MPLS: The Services Enabler
AfNOG 2006 Tutorial

May 14, 2006

Nairobi, Kenya

By: Tamrat Yossef (tyossef@cisco.com)

Agenda

- Why MPLS?
- Label Distribution Protocol (LDP)
- MPLS Layer 3 Virtual Private Network (L3VPN)
- MPLS Layer 2 VPN (L2VPN)
- MPLS Traffic Engineering (TE)
- MPLS Quality of Service (QoS)
- MPLS Security
- MPLS Management
- Lab

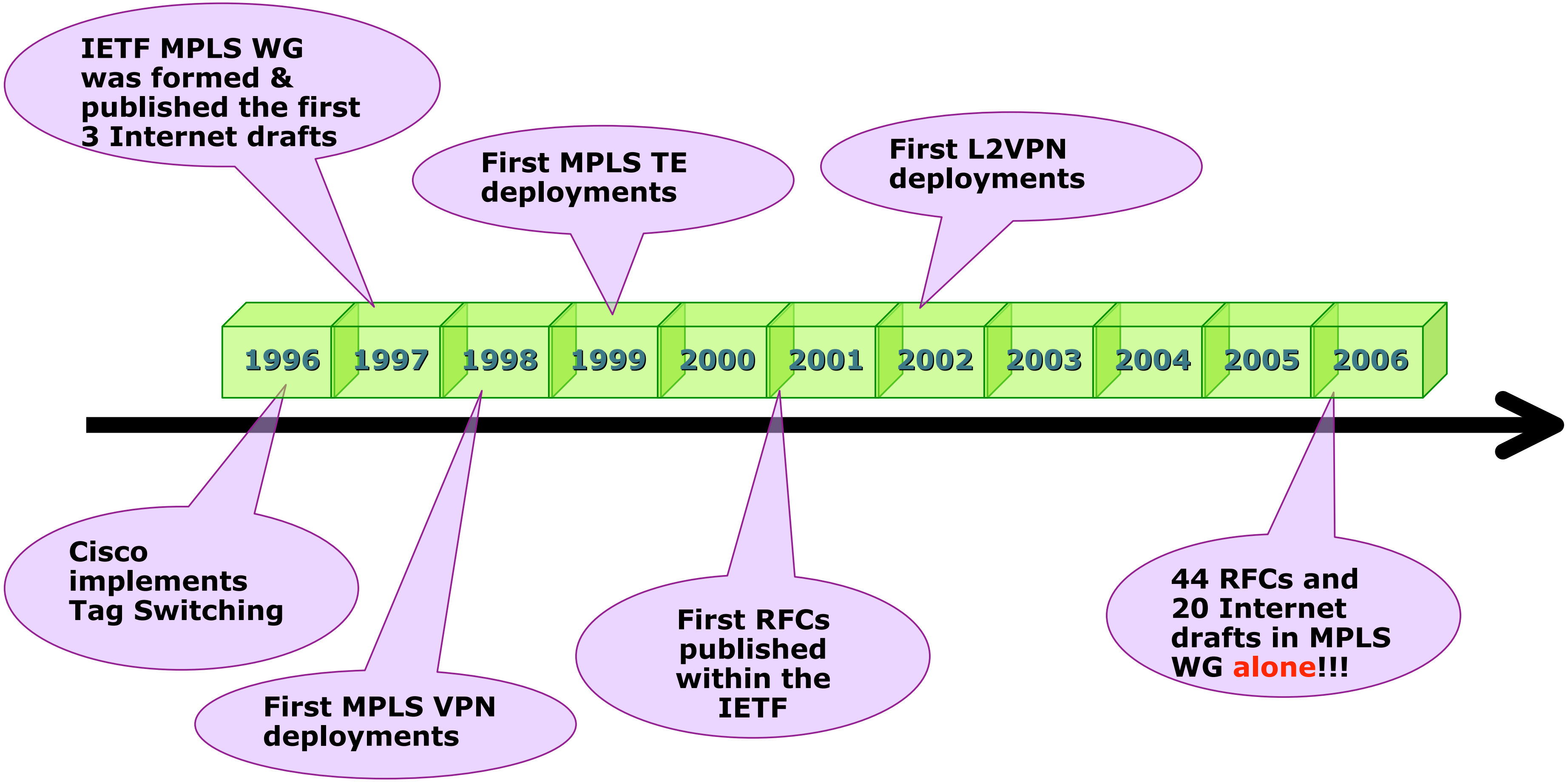
Acknowledgement

- **Muhammad Sagheer (Waris) – too bad he was not able to make the trip ☹**

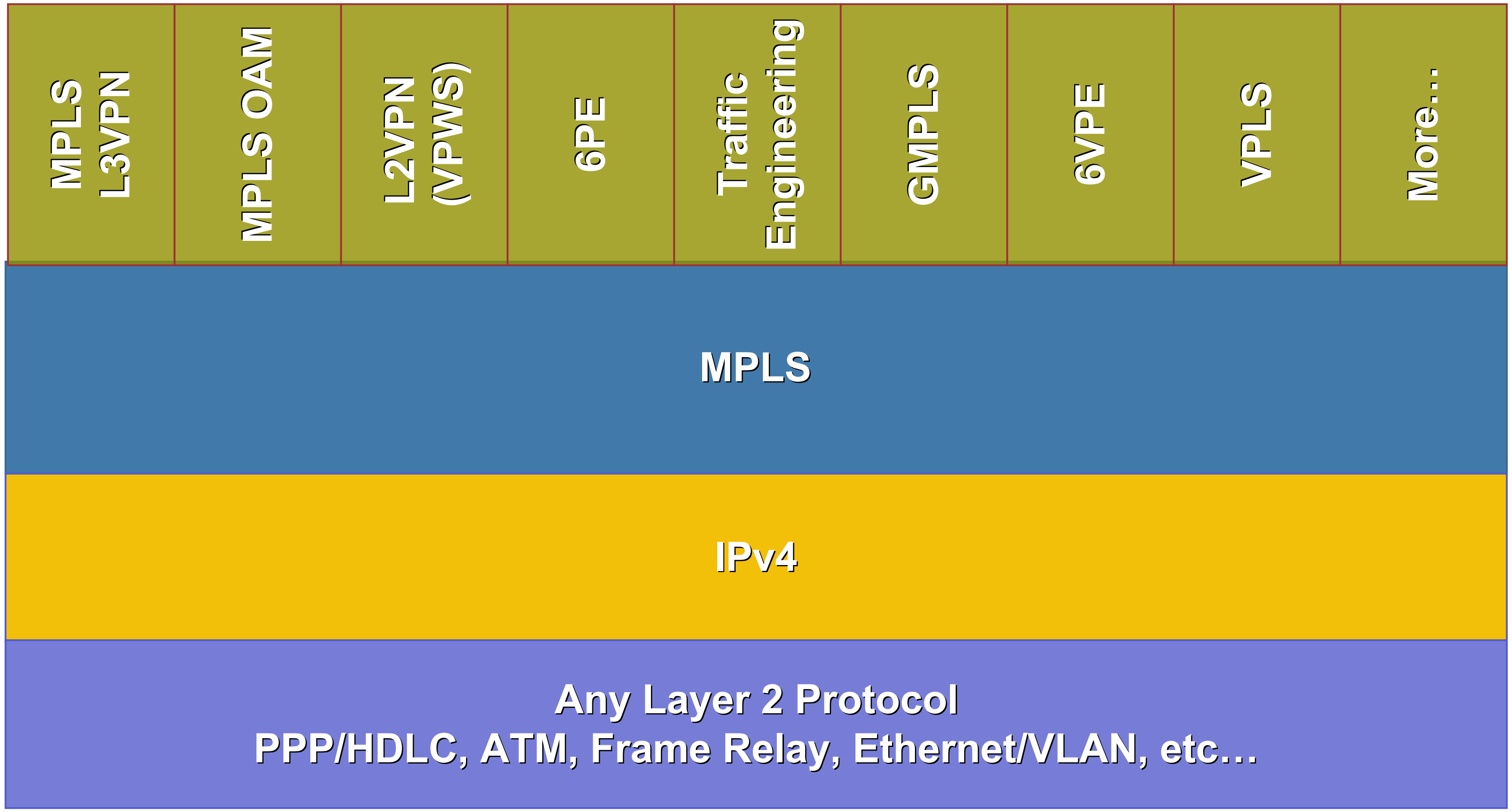
Why MPLS? A brief history

- MPLS stands for Multi Protocol Label Switching. The motivation behind MPLS was to emulate the speed of Layer 2 switching (in particular ATM) for Layer 3 forwarding.
- As the name suggests, it was designed to support several Layer 3 protocols; however, IPv4 is the only protocol that MPLS is used for today.
- MPLS evolved from Cisco's "Tag Switching", IBM's "ARIS", and Toshiba's "Cell-Switched Router” in the mid 1990s.

Why MPLS? A brief history


IETF MPLS WG was formed & published the first 3 Internet drafts
First MPLS TE deployments
First L2VPN deployments
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
Cisco implements Tag Switching
First MPLS VPN deployments
First RFCs published within the IETF
44 RFCs and 20 Internet drafts in MPLS WG alone!!!

MPLS Services



MPLS L3VPN

MPLS OAM

L2VPN (VPWS)

6PE

Traffic Engineering

GMPLS

6VPE

VPLS

More…

MPLS

IPv4

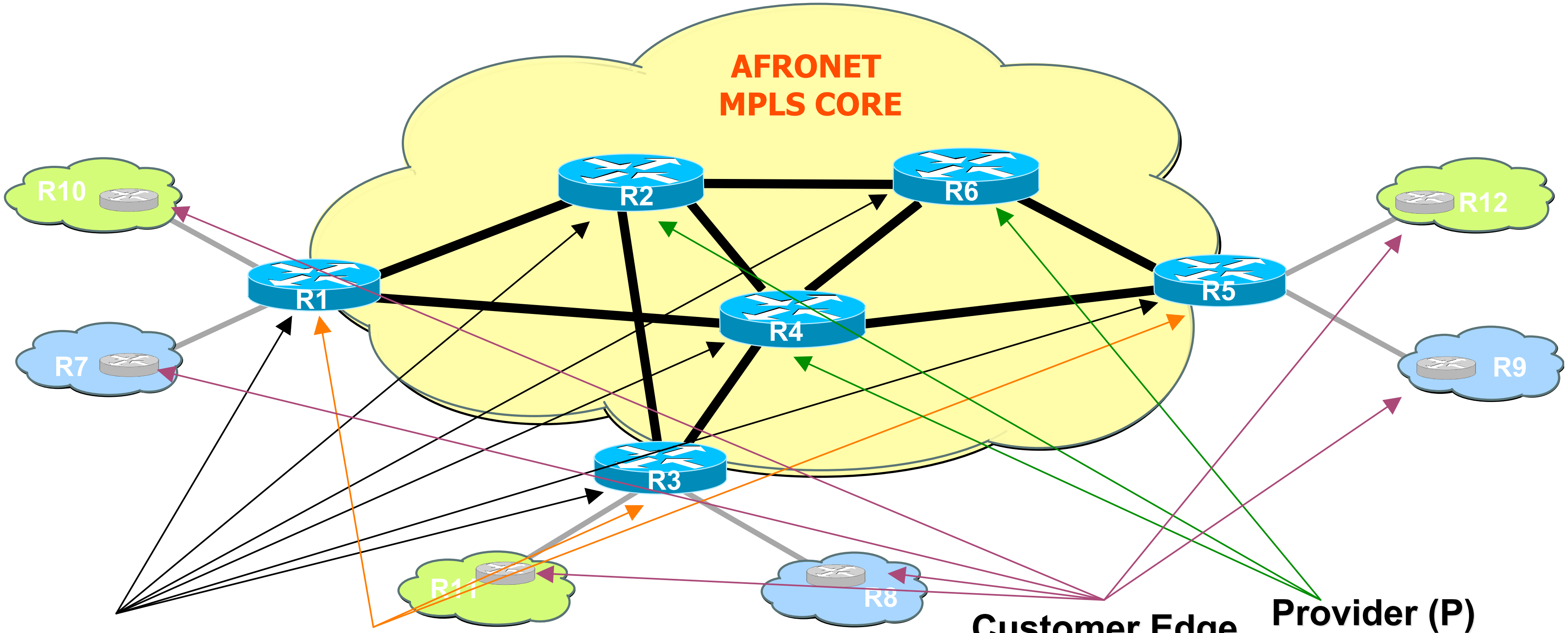
Any Layer 2 Protocol
PPP/HDLC, ATM, Frame Relay, Ethernet/VLAN, etc…

Standards documents

Some of the most important IETF standards track RFCs are:

Base MPLS	– **RFC 3031: Multiprotocol Label Switching Architecture** – **RFC 3032: MPLS Label Stack Encoding** – **RFC 3036: LDP Specification**
L3VPN	– **RFC 4364: BGP/MPLS IP Virtual Private Networks (VPNs) Obsoletes RFC2547** – **RFC 3107: Carrying Label Information in BGP-4**
L2VPN	– **RFC 3985: Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture**
MPLS TE	– **RFC 3209: RSVP-TE: Extensions to RSVP for LSP Tunnel**

MPLS Terminology



LSR (Label Switch Router) - A router that is participating in MPLS operations

LER (Label Edge Router) – A router at the edge of an MPLS network. Also known as a Provider Edge (PE) router.

Customer Edge (CE) router - A router that connects to MPLS PE router. Is not part of the MPLS network.

Provider (P) router – A router that is exclusively in the MPLS network. Connects to other P routers or PE routers.

What is a label?

According to RFC3031:

“A label is a short, fixed length, locally significant identifier which is used to identify a FEC. The label which is put on a particular packet represents the Forwarding Equivalence Class to which that packet is assigned.”

What is a label?

00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Label value																				EXP			S	TTL							

Four octets long string with the following four components:

Label (20 bits): Value range 0 – 1048575
Label 0 – 15 are reserved and have special meaning

EXP (3 bits): Experimental bits used for QoS purposes in the same fashion as IP Precedence.

S (1 bit): Bottom of Stack indicator. The bottommost label has the S bit set to 1. Simplifies the parsing of labels by routers.

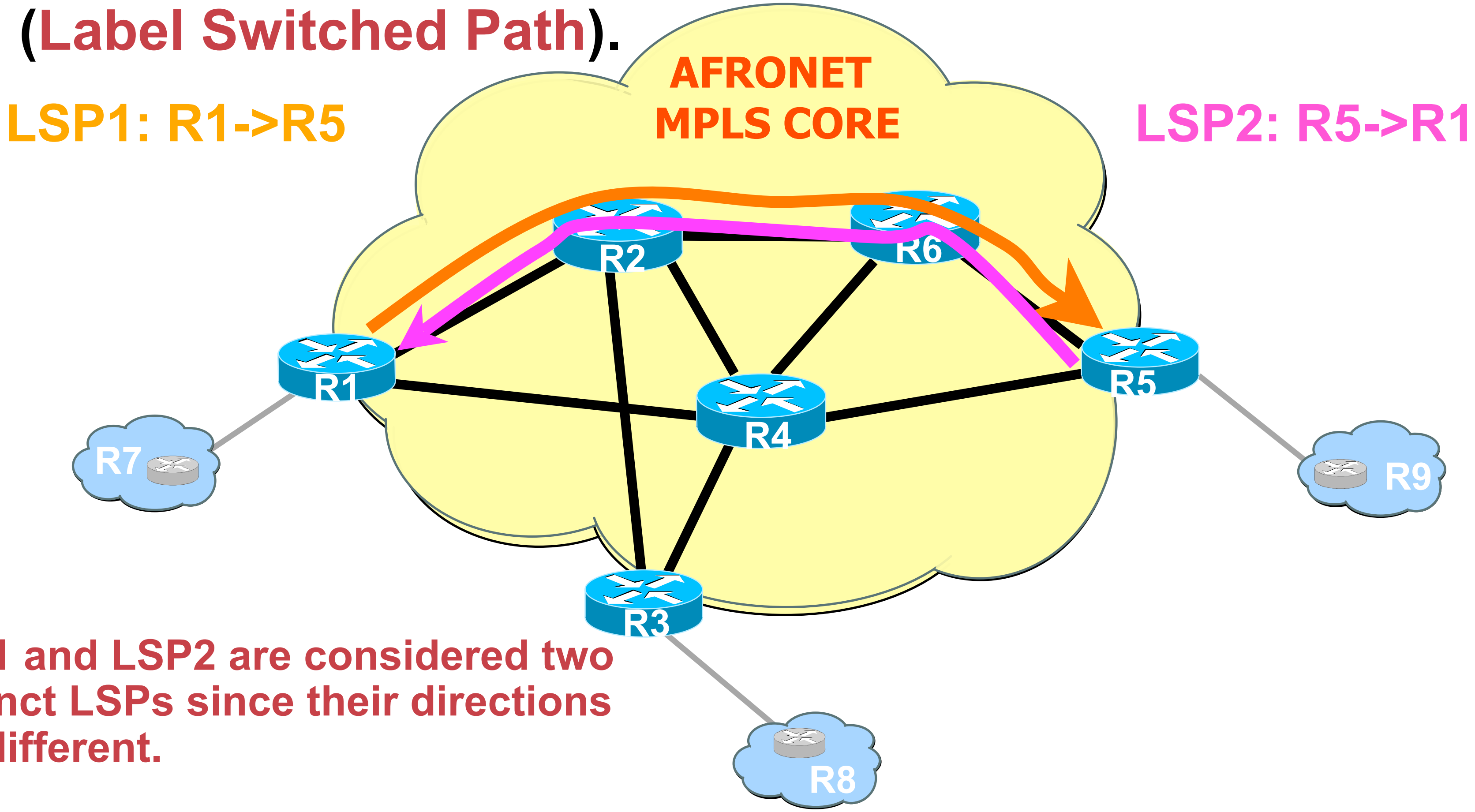
TTL (8 bits): Time to Live. Used to limit the scope of labeled packets, in case of routing loops.

What is a FEC?

- A FEC (Forwarding Equivalence Class) is a group of IP packets which are forwarded in the same manner, over the same path, and with the same forwarding treatment.
- Packet headers contain considerably more information than is needed simply to choose the next hop. Choosing the next hop can therefore be thought of as the composition of two functions.
  - The first function partitions the entire set of possible packets into a set of "Forwarding Equivalence Classes (FECs)".
  - The second maps each FEC to a next hop.
- Each FEC is assigned a label.

What is an LSP?

- A uni-directional path through an MPLS network from ingress to egress router corresponds to a LSP (Label Switched Path).



LSP1 and LSP2 are considered two distinct LSPs since their directions are different.

MPLS Modes

MPLS can operate in one of two modes:

- frame mode

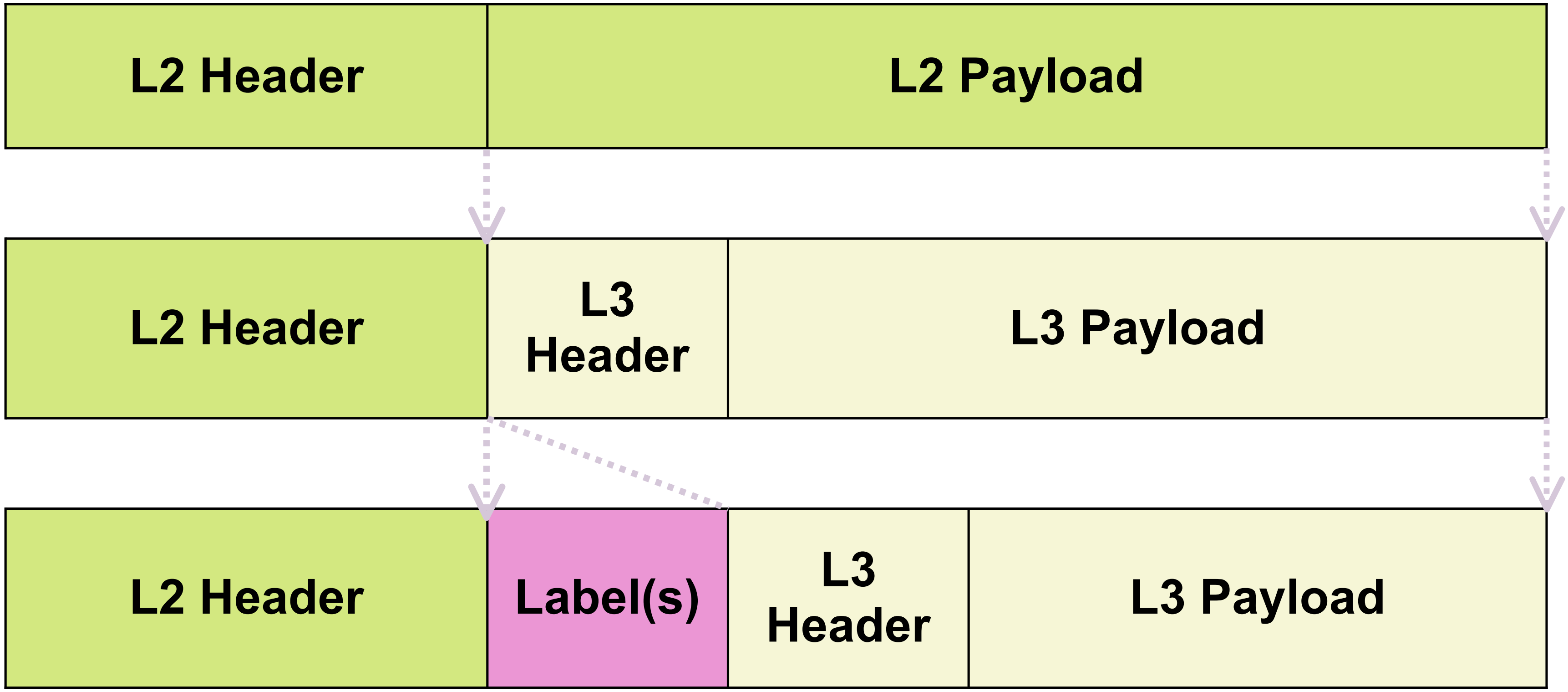
  In this mode, the incoming packet has one or more labels inserted just before the Layer3 header. That is why a label is sometimes referred to as *shim header*.

- Cell mode

  In ATM networks running MPLS, the label is encoded within the VPI/VCI fields of the ATM cell header.

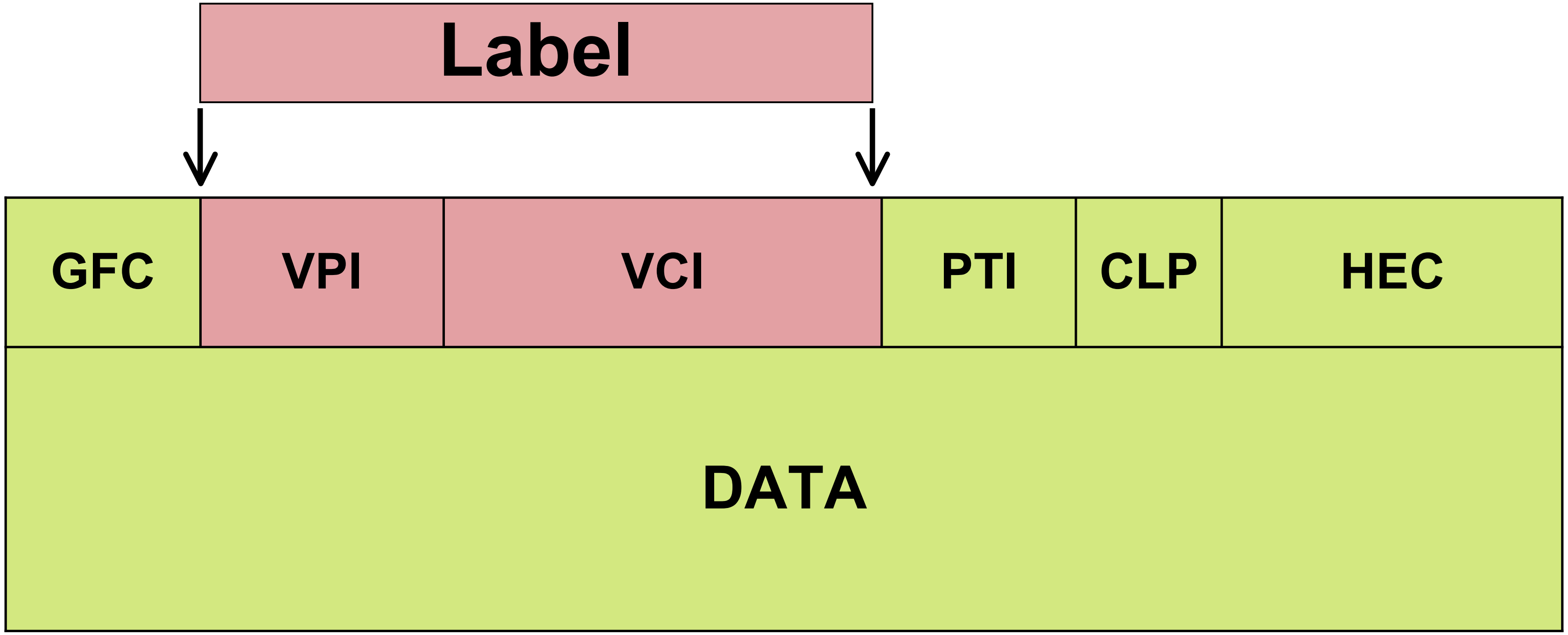
Frame mode is by far the most popular mode, thus all subsequent discussion will be based on frame mode.

MPLS Frame Mode



Label is added between the L2 and L3 headers (also called a ‘shim’ header)

MPLS Cell Mode



Label is encoded in the ATM VPI/VCI.

Other MPLS concepts

The following MPLS concepts will be covered:

- **LSR location**

  upstream vs. downstream

- **Distribution mode**

  unsolicited distribution vs. downstream on demand

- **Label retention**

  liberal vs. conservative label retention

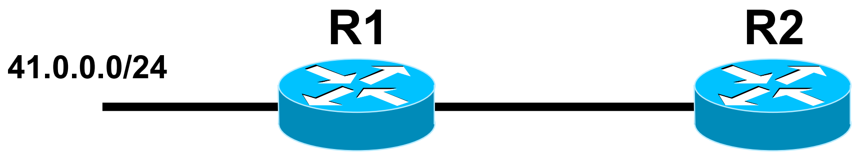
- **Label space**

  per platform vs. per interface

Upstream vs. Downstream

Upstream and downstream is relative to a reference router.

- **An upstream MPLS router that is one that is closer to the source of a packet, relative to another MPLS router.**
- **A downstream MPLS router is one that is farther from the source of a packet, relative to another MPLS router.**



In the example above, for prefix 41.0.0.0/24, R1 is downstream relative to R2. Conversely, R2 is upstream relative to R1.

Label distribution modes

Label distribution can operate in one of two modes:

- Unsolicited Downstream (UD)

  UD is used when an LSR sends label binding for a FEC to its downstream neighbor, whether the downstream LSR requested the binding or not.

- Downstream on demand (DoD)

  When an LSR explicitly requests a neighboring LSR for the label binding for a particular FEC.

Most routers use Unsolicited Downstream mode.

Label retention modes

Label retention can be in one of two modes:

– Liberal

If a LSR R1 receives label binding for a FEC from a downstream LSR R2, even though the R2 is not the next hop for the FEC, R1 will still retain the label binding information.

– Conservative

In this case, R1 will discard the label binding it has received from R2.

Liberal retention mode is the most common one of the two.

Label space

The MPLS label space can have the following scopes:

– per platform

The MPLS labels are unique on a platform basis. For example, if FEC1 is assigned label 16, that label 16 cannot be used by any other FEC.

– per interface

In this scenario, each interface running MPLS will have its own label space. In the same example as above, FEC1 on interface1 has label 16, but FEC2 on interface2 can also have a label 16, since the label spaces are unique, on a per interface basis.

Per platform label space is what is commonly on most routers. Cell mode MPLS uses per interface label space.

Label Distribution

- **MPLS labels are distributed via:**
  - **Label Distribution Protocol (LDP)**
  - **Multi Protocol BGP (MP-BGP) when using MPLS VPN**
  - **Resource Reservation Protocol (RSVP) when Traffic Engineering is used**

Label Distribution Protocol

LDP

- LDP, as its name implies, is designed to do label distribution for MPLS.
- LDP is the set of procedures and messages by which Label Switched Routers (LSRs) establish Label Switched Paths (LSPs) through a network by mapping network-layer routing information directly to data-link layer switched paths.
- LDP has four major functions:
  - Neighbor discovery
  - Session establishment and maintenance
  - label advertisement
  - notification

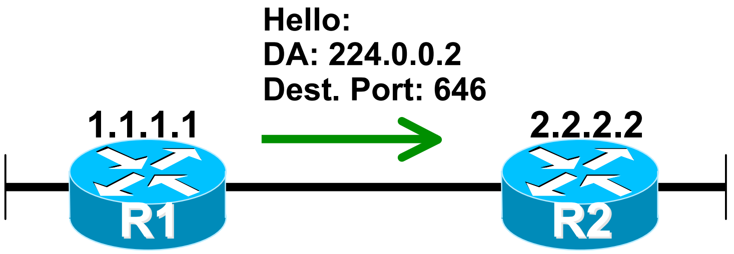
LDP neighbor discovery

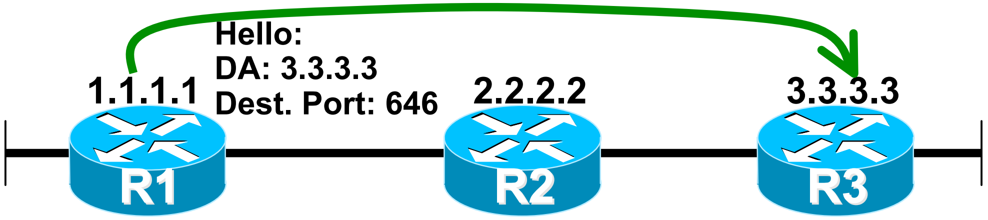
- LDP can have two types of neighbors:
  1. directly connected neighbor

     LDP uses UDP hello messages sent to port 646 over the all routers Multicast address (224.0.0.2) to discover directly connected neighbors.

  2. non-directly connected neighbor

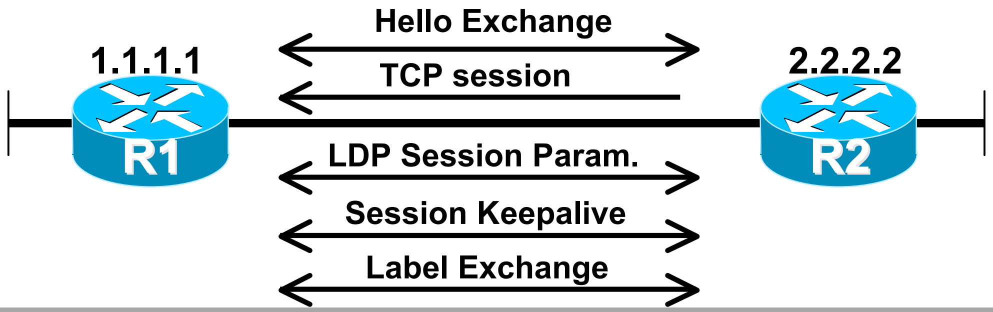
     LDP has the ability to establish what is called a directed (or targetted) LDP session with a router two or more hops away. Hellos in this case are unicast to the peer router using UDP port 646.





LDP Session establishment & maintenance

- Once both routers send and receive hello messages, a Hello Adjacency is established.
- The router with the highest LDP ID becomes the active router (other becomes passive). Active LSR initiates a TCP session over port 646.
- Once TCP is established, LDP is initialized by exchanging session parameters, such as version number, label distribution method, timer values, etc..
- When both routers are able to exchange Keepalive messages, the LDP session goes into Established state.
- Session is maintained by periodic exchange of Keepalive messages.
- Once the LDP session is established, LSRs start exchanging label bindings with one another.



LDP Label Advertisement

- There exists two Label Distribution Control Modes:
  - Independent LSP control mode

    LSR will freely distribute label bindings to all its neighbors
  - Ordered LSP control mode

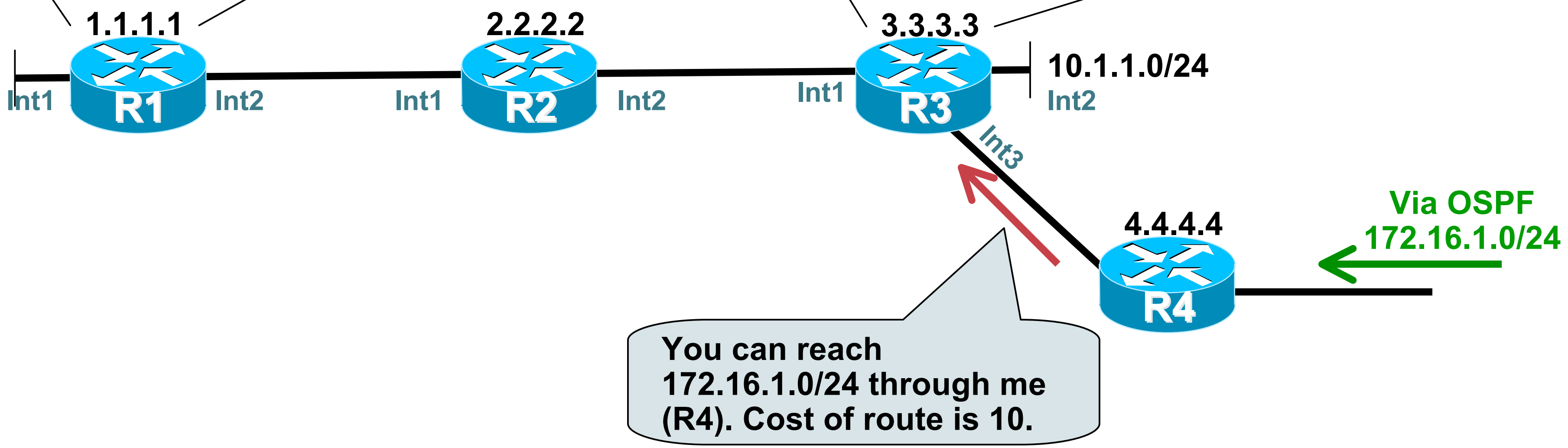
    LSR will wait to receive label bindings from its downstream neighbor(s) before sending its own bindings to its upstream neighbor(s).
- Independent mode is used for frame mode MPLS, whereas ordered mode is used in cell mode MPLS.
- In our case, Independent mode is what we will see used.

LDP label assignment example

How does the routing table on R1 look?

Route	Outgoing Int	Cost	Next Hop
172.16.1.0/24	Int2	30	R2
10.1.1.0/24	Int2	16	R2

Route	Outgoing Int	Cost	Next Hop
172.16.1.0/24	Int3	15	R4
10.1.1.0/24	Int2	1	Directly Connected



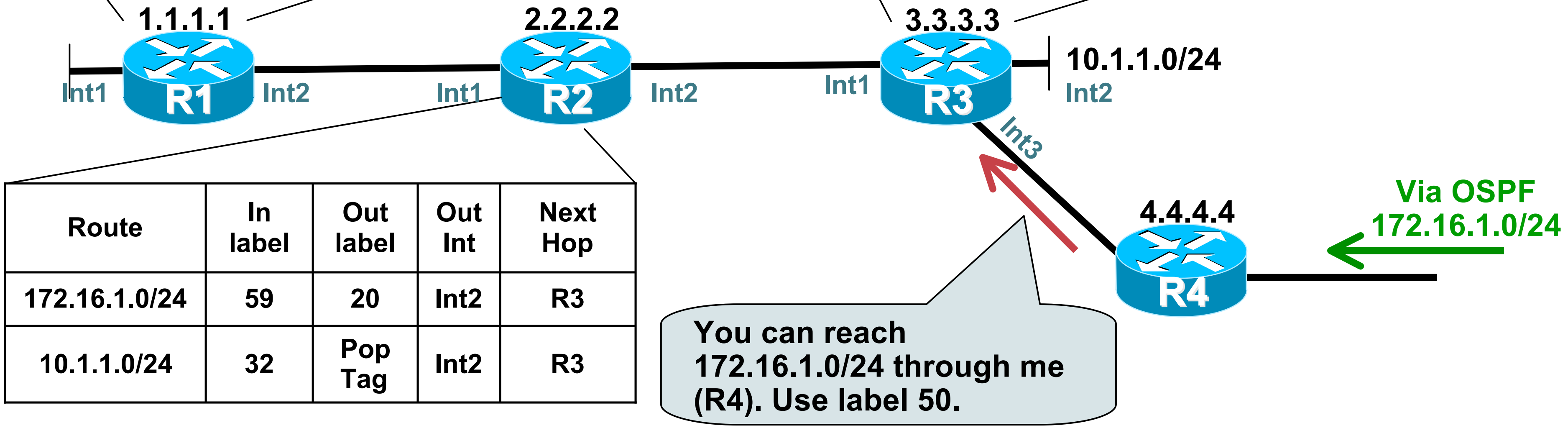
You can reach 172.16.1.0/24 through me (R4). Cost of route is 10.

LDP label assignment example

How does the label assignment look like?

Route	In label	Out label	Outgoing Int	Next Hop
172.16.1.0/24	-	59	Int2	R2
10.1.1.0/24	-	32	Int2	R2

Route	In label	Out label	Out Int	Next Hop
172.16.1.0/24	20	50	Int3	R4
10.1.1.0/24	Untagged	-	Int2	Directly Connected



You can reach 172.16.1.0/24 through me (R4). Use label 50.

Route	In label	Out label	Out Int	Next Hop
172.16.1.0/24	59	20	Int2	R3
10.1.1.0/24	32	Pop Tag	Int2	R3

LDP Configuration steps

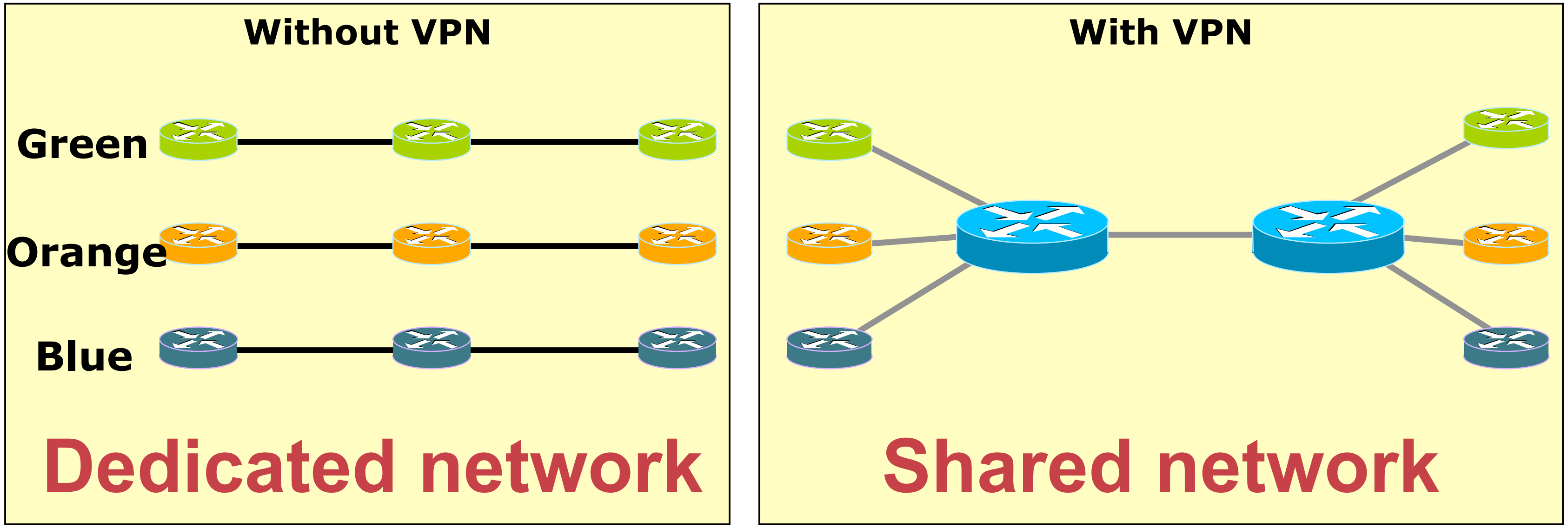
	Router# **configure terminal**	**Enables configuration mode**
Step 1	Router(config)# **ip cef** [**distributed**]	**Pre-requisite: Configures Cisco Express Forwarding**
Step 2	Router(config)# **mpls ip**	**Enable MPLS globally**
Step 3	Router(config)# **mpls label protocol ldp**	**Set the label distribution protocol to LDP (TDP is the default)**
Step 4	Router(config-if)# **mpls ldp router-id loopback 0 force**	**Set the LDP router ID to use IP address of one of your logical interfaces**
Step 5	Router(config)# **interface** *interface*	**Specifies the interface to configure**
Step 6	Router(config-if)# **mpls ip**	**Enable MPLS forwarding on the interface. Repeat for all interfaces that should be running MPLS.**

Repeat these simple steps on all LSRs on your network.

MPLS VPN

VPN Concept

- A VPN, as the name suggests, is a mechanism that creates a private network on a public (shared) network.



- This is a generic VPN scenario and it is not new: ATM, Frame Relay, VLANs use this concept at Layer 2.

MPLS VPN vs. traditional VPN

Overlay vs. peer to peer models

- Overlay network

  A customer IP network is overlaid on top of the provider network. The Provider network consists of Layer 2 (circuit based) network.

  For enhanced scalability, hub and spoke is usually used

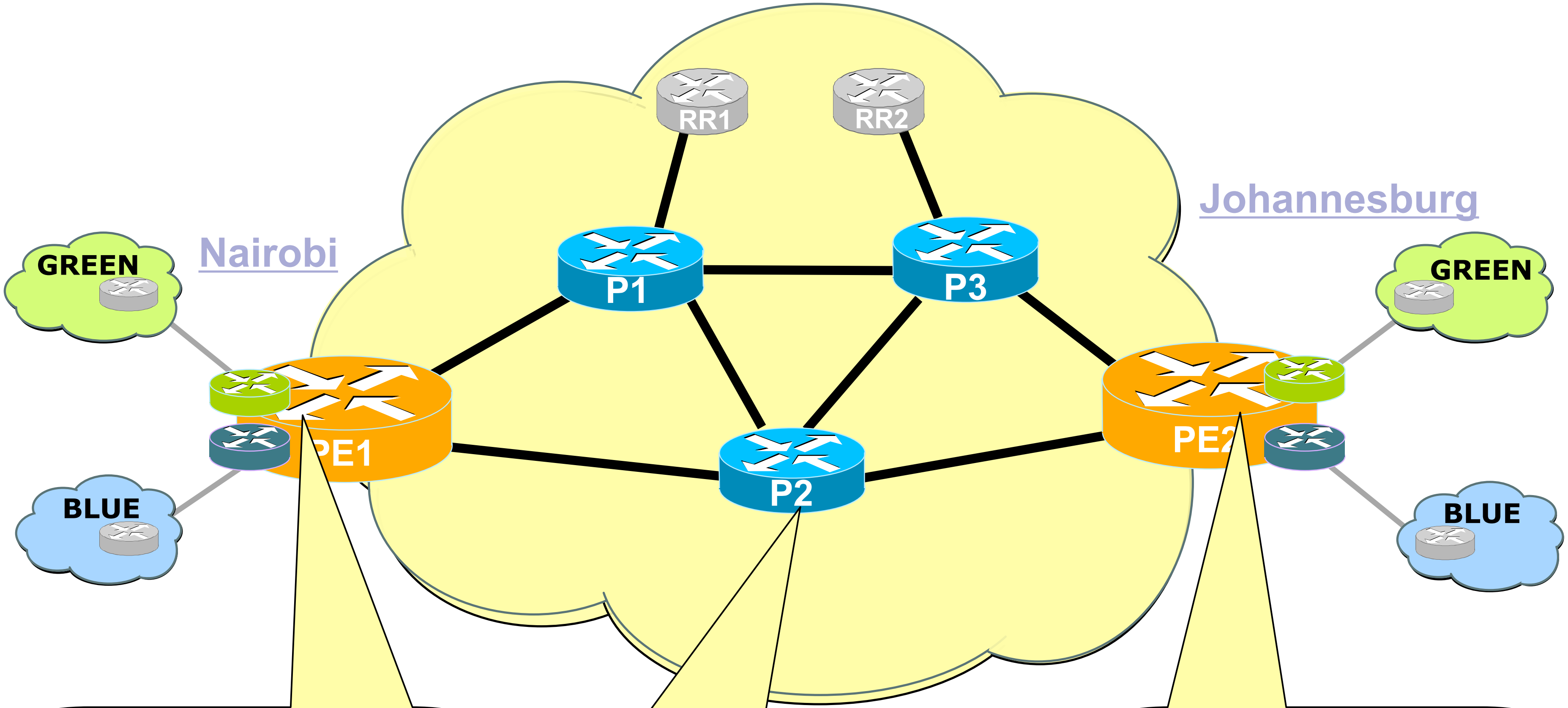
- Peer Network

  Provider and customer exchange IP routing information directly.

  Customer only has one routing peer per site.

MPLS VPN combines the benefits of both models

MPLS VPN



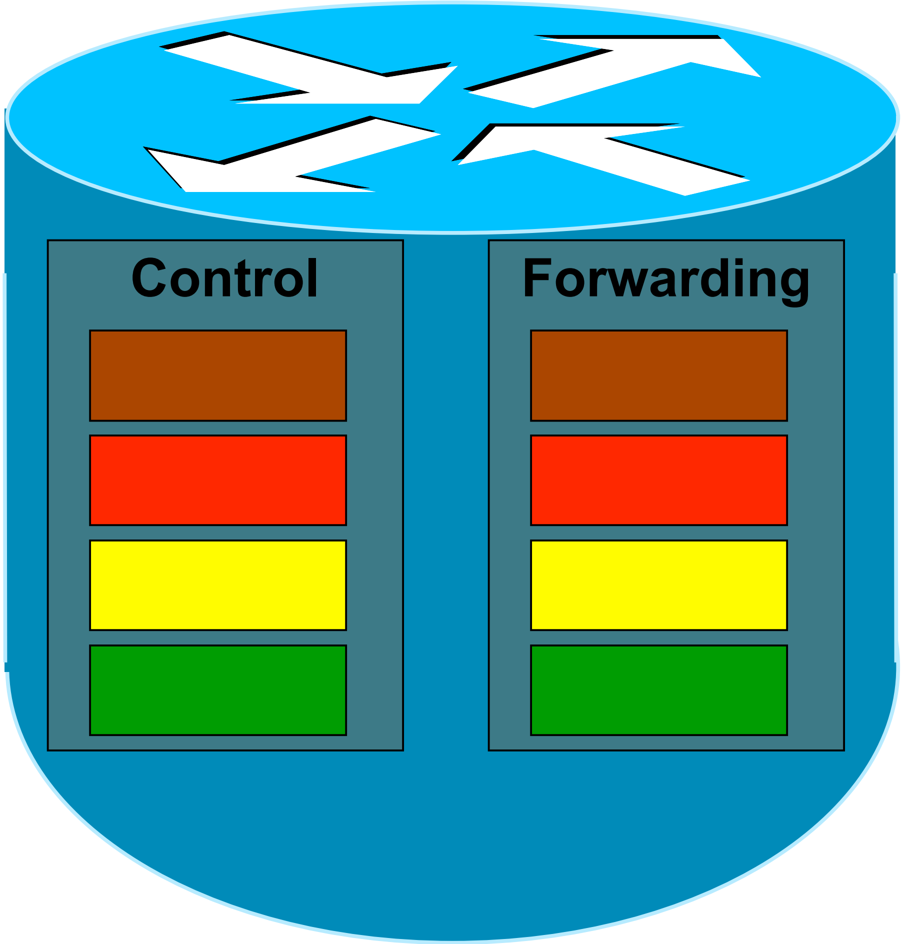
Each VRF emulates a separate router, within the PE router.

Core routers have no knowledge of any VPN. They switch the traffic transparently.

The routing and forwarding table within each VRF is not visible to other VRFs.

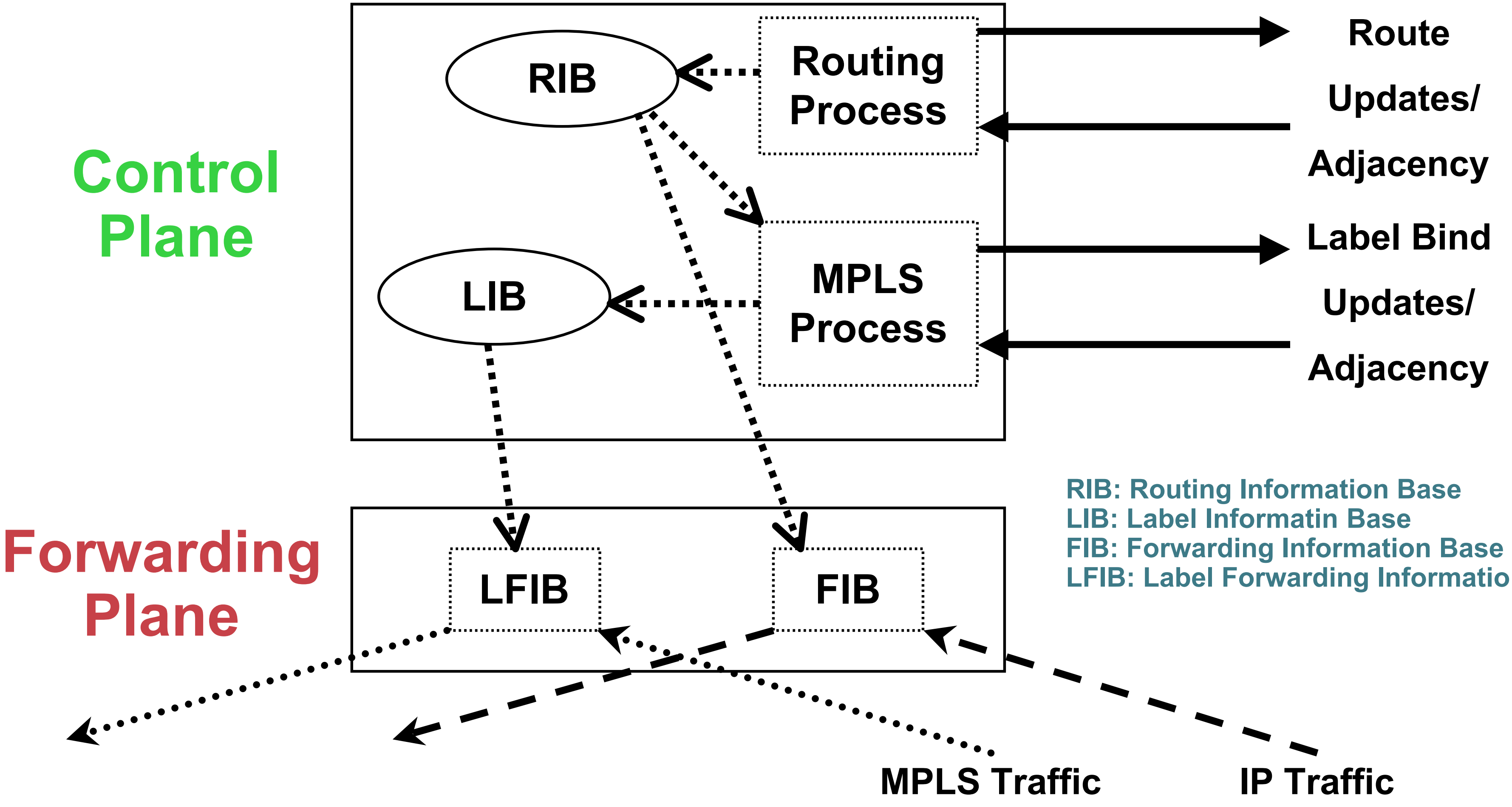
Control and Forward Planes

Each VPN has its own control (routing) and forwarding plane, independent of other VPNs, and independent of the global control and forwarding planes.



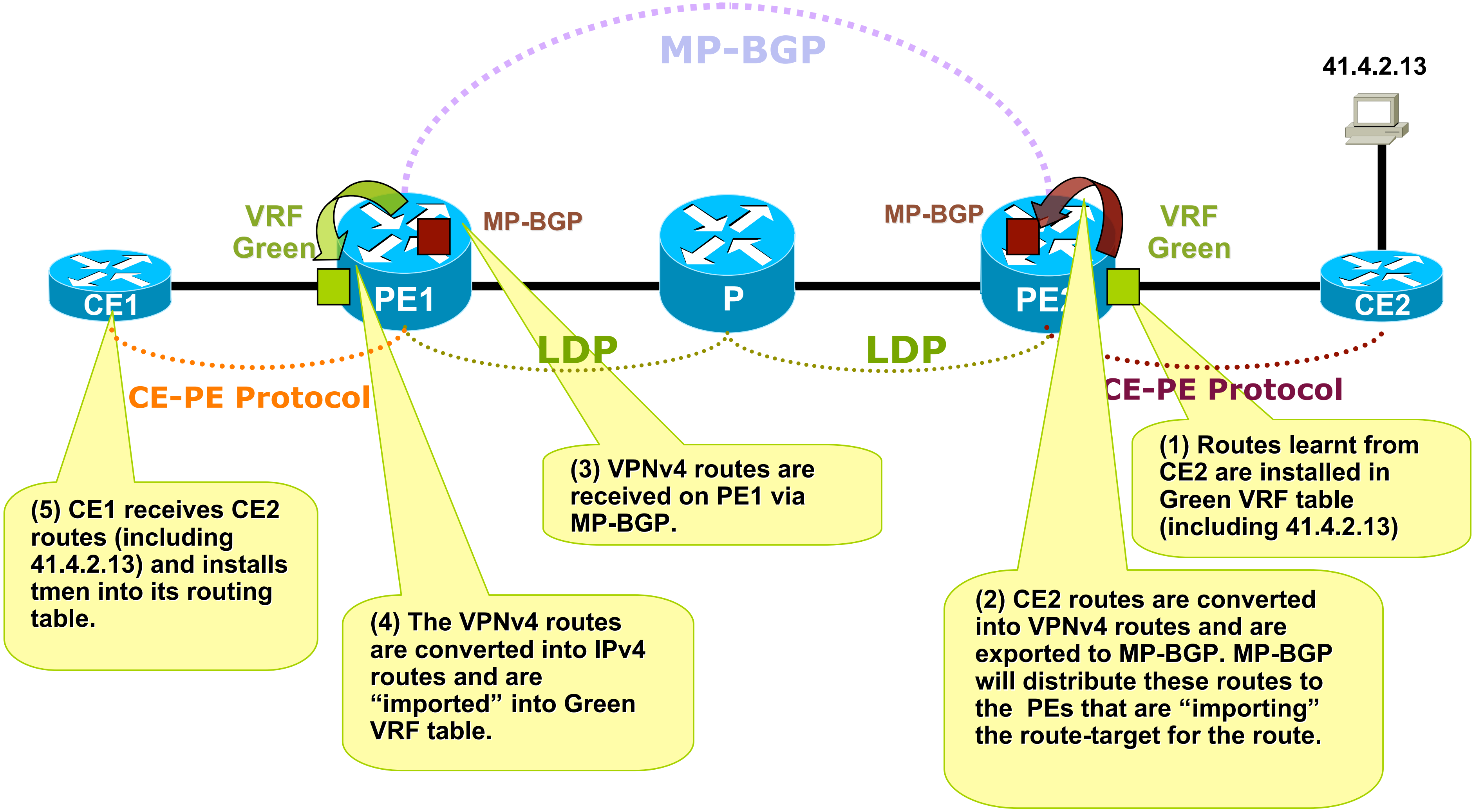
Global
VPN Red
VPN Yellow
VPN Green

Control and Forward Planes



RIB: Routing Information Base
LIB: Label Informatin Base
FIB: Forwarding Information Base
LFIB: Label Forwarding Information Base

VPN route propagation



(1) Routes learnt from CE2 are installed in Green VRF table (including 41.4.2.13)

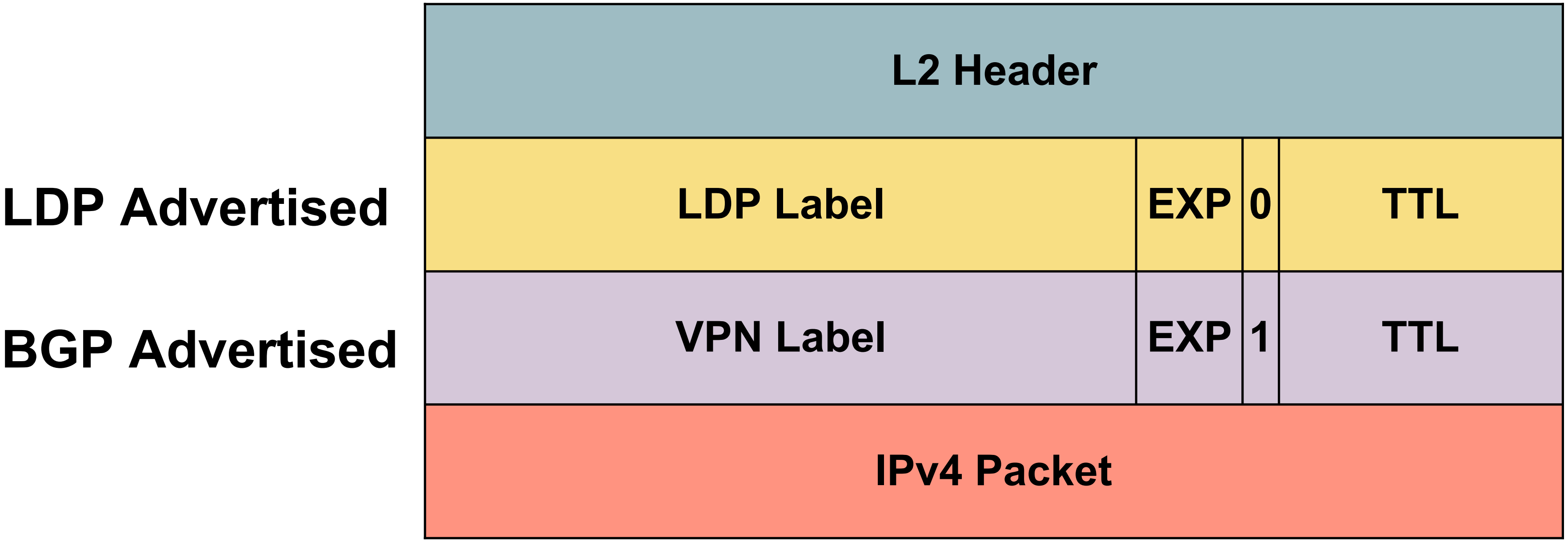
(2) CE2 routes are converted into VPNv4 routes and are exported to MP-BGP. MP-BGP will distribute these routes to the PEs that are “importing” the route-target for the route.

(3) VPNv4 routes are received on PE1 via MP-BGP.

(4) The VPNv4 routes are converted into IPv4 routes and are “imported” into Green VRF table.

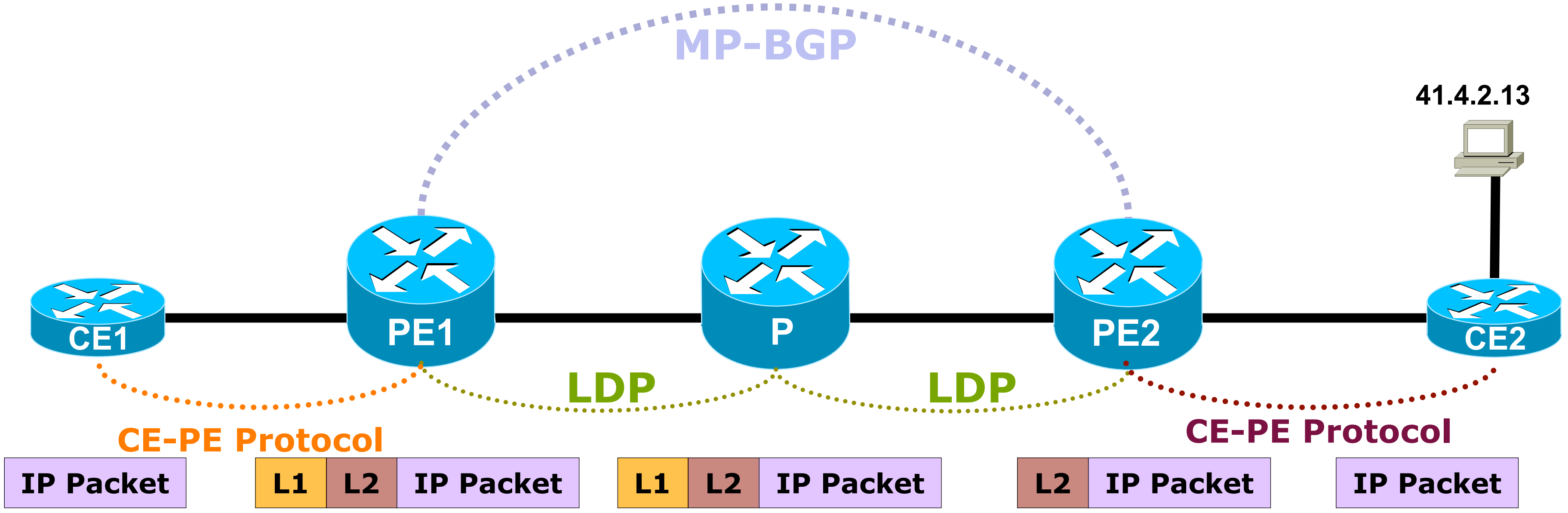
(5) CE1 receives CE2 routes (including 41.4.2.13) and installs tmen into its routing table.

MPLS VPN label stack



	L2 Header			
LDP Advertised	LDP Label	EXP	0	TTL
BGP Advertised	VPN Label	EXP	1	TTL
	IPv4 Packet			

MPLS VPN traffic forwarding



Step 1:
CE1 sends traffic to address 41.4.2.13 (which is being CE2)

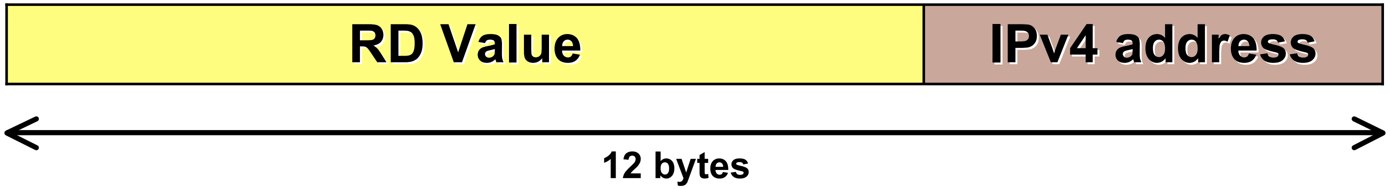
Step 2:
PE1 imposes a BGP/VPN label (L2) & an IGP/LDP label (L1) onto the received IP packet. L1 is learnt from P1; L2 from PE2.

Step 3:
P1 being the penultimate hop router, it pops the IGP/LDP label and sends the single labeled packet to PE2.

Step 4:
PE2 pops the L2 label, looks up the forwarding Table and sends the IP packet to CE2.

MPLS VPN terms

- **Route Distinguisher**
  - **identifies a set of VRFs**
  - **8 bytes long**
  - **Extended community in BGP.** (AFI=1, SAFI=128)
  - **when combined with an IPv4 address, it creates a 12 bytes (96 bits long VPNv4 address)**



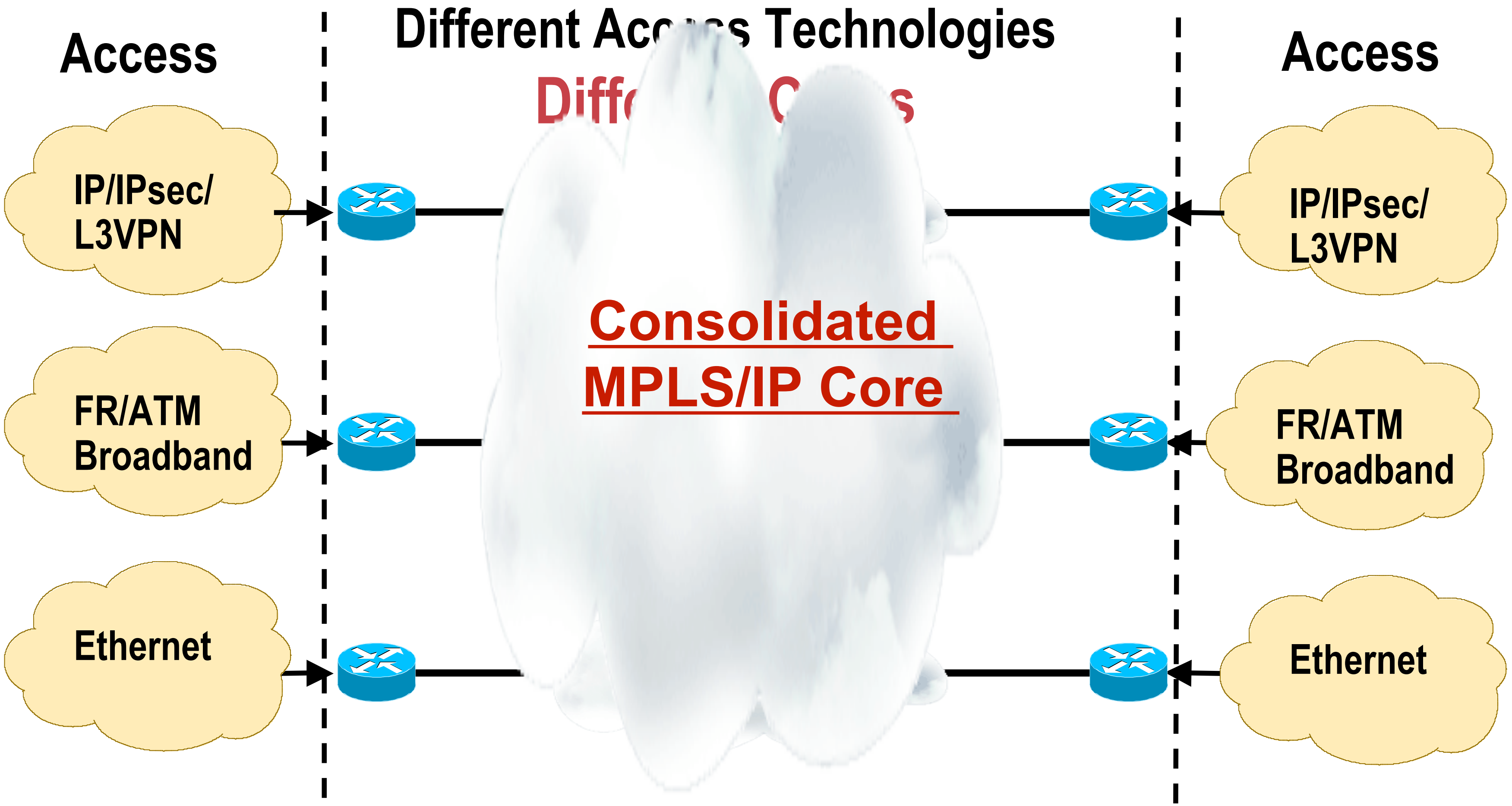
- **If a PE router receives an UPDATE message with an export target that matches the import target of any of its VRFs, the route is installed in that VRF.**
- **Intranet is when you import and export the same RT**
- **Extranet is when you import a different RT into a VRF.**

MPLS VPN Configuration steps

Step 1	Define and configure VRFs	ip vrf my_VRF
Step 2	Set the import and export policies for each VRF	rd 1:1 route-target both 1:1
Step 3	Assign interfaces to VRFs	interface Serial 1/0 ip vrf forwarding my_VRF ip address 10.1.1.1 255.255.255.0
Step 4	Configure MP-BGP for route distribution between PEs	router bgp 1 neighbor 2.2.2.2 remote-as 1 neighbor 2.2.2.2 update-source loopback 0 ! address-family vpnv4 neighbor 2.2.2.2 activate neighbor 2.2.2.2 send-community both
Step 5	Configure CE-PE routing	(Depends on the CE-PE protocol used – Example is with RIPv2) router rip version 2 ! address-family ipv4 vrf my_VRF version 2 network 10.1.1.0
Step 6	Mutually redistribute between iBGP and your CE-PE protocol used.	router bgp 1 address-family ipv4 vrf my_VRF redistribute connected redistribute rip ! router rip address-family ipv4 vrf my_VRF redistribute bgp 1 metric 1

Introduction to L2VPNs

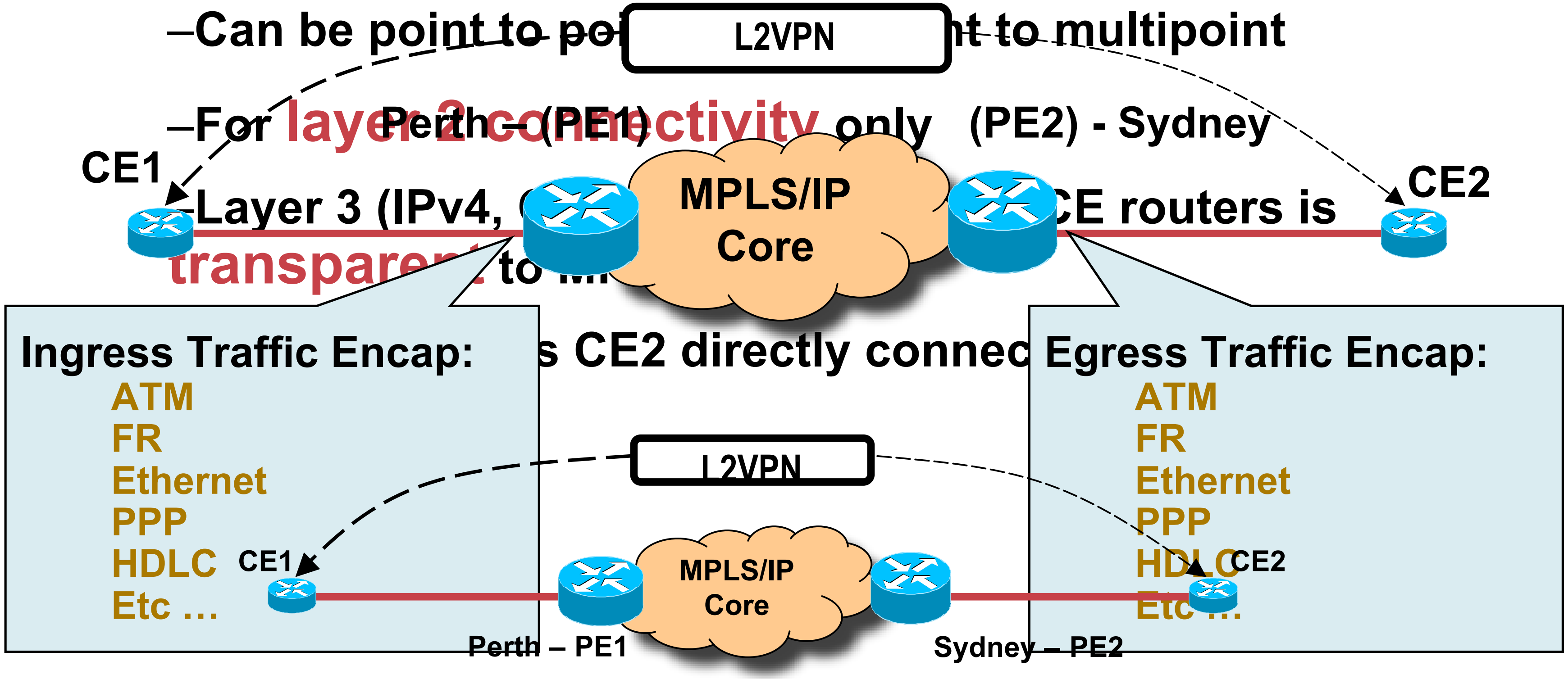
Motivation for L2VPNs – *Carrier Challenges at the Edge*



Multiple Access Services Require Multiple Core Technologies = $$$ High Costs / Complex Management

L2VPN - Simple definition

- L2VPN provides an end-to-end layer 2 connection to an enterprise office in Perth and Sydney over a SP's MPLS or IP core
  - –Can be point to poi nt to multipoint
  - –For layer 2 connectivity only
  - –Layer 3 (IPv4, CE routers is transparent to M
  - s CE2 directly connec



Ingress Traffic Encap:
ATM
FR
Ethernet
PPP
HDLC
Etc …

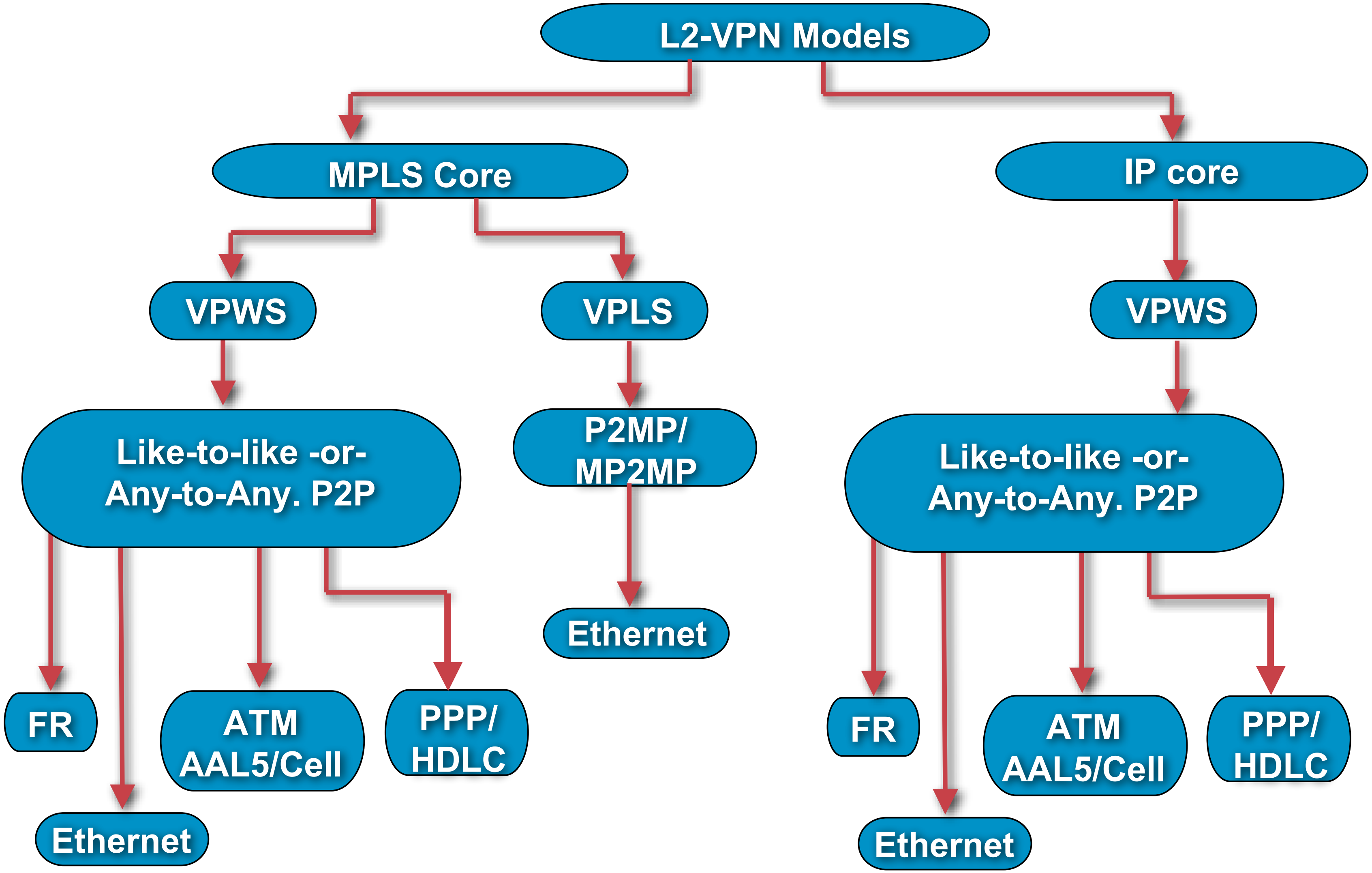
Egress Traffic Encap:
ATM
FR
Ethernet
PPP
HDLC
Etc …

L3 and L2 VPN Characteristics
Customer View

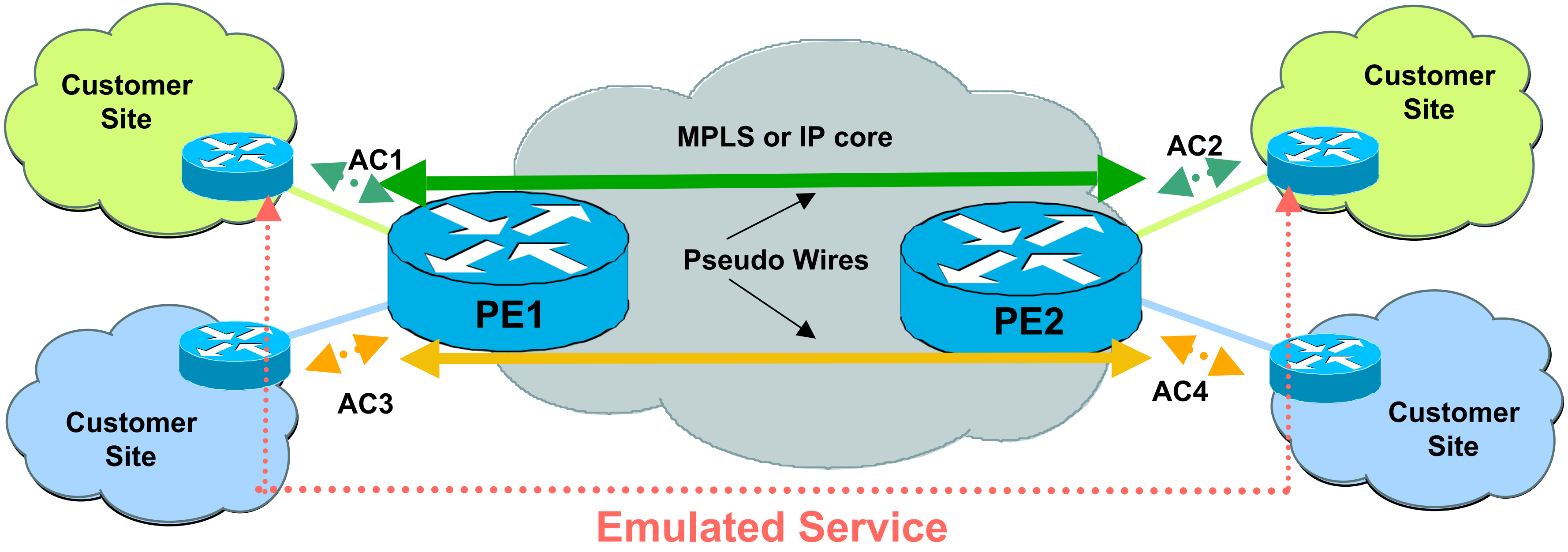
LAYER 3 VPNS	LAYER 2 VPNS
1. Layer 3 IP packet based forwarding 2. SP is involved in routing using VRF tables 3. SP involved in security decisions 4. Example: RFC 2547bis VPNs (L3 MPLS-VPN)	1. Layer 2 frame based forwarding e.g. DLCI, VLAN, VPI/VCI, etc … 2. No SP routing involvement 3. No SP security involvement 4. Medias: FR, ATM, Ethernet, PPP, HDLC, POS, etc …

The Choice of L2VPN over L3VPN Will Depend on How Much Routing/Security Control the Enterprise Wants to Retain.
L2 VPN Services Are Complementary to L3 VPN Services

L2VPN Models

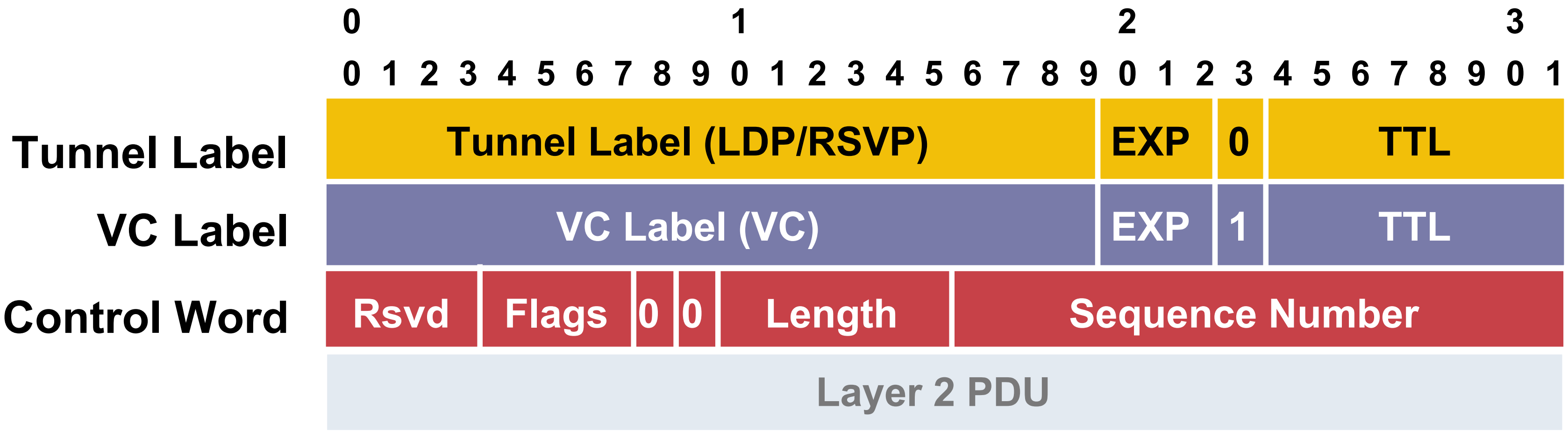


L2VPN - Pseudo Wire Reference Model



- A pseudo-wire (PW) is a connection between two provider edge (PE) devices which connects two attachment circuits (ACs)
- Emulates essential attributes of a service

L2VPNs –
Label Stacking

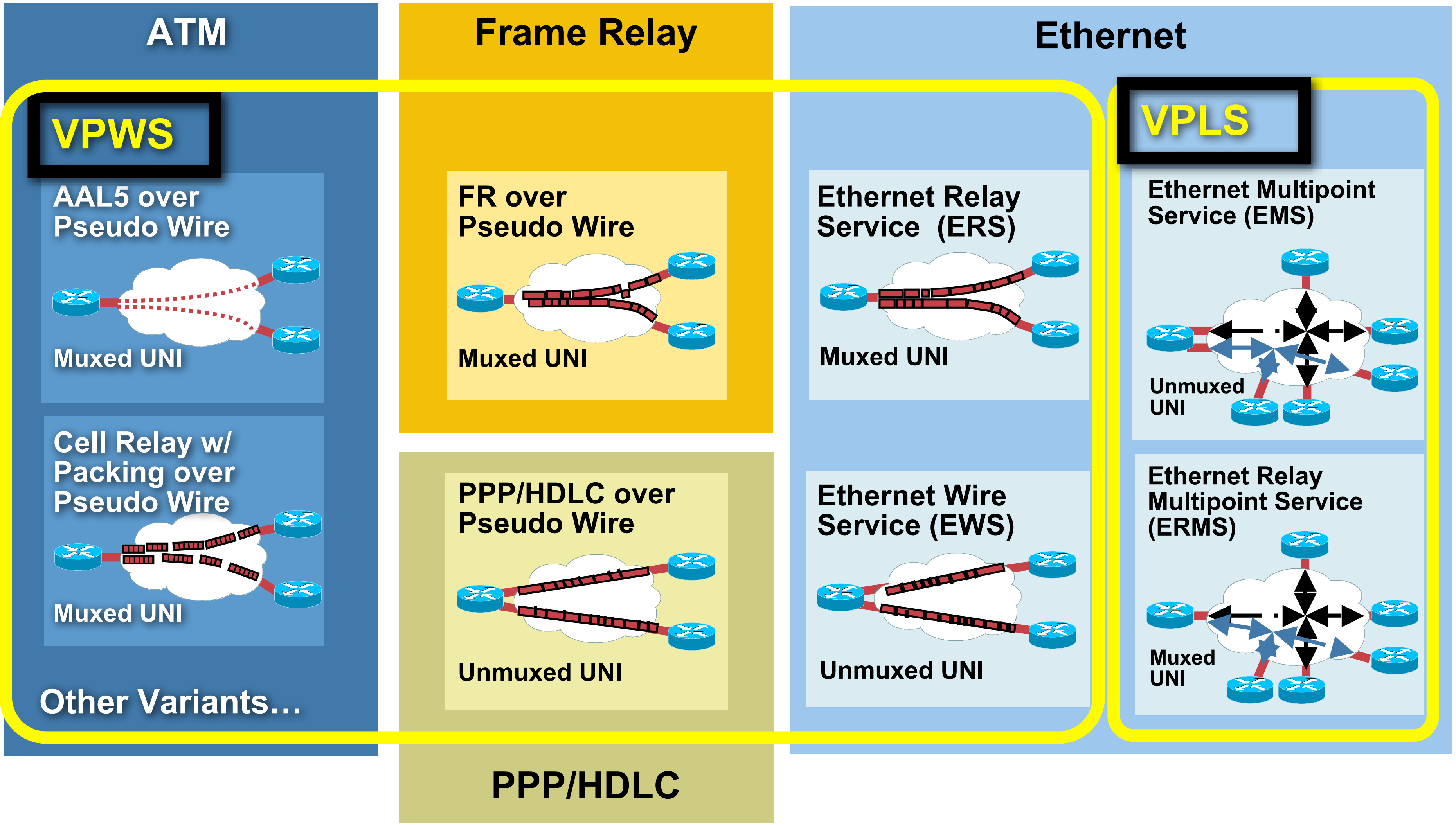


	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
Tunnel Label	Tunnel Label (LDP/RSVP)		EXP 0 TTL	
VC Label	VC Label (VC)		EXP 1 TTL	
Control Word	Rsvd Flags 0 0 Length	Sequence Number		
	Layer 2 PDU			

- **Three Layers of Encapsulation**
  - **Tunnel Label – Determines path through network**
  - **VC Label – Identifies VC at endpoint**
  - **Control Word – Contains attributes of L2 payload (optional)**

Control Word	
Encap.	**Required**
CR	No
AAL5	Yes
Eth	No
FR	Yes
HDLC	No
PPP	No

A Look at L2 VPN Services



ATM

Frame Relay

Ethernet

VPWS

AAL5 over Pseudo Wire

Muxed UNI

Cell Relay w/ Packing over Pseudo Wire

Muxed UNI

Other Variants…

FR over Pseudo Wire

Muxed UNI

PPP/HDLC over Pseudo Wire

Unmuxed UNI

PPP/HDLC

Ethernet Relay Service (ERS)

Muxed UNI

Ethernet Wire Service (EWS)

Unmuxed UNI

VPLS

Ethernet Multipoint Service (EMS)

Unmuxed UNI

Ethernet Relay Multipoint Service (ERMS)

Muxed UNI

Pseudo Wire –
IETF Technology Adoption

Virtual Private Wire Service (VPWS) P2P

RFC3916 Pseudo Wire Emulation Edge-to-Edge (PWE3) Requirements

RFC3985 Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture

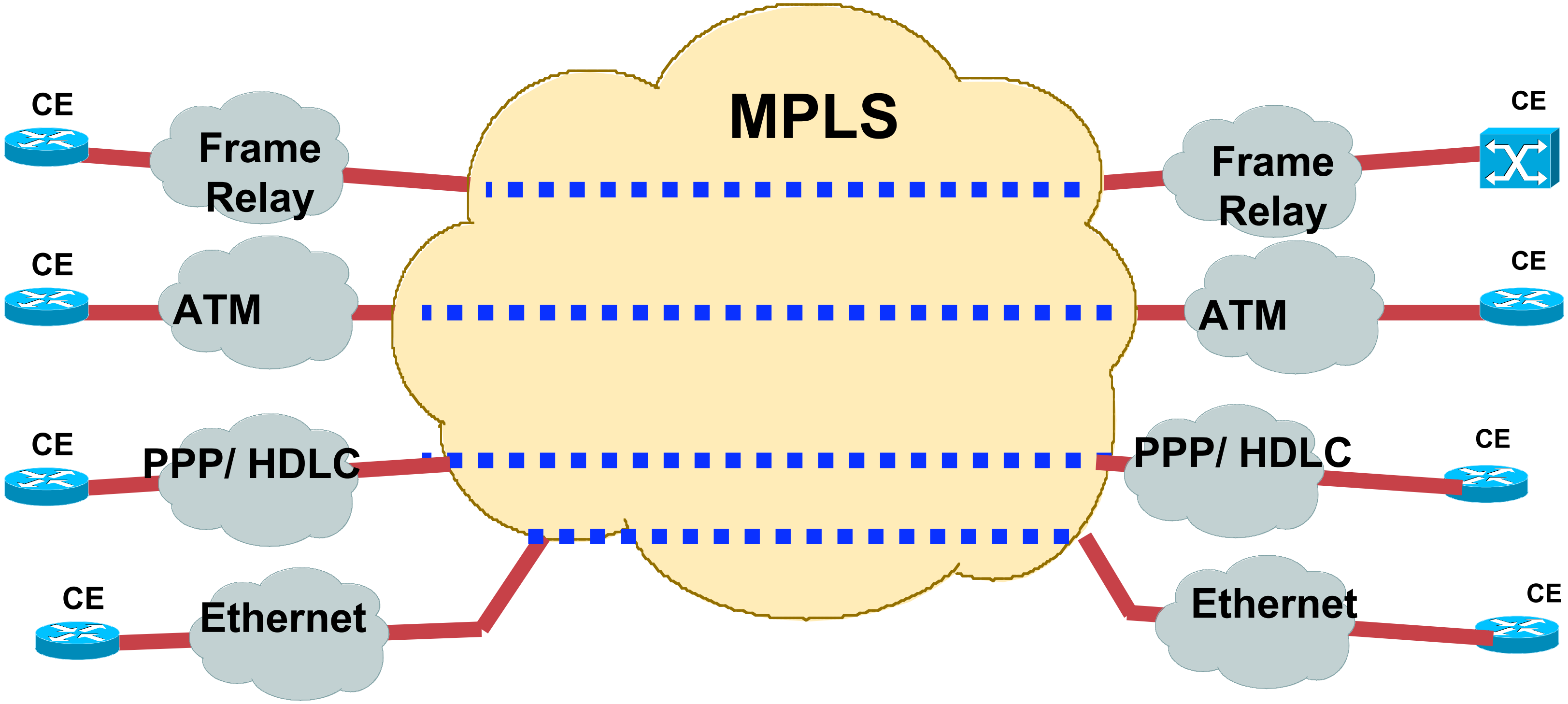
RFC4385 Pseudo wire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN

draft-ietf-pwe3-[atm, frame-relay, ethernet, etc.]

Virtual Private LAN Services (VPLS) P2M

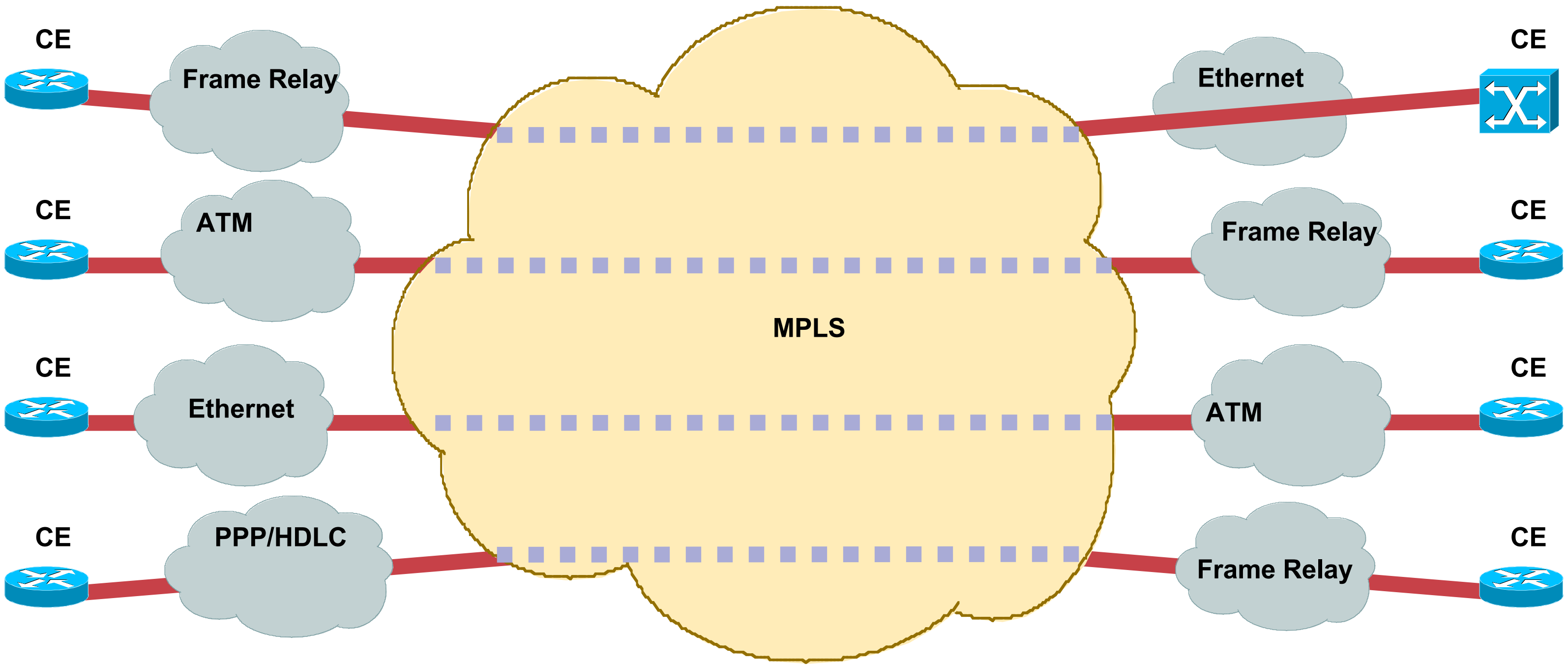
draft-ietf-l2vpn-vpls-ldp-xx
draft-ietf-l2vpn-vpls-bgp-xx

VPWS Transports –
Customers Viewpoint



L2VPN: Service Interworking
Customers Viewpoint

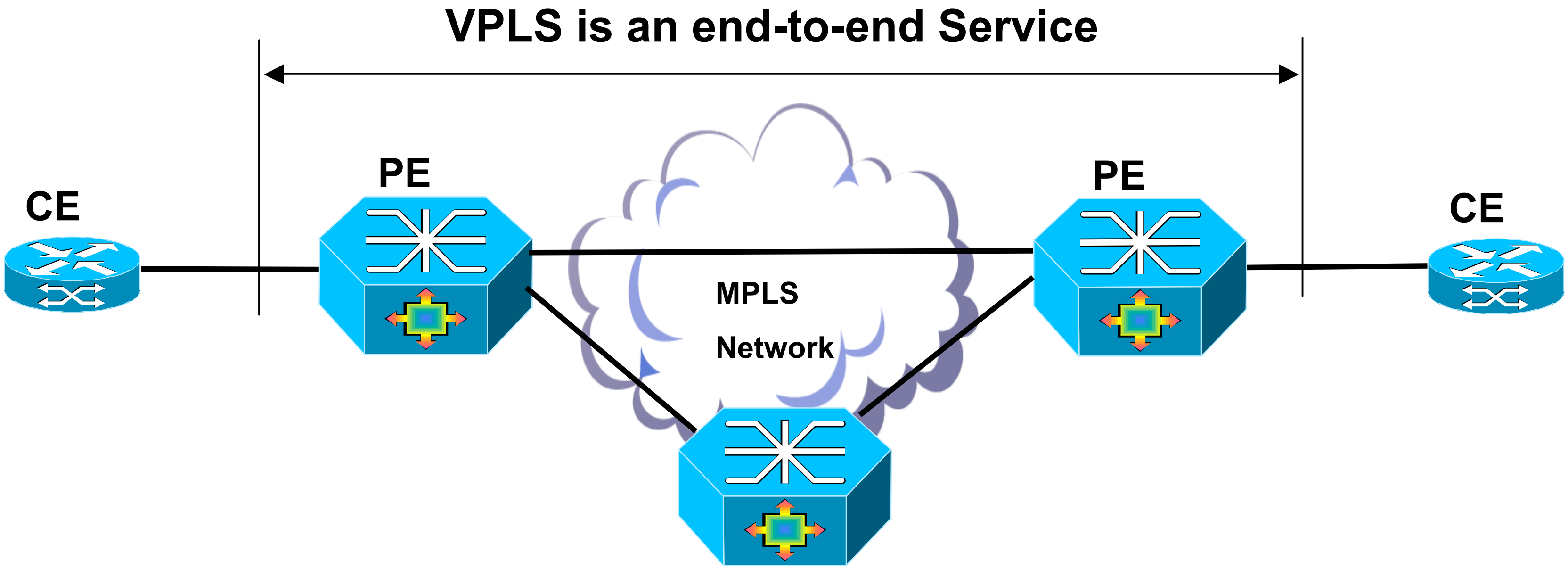
MPLS Service Inter-working



MPLS Service Inter-working

Service inter-working allows Customer Edge (CE) devices to exchange service layer PDUs transparently across different link layer technologies.

Virtual Private LAN Services (VPLS)



- Provides Ethernet Multipoint Services (EMS) over MPLS network
- VPLS operation emulates an IEEE Ethernet bridge
- Two VPLS drafts in existance

  Draft-ietf-l2vpn-vpls-ldp-xx

  draft-ietf-l2vpn-vpls-bgp-xx

VPLS: Requirements

A Virtual Switch MUST operate like a conventional L2 bridge!

Flooding / Forwarding:

- MAC table instances per customer and per customer VLAN (L2-VRF idea) for each PE
- VSI will participate in learning, forwarding process
- Flood unknown MAC addresses

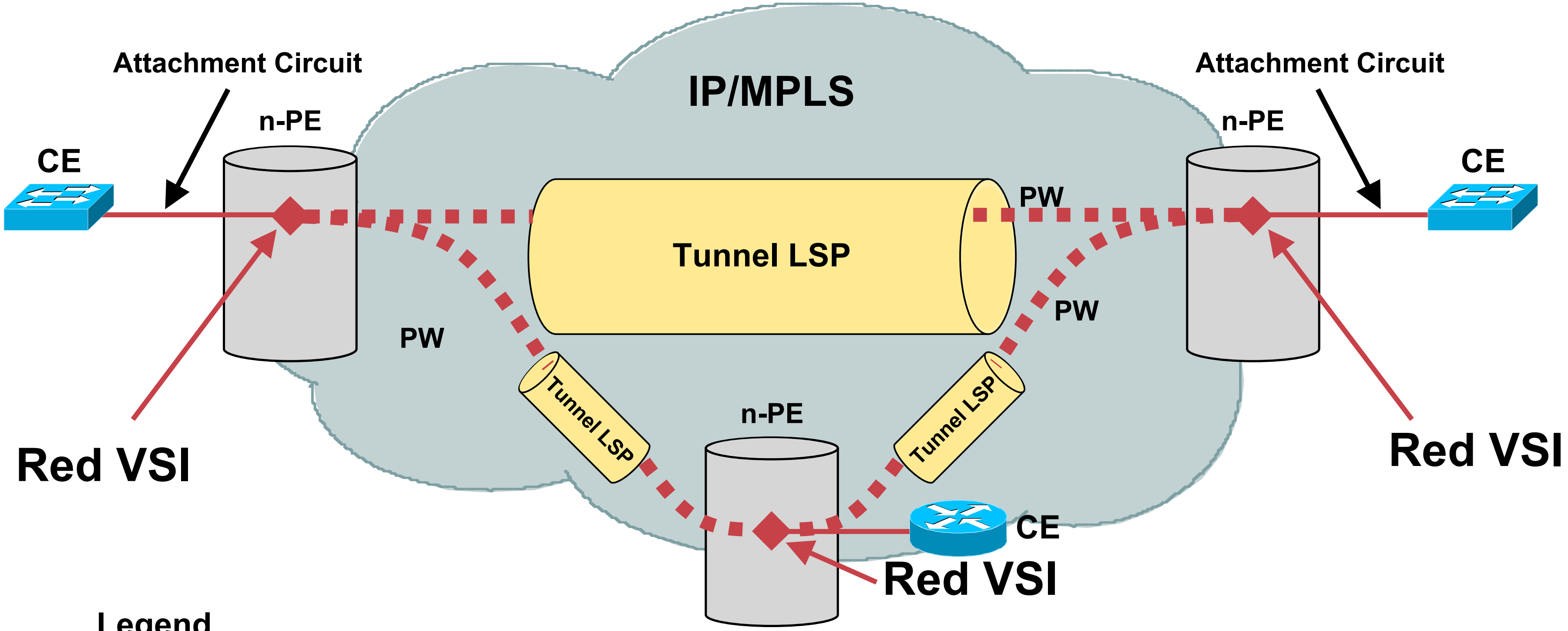
Address Learning / Aging:

- Self Learn Source MAC to port associations
- Refresh MAC timers with incoming frames
- New additional MAC TLV to LDP

Loop Prevention:

- Create partial or full-mesh of EoMPLS VCs per VPLS
- Use “split horizon” concepts to prevent loops
- Announce EoMPLS VPLS VC tunnels

VPLS Components



Legend

- CE—Customer Edge Device; used to connect to the SP’s network
- n-PE—Network facing-Provider Edge; acts as a gateway between the MPLS core and edge domain
- VSI—Virtual Switching Instance; describes an Ethernet bridge function within the n-PE; the VFI terminates the Pseudowire
- PW—Pseudowire; a PW connects two VSI’s; Consists of a pair of MPLS uni-directional VC’s
- AC—Attachment Circuit; a customer connection to the service provider; may be a physical port or Ethernet VLAN
- Tunnel LSP—Tunnel Label Switch Path is used to tunnel PW’s between VFI’s

Introduction to MPLS Traffic Engineering

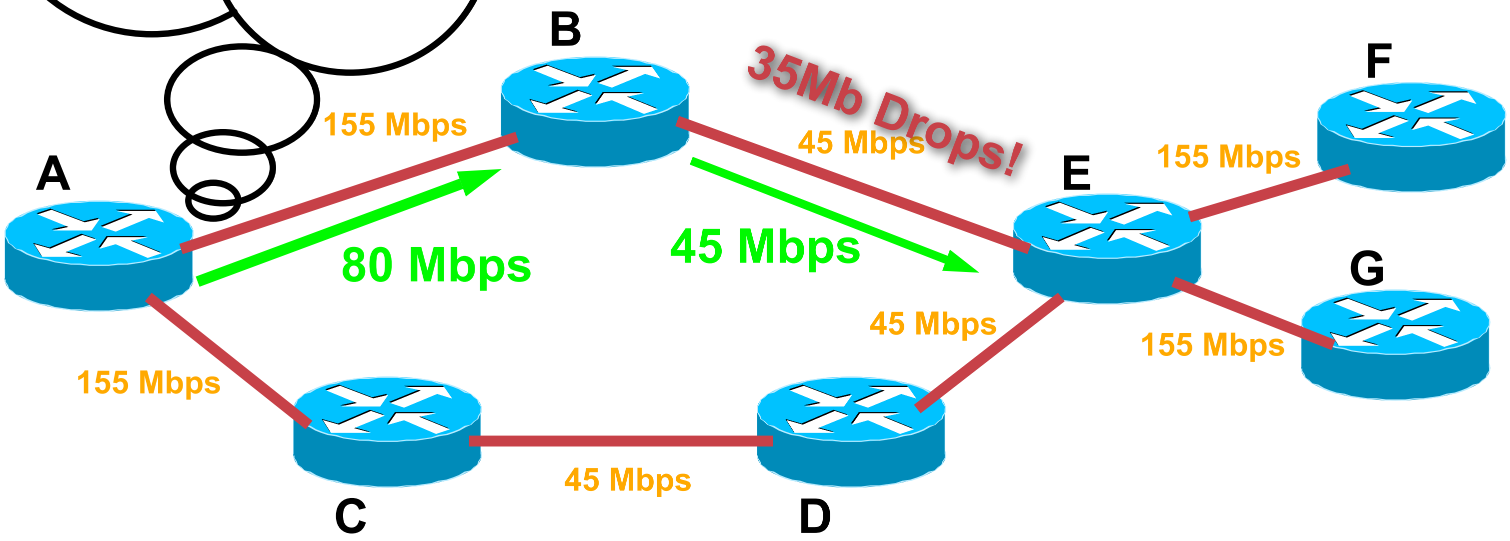
Motivation for Traffic Engineering

- Increase efficiency of bandwidth resources
  - Prevent over-utilized (congested) links whilst other links are under-utilized
- Ensure the most desirable/appropriate path for some/all traffic
  - Explicit-Path overrides the shortest path selected by the IGP
- Replace ATM/FR cores
  - PVC-like traffic placement without IGP full mesh and associated O(N^2) flooding
- The ultimate goal is COST SAVING
  - Service development also progressing

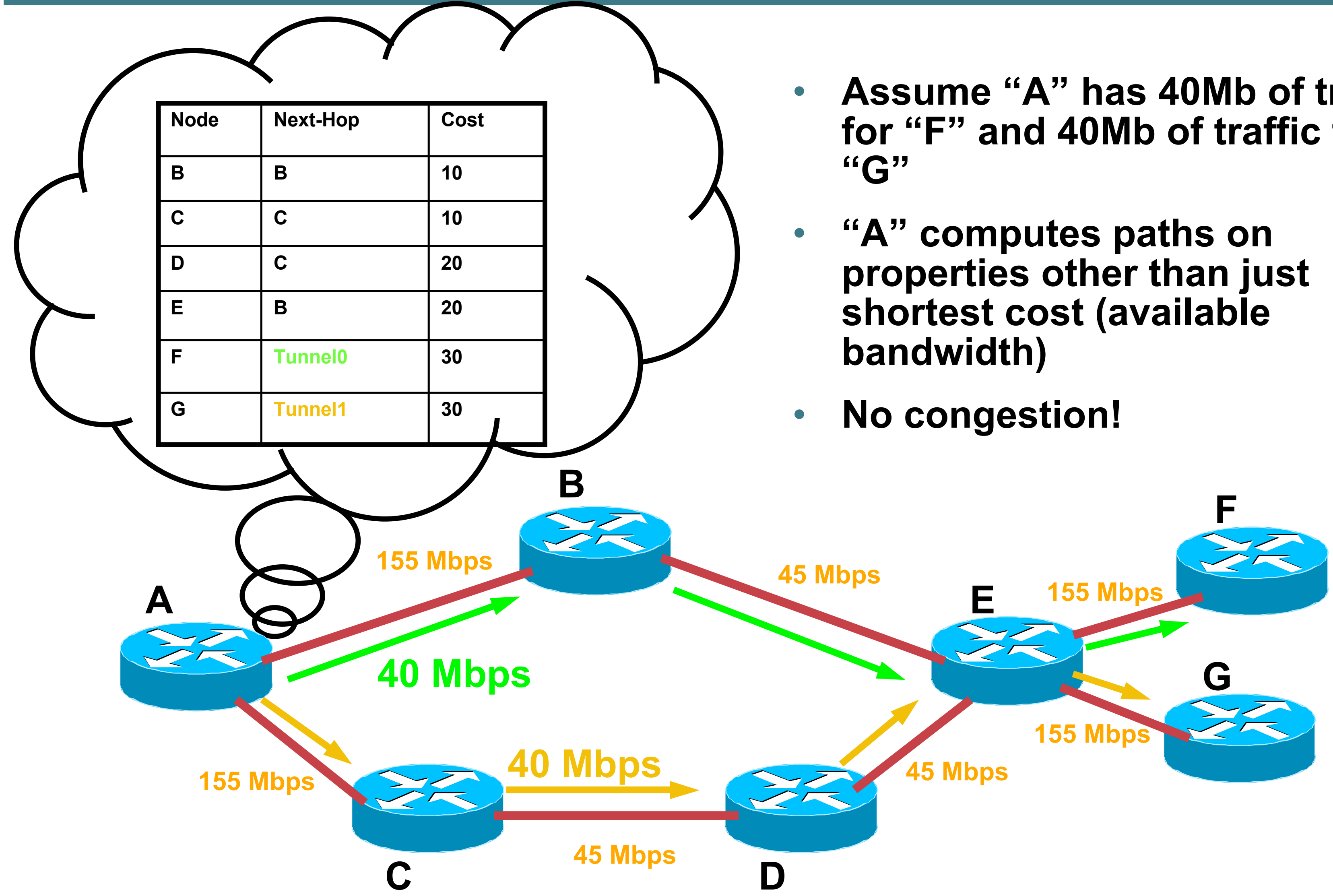
The Problem With Shortest-Path

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	B	30

- Assume “A” has 40Mb of traffic for “F” and 40Mb of traffic for “G”
- Some links are 45 Mbps, some are 155 Mbps
- Massive (44%) packet loss between “B” and “E”
- Changing path to A->C->D->E won’t help



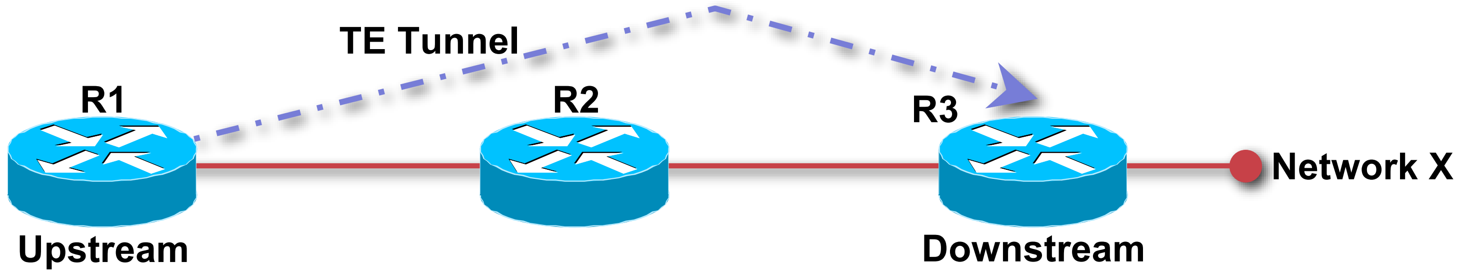
MPLS-TE Example



Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	Tunnel0	30
G	Tunnel1	30

- Assume “A” has 40Mb of traffic for “F” and 40Mb of traffic for “G”
- “A” computes paths on properties other than just shortest cost (available bandwidth)
- No congestion!

A Terminology Slide—Head, Tail, LSP, etc.

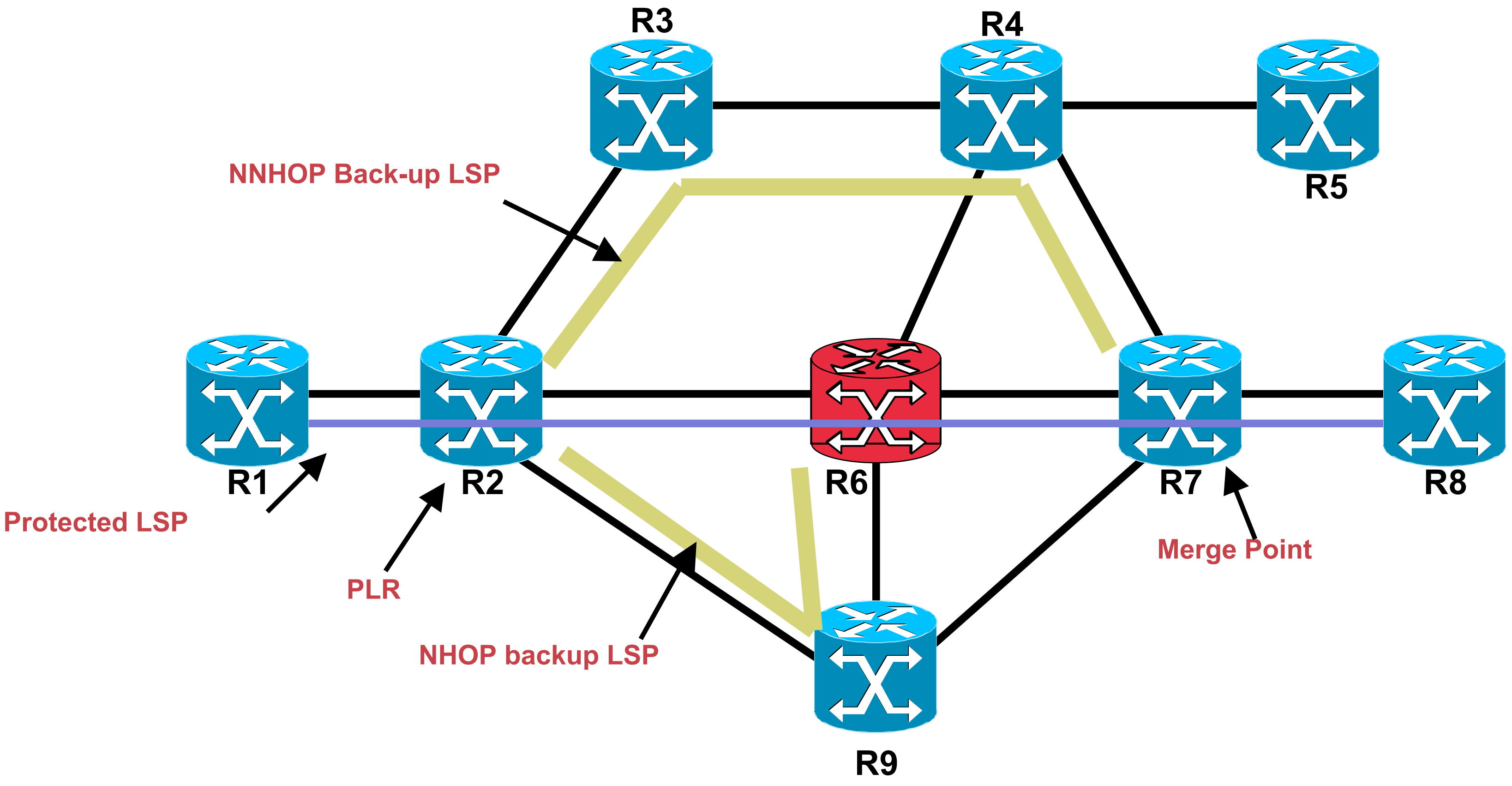


- Head-End is a router on which a TE tunnel is configured (R1)
- Tail-End is the router on which TE tunnel terminates (R3)
- Mid-point is a router thru which the TE tunnel passes (R2)
- LSP is the Label Switched Path taken by the TE tunnel, here R1-R2-R3
- Downstream router is a router closer to the tunnel tail
- Upstream router is farther from the tunnel tail (so R2 is upstream to R3's downstream, R1 is upstream from R2's downstream)

Fast ReRoute

- Fundamental point from earlier: “you can use MPLS-TE to forward traffic down a path other than that determined by your IGP cost”
- FRR builds a path to be used in case of a failure in the network
- Minimize packet loss by avoiding transient routing loops

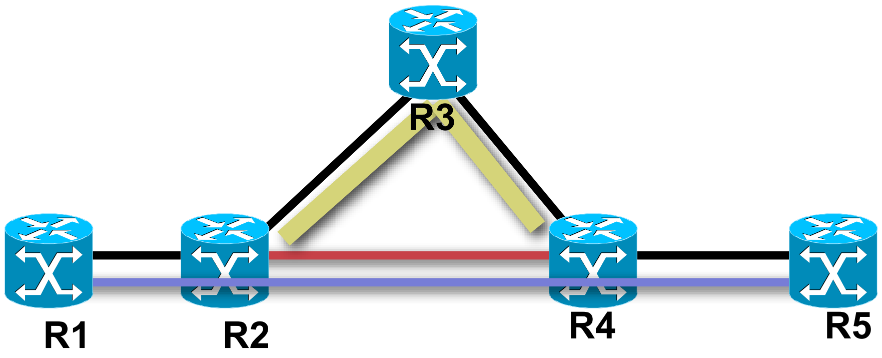
Terminology

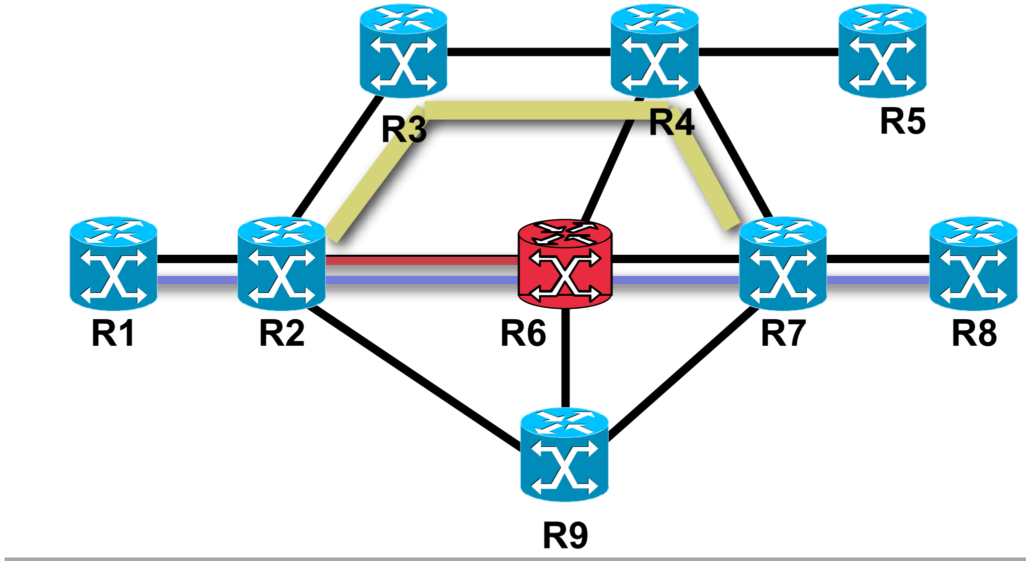


Fast ReRoute

MPLS Fast Reroute Local Repair

- **Link protection**: the backup tunnel tail-end (MP) is one hop away from the PLR





- **Node protection**: the backup tunnel tail-end (MP) is two hops away from the PLR

FRR Procedures

1. Pre-establish backup paths
2. Failure happens, protected traffic is switched onto backup paths
3. After local repair, tunnel headends are signaled to recover if they want; no time pressure here, failure is being protected against
4. Protection is in place for hopefully ~10-30+ seconds; during that time, data gets through

MPLS Quality of Service

What is Quality of Service?

“

QoS represents the set of techniques necessary t manage network bandwidth, delay, jitter, and pac loss.

From a business perspective, it is essential to ass that the critical applications are guaranteed the network resources they need, despite varying network traffic load.

”

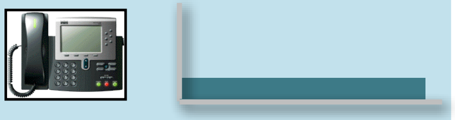
Traffic Characterization

- Identify traffic sources and types
- Need for appropriate handling
  - Realtime and Non-realtime
    - Voice (Delay sensitive)
    - Video (Bandwidth intensive)
    - Data (Loss sensitive)
      - HTTP, FTP, SMTP
  - Bursty and Constant type
  - Multi-service traffic: IP, MPLS
  - Single or Multiple flows of the same type

Traffic Profiles and basic QoS requirements

Voice, Video and Data

Voice



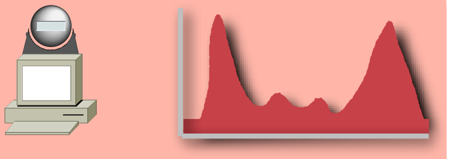
- Smooth
- Benign
- Drop Sensitive
- Delay Sensitive
- UDP Priority

Bandwidth per call depends on codec, sampling-rate, and Layer 2 media

- Latency ≤ 150 ms
- Jitter ≤ 30 ms
- Loss ≤ 1%

One-way requirements

Video-Conf



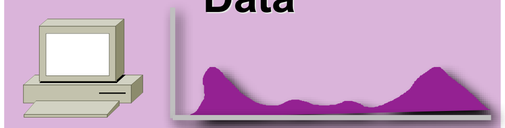
- Bursty
- Greedy
- Drop Sensitive
- Delay Sensitive
- UDP Priority

IP/VC has the same requirements as VoIP, but has radically different traffic patterns

- Latency ≤ 150 ms
- Jitter ≤ 30 ms
- Loss ≤ 1%

One-way requirements

Data



- Smooth/Bursty
- Benign/Greedy
- Drop Insensitive
- Delay Insensitive
- TCP Retransmits

Traffic patterns for Data vary among applications (and even among different *versions* of the same application)

Data Classes:

Mission-Critical Apps

Transactional/Interactive Apps

Bulk Data Apps

Best Effort Apps (Default)

Optional: Scavenger Apps

Different QoS Requirements

	Voice	FTP	ERP and Mission-Critical
Bandwidth	Low to Moderate	Moderate to High	Low
Random Drop Sensitive	Low	High	Moderate To High
Delay Sensitive	High	Low	Low to Moderate
Jitter Sensitive	High	Low	Moderate

Traffic Is Grouped into Classes that Have Similar QoS Requirements (or Part of Same SLA)

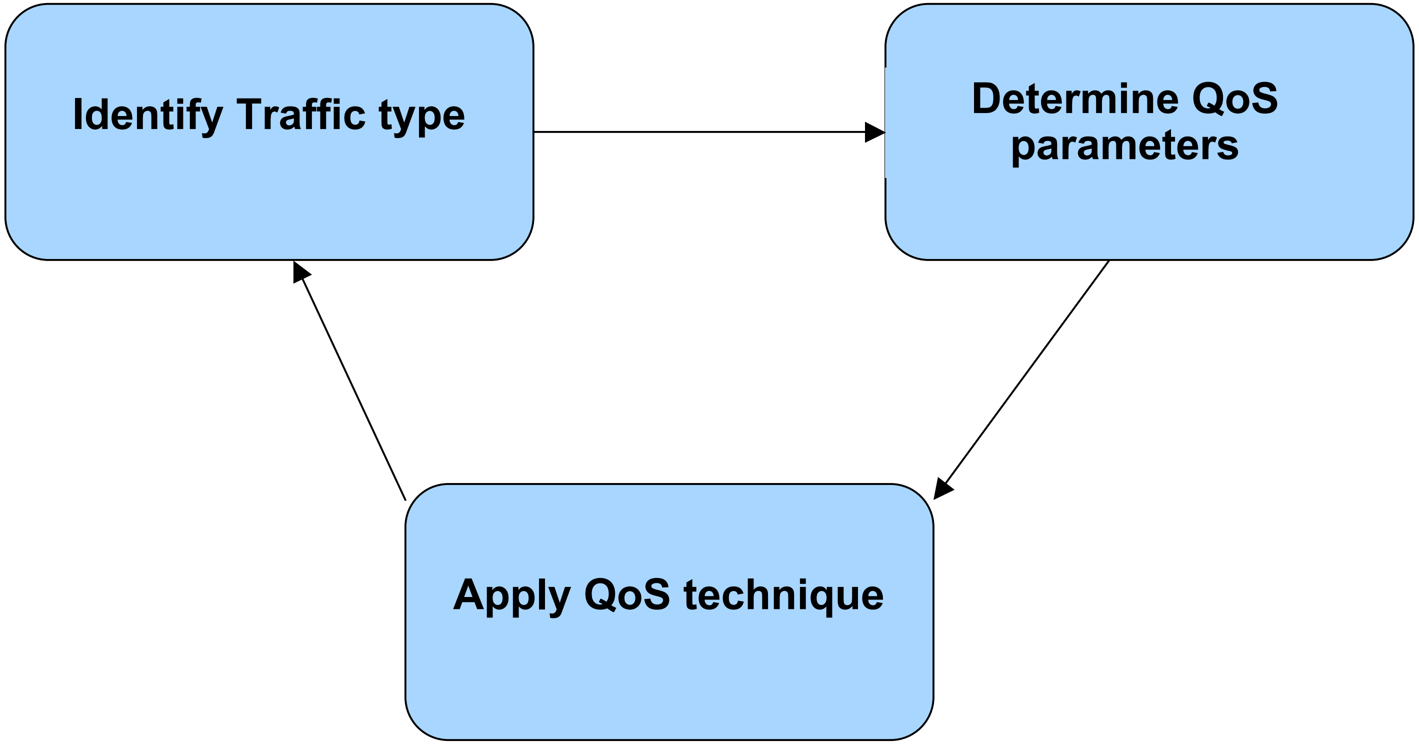
QoS Requirements

- **Traffic influencing parameters**
  - **Latency, Jitter, Loss**

- **Management of finite resources**
  - **Rate Control**
  - **Queuing and Scheduling**
  - **Congestion Management**
  - **Admission Control**
  - **Routing Control Traffic protection**

- **Service Level Agreement (SLA)**
  - **per-flow**
  - **aggregated**

QoS Triangle

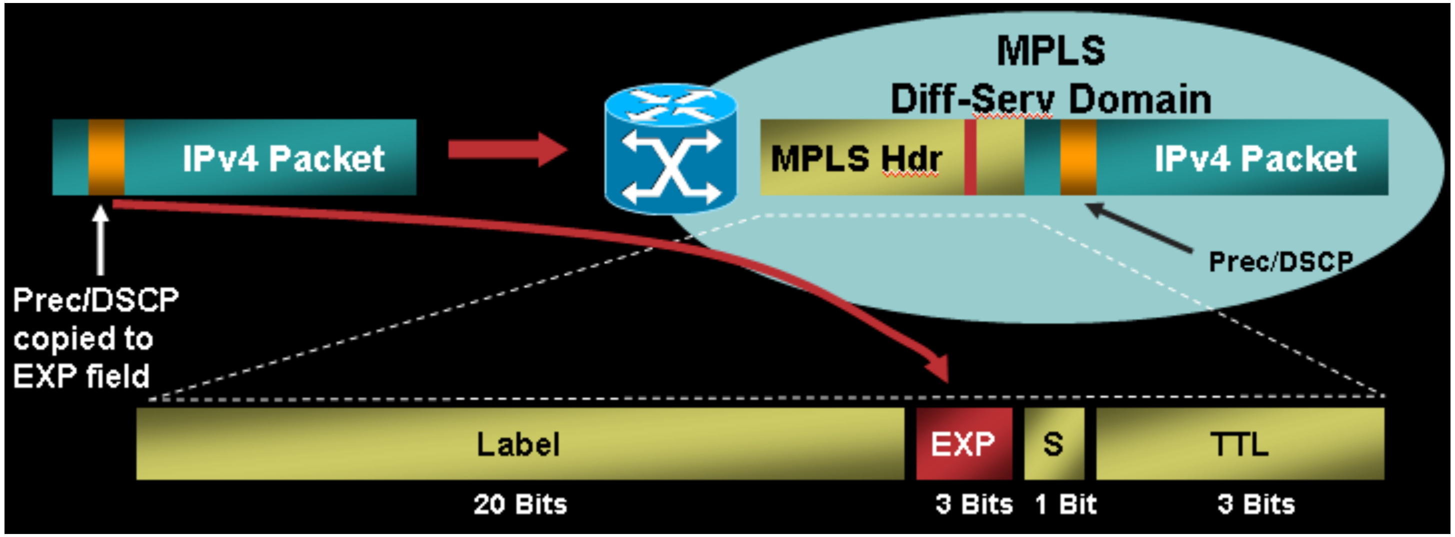


MPLS DiffServ Architecture

- **MPLS does NOT define new QoS architectures**
- **MPLS QoS uses Differentiated Services (DiffServ) architecture defined for IP QoS (RFC 2475)**
- **MPLS DiffServ is defined in RFC3270**

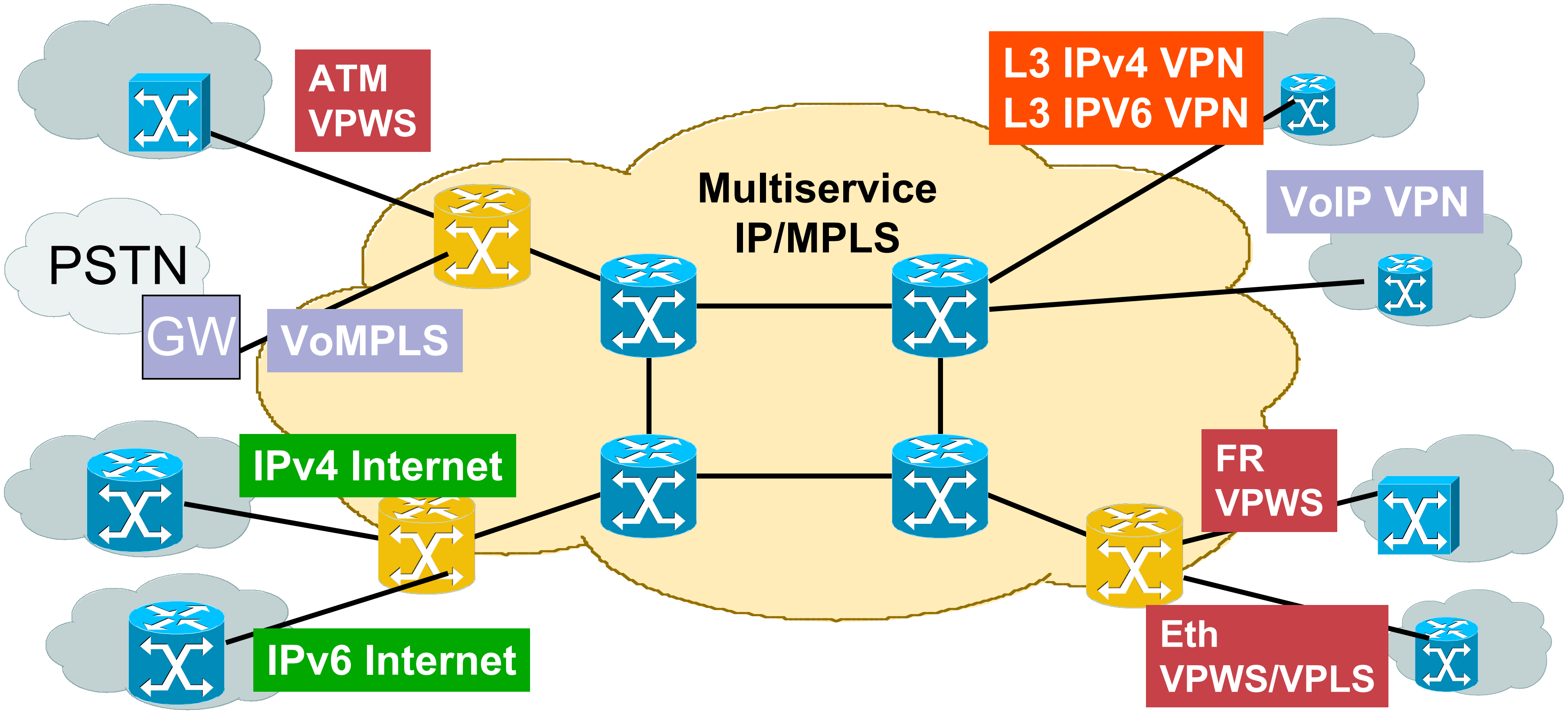
What’s new in MPLS DiffServ ?

IP DiffServ Domain



- **Prec/DSCP field is not directly visible to MPLS Label Switch Routers (they forward based on MPLS Header and EXP field)**
- **Information on DiffServ must be made visible to LSR in MPLS Header using EXP field / Label.**

Multiservice MPLS QoS Architecture



- MPLS typically the basis for next generation Multiservice Infrastructure supporting many services
- MPLS QoS architecture must fit in Multiservice strategy

Quality of Service Operations
How Do QoS Tools Work?

CLASSIFICATION AND MARKING

QUEUING AND
(SELECTIVE) DROPPING

SHAPING

Conclusions

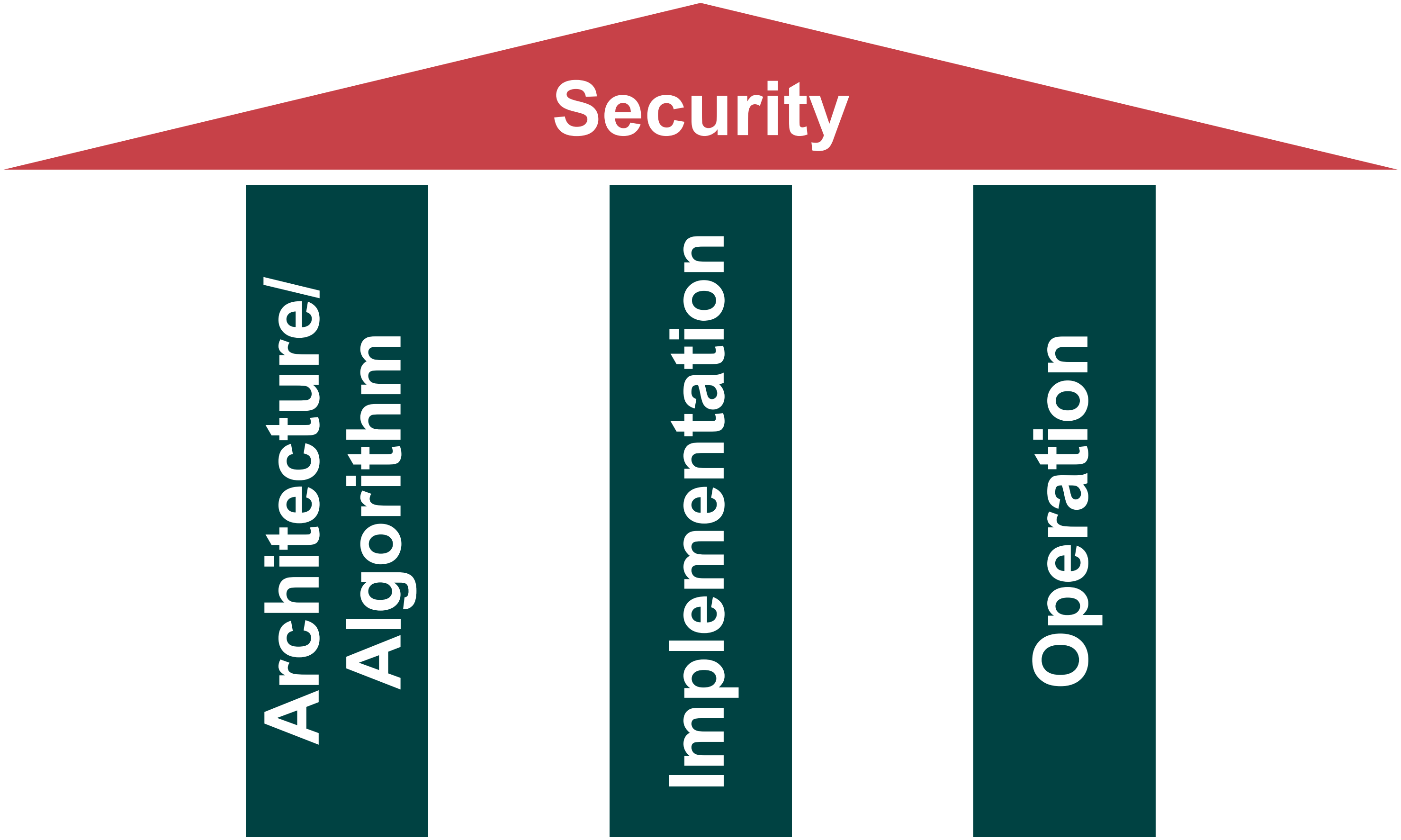
- QoS is fundamental element in enabling IP/MPLS to as the next generation Multiservice infrastructure
- Generic scalable Core QoS design with set of tools selected based on level of bandwidth optimisation and service differentiation required
- Core QoS designs ready and used for the most stringent applications like PSTN/3G trunking, ATM Trunking, VoIP etc
- Multiple sophisticated service-specific Edge QoS designs (L3VPN, VPWS, PSTN Trunking, …)
- L3VPN QoS offering is rich and ready for multimedia intranet
- VPWS QoS offering becoming very rich too

MPLS Security

Comparison with ATM/FR

	ATM/FR	MPLS
Address Space Separation	Yes	Yes
Routing Separation	Yes	Yes
Resistance to Attacks	Yes	Yes
Resistance to Label Spoofing	Yes	Yes
Direct CE-CE Authentication (Layer 3)	Yes	With IPsec

Security Relies on Three Pillars



Security

Architecture/ Algorithm

Implementation

Operation

Break One, and All Security Is Gone!

Secure MPLS/VPN Core Design

- Don’t let packets into the core!

  No way to attack core, except through routing, thus:

- Secure the routing protocol

  Neighbor authentication, maximum routes, dampening,…

- Design for transit traffic

  QoS to give VPN priority over Internet

  Choose correct router for bandwidth

  Separate PEs where necessary

- Operate Securely

Still “Open”: Routing Protocol

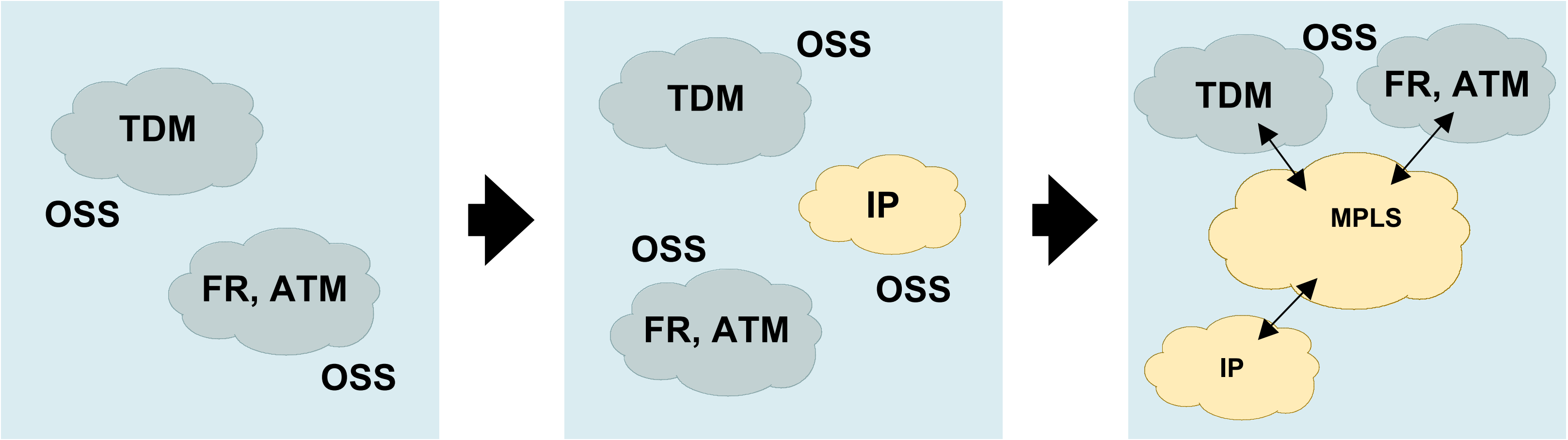
Only Attack Vector: Transit Traffic

Now Only Insider Attacks Possible

Avoid Insider Attacks

MPLS Operation, Administration and Management (OAM)

Service Provider Network Operation



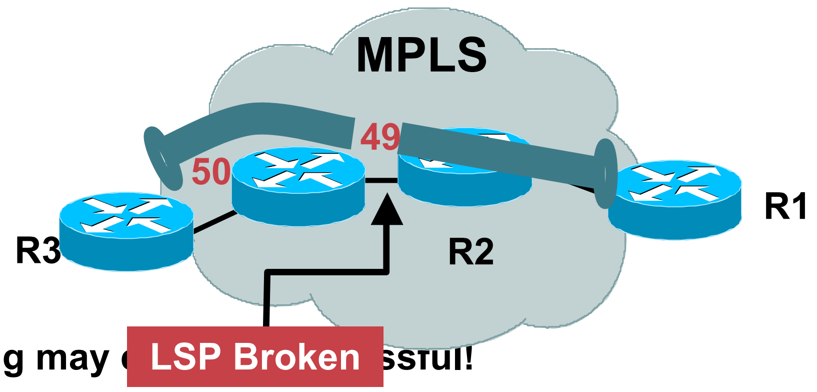
- Create operational efficiencies and increase automation in a highly technology-intensive market
- Enable competitive differentiation and customer retention through high-margin, bundled services
- Progressively consolidate disparate networks
- Sustain existing business while rolling out new services

Some MPLS Transport Problems…

- LDP/RSVP-TE fails but not IGP protocol (control plane)-LDP session fails to start
- Data plane fails -IGP protocol doesn't ("Black Holes")
- Figuring out the root cause of performance degradation on a specific link/ path
- Connectivity problem diagnosis: PE–PE…is the SP network functioning? aka is it a customer premises error?
- Visualizing the MPLS packet flow path through a complex network: what path is being taken? (namely with TE)
- Packets not being labelled…

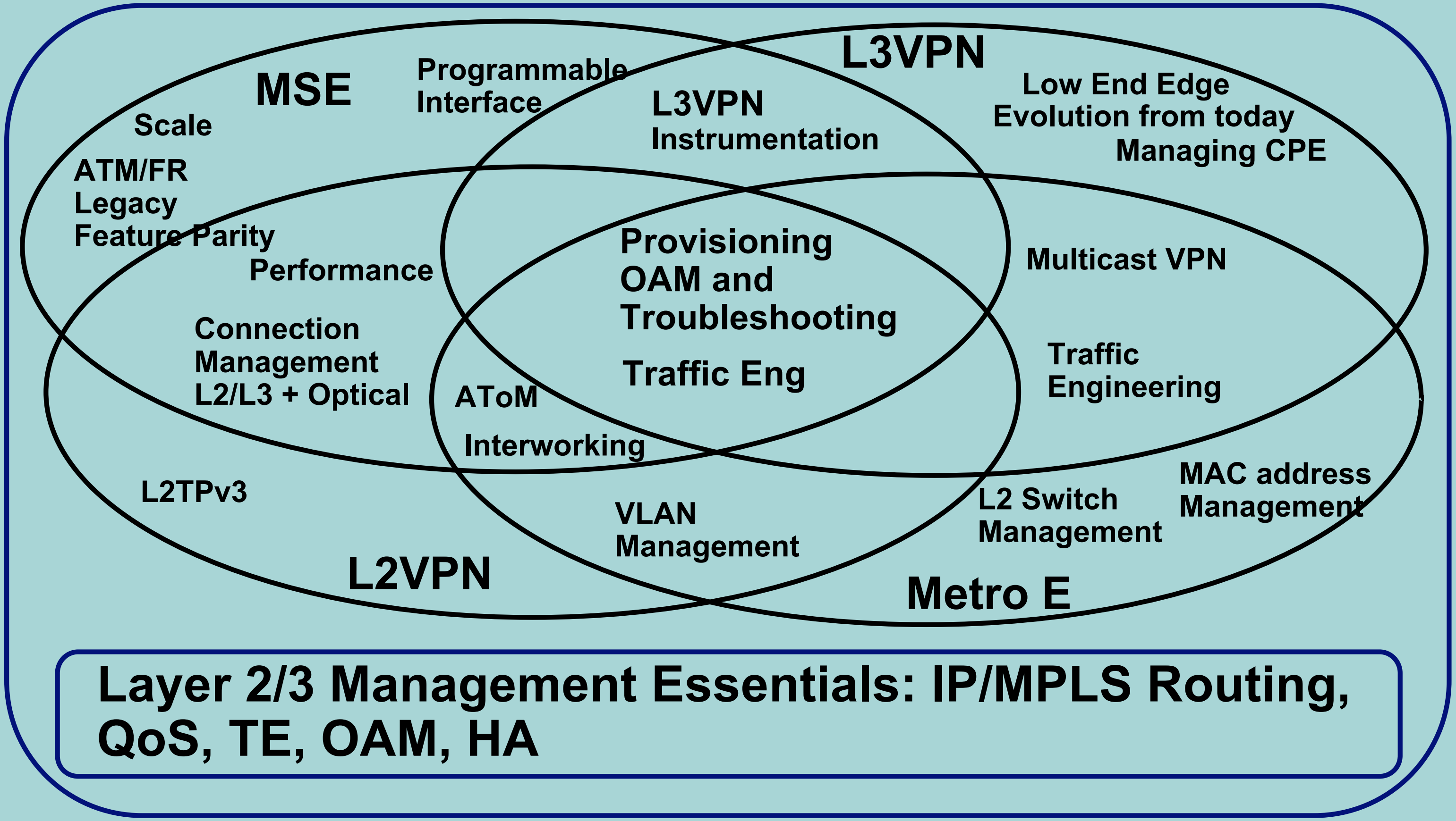
MPLS OAM Leveraged in the Workflow

- A broken LSP will affect end to end connectivity and services, e.g., MPLS VPN PE-PE connectivity => several customers affected
- Various reasons for an LSP to break:
  - Broken LDP adjacency
  - MPLS not enabled (globally or per interface)
  - Mismatched labels
  - Software/hardware corruption…
- It is difficult to detect an MPLS failure:
  - In some instances, traditional ICMP-based ping may ssful!
  - If not successful, no diagnostic given—just timeout
- Difficult to troubleshoot an MPLS failure:
  - Requires the operator to do manual/hop-by-hop work



MPLS OAM Facilitates and Speeds Up Troubleshooting of MPLS Failures

MPLS Services and Transport Network Management

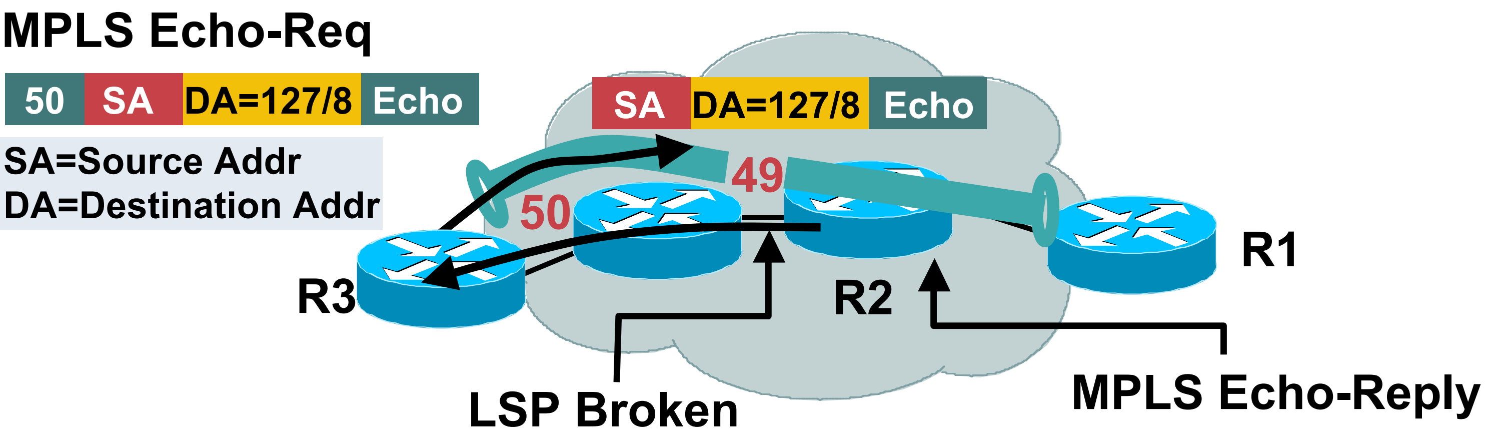


Layer 2/3 Management Essentials: IP/MPLS Routing, QoS, TE, OAM, HA

MPLS LSP Ping/Traceroute

Requirement	• Detect MPLS traffic black holes or misrouting • Isolate MPLS faults • Verify data plane against the control plane • Detect MTU of MPLS LSP paths
Solution	• MPLS LSP Ping (ICMP) for connectivity checks • MPLS LSP Traceroute for hop-by-hop fault localization • MPLS LSP Traceroute for path tracing
Applications	• IPv4 LDP prefix, VPNv4 prefix • TE tunnel • MPLS PE, P connectivity for MPLS transport, MPLS VPN, MPLS TE applications
IETF Standards	• Draft-ietf-mpls-lsp-ping-09.txt

How to Use the MPLS OAM Options and Features for Troubleshooting



- Same label stack: takes the exact same path as MPLS data
- The IP header destination address field of the echo request is a 127/8 address (127.0.0.1 by default)
- Presence of the 127/8 address in the IP header destination address field causes the packet to be consumed by any router trying to forward the packet using the IP header; also prevents the probe packet to exit the MPLS cloud
- In this case R2 would not forward the echo-req to R1 but rather consumes the packet and reply to it accordingly
- LSP reply will be generated as an IP packet which may use an LSP path back if available

Diagnostic Capability at Failure Point for Fault Localization and more Options to Provide for Efficient Troubleshooting Information

This is only a starting point…

- For further help while you work with MPLS, email the email alias afnog-help@cisco.com.
- Tons of materials are available on the Internet, on cisco.com and other sites.

Q & A
LAB