MPLS Layer 2 and Layer 3 Deployment Best Practice Guidelines

Monique Morrow

mmorrow@cisco.com

February 21 2005

Prerequisites and Scope

- **Must understand fundamental MPLS principles**

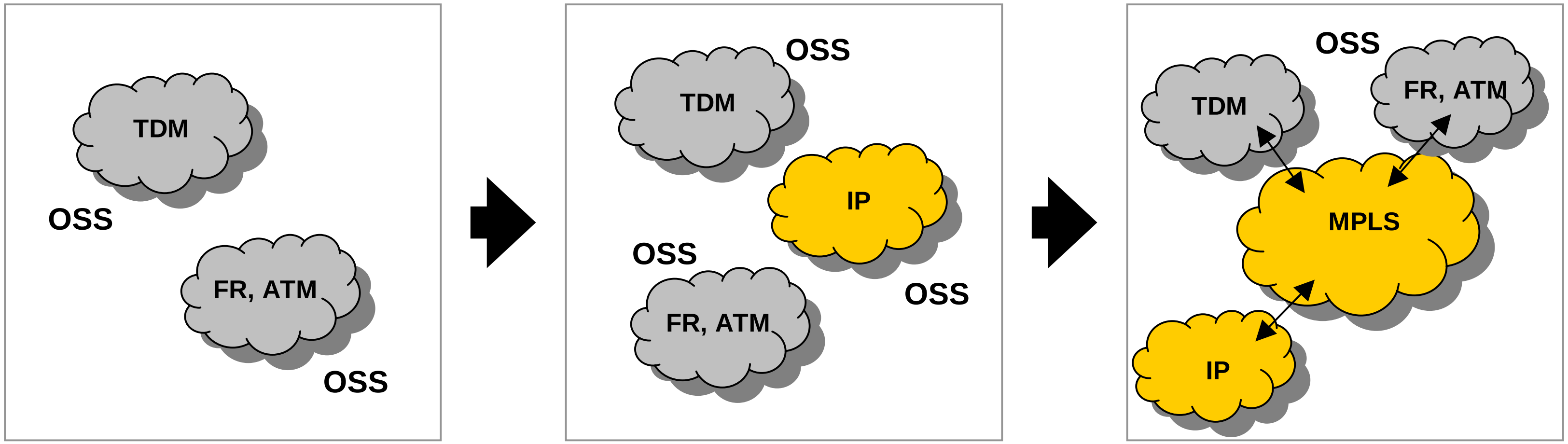
- **Must understand basic routing especially BGP**

Agenda

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider Considerations
- Layer 2 Deployment Considerations
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Service Provider Network Operation

Cisco.com



- **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**

MPLS's Momentum in Convergence & Service Creation

- **IDC, July 2004:**

  Increasingly, service providers use MPLS as the cornerstone for traffic routing capabilities for converged frame, ATM, and packet based networks to improve QoS visibility and assure service level guarantees.

- **CIBC World Markets, June 2004:**

  The most significant trend was a wholesale shift to IP-MPLS as the new foundation technology for carriers' data networks. This transition appears irreversible and is gaining momentum surprisingly fast.

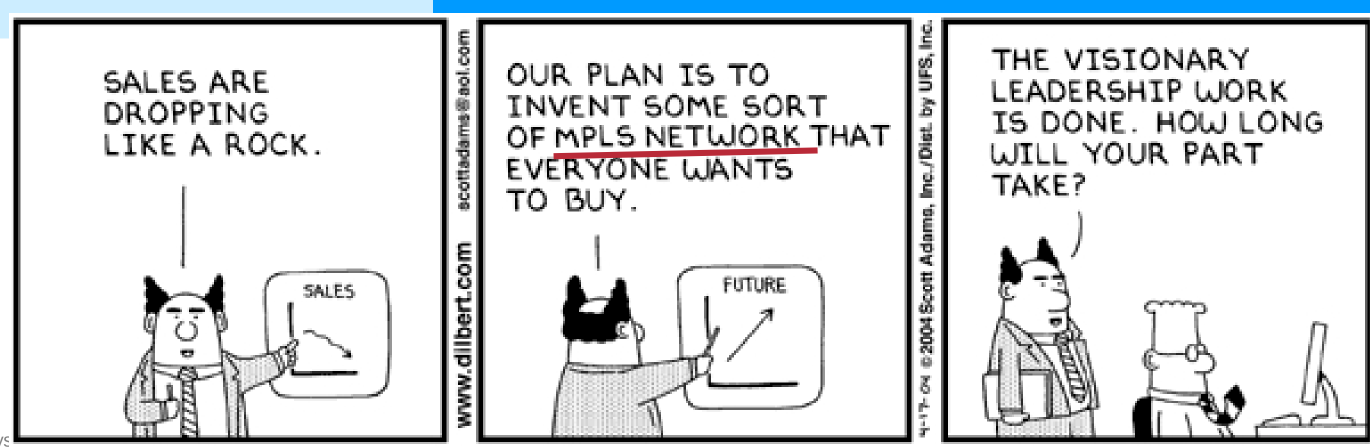
- **Heavy Reading Jan. 2004:**

  Most of the world's telecom service providers now agree in principle that they must migrate to converged backbones, and that MPLS (Multiprotocol Label Switching) technology will enable this migration.

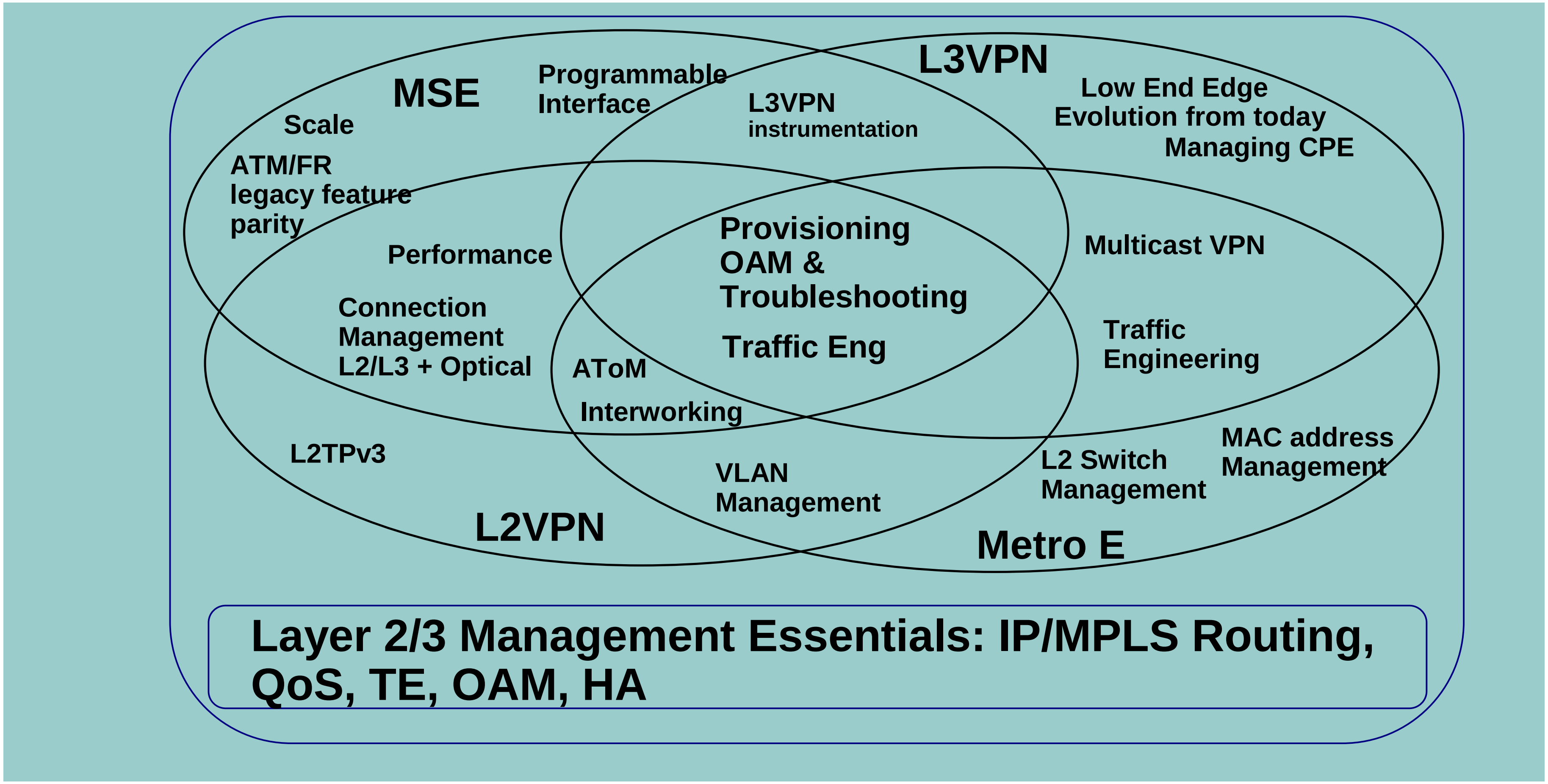
- **Heavy Reading Sep. 2003:**

  MPLS is gaining support from MSPP vendors as a key mechanism for enabling packet services, QoS, and traffic engineering in the metro.

- **Even in Dilbert Comic Strip, May 2004:**



MPLS Services and Transport Network Management



Agenda

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider Considerations
- Layer 2 Deployment Considerations
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Why Half Duplex VRFs? Problem

- **Only way to implement hub and spoke topology is to put every spoke into a single and unique VRF**

    Ensures that spokes do not communicate directly

- **Single VRF model, which does not include HDV, impairs the ability to bind traffic on the upstream ISP Hub**

Why Half Duplex VRFs? Solution

- **HDV allows the wholesale Service Provider to provide true hub and spoke connectivity to subscribers, who can be connected to the:**

    Same or different PE-router(s)

    Same or different VRFs, via the upstream ISP

Technical Justification

- **Problem**

  PE requires multiple VRF tables for multiple VRFs to push spoke traffic via hub

  If the spokes are in the same VRF (no HDV), traffic will be switched locally and will not go via the hub site

- **Solution**

  HDVs allows all the spoke site routes in one VRF

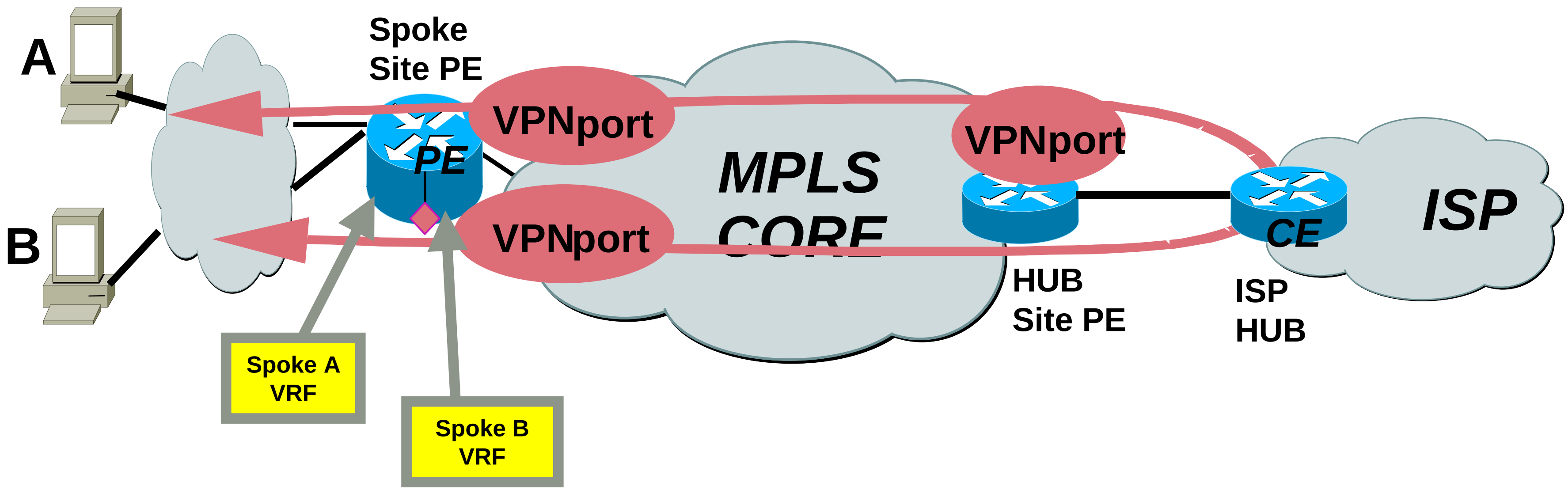
- **Benefit**

  Scalability for Remote Access to MPLS connections

  Reduces memory requirements by using just two VRF tables

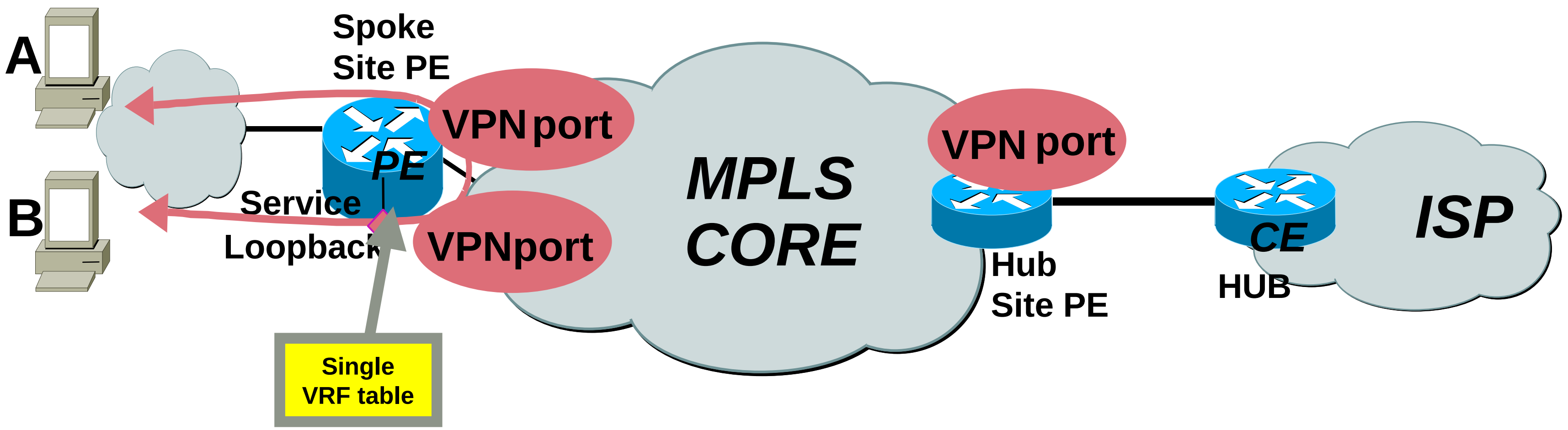
  Simplifies provisioning, management, and troubleshooting by reducing the number of Route Target and Route Distinguisher configuration

Hub & Spoke Connectivity Without HDV Requires Dedicated VRF Tables Per Spoke



- **All the spokes in the same VPN (yellow)**
- **Dedicated (separate) VRF per spoke is needed to push all traffic through upstream ISP Hub**

Hub & Spoke Connectivity Without HDV Using A Single VRF

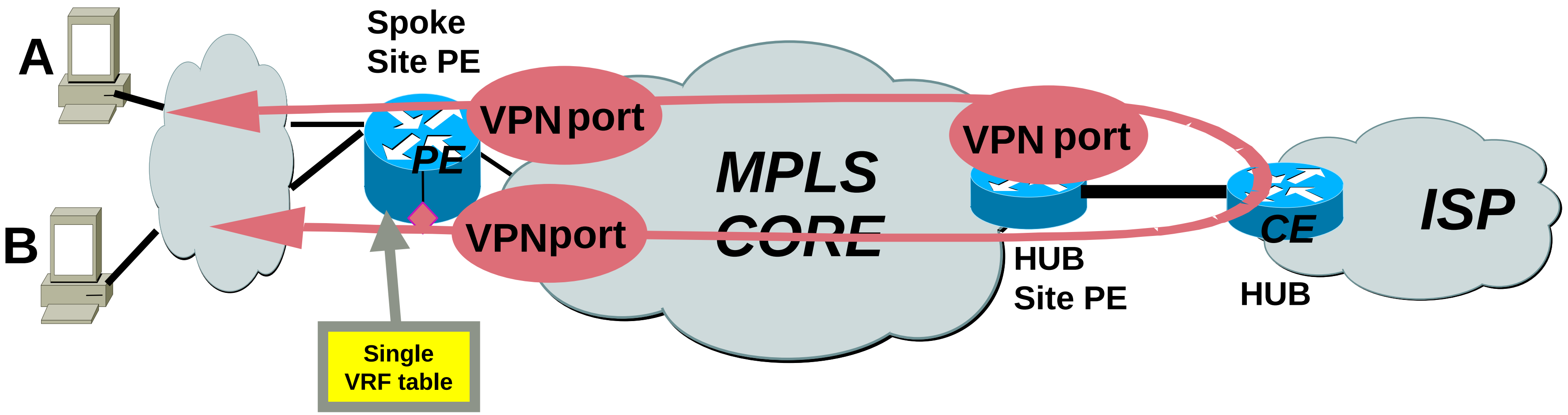


- If two subscribers of the same service terminate on the same PE-router, then traffic between them can be switched locally at the PE-router (as shown), which is undesirable
- All inter-subscriber traffic needs to follow the default route via the Home Gateway (located at upstream ISP).

Terminology

- **Upstream VRF**
  - **Used to forward packets from Spokes to Hub**
  - **Contains a static default route**
- **Downstream VRF**
  - **Use to forward packets from Hub to Spoke**
  - **Contains a /32 route to a subscriber (installed from PPP)**

Hub & Spoke Connectivity With HDV Using A Single VRF



- If two subscribers of the same service terminate on the same PE-router, traffic between them is not switched locally
- All inter-subscriber traffic follows the default route via the Home Gateway (located at upstream ISP)

Half Duplex VRF Functionality

1. HDVs are used in only one direction by incoming traffic

   Ex: upstream toward the MPLS VPN backbone or downstream toward the attached subscriber

2. PPP client dial, and is authenticated, authorized, and assigned an IP address.

3. Peer route is installed in the downstream VRF table

   One single downstream VRF for all spokes in the single VRF

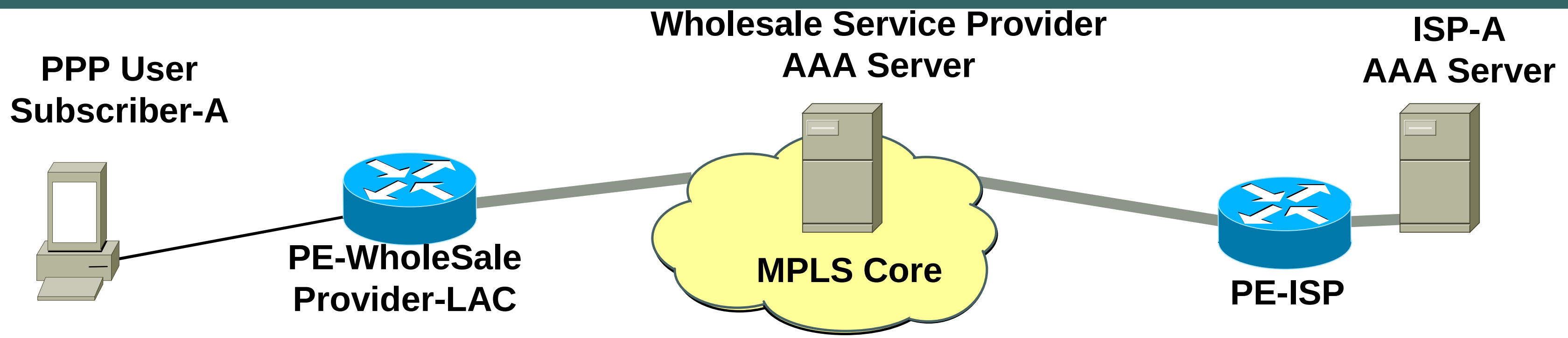
4. To forward the traffic among spokes (users), upstream VRF is consulted at the Spoke PE and traffic is forwarded from a Hub PE to Hub CE

   Return path: downstream VRF is consulted on the Hub PE before forwarding traffic to appropriate spoke PE and to the spoke (user)

5. Source address look up occurs in the downstream VRF, if unicast RPF check is configured on the interface on which HDV is enabled

Subscriber Connection Process

Cisco.com



1. PPP user initiates a session with PPP session using a name Subscriber-A@ISP-A.com and password
2. LAC/PE-router sends username information to the WholesaleServiceProvider Radius Server
3. ISP-A (service name) is used to index into a profile that contains information on the IP address of the Radius server of the ISP-A
4. Subscriber-A@ISP-A.com and password is then forwarded from the Wholesale Provider Radius server (which acts as a "proxy-radius"), towards the ISP Radius server
5. ISP-A Radius server authenticates and assigns IP address
6. ISP-A Radius server sends "Access-Accept" to Wholesale Service Provider Radius Server
7. The wholesale Service Provider Radius server adds authorization information to the Access-Accept, (based on the domain or servicename)and the VRF to be used by Subscriber-A, and forwards it to PE-WholesaleProvider-LAC router
8. PE-WholesaleProvider-LAC router creates temporary Virtual-Access interface (with associated /32 IP address) and places it into the appropriate VRF

Reverse Path Forwarding Check

- **Reverse Path Forwarding (RPF)**

  Used by Service Provider determine the source IP address of an incoming IP packet and ascertain whether it entered the router via the correct inbound interface

- **Concern**

  HDV populates a different VRF than the one used for “upstream” forwarding

- **Solution**

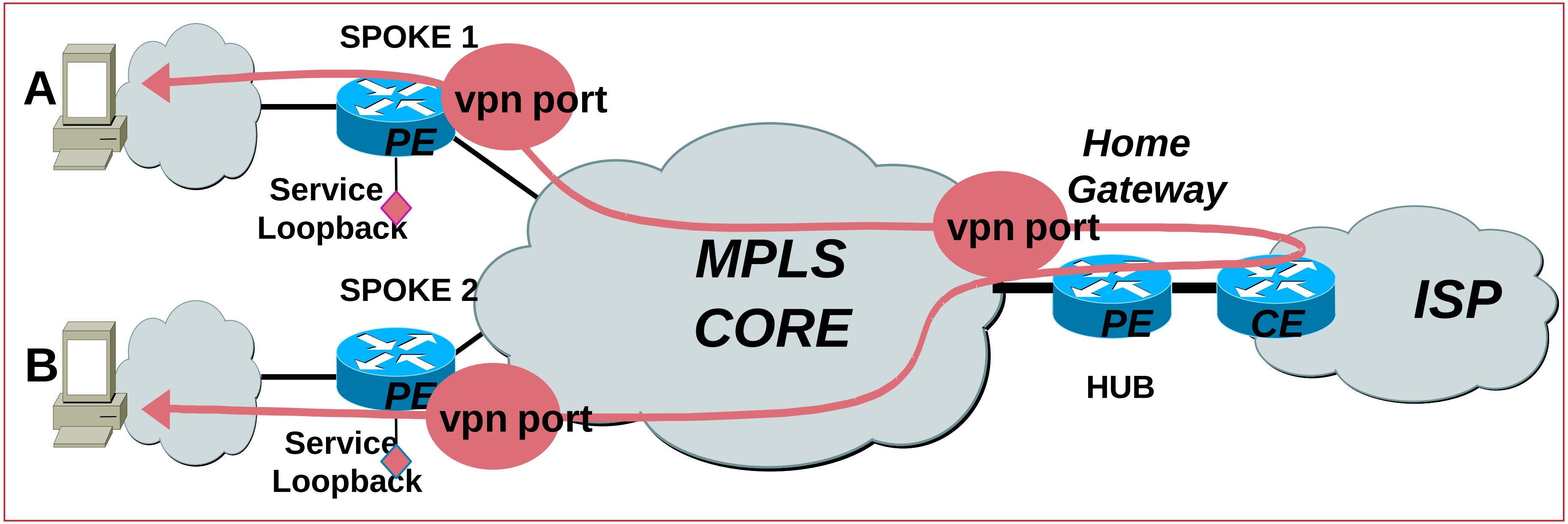
  Extend the RPF mechanism so the “downstream” VRF is checked

- **To enable RPF extension, configure:**

```
ip verify unicast reverse-path <downstream vrfname>
```

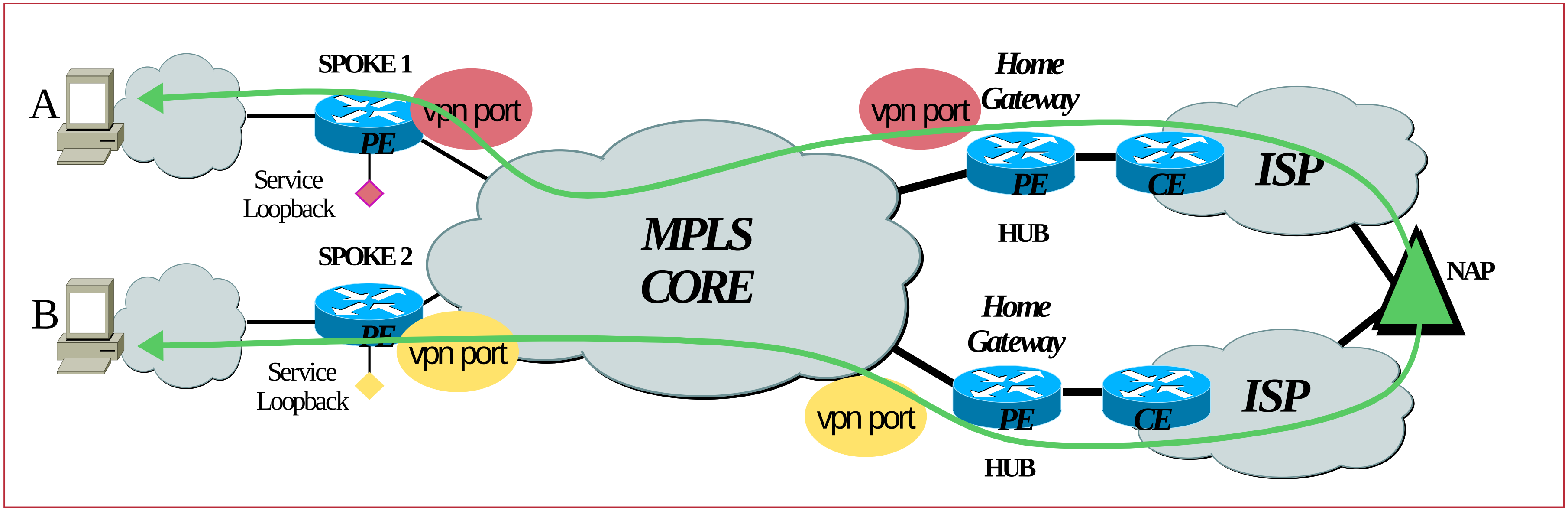
Topology I: Hub and Spoke Connectivity Between Distributed PE-Routers

- Upstream traffic (ie: traffic toward the upstream ISP or toward another subscriber) is sent to the hub PE-router and forwarded across the link between the wholesale SP and the ISP
- Subscriber traffic follows a default route within the VRF
- Traffic is forwarded towards and received from the wholesale Service Providers PE-router and the subscriber



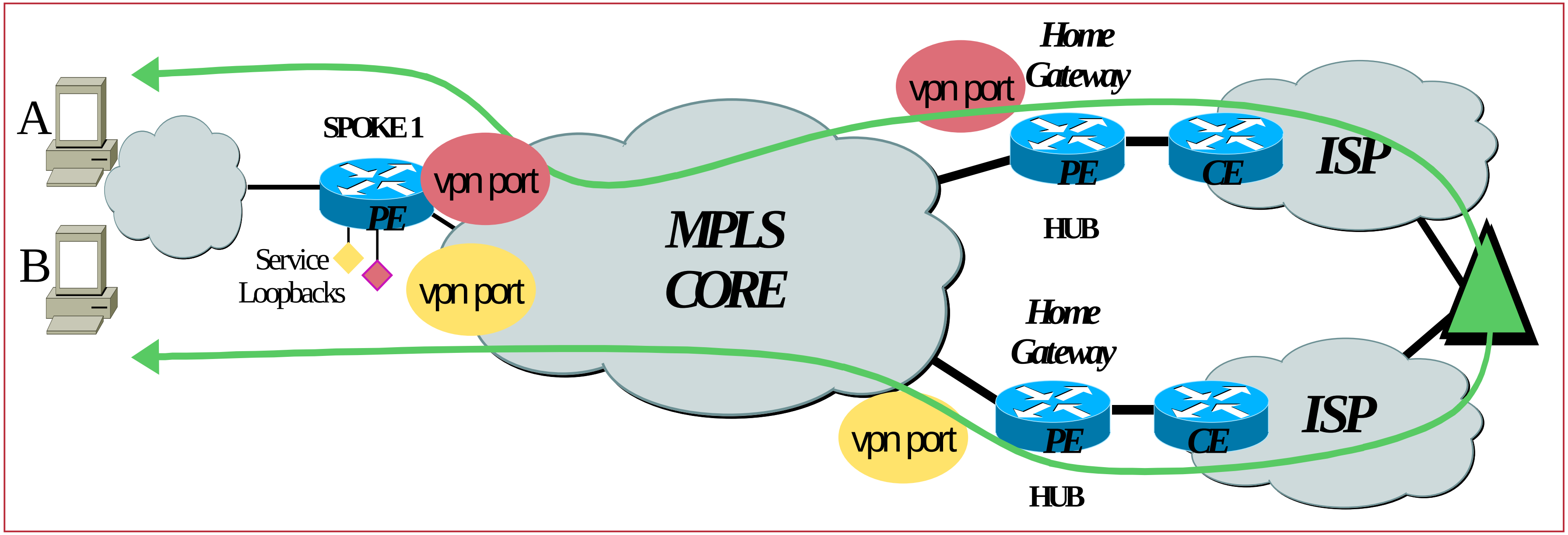
Topology II: Hub and Spoke Connectivity Between Subscribers Of Different Services

- **Data flow between two subscribers that belong to different services goes through the hub location of the Service Provider**
- **Data will traverse through a network exchange point, either public or private, by following a default route within the subscriber VRF**



Topology III: Hub and Spoke Connectivity Via the Same PE-Router (Different Services)

- **If two subscribers are terminated on the same PE-router and belong to different services, the data is required to traverse through the home gateways of both services.**



Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

VPN Connectivity between AS#s

- **VPN sites may be geographically dispersed**

  Requiring connectivity to multiple providers, or different regions of the same provider

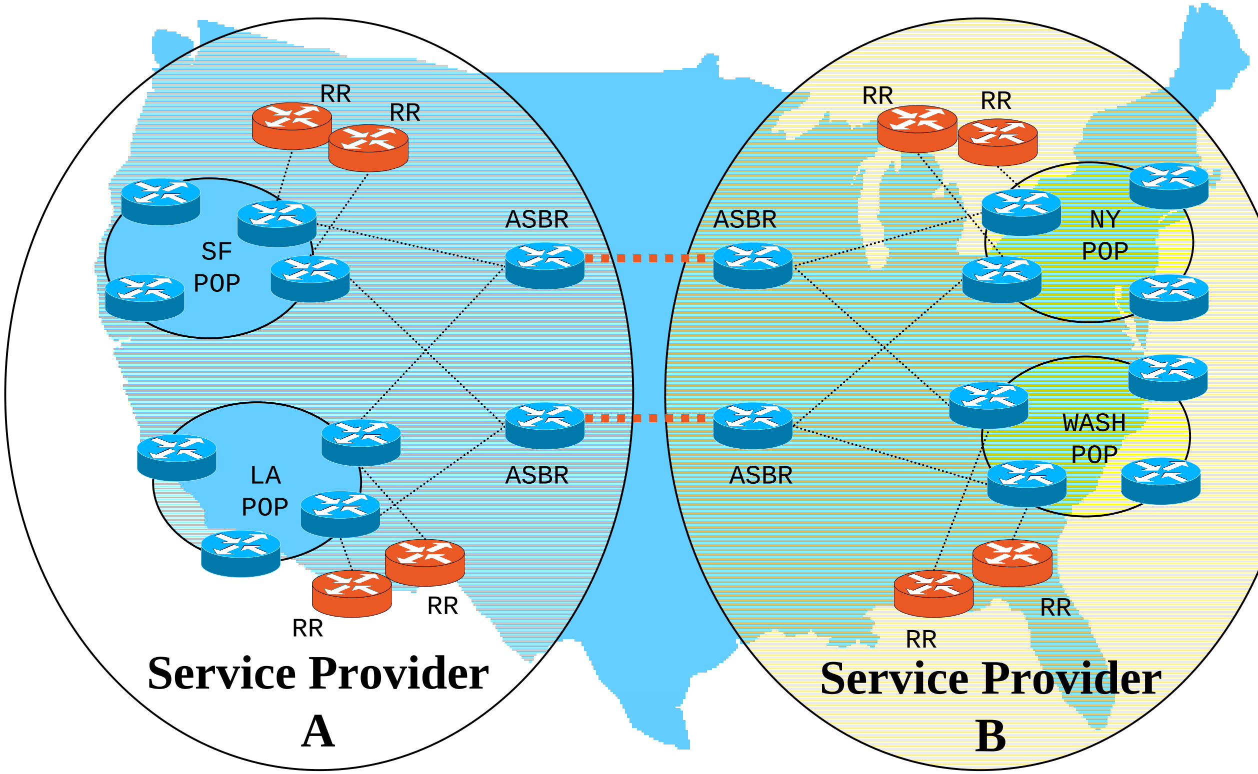
- **Transit traffic between VPN sites may pass through multiple AS#s**

  This implies that routing information MUST be exchanged across AS#s

- **Distinction drawn between Inter-Provider & Inter-AS**

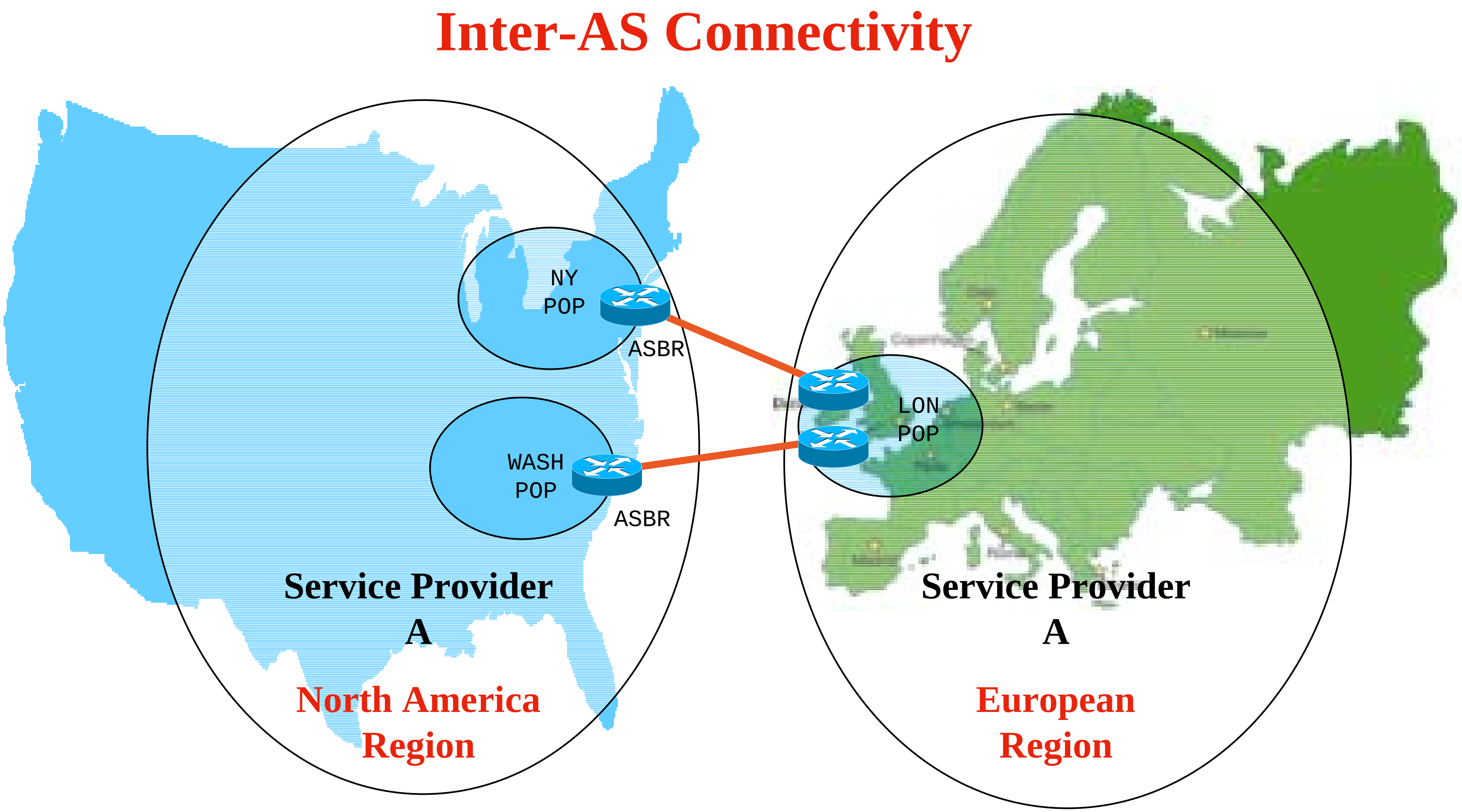
Inter-Provider Vs. Inter-AS

Inter-Provider Connectivity

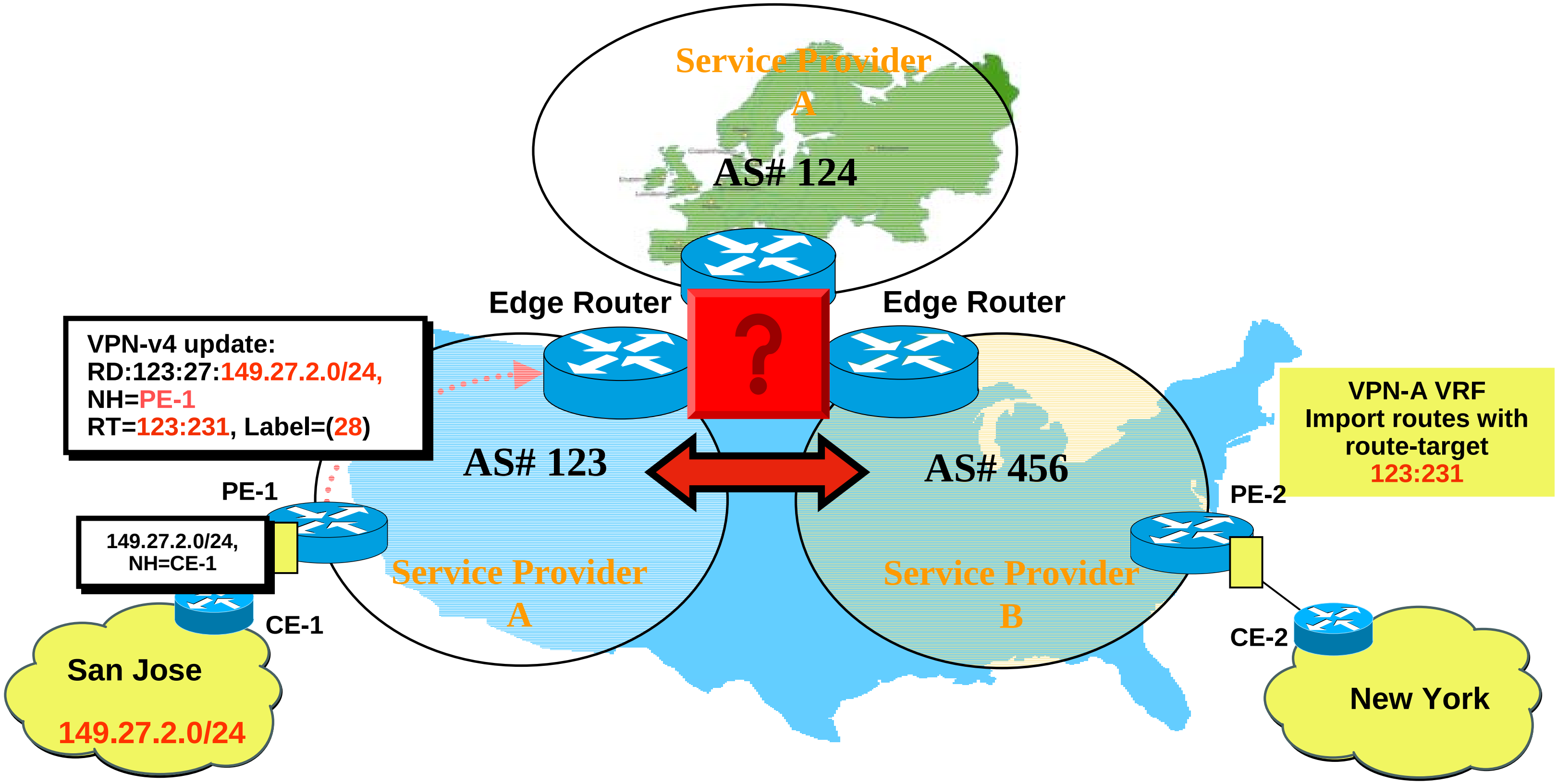


Inter-Provider Vs Inter-AS

Cisco.com

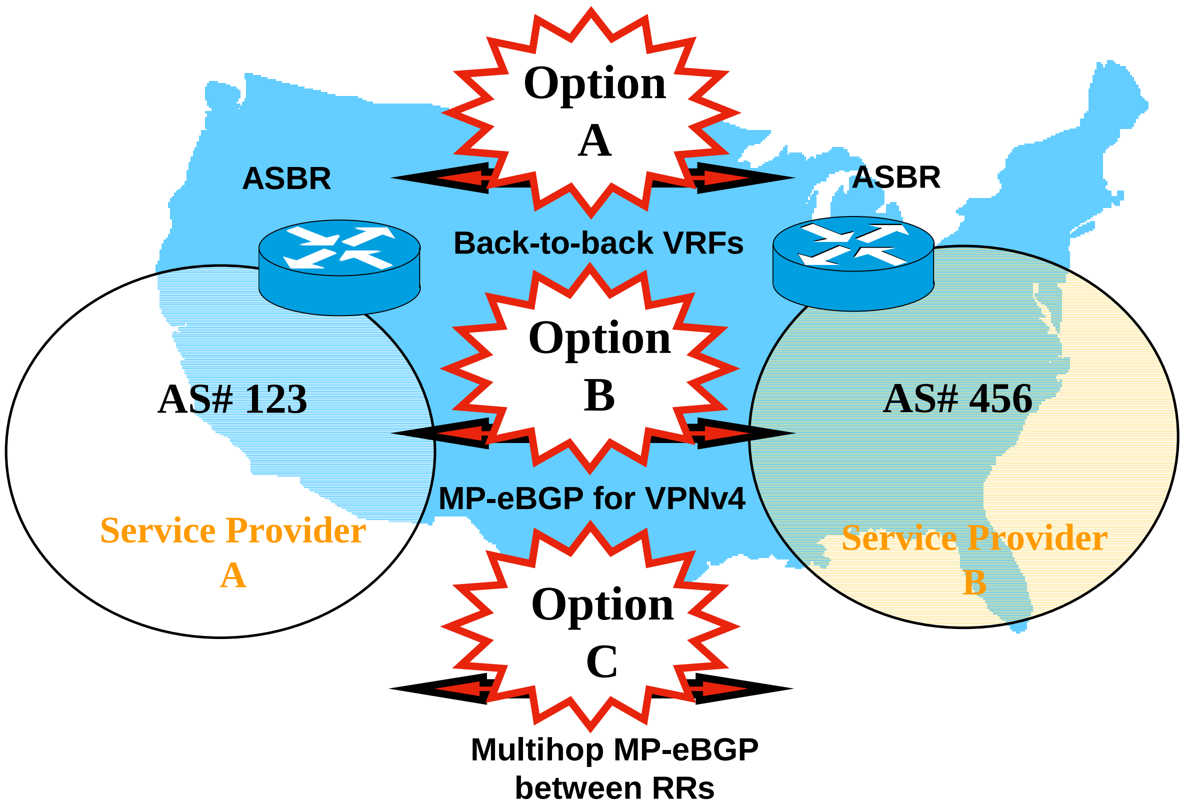


VPN Route Distribution



How to distribute VPNv4 routes between different AS's ?

VPN Route Distribution Options

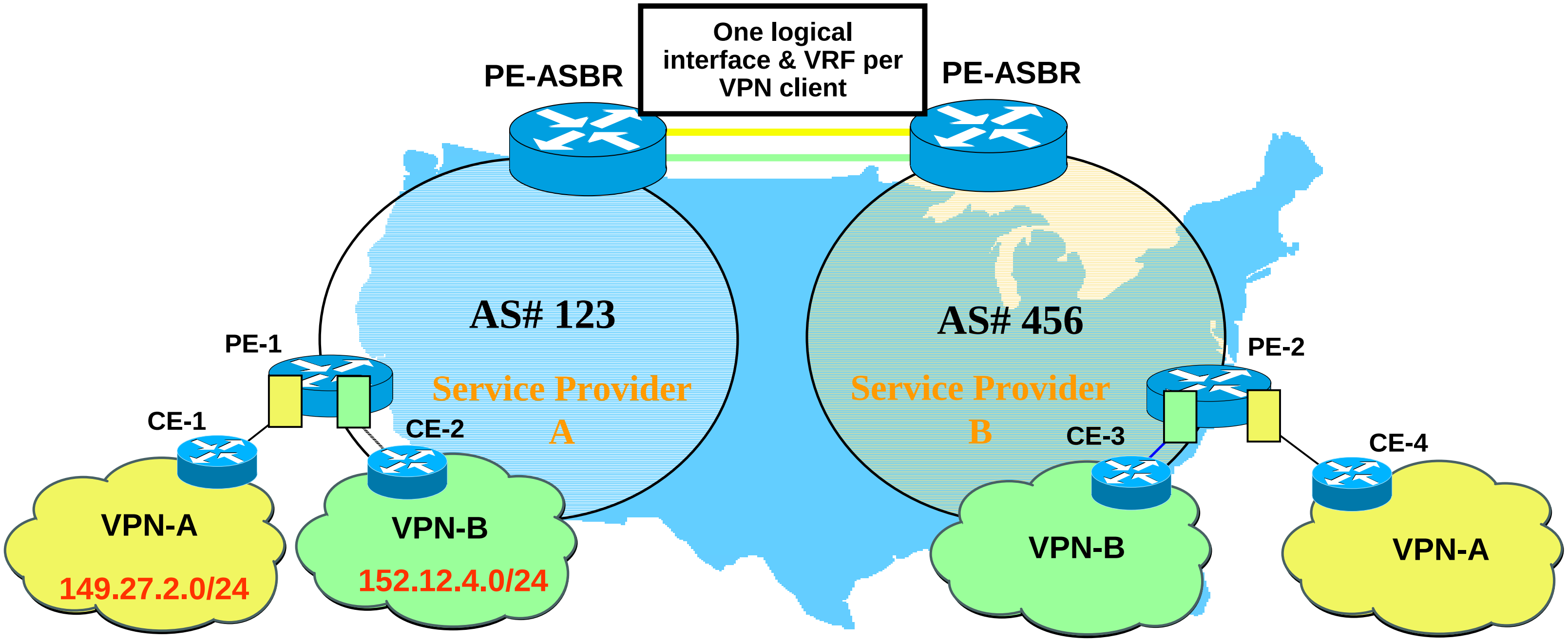


Several options available for route distribution

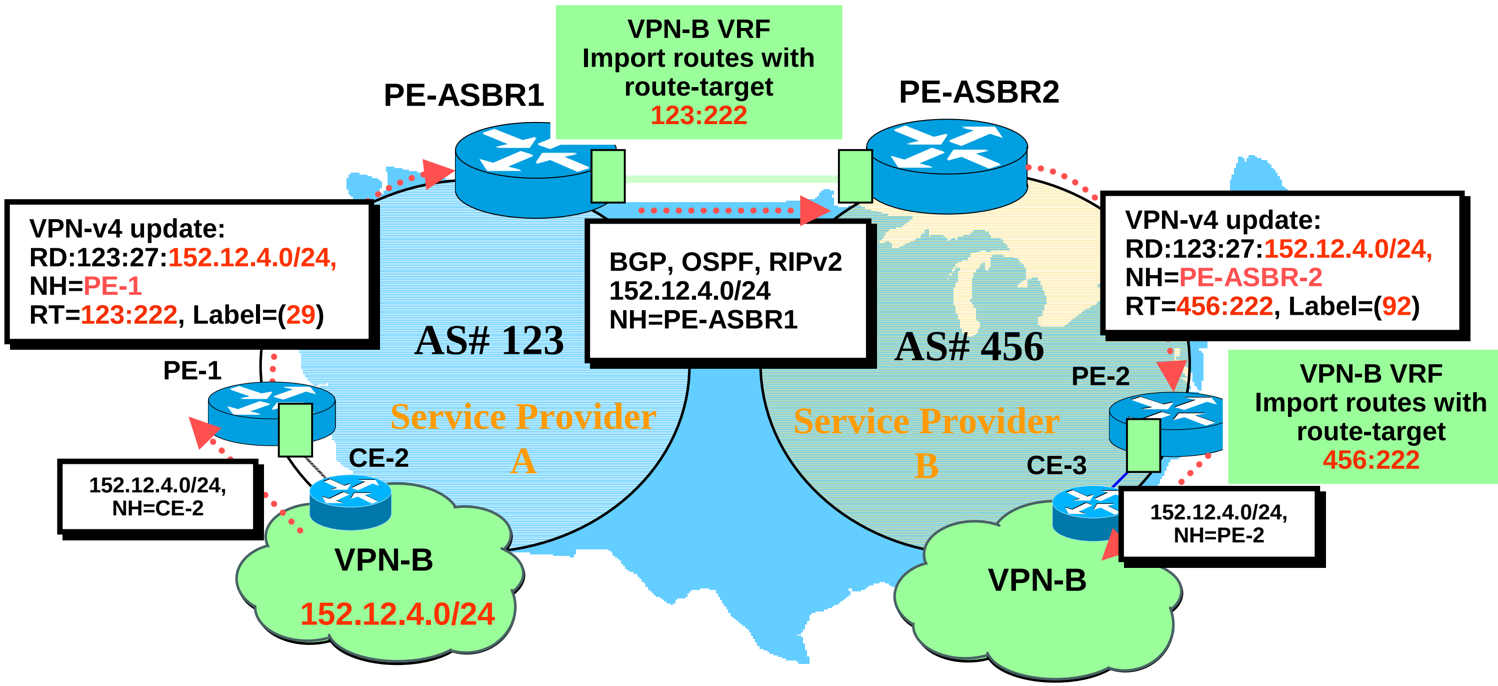
Option A – Back-to-back VRFs

- **2547 providers exchange routes between ASBRs over VRF interfaces**
  - **Hence ASBR is known as a PE-ASBR**
- **Each PE-ASBR router treats the other as a CE router**
  - **Although both provider interfaces are associated with a VRF**
- **Provider edge routers are gateways used for VPNv4 route exchange**
- **PE-ASBR link may use any PE-CE routing protocol**

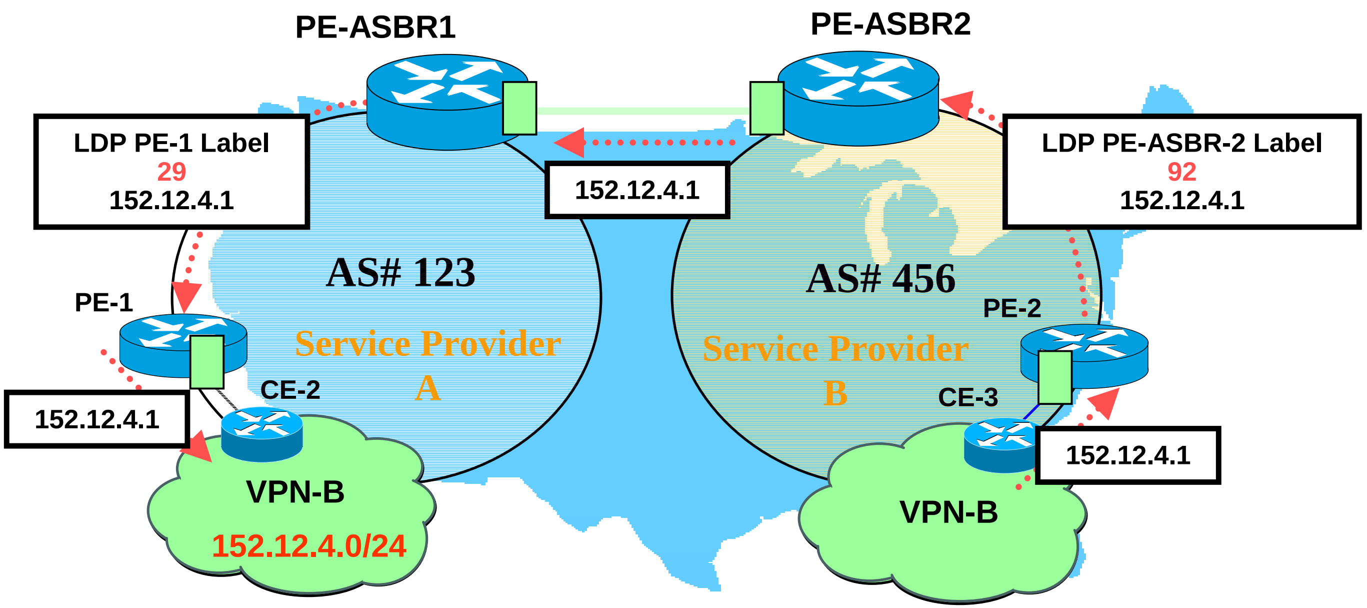
Back-to-back VRF Connectivity Model



Back-to-back Prefix Distribution



Back-to-back Packet Flow



Back-to-back VRFs Summary

- Scalability is an issue with many VPNs
  - 1 VRF & logical interface per VPN
  - Gateway PE-ASBR must hold ALL routing information
- PE-ASBR must filter & store VPNv4 prefixes
- No MPLS label switching required between providers
  - Standard IP between gateway PE-ASBRs
  - No exchange of routes using External MP-BGP
  - Simple deployment but limited in scope
  - However, everything just works

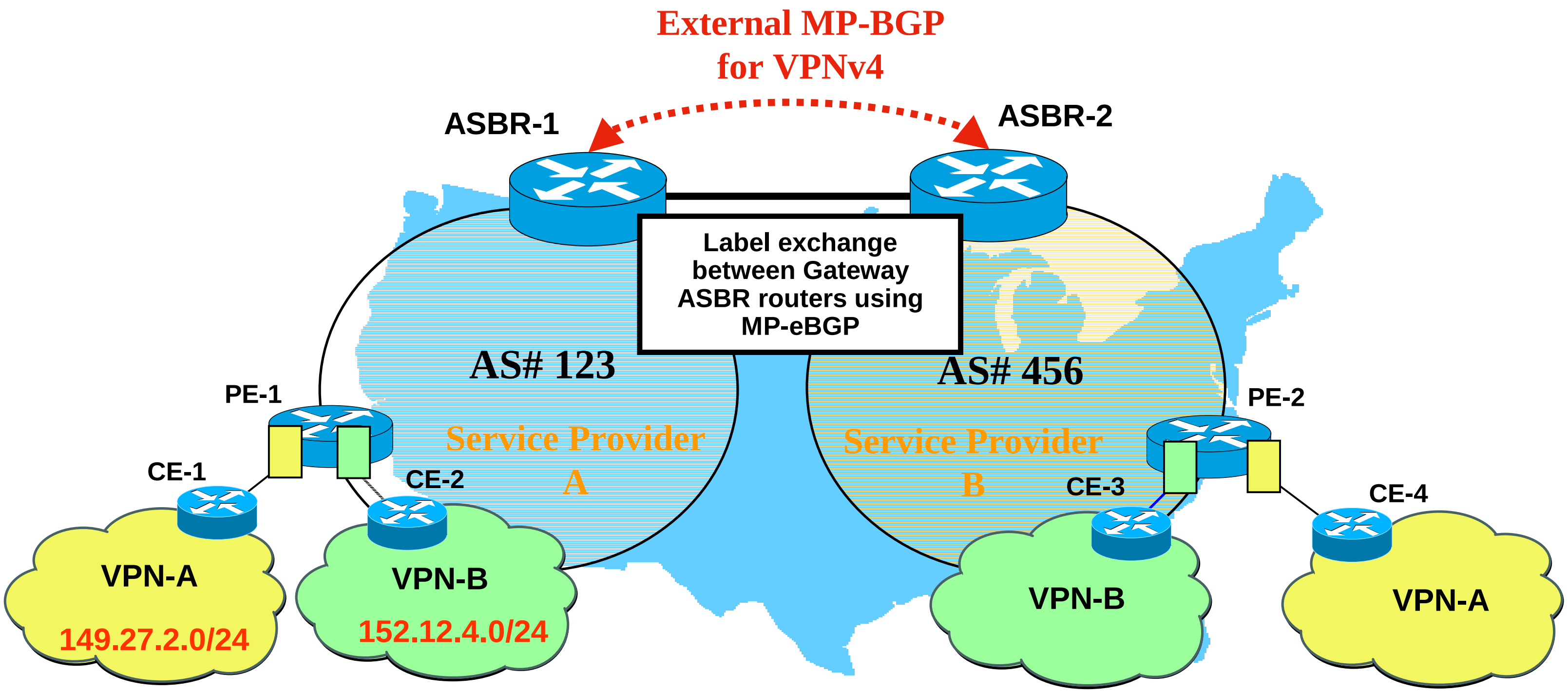
Option B – External MP-BGP

- **Gateway ASBRs exchange VPNv4 routes directly**
  - External MP-BGP for VPNv4 prefix exchange. No LDP/IGP
- **BGP next-hop set to advertising ASBR**
  - Next-hop/labels are rewritten when advertised across ASBR-ASBR link
- **ASBR stores all VPN routes that need to be exchanged**
  - But only within the BGP table. No VRFs. Labels are populated into LFIB at ASBR

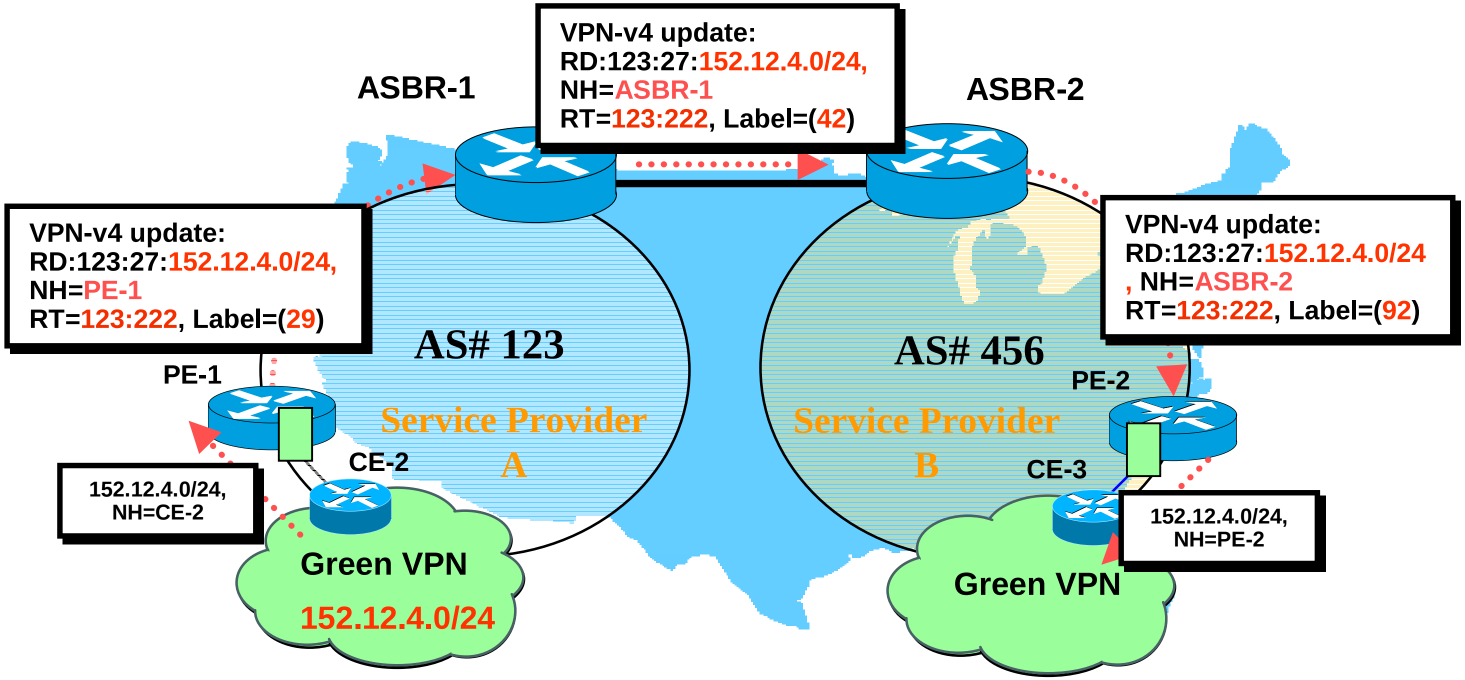
Label allocation at receiving PE-ASBR

- **Receiving gateway ASBR may allocate new label**
  - Controlled by configuration of next-hop-self
  - LFIB holds new label allocation
- **Receiving ASBR automatically creates a /32 host route for its ASBR neighbor**
  - Which must be advertised into receiving IGP if next-hop-self is not in operation (to maintain the LSP)

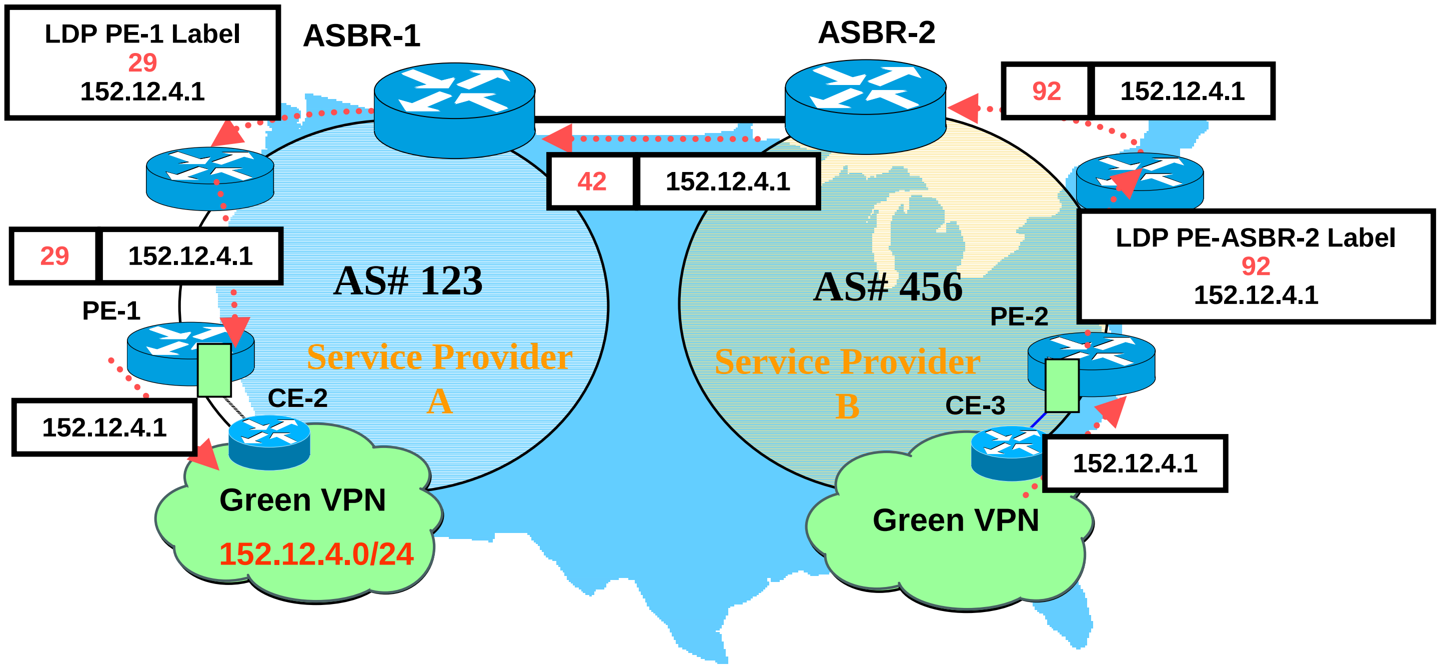
External MP-BGP Connectivity Model



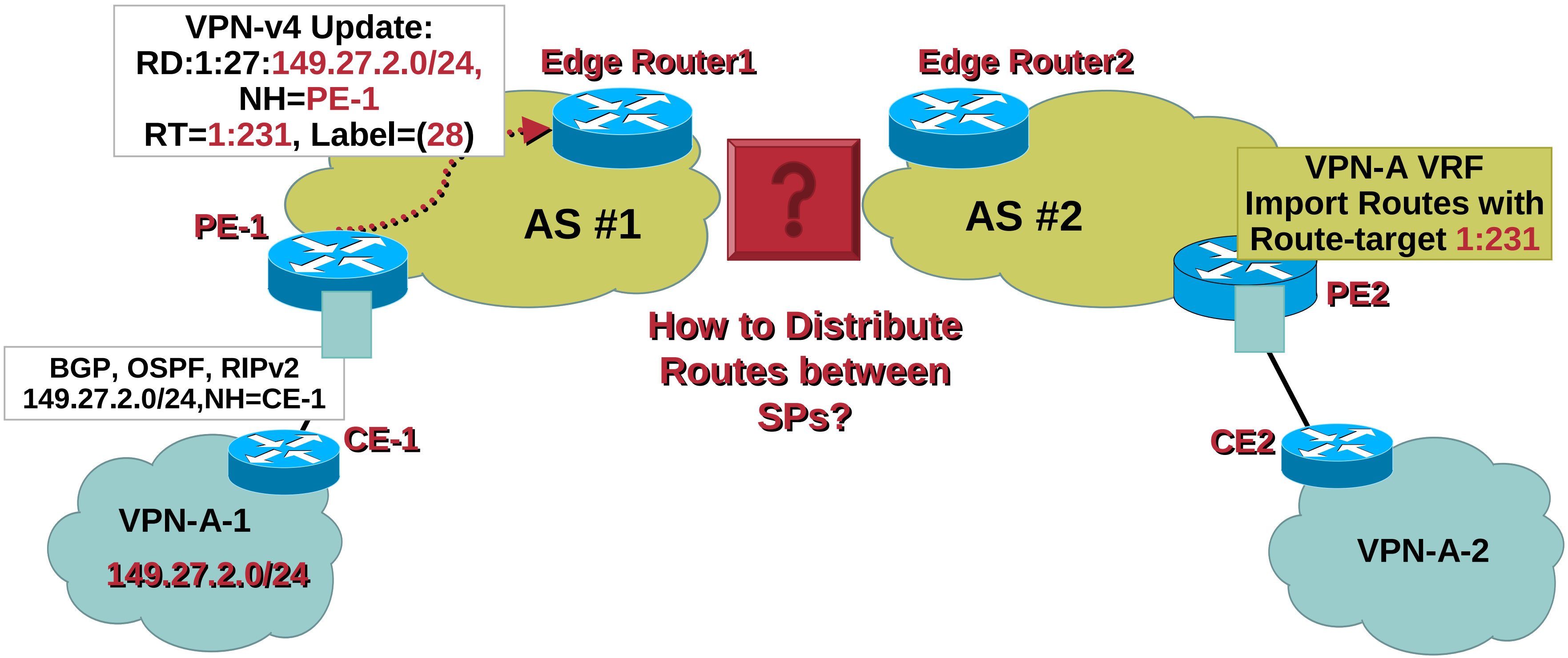
External MP-BGP Prefix Distribution



External MP-BGP Packet Flow



VPN Client Connectivity



VPN Sites Attached to Different MPLS VPN Service Providers

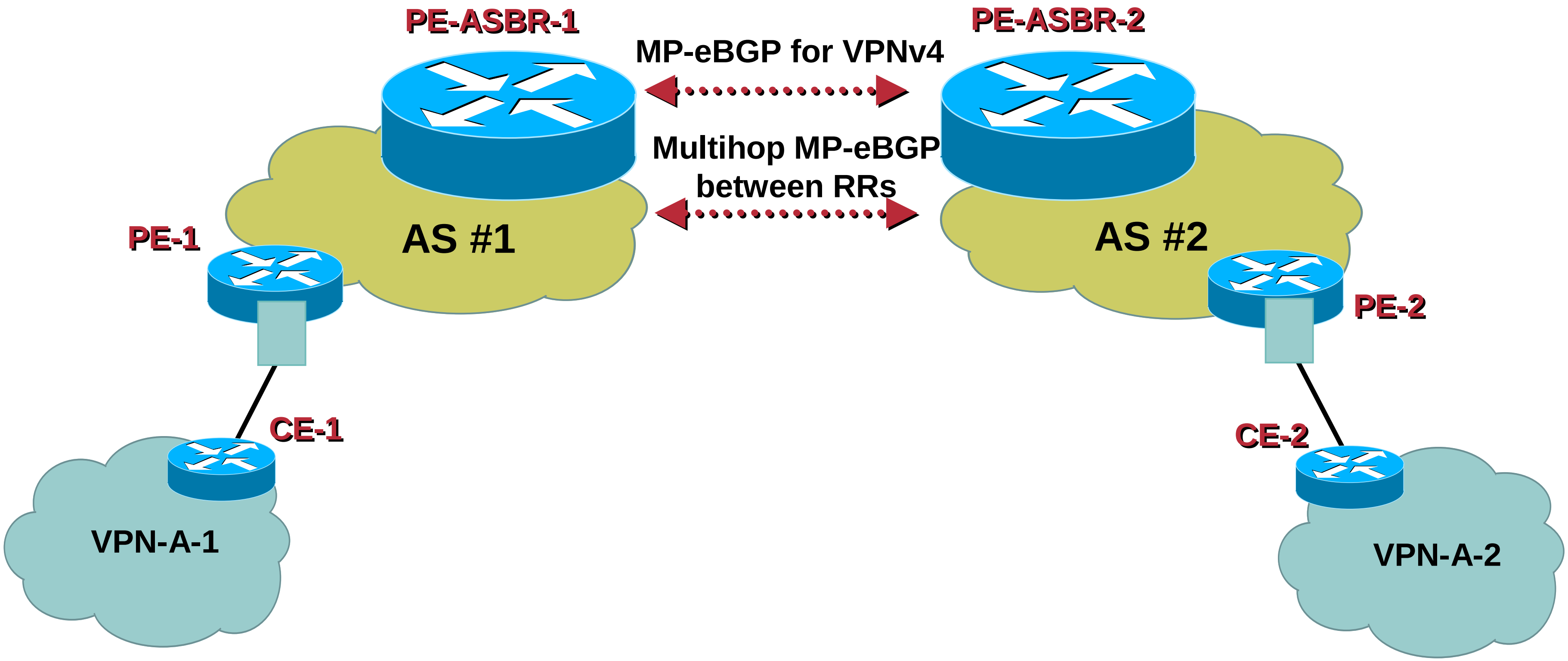
External MP-BGP Summary

- Scalability less of an issue when compared to back-to-back VRF connectivity
  - Only 1 interface required between ASBR routers
  - No VRF requirement on any ASBR router
- Automatic route filtering must be disabled
  - Hence filtering on RT values essential
  - Import of routes into VRFs is NOT required (reduced memory impact)
- Label switching required between ASBRs

External MP-BGP Summary (Cont).

- **Preferred option for Inter-Provider connectivity**
  - No IP prefix exchange required between providers
  - Security is tighter
  - Peering agreements specify VPN membership

VPNv4 Distribution Options



Other Options Available,
These Two Are the Most Sensible

ASBR Router Protection/Filtering

- MP-eBGP session is authenticated with MD5
  - Potentially also IPSec in the data plane
- Routing updates filtered on ingress based on extended communities
  - Both from internal RRs and external peerings
  - ORF used between ASBRs and RRs.
  - Maximum-prefix on MP-BGP session
- Per-interface label space for external facing links to avoid label spoofing

Option C – Multihop MP-eBGP between RRs

- 2547 providers exchange VPNv4 prefixes via RRs

  Requires multihop MP-eBGP session

- Next-hop-self MUST be disabled on the RRs

  Preserves next-hop/label as allocated by originating PE router

- Providers exchange IPv4 routes with labels between directly connected ASBRs using External BGP

  Only PE router BGP next-hop addresses exchanged

  RFC3107 "Carrying Label Information in BGP-4"

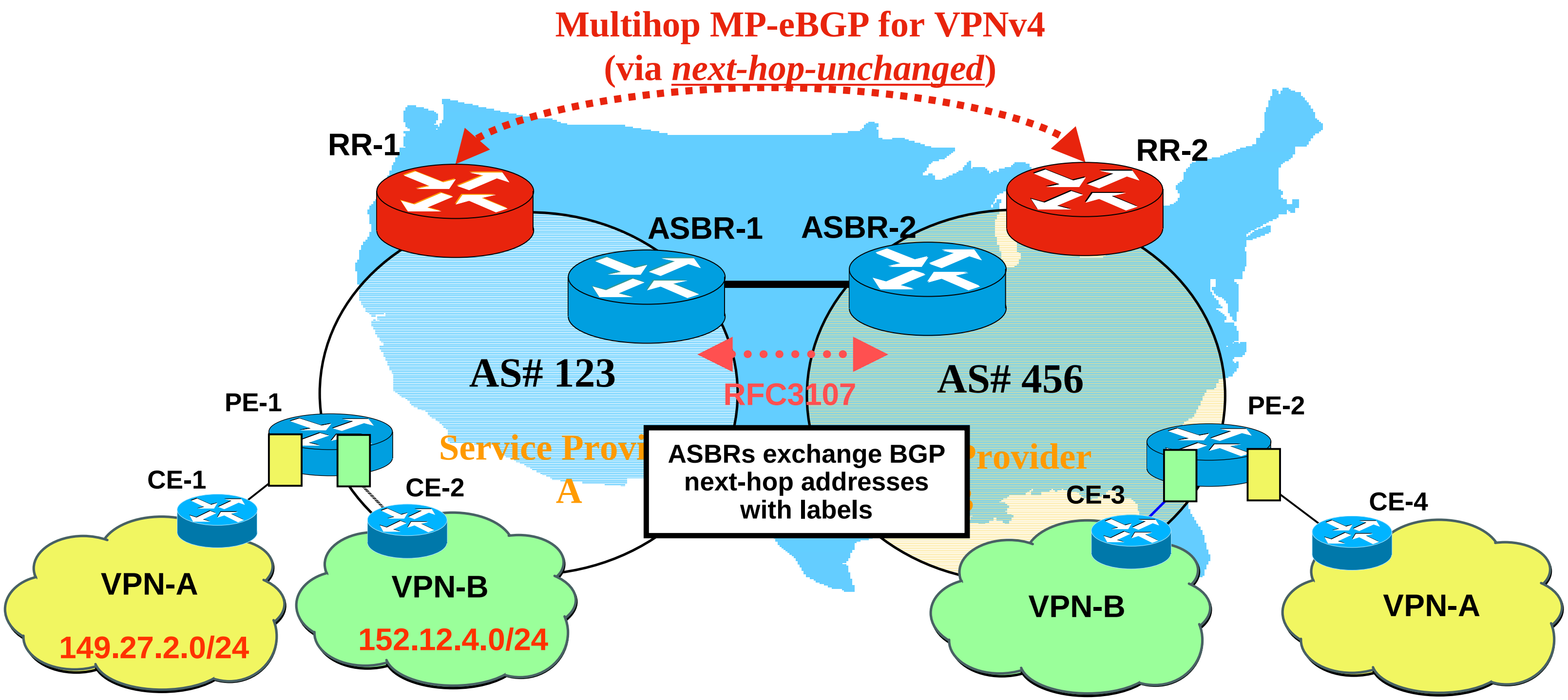
RFC3107 – Carrying labels with BGP-4


```
0                   1                   2                   3
  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
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | Address Family Identifier (1) |   SAFI (4)    | Next-hop Lth  |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |          Network Address of next-hop (variable)               |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | # of SNPAs    |  Network Layer Reachability Info (variable)   |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |    Length     |                    MPLS Label                 |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |               |          Prefix (variable)                    |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

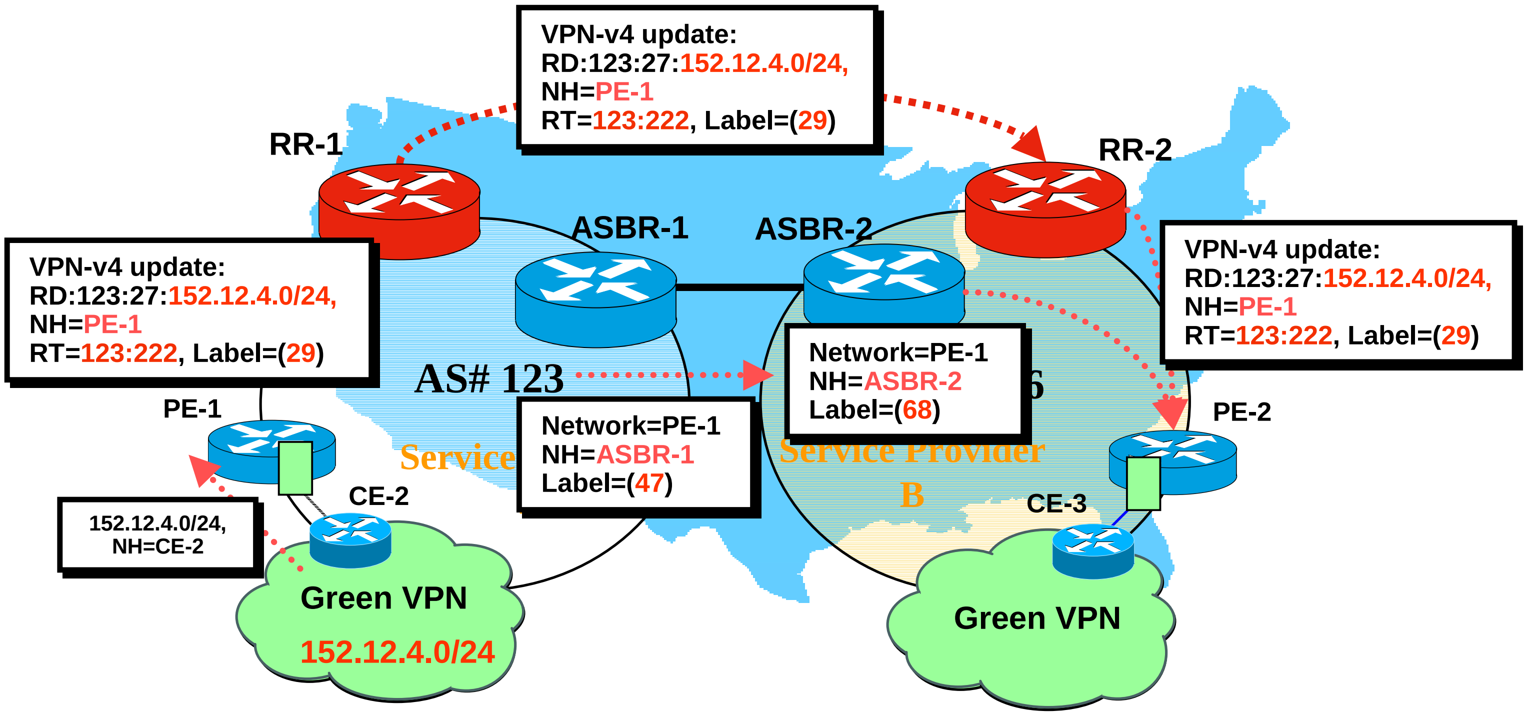
MP_REACH_NLRI Attribute
(Specified in RFC 2858)

Prefix plus MPLS label
(Specified in RFC 3107)

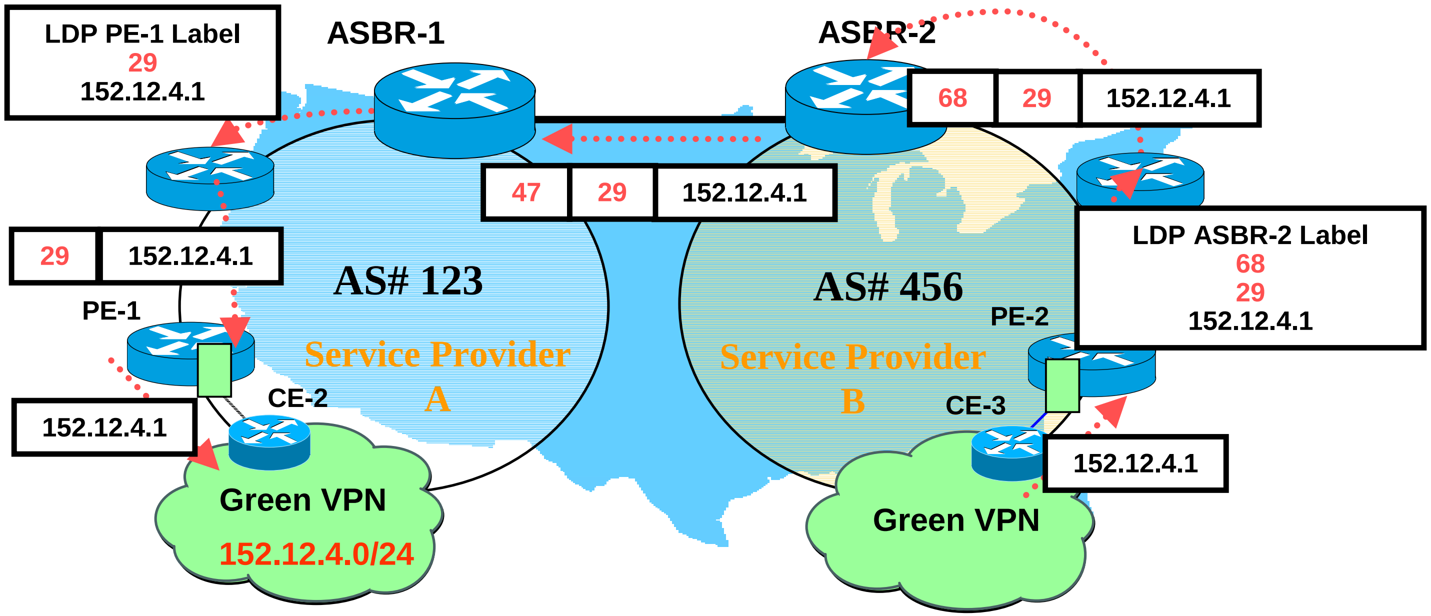
Multihop MP-eBGP Connectivity Model



Multihop MP-eBGP Prefix Distribution



Multihop MP-eBGP Packet Flow



Multihop MP-eBGP Summary

- **More scalable than previous options**

  As all VPNv4 routes held on route reflectors rather than the ASBRs

- **Route reflectors hold VPNv4 information**

  Each provider utilizes route reflectors locally for VPNv4 prefix distribution

  External BGP connection added for route exchange

- **BGP next-hops across ASBR links using RFC3107**

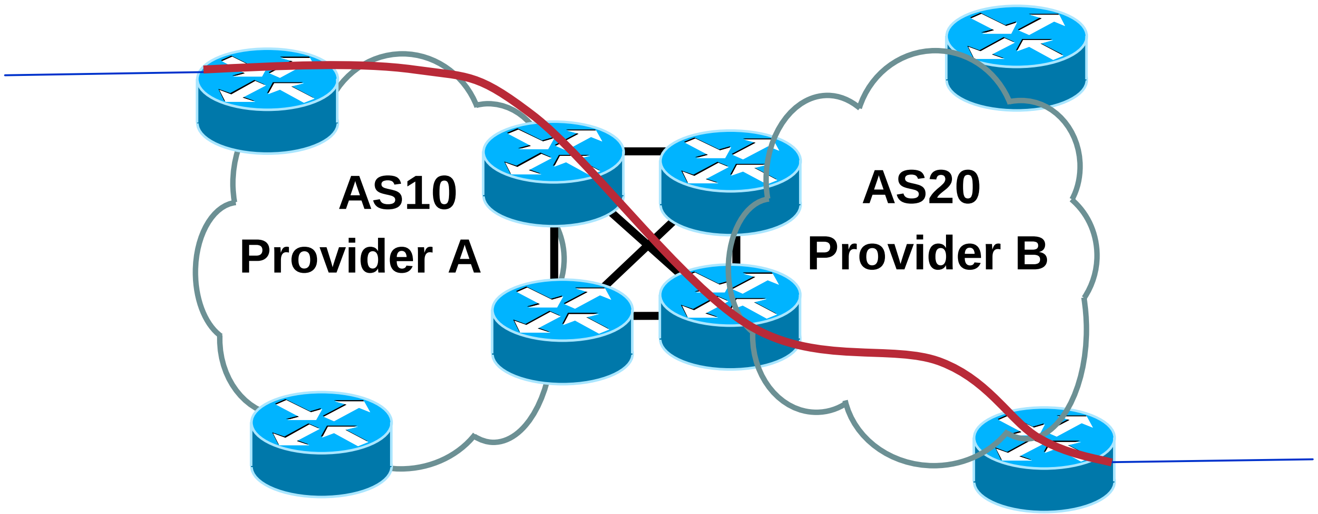
  Separation of forwarding/control planes

Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Inter-provider PW



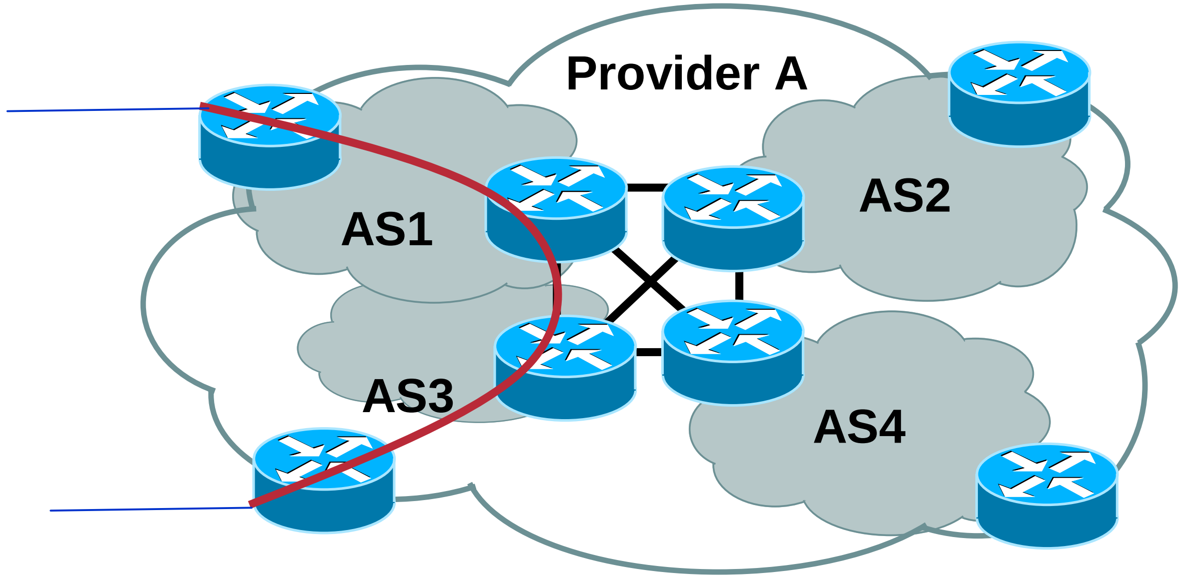
We will refer to an Inter-provider model when a pseudo-wire circuit will span across 2 different service providers domains or AS's

- In this model, the SP will have "no" or "very limited" trust between people managing different AS's…

- Different providers will certainly apply different QoS policies, definition and implementation.

- Inter-provider model will have to have mechanisms for Security and QoS mediation

Inter-AS PW

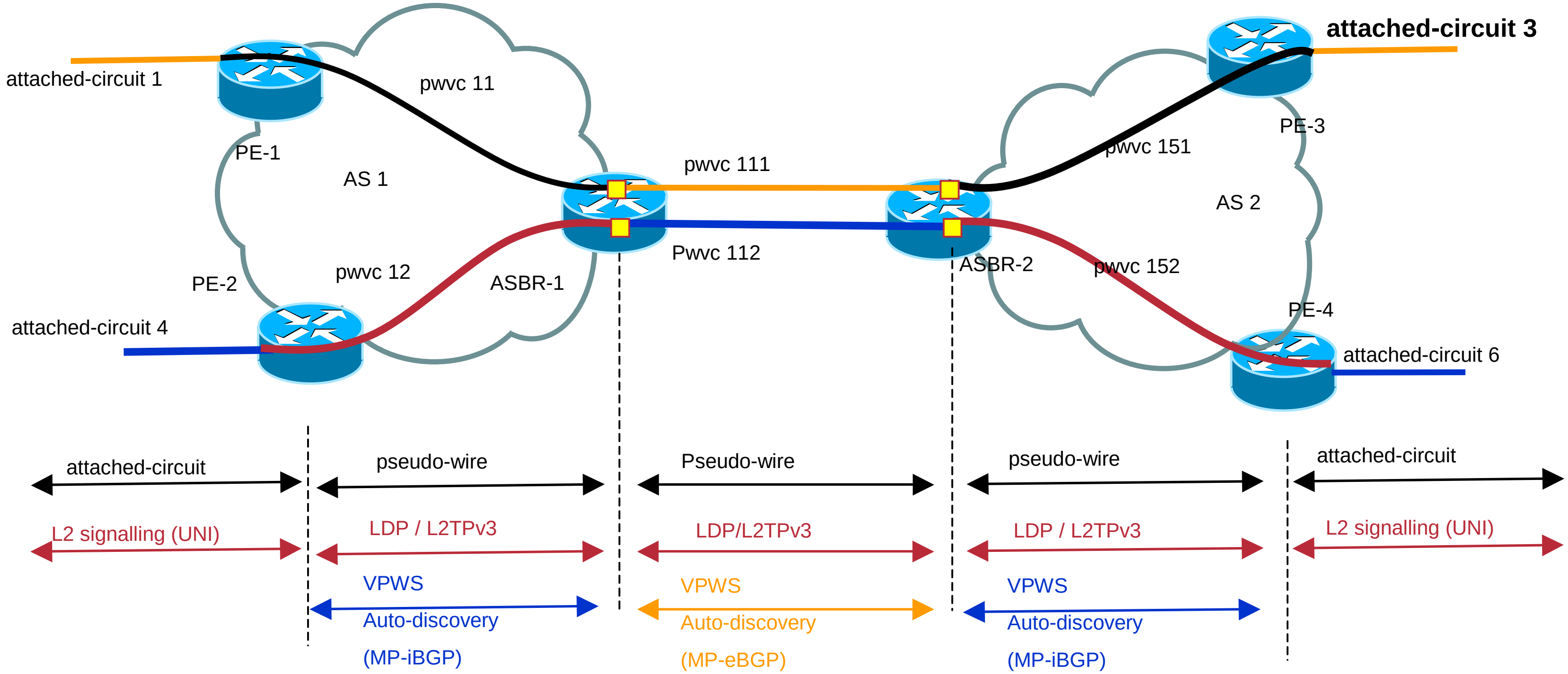


We will refer to Inter-AS model when a provider (Provider A) has, divided its network within multiple domain or ASes.

- In this model, degree of trust between people managing different ASes,

- In general QoS definition and implementation will be consistent across ASes

Pseudo-wire Stitching /Switching Model



Pseudo-wire stitching mechanism is the mechanism that permits a service provider to extend an existing pseudo-wire with an other pseudo-wire. In an other words, to replace the attached circuit by an other pseudowire from same type (atom pw with atom pw) or different type (atom pw with l2tpv3 pw).

Pseudo-wire Stitching model

Pro

-QoS model : Re-coloring of EXP value will works

- Security model : light trustiness (LDP, IGP cross boundary of SP’s but is limited to neighbour ASBR)

- Link between ASBR’s is independent of attached-circuit media, on same link, we could have ATM, FR, Ethernet pseudowire, and/or other services (IP, MPLS-VPN, …)

- De-jitter mechanism of De-cell-packing mechanism could occur only at egress PE’s

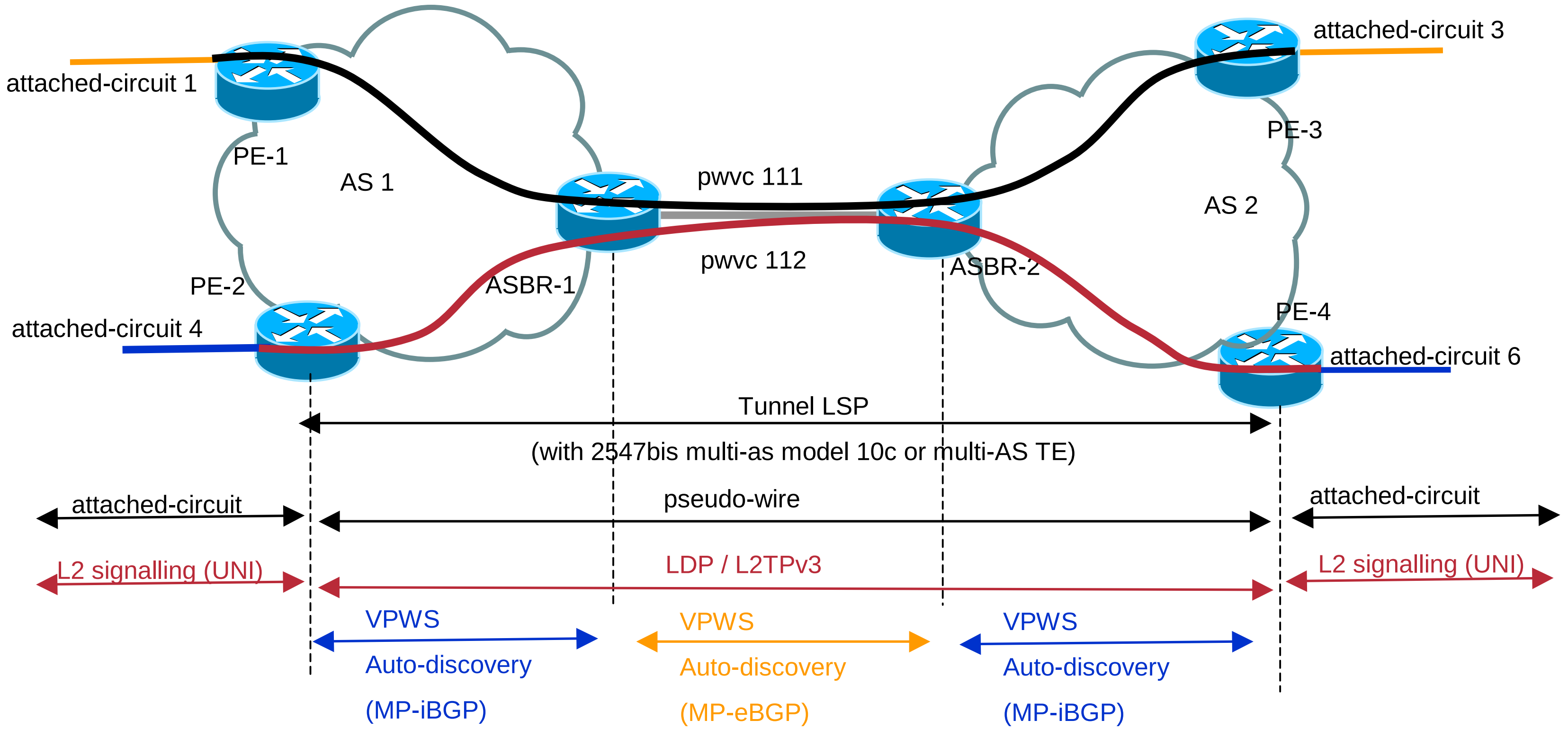
Cons

- Required to develop pseudowire stitching mechanism and/or to extend auto-discovery mechanism to support multi-as signalling.

- QoS Model: Lot’s of function like shaping and policing function on per pseudowire will required to be developed

- PW redundancy not optimized when NOT USING auto-discovery mechanism

Multi-AS tunnel LSP model



In this model we ruse existing RFC2457bis Multi-AS 10c or Multi-AS TE to build end-end tunnel LSP and to build end-end pseudowire VC's

Inter-AS tunnel LSP model

Pro

- Multi-AS model 10c or Inter-AS TE is developed.

- Link between ASBR's is independent of attached-circuit media, on same link, we could have ATM, FR, Ethernet pseudowire, and/or other services (IP, MPLS-VPN, …)

- PW redundancy can be optimized by optimizing end-end tunnel LSP technique

- De-jitter mechanism of De-cell-packing mechanism could occur only at egress PE's

- Ease to provisioning

Cons

- Security model : Untrusted (LDP, IGP cross boundary of ASes)

- QoS Model: Lot's of functions like CoS re-coloring, shaping and policing will not be possible at ASBR (VC labels have NO signification for ASBR).

In summary (what to deploy ?)

- When SP will connect 2 or more of their ASes together (Inter-AS model), the 2nd & 3th model will be certainly the most popular one.
- When the SP will connect to other SPs (Inter-Provider model), the 1st model will be certainly the most popular model to start with.
- If SP's start to have numerous circuits with some specific partners, then the second model may be interesting to consider.

Deployment/Architecture Challenges

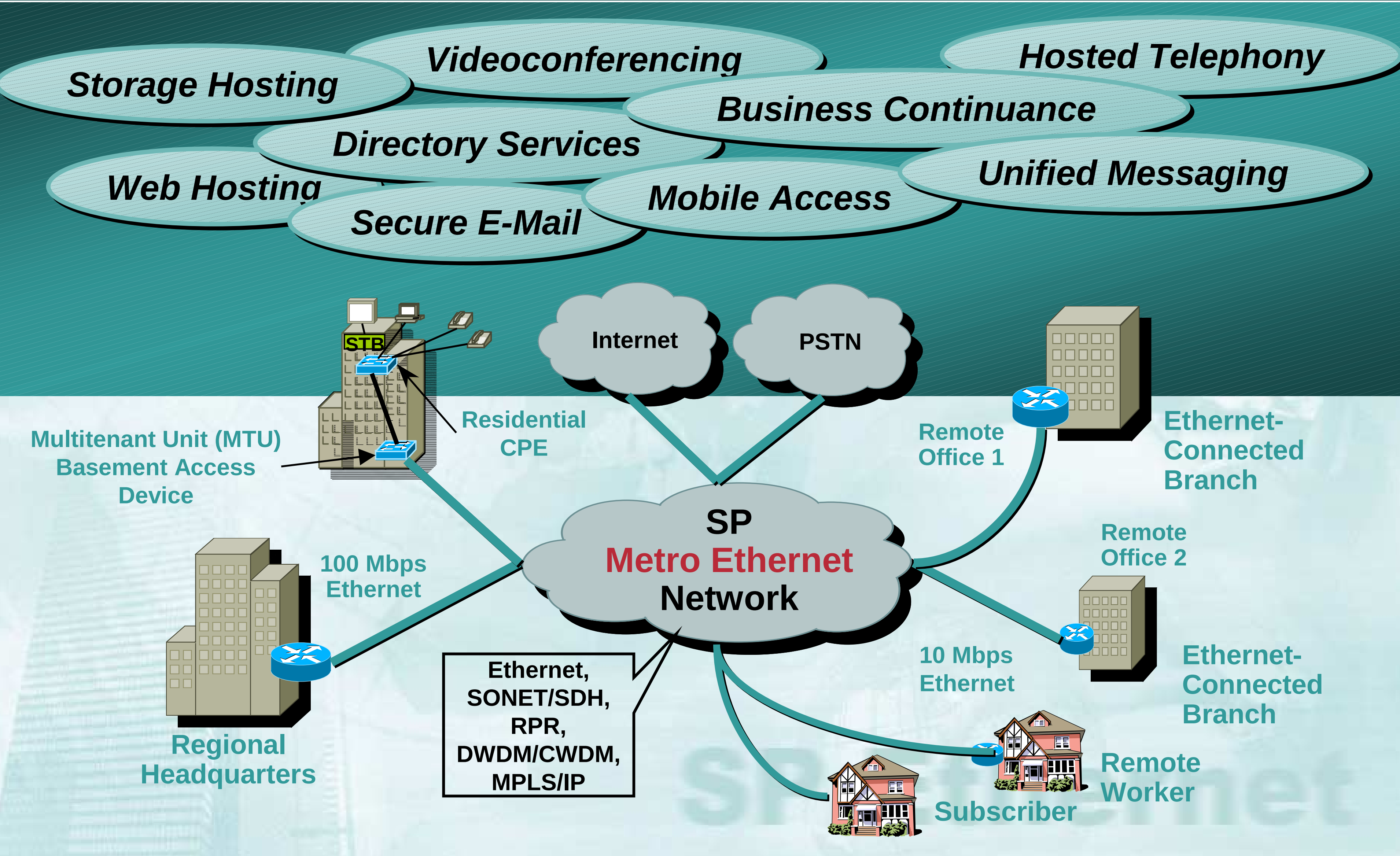
- As with all technologies there are challenges
  - Control-plane Scale
  - Filtering & route distribution
  - Security
  - Multicast
  - QOS/End-to-end SLA's
  - Integration of services e.g. Layer-2/Layer-3
  - Network Management
  - Traffic Engineering
- Opportunity for industry collaborative development!

Agenda

Cisco.com

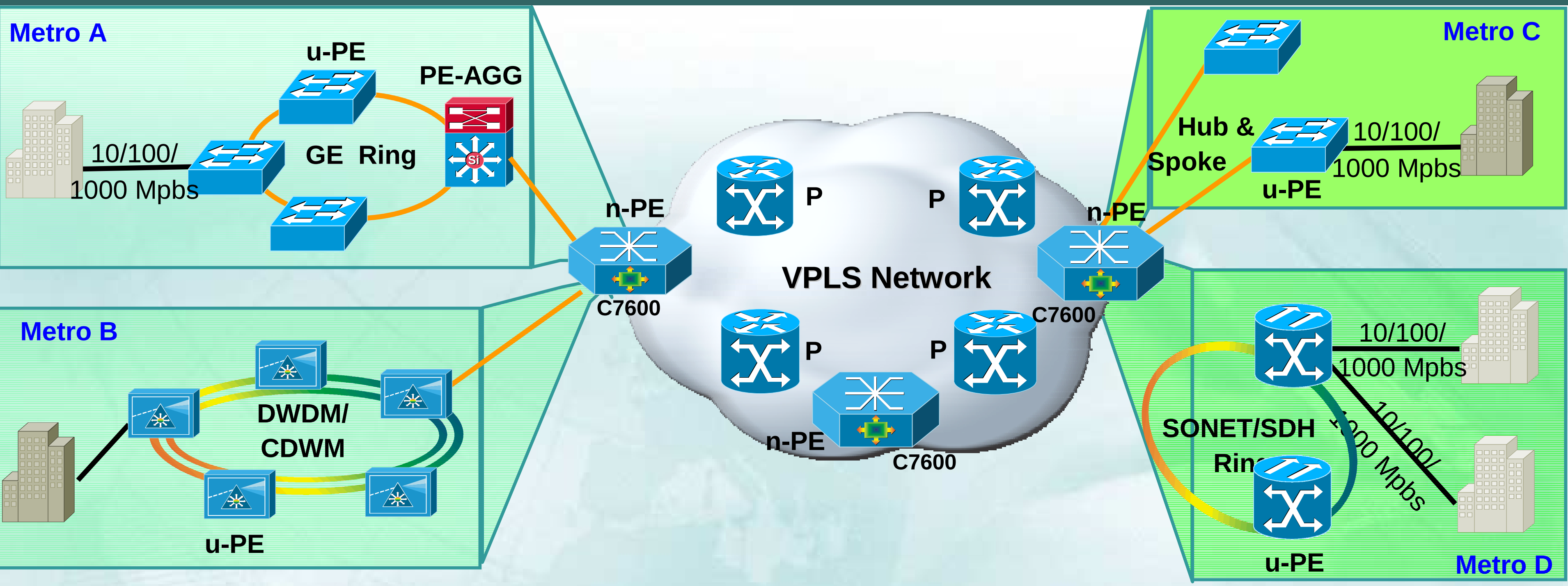
- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Metro Ethernet: Emerging Multiservice Access Opportunity



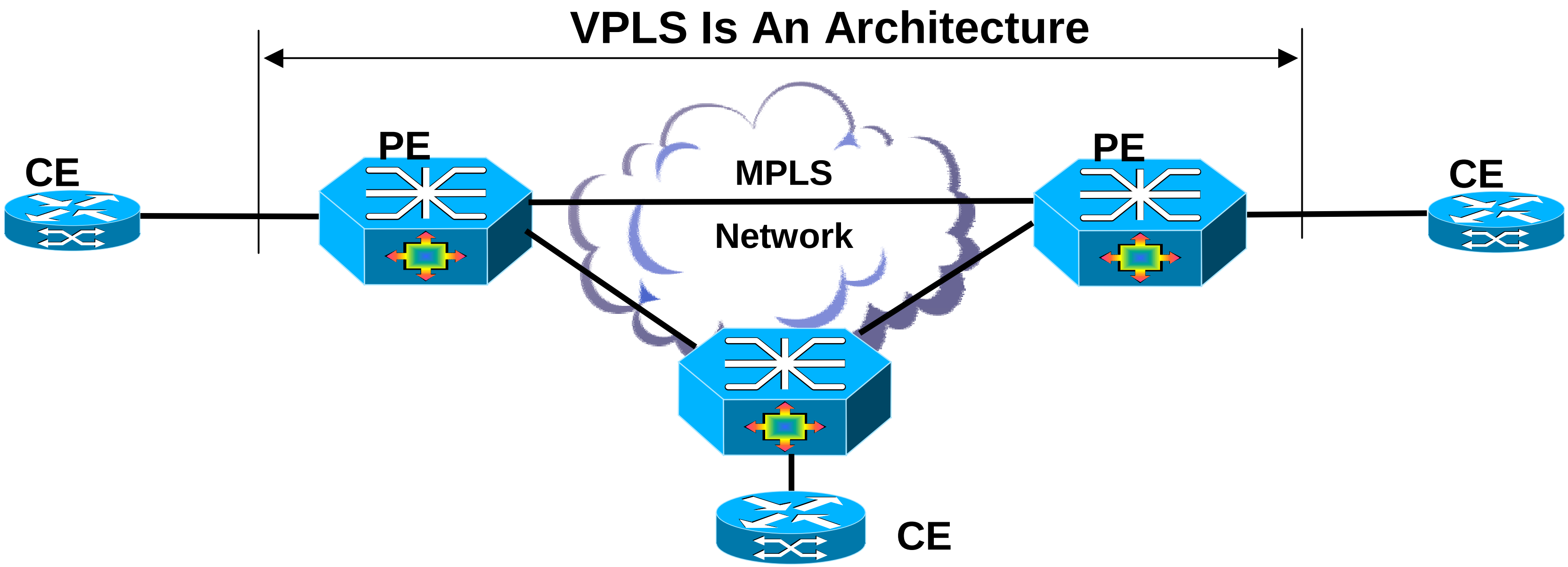
VPLS Overview for Metro Ethernet

Cisco.com



- Delivers Ethernet-based multipoint L2 VPN service
- Enhances L2 VPN scalability (geographic sites & no. of customers)
- Leverages existing SP MPLS Core
- Supports operational speeds of GB to 10 GB
- On track for IETF standardization: Draft Lasserre-Kompella
- Uses familiar Ethernet user network interface

Virtual Private LAN Services (VPLS)

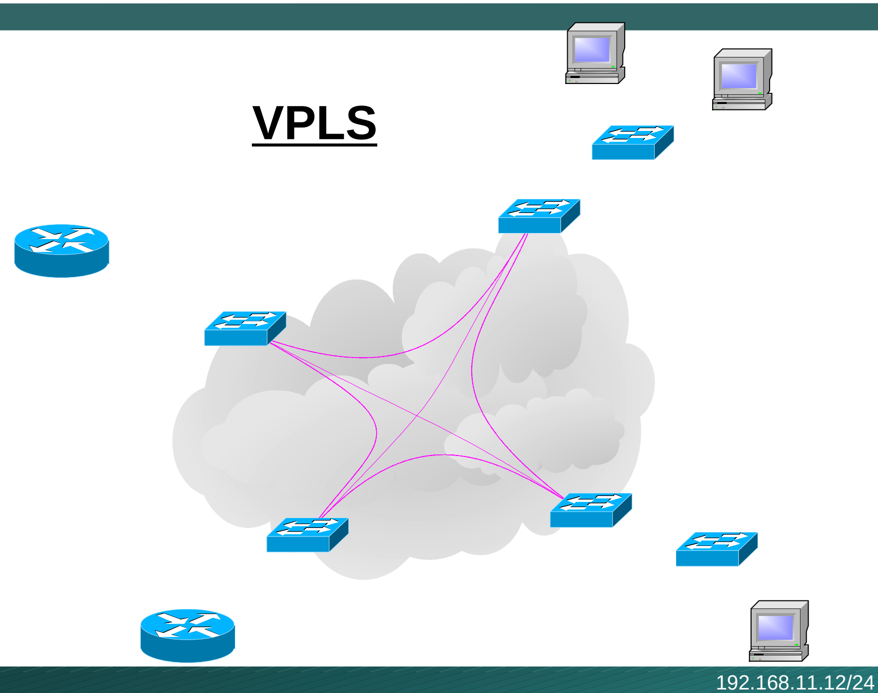


- VPLS defines an architecture that delivers Ethernet Multipoint Services (EMS) over an MPLS network
- VPLS operation emulates an IEEE Ethernet bridge
- Two VPLS drafts in existance

  Draft-ietf-l2vpn-vpls-ldp-01

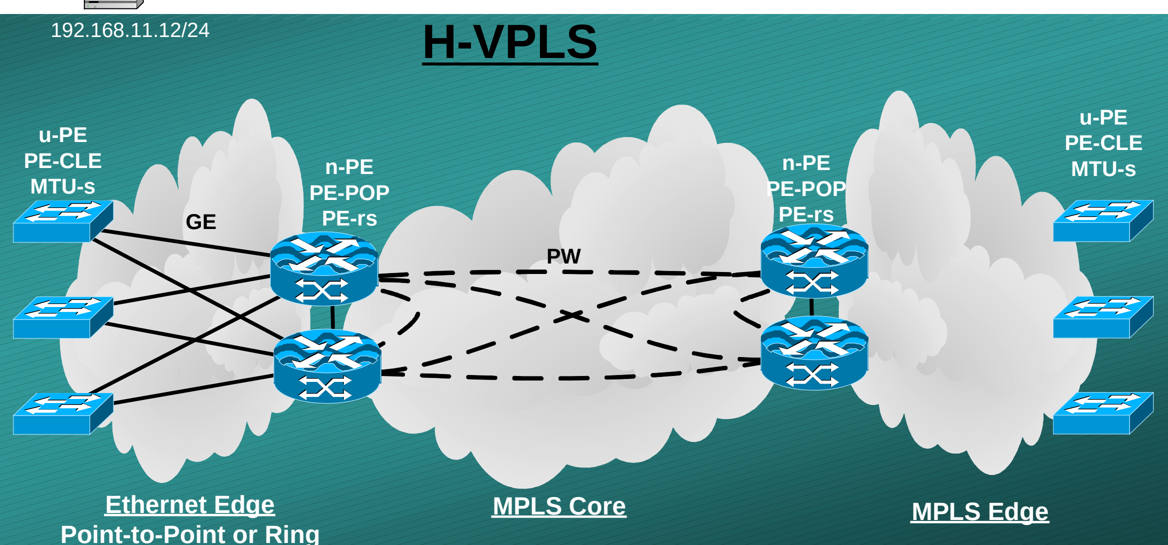
  draft-ietf-l2vpn-vpls-bgp-01

VPLS & H-VPLS

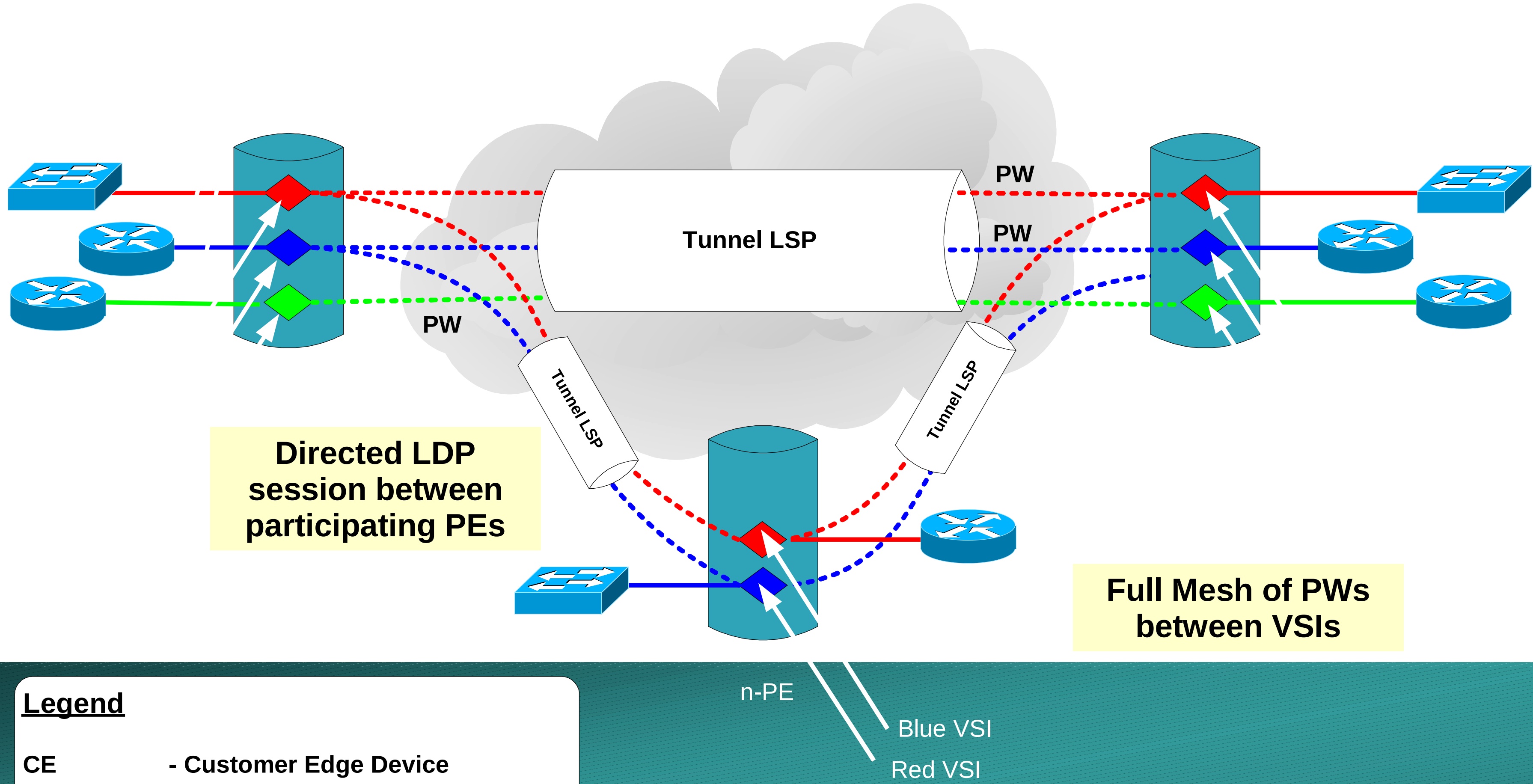


- **VPLS Direct Attachment**
  - Single Flat Hierarchy
  - MPLS to the Edge

- **H-VPLS**
  - Two Tier Hierarchy
  - MPLS or Ethernet Edge
  - MPLS Core



VPLS Components



Legend

CE	- Customer Edge Device
n-PE	- network facing-Provider Edge
VSI	- Virtual Switch Instance
PW	- Pseudo-Wire
Tunnel LSP	- Tunnel Label Switch Path that provides PW transport

VPN & VPLS Desirable Characteristics

- **Auto-discovery of VPN membership**
  - Reduces VPN configuration and errors associated with configuration
- **Signaling of connections between PE devices associated with a VPN**
- **Forwarding of frames**
  - AToM uses Interface based forwarding
  - VPLS uses IEEE 802.1q Ethernet Bridging techniques
- **Loop prevention**
  - MPLS Core will use a full mesh of PWs and “split-horizon” forwarding
  - H-VPLS edge domain may use IEEE 802.1s Spanning Tree, RPR, or SONET Protection

VPLS: Layer 2 Forwarding Instance Requirements

A Virtual Switch MUST operate like a conventional L2 switch!

Flooding / Forwarding:

- **MAC table instances per customer and per customer VLAN (L2-VRF idea) for each PE**
- **VSI will participate in learning, forwarding process**
- **Uses Ethernet VC-Type defined in pwe3-control-protocol-xx**

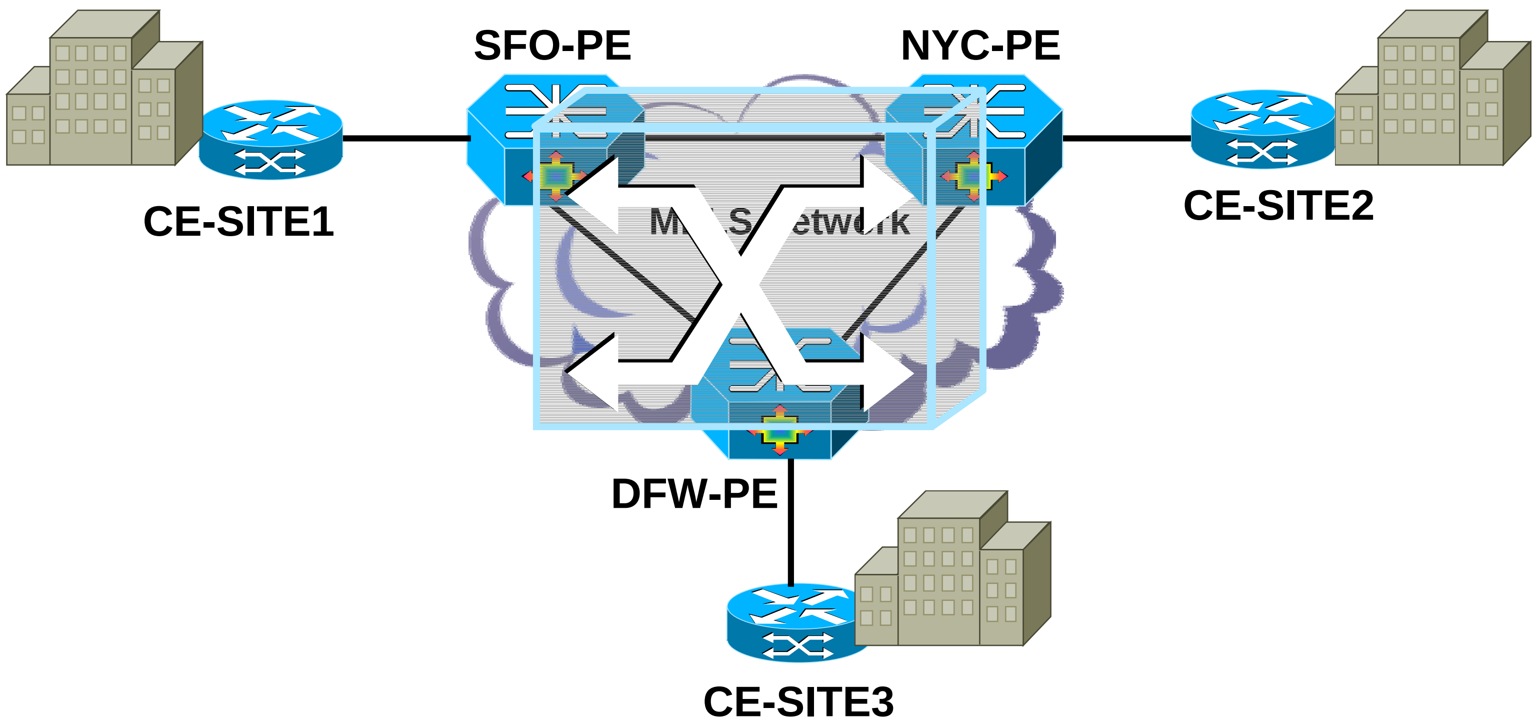
Address Learning / Aging:

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

Loop Prevention:

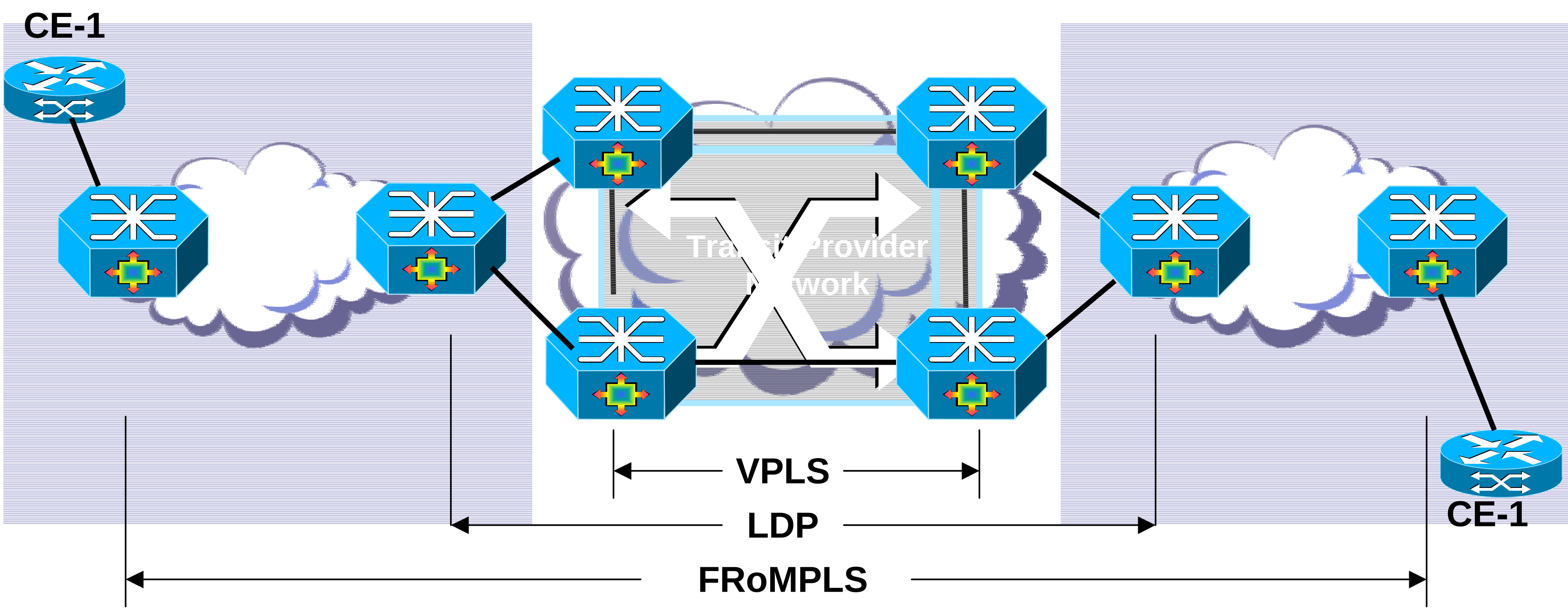
- **Create partial or full-mesh of EoMPLS VCs per VPLS**
- **Use "split horizon" concepts to prevent loops**
- **Announce EoMPLS VPLS VC tunnels**

VPLS Deployment: SMB Connectivity



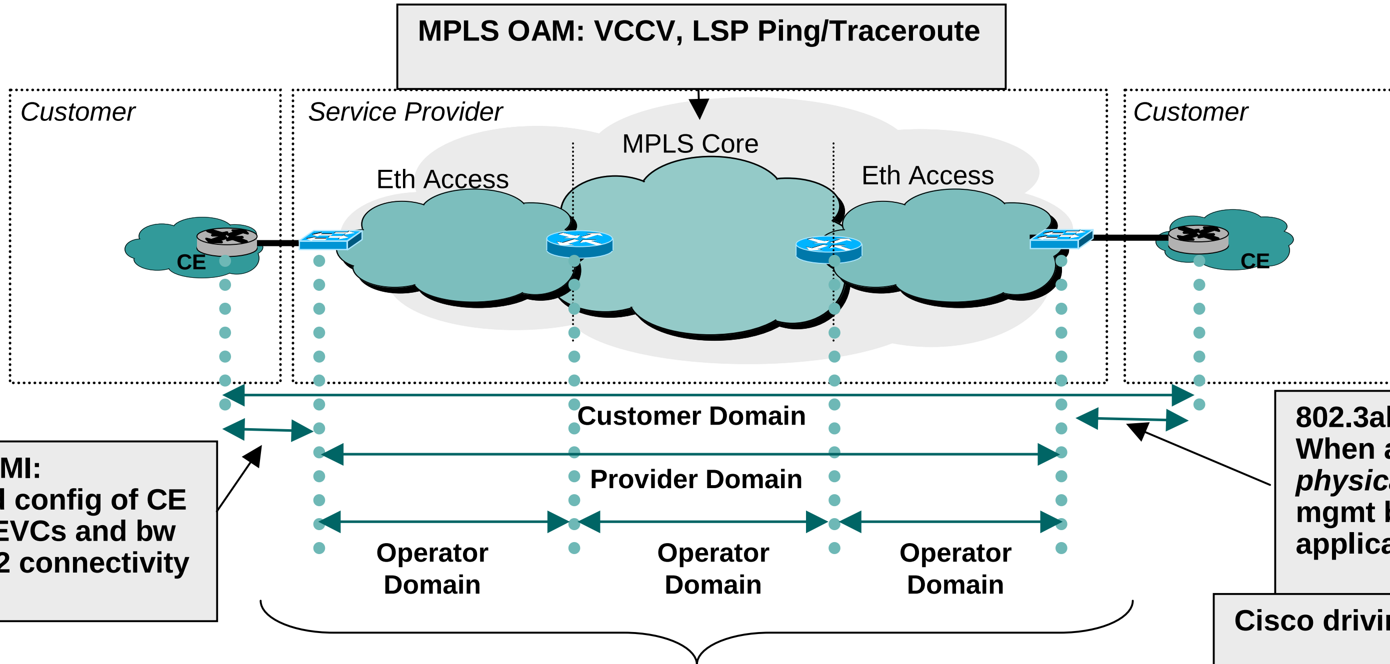
- New Layer 2 multipoint service offering
- Enterprise maintains routing and administrative autonomy
- Layer 3 protocol independence
- Full mesh between customer sites

VPLS Deployment: Layer 2 Multipoint Transit Provider



- SP-As PEs appear back to back and packets are forwarded
- No LDP or Route exchange with transit provider
- Provides optimal traffic path to carrier’s PE

Ethernet OAM – Future



Ethernet-LMI: Automated config of CE based on EVCs and bw profiles; L2 connectivity mgmt

802.3ah Eth in First Mile: When applicable, *physical* connectivity mgmt betw devices. Most applicable to "first mile"

802.1ag Connectivity Fault Management:

- Uses Domains to contain OAM flows & bound OAM responsibilities
- Provides per EVC connectivity mgmt and fault isolation
- Three types of packets: Continuity Check, L2 Ping, L2 Traceroute

Cisco driving standards

ITU-T SG 13 and SG 15:
- Ethernet Layer Netw Arch (G.8010 SG 15)
- Ethernet OAM Functionality (Y.ethoam SG 13)
- Req'ts for OAM in Ethernet based netw (Y.1730 – SG 13)

IEEE:
- 802.3ah – Ethernet in First Mile (Physical OAM);
- 802.1ad – Provider Bridges
- 802.1ag – Connectivity Mgmt (Per VLAN OAM)

MEF:
- E-LMI

Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

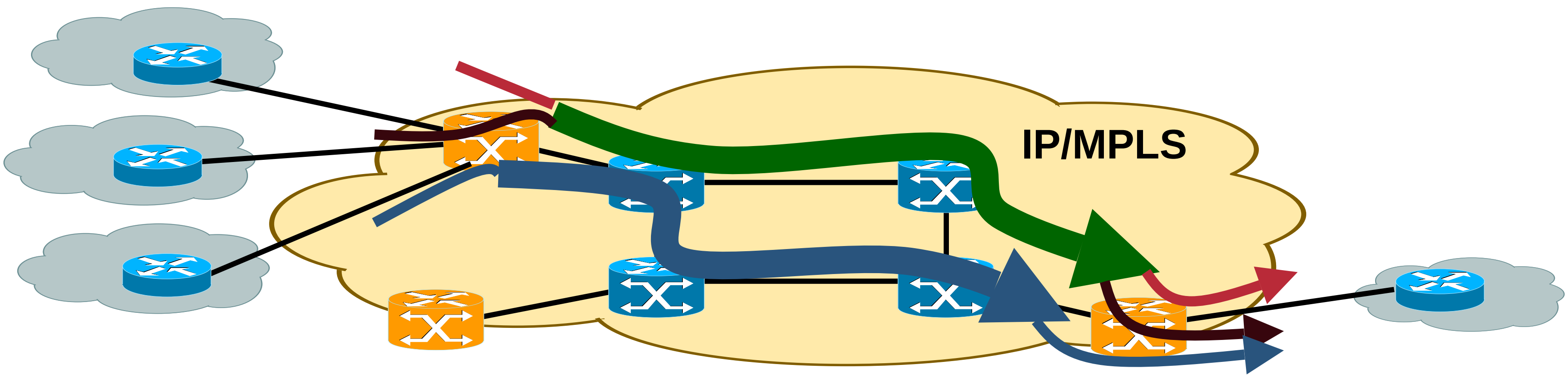
Why Traffic Engineering?

- Congestion in the network due to changing traffic patterns
  - Election news, online trading, major sports events
- Better utilization of available bandwidth
  - Route on the non-shortest path
- Route around failed links/nodes
  - Fast rerouting around failures, transparently to users
  - Like SONET APS (Automatic Protection Switching)
- Build New Services—Virtual leased line services
  - VoIP Toll-Bypass applications, point-to-point bandwidth guarantees
- Capacity planning
  - TE improves aggregate availability of the network

Background – Why Have MPLS-TE?

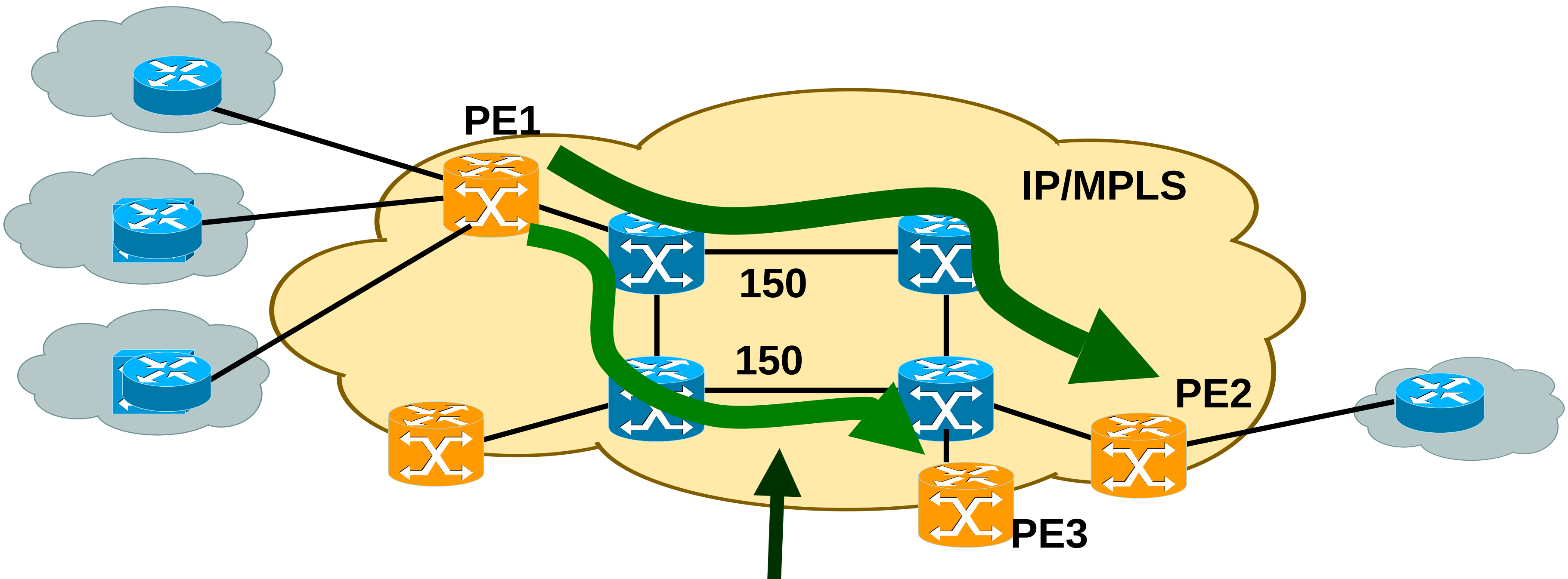
- IP networks route based only on destination (route)
- ATM/FR networks switch based on both source and destination (PVC, etc)
- Some very large IP networks were built on ATM or FR to take advantage of src/dst routing
- Overlay networks inherently hinder scaling (see “The Fish Problem”)
- MPLS-TE lets you do src/dst routing while removing the major scaling limitation of overlay networks
- MPLS-TE has since evolved to do things other than bandwidth optimization

Traffic Engineering services



- **Traffic engineering offers the carrier mechanisms to optimise their infrastructure.**
  - **Distributing traffic**
  - **Pre-built back-up paths**
  - **Traffic separation over different TE paths**
- **Solution Examples**
  - **Basic Traffic engineering**
  - **Diffserv aware TE**
  - **TE optimisation tools**
  - **FRR using TE**

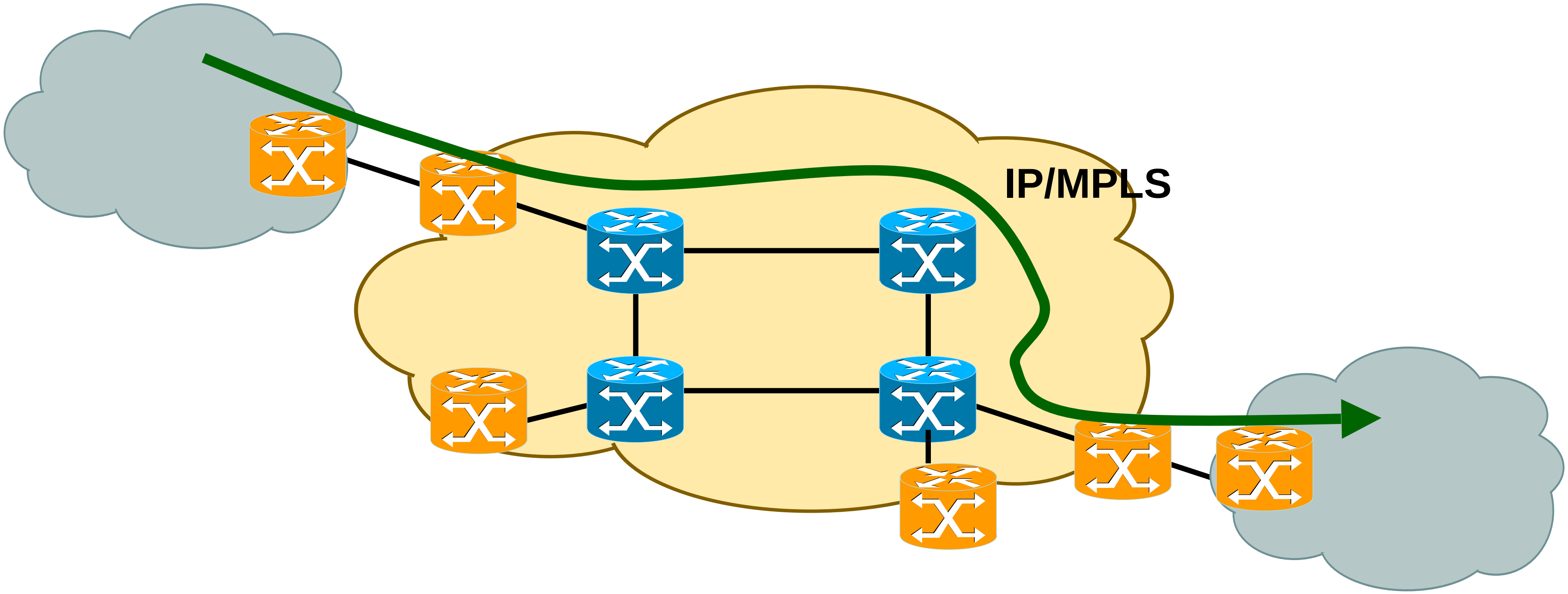
MPLS Traffic Engineering in Core



- MPLS TE Tunnels MAY be used to distribute aggregate load via Constraint Based Routing
- avoid congestion
- in this example, routing PE1→PE2 traffic (80Mb/s) and PE1→PE3 traffic (90Mb/s) on separate path in the core avoids congestion

RFC2702 Requirements for MPLS Traffic Engineering
RFC3209 RSVP extensions for LSP Tunnels

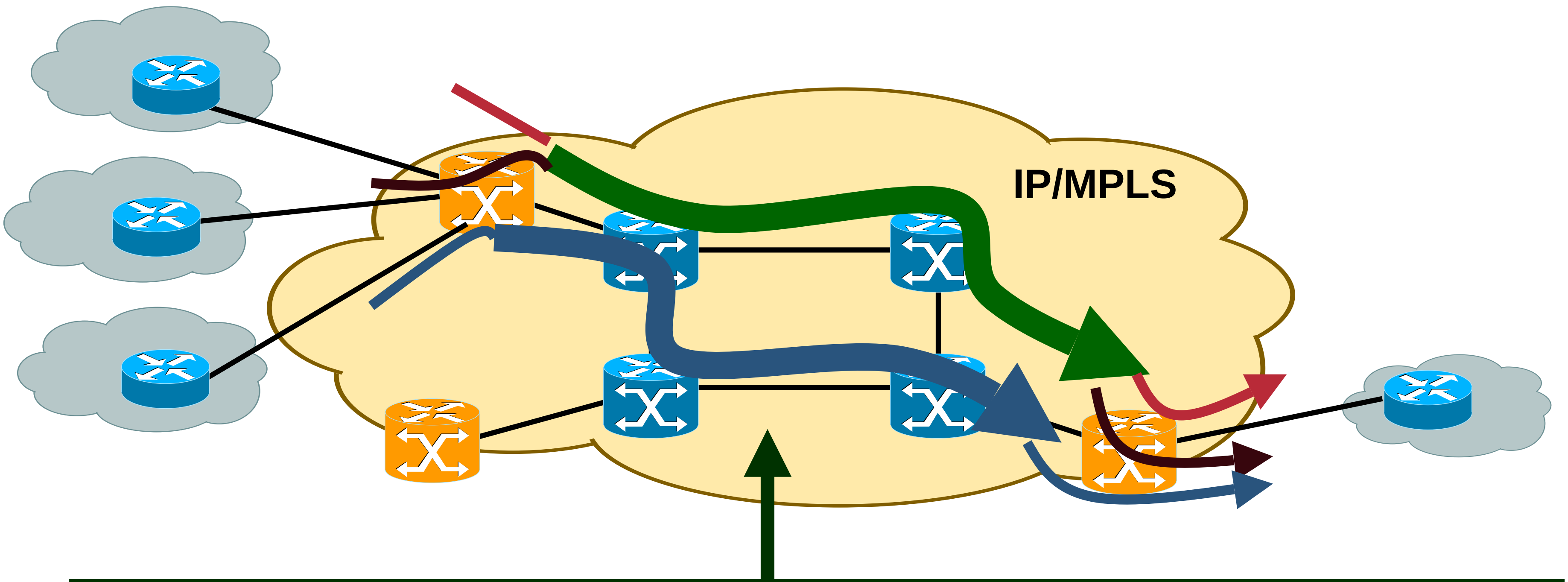
InterAS TE



TE Tunnel spanning multiple Autonomous Systems
Allows bandwidth reservations to span multiple domains

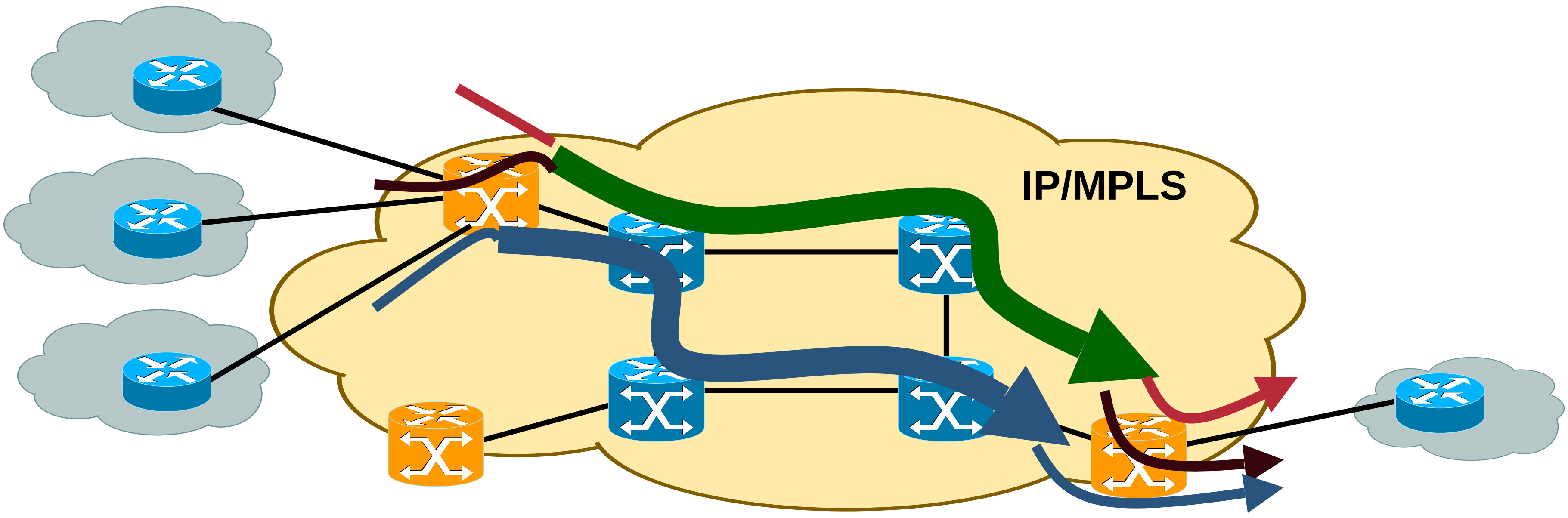
draft-zhang-mpls-interas-te-req-xx, draft-vasseur-inter-as-te-xx
draft-vasseur-mpls-loose-path-reopt-xx, draft-vasseur-mpls-nodeid-subobject-xx

Diff-Serv-aware Traffic Engineering (DS-TE) in Core



- MPLS DS-TE Tunnels MAY be used to carry separately different classes of service
- canonical example is separate tunnels for Voice and for Data
-facilitates strict enforcement of different QoS objectives for differnet classes WITHOUT over-engineering
- per class CAC (eg. route Voice tunnels taking into account the EF queue capacity – and not just the link capacity)
- per class C-SPF (eg. Use a “hop/Bw based metric“ for data tunnels and a “delay-based metric” for voice tunnels)

Diff-Serv-aware Traffic Engineering (DS-TE) in Core



RFC3564 Requirements for Diff-Serv-aware MPLS Traffic Engineering

draft-ietf-tewg-diff-te-proto-xx

draft-ietf-tewg-diff-te-russian-xx

draft-ietf-tewg-diff-te-mam-xx

***Path Computation Element (PCE) WG Now*.**

The PCE Working Group is chartered to specify a Path Computation Element(PCE) based architecture for the computation of paths for MPLS and GMPLSTraffic Engineering LSPs

Applicability of Core QoS mechanims

What should be deployed: ???

- Nothing

- MPLS TE

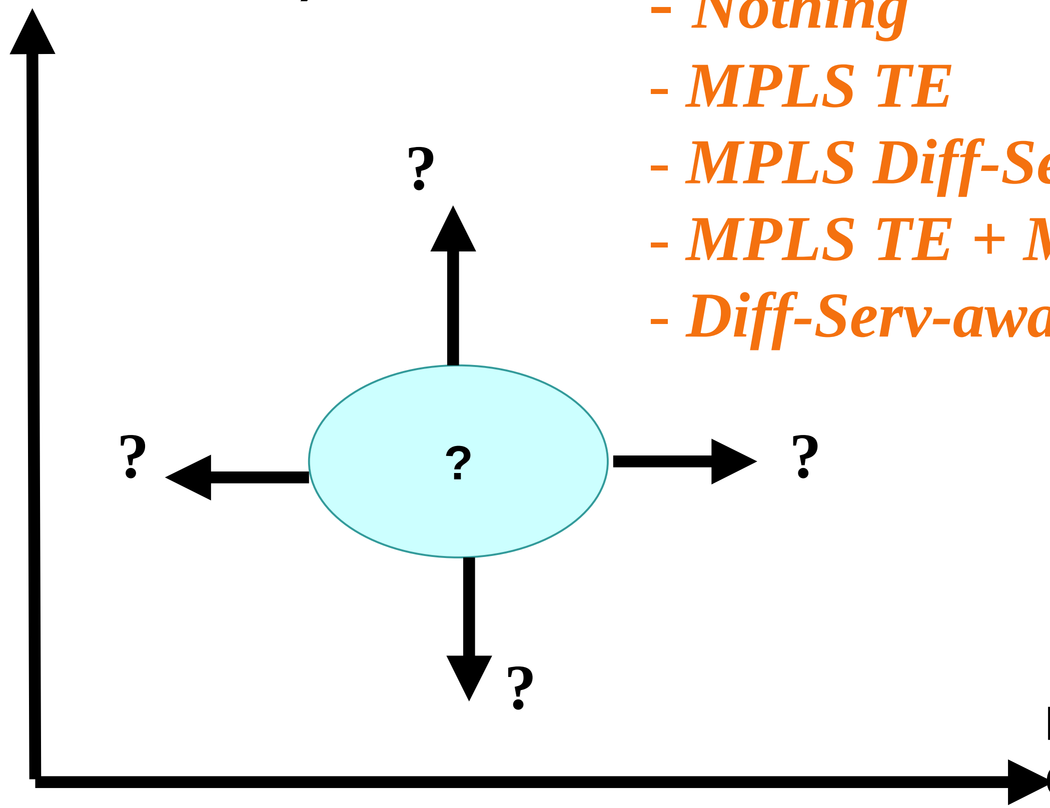
- MPLS Diff-Serv

- MPLS TE + MPLS Diff-Serv

- Diff-Serv-aware TE

Applicability of Core QoS mechanims

Service
Differentiation
(increase revenue)



Resource
Optimisation
(reduce spending)

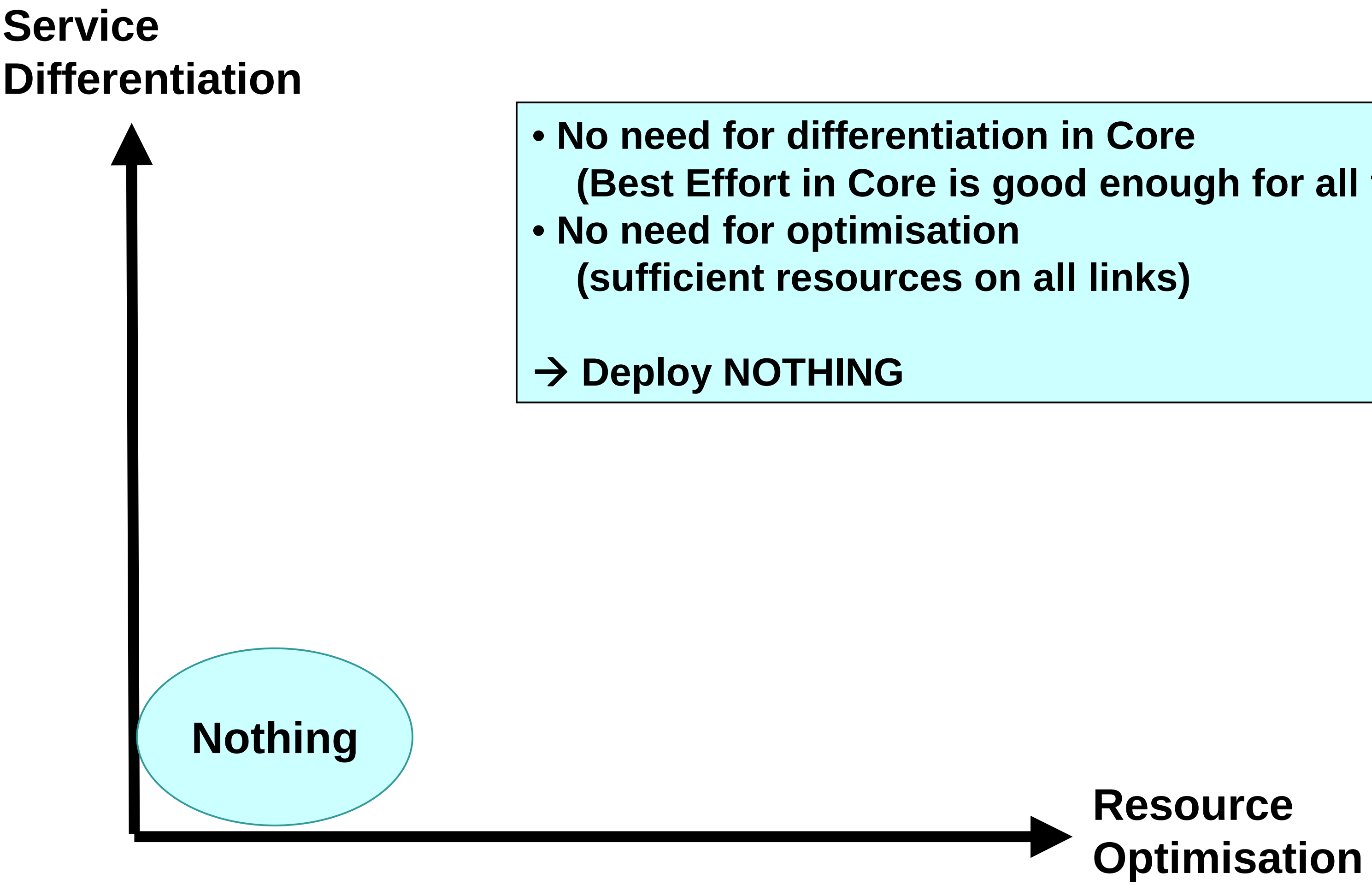
What should be deployed: ???

- Nothing
- MPLS TE
- MPLS Diff-Serv
- MPLS TE + MPLS Diff-Serv
- Diff-Serv-aware TE

Applicability of Core QoS mechanims

- No need for differentiation in Core
  (Best Effort in Core is good enough for all traffic)
- No need for optimisation
  (sufficient resources on all links)

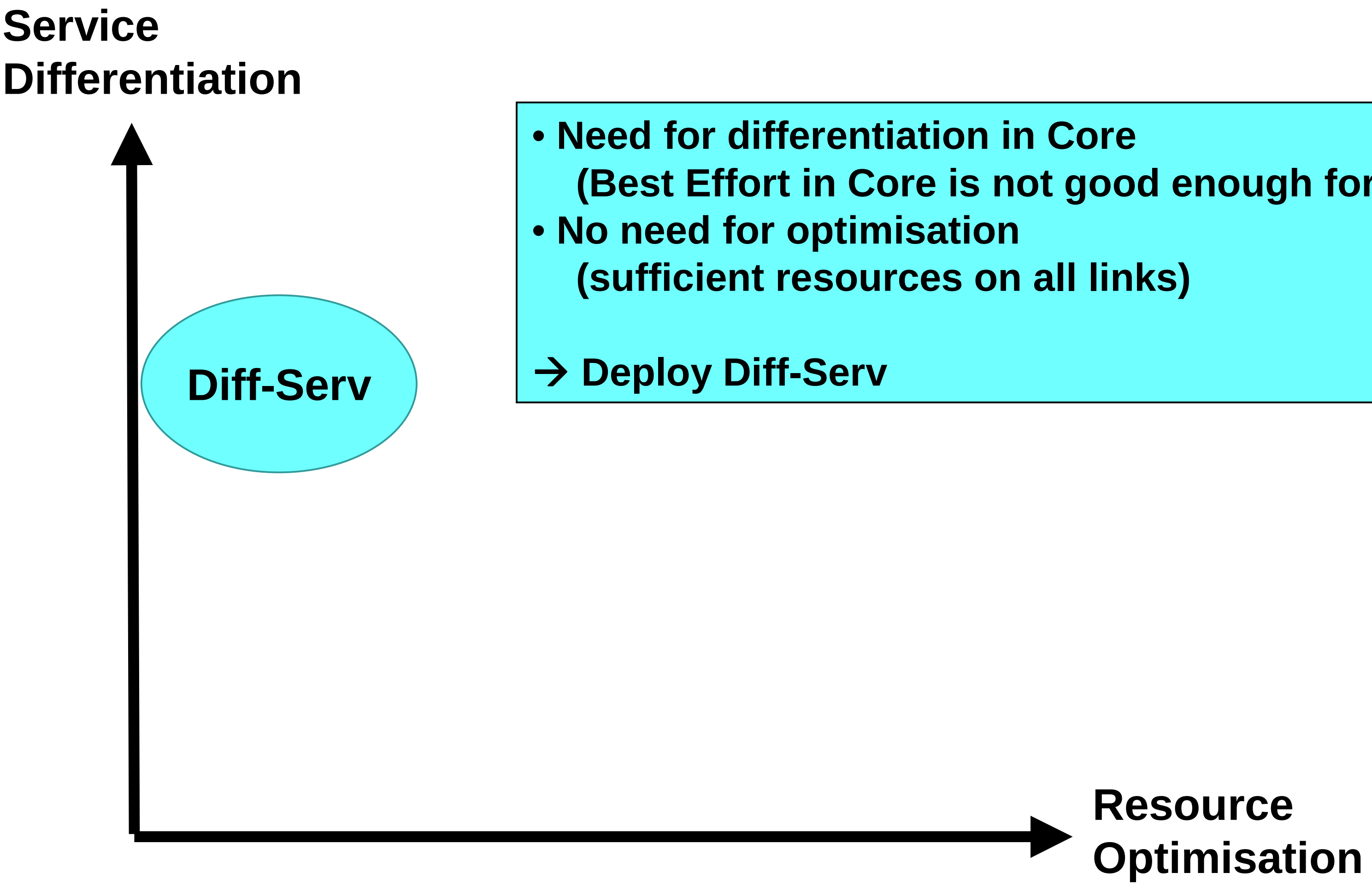
→ Deploy NOTHING



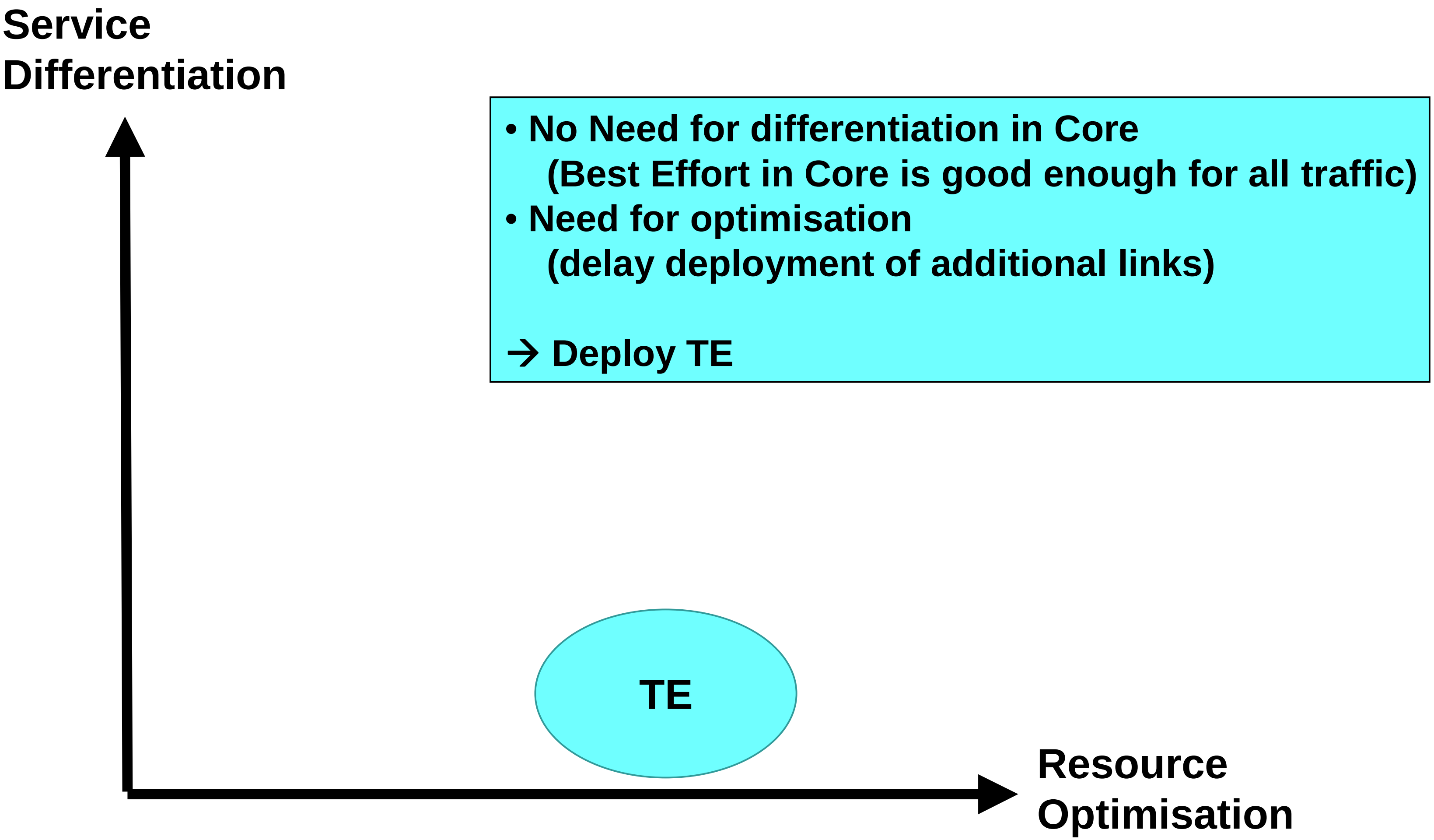
Applicability of Core QoS mechanims

- Need for differentiation in Core
  (Best Effort in Core is not good enough for voice)
- No need for optimisation
  (sufficient resources on all links)

→ Deploy Diff-Serv



Applicability of Core QoS mechanims



Applicability of Core QoS mechanims

Service Differentiation

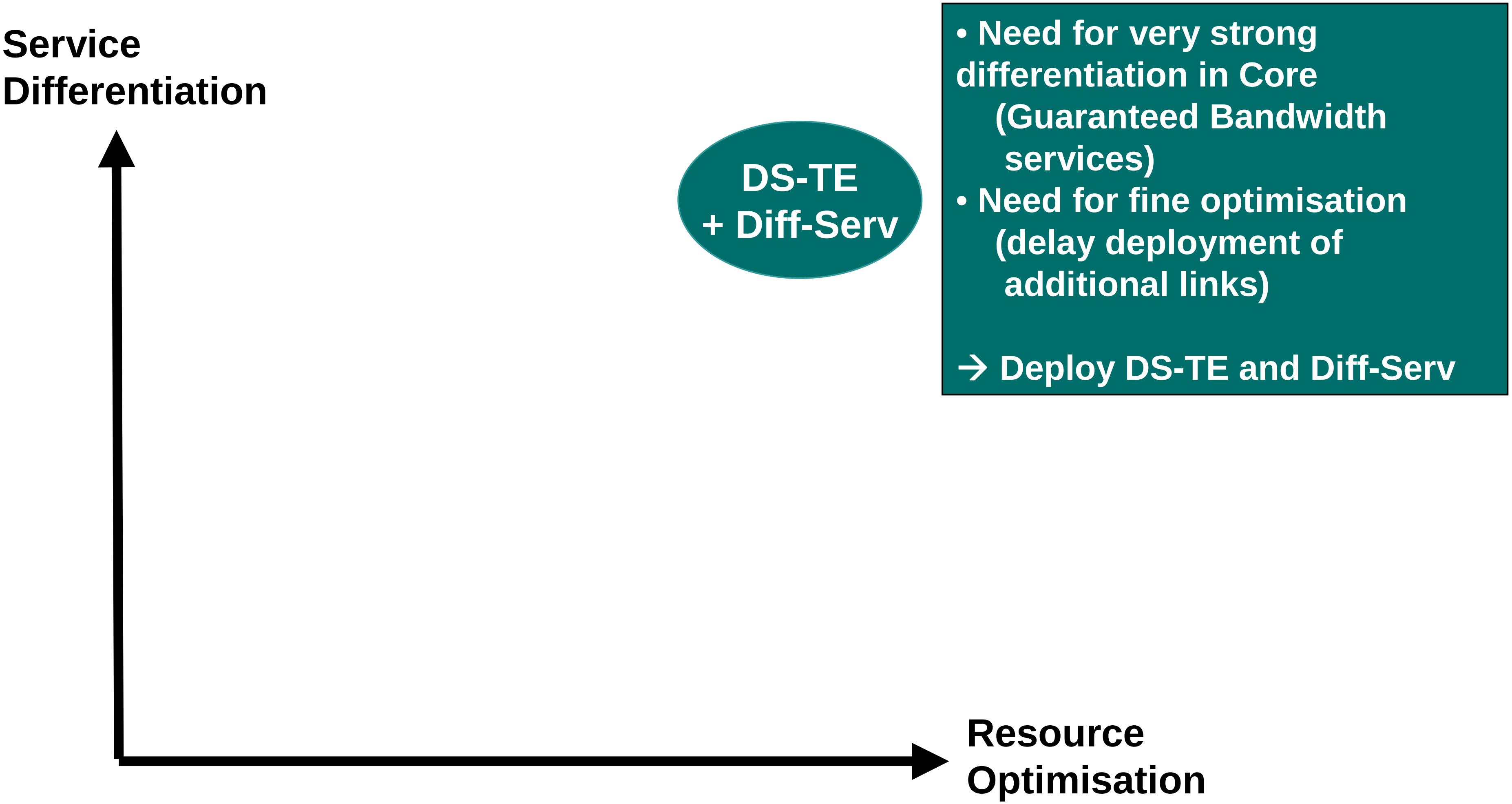
- Need for differentiation in Core
  (Best Effort in Core is not good enough for Voice)
- Need for optimisation
  (delay deployment of additional links)

→ Deploy TE and Diff-Serv

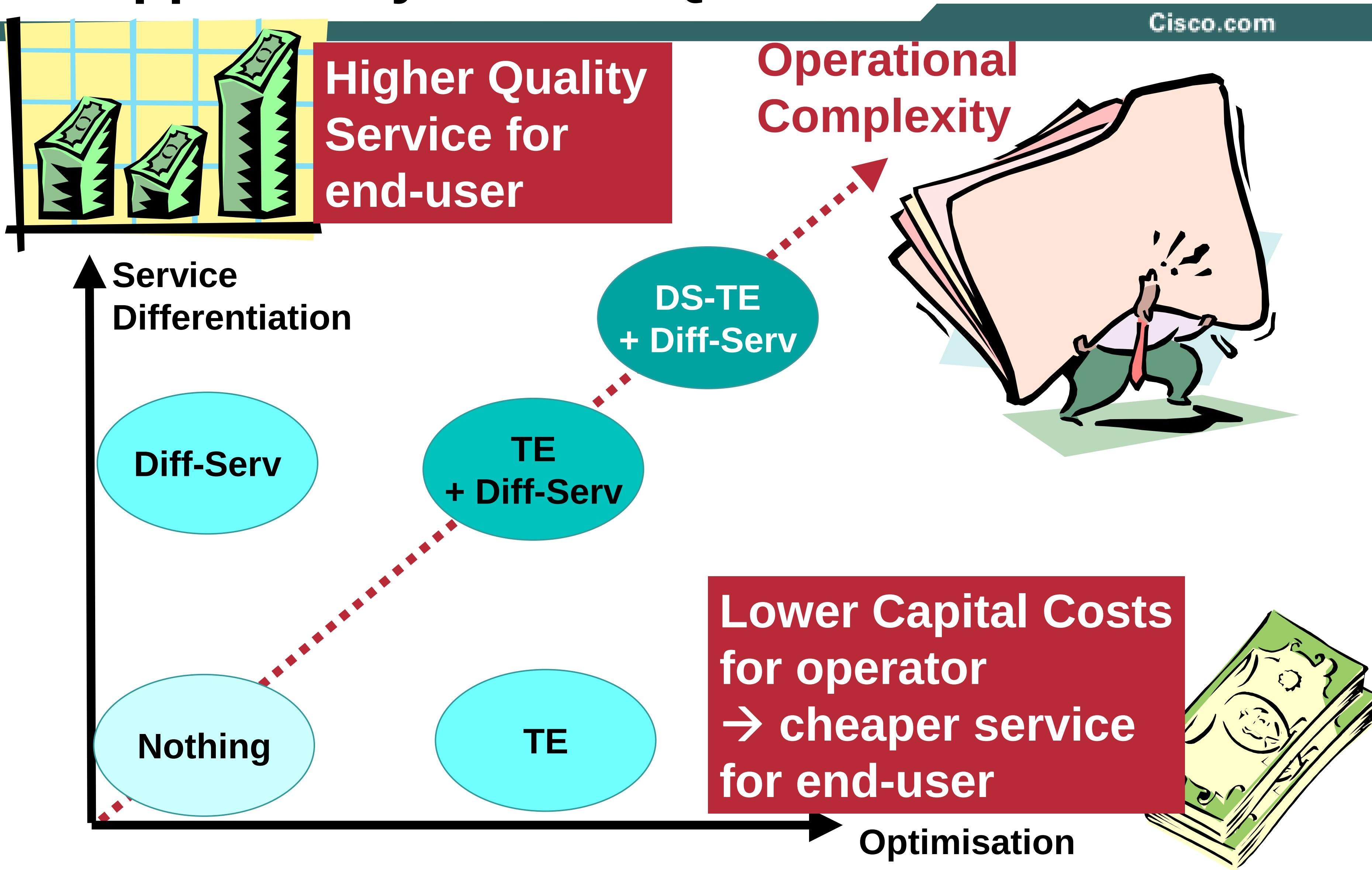
TE
+ Diff-Serv

Resource Optimisation

Applicability of Core QoS mechanims



Appicability of Core QoS mechanisms

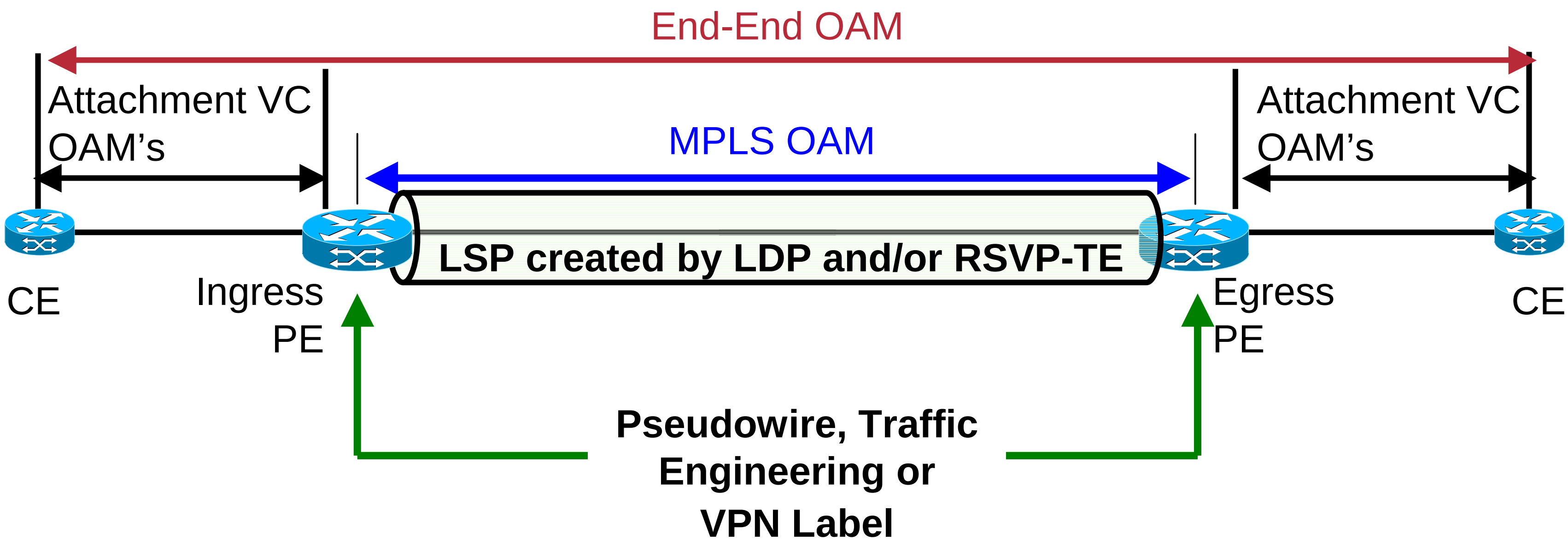


Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Where does MPLS OAM fit



- **MPLS OAM mechanisms applicable between Ingress and Egress Provider Edges;**
- **Label Switched Path (LSP) created by Control protocols such as Label Distribution Protocol and/or RSVP-TE**

MPLS LSP Ping/Traceroute

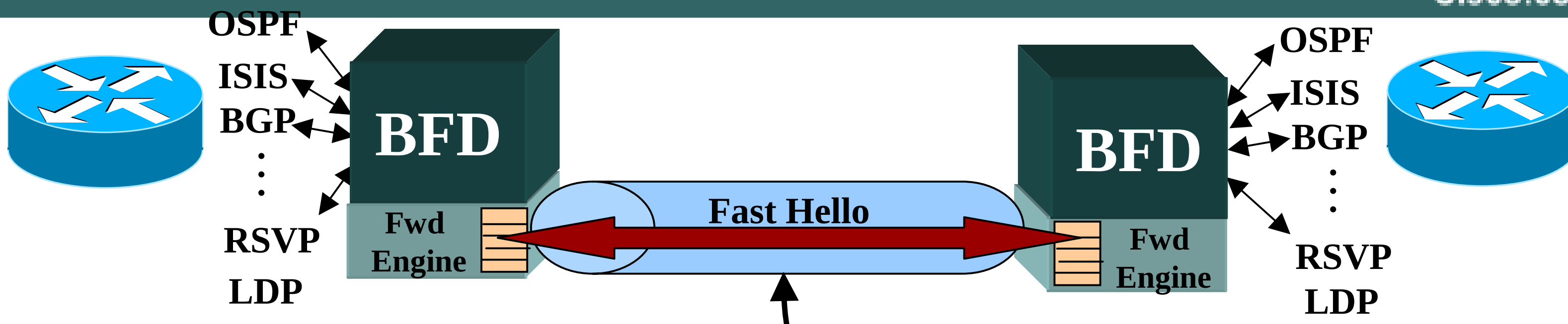
Requirement	• Detect MPLS traffic black holes or misrouting • Isolate MPLS faults • Verify data plane against the control plane • Detect MTU of MPLS LSP paths
Solution	• MPLS LSP Ping (ICMP) for connectivity checks • MPLS LSP Traceroute for hop-by-hop fault localization • MPLS LSP Traceroute for path 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-06.txt

MPLS AToM Virtual Circuit Connection Verification (VCCV)

Requirement	• Ability to provide end-to-end fault detection and diagnostics for an emulated pseudowire service One tunnel can serve many pseudowires. MPLS LSP ping is sufficient to monitor the PSN tunnel (PE-PE connectivity), but not VCs inside of tunnel
Solution	• AToM VCCV allows sending control packets in band of an AToM pseudowire. Two components: Signaled component to communicate VCCV capabilities as part of VC label Switching component to cause the AToM VC payload to be treated as a control packet Type 1: uses Protocol ID of AToM Control word Type 2: use MPLS router alert label
Applications	• Layer 2 transport over MPLS FRoMPLS, ATMoMPLS, EoMPLS
IETF Standards	• Draft-ietf-pwe3-vccv-xx.txt

Attributes of BFD

Cisco.com



- Direct physical links
- Multi-hop routed paths
- Virtual circuits, Tunnels
- MPLS LSPs
- Bi/uni-directional links

- **Protocol Independence**
- **Media Independence**
- **Fast failure detection**

  Light Weight, Fixed Length; simple to parse

- **Forwarding plane liveliness**

  E.g., Link may be up but forwarding engine may be down or an entry may be incorrectly programmed.

- No discovery mechanism in BFD

  Applications bootstrap a BFD session

MPLS BFD Vs. LSP Ping

Method	Data Plane Failure Detection	Control Plane Consistency	Protocol Overhead
LSP Ping	YES	YES	Higher than BFD
MPLS-BFD	YES	NO	Low

MPLS-BFD can <u>complement</u> LSP Ping to detect a data plane failure in the forwarding path of a MPLS LSP

Supported FECs:
RSVP IPv4/IPv6 Session, LDP IPv4/IPv6 prefix
VPN IPv4/IPv6 prefix, Layer 2 VPN, Layer 2 Circuit ID

VCCV BFD Vs. VCCV Ping

Method	Data Plane Failure Detection	Control Plane Consistency	Protocol Overhead
VCCV Ping	YES	YES	Higher than BFD
VCCV-BFD	YES	NO	Low

VCCV-BFD can complement VCCV-LSP Ping to detect a data plane failure in the forwarding path of a PW

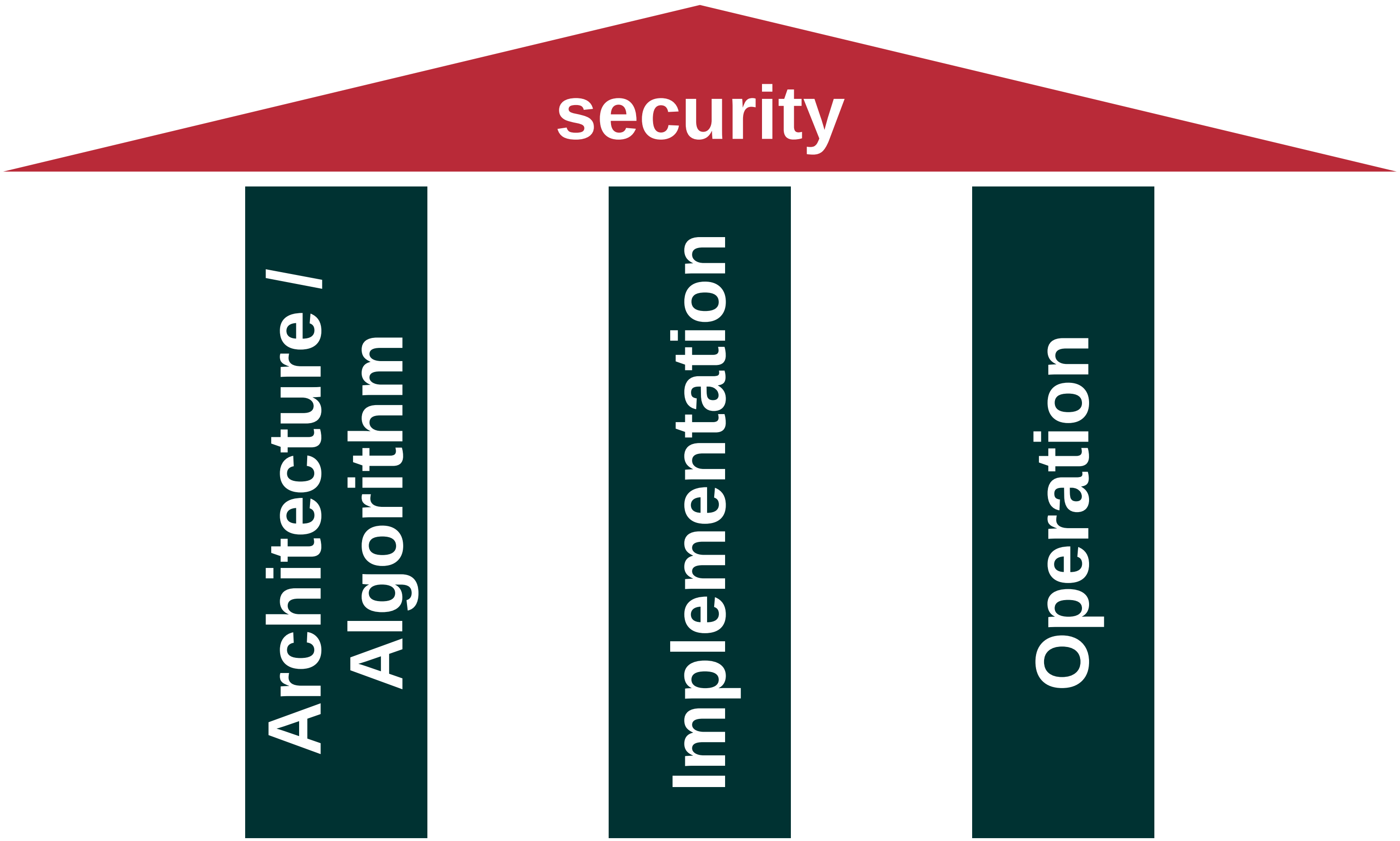
VCCV-BFD works over MPLS or IP networks; Multiple PSN Tunnel Type MPLS, **IPSEC, L2TP, GRE, etc.**

Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Three Pillars of Security



Break one, and all security is gone!

What Kind of Threats?

- **Threats from Outside the Backbone**
  - From VPN customers
  - From the Internet
- **Threats from Inside the Backbone**
  - SP misconfigurations (error or deliberate)
  - Hacker “on the line” in the core
- **Threats that are independent of MPLS**
  - Customer network security

Reference model for best practice deployments

Why is MPLS Security Important?

- **Customer buys "Internet Service":**

  Packets from SP are not trusted

  → Perception: Need for firewalls, etc.

- **Customer buys a "VPN Service":**

  Packets from SP are trusted

  → Perception: No further security required

SP Must Ensure Secure MPLS Operations

Protecting an MPLS/VPN Core—Overview

Cisco.com

1. **Don't let packets into (!) the core**

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

2. **Secure the routing protocol**

   Neighbor authentication, maximum routes, dampening, …

3. **Design for transit traffic**

   QoS to give VPN priority over Internet

   Choose correct router for bandwidth

   Separate PEs where necessary

4. **Operate Securely**

→ Still "Open": Routing Protocol

→ Only Attack Vector: Transit Traffic

→ Now Only Insider Attacks Possible

→ Avoid Insider Attacks

Best Practice Security Overview (1)

- Secure devices (PE, P): They are trusted!
- Core (PE+P): Secure with ACLs on all interfaces

  Ideal: deny ip any <core-networks>
- Static PE-CE routing where possible
- If routing: Use authentication (MD5)
- Separation of CE-PE links where possible (Internet / VPN)
- LDP authentication (MD5)
- VRF: Define maximum number of routes

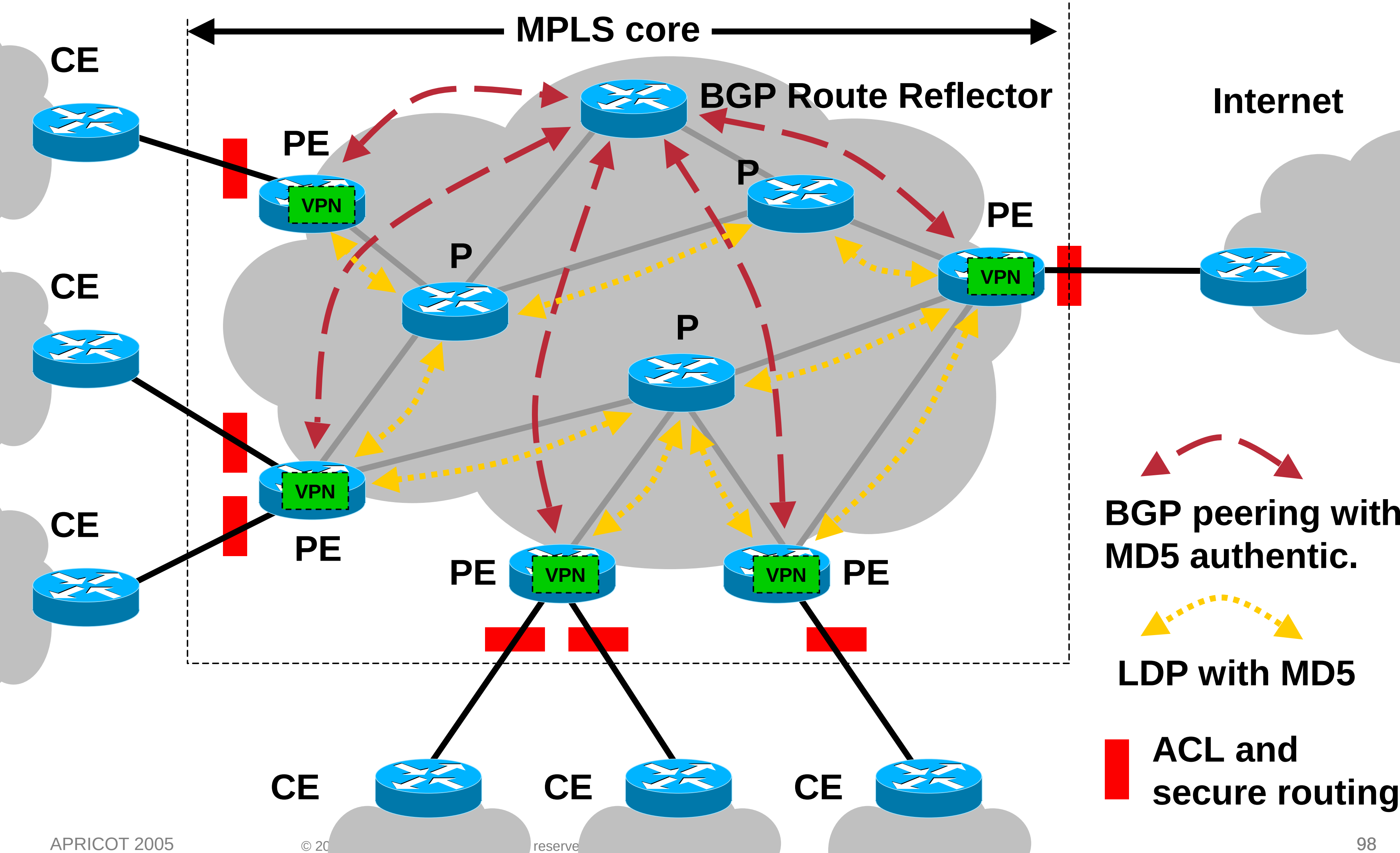
Note: Overall security depends on weakest link!

PE-CE Routing Security

In order of security preference:

1. **Static**: If no dynamic routing required (no security implications)
2. **BGP**: For redundancy and dynamic updates (many security features)
3. **IGPs**: If BGP not supported (limited security features)

Securing the MPLS Core



Use IPsec if you need:

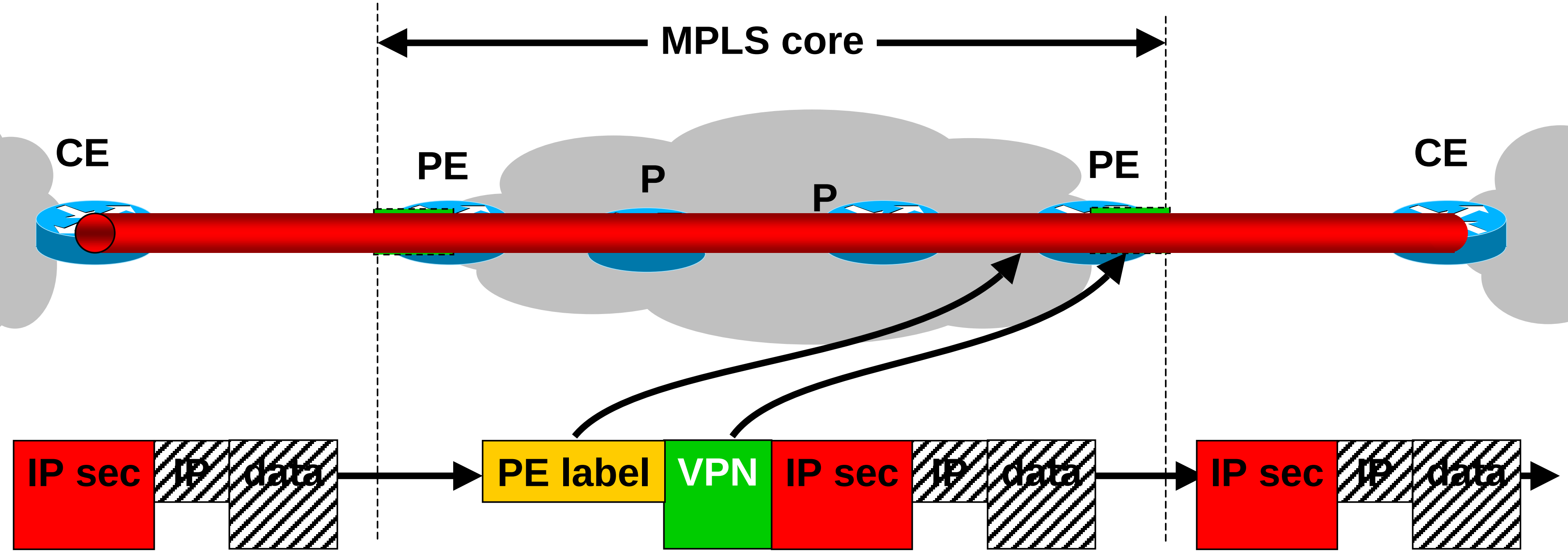
- Encryption of traffic
- Direct authentication of CEs
- Integrity of traffic
- Replay detection

Maybe more important than encryption?

- Or: If you don’t want to trust your ISP for traffic separation!

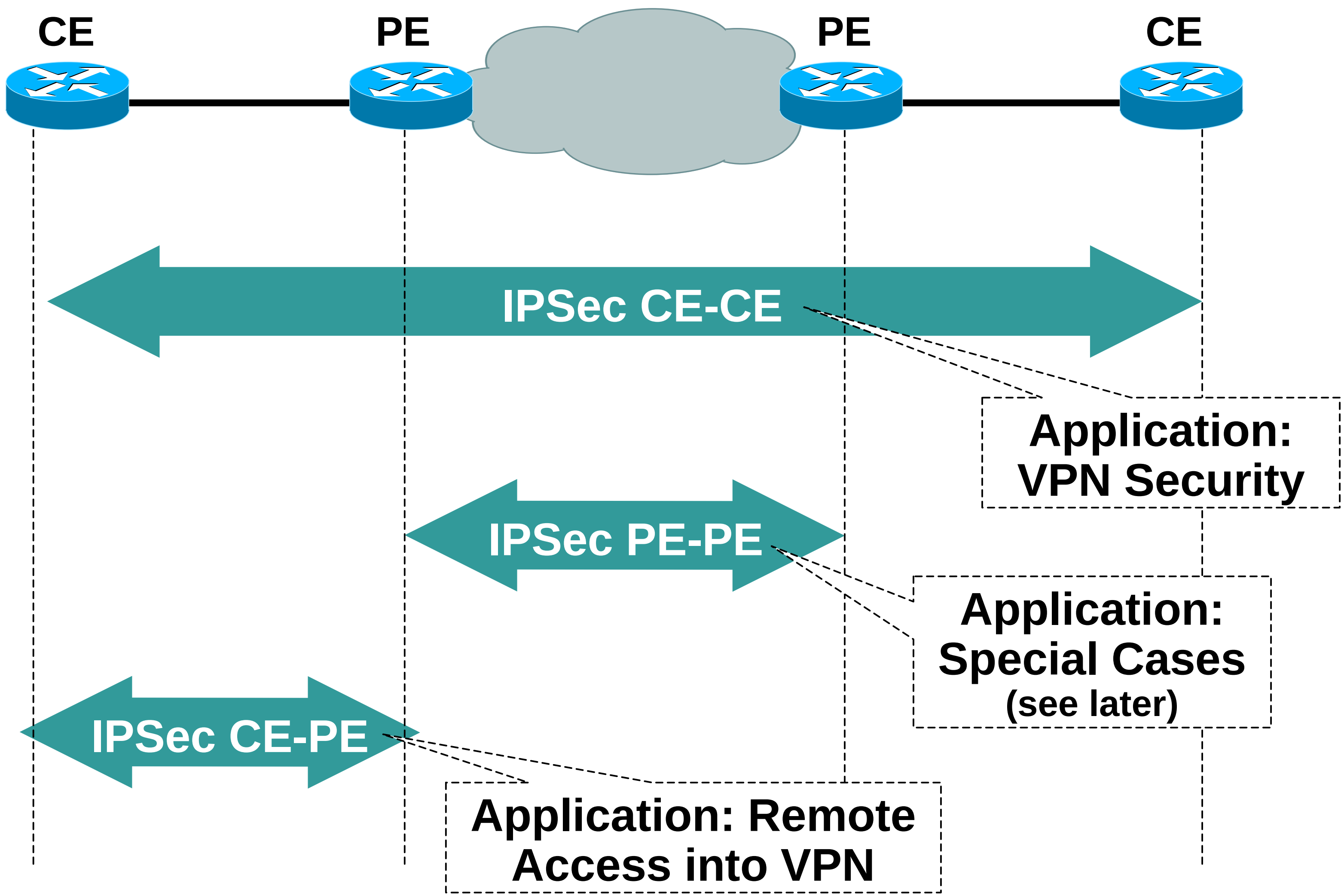
End-to-End Security with IPsec

Cisco.com



- **Encryption: Data invisible on core**
- **Authentication: Only known CEs**
- **Integrity: Data not changed in transit**

Where to Apply IPSec



Where to do IPsec

Cisco.com

1. CE to CE

    SP not involved (unless manages CEs)

    MPLS network only sees IPsec traffic

    Very secure

2. PE to PE

    Does not prevent sniffing access line

    Not very secure for the customer

    There are some specific applications for this (US ILECs)

Mixtures

Need to trust SP

Mostly for access into VPN

Applications of PE-PE IPSec

- If core is not pure MPLS, but IP based
    - Standard 2547bis requires MPLS core, PE-PE IPSec does not
    - Alternative: MPLS in IP/GRE/L2TPv3, but with PE-PE IPSec spoofing impossible
- Protect against misbehaving transit nodes
- Protection against sniffing on core lines
- Protection of pseudowire construct in Inter-AS

Non-Application: Customer Security

Cisco.com

Hacker wants to …	IPSec CE-CE	IPSec PE-PE
… read VPN traffic	Protects Fully	Protects Partially
… insert traffic into VPN	Protects Fully	Protects Partially
… join a VPN	Protects Fully	Doesn’t Protect
… DoS a VPN / the core	Doesn’t Protect	Doesn’t Protect

MPLS doesn’t provide:

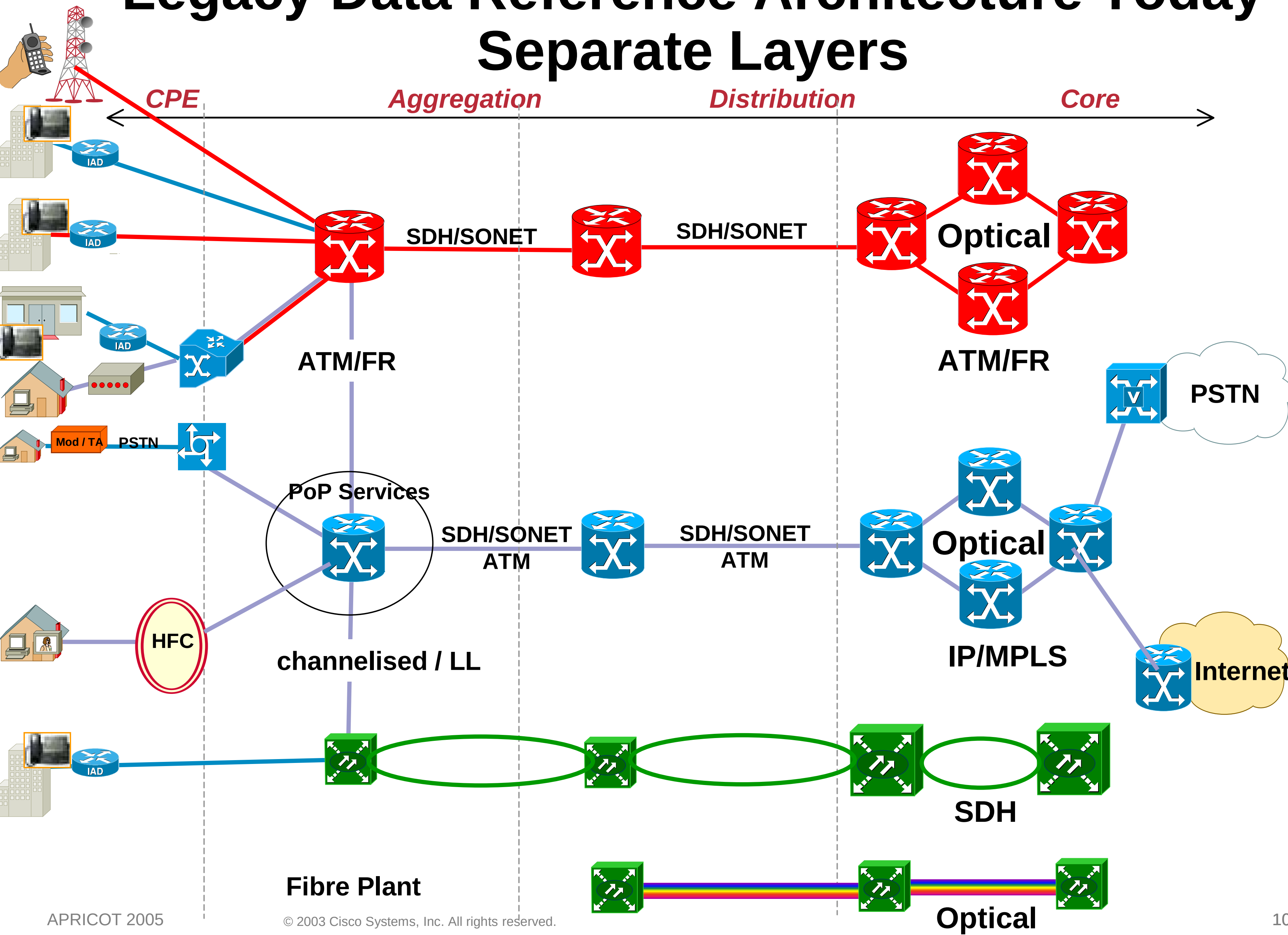
- Protection against mis-configurations in the core
- Protection against attacks from within the core
- Confidentiality, authentication, integrity, anti-replay
- Use IPsec if required
- Customer network security

Agenda

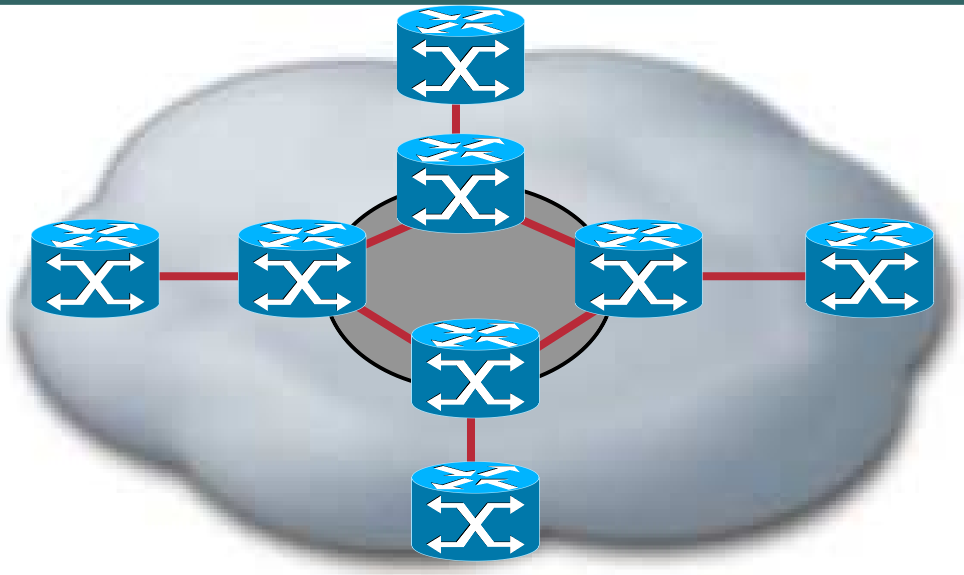
Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

Legacy Data Reference Architecture Today
Separate Layers

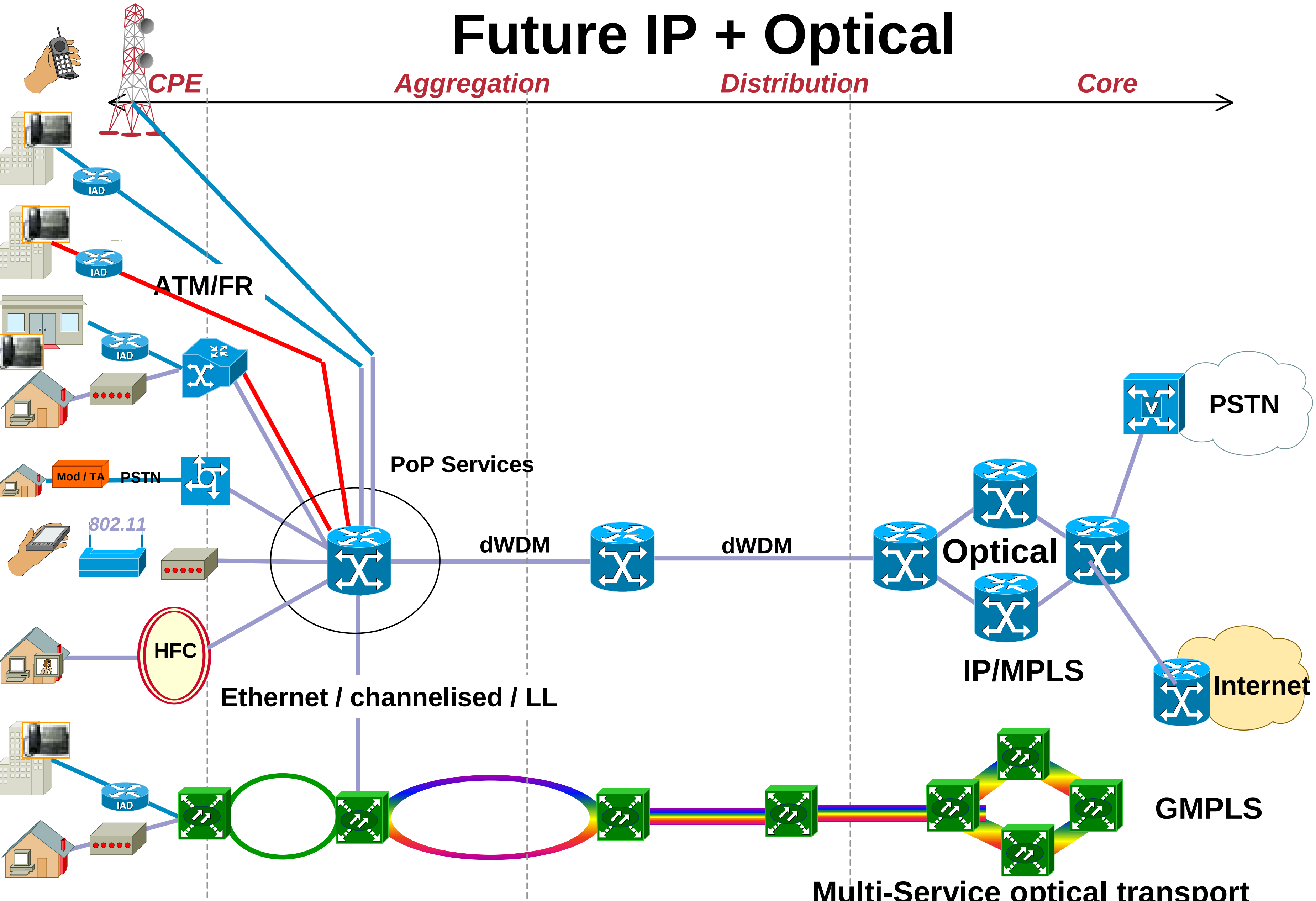


What is Happening in Core ?

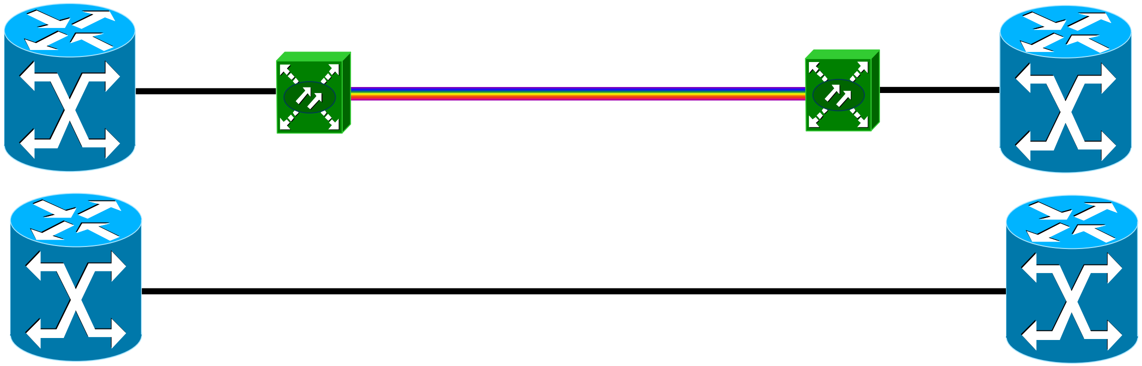


- Core bandwidth is increasing
  - Broadband based
  - New Business services
- Slot count pressure
- 10 Gbps in production in larger PTT networks
- 40 Gbps requirement appearing
- 100 Gbps under discussion !

Data Reference Architecture
Future IP + Optical

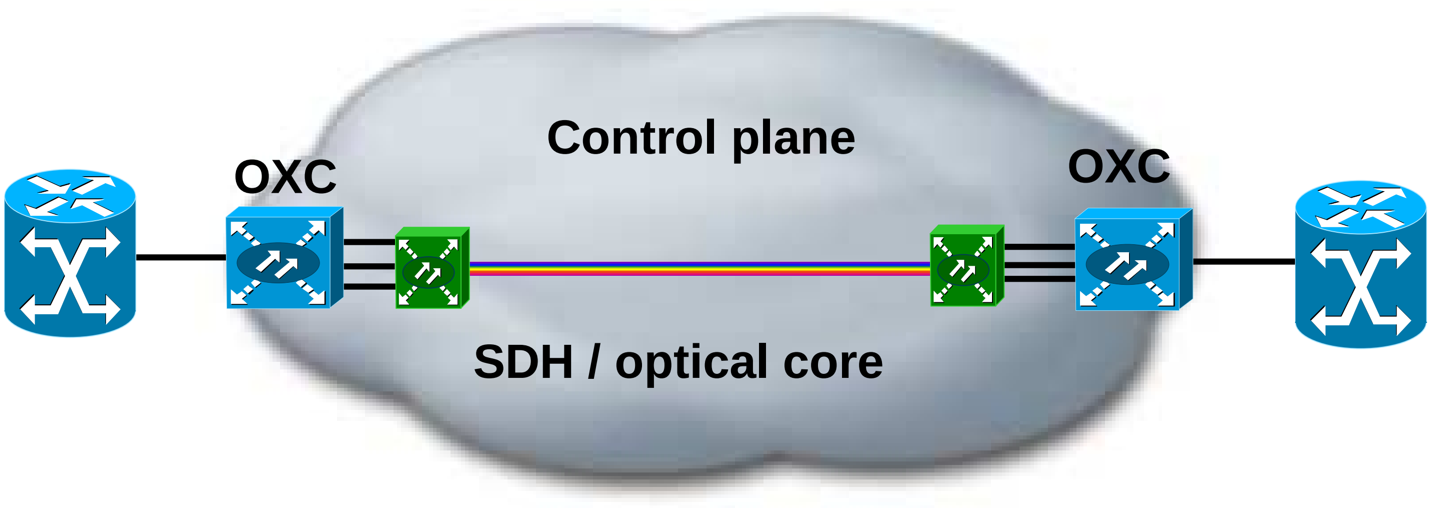


Core Infrastructures Option 1
P-to-P DWDM / Dark Fibre / GE Switches



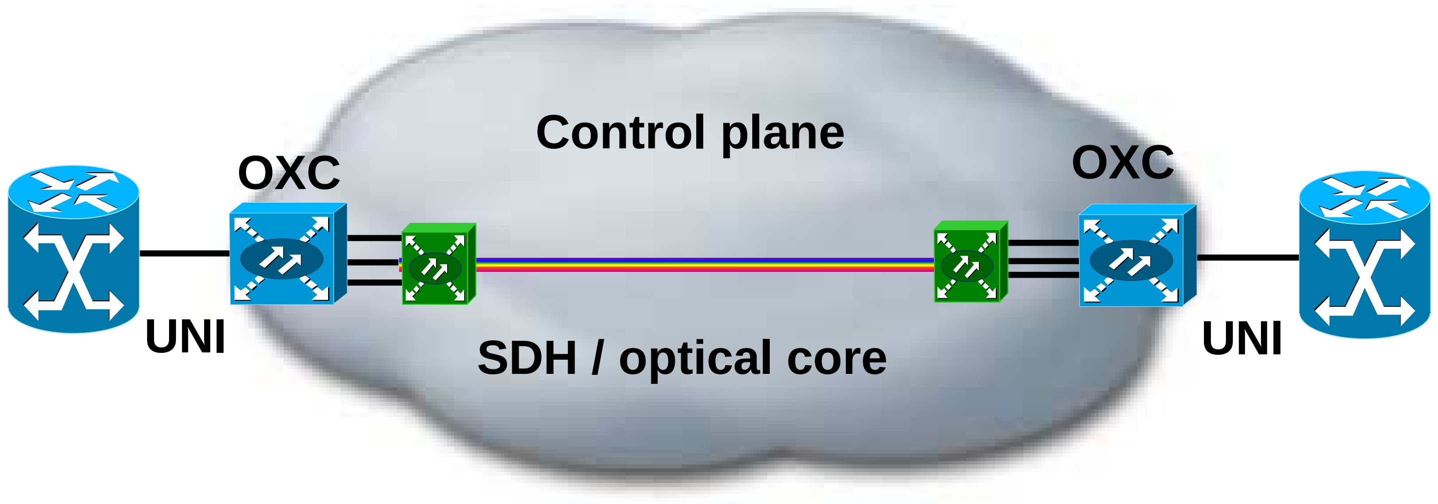
- **Simplest model**
- **Very high BW connections**
  - **STM-16c – STM-256c, RPR, GE, 10GE**
  - **WAN PHY & LAN PHY Long Distance**
- **Static - Does it matter ?**
- **No layer 1 recovery**
  - **L3 or FRR**
- **Cheap and efficient solution**

Core Infrastructures Option 2
Overlay without Signalling



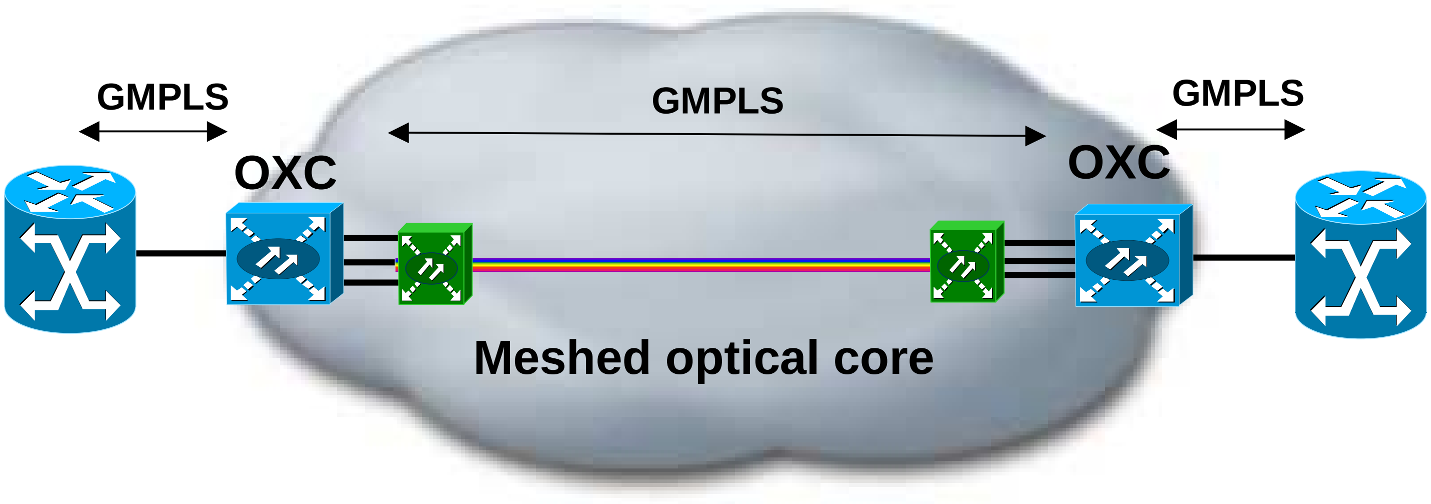
- Router connected to optical network
- No signalling interaction
- Limited interaction between Router and optical layer
- Backup at either L1 or L3
- More dynamic / more cost
- Bandwidth capabilities determined by SDH / Optical layer

Core Infrastructures Option 3 Overlay with UNI



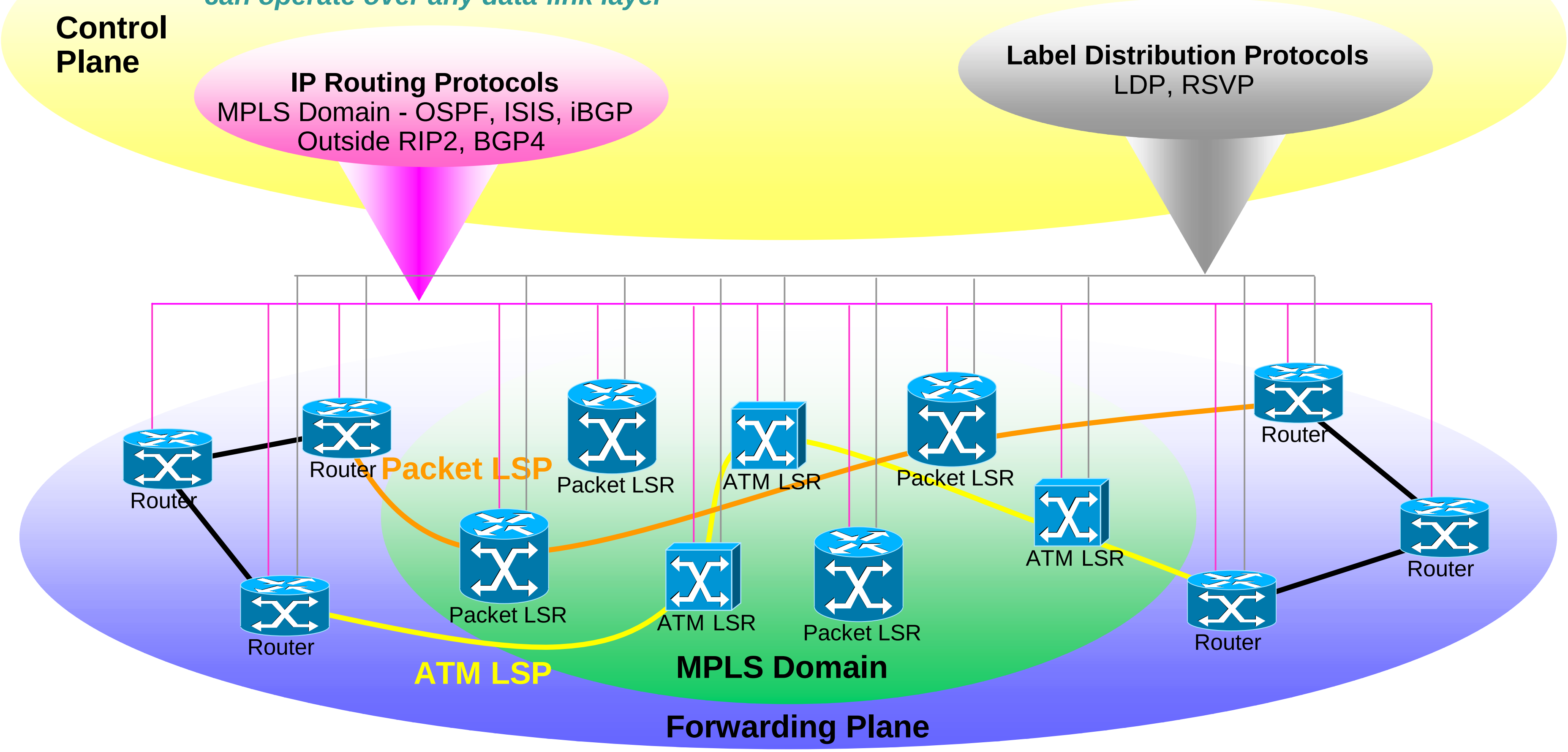
- Optical UNI interface between Router and Optical Layer
- Overlay model
- Dynamic bandwidth / BW on demand
    - Initiated from the edge
- Bandwidth capabilities determined by Optical Layer

Core Infrastructures Option 4
Peer Model – GMPLS / G.ASON / …



…. when MPLS started …

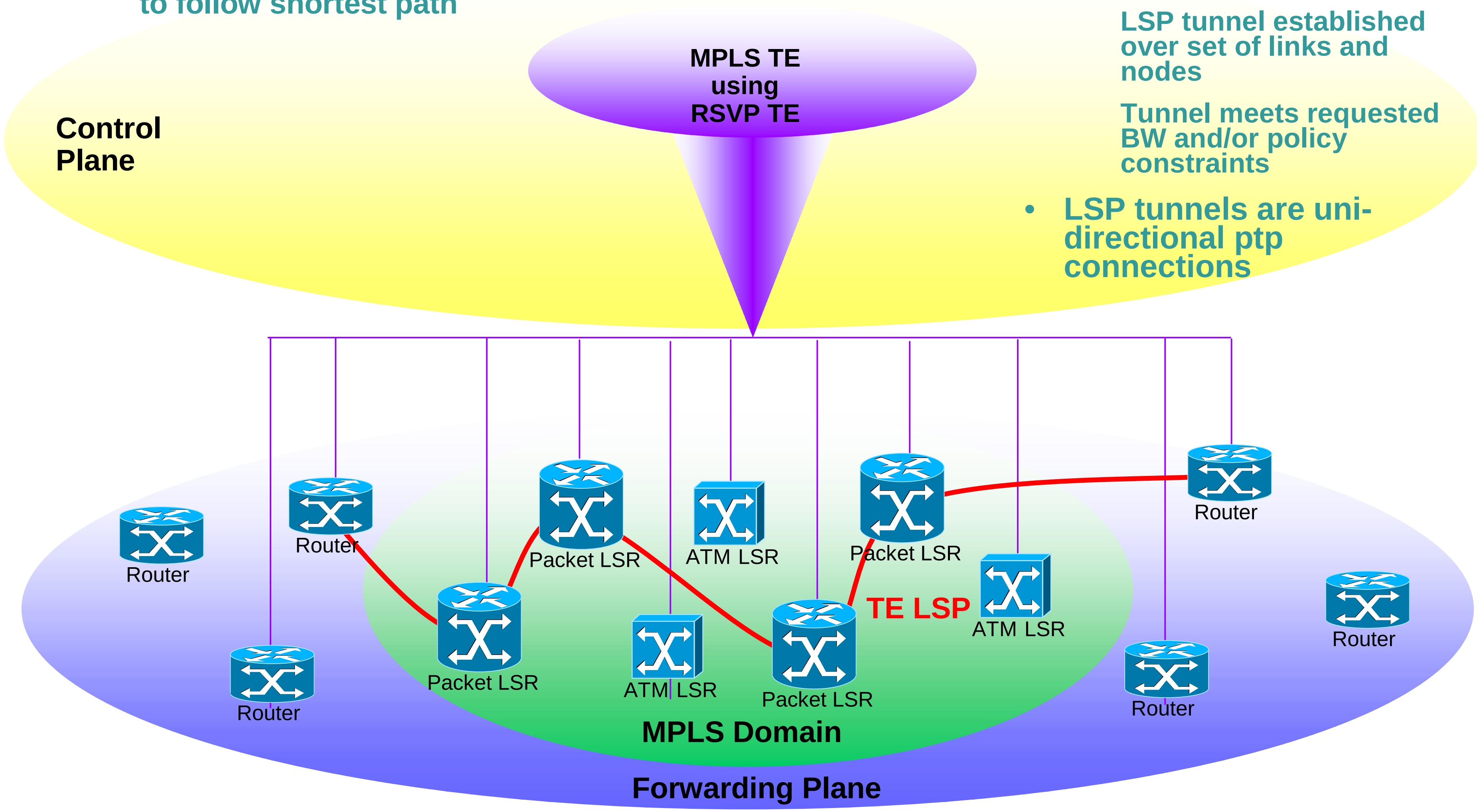
- *General-purpose tunneling mechanism*
  - *carry IP and non-IP payloads*
  - *uses label switching to forward packets/cells through the network*
  - *can operate over any data-link layer*
- *Separate Control Plane from Forwarding Plane*
- *Effort began 1996 ….. RFCs out 2001*
- *RFC 3031 MPLS Architecture*



…. MPLS TE emerged …

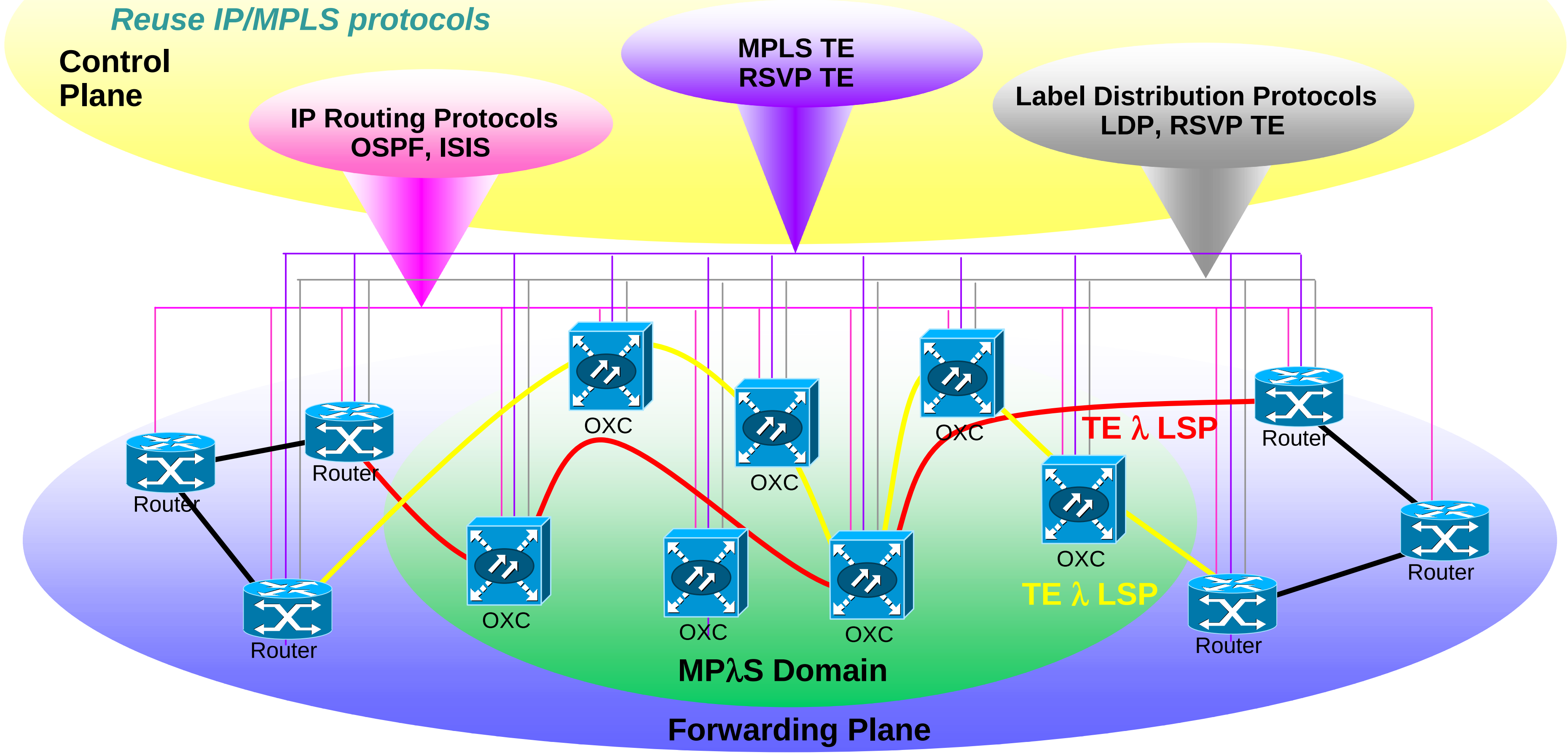
- Packets no longer need to follow shortest path

- Constraint-based routing
  - LSP tunnel established over set of links and nodes
  - Tunnel meets requested BW and/or policy constraints
- LSP tunnels are uni-directional ptp connections



…. then came MPλS …

- *Extend MPLS TE protocols to control optical cross-connect (OXC)*
  - *LSRs are like OXC*
  - *LSPs are like optical connections*
  - *Reuse IP/MPLS protocols*
- *Advantages*
  - *fast provisioning of optical connections*
  - *Unified IP/Optical Control Plane*
- *draft-awduche-mpls-te-optical-03.txt Q2 2001*

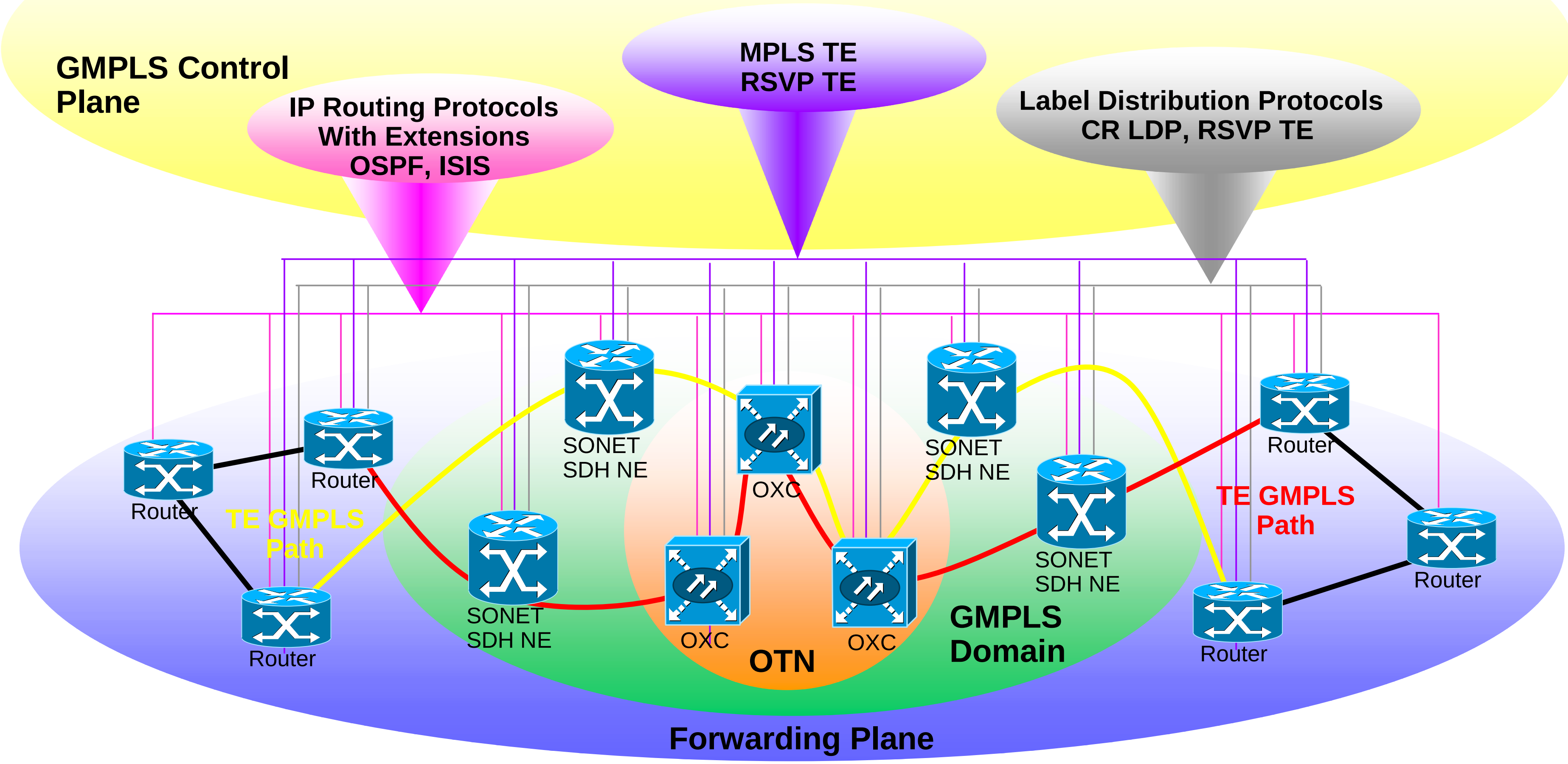


…. finally Generalized MPLS - GMPLS …

- *GMPLS control plane supports multiple switching and forwarding planes*
- *Introduces new functions to accommodate circuit-oriented optical network regimes*

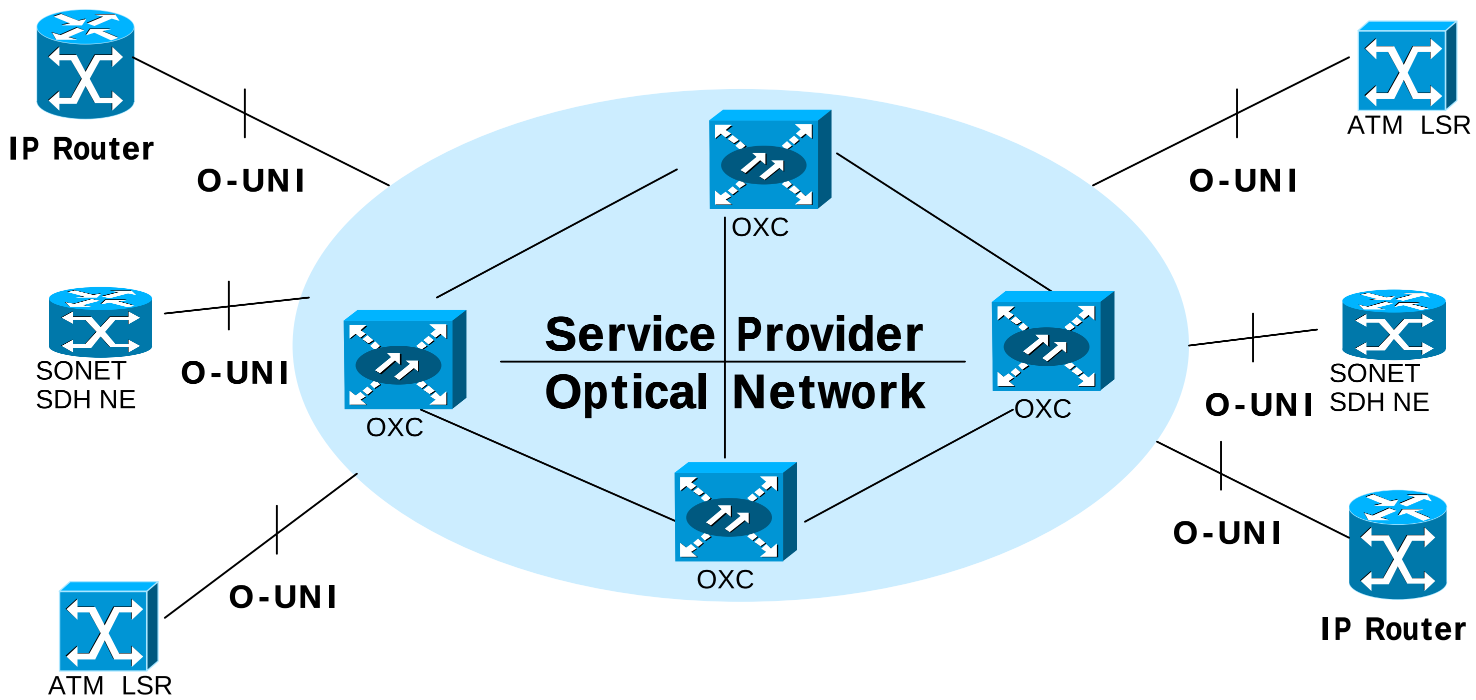
GMPLS = MPLS + MPλS + N

- where N is MPLS control of new switching planes
- draft-ietf-ccamp-gmpls-architecture-07.txt



O-UNI Multi-Service Network Applications

Cisco.com

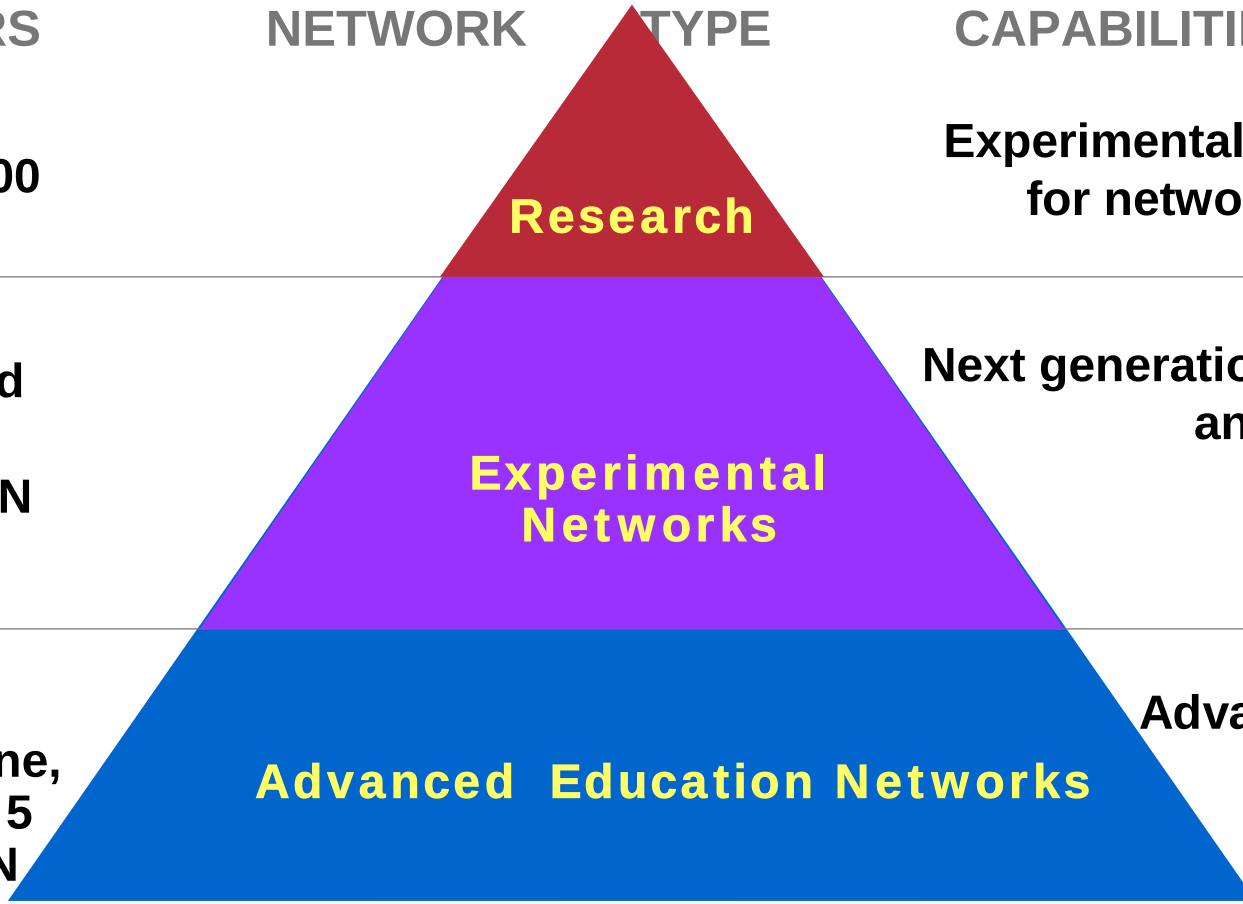


Service Provider offering dynamic optical paths for myriad of optical client equipment and networks

Offer Bandwidth On Demand, OVPN, and new Transport classes of services

Research & Education Network Tiers

LEADERS	NETWORK TYPE	CAPABILITIES/USERS
Web100 NLR	Research	Experimental environments for network researchers
Teragrid WIDE CALREN NLR	Experimental Networks	Next generation architecture and applications for research community
I2-Abilene, SurfNet 5 CALREN	Advanced Education Networks	Advanced services for education
ISPs	Commodity Internet	General Use



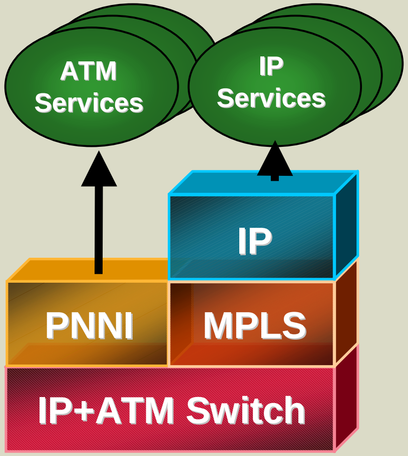
Agenda

Cisco.com

- Dynamics and Background
- Layer 3 : Half-Duplex VRF
- Inter-Provider : Layer 3
- Inter-Provider: Layer 2
- A Word on VPLS
- A Word on Traffic Engineering
- Management Considerations and MPLS OAM
- Security Considerations
- What About G-MPLS?
- Summary

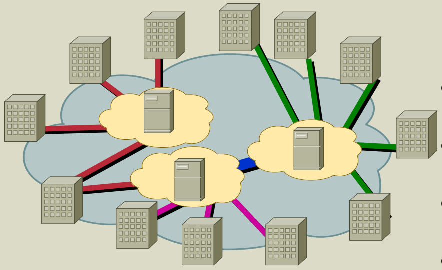
MPLS: The Key Technology for the delivery of L2 & L3 Services

Cisco.com



IP+ATM: MPLS Brings IP and ATM Together

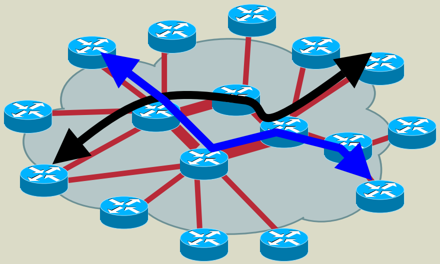
- eliminates IP “over” ATM overhead and complexity
- one network for Internet, Business IP VPNs, and transport



Network-Based VPNs with MPLS: a Foundation for Value Added Service Delivery

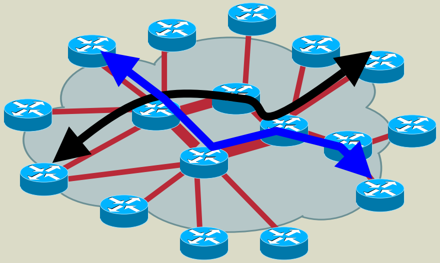
- flexible user and service grouping (biz-to-biz)
- flexibility of IP and the QoS and privacy of ATM
- enables application and content hosting inside each VPN
- transport independent
- low provisioning costs enable affordable managed services

MPLS: The Key Technology for the delivery of L2 & L3 Services



MPLS Traffic Engineering

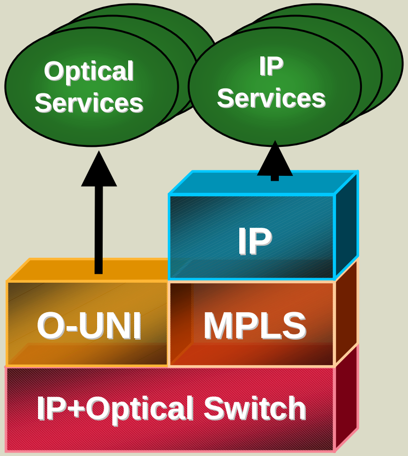
- Provides Routing on diverse paths to avoid congestion
- Better utilization of the network
- Better availability using Protection Solution (FRR)



Guaranteed Bandwidth Services

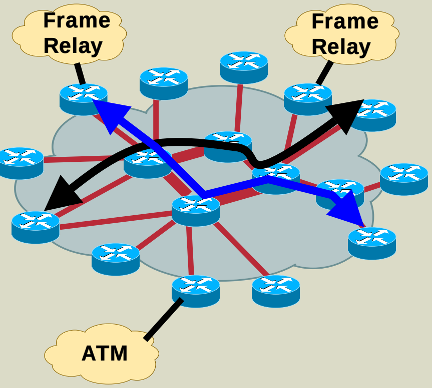
- Combine MPLS Traffic Engineering and QoS
- Deliver Point-to-point bandwidth guaranteed pipes
- Leverage the capability of Traffic Engineering
- Build Solution like Virtual leased line and Toll Trunking

MPLS: The Key Technology for the delivery of L3 Services



IP+Optical Integration

- eliminates IP "over" Optical Complexity
- Uses MPLS as a control Plane for setting up lightpaths (wavelengths)
- one control plane for Internet, Business IP VPNs, and optical transport



Any Transport over MPLS

- Transport ATM, FR, Ethernet, PPP over MPLS
- Provide Services to existing installed base
- Protect Investment in the installed gear
- Leverage capabilities of the packet core
- Combine with other packet based services such as MPLS VPNs

Questions?