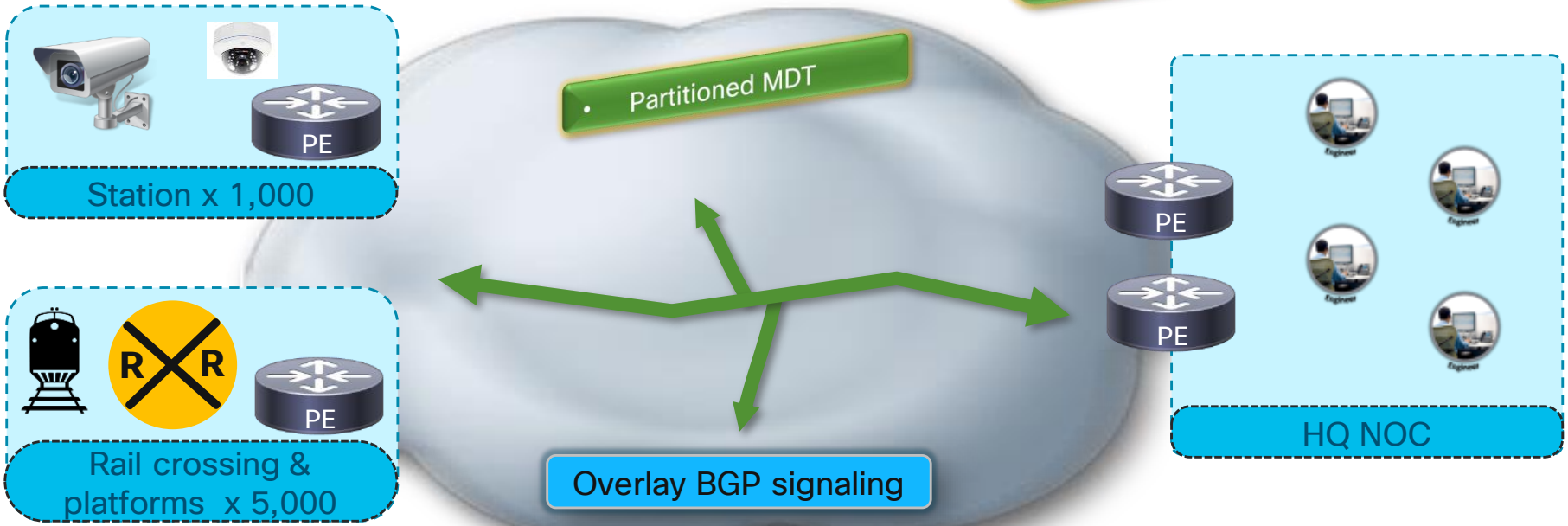



- P2MP TE
  - Disjoint paths
  - Bandwidth reservation
  - FRR protection

# Example 2: Scalability

- Railroad CCTV, 10,000 PEs

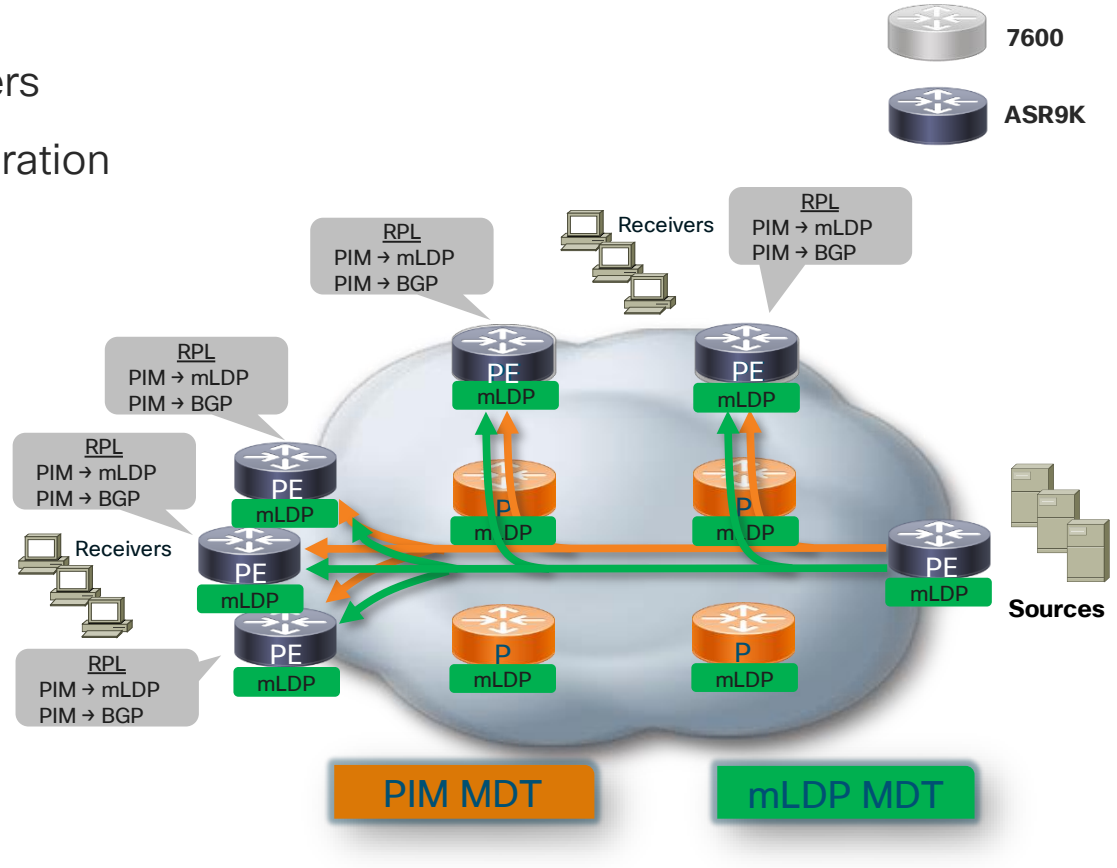
- Partitioned MDT = on-demand MDT
- Partitioned MDT = one-way PIM neighbor
- BGP overlay = very scalable signaling



# Example 3: Migration of Overlay Signaling and Core Tree

- Replacing 7600 with ASR9K routers
- PIM to mLDP and PIM to BGP migration

- PIM in core and PIM as overlay
- Enable mLDP on P and PE
- Enable MDT (mLDP) on ingress PE
- Enable MDT (mLDP) on egress PE
- Replace 7600 with ASR9K
- Pull multicast over new MDT (mLDP): RPL
- Remove MDT (PIM) on migrated egress PE
- Remove MDT (PIM) and PIM in core when migration ends

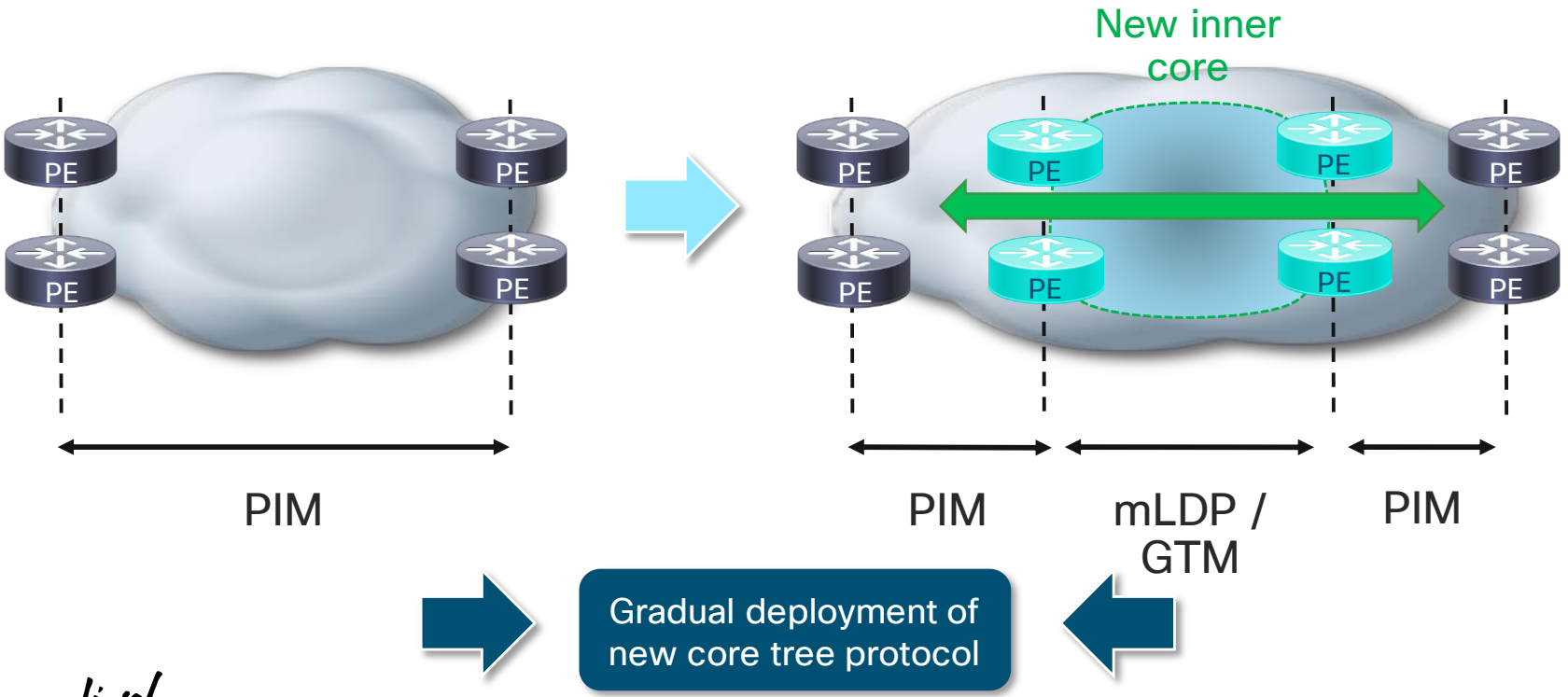


# Example 3: Migration of Overlay Signaling and Core Tree – Notes

- Ingress PE must be connected to core tree of both core protocols
  - Ingress PE sends traffic on both MDTs
- Data Join TLV (signaling Data MDT) must also be sent as PIM message
- RPL used to choose between core tree protocol and overlay protocol

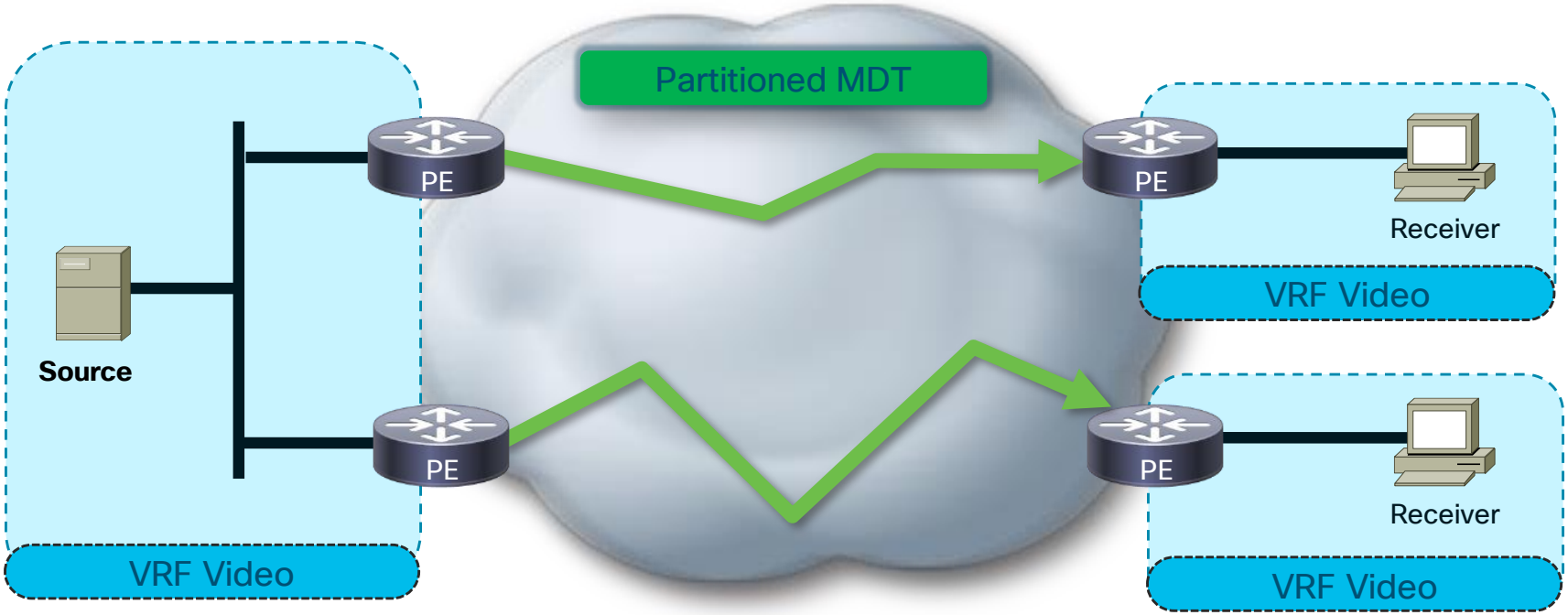
# Example 4: SP: Migration Rosen to mLDP

- For example: Using GTM to expand mLDP in old PIM core



# Example 5: Anycast Source

- Same (S, G) over two separate paths
- Default/Data MDTs would cause asserts



# Key Takeaways

- Default MDT (Rosen) is still ok
- Go LSM if you do not want/have to use PIM / IP Multicast routing in the core
  - Better manageability
  - FRR
  - More models, better fit of requirements to model
- Chose a model based on requirements
  - Scalability
  - Application
  - PIM mode
- All and everything is per VPN

# Complete your online session evaluation

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live water bottle.
- All surveys can be taken in the Cisco Live Mobile App or by logging in to the Session Catalog on ciscolive.cisco.com/us.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.cisco.com.

# Continue your education

Demos in the Cisco campus

Walk-in self-paced labs

Meet the engineer 1:1 meetings

Related sessions

# Thank you

You make **possible**