# Introduction to MPLS-based VPNs

Ferit Yegenoglu, Ph.D.
ISOCORE

ferit@isocore.com

Internetworking 2003

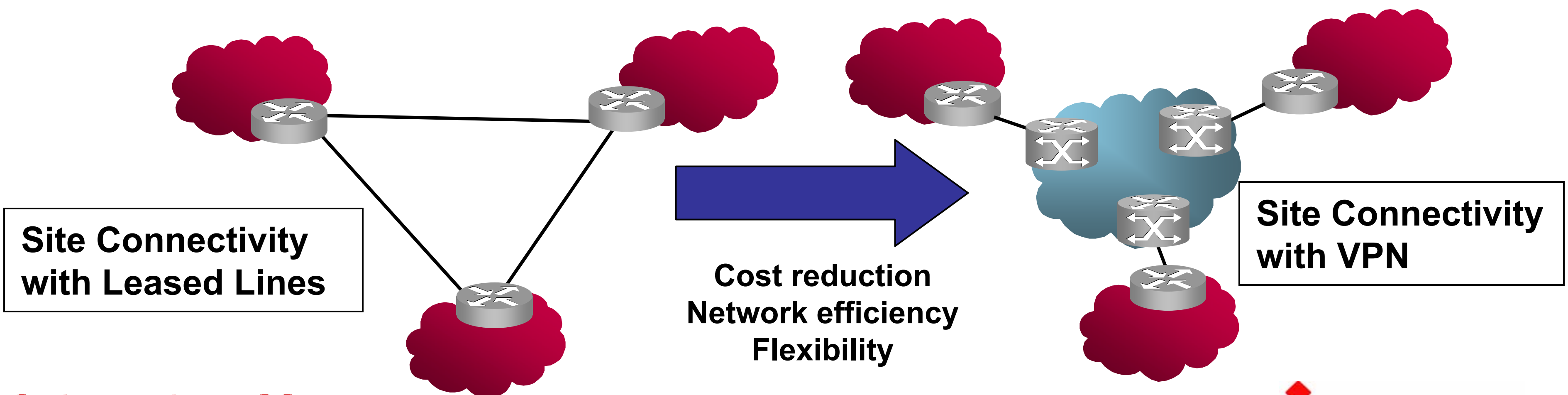
ISOCORE

# Outline

- Introduction
- BGP/MPLS VPNs
  - Network Architecture Overview
  - Main Features of BGP/MPLS VPNs
  - Required Protocol Extensions
  - Route Distribution and Packet Forwarding
  - Building Different VPN Topologies
  - Hierarchical BGP/MPLS VPNs
  - Security
- Layer 2 VPNs
  - Point-to-point
  - Point-to-multipoint

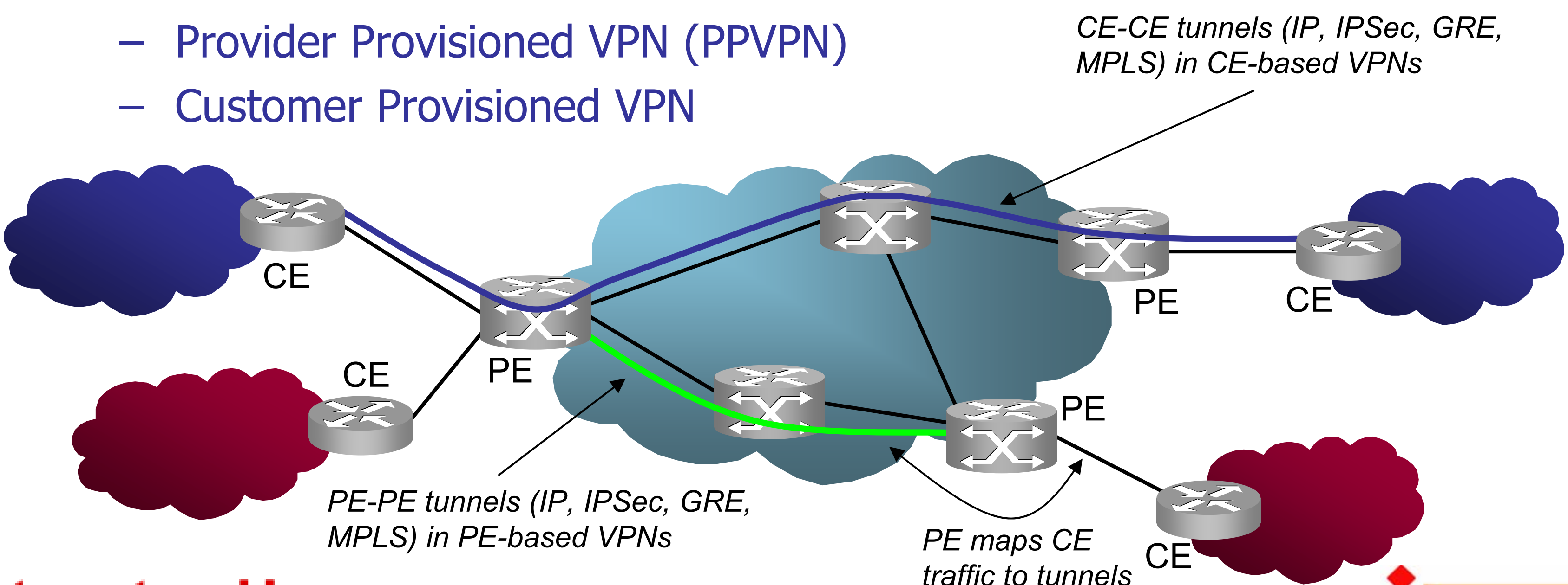
# Virtual Private Networks (VPNs)

- Virtual
  - Emulated connectivity over a public network
- Private
  - Access limited to VPN members
  - Total address and route separation
- Network
  - A collection of customer sites



# Classification of IP VPNs

- Classification based on where VPN functions are implemented
  - Customer Edge (CE) – based VPN
  - Provider Edge (PE) – based VPN
- Classification based on service provider's role in provisioning the VPN
  - Provider Provisioned VPN (PPVPN)
  - Customer Provisioned VPN



# Classification of IP VPNs (2)

- Classification based on protocol layer

  - Layer 2 VPNs
    - SP network switches customer Layer-2 frames based on Layer-2 header
    - SP delivers layer 2 circuits to the customer, one for each remote site
    - Customer maps their layer 3 routing to the circuit mesh
    - Customer routes are transparent to provider

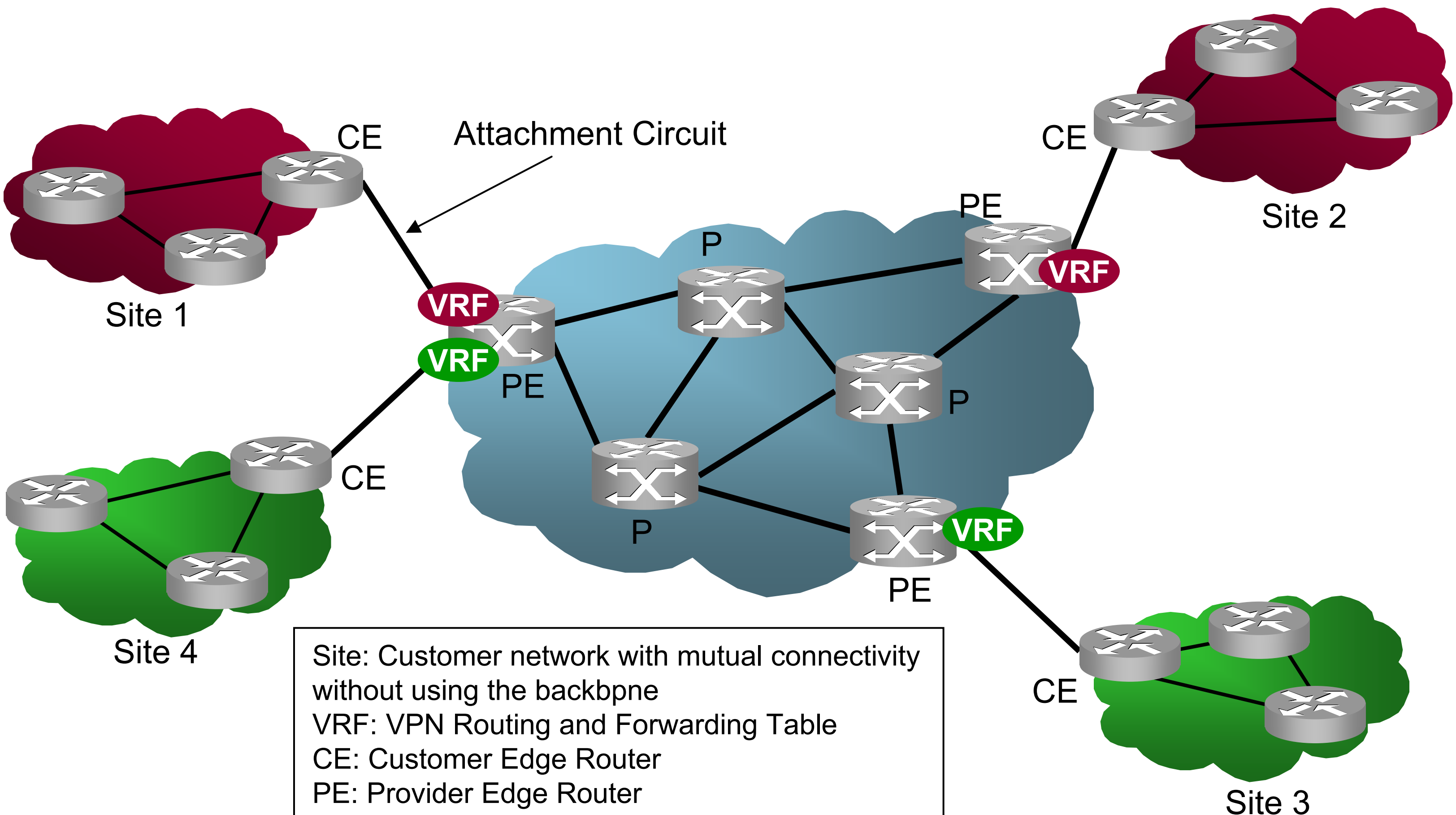
  - Layer 3 VPNs
    - SP network routes incoming customer packets based on the destination IP address
    - SP network participates in customer's layer 3 routing
    - SP network manages VPN-specific routing tables, distributes routes to remote sites
    - CPE routers advertise their routes to the provider

# MPLS-based VPNs

- MPLS can provide the required tunneling mechanism
  - MPLS can be used to provide traffic engineered PE-PE tunnels
  - An additional MPLS label can also used to associate packets with a VPN
- Layer 3 MPLS-based VPNs
  - BGP/MPLS VPNs (RFC 2547bis)
- Layer 2 MPLS-based VPNs
  - Virtual Private Wire Service (VPWS)
  - Virtual Private LAN service (VPLS)

# BGP/MPLS – Based VPNs

# BGP/MPLS VPN Network Overview



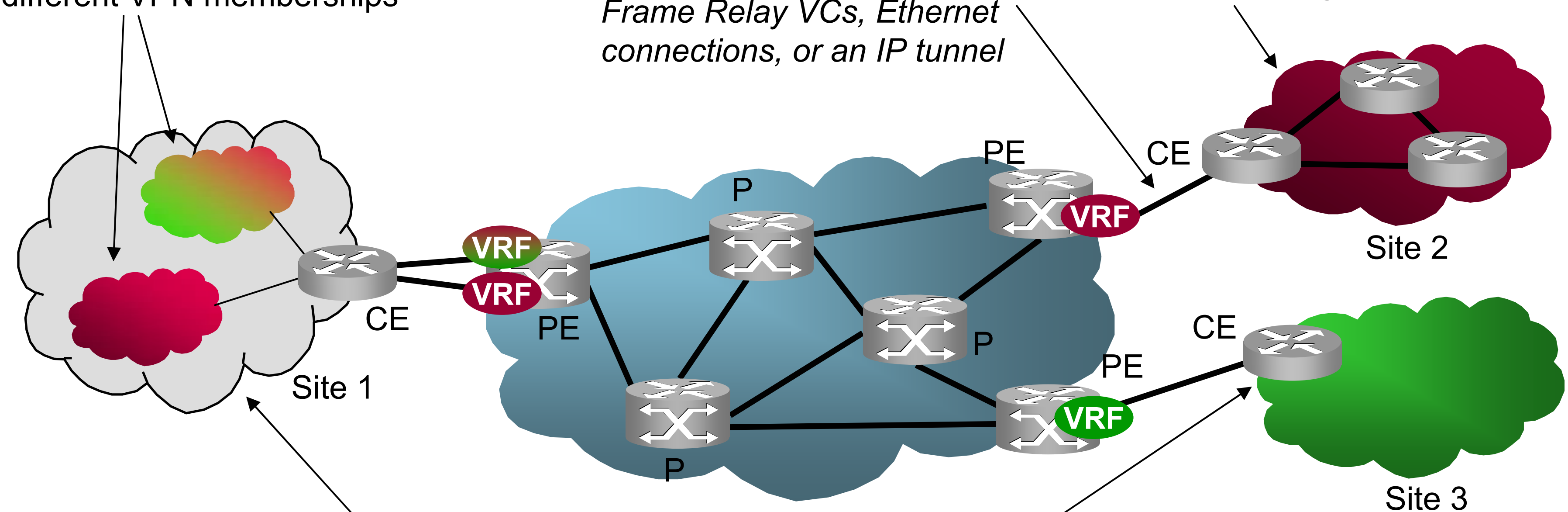

Site: Customer network with mutual connectivity without using the backbpne
VRF: VPN Routing and Forwarding Table
CE: Customer Edge Router
PE: Provider Edge Router
P: Provider Router

# Sites and Customer Edge Devices

Systems within a site may have different VPN memberships

*Attachment circuits can be PPP connections, ATM VCs, Frame Relay VCs, Ethernet connections, or an IP tunnel*

*A site is a set of IP systems that have mutual connectivity without using the backbone*

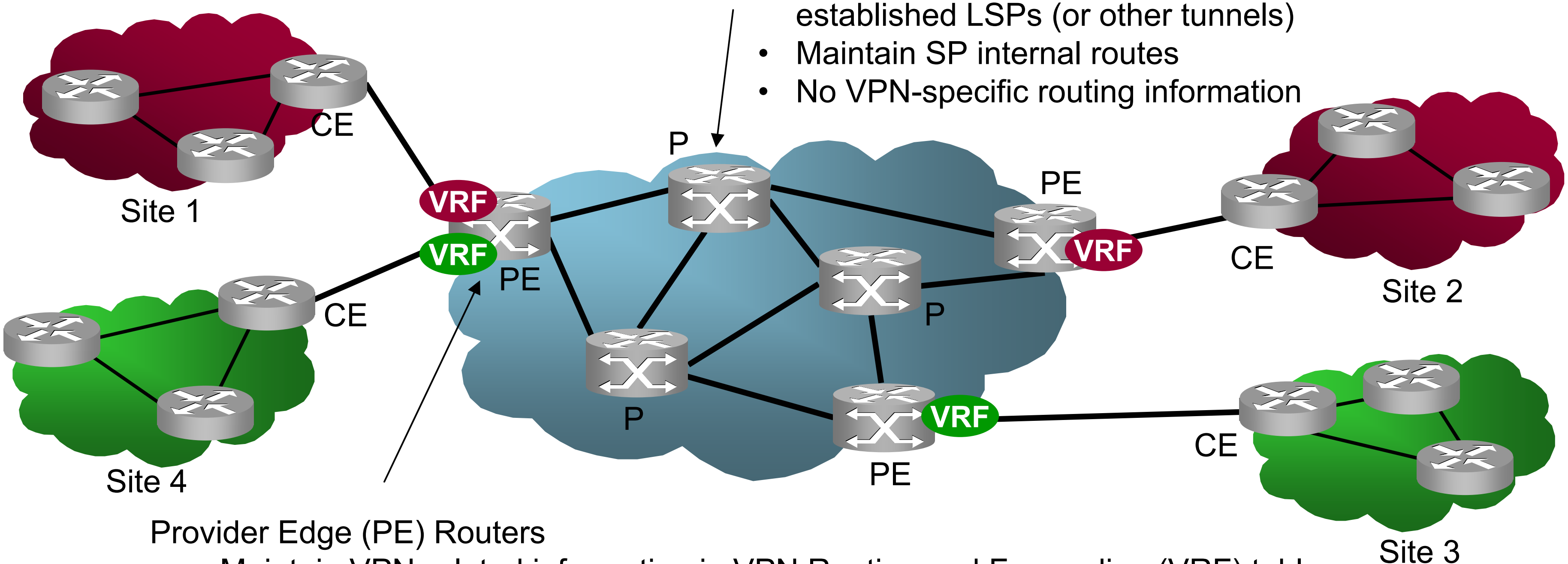



*A site may belong to multiple VPNs*

*CE devices are hosts or routers that are connected to PE routers by an attachment circuit - Each VPN must contain at least one CE*

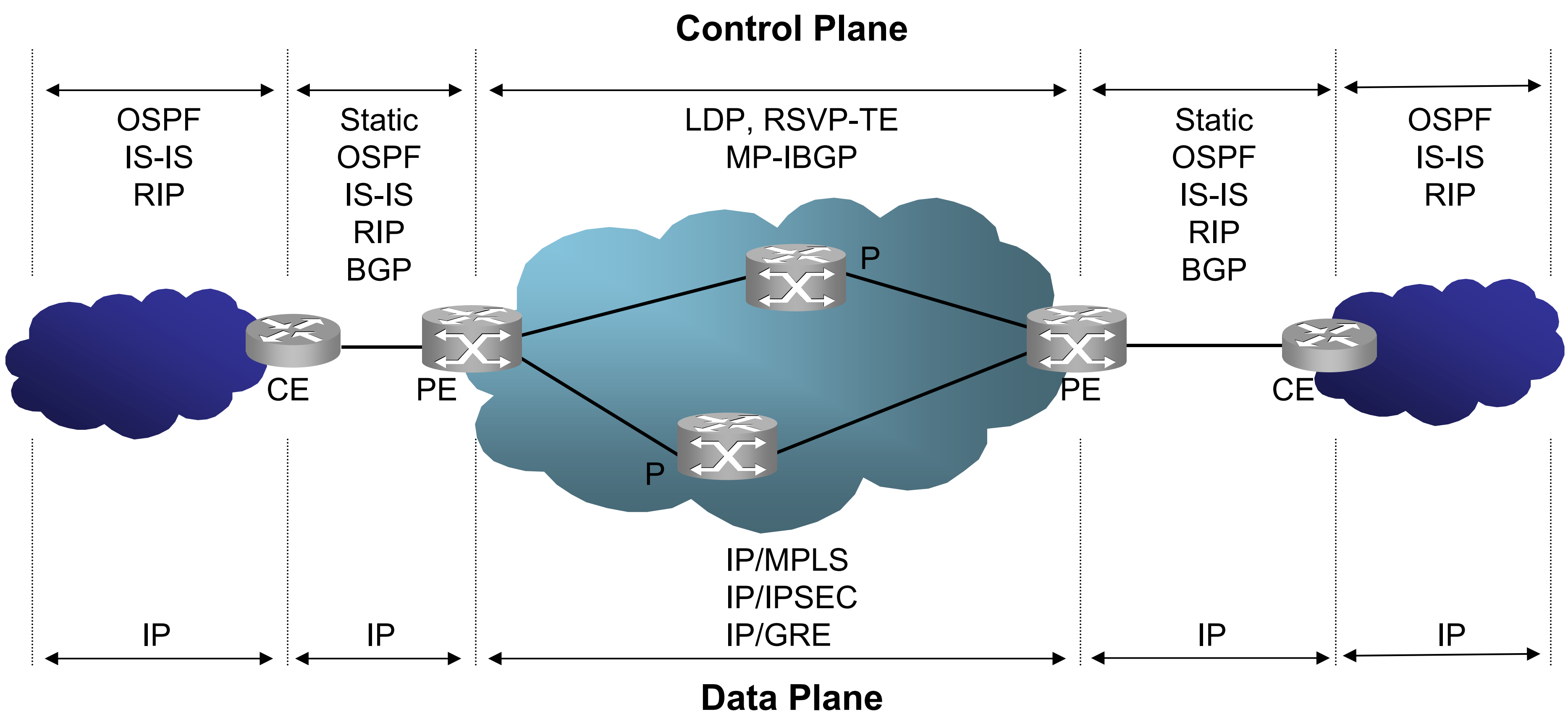
# Provider Edge and Core Routers

Provider (P) Routers
- Forward VPN traffic transparently over established LSPs (or other tunnels)
- Maintain SP internal routes
- No VPN-specific routing information



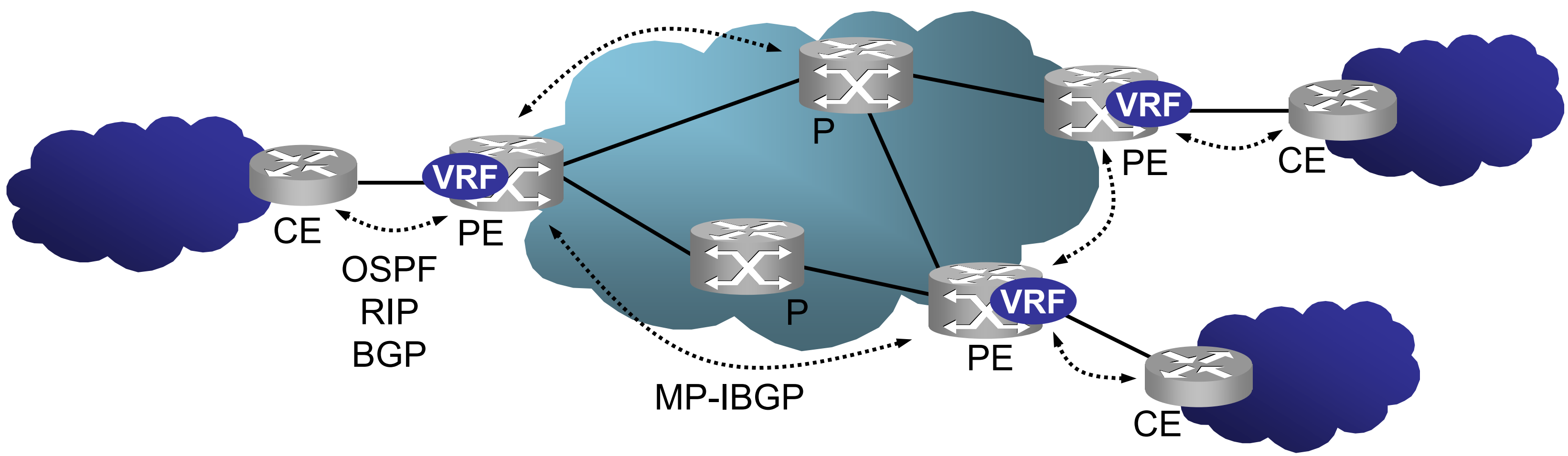

Provider Edge (PE) Routers
- Maintain VPN-related information in VPN Routing and Forwarding (VRF) tables
- Exchange routing information with the CE devices
- Exchange VPN-related information with other PEs
- Forward VPN traffic based on IP header and VPN information

# BGP/MPLS VPN Protocols



# Main Features of BGP/MPLS VPNs

- SP assisted exchange of VPN routes without requiring a full-mesh overlay network
  - Each customer sites peers only with a SP edge router
  - VPN routes can be exchange between customer sites and the SP edge routers using OSPF, RIP, or BGP or routes can be configured statically
  - SP edge routers use full-mesh MP-IBGP to exchange routing information

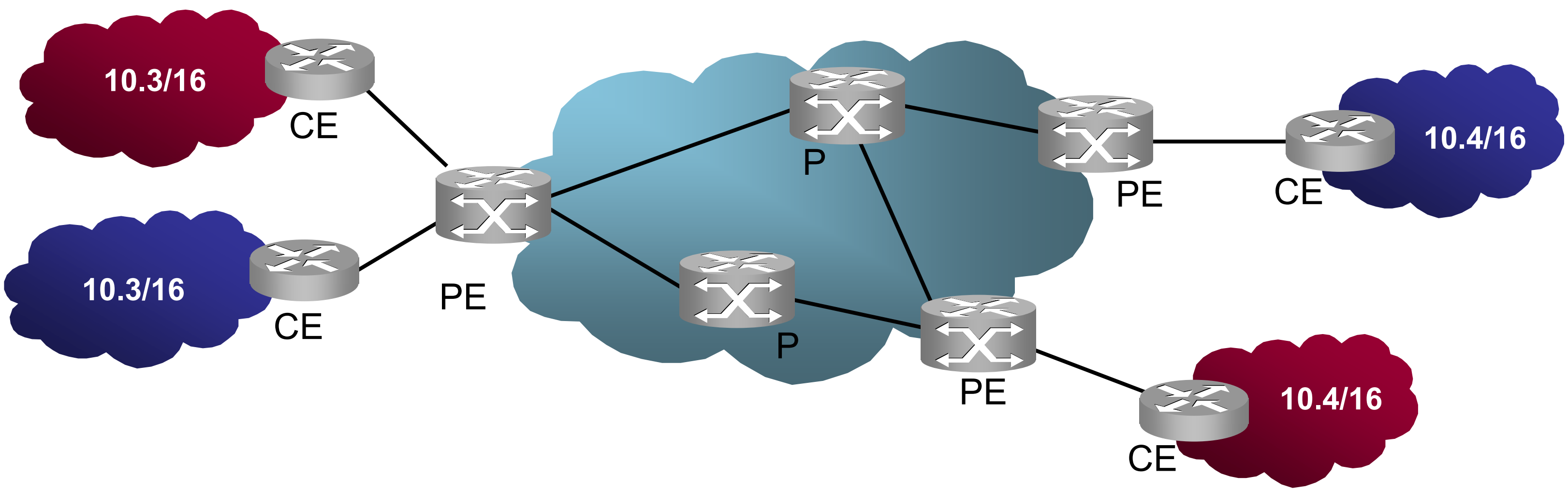


# Main Features of BGP/MPLS VPNs

- Scalability in the VPN Service Provider Network
  - Customer routing information is maintained at the PE routers
  - P routers are aware of only internal routes
  - Route reflectors can be used to reduce full-mesh MP-IBGP
  - Outbound route filtering can be used to reduce route updates
- Scalability in the Customer VPN
  - Each CE router peers with only a service provider PE router

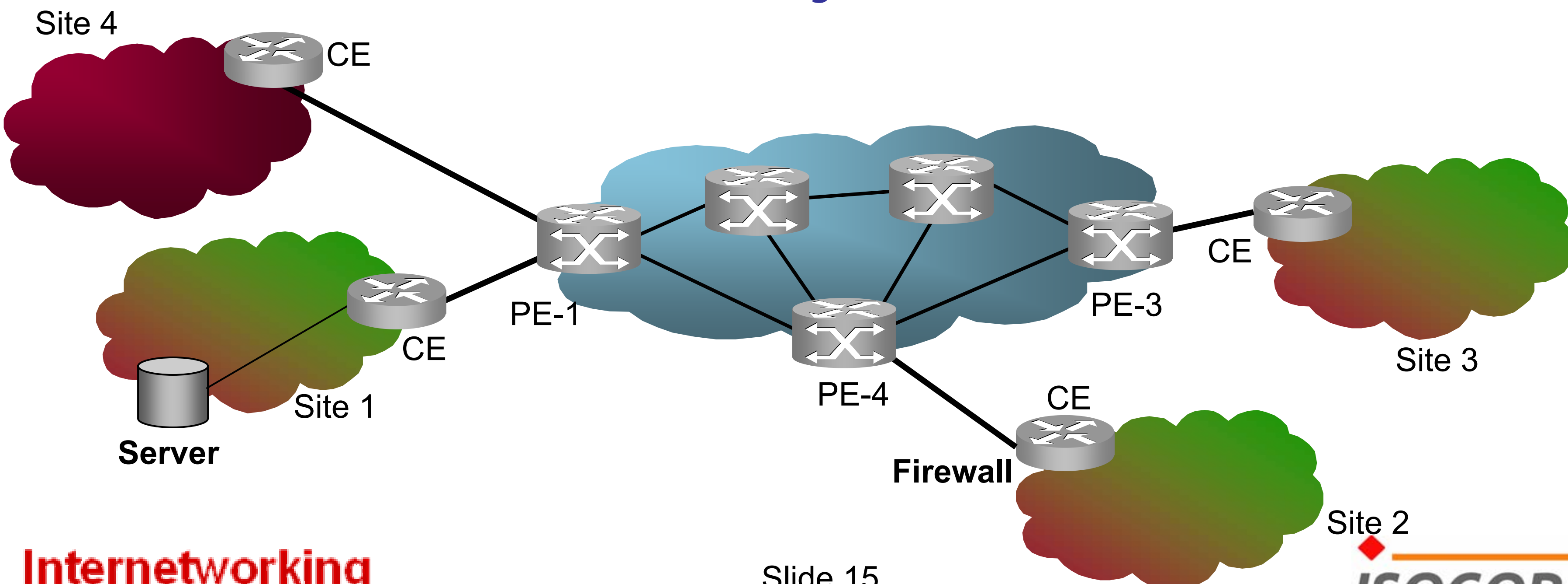
# Main Features of BGP/MPLS VPNs

- Address separation
  - Two sites can have an overlapping address space if they are members of different VPNs



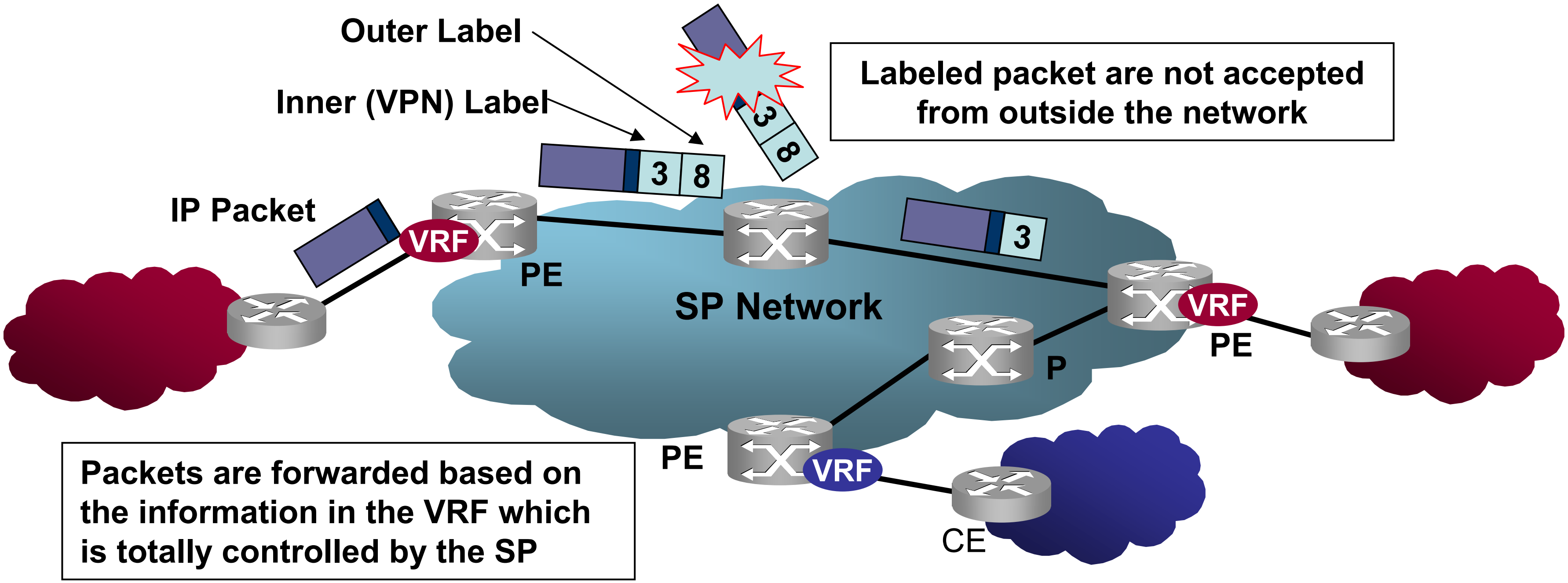
# Main Features of BGP/MPLS VPNs

- Alternative routes to the same system based on VPN membership
  - Sites 1, 2, and 3 form an intranet
  - Sites 1, 2, 3, and 4 form an extranet
  - Sites 2 and 3 can access the server directly
  - Site 4 accesses the server through the firewall on Site 2



# Main Features of BGP/MPLS VPNs

- Security is equivalent to those of ATM/Frame Relay Networks
  - Access to VPN sites is possible only from the PE routers
  - PE routers control how incoming packets from customer sites are routed
  - SP network does not accept packets or routes from untrusted sources

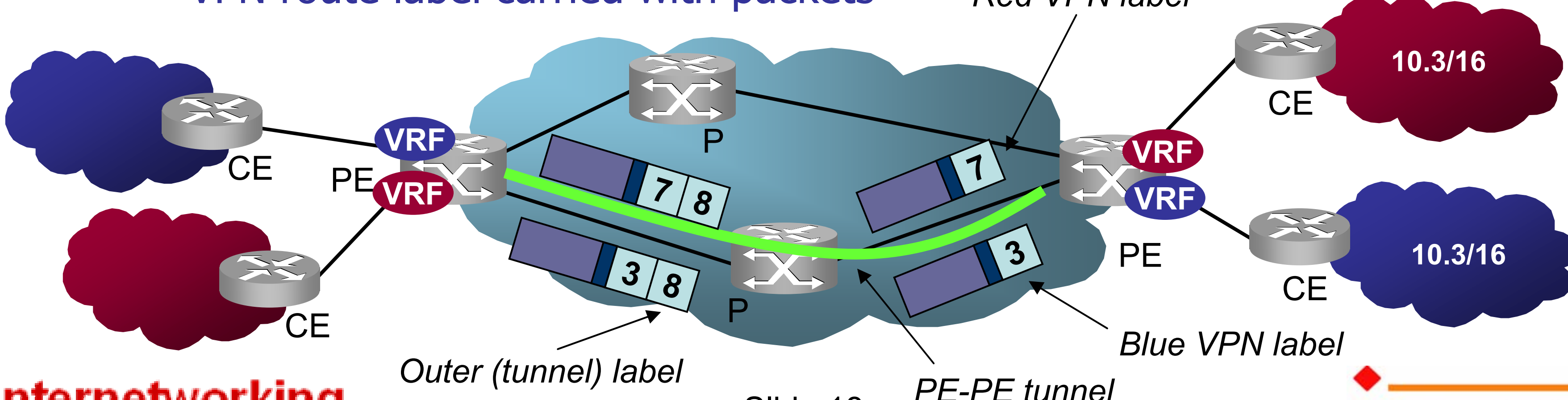


# Main Features of BGP/MPLS VPNs

- Simple control of VPN membership, topology, and route exchange
  - Full mesh, hub-and-spoke VPNs
  - Hierarchical BGP/MPLS VPNs
- QoS support through the use of traffic engineered tunnels as well as experimental bits in the MPLS shim header

# Functional Requirements

- PE routers must be able to route packets differently depending on the customer site that the packet is received from
  - Multiple Routing Tables – VPN Routing and Forwarding Table (VRF)
- Core routers must not have to maintain VPN routing information
  - Tunnels between Provider Edge Routers
- PE routers must be able to identify what VPN a packet received form the core belongs to
  - VPN route label carried with packets

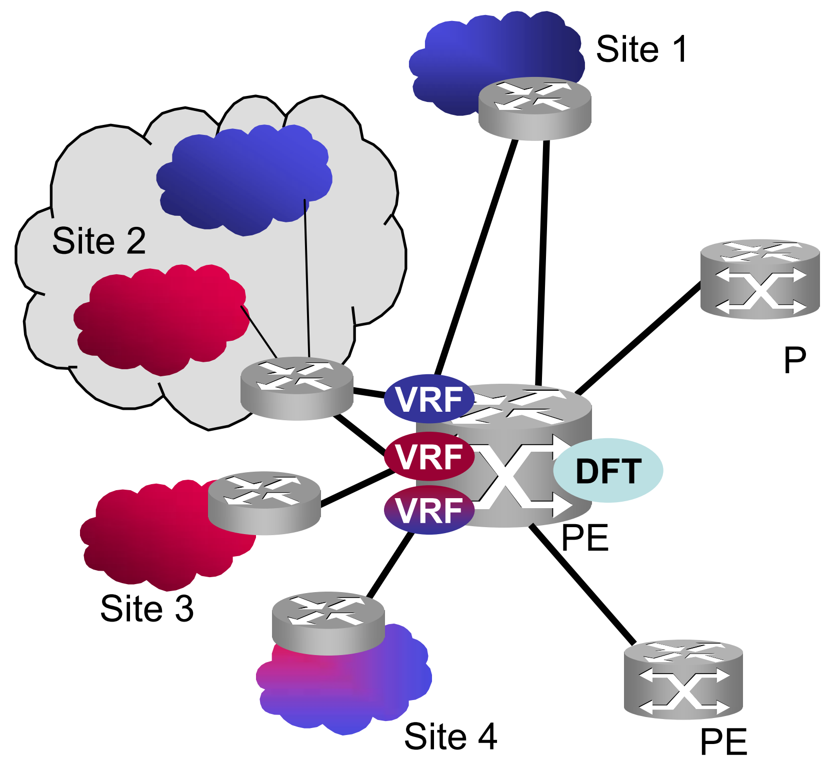


# Protocol Requirements

- Routing protocols must have a means to differentiate between routes with identical IP address prefixes but in different VPNs
  - BGP VPN IPv4 Addresses
  - BGP Multiprotocol Extensions allow BGP to carry routes from multiple address families
- There has to be a mechanism to associate advertised routes with the VPNs that they belong
  - BGP Route Target Attribute
  - Carried as a BGP Extended Community Attribute
- VPN routes must be assigned a VPN Label
  - Labeled VPN IPv4 Routes
  - Label is carried as part of the Network Layer Reachability Information (NLRI)

# VPN and Default Routing and Forwarding Tables

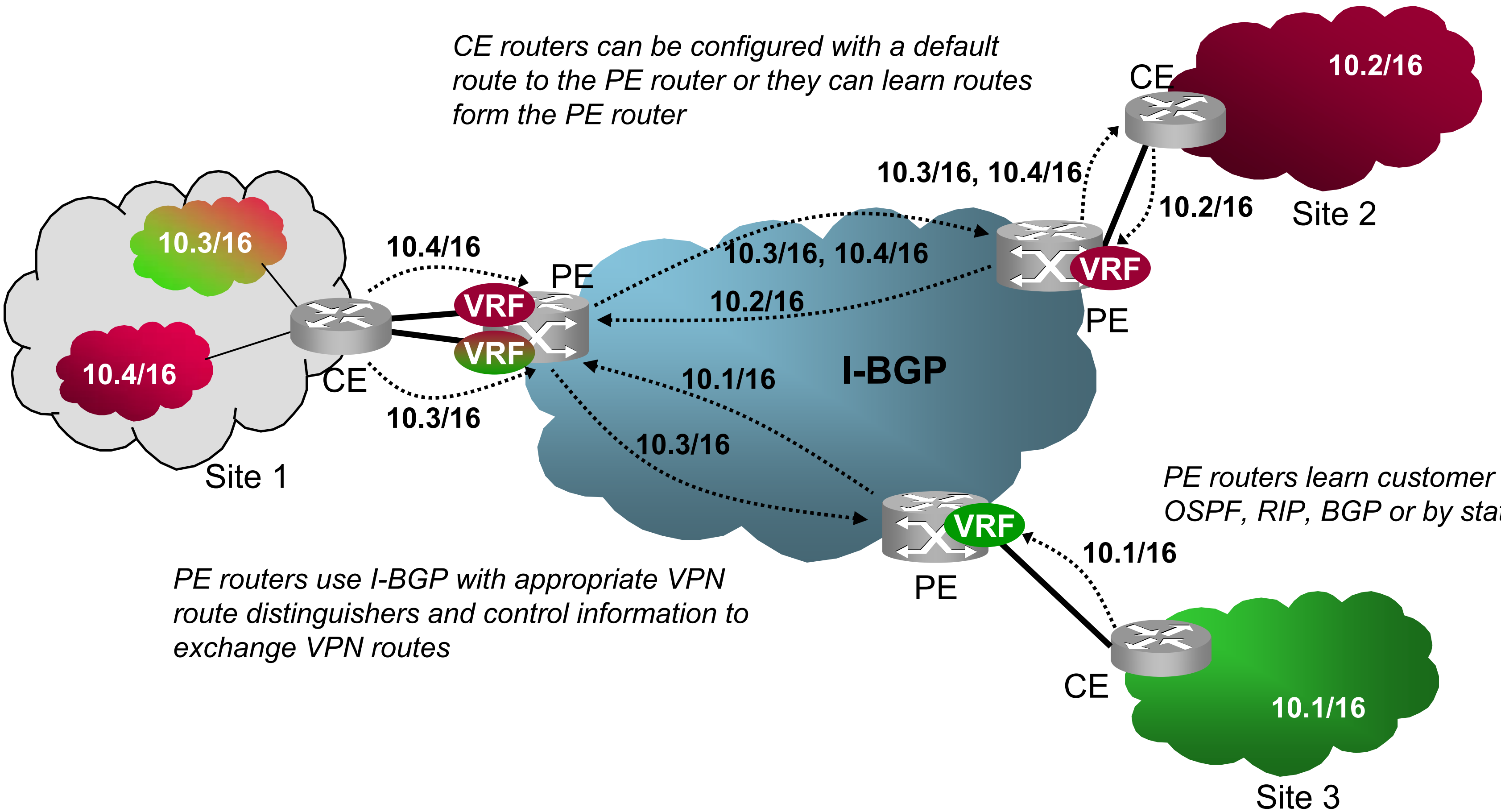
- Every PE-CE attachment circuit is associated, by configuration, with a VRF
- VRF is used to route customer packets associated with a VPN
- DFT is used to forward packets received from neighboring P and PE routers as well as packets from customer sites that are not associated with a VRF
- A CE may be associated with one or more VRFs
- Total separation between VRF and DFT
- Physical port, VC-ID, VLAN ID, or IP source address can be used to determine which VRF to use for an IP packet




VRF: VPN Routing and Forwarding Table
DFT: Default Forwarding Table

# Populating the VRFs

*CE routers can be configured with a default route to the PE router or they can learn routes form the PE router*

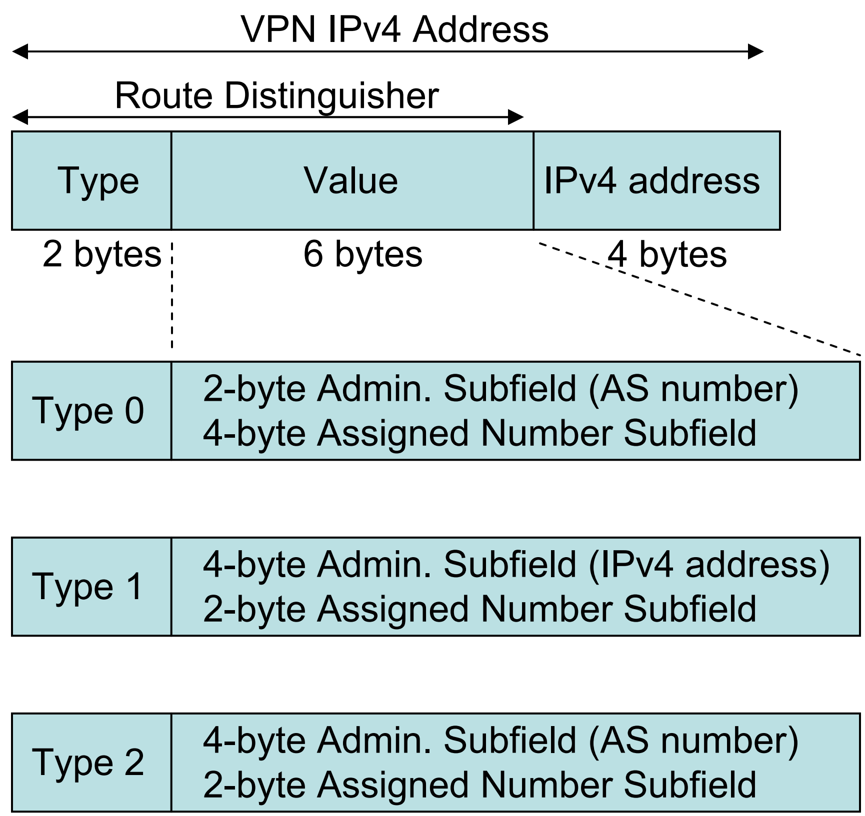



*PE routers learn customer routes by OSPF, RIP, BGP or by static configuration*

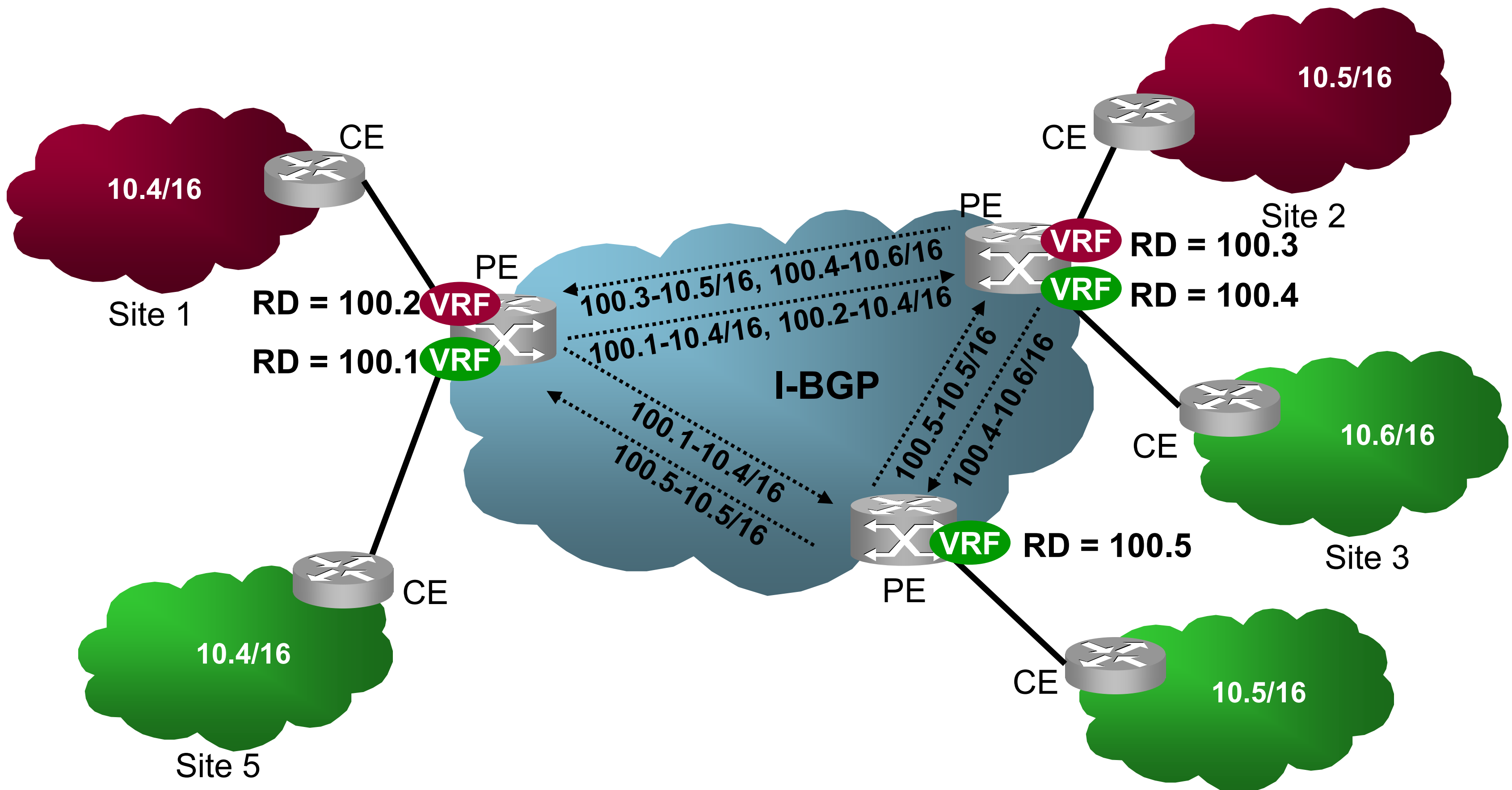
*PE routers use I-BGP with appropriate VPN route distinguishers and control information to exchange VPN routes*

# VPN-IPv4 Address Family

- Without a new address family BGP would not be able to carry identical IPv4 addresses from different VPNs
- A VPN-IPv4 address consists of an 8-byte Route Distinguisher (RD) and a 4-byte IPV4 address
- BGP Multiprotocol Extensions allow BGP to carry routes from multiple address families
- A PE router needs to be configured to associate routes that lead to a particular CE with one or more RDs
- Each VRF is associated with one RD

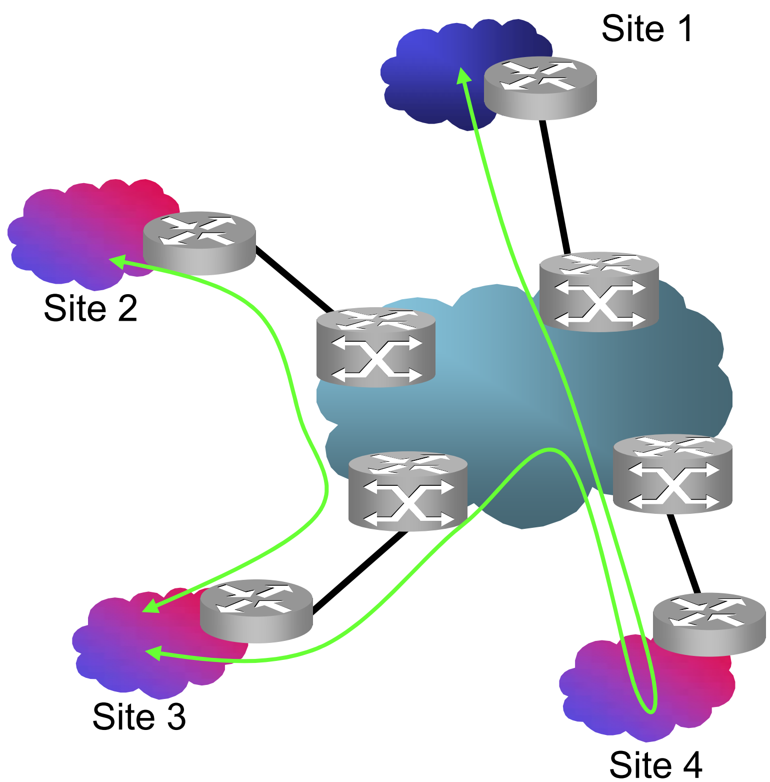


# Duplicate Addresses using RD



# Different Routes to Same System using RD

- Sites 2, 3, and 4 are members of VPN 1 (intranet)
- Sites 1, 2, 3, and 4 are members of VPN 2 (extranet)
- Sites 2, 3, and 4 have direct access to each other
- Site 1 can access sites 2, 3, and 4 only via Site 4 where there is a firewall
- Site 2, 3, and 4 routes are distributed twice with two RDs:
  - RD 1 – Used to establish direct routes between 2, 3 and 4 under intranet policy
  - RD 2 – Used to provide Site 1 access to sites 2, 3, and 4 via site 4 under extranet policy



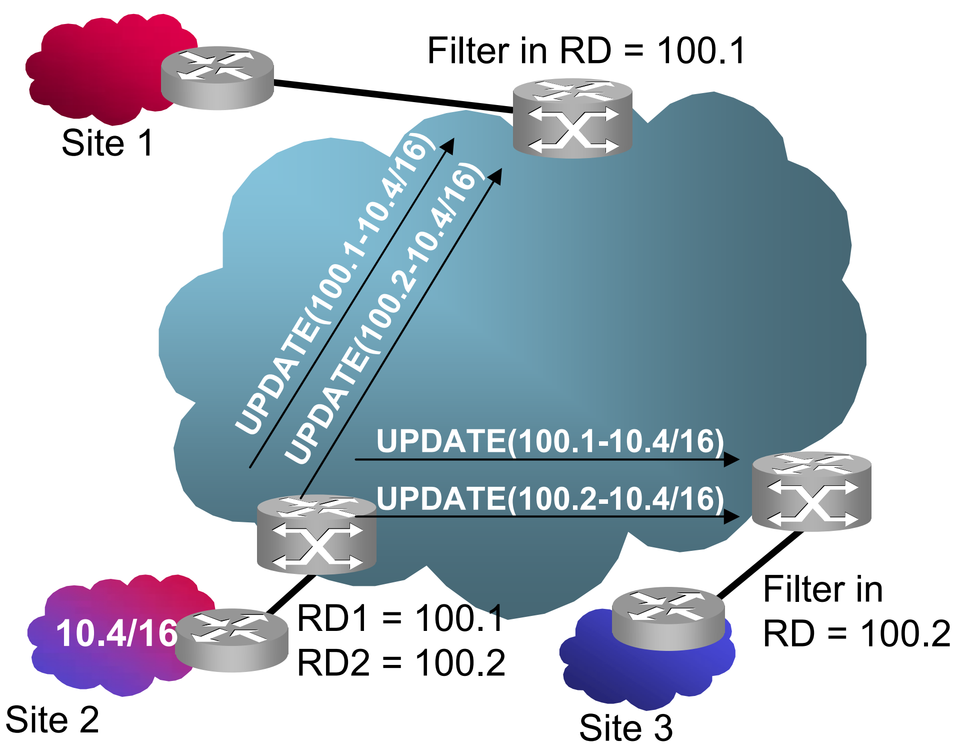
# Route Target Attribute

- Each VRF is associated with one or more Route Target (RT) attributes
  - Export Targets determine what other VRFs the routes in a particular VRF can be exported to and are carried in BGP route advertisements
  - Import Targets determine whether a route can be imported in a VRF
  - A route with an Export Target "X" gets installed in a VRF with an Import Target "X"
- RTs are carried in BGP as Extended Community Route Targets and structured similarly to the RDs
- Associating Export Targets with routes
  - All routes leading to a particular site are assigned the same RT
  - Different routes in a given site are assigned different RTs
  - Each route can be assigned multiple RTs

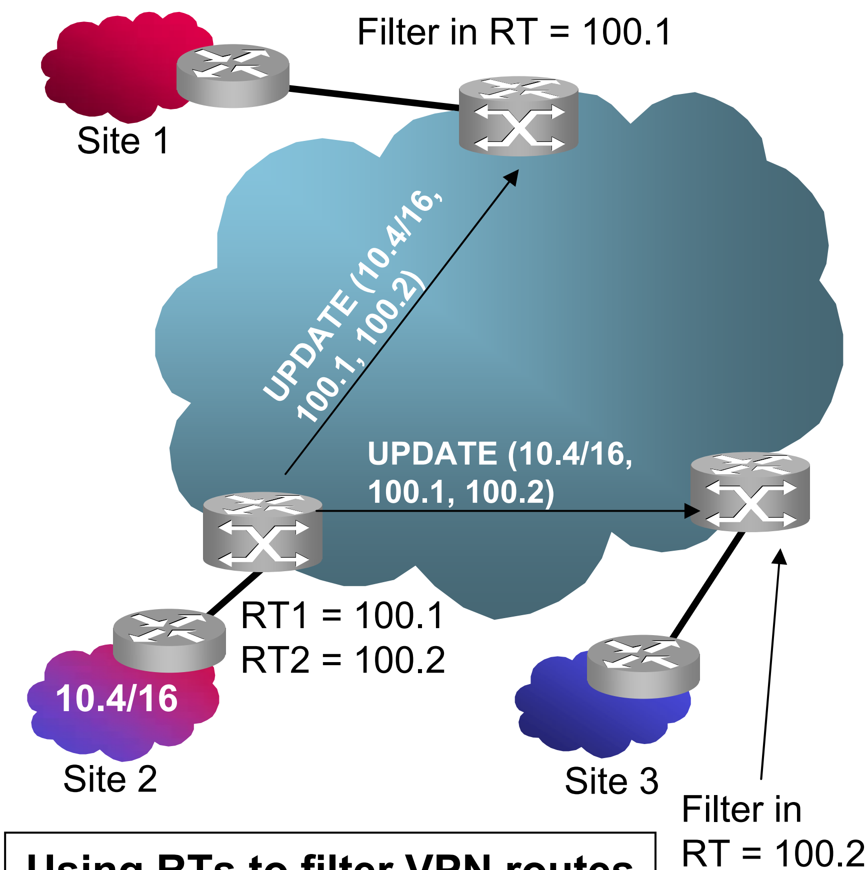
# Relationship between RDs and RTs

- RDs convert IPv4 addresses to unique VPN IPv4 addresses that can be carried in BGP
- RTs are attributes of VPN IPv4 routes that control which sites can access these routes
- In BGP, each route can have multiple attributes, therefore the fact that a route is a member of multiple VPNs can be conveyed in one UPDATE message
- An alternative design could have used RDs to determine VPN membership
  - When a site is in multiple VPNs, its routes would be advertised multiple times, each with a different RD
  - This would not be a scalable solution

# Using RDs versus RTs for Route Filtering



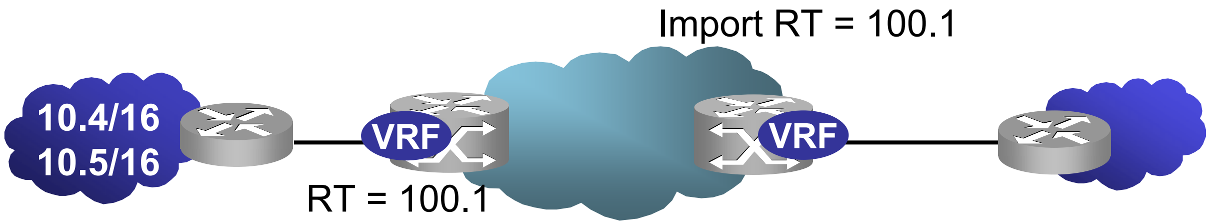

**Using RDs to filter VPN routes**



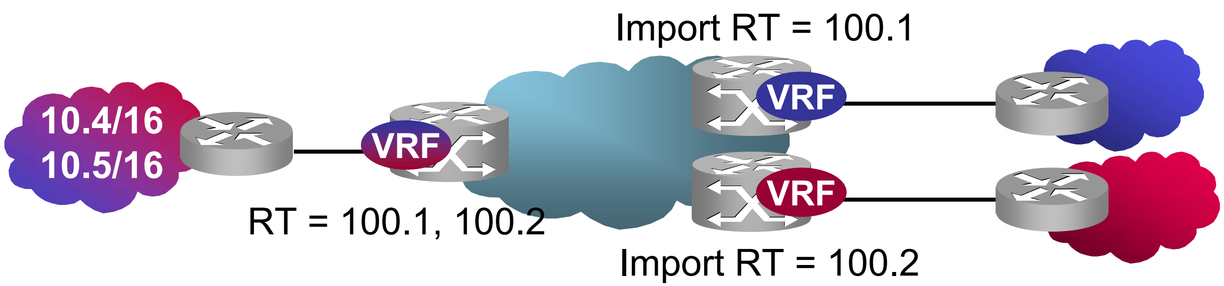

**Using RTs to filter VPN routes**

# Associating Export Targets with Routes (1)

- A PE can be configured to associate all routes of a site with one RT

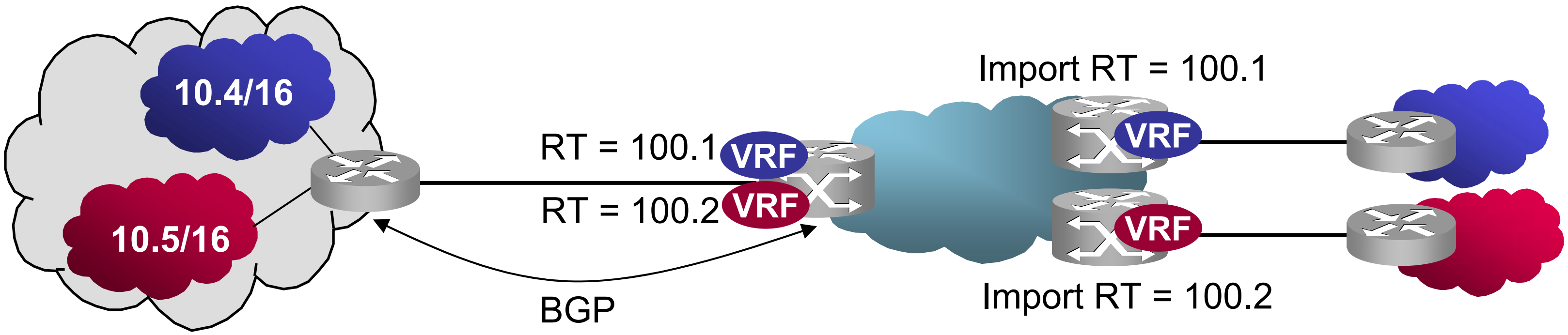



- A PE can be configured to associate all routes of a site with multiple RTs

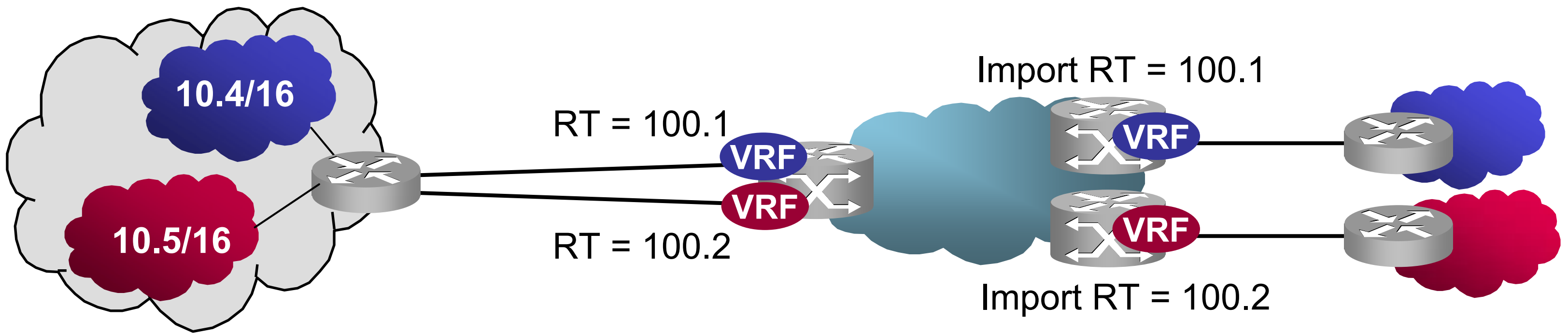


# Associating Export Targets with Routes (2)

- Different routes can be associated with different RTs
  - CE attaches RTs (within limits) to routes that it distributes to PE with BGP




  - CE is attached to PE by multiple attachment circuits, each configured with a different RT

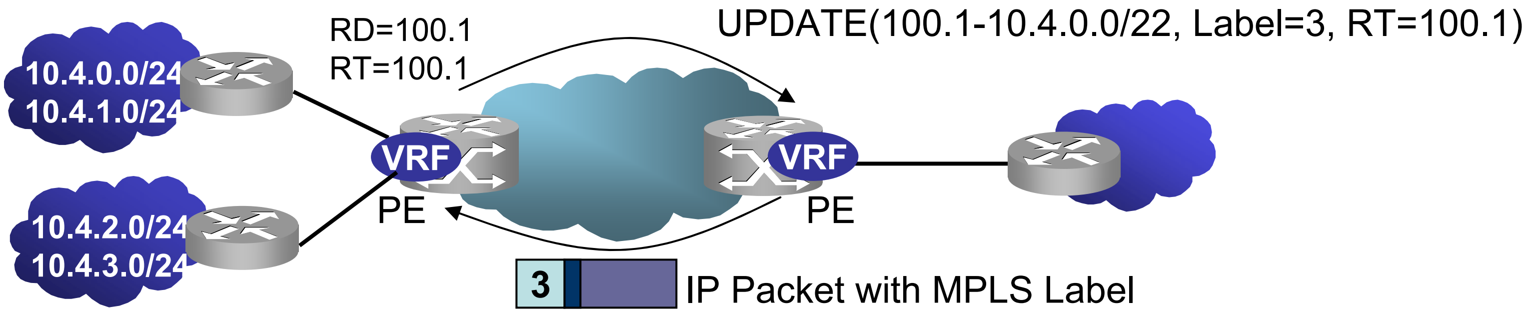


# Route Distribution Among PEs by BGP

- PEs can distribute VPN IPv4 routes using full-mesh I-BGP connections between them or via an I-BGP connection to a route reflector
- PEs may distribute the exact set of routes that appears in the VRF or perform aggregation
- PEs distribute routes with their address as the BGP next hop
- PEs assign and distribute MPLS labels with the routes
  - A single label may be used for the entire VRF
  - A single label may be used for each attachment circuit
  - Different labels may be used for each route
- Packets sent to VPN destinations are appended with the appropriate label
- An egress PE forwards the packet to one of its customer interfaces based on the label

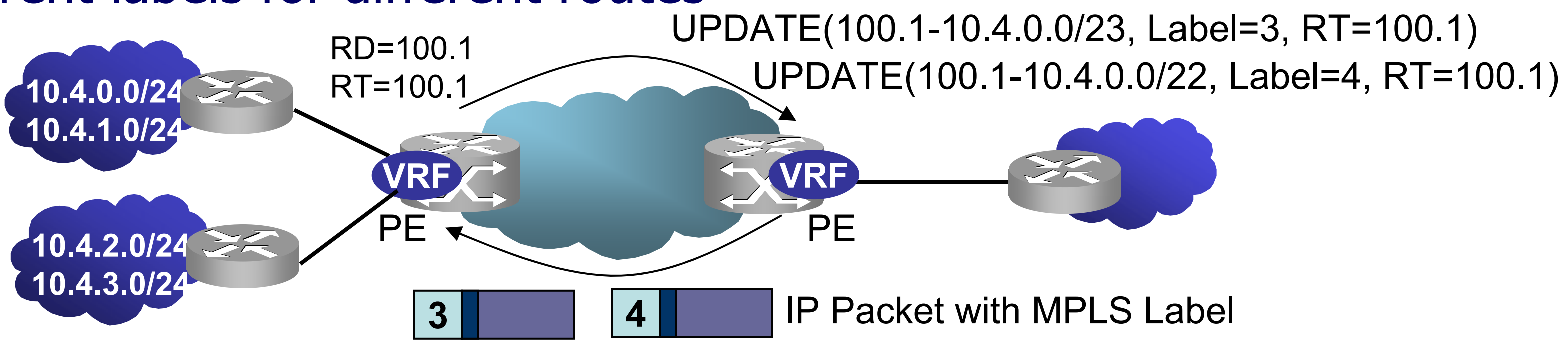
# Forwarding Packets based on VPN Labels

- A single label used for the entire VRF




**PE needs to look up packet's IP address in the VRF to determine packet's egress attachment circuit**

- Different labels for different routes




**PE can determine packet's egress attachment circuit based on the VPN label, without looking at VRF**

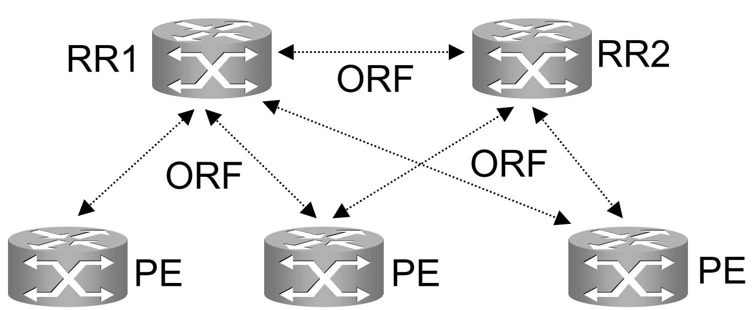
# Outbound Route Filtering

- If there is no outbound filtering, a PE router often receives unwanted routes from peers and filters them based on RTs
- The number of BGP VPN route updates can be reduced by using BGP cooperative route filtering capability
  - PE routers willing to send or receive ORFs advertise Cooperative Route Filtering Capability
  - PE routers send ORFs in BGP Refresh messages
  - By using Extended Communities ORF type, a PE router can request its peers to send VPN route updates for specific RT values
  - The peers use the received ORFs as well as locally configured export target policy to constrain and filter outbound route updates
- Cooperative route filtering conserves bandwidth and packet processing resources

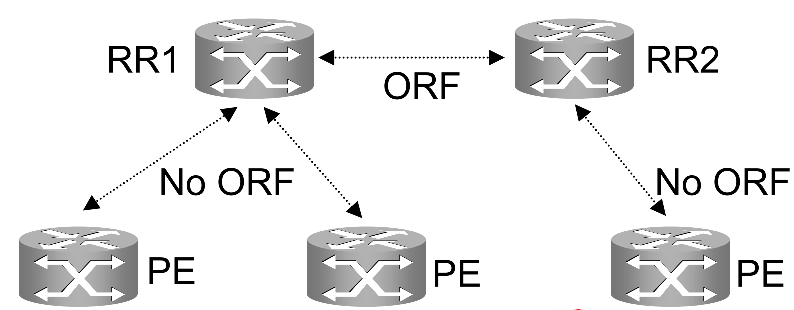
# Use of Route Reflectors

- Scalability of VPN route distribution can be increased by use of BGP Route Reflectors (RR)
- Two ways to partition VPN IPv4 routes among different RRs
  - Each RR is pre-configured with a list or Route Targets
  - Each PE is a client of a subset of RRs

- RR1 and RR2 perform inbound filtering based on pre-configured list of RTs
- They can use this list of RTs to install ORFs on their RR or PE peers

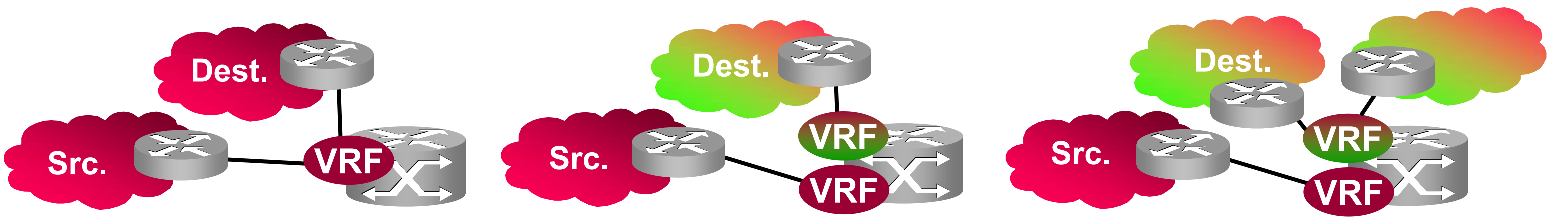



- RR1 and RR2 do not perform inbound filtering on routes received from PEs
- They generate an RT list based on routes received from the PEs
- This set is used to apply inbound filters to routes received from other RRs



# Packet Forwarding

- For packets received from a CE, the PE determines which VRF to use based on ingress attachment circuit
  - If the packet is destined to a site connected to the same PE, packet is forwarded without a VPN label
  - A second VRF look-up may be required when the two sites are attached to different VRFs

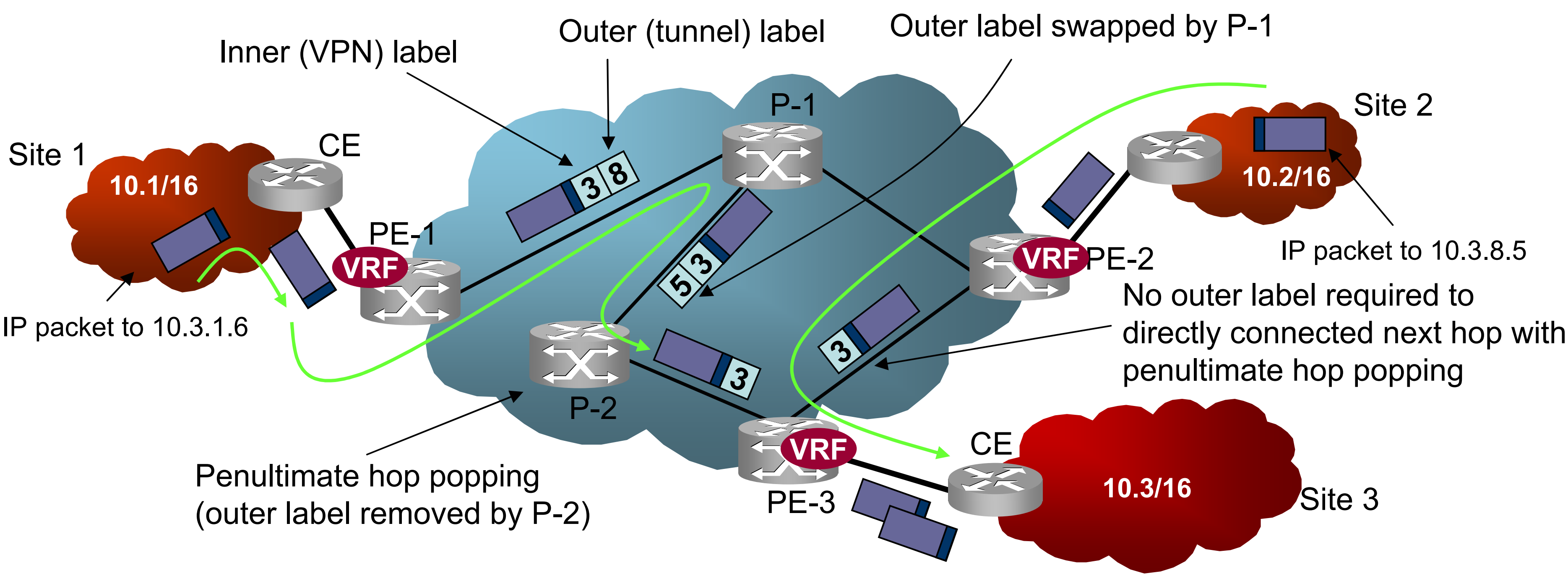



| | | |
|---|---|---|
| **Forward to egress with one VRF look-up** | **Forward to egress with one VRF look-up** | **Forward to egress with two VRF look-ups** |

  - For packets received from a PE, the egress attachment circuit is determined from the VPN label – a VRF look-up may be necessary

# Packet Forwarding (2)

- When a packet is received from a CE and when the destination site is connected to a different PE, a VPN label is attached to the packet
  - The resulting packet is tunneled to the destination PE (BGP-Next Hop) via an MPLS, GRE, IPSec, or IP tunnel



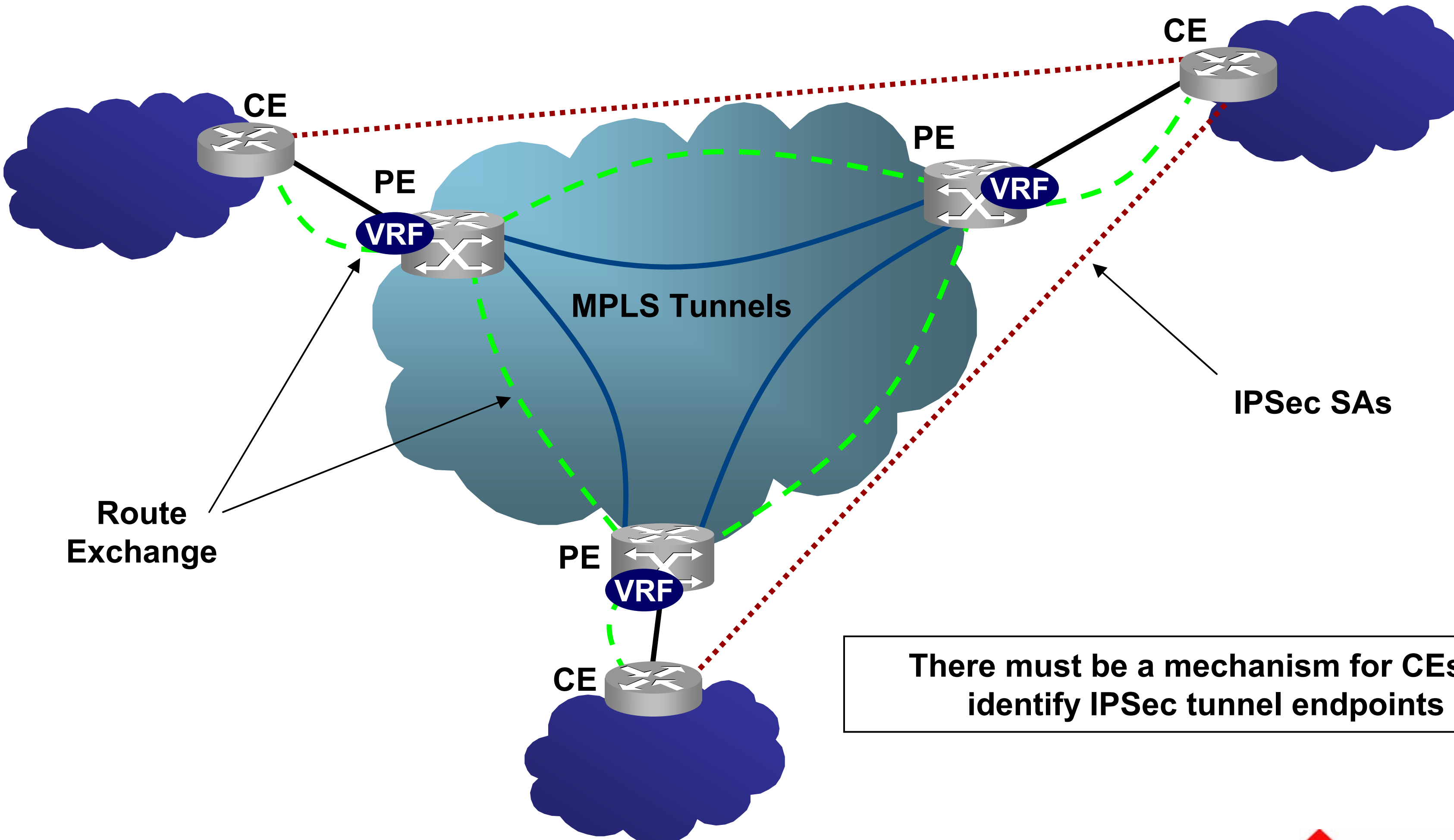
# Route Exchange Between PE and CE

- PE router may be configured with static routes to the CE router
- PE and CE routers may be RIP or OSPF peers
  - The CE router must not re-advertise VPN routes learned from a PE back to that PE or another PE
- PE and CE routers may be BGP peers
  - Does not require running multiple protocol instances
  - Makes it easier for the CE router to pass route attributes such as Route Targets to the PE router
  - The “Site of Origin” attribute can be used to ensure that routes learned from a site are not re-distributed to the site over a different connection
- PE router may distribute the VPN routes learned from other PE routers or just a default route to the CE router

# Security of BGP/MPLS VPNs

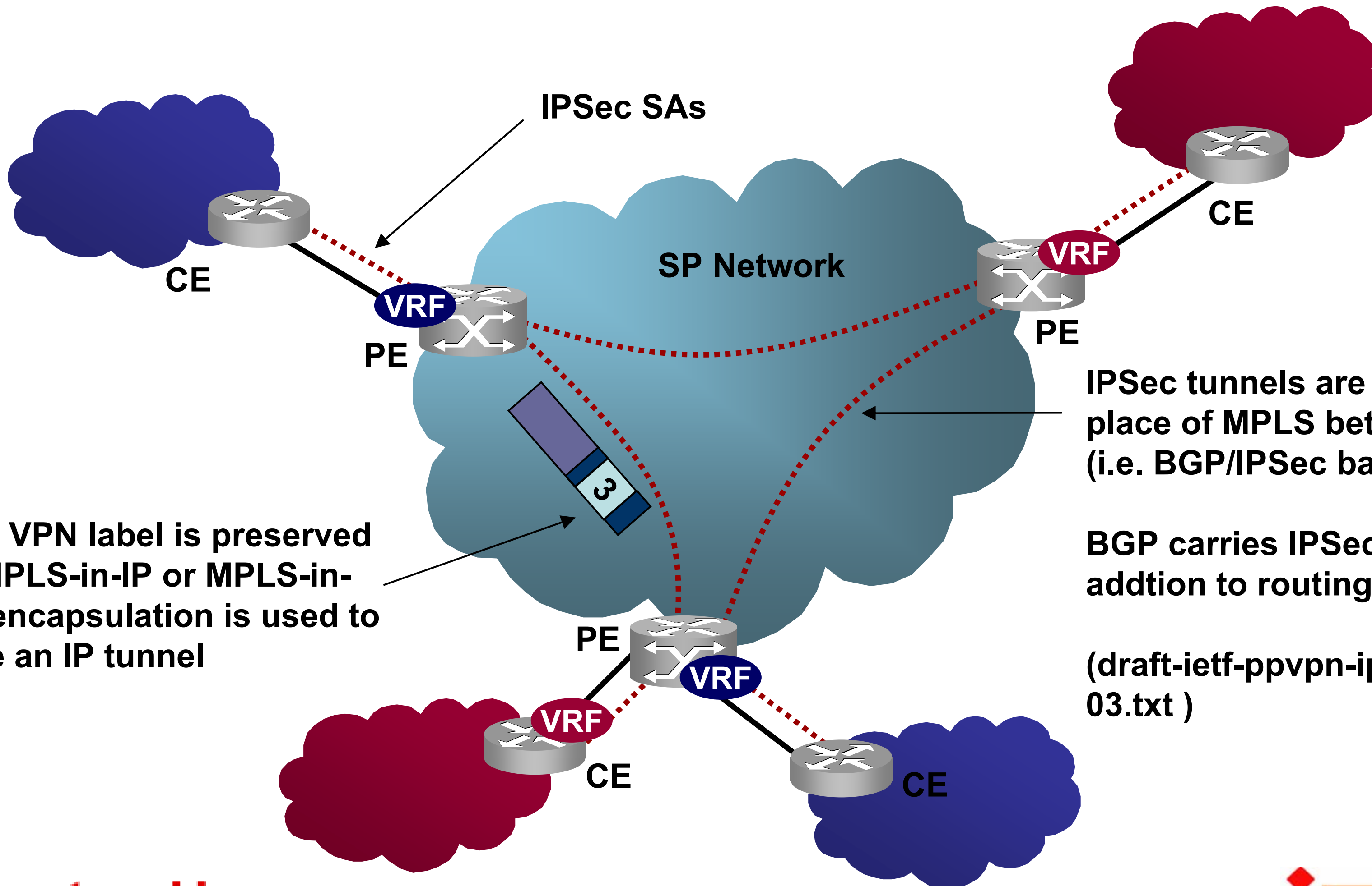
- Built-in security features
  - Access to VPNs is tightly controlled by the PEs
  - Total address separation by use of VPN IPv4 addresses
  - Separation of routing information by use of route targets
- Vulnerabilities
  - Misconfiguration of the core and attacks within the core
  - Security of the access network
- Additional Security can be provided by combining IPSec and MPLS
  - End-to-end IPSec overlaid on an MPLS VPN
  - IPSec in the core

# End-to-end IPSec Tunnels Overlaid on a BGP/MPLS VPN




There must be a mechanism for CEs to identify IPSec tunnel endpoints

# IPSec in the Core




**MPLS VPN label is preserved and MPLS-in-IP or MPLS-in-GRE encapsulation is used to create an IP tunnel**

**IPSec tunnels are used in place of MPLS between PEs (i.e. BGP/IPSec based VPN)**

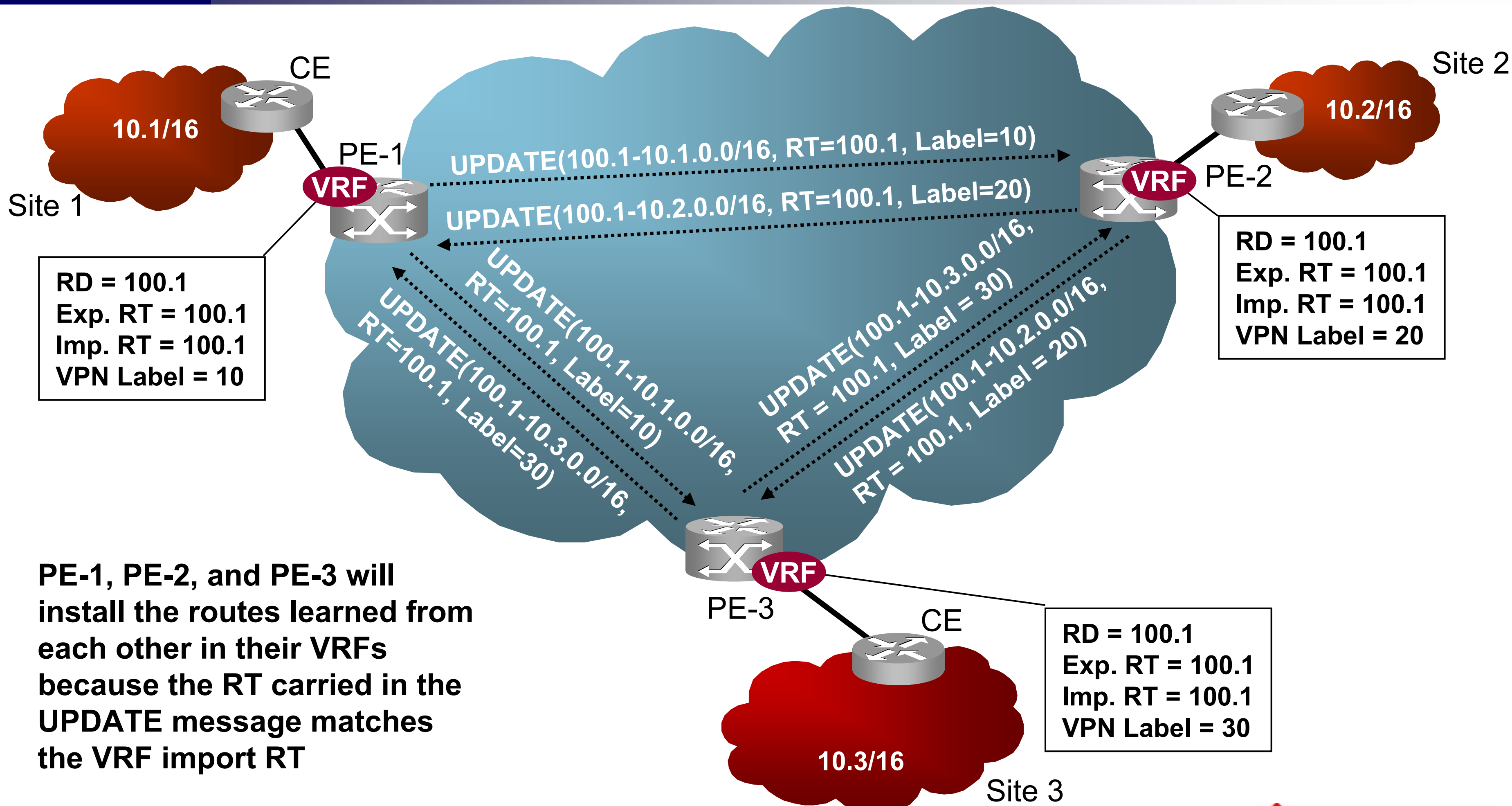
**BGP carries IPSec policy in addtion to routing information**

**(draft-ietf-ppvpn-ipsec-2547-03.txt )**

# Building BGP/MPLS VPNs

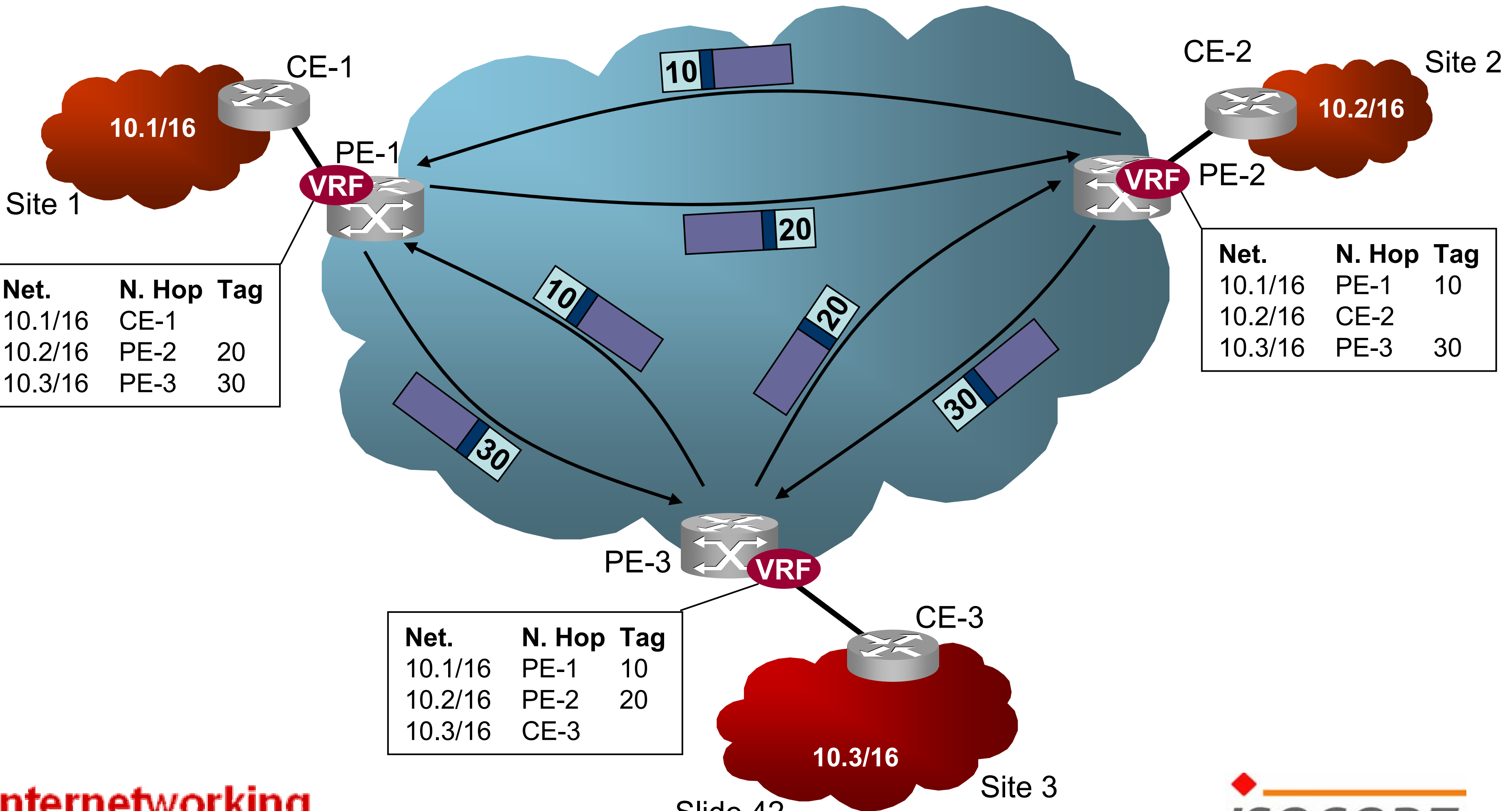
Full-Mesh VPNs
Sites with Multiple VPN Membership
Hub and Spoke VPNs
Overlapping Intranet and Extranet VPNs
Accessing Public Internet from a VPN
Hierarchical VPNs

# Building Full-Mesh VPNs (1)




**PE-1, PE-2, and PE-3 will install the routes learned from each other in their VRFs because the RT carried in the UPDATE message matches the VRF import RT**

# Building Full-Mesh VPNs (2)

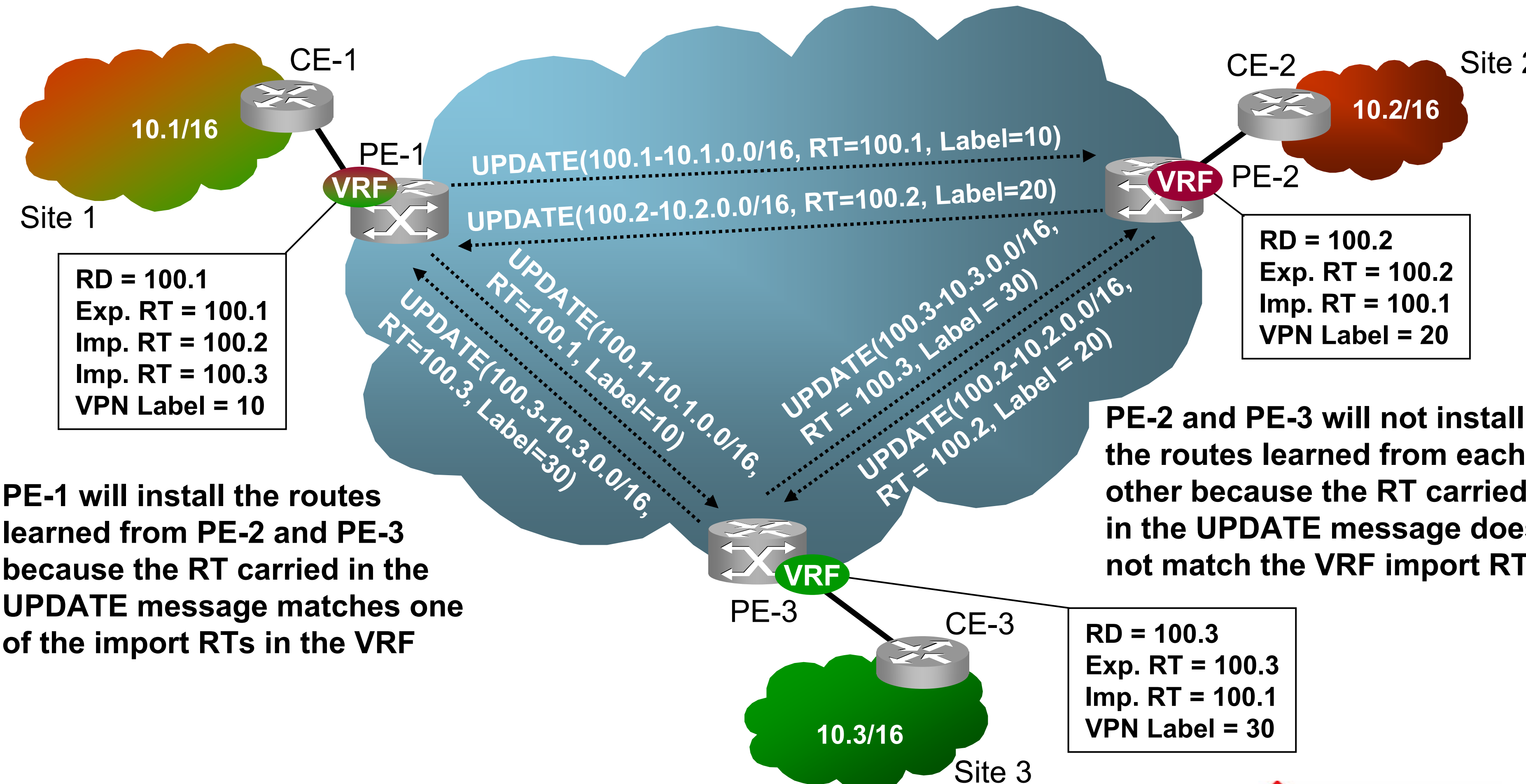



| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | CE-1 | |
| 10.2/16 | PE-2 | 20 |
| 10.3/16 | PE-3 | 30 |

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-1 | 10 |
| 10.2/16 | CE-2 | |
| 10.3/16 | PE-3 | 30 |

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-1 | 10 |
| 10.2/16 | PE-2 | 20 |
| 10.3/16 | CE-3 | |

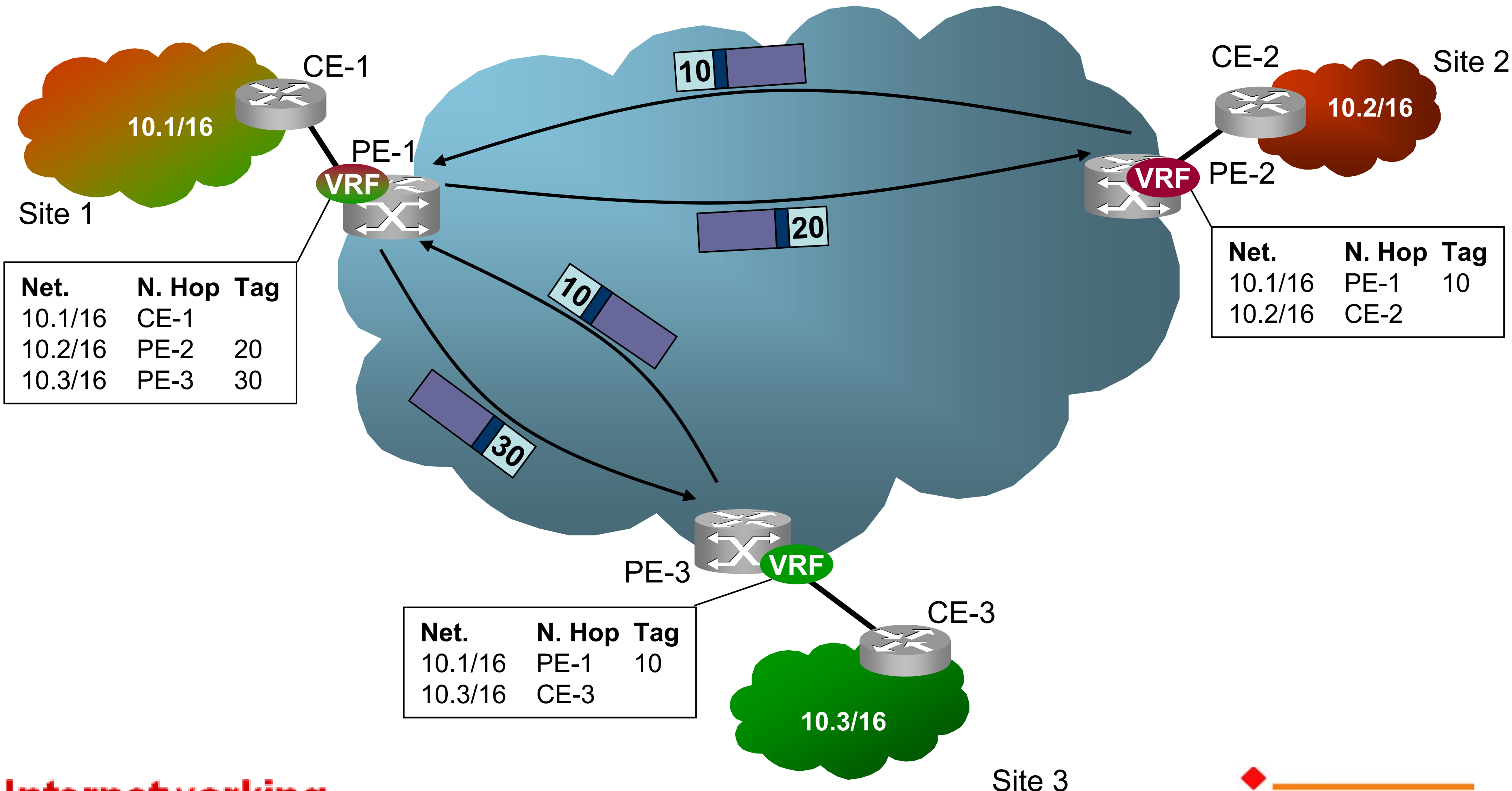
# Sites with Multiple VPN Membership (1)




**PE-1 will install the routes learned from PE-2 and PE-3 because the RT carried in the UPDATE message matches one of the import RTs in the VRF**

**PE-2 and PE-3 will not install the routes learned from each other because the RT carried in the UPDATE message does not match the VRF import RT**

# Sites with Multiple VPN Membership (2)

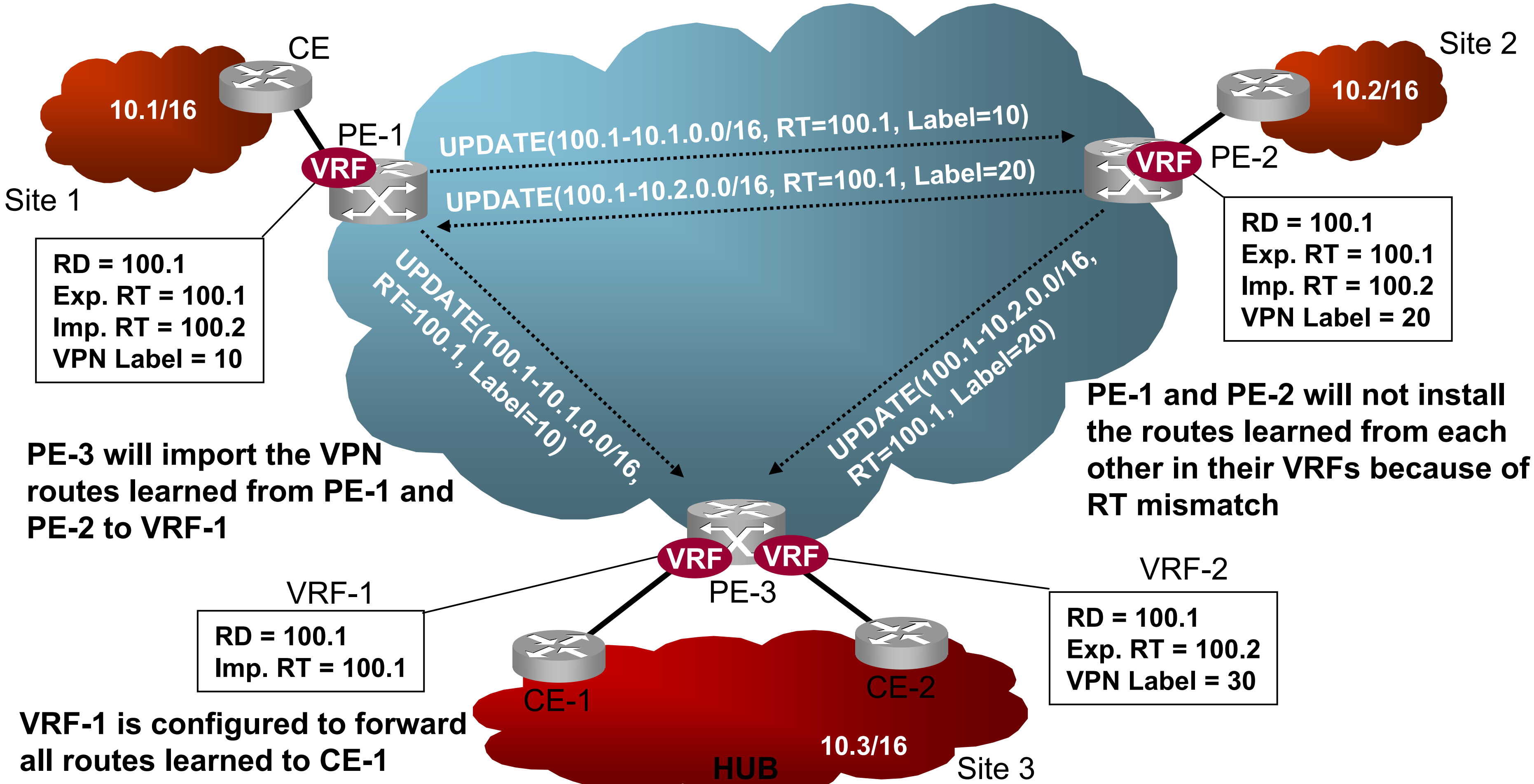



| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | CE-1 | |
| 10.2/16 | PE-2 | 20 |
| 10.3/16 | PE-3 | 30 |

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-1 | 10 |
| 10.2/16 | CE-2 | |

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-1 | 10 |
| 10.3/16 | CE-3 | |

# Building Hub and Spoke VPNs (1)

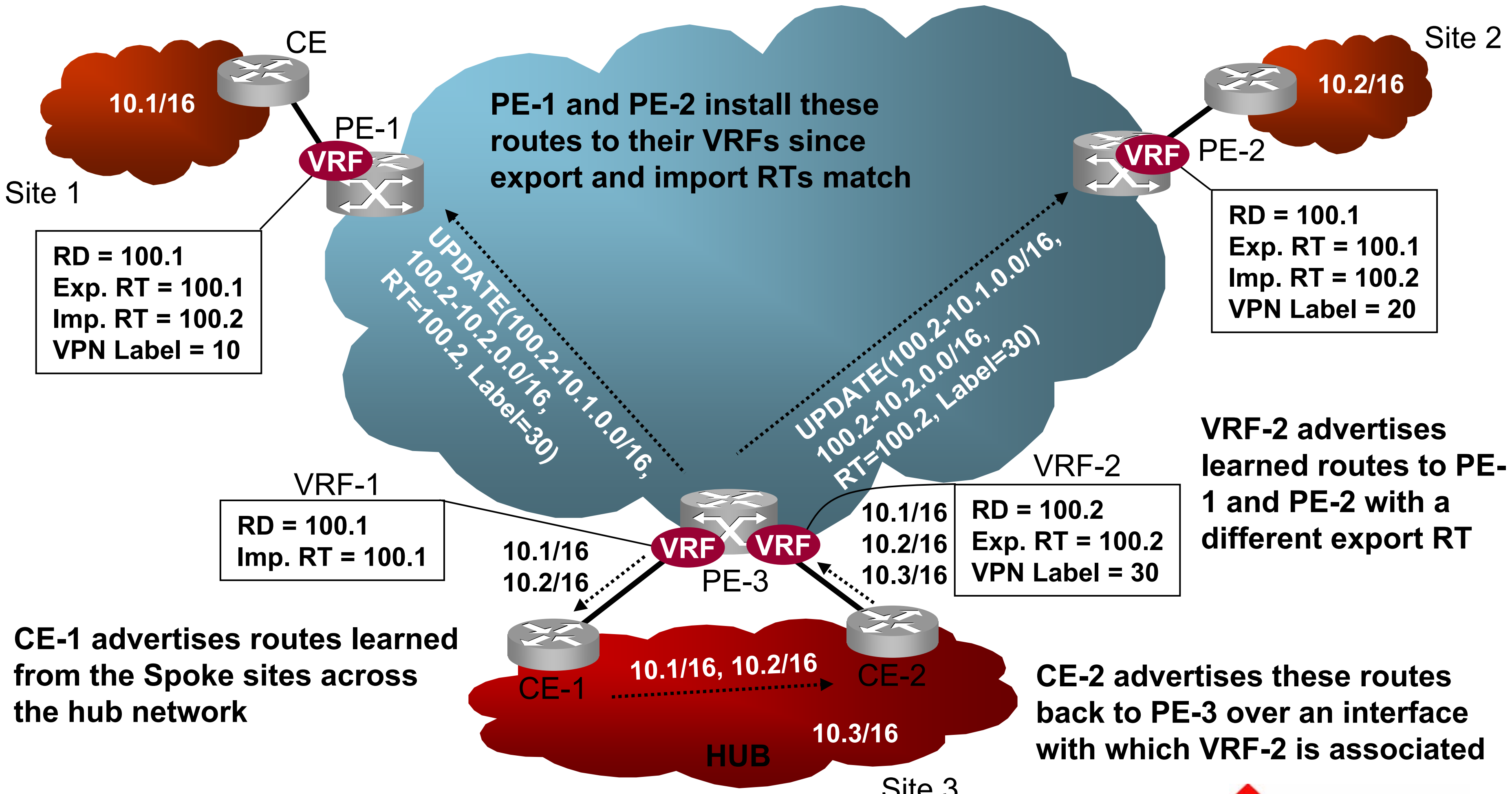



**PE-3 will import the VPN routes learned from PE-1 and PE-2 to VRF-1**

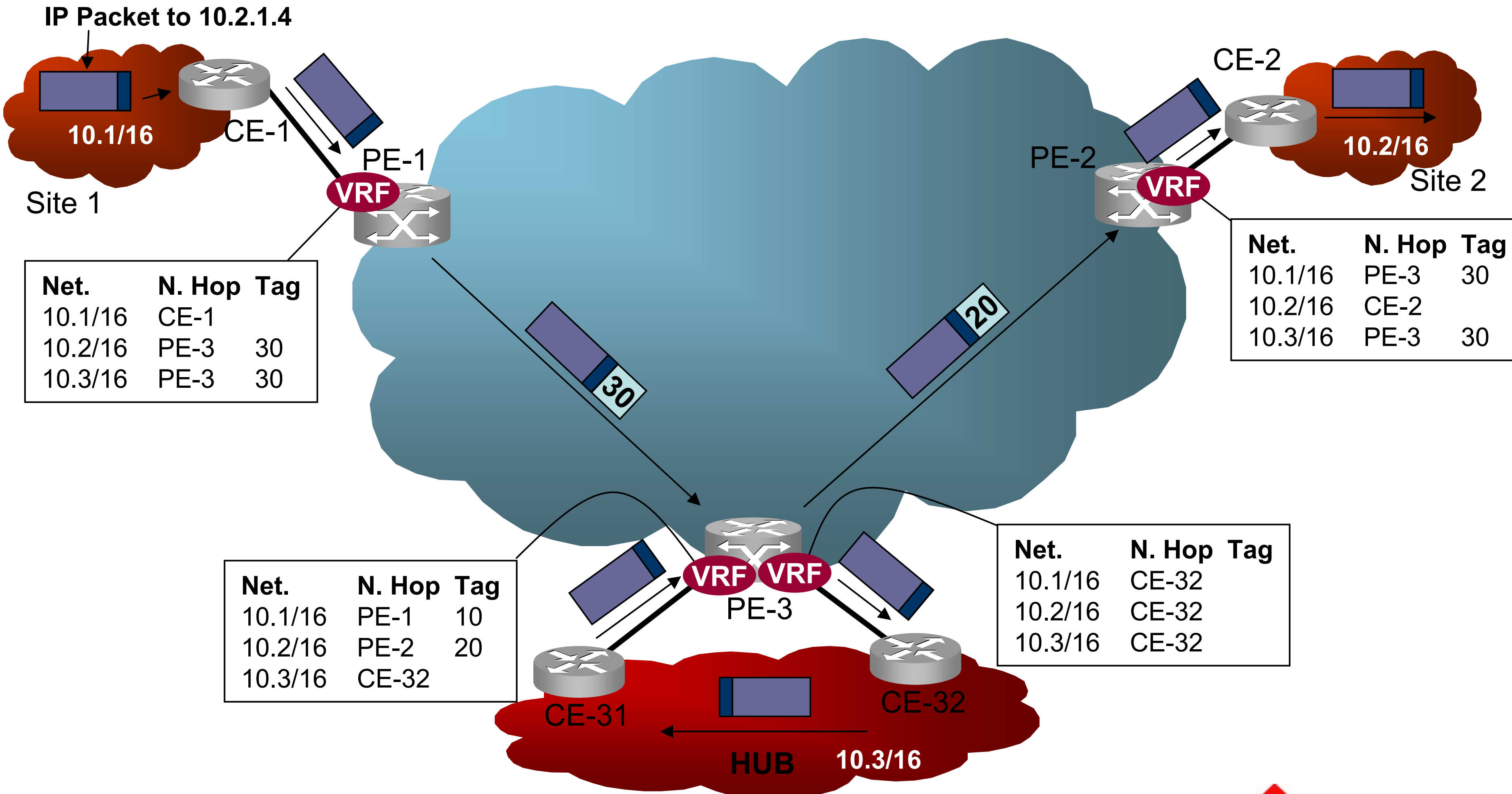
**PE-1 and PE-2 will not install the routes learned from each other in their VRFs because of RT mismatch**

**VRF-1 is configured to forward all routes learned to CE-1**

# Building Hub and Spoke VPNs (2)



# Building Hub and Spoke VPNs (3)




PE-1:

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | CE-1 | |
| 10.2/16 | PE-3 | 30 |
| 10.3/16 | PE-3 | 30 |

PE-2:

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-3 | 30 |
| 10.2/16 | CE-2 | |
| 10.3/16 | PE-3 | 30 |

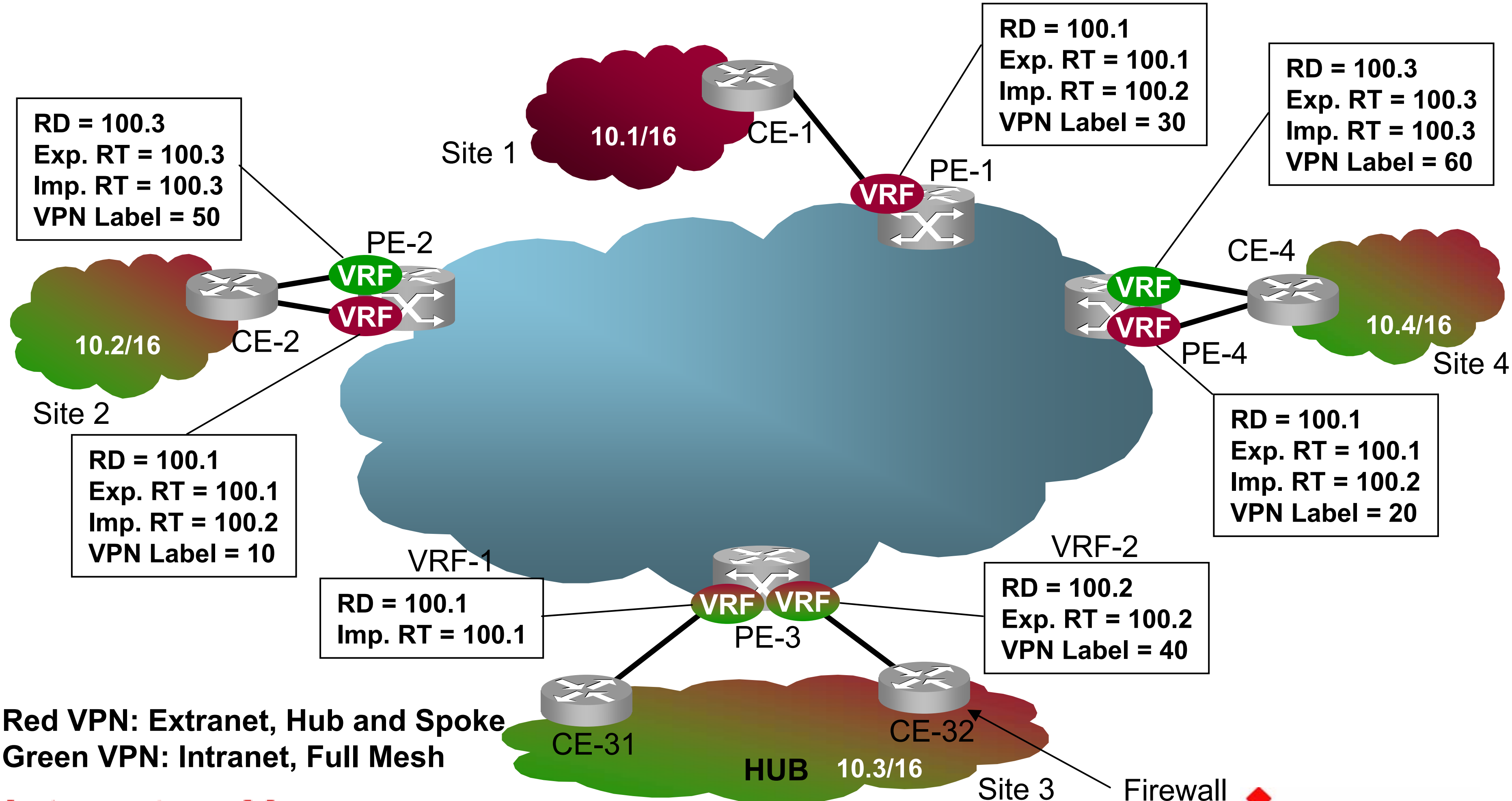
PE-3 (left VRF):

| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | PE-1 | 10 |
| 10.2/16 | PE-2 | 20 |
| 10.3/16 | CE-32 | |

PE-3 (right VRF):

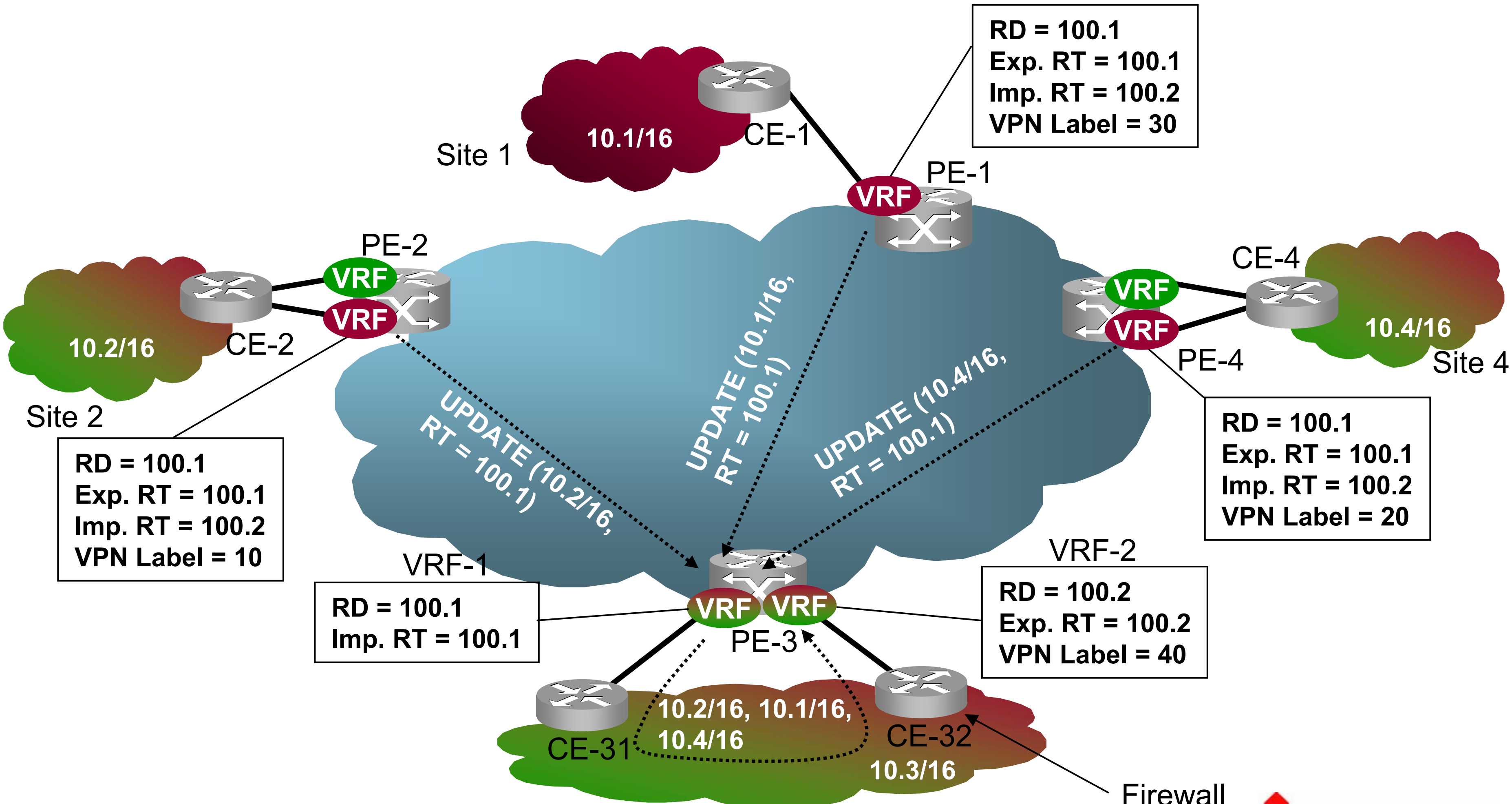
| Net. | N. Hop | Tag |
|---|---|---|
| 10.1/16 | CE-32 | |
| 10.2/16 | CE-32 | |
| 10.3/16 | CE-32 | |

# Overlapping Intranet and Extranet VPNs (1)

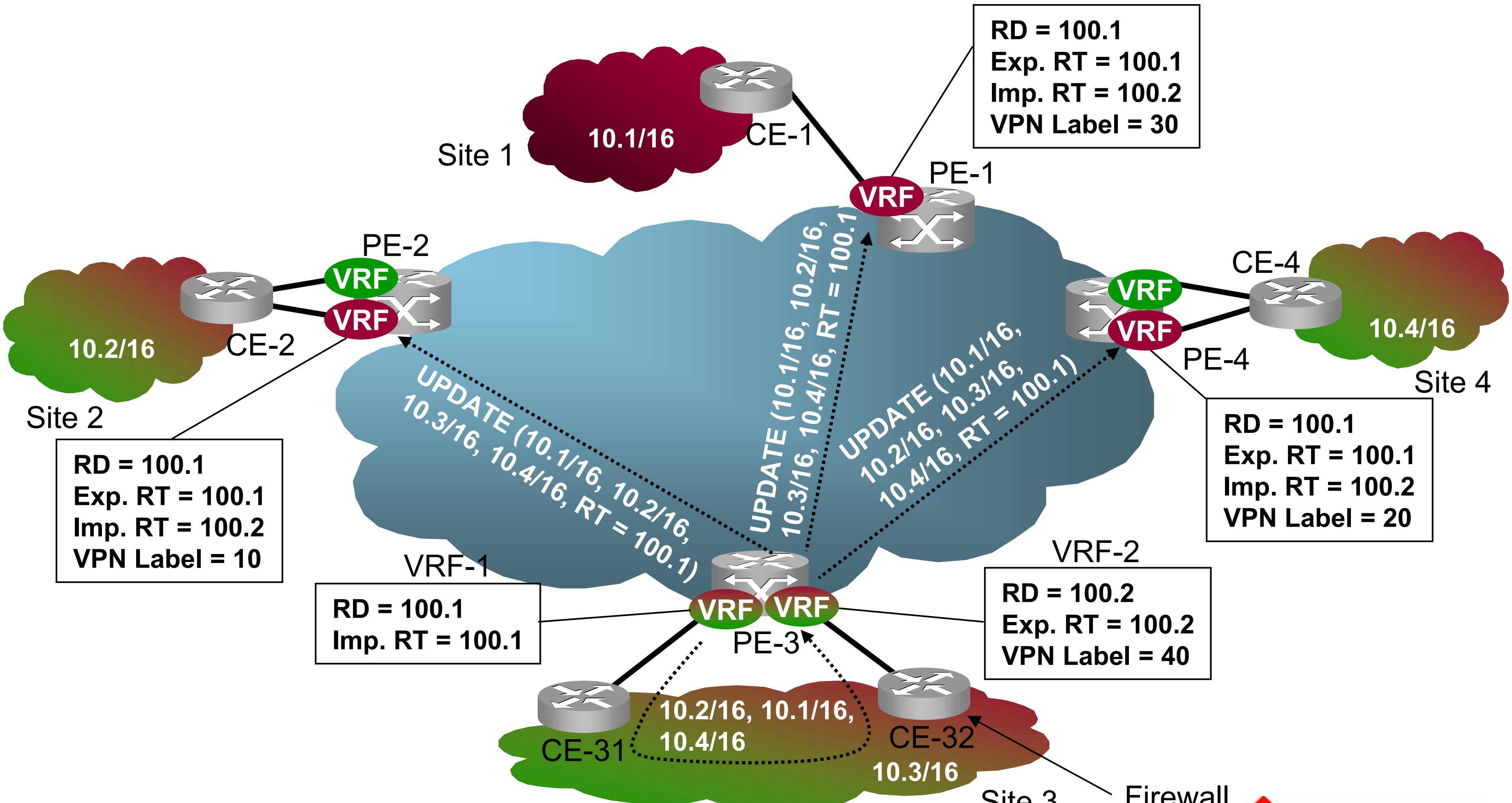



**Red VPN: Extranet, Hub and Spoke**
**Green VPN: Intranet, Full Mesh**

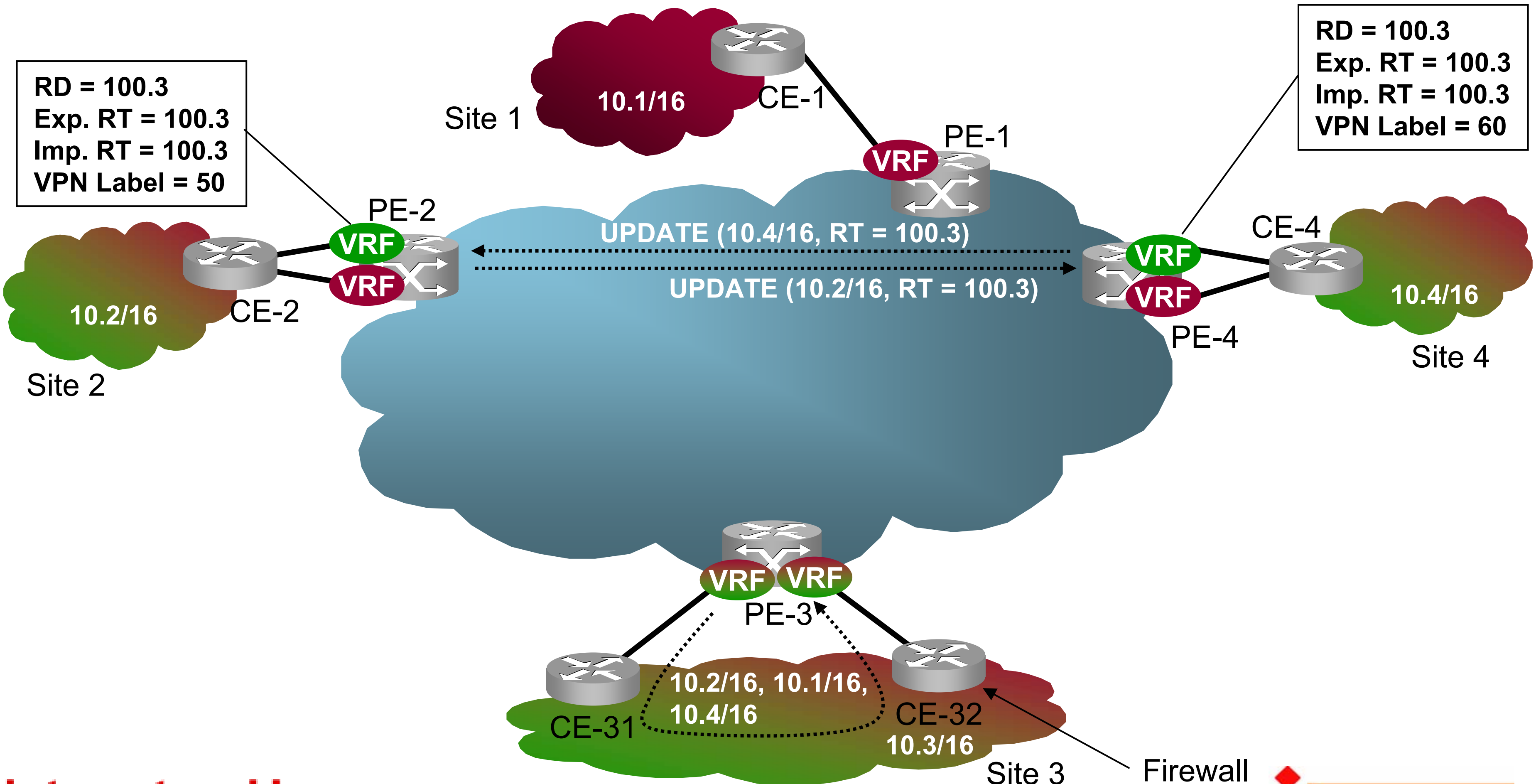
# Overlapping Intranet and Extranet VPNs (2)



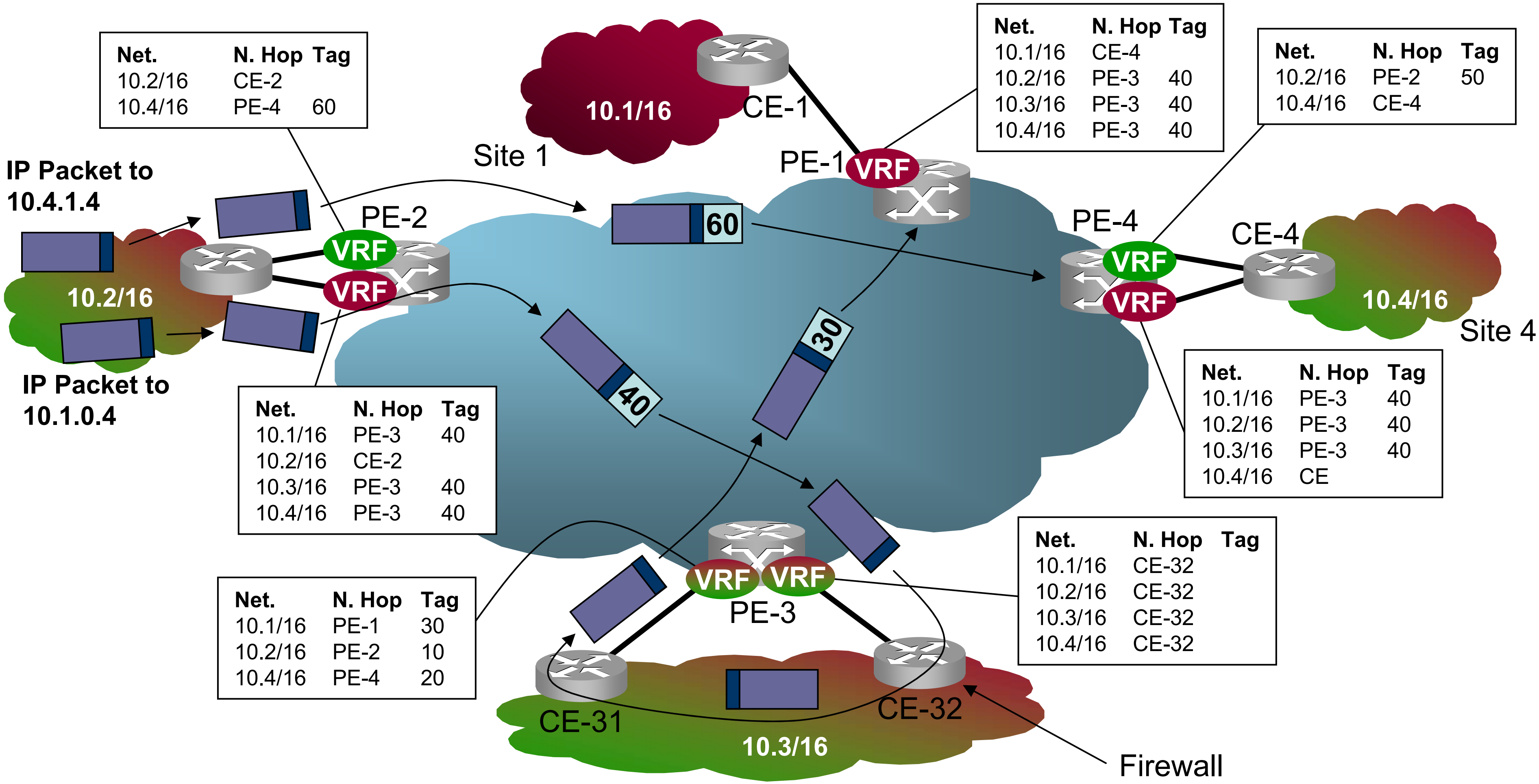
# Overlapping Intranet and Extranet VPNs (3)



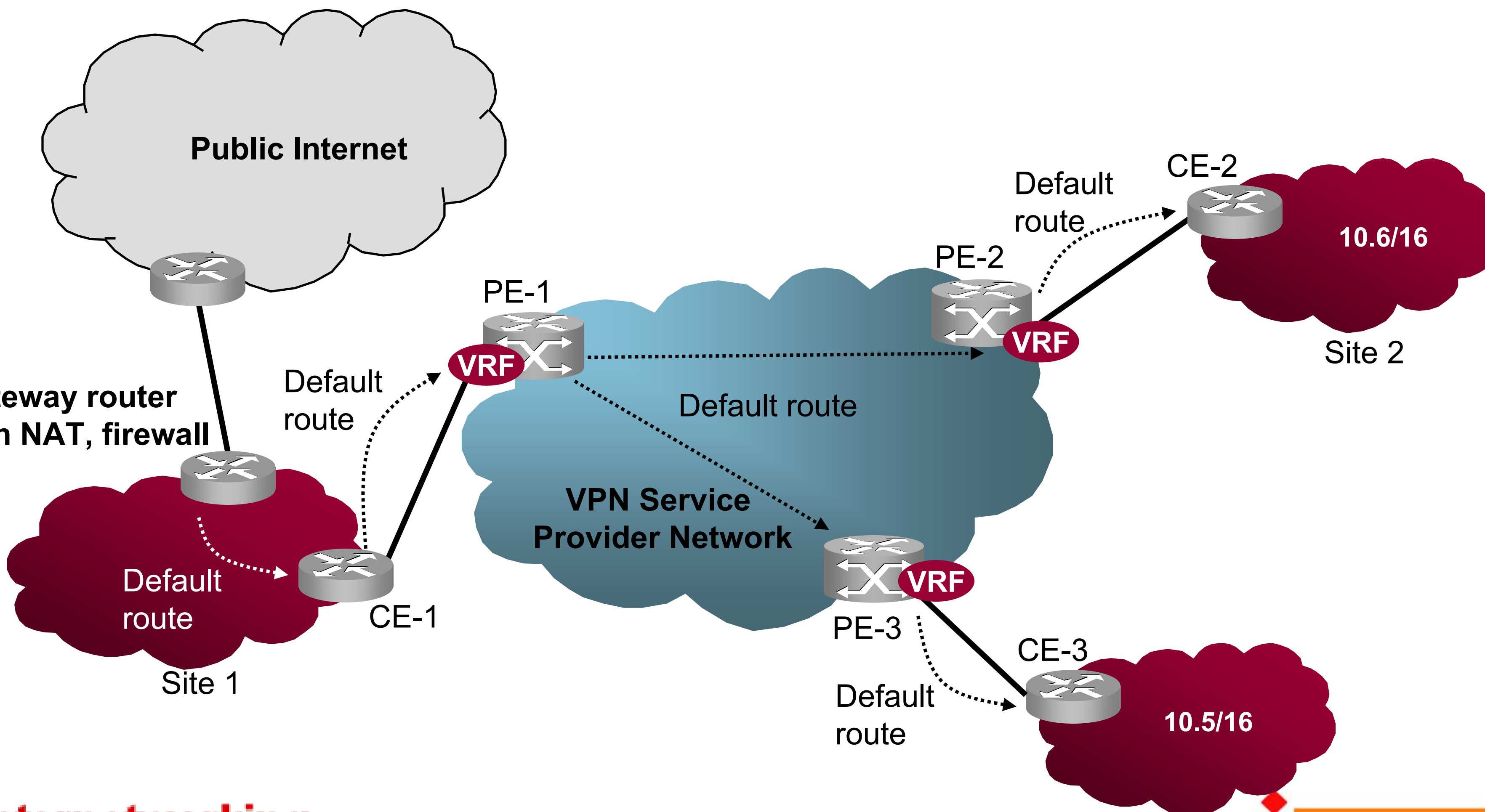
# Overlapping Intranet and Extranet VPNs (4)



# Overlapping Intranet and Extranet VPNs (5)
# Resulting VRFs

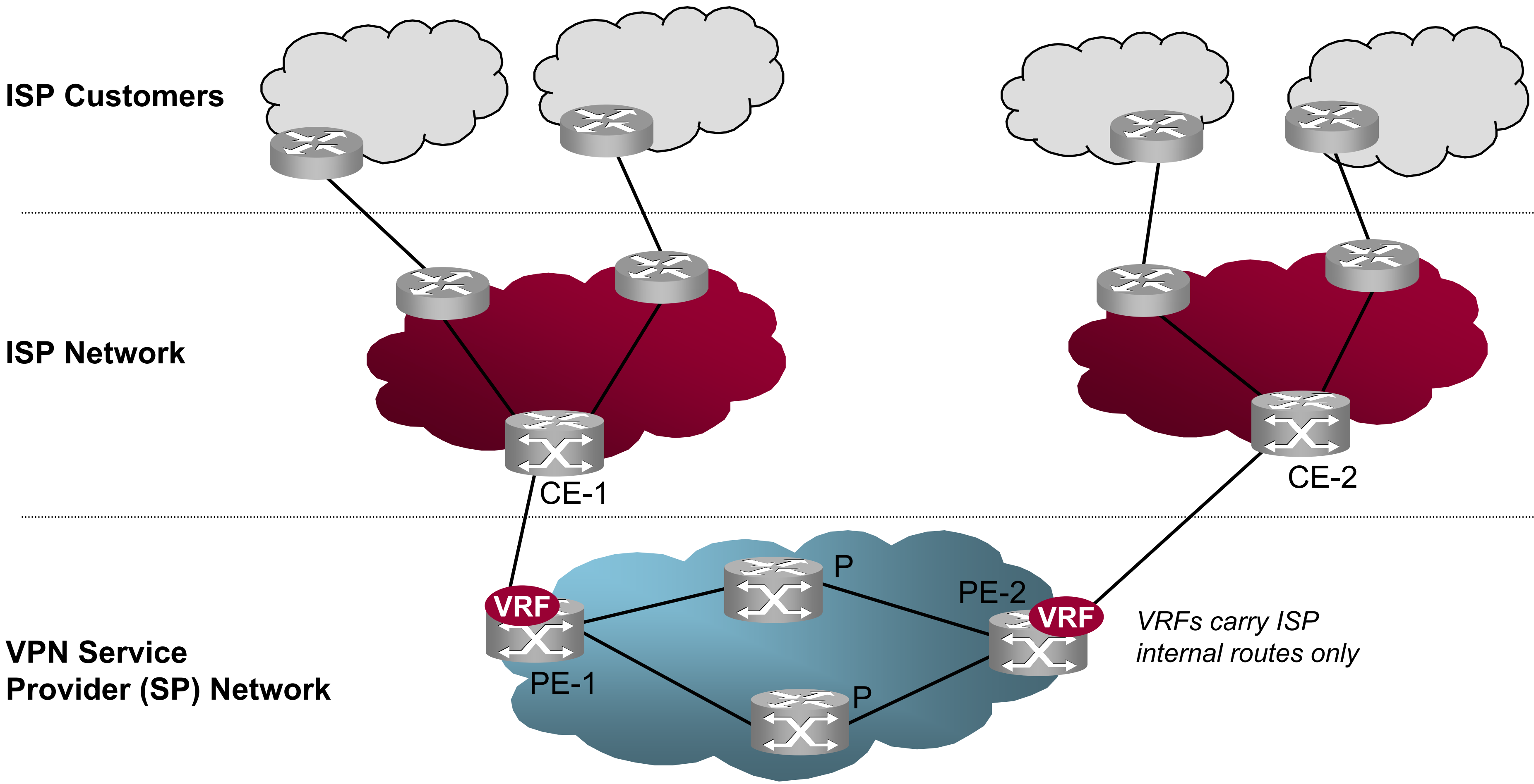


# Accessing Public Internet from a VPN

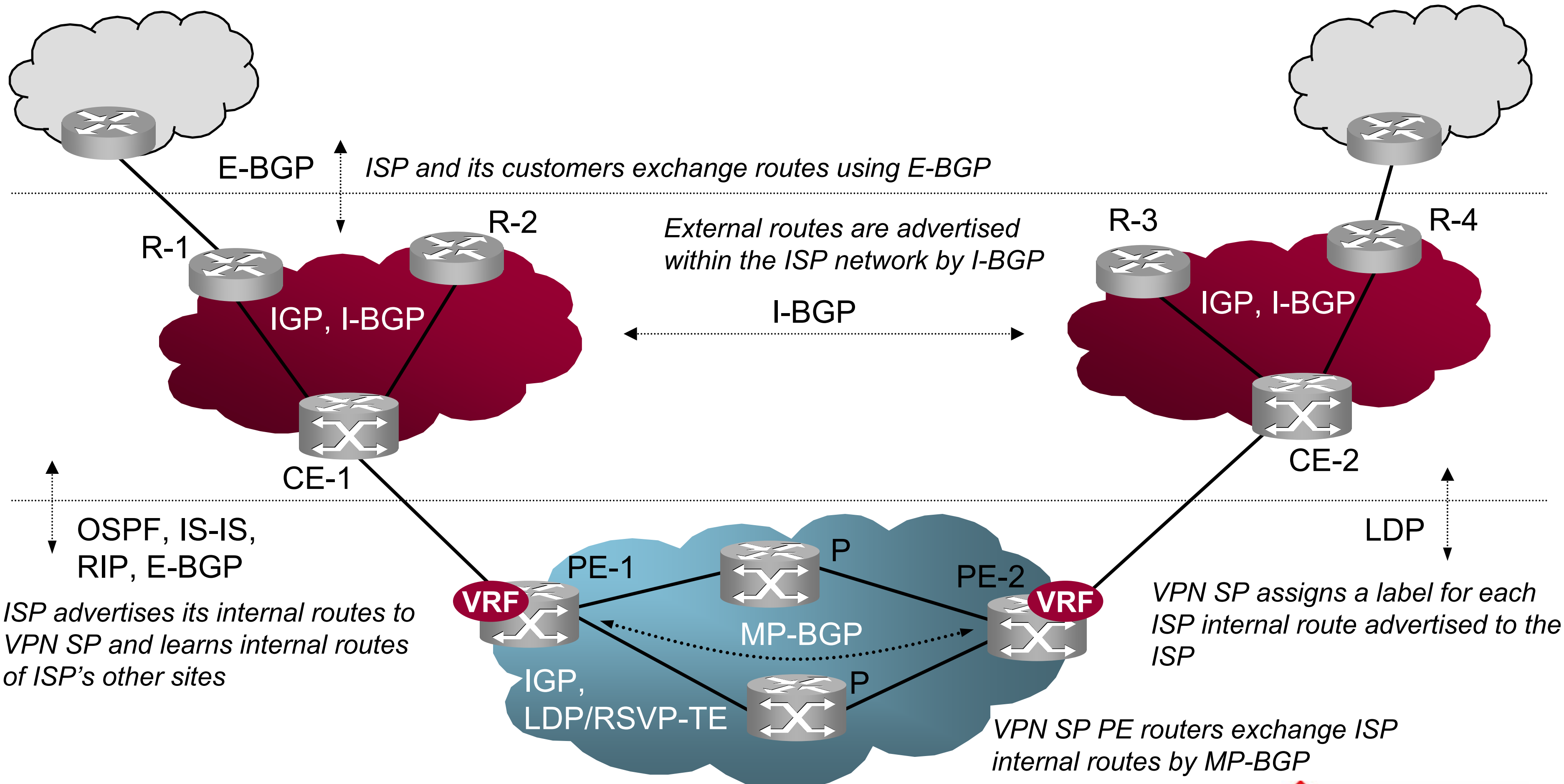


# Hierarchical BGP/MPLS VPNs

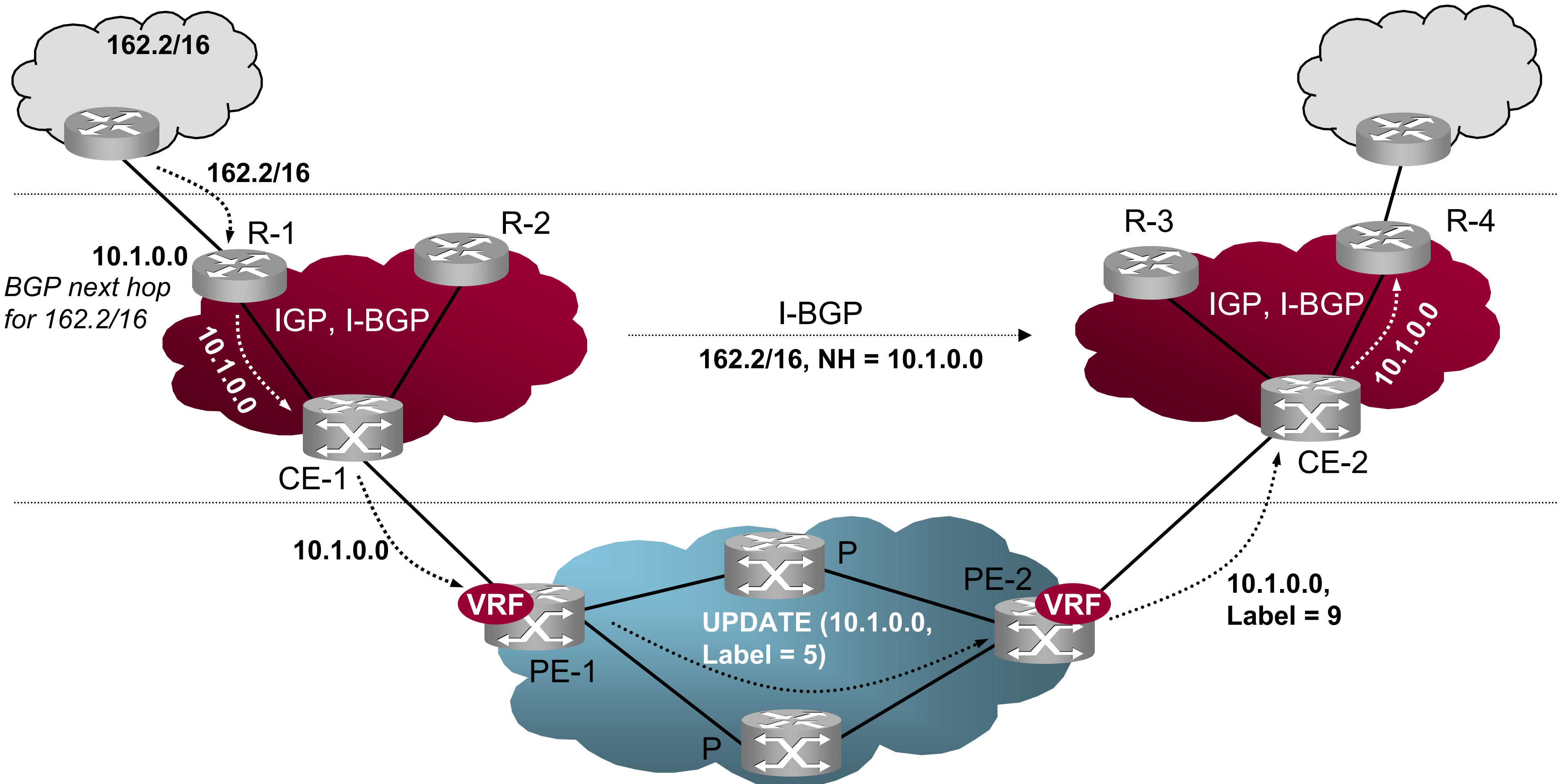
# ISP is a Customer of VPN Service Provider (1)



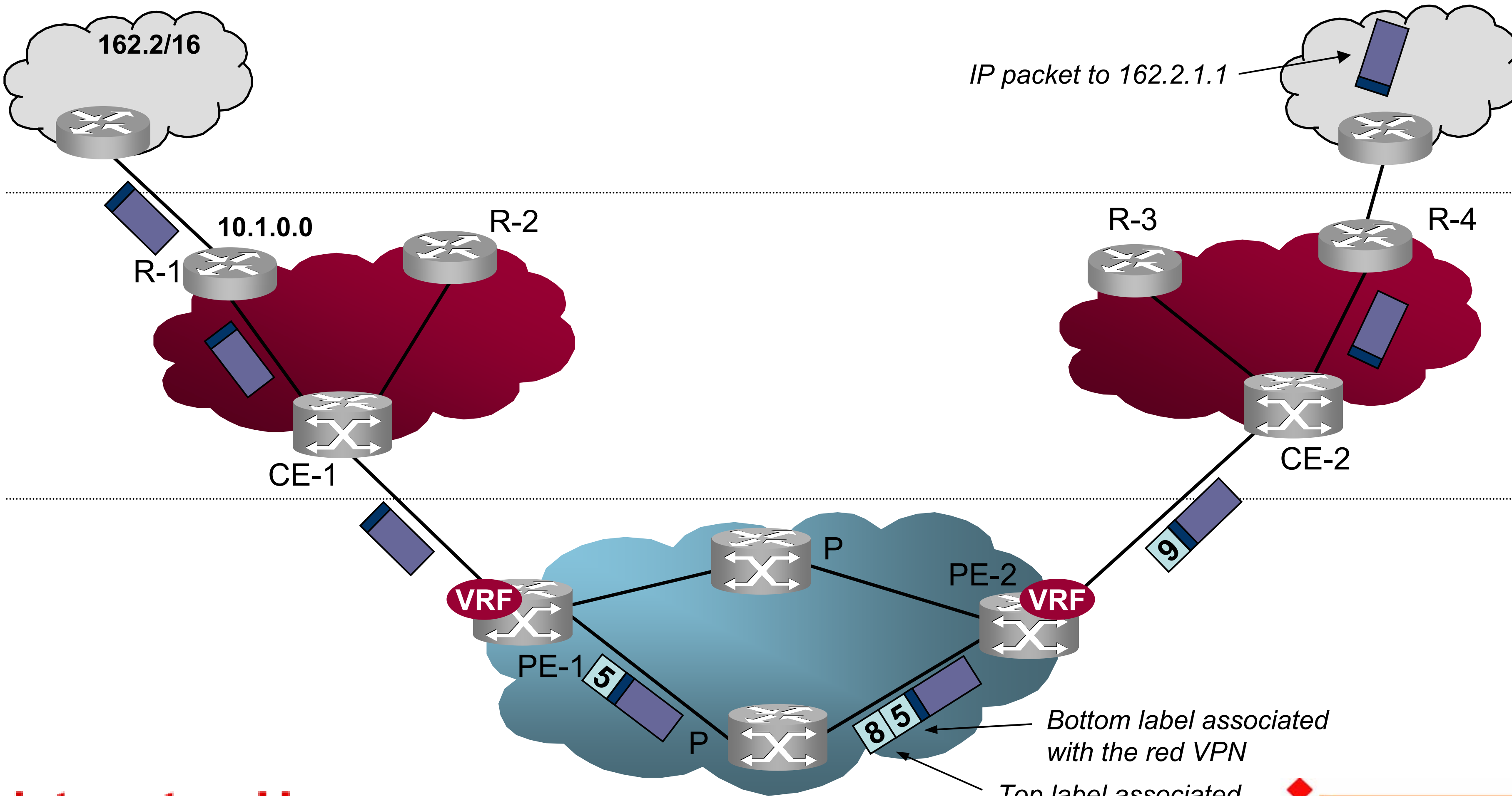
# ISP is a Customer of VPN Service Provider (2)



# ISP is a Customer of VPN Service Provider (3)



# ISP is a Customer of VPN Service Provider (4)

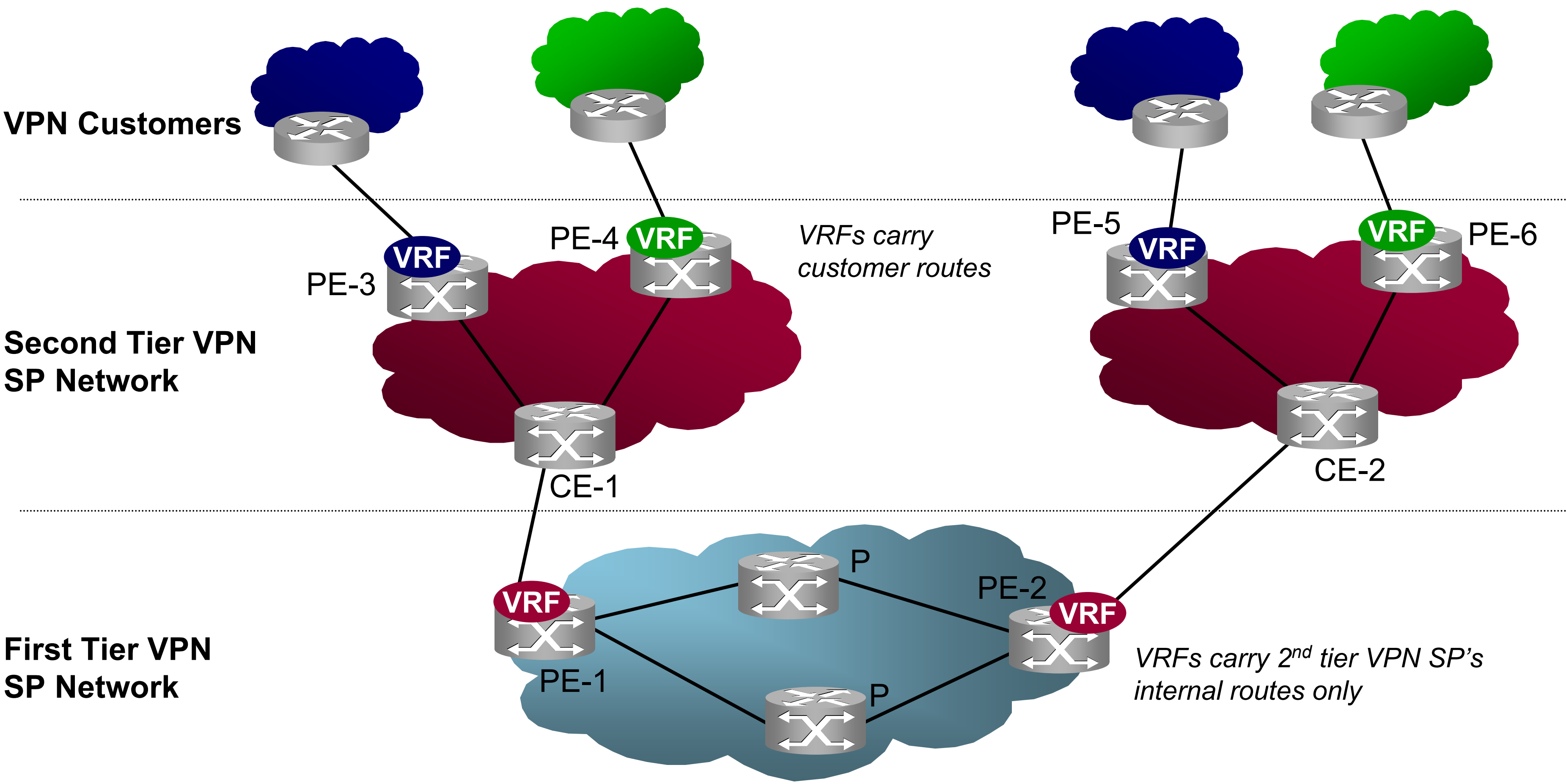




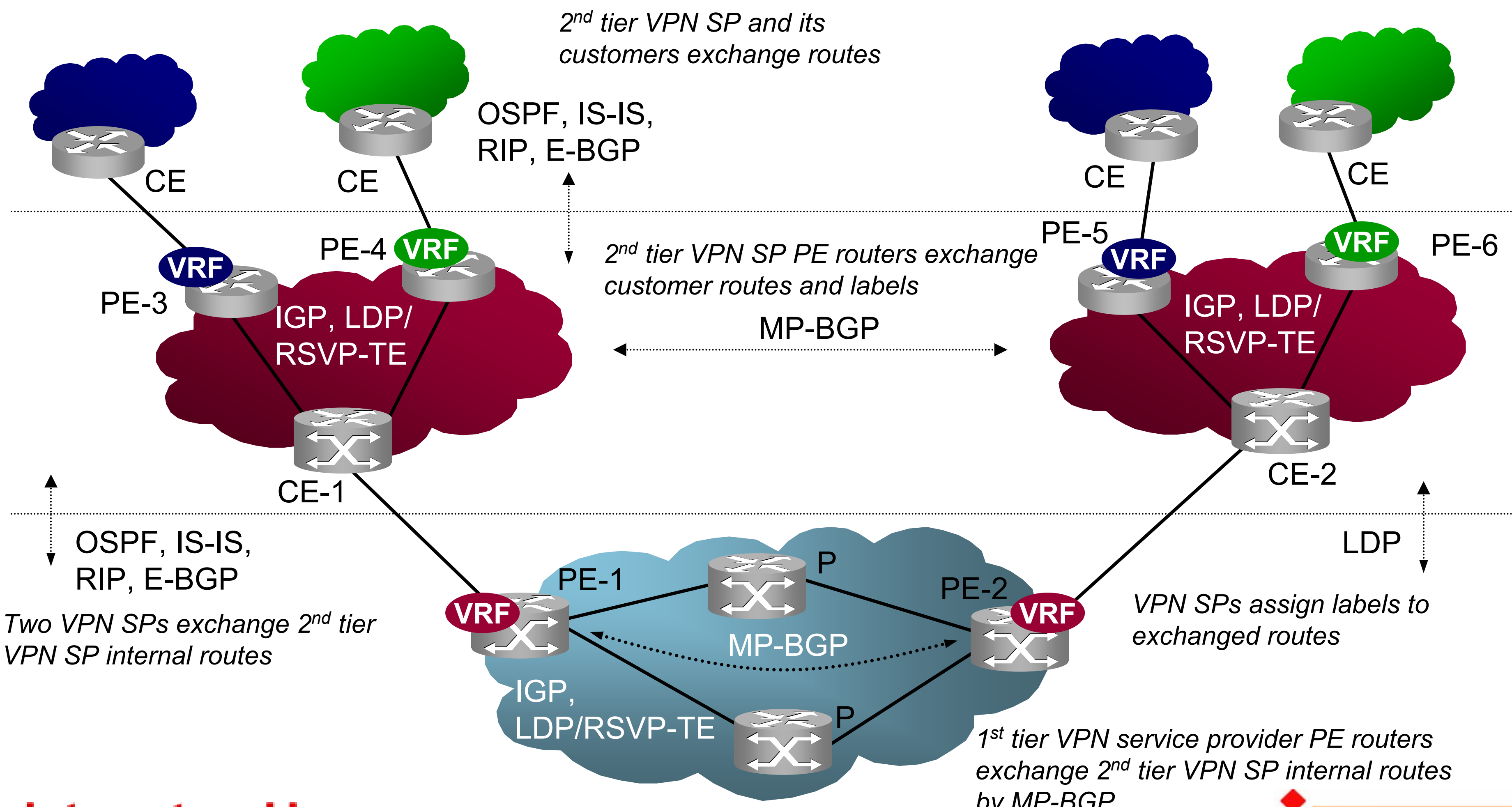

Internetworking 2003

ISOCORE

# VPN service Provider is a Customer of another VPN Service Provider (1)



# VPN service Provider is a Customer of another VPN Service Provider (2)



# References – BGP/MPLS-based VPNs

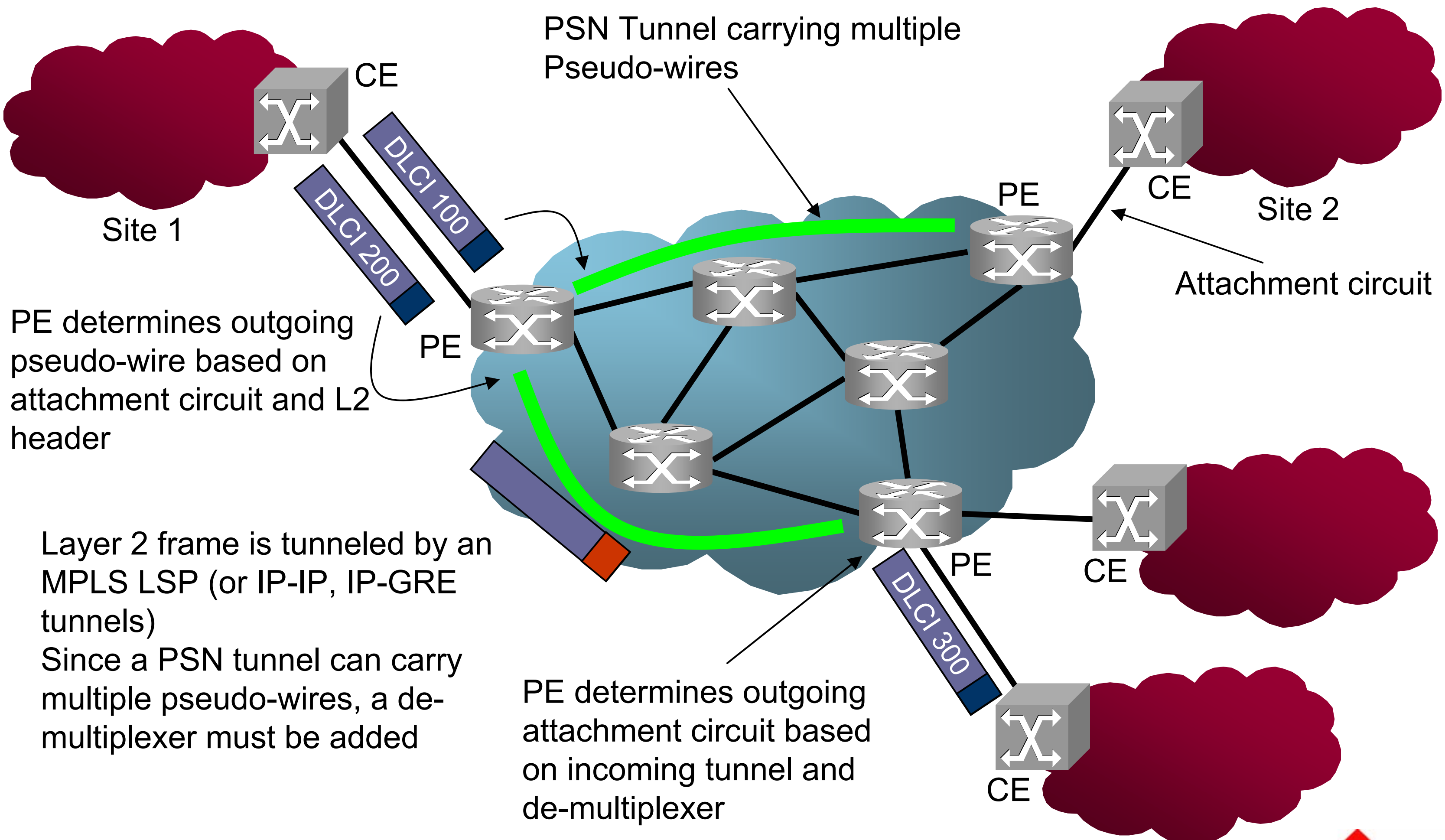

- draft-ietf-ppvpn-rfc2547bis-04.txt, "BGP/MPLS IP VPNs"
- draft-ietf-ppvpn-requirements-06.txt, "Service Requirements for Layer 3 Provider Provisioned Virtual Private Networks"
- draft-ietf-ppvpn-framework-08.txt, "A Framework for Layer 3 Provider Provisioned Virtual Private Networks"
- draft-ietf-ppvpn-as2547-01.txt, "Applicability Statement for VPNs Based on rfc2547bis"
- draft-ietf-ppvpn-ipsec-2547-03.txt, "Use of PE-PE IPsec in RFC2547 VPNs"
- draft-ietf-ppvpn-gre-ip-2547-02.txt, "Use of PE-PE GRE or IP in RFC2547 VPNs"
- J. Guichard and I. Pepelnjak, "MPLS and VPN Architectures," Cisco Press, 2000

# MPLS-based Layer 2 VPNs

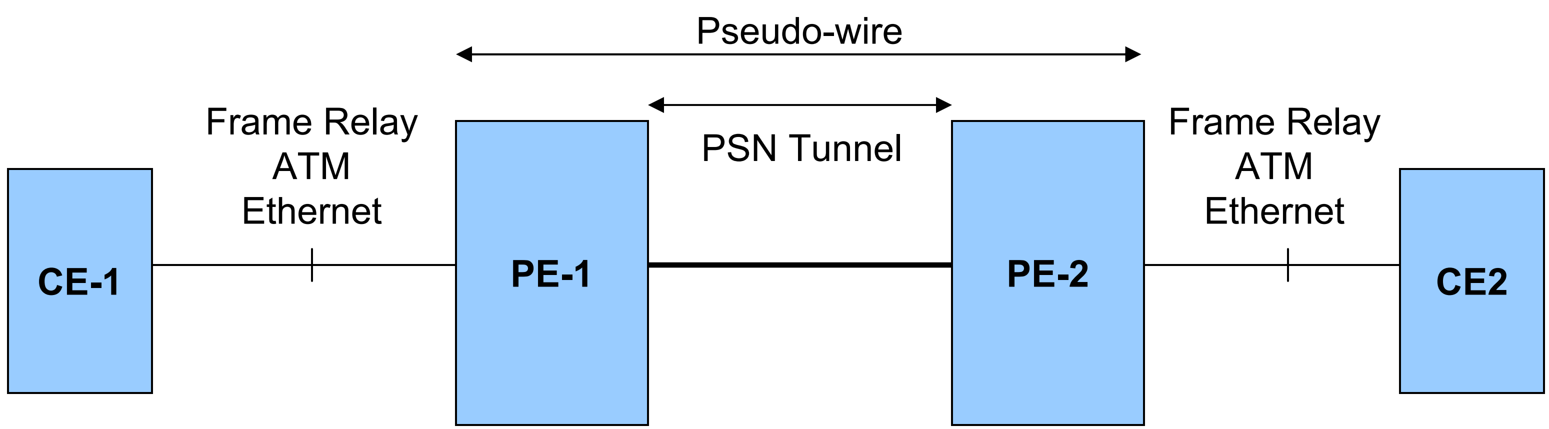
# MPLS-Based Layer 2 VPNs

- Virtual Private Wire Service (VPWS)
  - Point-to-point connectivity between CE devices by pseudo-wires over an IP network
  - SP network acts as a Layer 2 switch
  - Mapping to pseudo-wires can be based on incoming port or Layer 2 header
- Virtual Private LAN Service (VPLS)
  - Point-to-multipoint connectivity between CE devices
  - Forwarding of incoming packets is based on Ethernet addresses
  - SP network acts as a LAN bridge

# Virtual Private Wire Service (VPWS)

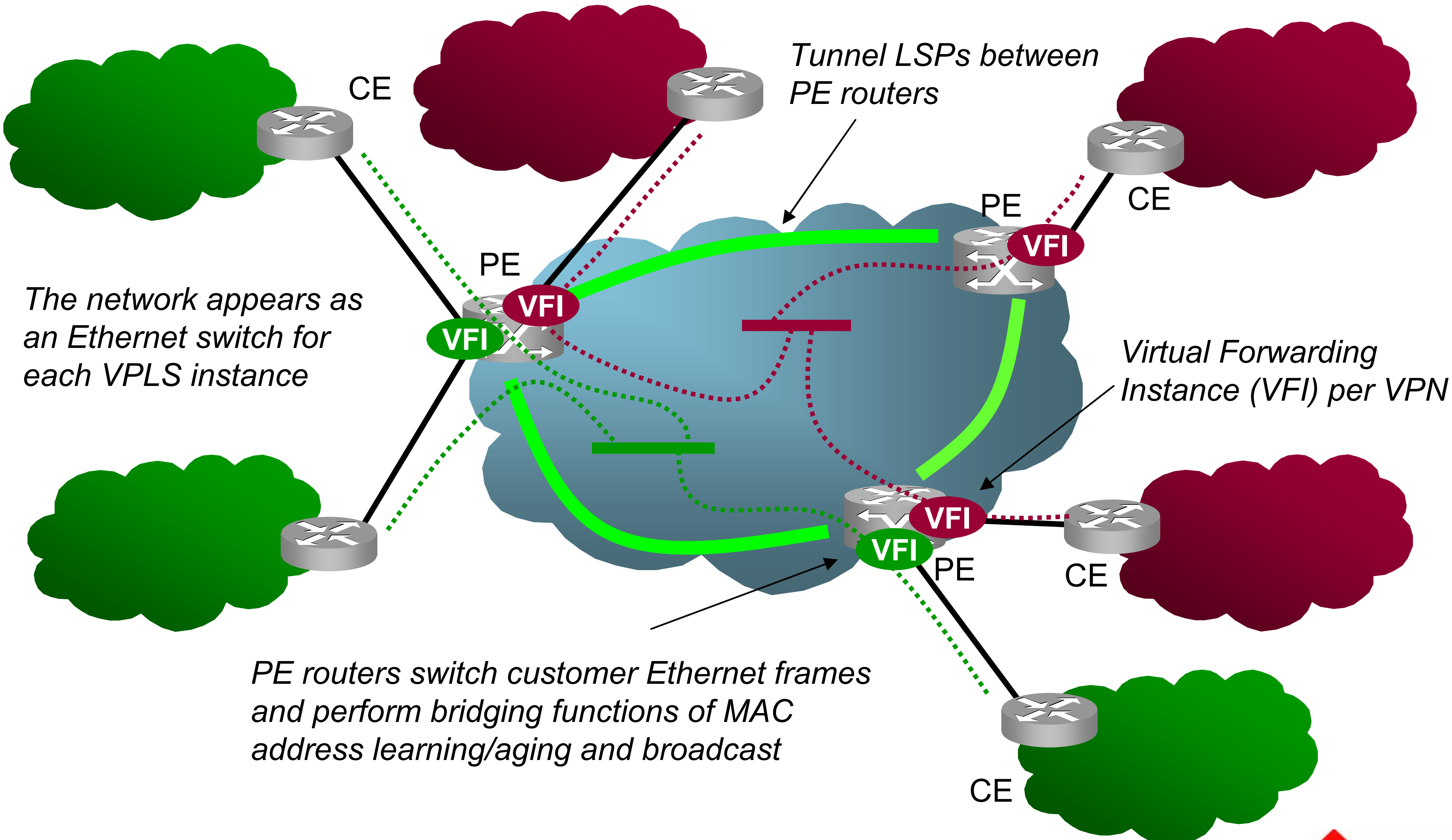


# VPWS Reference Model and Encapsulation




| | |
|---|---|
| **L2 Frame** | *Ethernet or Frame Relay packet, ATM cell, or ATM AAL-5 PDU L2 frame can be carried with or without original header* |
| **L2 Protocol Specific Information** | *Control word (sequence number, length, and L2 protocol flags)* |
| **Demultiplexer (VC label)** | *MPLS Label* |
| **PSN Header** | *MPLS Label* |

# Virtual Private LAN Service (VPLS)



# VPLS Issues(2)

- Scalability
  - N(N-1) VCs must be setup between PE devices for one VPLS service with N customer nodes
    - Signaling overhead
    - Packet replication requirements
  - Hierarchical VPLS can improve scalability
- Signaling
  - Currently LDP and BGP are being proposed for establishing VPLS pseudo-wires
- Node and Service Discovery
  - Capability for a PE router to discover other VPLS-capable routers
  - Proposed methods include LDP, BGP, DNS, and Radius

# References – MPLS-based Layer 2 VPNs


- draft-ietf-ppvpn-l2-framework-03.txt, “L2VPN Framework”
- draft-ietf-ppvpn-l2vpn-requirements-00.txt, “Service Requirements for Layer 2 Provider Provisioned Virtual Private Networks”
- draft-ietf-pwe3-ethernet-encap-02.txt, “Encapsulation Methods for Transport of Ethernet Frames Over IP/MPLS Networks”
- draft-ietf-pwe3-frame-relay-00.txt, “Frame Relay over Pseudo-Wires”
- draft-ietf-pwe3-atm-encap-01.txt, “Encapsulation Methods for Transport of ATM Cells/Frame Over IP and MPLS Networks”
- draft-ietf-pwe3-control-protocol-02.txt, “Pseudo-wire Setup and Maintenance using LDP”
- draft-lasserre-vkompella-ppvpn-vpls-04.txt, “Virtual Private LAN Services over MPLS”
- draft-ietf-ppvpn-vpls-bgp-00.txt, “Virtual Private LAN Service”