Introduction to MPLS

Tor Skeie

University of Oslo, Department of informatics

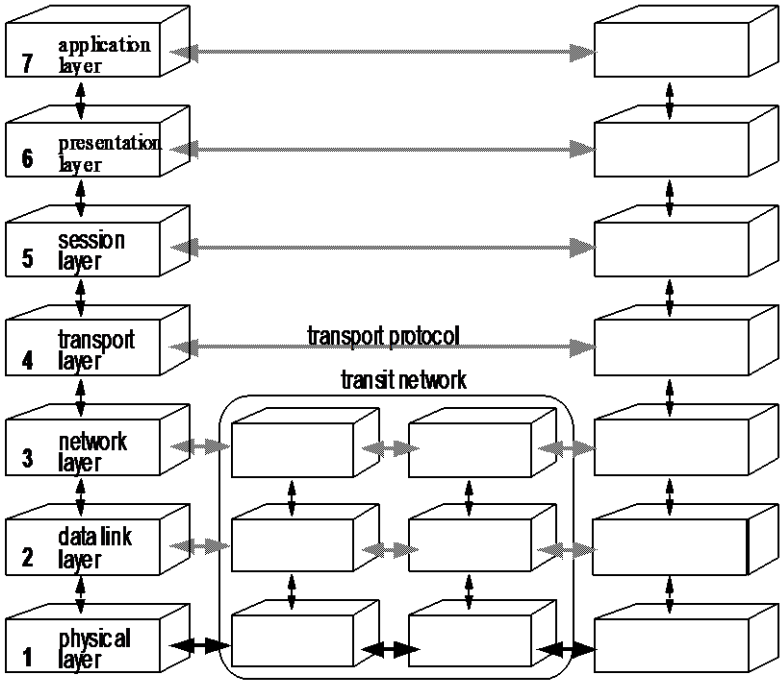
This presentation is based on Santanu Dasgupta´s slide series that describe the Multiprotocol Label Switching (MPLS) concept, Santanu Dasgupta is a Distinguished Architect at Cisco



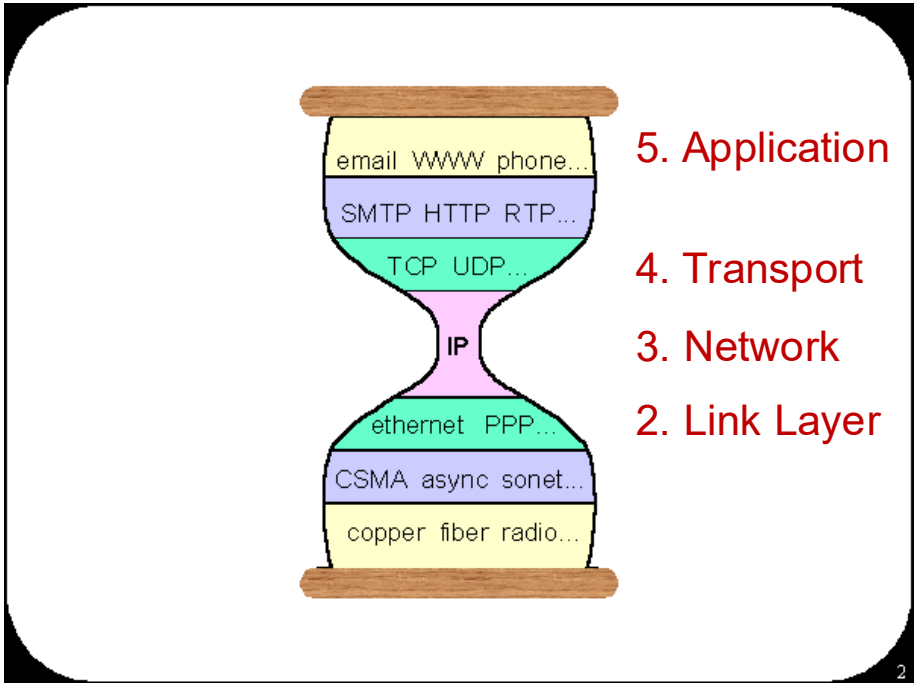
UiO : University of Oslo

Theory

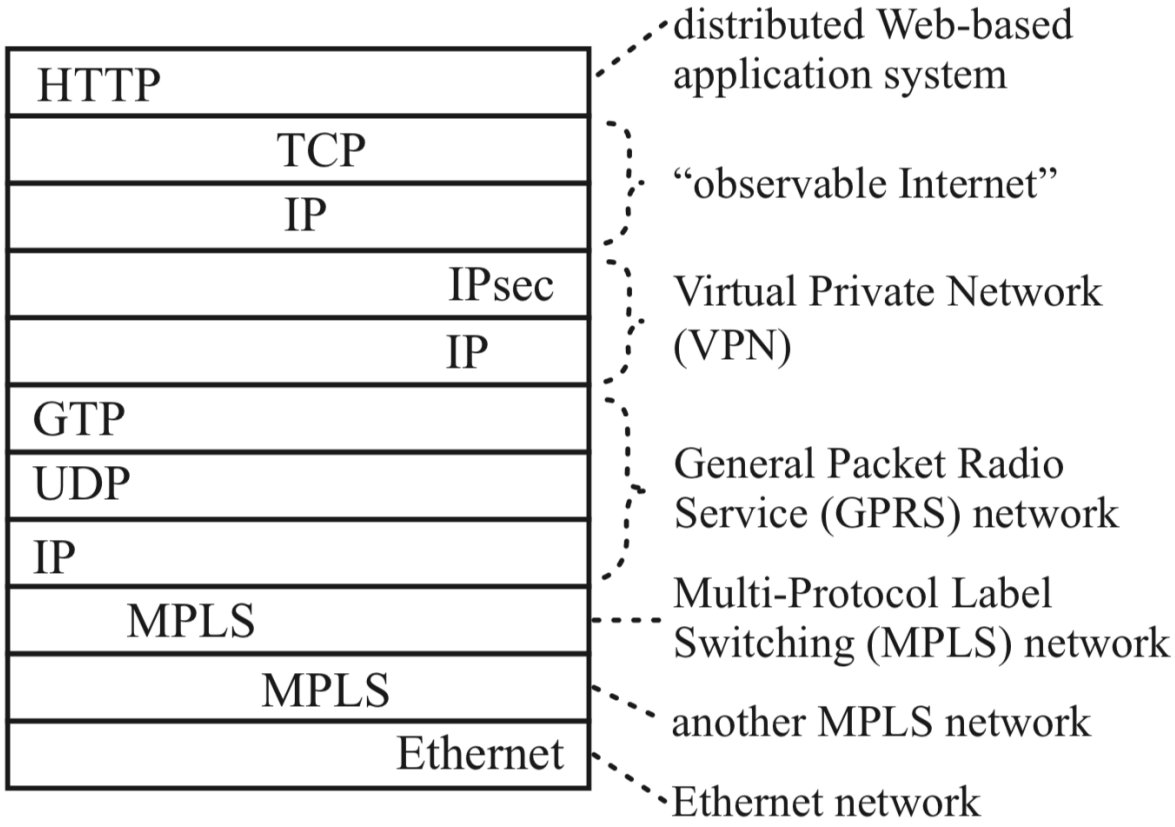
ISO/OSI model



DoD model



Reality



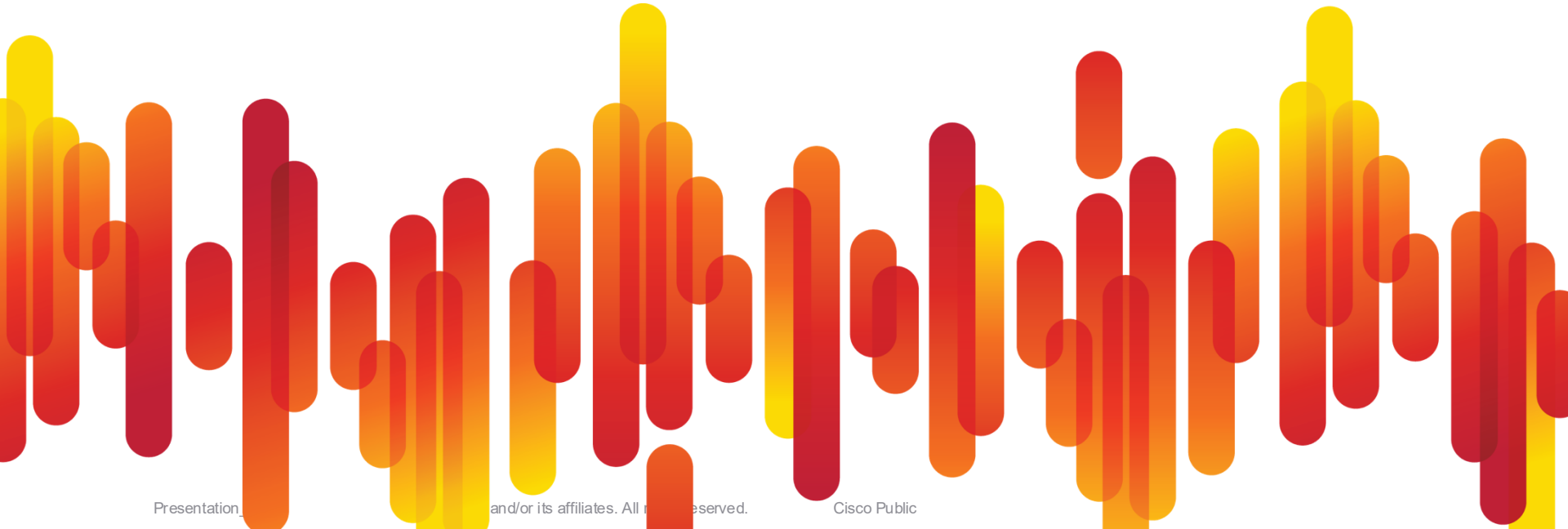
Typical packet from AT&T backbone network

Pamela Zave and Jennifer Rexford. 2019. The compositional architecture of the internet. Commun. ACM 62, 3 (February 2019), 78-87. DOI: https://doi.org/10.1145/3226588

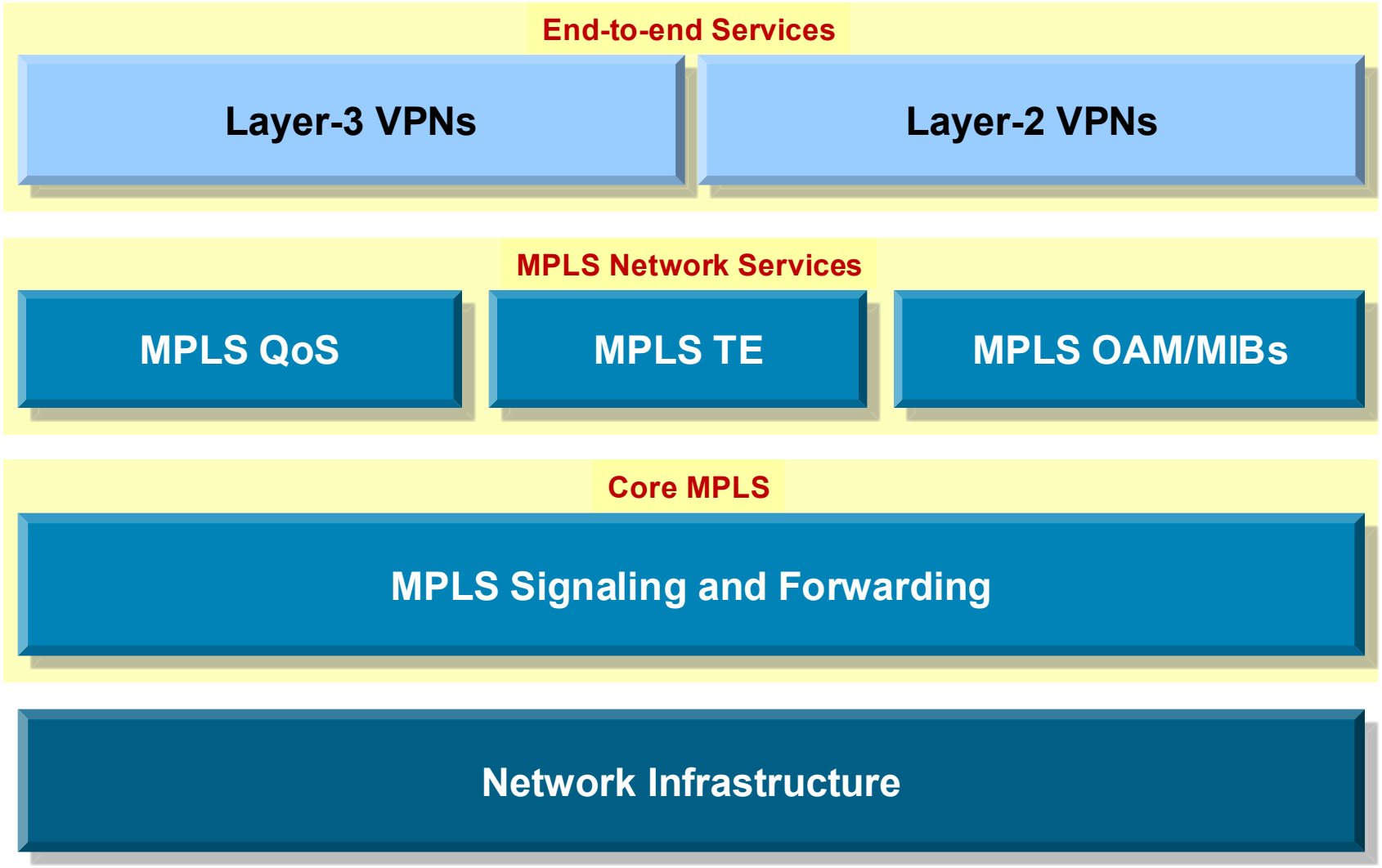
Goals of this Lecture

- Understand the “business” drivers for MPLS
  - Understand the problems MPLS is addressing, herein learn about MPLS customer and market segments
  - Understand benefits of deploying MPLS
- Understand the major MPLS technology components
- Understand typical applications of MPLS

Introduction



The Big Picture



Why Multi Protocol Label Switching?

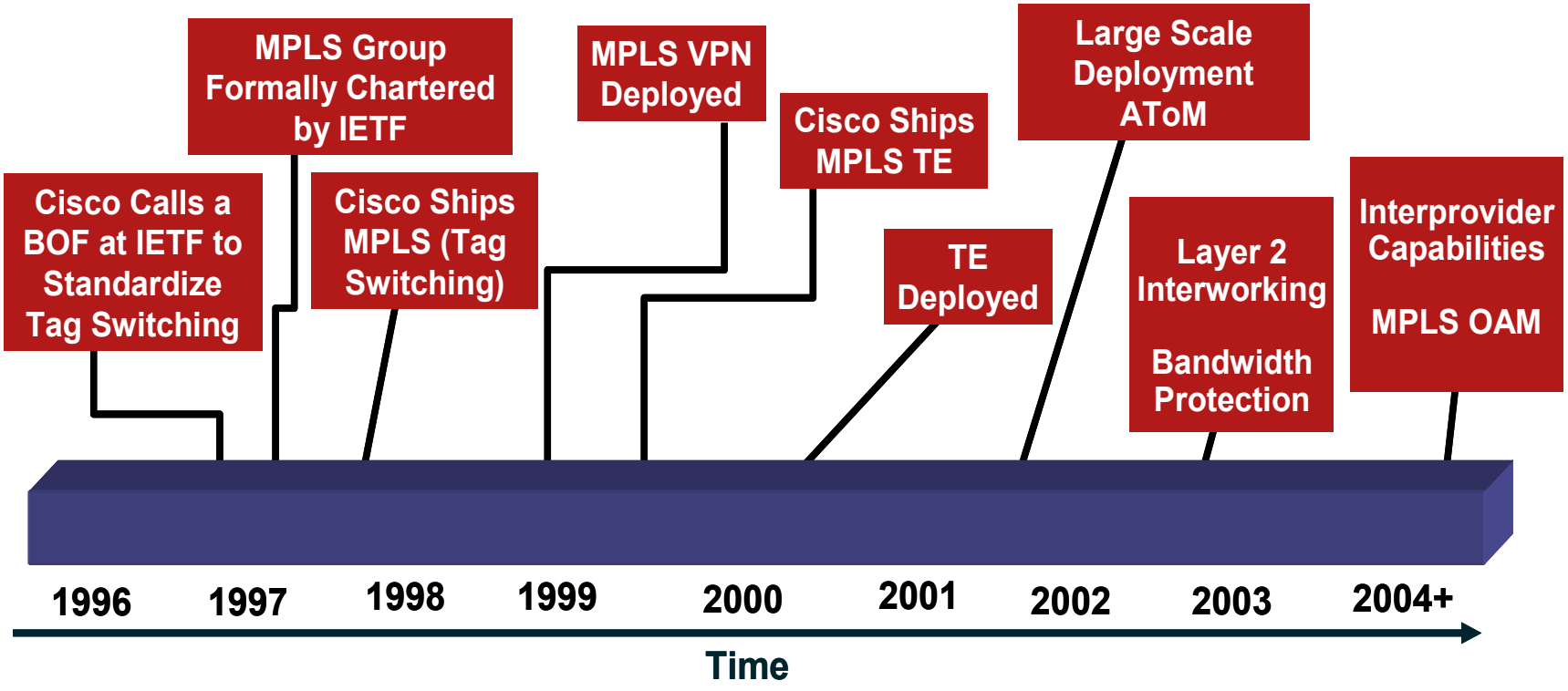
- SP/Carrier perspective
  - Reduce costs (CAPEX); consolidate networks
  - Consolidated network for multiple Layer-2/3 services
  - Support increasingly stringent SLAs
  - Handle increasing scale/complexity of IP-based services
- Enterprise/end-user perspective
  - Campus/LAN
    - Need for network segmentation (users, applications, etc.)
  - WAN connectivity (connecting enterprise networks)
    - Need for easier configuration of site-to-site WAN connectivity

What Is MPLS Technology?

- It's all about labels …
- Use the best of both worlds
  - Layer-2 (ATM/FR): efficient forwarding and traffic engineering
  - Layer-3 (IP): flexible and scalable
- MPLS forwarding plane
  - Use of labels for forwarding Layer-2/3 data traffic
  - Labeled packets are being switched instead of routed
    - Leverage layer-2 forwarding efficiency
- MPLS control/signaling plane
  - Use of existing IP control protocols extensions + new protocols to exchange label information
    - Leverage layer-3 control protocol flexibility and scalability

Evolution of MPLS

- Evolved from tag switching in 1996 to full IETF standard, covering over 130 RFCs
- Key application initially were Layer-3 VPNs, followed by Traffic Engineering (TE), and Layer-2 VPNs



Attractive MPLS Key Features …



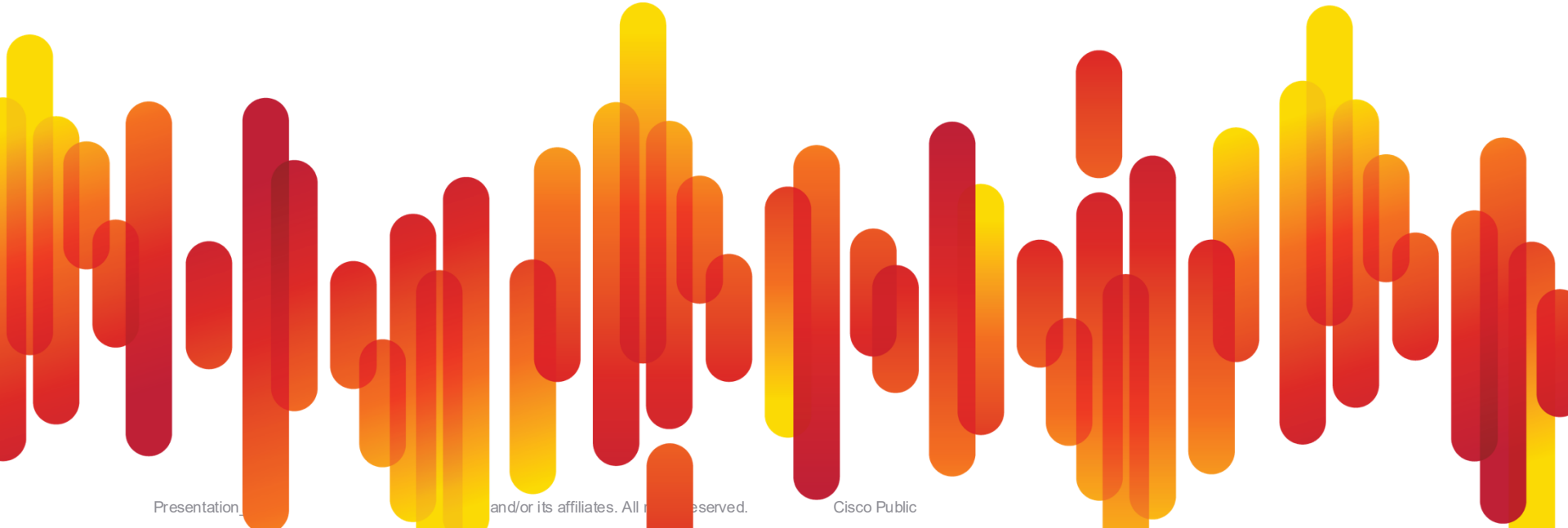
For your reference only

	Service Providers	Enterprise Data Center	Data center interconnects	EWAN Edge
Key Features	L2/L3VPN's TE/FRR QoS High Availability	VPN's TE/FRR High Availability	VPN's / VRF's VRF-Aware Security High Availability	VPN's / VRF's VRF Aware Security High Availability
Applications	Hosted Data centers Data center interconnect Segmentation for IT Mergers, Acquisitions, spinoffs	Departmental segmentation Service multiplexing Security Mergers, Acquisitions, spinoffs	Disaster Recovery Vmotion support Branch Interconnects	Internet Access Branch Connectivity

- **Network Consolidation** – Merging Multiple parallel networks into a shared infrastructure
- **Network segmentation** – By user groups or business function
- **Service and policy centralization** – Security policies and appliances at a central location
- **New applications readiness** – Converged multi-service network
- **Increased network security** – User groups segmentation with VPNs

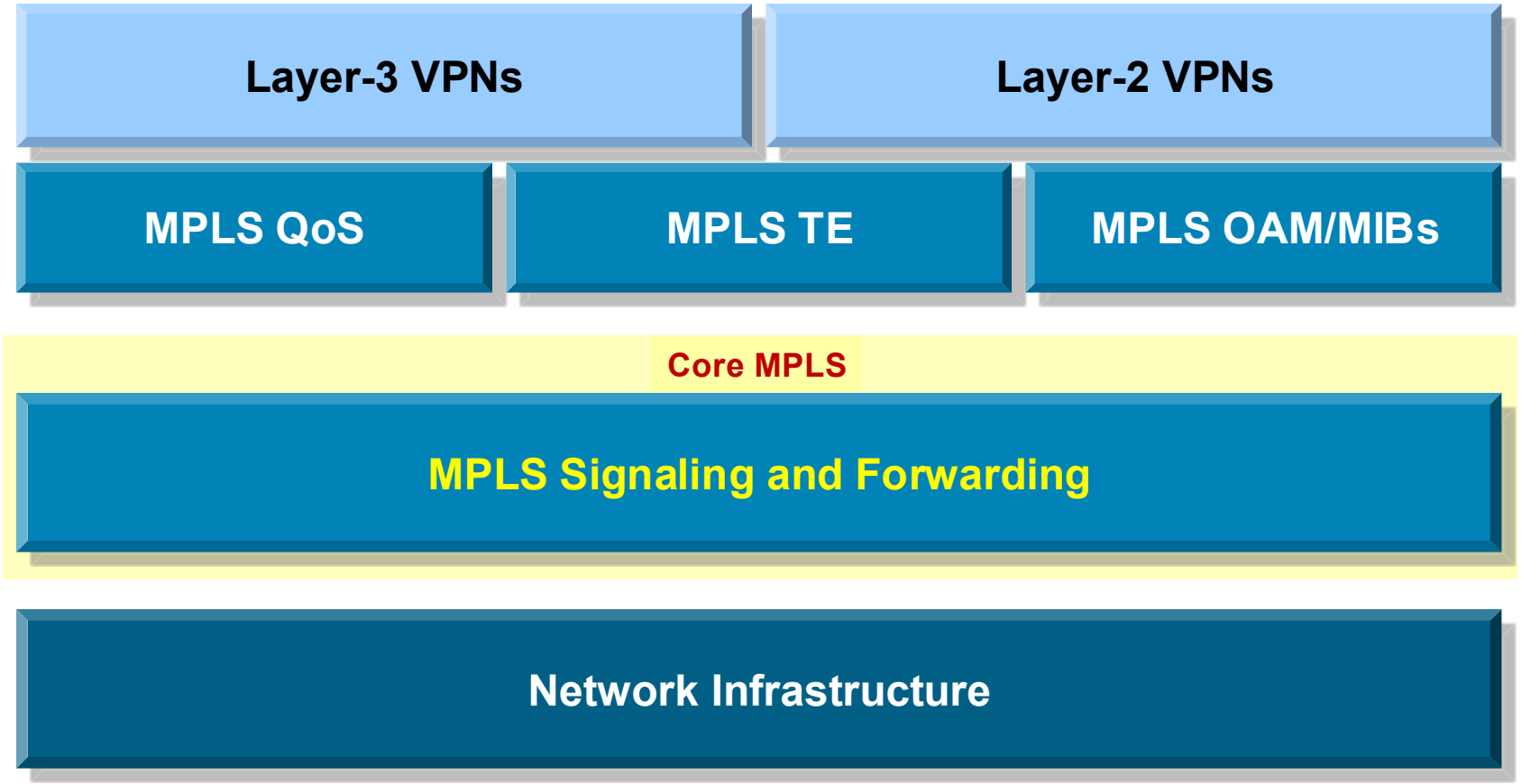
MPLS Technology Components

Basic building blocks of MPLS



MPLS Forwarding and Signaling

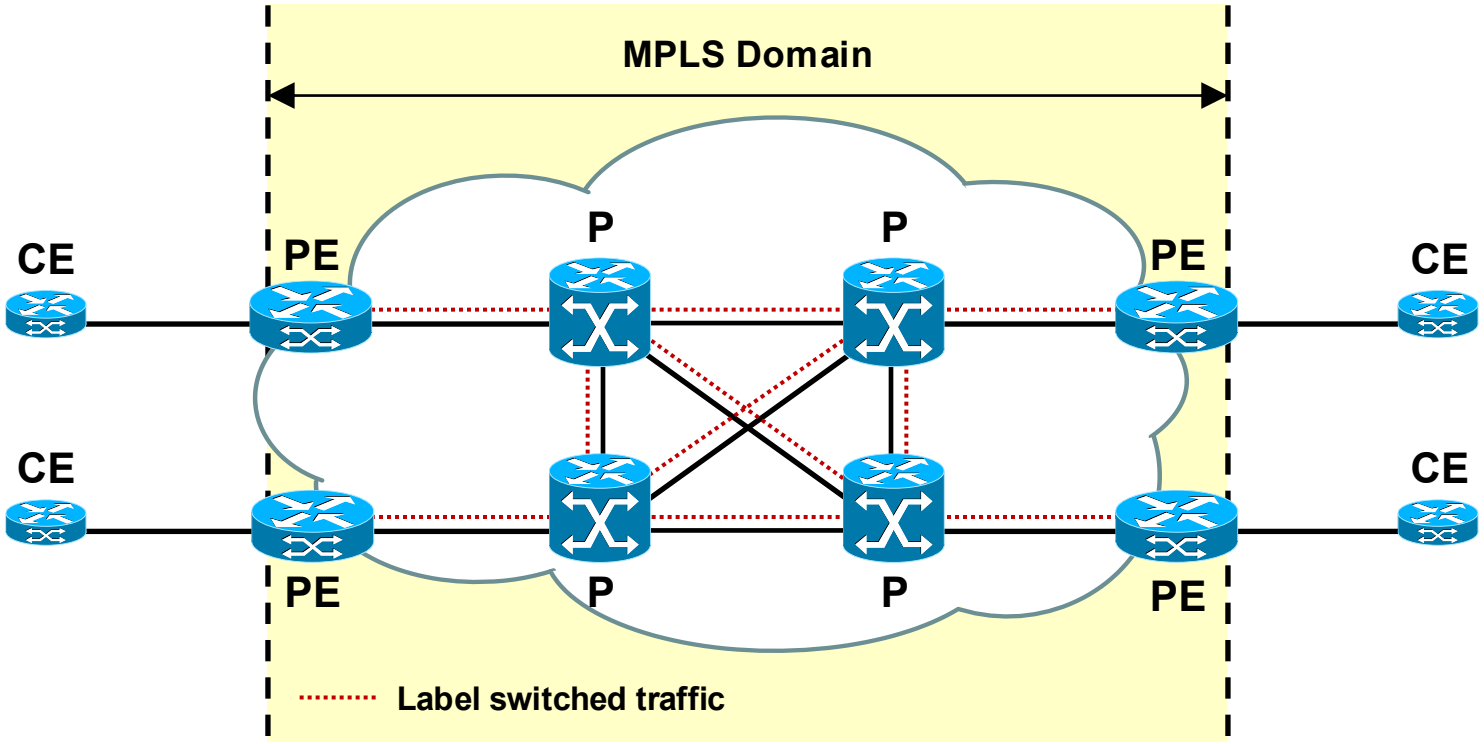
- MPLS label forwarding and signaling mechanisms



Basic Building Blocks

- The big picture
  - MPLS-enabled network devices
  - Label Switched Paths (LSPs)
- The internals
  - MPLS labels
  - Processing of MPLS labels
  - Exchange of label mapping information
  - Forwarding of labeled packets
- Other related protocols to exchange label information
  - Between MPLS-enabled devices

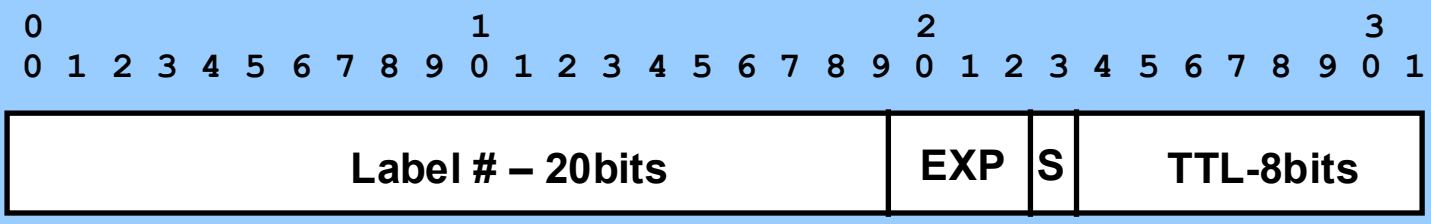
MPLS Network Overview



- CE (Customer Edge) router
  - Connects customer network to MPLS network
- PE (Provider Edge) router = edge router (LSR)
  - Imposes and removes MPLS labels
- P (Provider) router = label switching router = core router (LSR)
  - Switches MPLS-labeled packets

MPLS Label and Label Encapsulation

MPLS Label



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

MPLS Label Encapsulation

PPP Header (Packet over SONET/SDH)

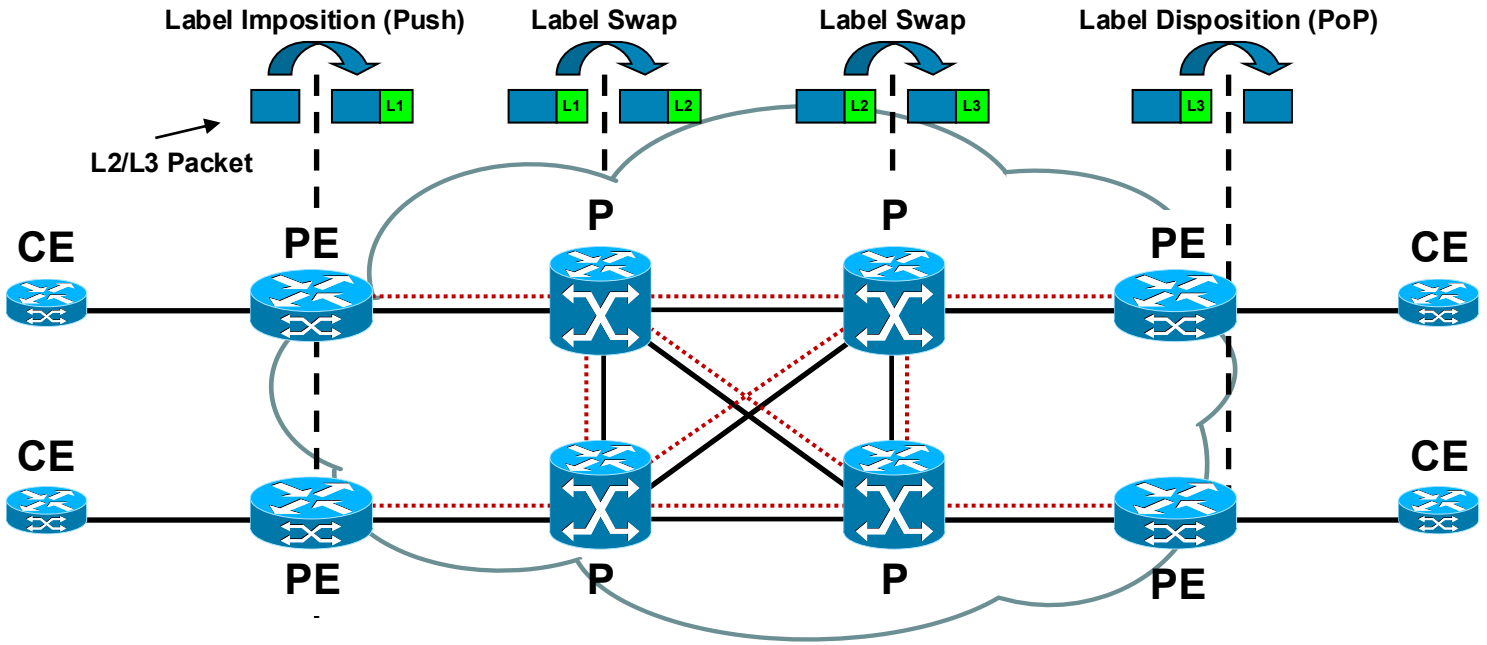


One or More Labels Appended to the Packet (Between L2/L3 packet and link layer header)

LAN MAC Label Header



MPLS Label Operations



- Label imposition (Push)

  By ingress PE router; classify and label packets

- Label swapping or switching

  By P router; forward packets using labels; indicates service class & destination

- Label disposition (PoP)

  By egress PE router; remove label and forward original packet to destination CE

Forwarding Equivalence Class

- Mechanism to map ingress layer-2/3 packets onto a Label Switched Path (LSP) by ingress PE router
  - Part of label imposition (Push) operation
- Variety of FEC mappings possible
  - IP prefix/host address
  - Groups of addresses/sites (VPN x)
    - Used for L3VPNs
  - Layer 2 circuit ID (ATM, FR, PPP, HDLC, Ethernet)
    - Used for Pseudowires (L2VPNs)
  - Tunnel interface
    - Used for MPLS traffic engineering (TE)

Label Distribution Protocol (LDP)

- MPLS nodes need to exchange label information with each other
  - Ingress PE node (Push operation)
    - Needs to know what label to use for a given FEC to send packet to neighbor
  - Core P node (Swap operation)
    - Needs to know what label to use for swap operation for incoming labeled packets
  - Egress PE node (Pop operation)
    - Needs to tell upstream neighbor what label to use for specific FEC type - **LDP used for exchange of label (mapping) information**
- Label Distribution Protocol (LDP)
  - Defined in RFC 3035 and RFC3036; updated by RFC5036
  - LDP is a superset of the Cisco-specific Tag Distribution Protocol
- Note that, in addition LDP, also other protocols are being used for label information exchange
  - Will be discussed later

MPLS Control and Forwarding Plane

- MPLS control plane
- Used for distributing (signaling) labels and building label-switched paths (LSPs) end-to-end

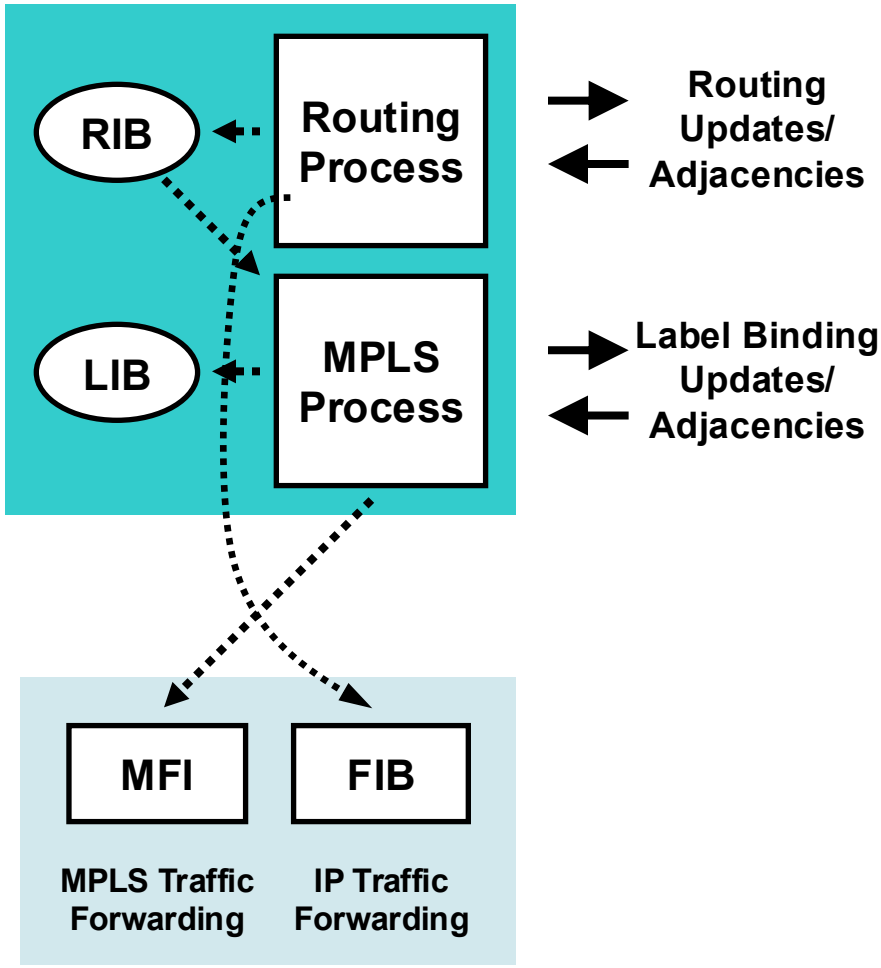
  Typically supported by LDP; also supported via BGP

  Labels define destination and service
- MPLS forwarding plane

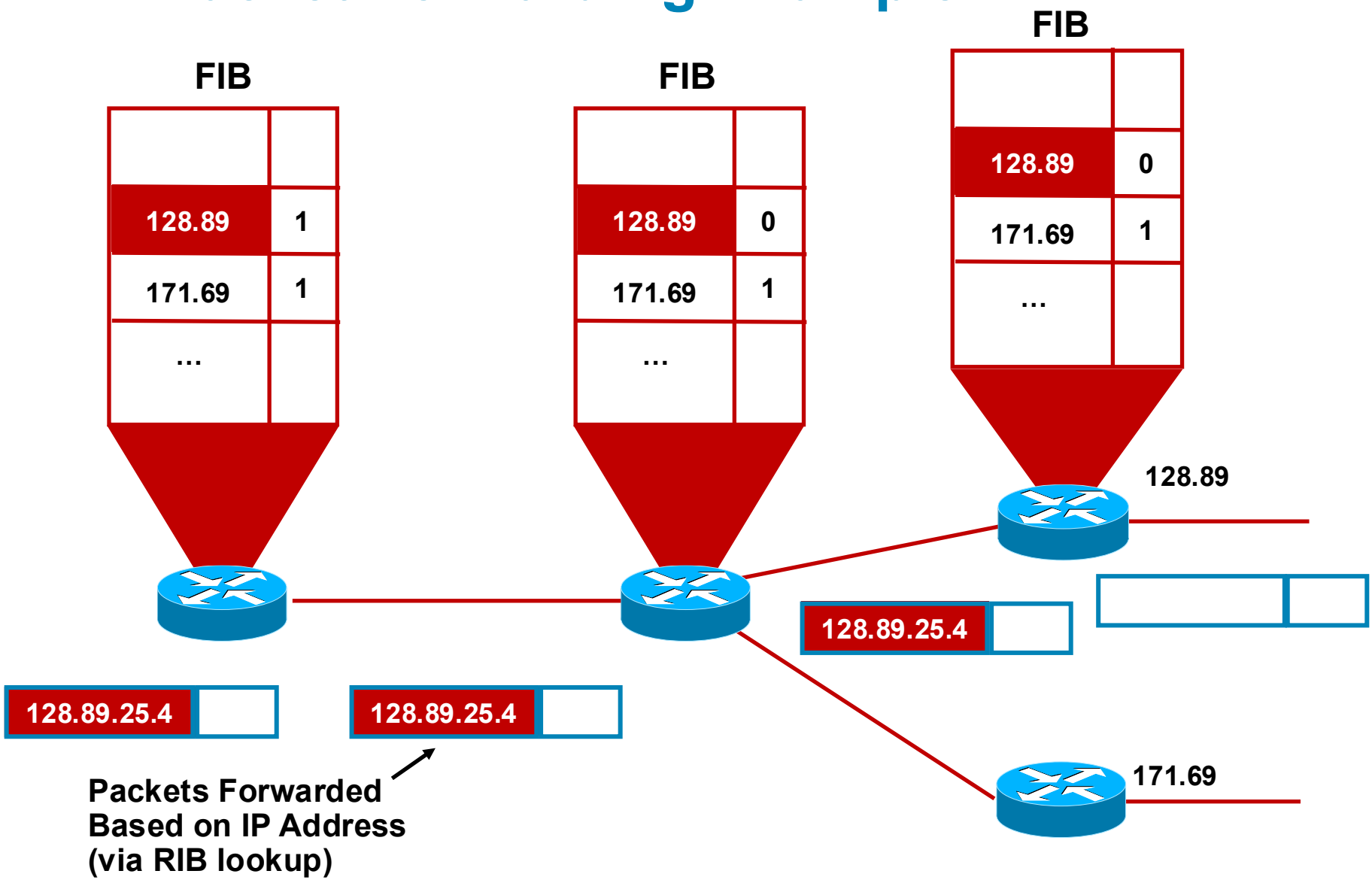
  Used for label imposition (PUSH), swapping, and disposition (POP) operation

  Labels separate forwarding from IP address-based routing

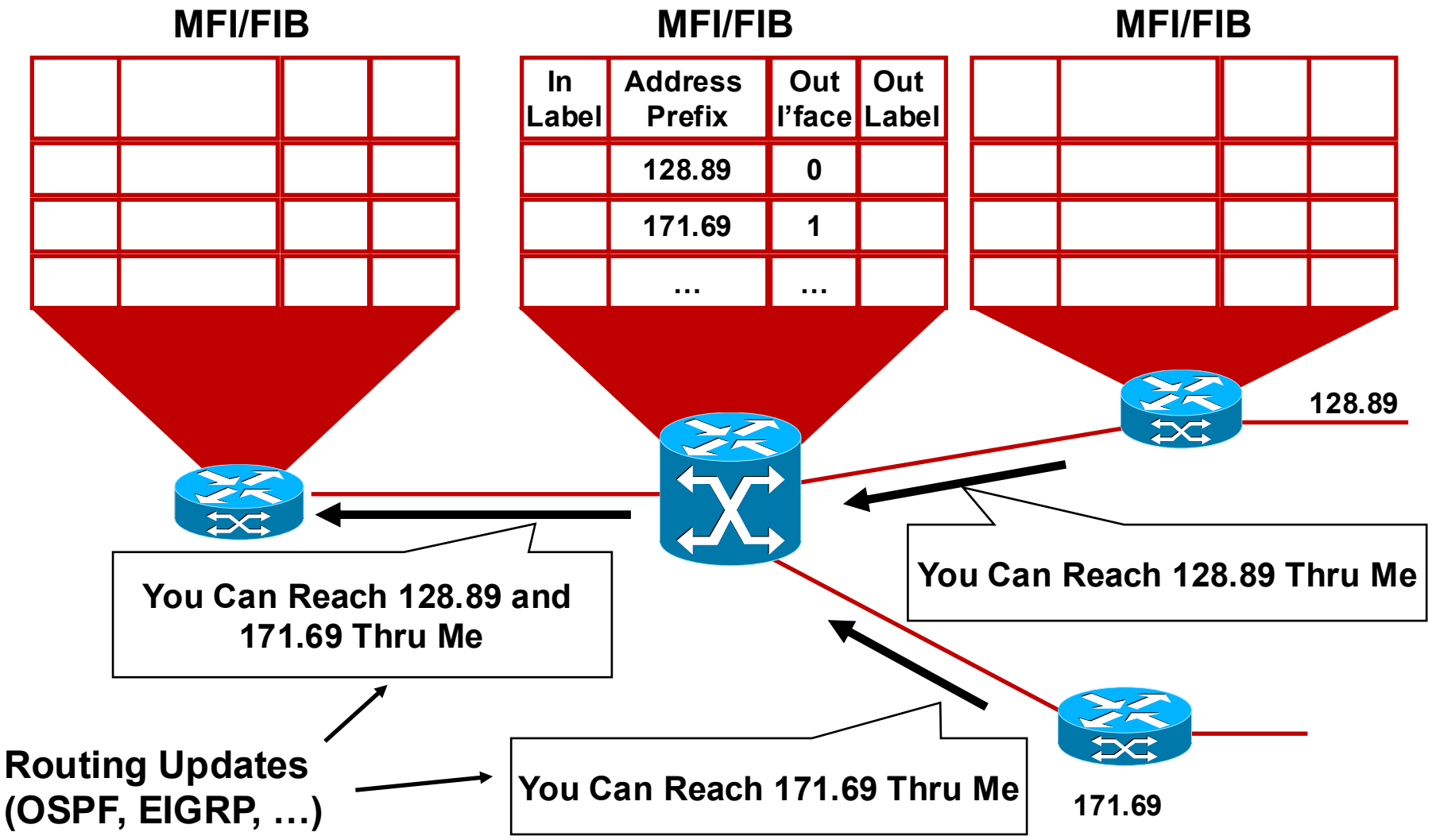
  MFI – MPLS Forwarding Infrastructure



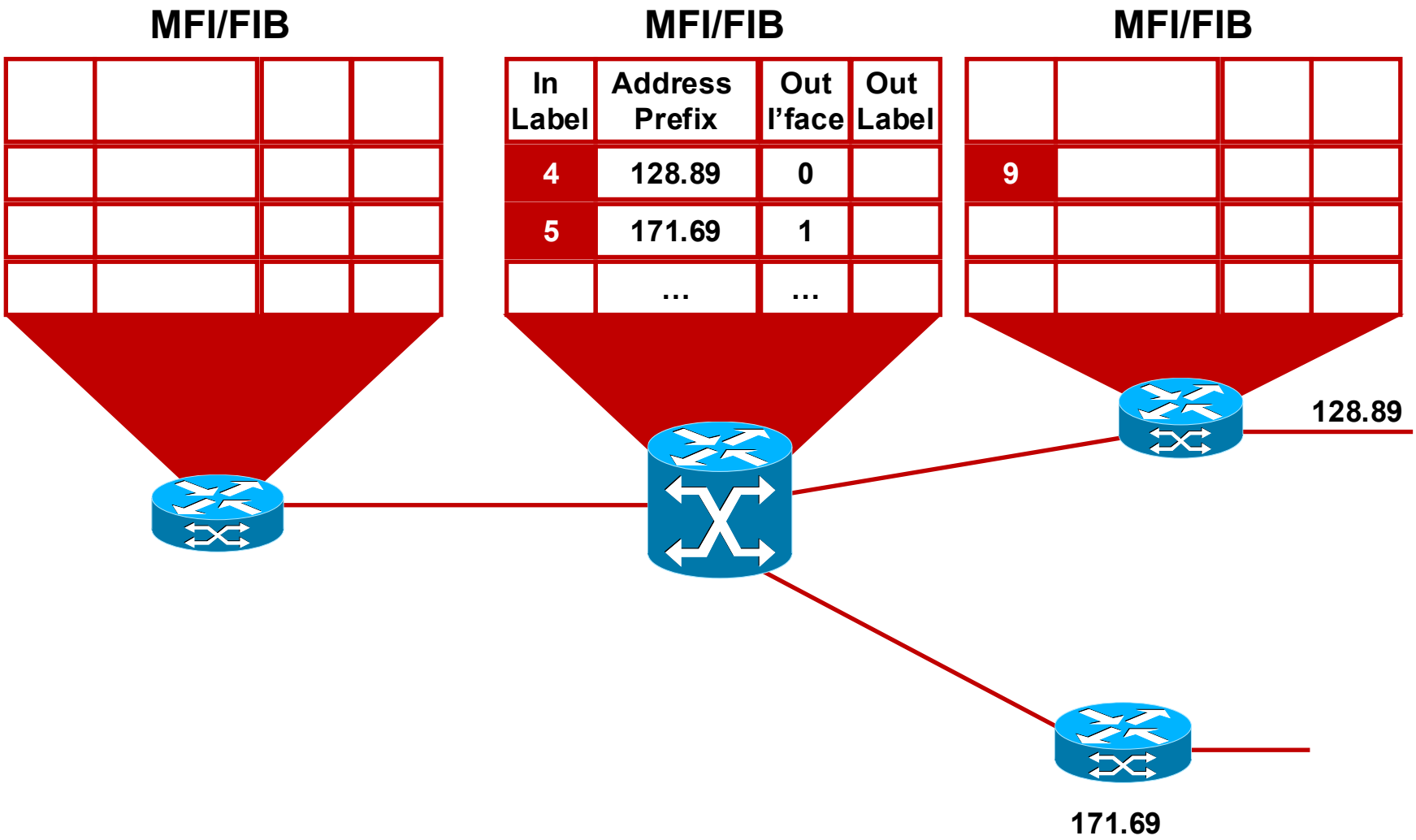
IP Packet Forwarding Example



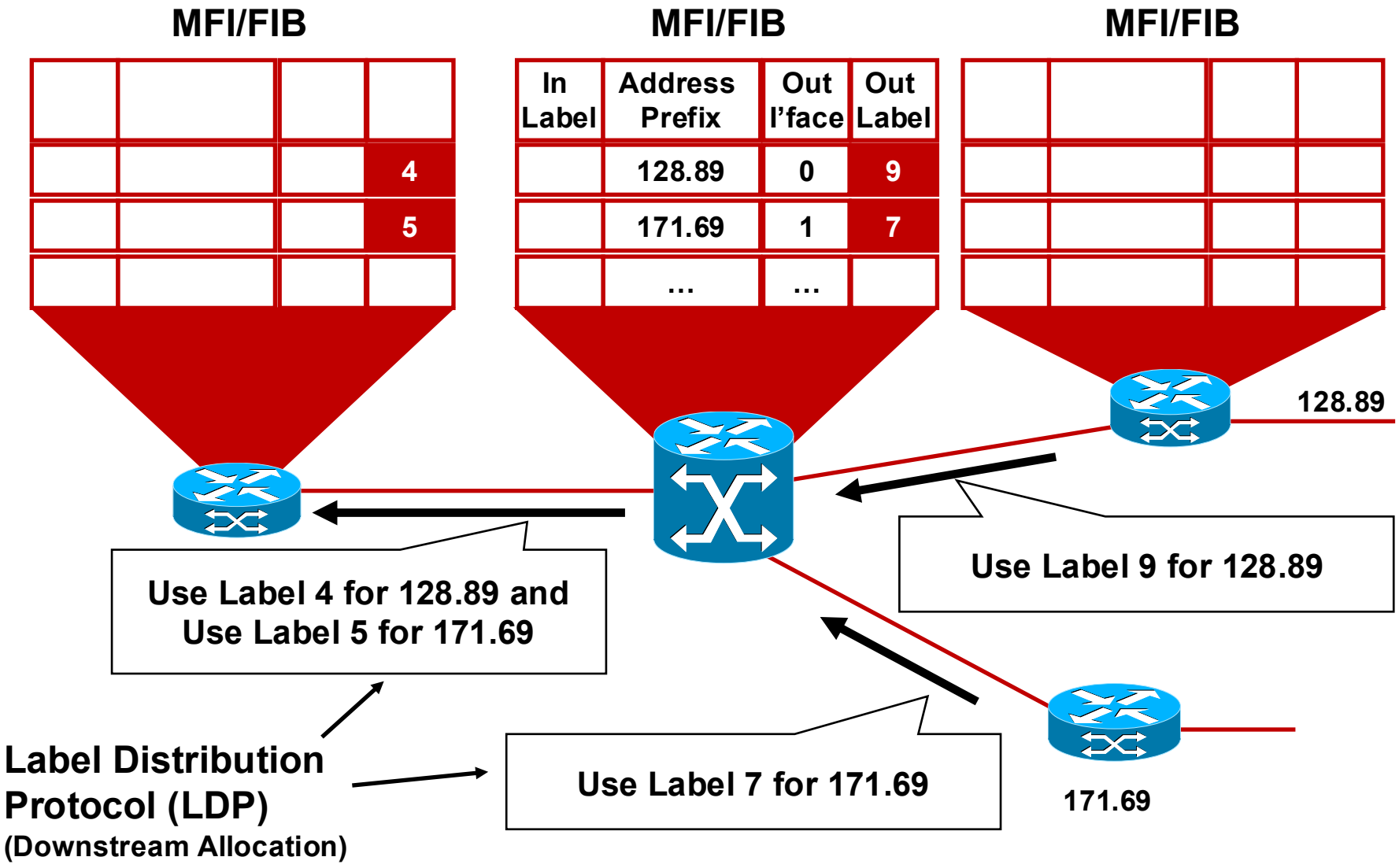
Step 1: IP Routing (IGP) Convergence



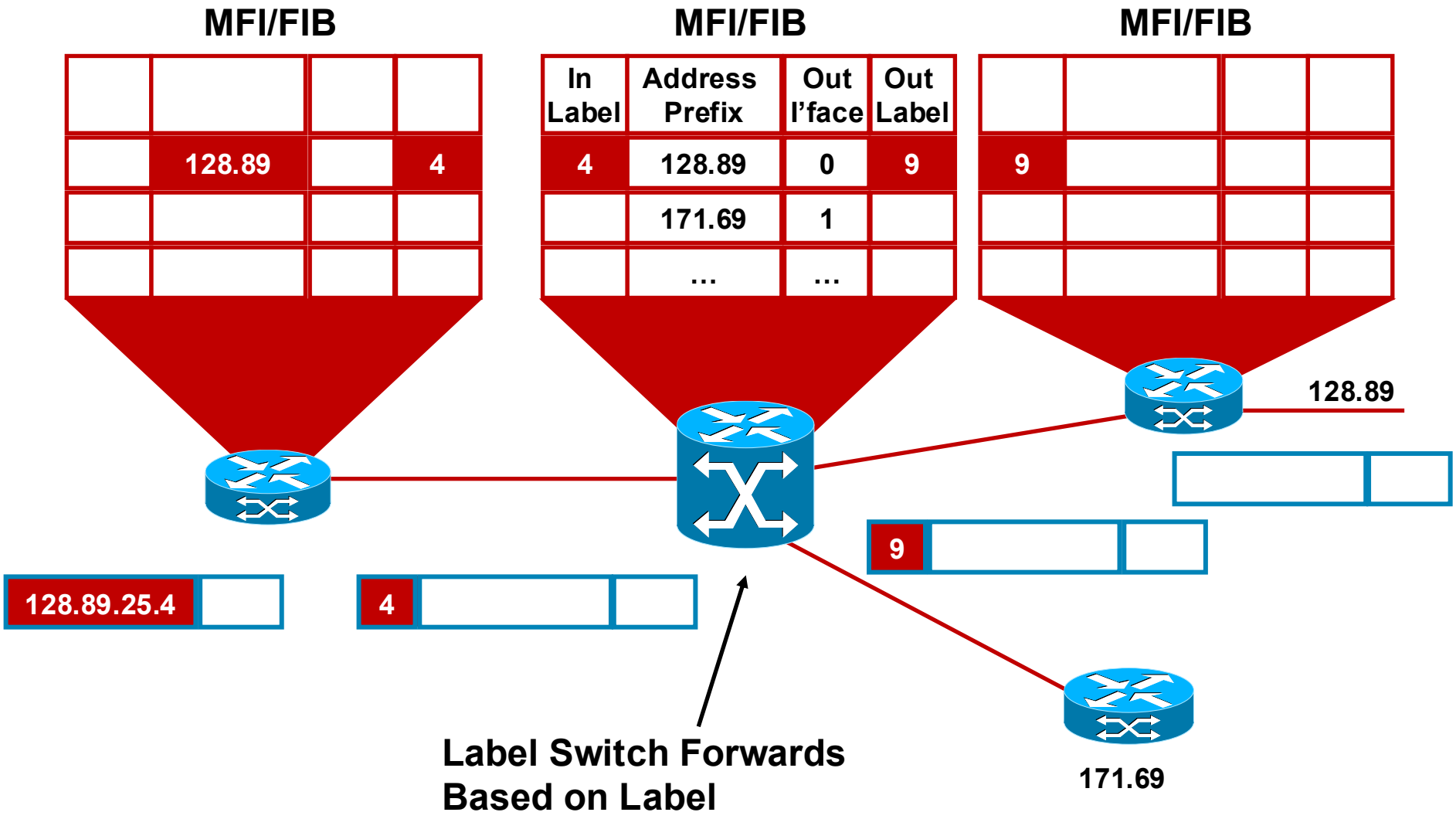
Step 2a: LDP Assigns Local Labels



Step 2b: LDP Assigns Remote Labels



Step 3: Forwarding MPLS Packets

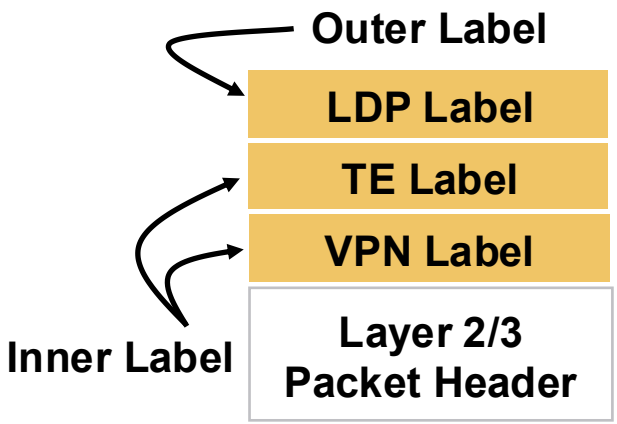


Summary Steps For MPLS Forwarding

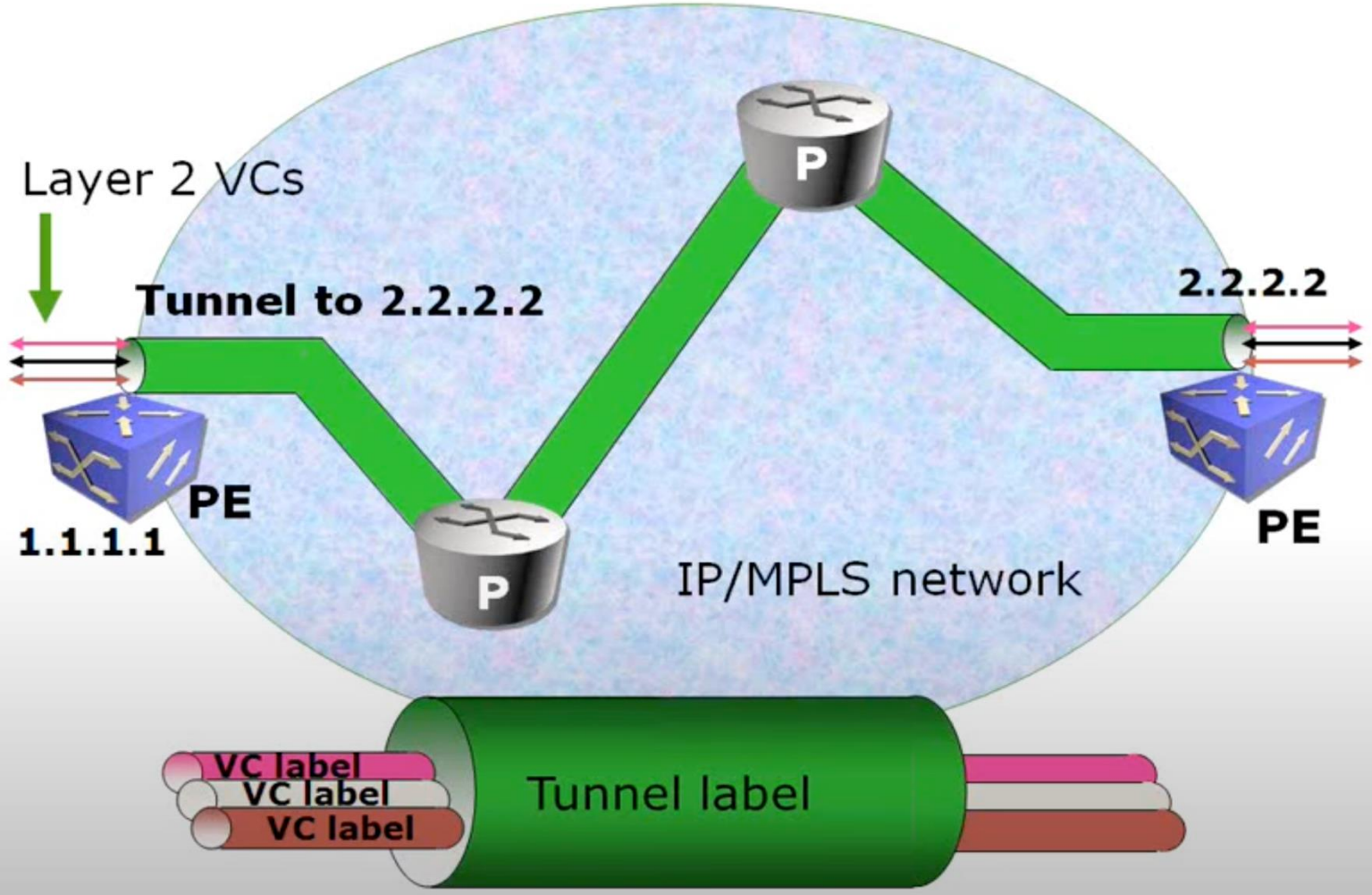
- Each node maintains IP routing information via IGP (Interior Gateway Protocol)
  - IP routing table (RIB) and IP forwarding table (FIB)
- LDP leverages IGP routing information
- LDP label mapping exchange (between MPLS nodes) takes place after IGP has converged
  - LDP depends on IGP convergence
  - Label binding information stored in LIB
- Once LDP has received remote label binding information MPLS forwarding is updated
  - Label bindings are received from remote LDP peers
  - MPLS forwarding via MFI (MPLS Forwarding Infrastructure)

Label Stacking

- More than one label can be used for MPLS packet encapsulation
  - Creation of a label stack
- Recap: labels correspond to Forwarding Equivalence Class (FEC)
  - Each label in stack used for different purposes
- Outer label always used for forwarding MPLS packets in network
- Remaining inner labels used to implement specific services/FECs, etc.
- Last label in stack marked with EOS bit
- Allows building services such as
  - MPLS VPNs; LDP + VPN label
  - Traffic engineering (FRR): LDP + TE label
  - VPNs over TE core: LDP + TE + VPN label
  - Any transport over MPLS: LDP + PW label

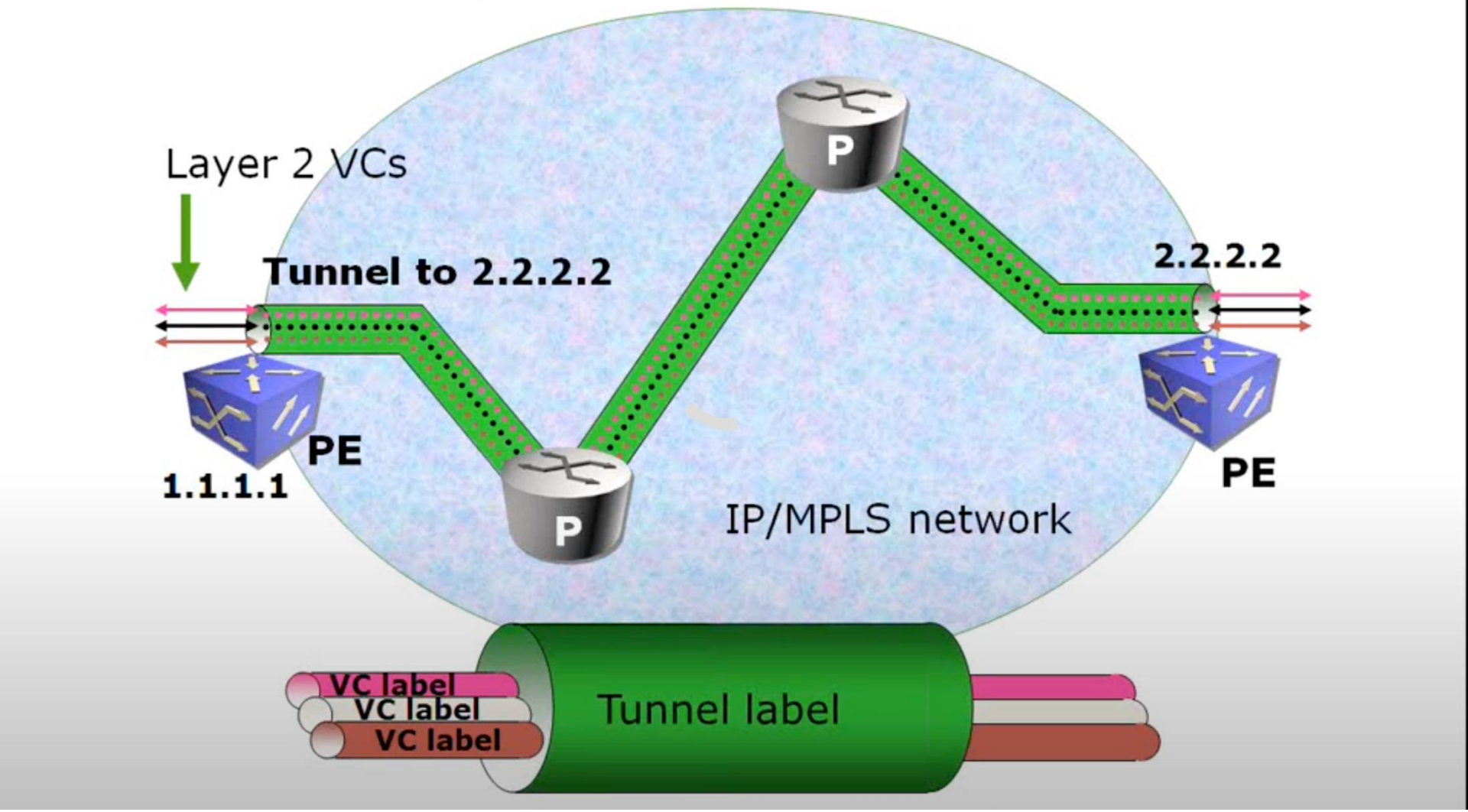


Scalability through label stacking



Tunnel (outer) label determines the path to remote edge
VC (inner) label designates connection at tunnel endpoint

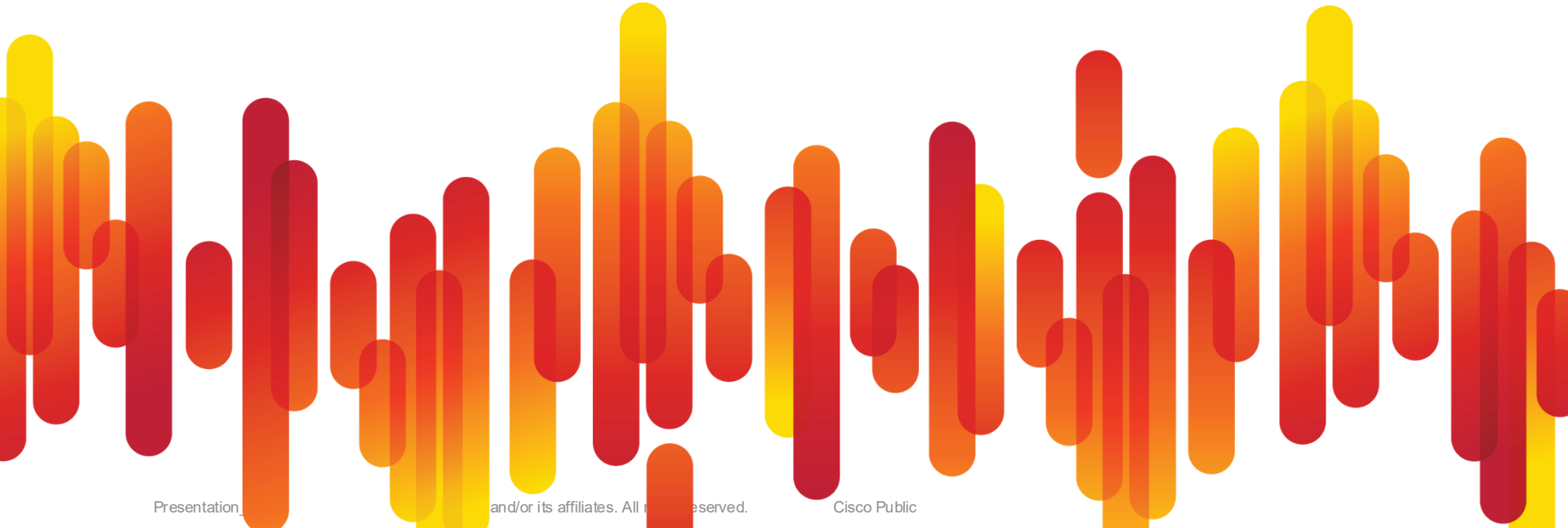
Scalability through label stacking



Tunnel (outer) label determines the path to remote edge
VC (inner) label designates connection at tunnel endpoint

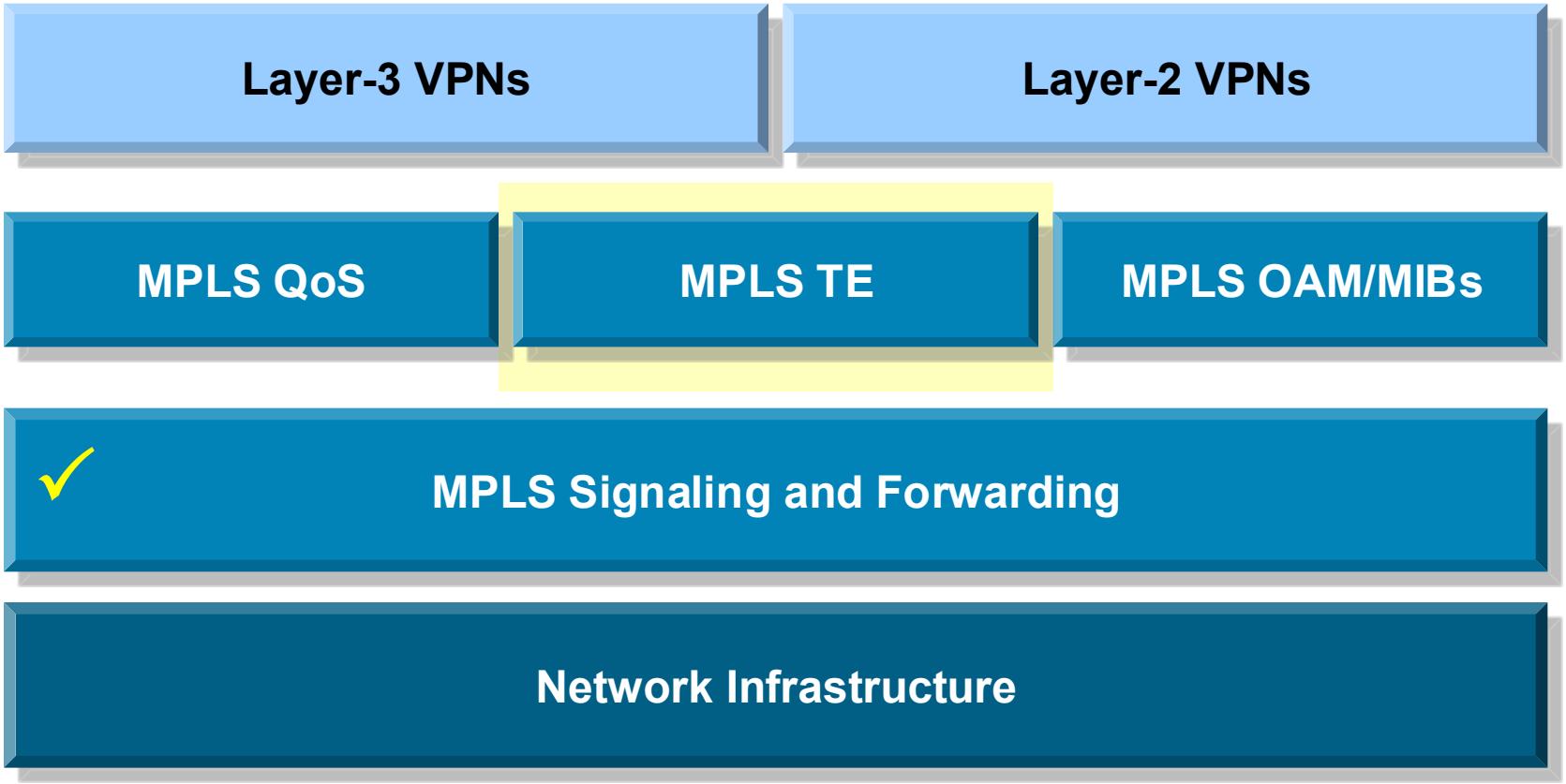
MPLS Traffic Engineering

Technology Overview and Applications



MPLS Technology Framework

- Traffic engineering capabilities for bandwidth management and network failure protection



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

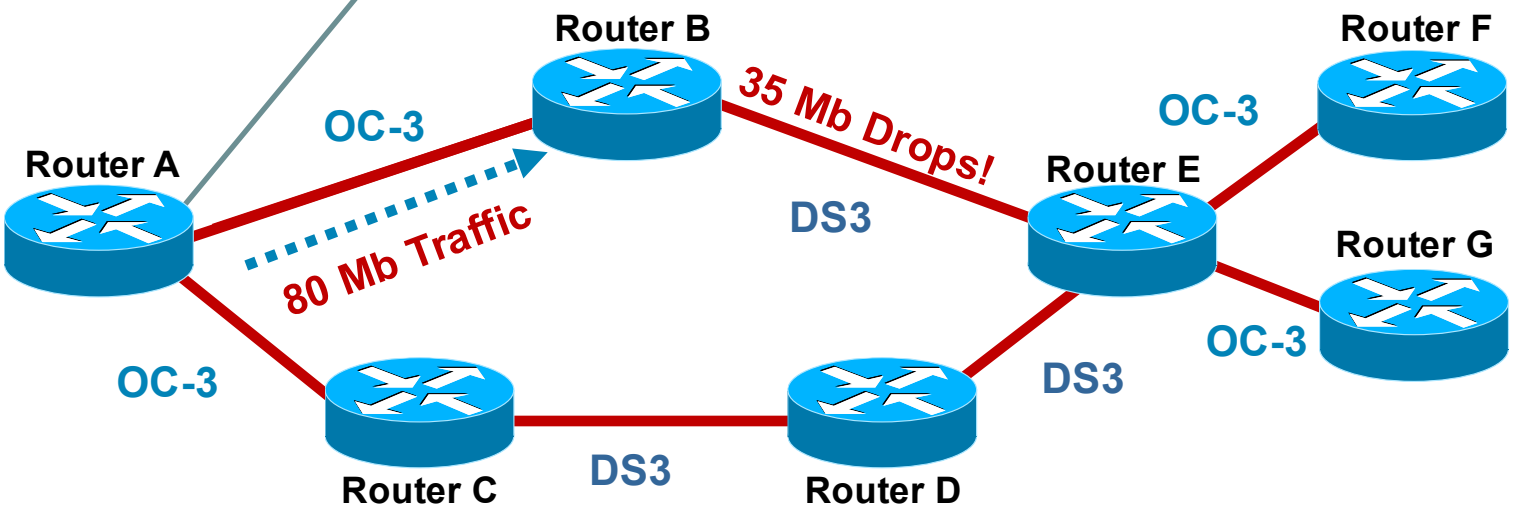
The Problem with Shortest-Path

IP (Mostly) Uses Destination-Based Least-Cost Routing, causing that Alternate Path Under Utilized

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 (45 Mbps), some are OC-3 (155 Mbps)
- Router A has 40M of traffic for router F, 40M of traffic for router G
- Massive (44%) packet loss at router B→router E!

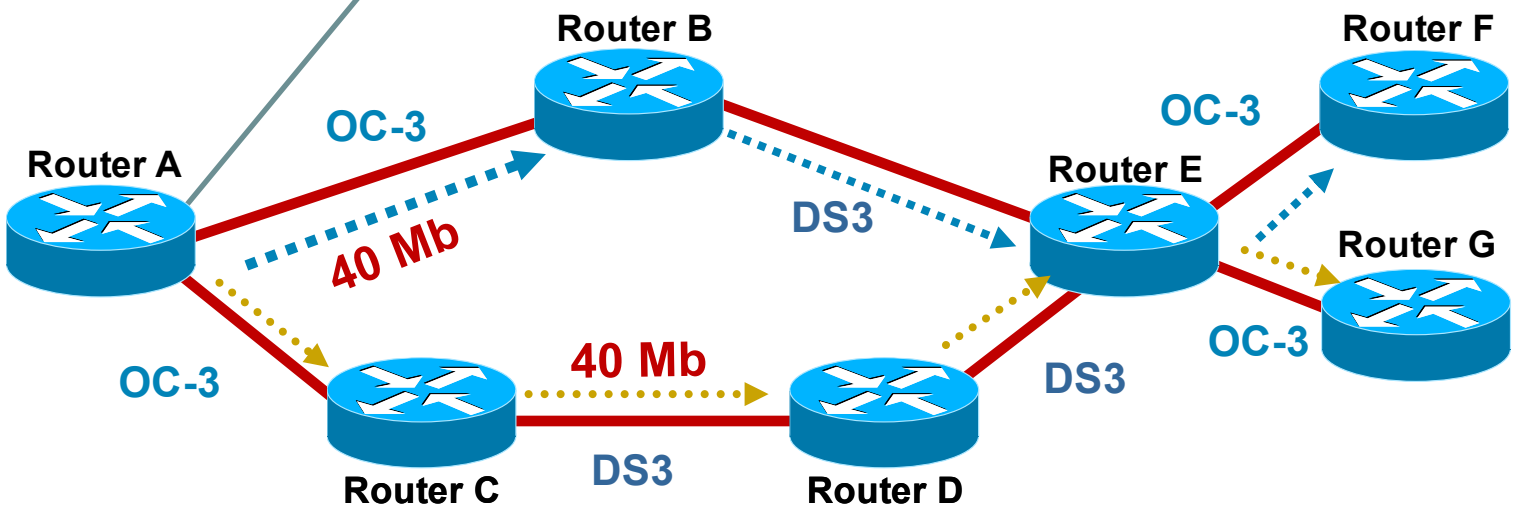
Changing to A->C->D->E won't help



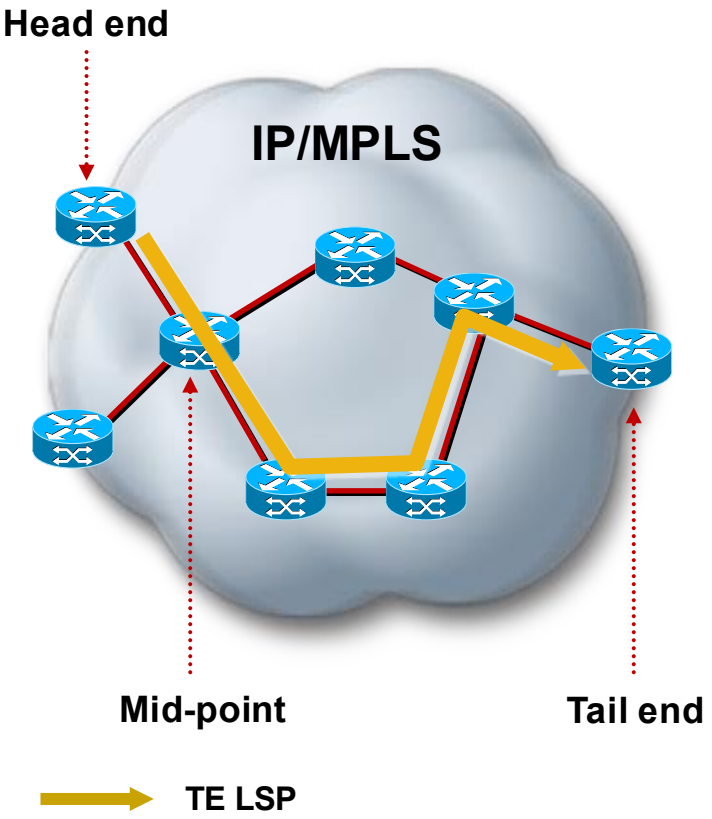
How MPLS TE Solves the Problem

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
G	Tunnel 0	40
F	Tunnel 1	30

- Router A sees all links
- Router A computes paths on properties other than just shortest cost; creation of 2 tunnels
- No link oversubscribed!



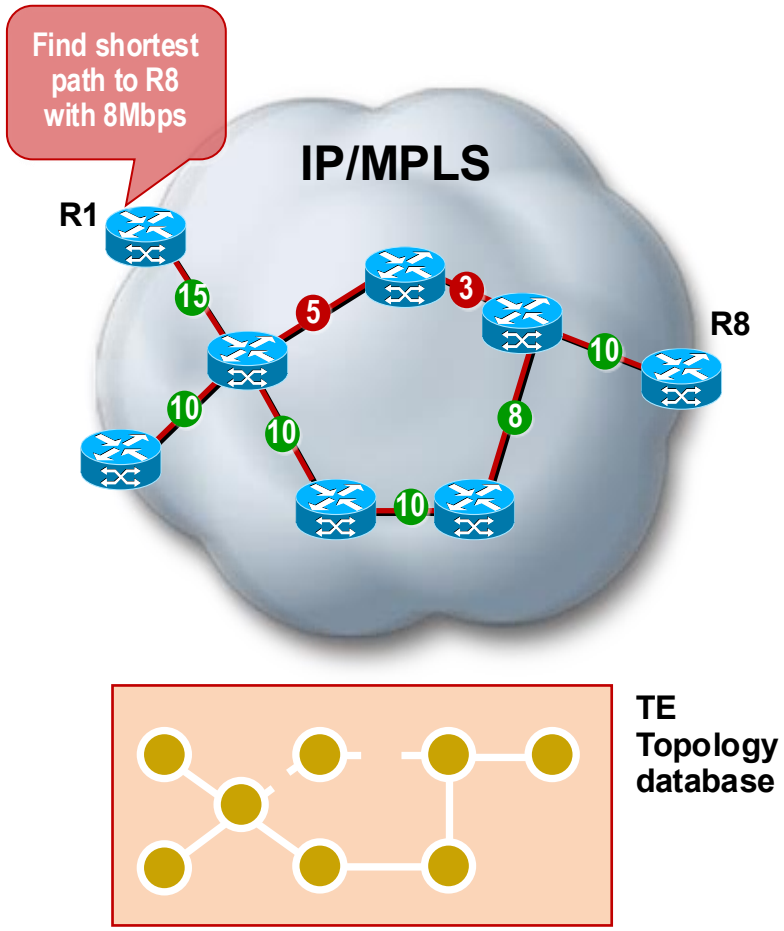
How MPLS TE Works



- Link information Distribution*
  - ISIS-TE
  - OSPF-TE
- **Path Calculation (CSPF)***
- **Path Setup**
- Forwarding Traffic down Tunnel
  - Auto-route
  - Static
  - PBR
  - CBTS / PBTS
  - Forwarding Adjacency
  - Tunnel select

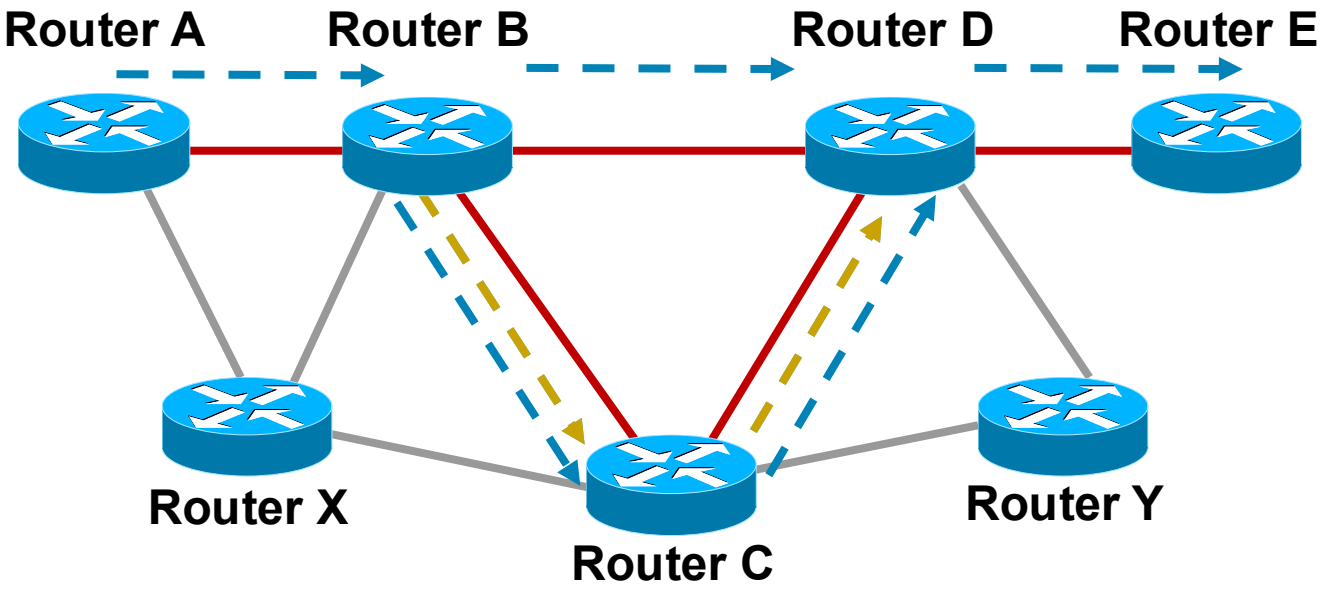
*** Optional**

Path Calculation



- TE nodes can perform constraint-based routing
- Constraints and topology database as input to path computation
- Shortest-path-first algorithm ignores links not meeting constraints
- Tunnel can be signaled once a path is found

MPLS TE FRR - Link Protection



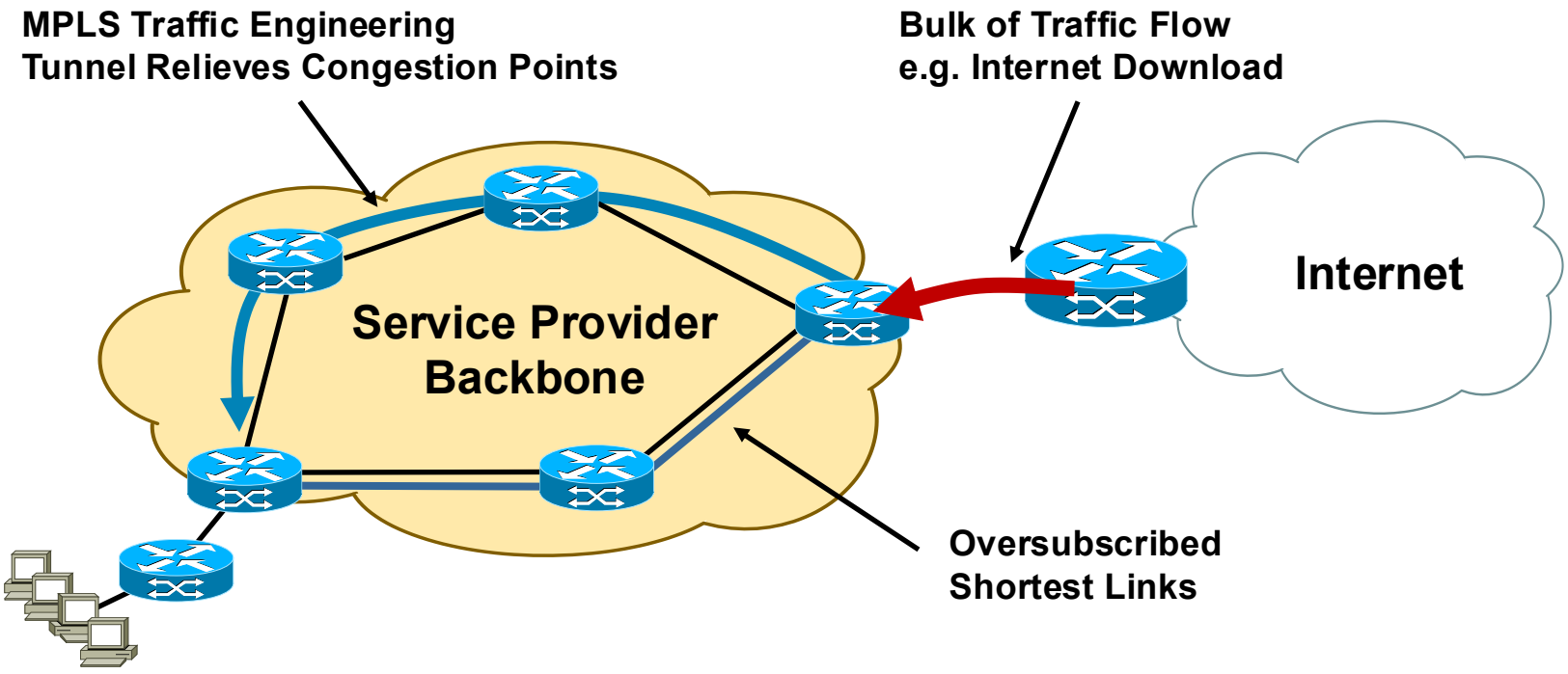
- Primary tunnel: A → B → D → E
- Backup tunnel: B → C → D (preprovisioned)
- Recovery = ~ 50 ms

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

Use Case 1: Tactical TE Deployment

Requirement: Need to Handle Scattered Congestion Points in the Network

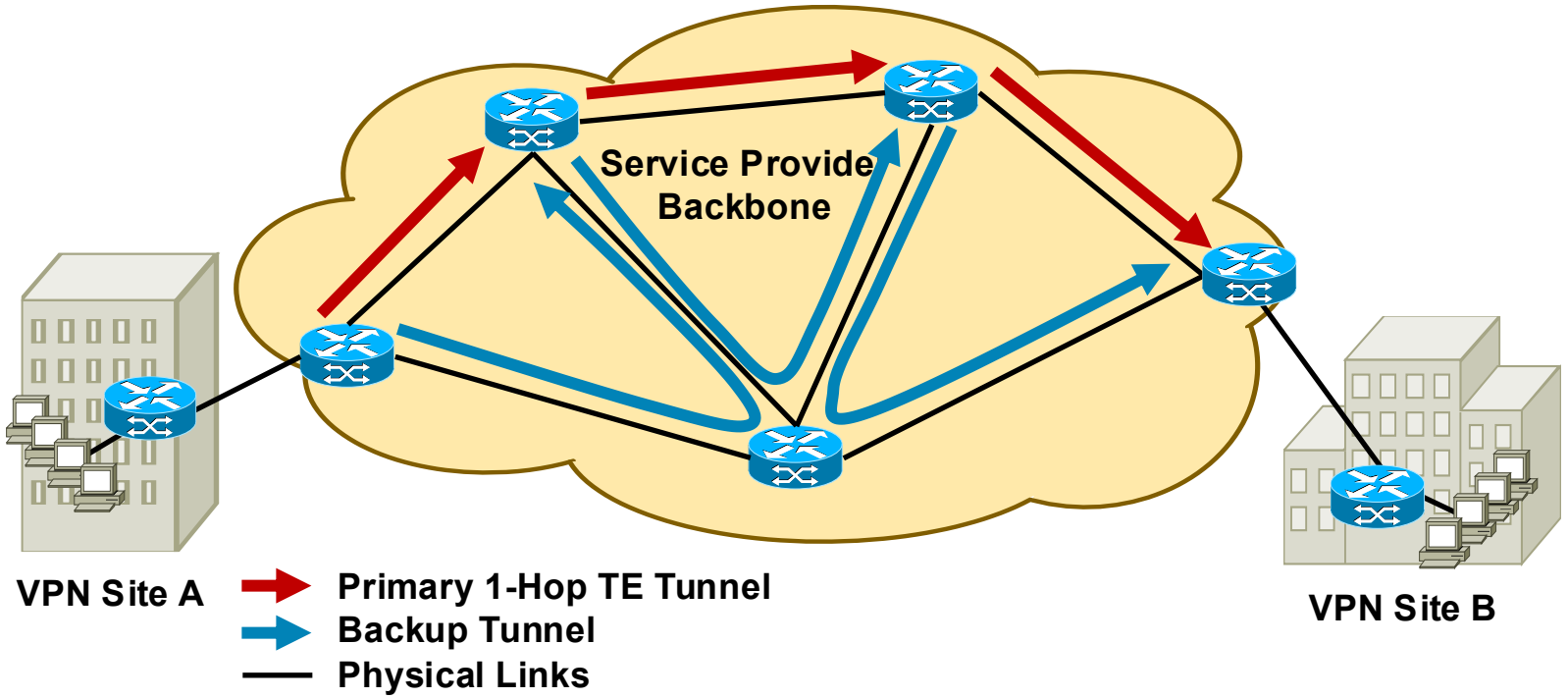
Solution: Deploy MPLS TE on Only Those Nodes that Face Congestion



Use Case 2: End-to-End Resilience

Requirement: Need Protection of all the links — Minimize Packet Loss 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

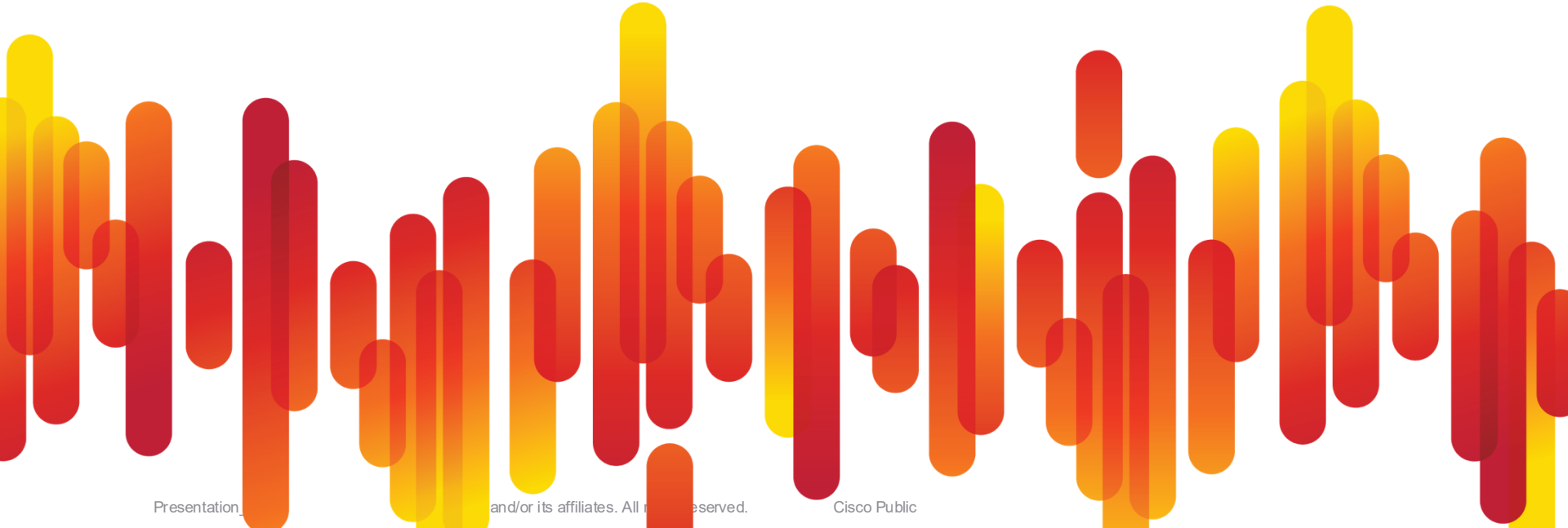


MPLS TE Summary

- MPLS TE can be used to implement traffic engineering to enable enhanced network availability, utilization, and performance
- Enhanced network availability can be implemented via MPLS TE Fast Re-Route (FRR)
  - Link, node, and path protection
  - Automatically route around failed links/nodes; like SONET APS
- Better network bandwidth utilization can be implemented via creation of MPLS TE tunnels using explicit routes
  - Route on the non-shortest path
- MPLS TE can be used for capacity planning by creation of bandwidth-specific tunnels with explicit paths through the network
  - Bandwidth management across links and end-to-end paths

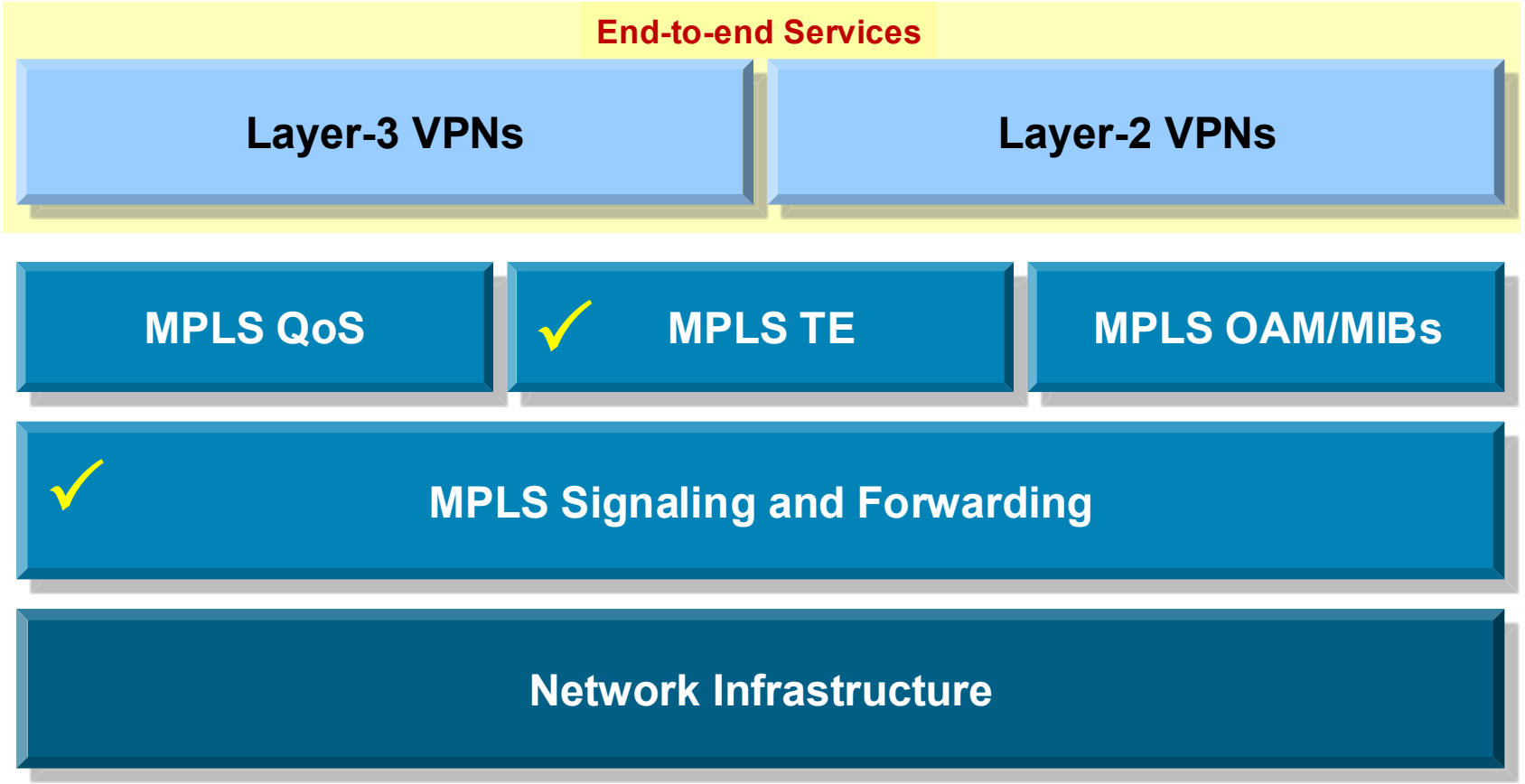
MPLS VPNs

Overviews



MPLS Technology Framework

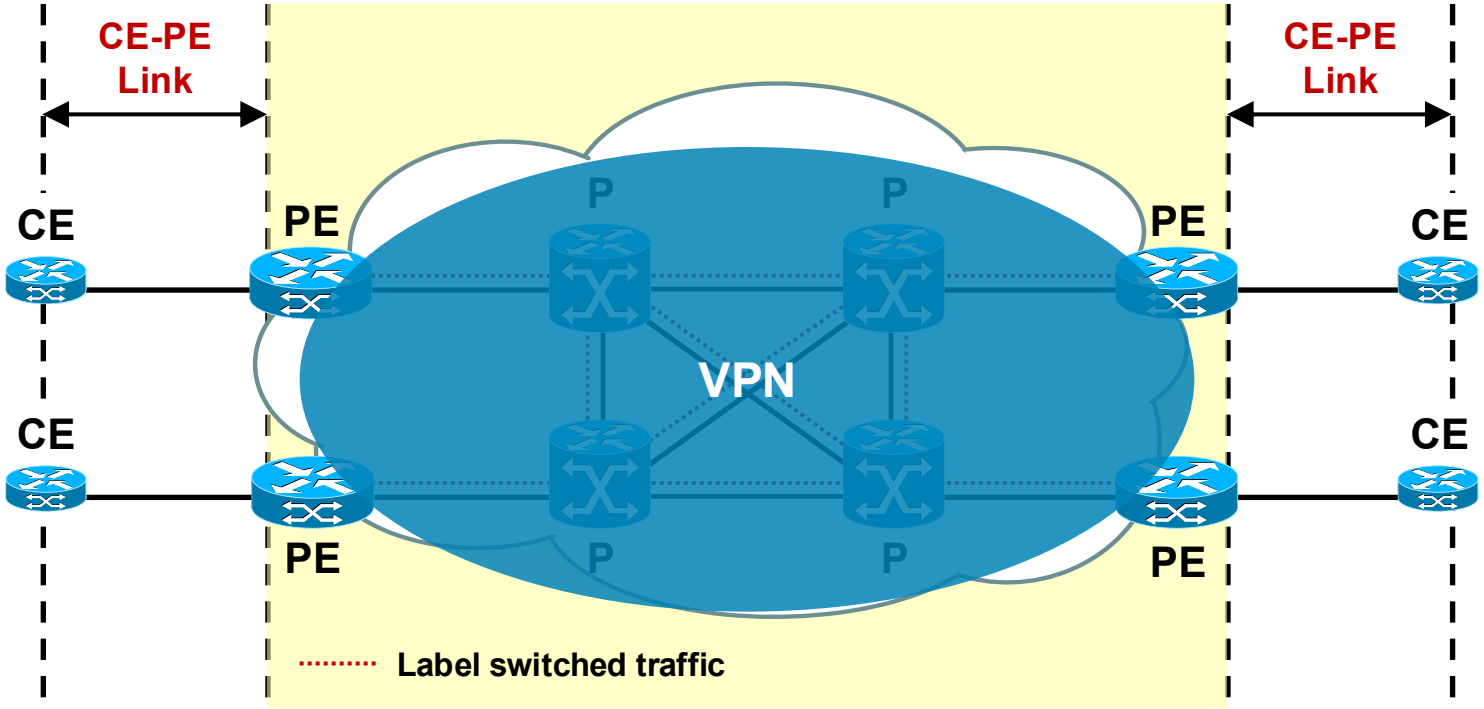
- End-to-end data connectivity services across MPLS networks (from PE to PE)



What Is a Virtual Private Network?

- VPN is a set of sites or groups which are allowed to communicate with each other in a secure way
  - Typically, over a shared public or private network infrastructure
- VPN is defined by a set of administrative policies
  - Policies established by VPN customers themselves (DIY)
  - Policies implemented by VPN service provider (managed/unmanaged)
- Different inter-site connectivity schemes possible
  - Ranging from complete to partial mesh
- Sites/users may be either within the same or in different organizations
  - VPN can be either intranet or extranet
- Sites/users 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

MPLS VPN Example



- CE-PE link

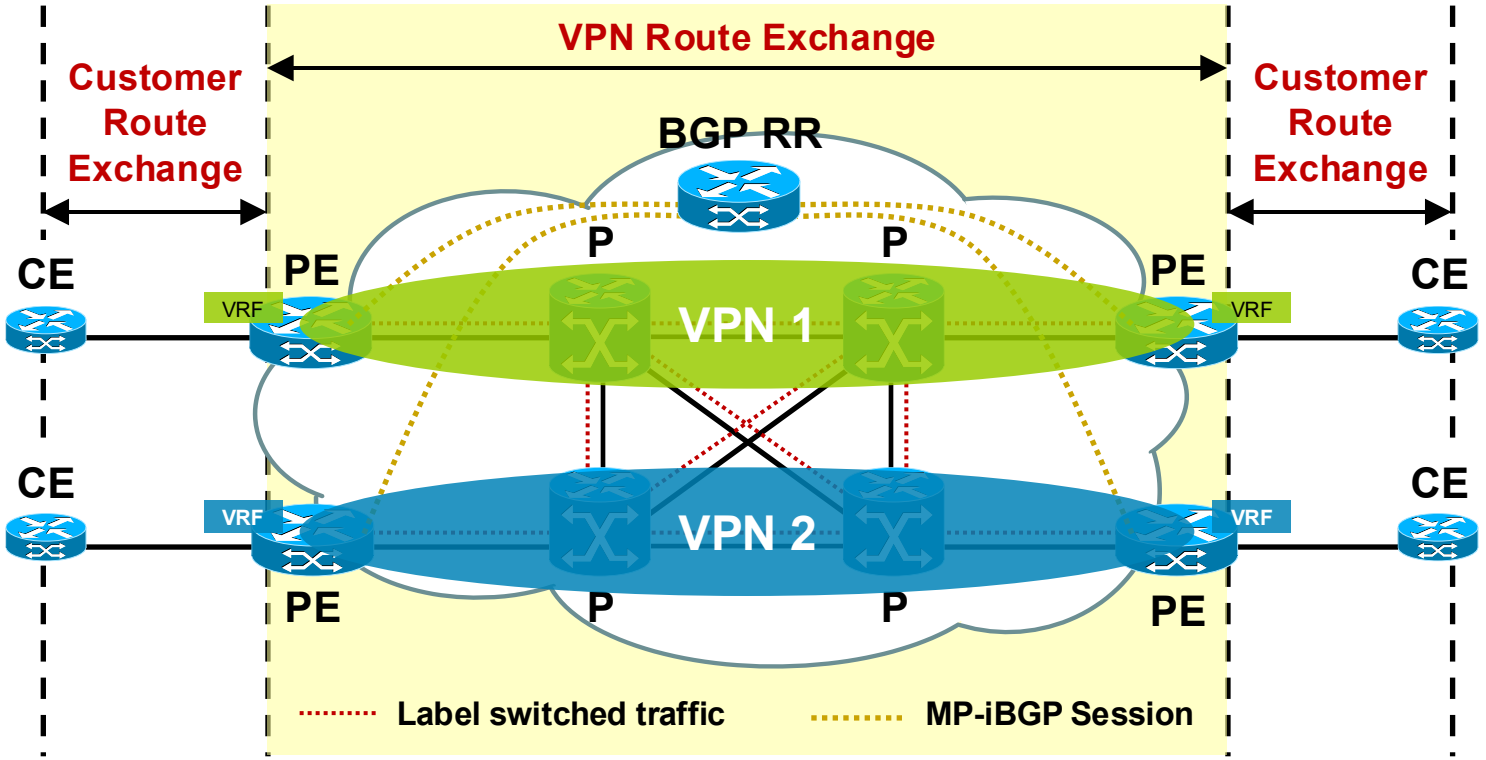
  Connect customer network to SP network; layer-2 or layer-3

- VPN

  Dedicated secure connectivity over shared infrastructure

VPN Route Distribution

Exchange of VPN Policies among PE Routers



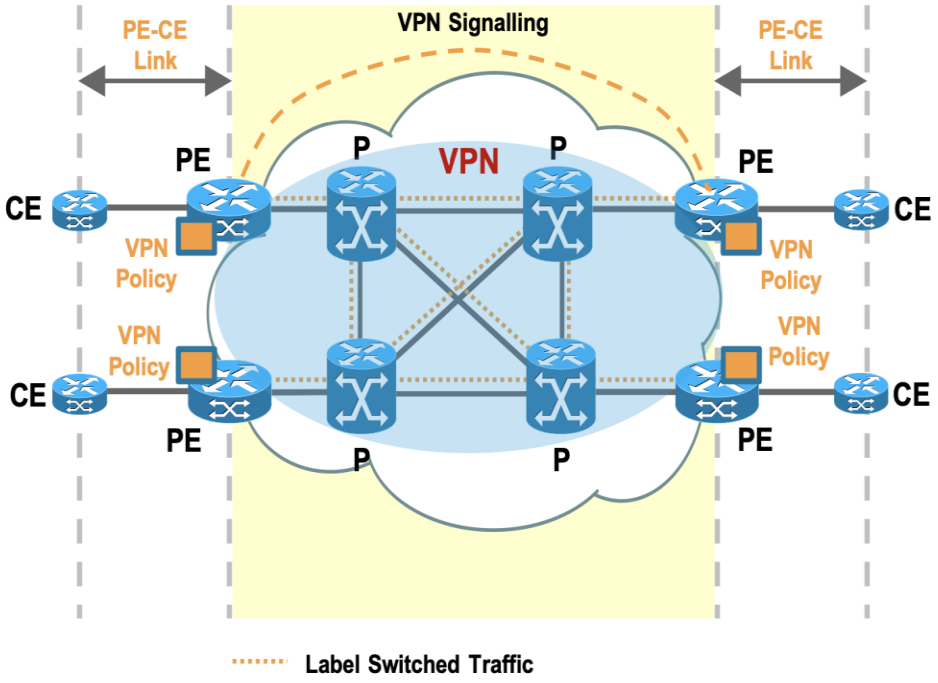
- Full or partial mesh of BGP sessions among all PE routers

  Typically BGP Route Reflector (RR) used for improved scalability

MPLS VPN Example

Basic Building Blocks

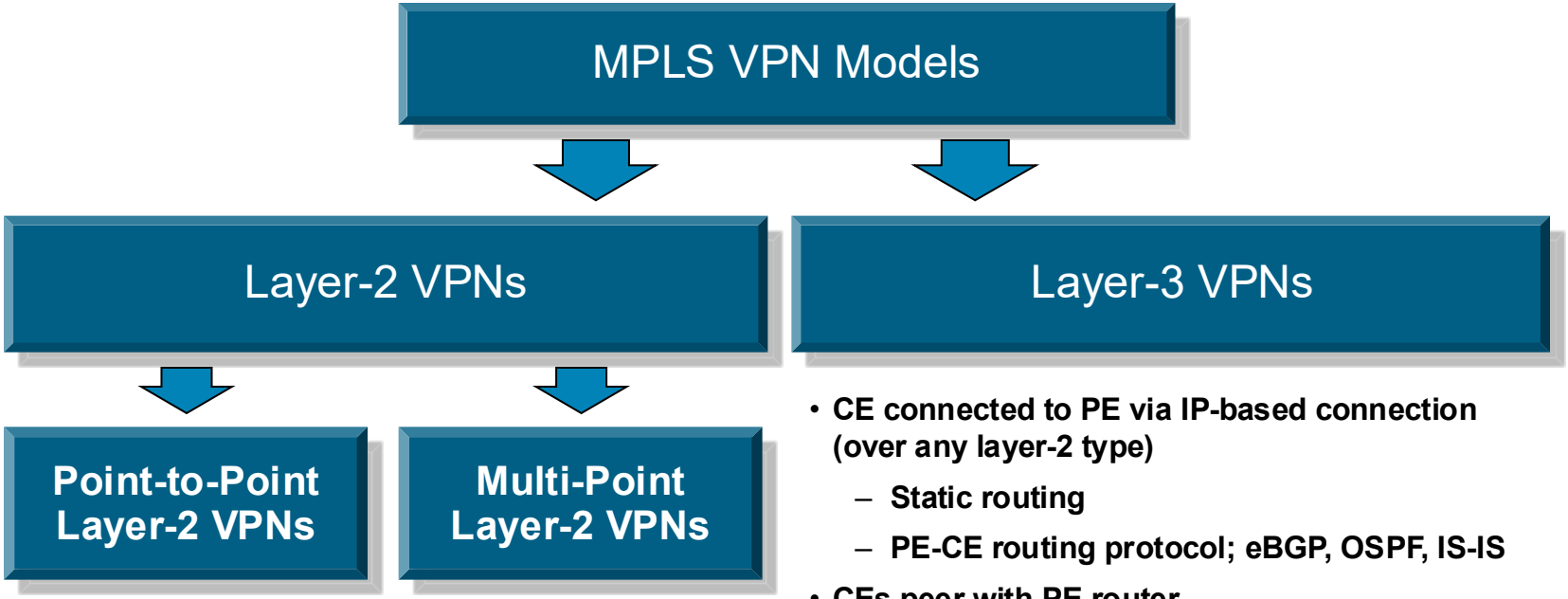
- **VPN policies**
  - Configured on PE routers (manual operation)
- **VPN signalling**
  - Between PEs
  - Exchange of VPN policies
- **VPN traffic forwarding**
  - Additional VPN-related MPLS label encapsulation
- **PE-CE link**
  - Connects customer network to MPLS network; either layer-2 or layer-3



MPLS VPN Benefits

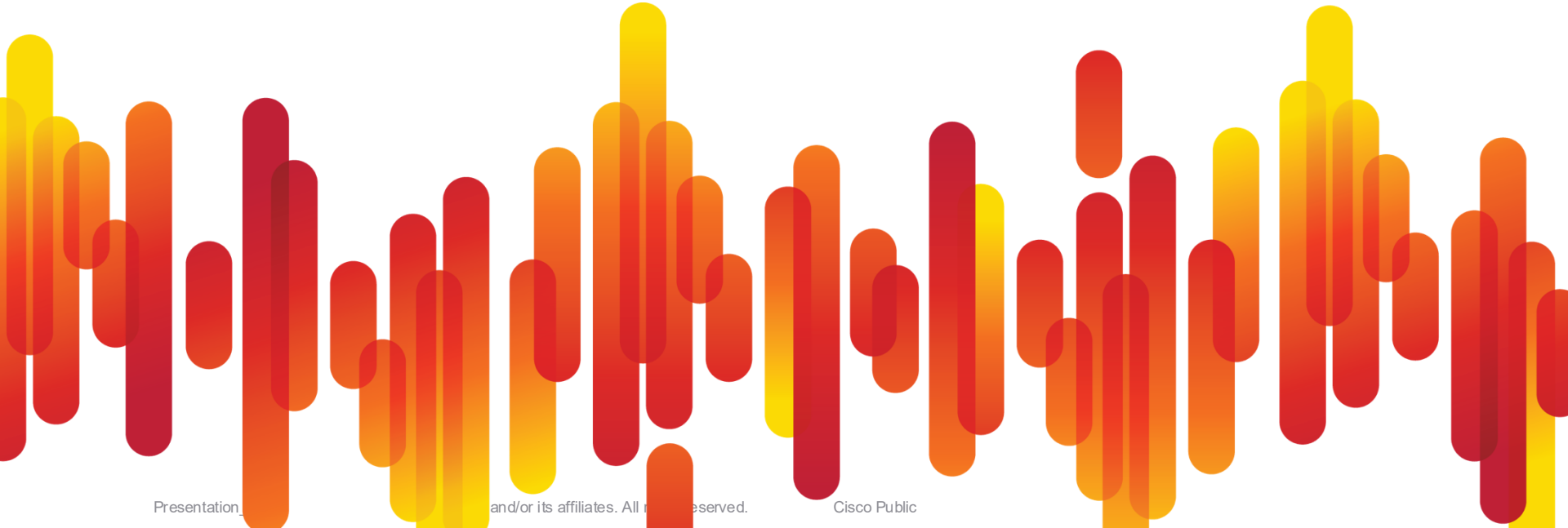
- SP/Carrier perspective
  - Reduce costs (CAPEX)
    - Leverage same network for multiple services and customers
    - Migrate legacy networks onto single converged network
  - Reduce costs (OPEX)
    - Easier service enablement; only edge node configuration
- Enterprise/end-user perspective
  - Enables site/campus network segmentation
    - Allows for dedicated and secture connectivity for users, applications, etc.
  - Enables easier setup of WAN connectivity
    - Easier configuration of site-to-site WAN connectivity (for L3VPN); only one WAN connection needed

MPLS VPN Options



MPLS Layer-3 VPNs

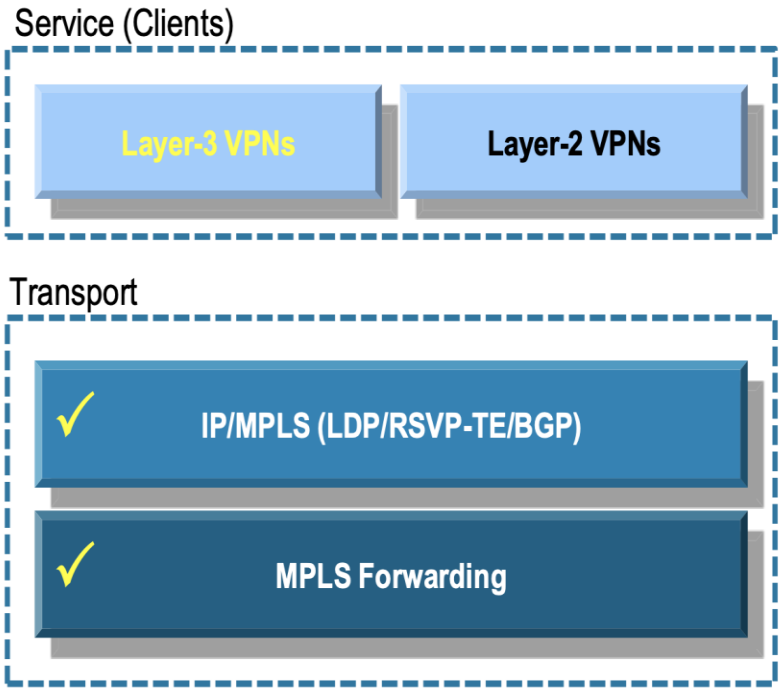
Technology Overview and Applications



MPLS Layer-3 Virtual Private Networks

Topics

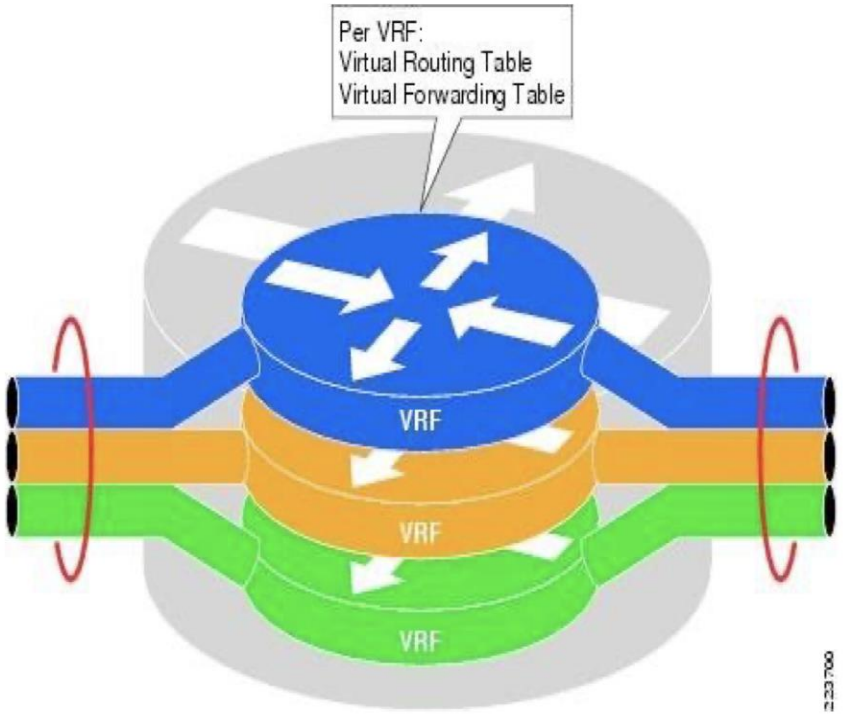
- Technology components
- VPN control plane mechanisms
- VPN forwarding plane
- Deployment use cases
  - Business VPN services
  - Network segmentation
  - Data Centre access



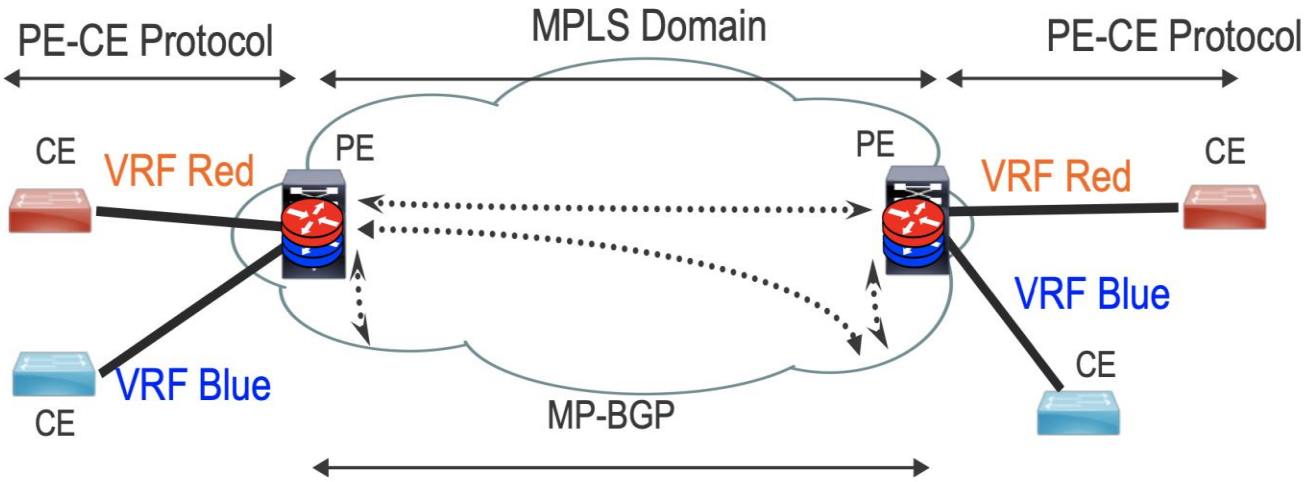
Virtual Routing and Forwarding Instance - VRF

Virtual Routing Table and Forwarding Separate to Customer Traffic

- Logical routing context within the same PE device
- Unique to a VPN
- Allows for customer overlapping IP addresses
- Deployment use cases
  - Business VPN services
  - Network segmentation
  - Data Centre access

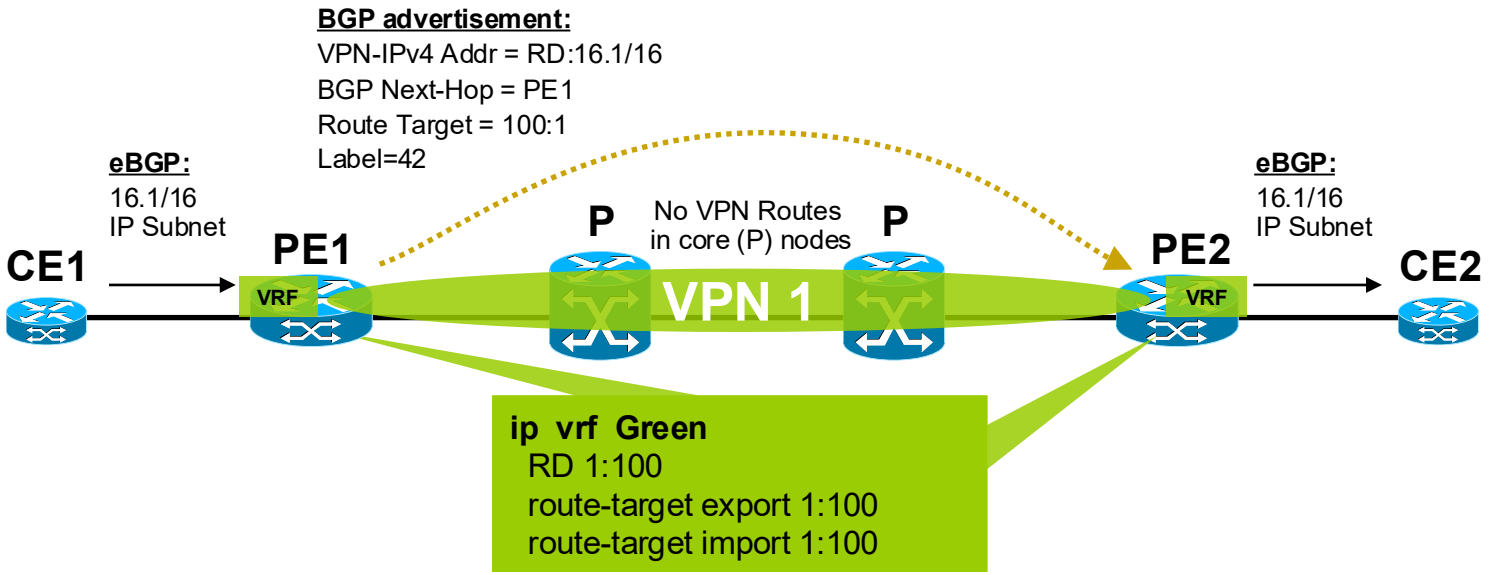


Virtual Routing and Forwarding Instance



- Typically one VRF created for each customer VPN on PE router
- VRF associated with one or more customer interfaces
- VRF has its own instance of routing table (RIB) and forwarding table (CEF)
- VRF has its own instance for PE-CE configured routing protocols

VPN Control Plane Processing



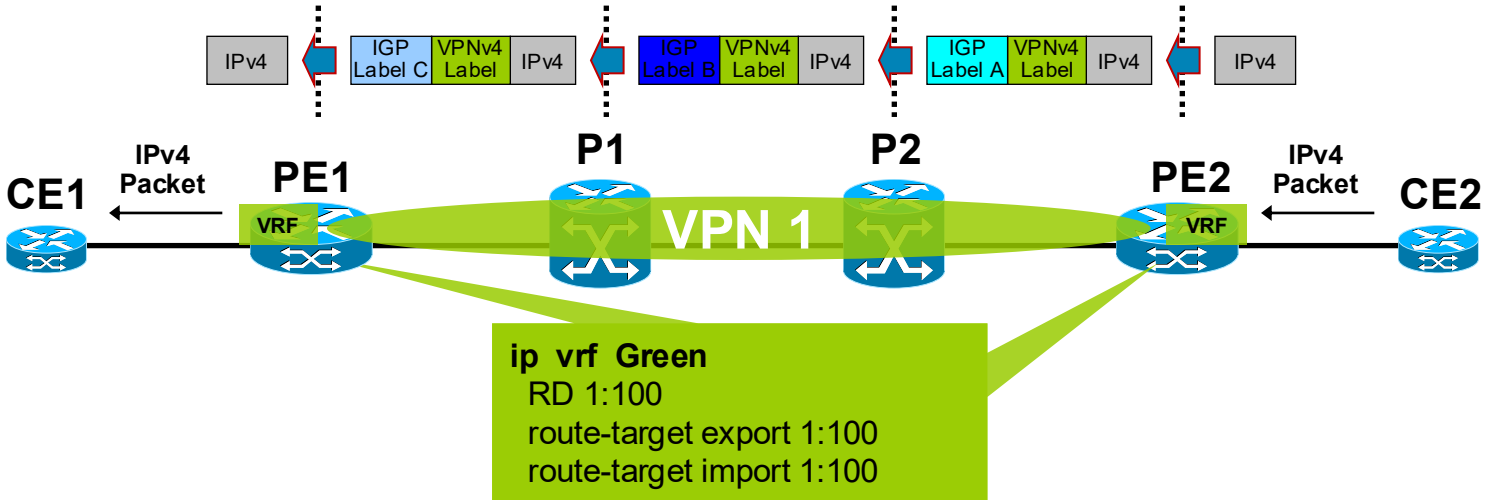
Make customer routes unique:

- Route Distinguisher (RD): 8-byte field, VRF parameters; unique value assigned by a provider to each VPN to make different VPN routes unique
- VPNv4 address: RD+VPN IP prefix

Processing Steps:

1. CE1 redistribute IPv4 route to PE1 via eBGP.
2. PE1 allocates VPN label for prefix learnt from CE1 to create unique VPNv4 route
3. PE1 redistributes VPNv4 route into BGP, it sets itself as a next hop and relays VPN site routes to PE2
4. PE2 receives VPNv4 route and, via processing in local VRF (green), it redistributes original IPv4 route to CE2.

VPN Forwarding Plane Processing



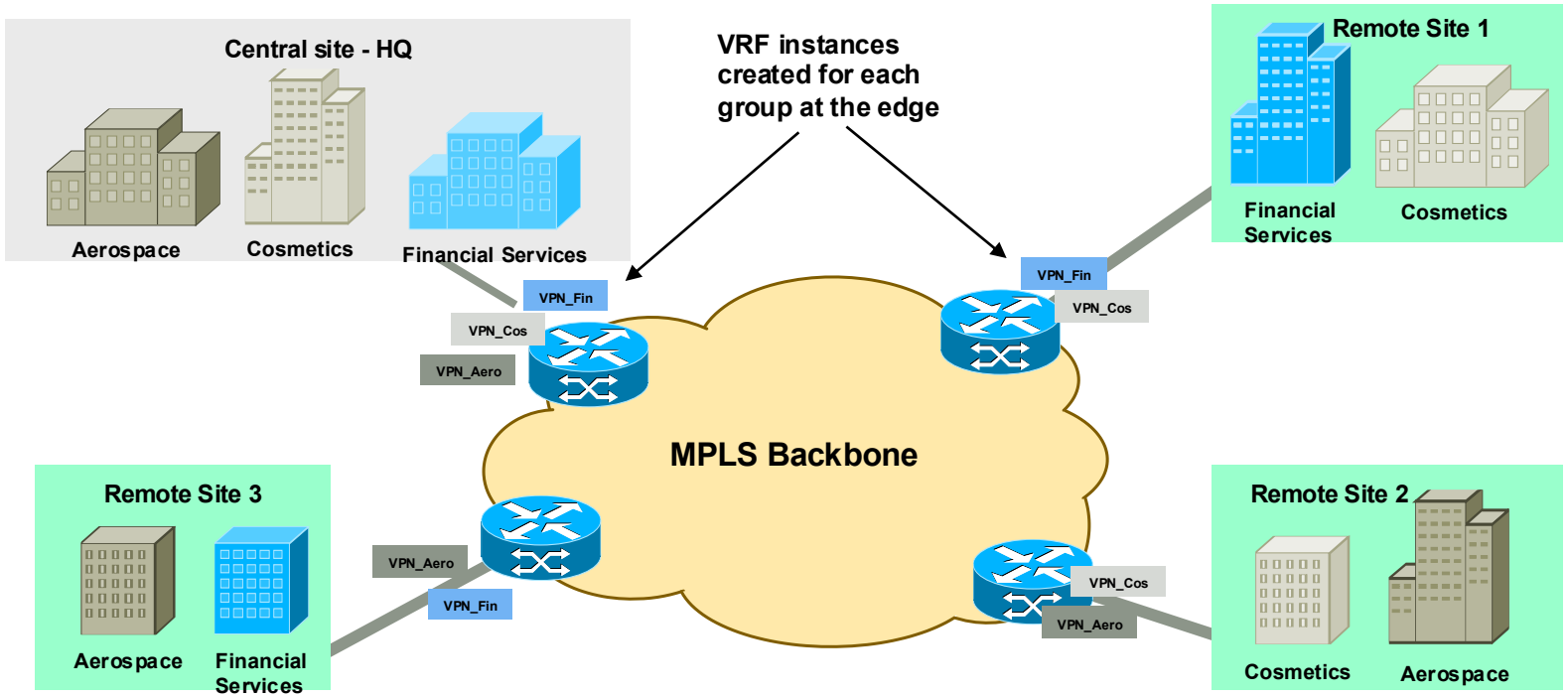
Processing Steps:

1. CE2 forwards IPv4 packet to PE2.
2. PE2 imposes pre-allocated VPN label (learned via BGP) to IPv4 packet received from CE2.
3. PE2 imposes outer IGP label (learned via LDP) and forwards labeled packet to next-hop P-router P2.
4. P-routers P2 and P1 swap outer IGP label and forward label packet to PE1.
5. Router PE1 strips VPN label and forwards IPv4 packet to CE1.

Use Case 1: Traffic Separation

Requirement: Need to ensure data separation between Aerospace, Cosmetics and Financial Services, while leveraging a shared infrastructure

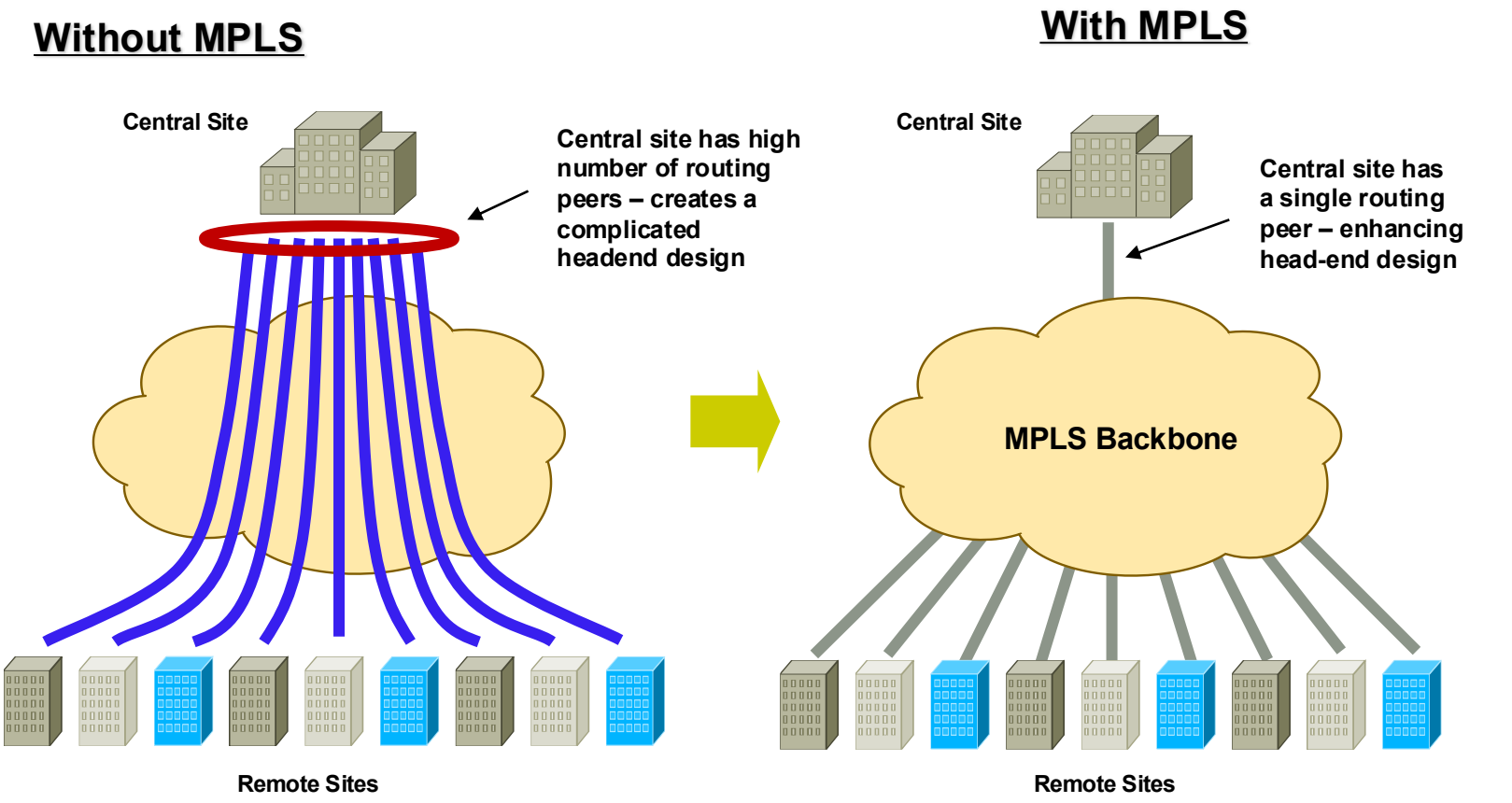
Solution: Create MPLS VPN for each group



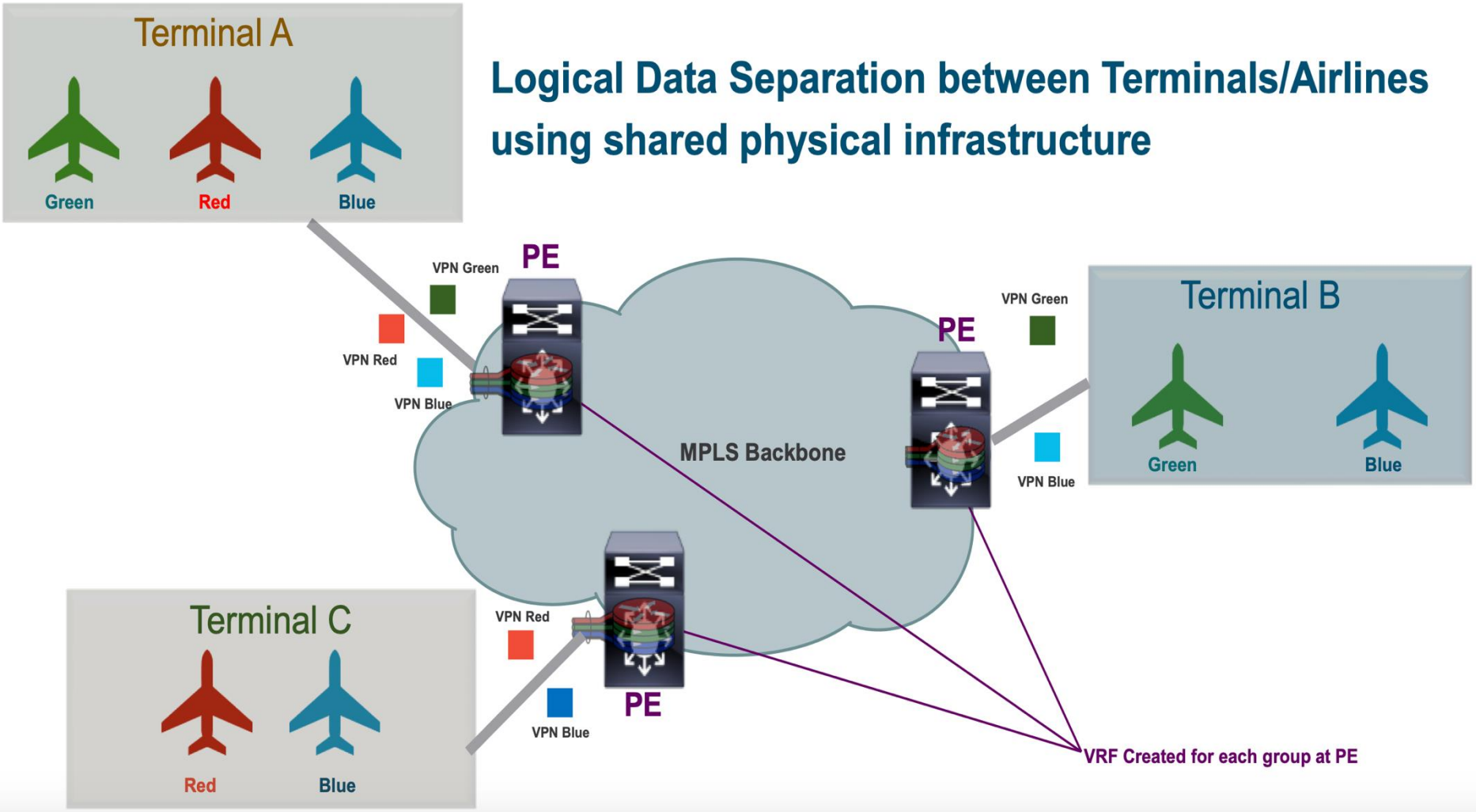
Use Case 2: Simplify Central Site Design

Requirement: To ease the scale and design of a central site

Solution: Implement MPLS Layer 3 VPNs, which reduces the number of routing peers of the central site

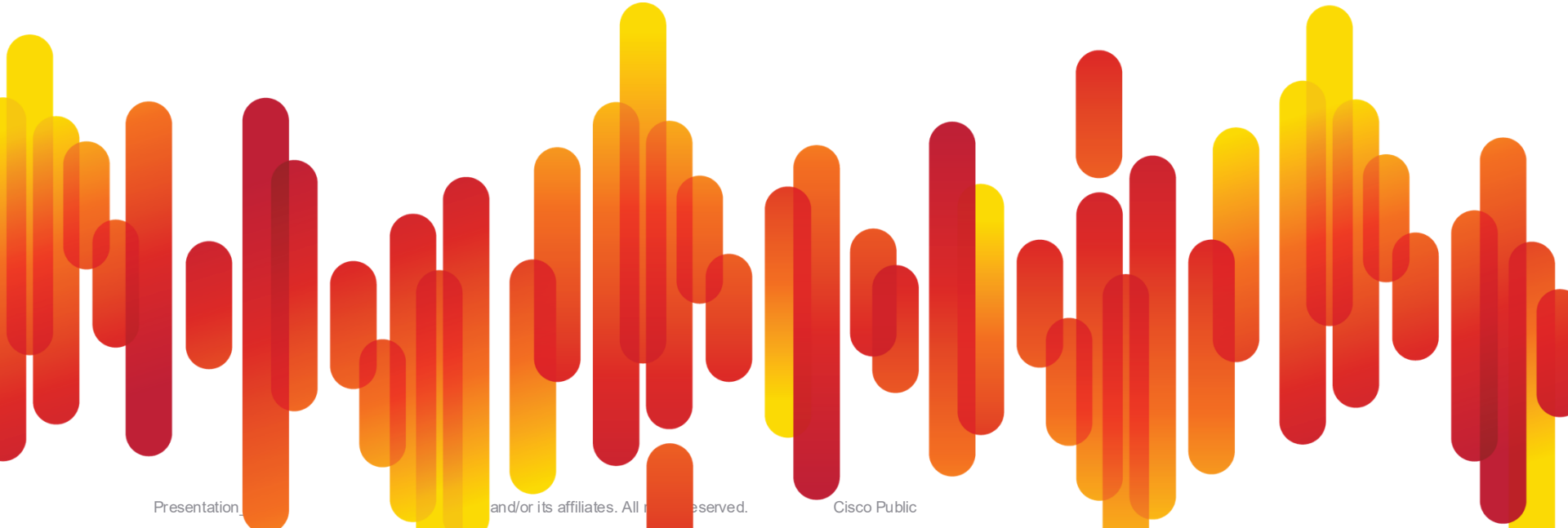


Use Case 3: Traffic Separation with L3 VPN in an Airport

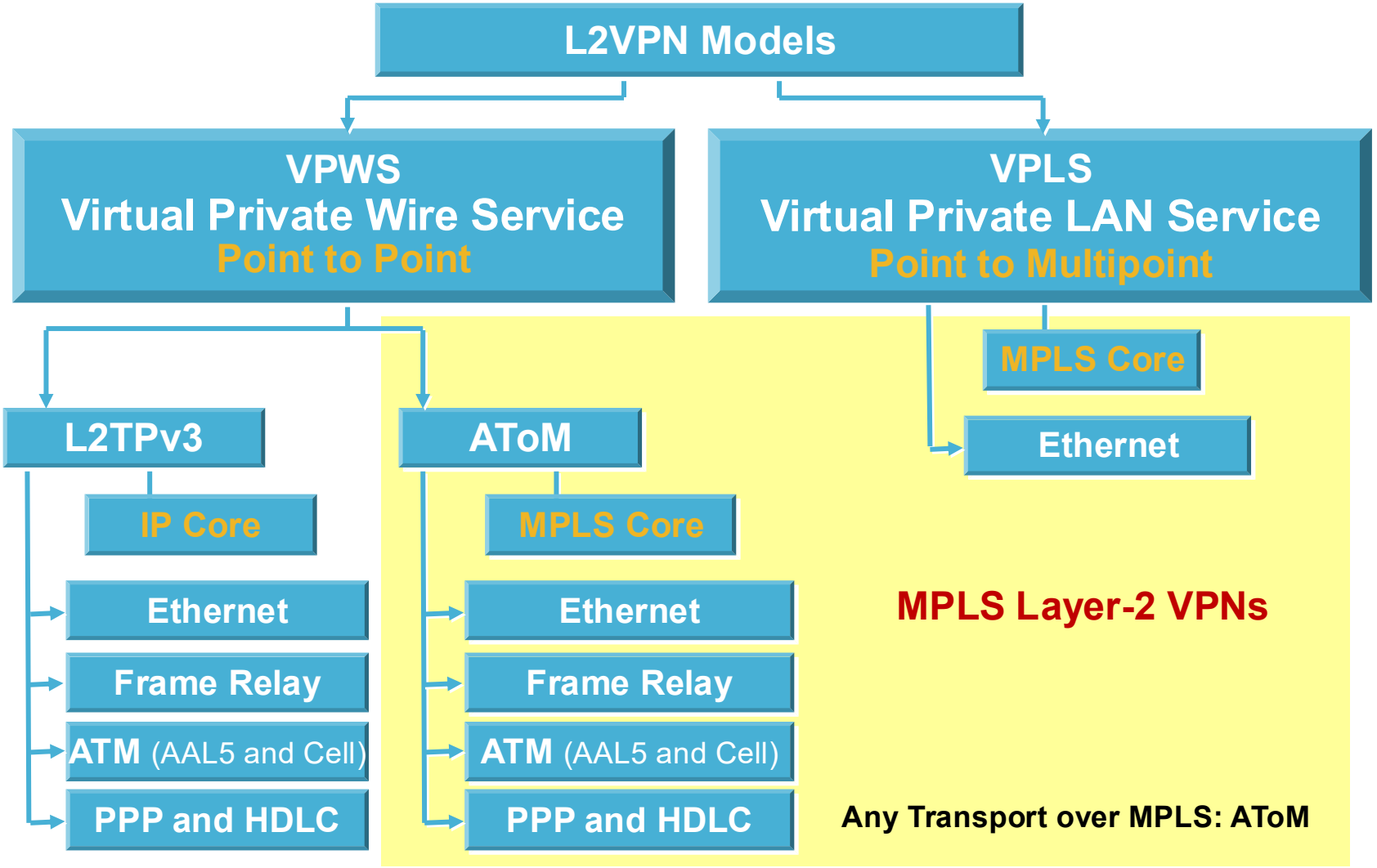


MPLS Layer-2 VPNs

Technology Overview and Applications

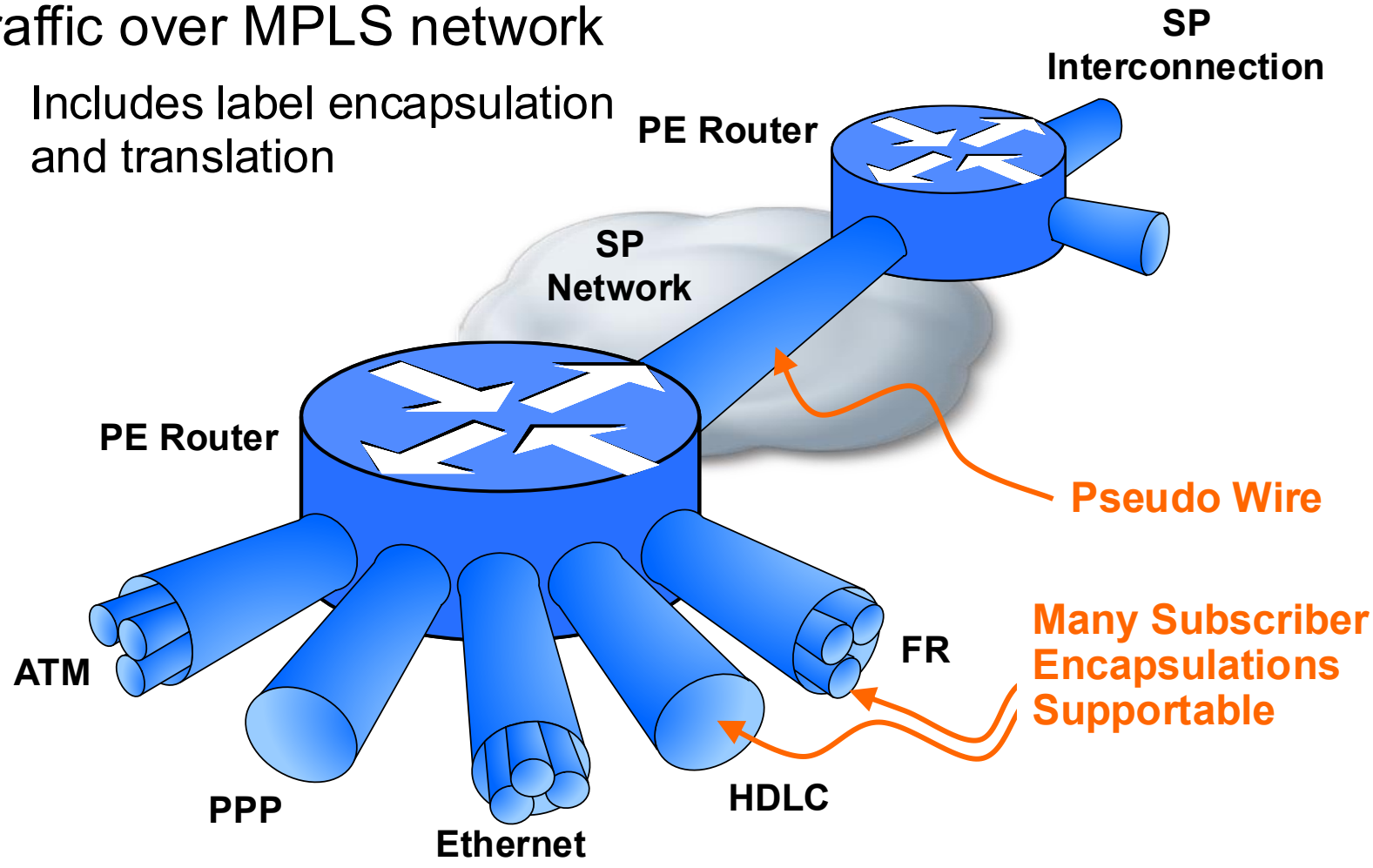


L2VPN Options



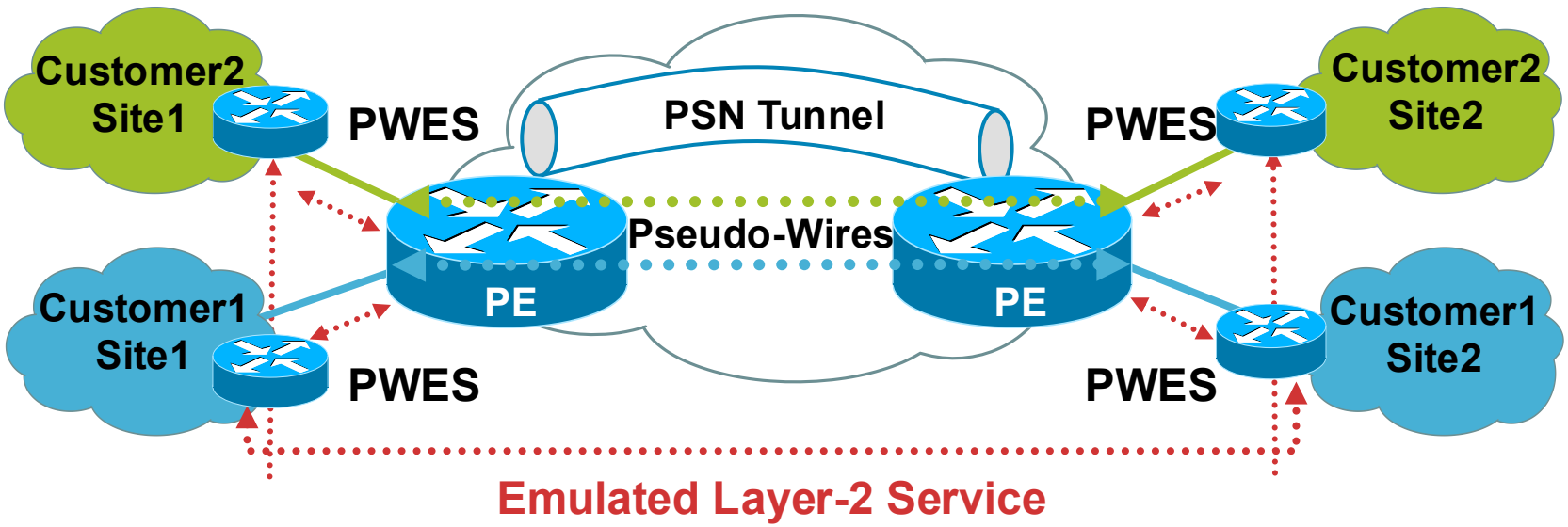
Layer-2 VPN Overview

- Enables transport of any Layer-2 traffic over MPLS network
  - Includes label encapsulation and translation



Any Transport over MPLS Architecture

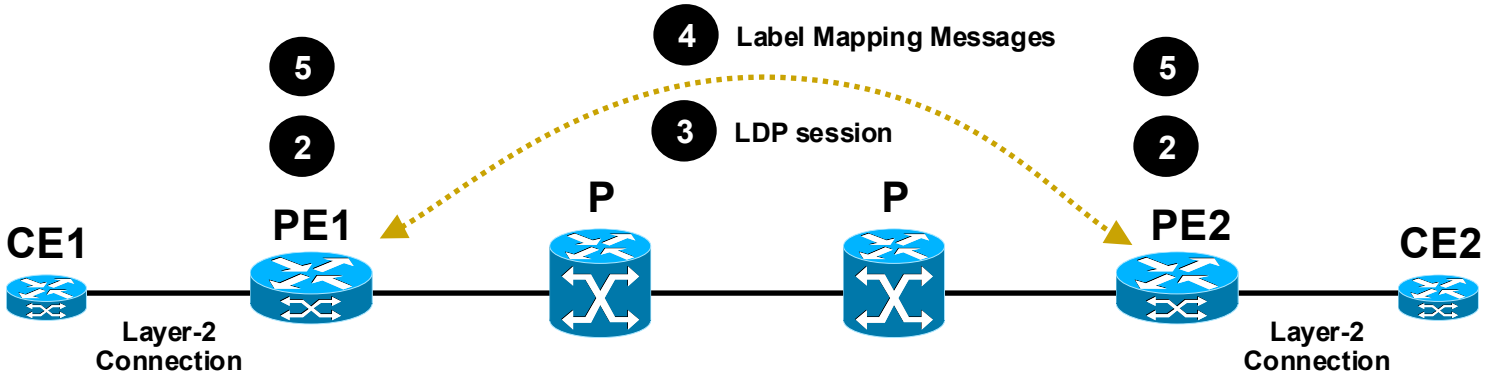
- Based on IETF's Pseudo-Wire (PW) Reference Model
- PW is a connection (tunnel) between 2 PE Devices, which connects 2 PW End-Services
  - PW connects 2 Attachment Circuits (ACs)
  - Bi-directional (for p2p connections)
  - Use of PW/VC label for encapsulation



AToM Technology Components

- PE-CE link
  - Referred to as Attachment Circuit (AC)
  - Can be any type of layer-2 connection (e.g., FR, Ethernet)
- AToM Control Plane
  - Targeted LDP (Label Distribution Protocol) Session
    - Virtual Connection (VC)-label negotiation, withdrawal, error notification
- AToM Forwarding Plane
  - 2 labels used for encapsulation
  - Outer tunnel (LDP) label
    - To get from ingress to egress PE using MPLS LSP
  - Inner de-multiplexer (VC) label
    - To identify L2 circuit (packet) encapsulated within tunnel label

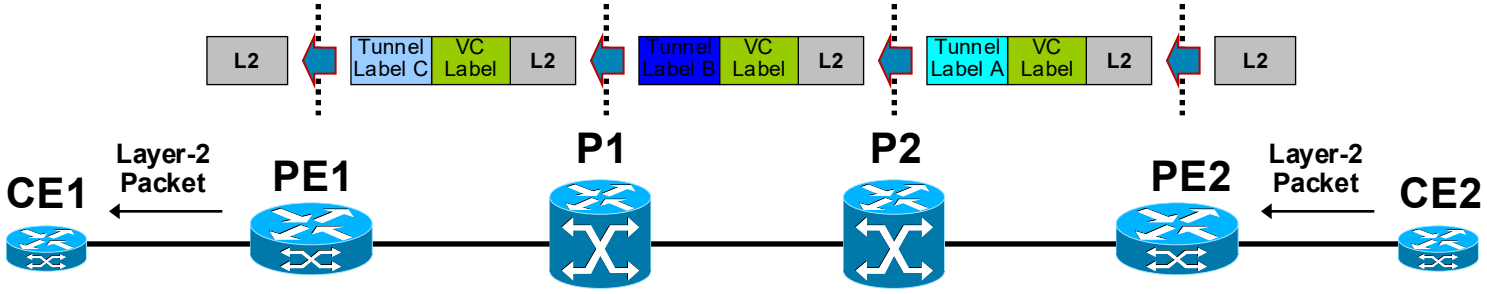
AToM Control Plane Processing



Processing Steps (for both P1 and P2):

1. CE1 and CE2 are connected to PE routers via layer-2 connections
2. Via CLI, a new virtual circuit cross-connect is configured, connecting customer interface to manually provided VC ID with target remote PE
3. New targeted LDP session between PE routers established, in case one does not already exist
4. PE binds VC label with customer layer-2 interface and sends label-mapping message to remote PE over LDP session
5. Remote PE receives LDP label binding message and matches VC ID with local configured cross-connect

AToM Forwarding Plane Processing



Processing Steps:

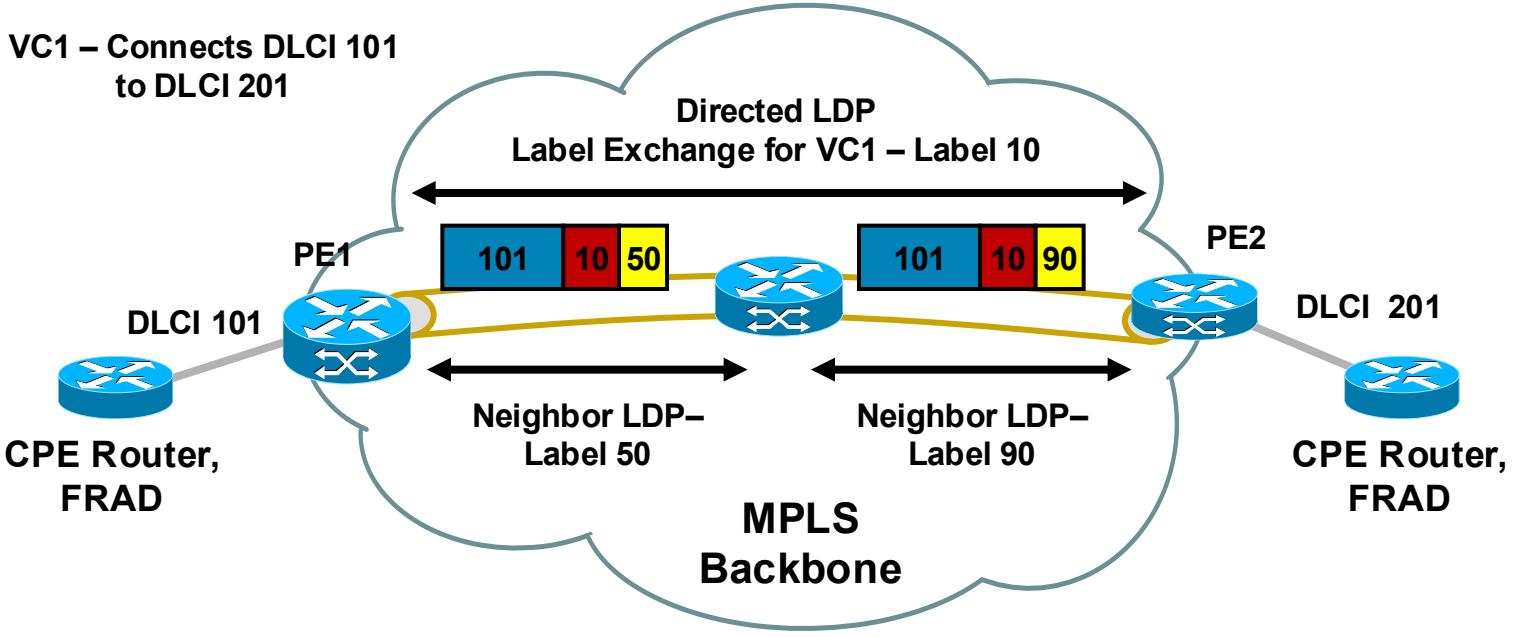
1. CE2 forwards layer-2 packet to PE2.
2. PE2 imposes VC (inner) label to layer-2 packet received from CE2 and optionally a control word as well (not shown).
3. PE2 imposes Tunnel outer label and forwards packet to P2.
4. P2 and P1 router forwards packet using outer (tunnel) label.
5. Router PE1 strips Tunnel label and, based on VC label, layer-2 packet is forwarded to customer interface to CE1, after VC label is removed

   In case control word is used, new layer-2 header is generated first.

Use Case: L2 Network Interconnect

Requirement: Need to create connectivity between remote customer sites, currently interconnected via Frame Relay WAN connectivity. Only point-to-point connectivity required.

Solution: Interconnect AToM PW between sites, enabling transparent Frame Relay WAN connectivity.



Layer-2 VPN Summary

- Enables transport of any Layer-2 traffic over MPLS network
- The most used L2 VPN model is AToM - Any Transfer over MPLS
- AToM suited for implementing transparent point-to-point connectivity / service between Layer-2 end-points
- AToM - Layer-2 VPN forwarding based on Pseudo Wires (PW), which use VC label for L2 packet encapsulation
  - LDP used for PW signaling