MPLS L2/L3
Virtual Private Networks (VPNs)

An IP/MPLS Forum Sponsored Tutorial

Dave Christophe
IP/MPLS Forum Education WG Chair
Director, Solutions Marketing
Alcatel-Lucent

MPLS VPN Tutorial Agenda

- Introduction to the IP/MPLS Forum
- Introduction to MPLS and MPLS VPNs
  - Defining Layer 2 and 3 VPNs
- Layer 3 MPLS VPN
  - Overview
  - BGP Review
  - RFC 4364 (2547bis) Key Characteristics
  - BGP/MPLS VPN Architecture Overview
    - VPN Routing and Forwarding (VRF) Tables
    - Overlapping VPNs
    - VPN Route Distribution
    - VPN Packet Forwarding
    - Scaling L3 VPNs and Route Reflectors

MPLS VPN Tutorial Agenda

- Layer 2 VPNs
  - Overview
  - Encapsulation and Label Stacking
  - Virtual Private Wire Services – VPWS
    - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
  - Virtual Private LAN Services – VPLS
- Introduction to Multi-Service Interworking over MPLS
  - Interworking History and Definition
  - Multi-Service Interworking of Ethernet over MPLS
  - Migration Scenarios and Benefits
- Summary

Introduction to the IP/MPLS Forum

- IP/MPLS Forum is an international, industry-wide, non-profit association of service providers, equipment vendors, testing centers and enterprise users
  - Created with the name change of the MFA Forum (Oct 2007) to reflect renewed focus on driving global industry adoption of IP/MPLS solutions in the market, by focusing on standards initiatives for IP/MPLS such as inter carrier interconnect (ICI), mobile wireless backhaul, and security.
- Objectives: Unify service providers, suppliers and end users on common vision of IP/MPLS based solutions

Awareness

- Promote global awareness of the benefits of IP/MPLS
- Empower the telecom industry to migrate from legacy technologies to IP/MPLS-based next generation networking

Migration

- Guide the telecom end user to make the leap from legacy technologies to IP/MPLS-based services

Systems-Level Solutions

- Drive implementation of standards for IP/MPLS based solutions
- Validate implementations and advance interoperability of standardized IP/MPLS based solutions

- Deliverables: Technical Specifications, Test Plans, Technical Tutorials, Collateral

Introduction to the IP/MPLS Forum

- **Current Work Items**
  - Framework and Reference Architecture for MPLS in Mobile Backhaul Networks
  - MPLS Inter-Carrier Interconnect
  - Packet Based GMPLS Client to Network Interconnect
  - Generic Connection Admission Control (GCAC) Requirements for IP/MPLS Networks
  - Layer 2 VPNs using BGP for Auto-discovery & Signaling (BGP L2 VPN)
  - MPLS Over Aggregated Interface
  - Voice Trunking format over MPLS
  - TDM Transport over MPLS using AAL1

  The Forum is also planning several industry-driven future Work Items.
- **Service Provider Council**
- **Public Interoperability Events**
- **Technical Tutorials -** to broaden the understanding of the technology and benefits of the solutions
- Next meeting: June 24-26, Vancouver, Canada
- Please join us!
  - **To join the Forum contact Alysia Johnson, Executive Director**

    E-Mail: ajohnson@ipmplsforum.org

    Phone: 510 492-4057

Technical Tutorials

Introduction to MPLS	**½ and full day**
MPLS L2/L3 VPNs	**½ day**
MPLS VPN Security	**½ day**
Traffic Engineering	**½ day**
GMPLS	**½ day**
Migrating Legacy Services to MPLS	**½ day**
MPLS OAM	**½ day**
Voice over MPLS	**½ day**
Multi-service Interworking over MPLS	**½ day**
Multicast in MPLS/VPLS Networks	**½ day**
IP/MPLS in the Mobile RAN	**½ day**
MPLS Inter-Carrier Interconnect	**½ day**

New tutorials based upon demand

Section 1

Introduction to MPLS and MPLS VPNs

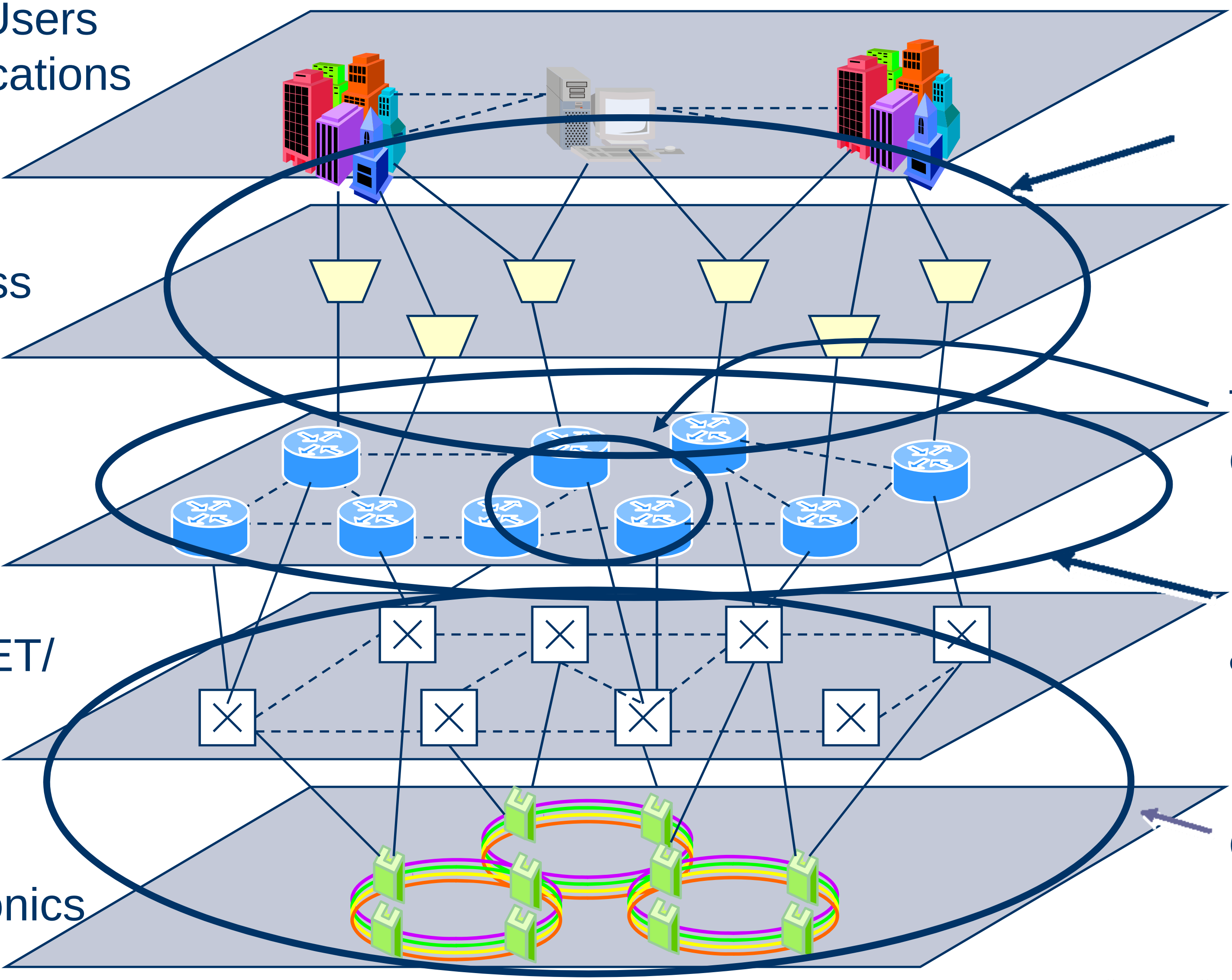
Why MPLS ?

A Common Control Plane

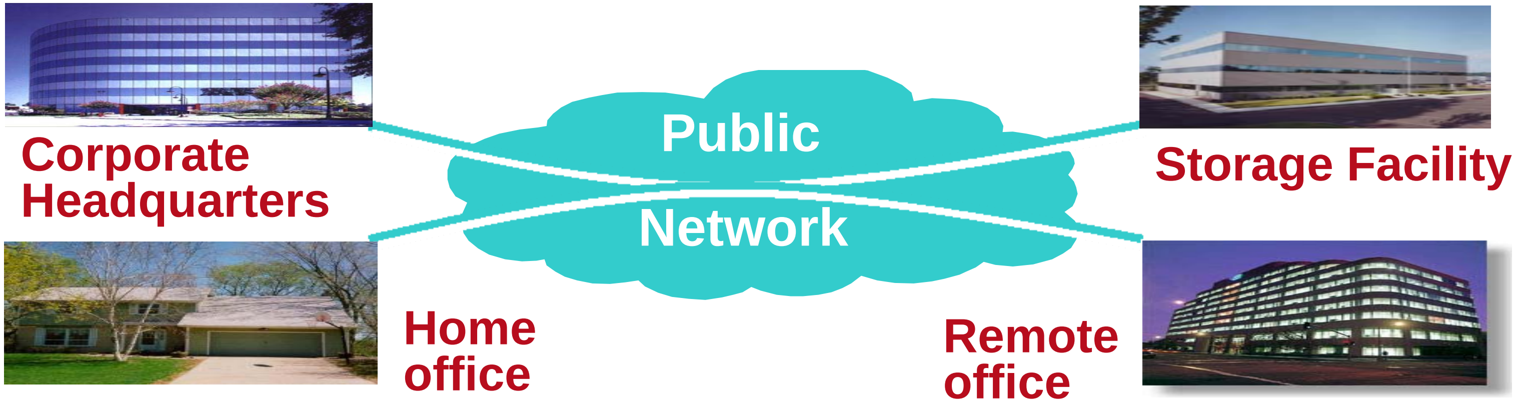
Best of the packet-switched and circuit-switched worlds

- Enhancement and scalability of IP
- Layer 2 and Layer 3 VPNs
- Link Resiliency and Path Protection
- Metro Ethernet Services
- Differentiated Services - CoS and QoS
- Legacy Network Migration

MPLS: Addresses many network needs



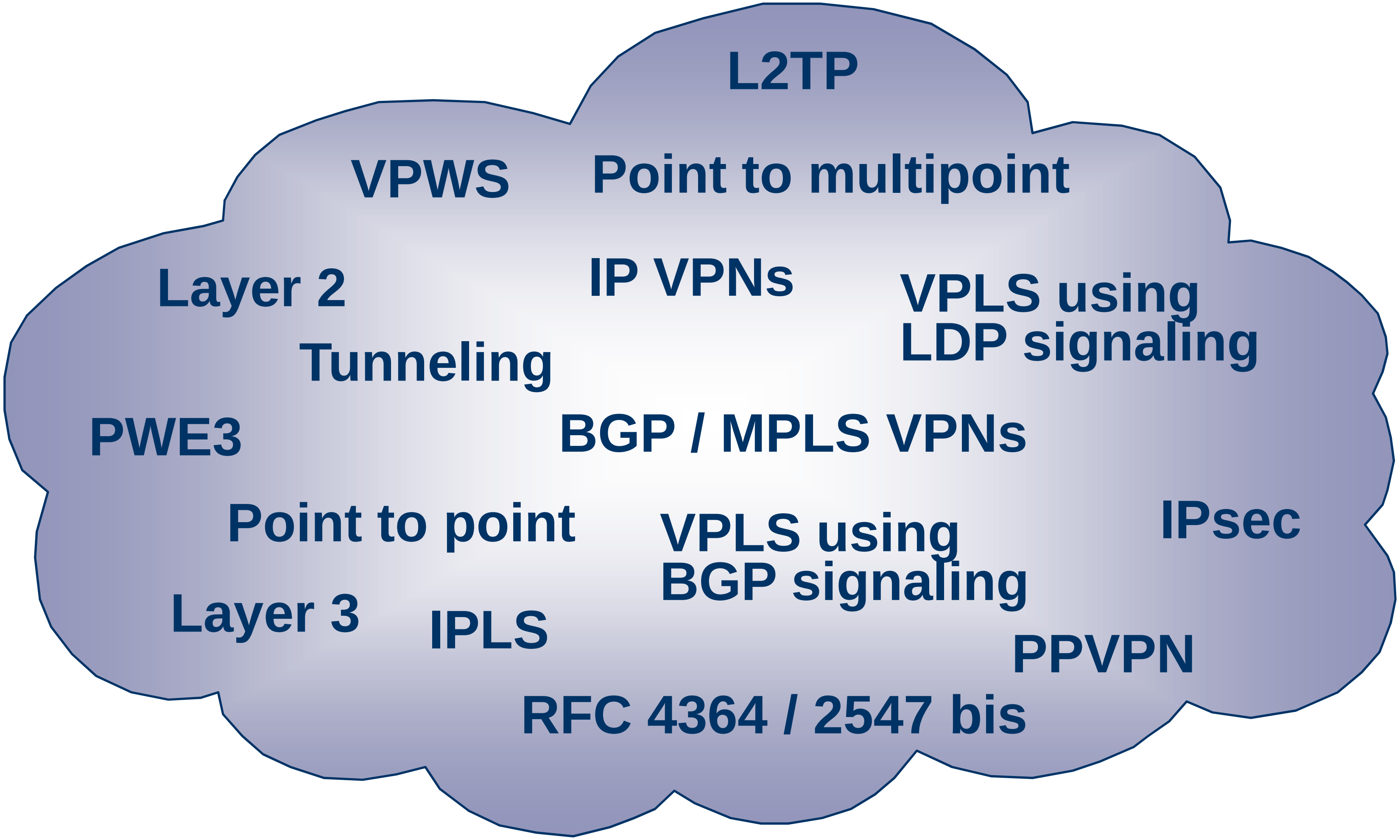
Virtual Private Networks



- VPN (Virtual Private Network) is simply a way of using a public network for private communications, among a set of users and/or sites
- Remote Access: Most common form of VPN is dial-up remote access to corporate database - for example, road warriors connecting from laptops
- Site-to-Site: Connecting two local networks (may be with authentication and encryption) - for example, a Service Provider connecting two sites of the same company over its shared network

MPLS, VPNs, and Standards

Many options



L2TP

VPWS

Point to multipoint

IP VPNs

Layer 2

VPLS using LDP signaling

Tunneling

PWE3

BGP / MPLS VPNs

Point to point

IPsec

VPLS using BGP signaling

Layer 3

IPLS

PPVPN

RFC 4364 / 2547 bis

VPNs

Types, Layers, and Implementations

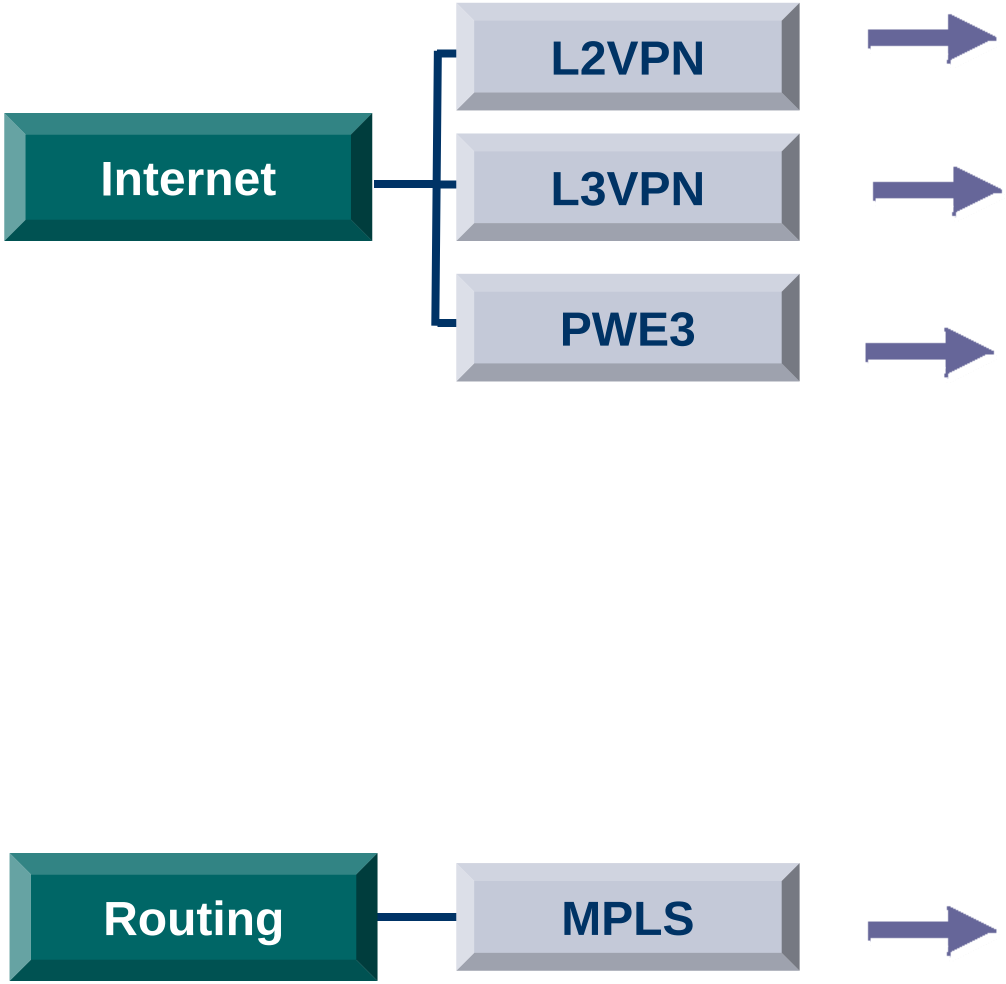
VPN Type	Layer	Implementation
Leased Line	1	TDM/SDH/SONET
Frame Relay	2	DLCI
ATM	2	VC
GRE/UTI/L2TPv3	3	IP Tunnel
Ethernet	2	VLAN / VPWS / VPLS
IP	3	RFC 4364 / VR
IP	3	IPsec

VPNs

How do they compare?

	FR or ATM	IPsec	L3 MPLS	L2 MPLS
Point-to-multipoint	×	×	√	√
Multi-protocol	√	×	×	√
QoS and CoS	√	×	√	√
Low latency	√	×	√	√
Security	√	√	√	√
SLAs	√	×	√	√

MPLS VPNs in the IETF

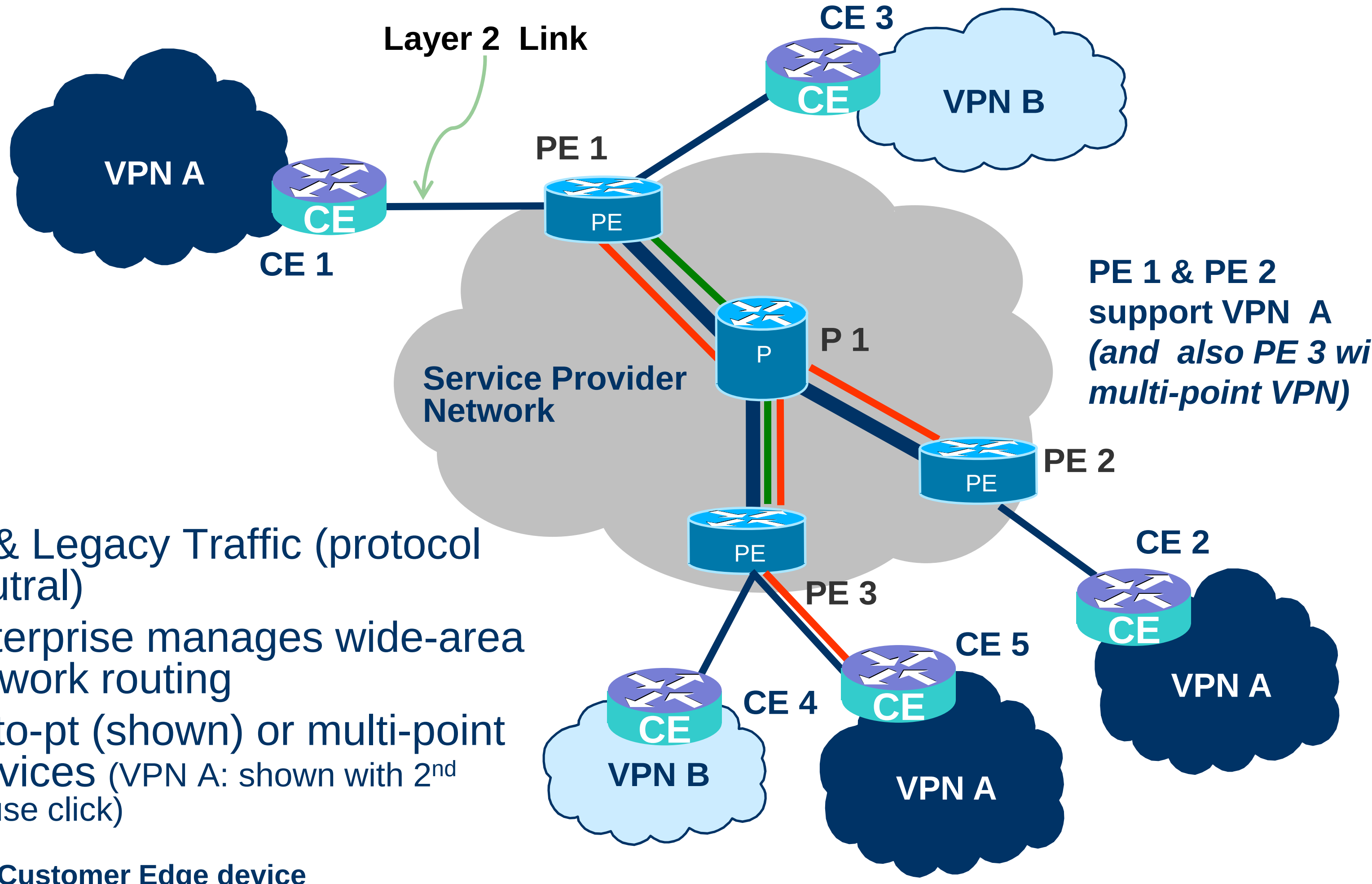


- Layer 2 VPNs
- Layer 3 VPNs
- Pt-to-Pt circuits
- Encapsulations
  - ATM
  - FR
  - Ethernet
  - PPP/HDLC
  - TDM
  - SONET/SDH
- Base Technology

What are Layer 2 and Layer 3 VPNs?

- VPNs based on a Layer 2 (Data Link Layer) technology and managed at that layer are defined as Layer 2 VPNs (MPLS, ATM, Frame Relay)
- VPNs based on tunneling at Layer 3 (Network Layer) are Layer 3 VPNs, (BGP/MPLS, VR, IPSec)

Visually - Layer 2 VPN



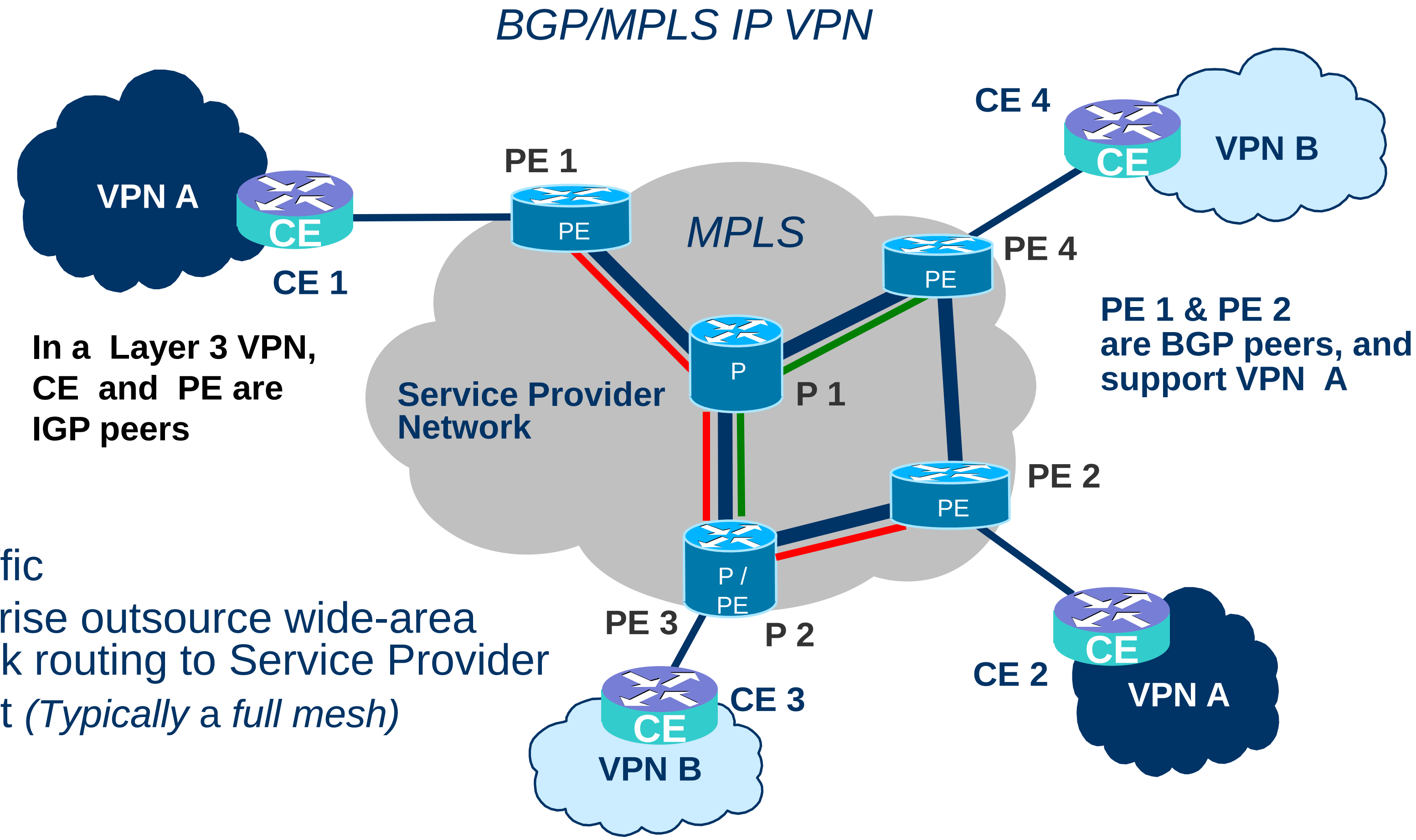
PE 1 & PE 2 support VPN A *(and also PE 3 with multi-point VPN)*

- IP & Legacy Traffic (protocol neutral)
- Enterprise manages wide-area network routing
- Pt-to-pt (shown) or multi-point services (VPN A: shown with 2nd mouse click)

CE: Customer Edge device
PE: Provider Edge router
P: Provider router not directly attached to a CE

VPN A 🟥
VPN B 🟩

Visually - Layer 3 VPN



In a Layer 3 VPN, CE and PE are IGP peers

PE 1 & PE 2 are BGP peers, and support VPN A

- IP Traffic
- Enterprise outsource wide-area network routing to Service Provider
- Pt-to-pt *(Typically a full mesh)*

CE: Customer Edge device
PE: Provider Edge router
P: Provider router not directly attached to a CE

VPN A	🟥
VPN B	🟩

*

Section 2

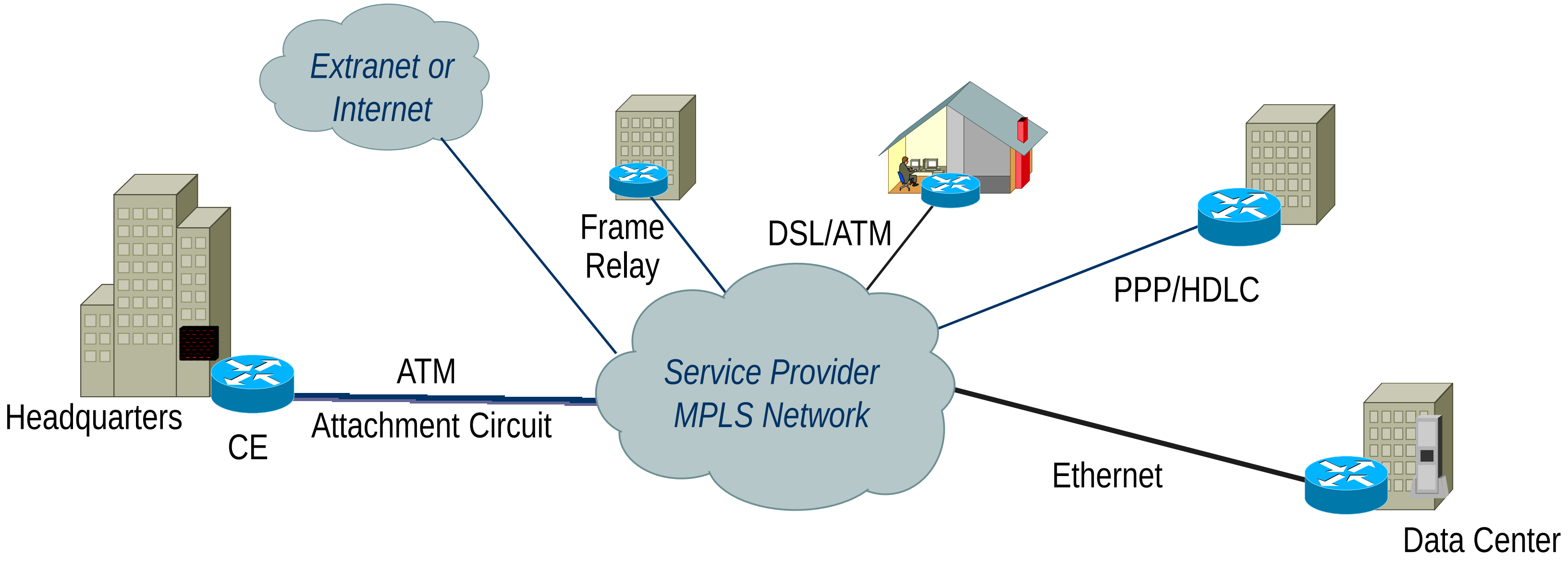
Layer 3 MPLS VPN

MPLS VPN Tutorial Agenda

Layer 3 MPLS VPN

- Overview
- BGP Review
- RFC 4364 / 2547bis Key Characteristics
- BGP/MPLS VPN Architecture Overview
  - VPN Routing and Forwarding (VRF) Tables
  - Overlapping VPNs
  - VPN Route Distribution
  - VPN Packet Forwarding
  - Scaling L3 VPNs and Route Reflectors

Layer 3 (BGP/MPLS) VPN Overview



- Cost effective full mesh connectivity between sites
- Utilize multiple VPNs at a site with different routes to control access
- Facilitates communications in dynamic organization & business application environments
- Leverages existing access options to preserve investment and effectively support a range of applications

What is BGP?

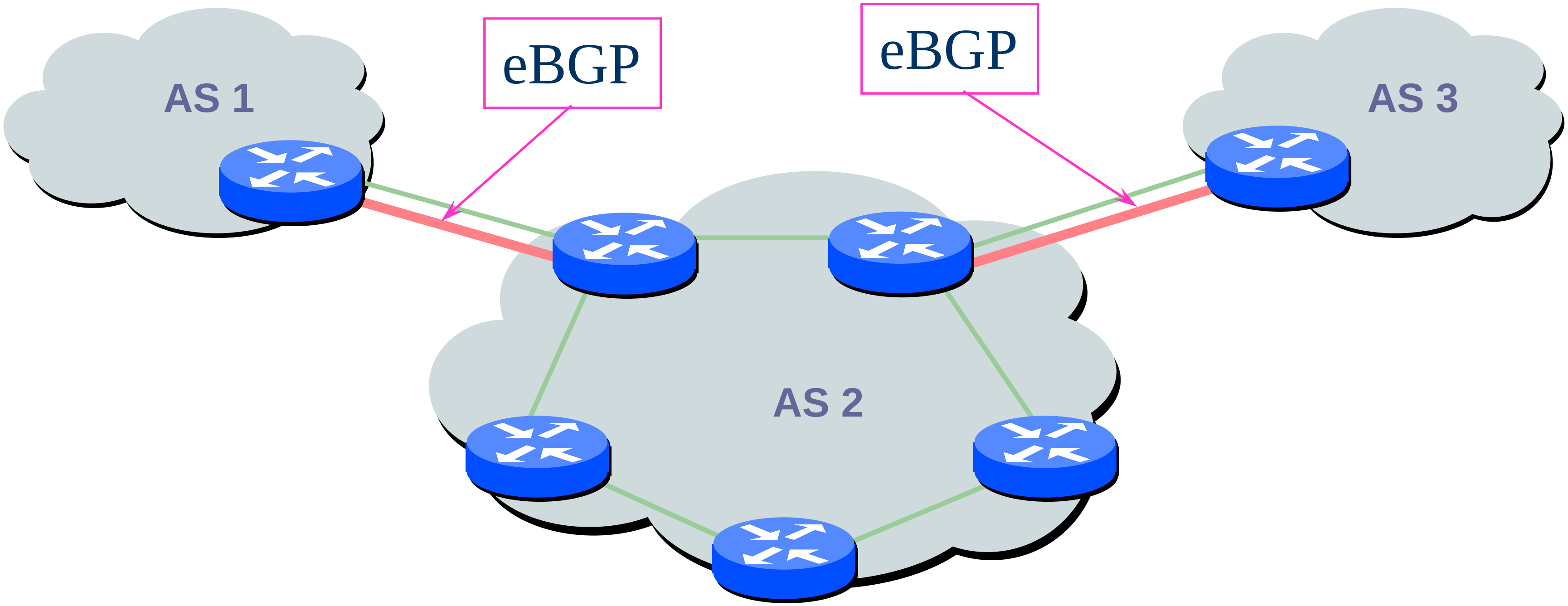
- BGP is an exterior gateway protocol that allows IP routers to exchange network reachability information
- BGP published as RFC 1105 in 1989, then after several updates as BGP-4 in 1995 with RFC 1771, and now as RFC 4271 (2006)
- Numerous other RFCs and Internet Drafts focus on various aspects and extensions including multi-protocol extensions, extended communities, carrying label information in BGP, etc

IGP vs. EGP

- Interior Gateway Protocols
  - RIP, OSPF, IS-IS
  - Dynamic, some more than others
  - Define the routing needed to pass data *within* a network
- Exterior Gateway Protocol
  - BGP
  - Less Dynamic than IGPs
  - Defines the routing needed to pass data *between* networks

External Border Gateway Protocol

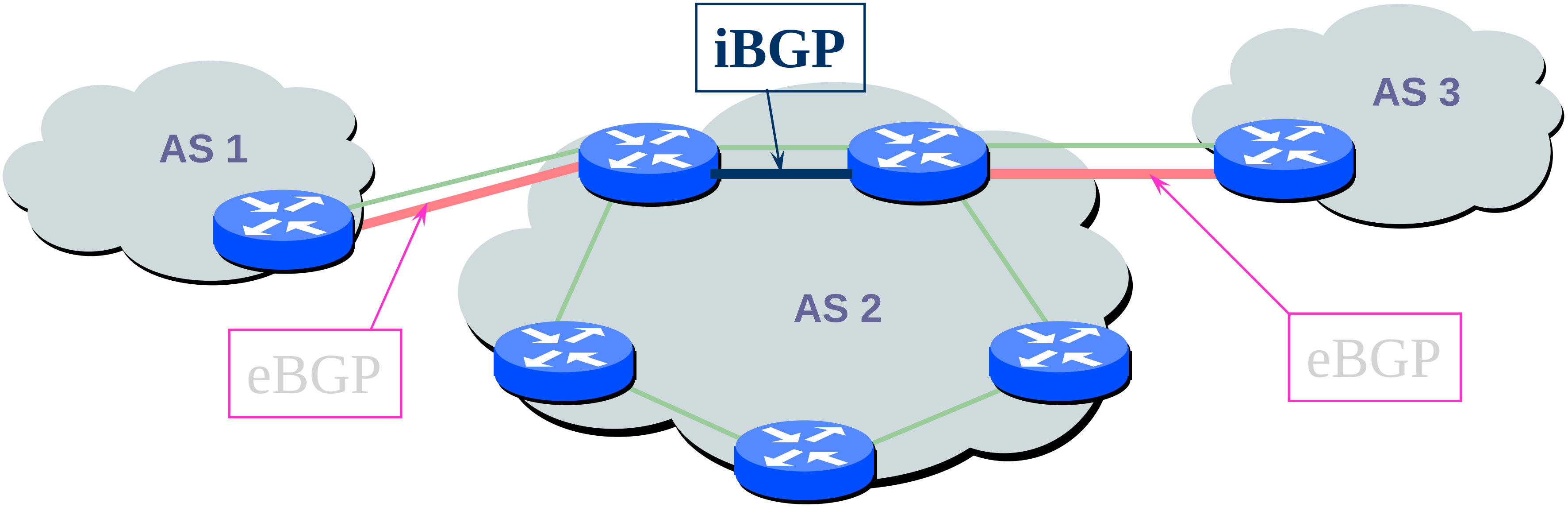
eBGP - **BGP between border routers in two different AS's.**



AS: Autonomous System
eBGP: External BGP

Internal Border Gateway Protocol

iBGP - **BGP between border routers in the same AS.**

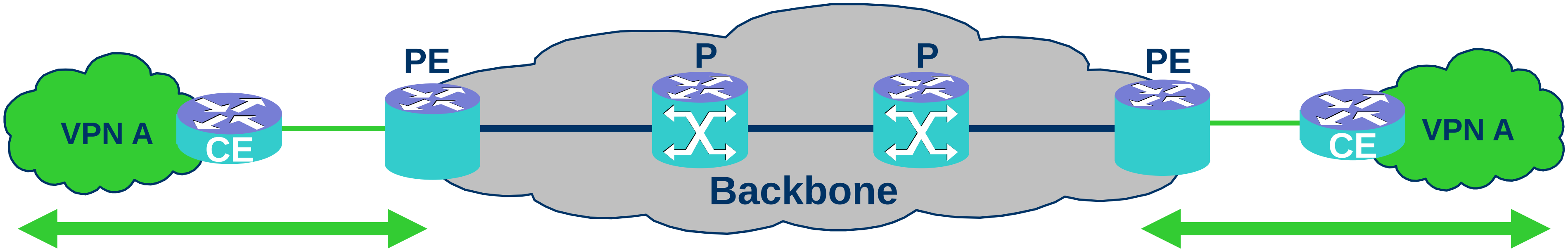


Provides a consistent view within the AS
of the routes exterior to the AS.

AS: Autonomous System
eBGP: External BGP
iBGP: Internal BGP

BGP/MPLS IP VPN (RFC 4364)

Key Characteristics



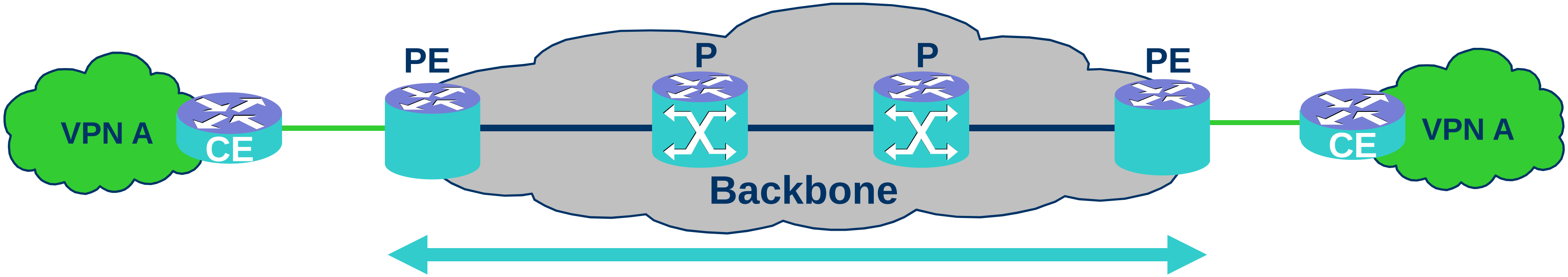
- **Requirements:**
  - **Support for overlapping, private IP address space**
  - **Different customers run different IGPs (i.e. RIP, OSPF, IS-IS)**
- **Solution:**
  - **VPN network layer is terminated at the edge (PE)**
    - **PE routers use plain IP with CE routers**

CE: Customer Edge router
PE: Provider Edge router
P: Provider router not directly attached to a CE

RFC 4364 obsoletes RFC 2547 and is updated by RFC 4577 & RFC 4684

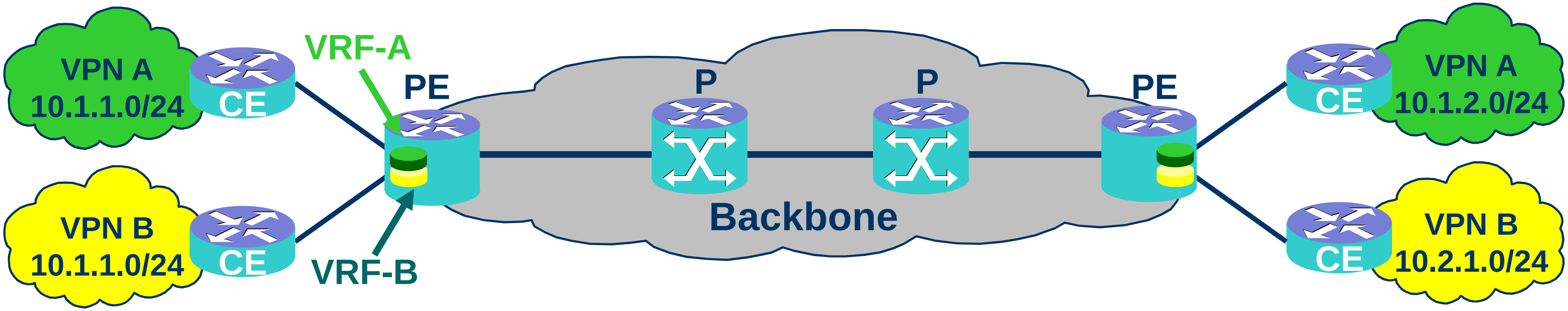
BGP/MPLS IP VPN

Key Characteristics



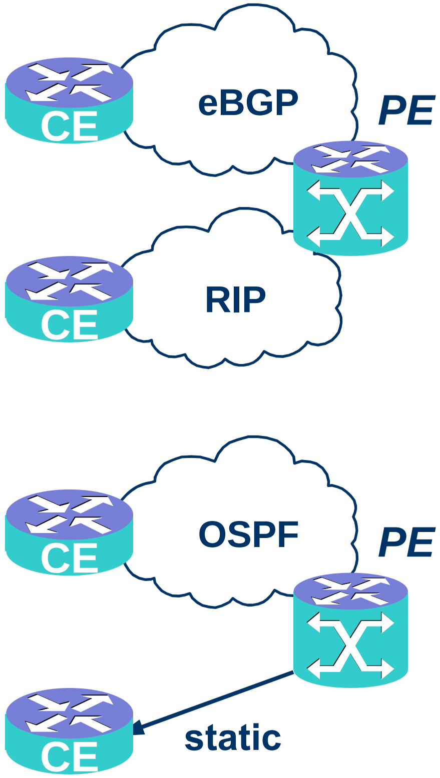
- P routers (LSRs) are in the core of the MPLS cloud
- P and PE (LERs) routers run an IGP and a label distribution protocol
  - Labelled VPN packets are transported over MPLS core
- PE routers are MP-iBGP fully meshed
  - for dissemination of **VPN membership and reachability information** between PEs

Virtual Routing and Forwarding (VRF) Tables



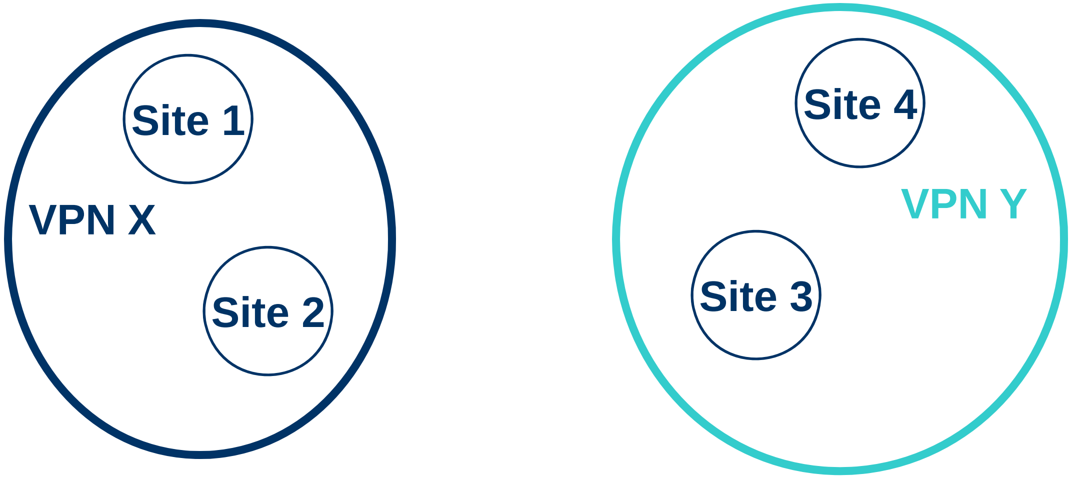
- Each VPN needs a separate **Virtual routing and forwarding instance (VRF)** in each PE router to
  - Provides VPN isolation
  - Allows overlapping, private IP address space by different organizations

Virtual Routing and Forwarding (VRF) *PE to CE Router Connectivity*



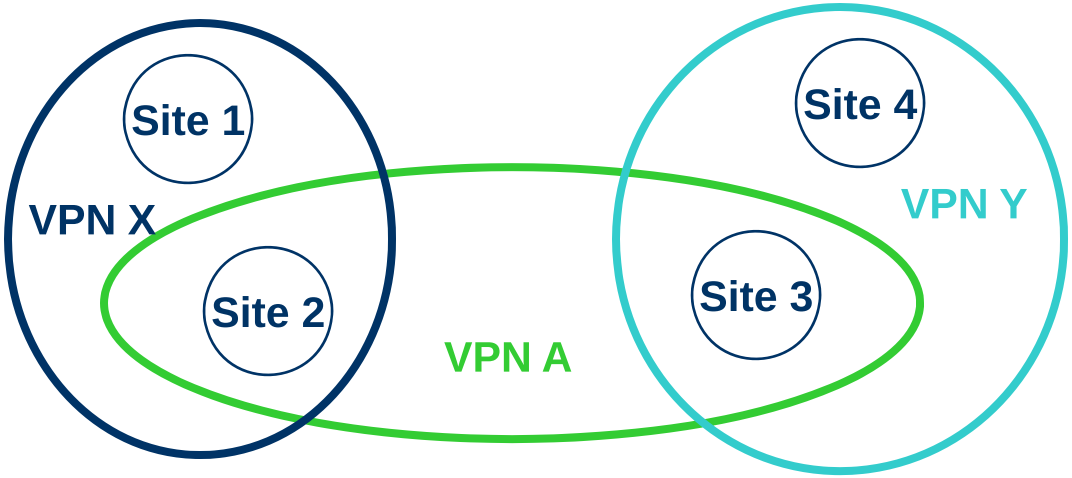
- Protocols used between CE and PE routers to populate VRFs with customer routes
  - BGP-4
    - Useful in stub VPNs and transit VPNs
  - RIPv2
  - OSPF
  - Static routing
    - Particularly useful in stub VPNs
- Note:
  - Customer routes need to be advertised between PE routers
  - Customer routes are not leaked into backbone IGP

Virtual Routing and Forwarding (VRF)



- A VPN is a collection of **sites** sharing a common routing information (routing table)
- A VPN can be viewed as a community of interest (or Closed User Group)

Virtual Routing and Forwarding (VRF) *Overlapping VPNs*



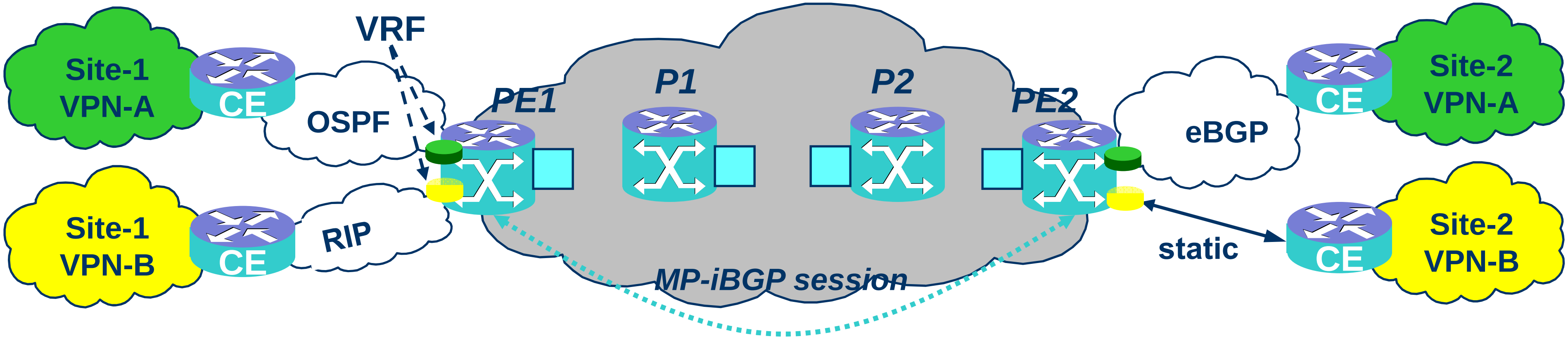
Examples:

- Extranet

- VoIP Gateway

- **A site can be part of different VPNs**
- **A site belonging to different VPNs *may* or *may not* be used as a transit point between VPNs**
- **If two or more VPNs have a common site, address space must be unique among these VPNs**

VRFs and Route Distribution

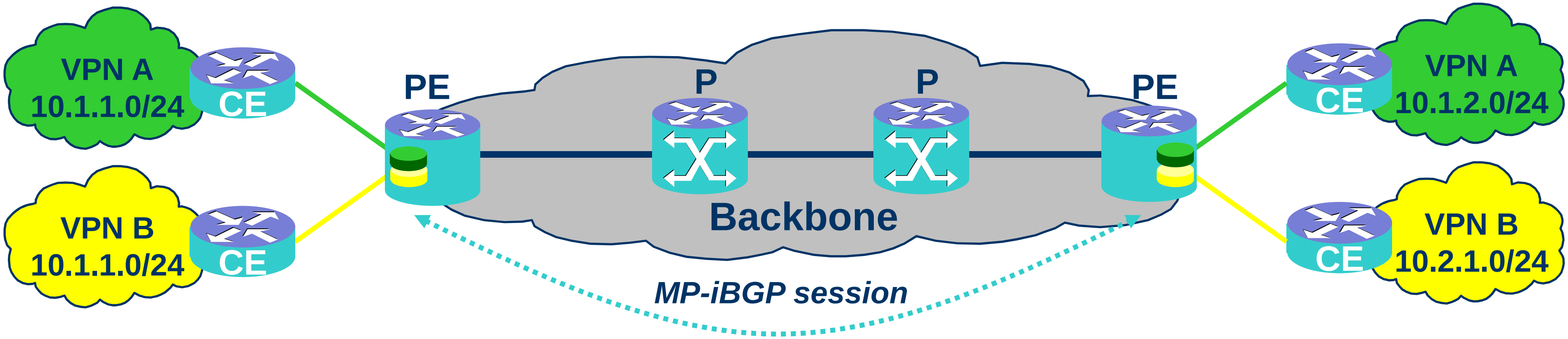


- Multiple VRFs are used on PE routers
- The PE learns customer routes from attached CEs
- Customer routes are distributed to other PEs with MP-BGP
- Different IGPs or eBGP supported between PE and CE peers
- Default forwarding table also exists – public routes

VRF: VPN Routing and Forwarding Table

VPN Route Distribution

Route Targets

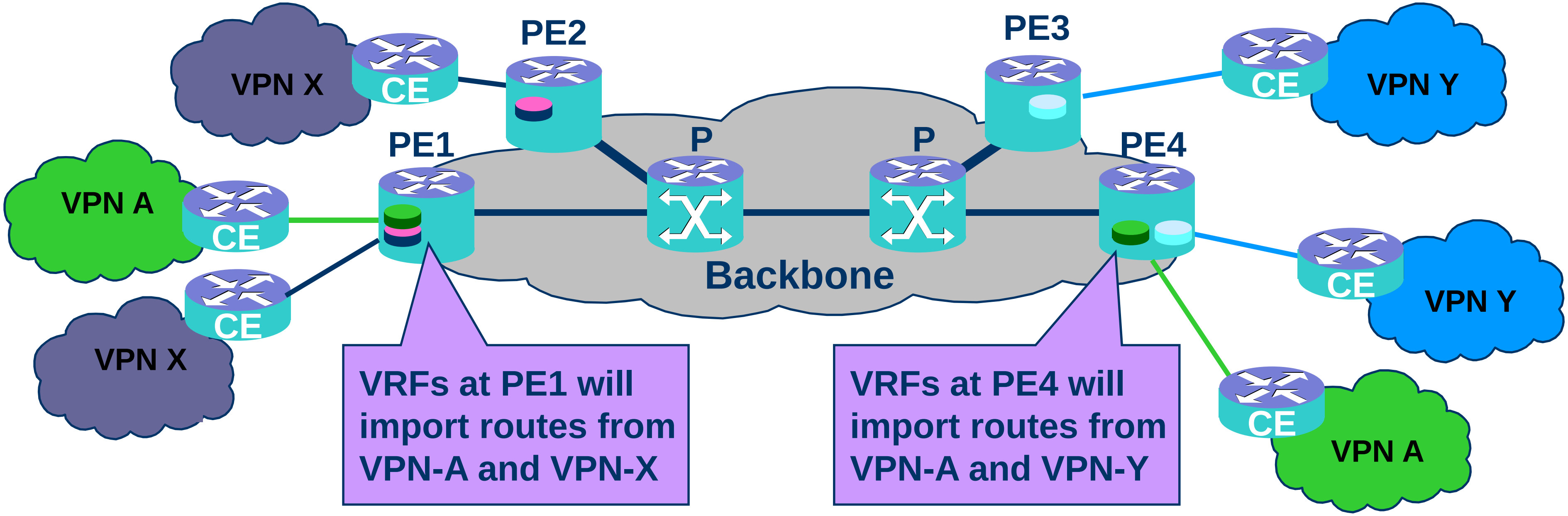


Route Target attributes:

- **"Export" Route Target: Every VPN route is tagged with one or more route targets when it is exported from a VRF (to be offered to other VRFs)**
- **"Import" Route Target: A set of routes targets can be associated with a VRF, and all routes tagged with at least one of those route targets will be inserted into the VRF**

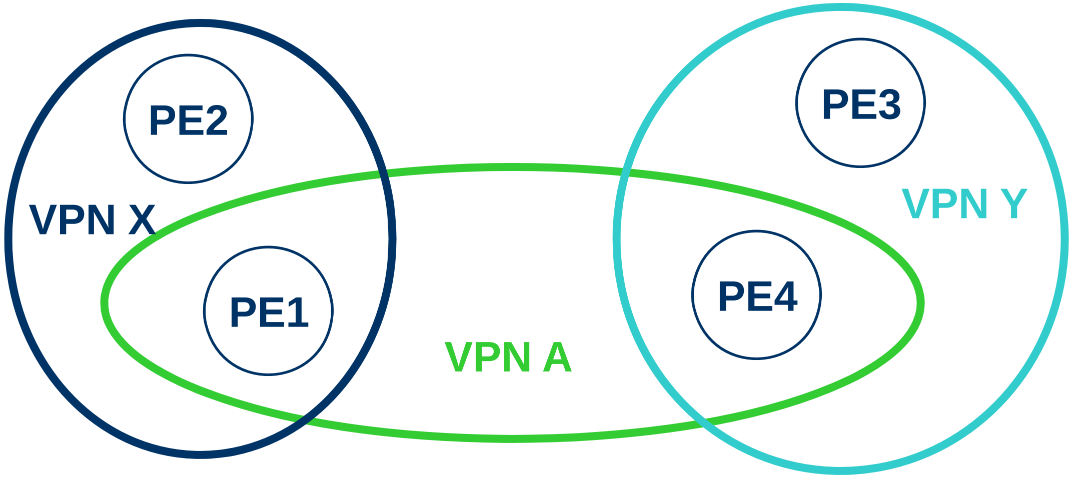
VPN Route Distribution

Route Targets

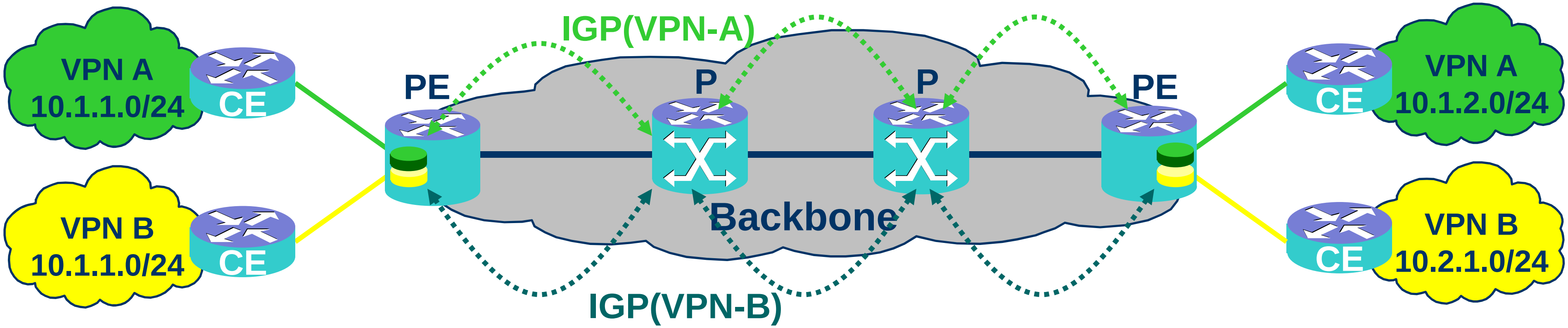


VRFs at PE1 will import routes from VPN-A and VPN-X

VRFs at PE4 will import routes from VPN-A and VPN-Y



VPN Route Distribution

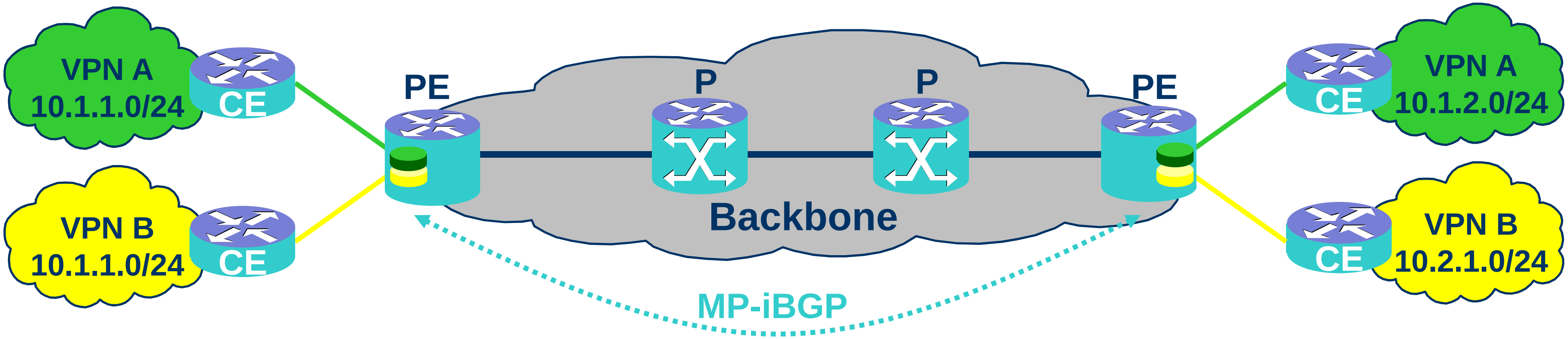


How will the PE routers exchange information about VPN customers and VPN routes between themselves?

Option #1: PE routers run a different routing algorithm for each VPN

- Scalability problems in networks with a large number of VPNs
- Difficult to support overlapping VPNs

VPN Route Distribution



How will the PE routers exchange information about VPN customers and VPN routes between themselves?

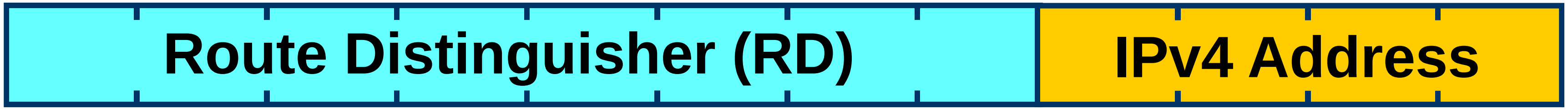
Option #2: BGP/MPLS IP VPN - PE routers run a single routing protocol to exchange all VPN routes

- Problem: <u>Non-unique IP addresses</u> of VPN customers. BGP always propagates one route per destination not allowing address overlap.

VPN Route Distribution

VPN-IPv4 Addresses

- **VPN-IPv4 Address**
  - **VPN-IPv4 is a globally unique, 96bit routing prefix**



64 bits

Creates a VPN-IPv4 address that is globally unique, RD is configured in the PE for each VRF, RD may or may not be related to a site or a VPN

32 bits

IP subnets advertised by the CE routers to the PE routers

VPN Route Distribution

VPN-IPv4 Addresses

Route Distinguisher format

00	00	ASN	nn

- ASN:nn
  - Autonomous System Number (ASN) assigned by Internet Assigned Number Authority (IANA)

00	01	IP address	nn

- IP-address:nn
  - Use only if the MPLS/VPN network uses a private AS number

00	02	BGP-AS4	nn

- BGP-AS4:nn
  - 4-byte Autonomous System Number (BGP-AS4)

nn: assigned number administered by Enterprise

VPN Route Distribution

BGP with Multiprotocol Extensions

- ***How are 96-bit VPN-IPv4 routes exchanged between PE routers?***
- **BGP with Multiprotocol Extensions (MP-BGP) was designed to carry such routing information between peer routers (PE)**
  - **Propagates VPN-IPv4 addresses**
  - **Carries additional BGP route attributes (e.g. route target) called extended communities**

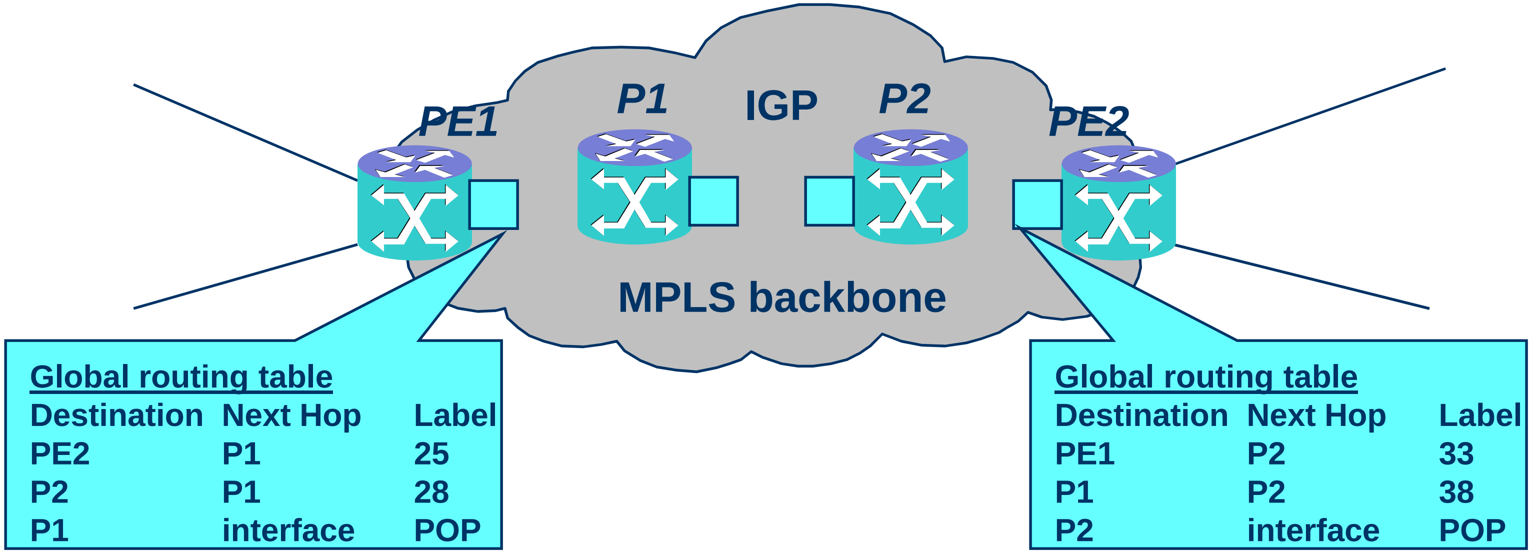
VPN Route Distribution

BGP with Multiprotocol Extensions

- A BGP route is described by:
  - Standard BGP Communities attributes (e.g. Local Preference, MED, Next-hop, AS_PATH, Standard Community, etc.)
  - Extended BGP Communities attributes
- Extended Communities
  - Route Target (RT)
    - Identifies the set of sites the route has to be advertised to
  - Route Origin (RO)/Site of Origin
    - Identifies the originating site
    - Prevents routing loops with multi-homed customer sites

MED: Multi_Exit_Disc

IGP Label Distribution



Global routing table

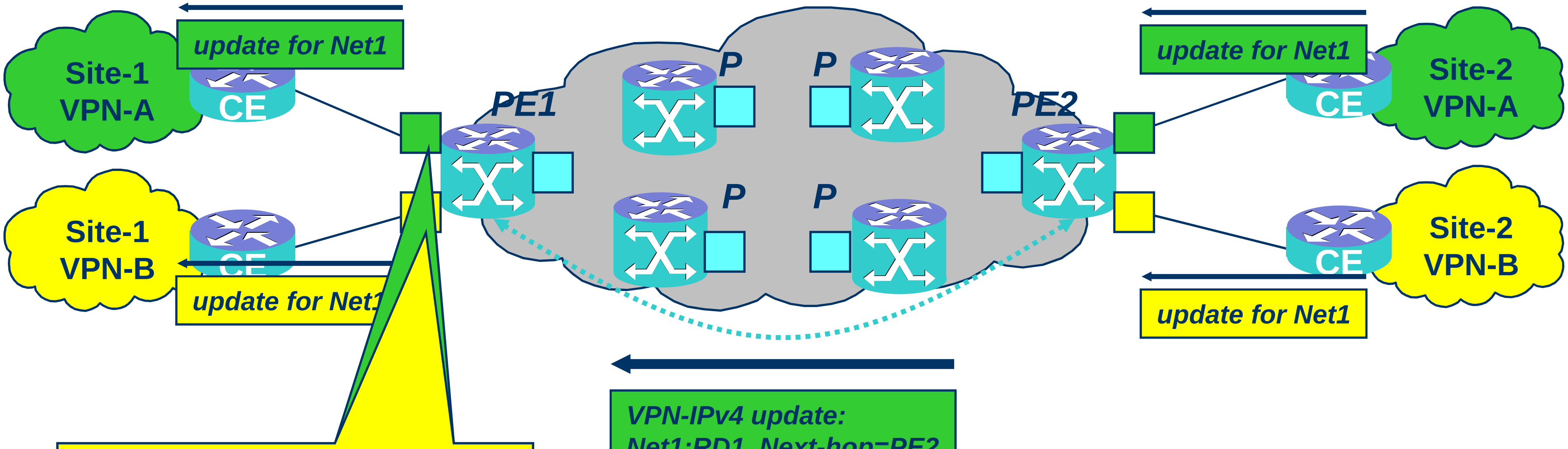
Destination	Next Hop	Label
PE2	P1	25
P2	P1	28
P1	interface	POP

Global routing table

Destination	Next Hop	Label
PE1	P2	33
P1	P2	38
P2	interface	POP

- All routers (P and PE) run an IGP and a label distribution protocol
- Each P and PE router has routes for the backbone nodes and a label is associated to each route
- MPLS forwarding is used within the backbone

MP-BGP Route Distribution



VPN-IPv4 updates are translated into IPv4 address and inserted into the VRF corresponding to the RT value

VPN-IPv4 update:
Net1:RD1, Next-hop=PE2
RO=Site-2, RT=Green
Label=10

VPN-IPv4 update:
Net1:RD2, Next-hop=PE2
RO=Site-2, RT=Yellow
Label=12

“Net1” is the provider’s autonomous system

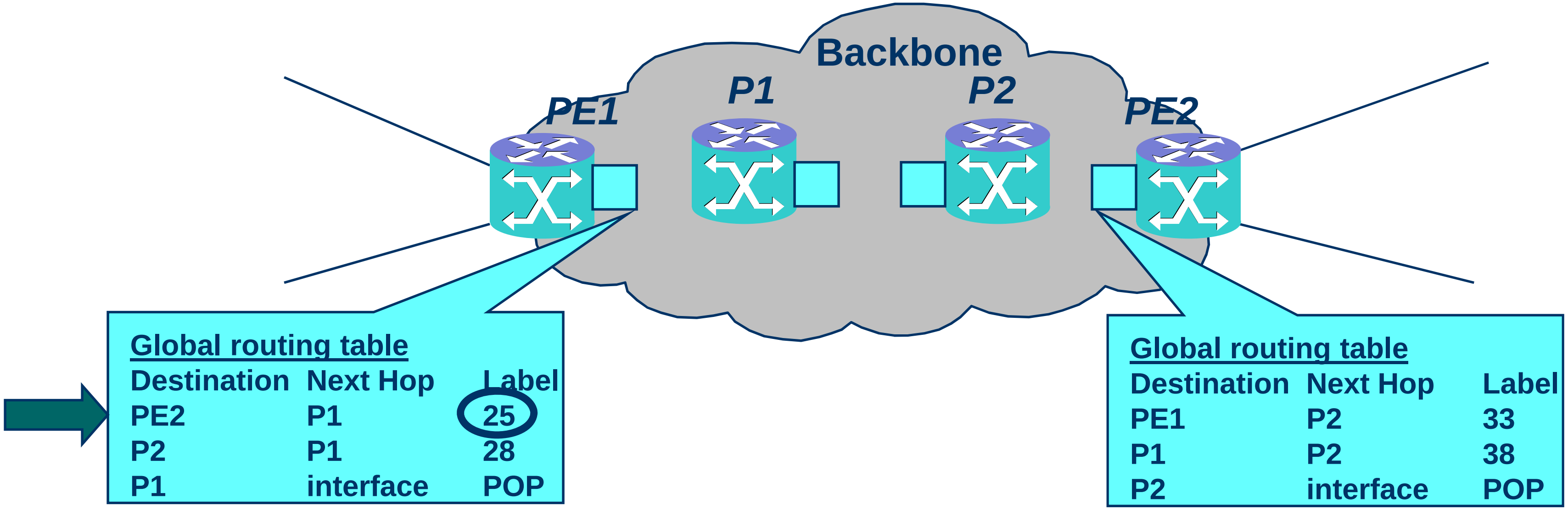
MP-BGP Route Distribution

Summary

- VPN Routing and Forwarding (VRF) Table
  - Multiple routing tables (VRFs) are used on PEs
    - VPNs are isolated
- Customer addresses can overlap
  - Need for unique VPN route prefix
  - PE routers use MP-BGP to distribute VPN routes to each other
  - For security and scalability, MP-BGP only propagates information about a VPN to other routers that have interfaces with the same Route Target value
- BGP-MPLS VPN extensions for IPv6 (RFC 4659)

MP-BGP: BGP with Multiprotocol Extensions

VPN Packet Forwarding



Global routing table

Destination	Next Hop	Label
PE2	P1	25
P2	P1	28
P1	interface	POP

Global routing table

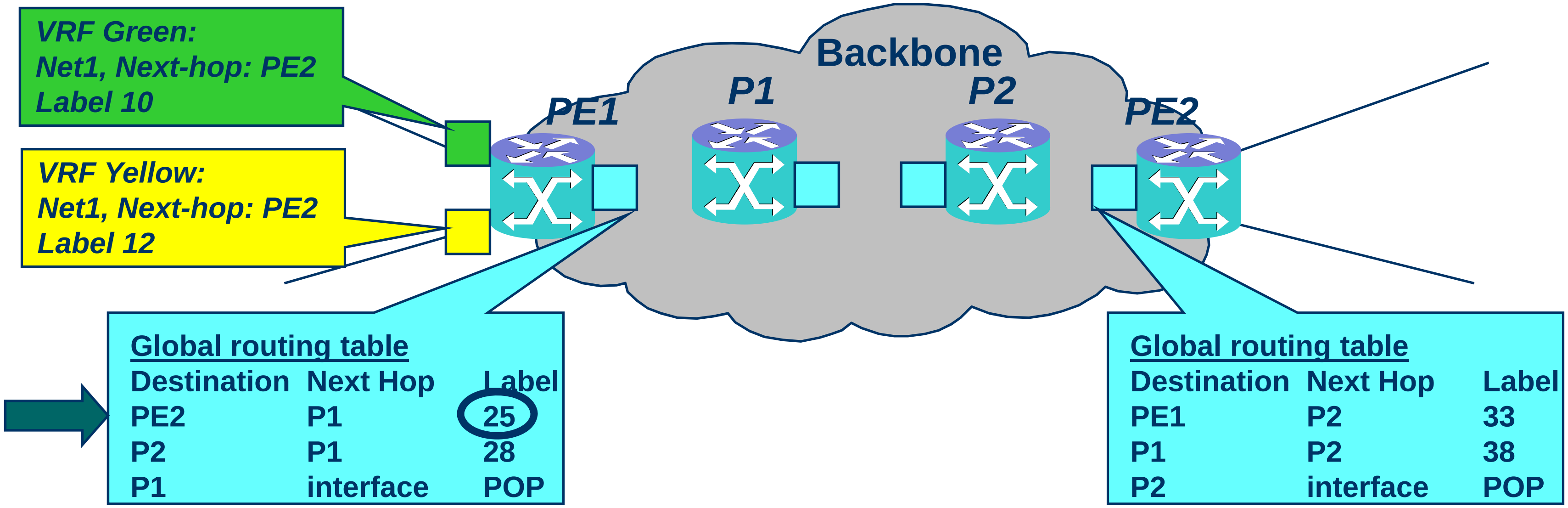
Destination	Next Hop	Label
PE1	P2	33
P1	P2	38
P2	interface	POP

PE-to-PE connectivity via LSPs

- All routers (P and PE) run an IGP and a label distribution protocol
- Each P and PE router has routes for the backbone nodes and a label is associated to each route
- MPLS forwarding is used within the backbone

VPN Packet Forwarding

Label Stacking



VRF Green:
Net1, Next-hop: PE2
Label 10

VRF Yellow:
Net1, Next-hop: PE2
Label 12

Global routing table

Destination	Next Hop	Label
PE2	P1	25
P2	P1	28
P1	interface	POP

Global routing table

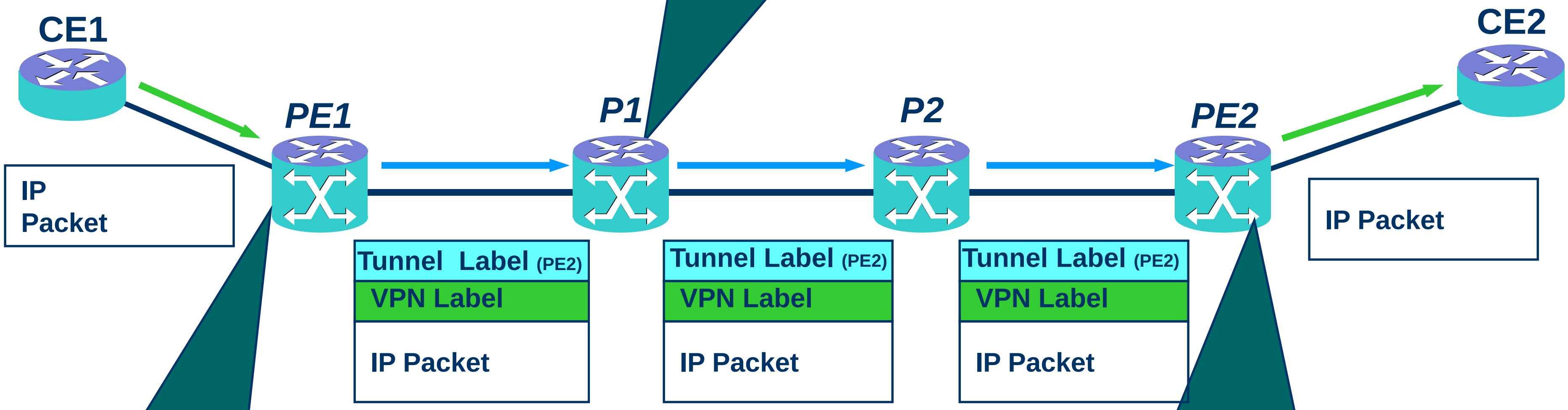
Destination	Next Hop	Label
PE1	P2	33
P1	P2	38
P2	interface	POP

- **Ingress PE router uses <u>two-level label stack</u>**
  - VPN label **(inner label) assigned by the egress PE router**
  - Tunnel (IGP) label **(top label) identifying the PE router**
- **Label stack is attached in front of the IP packet that belongs to a VPN**
- **The MPLS packet is forwarded across the P routers in the backbone network**

VPN Packet Forwarding

Label Stacking

P routers switch the packet based on the Tunnel Label (top label)



PE1 router receives normal IP packet from CE1 router.

PE router does “IP Longest Match” from VRF, finds iBGP next hop PE2 and imposes a stack of labels

Egress PE router removes top label,

uses inner label to select which VPN/CE to forward the packet to.

Inner label is removed and packet sent to CE2 router

VPN Packet Forwarding

Penultimate Hop Popping

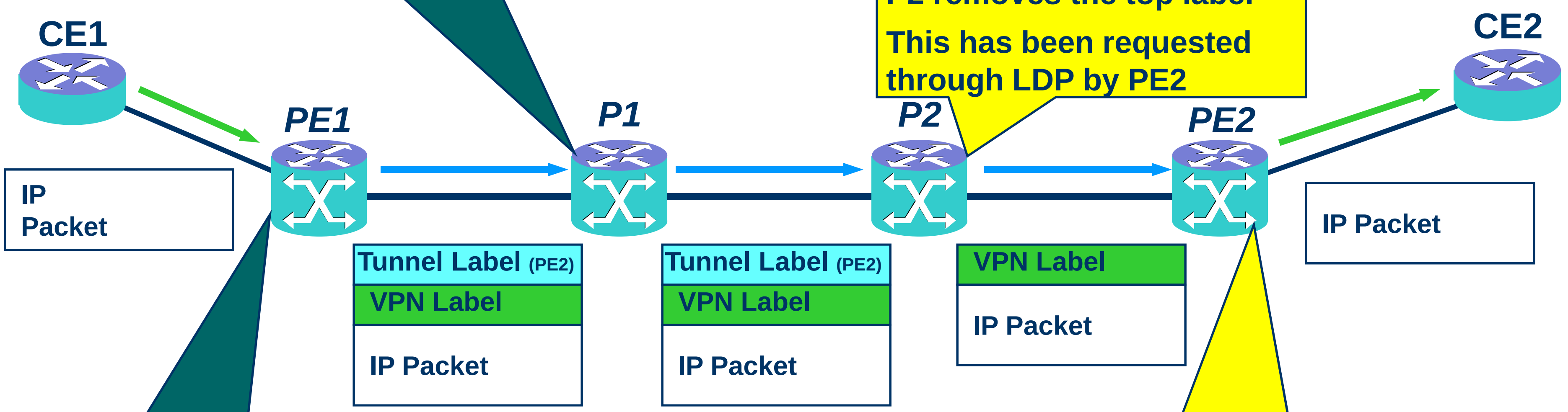
P routers switch the packet based on the Tunnel Label (top label)

Penultimate Hop Popping

P2 is the penultimate hop for the BGP next-hop

P2 removes the top label

This has been requested through LDP by PE2



PE1 router receives normal IP packet from CE1 router.

PE router does “IP Longest Match” from VRF, finds iBGP next hop PE2 and imposes a stack of labels

PE2 receives packet with the label corresponding to the outgoing VRF

One single lookup

Label is popped and packet sent to CE2 router

Core Routers (P Routers)

- Not involved in MP-BGP
- Does not make routing decision based on VPN addresses
- Forwards packet based on the top label value

⬇

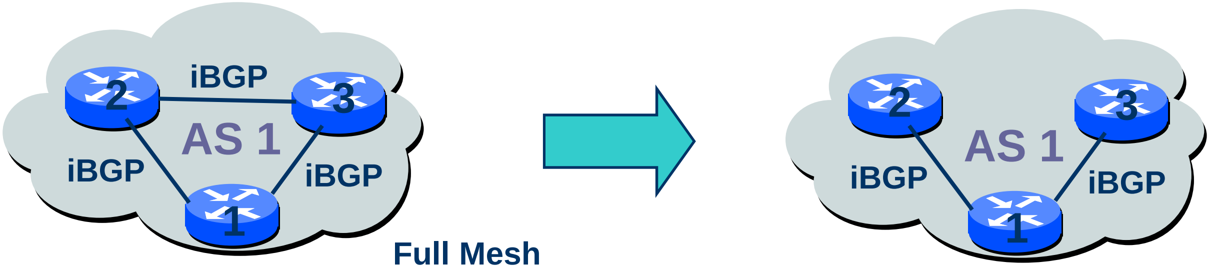
- P routers do not need to carry VPN routing information or Internet routing information, thus providing better network scalability

Scaling BGP/MPLS VPNs

- Scalability of BGP/MPLS VPNs
  - Expanding the MPLS core network
    - Without impact on the VPN services, e.g. adding P routers (LSRs), new or faster links
  - Label stacking
    - Allows reducing the number of LSPs in the network core and avoiding LSP exhaustion
  - VPN Route Distribution
    - Route Reflectors

Scaling BGP/MPLS VPNs

Route Reflectors



Full Mesh
iBGP

n*(n-1)/2

Route
Reflector

Use redundant
Route Reflectors to
eliminate single
point of failure

BGP Route Reflectors

- Existing BGP technique, can be used to scale VPN route distribution
  - PEs don’t need full mesh of BGP connections, only connect to RRs
  - By using multiple RRs, no one box needs to have all VPN routes
- Each edge router needs only the information for the VPNs it supports
  - Directly connected VPNs

RR: Route Reflector

Section 3

Layer 2 VPNs

MPLS VPN Tutorial Agenda

Layer 2 VPNs

- Overview
- Encapsulation and Label Stacking
- Virtual Private Wire Services – VPWS
    - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
- Virtual Private LAN Services – VPLS

MPLS L2 VPN Market Drivers

What can we conclude?

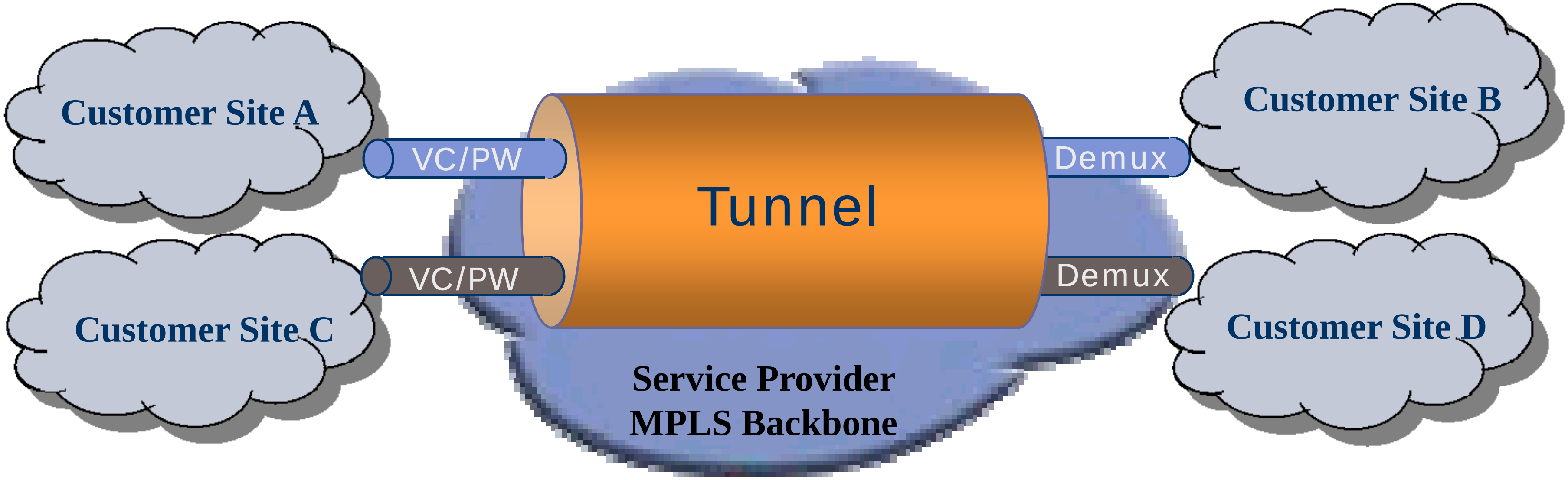
- Layer 3 IP is not the only traffic
  - Still a lot of legacy SNA, IPX, etc
  - Large enterprises have legacy protocols
- Layer 3 IP VPNs are not the whole answer
  - IP VPNs cannot handle legacy traffic
- Layer 2 legacy traffic widely deployed

Need for Layer 2 and Layer 3 VPNs to support the broad range of applications

MPLS Layer 2 VPNs

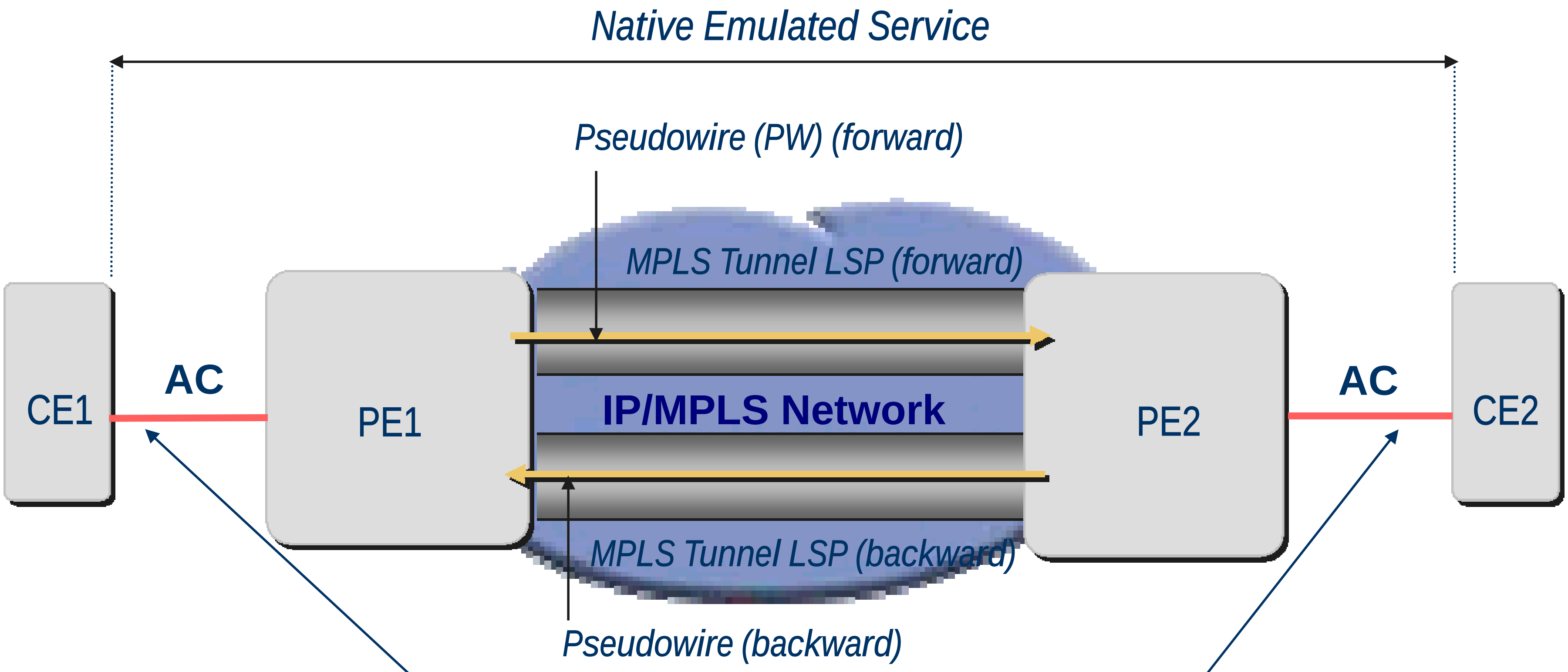
- **Point-to-point Layer 2 solutions**
  - Virtual Private Wire Services **- VPWS**
  - **Similar to ATM / FR services, uses tunnels and connections (LSPs)**
  - **Customer gets connectivity only from provider**
  - **Ongoing work to encapsulate Ethernet, ATM, FR, TDM, SONET, etc**
- **Multi-point Layer 2 solutions**
  - Virtual Private LAN Services **- VPLS**
  - **Virtual Private LAN Services aka Transparent LAN Service (TLS)**
  - **Ethernet Metro VLANs / TLS over MPLS**
  - **Independent of underlying core transport**
  - **Ethernet encapsulation for transport over MPLS (RFC 4448)**
  - **Two approaches to signaling (RFC 4761 & RFC 4762)**

Virtual Private Wire Service (VPWS)



- Point-to-Point Service
- Tunnel Label determines path through network
- VC/PW Label identifies VLAN, VPN, or connection at the end point

MPLS Pseudowire
Reference Model



Native Emulated Service
Pseudowire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
AC
PE1
IP/MPLS Network
PE2
AC
CE2
MPLS Tunnel LSP (backward)
Pseudowire (backward)

ATM, Ethernet , FR, IP, TDM, etc
Attachment Circuit (AC)
- Same at each end

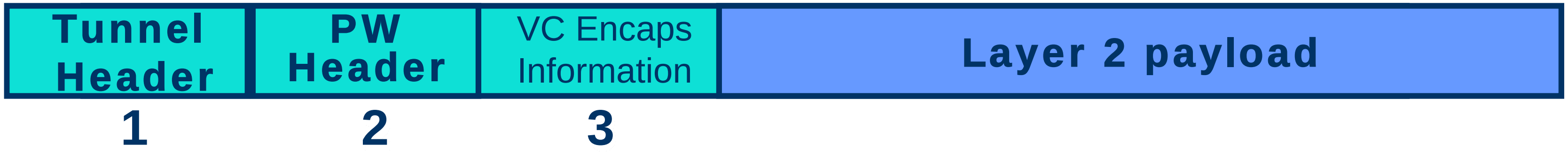
AC: Attachment Circuit
CE: Customer Edge
PE: Provider Edge

Pseudowire Emulation Edge-to-Edge (PWE3)

- Requirements for PWE3 (RFC 3916):
  - Base requirements for Pseudowire Emulation Edge-to-Edge (PWE3) WG
- PWE3 Architecture (RFC 3985):
  - Describes architecture for Pseudowire Emulation Edge-to-Edge Emulation of services (such as Frame Relay, ATM, Ethernet TDM and SONET/SDH) over packet switched networks (PSNs) using IP or MPLS
  - Architectural framework for pseudowires (PWs), defines terminology, specifies the various protocol elements and functions
- Pseudowire Set-up and Maintenance using LDP (RFC 4447)

MPLS Point-to-Point Services

Label Stacking



- Three Layers of Encapsulation
  1) Tunnel Header: Contains information needed to transport the PDU across the IP or MPLS network
  2) Pseudo wire Header (PW): Used to distinguish individual emulated VCs within a single tunnel
  3) Emulated VC Encapsulation: Contains the information about the enclosed PDU (known as Control Word)
- Tunnel Header determines path through network
- Pseudo wire Header identifies VLAN, VPN, or connection at the end point
- All services look like a Virtual Circuit to MPLS network

PDU: Protocol Data Unit

Encaps Information Field

bits	0-3	4-7	8-9	10-15	16-31
	Reserved	Flags	FRG	Length	Sequence Number

FRG: Fragmentation

Generic Control Word

- **Layer 2 header fields may be discarded at ingress**
- **Control word carries “flag” bits depending on encapsulation**
  - **(FR: FECN, BECN, C/R, DE, ATM: CLP, EFCI, C/R, etc)**
- **Length required when padding small frames on links which have a minimum frame size**
- **Sequence number is optional. It is used to detect out of order delivery of frames.**

bits	0-3	4-7	8-15	16-31
	Reserved	Version	Reserved	Channel Type

Control Word for PW Associated Channel

RFC 4385

LDP - Label Mapping Message

Label Mapping	Message Length
Message ID	
FEC TLV	
Label TLV	
Label Request Message ID TLV	
LSPID TLV (optional)	
Traffic TLV (optional)	

FEC: Forwarding Equivalence Class

TLV: Type-Length-Value

New VC FEC Element Defined

VC TLV	C	VC Type	VC Info Length
Group ID			
VC ID			
Interface Parameters			

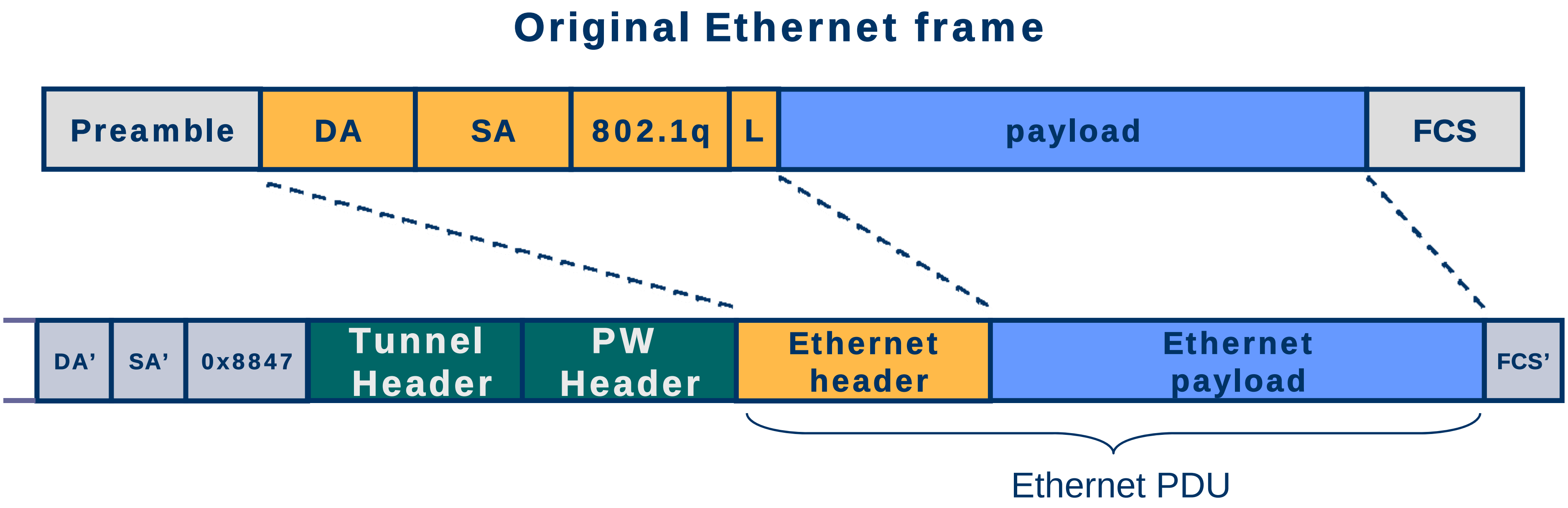
- **Virtual Circuit FEC Element**
  - **C - Control Word present**
  - **VC Type - FR, ATM, Ethernet, HDLC, PPP, ATM cell**
  - **VC Info Length - length of VCID field**
  - **Group ID - user configured - group of VCs representing port or tunnel index**
  - **VC ID - used with VC type to identify unique VC**
  - **Interface Parameters - Specific I/O parameters**

Layer 2 Encapsulation *PWE3*

PWE3 Work

- Ethernet / 802.1q VLAN
  - RFC 4448
- ATM AAL5 and ATM cell
  - RFC 4717
- Frame Relay
  - RFC 4619
- PPP/HDLC
  - RFC 4618
- TDM
  - RFC 4553
- Pseudowire Set-up and Maintenance using LDP
  - RFC 4447

Ethernet Encapsulation for Transport over MPLS



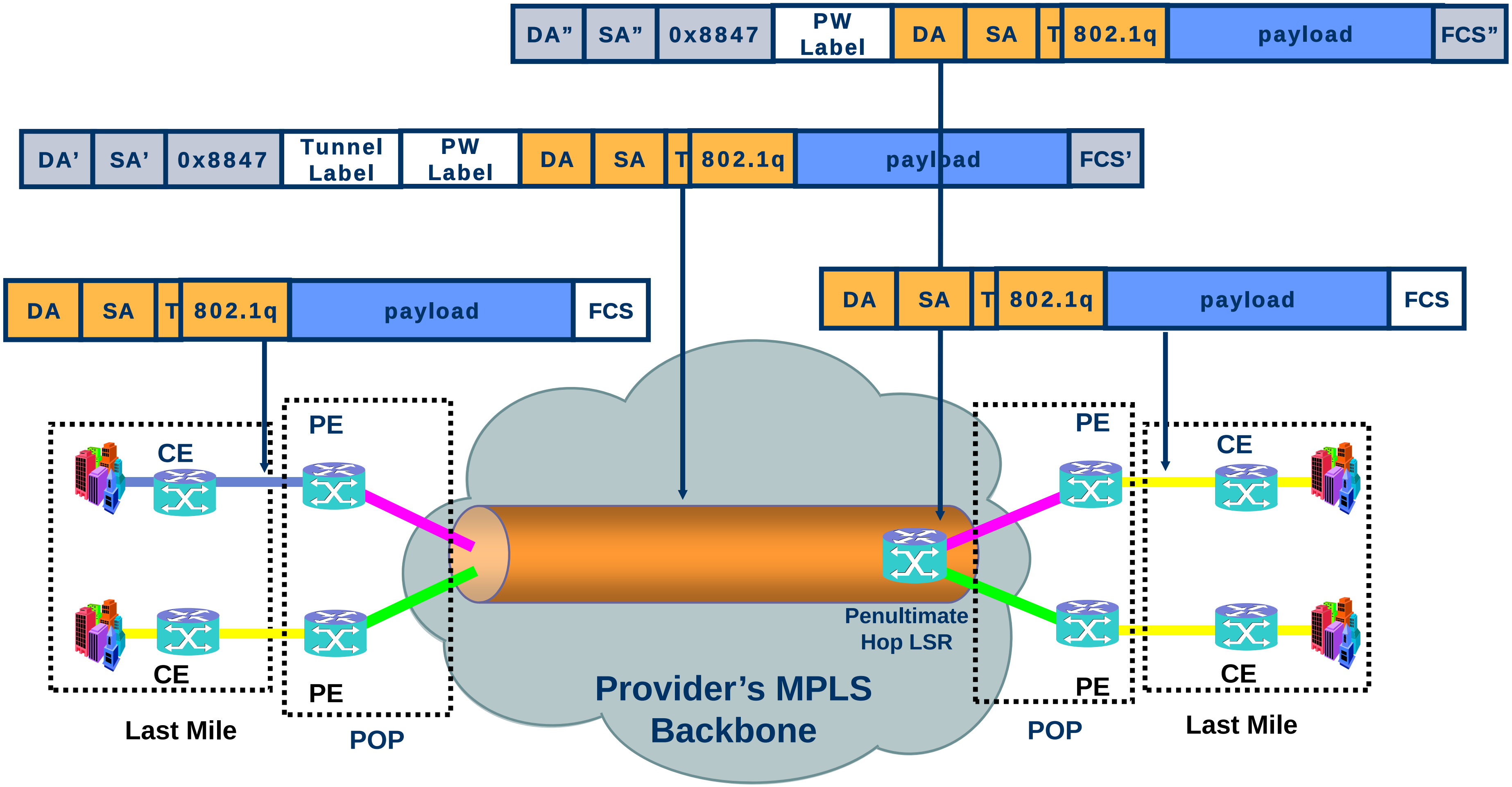
Encapsulated Ethernet over MPLS <u>over Ethernet Transport</u>

- Ingress device strips the Ethernet preamble and FCS
- Raw or Tagged mode
- Optional Control Word 0000 | Reserved | Sequence #
- New MPLS Ethernet header (type 0x8847) and new FCS is added to MPLS Ethernet packet

RFC 4448

Life of a Frame

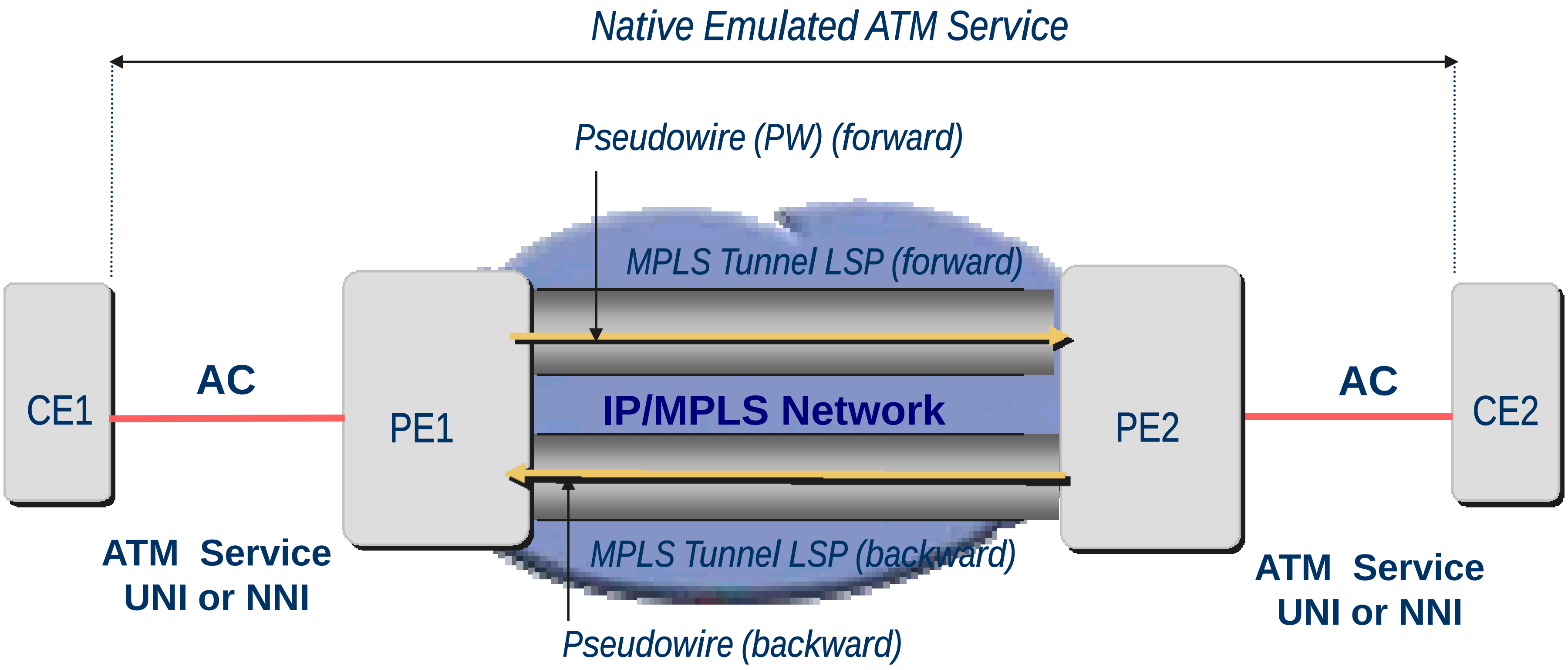
Ethernet over Ethernet MPLS



DA” | SA” | 0x8847 | PW Label | DA | SA | T | 802.1q | payload | FCS”
DA’ | SA’ | 0x8847 | Tunnel Label | PW Label | DA | SA | T | 802.1q | payload | FCS’
DA | SA | T | 802.1q | payload | FCS
DA | SA | T | 802.1q | payload | FCS
PE
CE
CE
PE
Last Mile
POP
Penultimate Hop LSR
Provider’s MPLS Backbone
PE
CE
PE
CE
POP
Last Mile

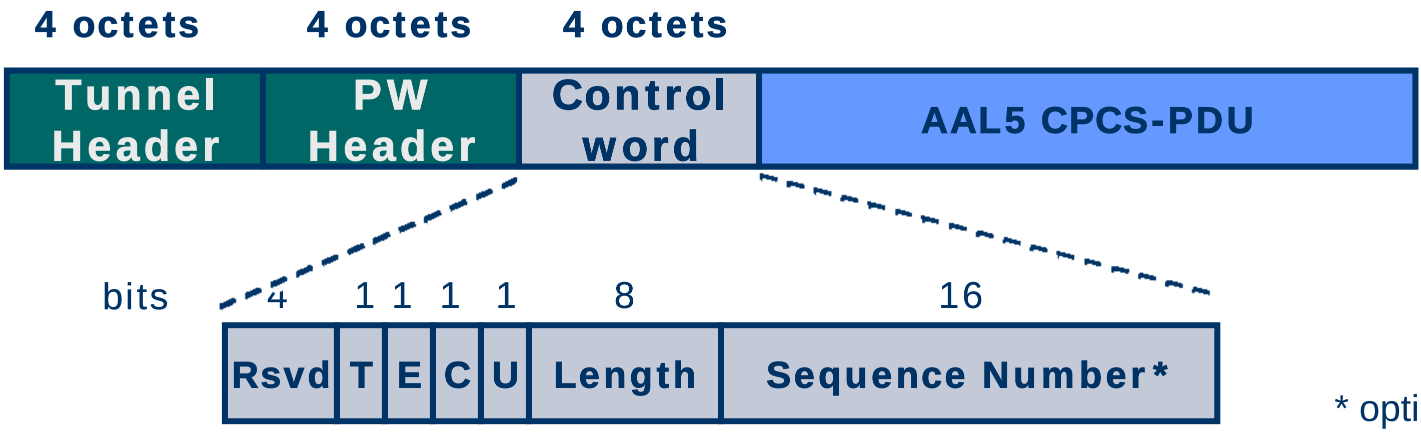
ATM Service Transport with a PW

Reference Model


Native Emulated ATM Service
Pseudowire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
AC
PE1
IP/MPLS Network
PE2
AC
CE2
MPLS Tunnel LSP (backward)
ATM Service UNI or NNI
ATM Service UNI or NNI
Pseudowire (backward)

AC: Attachment Circuit
CE: Customer Edge
PE: Provider Edge

ATM AAL5 Encapsulation for Transport over MPLS



4 octets | 4 octets | 4 octets

Tunnel Header | PW Header | Control word | AAL5 CPCS-PDU

bits: 4, 1, 1, 1, 1, 8, 16

Rsvd | T | E | C | U | Length | Sequence Number*

* optional

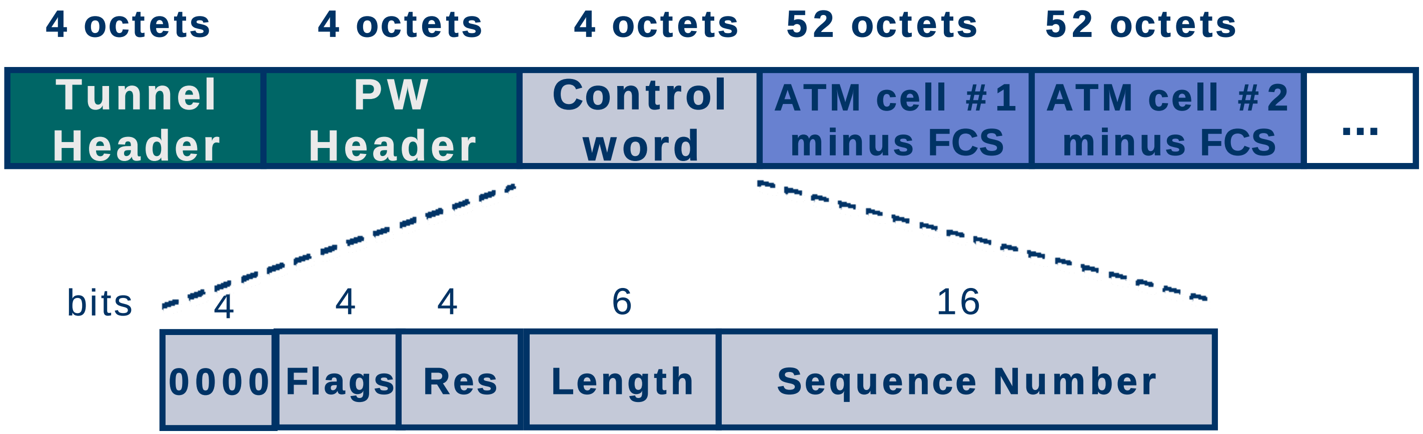
ATM Control Word

- **2 modes:**
  - **PDU Frame Mode – encapsulates PDU payload, pad and trailer**
  - **SDU Frame Mode – encapsulates PDU payload** *(shown above)*
- **Ingress reassembles AAL5 frames**
- **SDU Frame mode required control word includes:**
  - **T = Transport type bit identifies whether packet contains an AAL5 payload or ATM admin cell**
  - **E = EFCI bit - Explicit Forward Congestion Indication**
  - **C = CLP bit - Cell Loss Priority**
  - **U = Command / Response bit**

PDU: Protocol Data Unit

RFC 4717

ATM Cell Mode Encapsulation for Transport over MPLS



Control Word

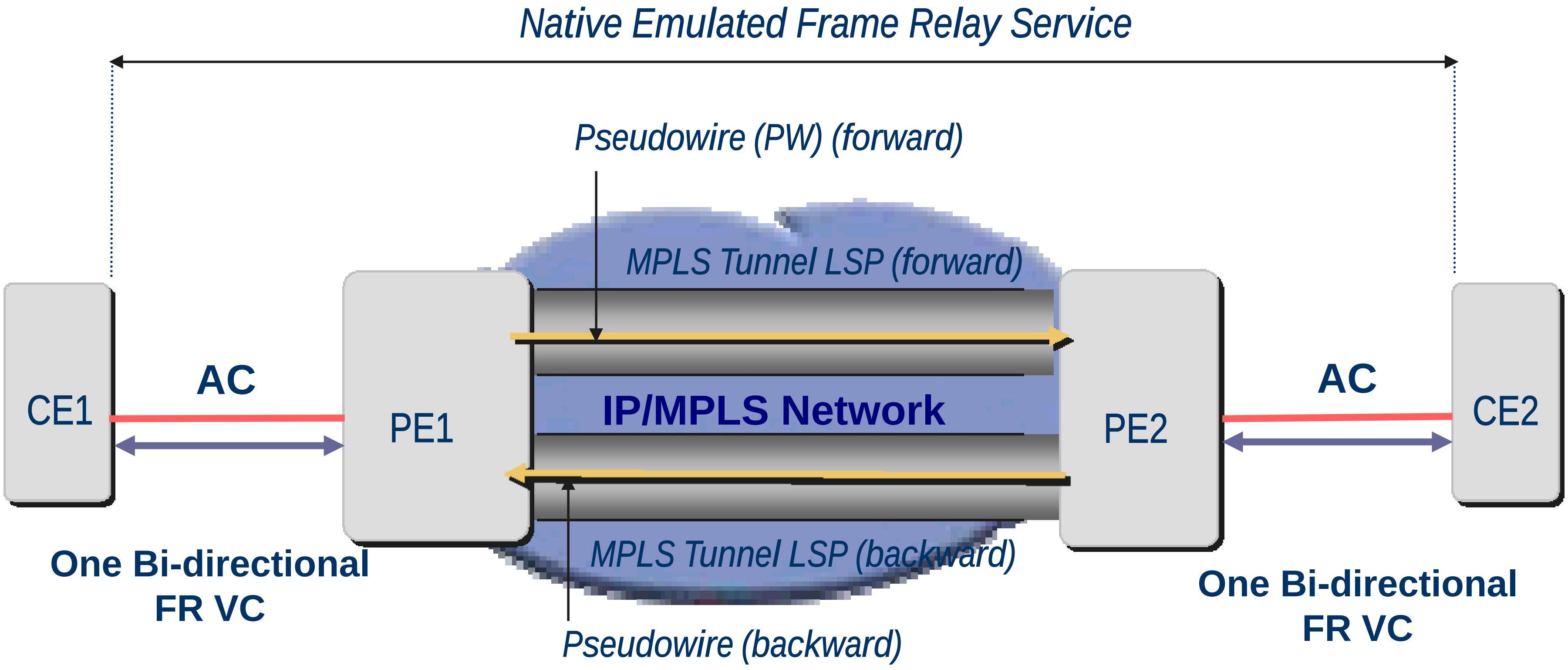
N-to-One Cell Mode Multiple Cell Encapsulation

Control Word (optional)			
VPI	VCI	PTI	C
ATM Payload (48 bytes) “ “			
VPI	VCI	PTI	C
ATM Payload (48 bytes) “ “			

- **2 modes:**
  - **One-to-One Cell Mode - maps one ATM VCC (or VPC) to one PW**
  - **N-to-One Cell Mode - maps one or more ATM VCCs (or VPCs) to one PW** *(shown above);* **only required mode for ATM support**
- **Ingress performs no reassembly**
- **Control word is <u>optional</u>: If used, Flag and Length bits are not used**

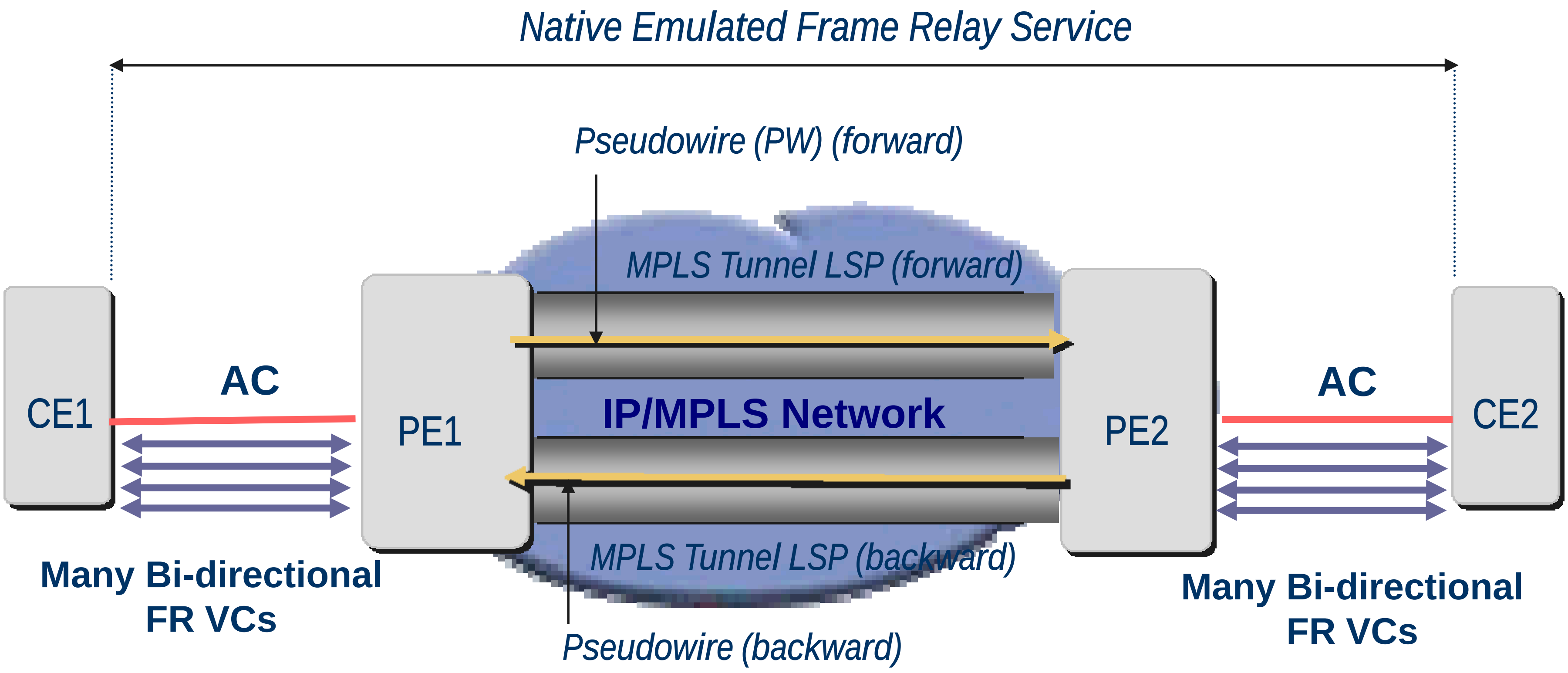
RFC 4717

Frame Relay Encapsulation for Transport over MPLS


Native Emulated Frame Relay Service
Pseudowire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
AC
PE1
IP/MPLS Network
PE2
AC
CE2
MPLS Tunnel LSP (backward)
One Bi-directional FR VC
Pseudowire (backward)
One Bi-directional FR VC

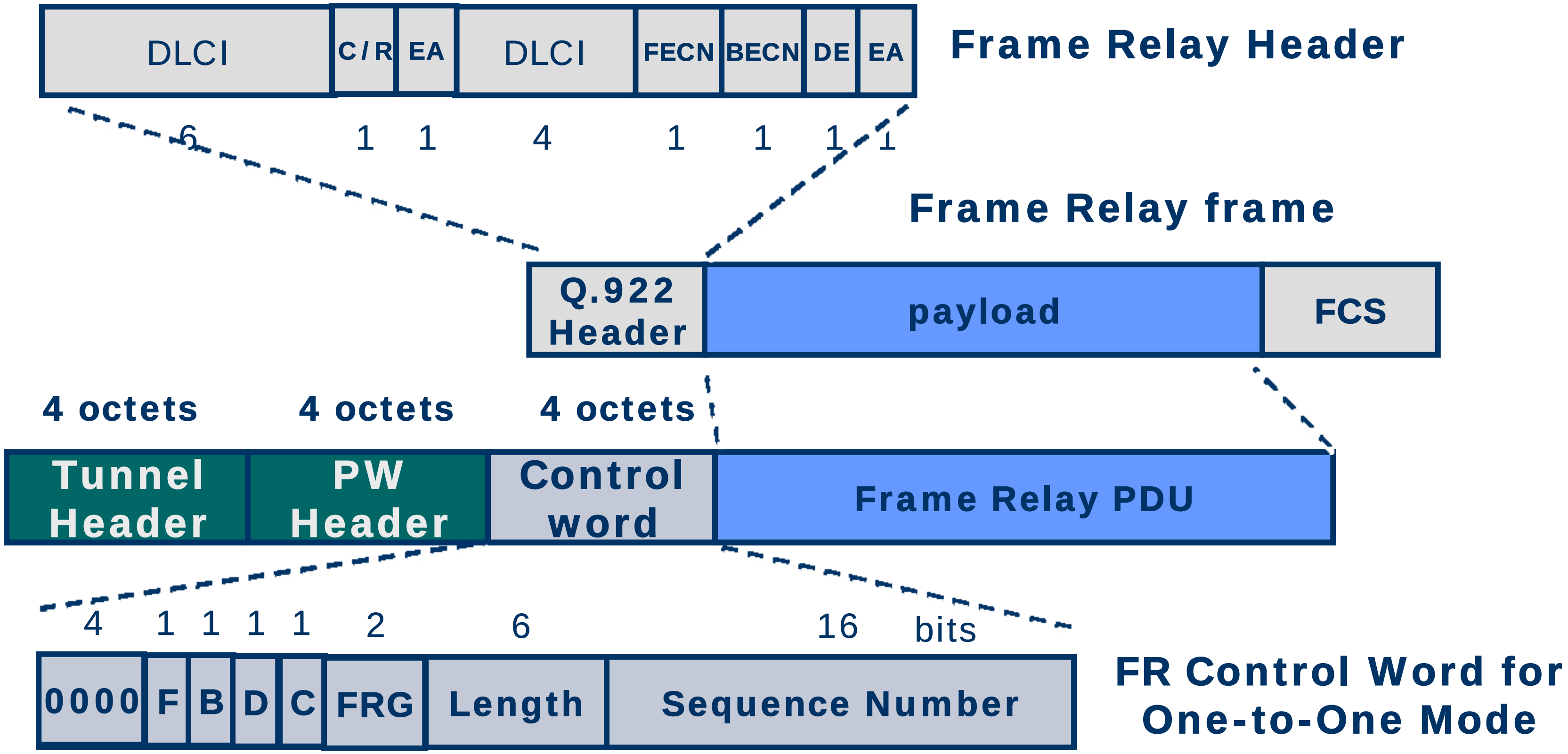
- **Frame Relay (FR) Transport Service application**
- **Two Mapping modes:**
  - **One-to-one mapping: One FR VC mapped to a pair of unidirectional PWs** *(shown above)*
  - **Many-to-one or port mode mapping: Many FR VCs mapped to a pair of Unidirectional PWs**

Frame Relay Encapsulation for Transport over MPLS


Native Emulated Frame Relay Service
Pseudowire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
AC
PE1
IP/MPLS Network
PE2
AC
CE2
MPLS Tunnel LSP (backward)
Many Bi-directional FR VCs
Pseudowire (backward)
Many Bi-directional FR VCs

- **Two Mapping modes:**
  - **One-to-One Mapping: One FR VC mapped to a pair of unidirectional PWs**
  - **Many-to-One or Port Mode Mapping: Many FR VCs mapped to a pair of Unidirectional PWs** *(shown above)*

Frame Relay Encapsulation for Transport over MPLS



- F = FECN (Forward Explicit Congestion Notification)
- B = BECN (Backward Explicit Congestion Notification)
- D = DE (Discard Eligibility Indicator)
- C = C/R (Command / Response Field)

RFC 4619

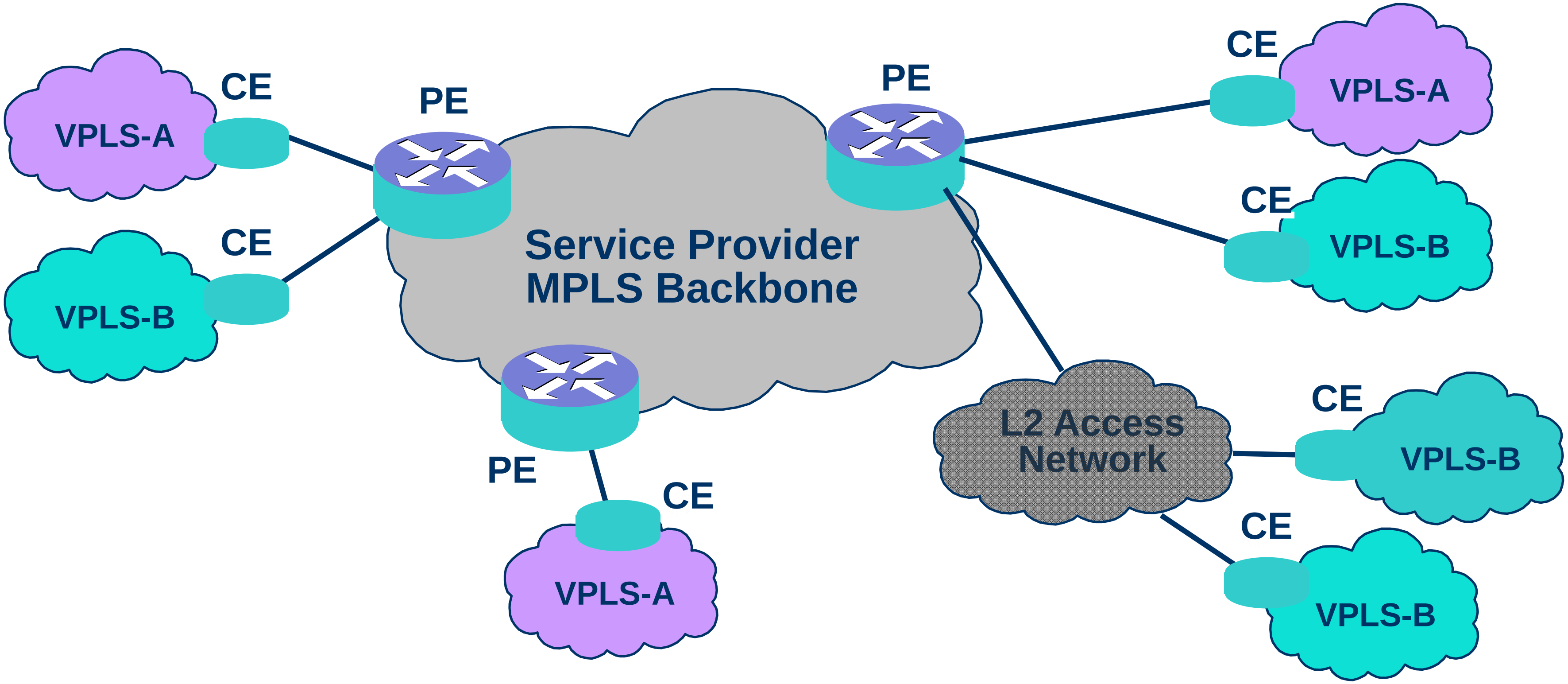
MPLS VPN Tutorial Agenda

Layer 2 VPNs

- Overview
- Encapsulation and Label Stacking
- Virtual Private Wire Services – VPWS
  - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
- Virtual Private LAN Services – VPLS

MPLS VPLS

Reference Model



Creates an emulated Ethernet LAN Segment across a wide-area network for a set of users

RFC 4664, RFC 4026

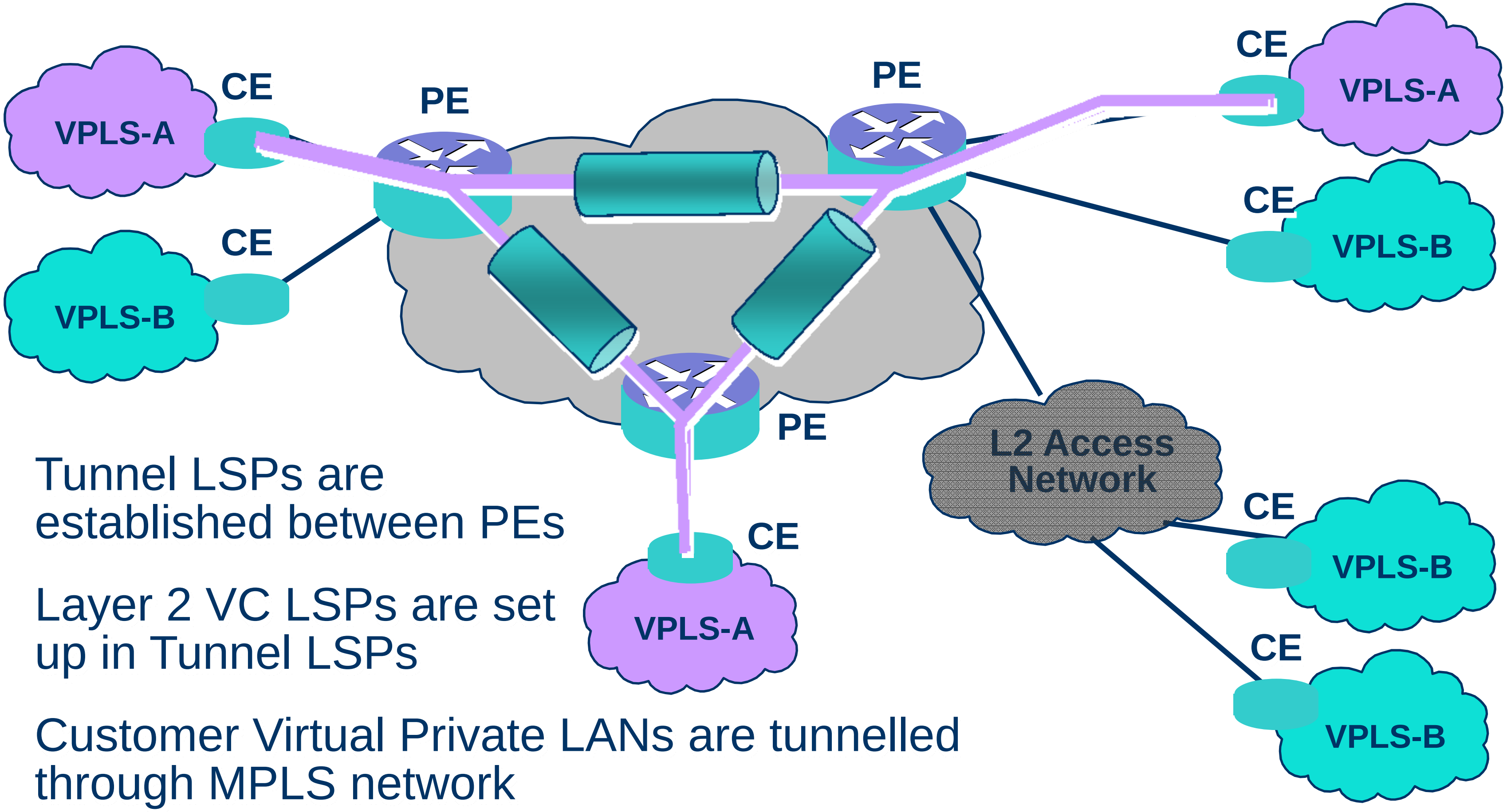
Virtual Private LAN Services

- Defines an Ethernet (IEEE 802.1D) learning bridge model over MPLS Ethernet PWs
- Defines the PE function for an MPLS VPLS network
- Creates a layer 2 broadcast domain for a closed group of users
- MAC address learning and aging on a per LSP basis
- Packet replication across LSPs for multicast, broadcast, and unknown unicast traffic
- Hierarchical VPLS for scalability

MPLS VPLS

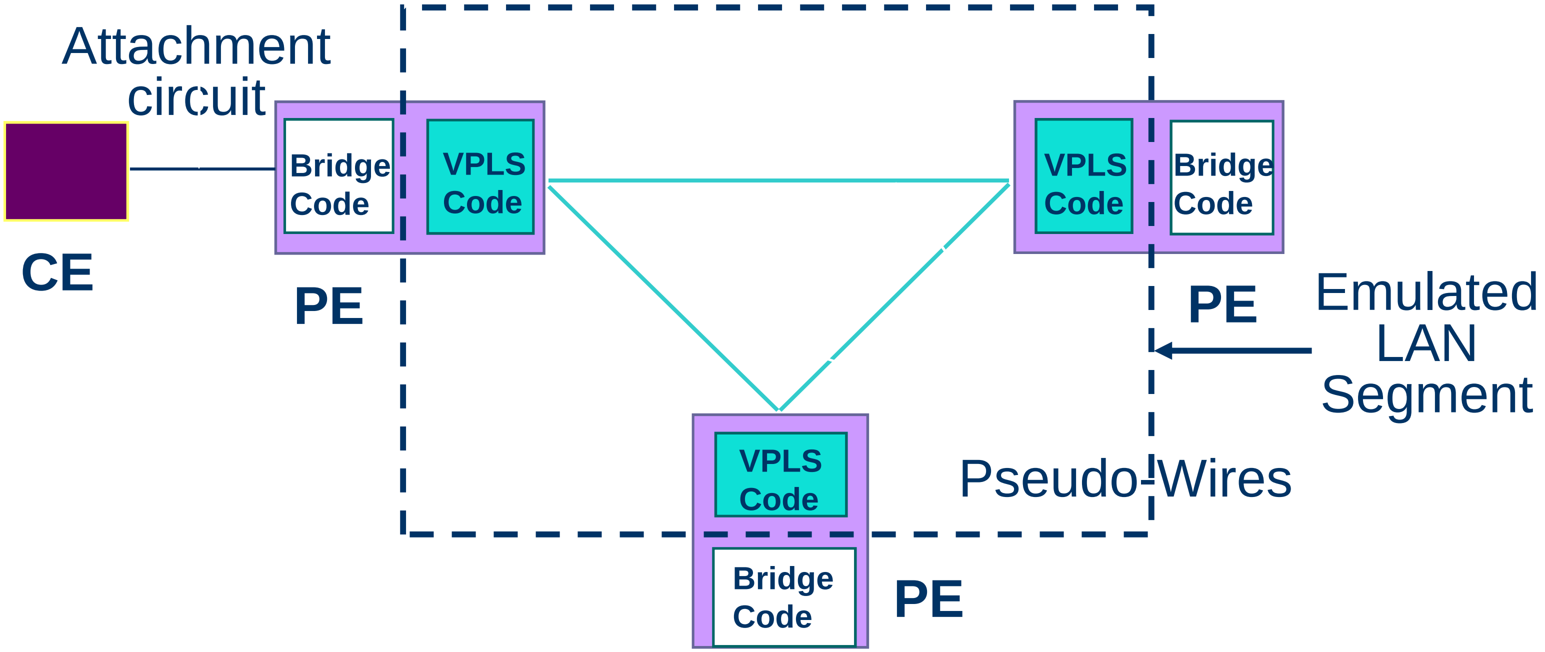
Reference Model

Emulates LAN Segment across a wide-area network



- Tunnel LSPs are established between PEs
- Layer 2 VC LSPs are set up in Tunnel LSPs
- Customer Virtual Private LANs are tunnelled through MPLS network
- Core MPLS network acts as a LAN switch

VPLS Internal PE Architecture

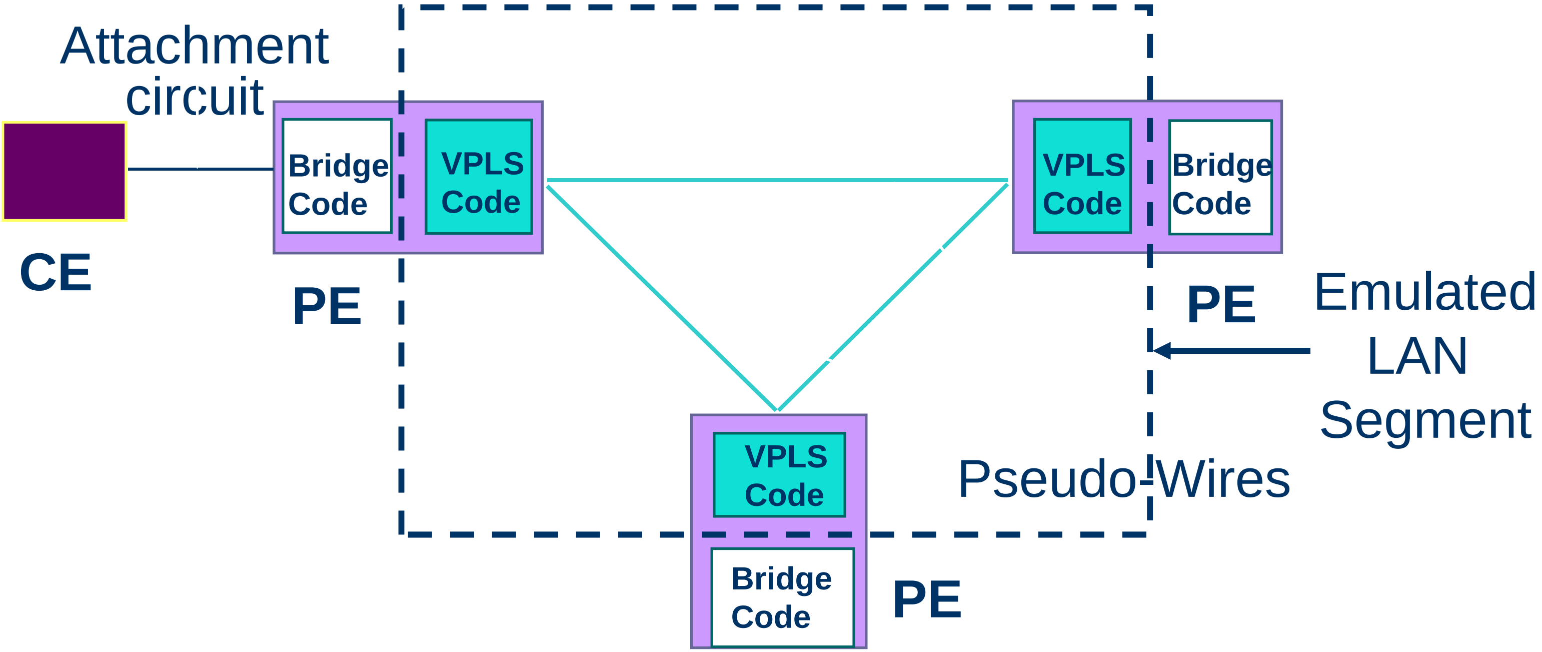


□ IEEE 802.1D bridging code

▨ IETF VPLS code

⬚ Emulated LAN instance

PE Bridging Code



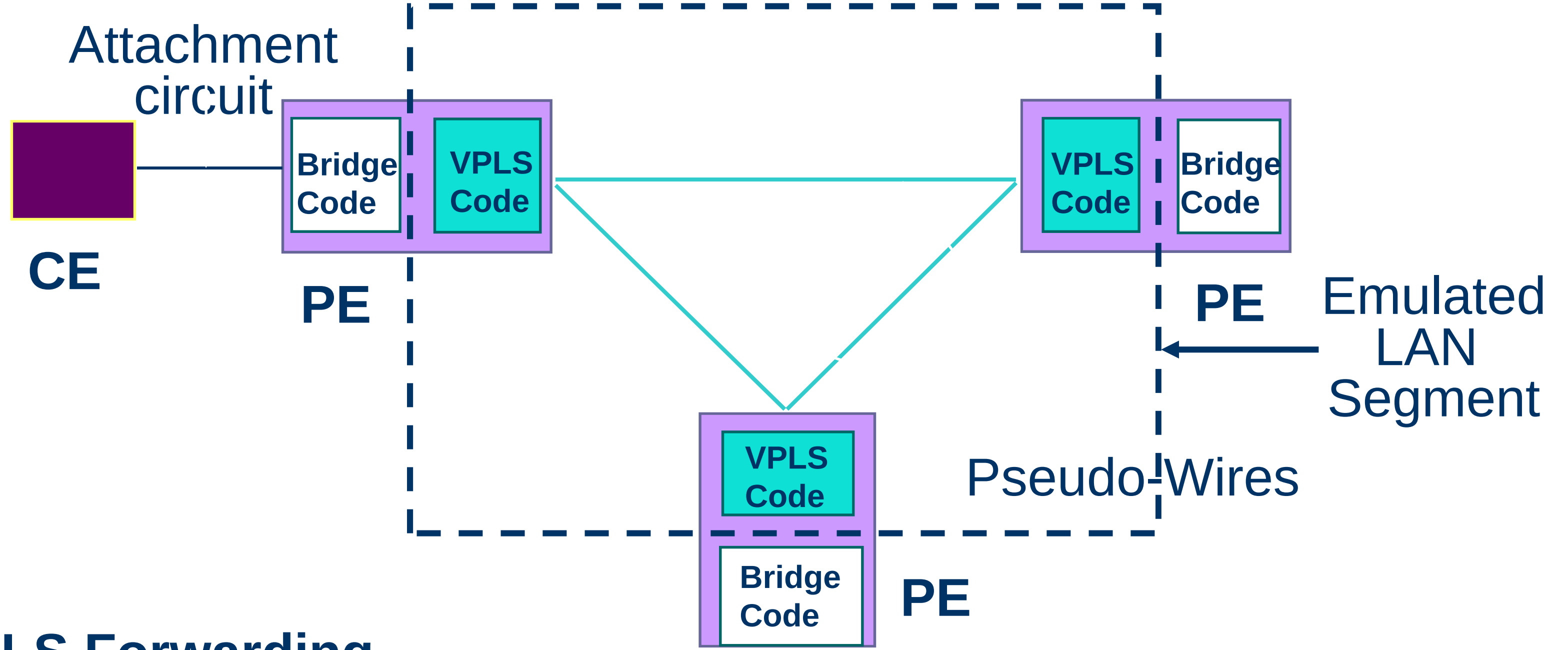
Standard IEEE 802.1D Bridging code

- Used to interface with CE facing ports
- Learn MAC addresses and aging
- Might run STP with CEs
- Used to interface with VPLS
- Might run STP between PEs

□ IEEE 802.1D bridging code

■ IETF VPLS code

PE VPLS Code



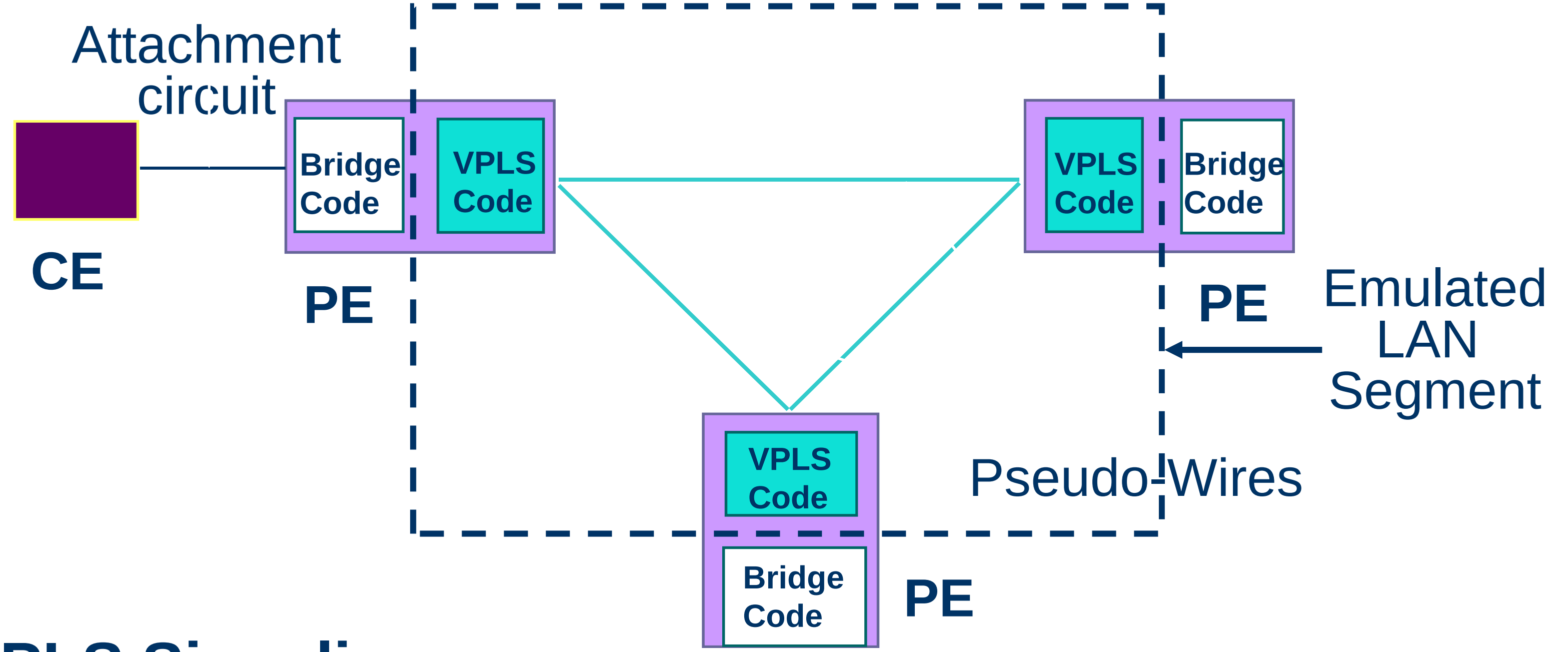
VPLS Forwarding

- Learns MAC addresses per pseudo-wire (VC LSP)
- Forwarding based on MAC addresses
- Replicates multicast & broadcast frames
- Floods unknown frames
- Split-horizon for loop prevention

□ IEEE 802.1D bridging code

🟦 IETF VPLS code

PE VPLS Code



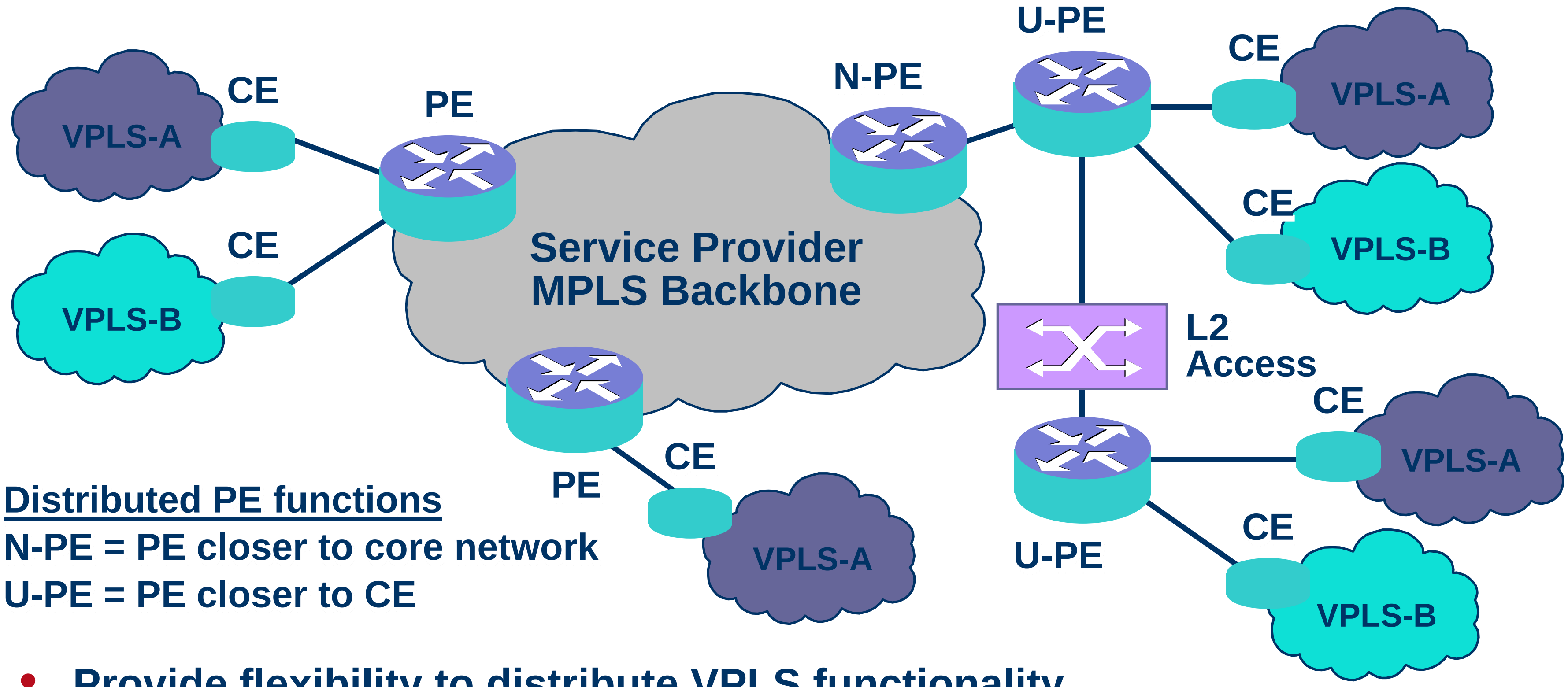
- **VPLS Signaling**
  - **Establishes pseudo-wires per VPLS between relevant PEs**
  - **Two signaling protocol options:**
    - **LDP – RFC 4762**
    - **BGP – RFC 4761**
- **VPLS Discovery (Manual, LDP, BGP, DNS)**

□ IEEE 802.1D bridging code

🟦 IETF VPLS code

MPLS VPLS

Reference Model – Distributed PE Functions



Distributed PE functions

N-PE = PE closer to core network

U-PE = PE closer to CE

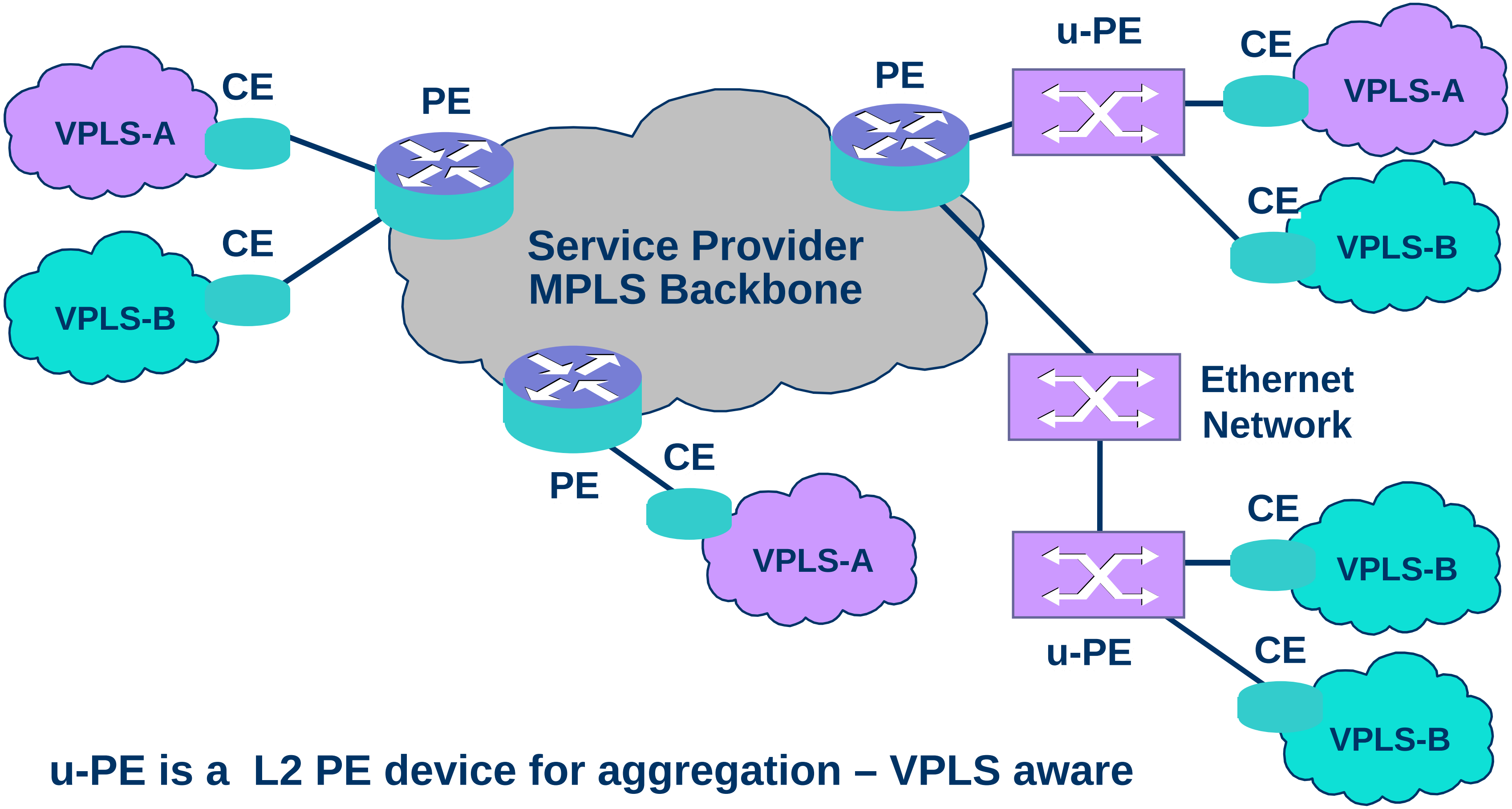
- **Provide flexibility to distribute VPLS functionality**
  - **Ex: U-PE might provide L2 aggregation and L2 functions such as MAC address learning and flooding and have limited L3 functions; N-PE might provide discovery, PE-PE signaling and establish tunnels/PWs/VCs**
- **Reduce solution cost: low cost L2 aggregation devices and utilize embedded equipment**

N-PE: Network-Facing PE
U-PE: User-Facing PE

RFC 4664, RFC 4026

MPLS VPLS
Reference Model

Virtual Private LAN Service (VPLS)
Using BGP for Auto-Discovery and Signaling



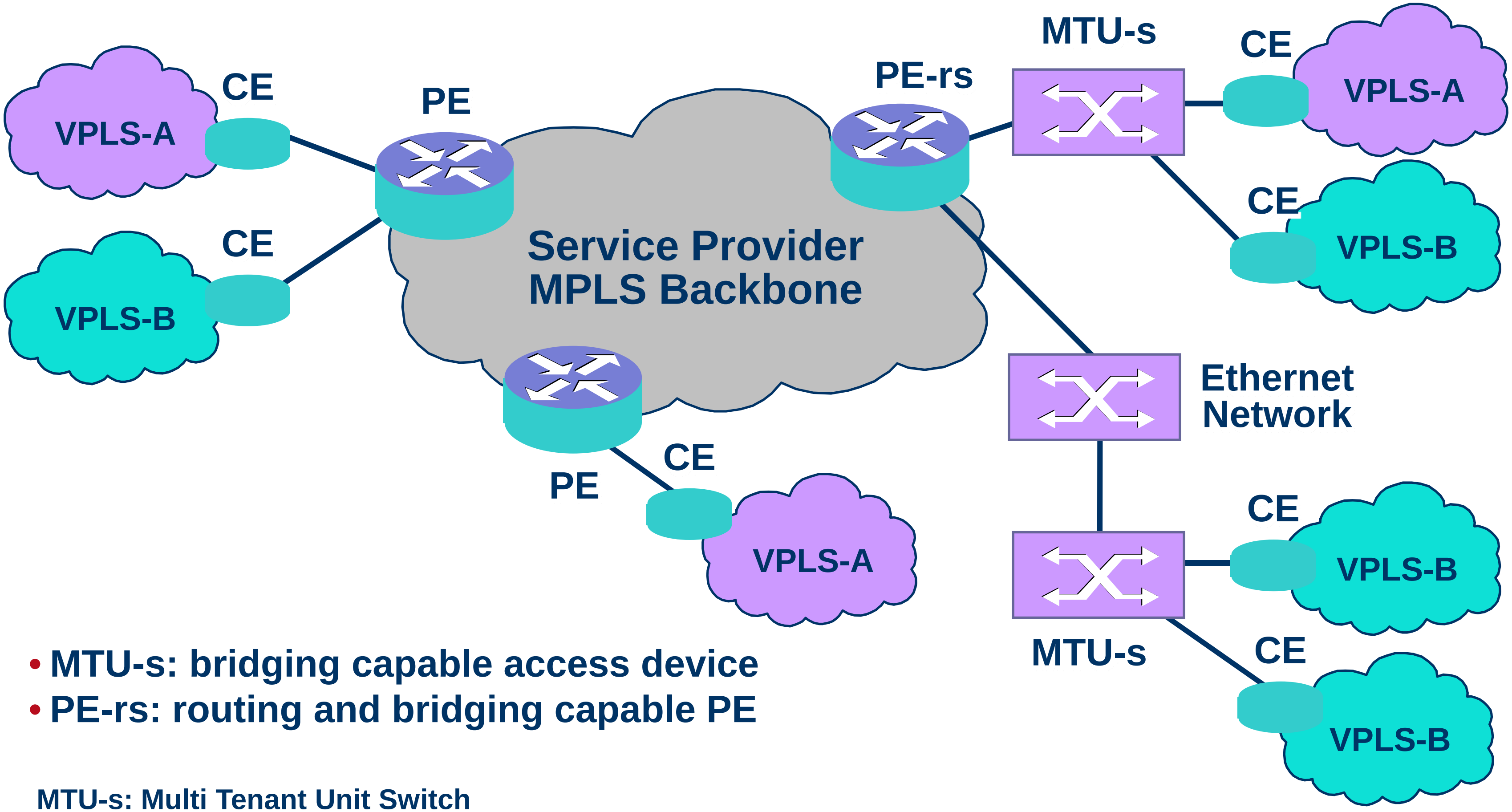
u-PE is a L2 PE device for aggregation – VPLS aware

RFC 4761

MPLS VPLS

Reference Model

Virtual Private LAN Service (VPLS)
Using Label Distribution Protocol (LDP) Signaling



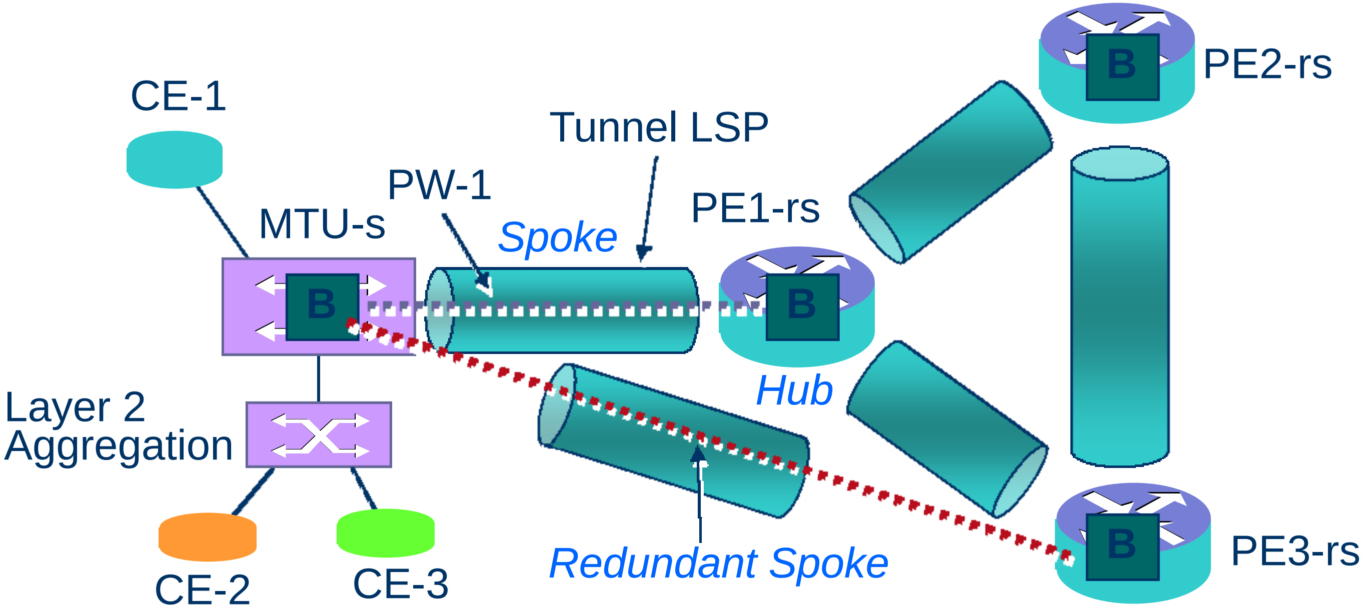
- MTU-s: bridging capable access device
- PE-rs: routing and bridging capable PE

MTU-s: Multi Tenant Unit Switch

RFC 4762

Virtual Private LAN Services

RFC 4762



- Reduce signaling and packet replication to allow large scale deployment of VPLS - Hub and spoke
- Uses single spoke PW for each VPLS service between edge MTU-s and VPLS aware PE-rs devices
- Redundant spoke to avoid single point of failure

MTU-s: bridging capable access device
PE-rs: routing and bridging capable PE

B **= Virtual VPLS (Bridge) Instance**

VPLS Scalability

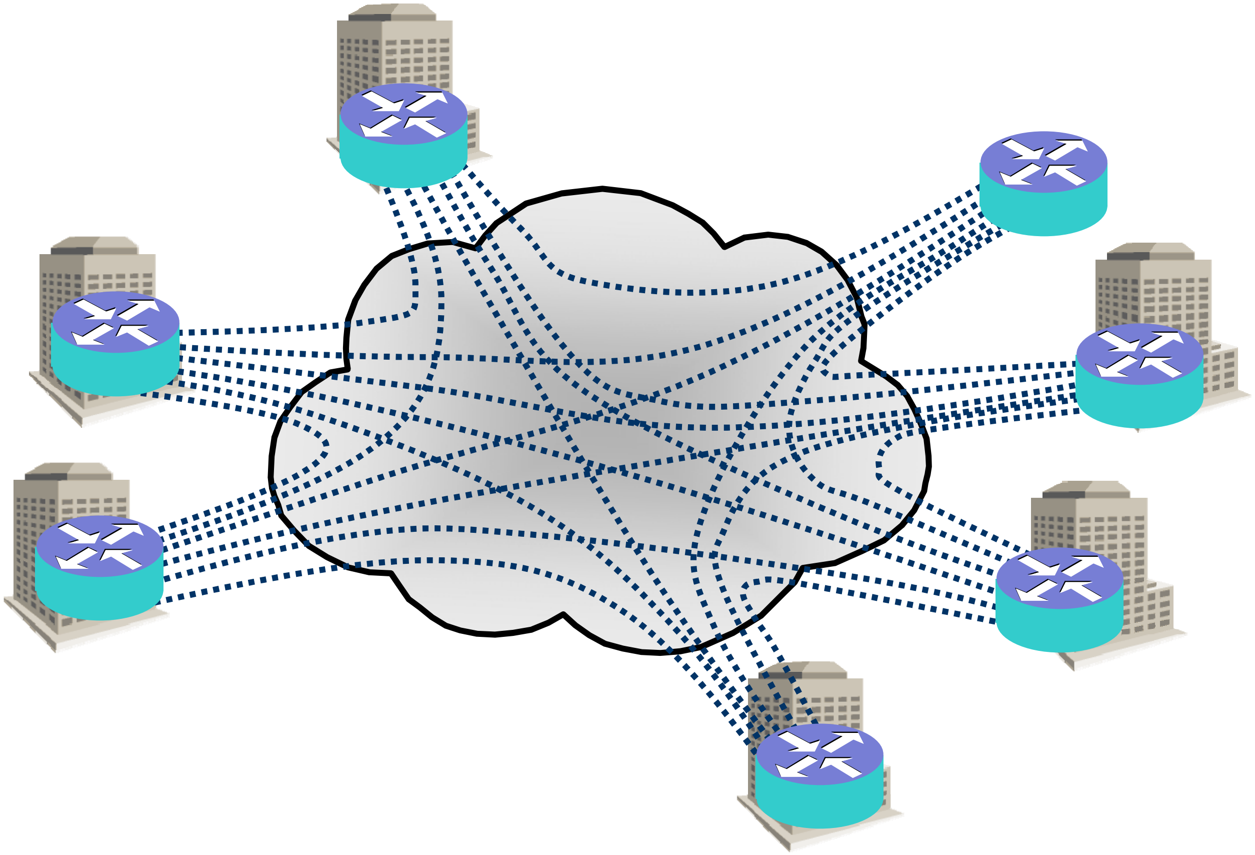
Parameters

- Number of MAC Addresses
- Number of replications
- Number of LSPs
- Number of VPLS instances
- Number of LDP peers
- Number of PEs

VPLS Scalability

Signaling Overhead – Flat Topology

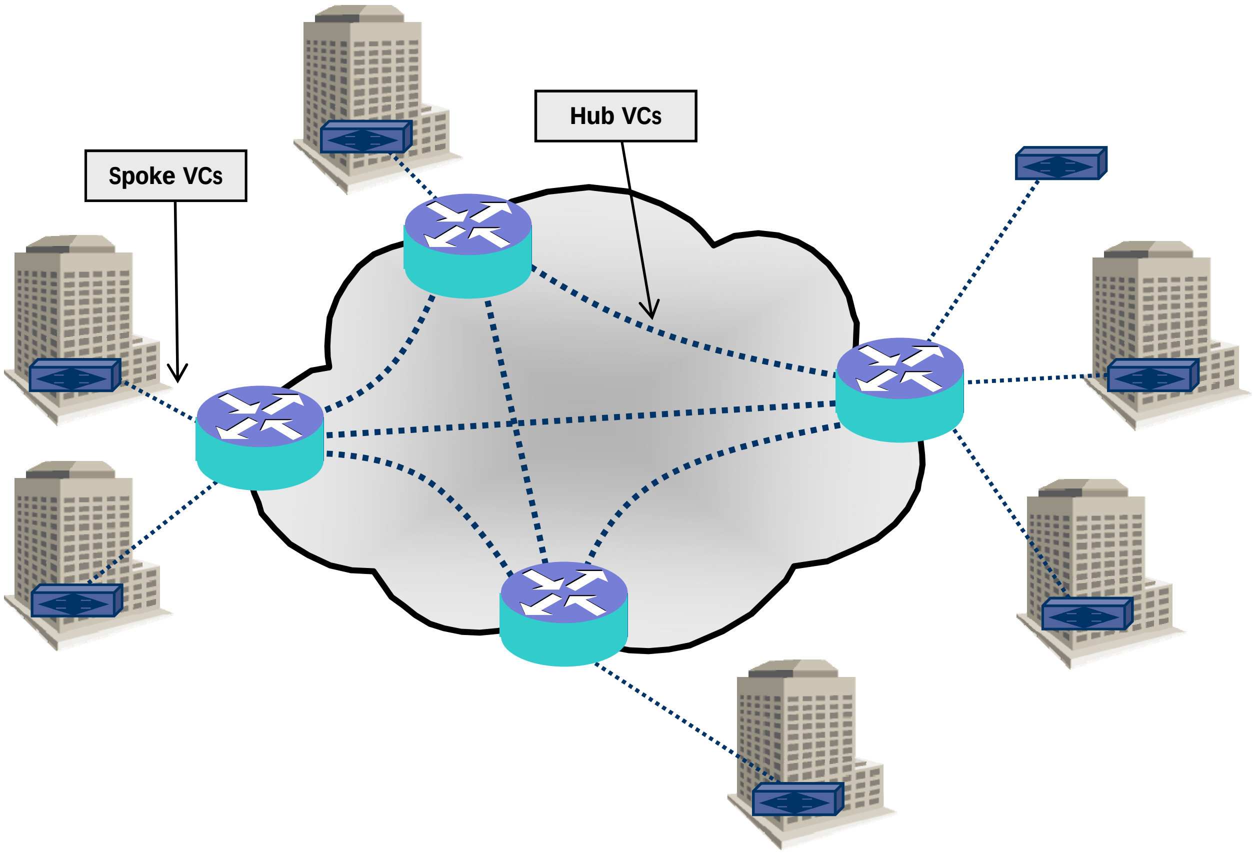
- Architecture has a direct impact on the Signaling Overhead (control plane)



VPLS Scalability

Signaling Overhead – Hierarchical Topology

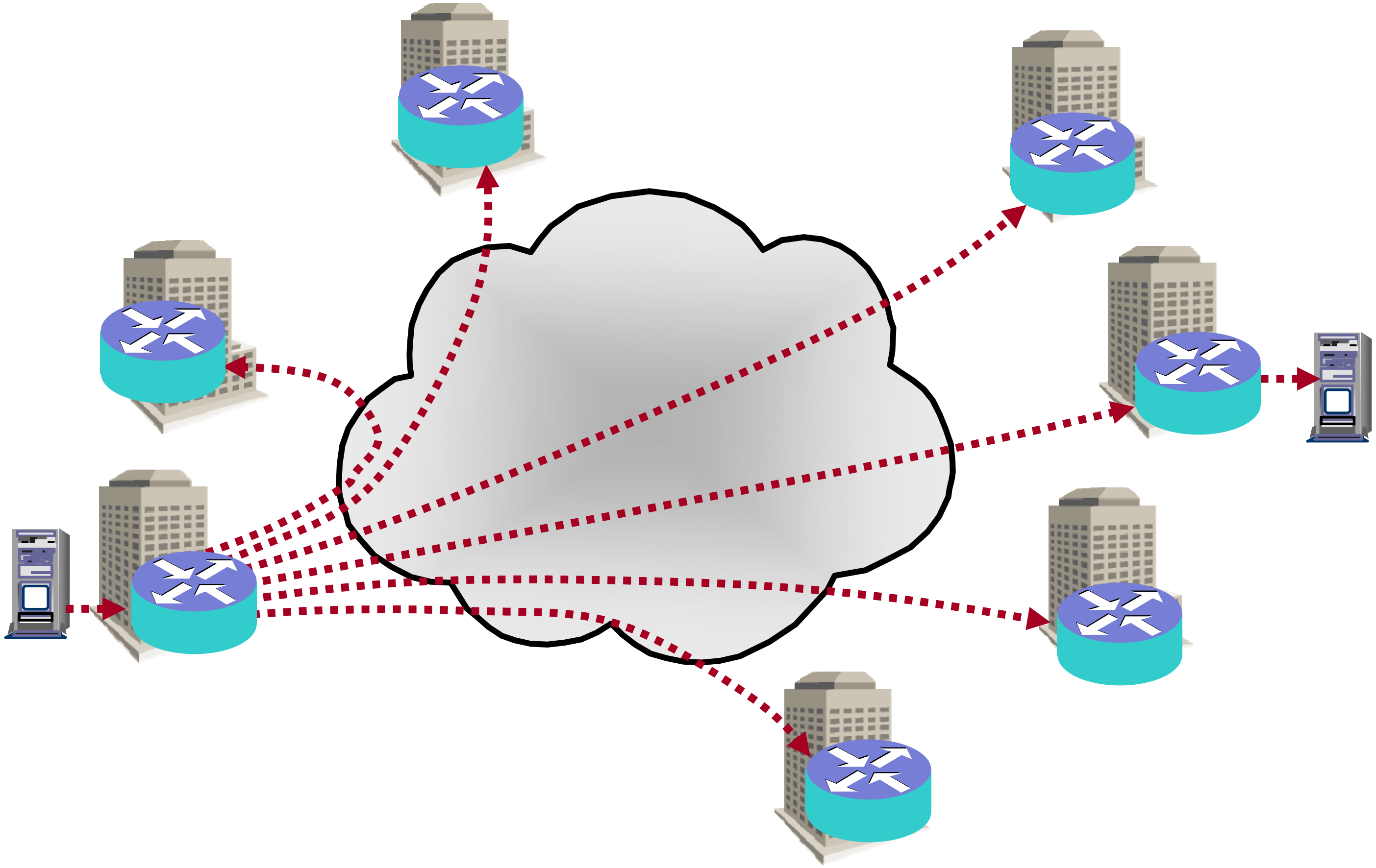
- **Architecture has a direct impact on the Signaling Overhead (control plane)**



VPLS Scalability

Replication Overhead – Flat Topology

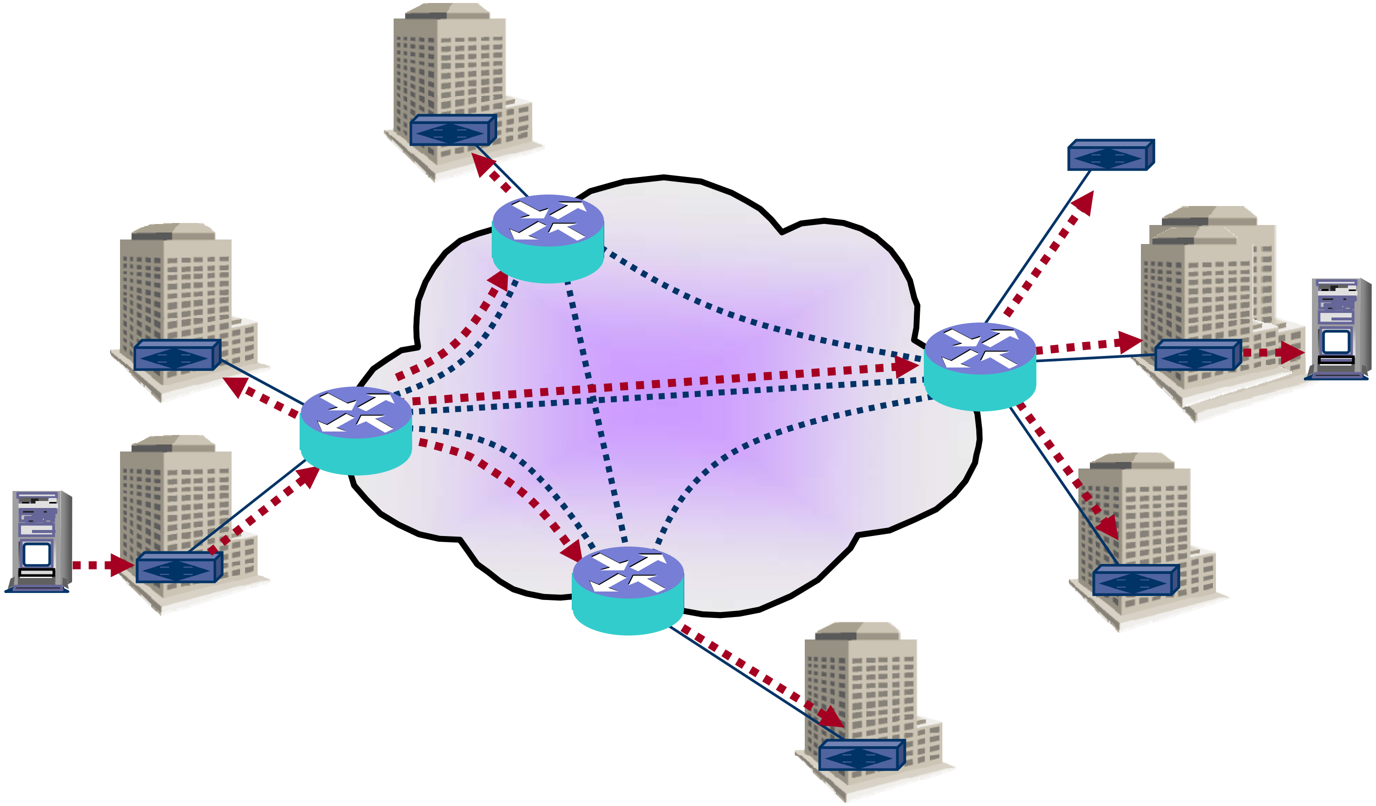
- **Architecture has a direct impact on Replication Overhead (forwarding plane)**



VPLS Scalability

Replication Overhead – Hierarchical Topology

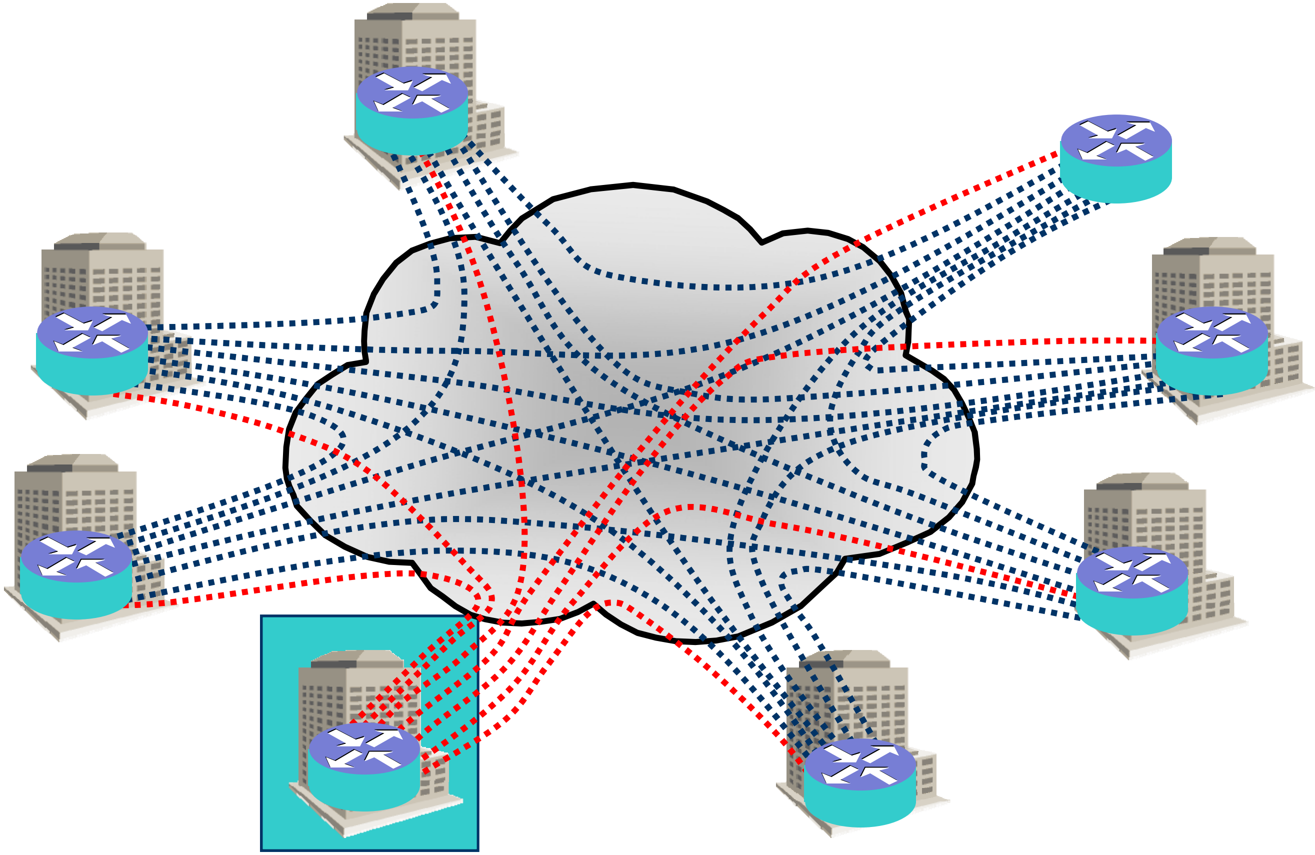
- **Architecture has a direct impact on Replication Overhead (forwarding plane)**



VPLS Scalability

Adding a New Site – <u>Flat</u> Topology

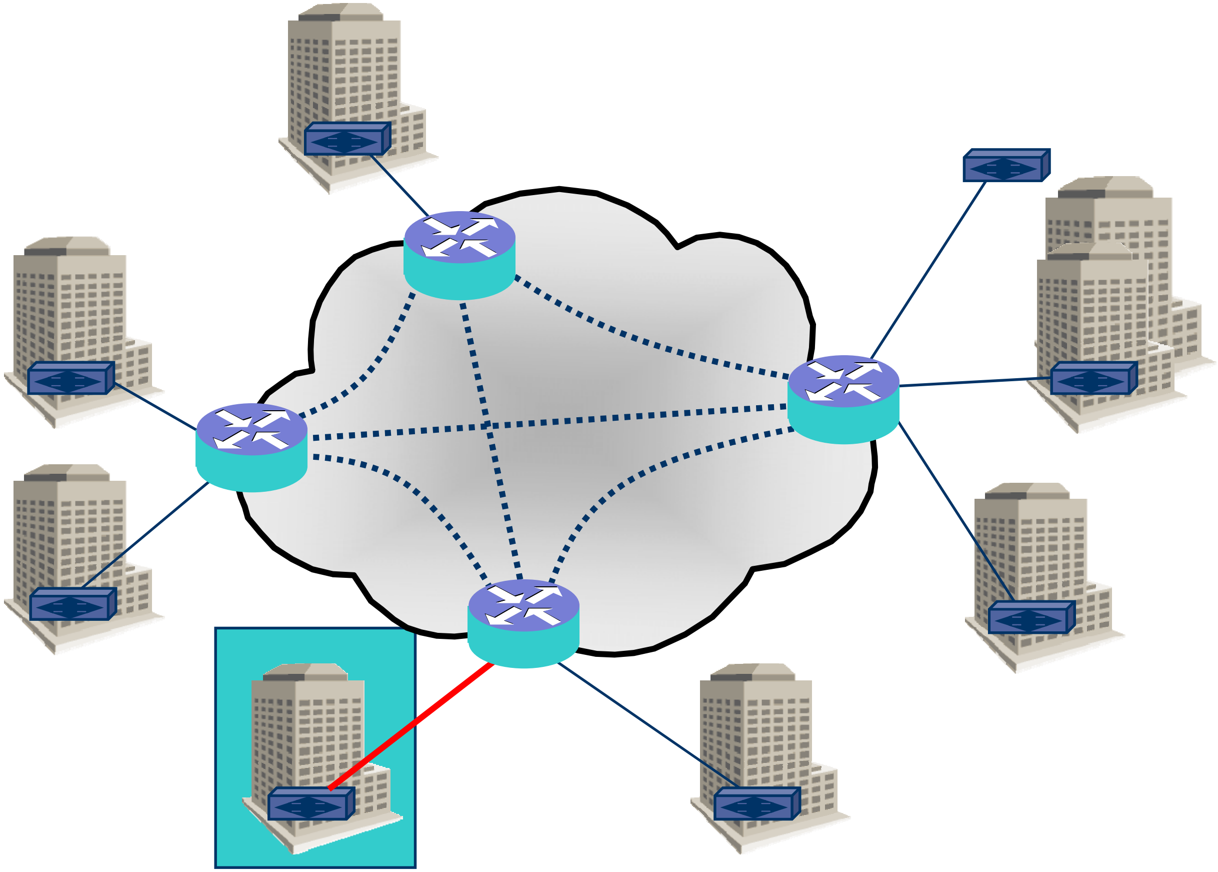
- Architecture affects <u>Provisioning & Signaling</u> between all nodes



VPLS Scalability

Adding a New Site – Hierarchical Topology

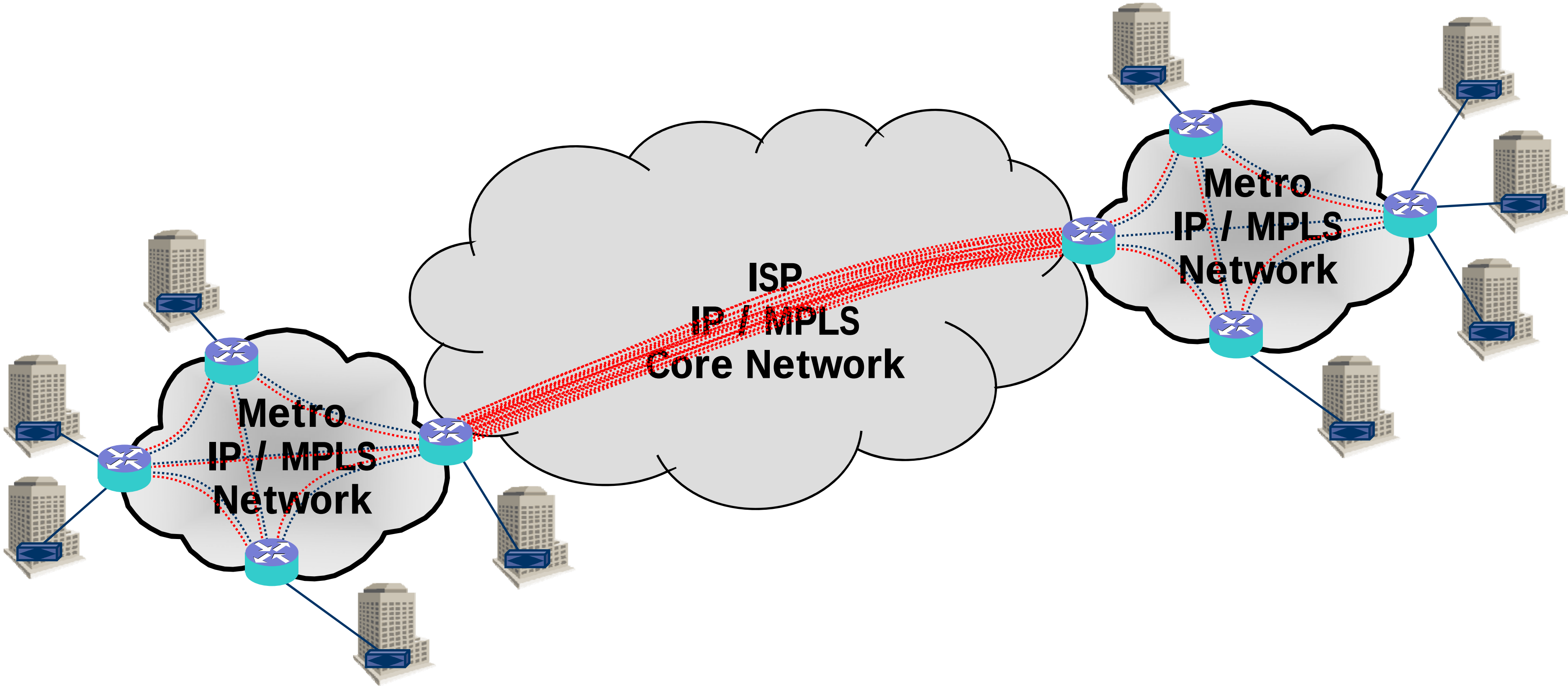
- **Architecture affects Provisioning & Signaling between all nodes**



VPLS Scalability

Inter-Metro Service

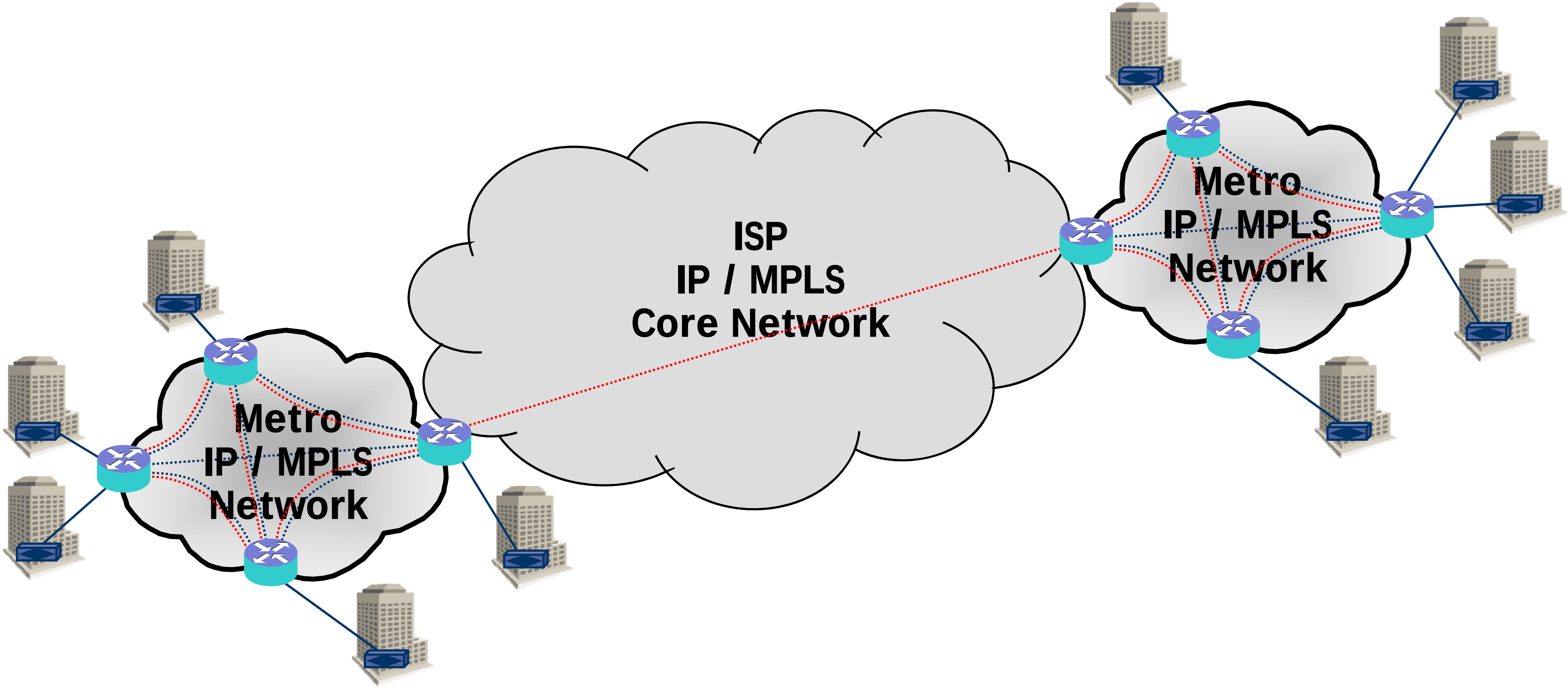
- Architecture has a direct impact on ability to offer Inter-Metro Service



VPLS Scalability

Inter-Metro Service

- Architecture has a direct impact on ability to offer Inter-Metro Service



VPLS Scalability

FIB Size

- VPLS FIB size depends on the type of Service Offering:
  - Multi-protocol Inter-connect service
    - Mimics the DSL Tariff Model
    - Customers are charged per site per block of MAC addresses
  - Router Inter-connect
    - One MAC address per site
- Same Network Design principles apply for
  - MAC FIB Size of VPLS Service and,
  - Route Table Size of Virtual Private Routed Network (VPRN) Service

FIB: Forwarding Information Base

IETF Layer 2 VPNs

RFC 4665

- Service requirements for L2 VPNs
  - Virtual Private Wire Services (VPWS) - point-to-point VPNs
  - Virtual Private LAN services (VPLS) - multipoint-to-multipoint VPNs
  - Service Provider and Enterprise Views'
- Checklist of requirements to help evaluate how an approach satisfies specific requirements
- Service Level Specification (SLS)

Section 4

Introduction to Multi-Service Interworking

MPLS VPN Tutorial Agenda

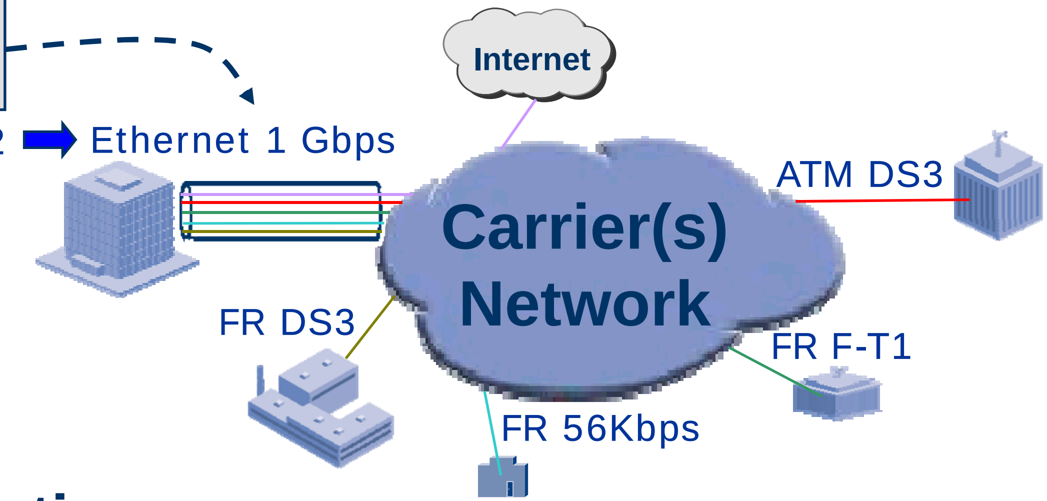
Introduction to Multi-Service Interworking over MPLS

- Interworking History and Definition
- Multi-Service Interworking of Ethernet over MPLS
- Migration Scenarios and Benefits

Why Interwork?

New business applications driving increase in bandwidth

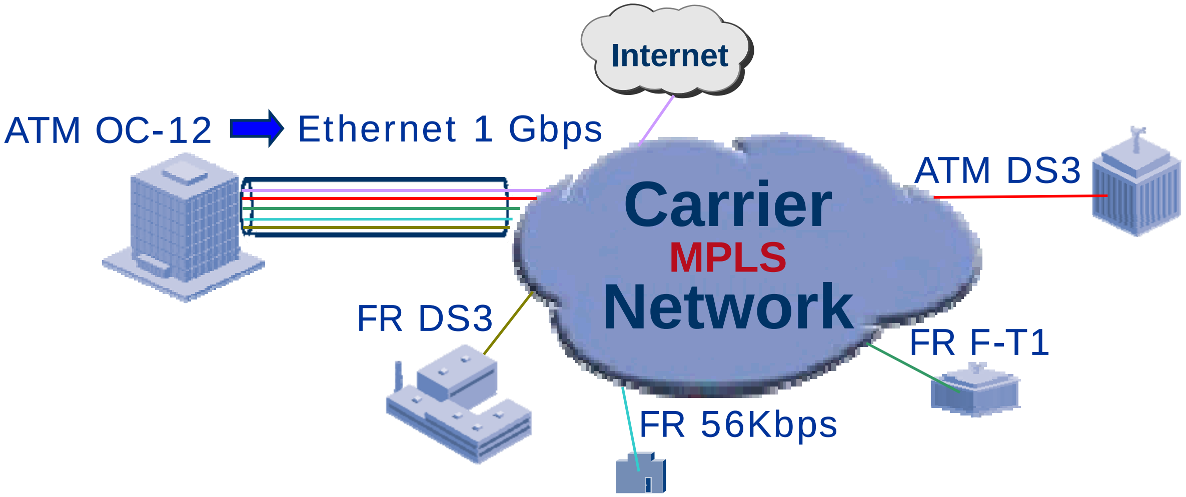
L2 vs. L3 service preference



Enterprise perspective:

- Many have an embedded Frame Relay and/or ATM network
- Need to cost effectively scale bandwidth at select sites to support new business applications
- Maintain a network with mixture of services, bandwidths to match application needs at specific sites
- Reduce cost, time and risk to address emerging needs

Why Interwork?



Carrier Perspective:

- Want a common edge infrastructure to support and “Interwork” with legacy and new services
- Support all legacy transport technologies and services
- Planning to converge on an IP / MPLS core
- Want to seamlessly introduce Metro Ethernet services and IP VPNs

Interworking

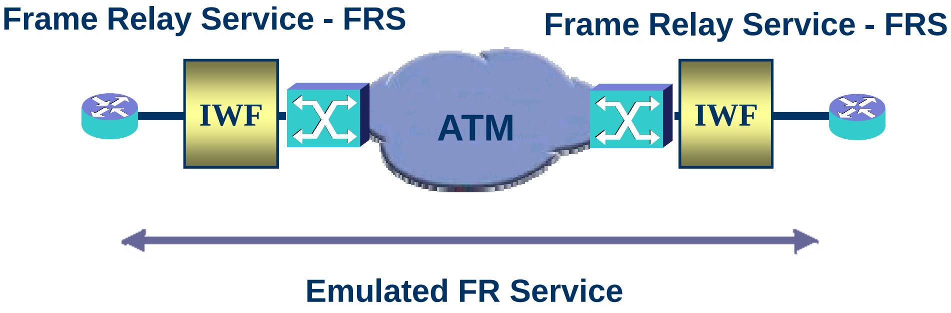
History

- The Frame Relay Forum defined the Network Interworking function between Frame Relay and ATM in the FRF.5 document finalized in 1994
- The Frame Relay Forum defined the Service interworking function between Frame Relay and ATM in the FRF.8.2 document finalized in 2004
- Why define FR and ATM interworking?
  - ATM cores with FR/ATM access services deployed
  - ATM and Frame Relay circuits are point-to-point
  - Both data links have services that are somewhat similar (ie. FR to AAL5) in nature even though the signaling is different

Interworking Function - IWF

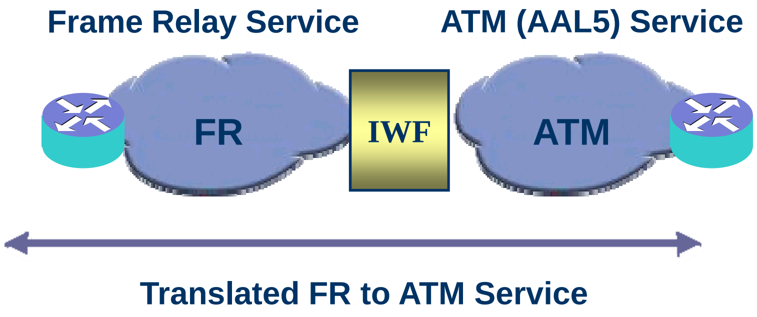
Network vs Service IWF

Network Interworking



- Network Interworking is used when one protocol is “tunneled” across another “intermediary” network / protocol
- The Network Interworking (IWF) function “terminates” and “encapsulates” the protocol over a Pt-to-Pt connection
- Service at end points ***has to be*** the same

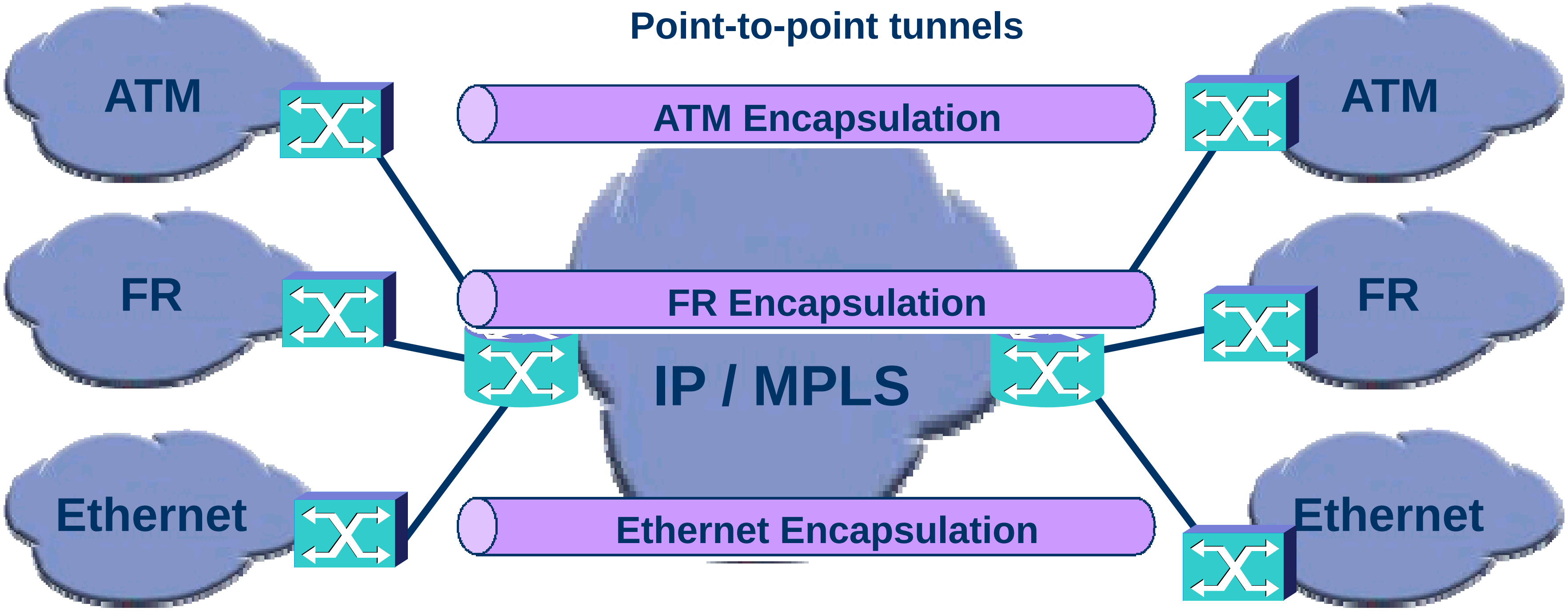
Service Interworking



- Service Interworking is required to “translate” one protocol to another protocol – used between two unlike protocols
- The Service Interworking function “translates” the control information transparently by an interworking function (IWF)
- Services at the end points ***are not*** the same

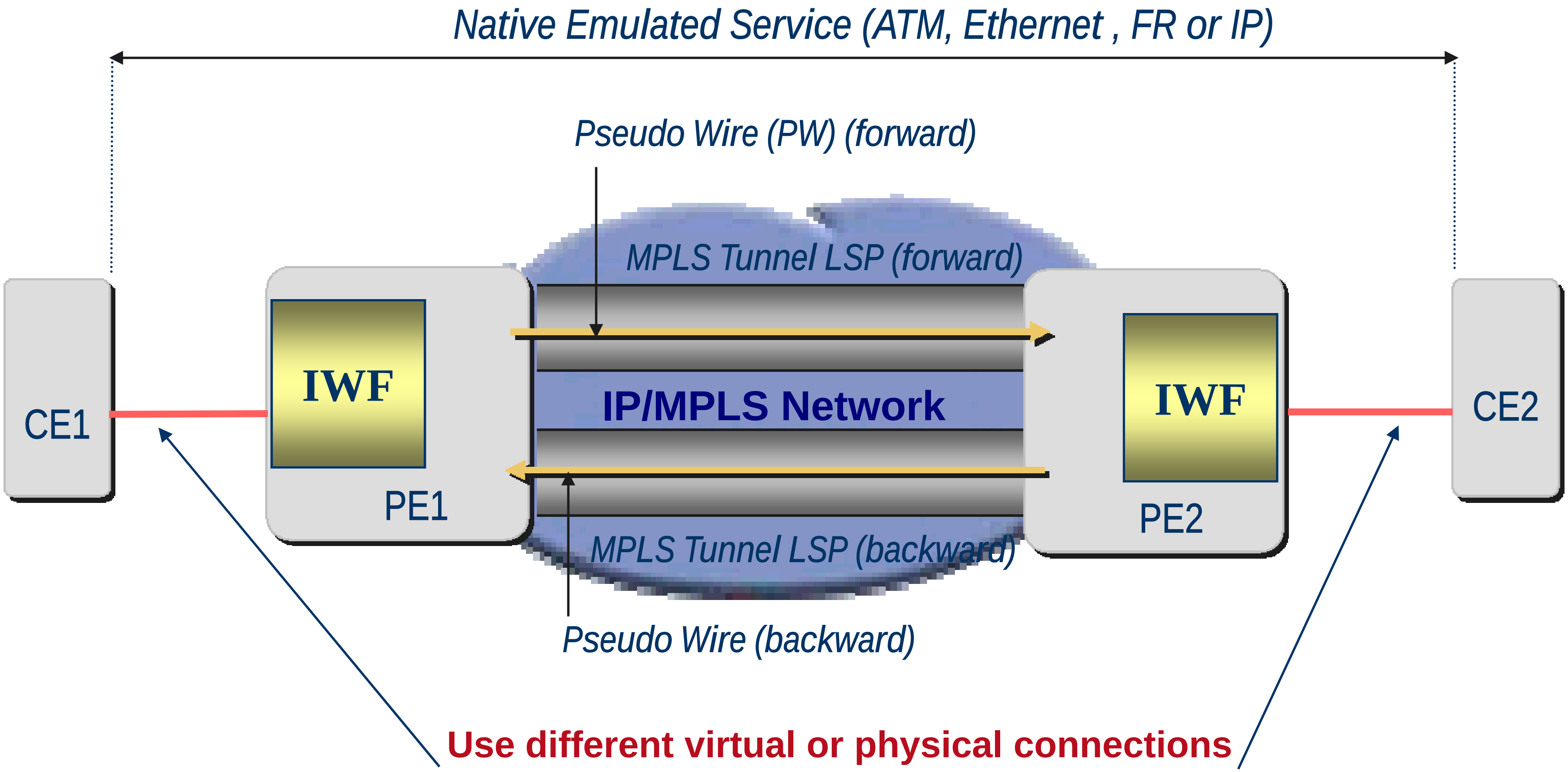
MPLS Network Interworking

IETF PWE3 Pt-to-Pt Encapsulation



Service has to be pt-to-pt between like services: ATM to ATM, FR to FR, Ethernet to Ethernet, etc

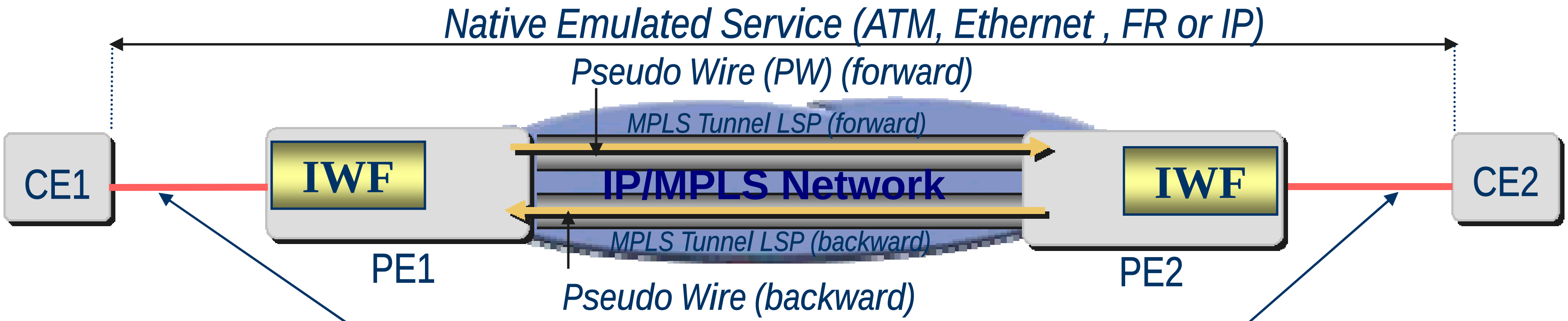
MPLS Multi-Service Interworking
Reference Model


Native Emulated Service (ATM, Ethernet , FR or IP)
Pseudo Wire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
IWF
PE1
IP/MPLS Network
IWF
PE2
CE2
MPLS Tunnel LSP (backward)
Pseudo Wire (backward)

Use different virtual or physical connections (Ex: ATM, Ethernet, FR, etc) at each end Attachment Circuit (AC)

PE: Provider Edge
CE: Customer Edge
IWF: Interworking Function
Multi-Service: Services are ATM, Ethernet, FR and IP

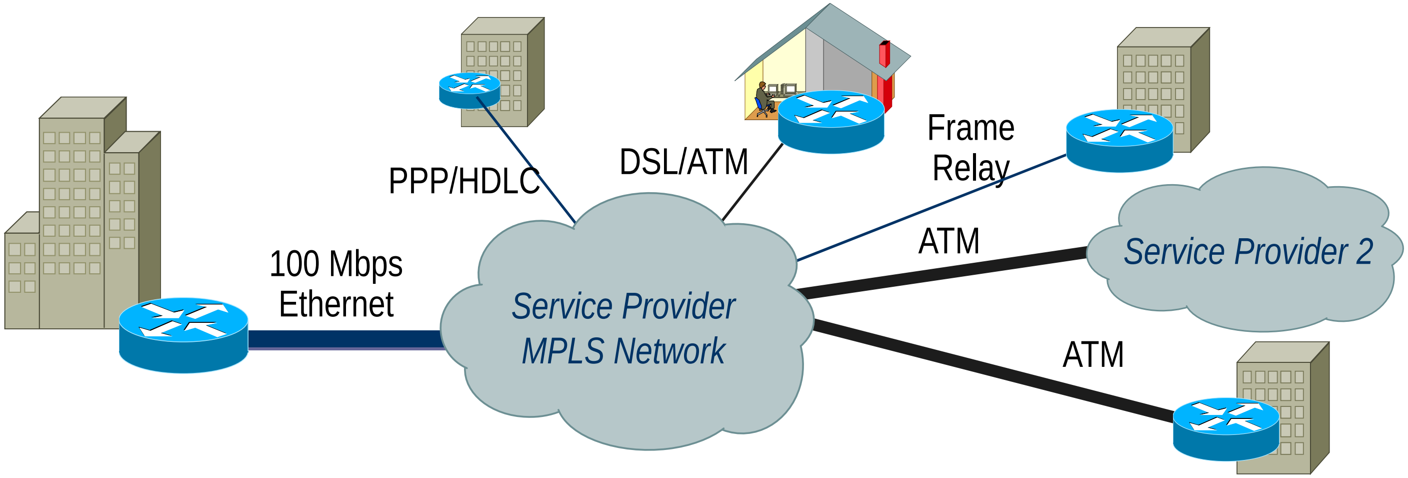
Multi-Service Interworking


Native Emulated Service (ATM, Ethernet , FR or IP)
Pseudo Wire (PW) (forward)
MPLS Tunnel LSP (forward)
CE1
IWF
IP/MPLS Network
IWF
CE2
MPLS Tunnel LSP (backward)
PE1
PE2
Pseudo Wire (backward)

Use different virtual or physical connections (Ex: ATM, Ethernet, FR, etc) at each end Attachment Circuit (AC)

- **Multi-Service Interworking of Ethernet over MPLS**
- Multi-Service Interworking of IP over MPLS
  - MFA Forum Multi-Service Interworking – IP over MPLS Implementation Agreement 16.0
- Frame Relay and ATM Service Interworking over MPLS
  - MFA Forum Multi-Service Interworking – Frame Relay and ATM Service Interworking over MPLS Implementation Agreement 15.0
- Fault Management for Multi-Service Interworking
  - MFA Forum Fault Management for Multi-Service Interworking over MPLS Implementation Agreement 13.0

Multi-Service Interworking - Ethernet over MPLS

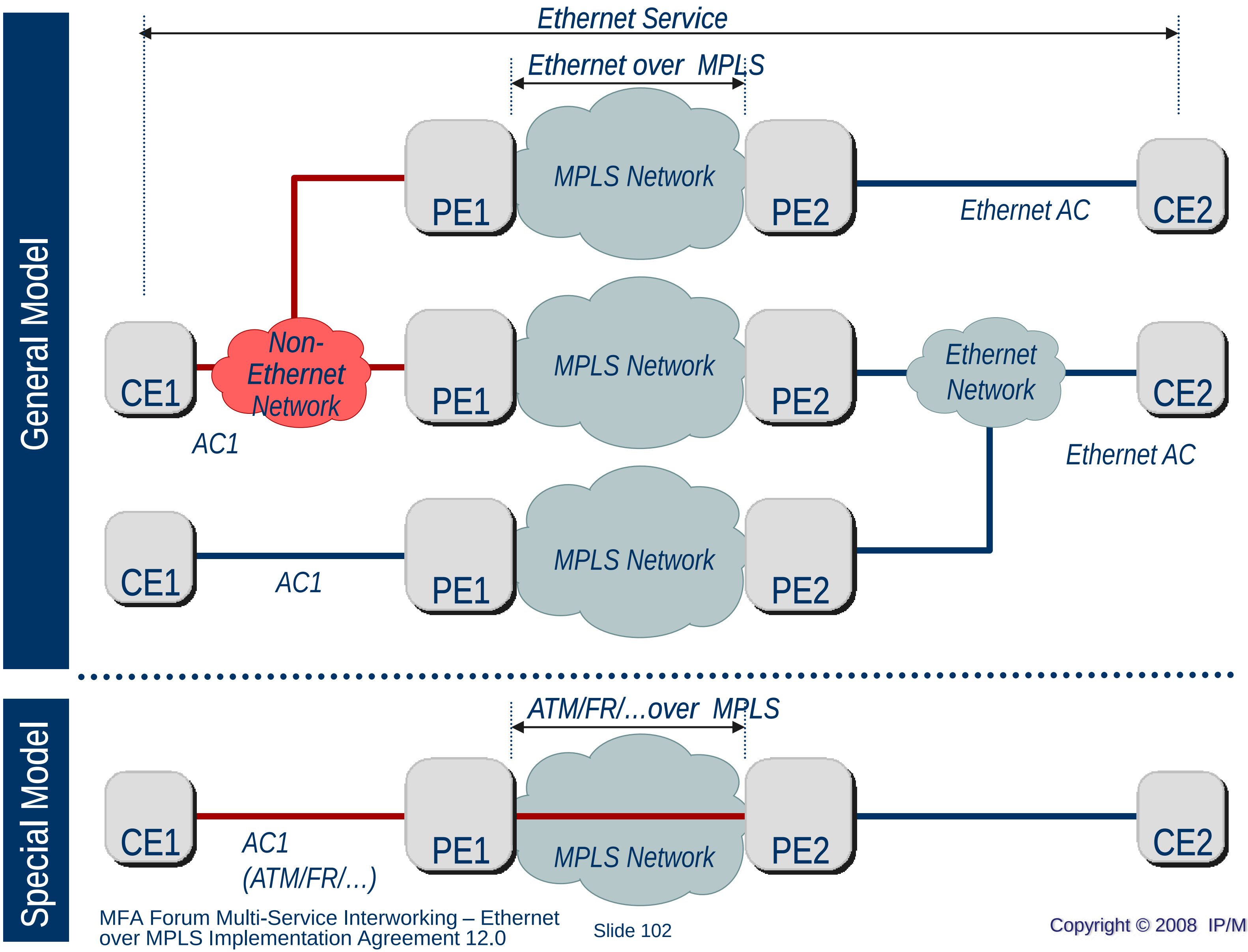


- **Ubiquitous Ethernet-Service offering requires that different UNI/NNIs are supported – Ethernet as well as ATM, FR, PPP, …**
  - **SPs expand their existing Ethernet UNI/NNI offering**

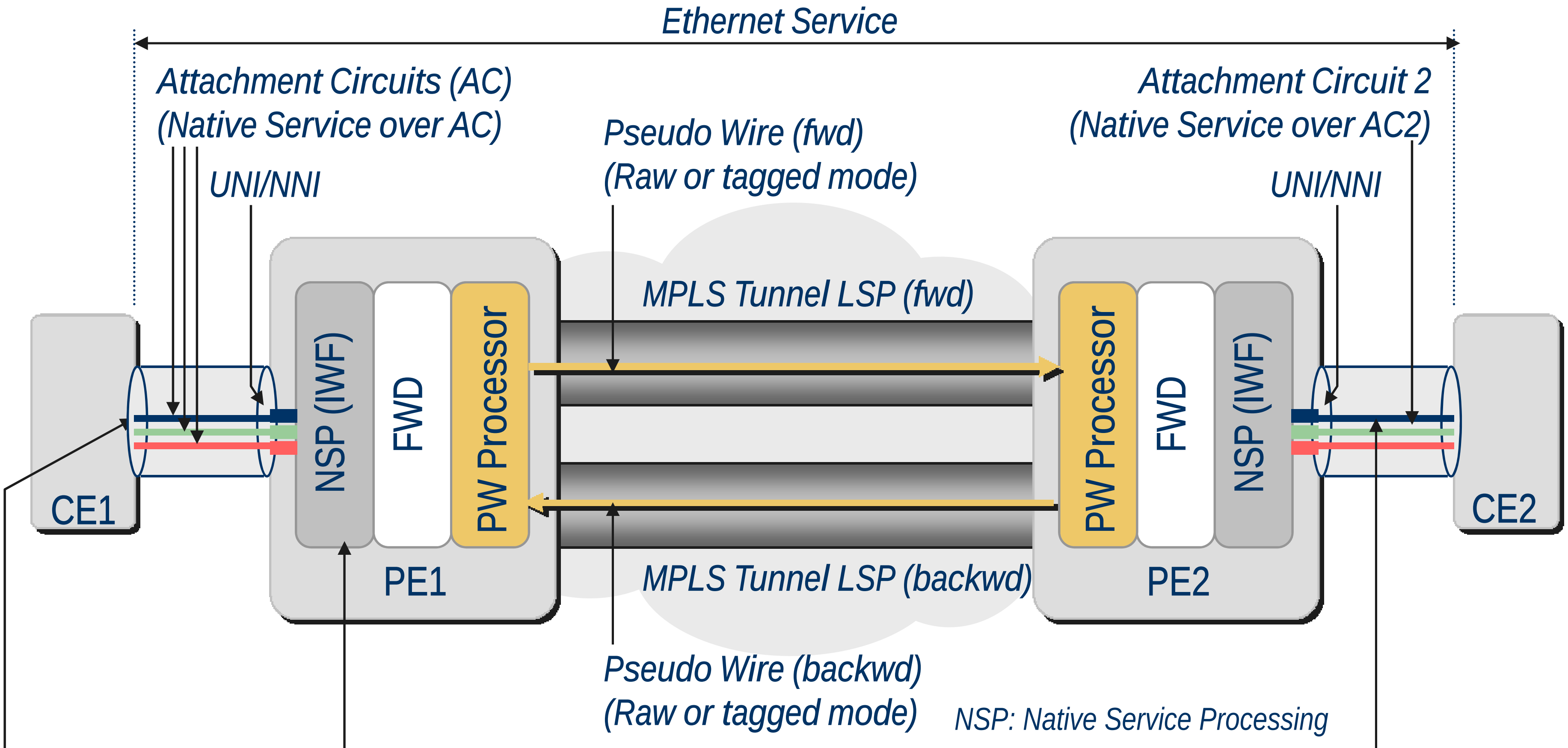
- **Characteristics**
  - **Native Service: Ethernet**
  - **Consistent service definitions across technology boundaries**
  - **Point-to-Point and Multipoint**
  - **Independence from CE protocol processing (address resolution, L3-protocols,…)**

MFA Forum Multi-Service Interworking – Ethernet over MPLS Implementation Agreement 12.0

Models for Ethernet Interworking



Interworking Reference Model



Ethernet Service

Attachment Circuits (AC)
(Native Service over AC)

UNI/NNI

Pseudo Wire (fwd)
(Raw or tagged mode)

Attachment Circuit 2
(Native Service over AC2)

UNI/NNI

MPLS Tunnel LSP (fwd)

MPLS Tunnel LSP (backwd)

Pseudo Wire (backwd)
(Raw or tagged mode)

NSP: Native Service Processing

Ethernet/FR/ATM/PPP
for Non-Ethernet:

- *Bridged Interface*
- *Routed Interface with bridged encapsulation*

- *Encapsulation Mapping AC <-> PW*
- *Traffic Management Interworking*
- *PVC Management Interworking*

- *Ethernet*
- *ATM: RFC 2684-B*
- *FR: RFC 2427-B*
- *PPP: RFC 2878*

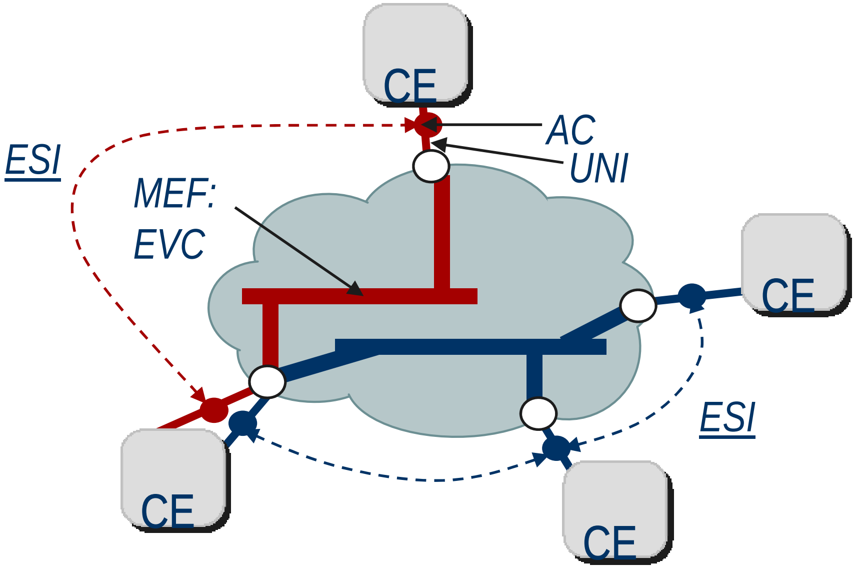
MFA Forum Multi-Service Interworking – Ethernet over MPLS Implementation Agreement 12.0

Multi-Service Interworking of Ethernet over MPLS - Observations

- Interworking is a local function to the PE
  - PE only needs to implement procedures for those interfaces it supports (e.g. PE with ATM: RFC2684 bridged only)
  - PE only needs to support PW of type Ethernet – irrespective of the other end. Set of translations limited to (to/from) Ethernet
  - AC configuration local to the PE
  - AC termination on PE supports VPLS (and VPWS) – MAC-addresses are visible to the PE
- CPE uses bridged encapsulation (native Service is Ethernet)
  - Implicit support for any L3 Network protocol
  - ARP resolution done by both end CPEs – no handling of protocol specific address resolution required
  - Integrated Routing and Bridging for Frame-Relay AC, IRB/Routed Bridge Encapsulation for ATM AC
  - Required configuration changes for CE devices that have routed interfaces
- Consider hidden complexities, e.g. IP-routing protocols behave differently over broadcast & non-broadcast media

Ethernet Service Instance (ESI)

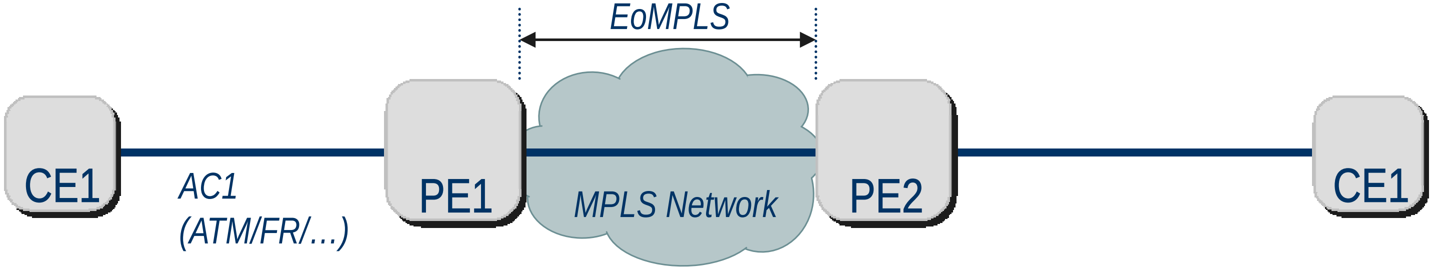
- **Ethernet Service Instance**
  - **"Association of two or more AC over which an Ethernet Service is offered to a given customer"**
- **Corresponding concepts**
  - **ESI can correspond to VPLS/VPWS (IETF L2VPN WG), S-VLAN (IEEE 802.1ad)**
  - **Note: MEF EVC associates a set of UNI, while ESI associates a set of AC**
- **Multiple Mappings options at individual AC to the corresponding Service Instance**

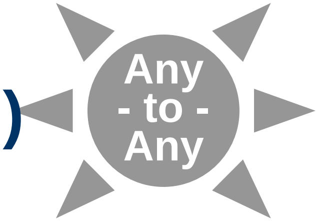


Mapping at an AC (per ESI)	Ethernet Interface	ATM/ FR VC	PPP/HDLC Interface
Port based (untagged only)	✓	✓	✓
Port based (tagged & untagged)	✓	✓	✓
VLAN mapping	✓	NS	NS
VLAN bundling	✓	NS	NS

NS: Not specified in this version

Ethernet Service Interworking Encapsulation Formats





Native Ethernet

Ethernet VLAN

Bridged Ethernet over ATM (RFC 2684-B)

Bridged Ethernet over FR (RFC 2427-B)

Bridged Ethernet over HDLC/PPP (RFC 2878)

Native Ethernet

Ethernet VLAN

Bridged Ethernet over ATM (RFC 2684-B)

Bridged Ethernet over FR (RFC 2427-B)

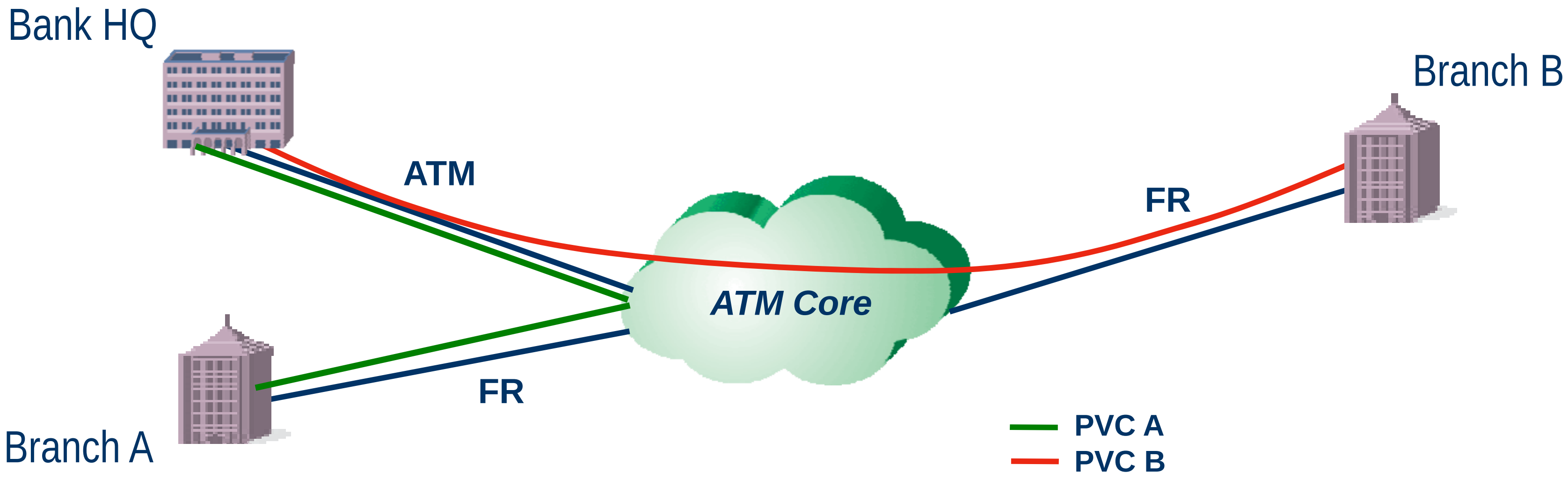
Bridged Ethernet over HDLC/PPP (RFC 2878)

MFA Forum Multi-Service Interworking – Ethernet over MPLS Implementation Agreement 12.0

Multi-Service Interworking of Ethernet over MPLS Summary

- Layer 2 Service Interworking is critically important to Ethernet WAN services
  - Limited Ethernet footprint
  - Leverages installed base of ATM/Frame Relay, and HDLC copper based circuits
- General Interworking Model
  - Concept of Ethernet Service Instance
  - Local Termination of the AC – keep complexities low
- Standards Evolution to support comprehensive service interworking
  - Ethernet OAM standards work (ITU, IEEE)

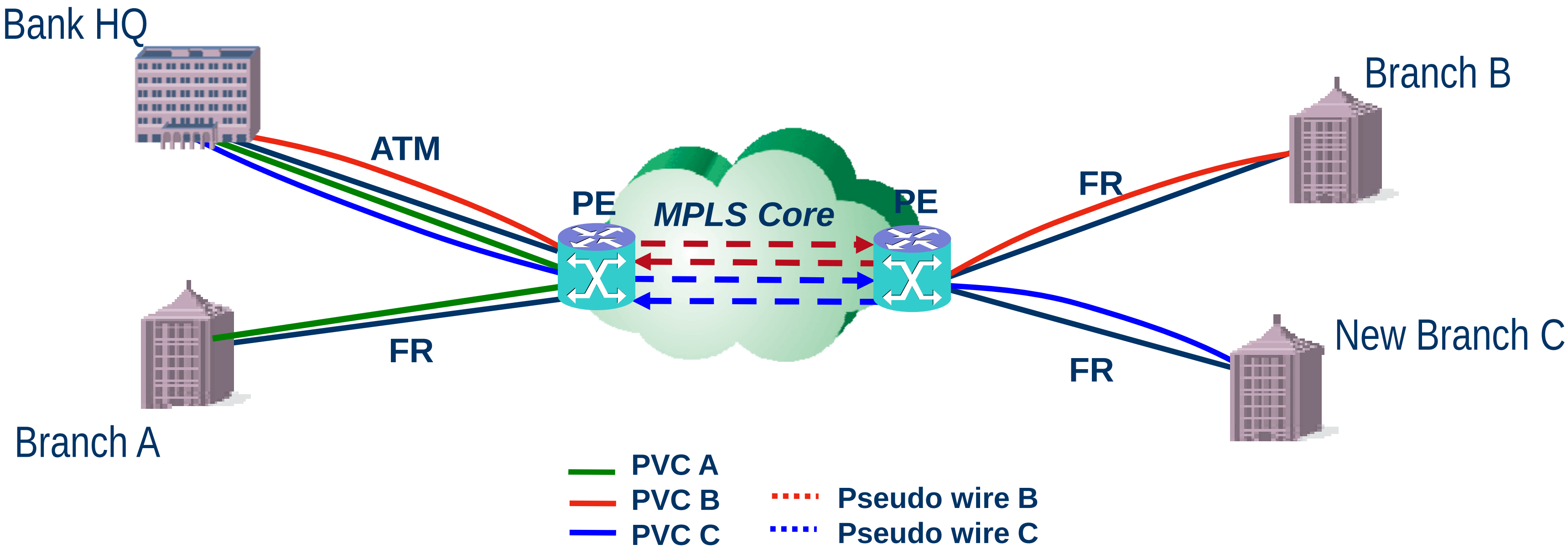
Enterprise Network Today



FRF.8.2 Service interworking is a key enabler

- Connecting branch offices with low-speed FR access to the Headquarter with a high-speed ATM connection

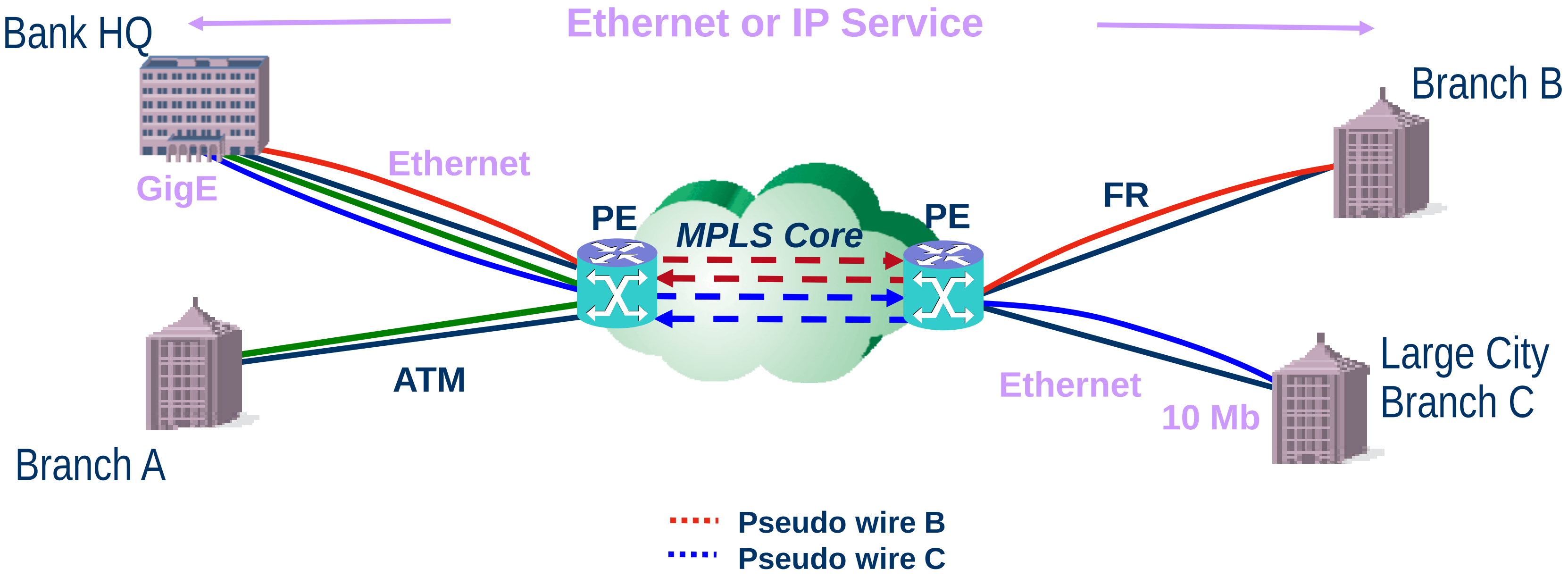
Network Migration Scenario 1:
- ATM/FR Interworking over MPLS



Enables graceful traffic migration from ATM to MPLS core

- Preserves existing ATM and FR service SLAs and revenues
- Transparent to Enterprise
- Enables service provider MPLS network investment for new FR/ATM endpoints

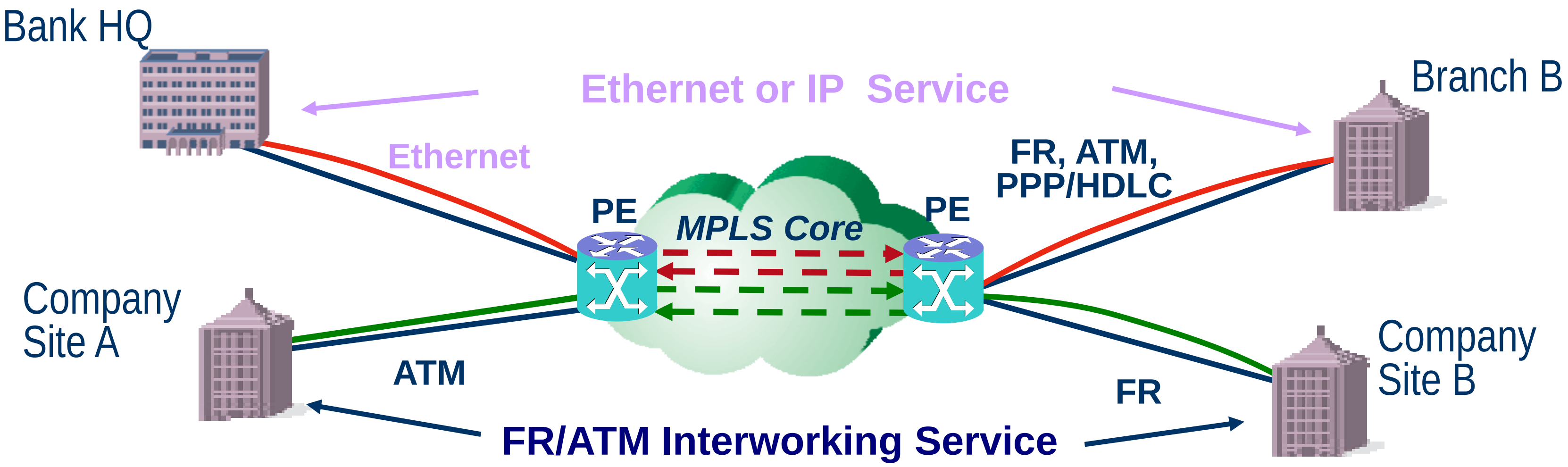
Network Migration Scenario 2:
- Ethernet or IP Interworking over MPLS



Introduce Ethernet connectivity to existing ATM/FR infrastructure

- Cost effectively scale bandwidth at select sites to support new business applications
- Graceful migration of legacy ATM/FR service to Ethernet services
- Ethernet and IP pt-pt *(shown)* and multipoint (Ethernet only) VPN services

Benefits of Multi-Service Interworking over MPLS



Carrier Benefits

- Increases addressable market
- Lowers capital expenses
- Increases flexibility
- Preserves revenues from legacy services

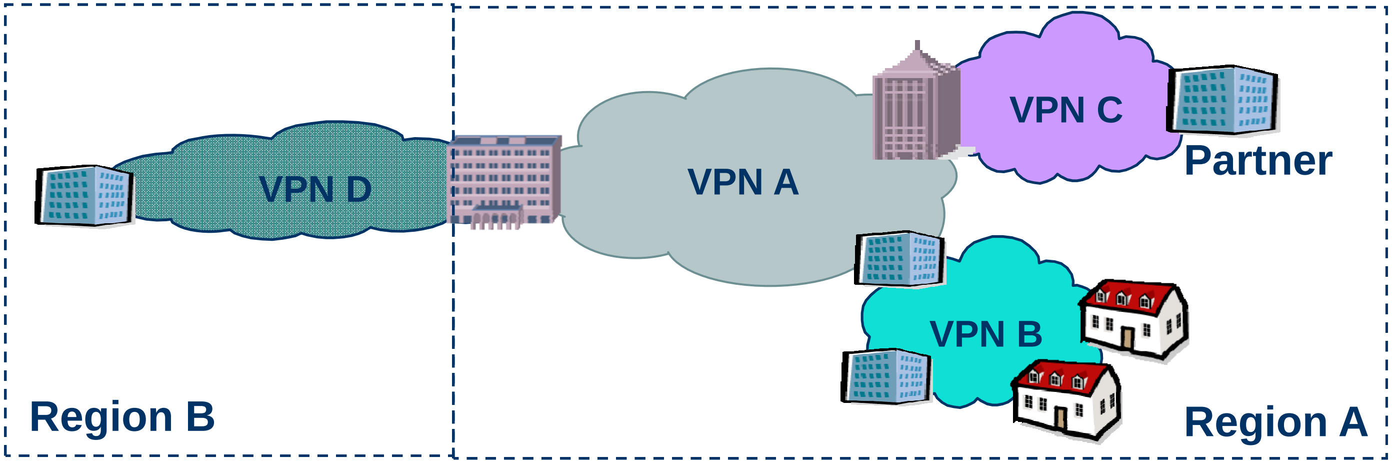
Enterprise Benefits

- Cost effectively scale bandwidth to support new applications
- Flexible support for sites with different access technologies
- Seamless integration of new sites on to network

Enables a smooth, cost effective evolution for both Enterprises and Carriers to new services

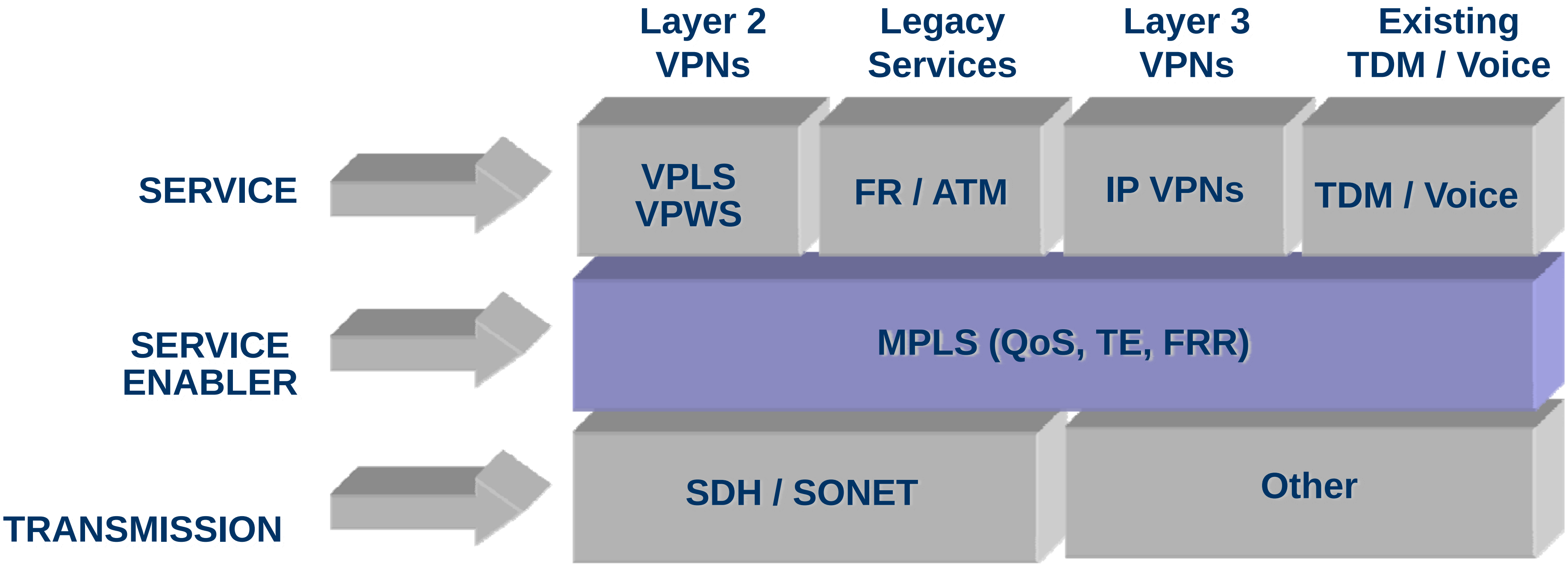
Summary

MPLS VPNs Summary



- Layer 2 and Layer 3 VPNs each address specific needs (traffic types, business applications, CPE investment, level of Service Provider participation in routing, etc)
- Both are standards based and widely deployed
- Solutions today include a combination of Layer 2 and Layer 3 VPNs

MPLS as a Service Enabler



VPLS = Virtual Private LAN Services

VPWS = Virtual Private Wire Services

L3 IP VPN = BGP/MPLS VPN RFC4364

For More Information. . .

- http://www.ipmplsforum.org
- http://www.ietf.org
- http://www.itu.int
- http://www.mplsrc.com

For questions, utilize the IP/MPLS Forum Message Board

Website: http://www.ipmplsforum.org/board/

Thank you **for attending the**

MPLS L2/L3
Virtual Private Networks Tutorial

Please visit the IP/MPLS Forum Booth in the Exhibit Area