Fundamentals
Multiprotocol Label Switching – MPLS III

Design of Telecommunication Infrastructures
2008-2009

Rafael Sebastian
Departament de tecnologies de la Informació i les Comunicaciones
Universitat Pompeu Fabra

Table of Contents

Goals of the section

- ☑ The operation of MPLS VPN
- ☑ What RDs and RTs are
- ☑ The role of Multiprotocol BGP in MPLS VPN networks
- ☑ The packet forwarding through the MPLS VPN network
- ☑ The operation of PE-CE routing protocols and their specifics for MPLS VPNs

MPLS VPN

- Architecture of MPLS VPN
- BGP Review
- Packet Forwarding
- PE-CE Protocols
- Topologies in MPLS VPNs

- Review Questions

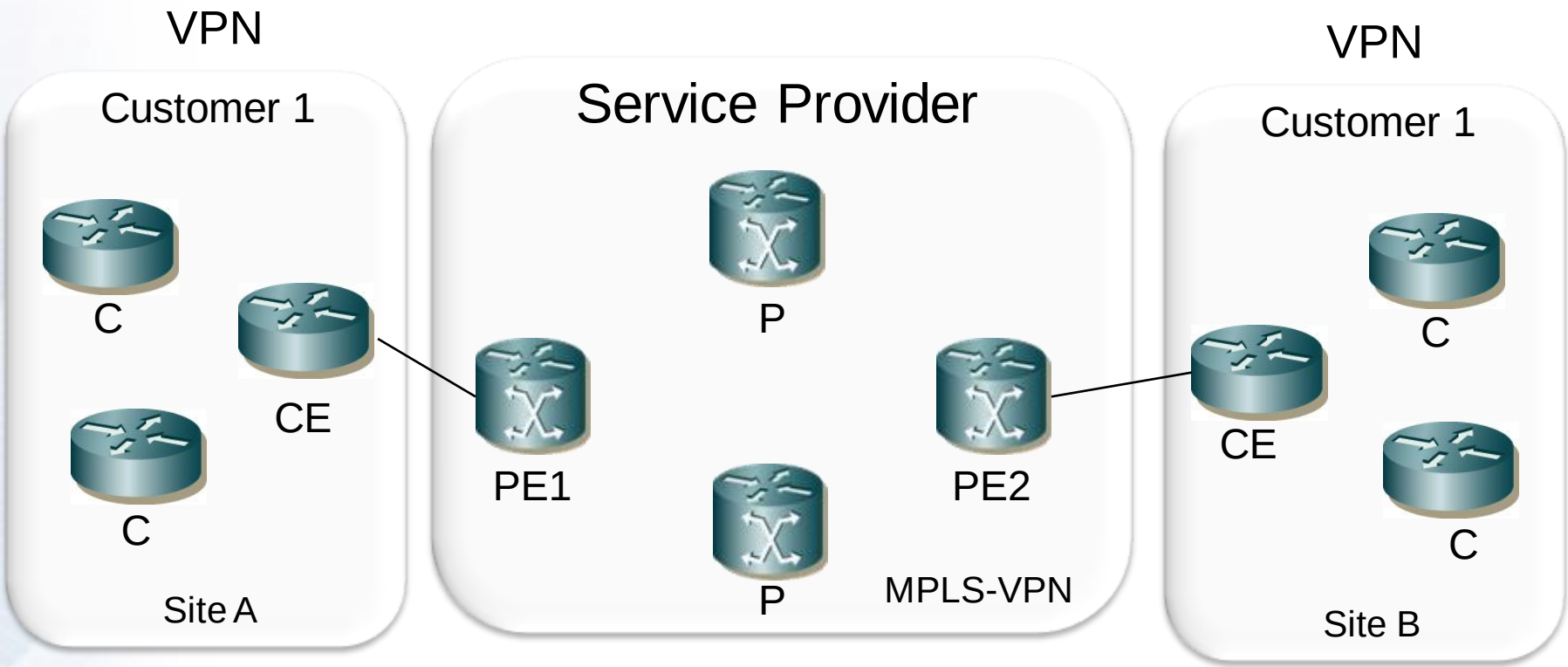
VPN Facts

Review

- VPN *emulates* a virtual private network over a common infrastructure
- They provide Layer 2 or 3 communications
- All customer inside the VPN must have connectivity
  - Connectivity to other VPN might be required
  - Connectivity to the Internet is desired

MPLS VPN Terminology

Review



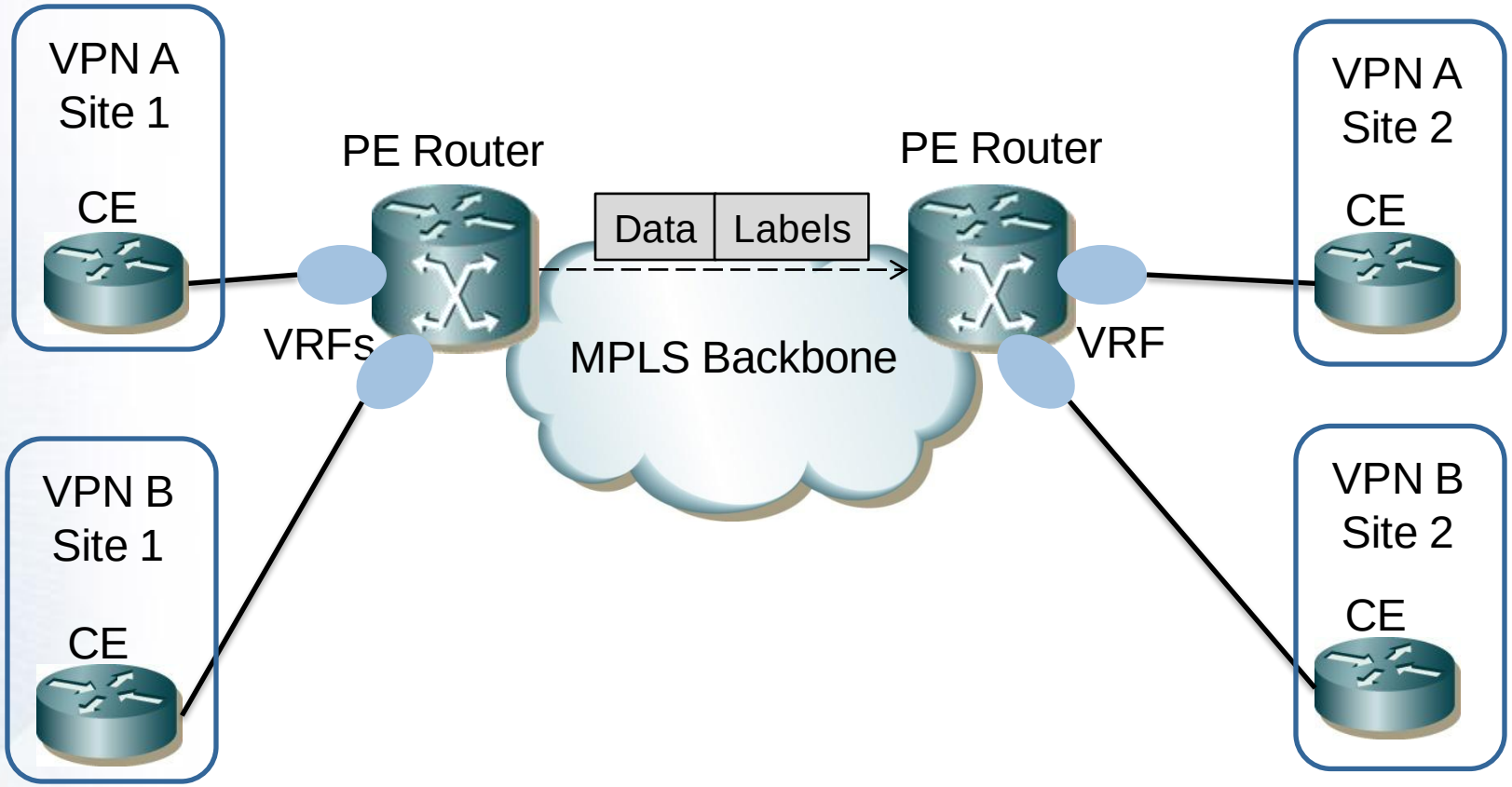
MPLS VPN

Review

- If P would forward looking at IP address
  - Addresses in customers cannot be private (possible overlap)
  - P and PE must have complete routing tables → large routing table for many customers
    - If that's the case: Use of iBGP mandatory in all P-routers
- If P have a separate routing table/Customer?
  - One routing process/VPN at each P (not scalable)
  - P cannot distinguish between VPN
- SOLUTION → P uses Labels
  - P completely unaware of VPNs

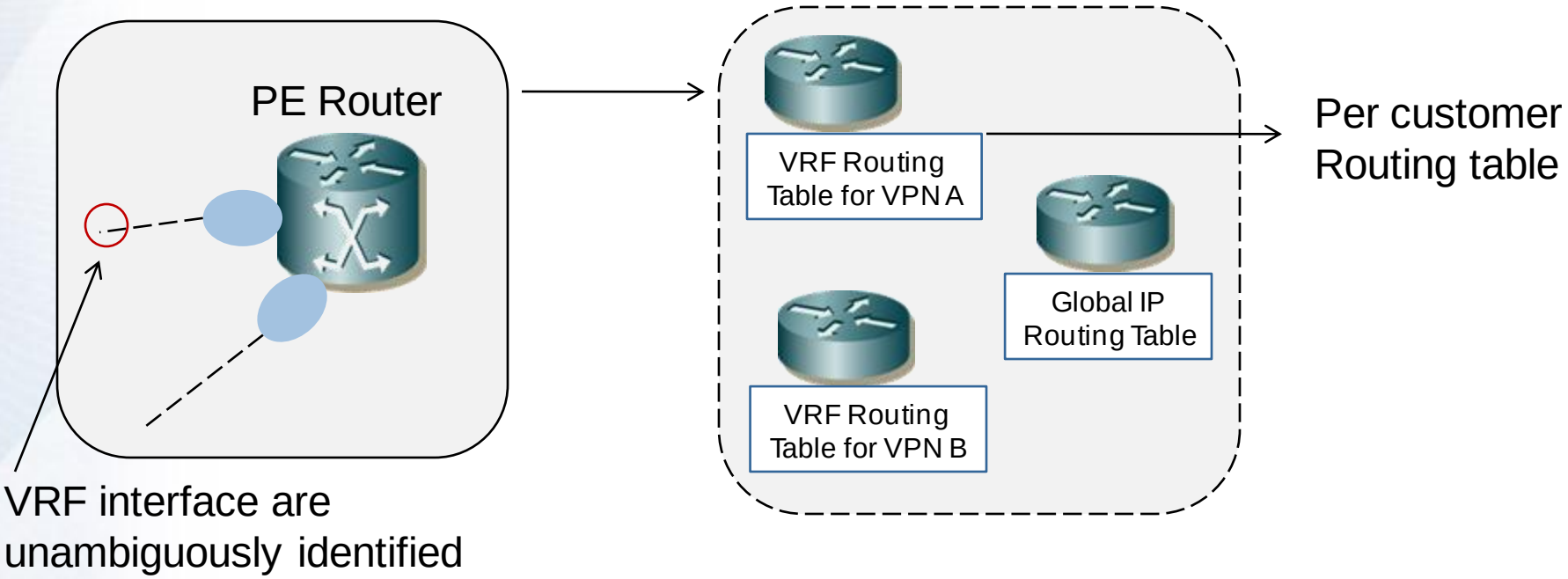
MPLS VPN model

Review



Virtual Routing Forwarding

- A virtual routing/forwarding (VRF) is a VPN routing and forwarding instance
- A PE router has a VRF instance for each attached VPN



Route Distinguisher - RD

- VPN prefixes carried into MPLS by MP-BGP
  - Prefixes across ISP must be unique (no overlap)
- In case same prefixes are in different customers
  - Use of RD → Unique ID (64-bit field)

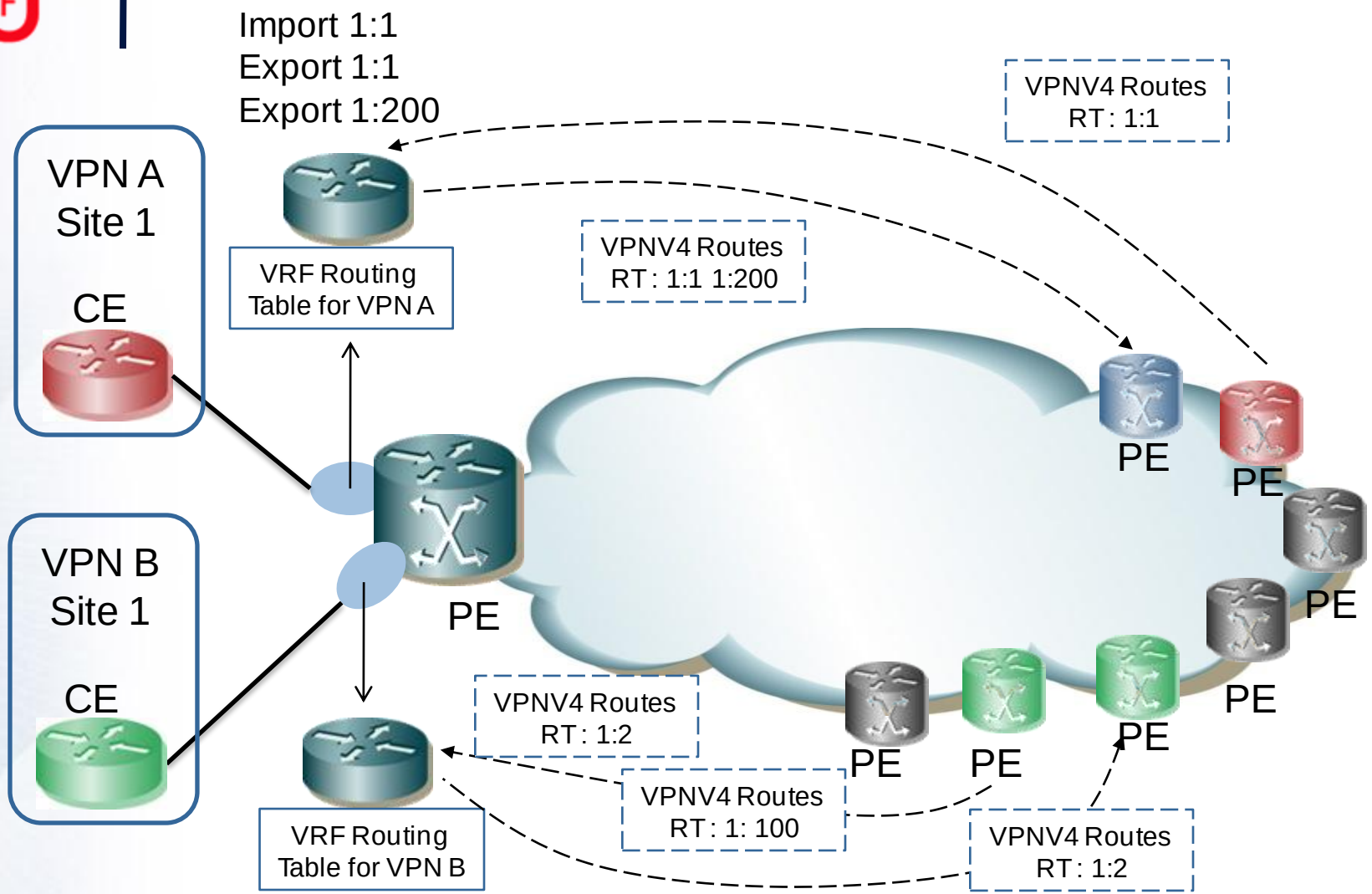
Now prefixes = IPv4 + RD = vpnv4

- MP-BGP carries vpnv4 prefixes between PEs
- RD (e.g. 1:1:10.100.1.0/24)
  - Option 1: ASN:nn (ASN = Autonomous system)
  - Option 2 IP-address:nn

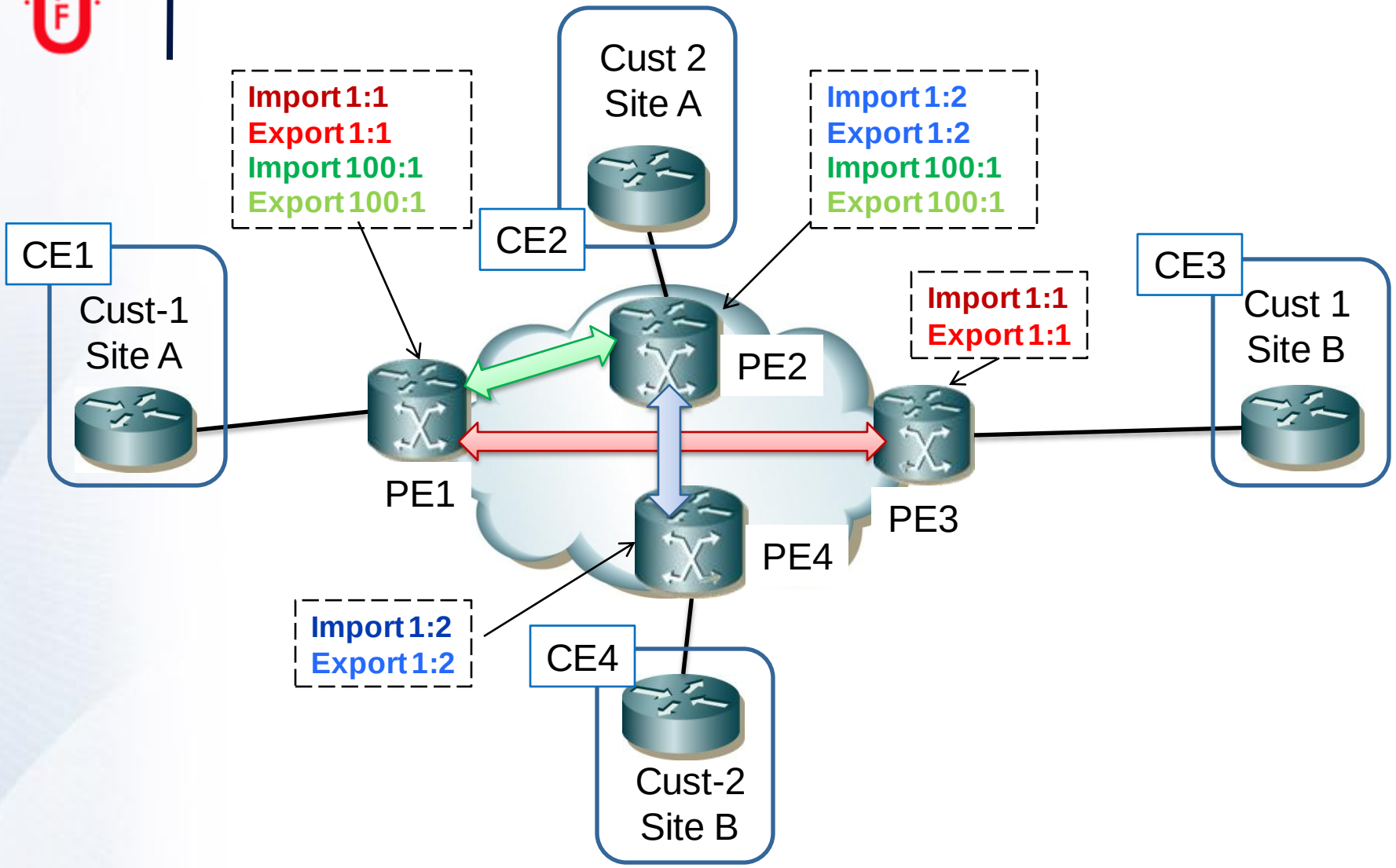
Route Targets - RT

- What if I want to communicate two different VPNs?
  - RD will not match!! → not sharing routes
- RT → routes that should be imported from MP-BGP into the VRF
- Exporting an RT
  - Export vpnv4 routes with the RT on the PE from VRF
- Import an RT
  - Received vpn4 route from MP-BGP imported by VRF

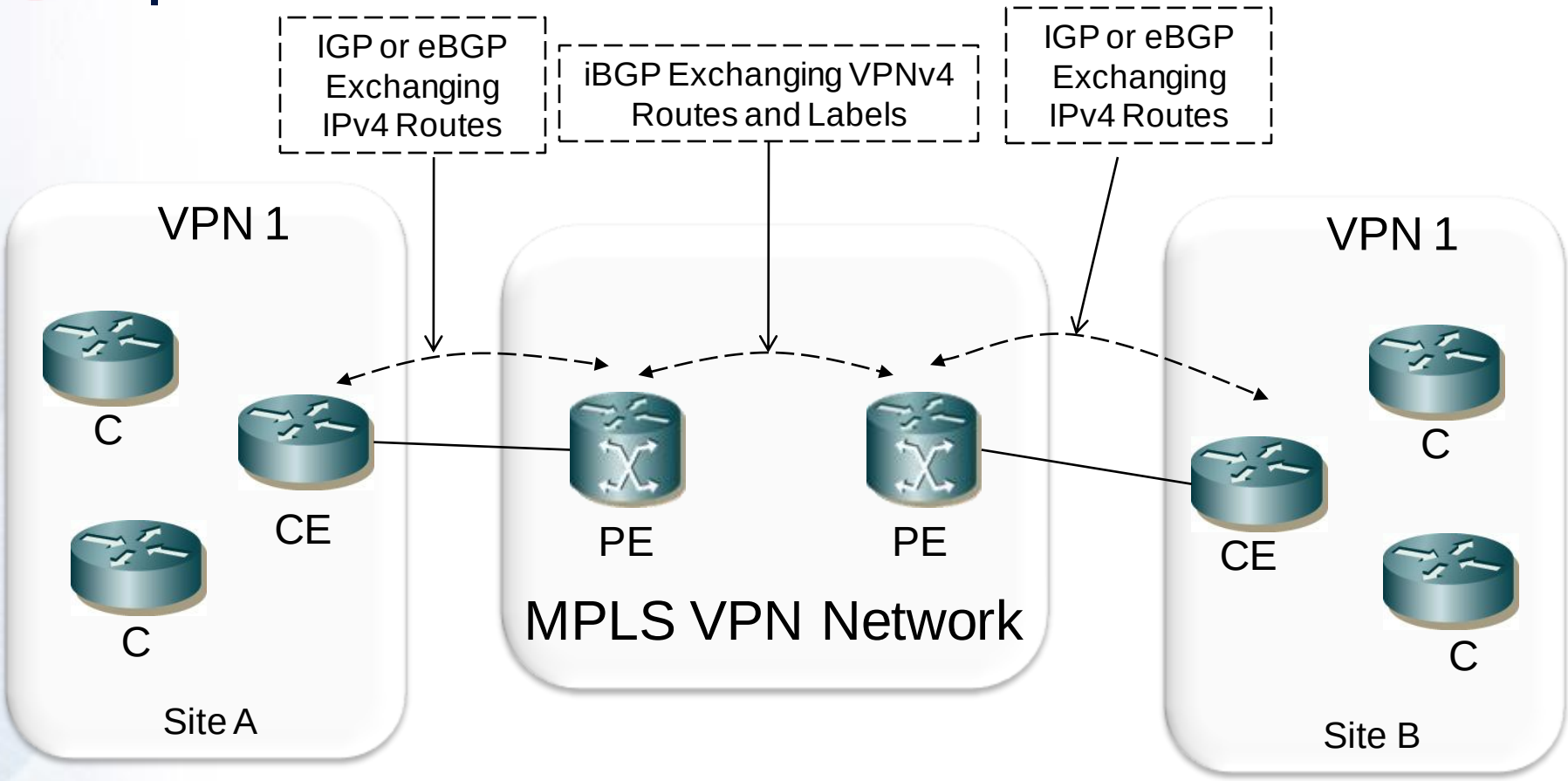
Exporting and Importing RTs



Extranet Example with RTs

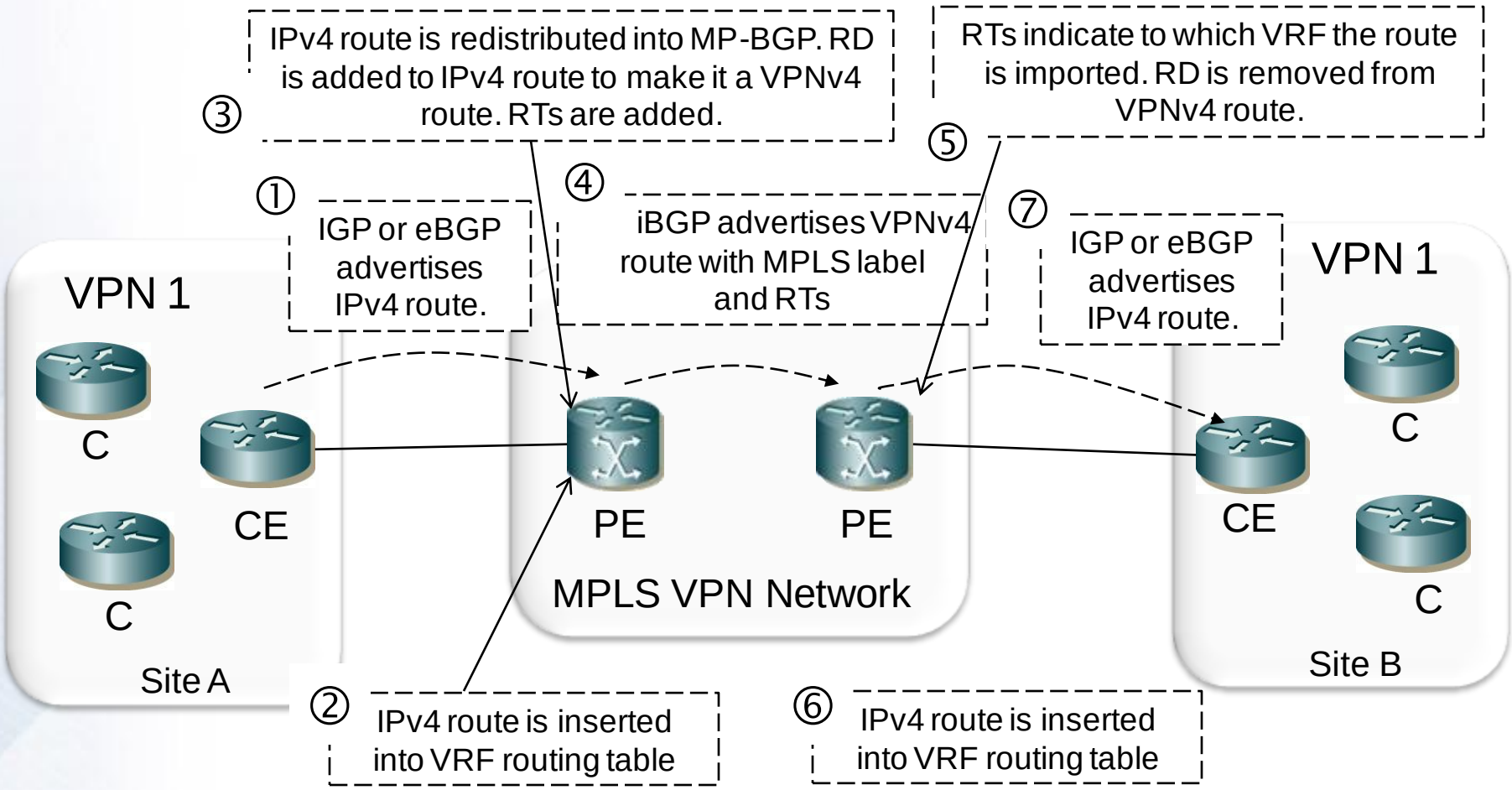


VPNv4 Route Propagation



Prefixes are transported inside MPLS Network → **BGP**

VPNv4 Route Propagation
In detail



Packet forwarding

- How does the egress PE router know which VRF the packet belongs to?
  - Not in IP Header or Label
- SOLUTION
  - Use another Label associated to VRF
  - All packets in MPLS VPN have two labels
    - IGP Label (top) and VRF Label (bottom)
- How egress PE signal ingress PE router the label to use for a VRF prefix?
  - MP-BGP signals VPN labels

Packet forwarding

- In Summary
  - VRF-to-VRF traffic has two labels in the MPLS VPN
  - Top label is the IGP label distributed by LDP or RSVP for TE between all P and PE routers
    - P routers use the IGP label to forward the packet to the correct egress PE router
  - Bottom label is the VPN label that is advertised by MP-iBGP from PE to PE
    - Egress PE router uses the VPN label to forward the IP packet to the correct CE router

BGP Review

- BGPv4 is an established inter-domain routing protocol
- eBGP used to peer with other ISPs
- iBGP run inside the ISP core
- Enabled to enforce policies
- iBGP is the best option to vpnv4 prefixes between PE routers

BGP Multiprotocol

- Original definition (RFC 1771) → For carrying IPv4 prefixes
- Extended (RFC 2858) → Multiprotocol Ext.
  - Negotiation: Routers capabilities exchanged

sydney-ce#show ip bgp neighbors
BGP neighbor is 10.10.4.1, **remote AS 1**, external link
BGP version 4, remote router ID 10.200.254.5
BGP state = Established, up for 00:00:37
Last read 00:00:30, hold time is 180, keepalive interval is 60 seconds
Neighbor capabilities:
Route refresh: advertised and received(new)
Address family IPv4 Unicast: advertised and received
ipv4 MPLS Label capability: advertised and received

BGP Extended Community RT

- The extended community is a optional BGP attribute (required for MPLS VPN)
  - Indicates to BGP speakers (PE routers) if the route should be imported into a VRF

london#show ip bgp vpnv4 all
BGP table version is 31, local router ID is 10.200.254.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
Route Distinguisher: 1:1 (default for vrf cust-one)					
*> 10.10.2.0/24	0.0.0.0	0		32768	?
*> 10.10.100.1/32	10.10.2.1	0		0	65001 i
*> 10.99.1.1/32	0.0.0.0	0		32768	?
Route Distinguisher: 2:2 (default for vrf cust-two)					
*> 10.140.1.1/32	0.0.0.0	0		32768	?

BGP carrying the label

- BGP advertises the vpnv4 prefixes in the MPLS VPN network
- PE Routers needs a ***new label*** to forward VPN traffic to the correct CE router
  - The label is simply **piggybacked** along with the vpnv4 prefix and advertised **by BGP**

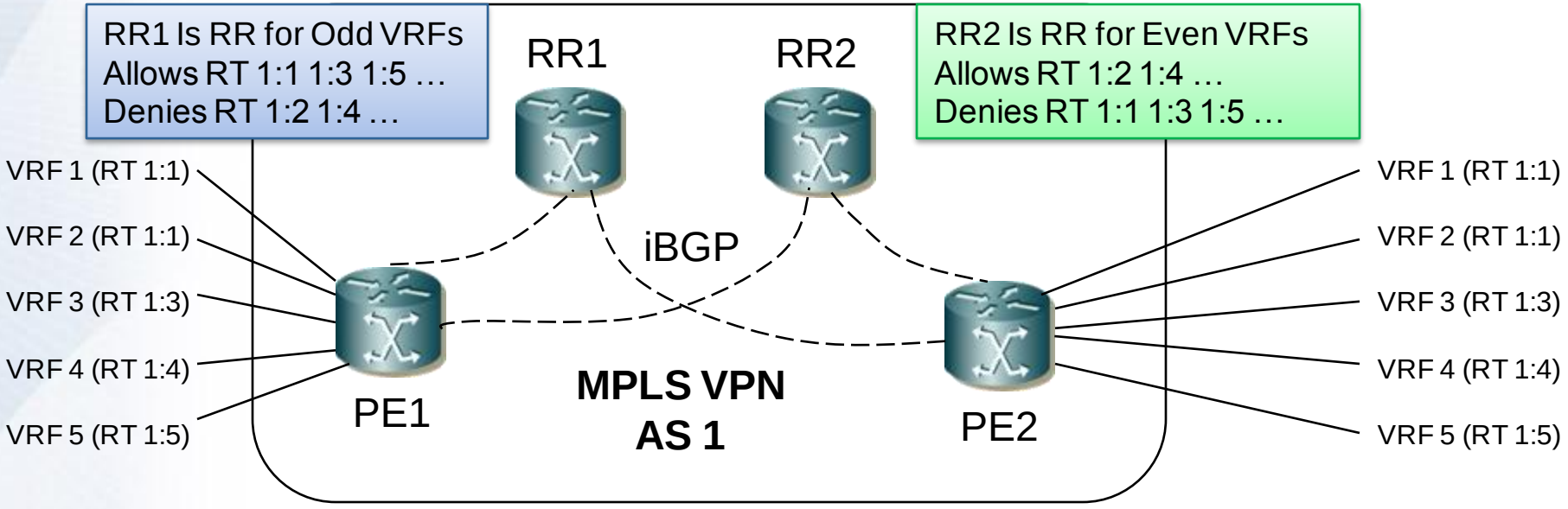
BGP label advertisement

```
sydney#show ip bgp vpnv4 rd 1:1 labels
Network          Next Hop        In label/Out label
Route Distinguisher: 1:1 (cust-one)
10.10.2.0/24     10.200.254.2    29/36
10.10.4.0/24     0.0.0.0         26/nolabel
10.10.4.2/32     0.0.0.0         37/nolabel
10.10.100.1/32   10.200.254.2    32/35
10.10.100.3/32   10.10.4.2       38/exp-null
10.88.1.1/32     10.200.254.2    34/34
10.99.1.1/32     10.200.254.2    28/33
10.99.1.2/32     0.0.0.0         27/nolabel
10.200.200.1/32  10.200.254.2    30/32
```

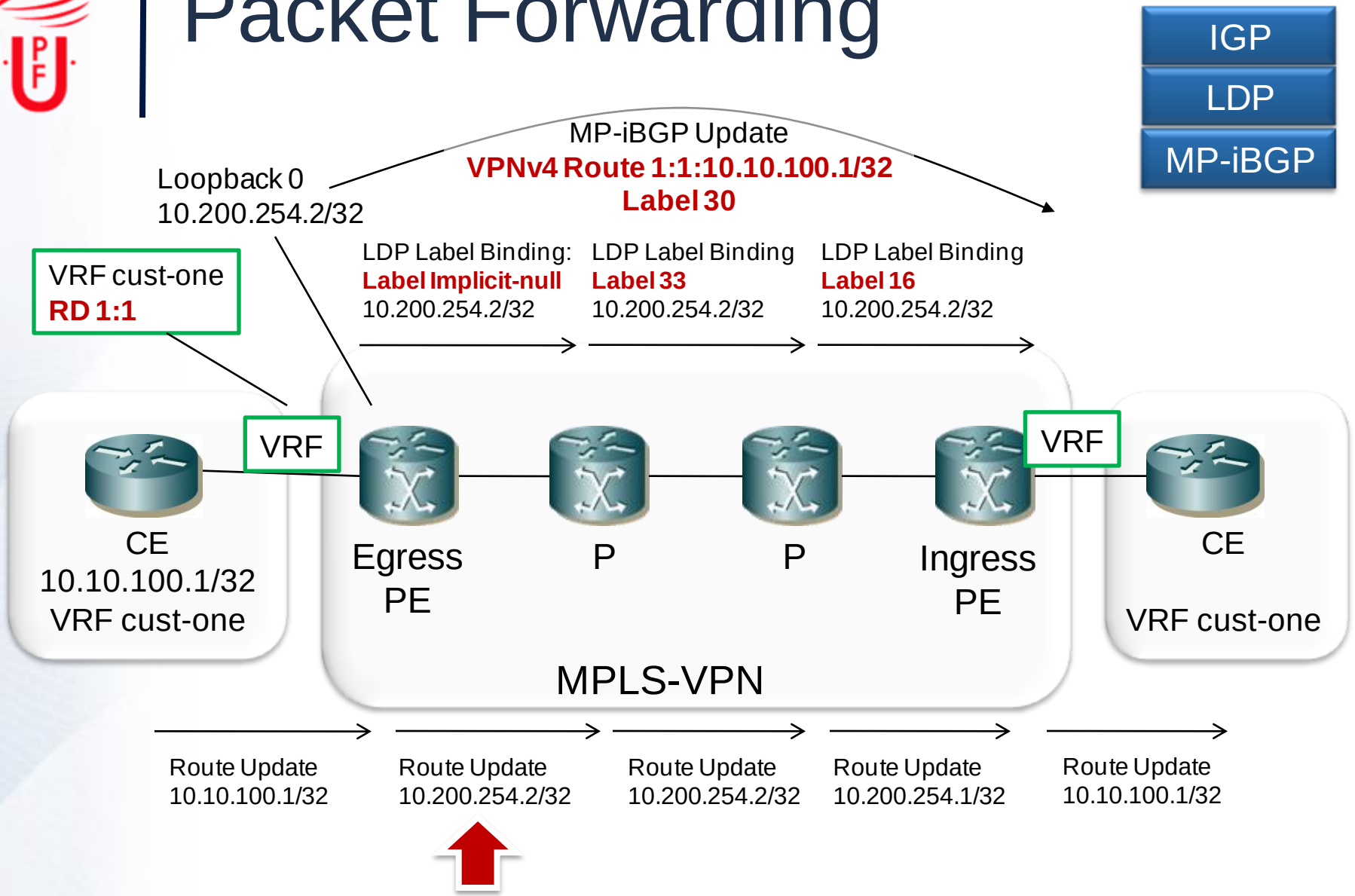
NOTE: Each vpnv4 prefix is assigned a unique MPLS label

Route Reflectors – RRs
Review Note

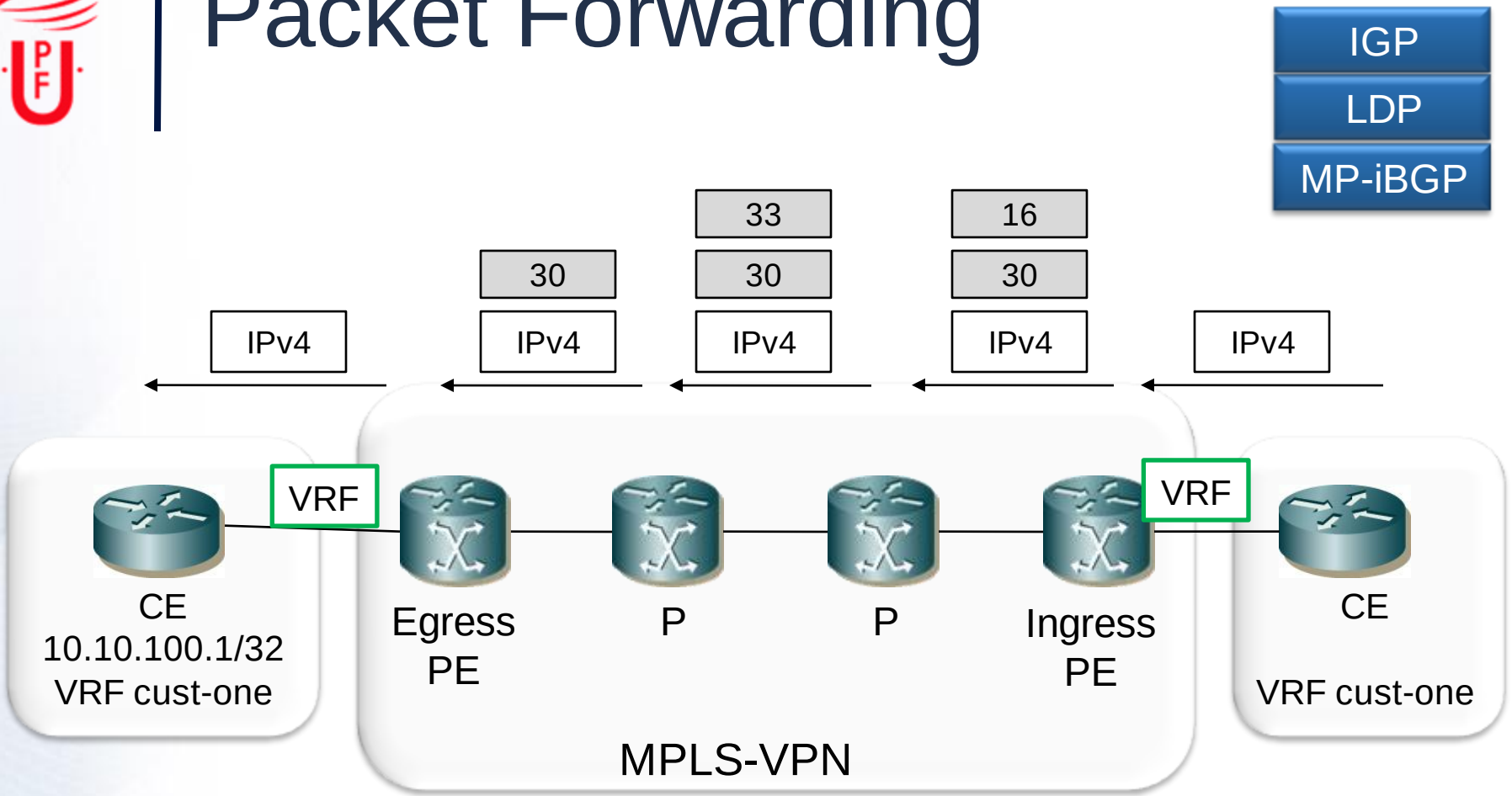
- An RR is a BGP speaker that reflects routes from other BGP speakers
- iBGP requires a full-mesh topology
- RR reflects BGP routes but not forward traffic



Packet Forwarding



Packet Forwarding

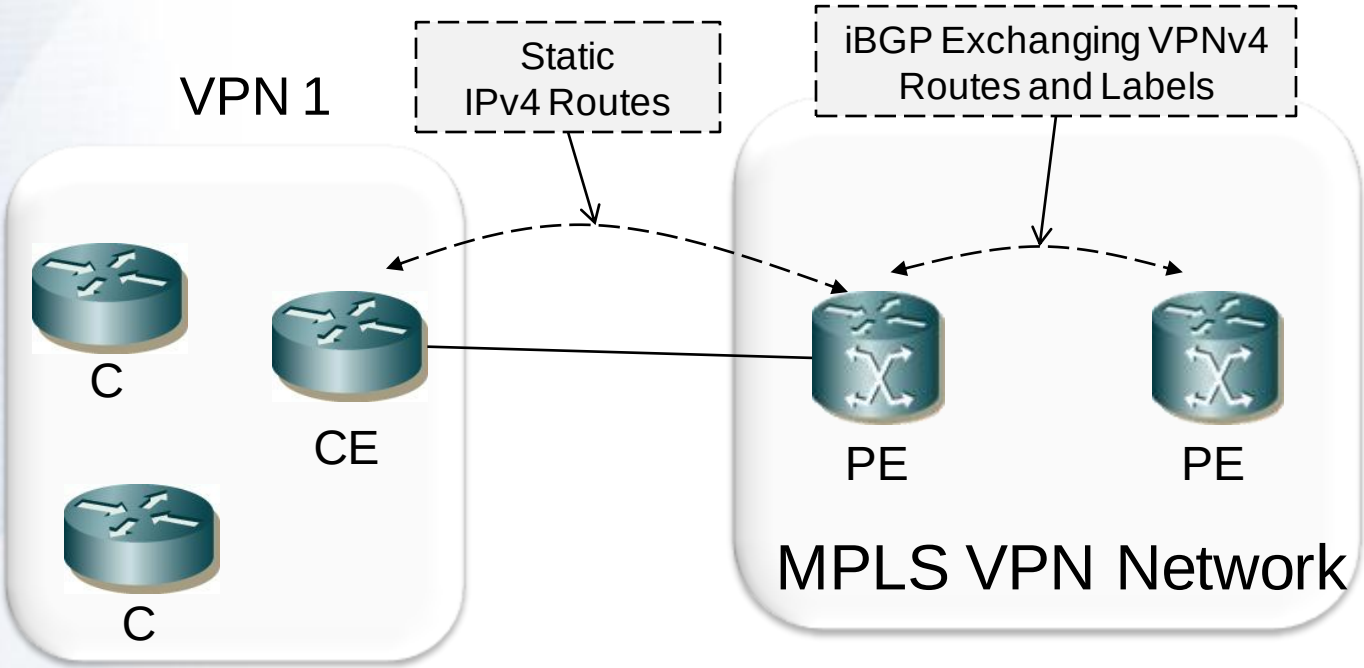


PE-CE Routing Protocols

- Routing is necessary between the PE and CE
  - Static routing
  - OSPF
  - EIGRP
  - IS-IS
  - eBGP

Static Routing PE-CE

- It is the simplest but it is tedious
- Redistribution of static routes (as vpnv4) into BGP → All PE learn them

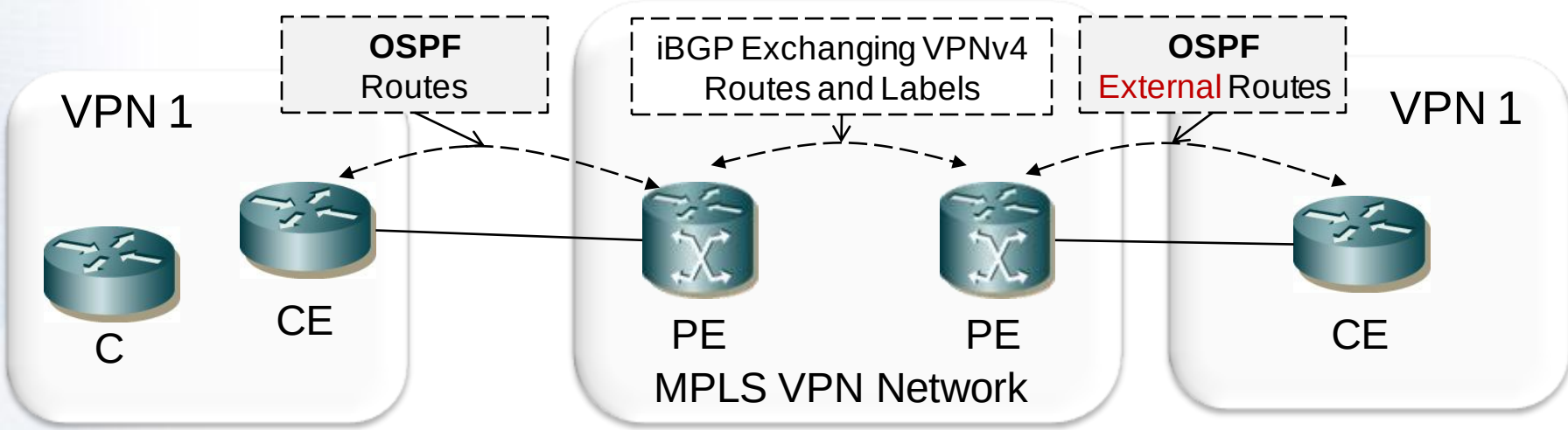


Static Routing PE-CE

- Used in environments where a customer site has a single connection to P-network and uses a single IP prefix
- Recommended in environments where the Service Provider needs tight control (some Central Services)
- Use default routes on CE routers in combination with static routes on PE routers
- Note: static routes increase the management burden on Service Provider

Dynamic Routing PE-CE
OSPF

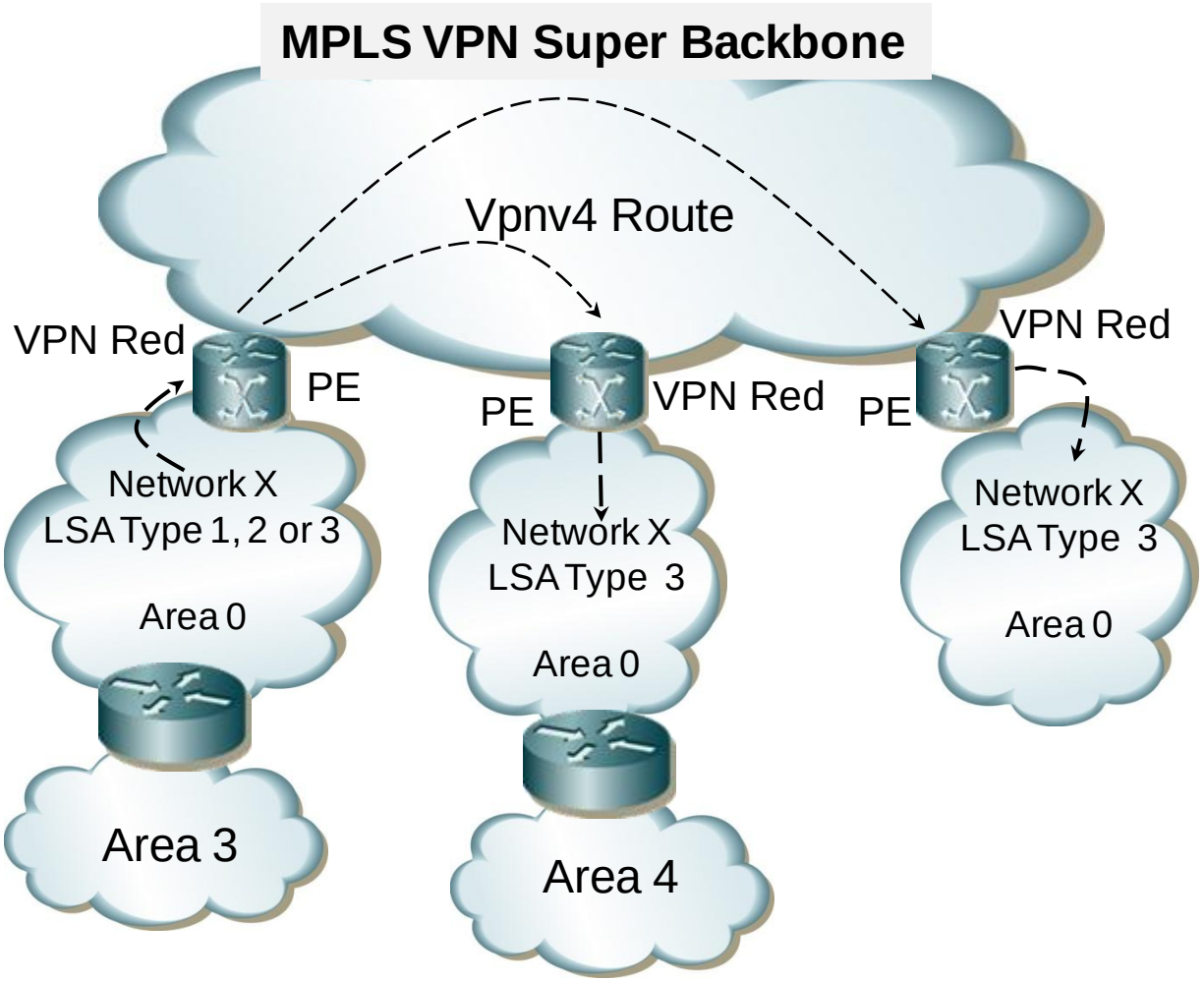
- Customer routes from PE to PE
  - OSPF is redistributed into iBGP and vice versa on the PE routers
- Use in cases where every CE router needs to know all of the routes



Dynamic Routing PE-CE
OSPF redistribution

- SOLUTION
  - **BUT** from PE to customer OSPF routes are marked as "*External*" → worse metric!!
  - Routes are advertised as "Summary routes" (LSA Type 3). i.e inter-area routes
  - This is not the normal way of working

Dynamic Routing PE-CE
OSPF redistribution



Dynamic Routing PE-CE
OSPF

- OSPF Metric Propagation
  - Redistribution from OSPF → MP-BGP on the PE
    - OSPF Metric used to set BGP MED (*external metric*)
    - BGP MED also use to redistribute MP-BGP→OSPF
- BGP Extended Communities for OSPF
  - Router type, Area number, OSPF Router ID Domain ID,

Dynamic Routing PE-CE
Recommendations

- RIPv2, OSPF and Exterior BGP are supported
- Use RIP for stub sites and when convergence is not an issue
- Use OSPF only as an exception
  - Very large customer network
  - Migrating existing large OSPF customer
- Use eBGP in complex PE-CE routing scenarios
  - Many routes exchanged between PE and CE
  - Multi-homed sites
  - No redistribution involved

MPLS VPN Topologies

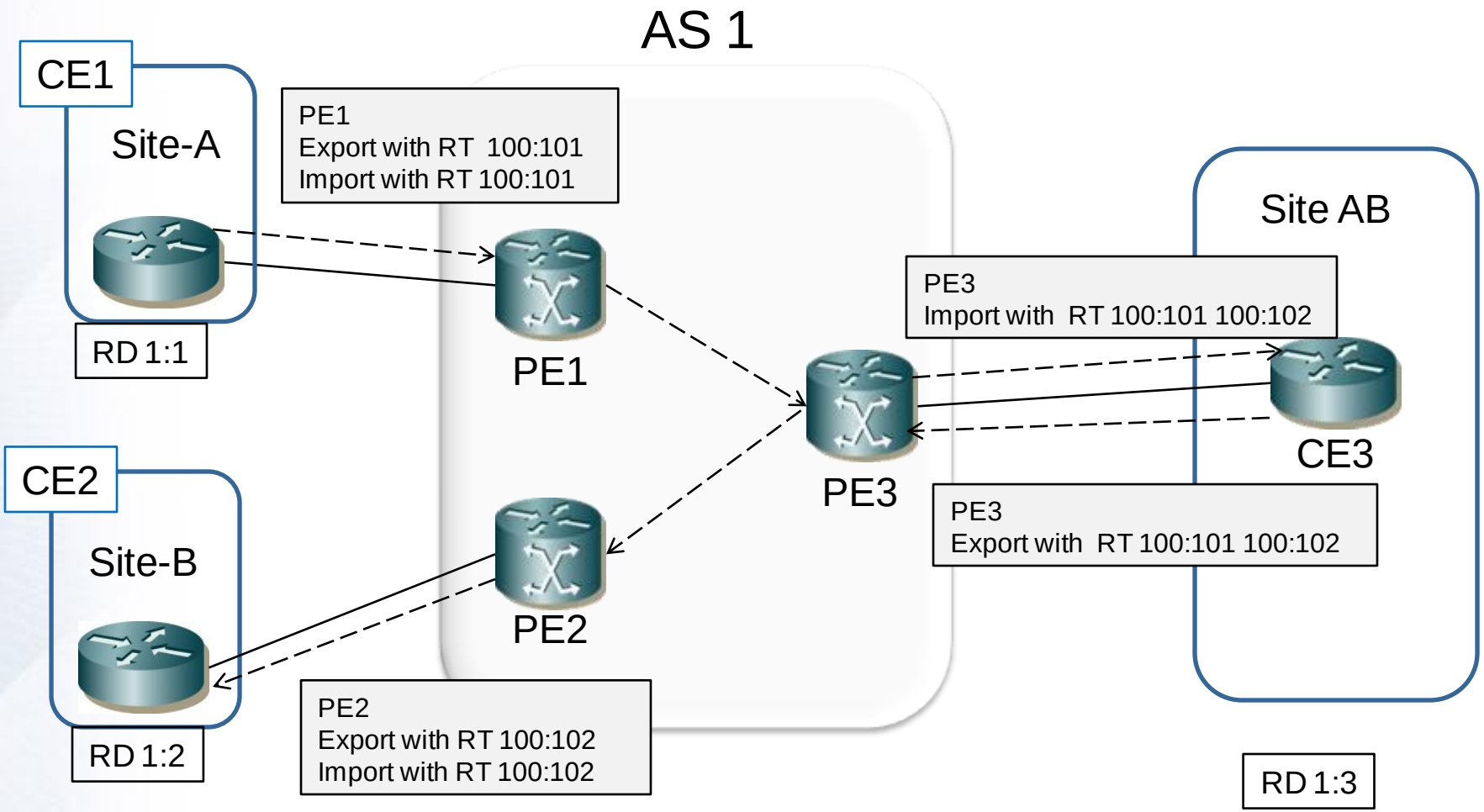
- Overlapping
- Central Site
- Hub-and-Spoke

MPLS VPN Topologies

- Overlapping scenario
  - Can be chosen for security reason
  - Companies where central sites participate in corporate network and in an extranet
  - Company with several security conscious departments that exchange data between their servers

MPLS VPN Topologies

Overlapping scenario

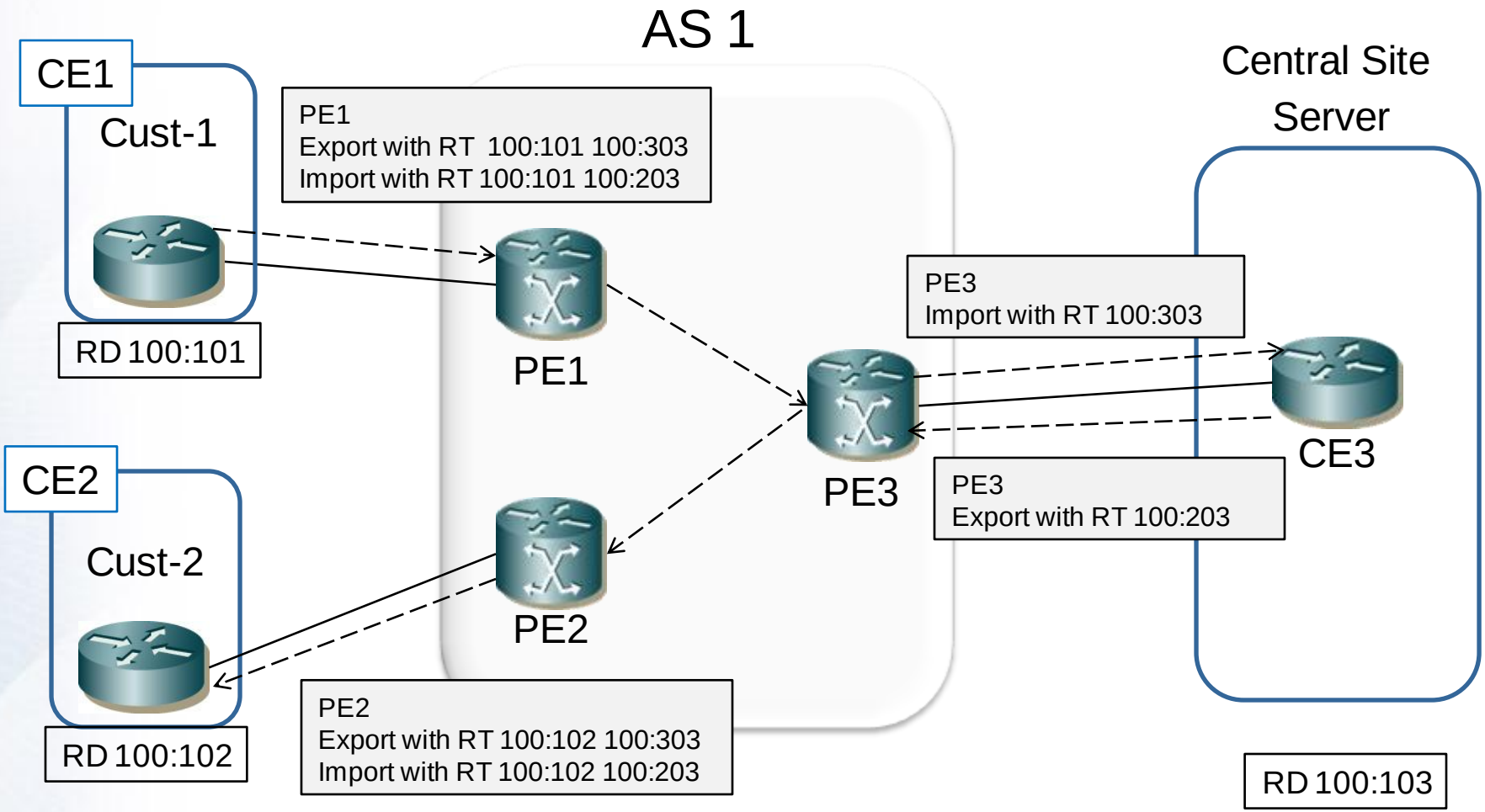


MPLS VPN Topologies

- Central Site
  - Clients need access to central servers
  - Servers can communicate with each other
  - Clients can communicate with all servers, but not with each other
  - Companies where central sites participate in corporate network and in an extranet
  - Company with several security conscious departments that exchange data between their servers

MPLS VPN Topologies

Central Site

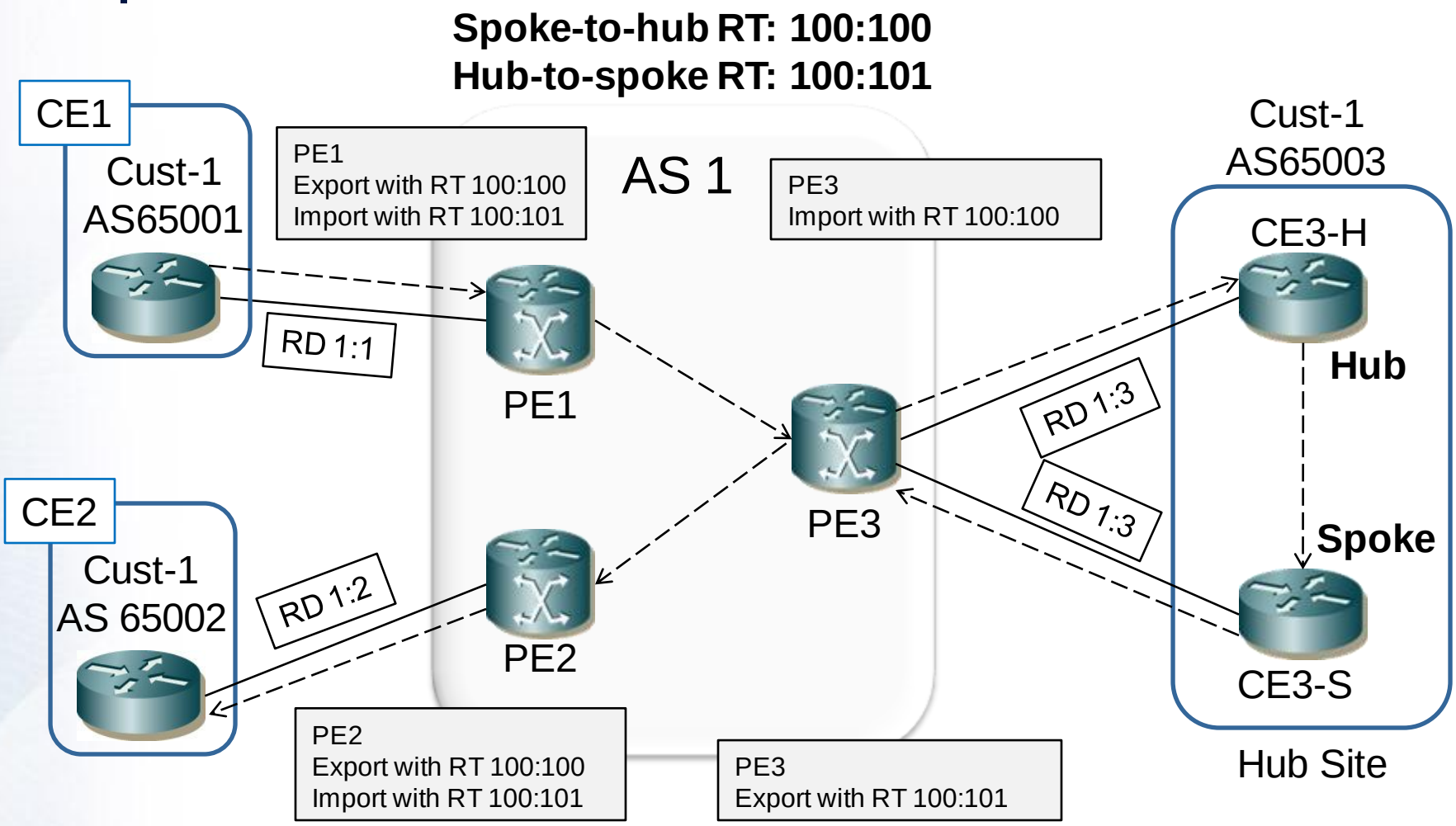


MPLS VPN Topologies

- Hub-and-Spoke scenario
  - One central site has full routing knowledge of all other sites of the same VPN
    - Hub-Site
  - Other sites will send traffic to the Hub-Site for any destination
    - Spoke-Sites
  - The Hub-Site is the central transit point between Spoke-Sites

MPLS VPN Topologies

Hub-and-Spoke



MPLS VPN - Internet Access

- Internet routing done via BGP table of the ISP
  - Global routing tables used for Internet routes
  - By default VRF and global routing separated
- **Internet in a VPN**
  - All Internet routes in VRFs
    - Bad decision → Huge number of routes
    - If all customers access same Internet VRF → Risk

MPLS VPN - Internet Access

- **Internet access through Global Routing**
  - PE-CE link for VPN traffic using VRF
  - PE-CE link for Internet using global routing
  - To avoid using two links
    - Sub-interfaces
    - GRE tunnel for the Internet link → default route

MPLS VPN - Internet Access

- **Internet access through Static Routing**
  - Forwarding Internet traffic to ISP gateway
  - Gateway known by all P routers through global routing tables
  - PE router use iBGP to peer ISP gateway

Review Questions

1. What is a route distinguisher?
2. How is a packet that is coming from the CE router identified as to which VRF it belongs?
3. What is the purpose of RTs? What is an RR group?
4. When would you use different route distinguishers for routes of the same VPN?
5. Why do MPLS VPN packets have two MPLS labels?

Test Questions

1. **How many bits long is a route distinguisher?**
   - A. 16
   - B. 32
   - C. 48
   - D. 64
2. **Which of the following protocols is used to propagate VPN labels between edge routers?**
   - A. TDP
   - B. LDP
   - C. Standard BGP with extended communities
   - D. MP-BGP
3. **To have a single router appear as many routers, which of the following mechanisms is used?**
   - A. RD
   - B. VPNv4
   - C. VPN
   - D. VRF

Test Questions

4. **In Multi-Protocol BGP (MP-BGP), neighbors need to be _____.**

   A. Configured
   B. Activated
   C. Sent standard communities
   D. Configured with VDP

5. **Which of the following mechanisms keeps overlapping addresses from doing so in MP-BGP?**

   A. RD
   B. VPNv4
   C. VPN
   D. VRF

6. **Which of the following is not a component of a VRF?**

   A. VRF-specific routes
   B. CEF
   C. Global routing table
   D. None of the above

Test Questions

7. **MPLS VPNs offer _____ security as traditional overlay VPNs.**
   - A. The same
   - B. Worse
   - C. Better
   - D. None of the above

8. **P routers _____ knowledge of a customer's VPN routes.**
   - A. Do have
   - B. Do not have

9. **A(n) _____ imposes the VPN label.**
   - A. LSR
   - B. LSP
   - C. Edge-LSR
   - D. None of the above

Test Questions

10. **What types of routes are in the PE router's global routing table?**
    - A. Customer routes
    - B. Service provider routes
    - C. Customer and service provider routes
    - D. None of the above
11. **What types of routes are in the PE router's VRF for a particular customer?**
    - A. Customer routes
    - B. Service provider routes
    - C. Customer and service provider routes
    - D. None of the above
12. **MP-BGP within an AS is called _____.**
    - A. MP-BGP
    - B. MP-IBGP
    - C. MP-EBGP
    - D. MP-MBGP

Table of Contents

Goals of the section

- ☑ The purpose and architecture of AToM
- ☑ The Layer 2 encapsulation types that can be carried across the MPLS backbone
- ☑ How to implement Ethernet over MPLS

AToM

- Understanding the need for AToM
- Transporting Layer 2 Frames
- AToM Architecture
- Transported Layer 2 Protocol

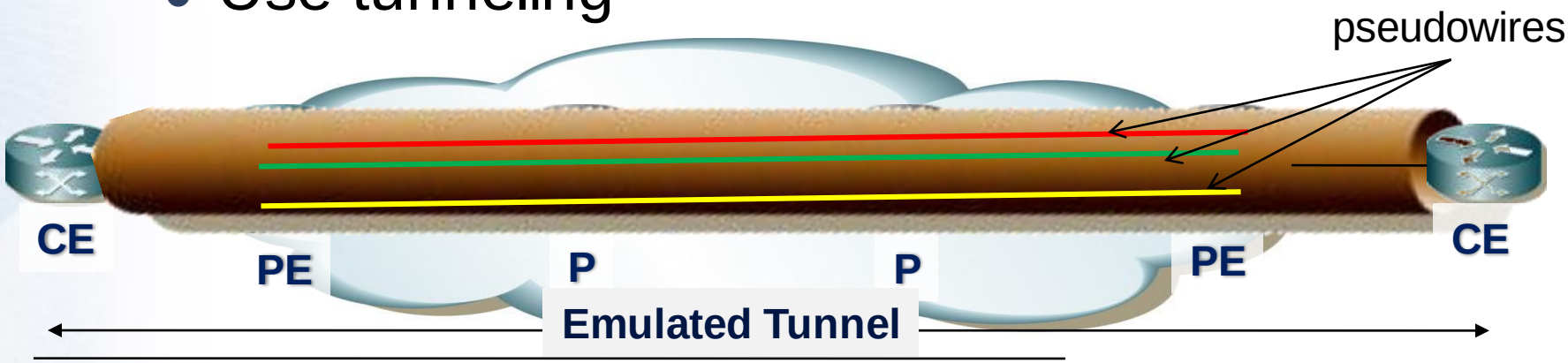
- Review Questions

Motivation

- MPLS VPN work over shared MPLS services
- Legacy leased lines, ATM and F/R still in use
- AToM
  - Limited to Layer 2 point-to-point services: *Virtual Private Wire Service* (**VPWS**)
  - Intelligence limited to the PE (edge technology)
  - Core MPLS P routers do not need extra conf.
  - Allows MPLS VPN (L3VPN) with legacy technologies (L2VPN) using the same infrastructure
  - Customers have full control of their network

Transporting Layer 2 Frames

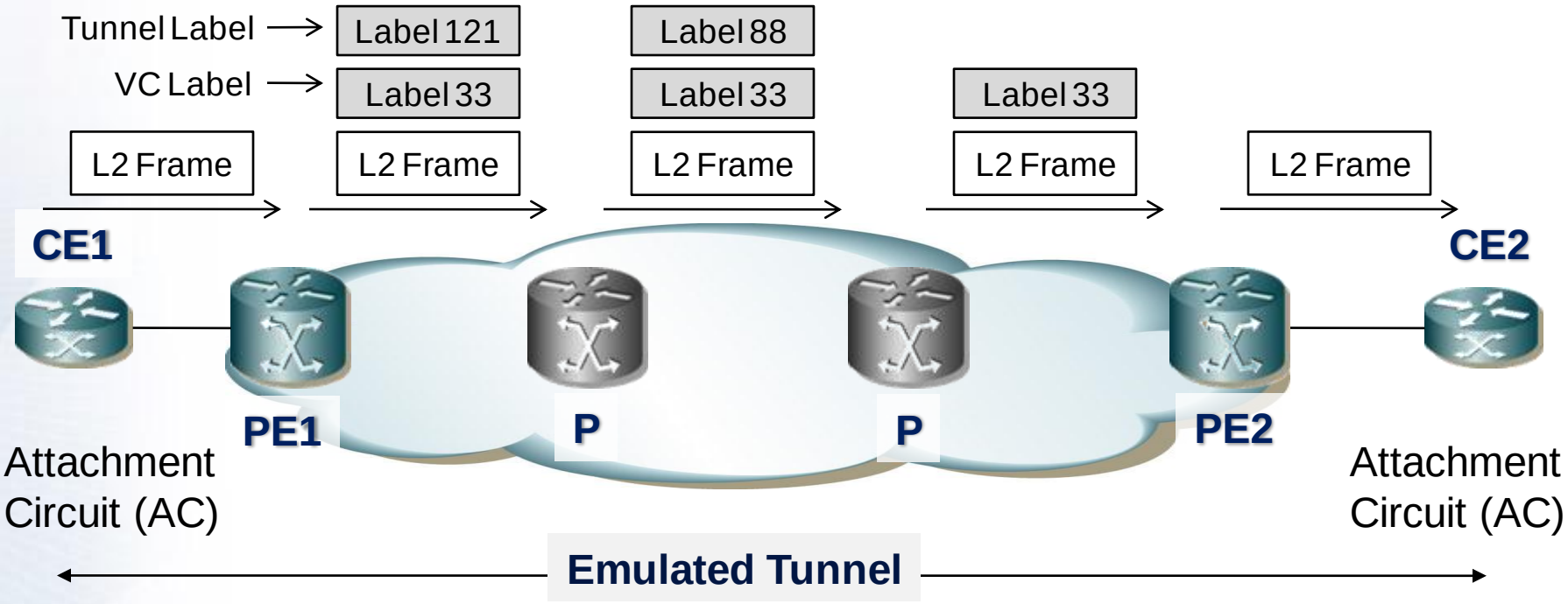
- Transporting L2 frames by:
  - Carry traffic across MPLS backbone (AToM)
  - Carry traffic across IP backbone (L2TPv3)
- AToM based on *pseudowires*
  - Connection PE-PE emulating a wire
  - Use tunneling



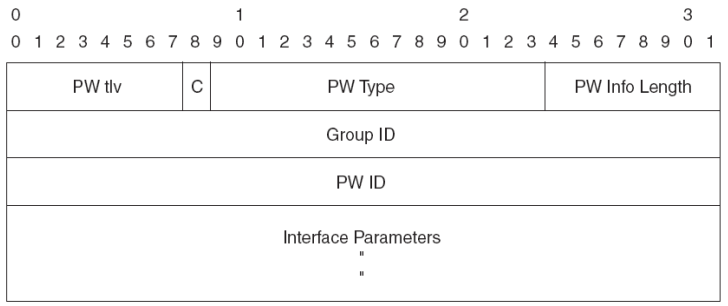
AToM Architecture

- **Core (**PE-PE): Common MPLS infrastructure with LDP or RSVP
- **Border** (PE-CE): Attachment circuits (AC)
  - AC can be F/R, ATM, HDLC, PPP, Ethernet
- **LSPs**: Defined for each direction →Tunnels
  - Inside each tunnel → Several pseudowires
  - Use of additional labels to identify pseudowires
  - Several pseudowires multiplexed in a tunnel
- Labels: 1 for tunnel + 1 for pseudowire

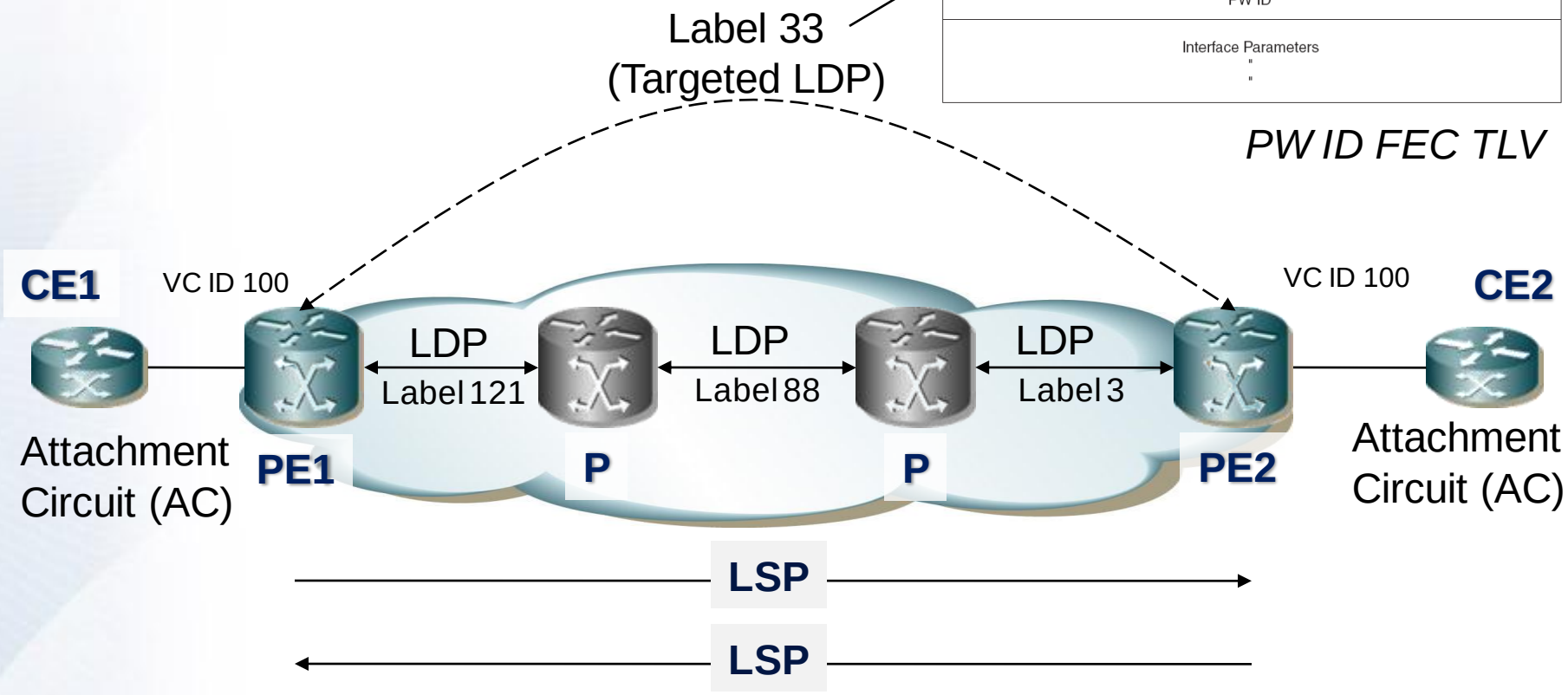
Data Plane of AToM



Signaling the pseudowires



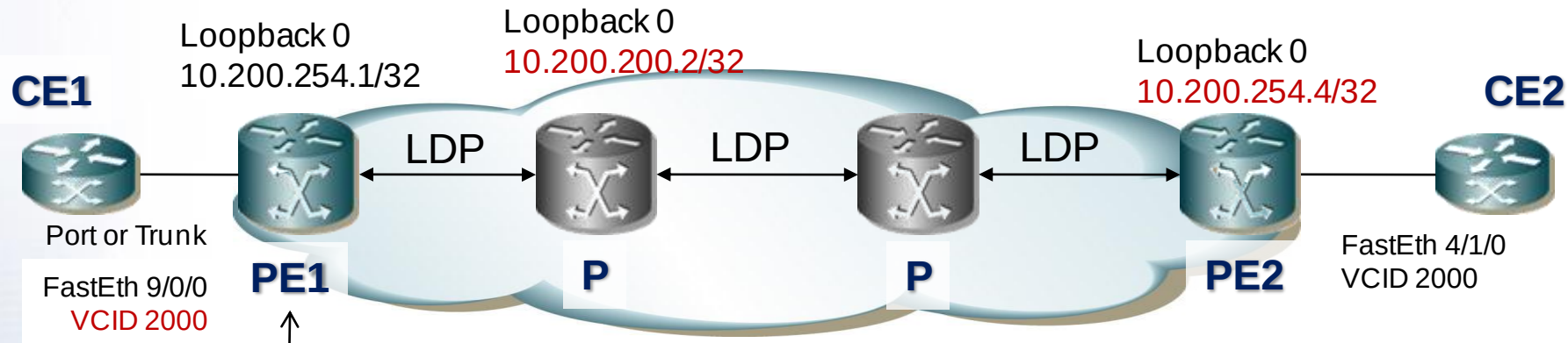
PW ID FEC TLV



Ethernet over MPLS
EoMPLS

- AToM solution for Ethernet is strictly point-to-point
- Equivalent to LAN-to-LAN bridging over point-to-point WANs
- VLAN header can be carried over the MPLS network transparently
- LAN-like solution over MPLS → VPLS
- Two modes can be signaled:
  - Port mode
  - VLAN mode

Scenarios – Simple Ethernet
EoMPLS



Label 23

Label 35

L2 Frame

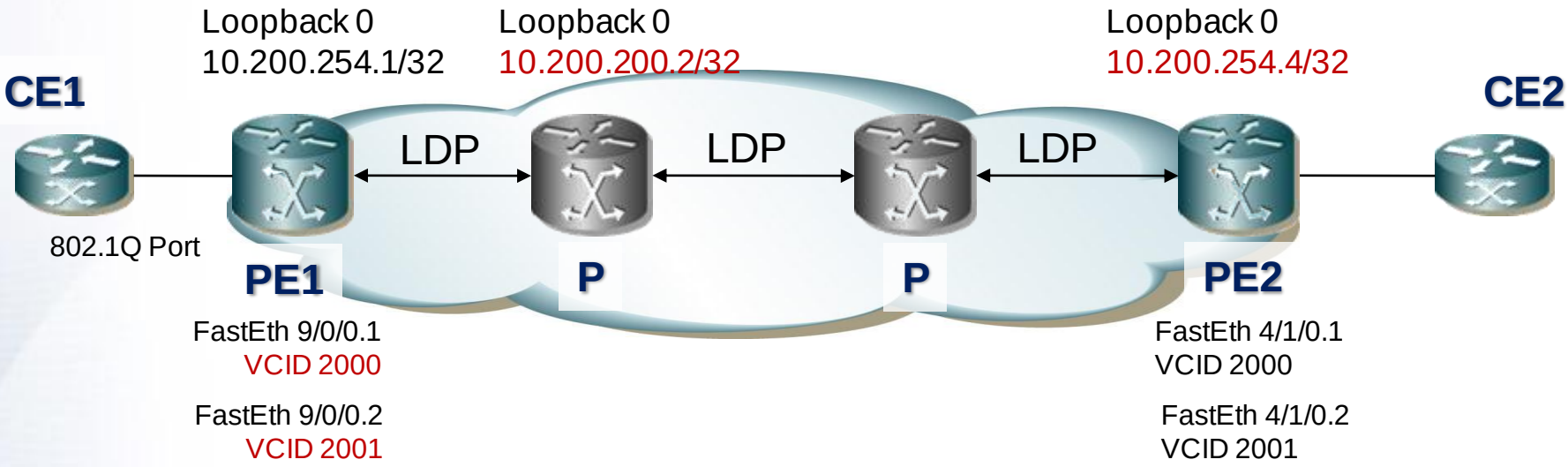
PE1#show mpls l2transport vc 2000 detail
Local interface: Fa9/0/0 up, line protocol up, Ethernet up
Destination address: 10.200.254.4, VC ID: 2000, VC status: up
Preferred path: not configured
Default path: active
Tunnel label: 23, next hop **10.200.200.2**
Output interface: Et0/0/0, imposed label stack **{23 35}**
Create time: 00:02:26, last status change time: 00:02:26
Signaling protocol: LDP, **peer 10.200.254.4:0** up
MPLS VC labels**: local 25, remote 35**
Group ID: local 0, remote 0
MTU: local 1500, remote 1500

Label ?

Label 25

L2 Frame

Scenarios – VLAN Ethernet
EoMPLS



PE1#show mpls l2transport vc

Local intf	Local circuit	Dest address	VC ID	Status
Fa9/0/0.1	Eth VLAN 100	10.200.254.4	2000	UP
Fa9/0/0.2	Eth VLAN 200	10.200.254.4	2001	UP

Dot1q Tunneling over AToM QinQ

- Double tagging (QinQ) → VLAN in VLAN
- Many VLANs customer on 1 VLAN provider

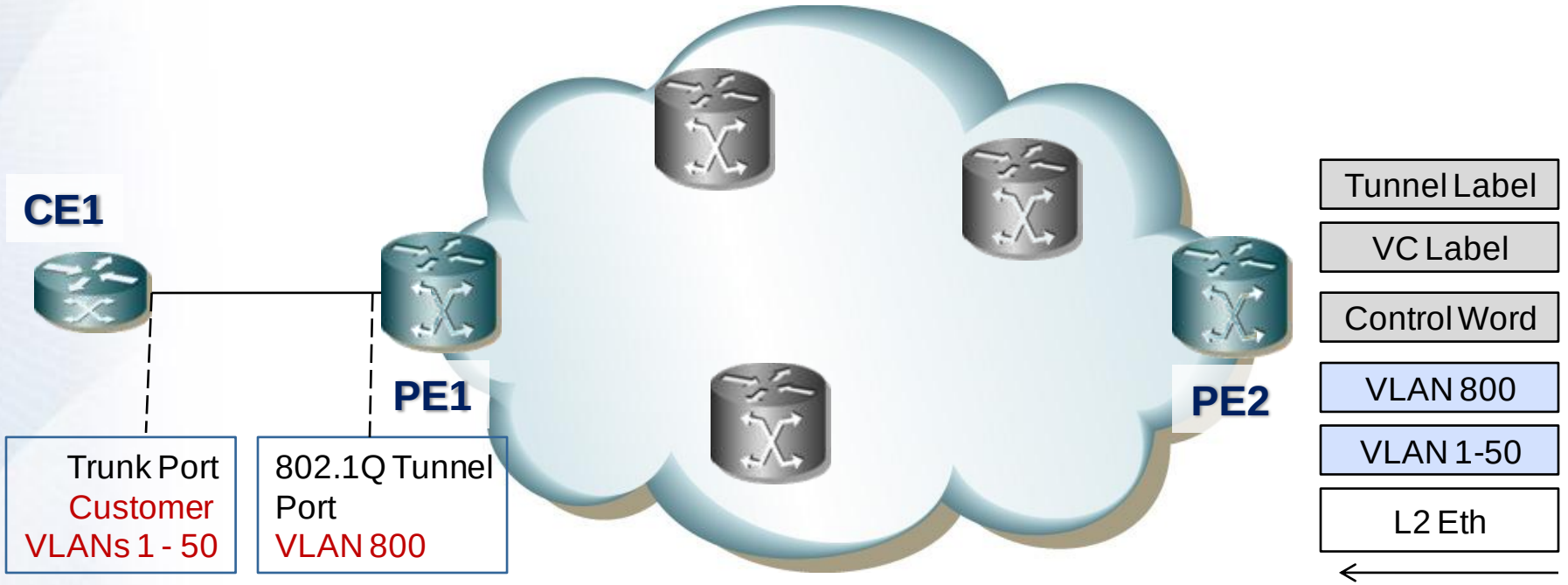


Table of Contents

- MPLS VPN
- Basics AoTM
- **VPLS**

- Test Questions

Goals of the section

- ☑ Explain what VPLS stands for
- ☑ Explain how VPLS emulates an Ethernet switched network over MPLS

VPLS

- Need for VPLS
- VPLS Architecture
- VPLS Forwarding
- H-VPLS

- Review Questions

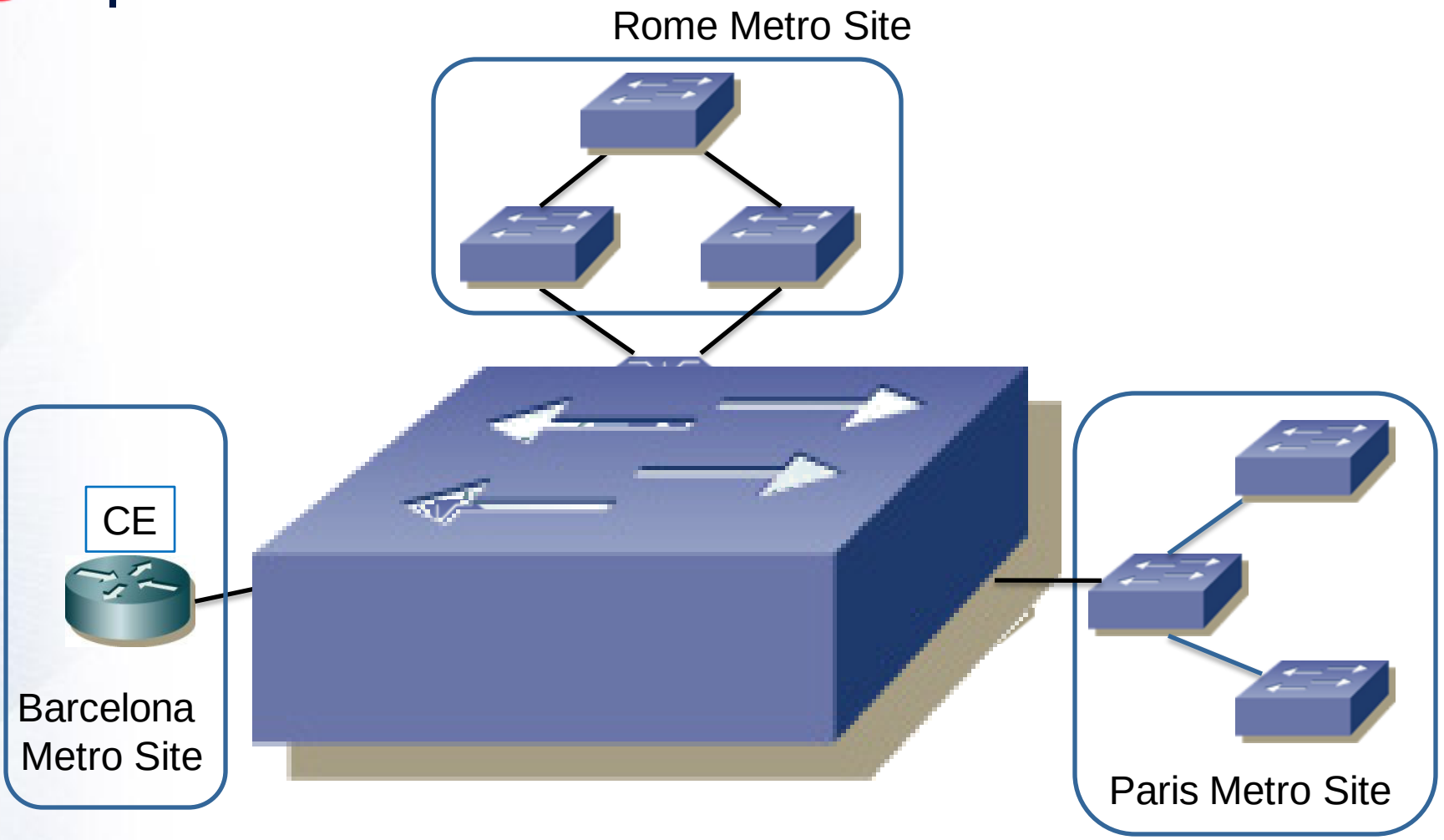
Introduction to VPLS

- VPLS emulates LAN segment across MPLS backbone using PW
- Each LAN is completely separated
- Spanning Tree Protocol (STP)
  - Option 1: end at PE
  - Option 2: crosses MPLS Backbone (virtual switch)
- Topology point-to-multipoint
- Ethernet features:
  - MAC Address learning, broadcast, multicast

Why deploy VPLS?

Feature	Benefits
Point-to-multipoint multiprotocol services	MPLS is only focused in IP. AToM can carry L2 frames (EoMPLS) but only delivers point-to-point services
MPLS core network emulates a flat LAN segment	**Overcomes distance limitations of Ethernet-switched networks** → Offer Virtual Private LAN Services *Formerly called Transparent LAN Services (TLS)*
Extends Ethernet broadcast capability across WAN → Point to Multipoint Connectivity	**Connects each customer site to many or all other customer sites** – A single CE-PE link transmits Ethernet packets to multiple remote CE routers – Fewer connections required to get full connectivity among customer sites
Multipoint plug-and-play Provisioning	Adding, removing or relocating a CE router requires configuring only the directly attached PE router

VPLS: Logical Bridge



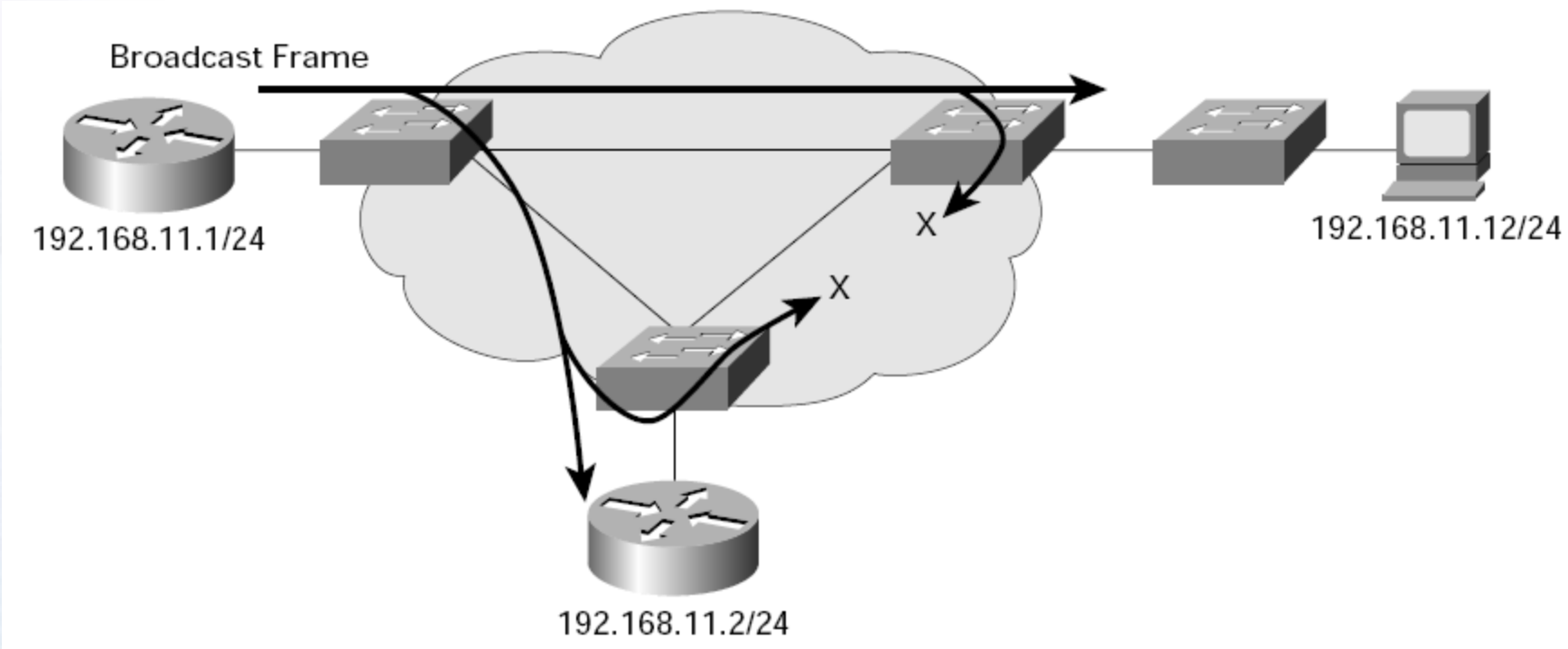
VPLS Characteristics

- Forwarding of Ethernet frames
- Forwarding of unicast frames with an unknown destination MAC address
- Replication of broadcast and multicast frames to more than one port
- Loop prevention (*split horizon*)
- Dynamic learning of MAC addresses
- MAC address aging

VPLS Characteristics

- Flooding / Forwarding:
  - MAC table instances per customer and per customer VLAN (similar to L3-VRF) for each PE
- Address Learning / Aging:
  - Self Learn Source MAC to port associations
  - Refresh MAC timers with incoming frames
- Loop Prevention:
  - Create partial or full-mesh of EoMPLS VCs per VPLS
  - Use “split horizon” concepts to prevent loops
  - Announce EoMPLS VPLS VC tunnels

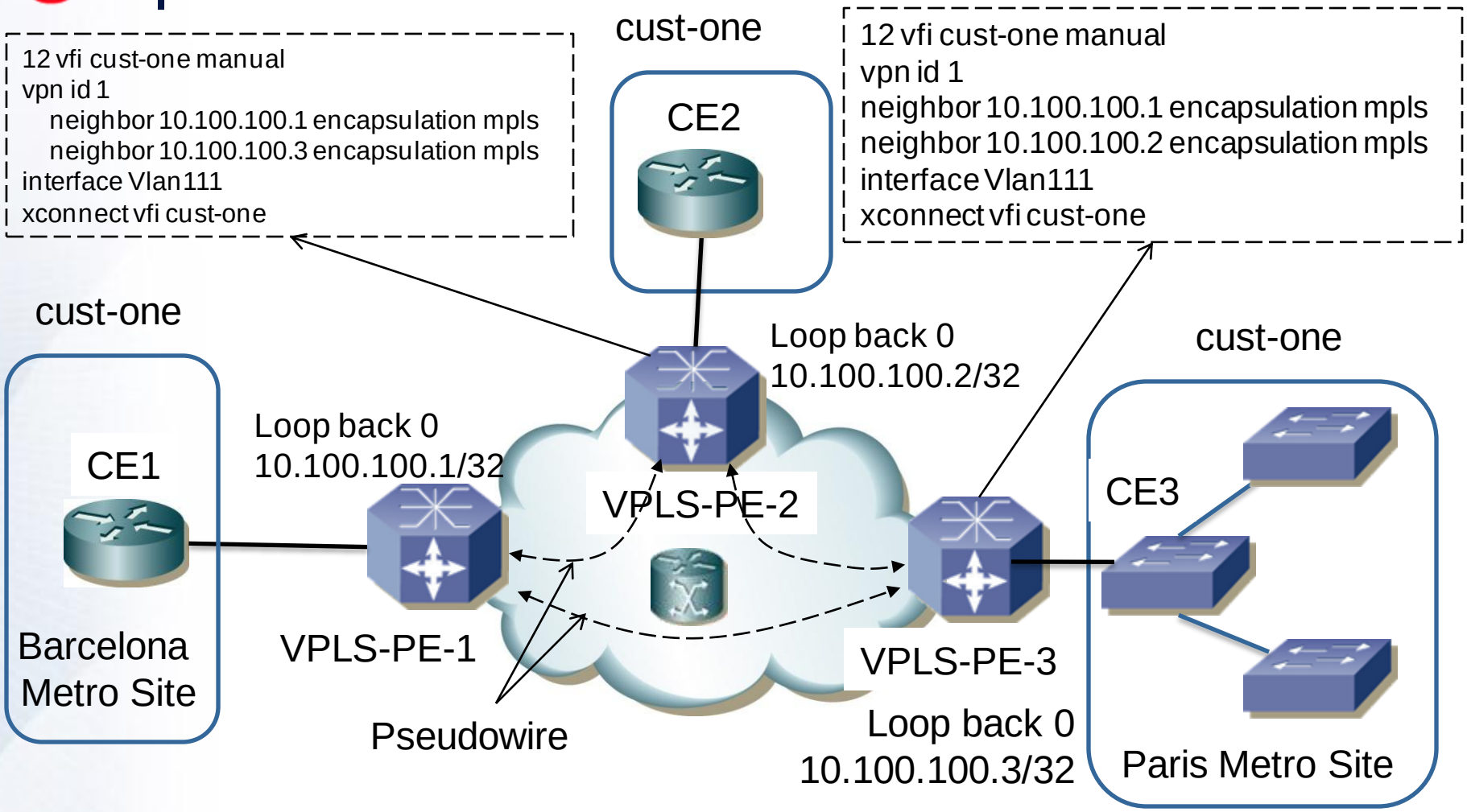
Split horizon
Review



VPLS Signaling

- Full-mesh of *pseudowires* between PE's for each VPLS instance
  - PE neighbors must be defined
  - Targeted sessions established → VC & PW
- If VPLS assigned to VLAN on PE
  - Local VC ID assigned to VPLS instance

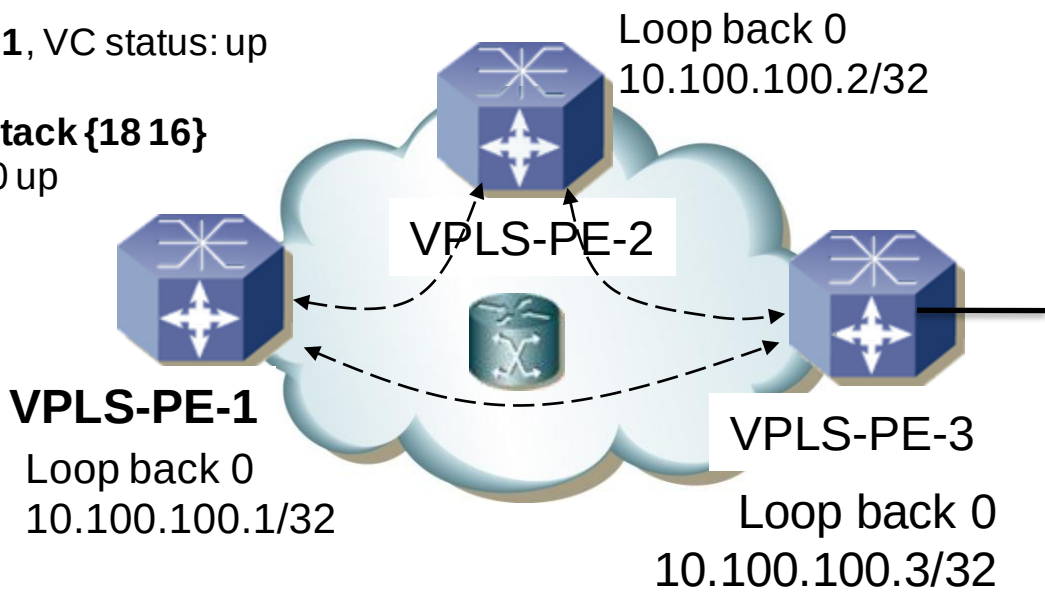
Example VPLS Signaling



Example VPLS Signaling

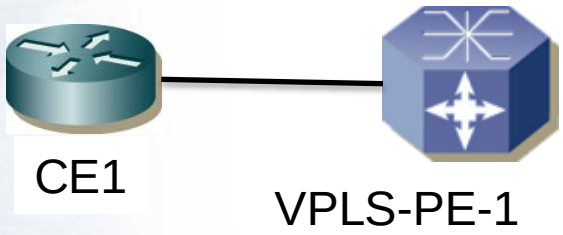
VPLS-PE-1#show mpls l2transport vc 1 detail
Local interface: VFI cust-one up
Destination address: 10.100.100.2, VC ID: 1, VC status: up
Tunnel label: 17, next hop point2point
Output interface: PO5/1, **imposed label stack {17 18}**
Signaling protocol: LDP, peer 10.100.100.2:0 up
MPLS VC labels: local 16, remote 18

Local interface: VFI cust-one up
Destination address: 10.100.100.3, VC ID: 1, VC status: up
Tunnel label: 18, next hop point2point
Output interface: PO5/1, **imposed label stack {18 16}**
Signaling protocol: LDP, peer 10.100.100.3:0 up
MPLS VC labels: local 17, remote 16



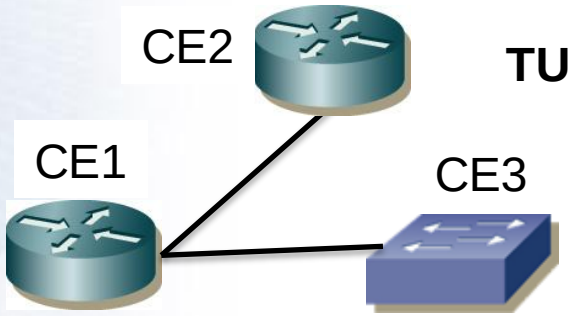
Tunneling L2 Protocols

- It is possible to tunnel L2 protocols
  - Protocols transparently tunnel without PE participating (e.g. STP, VTP, CDP)



CE1#show cdp neighbors

Device ID	Local Intrfce	Holdtme	Capability	Platform	Port ID
VPLS-PE-1	Fas 2/2	175	R S I	WS-C6506	Fas 4/2



TUNNELED CONFIGURATION ON

CE1#show cdp neighbors

Device ID	Local Intrfce	Holdtme	Capability	Platform	Port ID
CE3	Fas 2/2	146	S I	C2950-2	Fas 0/9
CE2	Fas 2/2	150	R	C10720	Fas 2/6

CDP = Cisco Discovery Protocol

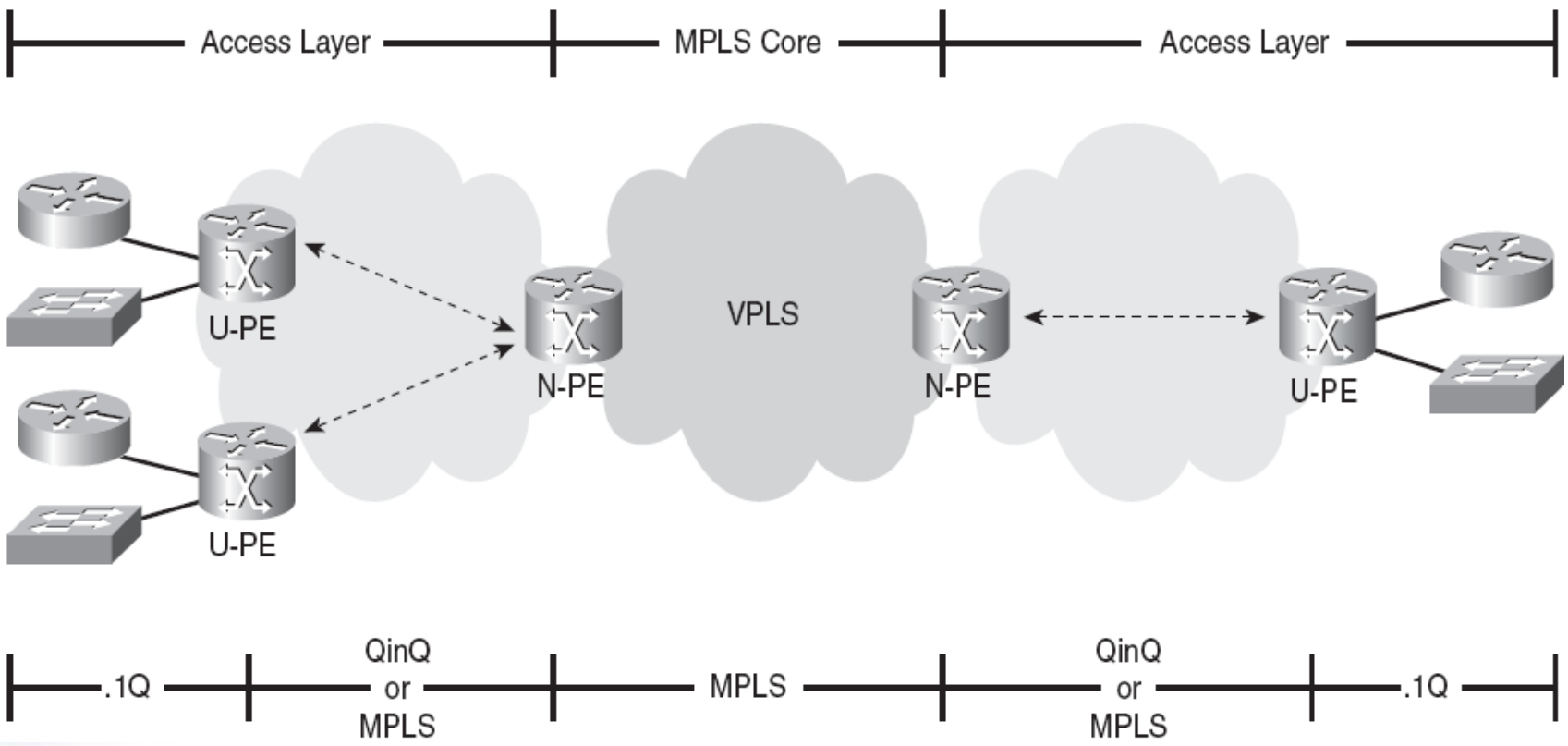
Hierarchical VPLS

H-VPLS

- The PE routers are no longer directly attached to the customer equipment
- Hierarchy introduced by adding another layer in the access layer toward the CE
- H-VPLS types
  - H-VPLS with dot1q tunneling in the access layer
  - H-VPLS with MPLS in the access layer

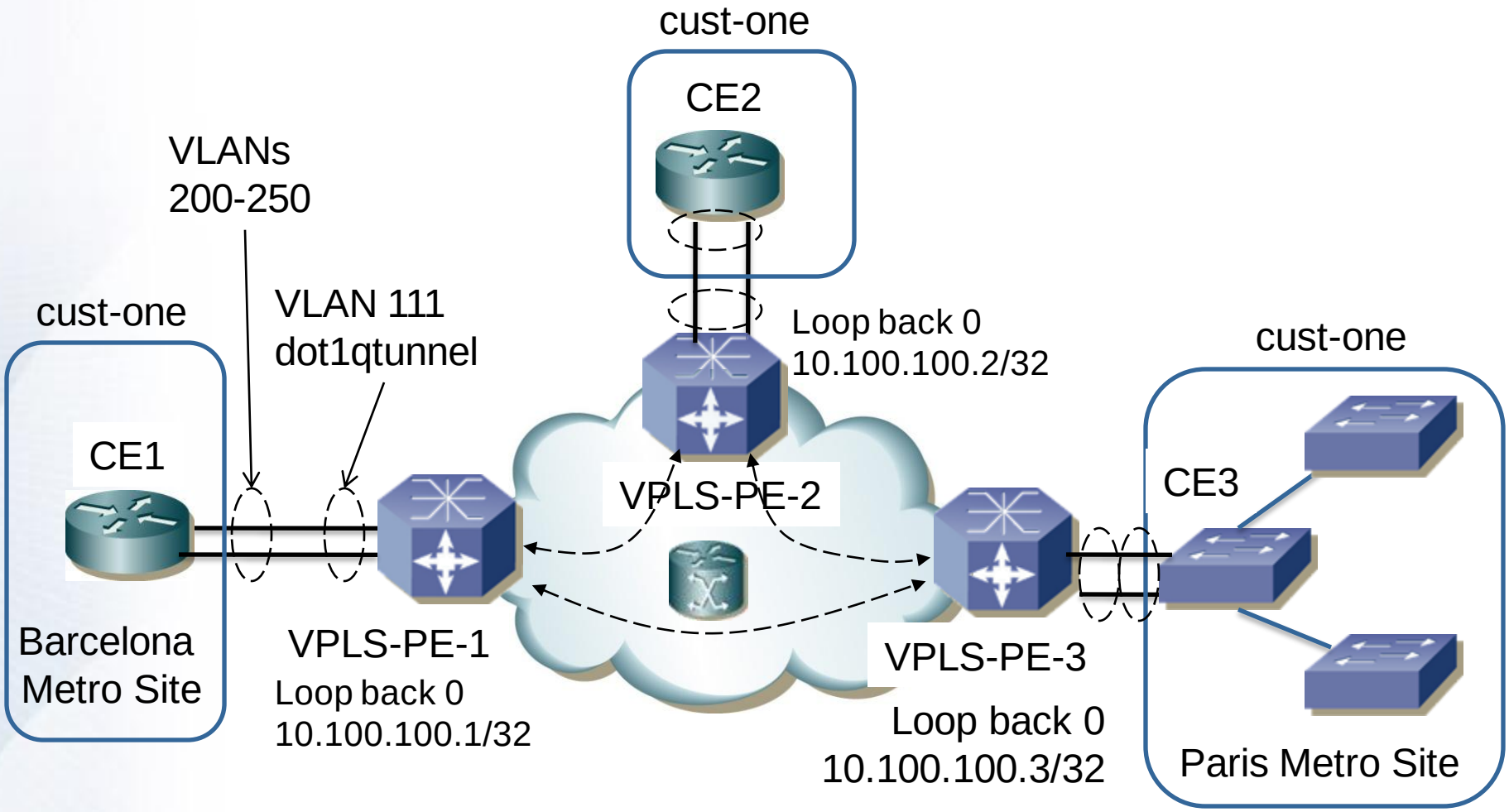
Hierarchical VPLS

H-VPLS



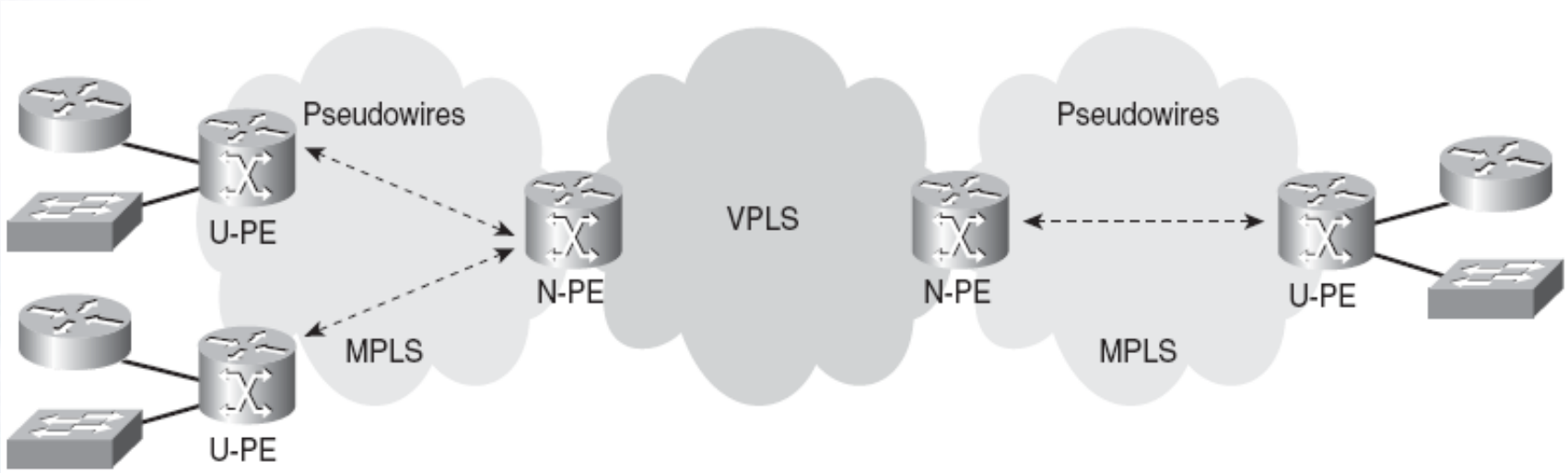
H-VPLS with dot1q

H-VPLS



H-VPLS with MPLS

H-VPLS



Review Questions

1. How many labels are used to forward VPLS traffic, and what is the use of each of those labels?
2. Which Layer 2 control protocols can be tunneled across the VPLS network? (name at least 2)
3. Why do the PE routers need to be in a full mesh of pseudowires in VPLS?
4. Name the six functions that VPLS performs in emulating an Ethernet switch.
5. In which two ways can H-VPLS be implemented?
6. Are there any differences in the encapsulation of Ethernet frames across the packet network between VPLS and Any Transport over MPLS (AToM)?

ANSWER TO REVIEW QUESTIONS

Review Questions

1. What is a route distinguisher?
2. How is a packet that is coming from the CE router identified as to which VRF it belongs?
3. What is the purpose of RTs? What is an RR group?
4. When would you use different route distinguishers for routes of the same VPN?
5. Why do MPLS VPN packets have two MPLS labels?

Test Questions

1. **How many bits long is a route distinguisher?**
   - A. 16
   - B. 32
   - C. 48
   - D. 64
2. **Which of the following protocols is used to propagate VPN labels between edge routers?**
   - A. TDP
   - B. LDP
   - C. Standard BGP with extended communities
   - D. MP-BGP
3. **To have a single router appear as many routers, which of the following mechanisms is used?**
   - A. RD
   - B. VPNv4
   - C. VPN
   - D. VRF

Test Questions

4. **In Multi-Protocol BGP (MP-BGP), neighbors need to be _____.**

   A. Configured
   B. Activated
   C. Sent standard communities
   D. Configured with VDP

5. **Which of the following mechanisms keeps overlapping addresses from doing so in MP-BGP?**

   A. RD
   B. VPNv4
   C. VPN
   D. VRF

6. **Which of the following is not a component of a VRF?**

   A. VRF-specific routes
   B. CEF
   C. Global routing table
   D. None of the above

Test Questions

7. **MPLS VPNs offer _____ security as traditional overlay VPNs.**
   - A. The same
   - B. Worse
   - C. Better
   - D. None of the above

8. **P routers _____ knowledge of a customer's VPN routes.**
   - A. Do have
   - B. Do not have

9. **A(n) _____ imposes the VPN label.**
   - A. LSR
   - B. LSP
   - C. Edge-LSR
   - D. None of the above

Test Questions

10. **What types of routes are in the PE router's global routing table?**
    - A. Customer routes
    - B. Service provider routes
    - C. Customer and service provider routes
    - D. None of the above
11. **What types of routes are in the PE router's VRF for a particular customer?**
    - A. Customer routes
    - B. Service provider routes
    - C. Customer and service provider routes
    - D. None of the above
12. **MP-BGP within an AS is called _____.**
    - A. MP-BGP
    - B. MP-IBGP
    - C. MP-EBGP
    - D. MP-MBGP

Review Questions

1. How many labels are used to forward VPLS traffic, and what is the use of each of those labels?
2. Which Layer 2 control protocols can be tunneled across the VPLS network? (name at least 2)
3. Why do the PE routers need to be in a full mesh of pseudowires in VPLS?
4. Name the six functions that VPLS performs in emulating an Ethernet switch.
5. In which two ways can H-VPLS be implemented?
6. Are there any differences in the encapsulation of Ethernet frames across the packet network between VPLS and Any Transport over MPLS (AToM)?