Introduction to MPLS Technologies

Santanu Dasgupta

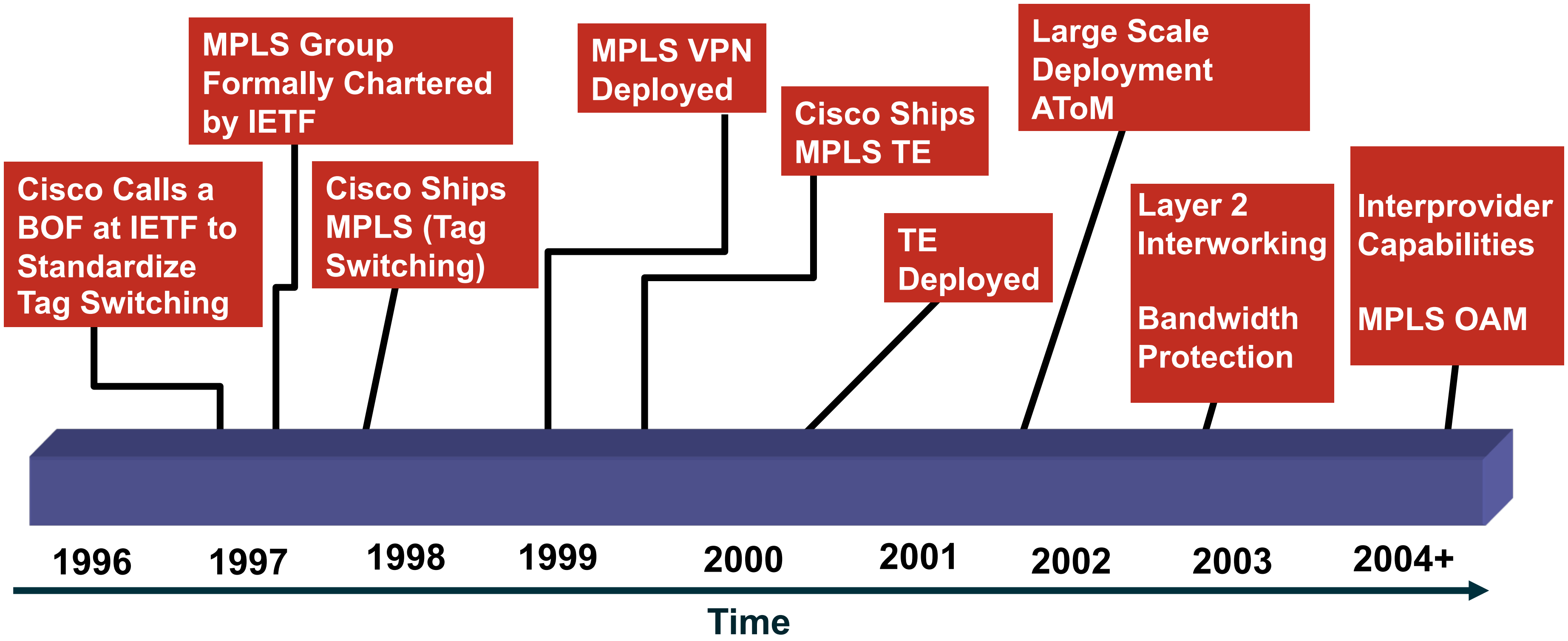
Why MPLS?

What Is MPLS?

- Multi Protocol Label Switching is a technology for delivery of IP services
- MPLS technology switches packets (IP packets, AAL5 frames) instead of routing packets to transport the data
- MPLS packets can run on other Layer 2 technologies such as ATM, FR, PPP, POS, Ethernet
- Other Layer 2 technologies can be run over an MPLS network

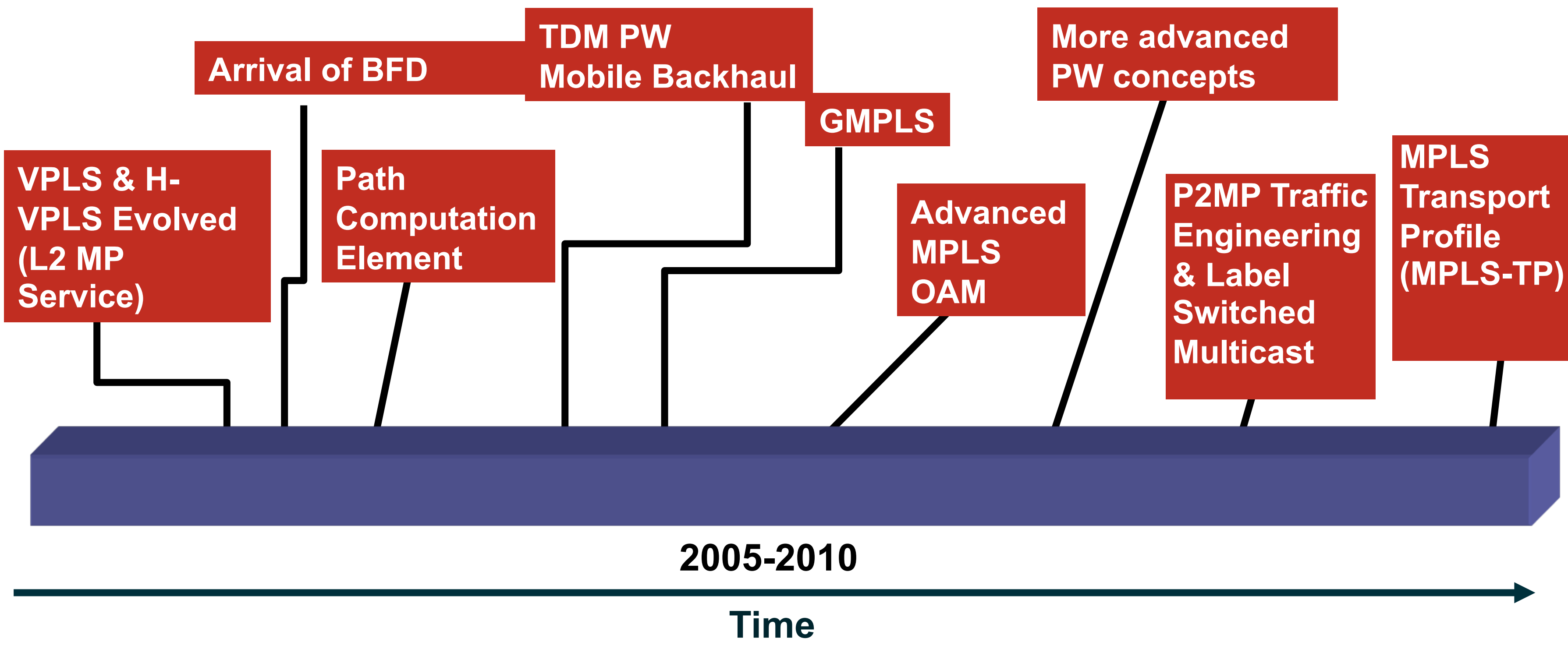
Evolution of MPLS

- It has evolved a long way from the original goal
- From tag switching
- Proposed in IETF—later combined with other proposals from IBM (ARIS), Toshiba (CSR)

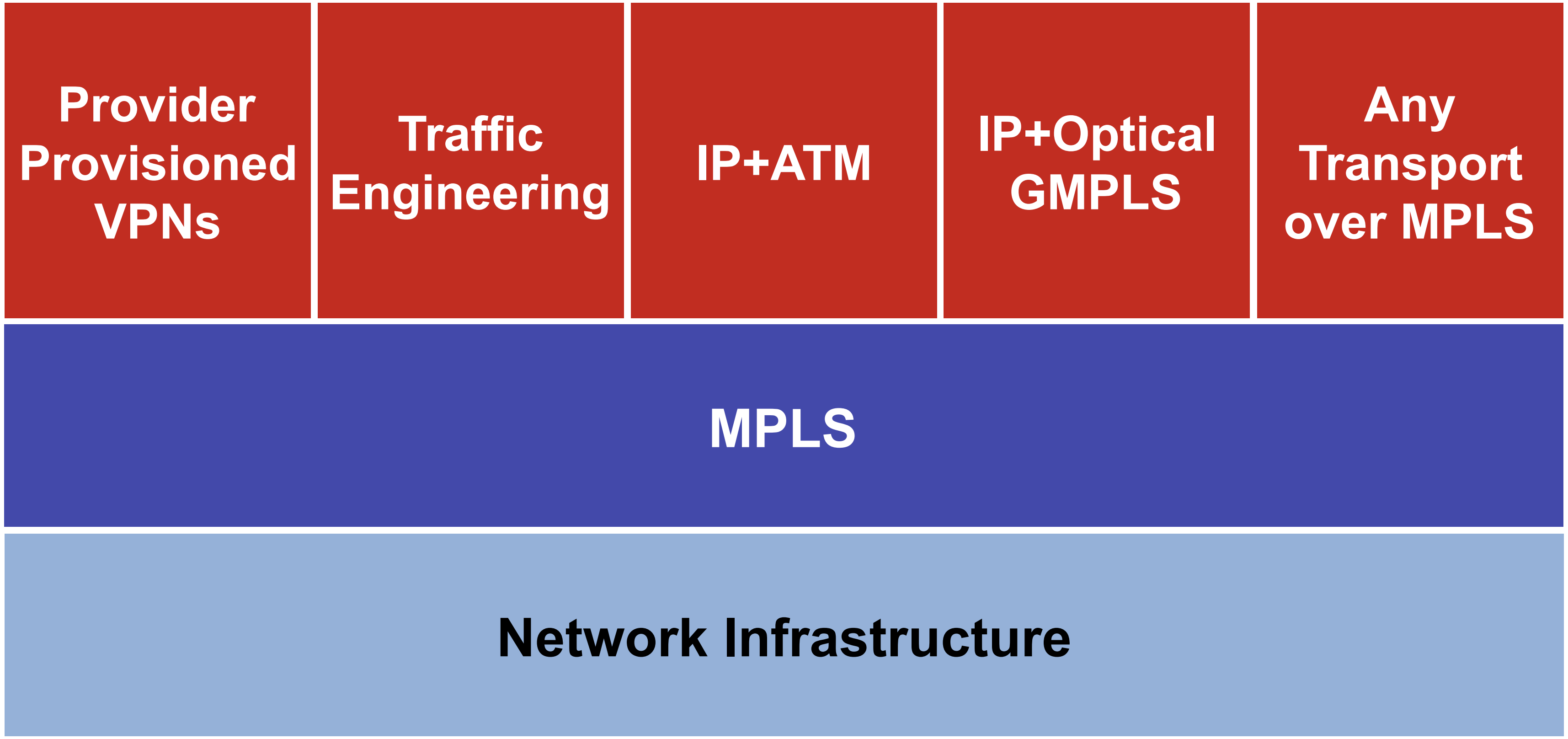


Evolution of MPLS

- Has been continuously evolving
- Multiple working groups at IETF are still focusing on more advancements
- Huge deployment across the world



MPLS as a Foundation for Value-Added Services



Technology Basics

MPLS Components

Few Components Play Role in Creating MPLS Network:

- IGP: Core Routing Protocol
- MPLS Label
- Encapsulation of MPLS label
- Forwarding Equivalence Class
- Label Distribution Protocol
- MPLS Applications related protocols: MP-BGP, RSVP…etc.

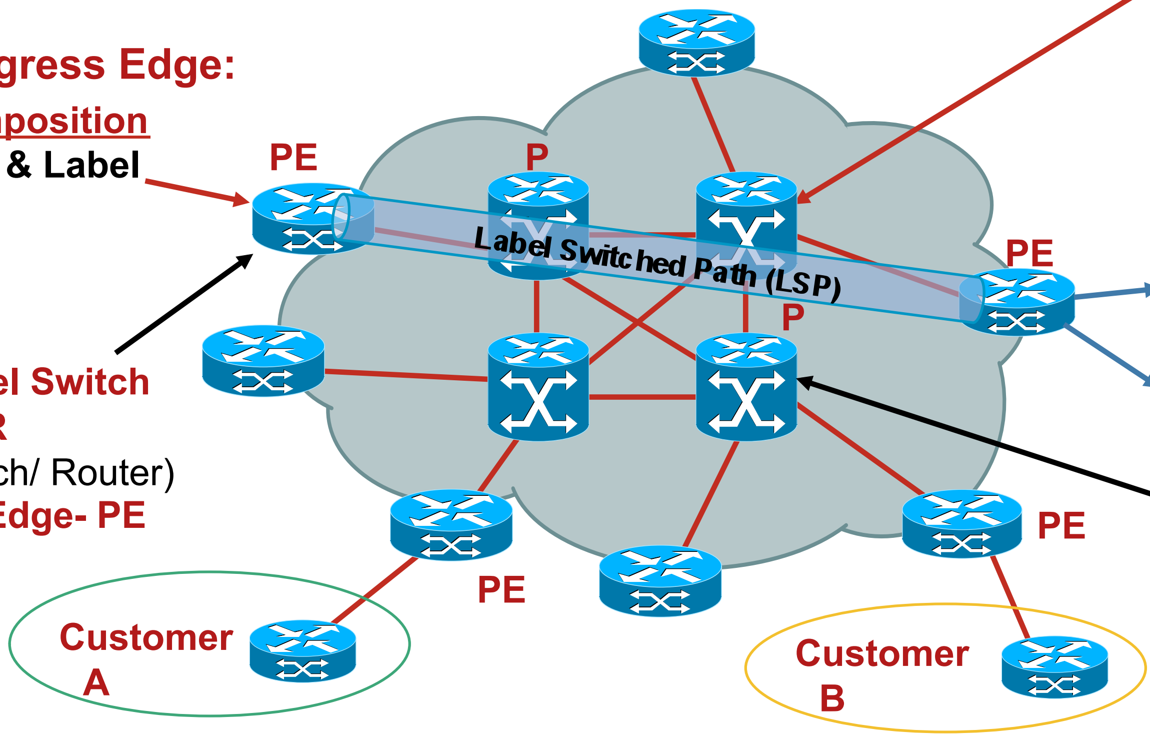
MPLS Network Overview
MPLS Core and Edge, Remote Customer Sites

1. At Ingress Edge:

Label imposition

Classify & Label packets

Edge Label Switch Router OR
(ATM Switch/ Router)
Provider Edge- PE



2. In the Core:

Label swapping or switching

Forward using labels (not IP addr). Label indicates service class and destination

3. At Egress Edge:

Label disposition

Remove labels and forward packets

Label Switch Router (LSR) or P (Provider) router

Router OR ATM switch + label switch controller

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

Label – 20bits	COS	S	TTL-8bits

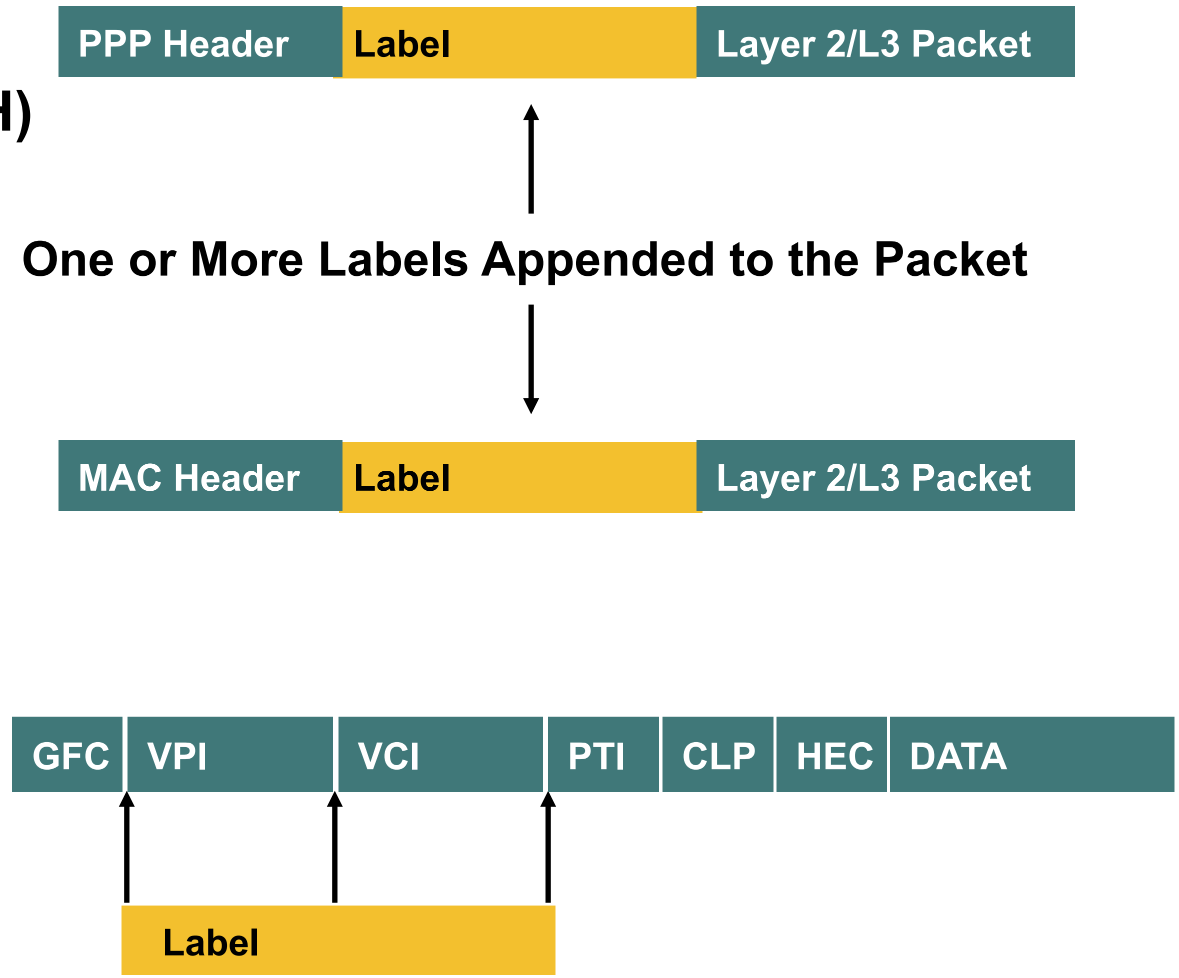
COS/EXP = Class of Service: 3 Bits; S = Bottom of Stack; TTL = Time to Live

MPLS Components Encapsulations

PPP Header (Packet over SONET/SDH)

LAN MAC Label Header

ATM MPLS Cell Header



MPLS Components
Forwarding Equivalence Class

FEC Is Used by Label Switching Routers to Determine How Packets Are Mapped to Label Switching Paths (LSP):

- IP prefix/host address
- Layer 2 circuits (ATM, FR, PPP, HDLC, Ethernet)
- Groups of addresses/sites—VPN x
- A bridge/switch instance—VSI
- Tunnel interface—traffic engineering

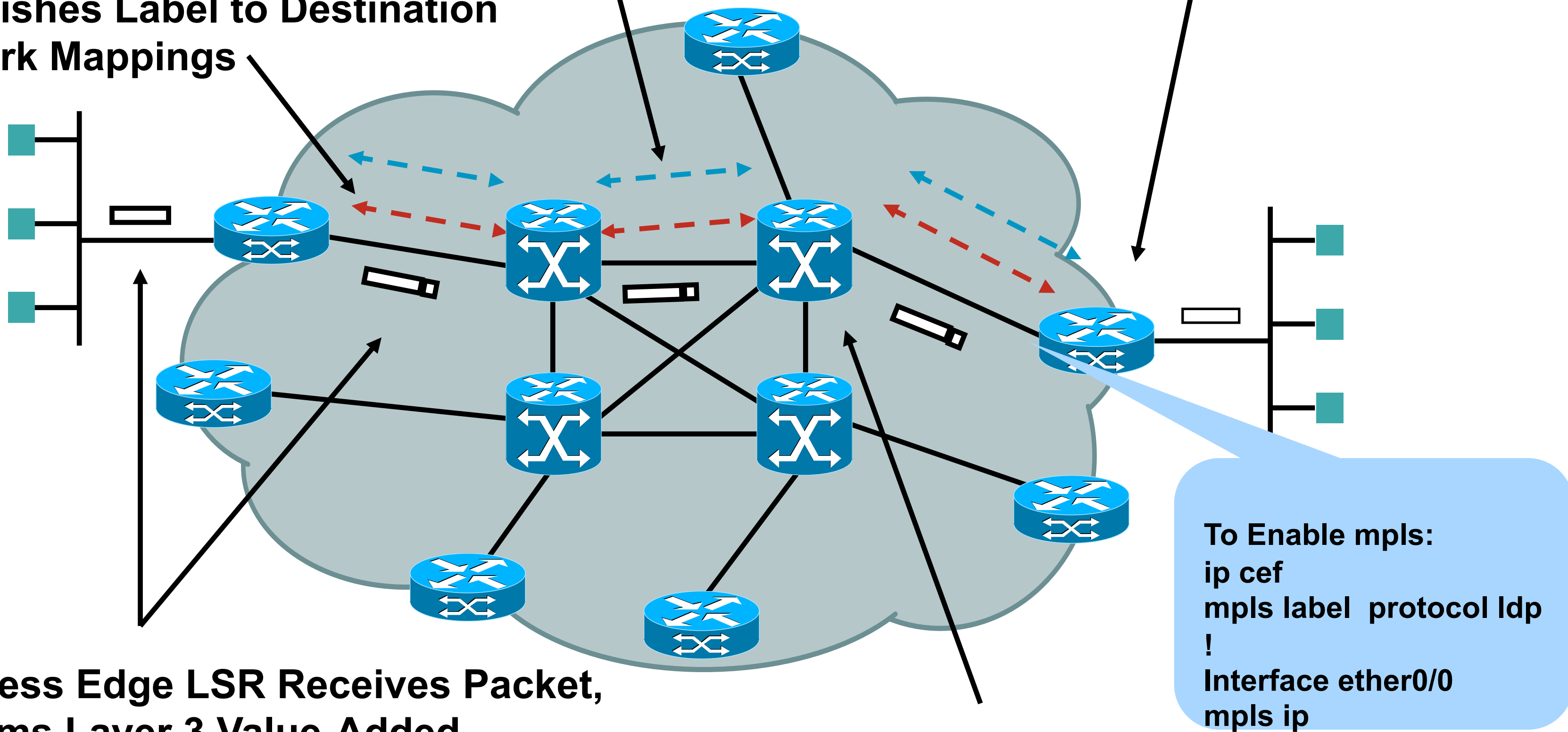
Label Distribution in MPLS Networks

MPLS Operation Overview

1a. Existing Routing Protocols (e.g. OSPF, IS-IS) Establish Reachability to Destination Networks

1b. Label Distribution Protocol (LDP) Establishes Label to Destination Network Mappings

4. Edge LSR at Egress Removes Label and Delivers Packet



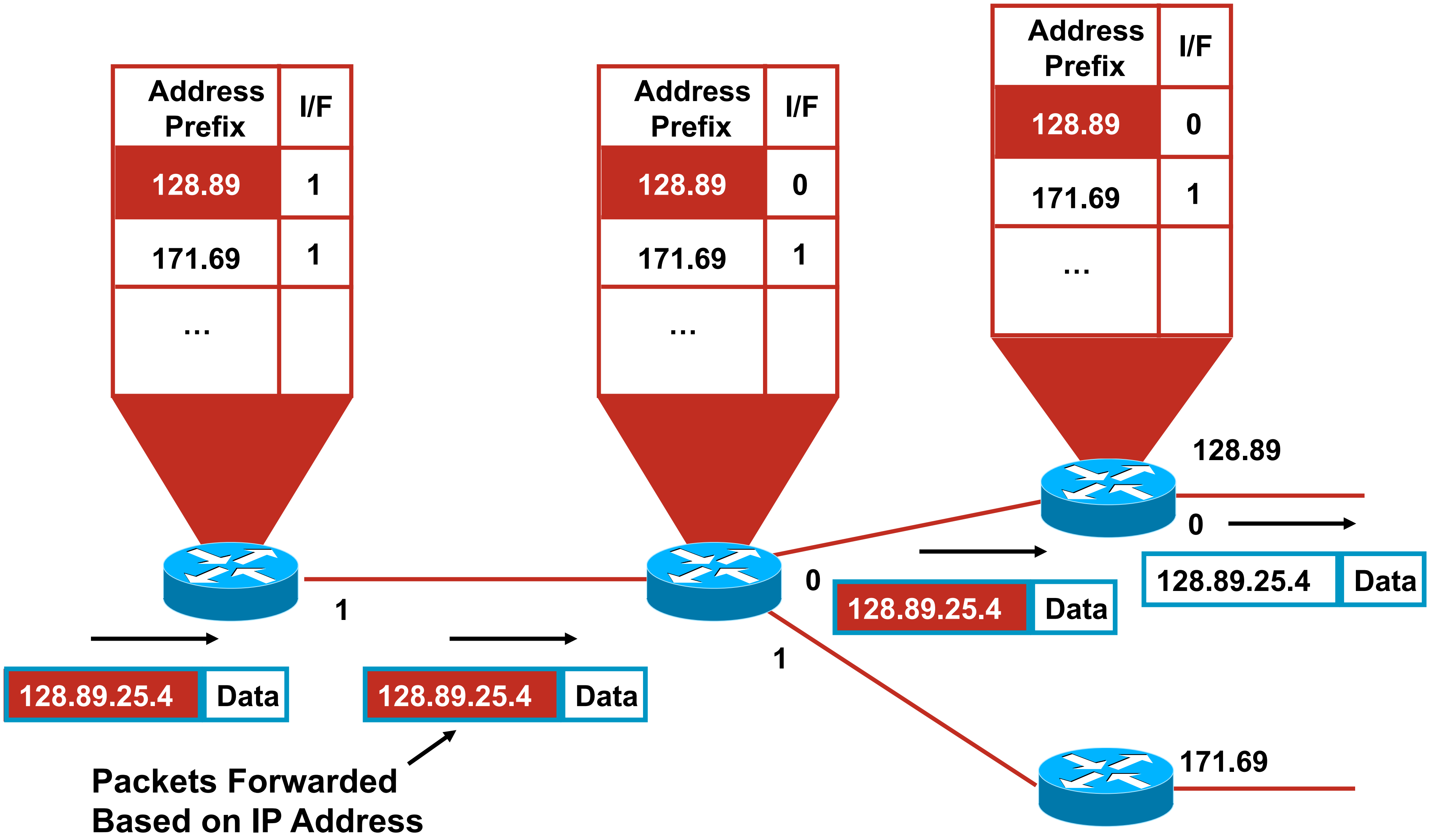
2. Ingress Edge LSR Receives Packet, Performs Layer 3 Value-Added Services, and "Labels" Packets

3. LSR Switches Packets Using Label Swapping

Label Advertisement Modes

- Downstream unsolicited
  - Downstream node just advertises labels for prefixes/FEC reachable via that device
- Downstream on-demand
  - Upstream node requests a label for a learnt prefix via the downstream node
  - Next example—ATM MPLS

IP Packet Forwarding Example

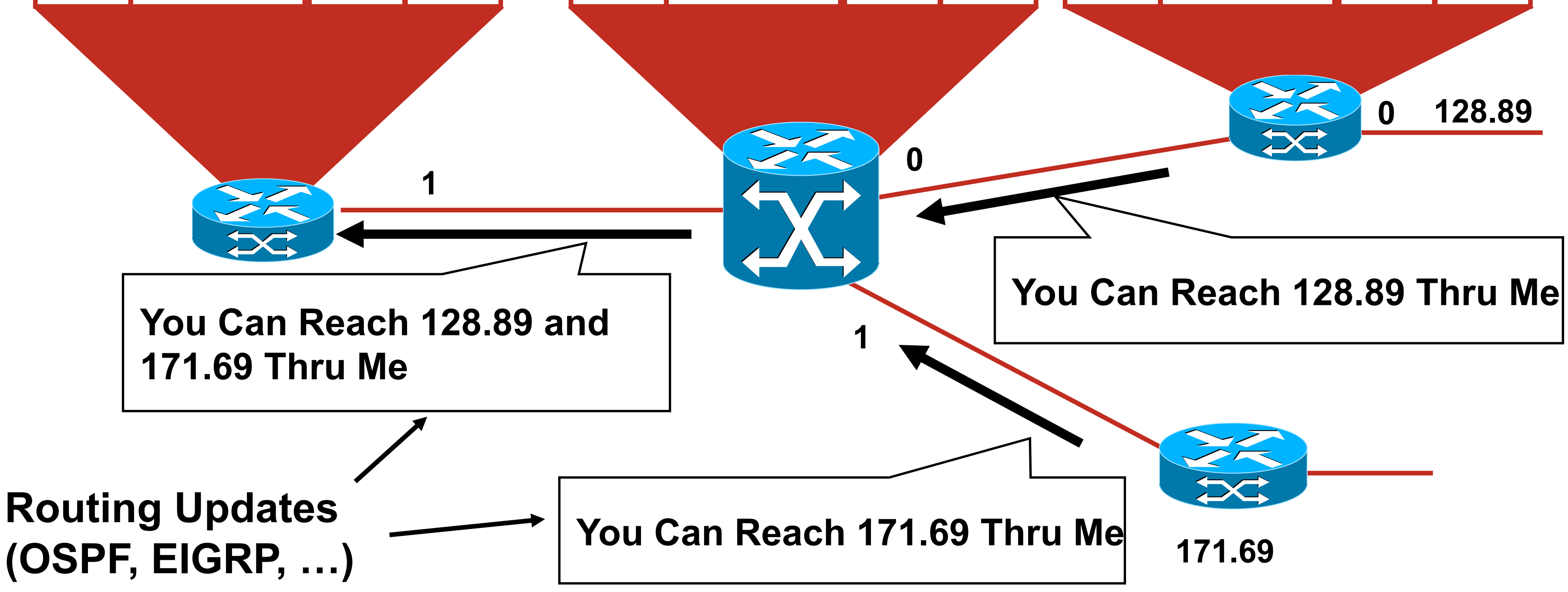


MPLS with Downstream Unsolicited Mode
Step I: Core Routing Convergence

In Labe	Address Prefix	Out I'face	Out Label
	128.89	1	
	171.69	1	
	…	…	

In Labe	Address Prefix	Out I'face	Out Label
	128.89	0	
	171.69	1	
	…	…	

In Labe	Address Prefix	Out I'face	Out Label
	128.89	0	
	…	…	

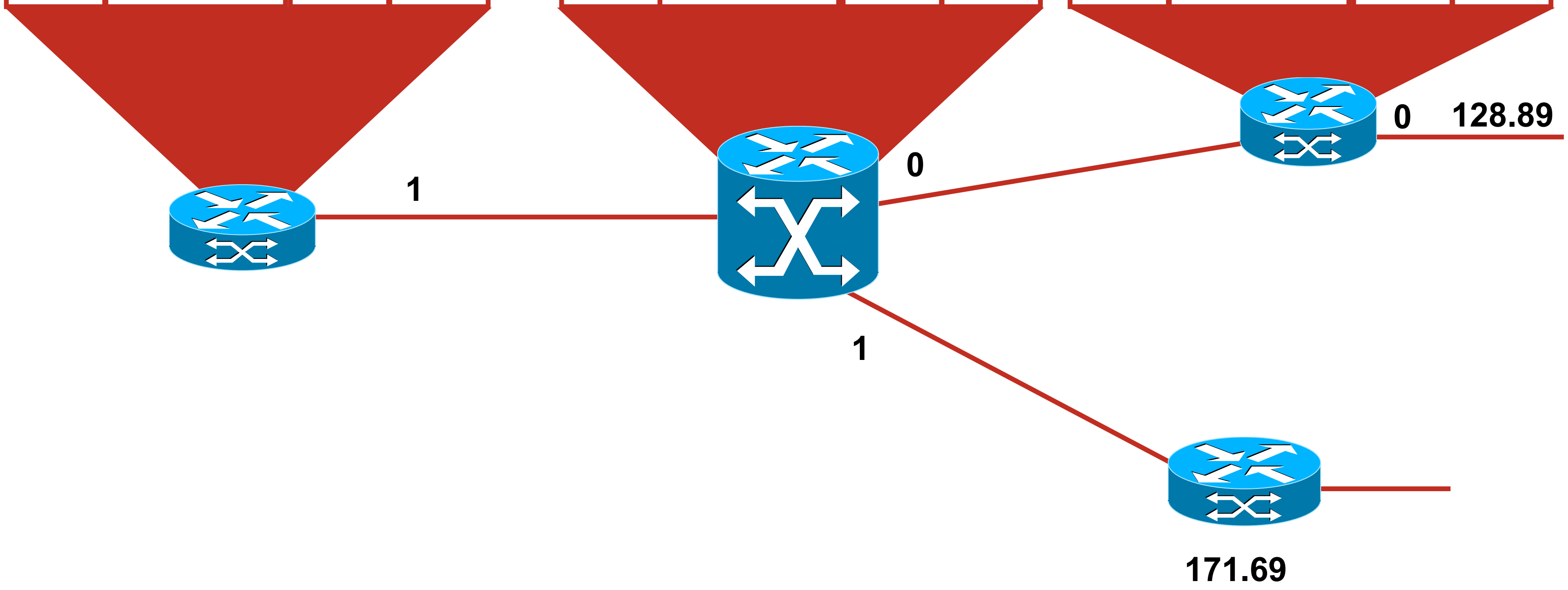


MPLS with Downstream Unsolicited Mode
Step II: Assigning Local Labels

In Labe	Address Prefix	Out I'face	Out Label
-	128.89	1	
-	171.69	1	
…	…	…	…

In Labe	Address Prefix	Out I'face	Out Label
4	128.89	0	
5	171.69	1	
…	…	…	…

In Labe	Address Prefix	Out I'face	Out Label
9	128.89	0	-
…	…	…	…

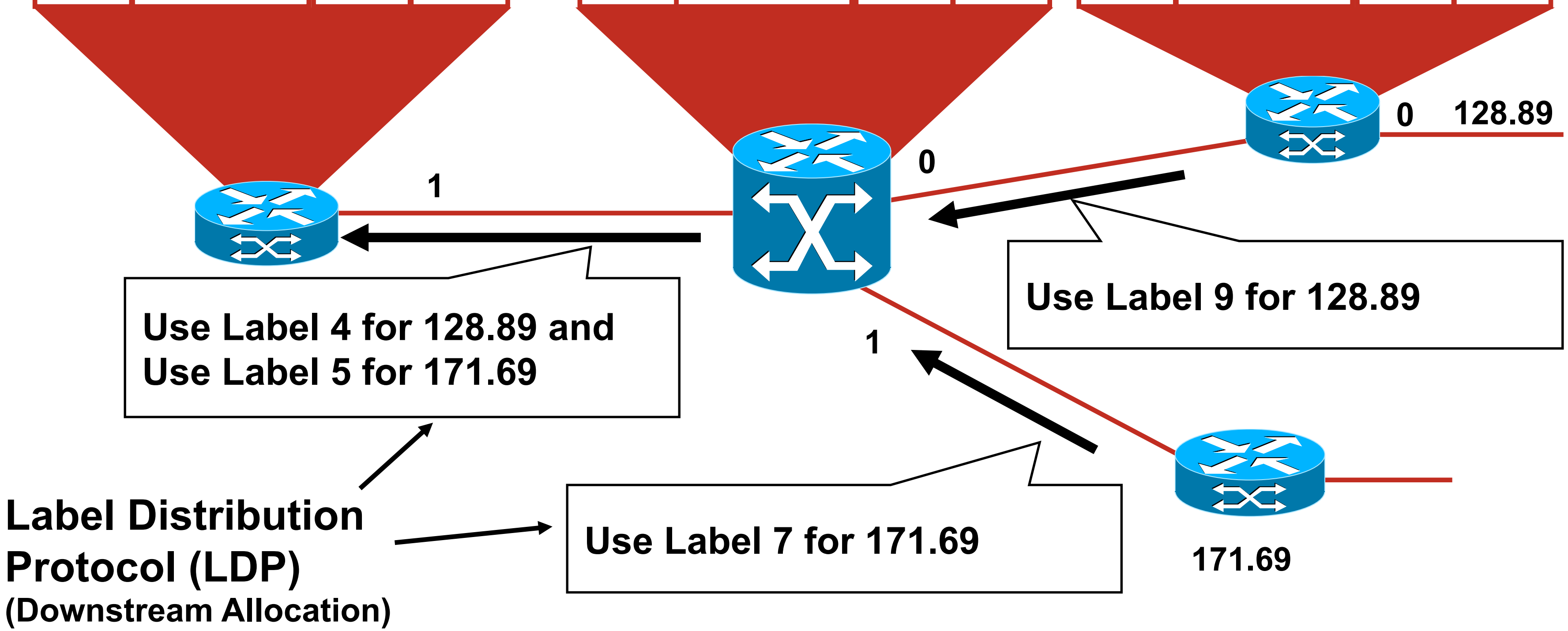


MPLS with Downstream Unsolicited Mode
Step II: Assigning Remote Labels

In Labe	Address Prefix	Out I’face	Out Label
-	128.89	1	4
-	171.69	1	5
…	…	…	…

In Labe	Address Prefix	Out I’face	Out Label
4	128.89	0	9
5	171.69	1	7
…	…	…	…

In Labe	Address Prefix	Out I’face	Out Label
9	128.89	0	-
…	…	…	…

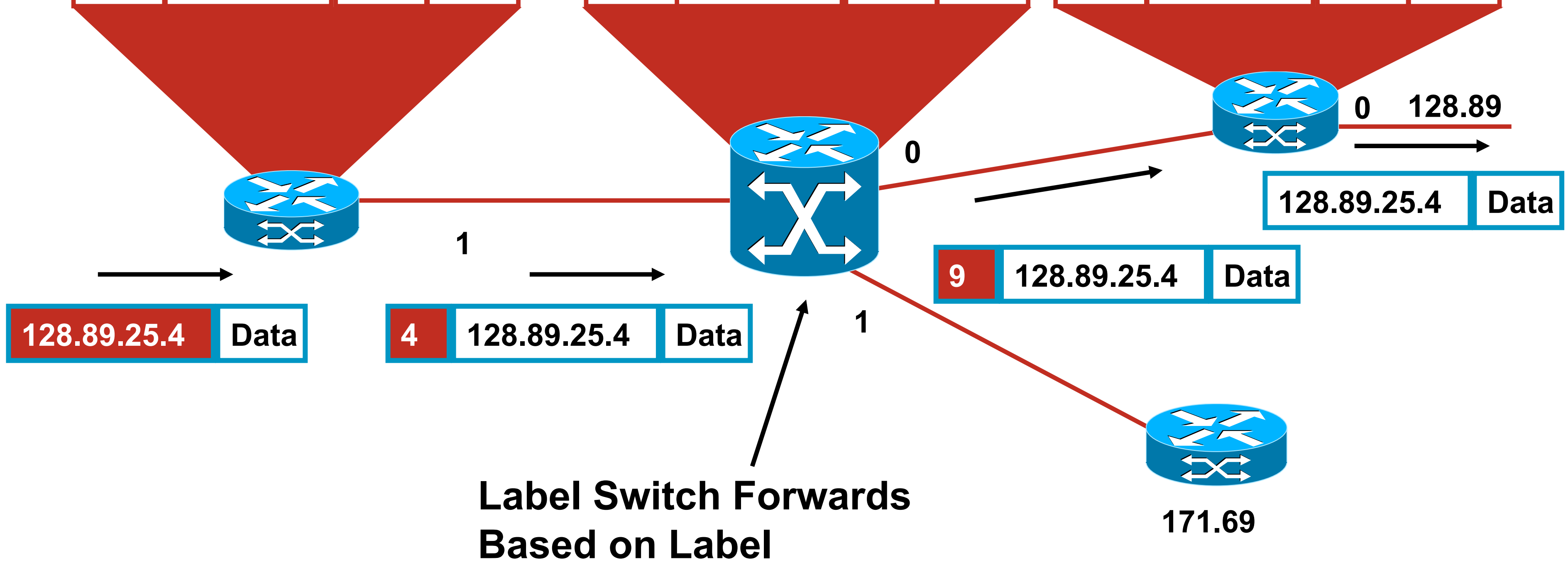


MPLS with Downstream Unsolicited Mode
Step III: Forwarding Packets

In Labe	Address Prefix	Out I'face	Out Label
-	128.89	1	4
-	171.69	1	5
...	...	...	...

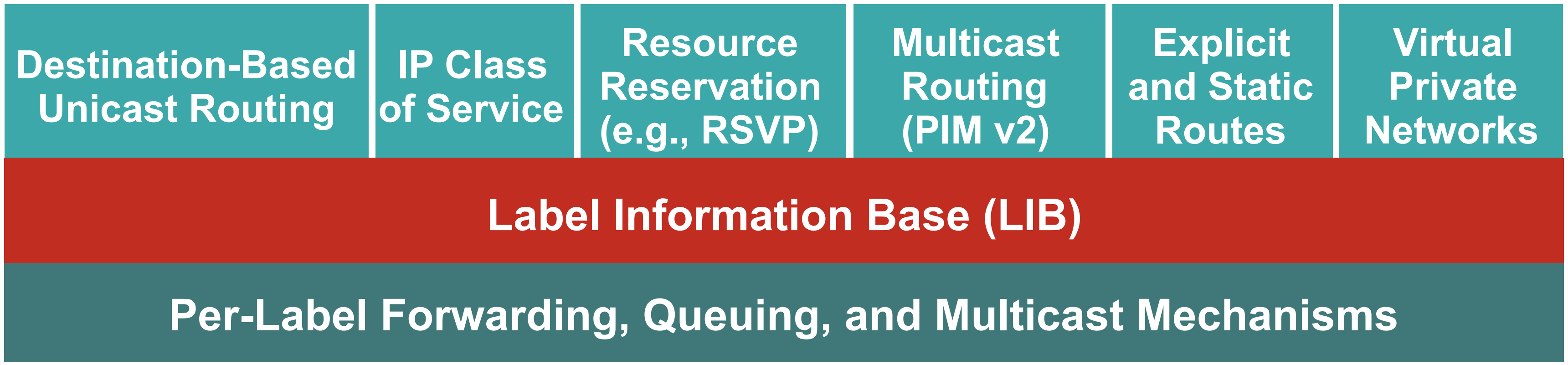
In Labe	Address Prefix	Out I'face	Out Label
4	128.89	0	9
5	171.69	1	7
...	...	...	...

In Labe	Address Prefix	Out I'face	Out Label
9	128.89	0	-
...	...	...	...

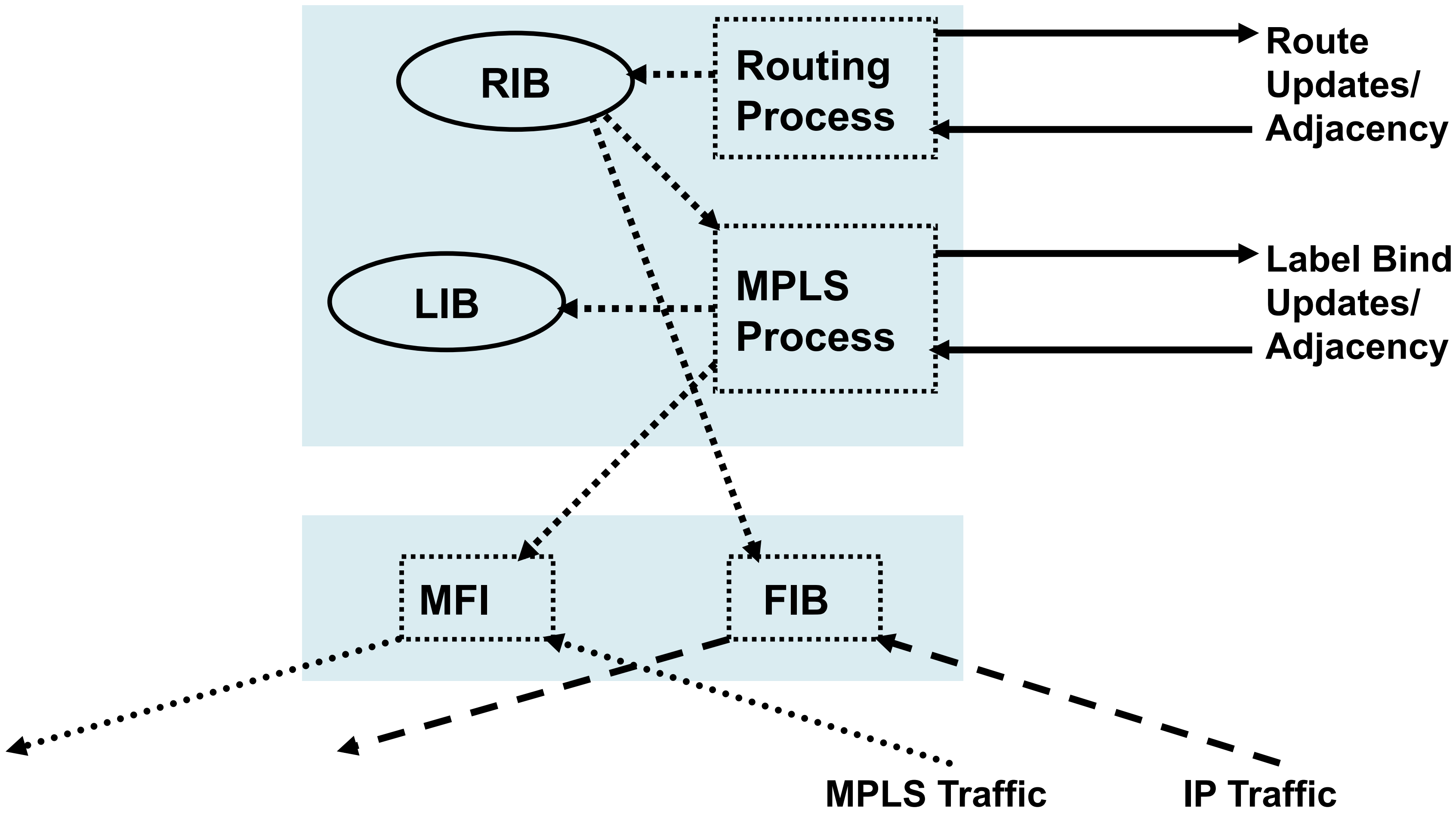


MPLS Control and Forwarding Planes

- Control plane used to distribute labels—BGP, LDP, RSVP
- Forwarding plane consists of label imposition, swapping and disposition—no matter what the control plane
- Key: there is a separation of control plane and forwarding plane
  - Basic MPLS: destination-based unicast
  - Labels divorce forwarding from IP address
  - Many additional options for assigning labels
  - Labels define destination and service

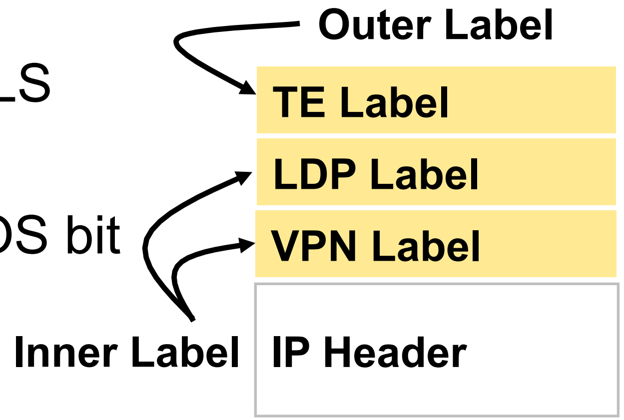


Control and Forward Plane Separation

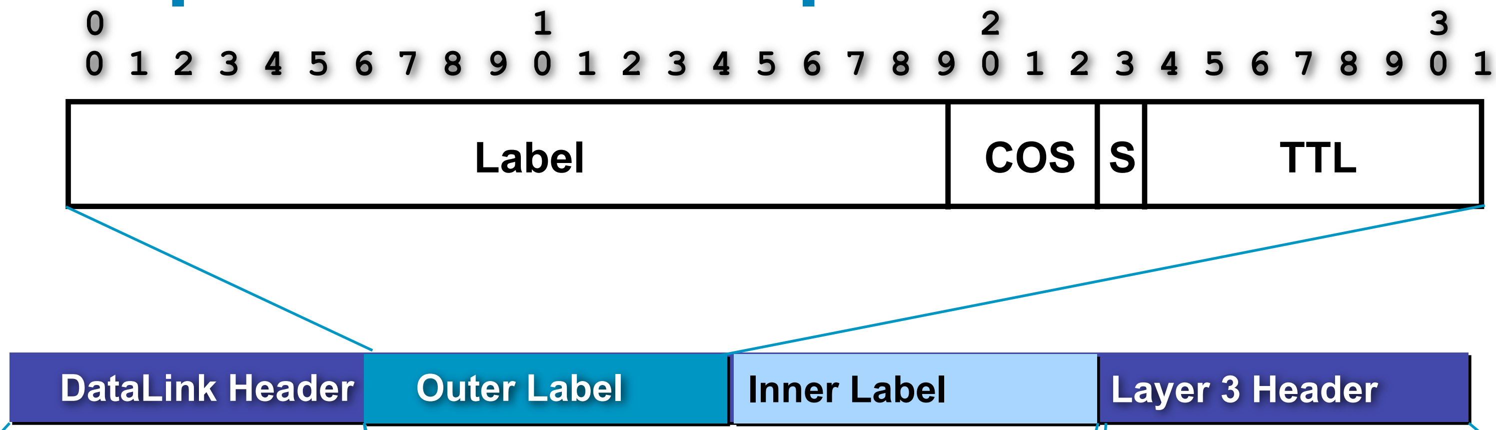


Label Stacking

- There may be more than one label in an MPLS packet
- As we know labels correspond to forwarding equivalence classes
  - Example—there can be one label for routing the packet to an egress point and another that separates a customer A packet from customer B
  - Inner labels can be used to designate services/FECs, etc.
    - e.g. VPNs, fast reroute
- Outer label used to route/switch the MPLS packets in the network
- Last label in the stack is marked with EOS bit
- Allows building services such as
  - MPLS VPNs
  - Traffic engineering and fast re-route
  - VPNs over traffic engineered core
  - Any transport over MPLS



Encapsulation Examples



Ethernet II
Destination: xx:xx:xx:xx:xx:xx
Source: yy:yy:yy:yy:yy:yy
eType: MPLS Unicast (0x8847)

WAN
HDLC, Frame Relay, ATM AAL5, etc

MultiProtocol Label Switching Header (Outer)
MPLS Label: 16
MPLS Experimental Bits: 0
MPLS Bottom Of Label Stack: 0
MPLS TTL: 255

MultiProtocol Label Switching Header (Inner)
MPLS Label: 100
MPLS Experimental Bits: 3
MPLS Bottom Of Label Stack: 1
MPLS TTL: 2

Internet Protocol
Version: 4
Header length: 20 bytes
[snip]
Time to live: 255
Protocol: ICMP (0x01)
Header checksum: 0xa3fd (correct)
Source: 10.1.1.2 (10.1.1.2)
Destination: 172.16.255.2 (172.16.255.2)

Label Stack

```
[PE1]#show ip cef vrf blue 11.2.1.3
11.2.1.3/32, version 13, epoch 0, cached adjacency to Serial1/0
0 packets, 0 bytes
  tag information set, all rewrites owned
    local tag: VPN route head
    fast tag rewrite with Se1/0, point2point, tags imposed {46 67}
  via 172.16.255.2, 0 dependencies, recursive
    next hop 172.16.1.1, Serial1/0 via 172.16.255.2/32 (Default)
    valid cached adjacency
    tag rewrite with Se1/0, point2point, tags imposed {46 67}
[PE1]#
```

46: IGP/LDP Label

67: VPN Label

2-2

MPLS VPNs

Layer 3 and Layer 2

What Is a Virtual Private Network?

- VPN is a set of sites or groups which are allowed to communicate with each other
- VPN is defined by a set of administrative policies
  - Policies established by VPN customers
  - Policies could be implemented completely by VPN service providers
- Flexible inter-site connectivity
  - Ranging from complete to partial mesh
- Sites may be either within the same or in different organizations
  - VPN can be either intranet or extranet
- Site may be in more than one VPN
  - VPNs may overlap
- Not all sites have to be connected to the same service provider
  - VPN can span multiple providers

L2 vs. L3 VPNs

Layer 2 VPNs

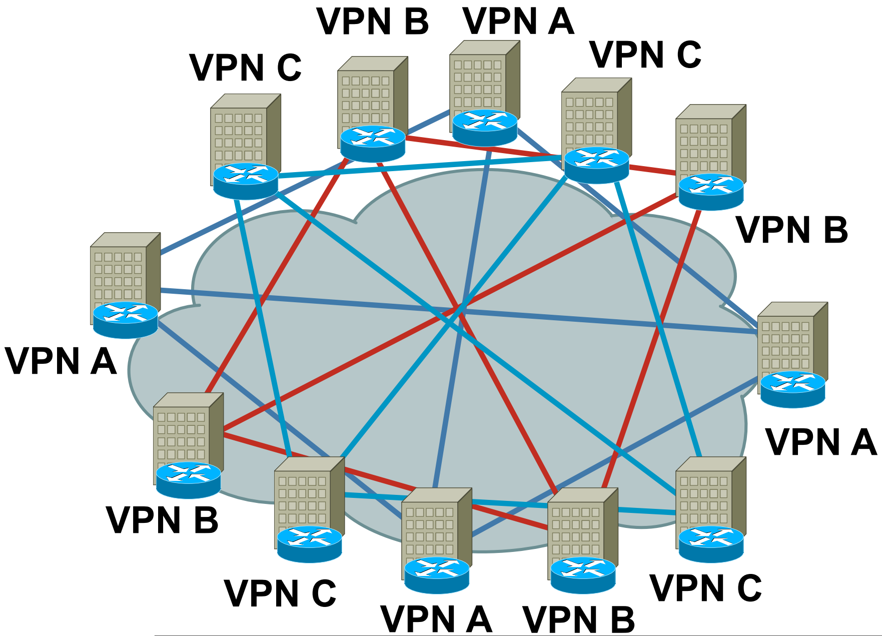
- Customer endpoints (CPE) connected via Layer 2 such as Frame Relay DLCI, ATM VC or point-to-point connection
- Provider network is not responsible for distributing site routers as routing relationship is between the customer endpoints
- Good for point to point L2 connectivity, provider will need to manually fully mesh end points if any-to-any connectivity is required

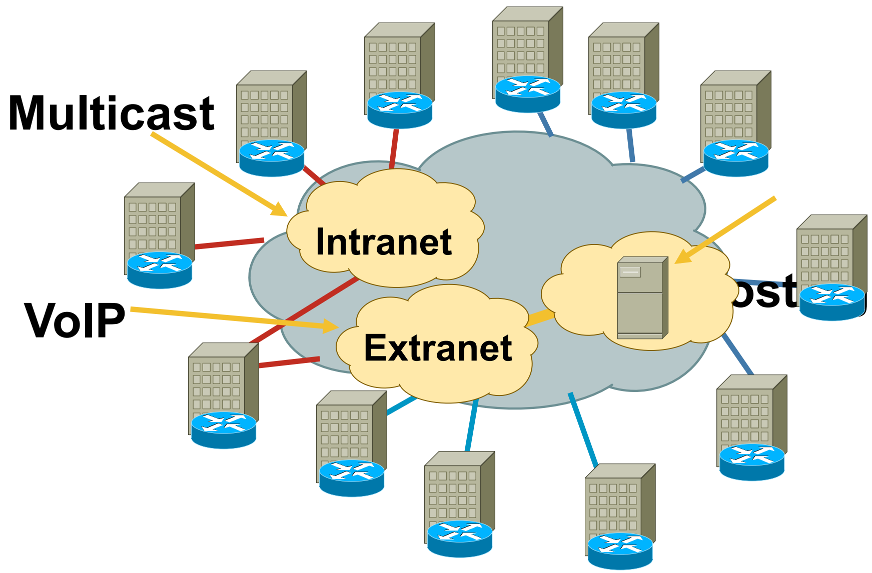
Layer 3 VPN

- Customer end points peer with providers' routers @ L3
- Provider network responsible for distributing routing information to VPN sites
- Don't have to manually fully mesh customer endpoints to support any-to-any connectivity

Layer 3 VPNs

IP L3 vs. MPLS L3 VPNs





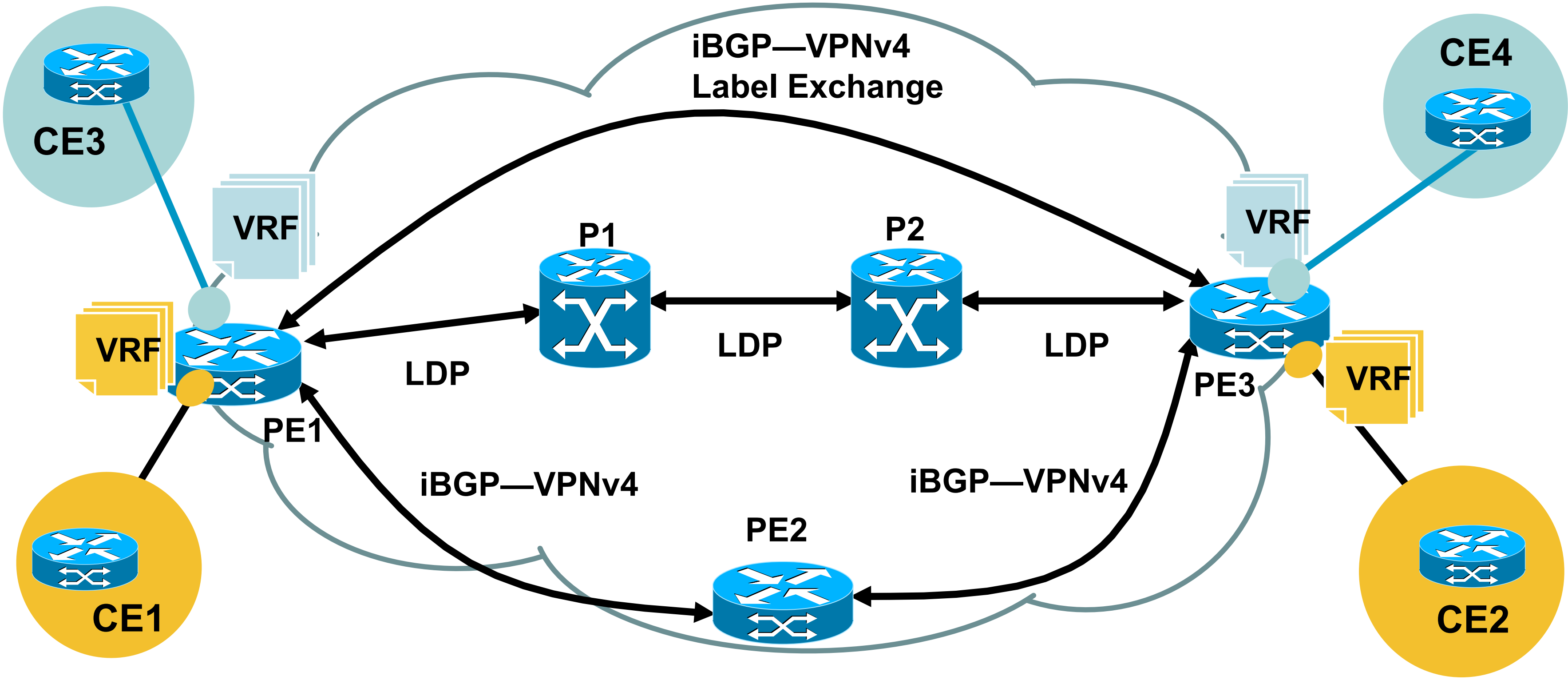
Overlay VPN

- ACLs, ATM/FR, IP tunnels, IPSec, …etc. requiring n*(n-1) peering points
- Transport dependent
- Groups endpoints, not groups
- Pushes content outside the network
- Costs scale exponentially
- NAT necessary for overlapping address space
- Limited scaling
- QoS complexity

MPLS-Based VPNs

- Point to Cloud single point of connectivity
- Transport independent
- Easy grouping of users and services
- Enables content hosting inside the network
- "Flat" cost curve
- Supports private overlapping IP addresses
- Scalable to over millions of VPNs
- Per VPN QoS

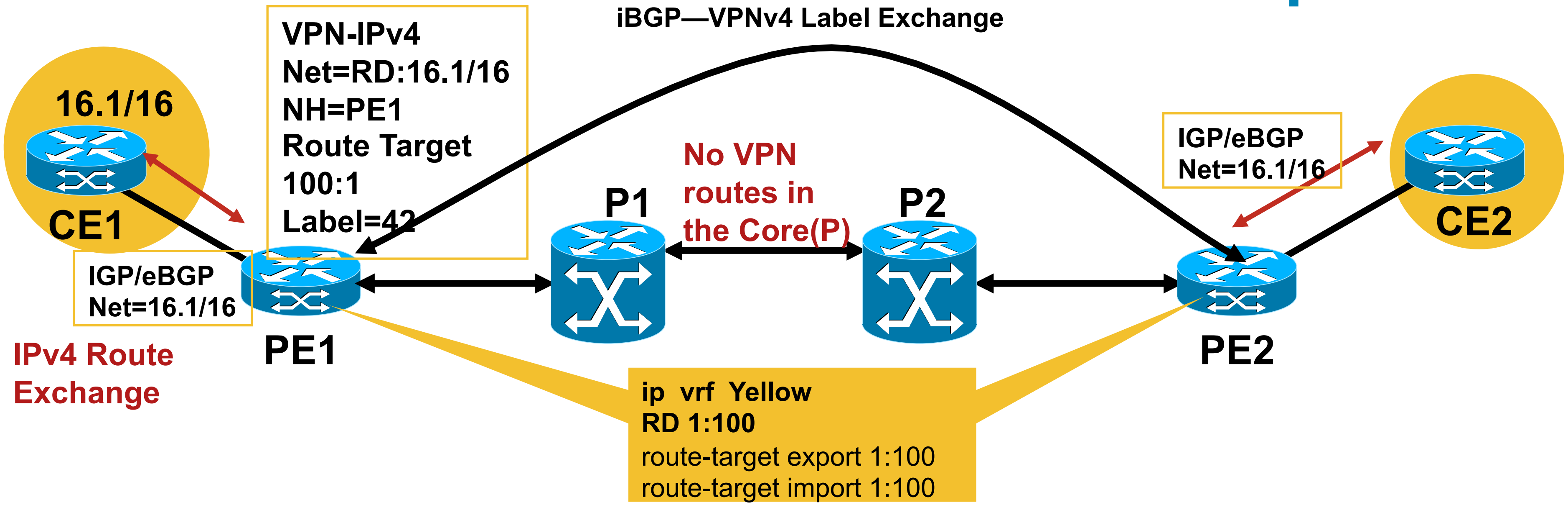
How Does It Work?
MPLS L3 VPN Control Plane Basics



1. VPN service is enabled on PEs (VRFs are created and applied to VPN site interface)
2. VPN site's CE1 connects to a VRF enabled interface on a PE1
3. VPN site routing by CE1 is distributed to MP-iBGP on PE1
4. PE1 allocates VPN label for each prefix, sets itself as a next hop and relays VPN site routes to PE3
9. PE3 distributes CE1's routes to CE2

(Similar happens from CE2 side…)

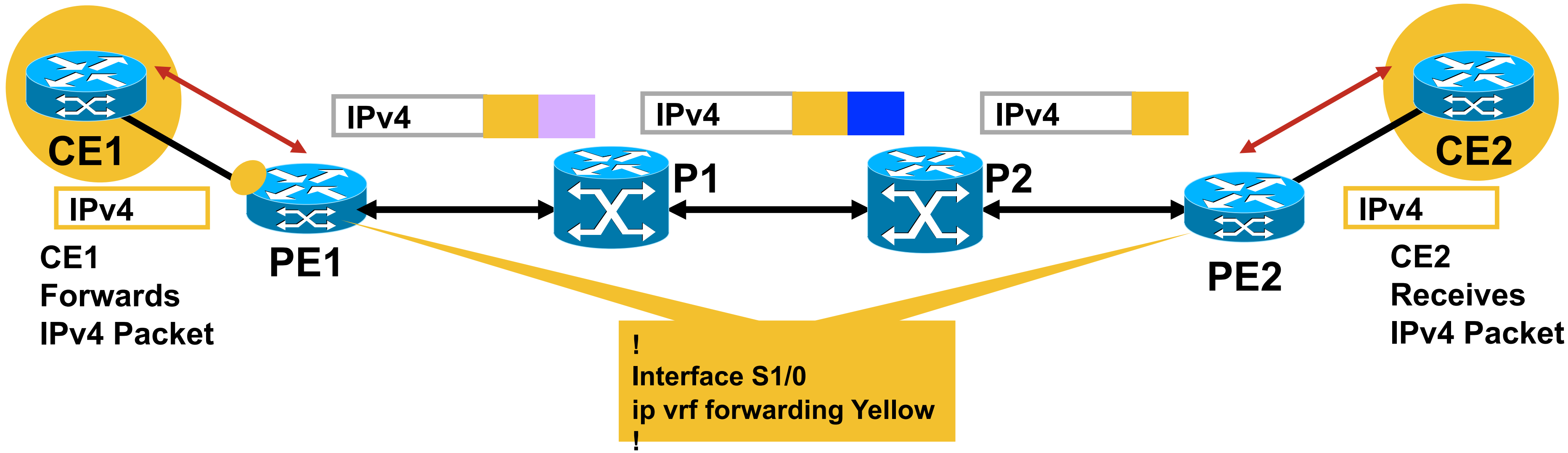
How Does It Work?
How Control Plane Information Is Separated



MPLS VPN Control Plane Components:

- Route Distinguisher: 8 byte field—unique value assigned by a provider to each VPN to make a route unique so customers don't see each other's routes
- VPNv4 address: RD+VPN IP prefix;
- Route Target: RT-8bytes field, unique value assigned by a provider to define the import/export rules for the routes from/to each VPN
- MP-BGP: facilitates the advertisement of VPNv4* prefixes + labels between MP-BGP peers
- Virtual Routing Forwarding Instance (VRF): contains VPN site routes
- Global Table: Contains core routes, Internet or routes to other services

How Does It Work?
How Data Plane Is Separated



1. PE1 imposes pre allocated label for the prefix
2. Core facing interface allocates IGP label
3. Core swap IGP labels
4. PE2 strips off VPN label and forwards the packet to CE2 as an IP packet

MPLS Security (1)
Comparison with ATM/FR

- MPLS VPN security is comparable to that provided by FR/ATM-based VPNs without providing data encryption
- Customer may still use IPSec-based mechanisms e.g., CE-CE IPSec-based encryption

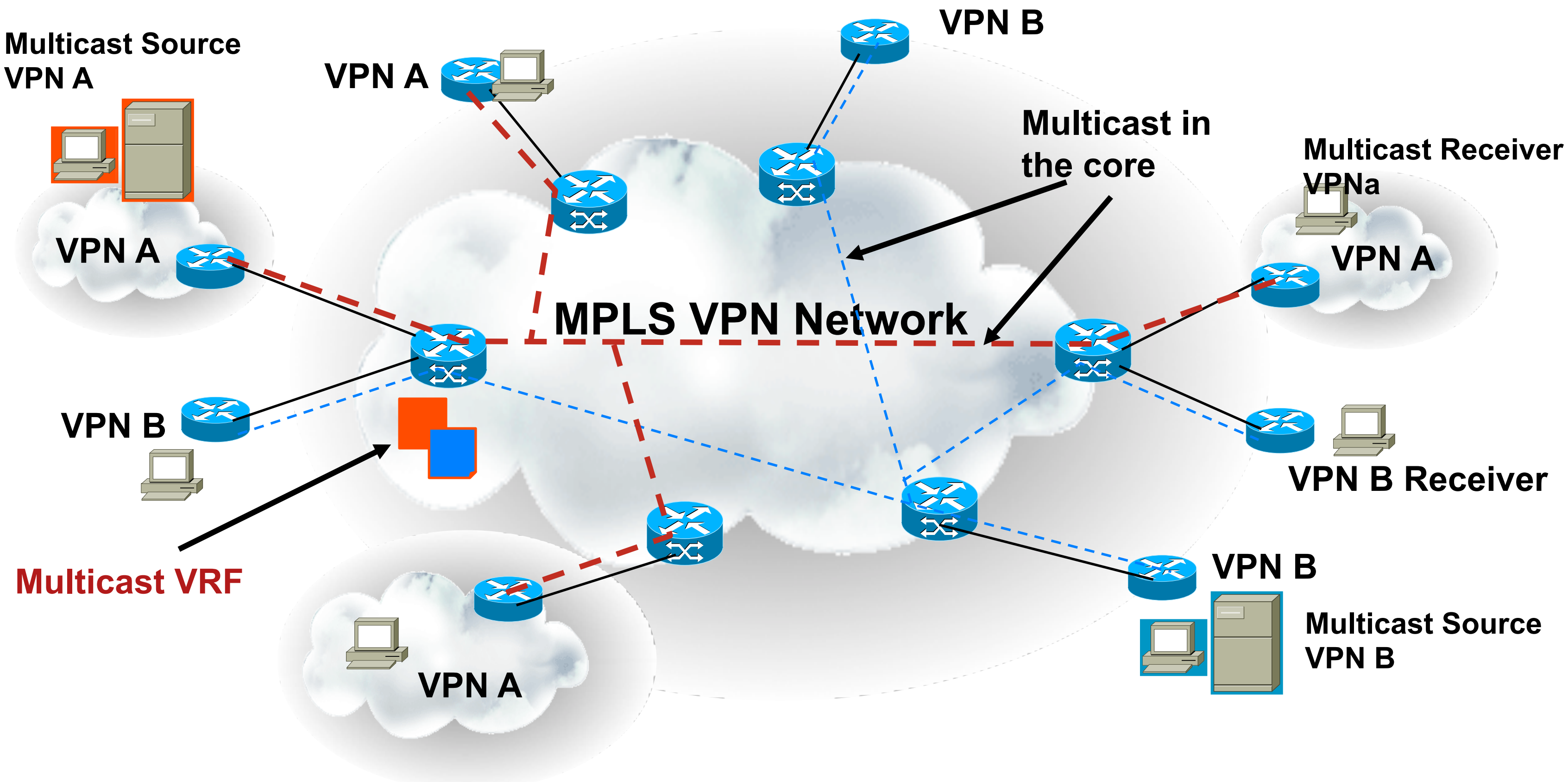
	ATM/FR	MPLS
Address Space	Yes	Yes
Routing Separation	Yes	Yes
Resistance to Attacks	Yes	Yes
Resistance to Label Spoofing	Yes	Yes

“CISCO MPLS-BASED VPNS: EQUIVALENT TO THE SECURITY OF FRAME RELAY AND ATM”

MIERCOM STUDY

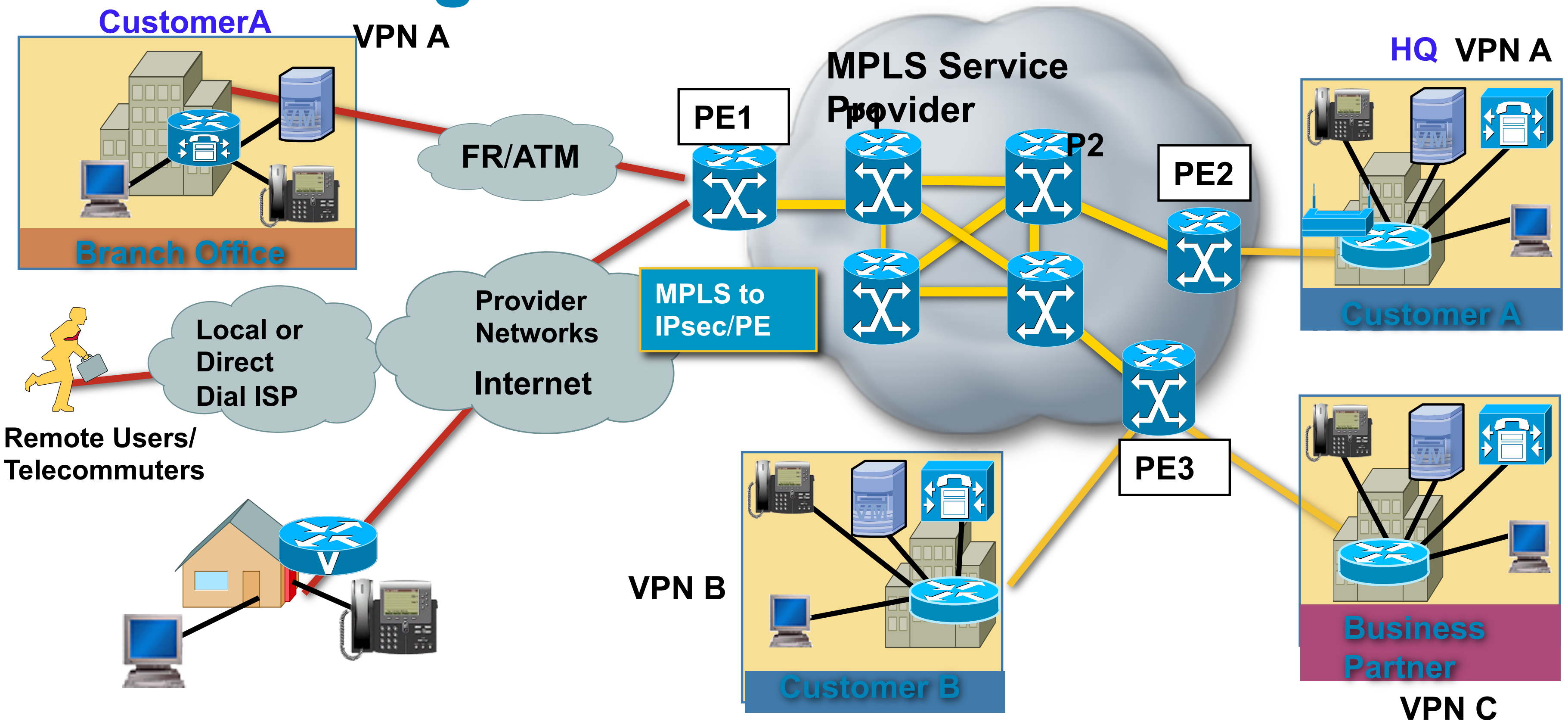
Miercom

MPLS VPN Services (2): Multicast VPNs



- Criticality of more than selling connectivity
- Run multicast within an MPLS VPN
- native multicast deployment in the core
- Simplified CE provisioning
- Highly Efficient – Multicast trees built dynamically in the core as needed

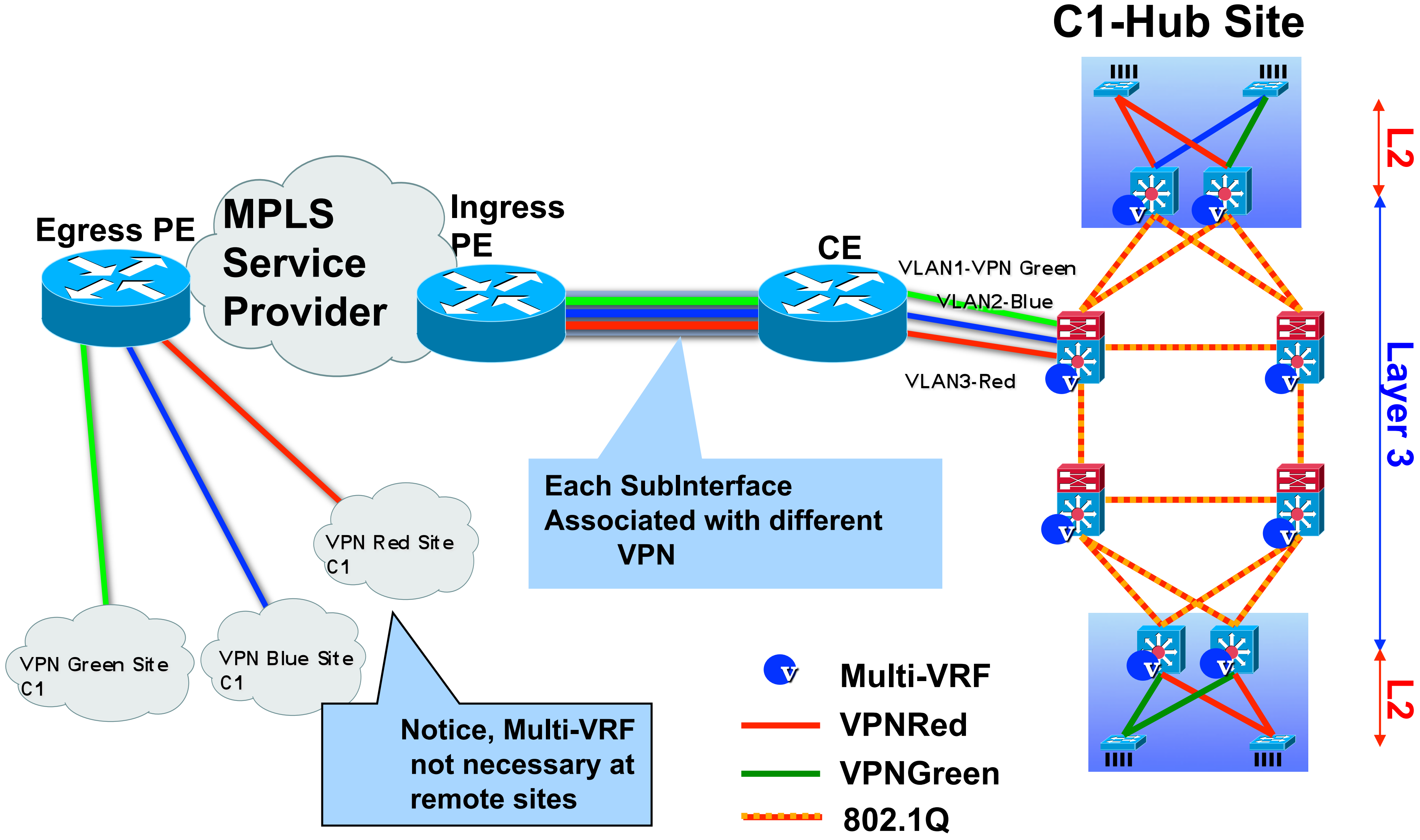
Deployment Example I: Service Provider Providing MPLS Services to Subscribers



Services Covering MAN and WAN areas:

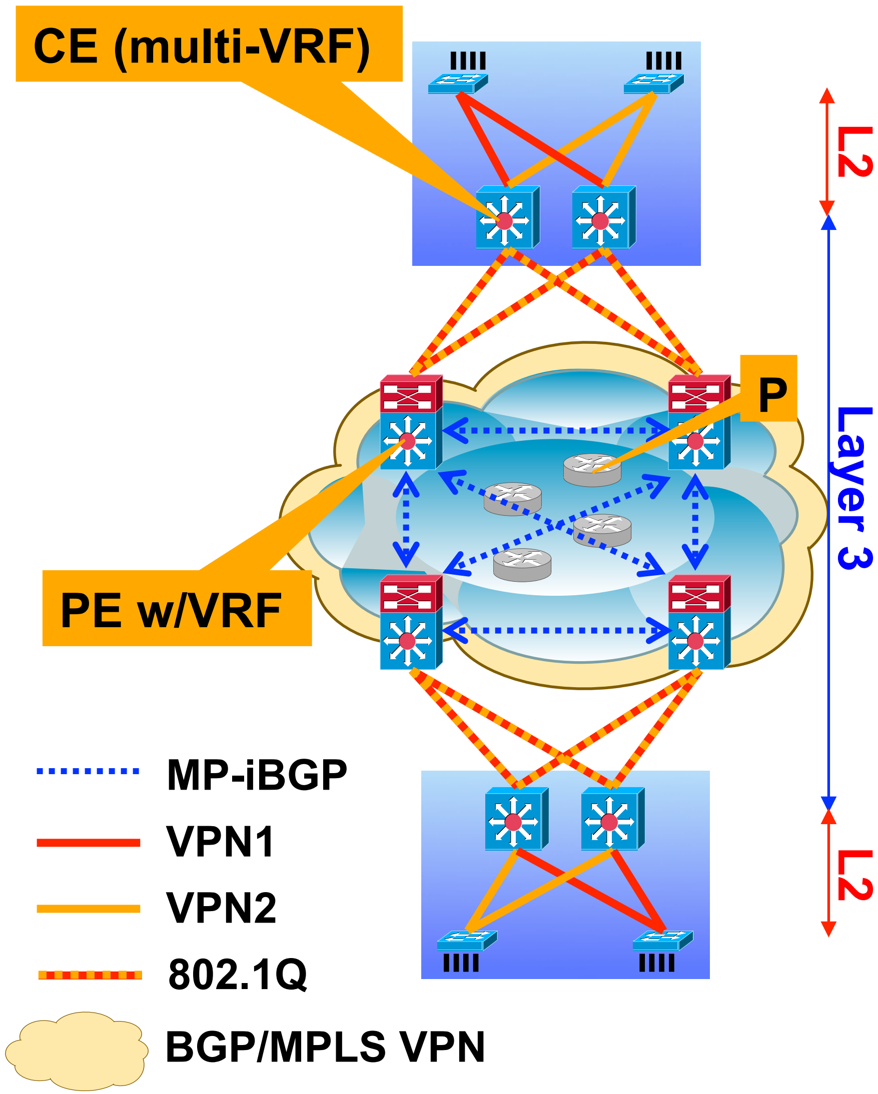
Intranet and Extranet L3 VPNs, Multicast VPNs, Internet VPN, Encryption & Firewall Services, Remote Access to MPLS Services….etc.

Deployment Example II: MPLS VPN Subscriber with VPNs in Campus That Spans Across SP's MPLS VPN Network

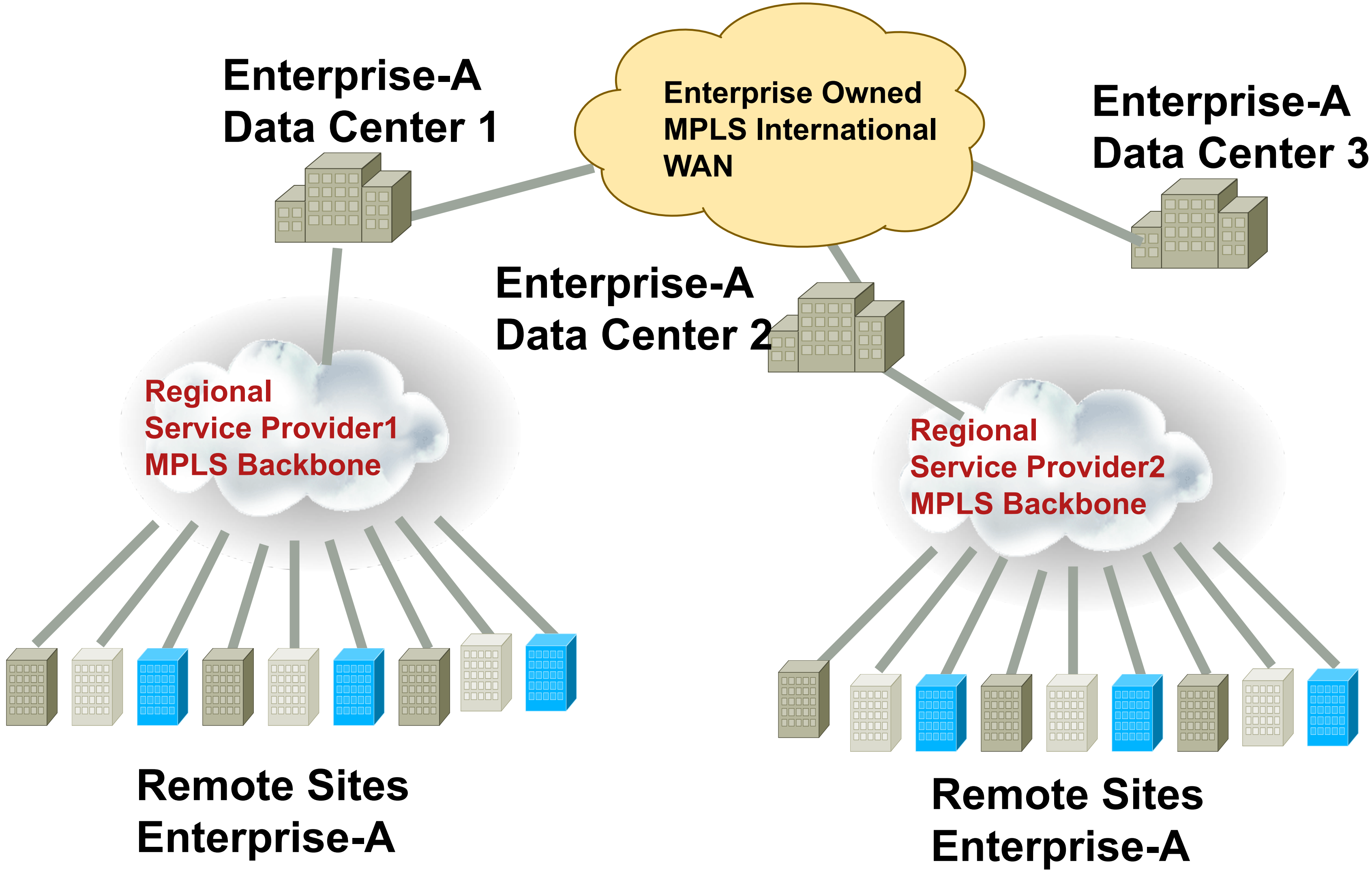


Deployment Example III: Full MPLS VPN in Enterprise Campus/LAN

- L2 Access
- Multi-VRF-CE at Distribution
- BGP/MPLS VPNs in core only
- Multi-VRF between core and distribution



Deployment Example IV: Full MPLS VPN in Enterprise WAN + Subscribed MPLS VPNs



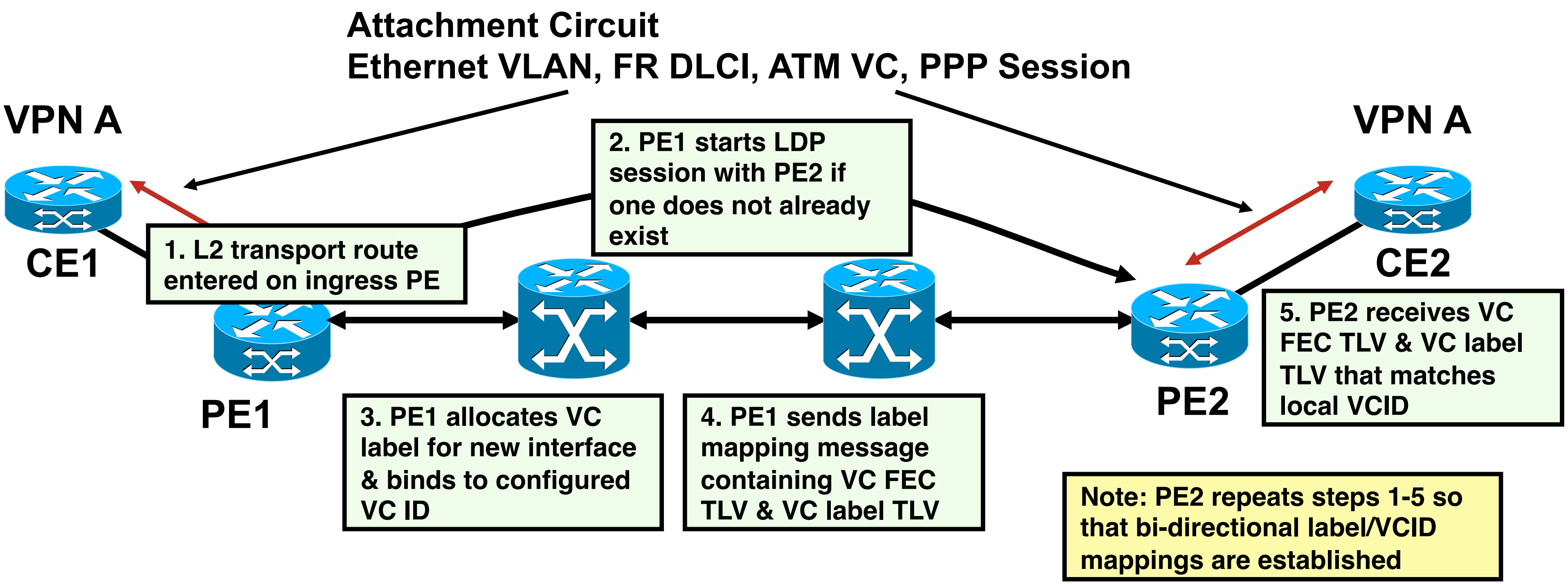
Layer 2 VPNs

Layer 2 VPNs

Similar to L3 VPN

- Designate a label for the circuit
- Exchange that label information with the egress PE
- Encapsulate the incoming traffic (Layer 2 frames)
- Apply label (learned through the exchange)
- Forward the MPLS packet (l2 encapsulated to destination on an LSP)
- At the egress
  - Look up the L2 label
  - Forward the packet onto the L2 attachment circuit

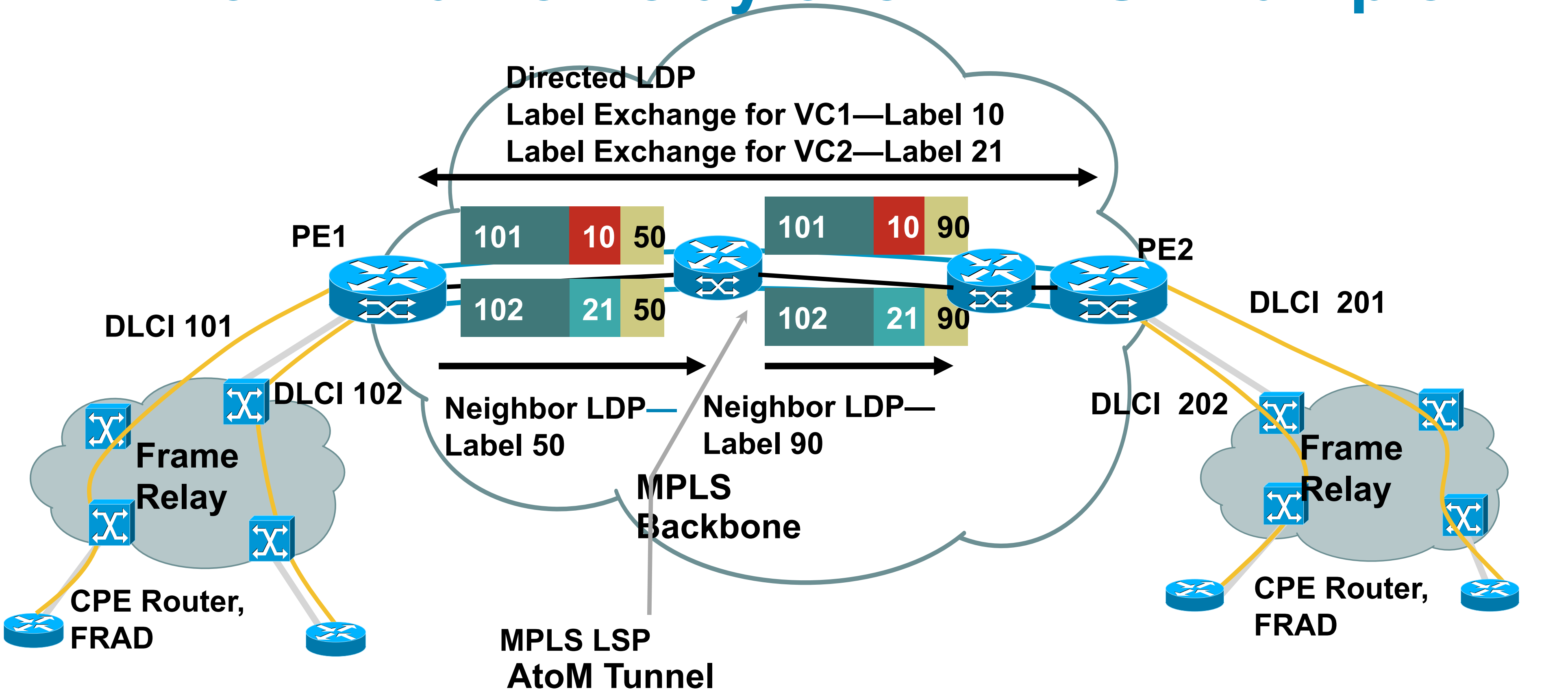
Any Transport over MPLS Architecture



Draft Martini compliant (point-to-point):

- draft-martini-l2circuit-trans-mpls
  - describes label distribution mechanisms for VC labels
- draft-martini-l2circuit-encap-mpls
  - describes emulated VC encapsulation mechanisms

AToM: Frame Relay over MPLS Example



PE1 Config:

```
connect FR1 serial5/0 101 l2transport
  mpls l2transport route 2.2.2.2    1
```

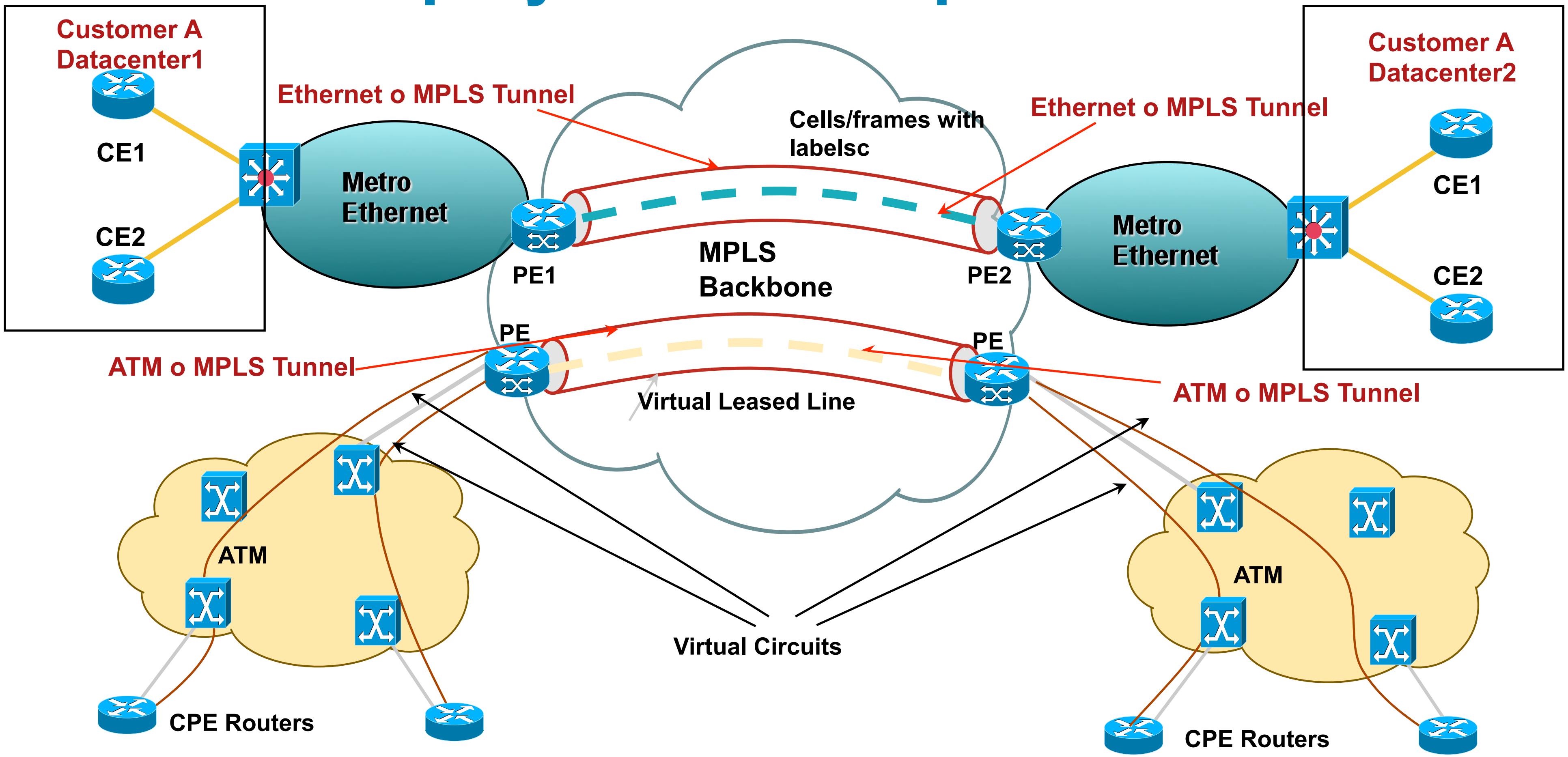
PE2 Config:

```
connect FR1 serial5/0 201 l2transport
  mpls l2transport route 1.1.1.1    1
```

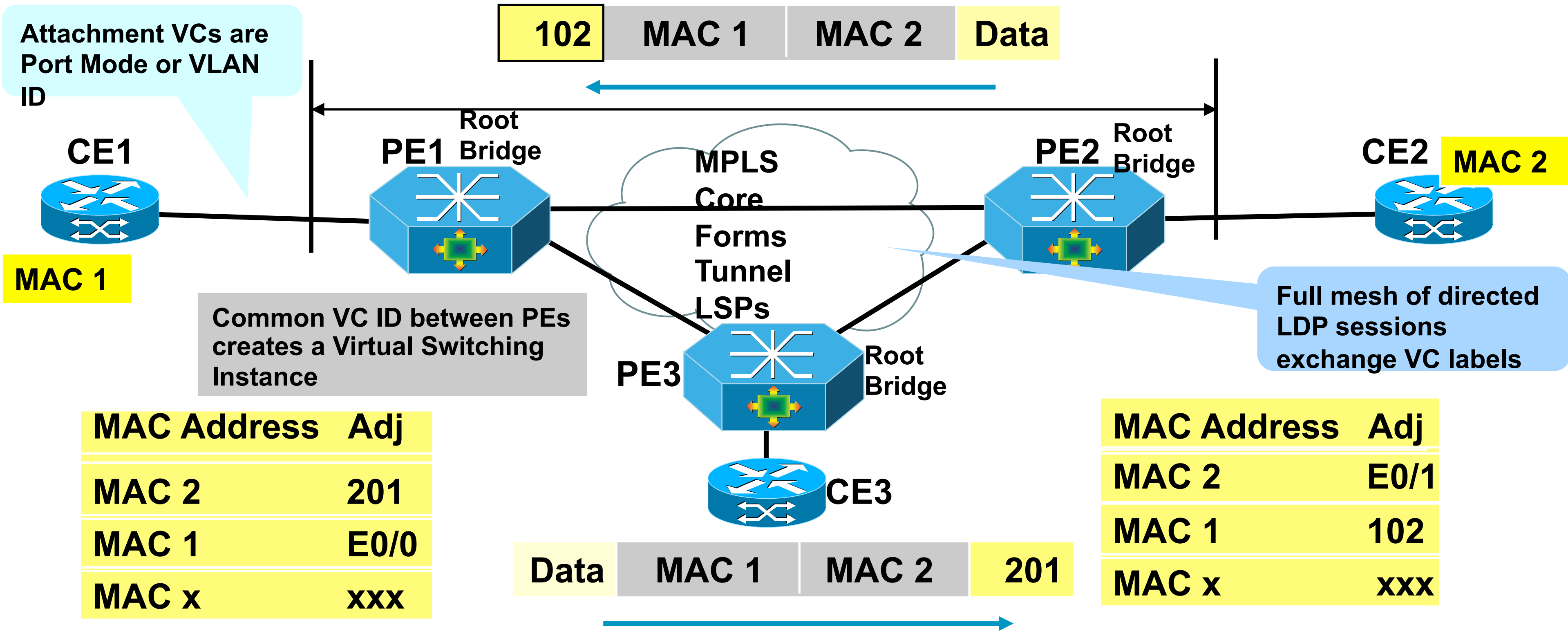
VC1—Connects DLCI 101 to DLCI 201

VC2—Connects DLCI 102 to DLCI 202

AToM Deployment Example

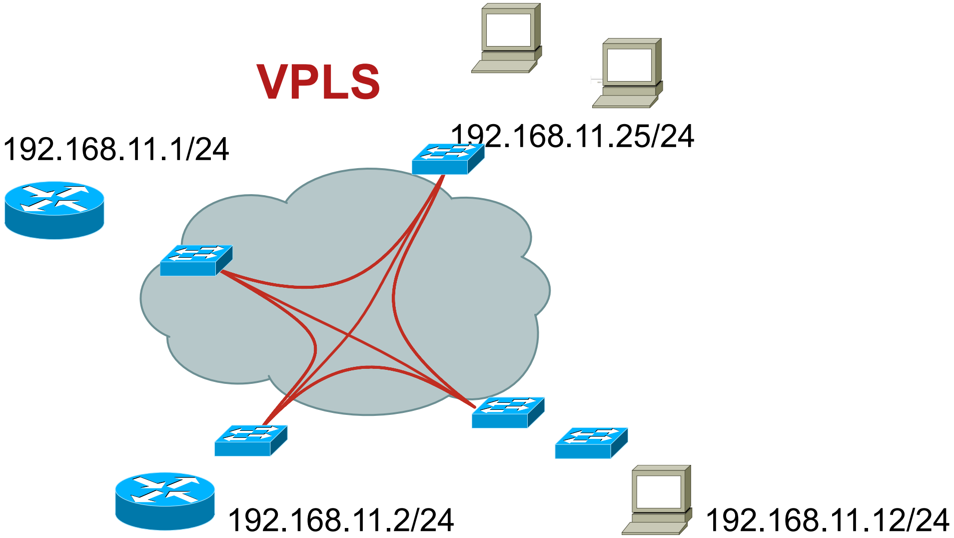


Virtual Private LAN Services (VPLS)



- VPLS defines an architecture that delivers Ethernet Multipoint Services (EMS) over an MPLS network
- VPLS operation emulates an IEEE Ethernet bridge. Two VPLS drafts in existence
  - Draft-ietf-l2vpn-vpls-ldp-01 ← Cisco's implementation
  - Draft-ietf-l2vpn-vpls-bgp-01

VPLS and H-VPLS

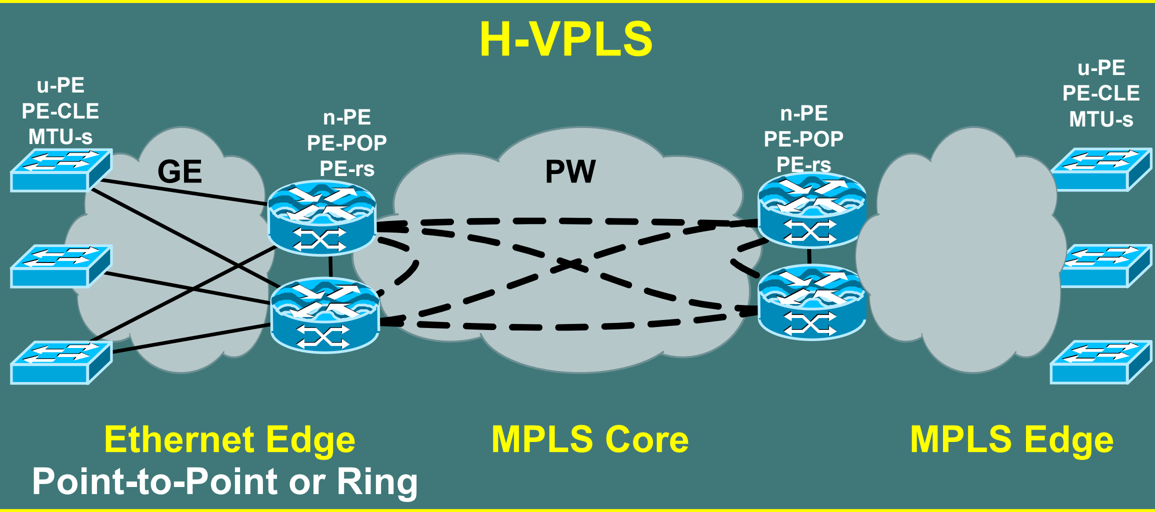


VPLS Direct Attachment

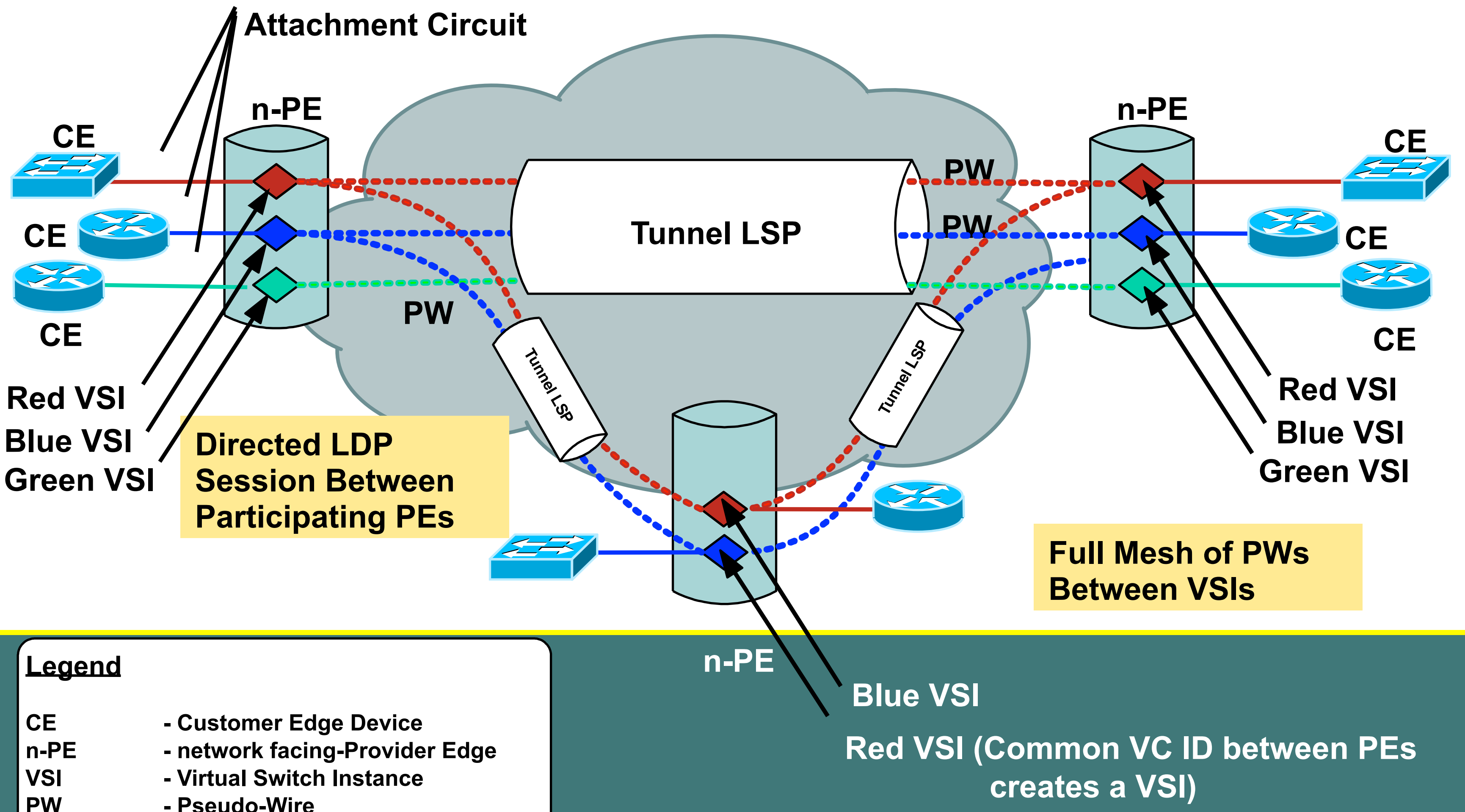
- Single flat hierarchy
- MPLS to the edge

H-VPLS

- Two tier hierarchy
- MPLS or Ethernet edge
- MPLS core



VPLS Components/Deployment Example



Legend

CE	- Customer Edge Device
n-PE	- network facing-Provider Edge
VSI	- Virtual Switch Instance
PW	- Pseudo-Wire
Tunnel LSP	- Tunnel Label Switch Path that provides PW transport

MPLS Traffic Engineering

Why Traffic Engineering?

- Congestion in the network due to changing traffic patterns
  - Election news, online trading, major sports events
- Better utilization of available bandwidth
  - Route on the non-shortest path
- Route around failed links/nodes
  - Fast rerouting around failures, transparently to users
  - Like SONET APS (Automatic Protection Switching)
- Build new services—virtual leased line services
  - VoIP toll-bypass applications, point-to-point bandwidth guarantees
- Capacity planning
  - TE improves aggregate availability of the network

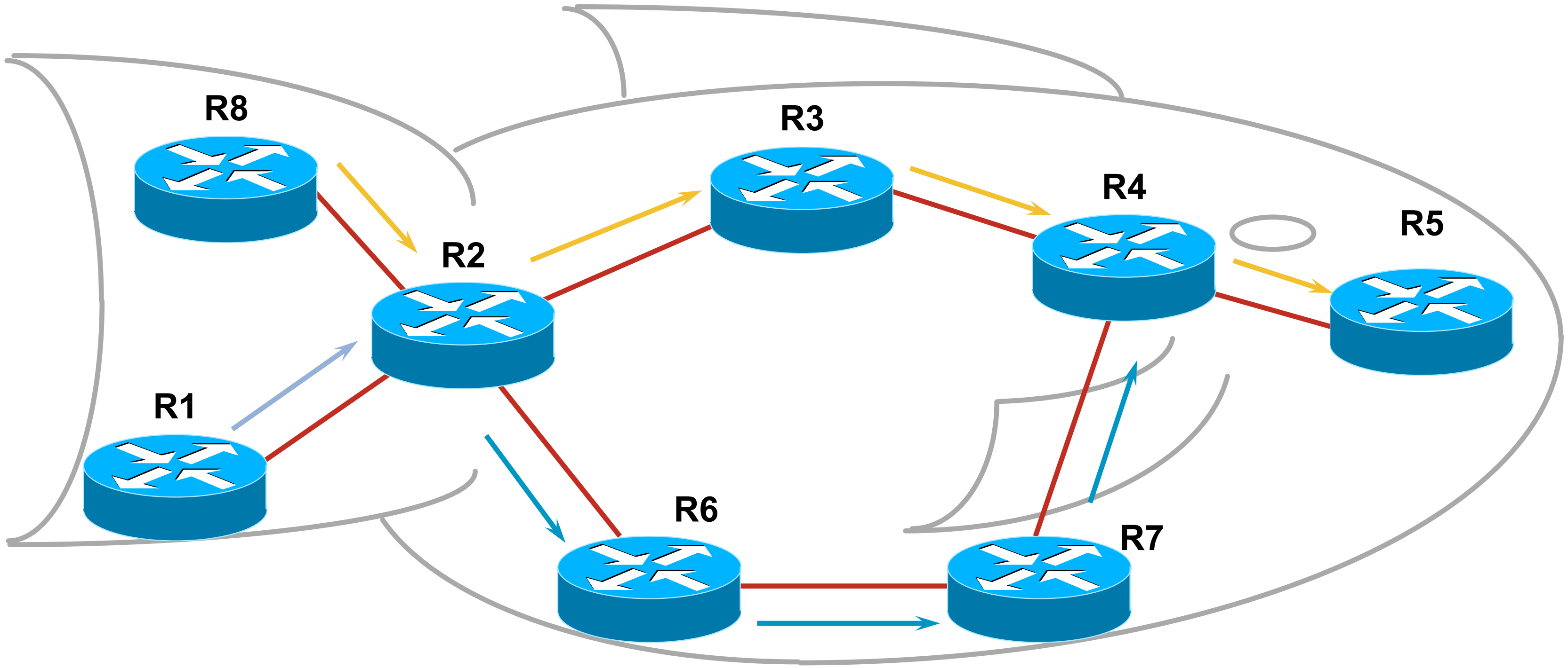
What Is MPLS Traffic Engineering?

- Process of routing data traffic in order to balance the traffic load on the various links, routers, and switches in the network
- Key in most networks where multiple parallel or alternate paths are available

Benefits of TE over Policy Routing

- Policy routing
  - Hop-by-hop decision making
  - No accounting of bandwidth
- Traffic engineering
  - Headend-based
  - Accounts for available link bandwidth
  - Admission control

IP Routing and the Fish



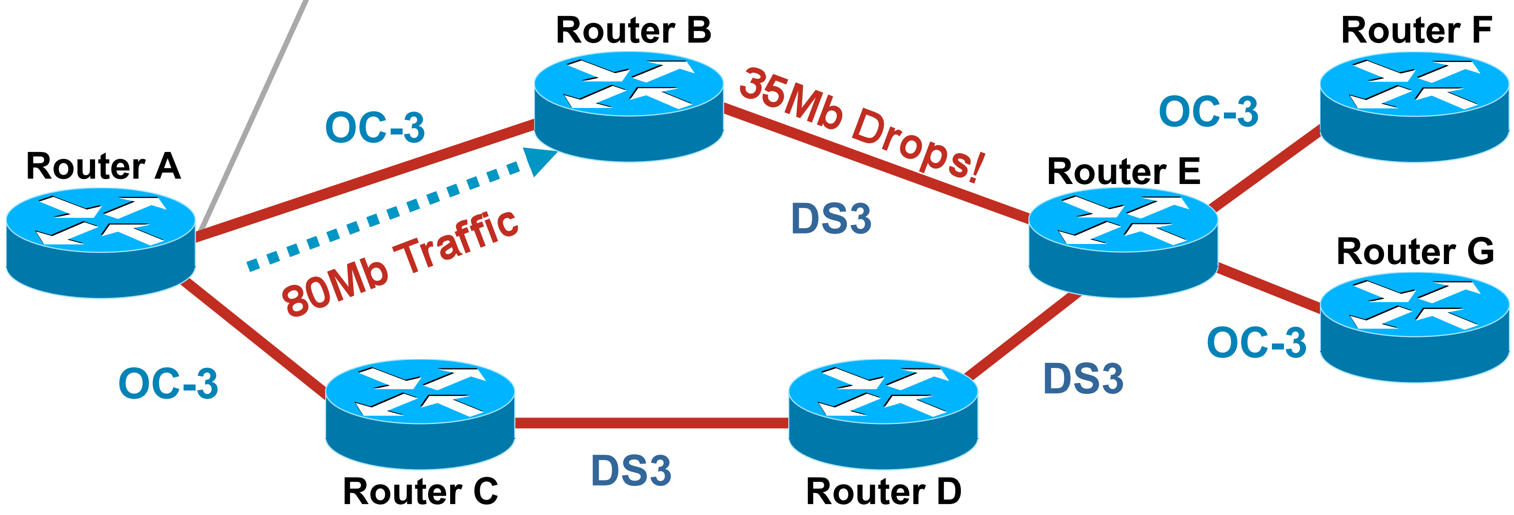
IP (Mostly) Uses Destination-Based Least-Cost Routing
Flows from R8 and R1 Merge at R2 and Become Indistinguishable
From R2, Traffic to R3, R4, R5 Use Upper Route

Alternate Path Under-Utilized

The Problem with Shortest-Path

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	B	30

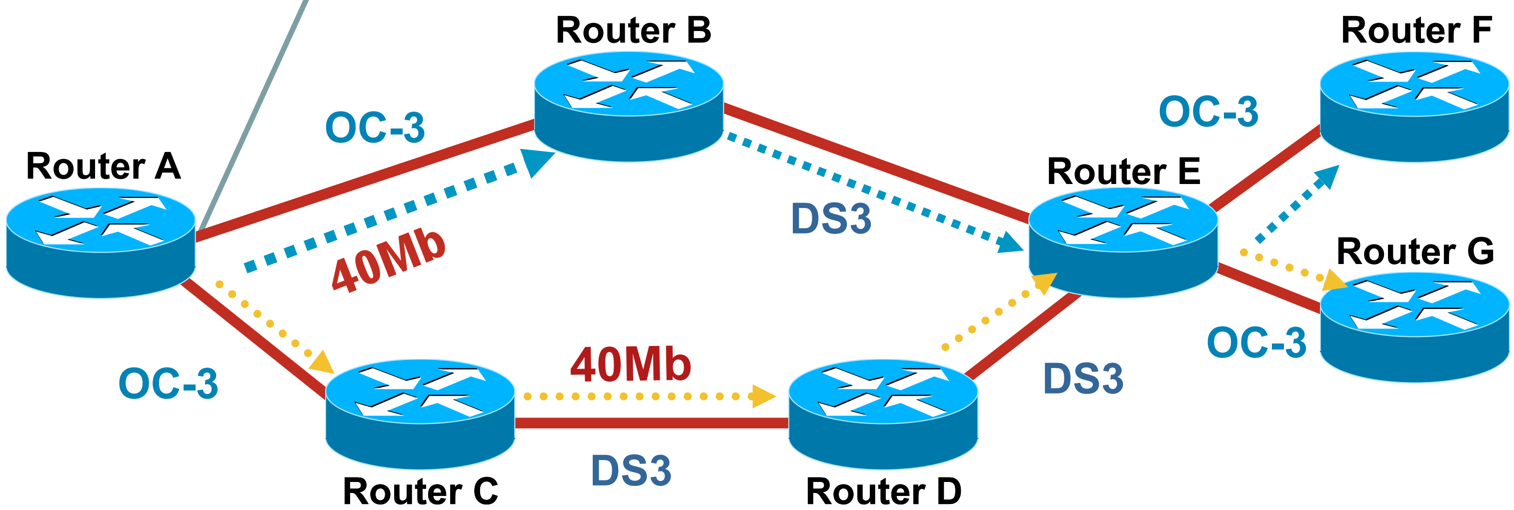
- Some links are DS3, some are OC-3
- Router A has 40mb of traffic for Router F, 40mb of traffic for Router G
- Massive (44%) packet loss at Router B→Router E!
  - Changing to A->C->D->E won't help



How MPLS TE Solves the Problem

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	Tunnel 0	30
G	Tunnel 1	30

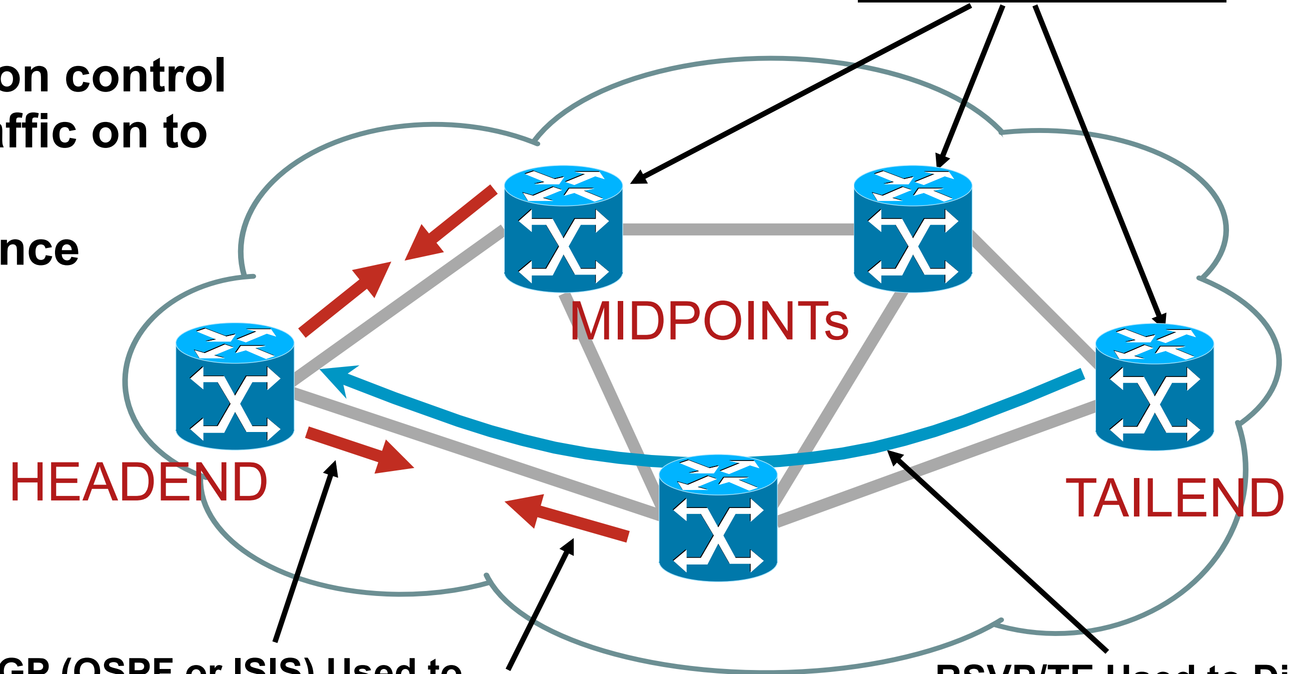
- Router A sees all links
- Router A computes paths on properties other than just shortest cost
- No link oversubscribed!



TE Fundamentals: "Building Blocks"

1. Information Distribution
2. Path selection/calculation
3. Path setup
4. Trunk admission control
5. Forwarding traffic on to tunnel
6. Path maintenance

Path Calculation—Uses IGP Advertisements to Compute <u>"Constrained" Paths</u>



IGP (OSPF or ISIS) Used to Flood Bandwidth Information Between Routers

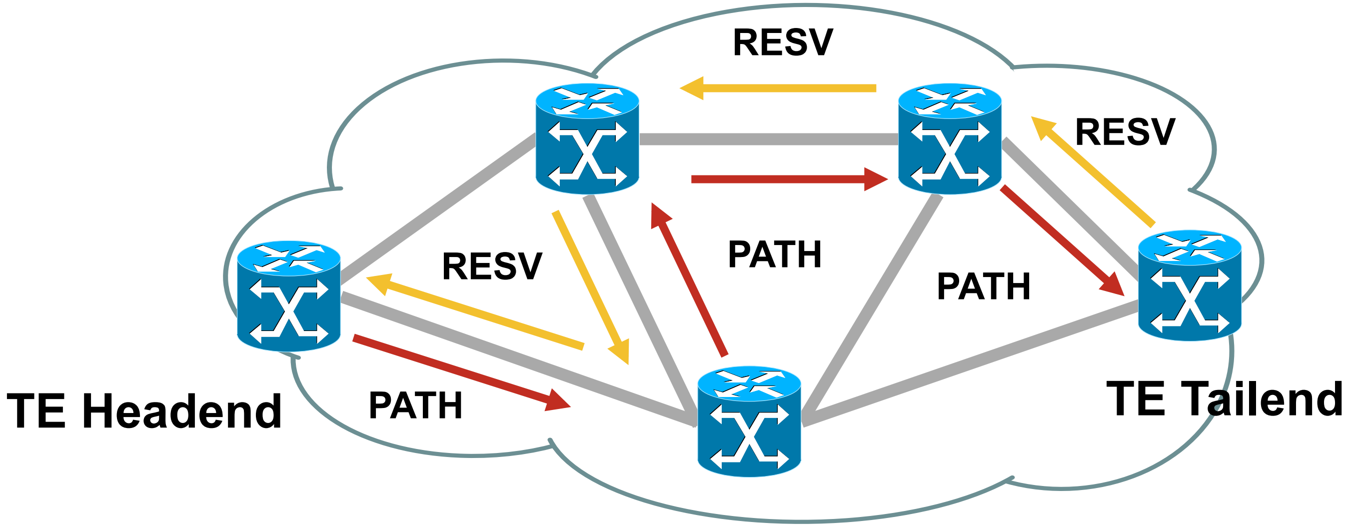
RSVP/TE Used to Distribute Labels, Provide CAC, Failure Notification, Etc.

Upstream **Unidirectional Tunnel** Downstream

Information Distribution

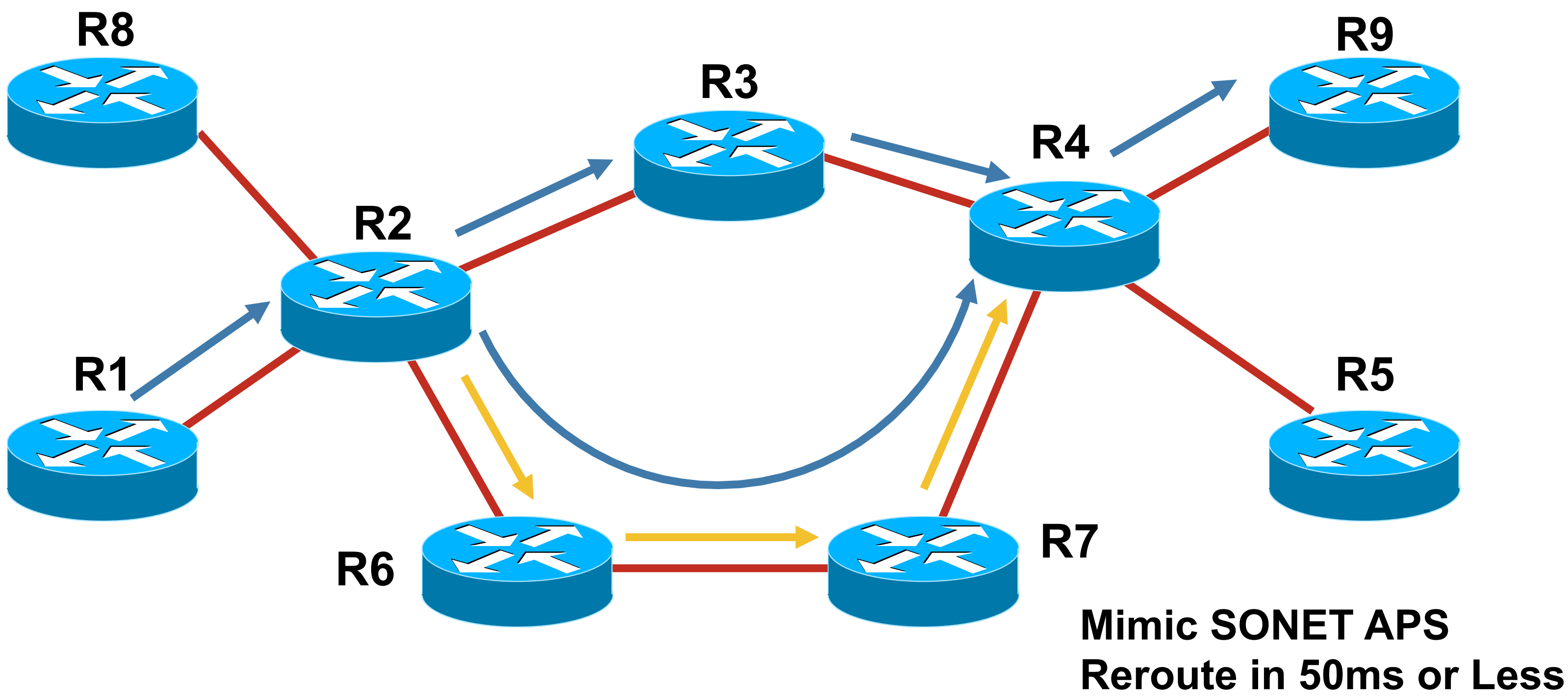
- You need a link-state protocol as your IGP
  - IS-IS or OSPF
- Link-state requirement is only for MPLS-TE!
  - Not a requirement for VPNs, etc.!
- Why do I need a link-state protocol?
  - To make sure info gets flooded
  - To build a picture of the entire network
- Information flooded includes link, bandwidth, attributes, etc.

Path Setup Example



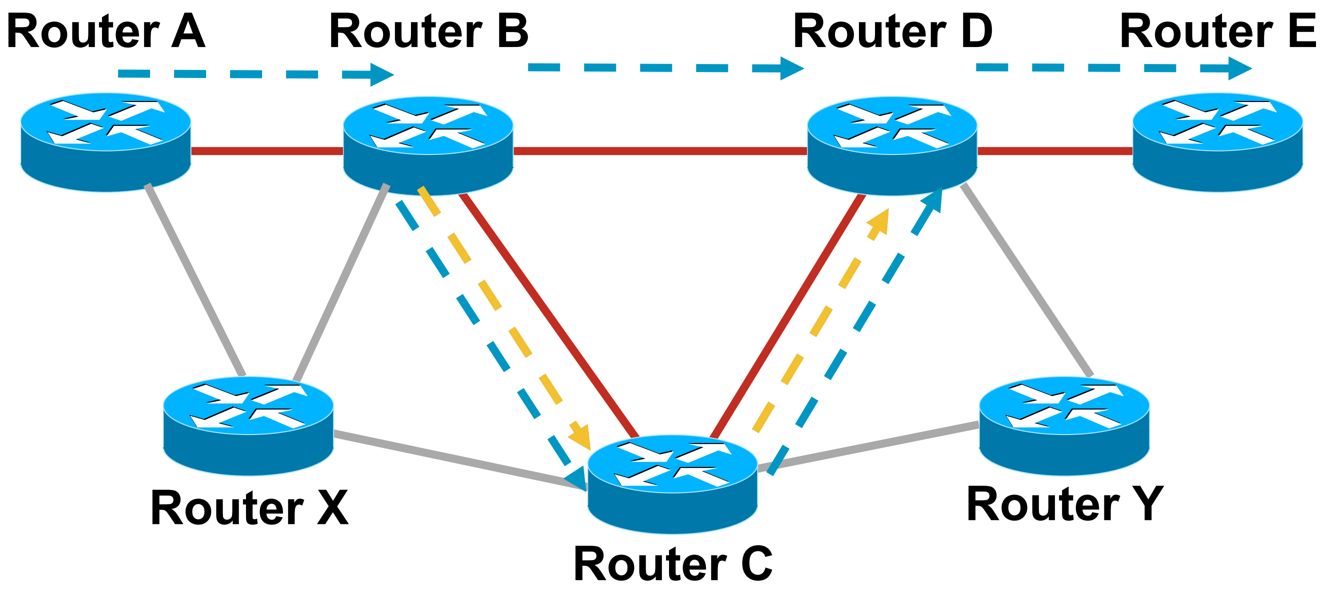
- PATH messages are sent with requested bandwidth (&label)
- RESV messages are sent with label bindings for the TE tunnel
- Tunnels can be explicitly routed
- Admission control at each hop to see if the bandwidth requirement can be met
- Packets are mapped to the tunnel via
  - Static routed
  - Autoroute
  - Policy route
- Packets follow the tunnel—LSP

Applications of MPLS TE: MPLS Fast Reroute

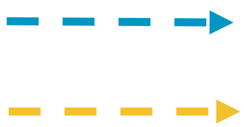


- Multiple hops can be by-passed; R2 swaps the label which R4 expects before pushing the label for R6
- R2 locally patches traffic onto the link with R6

Link Protection

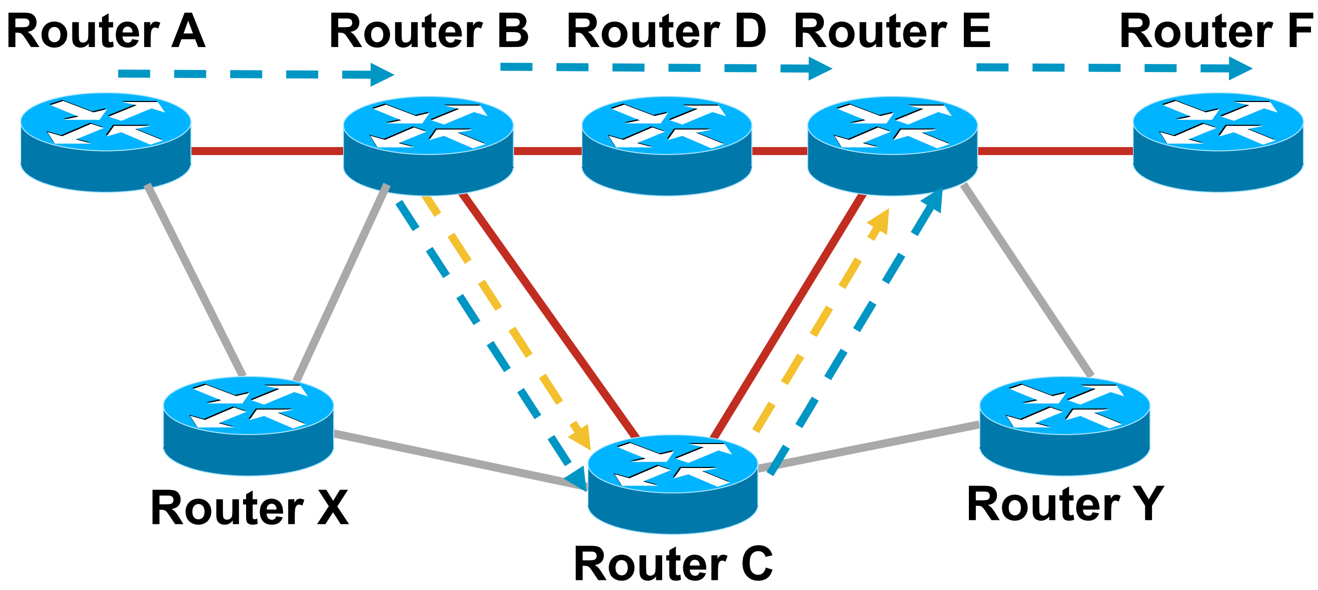


- Primary tunnel: A → B → D → E
- Backup tunnel: B → C → D (preprovisioned)
- Recovery = ~50ms



***Actual Time Varies—Well Below 50ms in Lab Tests, Can Also Be Higher**

Node Protection



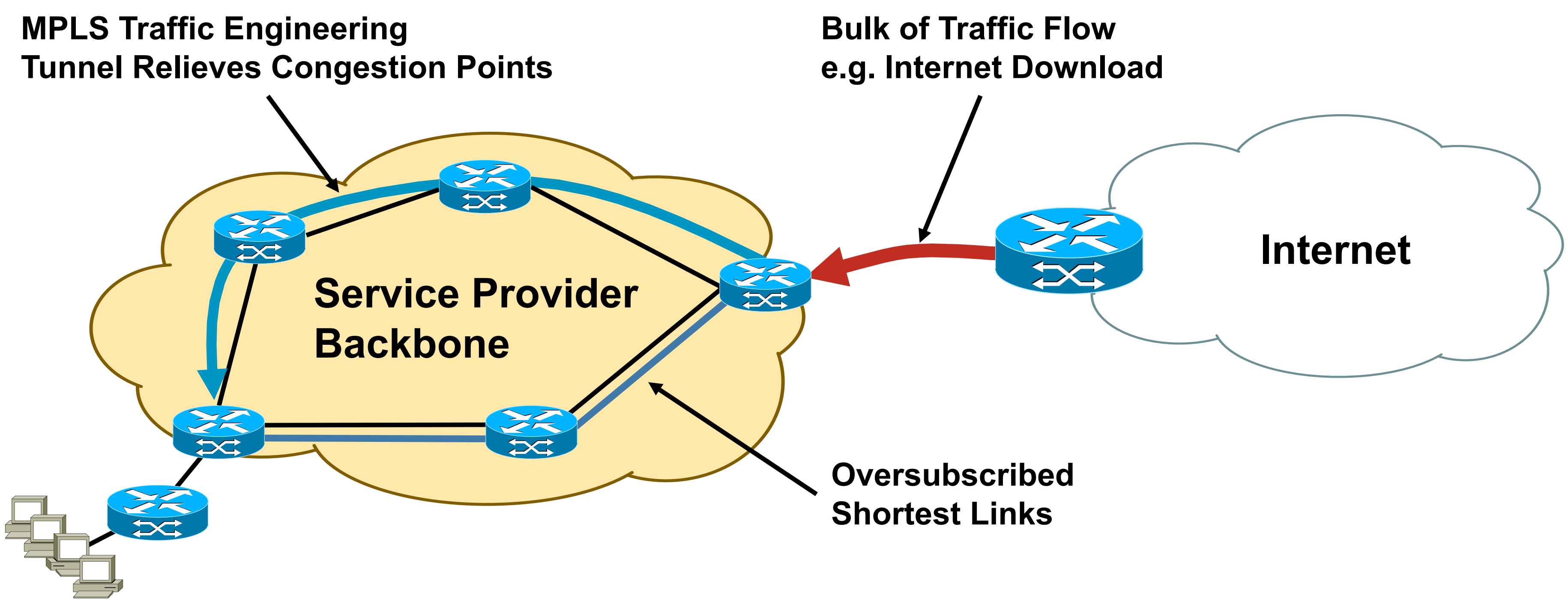
- Primary tunnel: A → B → D → E → F
- Backup tunnel: B → C → E (pre-provisioned)
- Recovery = ~100ms

TE Deployment Scenarios

Tactical TE Deployment

Requirement: Need to Handle Scattered Congestion Points in the Network

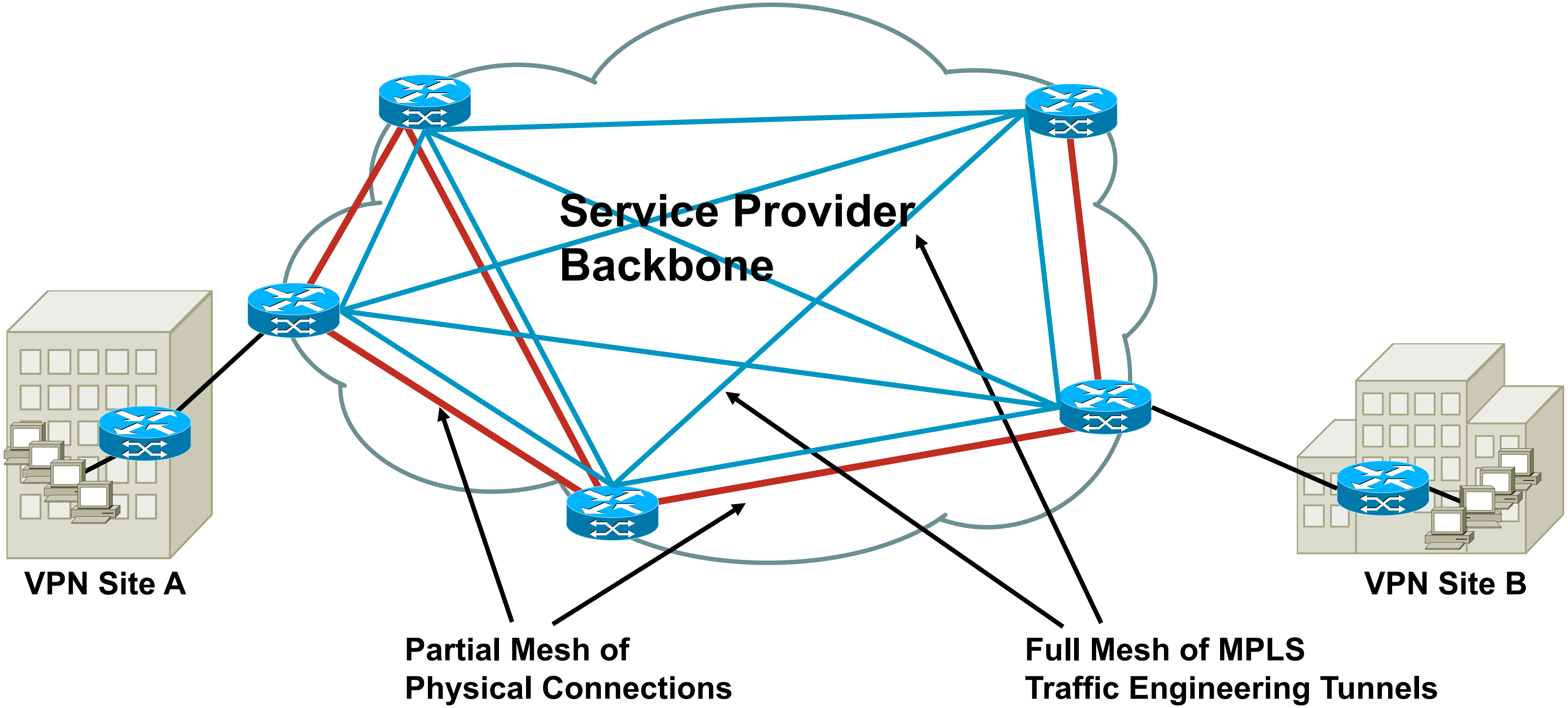
Solution: Deploy MPLS TE on Only Those Nodes That Face Congestion



Full Mesh TE Deployment

Requirement: Need to Increase "Bandwidth Inventory" Across the Network

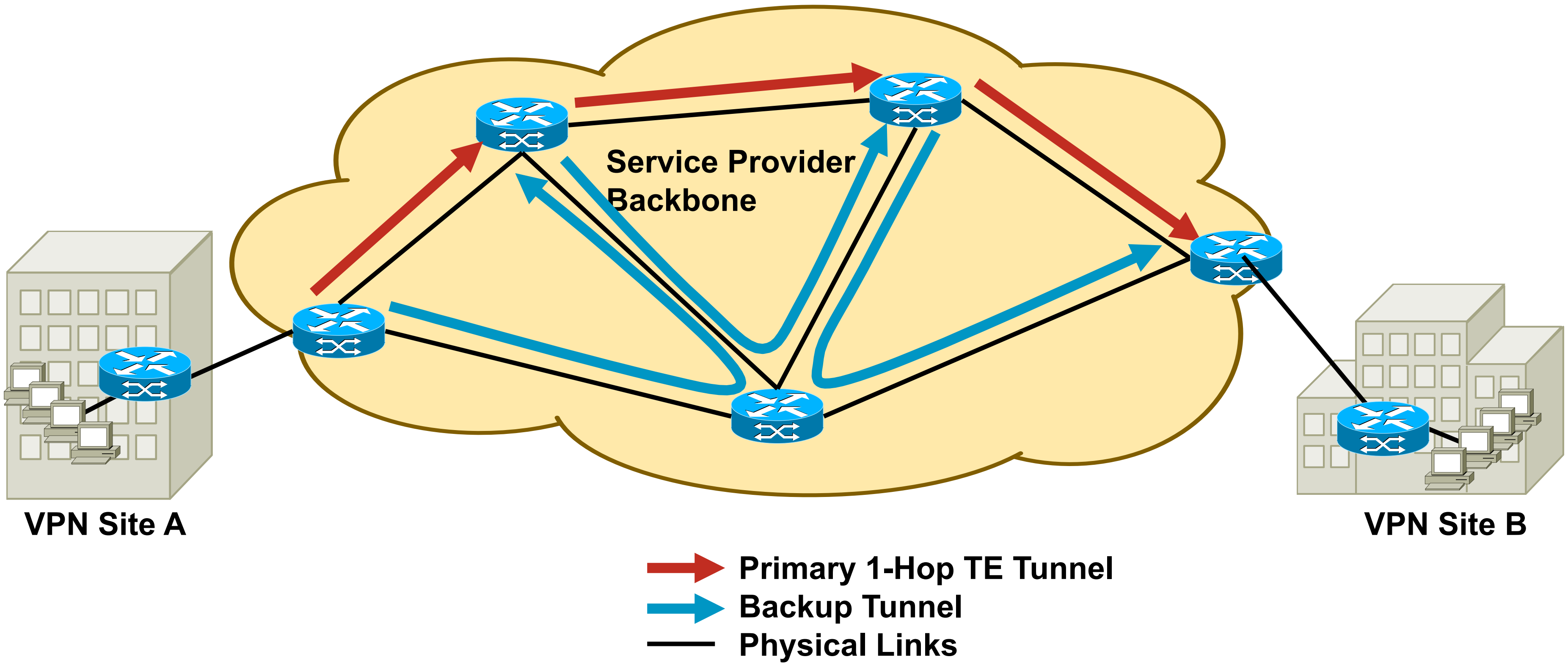
Solution: **Deploy MPLS TE with a Full Logical Mesh over a Partial Physical Mesh and Use Offline Capacity Planning Tool**



1-Hop TE Deployment

Requirement: Need Protection Only—Minimize Packet Loss
Lots of Bandwidth in the Core

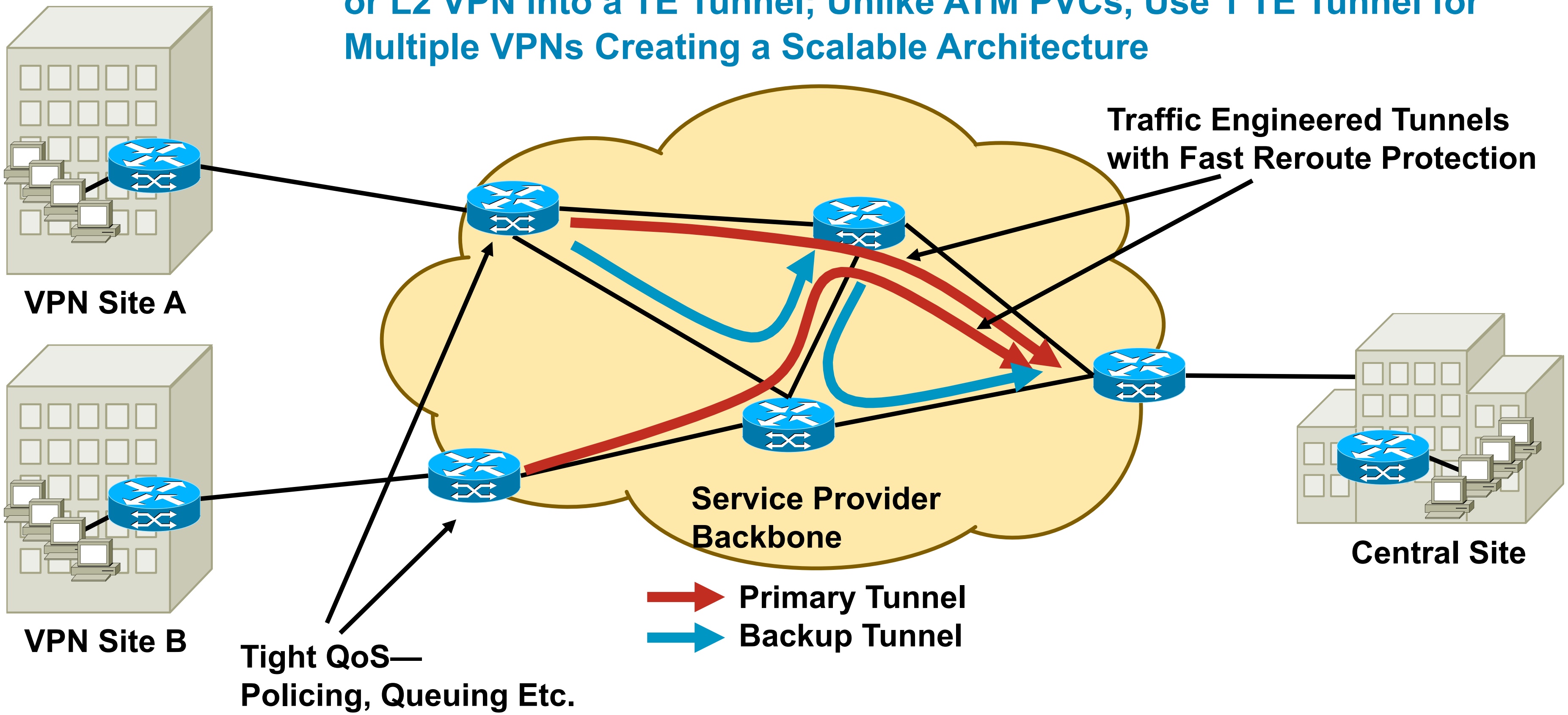
Solution: Deploy MPLS Fast Reroute for Less than 50ms Failover Time with 1-Hop Primary TE Tunnels and Backup Tunnel for Each



Virtual Leased Line Deployment

Requirement: Need to Create Dedicated Point-to-Point Circuits with Bandwidth Guarantees—Virtual Leased Line (VLL)

Solution: Deploy MPLS TE (or DS-TE) with QoS; Forward Traffic from L3 VPN or L2 VPN into a TE Tunnel; Unlike ATM PVCs, Use 1 TE Tunnel for Multiple VPNs Creating a Scalable Architecture

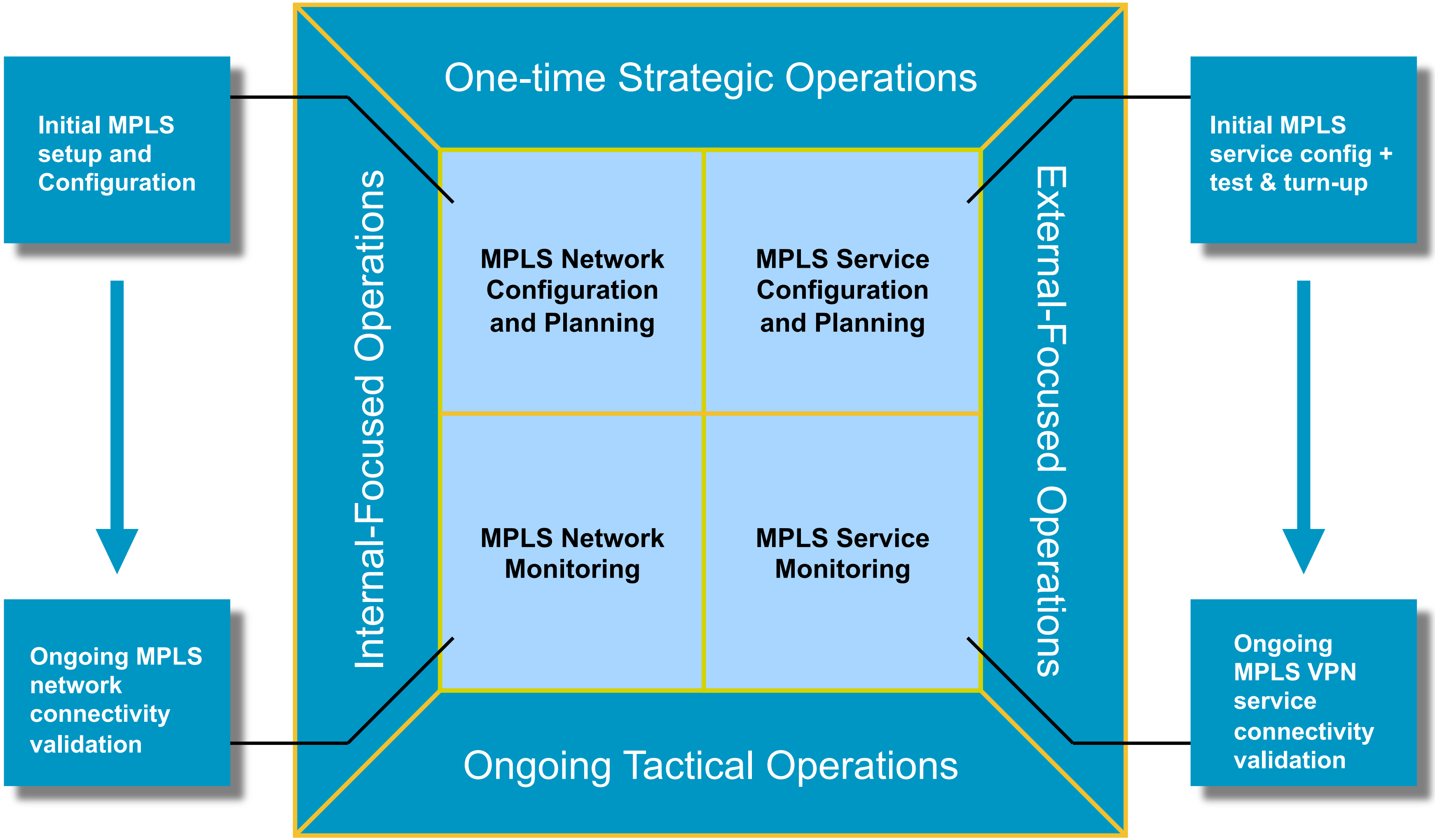


MPLS TE Summary

- Useful for rerouting traffic in congested environments
- Build innovative services like virtual leased line
- Build protection solutions using MPLS FRR

MPLS Management

MPLS Operations Framework



MPLS Embedded Management

- MPLS management capabilities integrated into routers
- IETF standards based + vendor-specific value adds
- MPLS embedded management feature areas
  - MPLS SNMP MIBs (Draft, RFC-based + vendor extensions)
  - MPLS OAM (Draft, RFC-based + Vendor-specific automation)
  - MPLS-aware Net Flow
- MPLS SNMP MIBs
  - MPLS LSR, LDP, TE, FRR, and L3VPN MIB
  - VRF-aware MIB support
- MPLS OAM
  - LSP Ping, Trace, and Multipath (ECMP) Tree Trace
  - IP SLA – LSP Health Monitor

LSP Ping

- Feature Functionality
  - Enables detailed MPLS data path validation between PE routers
- Benefits
  - Finds MPLS-specific forwarding errors not detected by regular IP ping operations
  - Enables detailed MPLS forwarding trouble shooting not available by other existing IP connectivity validations tools
- Key CLI Commands
  - **`ping mpls { ipv4 destination-address destination-mask | pseudowire ipv4-address vc-id vc-id | traffic-eng tunnel-interface tunnel-number }`**

LSP Trace

- Feature Functionality
  - Enables hop-by-hop trouble shooting (fault isolation) along PE-PE LSP path in MPLS network
- Benefits
  - Finds MPLS-specific forwarding failures along PE-PE LSP path, which can not be detected by regular IP traceroute operations
- Key CLI Commands
  - **`trace mpls {ipv4 destination-address destination-mask | traffic-eng tunnel-interface tunnel-number}`**

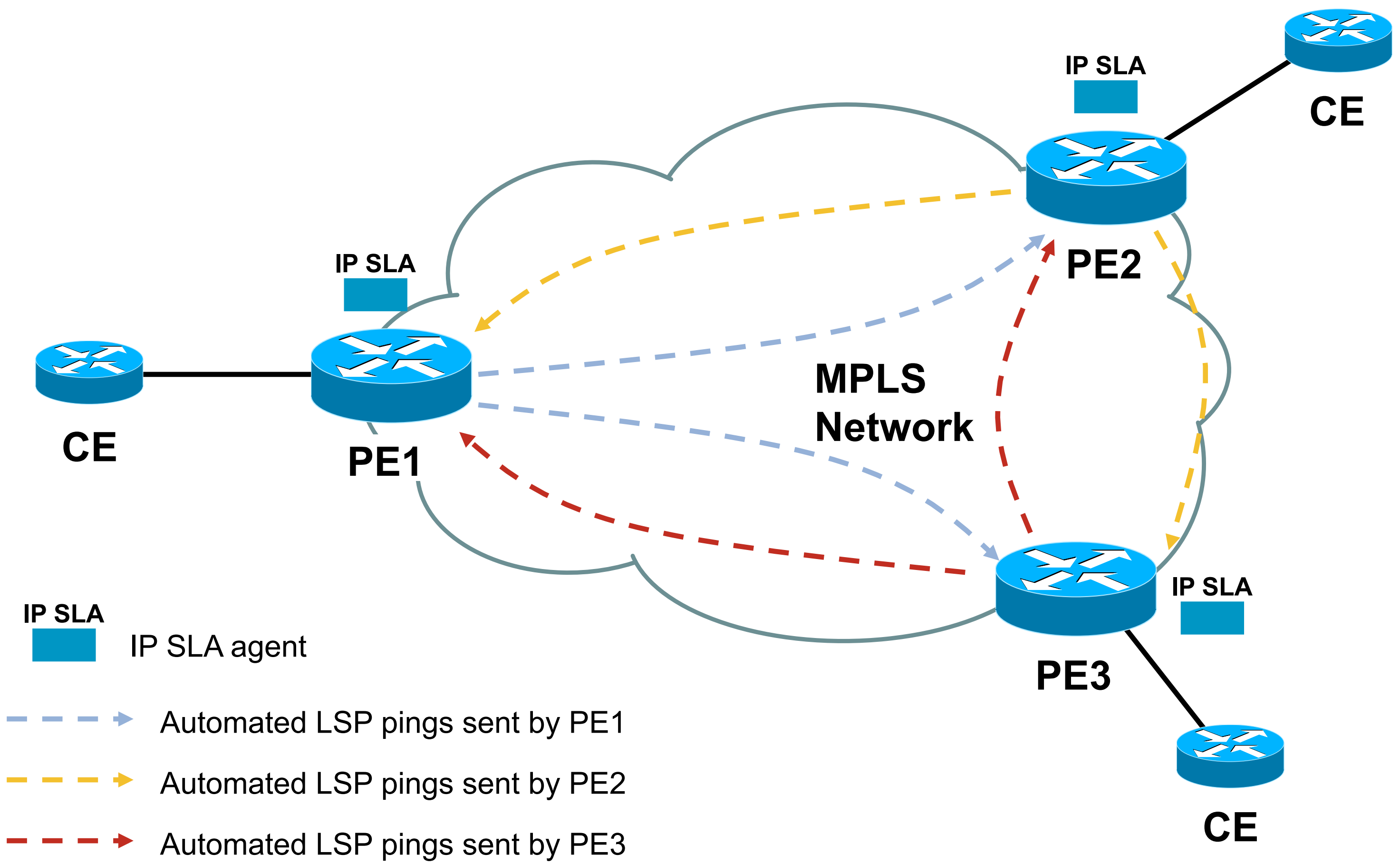
LSP Multi-Path (ECMP) Trace

- Feature Functionality
  - Enables discovery and hop-by-hop trouble shooting of all available MPLS (LSP) paths between two PE routers
- Benefits
  - Detailed discovery of all MPLS (LSP) paths between PE routers which can not be detected by regular IP traceroute operations
- Key CLI Commands
  - `trace mpls multipath ipv4 destination-address/destination-mask-length`

IP SLA – LSP Health Monitor

- Feature Functionality
  - Enables automation of LSP ping operation and generation/logging of SNMP Traps after consecutive MPLS LSP connectivity failures have been detected
- Benefits
  - Detailed control over LSP ping probe frequency (primary and secondary frequency) and event control (e.g., Traps, logging) after MPLS LSP connectivity failure has been detected
  - Automated discovery of remote PE target routers via BGP VPN next-hop discovery
- Key CLI Commands
  - `mpls discovery vpn next-hop`
  - `auto ip sla mpls-lsp-monitor [operation-number]`
  - `type echo | pathEcho`
  - `show ip sla mpls-lsp-monitor configuration [operation-number]`
  - `auto ip sla mpls-lsp-monitor schedule`

Automated MPLS OAM



Summary