



MPLS

Jim Raine

jraine@cisco.com



- **Background**

1

Original Motivation of MPLS

- **Allow core routers/networking devices to switch packets based on some simplified header**
- **Provide a highly scalable mechanism that was topology driven rather than flow driven**
- **Leverage hardware so that simple forwarding paradigm can be used**
- **It has evolved a long way from the original goal**
 - Hardware became better and looking up longest best match was no longer an issue**
 - By associating labels with prefixes, groups of sites or bandwidth paths or light paths new services such as MPLS VPNs and traffic engineering, GMPLS were now possible**

What Is MPLS?

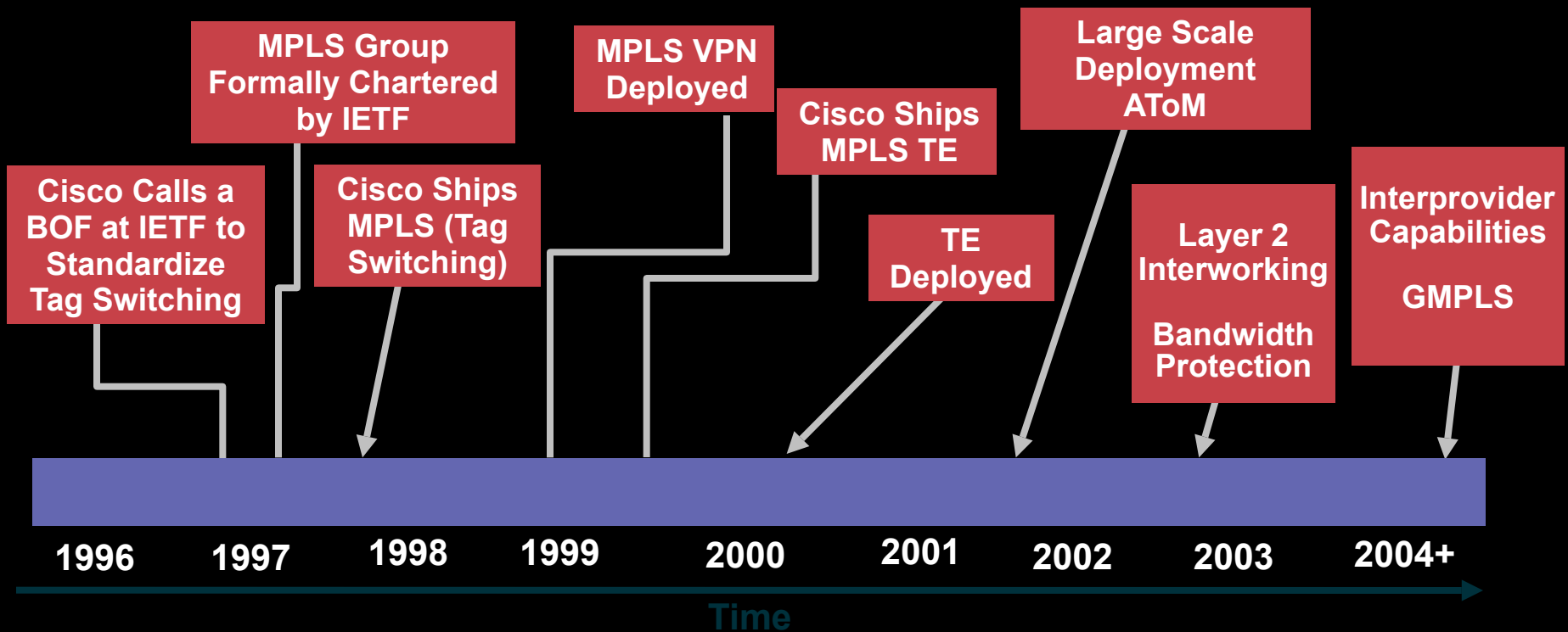
- **Multi Protocol Label Switching**
- **MPLS is an efficient encapsulation mechanism**
- **Uses “labels” appended to packets (IP packets, AAL5 frames) for transport of data**
- **MPLS packets can run on other Layer 2 technologies such as ATM, FR, PPP, POS, Ethernet**
- **Other Layer 2 technologies can be run over an MPLS network**
- **Labels can be used as designators**
 - For example—IP prefixes, ATM VC, or a bandwidth guaranteed path
- **MPLS is a technology for delivery of IP services**

Terminology

- **LSR** - Label switch router
- **LSP** - Label switched path
 - The chain of labels that are swapped at each hop to get from one LSR to another
- **VRF** - VPN routing and forwarding
 - Mechanism in Cisco IOS® used to build per-interface RIB and FIB
- **MP-BGP** - Multiprotocol BGP
- **PE** - Provider edge router interfaces with CE routers
- **P or PC** - Provider (core) router, without knowledge of VPN
- **VPNv4** - Address family used in BGP to carry MPLS-VPN routes
- **RD** - Route distinguisher
 - Distinguish same network/mask prefix in different VRFs
- **RT** - Route target
 - Extended community attribute used to control import and export policies of VPN routes
- **TE** - Traffic Engineering
- **LFIB** - Label forwarding information base
- **FIB** - Forwarding information base

Evolution of MPLS

- From tag switching
- Proposed in IETF—later combined with other proposals from IBM (ARIS), Toshiba (CSR)





• TECHNOLOGY BASICS

2

Unicast Routing Protocols

Cisco.com

- **OSPF, IS-IS, BGP are needed in the network**
- **They provide reachability**
- **Label distribution protocols distribute labels for prefixes advertised by unicast routing protocols using**

Either a dedicated Label Distribution Protocol (LDP)

Extending existing protocols like BGP to distribute labels

MPLS Control and Forwarding Planes

Cisco.com

- **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



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

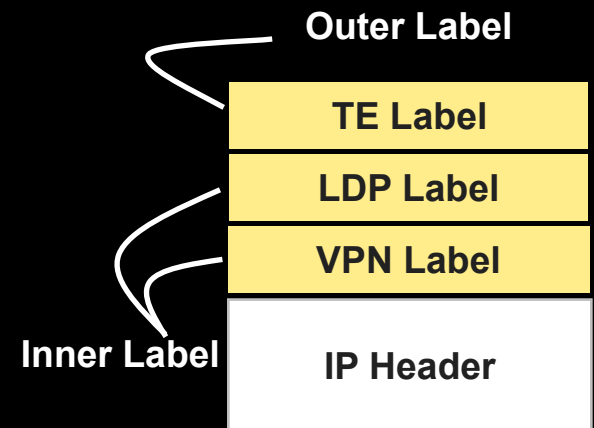
- Outer label used to route/switch the MPLS packets in the network
- Last label in the stack is marked with EOS bit
- Allows building services such as

MPLS VPNs

Traffic engineering and fast re-route

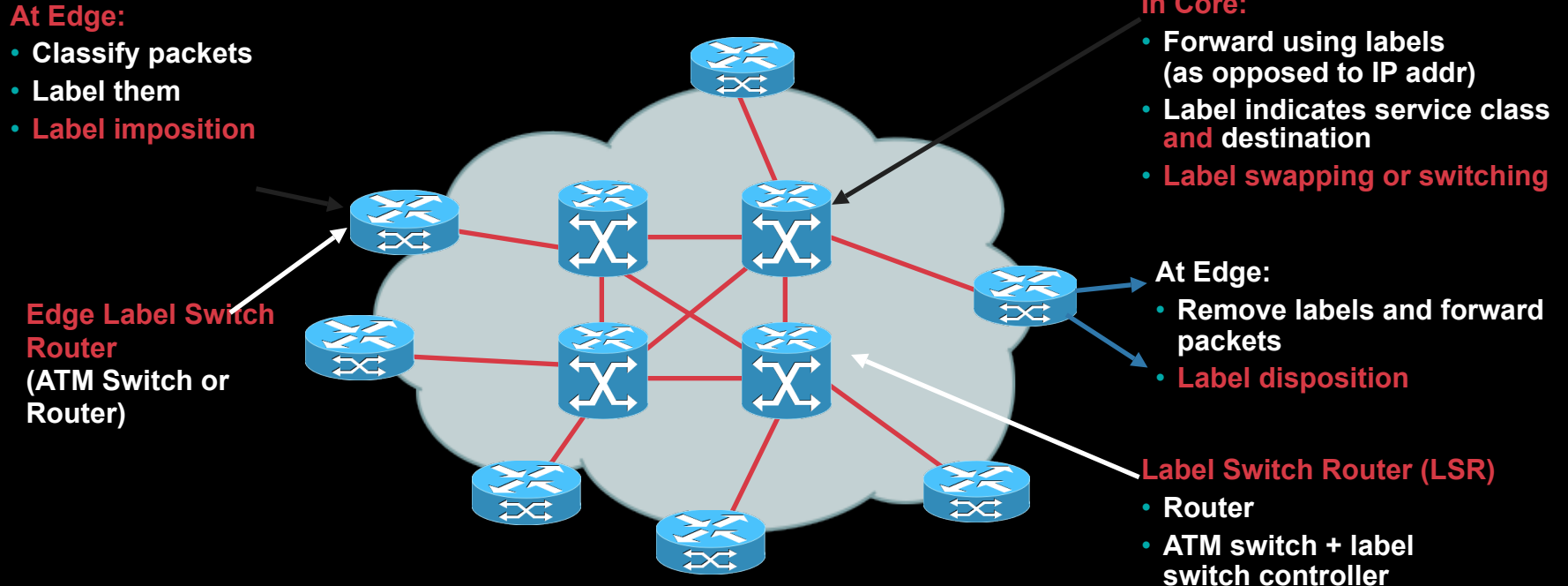
VPNs over traffic engineered core

Any transport over MPLS



MPLS Concepts

Cisco.com



Label Distribution Protocol

- Create new services via flexible classification
- Provide the ability to setup bandwidth guaranteed paths
- Enable ATM switches to act as routers

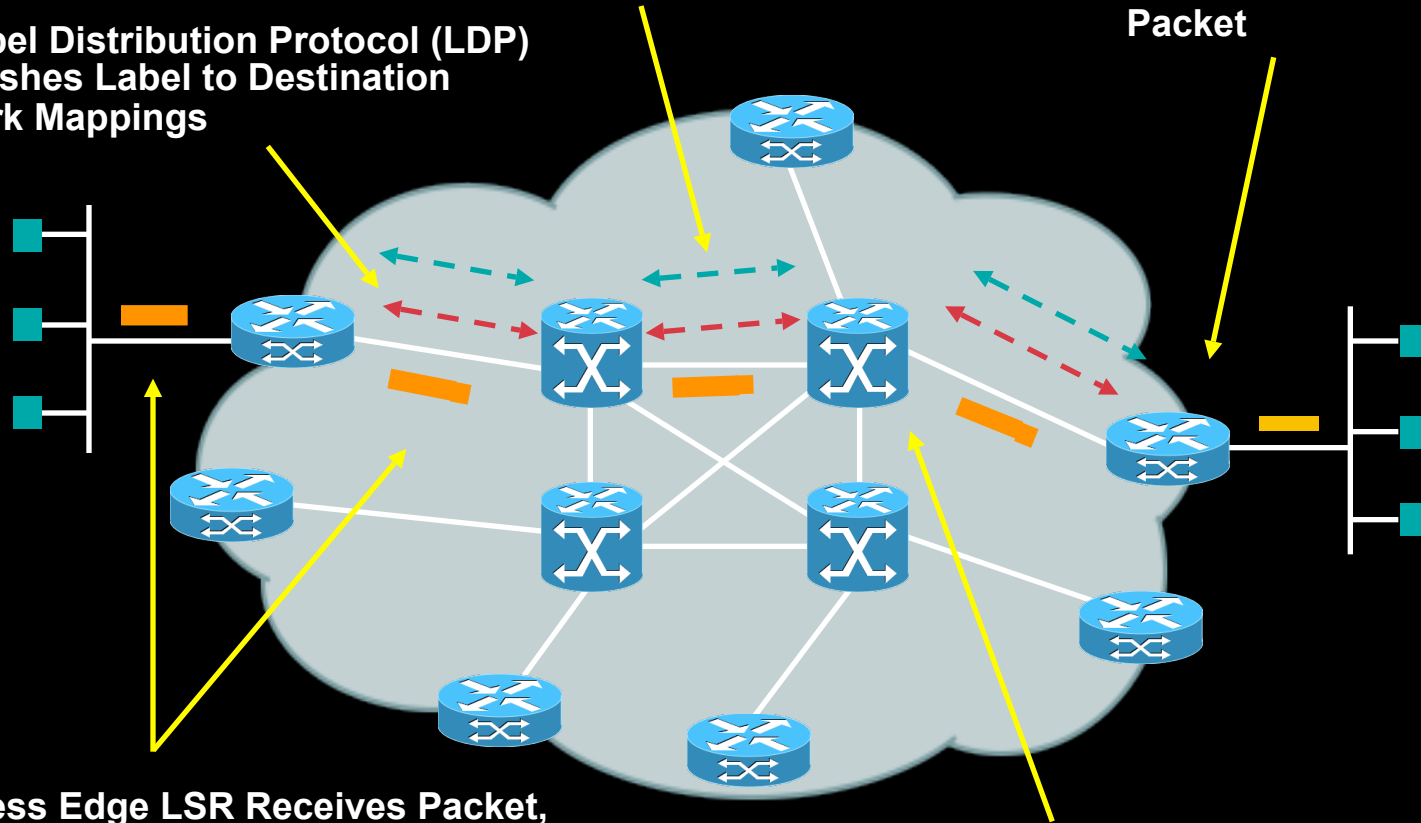
MPLS Operation

Cisco.com

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



- **MPLS VPNS**
- **LAYER 2 AND LAYER 3**

3

- **Layer 2 VPNs**

- Customer endpoints (CPE) connected via Layer 2 such as Frame Relay DLCI, ATM VC or point-to-point connection**

- If it connects IP routers then peering or routing relationship is between the endpoints**

- Multiple logical connections (one with each endpoint)**

- **Layer 3 VPNs**

- Customer end points peer with provider routers**

- Single peering relationship**

- No mesh of connections**

- Provider network responsible for**

- Distributing routing information to VPN sites**

- Separation of routing tables from one VPN to another**

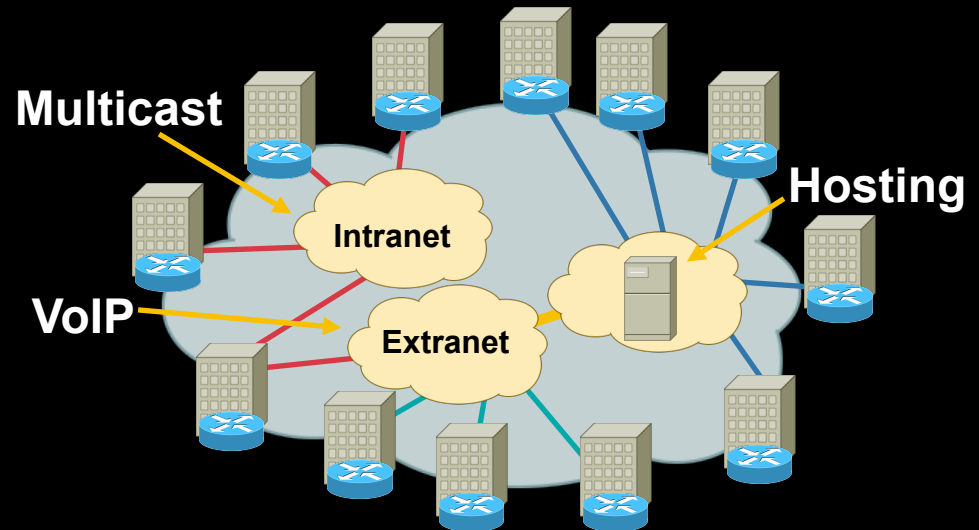
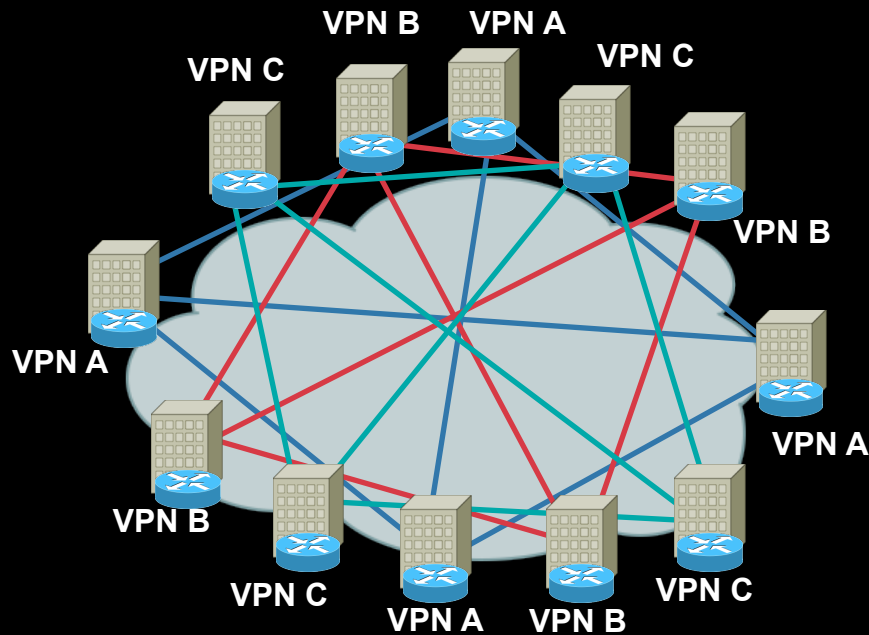


• LAYER 3 VPNS

4

Service Provider Benefits of MPLS-Based VPNs

Cisco.com



- **Overlay VPN**

- Pushes content outside the network
- Costs scale exponentially
- Transport dependent
- Groups endpoints, not groups
- Complex overlay with QoS, tunnels, IP

MPLS-Based VPNs

- Enables content hosting **inside** the network
- “Flat” cost curve
- Transport independent
- Easy grouping of users and services
- Enables QoS inside the VPNs

How Does It Work?

Cisco.com

- **Simple idea**

- Use a label to designate VPN prefix**

- Route that VPN packet to egress PE advertising that prefix**

- Use the IGP label to the VPN packet to the egress node**

- **How is it done?**

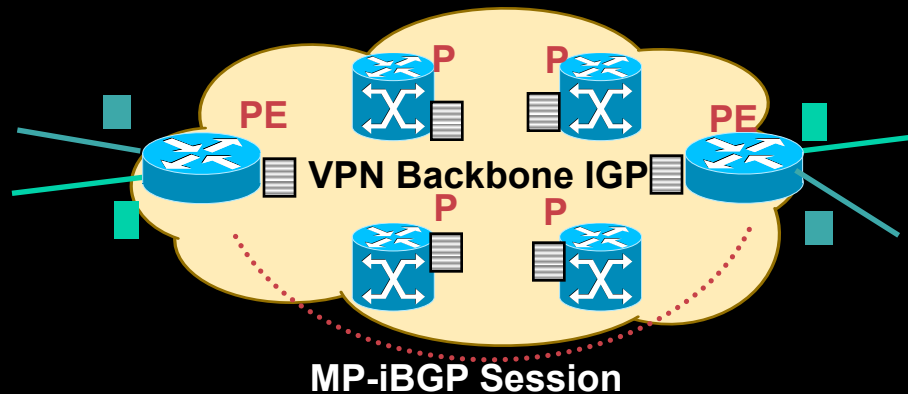
- Routers need to maintain separate VPN routing tables called VRFs (Virtual Routing and Forwarding tables)**

- Routers then export and import routes using BGP extensions to identify and separate one VPNs routes from another**

- Routers then exchange labels for VPN routes in addition to IGP routes**

MPLS-VPN Technology

MPLS VPN Connection Model



PE Routers

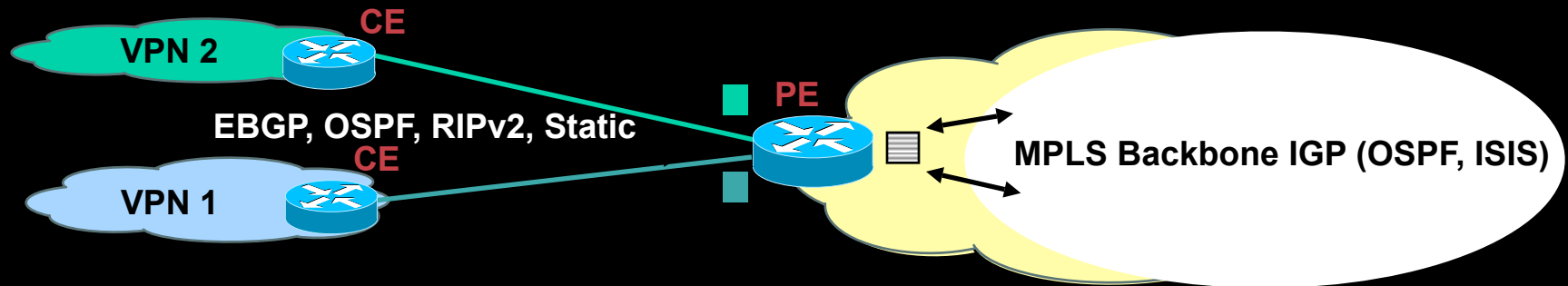
- Edge routers
- Use MPLS with P routers
- Uses IP with CE routers
- Connects to both CE and P routers
- Distribute VPN information through MP-BGP to other PE router with VPN-IPv4 addresses, extended community, label

P Routers

- P routers are in the core of the MPLS cloud
- P routers do not need to run BGP and doesn't need to have any VPN knowledge
- Forward packets by looking at labels
- P and PE routers share a common IGP

MPLS-VPN Technology

Separate Routing Tables at PE



VRF Routing Table

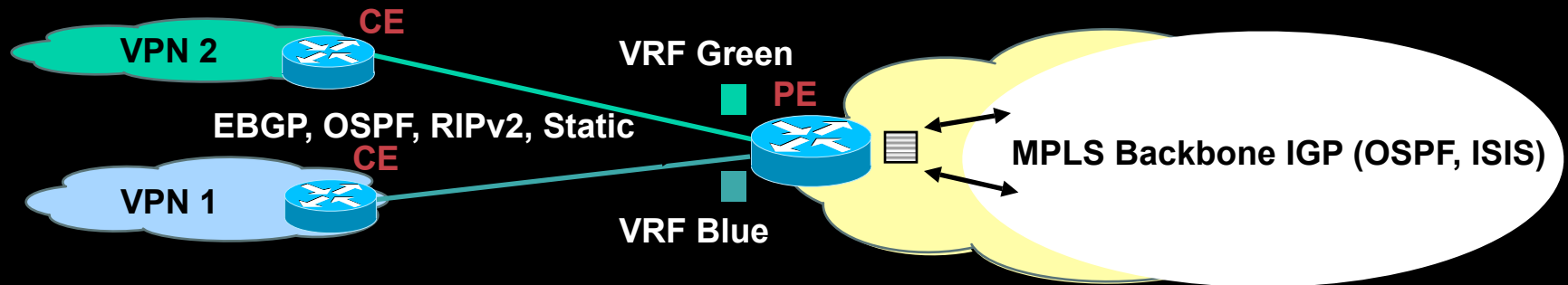
- Routing (RIB) and forwarding table (CEF) associated with one or more directly connected sites (CEs)
- The routes the PE receives from CE routers are installed in the appropriate VRF routing table(s)
 - **blue** VRF routing table or
 - **green** VRF routing table

The Global Routing Table

- Populated by the IGP within MPLS backbone

MPLS-VPN Technology

Virtual Routing and Forwarding Instance (1)

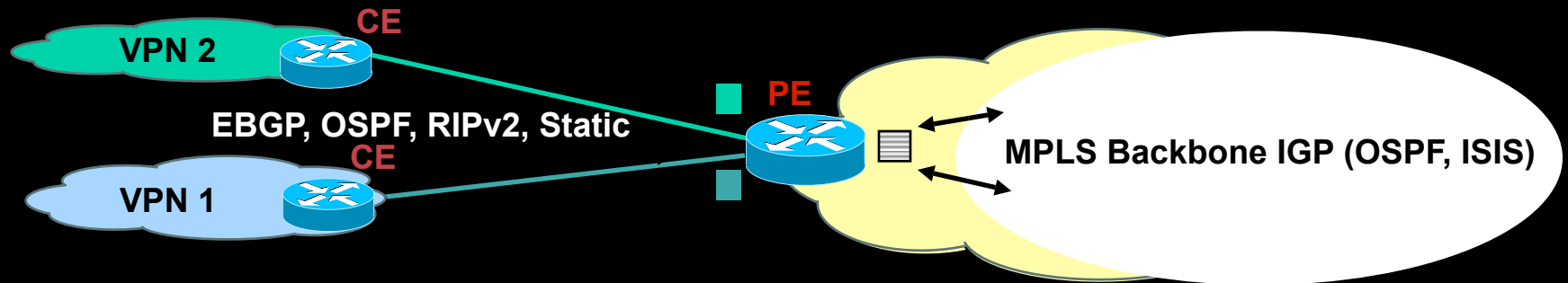


What's a VRF ?

- Associates to one or more interfaces on PE
 - Privatize an interface i.e., coloring of the interface
- Has its own routing table and forwarding table (CEF)
- VRF has its own instance for the routing protocol (static, RIP, BGP, EIGRP, OSPF)
- CE router runs standard routing software

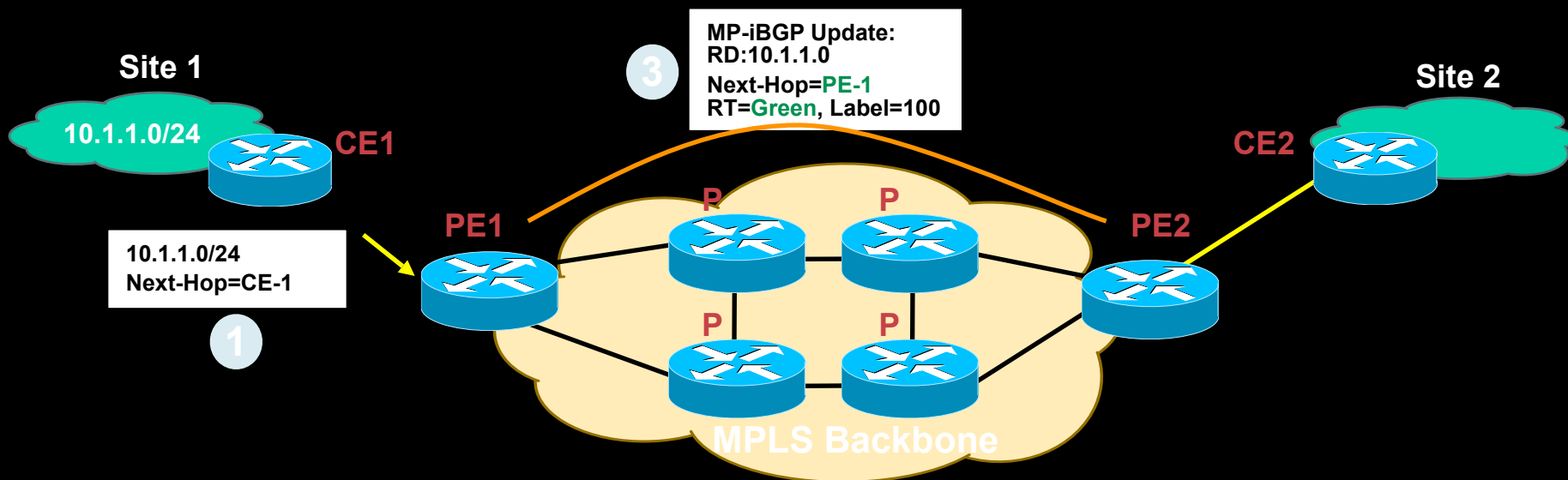
MPLS-VPN Technology

Virtual Routing and Forwarding Instance (2)



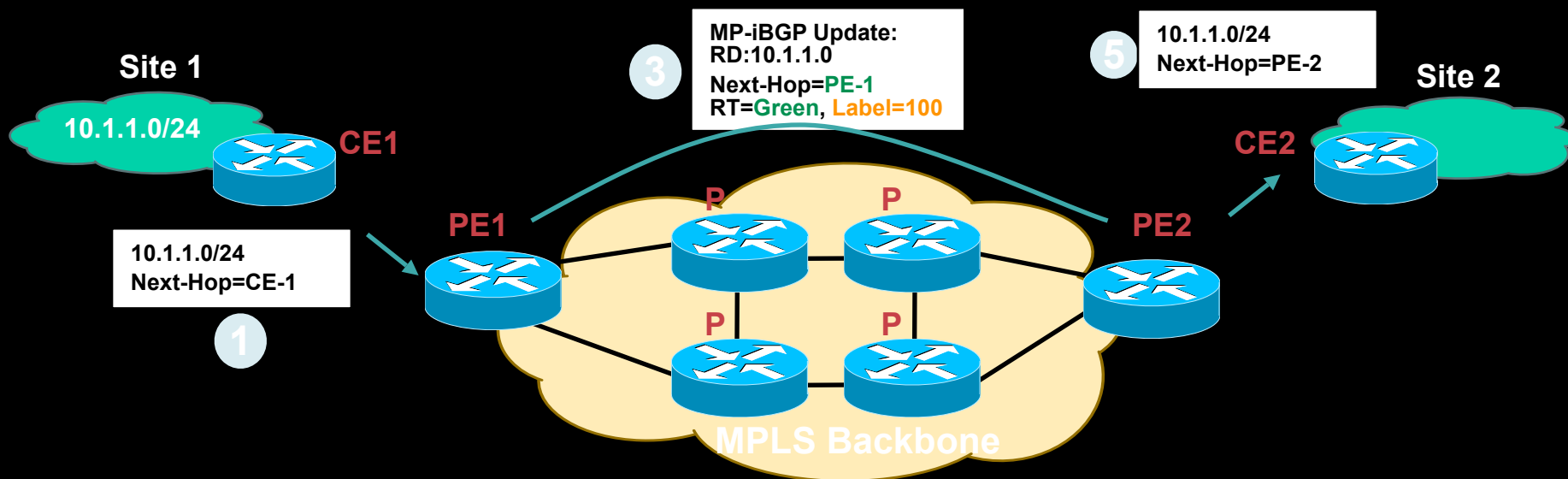
- PE installs the routes, learned from CE routers, in the appropriate VRF routing table(s)
- PE installs the IGP (backbone) routes in the global routing table
- **VPN customers can use overlapping IP addresses**

MPLS VPN Control Plane: Putting It All Together



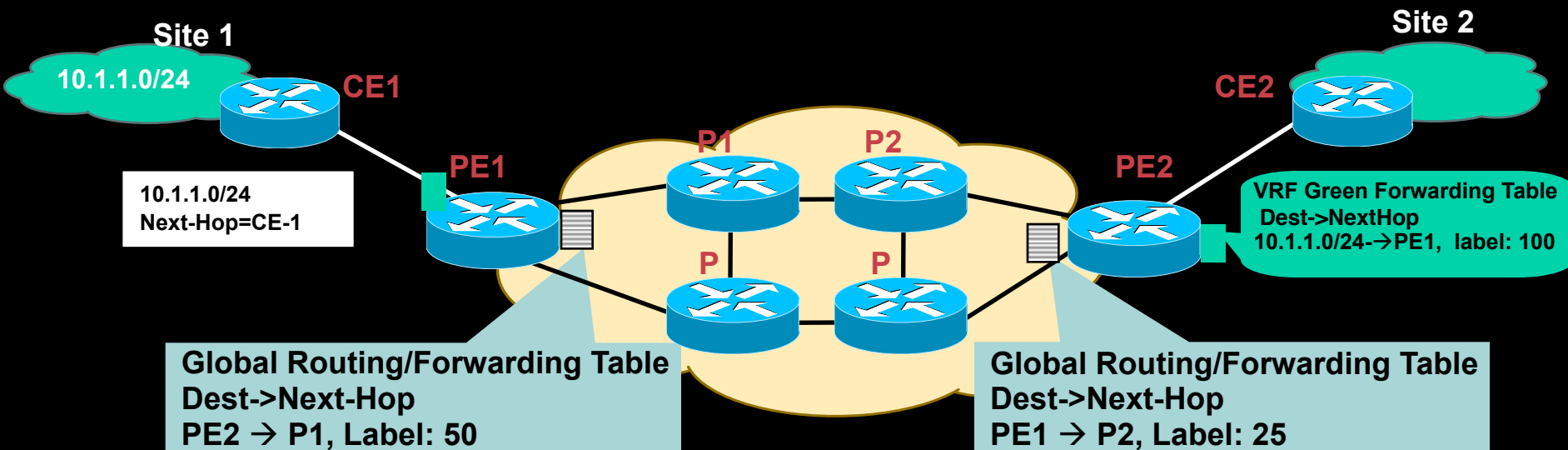
1. PE1 receives an IPv4 update (eBGP,OSPF,EIGRP)
2. PE1 translates it into VPNv4 address
 - Assigns an RT per VRF configuration
 - Rewrites next-hop attribute to itself
 - Assigns a label based on VRF and/or interface
3. PE1 sends MP-iBGP update to other PE routers

MPLS VPN Control Plane: Putting It All Together



1. PE2 receives and checks whether the RT=**green** is locally configured within any VRF, if yes, then
2. PE2 translates VPNv4 prefix back into IPv4 prefix,
Installs the prefix into the VRF routing table
Updates the VRF CEF table with label=100 for 10.1.1.0/24
Advertise this IPv4 prefix to CE2 (EBGP, OSPF, EIGRP)

MPLS-VPN Technology: Forwarding Plane



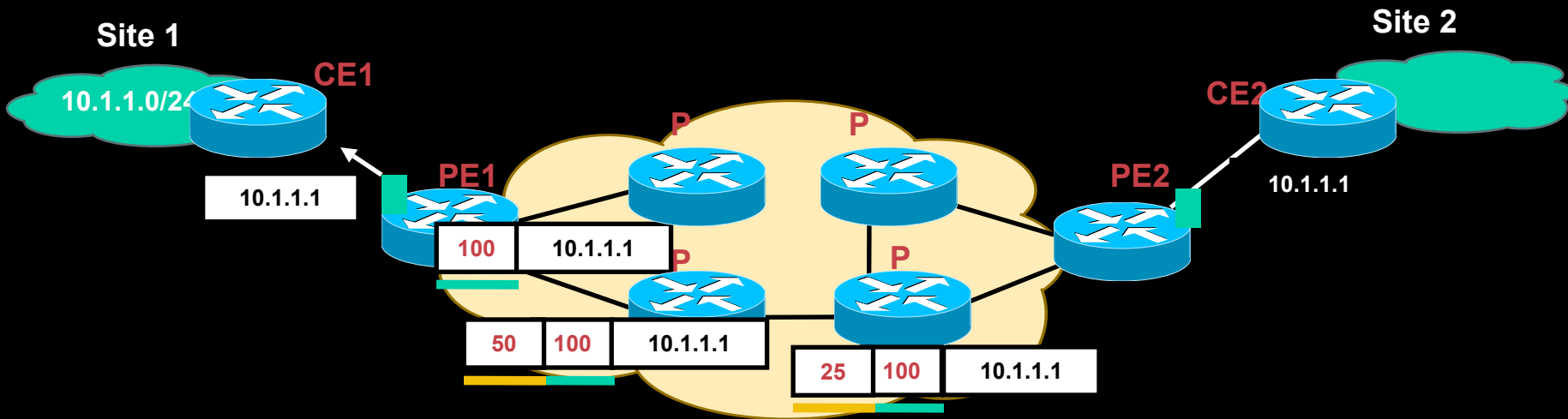
The Global Forwarding Table (show ip cef)

- PE routers store IGP routes
- Associated labels
- Label distributed through LDP/TDP

VRF Forwarding Table (show ip cef vrf <vrf>)

- PE routers store VPN routes
- Associated labels
- **Labels distributed through MP-BGP**

MPLS-VPN Technology: Forwarding Plane



- PE2 imposes TWO labels for each packet going to the **VPN destination 10.1.1.1**
- **The top label is LDP learned and derived from an IGP route**
Represents LSP to PE address (exit point of a VPN route)
- **The second label is learned via MP-BGP**
Corresponds to the VPN address

MPLS VPN Security: Comparison with ATM/FR

Cisco.com

- **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



Key Features

Cisco.com

- **No constraints on addressing plans used by VPNs**
—a VPN customer may:
 - Use globally unique and routable/non-routable addresses
 - Use private addresses (RFC1918)
- **Security:**
 - Basic security is comparable to that provided by FR/ATM-based VPNs without providing data encryption
 - VPN customer may still use IPSec-based mechanisms
 - e.g., CE-CE IPSec-based encryption

Key Features (Cont.)

Cisco.com

- **Quality of Service:**

 - Flexible and scalable support for a CoS-based networks

- **Scalability:**

 - Total capacity of the system isn't bounded by the capacity of an individual component

 - Scale to virtually unlimited number of VPNs per VPN Service provider and scale to thousands of sites per VPN

Key Features (Cont.)

Cisco.com

- **Connectivity to the Internet:**

- VPN service provider may also provide connectivity to the Internet to its VPN customers

- Common infrastructure is used for both VPN and the Internet connectivity services

- **Simplifies operations and management for VPN service providers:**

- No need for VPN service providers to set up and manage a separate backbone or “virtual backbone” for each VPN



- LAYER 2 VPNS

5

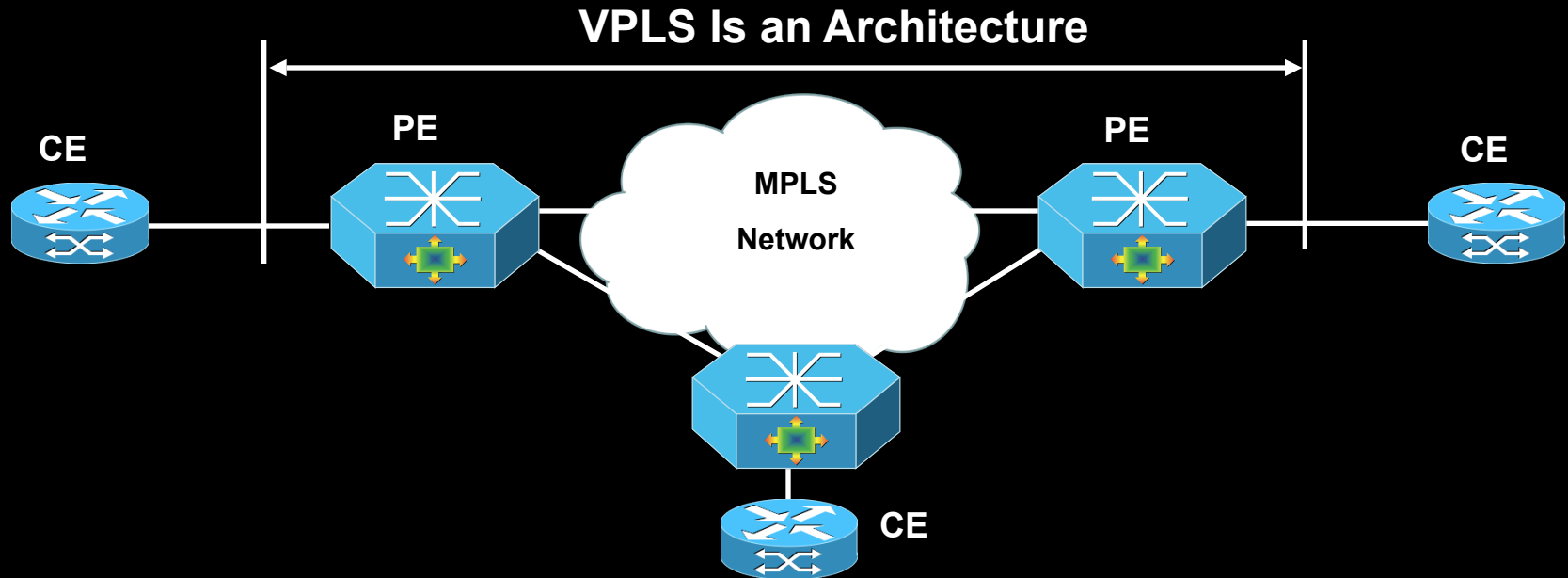
Layer 2 VPNs

Similar to L3VPN

- 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
 - Lookup the L2 label
 - Forward the packet onto the L2 attachment circuit

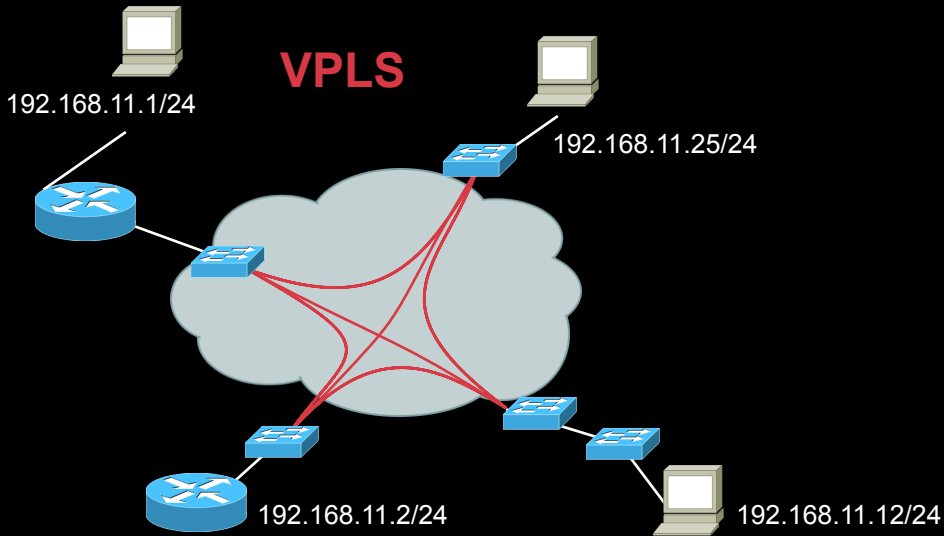
Virtual Private LAN Services (VPLS)

Cisco.com



- **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

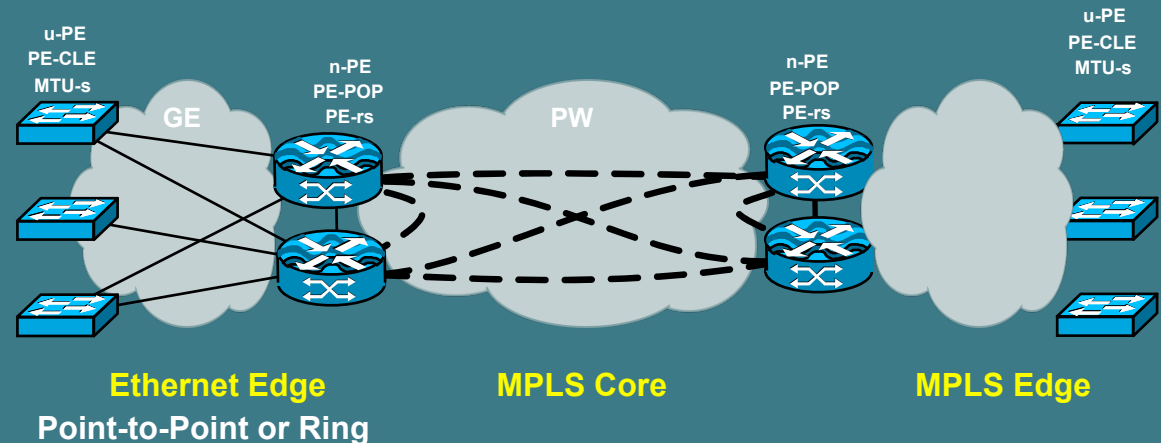


VPLS Direct Attachment

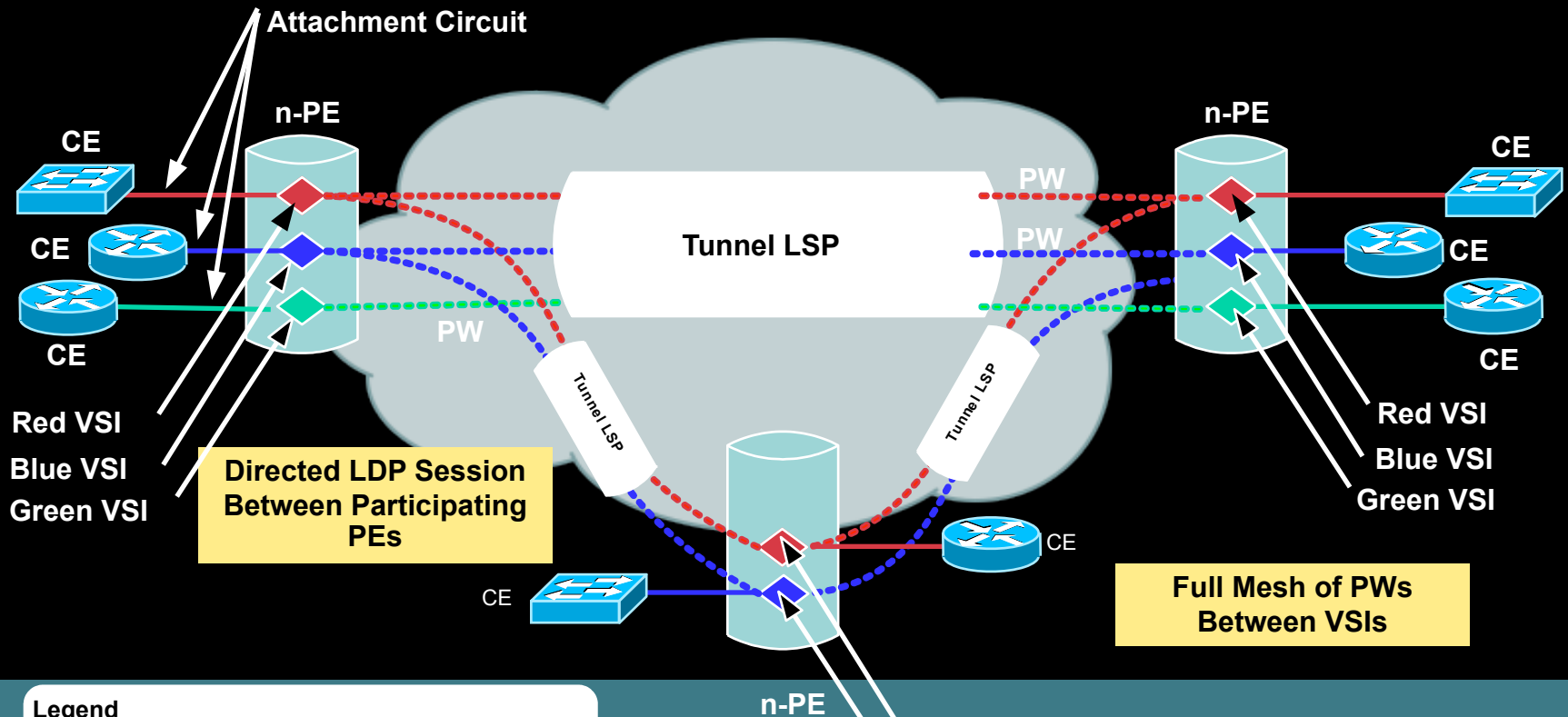
- Single flat hierarchy
- MPLS to the edge

- **H-VPLS**
- Two tier hierarchy
- MPLS or Ethernet edge
- MPLS core

H-VPLS



VPLS Components



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

n-PE
Blue VSI
Red VSI

Summary

Cisco.com

- **Easy way of transporting Layer 2 frames**
- **Can be used to transport ATM AAL5 frames, cells, FR DLCI, PPP sessions, Ethernet VLANs**
- **Point-to-point transport with QoS guarantees**
- **Combine with TE and QoS to emulate Layer 2 service over a packet infrastructure**
- **Easy migration towards network convergence**



• MPLS TRAFFIC ENGINEERING

7

What Is MPLS Traffic Engineering?

Cisco.com

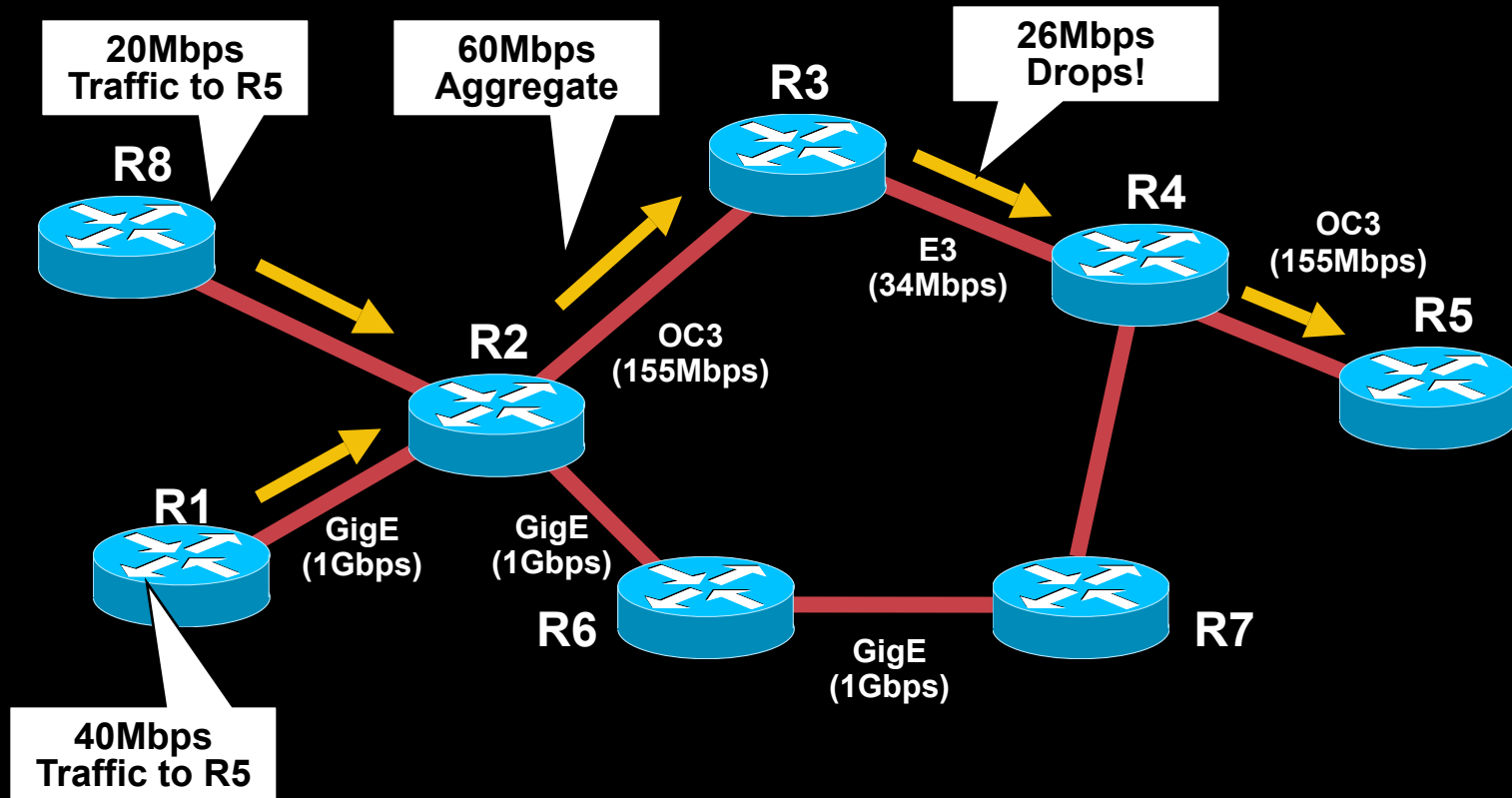
- **Process of routing data traffic in order to balance the traffic load on the various links, routers, and switches in the network**
- **Key in most networks where multiple parallel or alternate paths are available**

Why Traffic Engineering?

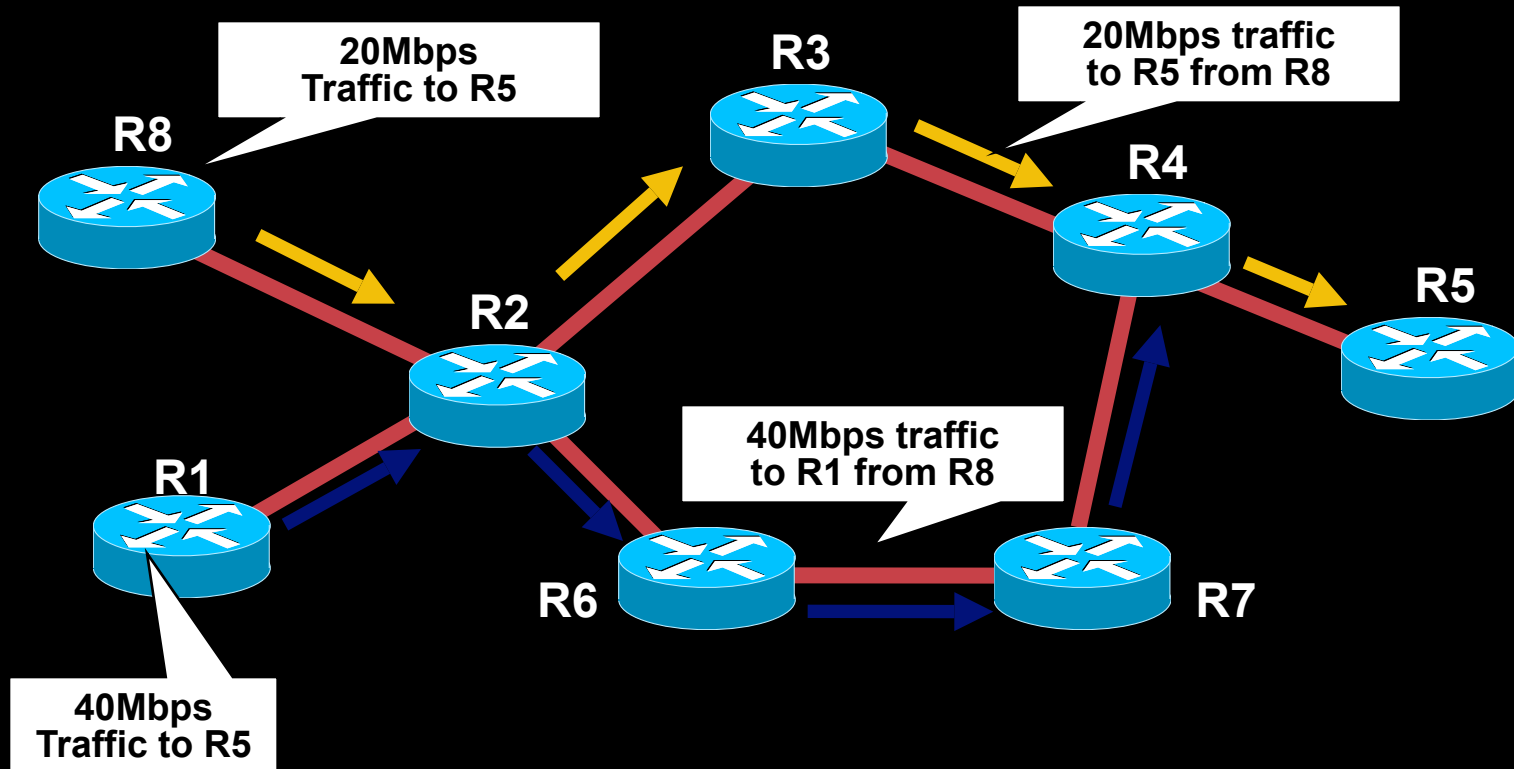
Cisco.com

- **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

Shortest Path and Congestion



The TE Solution



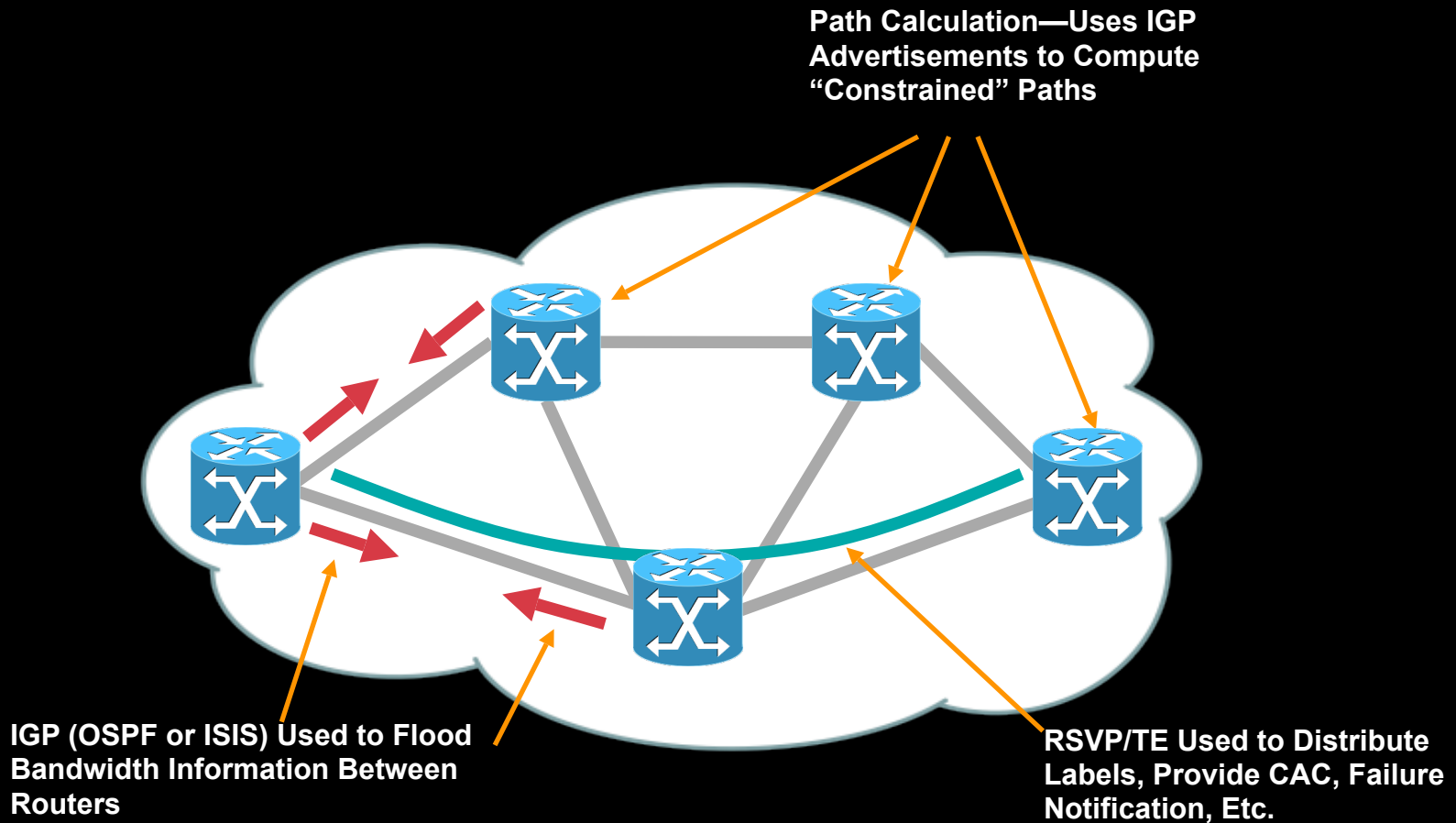
- MPLS Labels can be used to engineer explicit paths
- Tunnels are **UNI-DIRECTIONAL**

➔ Normal path: R8 → R2 → R3 → R4 → R5

➔ Tunnel path: R1 → R2 → R6 → R7 → R4

TE Fundamentals: “Building Blocks”

Cisco.com

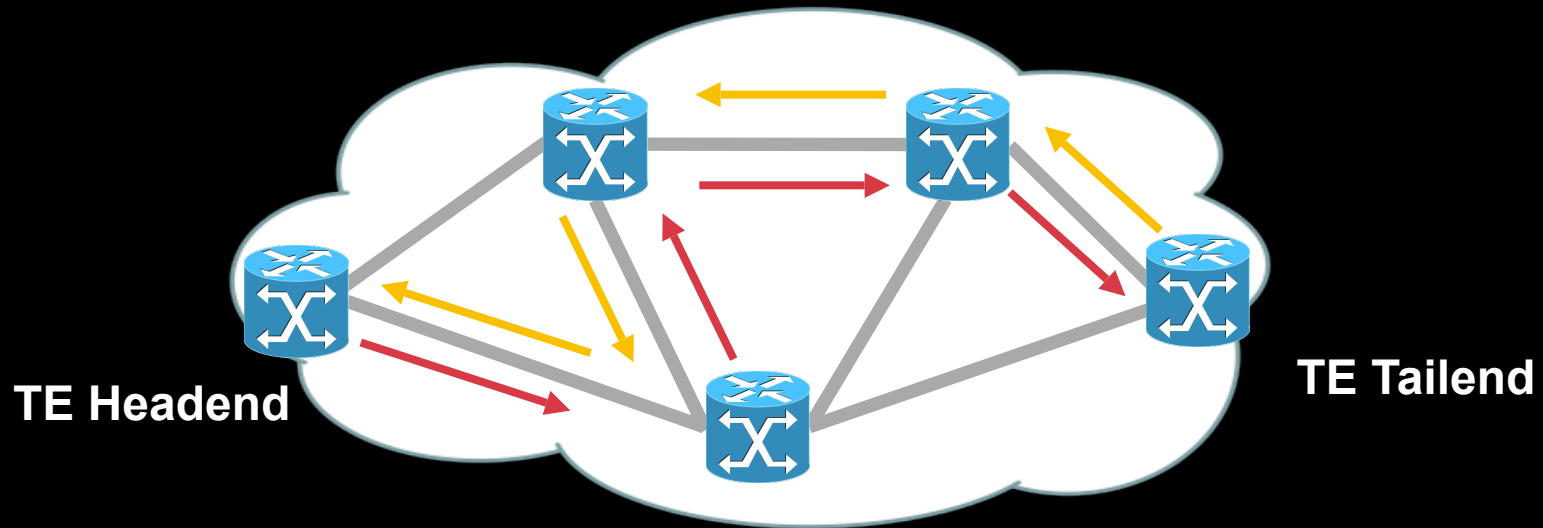


Information Distribution

Cisco.com

- 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.

Example



- **PATH messages are sent with requested bandwidth**
- **RESV messages are sent with label bindings for the TE tunnel**
- **Tunnels can be explicitly routes**
- **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**

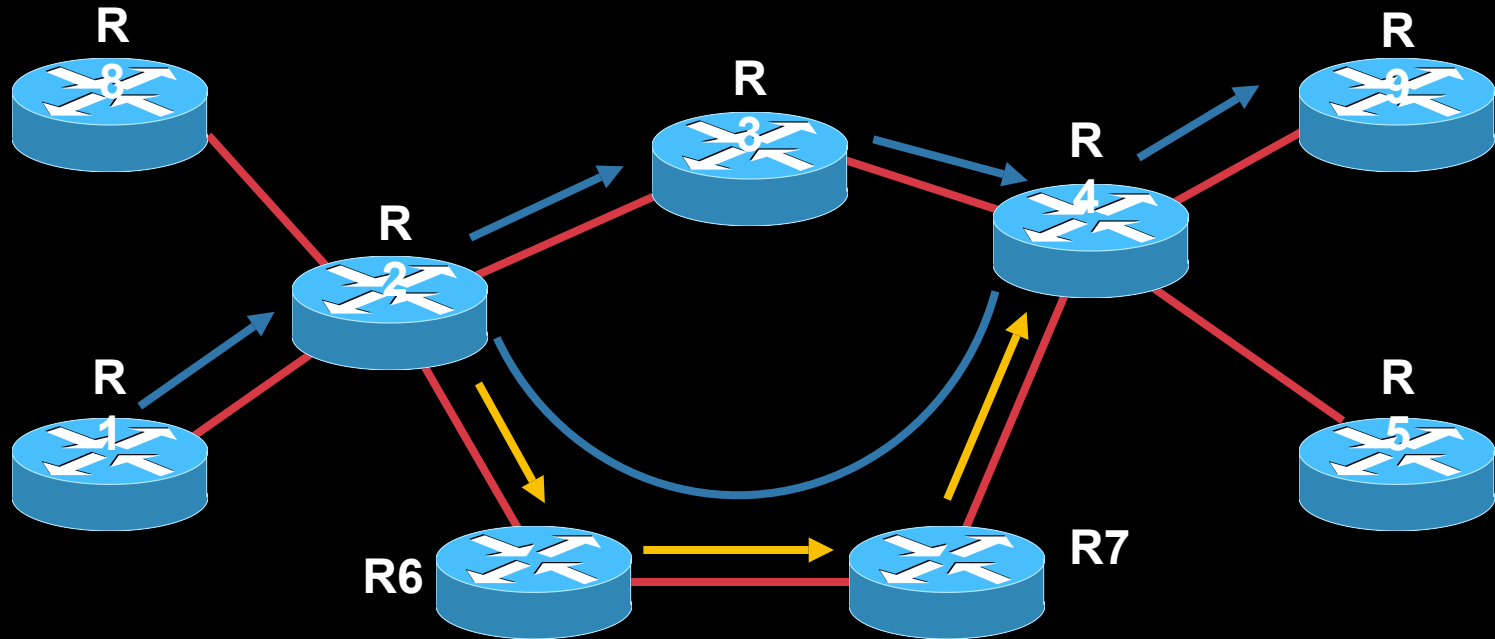
Benefits of TE over Policy Routing

Cisco.com

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

Applications of MPLS TE: MPLS Fast Reroute

Cisco.com

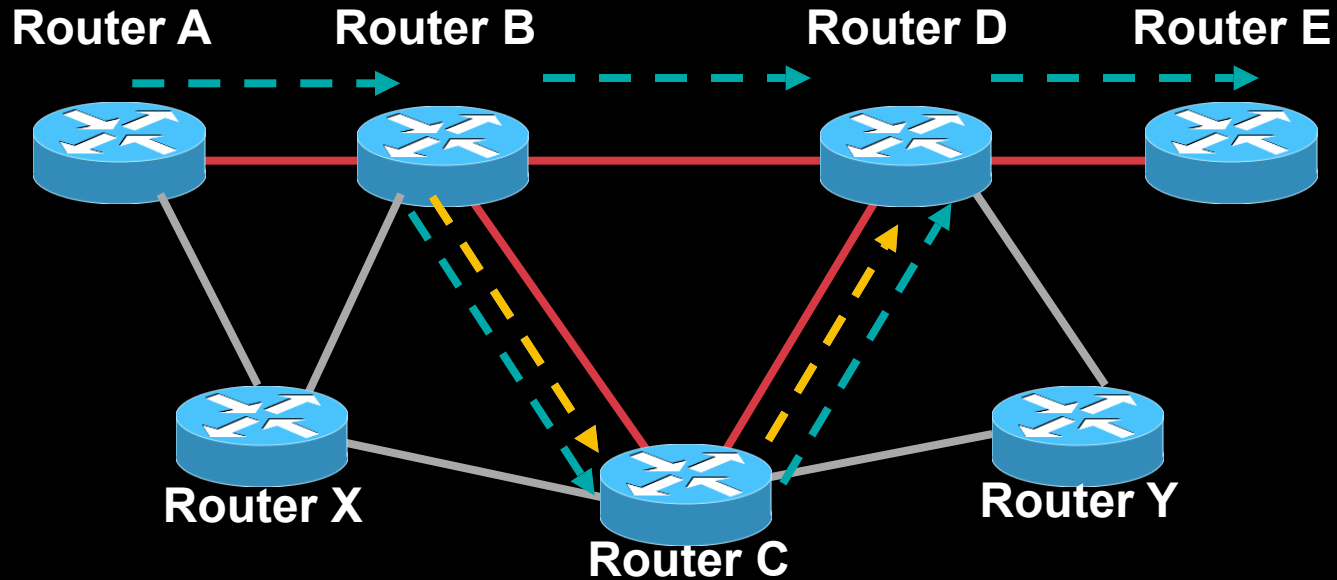


Mimic SONET APS
Reroute in 50ms or Less

- 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

Cisco.com



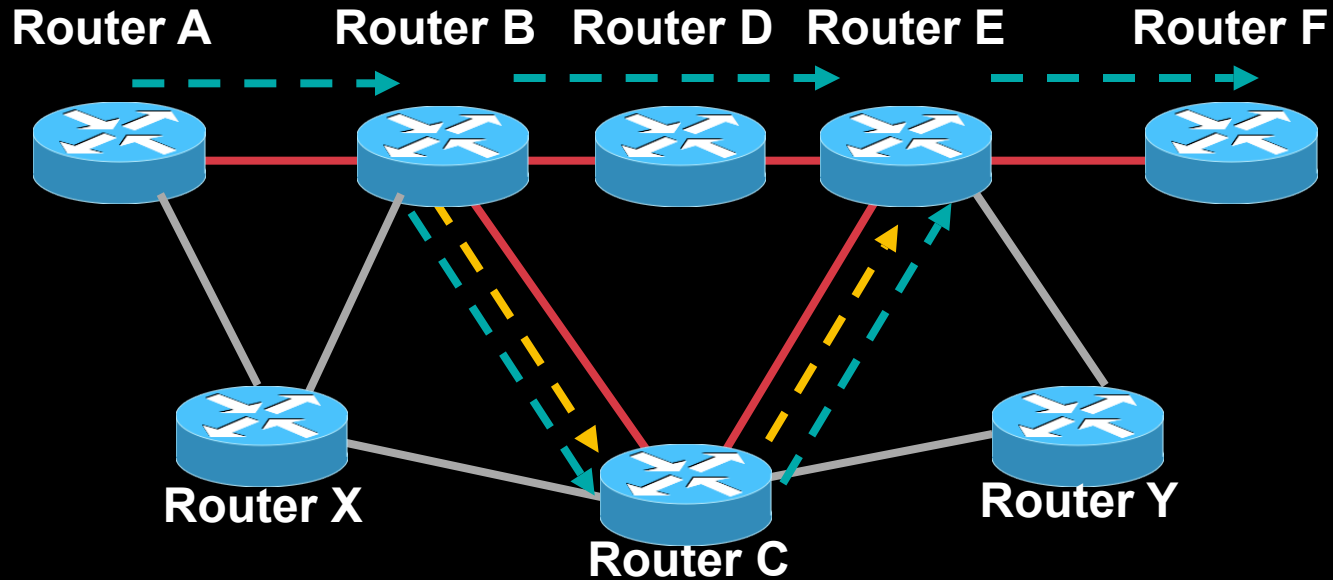
- Primary tunnel: A → B → D → E
- Backup tunnel: B → C → D (preprovisioned)
- Recovery = ~50ms



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

Node & Path Protection

Cisco.com



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

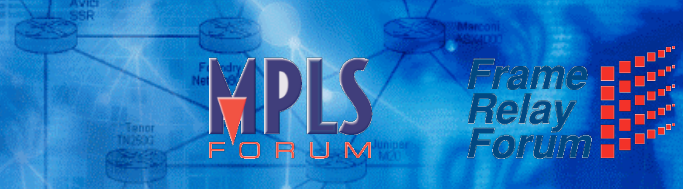




- **United Control Plane Interoperability Effort**
- **Generalized MPLS**

8

What is MPLS?

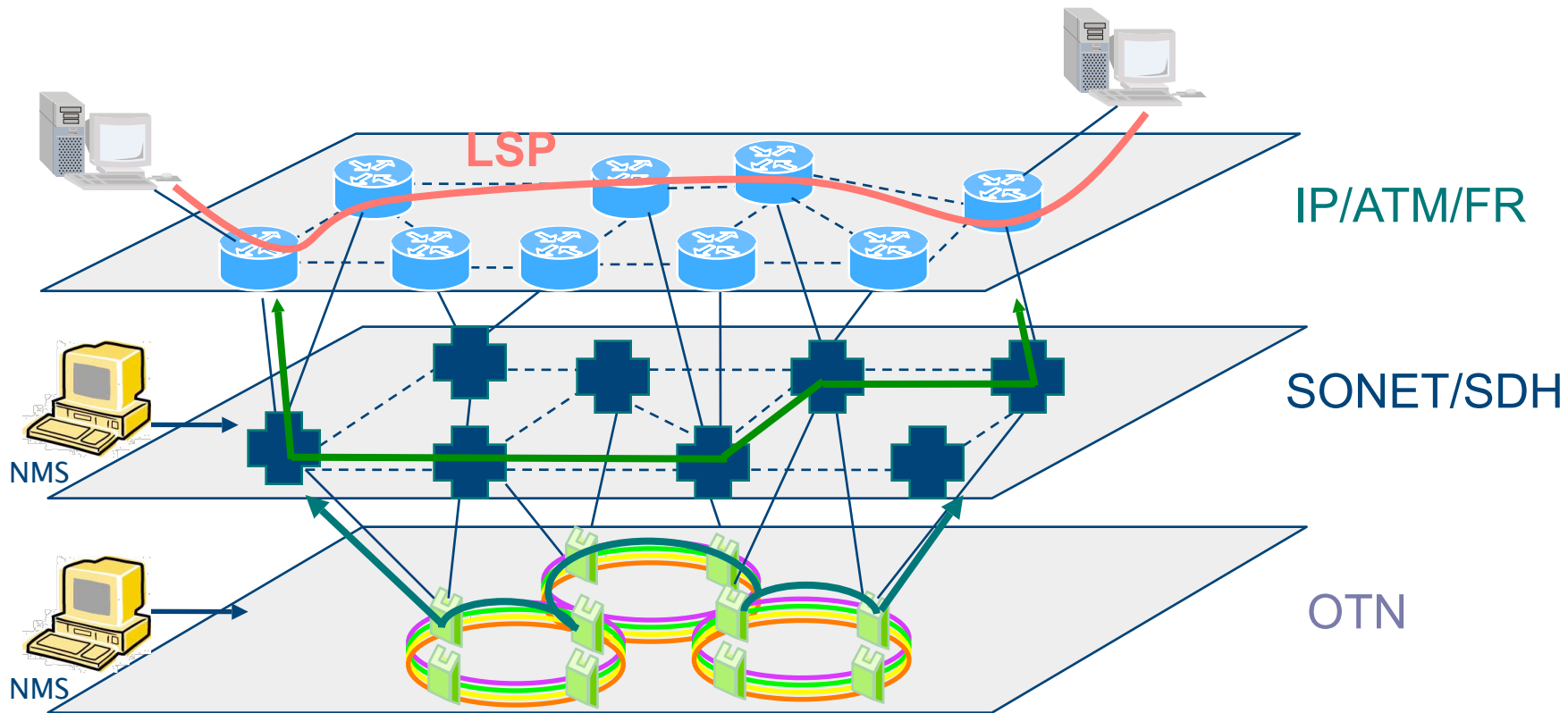


MPLS
FORUM

Frame
Relay
Forum

- **MPLS enables dynamic packet-switched virtual circuits**
- **MPLS brings connection-oriented paths to connectionless IP**
 - ✓ **Constraint Based Routing**
 - ✓ **Traffic Engineering**
- **MPLS provides fast restoration relative to today's IP networks**
- **MPLS enables multiple services**
 - ✓ **VPNs, Traffic Engineering, IPv6, Bandwidth on demand**

MPLS Creates Layer 3 Label Switched Paths - LSPs

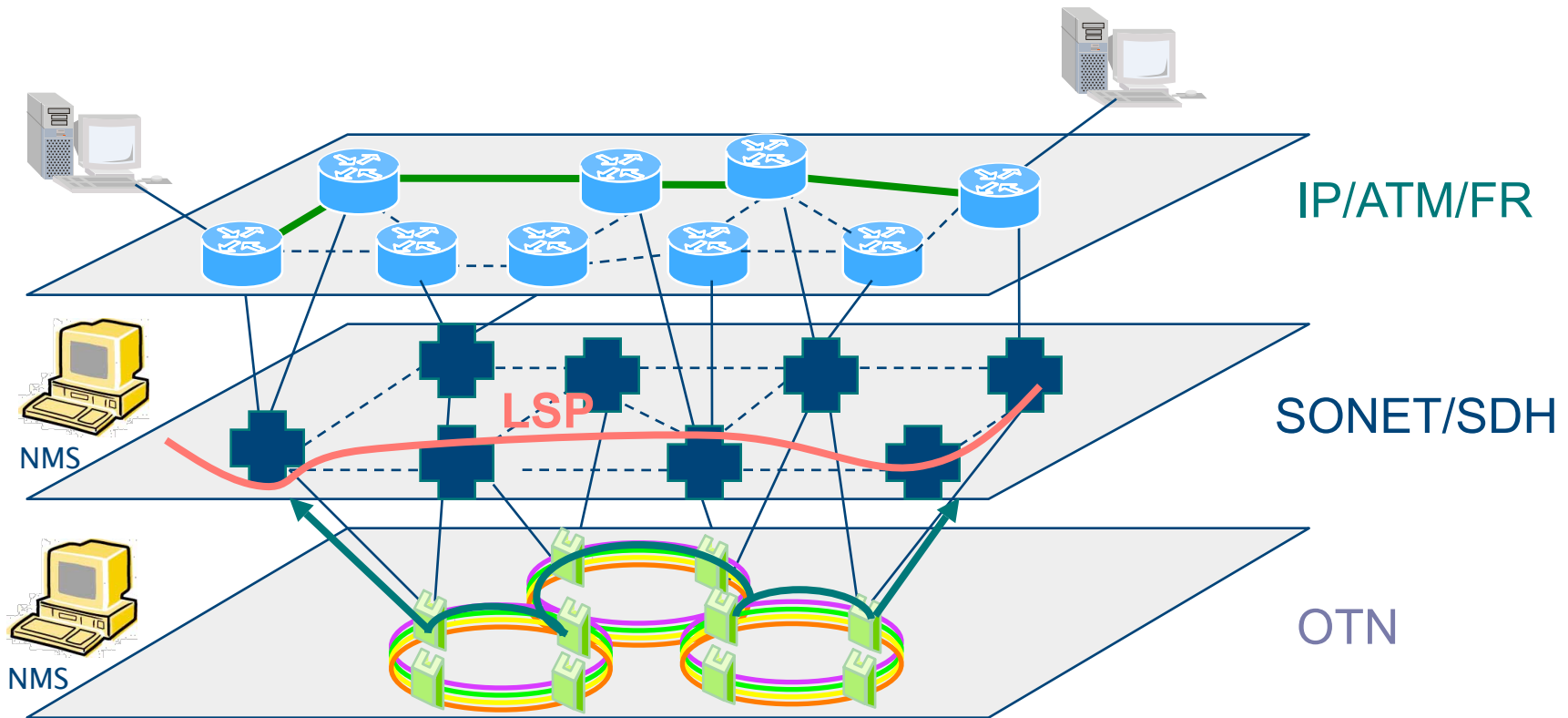


Generalized MPLS (GMPLS)

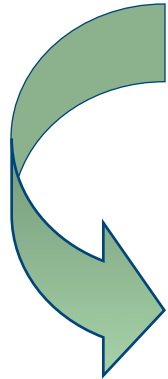


- **Extends MPLS signaling to support multiple switching types**
 - ✓ TDM switching (SDH/SONET)
 - ✓ Wavelength switching (Lambda)
 - ✓ Physical port switching (Fiber)
- **Uses existing and evolving technology**
- **Facilitates parallel evolution in the IP and optical transmission domains**
- **Enhances service provider revenues**
 - ✓ New service creation
 - ✓ Faster provisioning
 - ✓ Operational efficiencies

GMPLS Extends MPLS to Establish Optical Connections

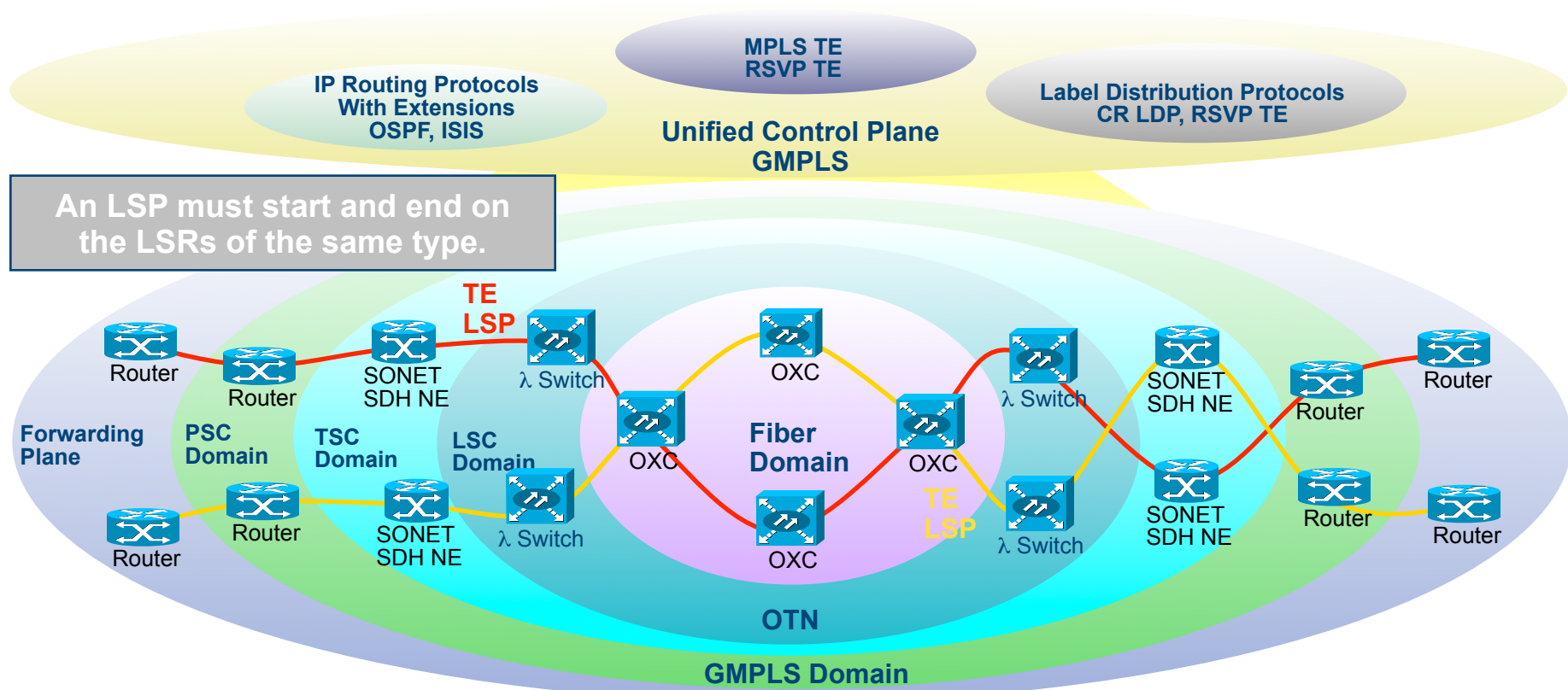


Broadening the scope of MPLS

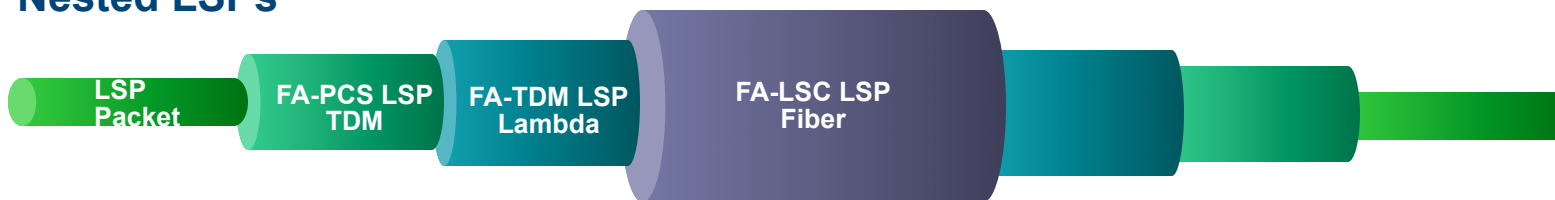


- **MPLS: MultiProtocol Label Switching**
 - ✓ MPLS control plane architecture that can be applied to packet switching
- **GMPLS: Generalized MPLS**
 - ✓ MPLS control plane architecture that can be applied to packet switching (IP, ATM, ...), circuits switching (SONET/SDN, PDH, G.709) and lambda switching (wavelengths/waveband)

Hierarchical LSPs



Nested LSPs





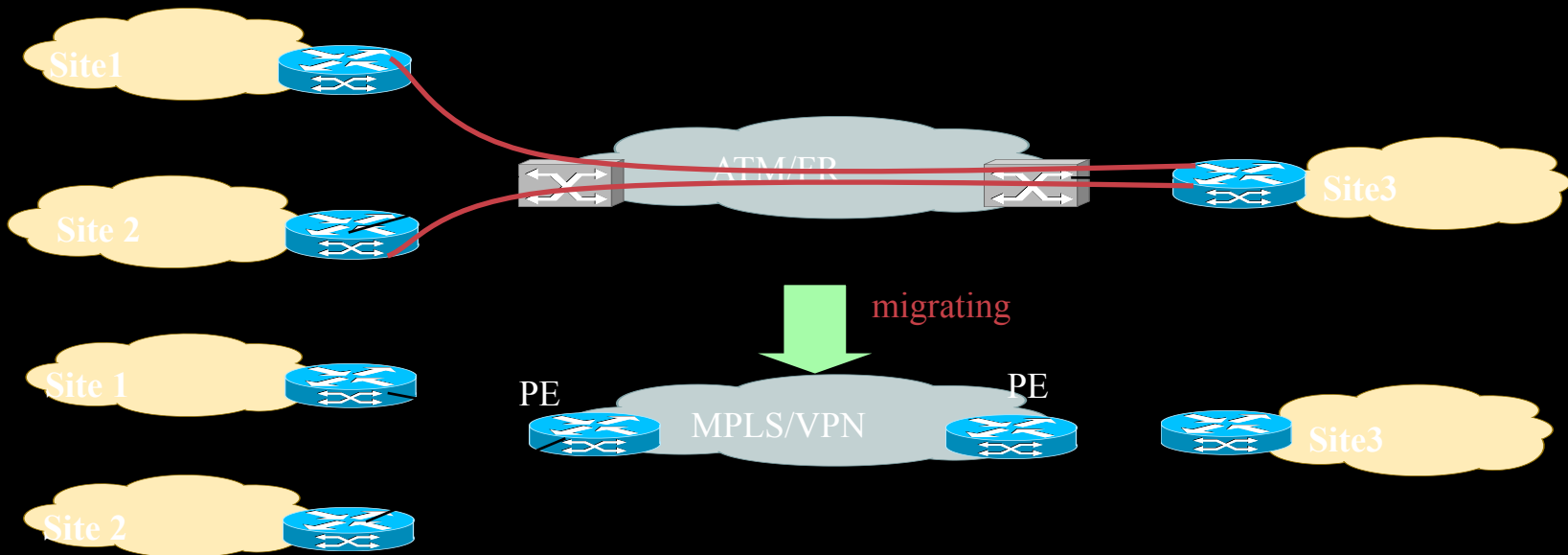
- **Physical Migration of existing WAN Core To SP MPLS VPN Backbone**

9

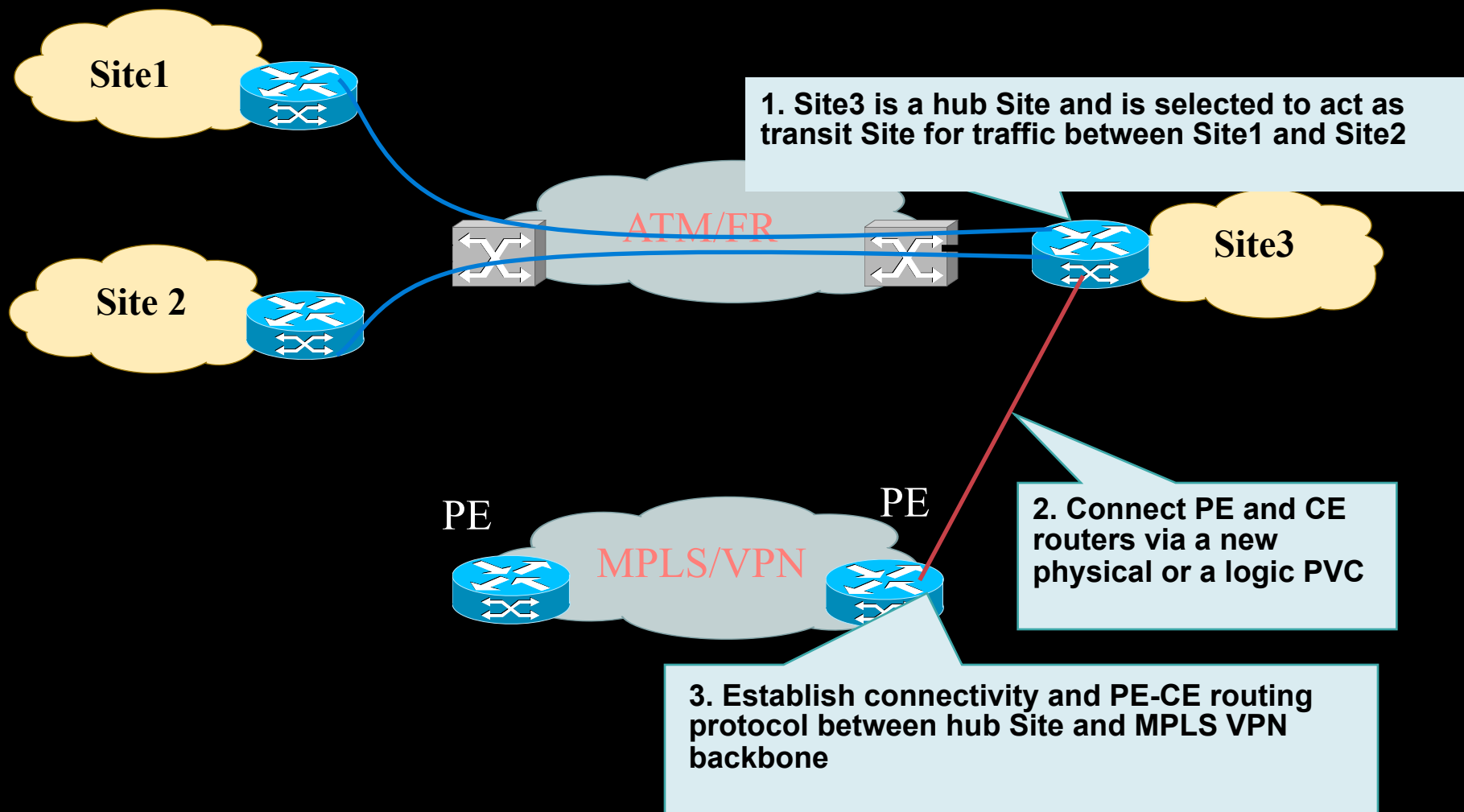
Migration Considerations

Cisco.com

- Minimize impact on customer connectivity and traffic forwarding as well as avoid potential Site isolation during migration.
- Routing interaction of PE-CE routing protocols with the Site local IGP
 - Customers may not use their existing internal routing protocol to exchange routing information with the provider.
- Need to make sure internal as well internet routing works as desired
- Migration of a large enterprise to MPLS VPN needs phased approach

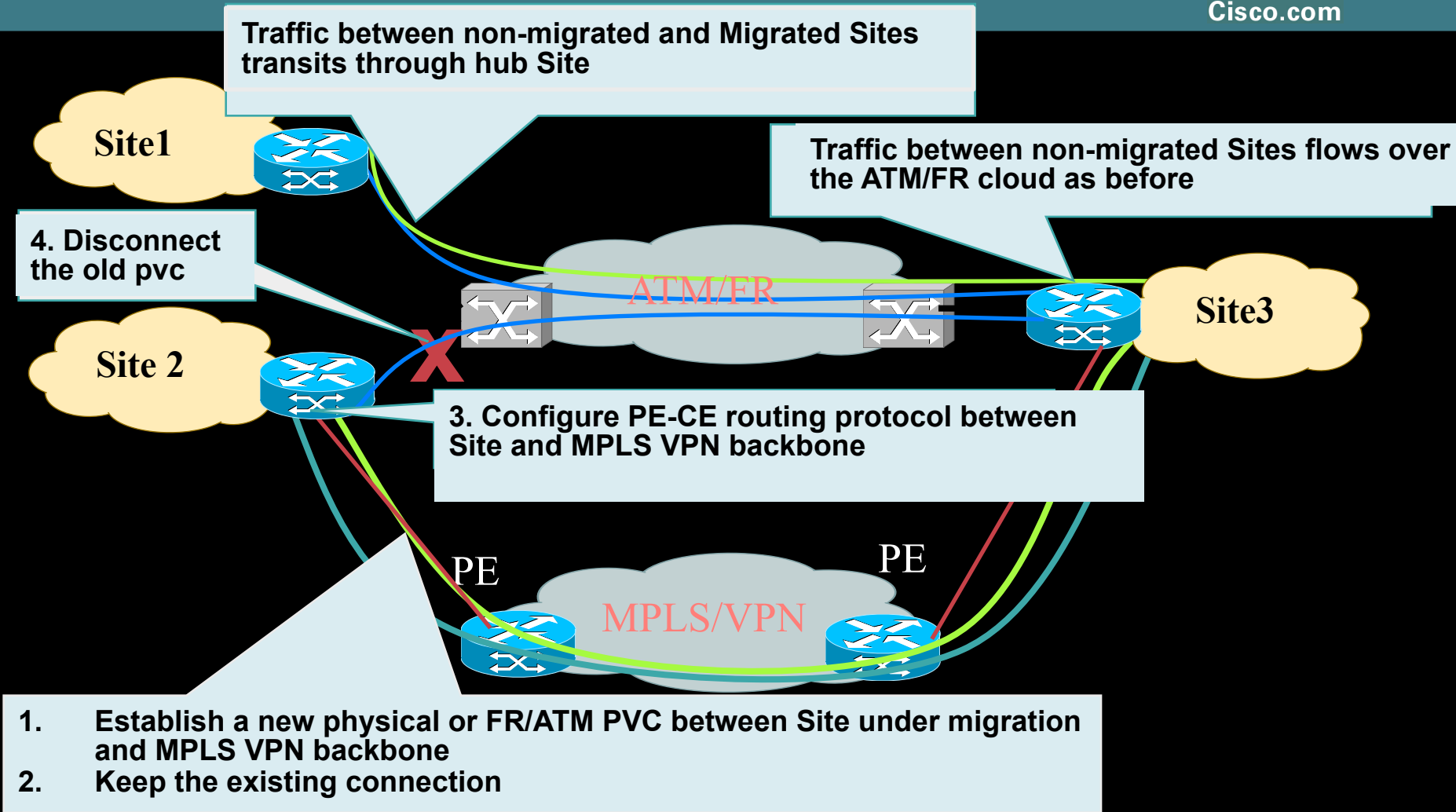


Migration steps: Hub Site Migration



Migration steps: Individual Sites Migration

Cisco.com



Depending on the routing protocol and the corresponding Admin distance and metrics, traffic will start flowing over MPLS VPN backbone



- **BGP as PE-CE Routing Protocol**
- **AS CONSIDERATIONS AND VPN TOPOLOGIES**

10

BGP AS Considerations

VPN Sites belong to same ASN

Cisco.com

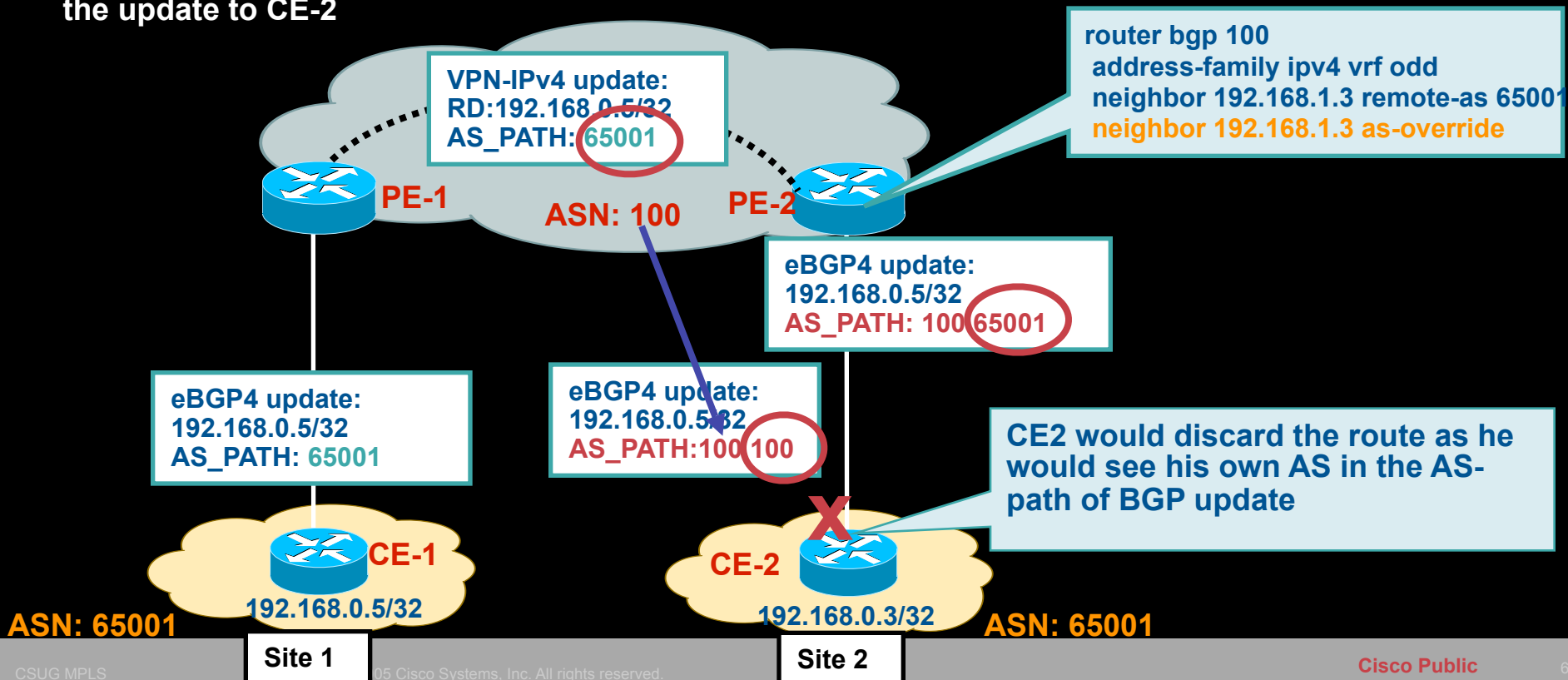
Customer may have same AS number in all its Sites

Default BGP behaviour would force the CE to drop the routing update because of the AS-path loop detection

“Allow-as in” can be used on the CE to accept the update even if it contains its own AS.

Service provider can re-write the customer AS using “AS-override” feature

PE-2 replaces all occurrences of customer ASN in the AS-Path with its own ASN and forwards the update to CE-2



ASN Override with AS_PATH Prepend

```
7200-1#sh ip bgp vpn all
Network      Next Hop      Metric LocPrf Weight Path
Route Distinguisher: 100:1 (default for vrf odd)
*>i192.168.0.3/32 192.168.0.7    0      0    250 i
*> 192.168.0.5/32 192.168.65.5  0      0    250 250 250 i
```

PE-2 Performs Following Actions:

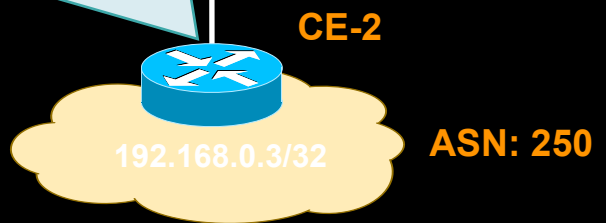
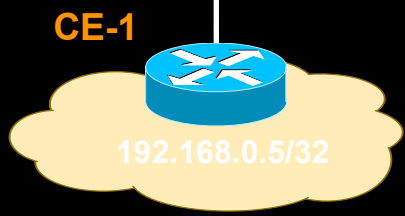
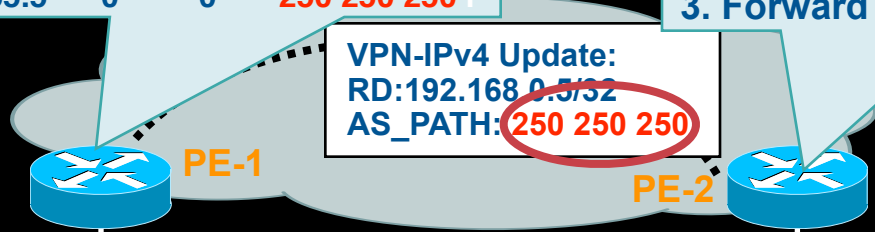
1. Replace All Occurrences of Last ASN with Its Own ASN
2. Update AS_PATH with Its Own ASN
3. Forward the Update to CE-2

```
VPN-IPv4 Update:
RD:192.168.0.5/32
AS_PATH:250 250 250
```

```
eBGP4 Update:
192.168.0.5/32
AS_PATH:100 100 100 100
```

```
eBGP4 Update:
192.168.0.5/32
AS_PATH:250 250 250
```

```
3640-5#sh ip b
Network      Next Hop      Metric LocPrf Weight Path
*> 192.168.0.5/32 192.168.73.7  0      0    100 100 100 100 i
*> 192.168.0.3/32 0.0.0.0      0      0
```

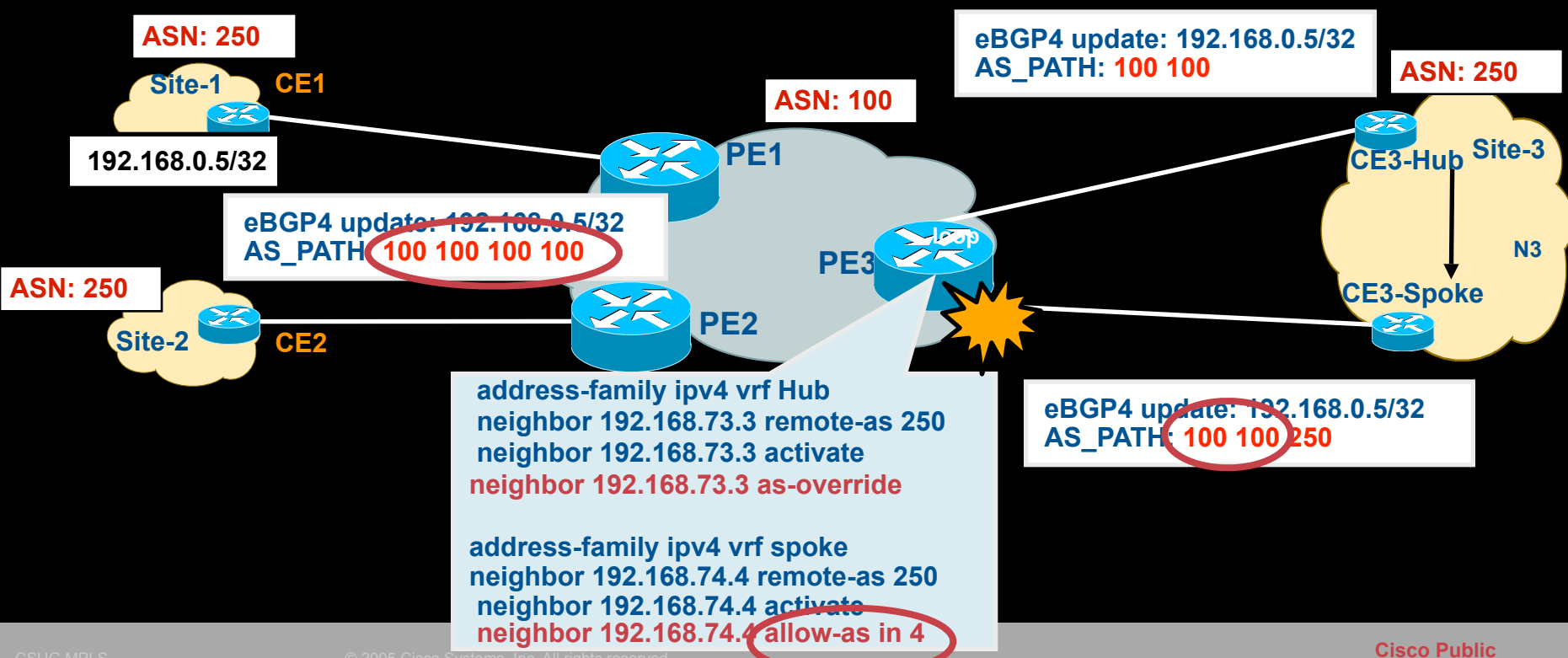


VPN Topology considerations

Hub and Spoke Model

Cisco.com

- PE3 sees its own AS in the AS-Path and rejects the update
- **“Allow-as in”** if configured at spoke Site, will allow the update at PE3 if it contains SP’s ASN



Summary

Cisco.com

- For large enterprises, migration to L3VPN service requires a **phased approach** so that disruption to existing services is minimal
- Existing **site local routing protocols policies** and their interaction with PE-CE routing protocols should be carefully analyzed
- **Topological considerations** such as backdoor links, multi-homing scenarios, summarization, OSPF areas placement and BGP AS number scheme etc should be taken into account to avoid sub-optimal routing or loops.
- **Default route and Summarization** is important for proper routing to the internet or to the central sites and could be coordinated with the service provider for optimal results.
- **Site-to-site VPN routing convergence** should be kept in mind while deploying delay sensitive application
- **Redundancy and Multi-provider** topologies may result in loops if not properly implemented.

CISCO SYSTEMS

