LSM: Overview and Applications

Shankar Vemulapalli

Agenda

- Need for Label Switched Multicast
- Solutions
  - mLDP
  - p2mp TE
- OAM Extensions for support of LSM
- Applications
  - PIM SSM Transit for IPv4/IPv6
  - mVPN Deployments (Default/Data)
  - Video Contribution & Distribution

Scope of the Presentation

- Presentation is going to cover the motivations for the Label Switched Multicast
- Goes over the building blocks of how both mLDP and P2MP-TE LSPs are set up
- Look at the applications of the LSM

What is LSM

- Label Switched Multicast

  MPLS Technology extensions to support multicast using Labels

  - ✓ Point-to-Multipoint LSPs
  - ✓ Multipoint-to-Multipoint LSPs

- Multicast Label Switched Paths

  Trees built using Labels

- Native Multicast Mapped onto Multicast LSPs

Drivers for LSM

- Customers want to leverage their MPLS infrastructure for transporting IP Multicast, so common data plane for unicast and multicast
- Service Providers asking for a tighter integration of Multicast with MPLS Traffic Engineering and GMPLS for their Triple Play Services
- Unify forwarding between VPN Unicast & Multicast for operational reasons in customer spaces using native MPLS/VPN.
- Simplification of Core Routers by removing PIM

Drivers for LSM

- Drivers for point-to-multipoint from Video Transport
  - ✓Contribution
    - Point-to-Multipoint Video feeds, e.g, sports events to multiple broadcasters
    - Desire to have Video quality probes at each network hop for service assurance and monitoring
    - Source Feed to Production Houses
  - ✓ Secondary Distribution
    - Implicitly required for IPTV BW Efficiency
    - Video Content to the end users

Work at the Standards

- Standardization work is happening at IETF for both LDP and RSVP signaling protocols to carry labels for the multicast along with extensions for OAM

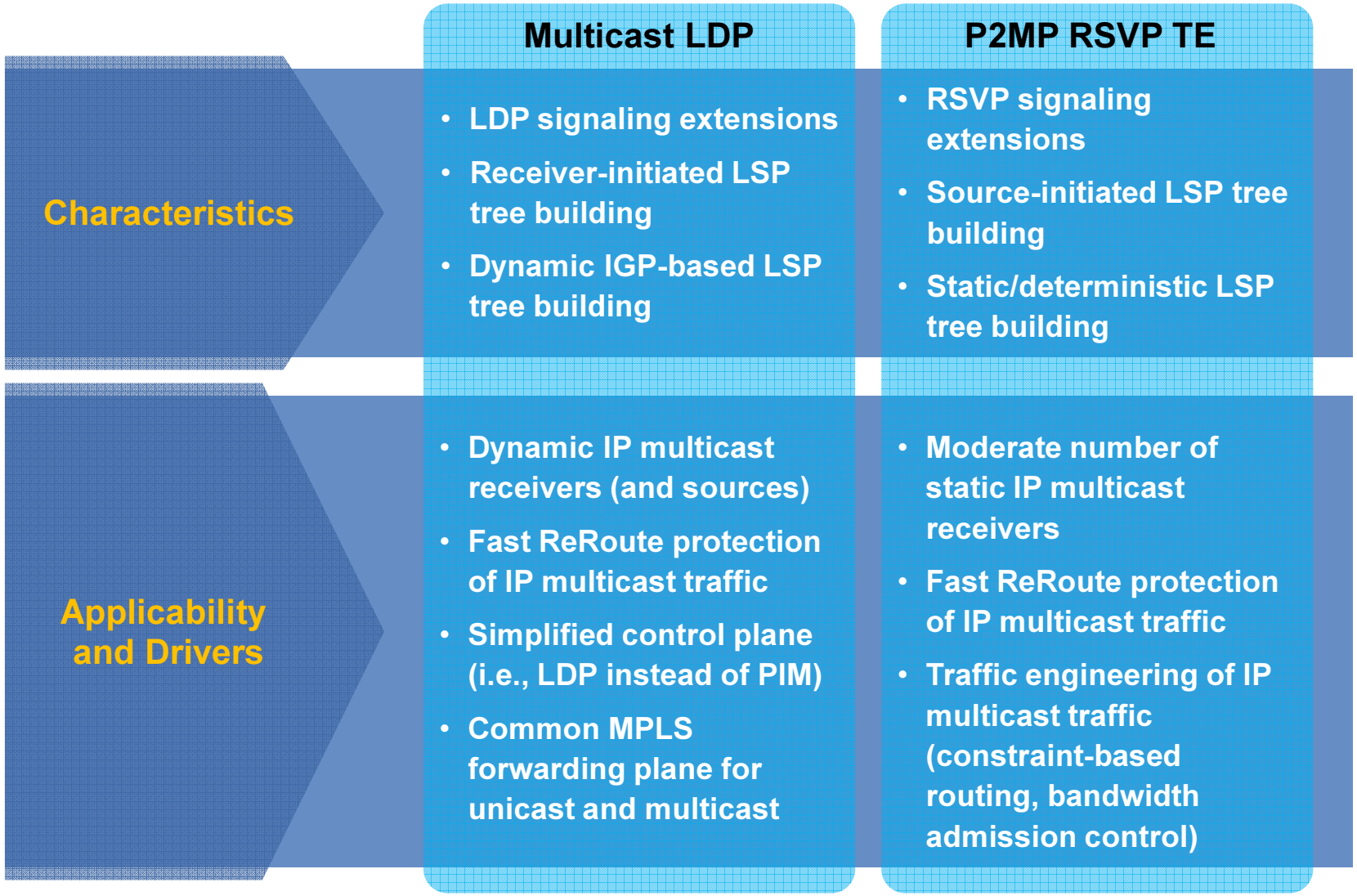
mLDP	
LDP Extensions to P2MP& MP2MP LSPs	draft-ietf-mpls-ldp-p2mp
LDP Capabilities	RFC 5561
In-Band Signaling	draft-ietf-mpls-mldp-in-band-signaling

P2MP TE	
Signaling Req. for P2MP-TE LSPs	RFC 4461
Extensions to RSVP	RFC 4875

LSM OAM	
P2MP LSP Extensions for LSP-Ping	draft-ietf-mpls-p2mp-lsp-ping
Proxy LSP Ping	draft-ietf-mpls-remote-lsp-ping
Connectivity Verification for Multicast LSPs	draft-ietf-mpls-mcast-cv

VPLS	
LSM Support for VPLS	RFC 5501

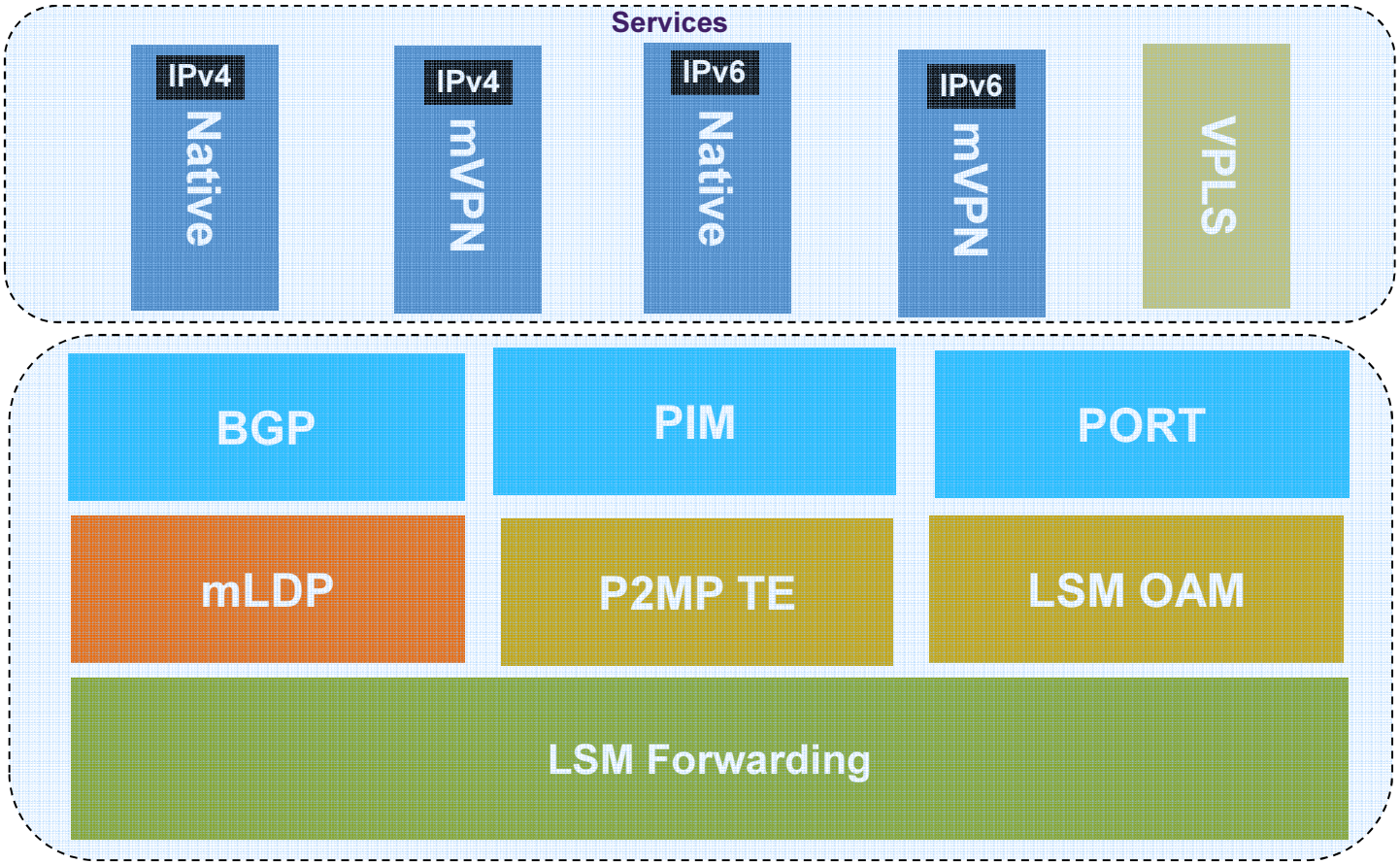
LSM Signaling Options



LSM Architecture

- LSM architecture supports a range of services or “clients”

  Over mLDP and P2MP TE control planes



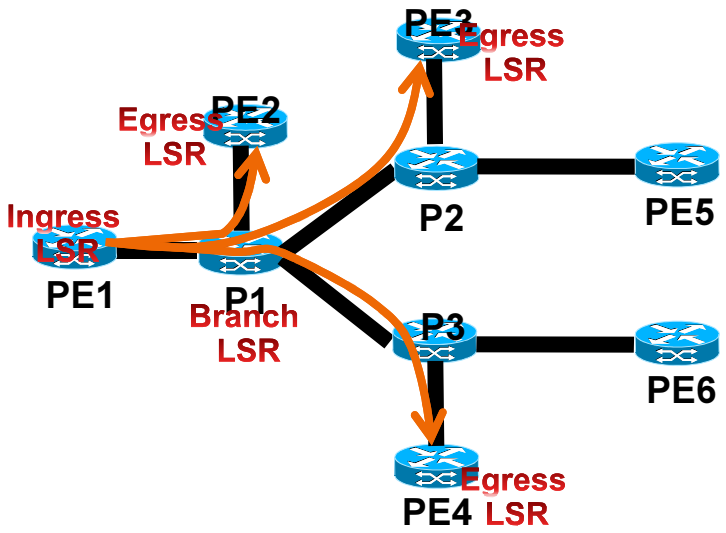
Terminology

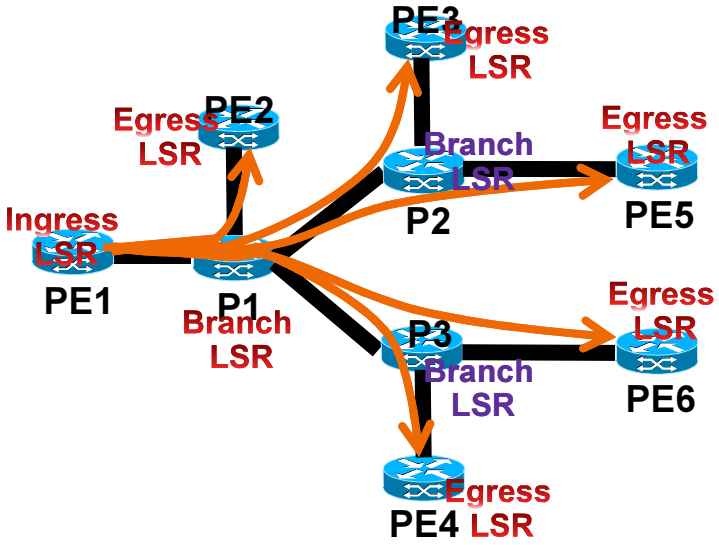
Terminology	Description
LSR	Label Switch Router
Ingress LSR	Router acting as a Sender of an LSP & is closest to multicast source (Root Node)
Egress LSR	Router acting as a Receiver of an LSP & is closest to the multicast receiver (Leaf Node)
P2P LSP	LSP with one Unique Ingress LSR & one Unique Egress LSR
P2MP LSP	LSP with one Unique Ingress LSR & one or more Egress LSRs
MP2MP LSP	LSP that has one or more Leaf LSRs acting as Ingress or Egress
MP LSP	Any type of Multipoint LSP

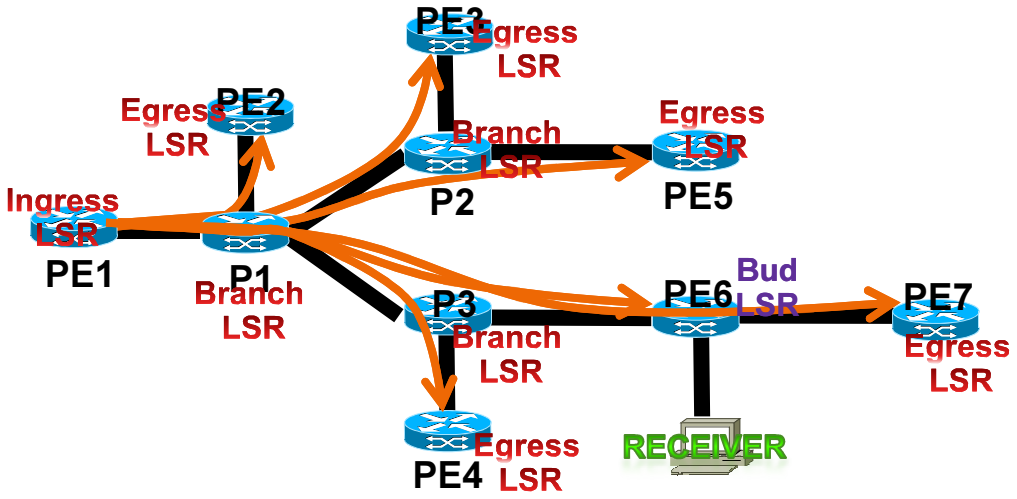
Terminology

Terminology	Description
P2MP Tree	The ordered set of LSRs & links that comprise the path of a P2MP LSP from its Ingress LSR to all of its Egress LSRs
Upstream	Direction of the Multicast packet received from (from Egress towards Ingress)
Downstream	Direction of the Multicast packet sent to (from Ingress towards Egress)
Branch LSR	LSR of a P2MP or MP2MP LSP that has more than ONE downstream LSR
Bud LSR	LSR of P2MP or MP2MP LSP that is an Egress but also has one or more directly connected downstream LSR(s)
Leaf LSR	Egress LSR of a P2MP or Ingress/Egress LSR of a MP2MP LSP

Terminology Mapping

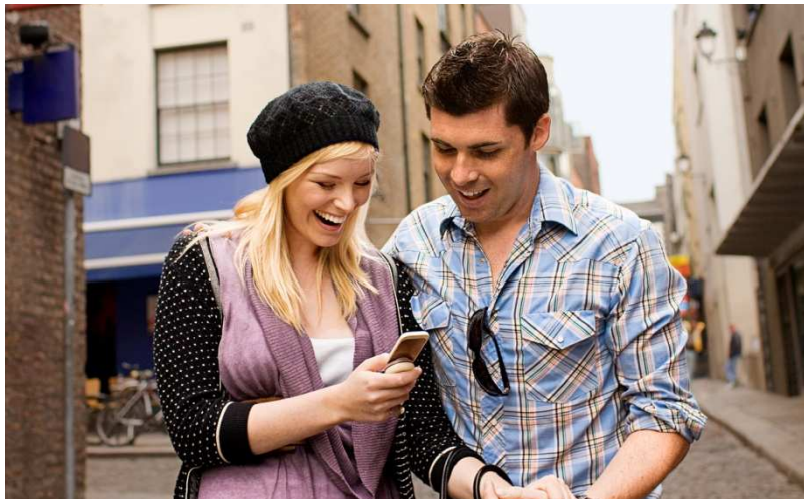






P2MP LSP

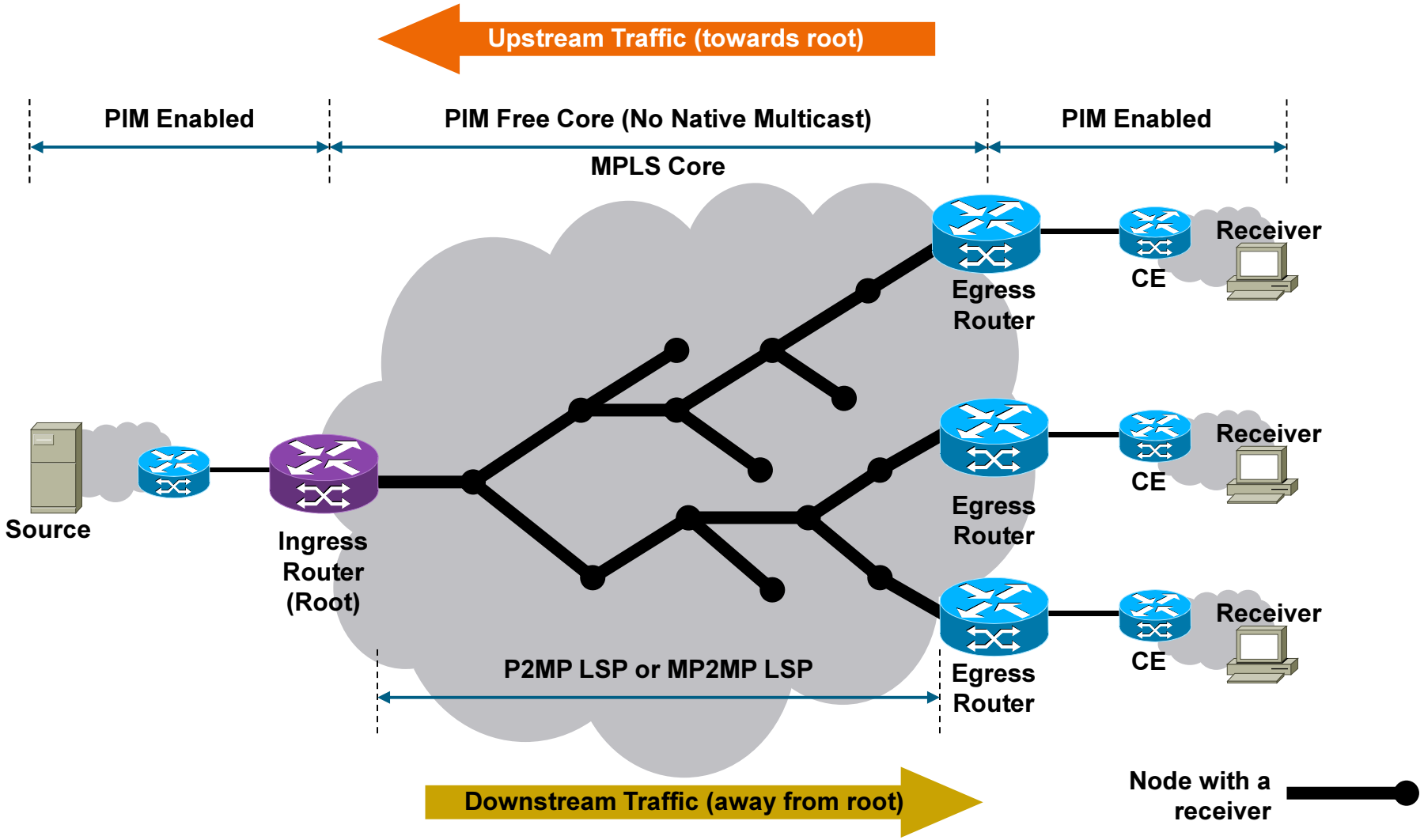
mLDP



mLDP

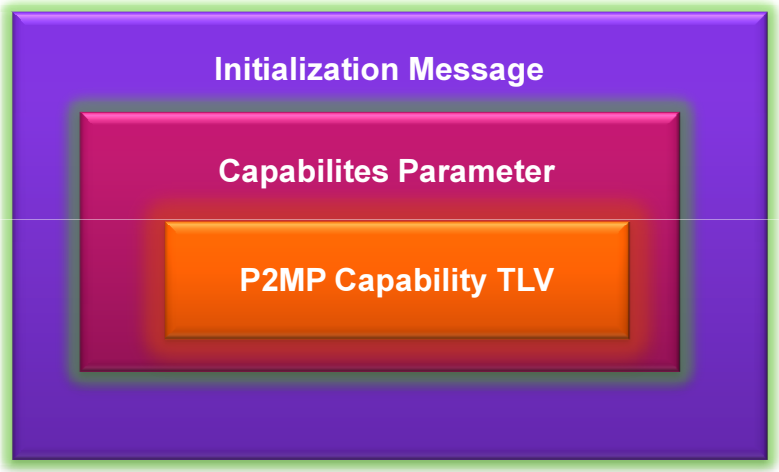
- Receiver driven (Egress LSR) MP LSP Setup
  - Labels are distributed from the Leaves towards the Root
- MP LSP Path Selection is based on Root Address
  - Derived from BGP NH of Source or Statically Configured
  - Supports P2MP & MP2MP Tree Construction
- Downstream on demand label allocation
  - Labels are not allocated unless there is a receiver interested
- Architecture supports In-Band & Overlay signaling
- No PHP – The top label is used to identify tree

mLDP Topology



mLDP – LDP Extensions

- A P2MP Capability TLV is defined which will be carried in the Capabilities Parameter as part of the INITIALIZATION Message

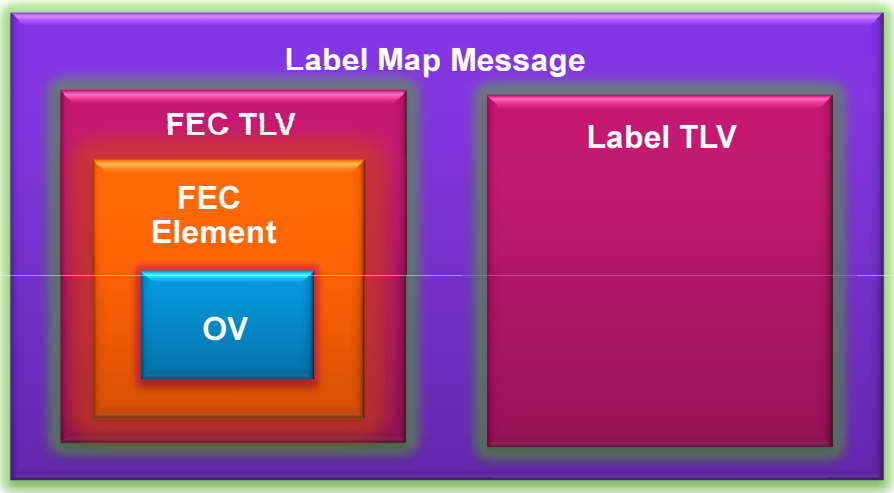


New mLDP Capabilities

New Capabilities	Value
P2MP Capability	0x0508
MP2MP Capability	0x0509
MBB Capability	0x050A

mLDP – LDP Extensions

- A new P2MP FEC Element will be advertised as part of the FEC TLV in the Label Mapping message



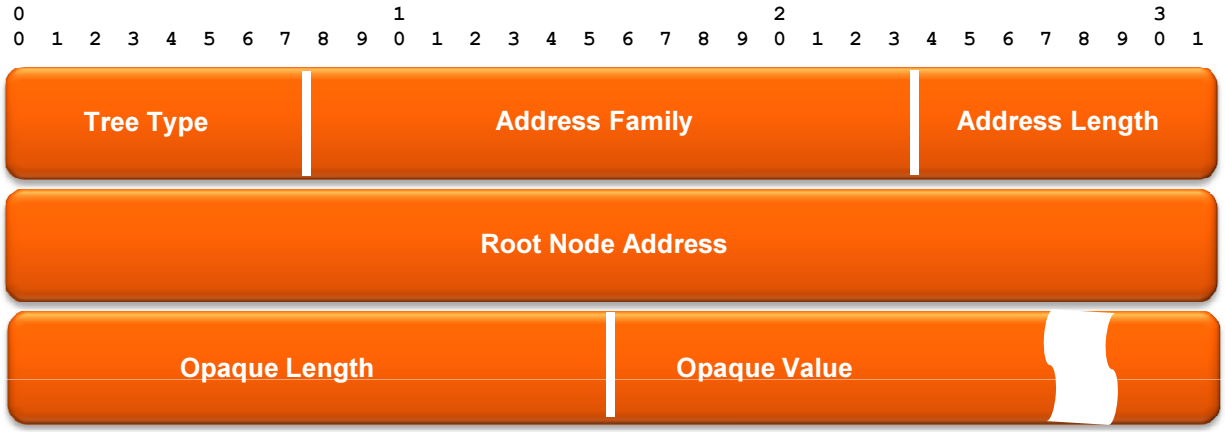
New LDP FEC Element Types

FEC Element Types	Value
P2MP FEC Type	0x06
MP2MP-UP FEC Type	0x07
MP2MP-Down FEC Type	0x08

mLDP – FEC Element

- **P2MP FEC Element**
- Consists of the Address of the Root of the P2MP LSP and Opaque Value
- Opaque Value consists of one or more LDP MP Opaque value Elements
- The Opaque Value is unique within the context of the Root Node.
- The combination uniquely identifies a P2MP LSP with in the MPLS Network
  - Root Node Address
  - Opaque Value

mLDP – FEC Element



Parameters	Description
Tree Type	P2MP, MP2MP Up, MP2MP Down
Address Family	Root node address format (IPv4 =1 or IPv6 = 2)
Address Length	Number of octets in Root Address (IPv4 = 4, IPv6 = 16)
Root Node Address	Host address of MP LSP Root (within MPLS core)
Opaque Value	One or more TLVs uniquely identifying MP LSP within the in context of the root

mLDP – Root Node Address

- Root Address is selected by the Egress Router
  - Automatically derived from BGP next-hop or statically configured
- Root address is used to build the MP LSP
- Each router in the path does a routing table lookup on the root to discover the next-hop.
  - Label mapping message then sent to that next-hop
- Resulting in a dynamically created MP LSP
  - No pre-computed, traffic engineered path

mLDP – Opaque Value Element

- **Opaque Value Element**
- Each MP LSP is identified by unique opaque value which is used to uniquely identify the MP LSP
- Carries information that is meaningful to Ingress LSRs and Leaf LSRs but need not be interpreted by Transit LSRs
- It can represent the (S, G) stream (PIM-SSM Transit) or can be an LSP Identifier to define the Default/Data MDTs in an mVPN application

mLDP – Opaque Value

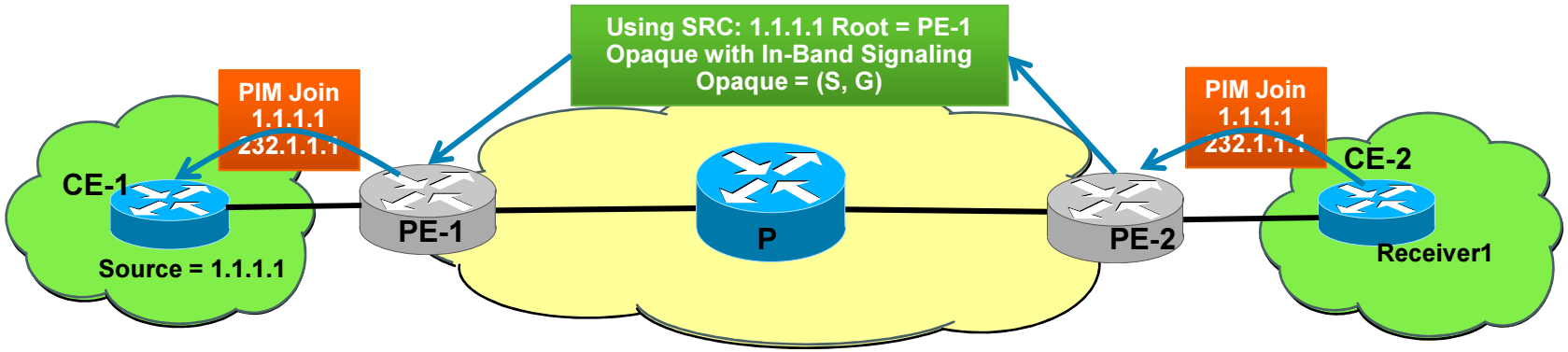
- 4 Multicast Applications are supported with each with its own Opaque Value

Applications	Description
IPv4 PIM-SSM Transit	Allows Global PIM-SSM Streams to be transported across the MPLS-Core. The Opaque Value contains the actual (S,G) which resides in the Global (mroute) table of the Ingress & Egress PE Routers
IPv6 PIM-SSM Transit	Same as Above but for IPv6
Multicast VPN	VPNv4 Traffic to be transported across Default-MDT (MI-PMSI) or Data-MDT (S-PMSI)
Direct-MDT or VPNv4 Transit	Allows VPNv4 streams to be directly built without the need for the Default-MDT to exist

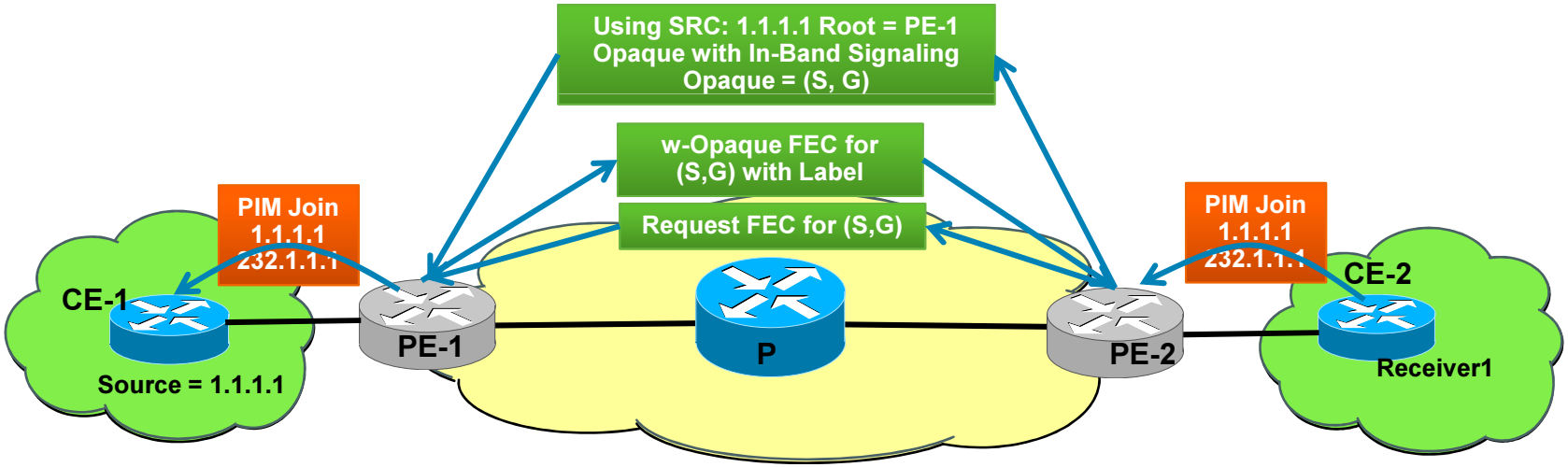
mLDP - Signaling

- mLDP Signaling provides TWO Functions:
  - To Discover the FEC & its associated Opaque Value for a MP LSP
  - To assign a multicast flow to a MP LSP
- mLDP uses two signalling methods:
- In-Band Signaling
  - All egress routers use the same algorithm to construct the opaque value based on the multicast stream they want to join.
  - That may include, Source, Group, RD, next-hop…
  - Egress routers interested in the same multicast stream will create the same FEC.
  - Ingress PE multicast component parses the FEC and knows what multicast stream to forward.
- Out-of-Band Signaling
  - Opaque value is assigned by the ingress Root PE.
  - Egress PE's use an out-of-band signaling protocol to request the opaque value that belongs to a multicast stream.
  - Egress routers use the opaque value to construct the FEC and build the tree.
  - Allows for aggregating multicast streams on a single MP-T.

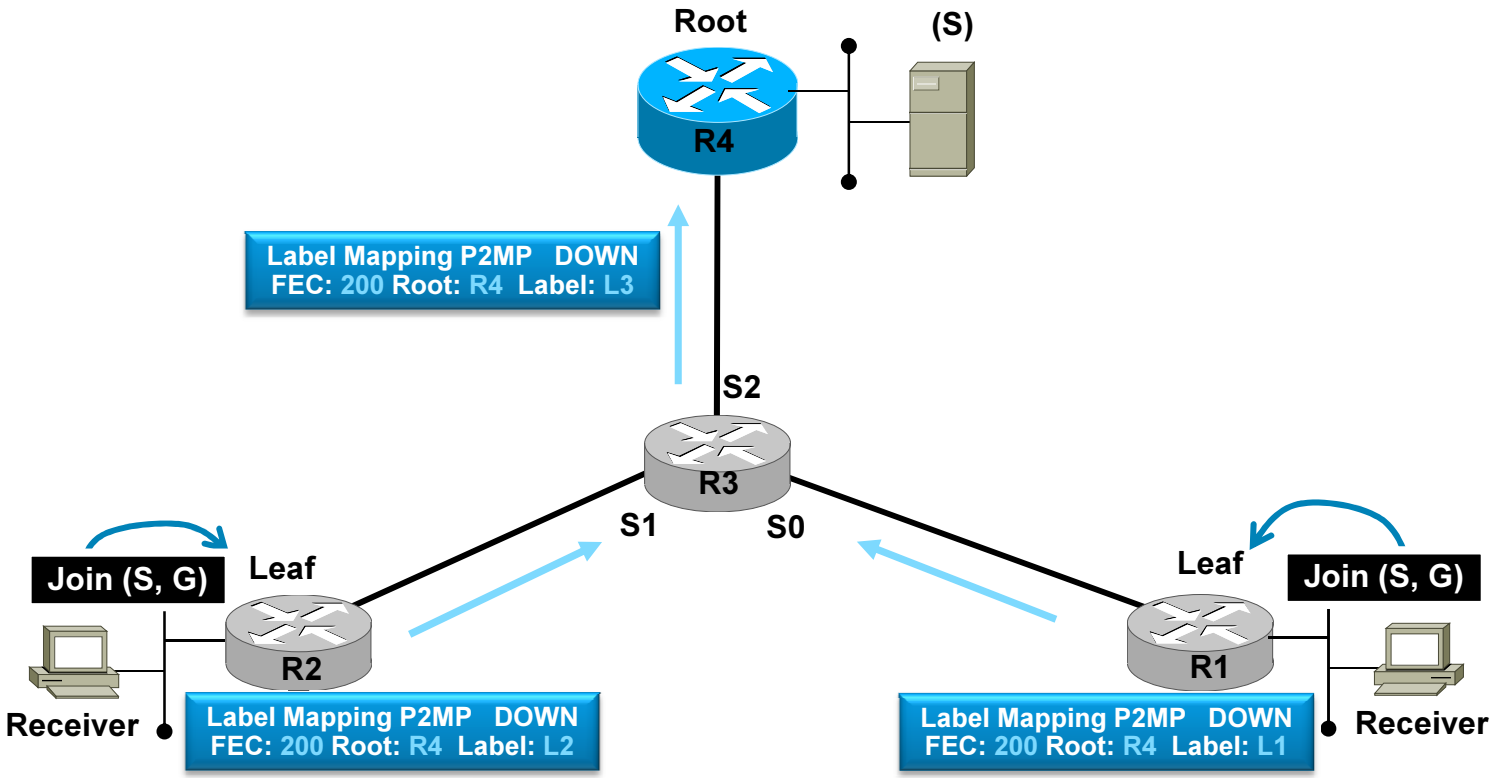
mLDP – In-Band Signaling Operation



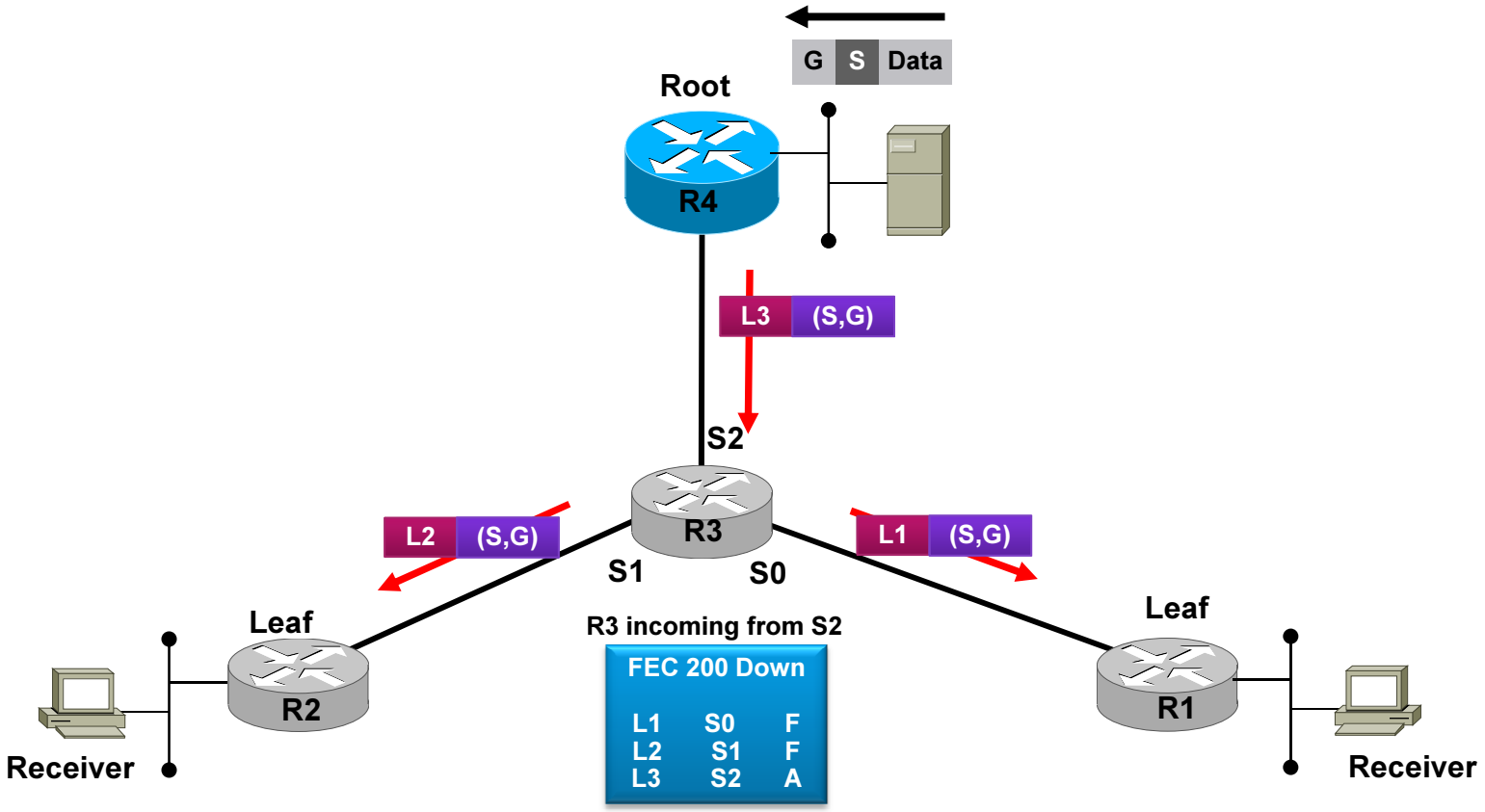
mLDP – Out-of-Band Signaling



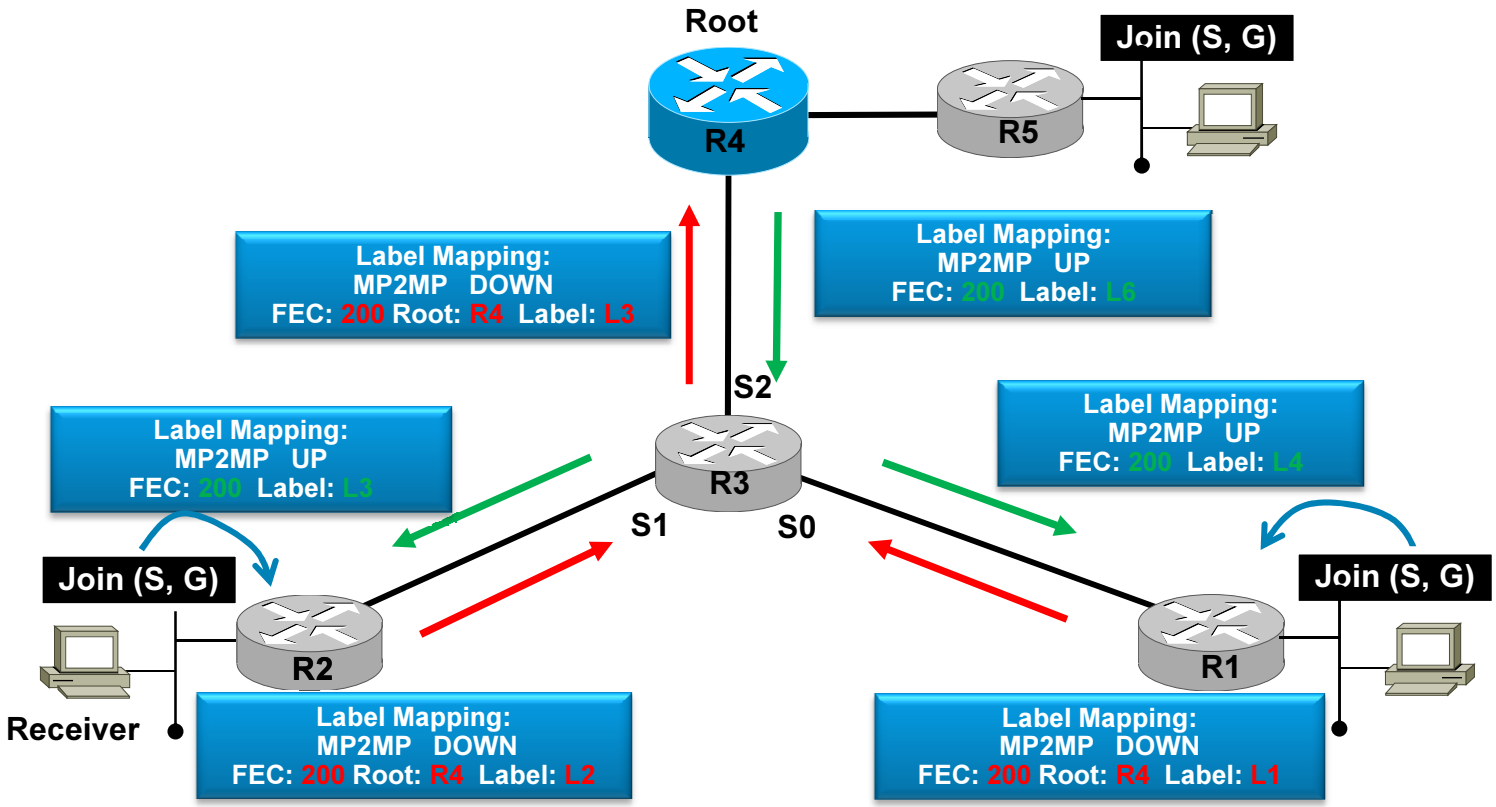
P2MP Basic Operation



P2MP Basic Operation

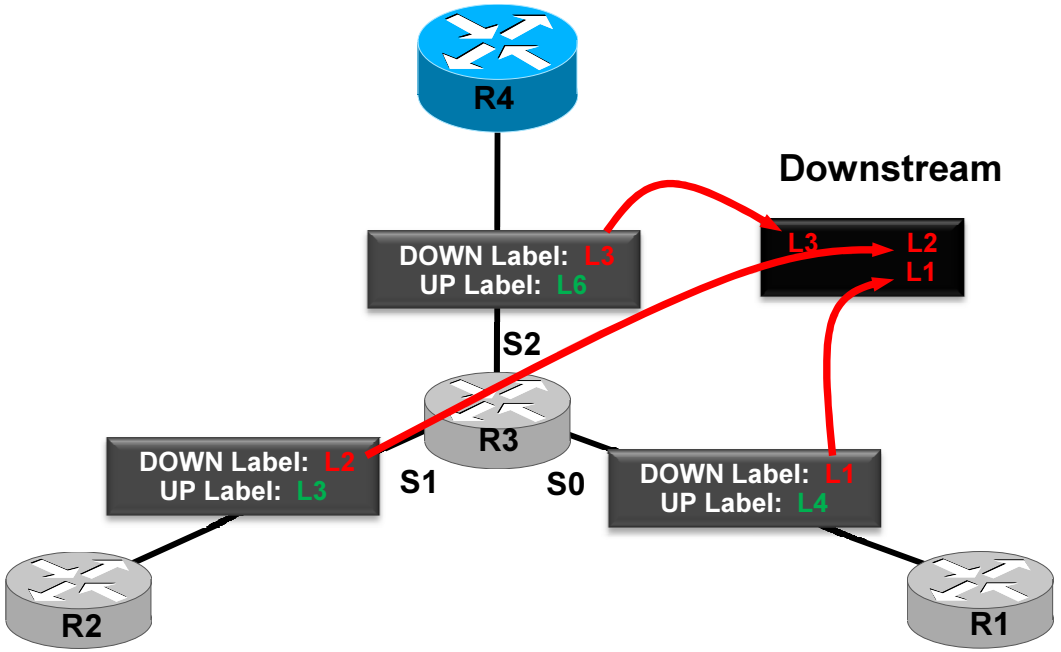


MP2MP Basic Operation



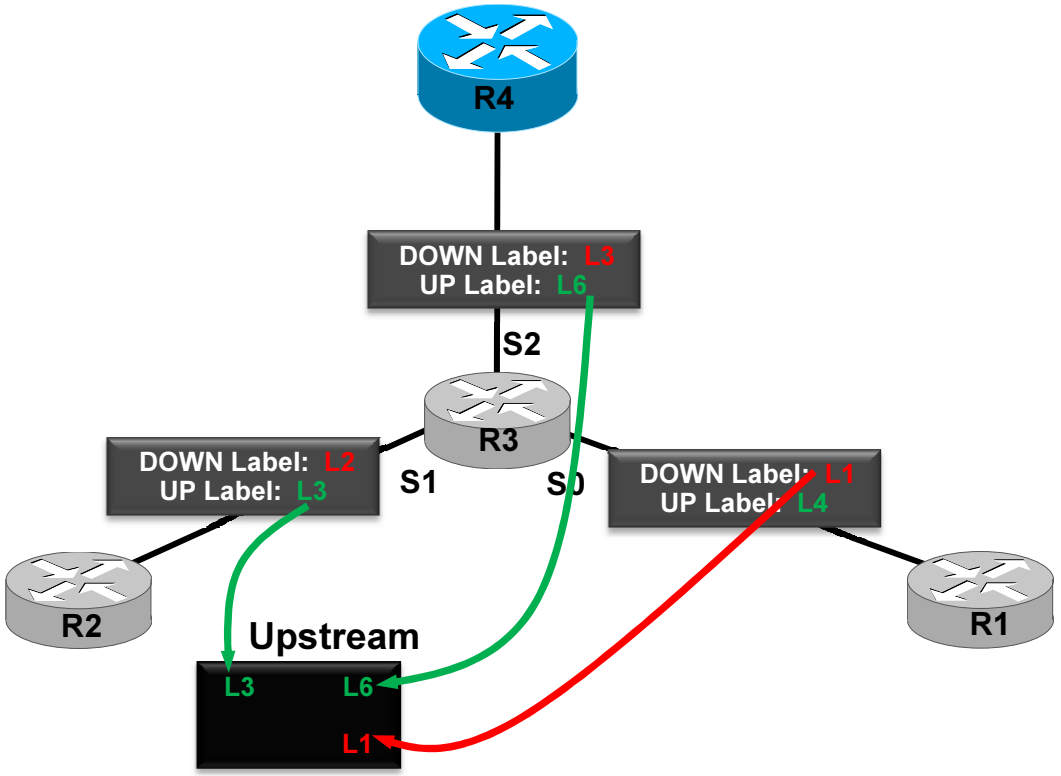
MP2MP Basic Operation

Look at R3 Downstream Label Replication Table from S2



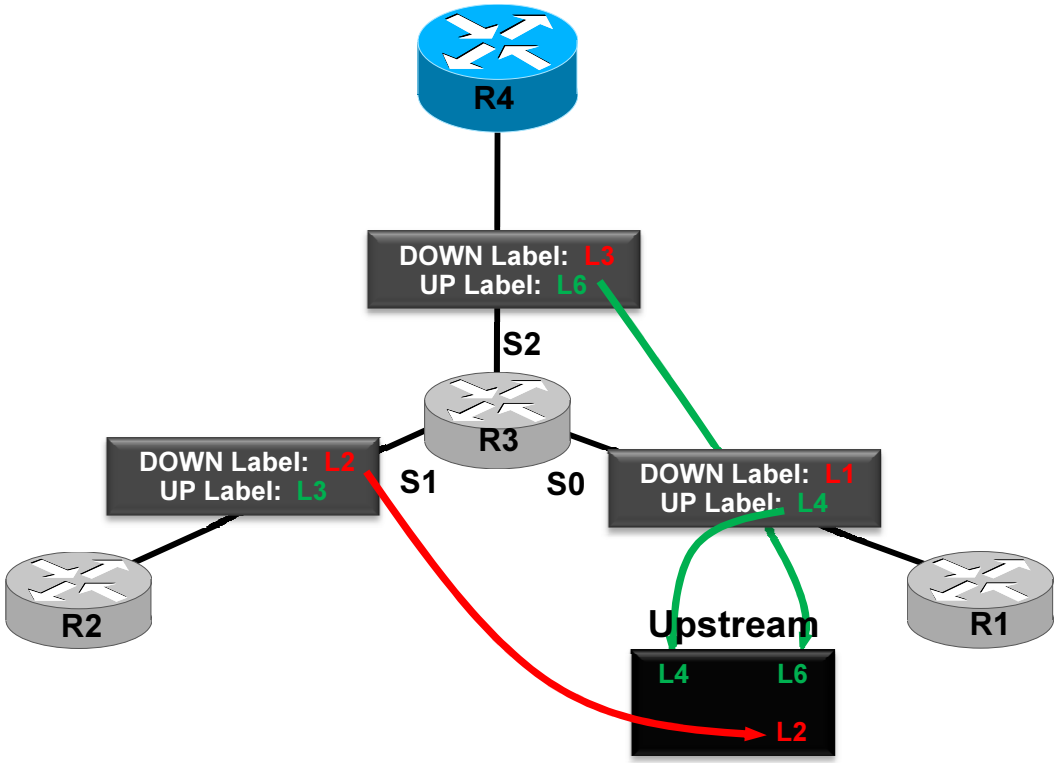
MP2MP Basic Operation

Look at R3 Upstream Label Replication Table from S1

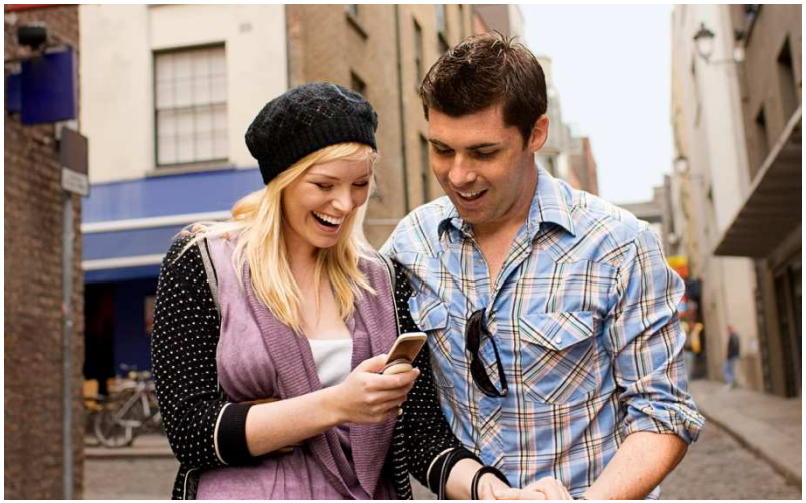


MP2MP Basic Operation

Look at R3 Upstream Label Replication Table from S0



P2MP TE



P2MP TE

- Extensions to RSVP-TE Protocol are defined via RFC 4875 to support P2MP TE LSPs
- P2MP TE LSP is initiated by the Ingress LSR towards the Egress LSRs
- Supports only P2MP LSPs
- Support Traffic Engineering
  - Explicit Routing
  - Fast ReRoute
  - BW Reservation

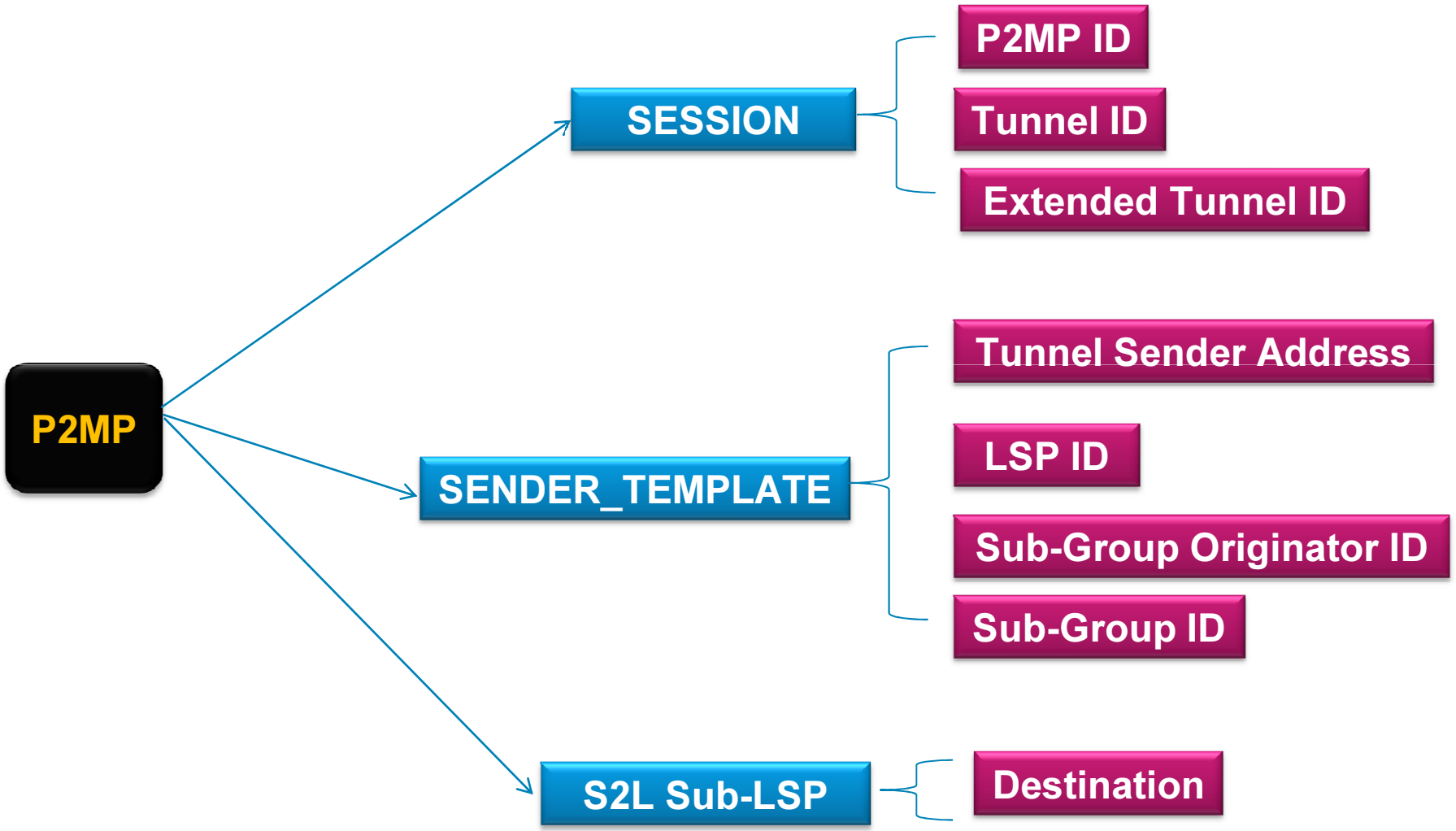
Terminology

Common Terms are covered earlier as part of the Introduction.

Ingress LSR Egress LSR P2P LSP P2MP LSP Upstream
Downstream Branch LSR Bud LSR Leaf LSR

Terminology	Description
Sub-LSP	A segment of a P2MP TE LSP that runs from one of the LSP's LSRs to one or more of its other LSRs
S2L Sub-LSP	Source to Leaf: A segment of a P2MP TE LSP that runs from HE to one Destination
Grafting	The operation of adding egress LSR(s) to an existing P2MP LSP
Pruning	An action where Egress LSR is removed from the P2MP LSP
Crossover	Crossover happens at an intersecting node when two or more incoming Sub-LSPs, belonging to the same LSP, have different input & different output interfaces
Remerge	Remerge happens at an intersecting node when two datastream belonging to the same P2MP LSP Merge into one datastream on output

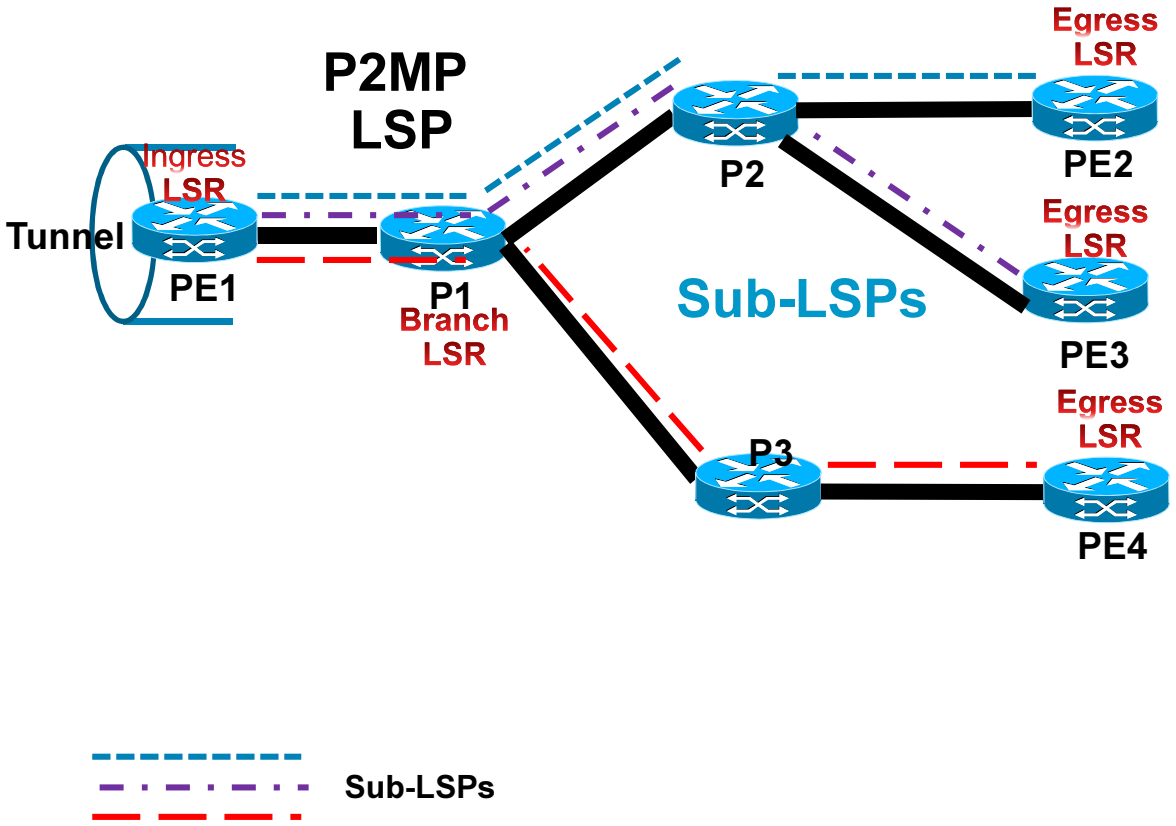
Extensions to RSVP for P2MP TE



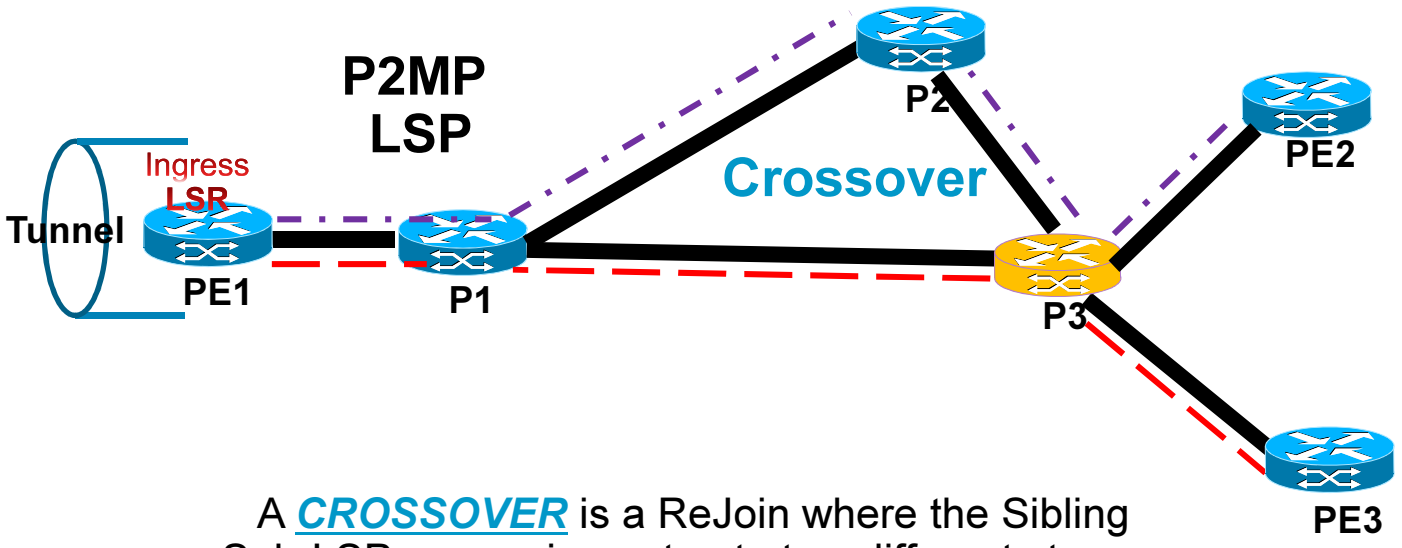
P2MP TE LSP

- It is ONE or MORE S2L Sub-LSPs
- It is a collection of all Sub-LSPs forms the P2MP LSP
- All Sub-LSPs belonging to the same P2MP LSP should share labels and resources when they share links
  - Share labels to prevent multiple copies of the same data being sent
- Identified by 5-Tuple Key

P2MP TE : Sub-LSPs



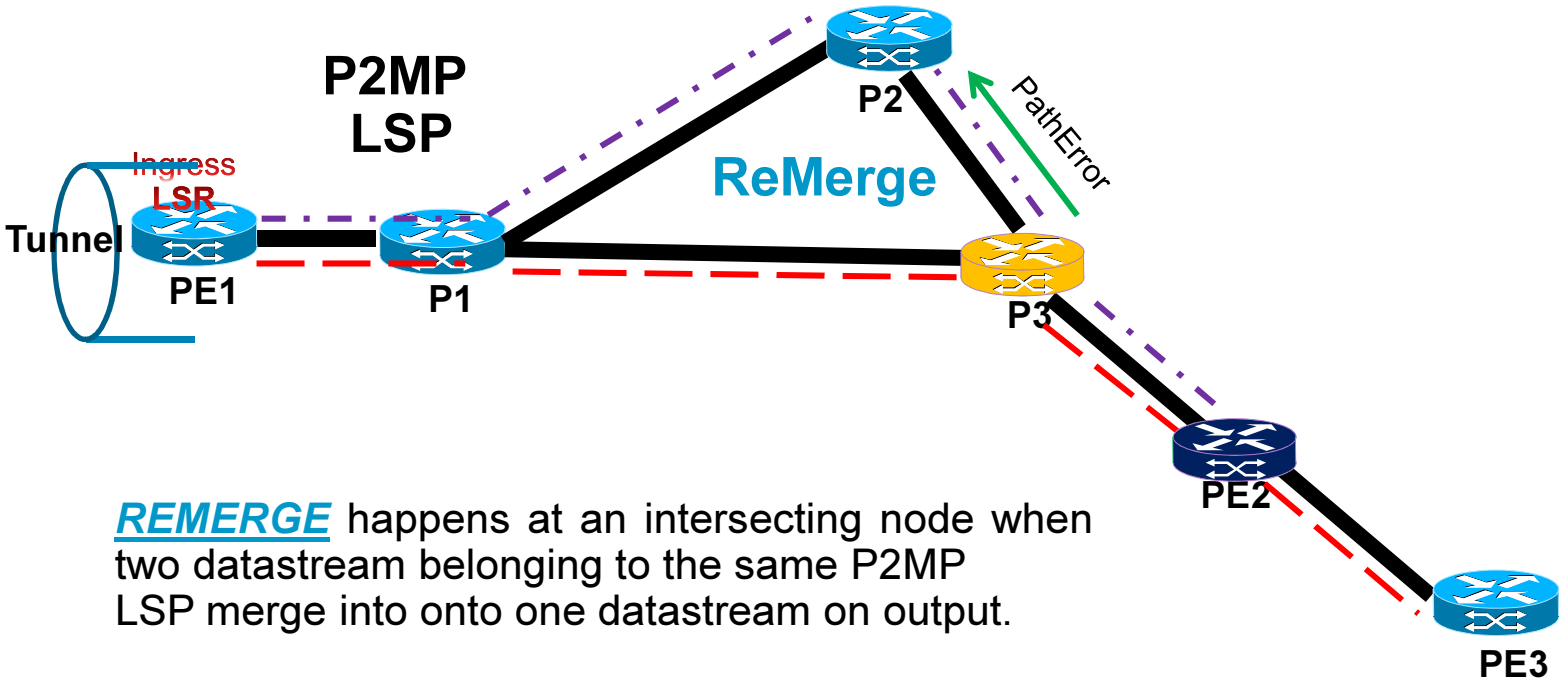
P2MP: Sub-LSPs Rejoin - Crossover



A ***CROSSOVER*** is a ReJoin where the Sibling Sub-LSPs are going out onto two different streams

Sub-LSPs

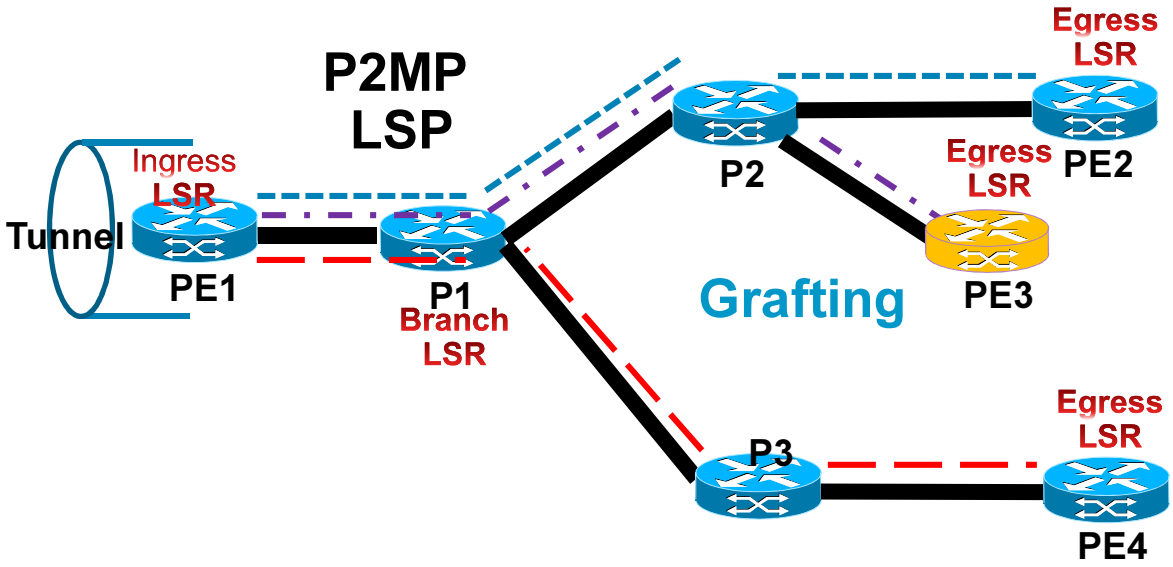
P2MP TE : Sub-LSPs Rejoin - ReMerge



REMERGE happens at an intersecting node when two datastream belonging to the same P2MP LSP merge into onto one datastream on output.

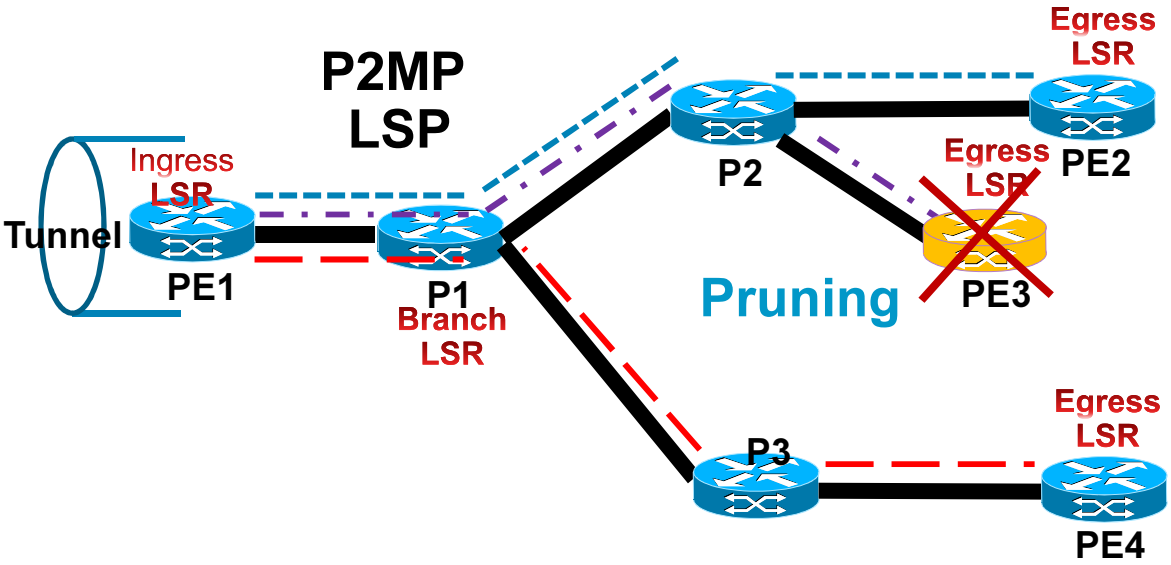
Sub-LSPs

P2MP TE: Sub-LSP Operation - Grafting



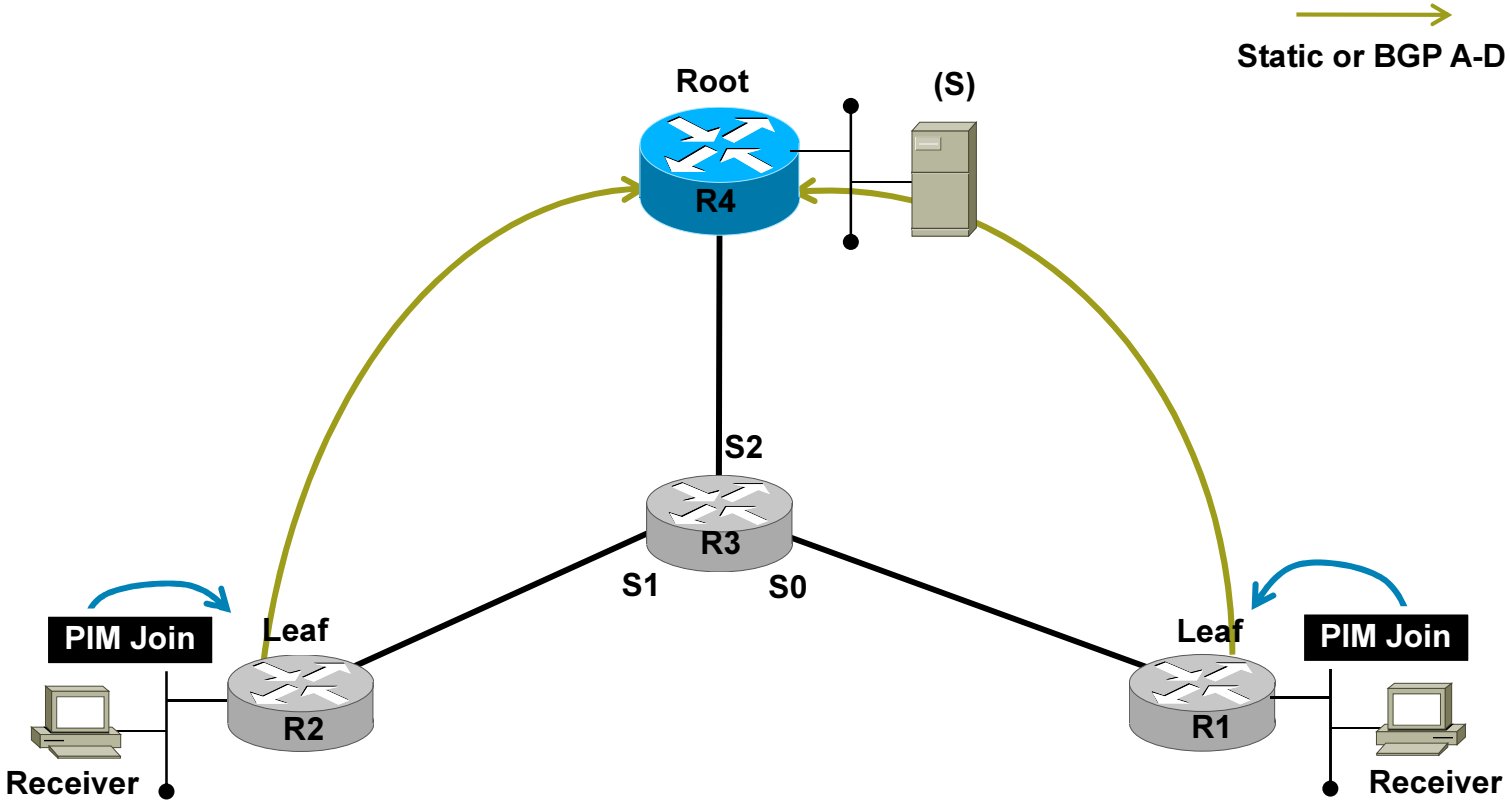
GRAFTING happens when a new Egress is added to an existing P2MP LSP. i.e., a new Sub-LSP (New SubGroup ID, New DST) is signaled with a new destination for an existing P2MP LSP (Same LSP ID)

P2MP TE: Sub-LSP Operation - Pruning

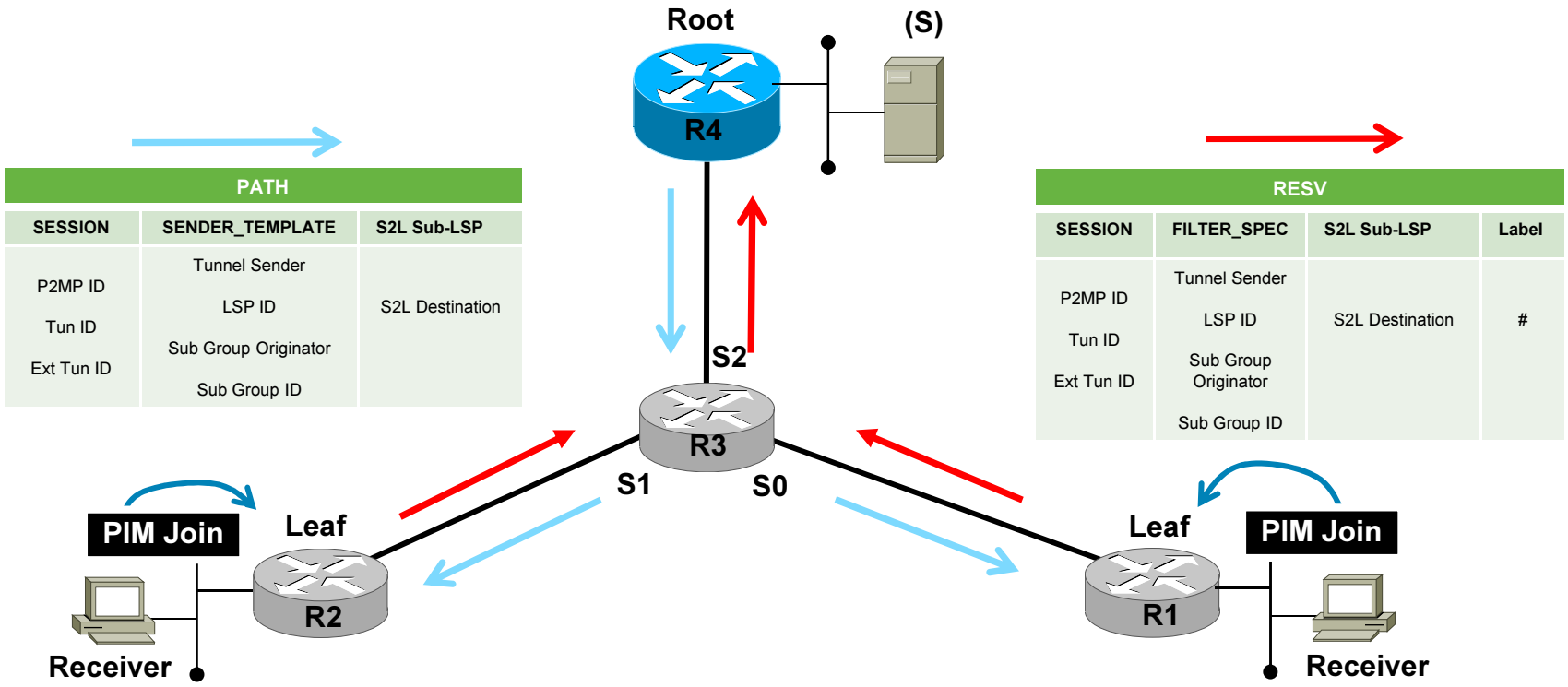


The operation of removing Egress LSRs from an existing P2MP LSP is termed ***PRUNING***

P2MP TE: Signaling



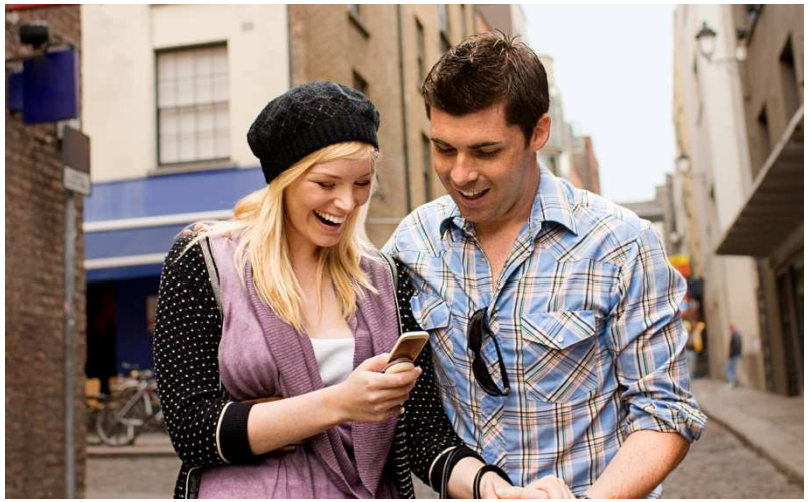
P2MP TE: Signaling



PATH		
SESSION	SENDER_TEMPLATE	S2L Sub-LSP
P2MP ID Tun ID Ext Tun ID	Tunnel Sender LSP ID Sub Group Originator Sub Group ID	S2L Destination

RESV			
SESSION	FILTER_SPEC	S2L Sub-LSP	Label
P2MP ID Tun ID Ext Tun ID	Tunnel Sender LSP ID Sub Group Originator Sub Group ID	S2L Destination	#

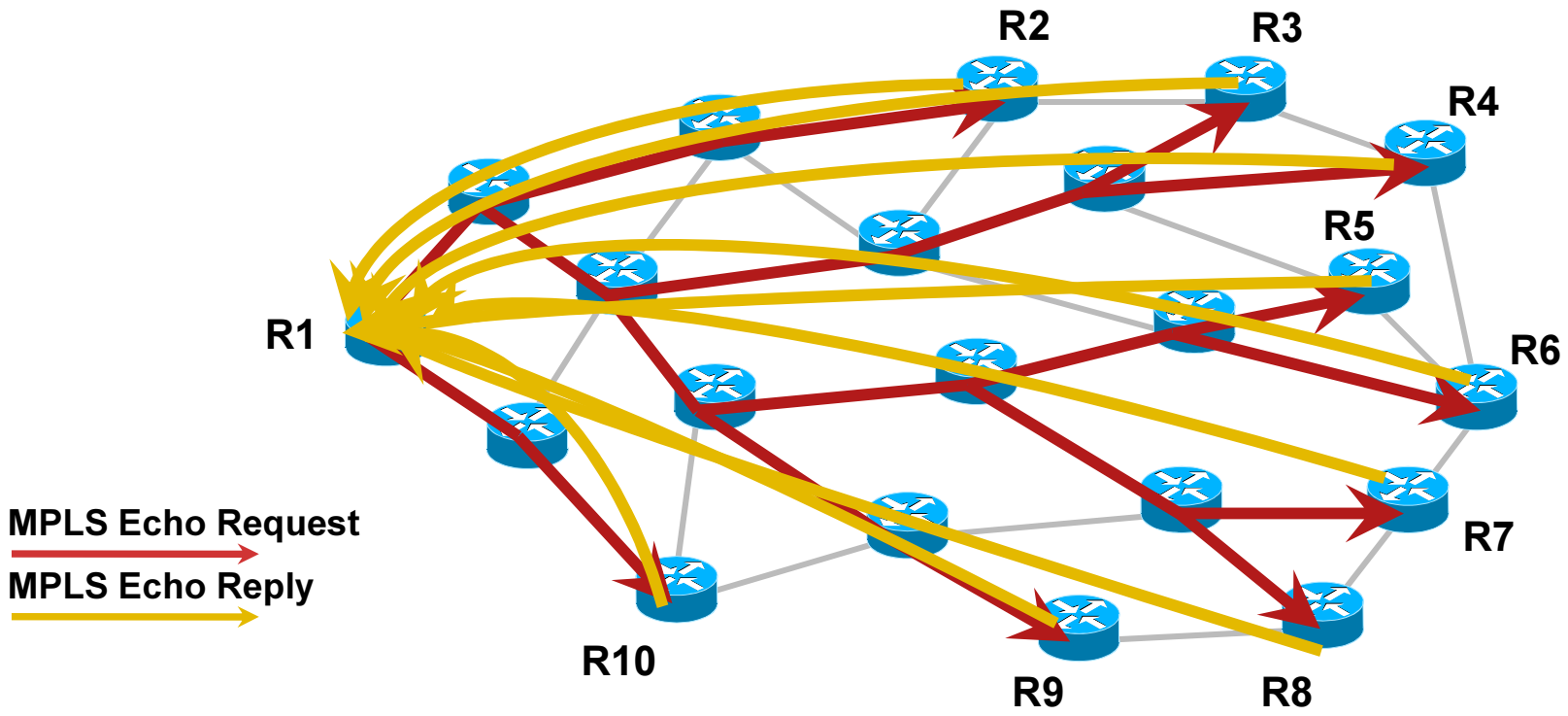
OAM Extensions for support of LSM



MPLS OAM for Multicast

- Multicast LSP Ping
- Proxy LSP Ping
- Multicast Connectivity Verification

Multicast LSP Ping



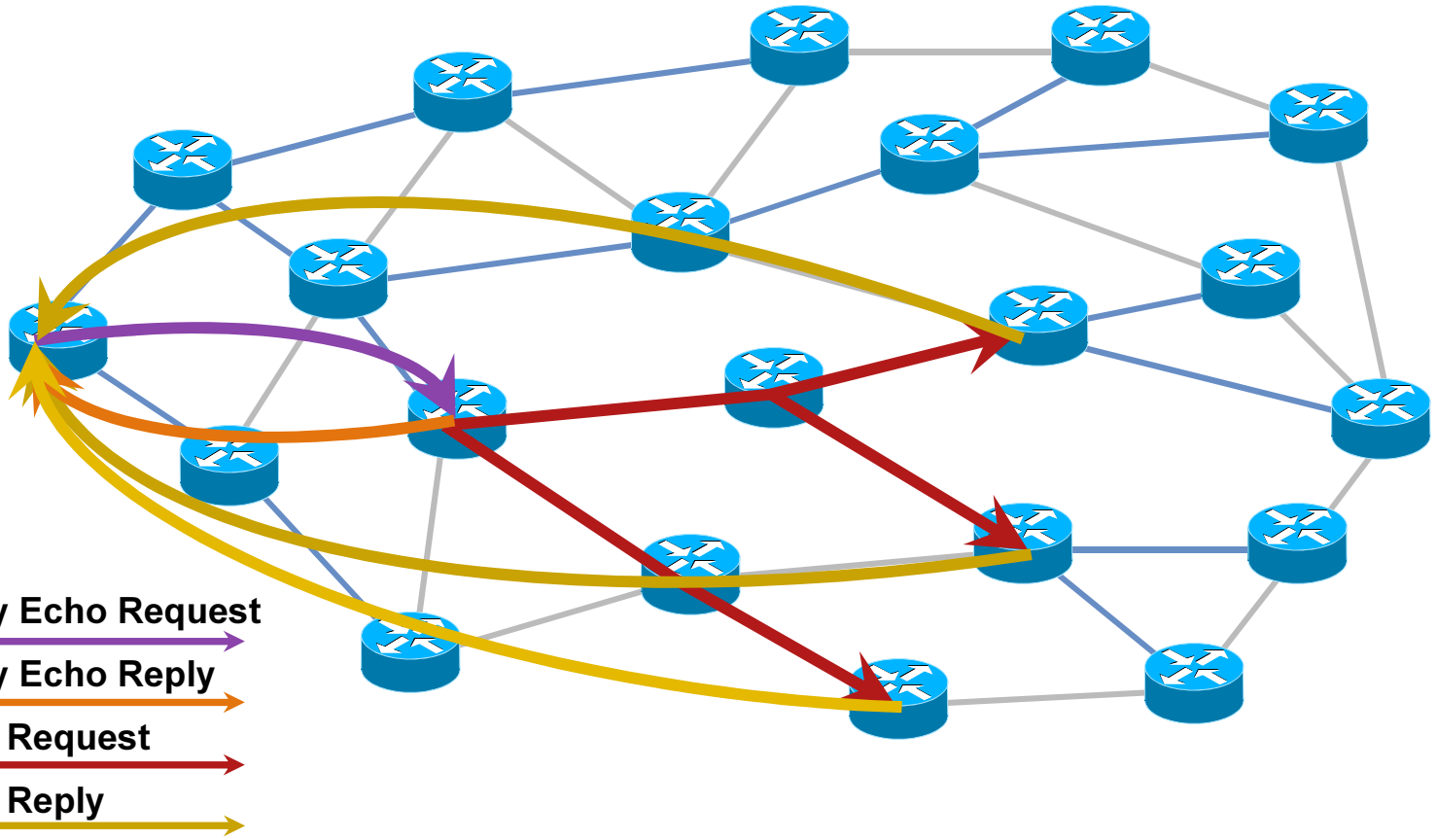
- Adds FEC Stack sub-TLVs for P2MP-TE and mLDP
- Adds capability to limit and jitter response
- Bud node indication (acting as both an egress and a mid-point)

draft-ietf-mpls-p2mp-lsp-ping

Proxy Ping Motivation

- Scalability
  - Reduce
    - number of replies
    - network wide processing of MPLS Echo Requests
  - Scoping can limit the number of replies but still requires processing at the ‘replying’ node
  - In some cases, scoping is not desired, but responses are only needed from a small region of the topology
- Previous Hop Information
  - Needed with mLDP

Proxy LSP Ping



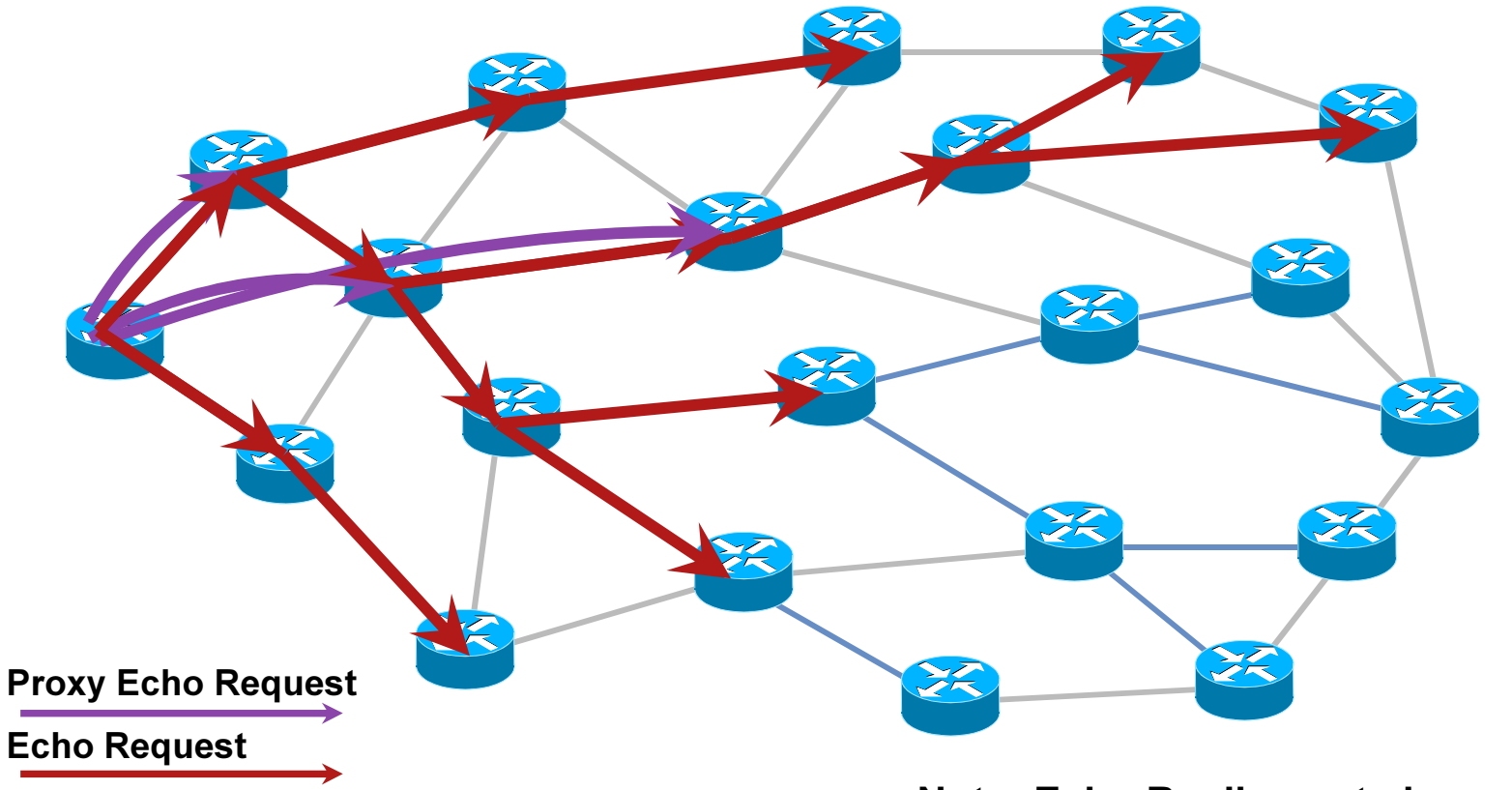
draft-ietf-mpls-remote-lsp-ping

P2MP TE Fault Localization

- In Traffic Engineering, the topology of the tree is known
- The branches which proceed towards the node reporting trouble can be determined
- By initiating trace further into the tree, traffic on the other branches can be avoided

P2MP TE Tracing with Proxy LSP Ping

Traffic on other branches avoided by initiating trace further into tree

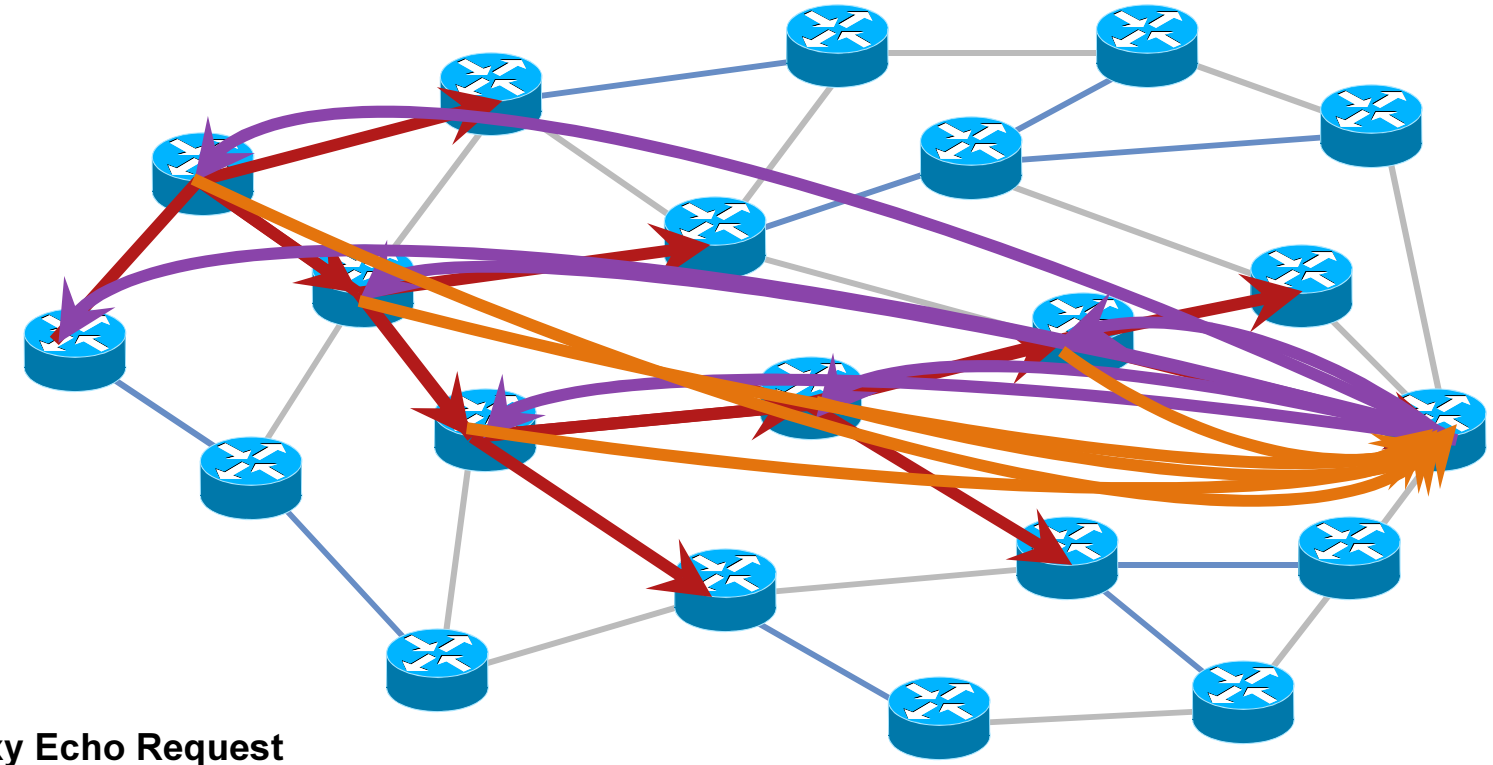


Note: Echo Replies not shown

Multicast LDP Tracing

- mLDP LSPs are initiated by the leaf nodes
- Root node may not know the tree topology
- Failure is most likely to be detected at leaf
- Leaf is a logical place to begin tracing

mLDP Traceroute



Proxy Echo Request

Proxy Echo Reply

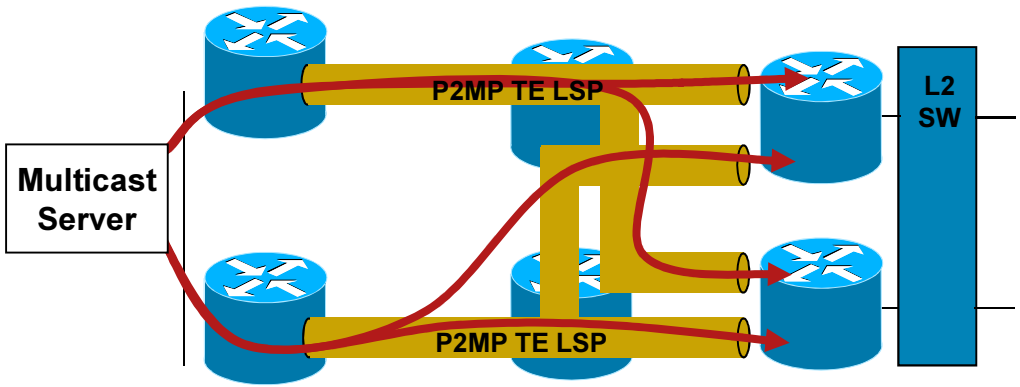
Echo Request

Echo Reply

- Tree is traced from leaf to root
- PHOPs are learned as trace progresses

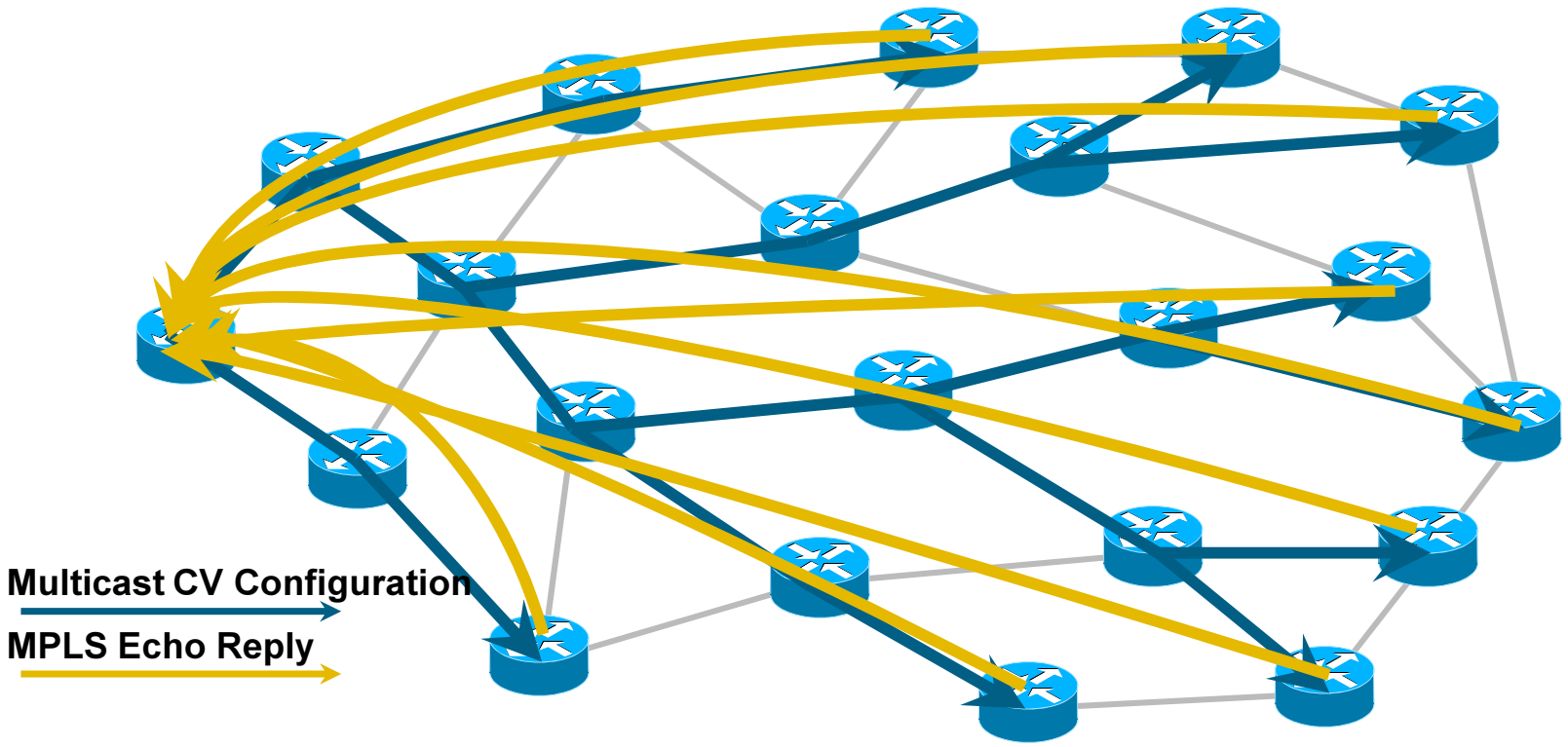
Requirements for Multicast Connectivity Verification

- Most P2MP-TE applications will require strong OAM
  - Television feeds
  - Stock ticker
- Scaling
  - Operators require low-impact on network.
  - Must deal with many endpoints (1000s)
- Auto-configuration
  - Simple initial configuration
  - Automatic OAM configuration on Prune/Join
- Many applications will use egress repair

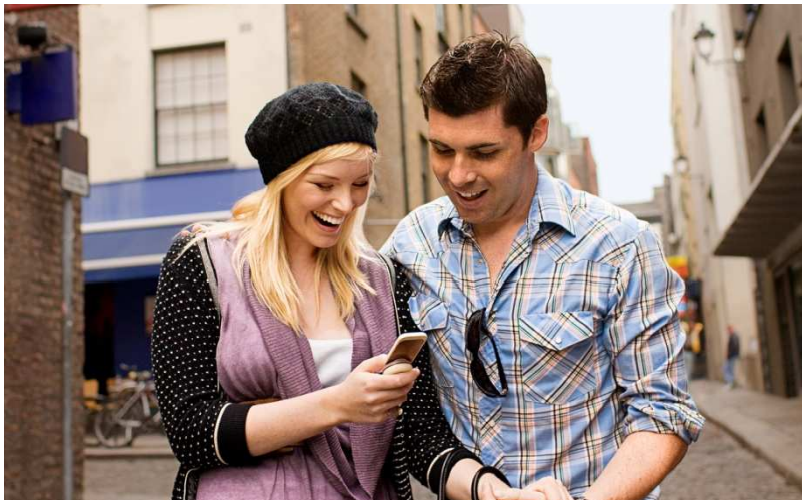


Failover requirement is 50 ms

Multicast CV Setup



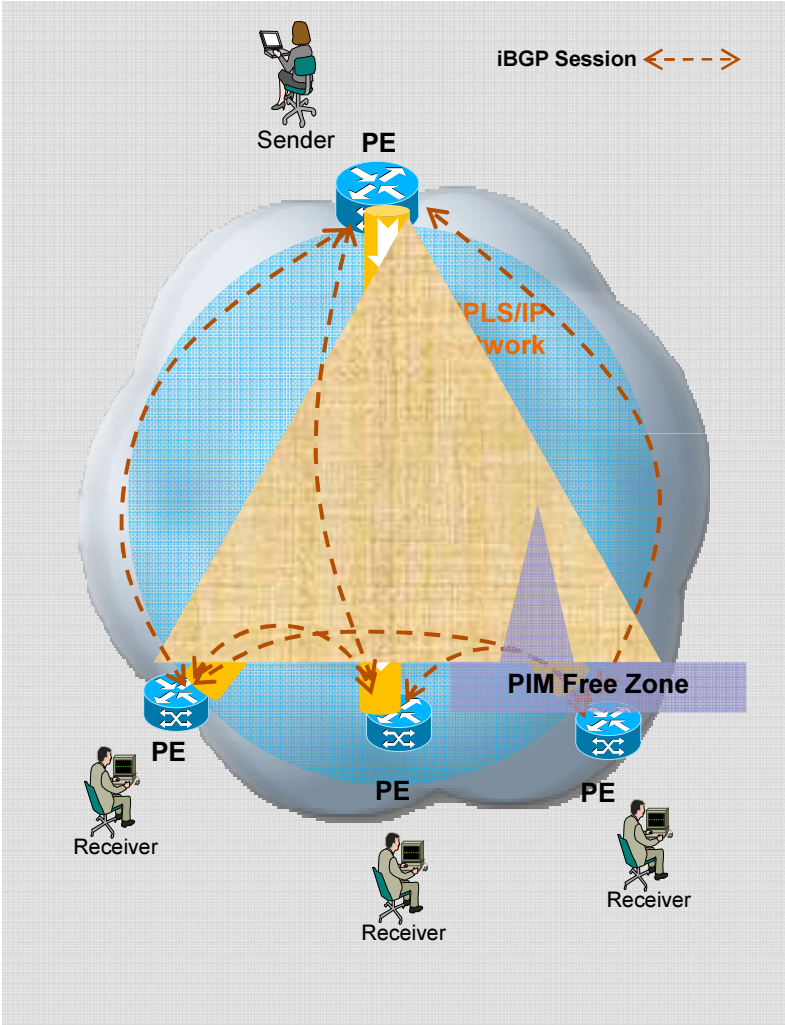
LSM Applications



Use Case: PIM-SSM Transit

- Supports IPv4 and IPv6 SSM multicast traffic
- Carried across core in P2MP LSP
- Source and Group are encoded into opaque value
  - Signalling of (S, G) state is done in-band
- Label Mapping message builds tree to root
  - Root is edge router connected to source (injects BGP route)
- Source prefixes distributed via BGP
- Root derived from BGP Next-Hop of Source
- PIM is present on the edge of the network
  - P2MP LSP in core of network

Use Case: PIM-SSM Transit



- PIM-SSM Transit

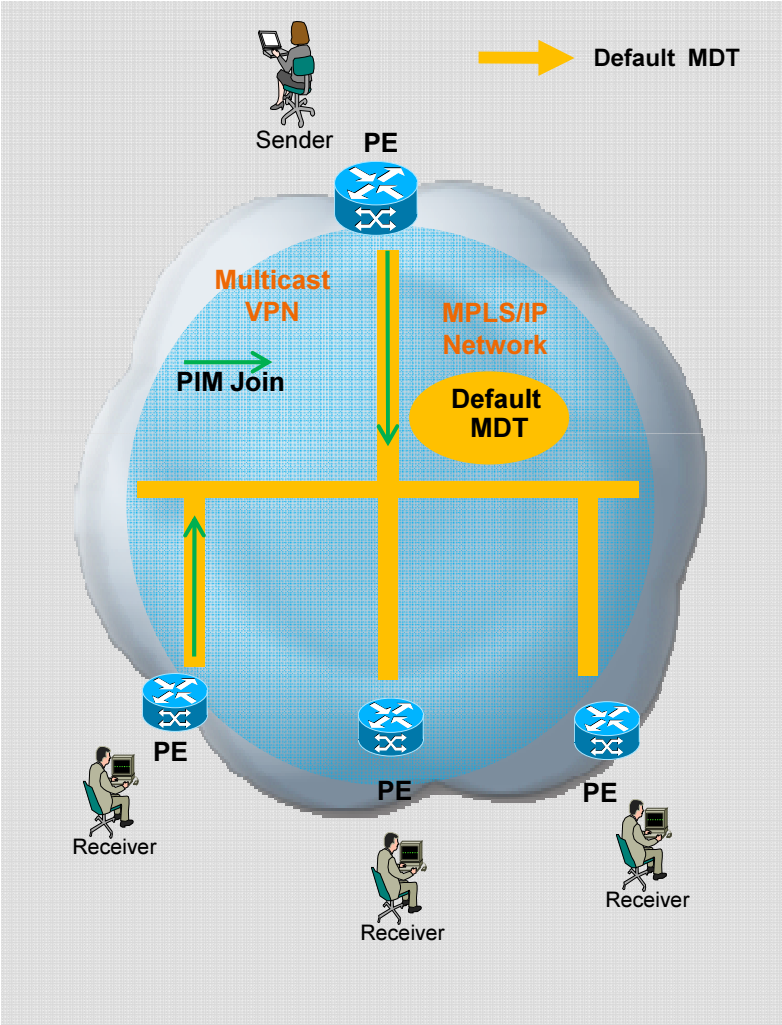
IPv4/v6 Opaque Value	
Source	Group

- PE Routers need to know the S and G (SSM)
- Solution for Multicast in Global Table
- Allows global PIM-SSM streams to be transported across the MPLS Core
- Source Prefixes distributed via BGP
- Useful when there is a need for carrying IP multicast over a MPLS network where there is an explicit requirement for MPLS encapsulation.

Use Case: Multicast VPN over mLDP

- mLDP supports Multicast Distribution Trees (mVPNs)
- mVPN solution is independent of the tunnelling mechanism
  - PIM with GRE encapsulation (Native Multicast)
  - mLDP with MPLS encapsulation
- Default-MDT uses MP2MP LSPs
  - Supports low bandwidth and control traffic between VRFs
- Data-MDT uses P2MP LSPs
  - Supports single high bandwidth source stream from a VRF
- All other operation of the mVPN remains the same
  - PIM neighbors in VRF seen across LSP-VIF
  - VPN multicast state signalling via PIM
- VPN-ID is used in place of MDT Multicast Group address

Use Case: Multicast VPN over mLDP

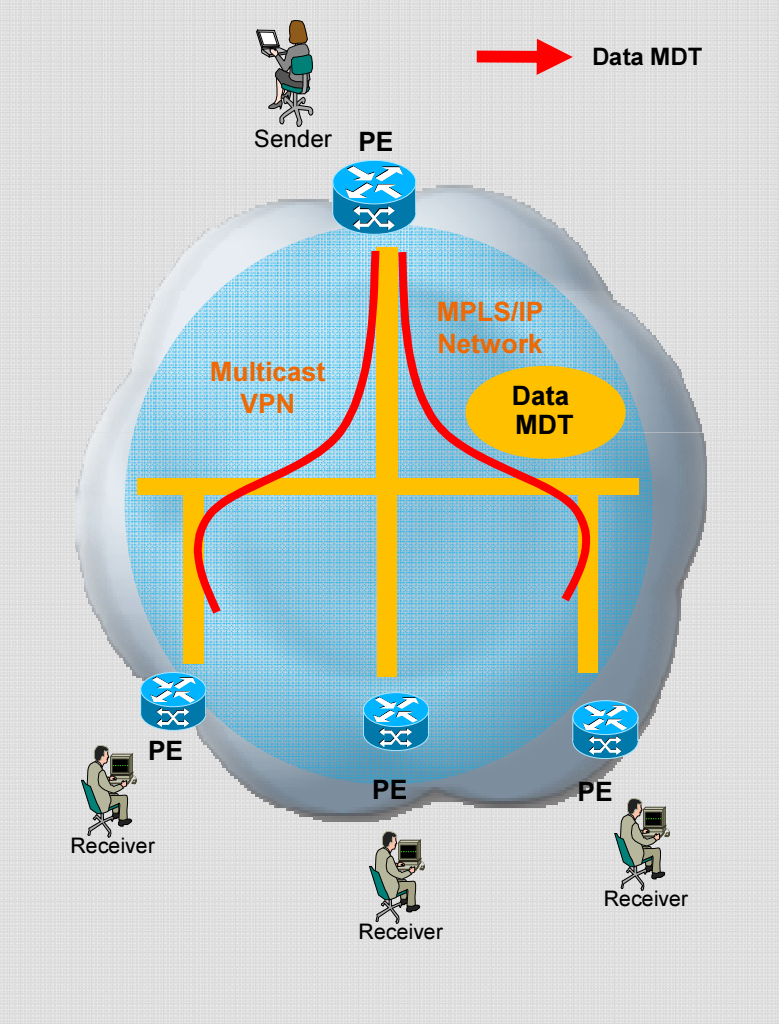


- Default-MDT

MDT Opaque Value	
VPN-ID	MDT# = 0

- Customer CE Devices joins the MPLS-Core through Provider's PE Devices
- The MPLS Core forms a Default-MDT for a given customer
- The Opaque Value used to signal a Default MDT
- It has two parameters:
  - ✓ VPN-ID
  - ✓ MDT number
- MVPN associates an interface for head and tail-end on the MP2MP LSP (just like a Tunnel interface).

Use Case: Multicast VPN over mLDP



- Data-MDT

MDT Opaque Value	
VPN-ID	MDT# > 0

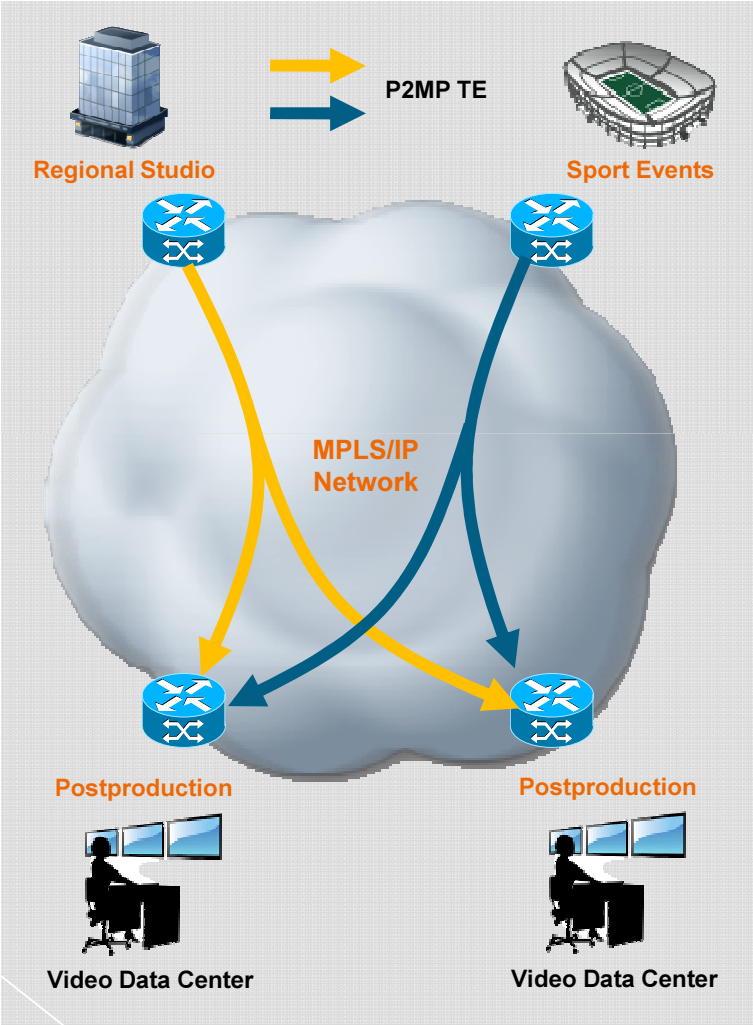
- Optionally a Data-MDT can be built based on traffic thresholds on sending PE when high BW source appears in the customer network.

- Data-MDT uses P2MP LSPs to support high BW (S,G) Stream

- Data-MDTs built for (S,G) in the mVPN

- The Opaque value is used to signal Data-MDT
  - ✓ VPN-ID
  - ✓ MDT #
  - ✓ (S,G)

Use Case: P2MP TE for Video Contribution



Typical Users:

- Broadcasters
- Content distribution providers

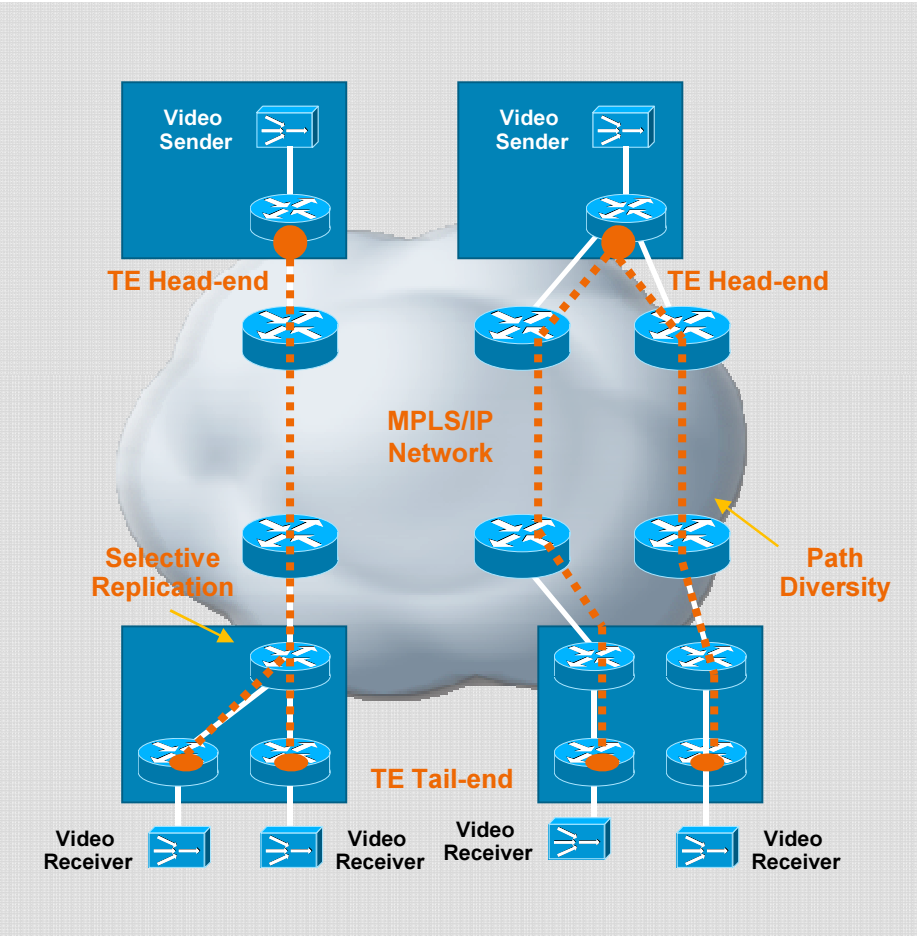
Deployment Requirements:

- Bandwidth management
- Explicit networks paths
- P2MP traffic distribution
- Network failure protection

LSM Application:

- Constrained-based P2MP TE tunnels (explicit paths, BW)
- TE FRR for link protection

Video Contribution Network



An typical P2MP TE Deployment for the Video distribution

Two Methods to push the Video

1. Push the replication as close to the receivers as possible
2. Two LSPs in the core with active and back up role

Use Case: FRR for Traffic Protection

- RSVP-TE P2MP supports FRR using unicast link protection.
- mLDP also supports FRR using RSVP TE unicast link protection
- Technology allows FRR protection on a per core tree basis for MLDP
- Ability to choose what type of traffic to go to the backup link (if there is concern like cost on backup link, etc.)

Common LSP Protocol in the Core

- RSVP-TE vs LDP

- Each protocol has its pros and cons
- MLDP and P2MP TE can coexist in a network
  - •MLDP for general purpose MVPN or global multicast routing;
  - •P2MP TE for application like studio to studio traffic which may require BW reservation

Summary

- Label Switched Multicast (LSM) offers RSVP and LDP extensions for optimized P2MP MPLS forwarding
- Depending up on the application requirements, either of the P2MP and MP2MP LSPs can be used
- OAM Extensions for support of LSM are being enhanced via Multicast LSP Ping, Proxy LSP Ping and Multicast CV
- Typical LSM Applications are for mVPN, PIM SSM Transit, Enterprise & Service Provider Video Contribution & Distribution