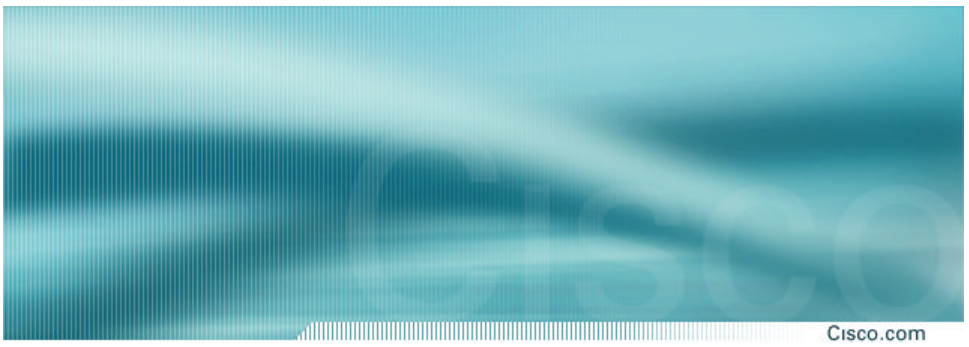


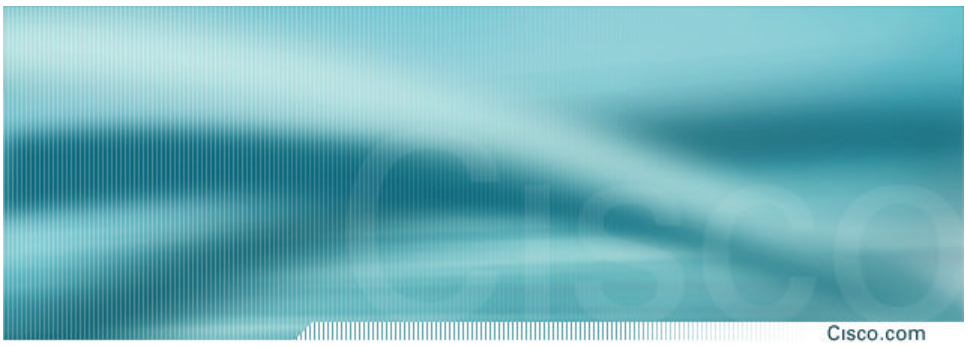
MPLS Implementation Status
Advanced MPLS VPNs

Peter Tomsu
Senior Consultant Cisco EMEA
ptomsu@cisco.com

Agenda

Cisco.com

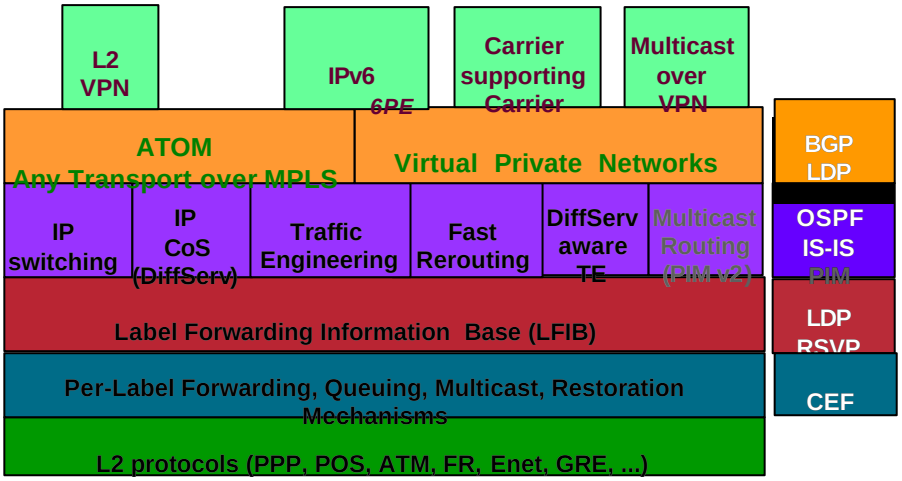
- MPLS Implementation
- MPLS VPN Concepts and Building Blocks
- Build a MPLS VPN
- MPLS VPN Scalability
- Advanced MPLS VPN Topologies
  - Extranet, Hub and Spoke, Internet Access, VPN Interconnect
- MPLS Dial VPNs
- Further Reading



MPLS Implementation

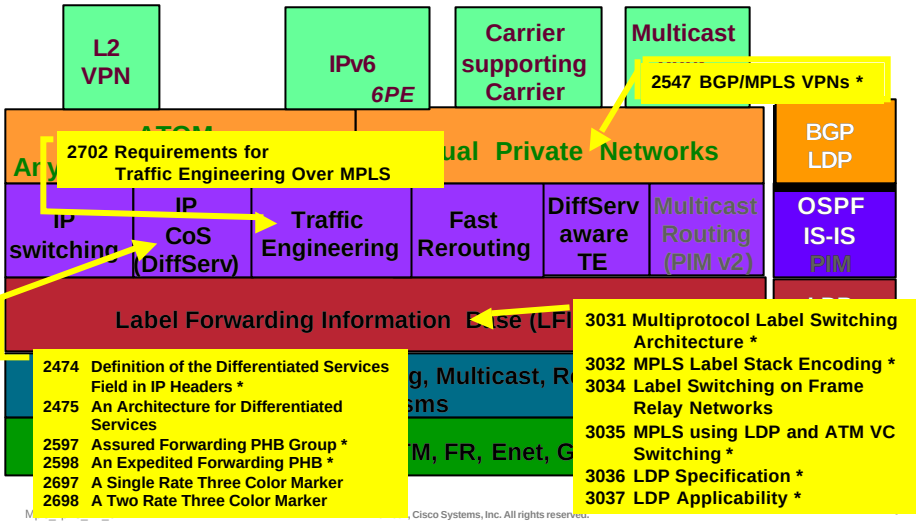
MPLS
Advanced Services

Cisco.com



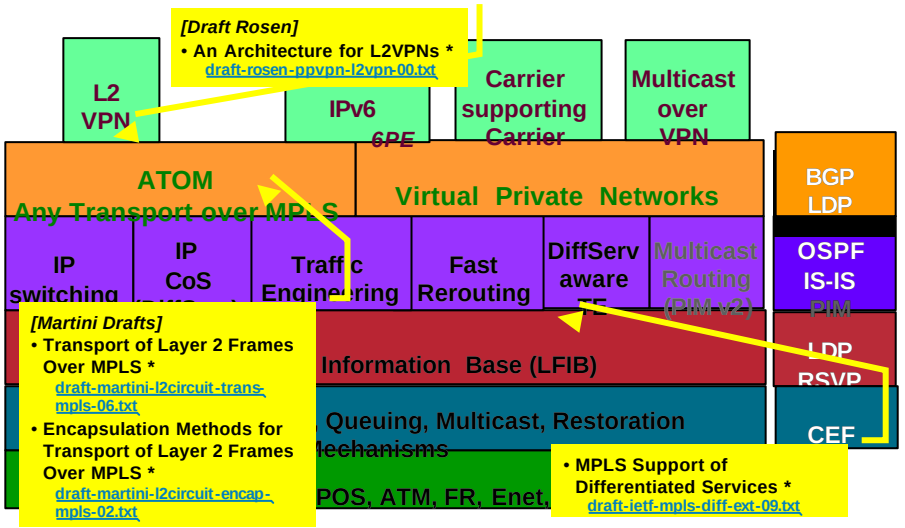
MPLS
Innovation & Standards

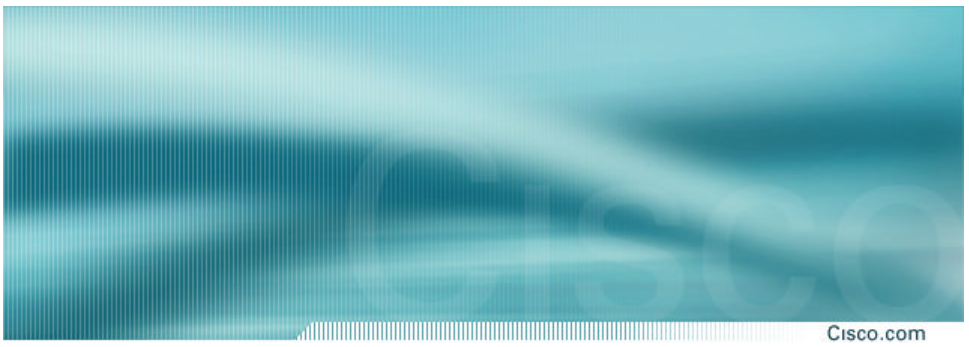
Cisco.com



MPLS
Innovation-in-Progress

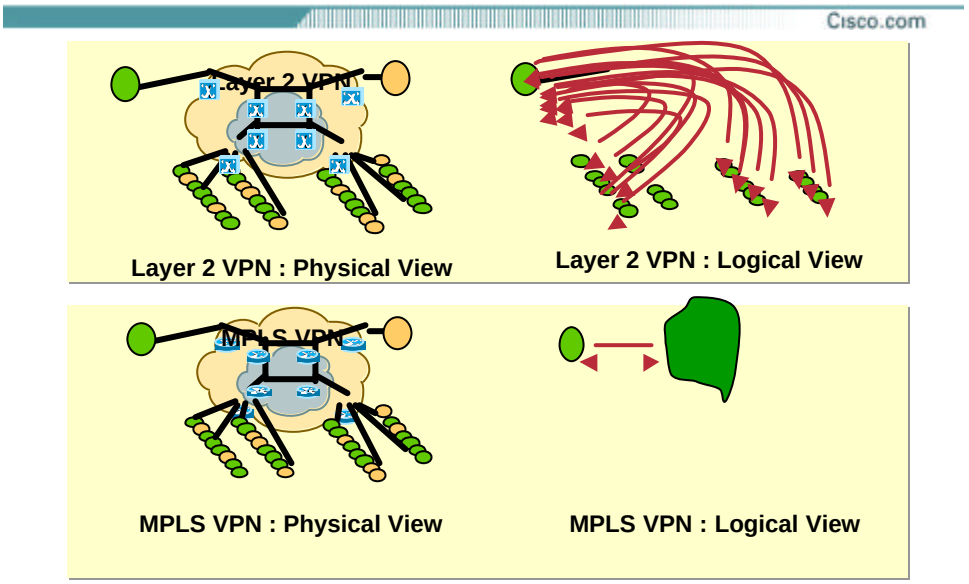
Cisco.com



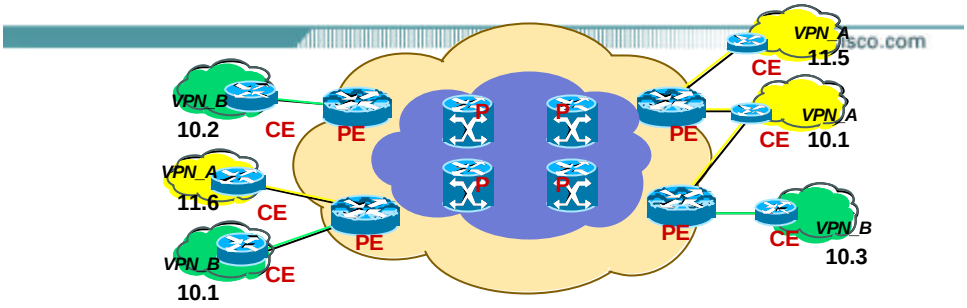


MPLS VPN Concepts and Building Blocks

MPLS VPN Model

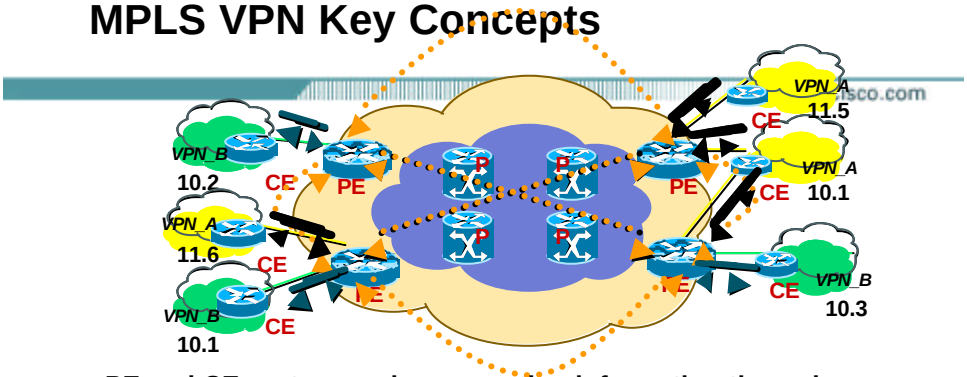


MPLS VPN Key Concepts



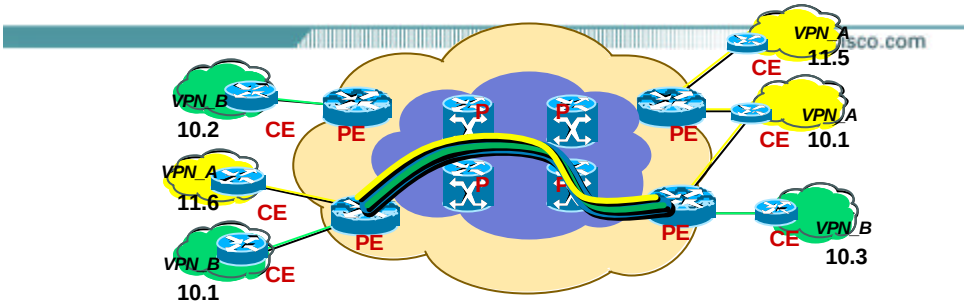
- Classical IP networks to connect
- CE router run standard routing software
  - No MPLS concept for customer
  - CE is basically in charge of network edge functions
  - Site may be a LAN directly attached to PE (Care to edge functions)
- An independant core IP network (MPLS)
- PE routers connect CE routers

MPLS VPN Key Concepts



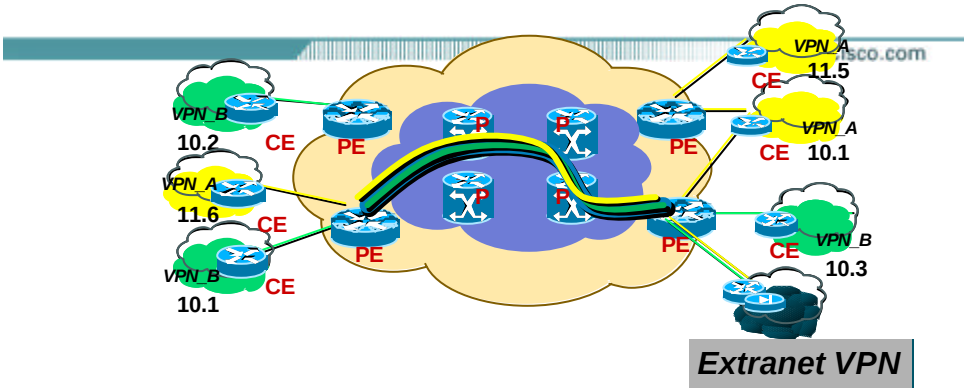
- PE and CE routers exchange routing information through:
  - eBGP, RIPv2, OSPF and Static routing
- Every PE builds an MP-iBGP adjacency to other PEs
  - Exchange { RD:Routing table }:
    - ie Signaling tunnelization
    - Per VPN private IP signaling plane
  - Blindly versus P network (no synchronization)
- Only interface VPN associated routes are advertised

MPLS VPN Key Concepts



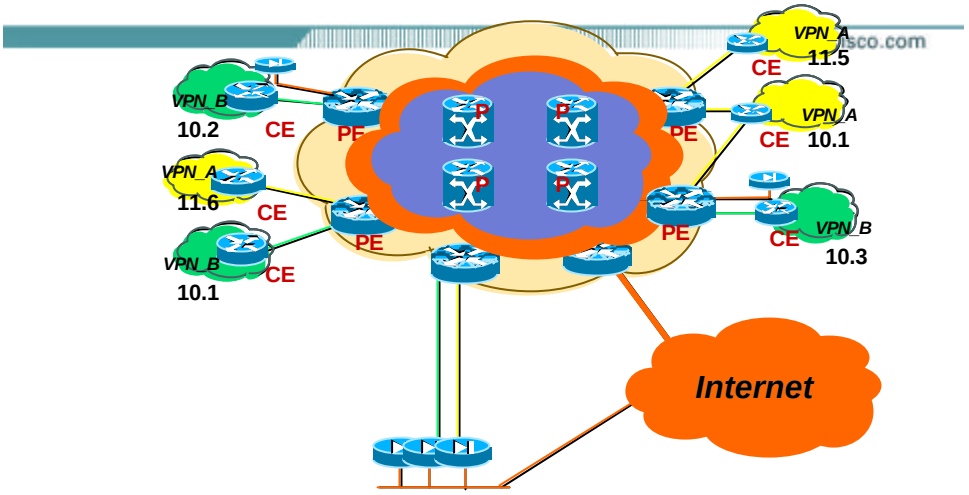
- Every PE establish a Label Switched Tunnel (MPLS) using LDP or RSVP
  - Connection-less
  - Follow OSPF / ISIS topology (or TE, or GB-TE)
  - Very fast (switched on label swapping)
- Packet exchange is done using dual stack of label
  - Switched all along the core using 'core label'
  - Switched to the destination interface using 'VPN label'
  - No IP exchange between VPN
  - ie Private switching at Data plane

MPLS *Flexible Architecture – Extranet VPN*

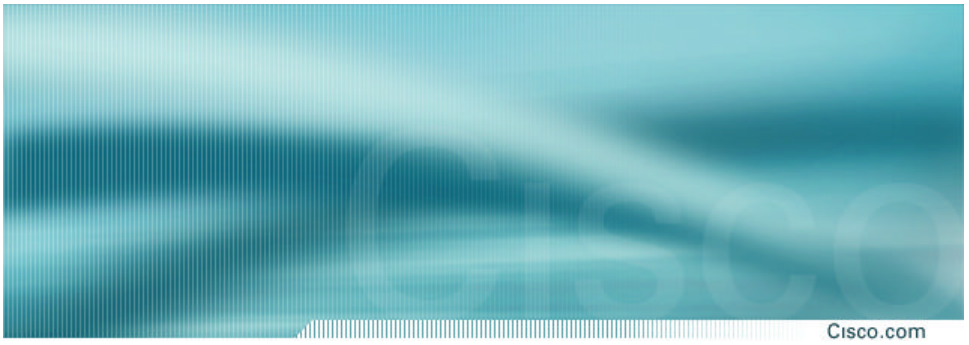


- A site may belong to many VPNs
  - Extranet services

MPLS *Flexible Architecture - Internet Access*

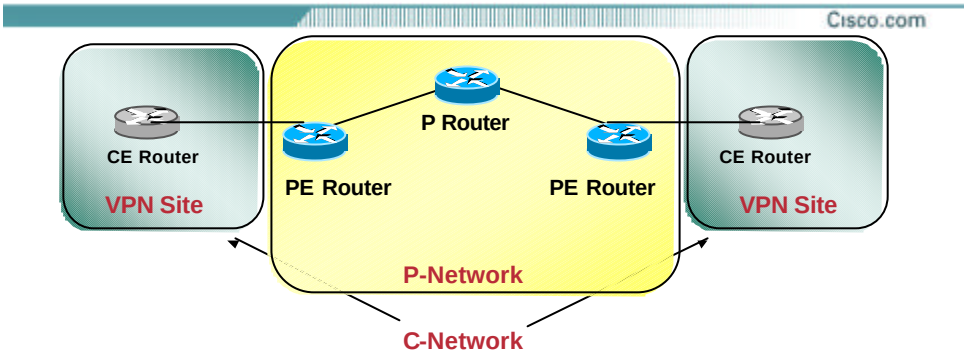


- Multiple Internet Access may be given at any point in the network
  - Per site through local VPN (sub-optimal routing toward Central site)
  - Down to site's DMZ



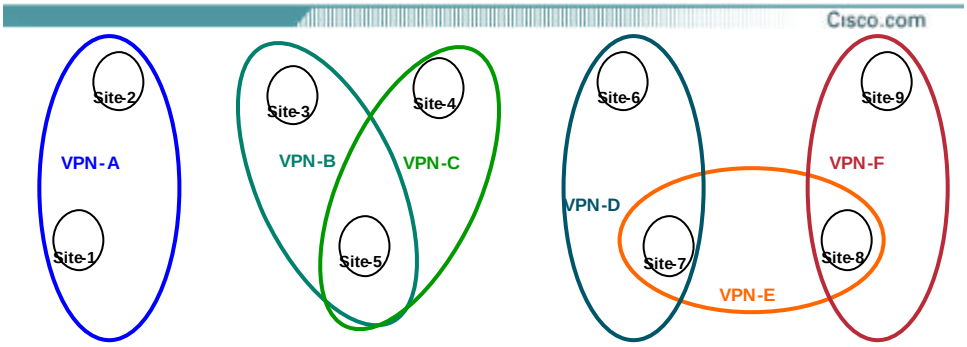
Build a MPLS VPN

MPLS/VPN Model



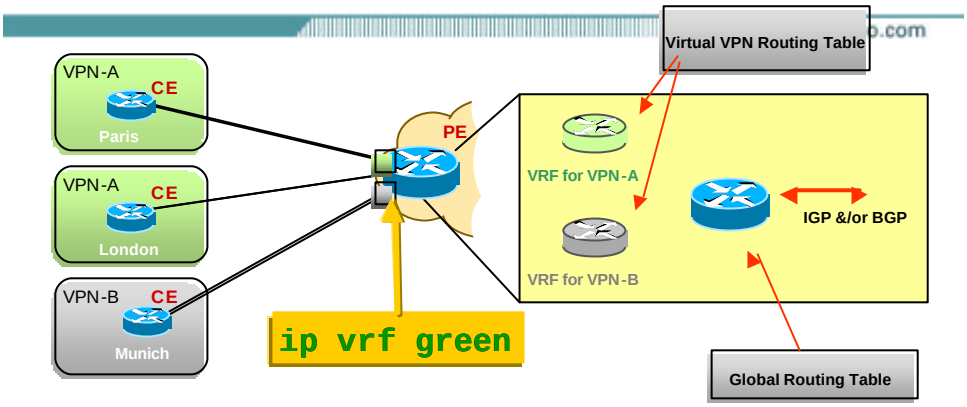
- A VPN is a collection of sites sharing common routing information – Common Routing Table
- A site can be part of different VPNs
- A VPN has to be seen as a community of interest (or Closed User Group)
- Multiple Routing/Forwarding instances (VRF) on PE routers

MPLS VPN Connection Model



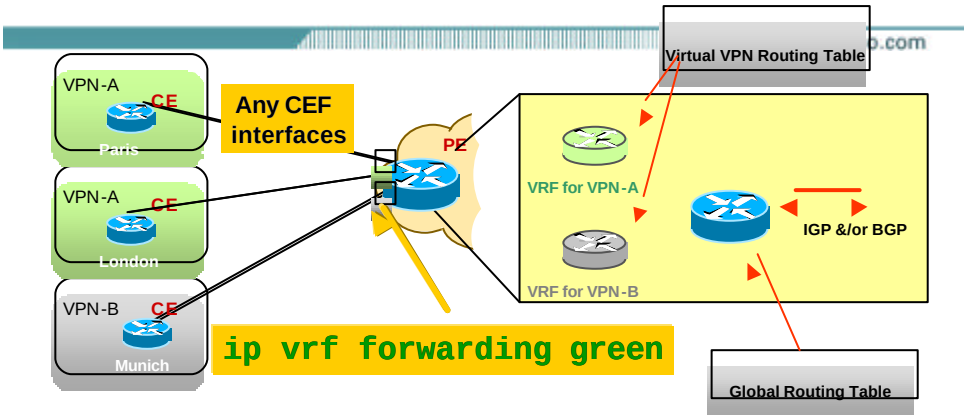
- A site belonging to different VPNs may or MAY NOT be used as a transit point between VPNs
- If two or more VPNs have a common site, address space must be unique among these VPNs

VPN Routing & Forwarding Instance - VRF



- **Multiple routing & forwarding instances (VRFs) provide the separation**
- **VRF can be seen as a Virtual Router's RT**

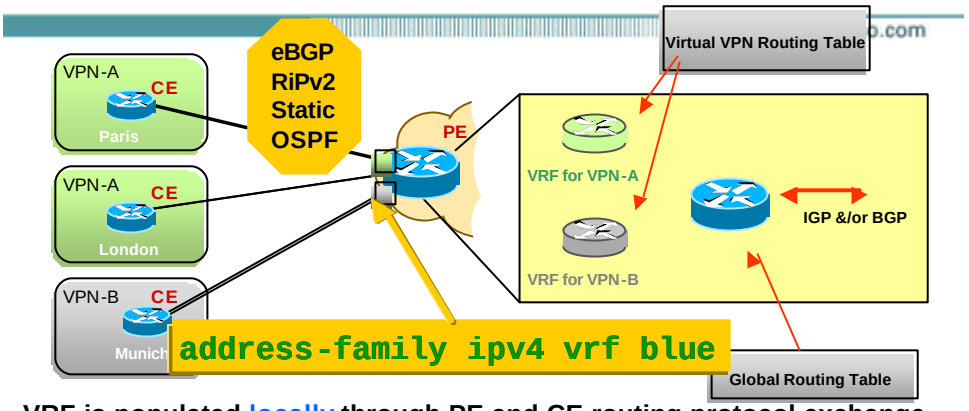
Assign Interfaces to VRF



Interface between CE - PE must be CEF

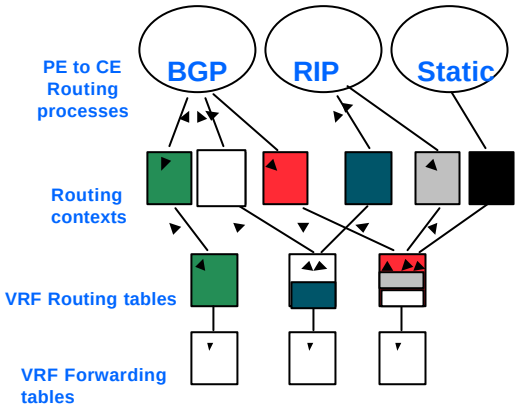
- POS, F/R, ATM, PPP, Ethernet
- GRE, L2TP
- Not for: (but may use GRE tunnel over)
  - X25, Token-ring
  - MLPPP and ISDN

Define Routing Exchange Between PE and CE



- **VRF is populated locally through PE and CE routing protocol exchange**
  - RIP Version 2, OSPF, BGP-4 & Static routing
- **Separate routing context for each VRF**
  - routing protocol context (BGP-4 & RIP V2)
  - separate process (OSPF)

Multiple Routing Protocols Filling 1 VRF



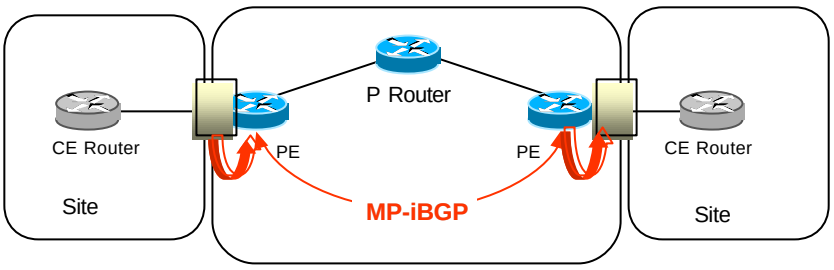
- Routing processes run within specific routing contexts
- Populate specific VPN routing table and FIBs (VRF)
- Interfaces are assigned to VRFs

VRF Route Distribution

Cisco.com

- **PE routers distribute local VPN information across the MPLS/VPN backbone**

  through the use of MP-iBGP & redistribution from VRF receiving PE **imports** routes into attached VRFs



Control Route Advertisements into VRF MP-iBGP Update *RFC2283*

Cisco.com

- **VPN-IPV4 address – transported over MP-iBGP**
  - **Route Distinguisher: Makes the IPv4 route globally unique**

    64 bits

    RD is configured in the PE for each VRF

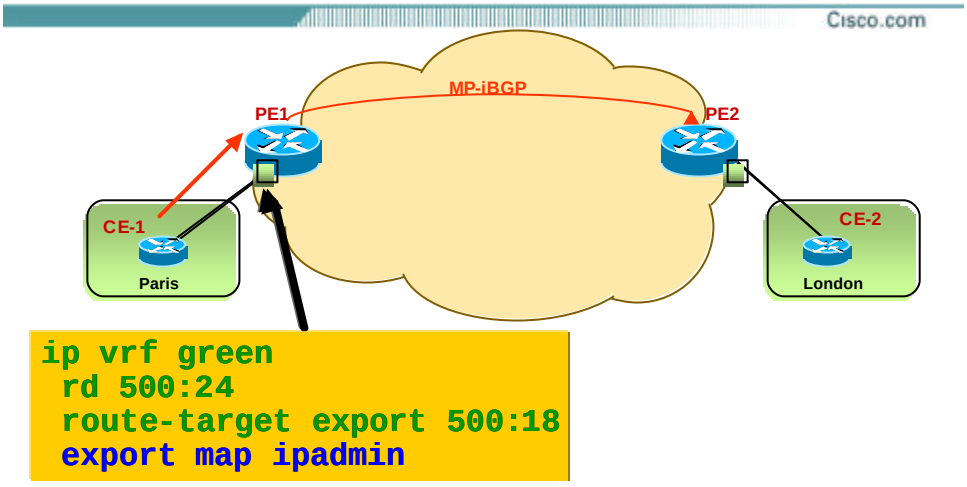
    RD may or may not be related to a site or a VPN
  - **IPv4 address (32bits)**
- **Extended Community attribute (64 bits)**
  - **Site of Origin (SOO): identifies the originating site**
  - **Route-target (RT): identifies the destination sites**

    RT acts as filter:

    – **RT export: Tag routes export criterias**

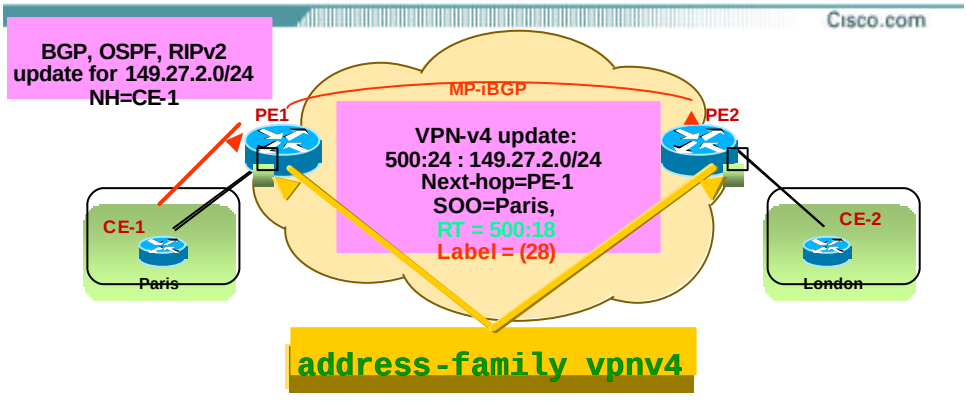
    – **RT import: Select the routes to import**

VRF Population via MP-iBGP



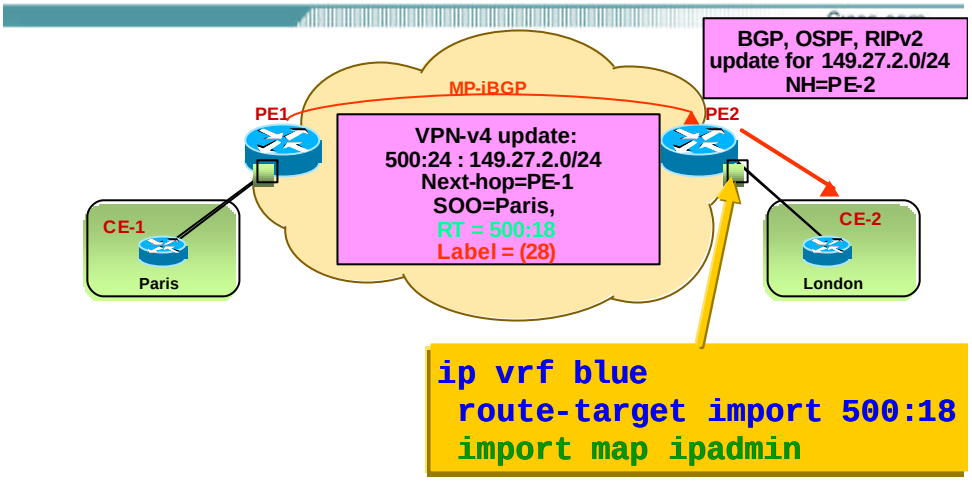
- **Give personality to a VRF**
  Group CE sites with same services in 1 VRF

VRF Population by MP-iBGP

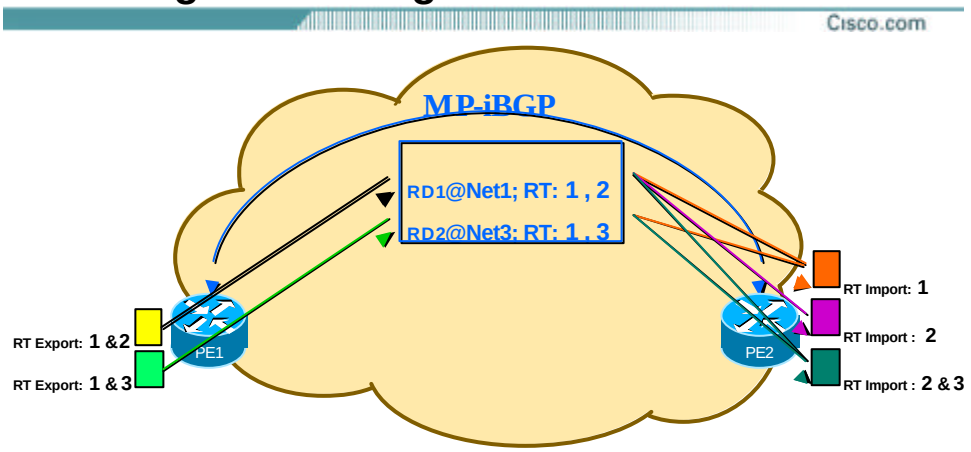


- **PE routers translate into VPN-V4 route**
  - Assign a RD, SOO and RT based on configuration
  - Re-write Next-Hop attribute (to PE loopback)
  - Assign a label based on VRF and/or interface
  - Send MP-iBGP update to all PE neighbors

VRF Population via MP-iBGP

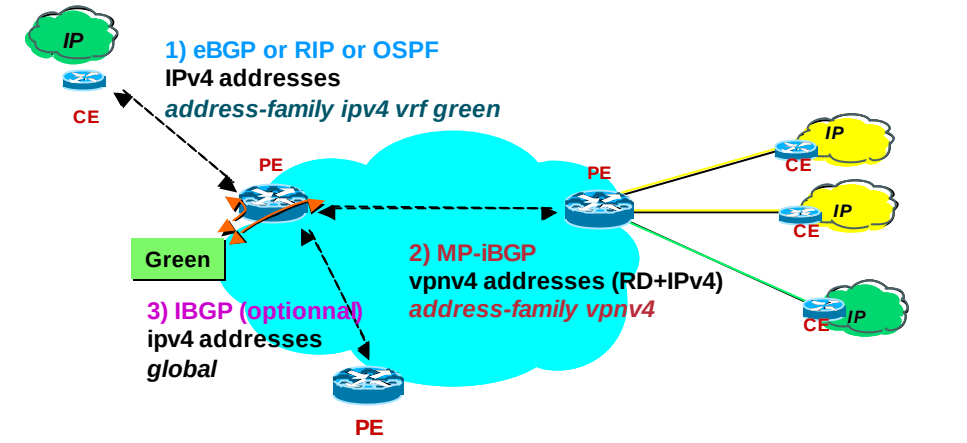


Populating VRF using Route-Target

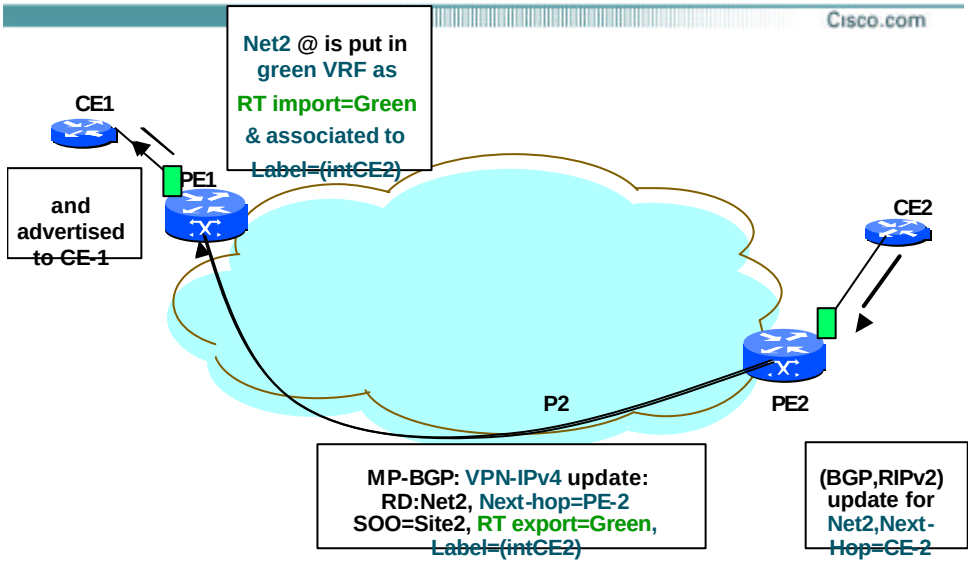


Route-Target (RT) are acting as Import/Export filters

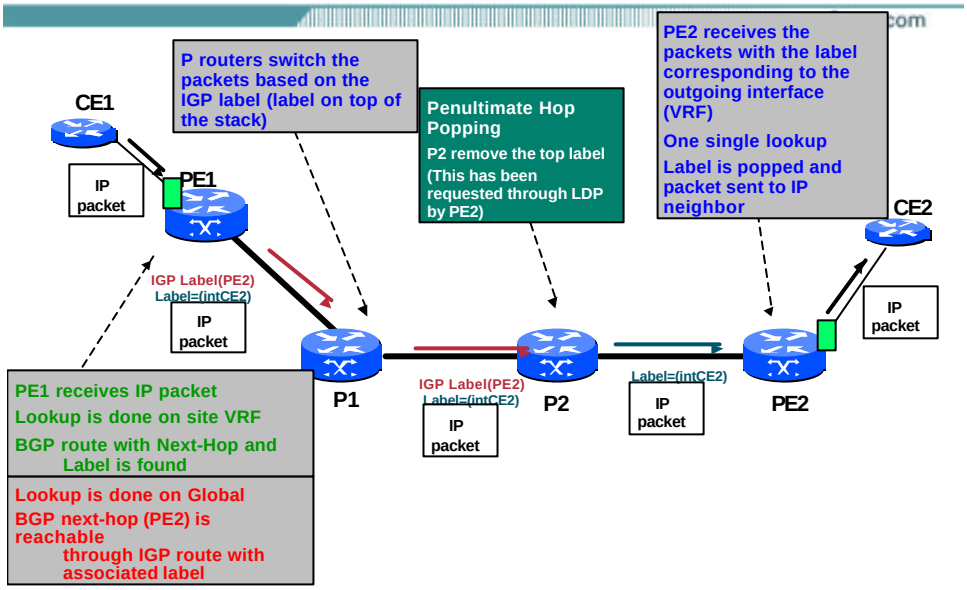
BGP Sessions



MPLS VPN
Signaling Layer



MPLS Forwarding
(& Penultimate Hop Popping)



MPLS VPN Connection Model
MP-BGP Update

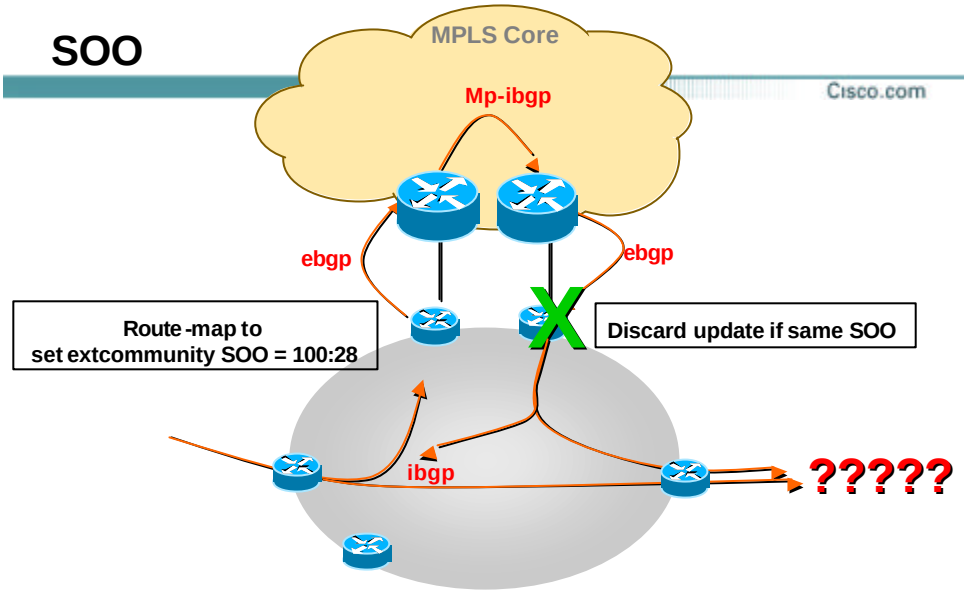
Any other standard BGP attribute

- Local Preference
- MED
- Next-hop
- AS_PATH
- Standard Community
- ...

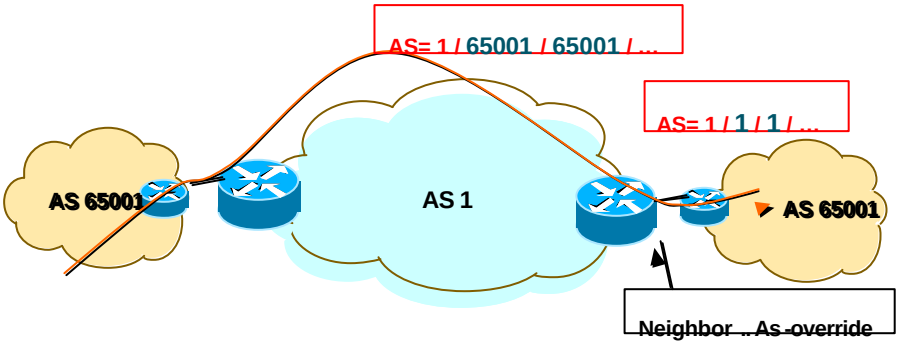
A Label identifying:

- The outgoing interface
- The VRF where a lookup has to be done (aggregate label)
- The BGP label will be the second label in the label stack of packets travelling in the core

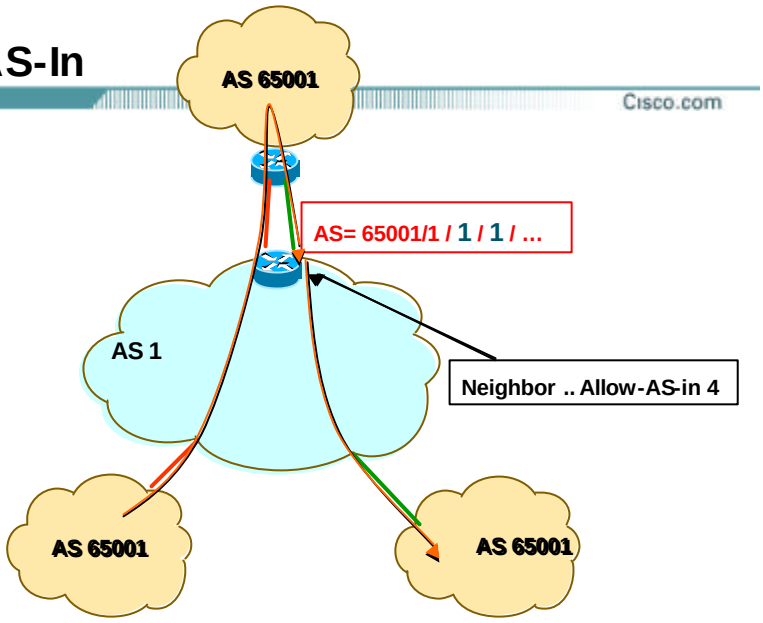
SOO



AS-Override



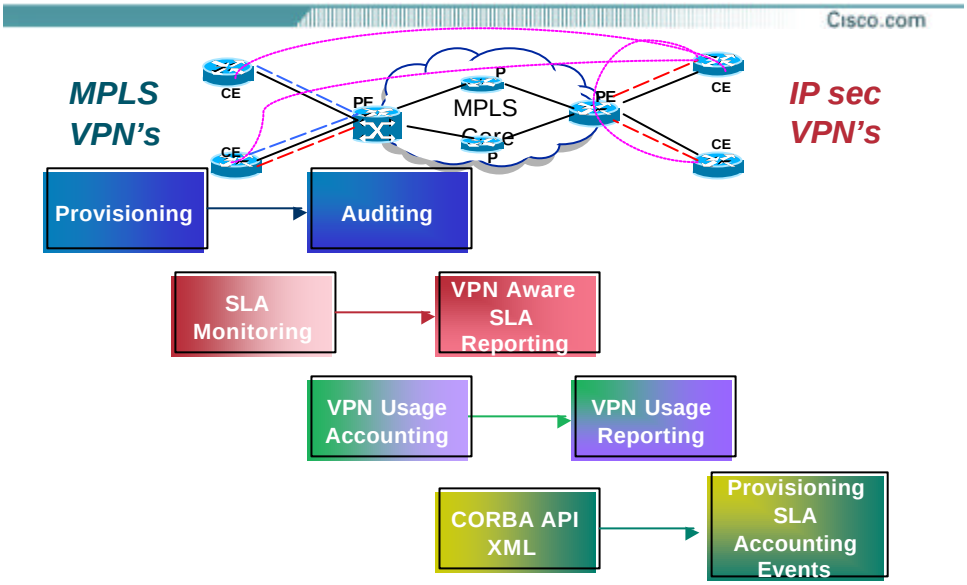
Allow-AS-In



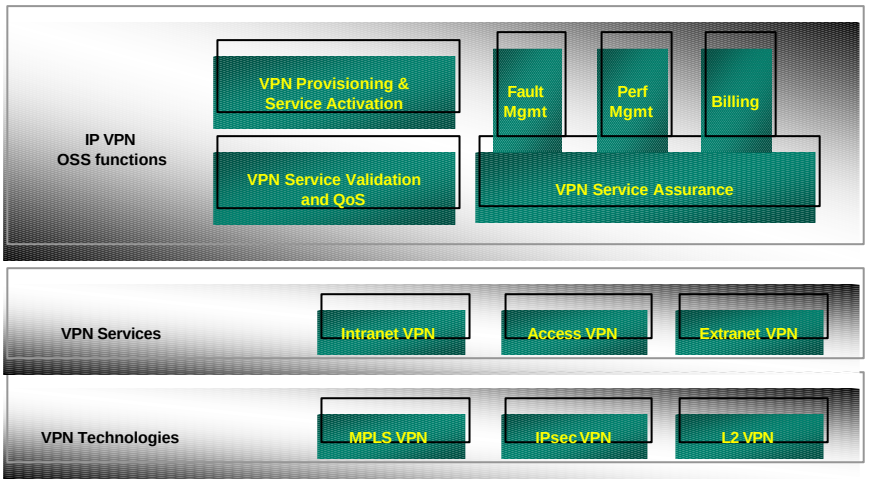
MPLS VPN Management

VPNSC

Cisco VPN Solutions Center VPNSC V2.0 Features

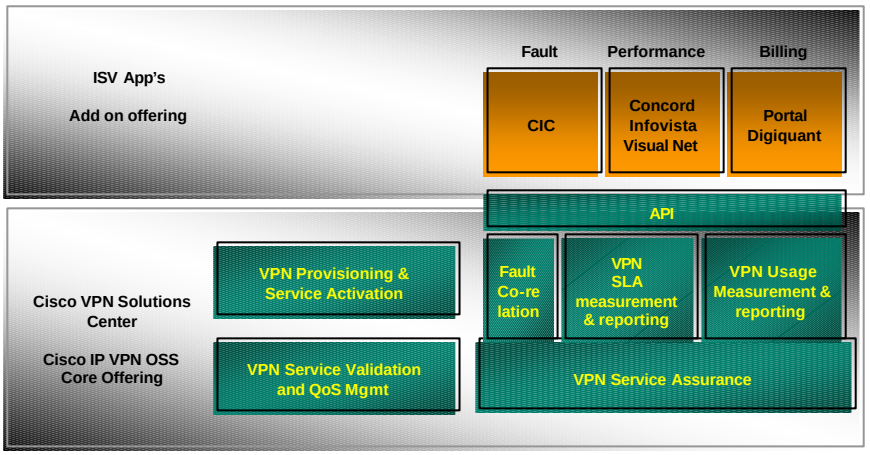


IP VPN OSS Requirements



IP VPN OSS

Cisco.com

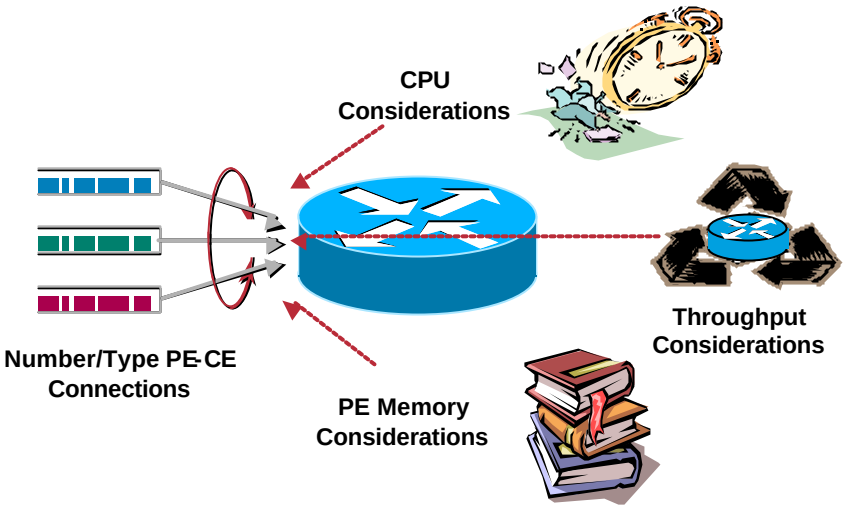


Cisco.com

MPLS VPN Scalability

Provider Edge (PE) Router Scalability

Cisco.com



VRF and Route Limits Summary

Cisco.com

- **VRF limits**
  - **Constrained mainly by CPU**
    - Between 500 and 1000 VRFs for static routing (depending on platform—10 routes per VRF)
    - Between 250 and 500 VRFs if using EBGP or RIPv2 (depending on platform—500 routes per VRF)
- **VPN and global route limits**
  - **Constrained mainly by available memory**

    With 256 Mb, 200,000 routes total (IPv4 and VPNv4)

    With 512 Mb, up to ~500,000 routes total

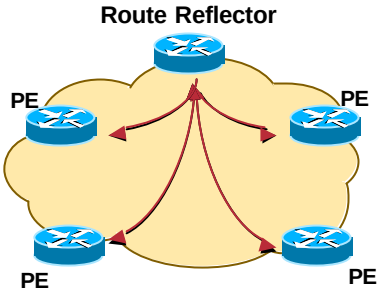
    If Internet table is present, this reduces the memory available for VPNs

    (current calculations are near 65 Mb for 100K Internet routes—with tightly packed attributes)

MP-BGP Deployment Requirements

Cisco.com

- Full iBGP mesh requirement between PE routers that require the same VPN information
- Easier to use Route Reflectors (and/or Confederations)
- Partition to further break up the topology and reduce processing overhead on PE routers



Remember – PE only needs routes for attached VPNs

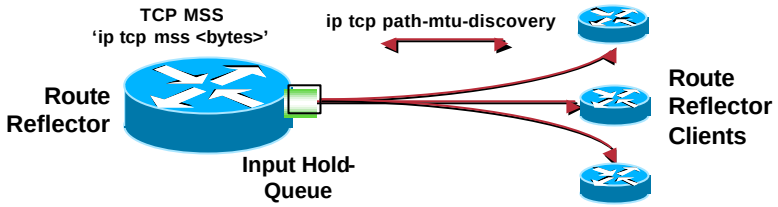
MP-BGP Used to Distribute VPN Prefix Information

RR Tuning and Recommendations

Cisco.com

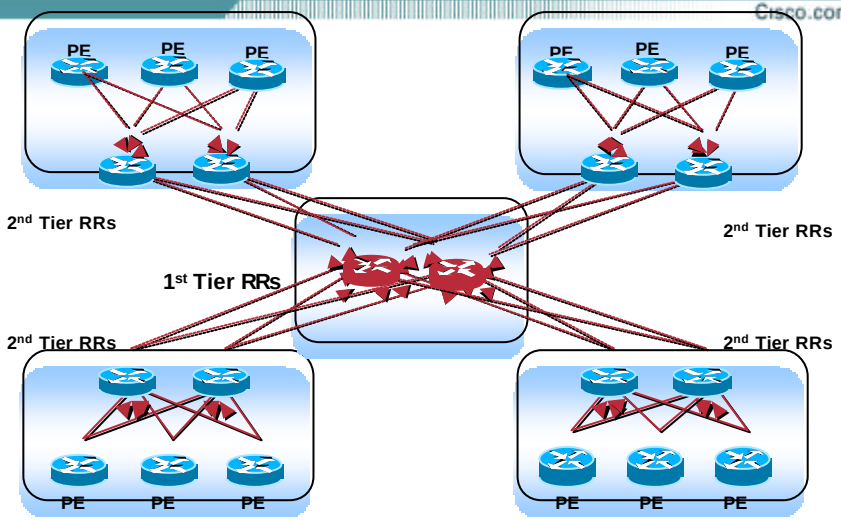
- **Keep RRs separate for IPv4 and VPNv4**
  - Better stability, faster convergence
  - Meets SLA requirements (with high number of routes)
- **Use RR Server Model**
  - Dedicated for RR function—not in forwarding path
  - Conserve CPU and memory for faster convergence
- **Recommended RR for VPN is NPE-400**
  - Highest CPU power
  - Availability of large DRAM memory—512 MB
- **Use Redundant RRs with Peer-Groups**
  - Lower CPU and I/O memory consumption
  - High improvement in convergence using same set of updates

RR Tuning and Recommendations



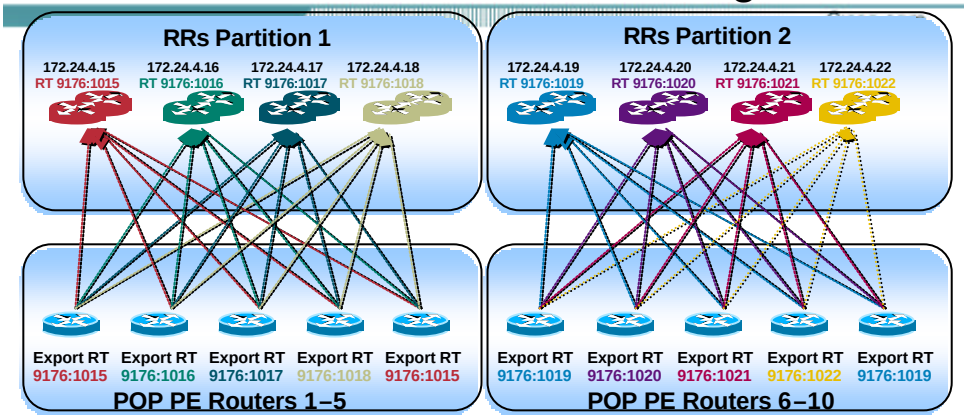
- On RR, increase input hold-queue and window-size
  - Window size/(MSS) * number of peers = hold-queue size
    - `hold-queue <packets> in`
  - Prevent input drops -> decrease convergence times
- Use ip tcp path-mtu-discovery
  - The packet rate will drop substantially;
    - default MSS size is 536 bytes

Distributed Route Reflection Design



DO NOT DO THIS FOR MPLS VPN !!!

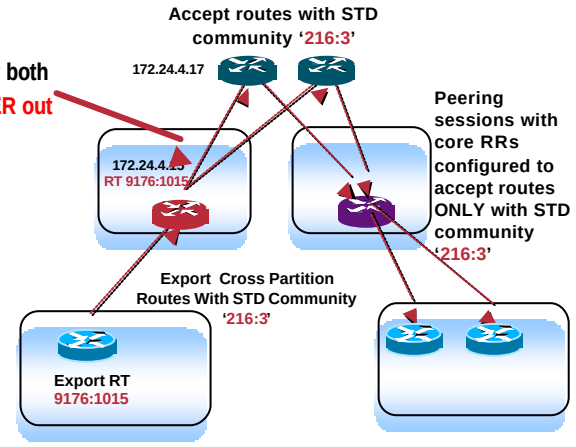
Centralized Route Reflection Design



- RT Value assigned for each core Route Reflector
- Each PE will export its routes with its own set of RT values but ALSO with the relevant Route Reflector RT value
- Allows the topology to be broken up based on number of peers/routes rather than any specific RT values which would be difficult to manage

Communication Between Partitions
Cross Partition Pollination Via 2nd Tier RRs


```
address-family vpnv4
 neighbor 172.24.4.15 activate
 neighbor 172.24.4.15 send-community both
 neighbor 172.24.4.15 route-map FILTER out
 exit-address-family
!
ip extcommunity-list 1 permit rt 100:1
!
route-map FILTER permit 10
 match extcommunity 1
 set community 216:3
```

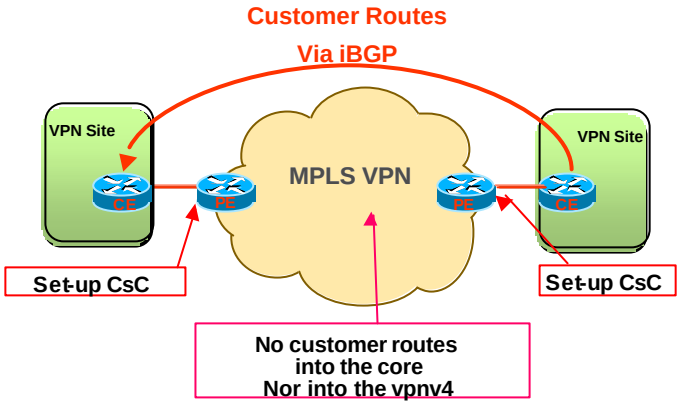


Cross Partition Pollination Via 2nd Tier RRs

The best way to scale
Use Carrier supporting Carrier

Cisco.com

- Achieved via a Fly-Over iBGP exchange between customer-CEs

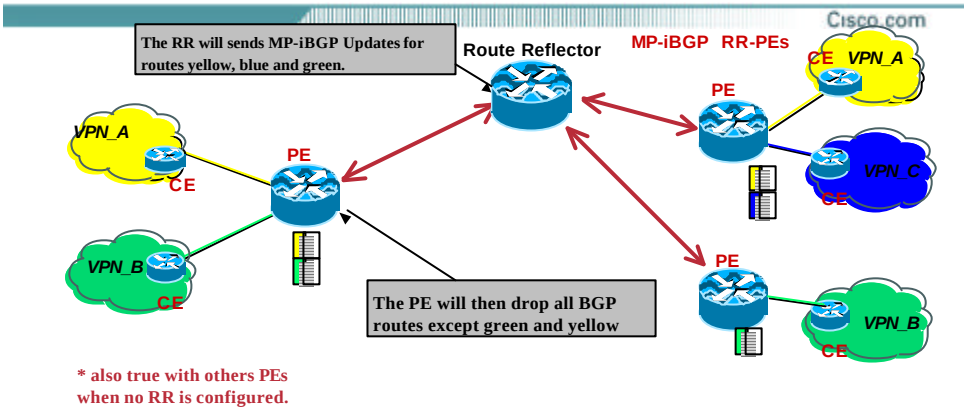


BGP Co-operative Route Filtering
ORF Entry

Cisco.com

- New BGP Capability: Route Refresh
- ORF Entry: Outbound Route Filter
  1. Type
     IP addresses (subnets) or Standard / Extended Communities
  2. Action
     ADD / DELETE / DELETE ALL
  3. Match
     PERMIT / DENY
  4. When-to-refresh
     IMMEDIATE / DEFER

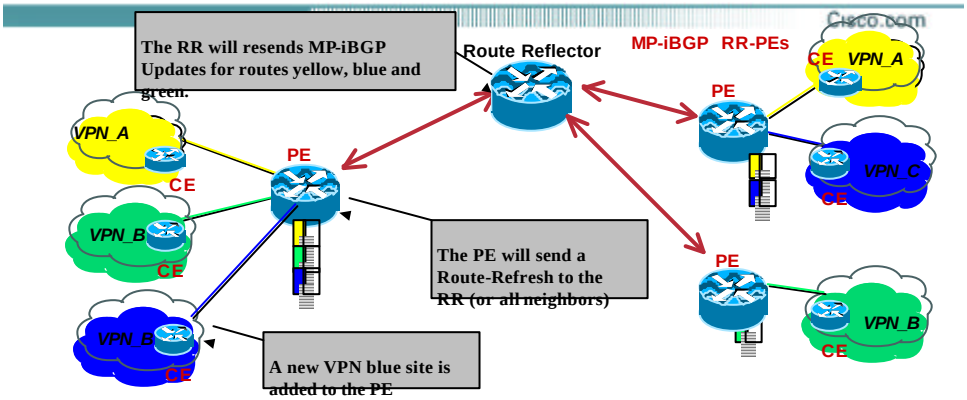
BGP Route Refresh / ORF



* also true with others PEs when no RR is configured.

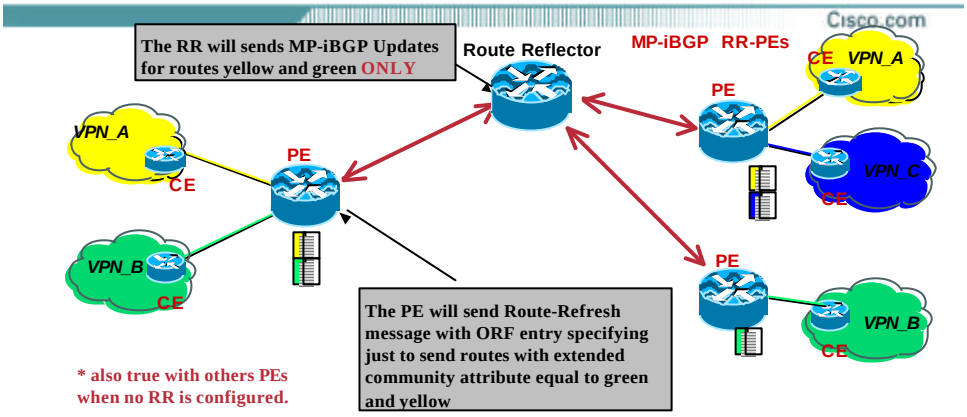
- For example, a PE has VFRs for green and yellow
- It receives VPN routes (through MP-iBGP updates) for VPN blue, the routes are simply discarded without any notification - Inbound Route Filtering

BGP Route Refresh / ORF

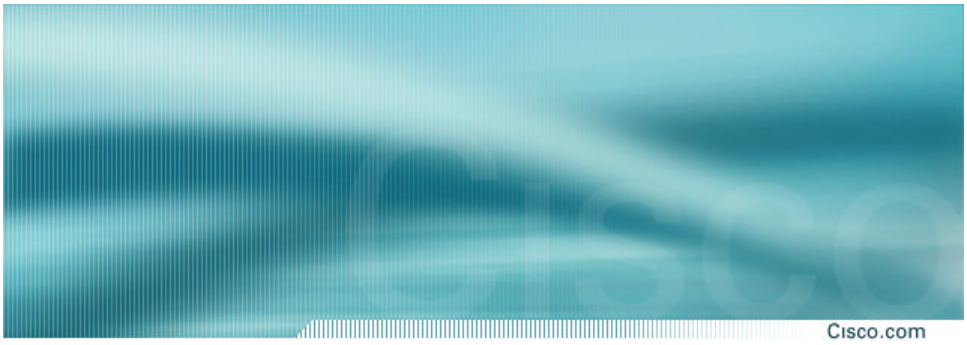


- New VRF for VPN blue is now added and the PE has dropped routes for VPN blue. The PE will send a Route-Refresh
- Another case is the removal of a VRF

BGP Route Refresh / ORF



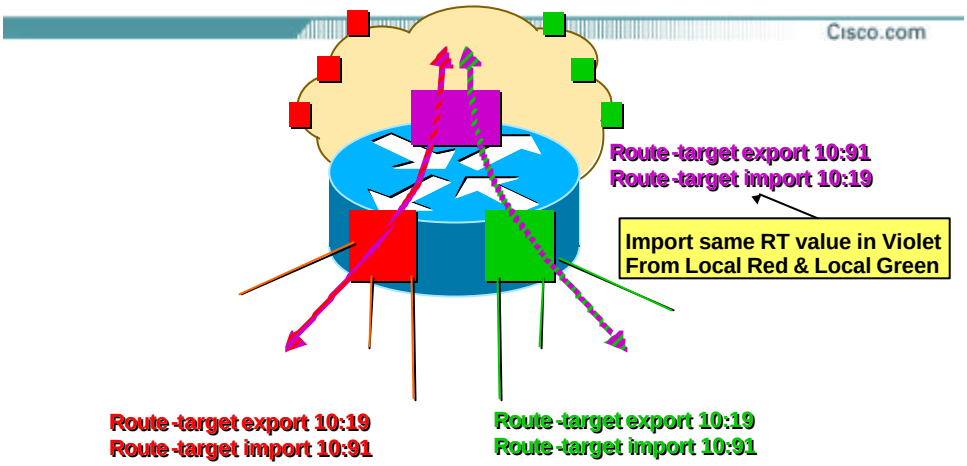
- Policy is changed again (I.e blue VPN site is removed)
- PE will send a new Route-Refresh message with new ORF entry
- Route Reflector will now send routes with extended community attribute “Route-Target is equal to Yellow and Green”



Advanced MPLS VPN Topologies

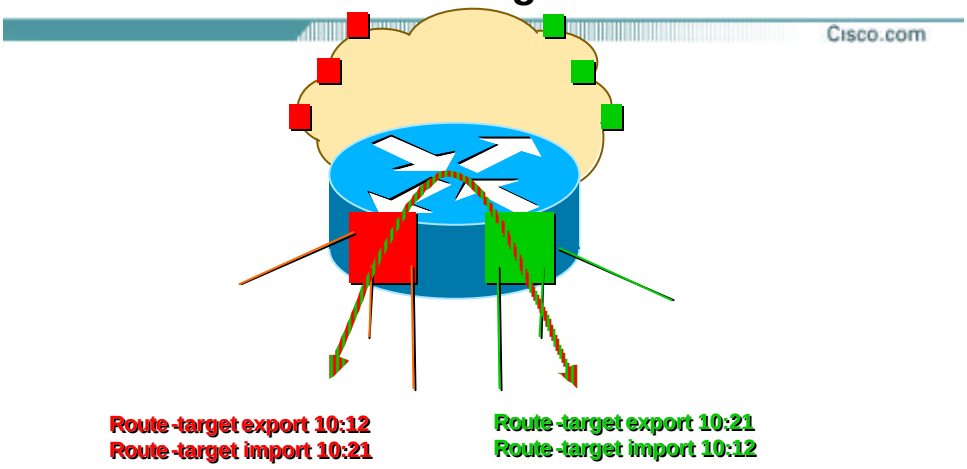
VRFs, Intranet, Extranet, Central Services, VLAN Interconnect

Common-VRF Routing



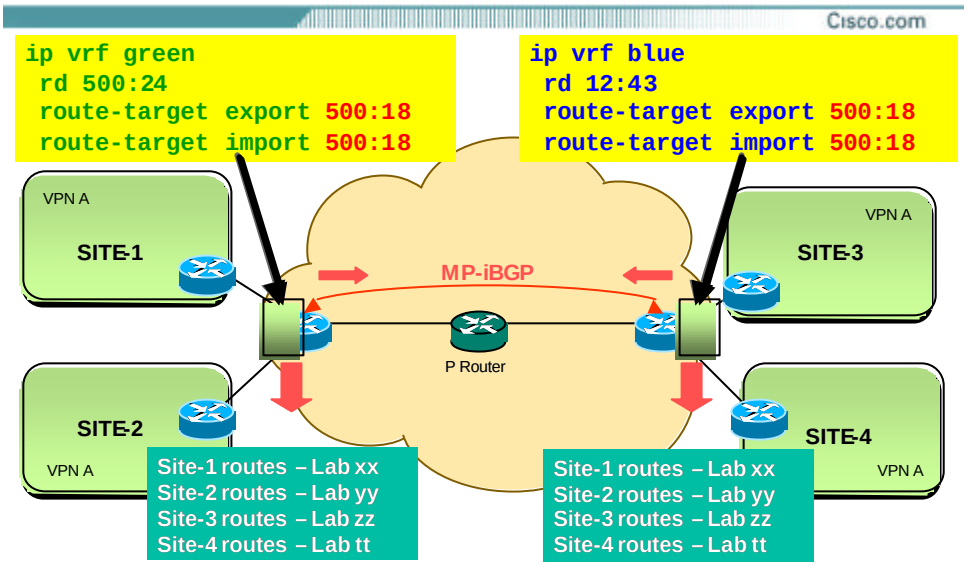
Violet is routed with Local Red & routed with Local Green
VPNs are still separated
(you can control export/import using Route-Map)

Local Inter-VRF Routing

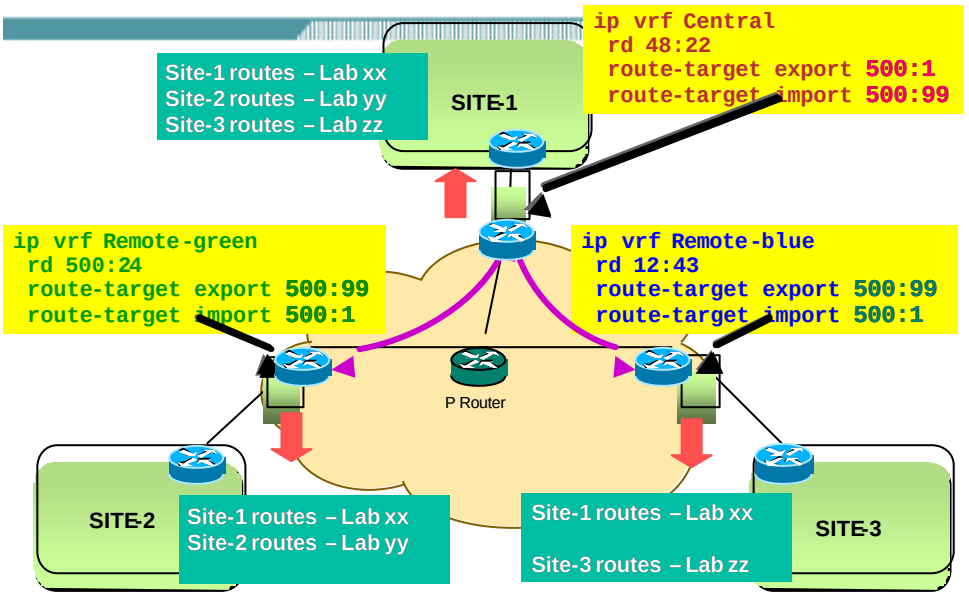


Local Red is routed with Local Green
VPNs are still separated
(you can control export/import using Route-Map)

Basic Intranet Model



Advanced Extranet Model



Central Services Model

- **Common topology is Central Services VPN**

  client sites may access central services but may not communicate directly with other client sites

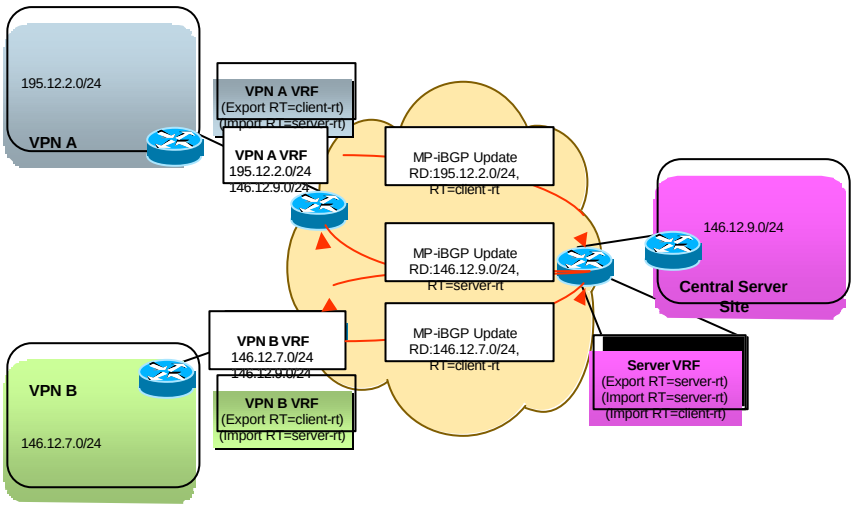
- **Once again controlled through the use of Route Target**

  client sites belong to unique VRF, servers share common VRF

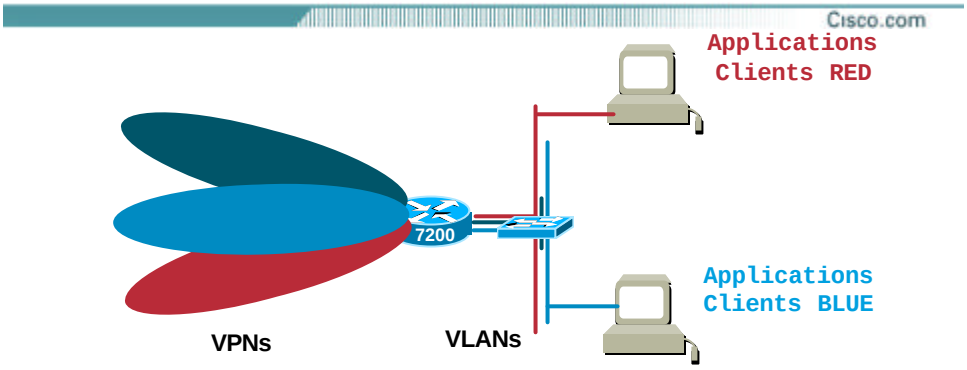
  client exports routes using **client-rt** and imports **server-rt**

  server exports routes using **server-rt** and imports **server-rt** & **client-rt**

Central Services Model



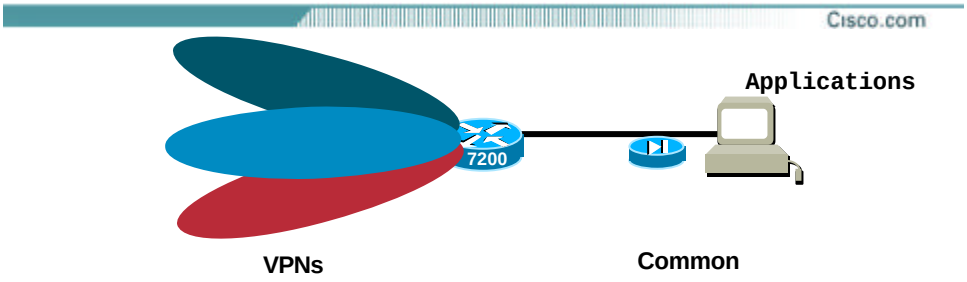
Server Hosting



VPN to VLAN mapping (one to one / or Many to one)

- **Common WAN network**
- **Common central site Lan Network**
- **Separated serveurs**

Application Sharing

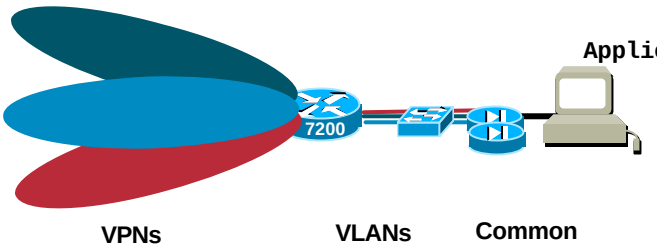


Multi-VPN Site

Interconnection services: NAT / Firewall
PIX
IOS Firewall

Beware of Address overlapping

Application Sharing

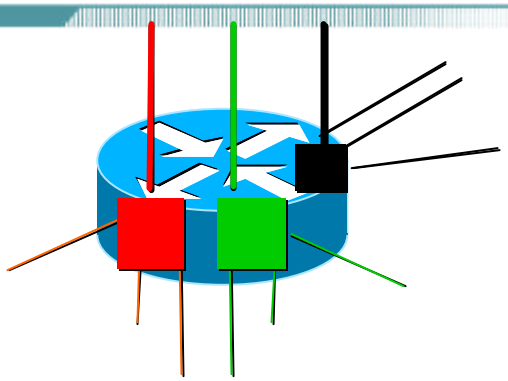


VPN to VLAN mapping (one to one / or Many to one)

Interconnection services: NAT / Firewall
- **PIX**
- **IOS Firewall**

Beware of Address overlapping versus Firewall

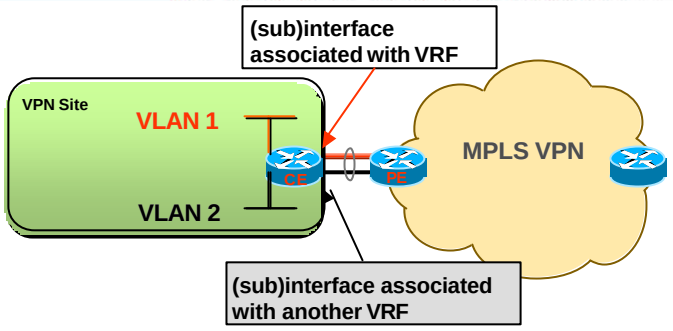
VRF-Lite
Standalone Virtual Router !



No MPLS, nor MP-iBGP

Local Inter-VRF routing is supported

VRF Lite
For Multi-VPN Extension to CE

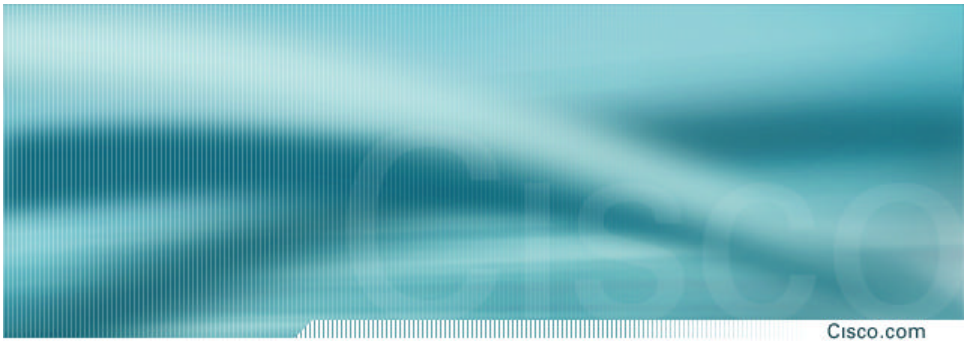


Allows to push 'PE-like' function to CE

- Independance of core versus edge (no peering between CE & all PEs)
- Using simple CE - no MP-BGP / no LDP

then some CE functions may not be supported (yet or not):

- » DHCP
- » NAT (But future NAT-PE)
- » CBAC / NBAR



Advanced MPLS/VPN Topologies

Internet Access

Internet Access

Cisco.com

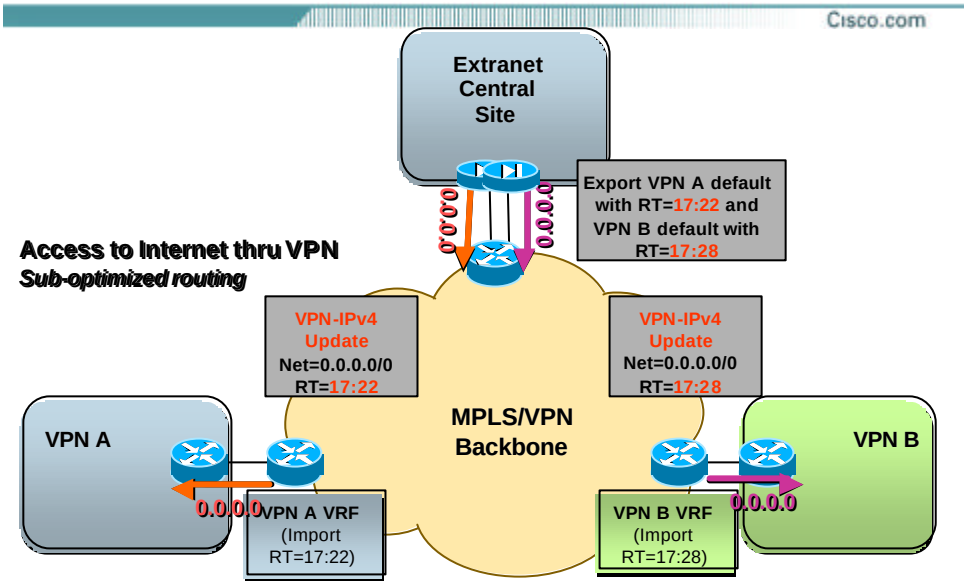
- **The problem:**
  - The Internet table is too big to be populated in many VRF
    - Ex: 100 VRF * 110.000routes = 11.000.000 !!!
  - It is not even recommended to push it into one only VRF
    - MP-iBGP is more consuming than iBGP
  - And even, it could be good not to distribute Internet in Global
    - The P routers have not to run BGP or to know Internet routes
    - No label is given to external BGP routes
  - Some PE are proxy to Internet gateway
    - They handle optimised access to Internet @

Internet Access
Four Ways Possible

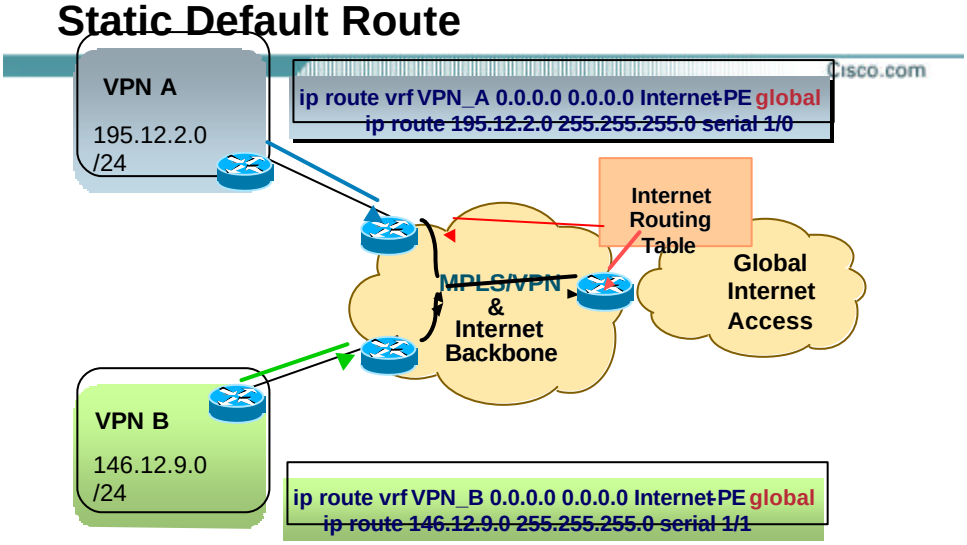
Cisco.com

- **Point to a Default-site to access Internet**
  - Hub & Spoke
  - Sub-optimised routing to Internet @
- **Push Internet flows from VRF to Global in PE**
  - Leak from VRF to Global
  - Optimise routing from PE to Internet
  - Security leakage (Use DMZ attached VRF for isolation/NAT/Firewall)
- **Build mixed MPLS VPN & Plain IP**
  - Use VRF for sub-interface with VPN service
  - Use plain IP for Internet access from DMZ
  - Optimal routing from CE to Internet
- **Use CsC**
  - To blindly transmit Internet route to CE

MPLS/VPN Internet Connectivity
Dynamic Default Route



MPLS/VPN Internet Connectivity
Static Default Route



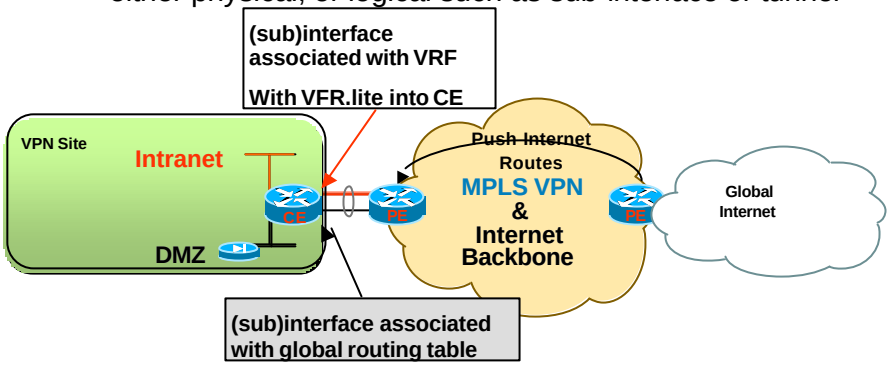
The PE router is also the Internet access router

MPLS/VPN Internet Connectivity
Dual Parallel Access

Cisco.com

- **Achieved by using a second interface to the client site**

  either physical, or logical such as sub-interface or tunnel

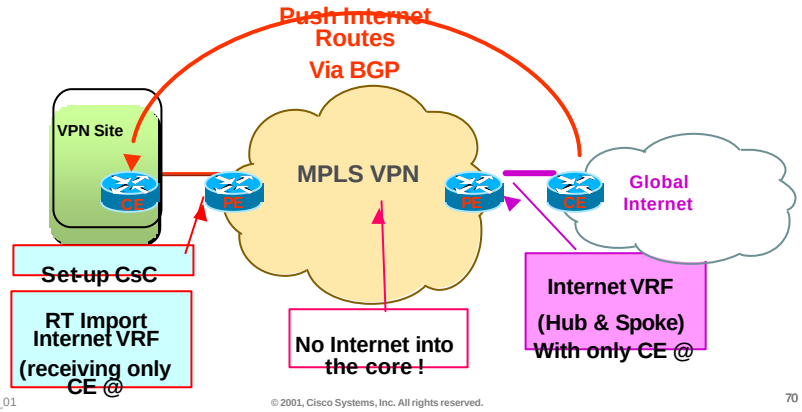


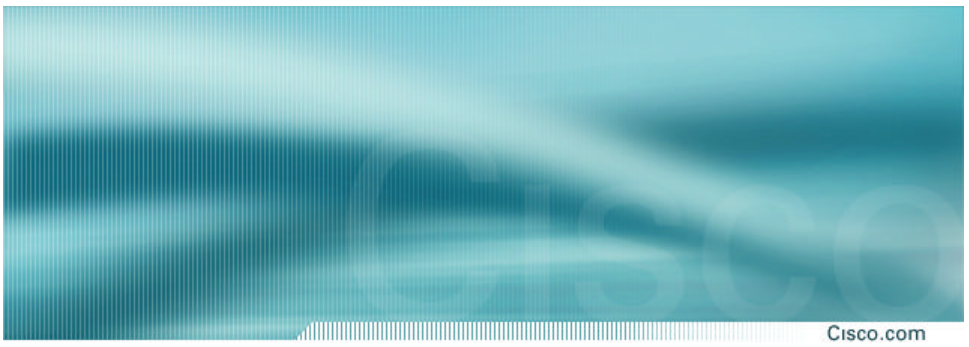
MPLS/VPN Internet Connectivity
Using CsC

Cisco.com

- **Achieved via a fly-over iBGP exchange between customer-CE and Internet-access-CE**

  Using a specific VRF for Internet-gateway access

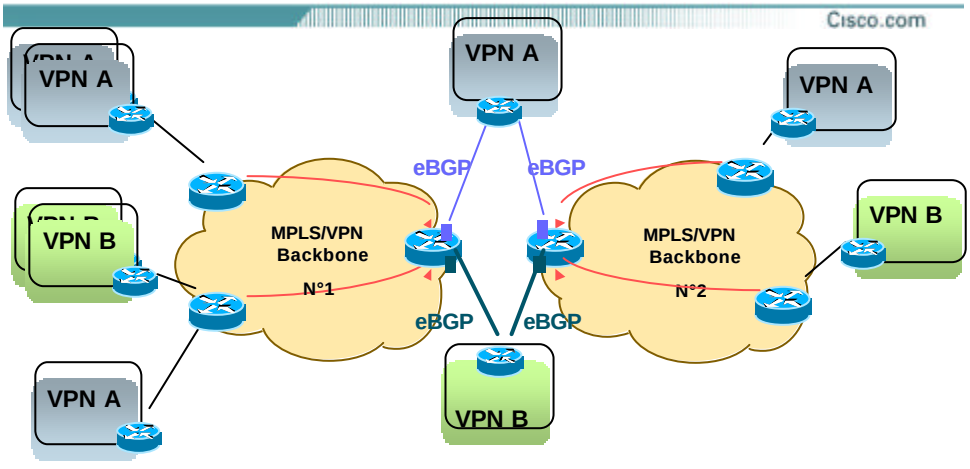




Advanced MPLS/VPN Topologies

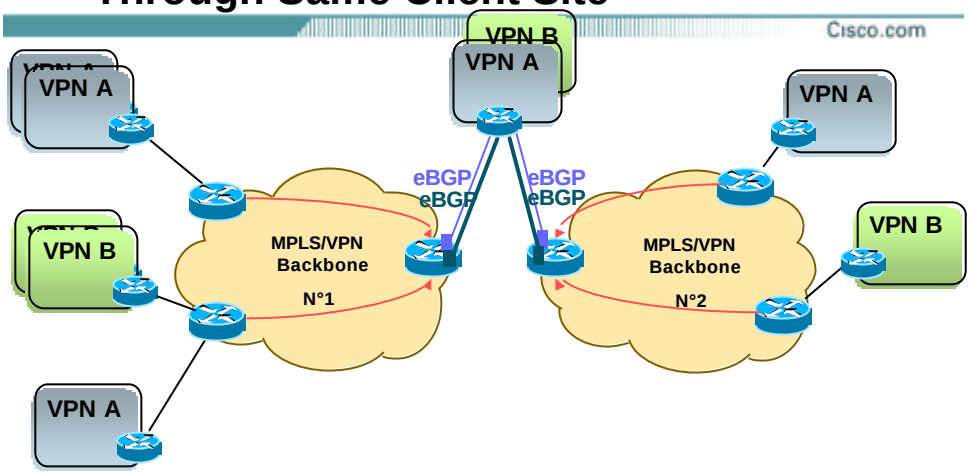
VPN Interconnection

MPLS/VPN Interconnection Through Different Client Sites



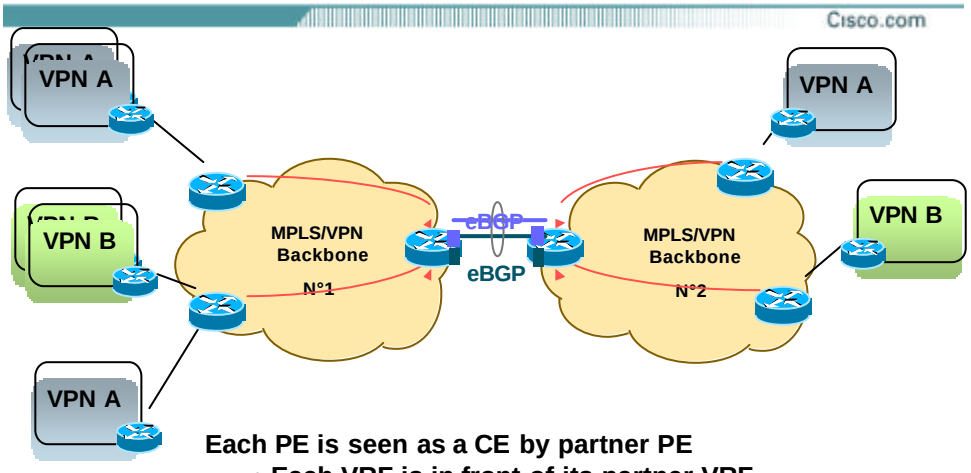
Each global VPN uses a different CPE site to interconnect VPNs

MPLS/VPN Interconnection Through Same Client Site



Each global VPN uses same CPE site to interconnect VPNs

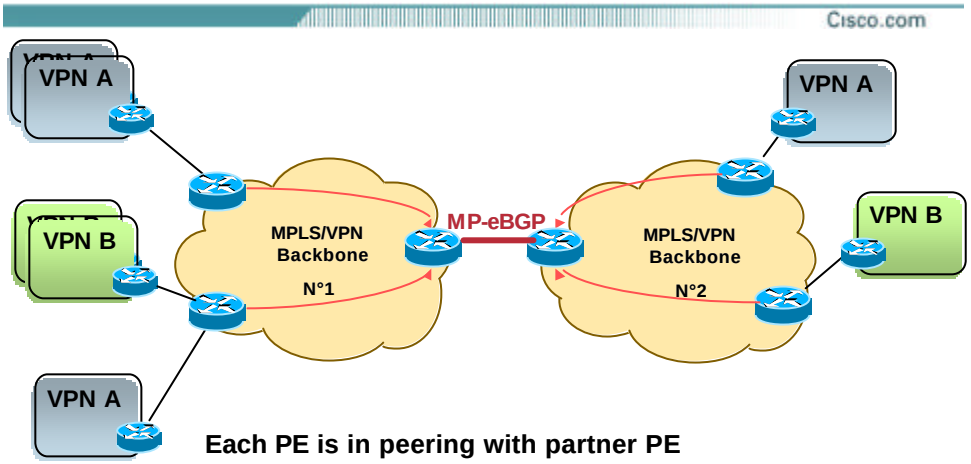
MPLS/VPN Interconnection Through Multi-Interfaces



Each PE is seen as a CE by partner PE

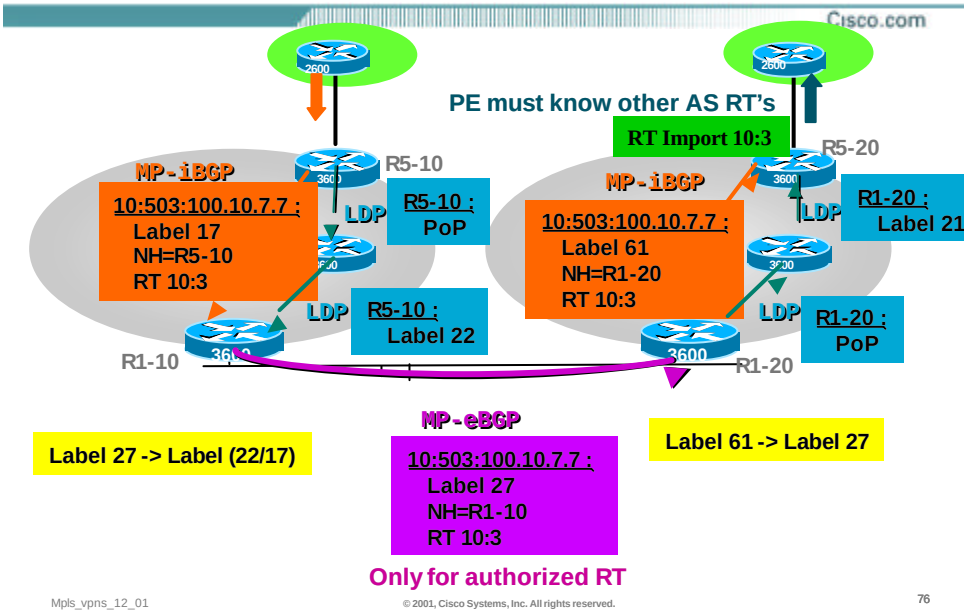
- **Each VRF is in front of its partner VRF**
- **One IPv4 eBGP session per VRF**
- **Use multi-VC or Multi-VLAN interface**

MPLS/VPN Interconnection Through Trunking

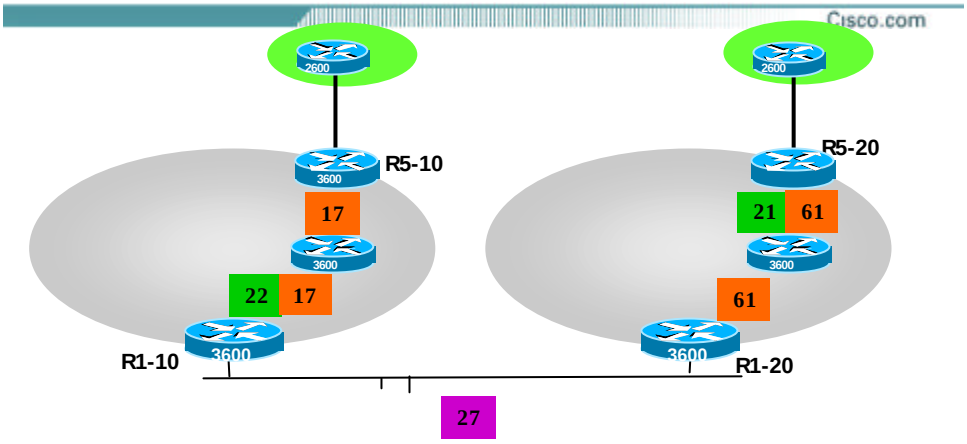


Each PE is in peering with partner PE
- RD used to unify addresses
- RT used to select services

Inter-Autonomous System MPLS VPN Interconnection – Control Plane



Inter-Autonomous System MPLS VPN Interconnection – Forwarding Plane



MP-eBGP for VPNv4

Cisco.com

- **Receiving Gateway PE-ASBRs may allocate new label if desired**

  Controlled by configuration of next-hop-self (default is on)

- **Receiving PE-ASBR will automatically create a /32 host route for its PE-ASBR neighbor**

  Which must be redistributed into receiving IGP if next-hop-self is NOT in operation

  /32 not created if iBGP session, eBGP multihop or if MP-eBGP exchange of VPNv4 capability not negotiated with neighbor

Multihop MP-eBGP for VPNv4

Cisco.com

- **MPLS VPN providers exchange VPNv4 prefixes via their Route Reflectors**

  Requires Multihop MP-eBGP (VPNv4 routes)

- **Next-hop-self MUST be disabled on Route Reflector**

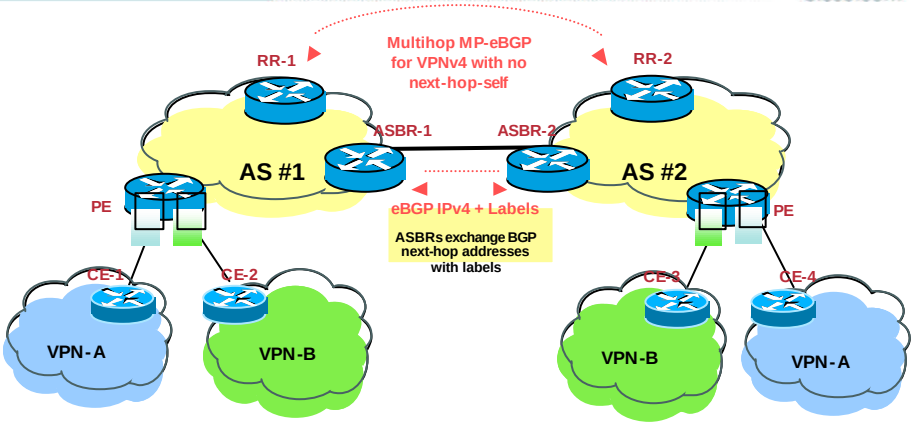
  Preserves next-hop and label as allocated by the originating PE router

- **Providers exchange IPv4 routes with labels between directly connected ASBRs using eBGP**

  Only PE loopback addresses exchanged as these are BGP next-hop addresses

Multihop MP-eBGP for VPNv4

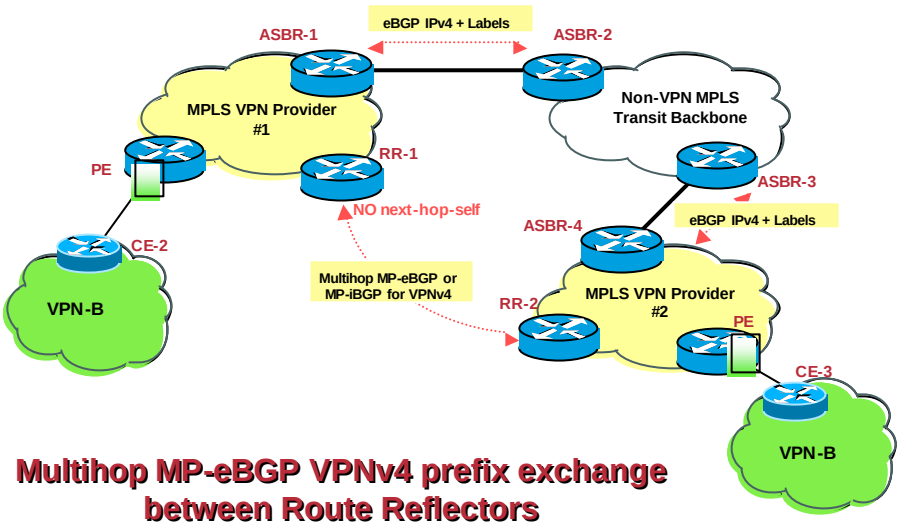
Cisco.com



Multihop MP-eBGP VPNv4 prefix exchange between Route Reflectors

Non-VPN Transit Provider

Cisco.com



Multihop MP-eBGP VPNv4 prefix exchange between Route Reflectors

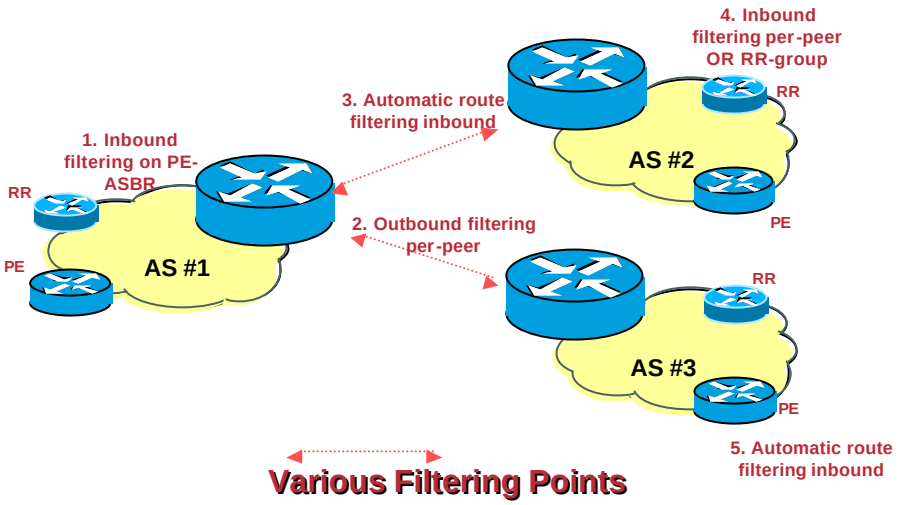
PE-ASBR Memory Scaling

Cisco.com

- **Potentially large amounts of VPN routing information**
  - That may or may not need to be carried between providers
  - Large percentage will be local VPN prefixes
- **PE-ASBRs must hold relevant VPN routing information**
  - But only Inter-AS VPN prefix details
- **Two methods available to aid scaling**
  - **ARF with local VRF import**
    - *If RT does not match locally configured import statement then drop the route*
  - **ARF disabled with inbound filtering**
    - *no default BGP route-target filter*
    - *Which implies filtering must occur to drop unwanted routes*

Inter-AS Filtering Points

Cisco.com



Load Balancing Between Backbones

Cisco.com

- **Balancing of Inter-AS traffic is an important issue**

  For distribution of traffic and redundancy of network design

- **All Inter-AS traffic must pass through PE-ASBRs**

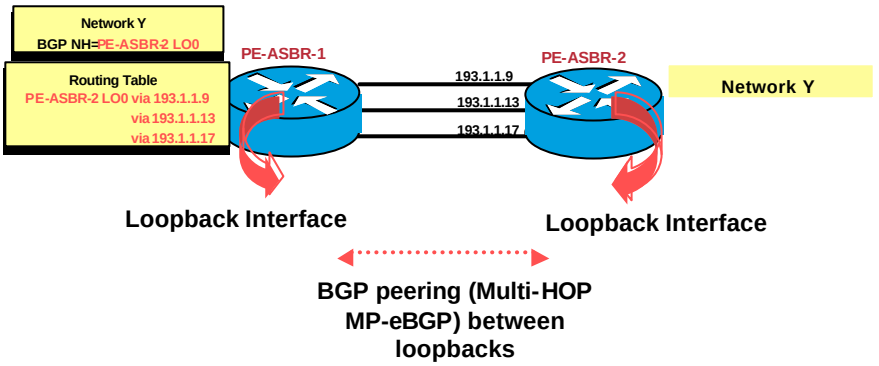
  As BGP next-hops are reachable via these routers

- **Multiple links provide traffic distribution**

  But do not provide redundancy due to single point of failure of the PE-ASBR

Load Balancing Between PE-ASBRs

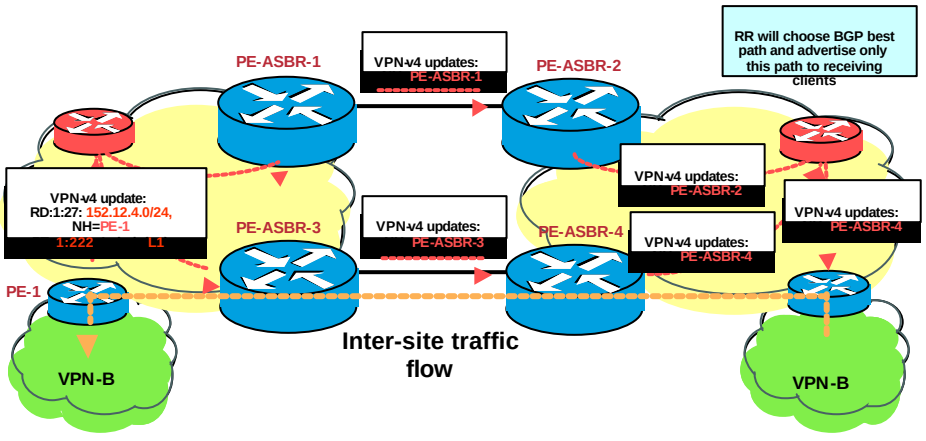
Cisco.com



Load Balancing across multiple PE-ASBR links

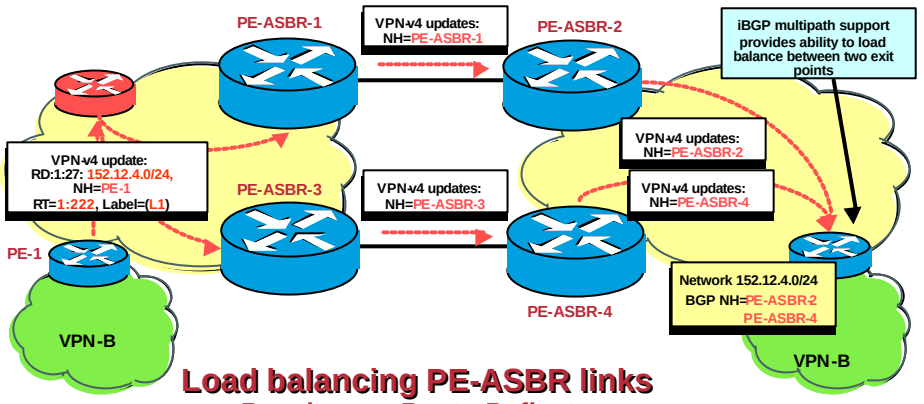
Redundant PE-ASBR Connections

Cisco.com



Redundant PE-ASBR used purely for backup

Redundant PE-ASBR Load Balancing

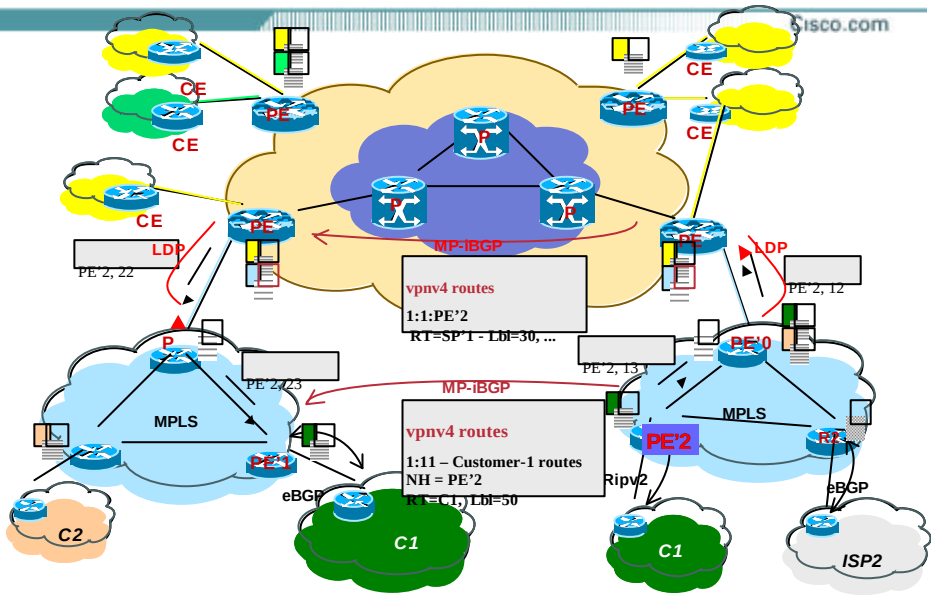


Load balancing PE-ASBR links

- *Requires no Route Reflectors*
- *Requires ASBRs to be located in same area*

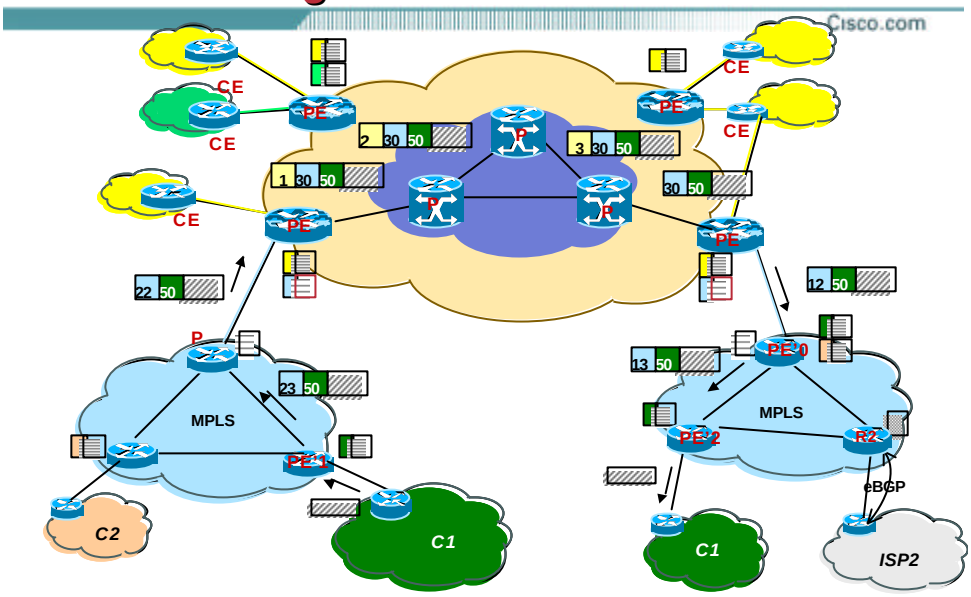
Carrier Supporting Carrier

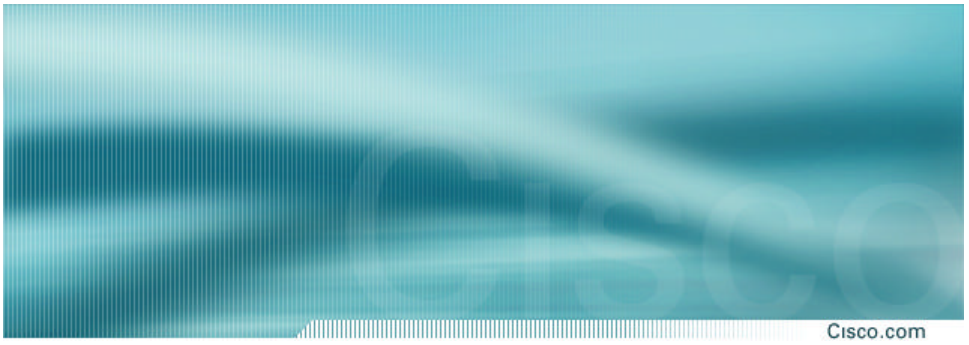
Control Plane



Carrier Supporting Carrier

Forwarding Plane

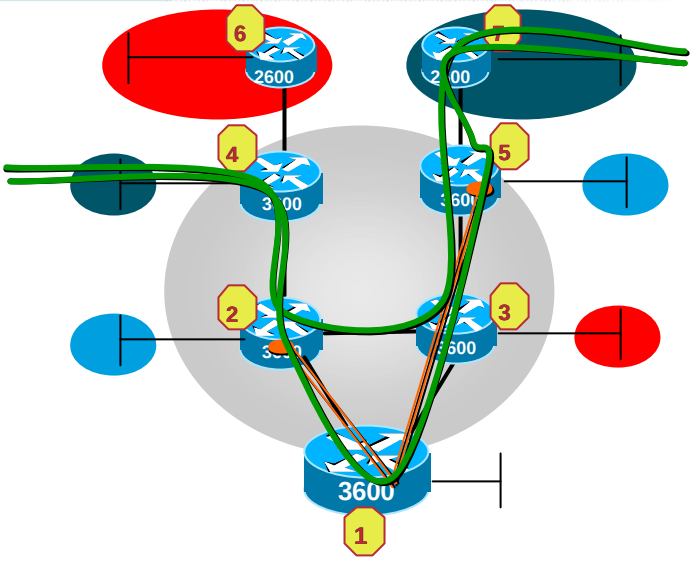




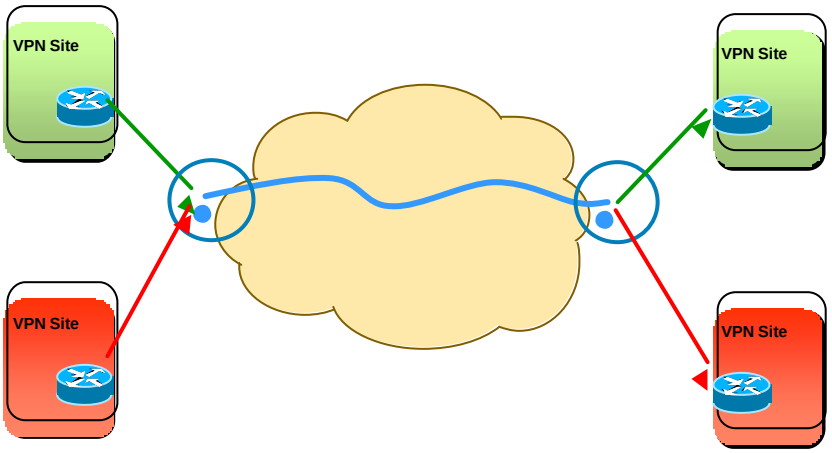
Advanced MPLS/VPN Topologies

MPLS VPN over MPLS TE

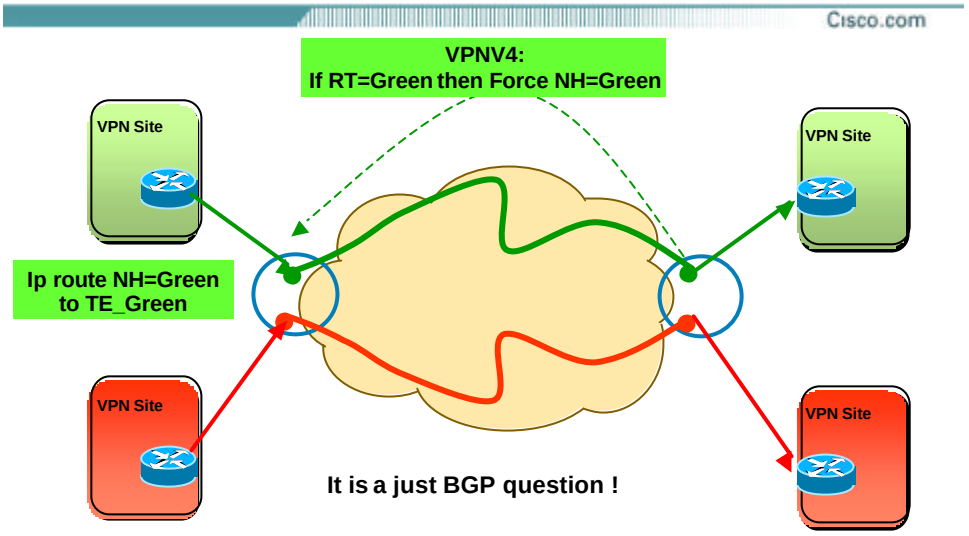
MPLS VPN over TE

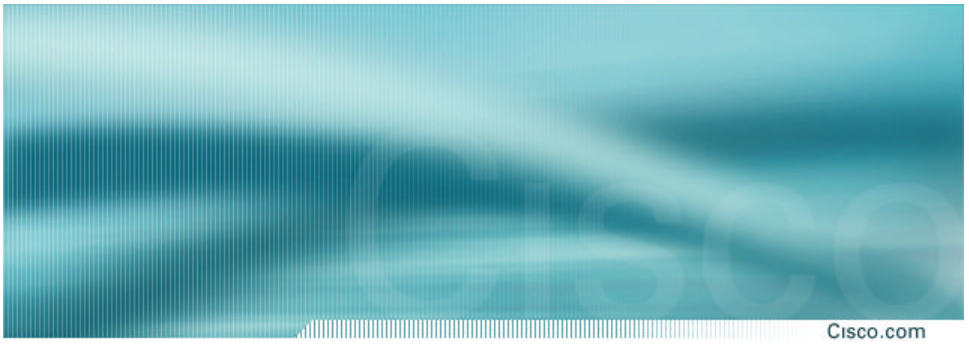


Per VPN TE



Per VPN TE





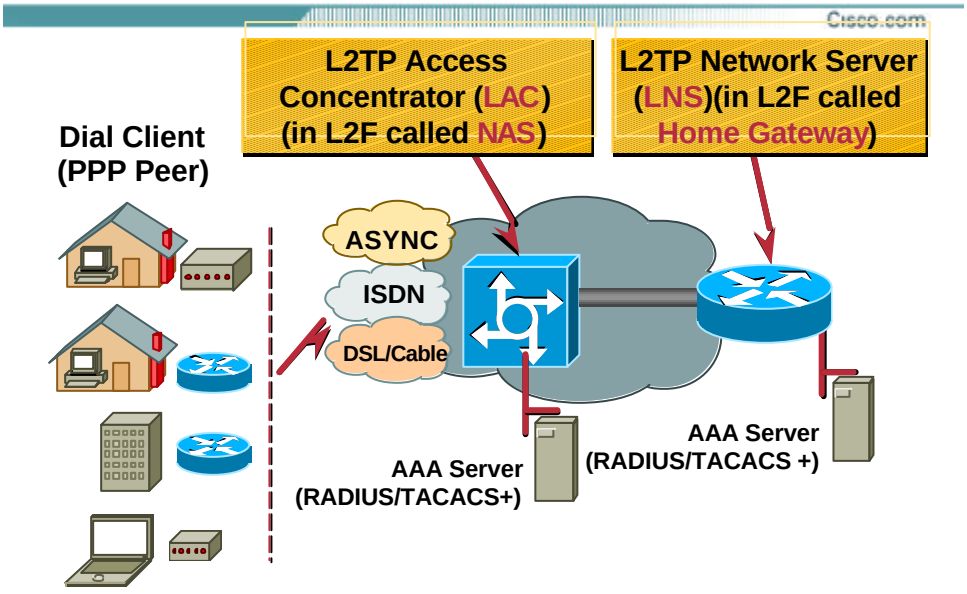
Dial MPLS VPNs

Access VPN Protocol History

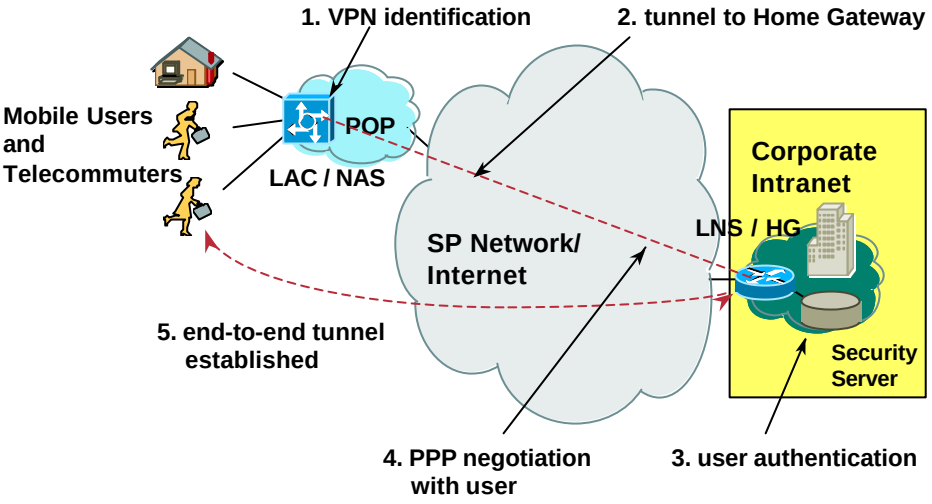
Cisco.com

- **PPTP** (Point-to-Point Tunneling protocol) Microsoft/Ascend/3COM *Proprietary*
- **L2F** (Layer 2 Forwarding) Cisco *Proprietary* (in Cisco IOS™ 11.2+)
- **L2TP** (Layer 2 Tunneling Protocol) IETF Draft combining the best of PPTP and L2F; Industry standard track

Access VPN Review

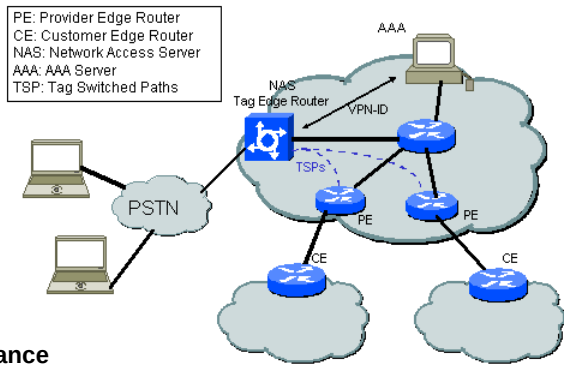


Access VPN Operation Review

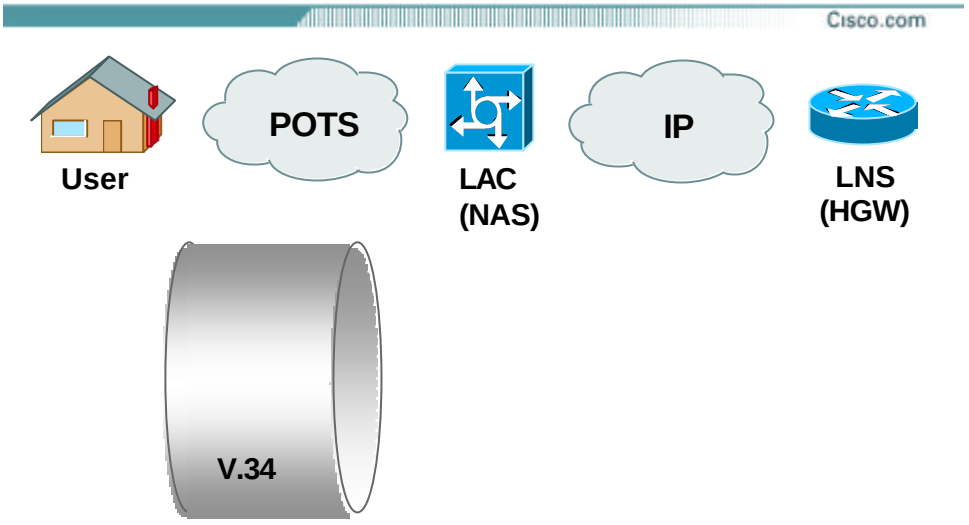


NAS-PE

- Authentication
- Security
- Authorization
  - CE-LNS
  - Remote Clients
- Routing
- Dial Backup
- Accounting
- Scalability & Performance
- Resiliency
- Conclusion

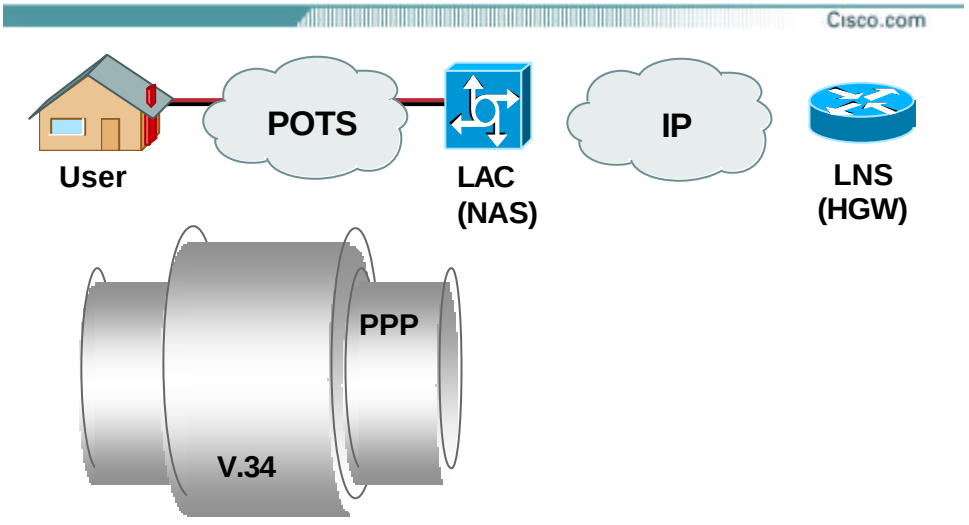


LAC-Initiated Tunneling "Call"



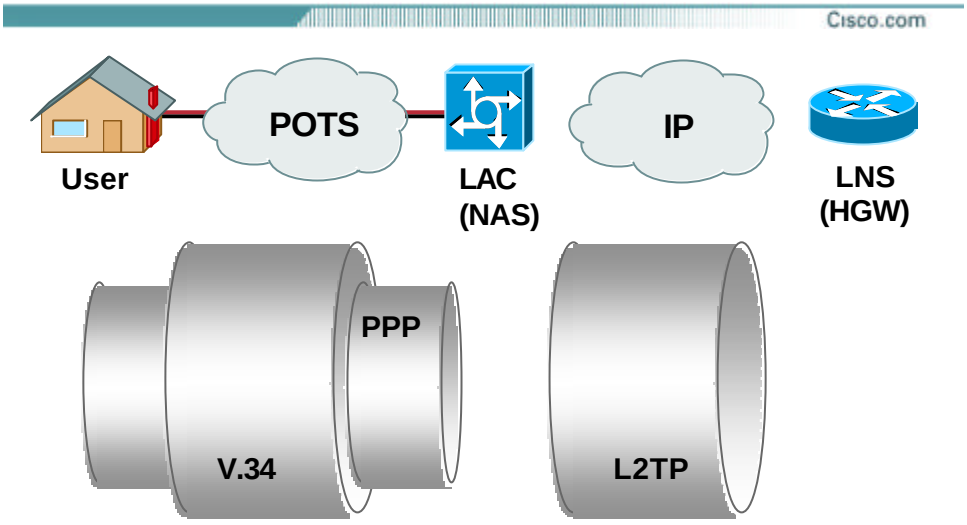
User starts things off by dialing into the LAC (NAS).

LAC-Initiated Tunneling "Connect"



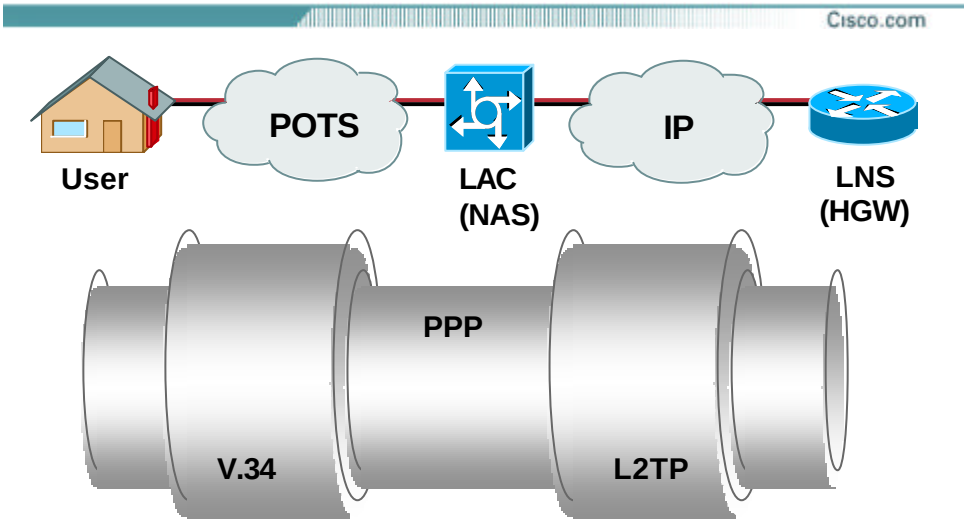
User's PC and the NAS negotiate PPP.

LAC-Initiated Tunneling “Tunnel”



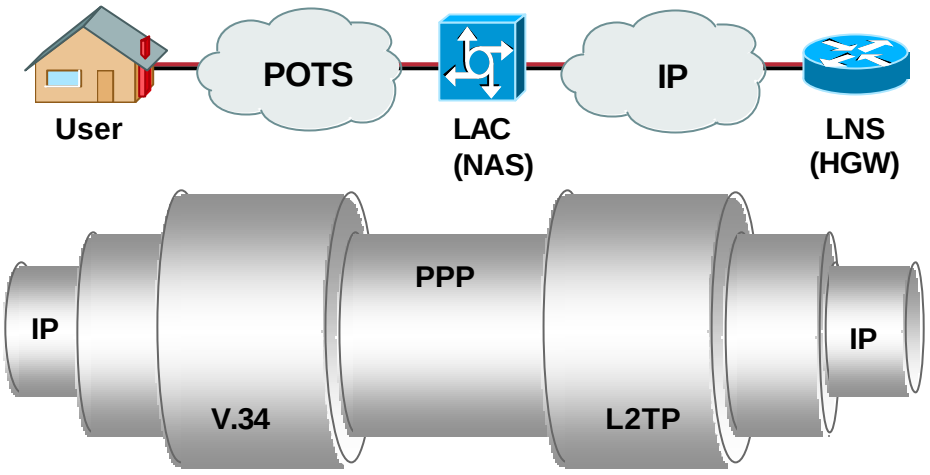
User must be forwarded, so a tunnel is created.

LAC-Initiated Tunneling “LCP”



The PPP connection is “stretched” to the home gateway.

LAC-Initiated Tunneling "NCP"



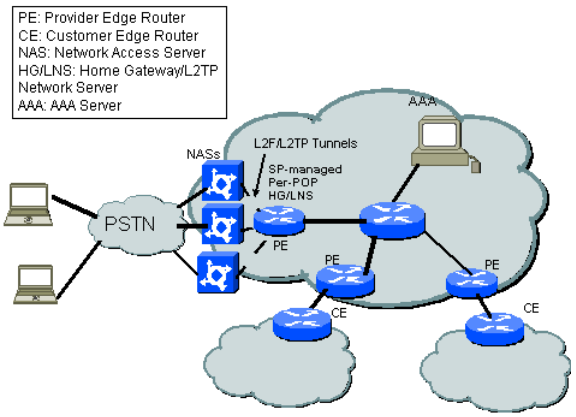
Finally, User's PC and Home Gateway negotiate IPCP.

PE-LNS

- **L2F or L2TP used as access method to the MPLS-VPN infrastructure**
- **DNIS or Domain Name used to identify VPN**
- **LAC configuration is unchanged**
- **LNS configuration need few changes**
  - **The Virtual-Access interface should be assigned to a specific VRF when cloned from a Virtual-Template.**

PE-LNS

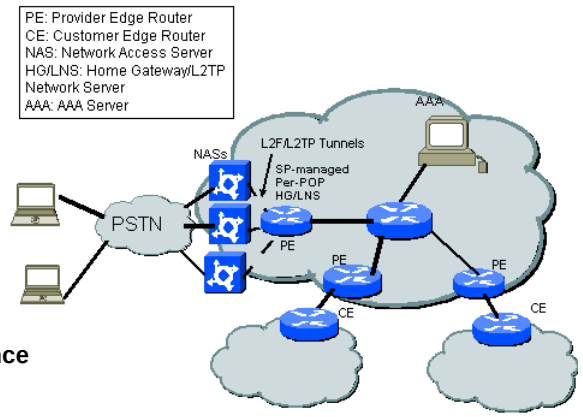
Cisco.com



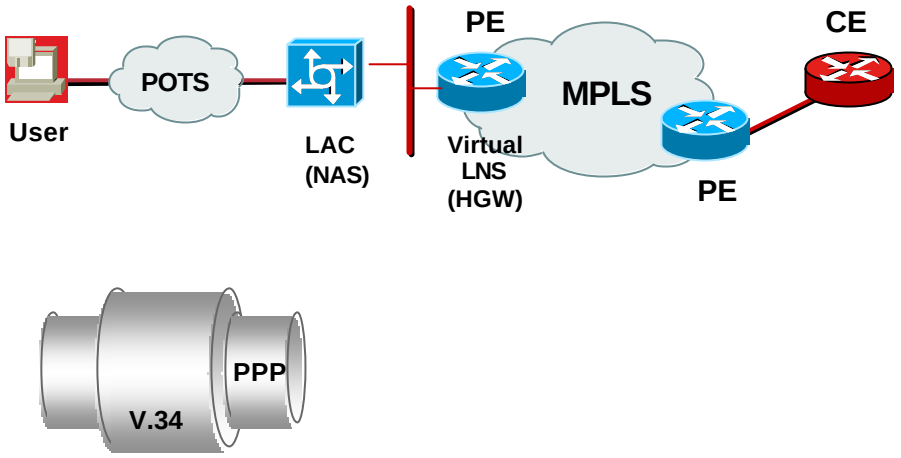
PE-LNS

Cisco.com

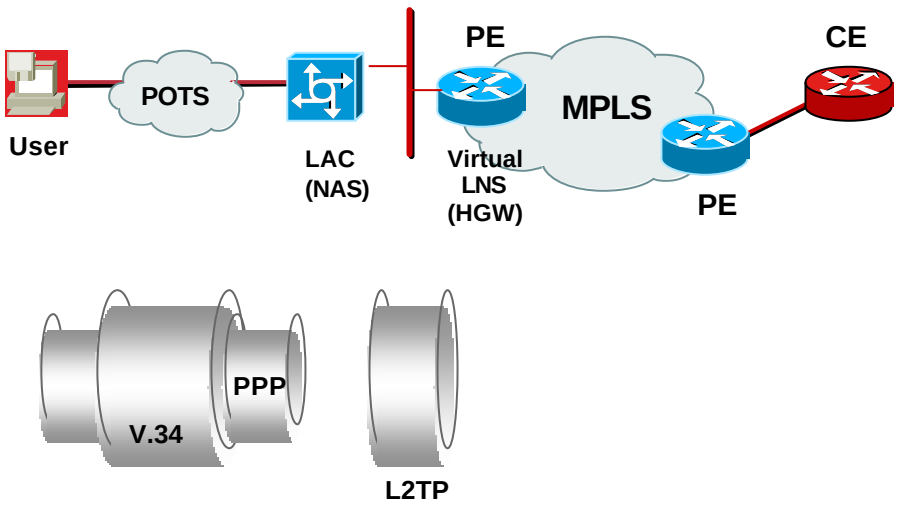
- Authentication
- Security
- Authorization
  - CE-LNS
  - Remote Clients
- Routing
- Dial Backup
- Accounting
- Scalability & Performance
- Resiliency
- Conclusion



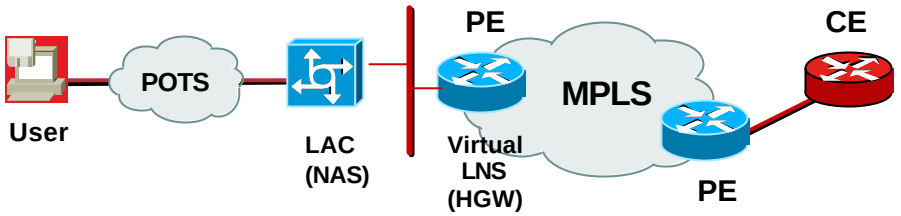
LAC-Initiated Tunneling “Connect”

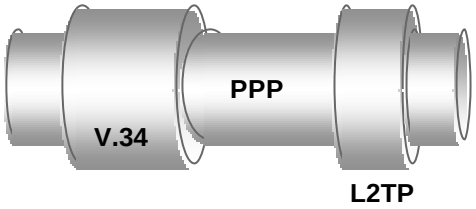


LAC-Initiated Tunneling “Tunnel”



LAC-Initiated Tunneling “NCP”

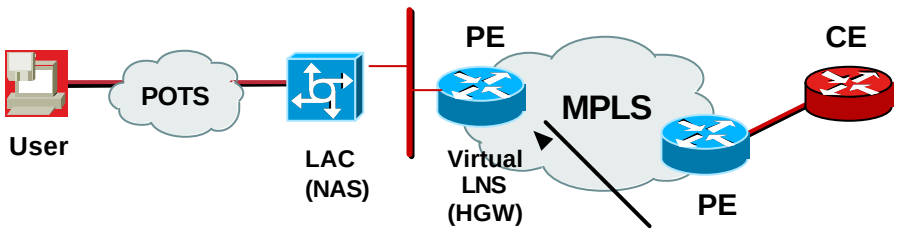


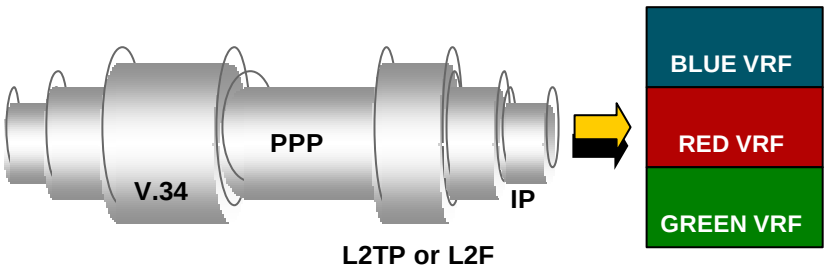


Virtual Template is used to clone a Virtual Access Interface

Since the VT is associated to a VRF AUTOMATICALLY the Virtual-Access will be associated to the same VRF

LAC-initiated Tunneling “VRF association”





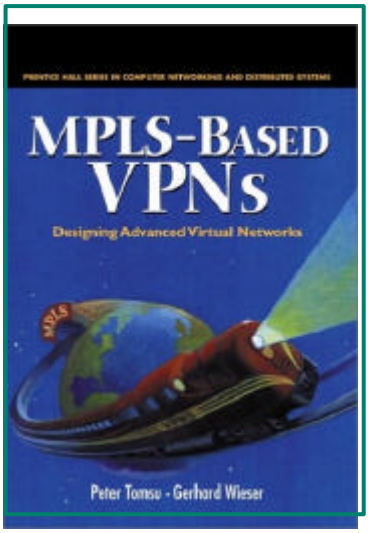
L2TP or L2F

Further Reading

MPLS-Based VPNs

Designing Advanced Virtual Networks

- **World wide available since End of October**
  - Amazon.com, Fatbrain.com, Prentice Hall
  - ISBN 0-13-028225-1
- **Covers**
  - MPLS Technology Basics
  - MPLS VPN Architecture
  - MPLS VPN Implementation
  - Advanced BGP Design Techniques
  - Application Scenarios

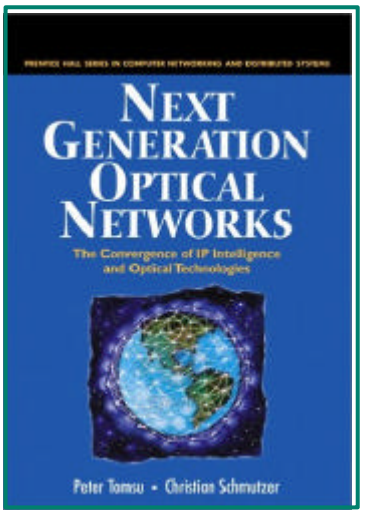


Next Generation Optical Networks

The Convergence of IP Intelligence and Optical Technologies

Cisco.com

- **World wide available since End of August**
  - Amazon.com, Fatbrain.com, Prentice Hall
  - ISBN 0-13-028226-x
- **Covers**
  - **Optical & Data Transmission Basics**
    - **Fibers, DWDM, POS, DPT, MPLS-TE**
  - **Optical Standardization**
    - **ITU, IEEE, OIF, IETF**
  - **IP+Optical Control Planes**
  - **Applications**



CISCO SYSTEMS

EMPOWERING THE INTERNET GENERATION℠