

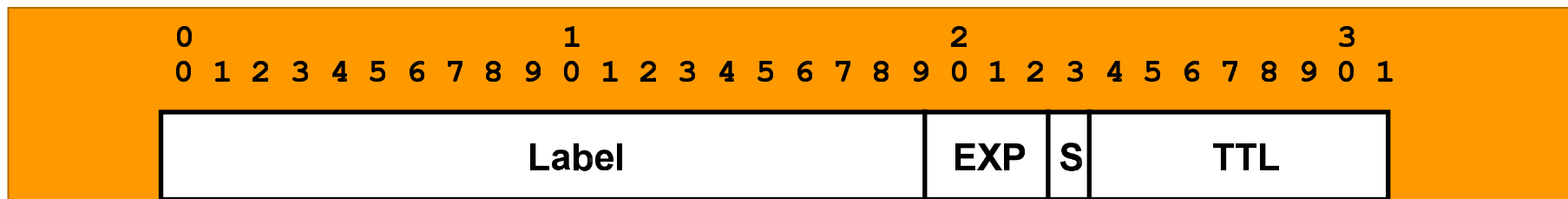
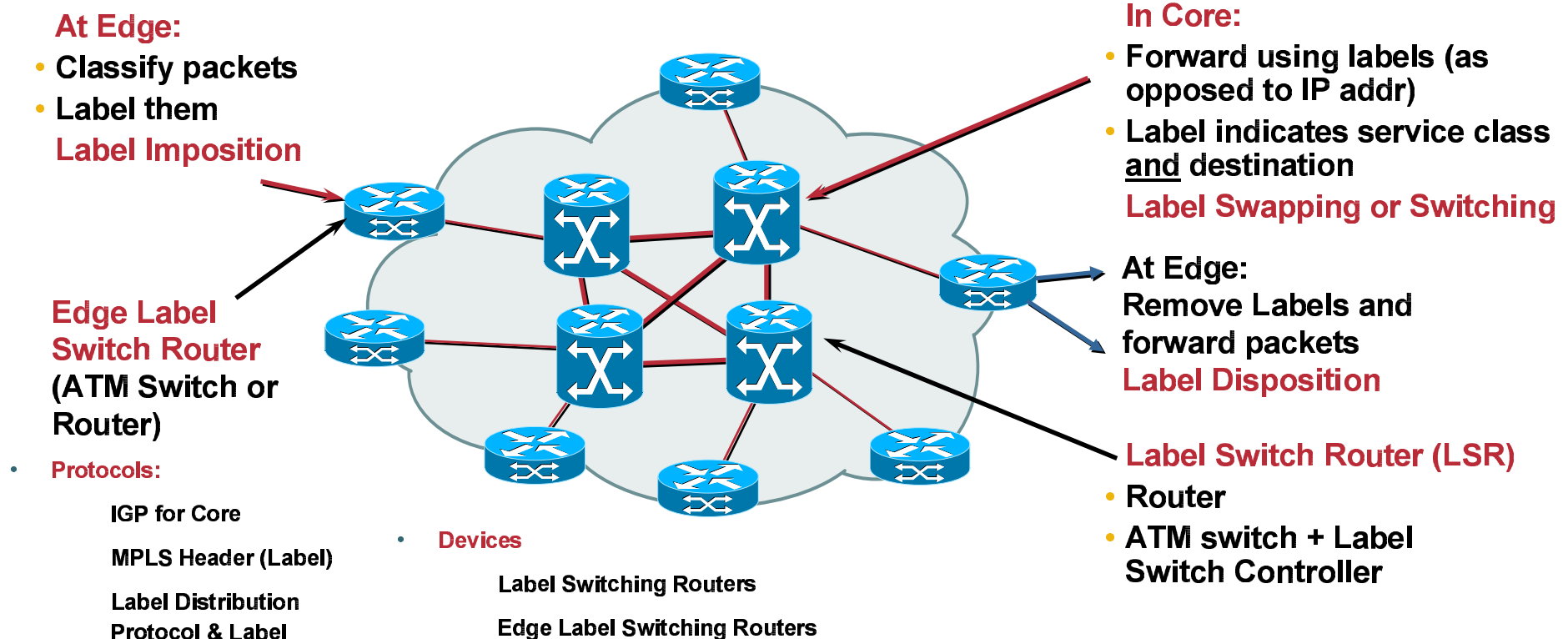


Building Enterprise MPLS Networks

Agenda

- **MPLS Basics**
- **MPLS L3 VPNs**
 - Architecture**
 - Security**
 - Services**
- **MPLS QoS**
- **Case Studies**

MPLS Concepts and Components



Label = 20 bits; COS/EXP = Class of Service, 3 bits; S = Bottom of Stack, 1 bit
 TTL = Time to Live, 8 bits

MPLS Operation

1a. Existing routing protocols (ie: OSPF, IS-IS, EIGRP) establish reachability to destination networks

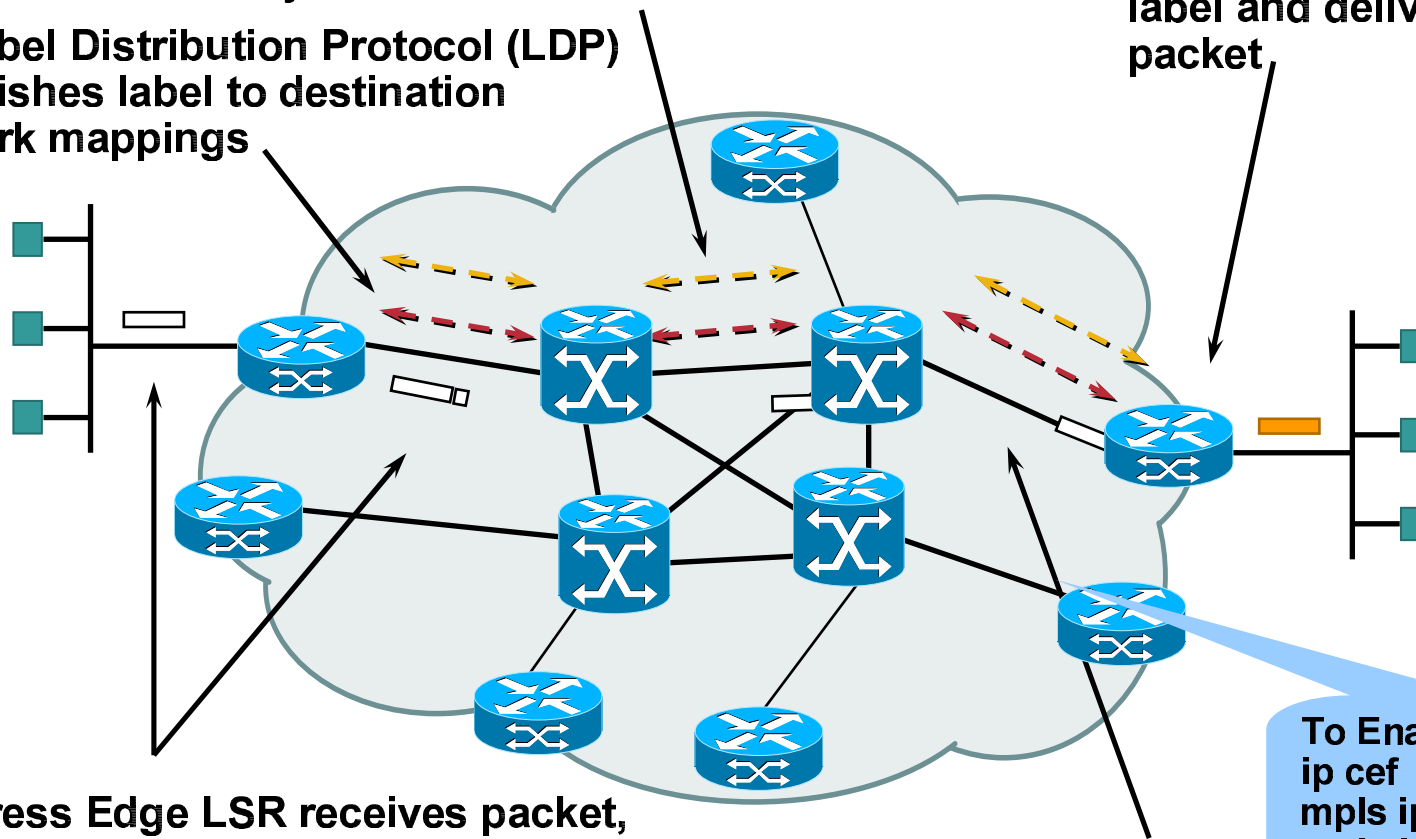
1b. Label Distribution Protocol (LDP) establishes label to destination network mappings

4. Edge LSR at egress removes label and delivers packet

2. Ingress Edge LSR receives packet, performs Layer 3 value-added services, and "labels" packets

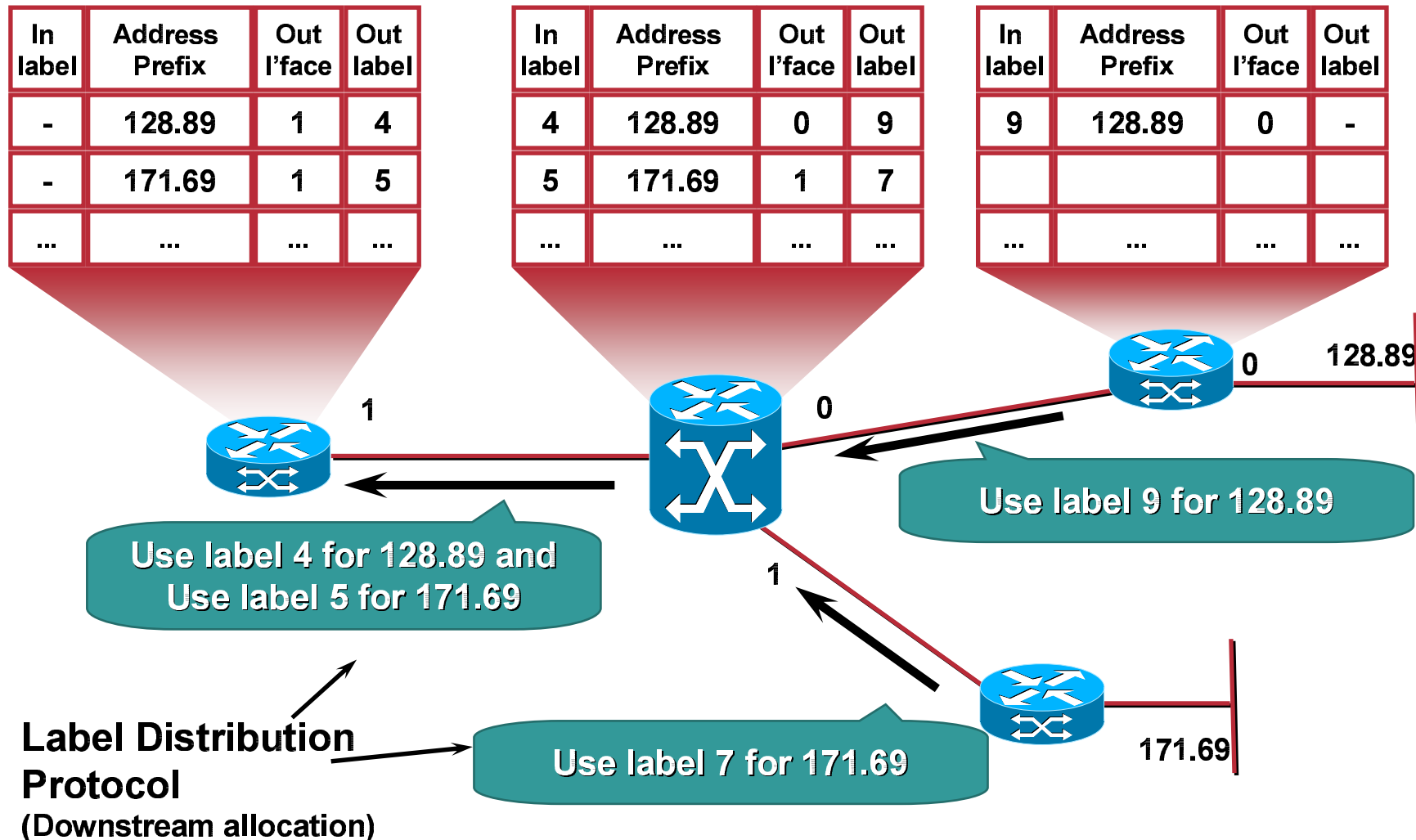
3. LSR switches packets using label swapping

To Enable mpls:
ip cef
mpls ip
mpls label protocol ldp



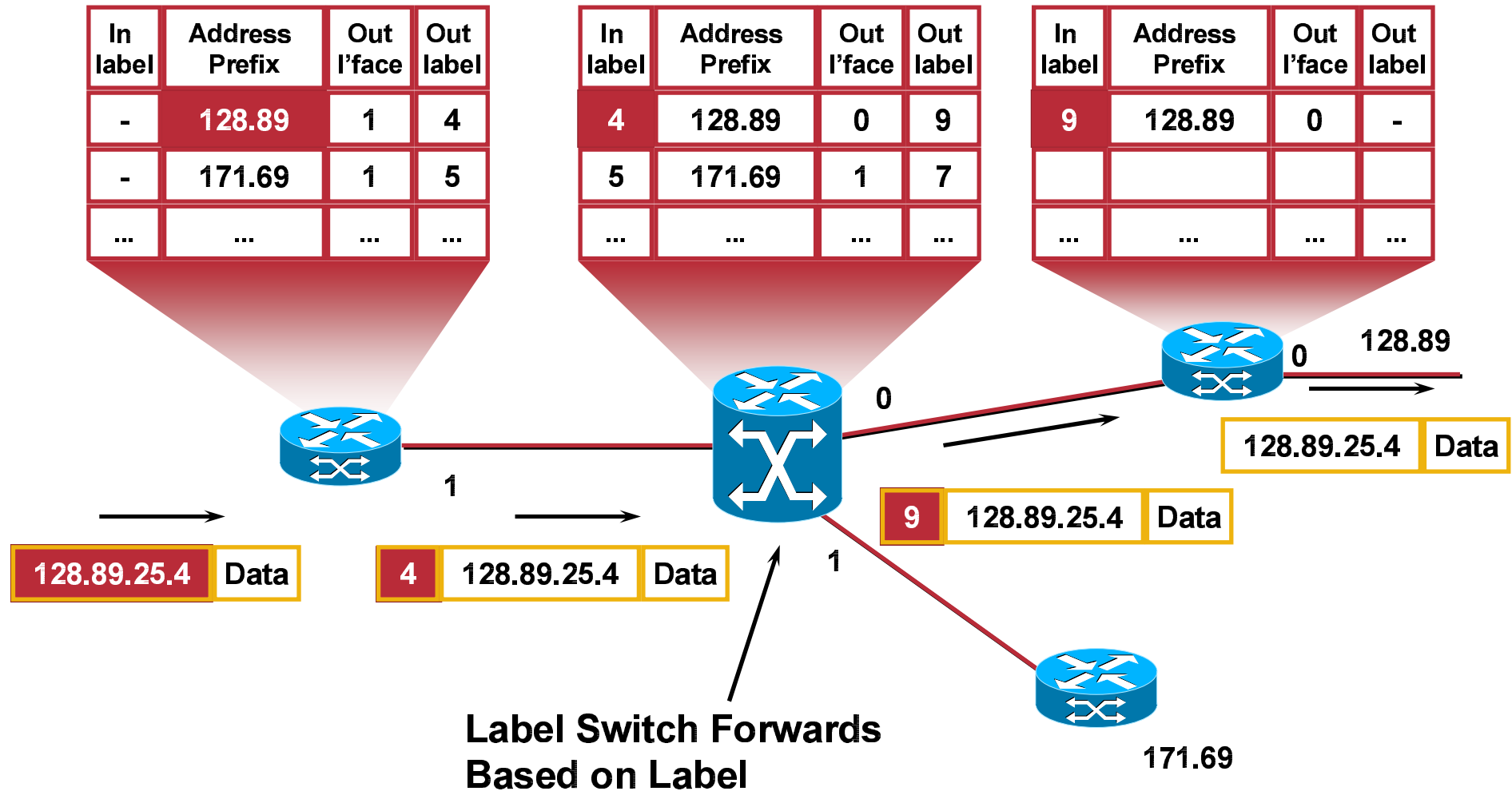
MPLS Control Plane

Assigning and Distributing Labels



MPLS Forwarding Plane

Appending Labels and Forwarding Packets



Agenda

- **MPLS Basics**
- **MPLS L3 VPNs**
- **MPLS QoS**
- **Case Studies**

MPLS L3 VPN Components

- **Control plane components**

Virtual Routing Forwarding (VRF) name

Route Target (RT)

Route Distinguisher (RD)

VRF table

MP-iBGP

IGP Core

IGP Edge

eBGP Edge

MPLS L3 VPN Components (Cont.)

- **Forwarding plane components**

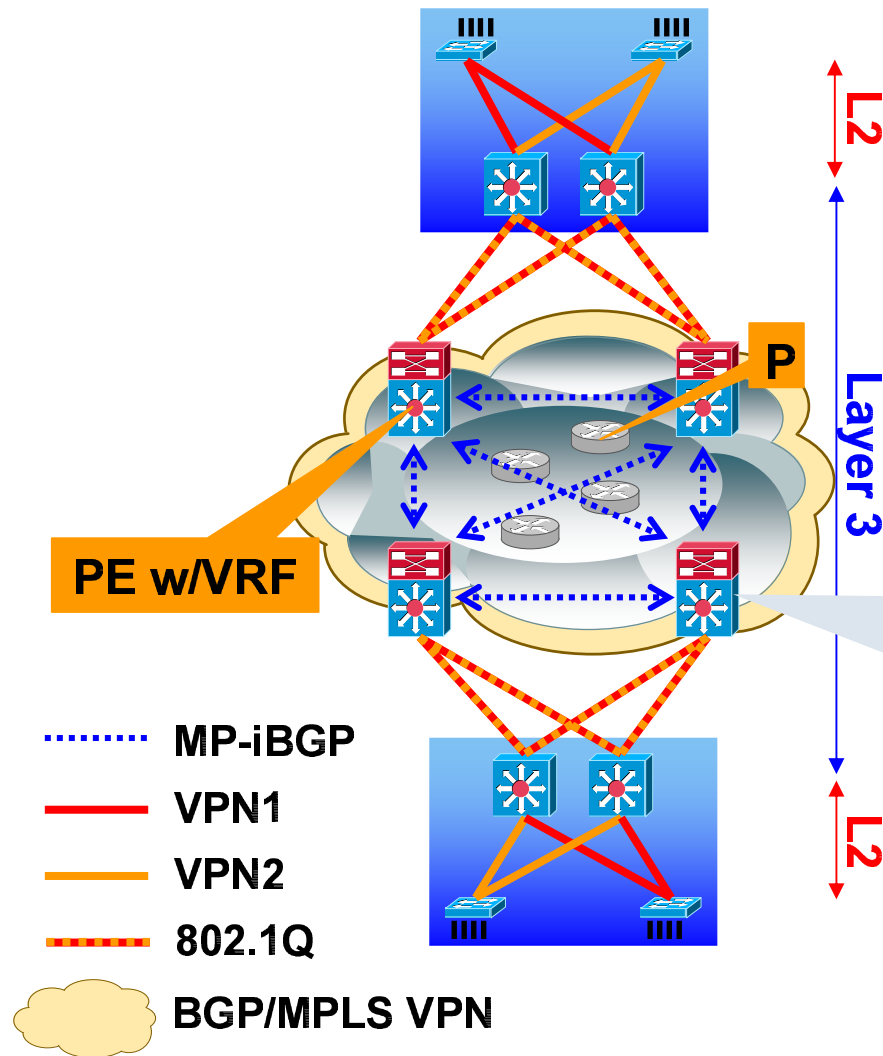
VRF name, VPN label

Routing Information Base (RIB)

Forwarding Information Base (FIB)

Label Forwarding Information Base (LFIB)

Building a VRF

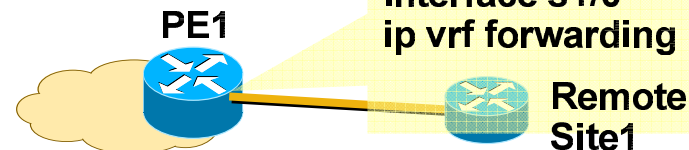


- Define a VRF
- Bind it to an interface for adjacent VLAN subnet or a VPN site

On a L3 device: router or a switch

```
!
ip vrf <Finance>
!
ip vrf <Aero>
!
Interface vlan 100
ip vrf forwarding <Finance>
!
Interface vlan 200
ip vrf forwarding <Aero>
```

```
!
ip vrf Aero
!
Interface s1/0
ip vrf forwarding <Aero>
```



Route Distinguisher

- Purely to make a route unique so customers don't see each other's routes

Example: differentiate 10.0.0.0/8 in VPN-A from 10.0.0.0/8 in VPN-B

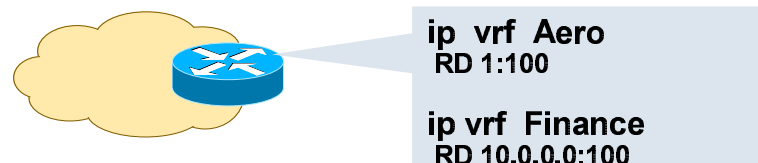
Makes IPv4 route a VPNv4 routes: VPNv4=RD:IPv4

Unique route is now **RD:IPaddr** (96 bits) plus a mask on the IPAddr portion

So route reflectors make a bestpath decision on something other than 32-bit network + 32-bit mask

- 64-bit quantity configured as **ASN:YY** or **IPADDR:YY**

Almost everyone uses Autonomous System Number (ASN)



Route Target

- **To control policy about who sees what routes**
- **Each VRF ‘imports’ and ‘exports’ one or more RTs**
 - Exported RTs are carried in VPNv4 BGP**
 - Imported RTs are local to the box**
- **A PE that imports an RT installs that route in its associated VRF table**
- **64-bit quantity (2 bytes type, 6 bytes value) carried as an extended community and typically written as written as ASN:YY**

Route Target (Cont.)

- **For deployment model:**

Full mesh:

All sites import X:Y and export X:Y

Hub-and-spoke:

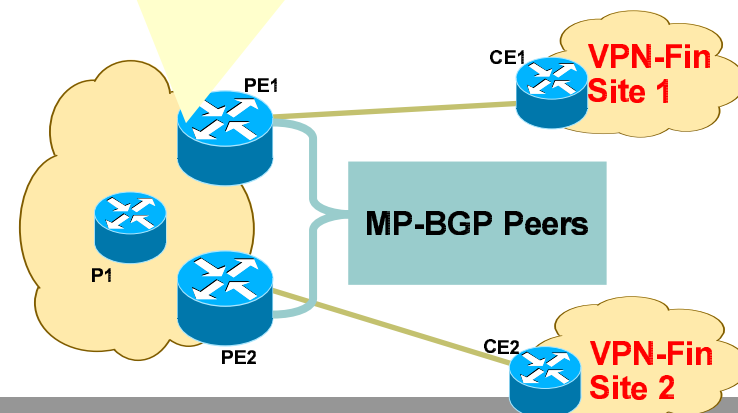
Hub exports X:H and imports X:S

Spokes export X:S and import X:H

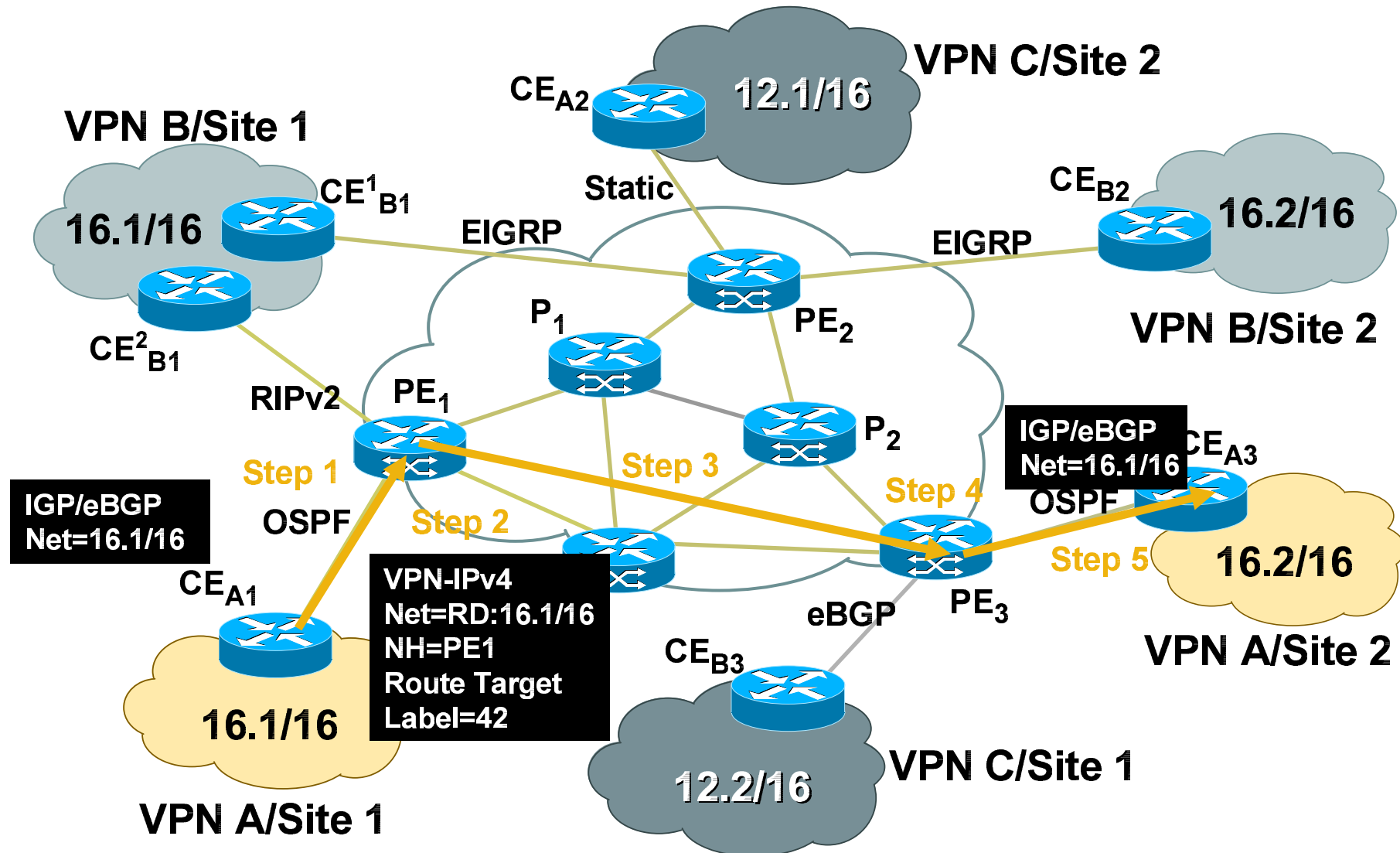
MP-BGP and VPNv4

- This is enhanced IPv4 BGP with additional communities; all the rules of IPv4 BGP applies
- MP-BGP session facilitates the advertisement of VPNv4* prefixes + labels between MP-BGP peers
- MP-BGP only on PE routers, not on the core routers
- All edge routers do need to be fully meshed; can use VPNv4 Route Reflectors for scalability
- At the advertising PE, BGP allocates labels for VPN prefixes and installs them in the LFIB (MPLS forwarding table)
- At the receiving PE, IF BGP accepts VPN prefixes with labels, THEN BGP installs VPN prefixes in the VRF FIB (CEF table)
- VPNv4 announcement carries a label with the route

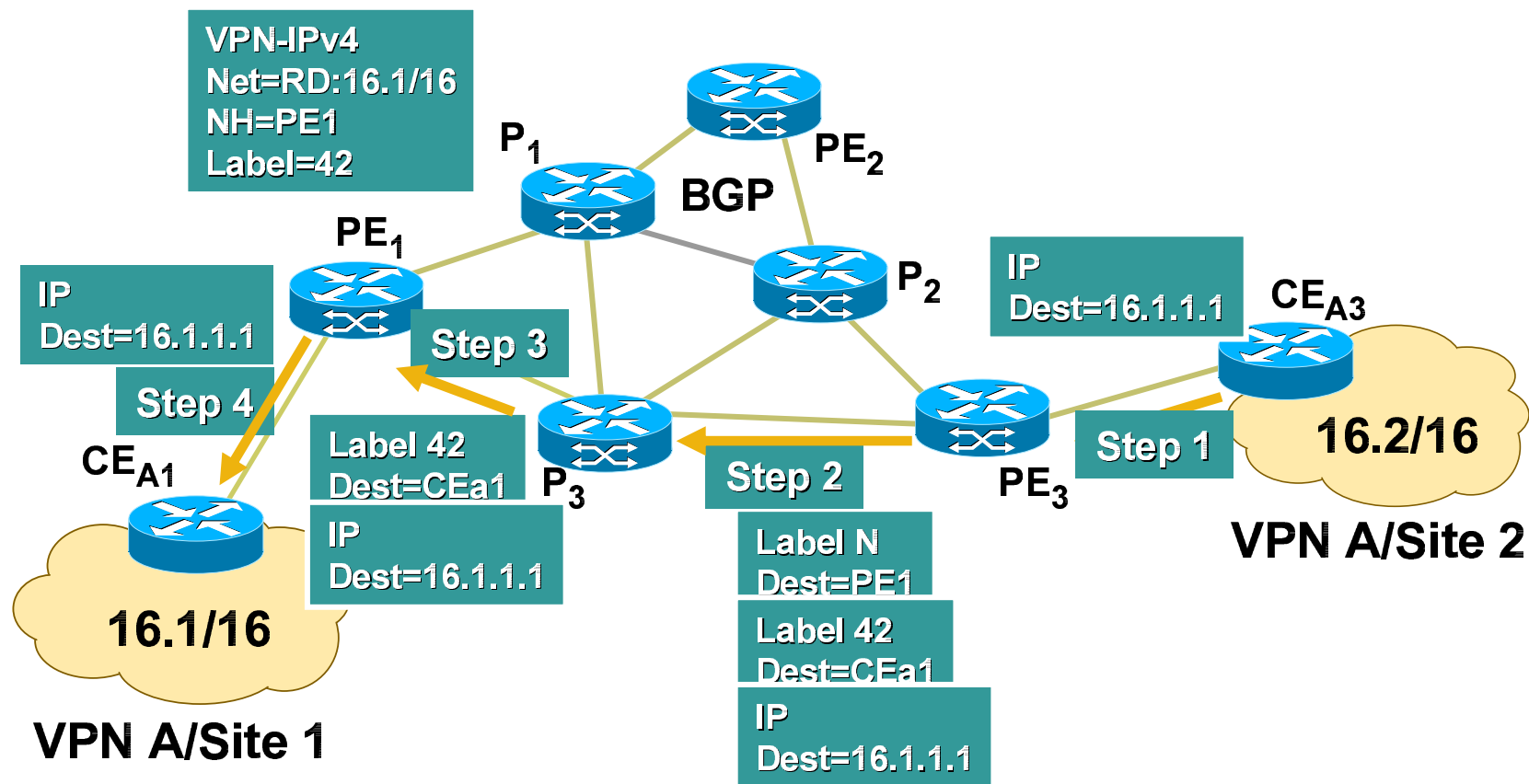
```
router eigrp 1
address-family ipv4 vrf VPN-Fin
  autonomous-system 1
  redistribute BGP 100 metric 10000 100 255 1 1500
  no auto-summary
  no eigrp log-neighbor-changes
!
router bgp 100
  neighbor PE2 remote-as 1
  neighbor PE2 update-source loop0
!
address-family vpnv4
  neighbor PE activate
  neighbor PE send-community extended
!
address-family ipv4 vrf VPN-Fin
  redistribute eigrp 1
  no auto-summary
  no synchronization
```



MPLS L3 VPN Control Plane Separates VPN control Plane Traffic

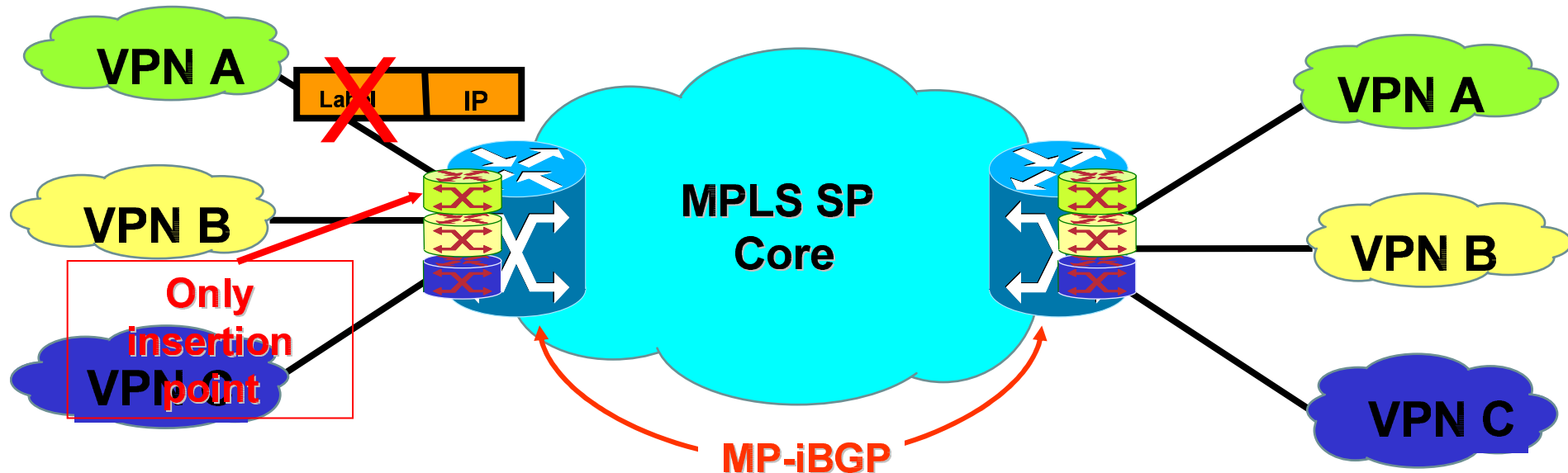


MPLS L3 VPN Forwarding Plane Separates VPN Forwarding Plane traffic



Label N is IGP label to switch traffic through the core from PE₃ to PE₁
 Label 42 is a VPN A Label for CE_{A1}-16.1.xx prefix.

How Secure are MPLS VPNs?



- Built in security in control and forwarding plane for not accepting traffic from unauthorized source for MPLS, still inherit security gaps from IP network
- Where can you attack?
Address and Routing Separation, thus:
Only Attack point: peering PE
- How?
 - Intrusions
(telnet, SNMP, ..., routing protocol)
 - DoS

Security Recommendations

- **Protect VPNs from it's own users**
 - Protect CE using privilege levels, acls, enable pws...etc.**
 - Use AAA to authenticate users**
 - ACLs on CE to protect PE**
- **Protect routing protocols from CE to PE**
 - Use static when possible**
 - Using ACL (source is only CE)**
 - MD5 authentication with LDP and MP-BGP and other CE-PE routing protocols**
 - BGP dampening, filtering, maximum-prefix**

Security Recommendations (Cont.)

- **Protect PE resources**

Number of routes limitations per VRF

No telnet access to vpn sites, privilege levels, acls, enable pws...etc.

Class Based Policing for rate limiting to control traffic (specially UDP)

Security Recommendations (Cont.)

- **Dedicate device/link per service**
 - Separate Links for Internet & VPN traffic**
 - Separate PEs for Internet & VPN Service**
 - Separate Internet Service from the MPLS VPN cloud**
- **Use encryption where required**
 - overlay IPsec VPNs**
 - VRF aware IPsec**
 - DGVPNs (in EFT)**

Security Recommendations (Cont.)

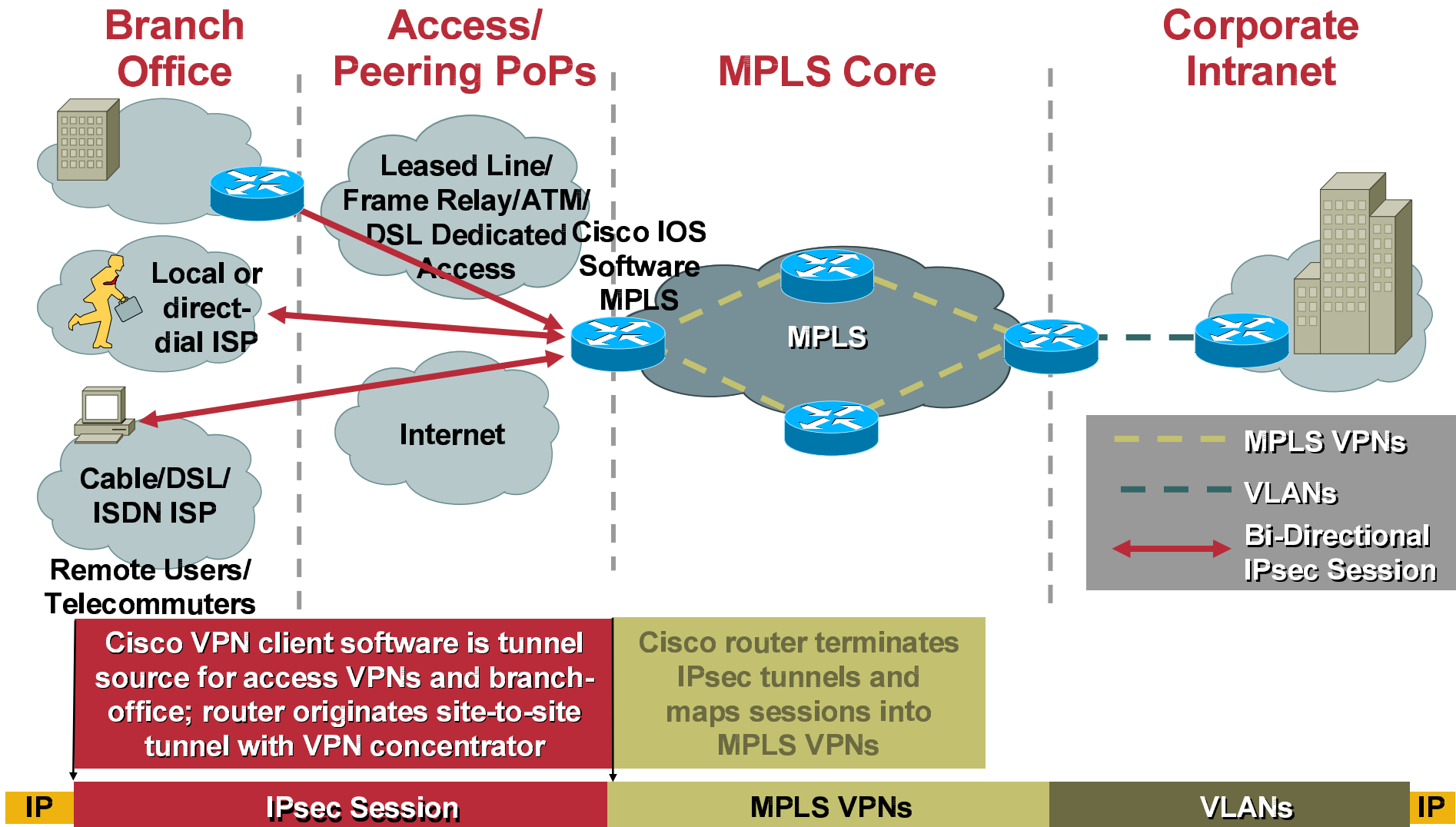
- **Use per VPN firewall**

Cisco 7600 with FWSM module

Others Cisco IOS Software based VRF aware FW

Distributed or centralized custom Firewall

VRF aware IPsec



CE-CE Encryption Solution in progress

Massively-scalable IPsec VPN IETF proposals

- **IPsec Group Keyed VPN (GKVPN)**

Group Key Management: GDOI (RFC 3547)

Data Encryption: IPsec (RFC 2401, etc.)

For IP or MPLS network

Will not completely replace DMVPNs

For offnet sites, DMVPNs or pt.pt. IPsec will still be desirable

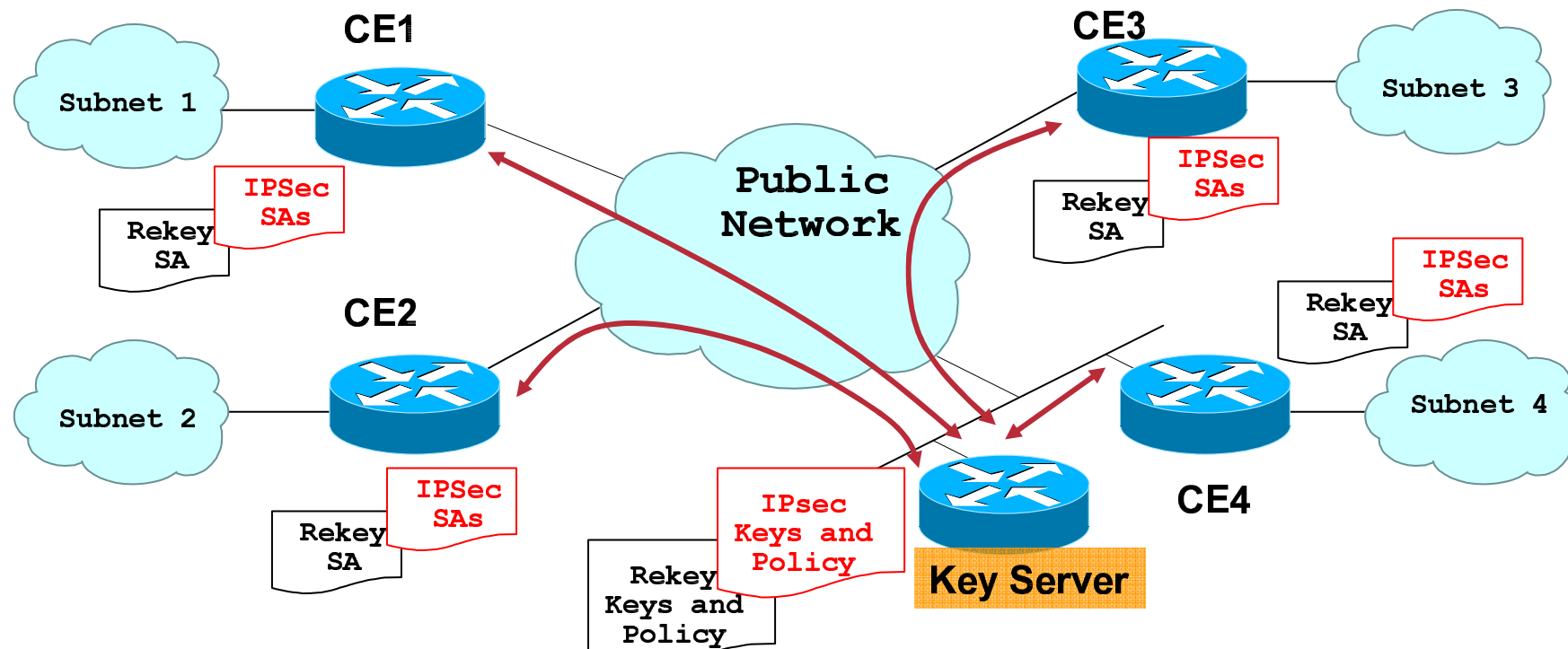
CE-CE Encryption Solution in progress

Massively-scalable IPsec VPN IETF proposals (Cont.)

- **Enhance GDOI to support Unicast traffic (static, BGP..routing..etc.)**
- **Leverage Any-to-Any characteristics of MPLS/VPN partition**
- **Eliminate Tunnel Overlay and IGP Overlay**
- **Mitigate Point-to-Point nature of IPsec**
- **Leverage Multicast replication of L3 VPN network**
- **Simplified VPN Key Management**
- **Convergence based on Routing, Not IPsec**
- **Communities of interest**

Group Domain of Interpretation (GDOI)

- Each router Registers with the Key Server; the Key Server authenticates the router, performs an authorization check, and downloads the policy and keys to the router.

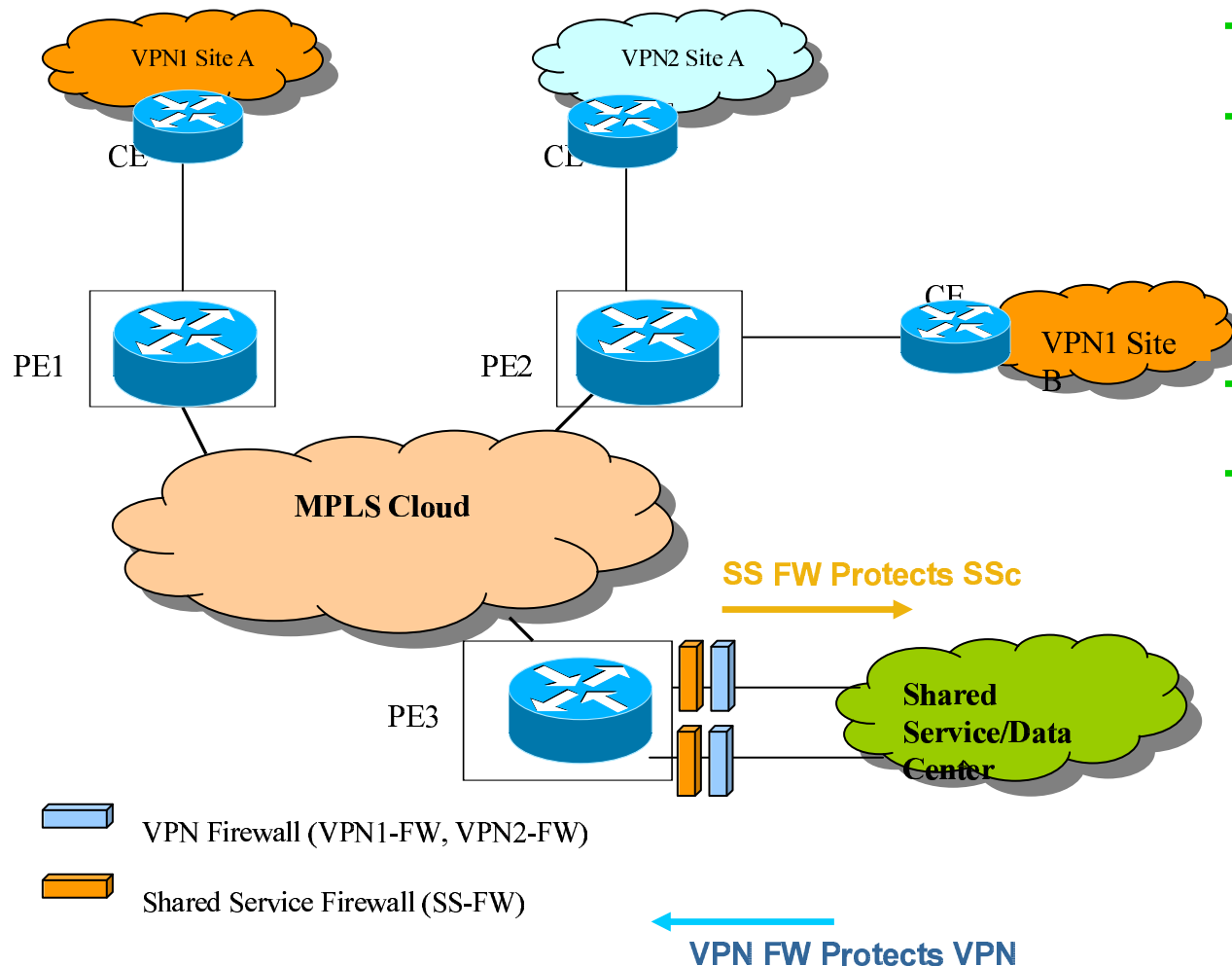


Each VPN CPE

- Registers to the GDOI key server to “receive” IPsec SAs
- Encrypts/Decrypts packets matching the SA
- Receives rekey messages, either to refresh the IPsec SAs or eject a group member

Firewall

VRF aware Cisco IOS or FWSM, PIX?



- + Central management
- + Different policies per subsidiary, VLANs, or organizations is possible
- + CEs not touched
- + One virtual firewall per VPN

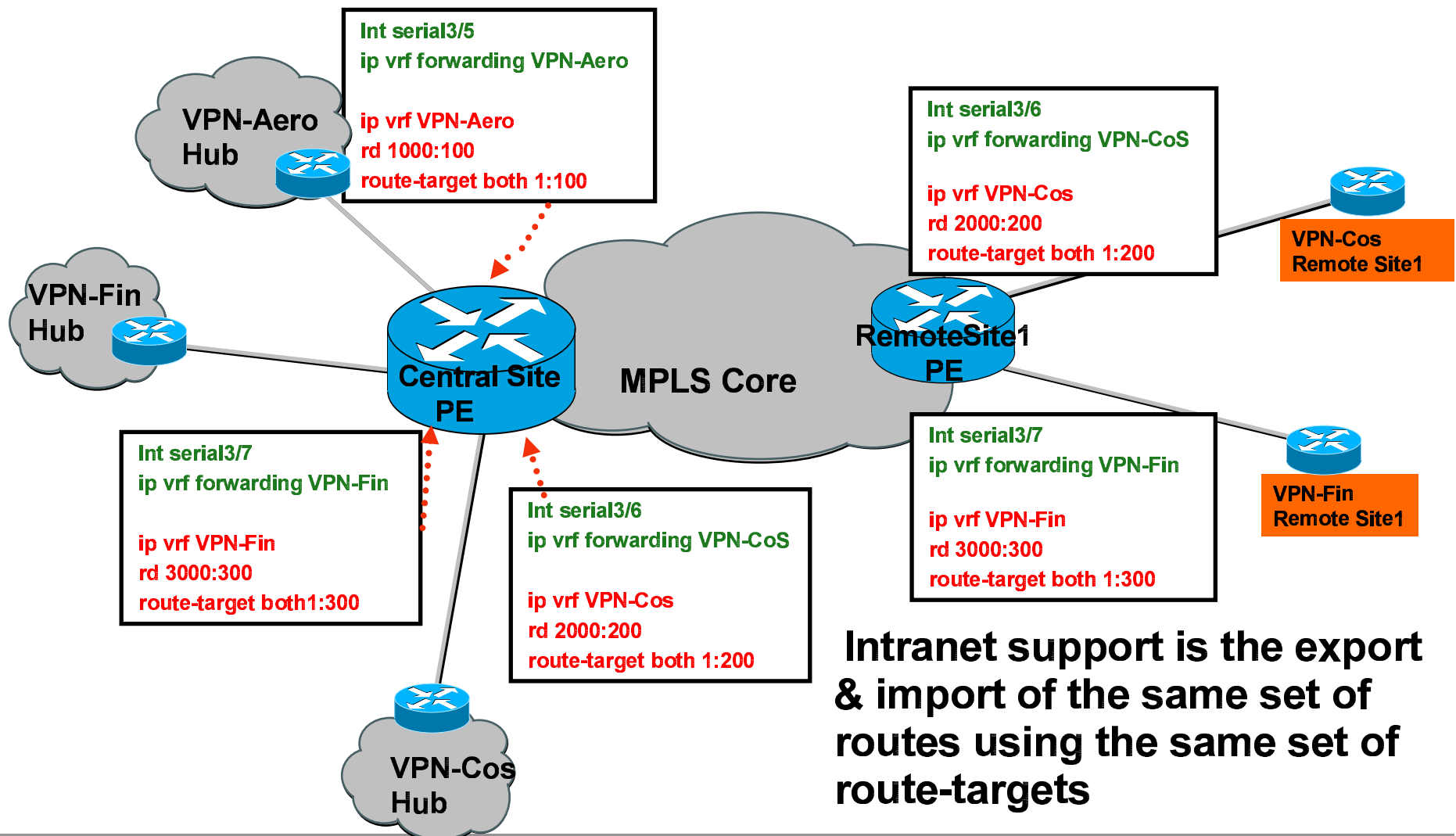
How to Build MPLS L3 VPN Services

- **Full mesh simple intranet service**
- **Advanced Extranet services (want two departments to communicate)**
- **Hub/spoke central site services**
 - Basic hub/spoke services
 - HDVRF
- **How to build Internet Services**
 - Dynamic Default Route
 - Using subinterfaces with FR switching
 - Using Multi-VRF

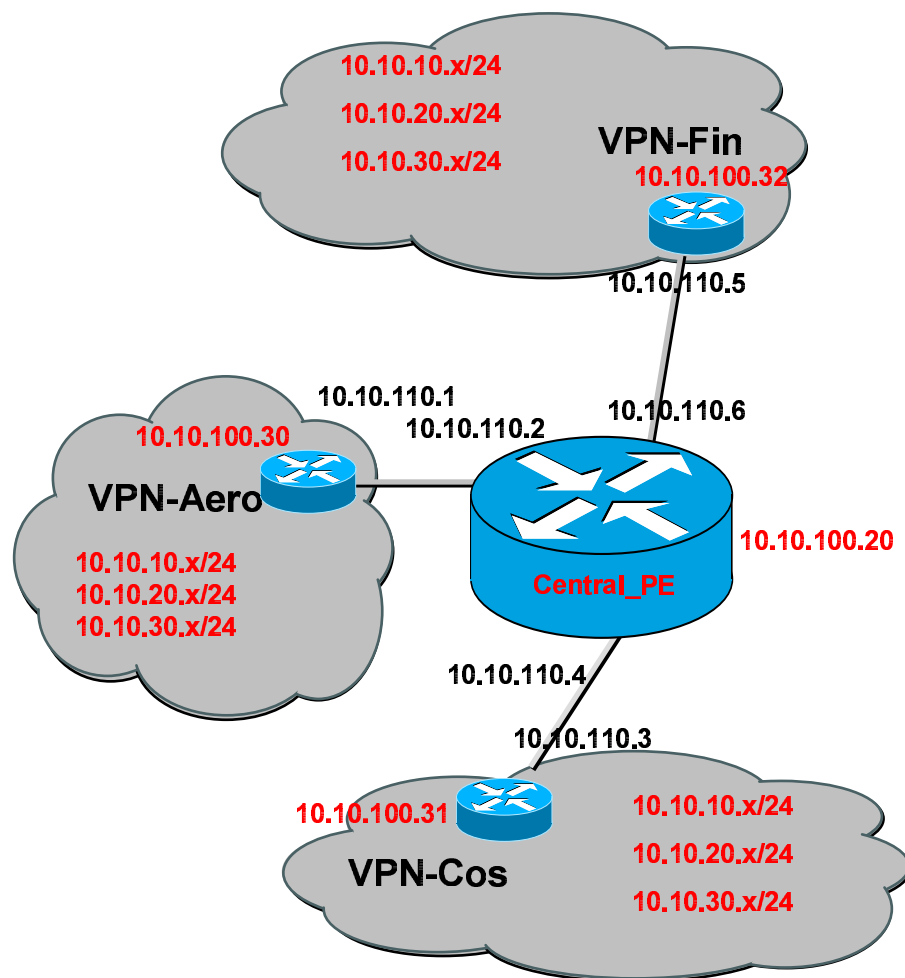
How to Build MPLS L3 VPN Services (Cont.)

- **Services Enabler**
- **Multicast VPNs**
- **Voice VPNs**
- **Multi-AS VPNs**
 - Inter-AS VPNs
 - Carrier Supporting Carrier VPNs

Basic Intranet Model – Full Mesh



Routing Protocol Configuration: Central Site, Routing Update CE-PE within VRFs and PE-PE (MP-iBGP)



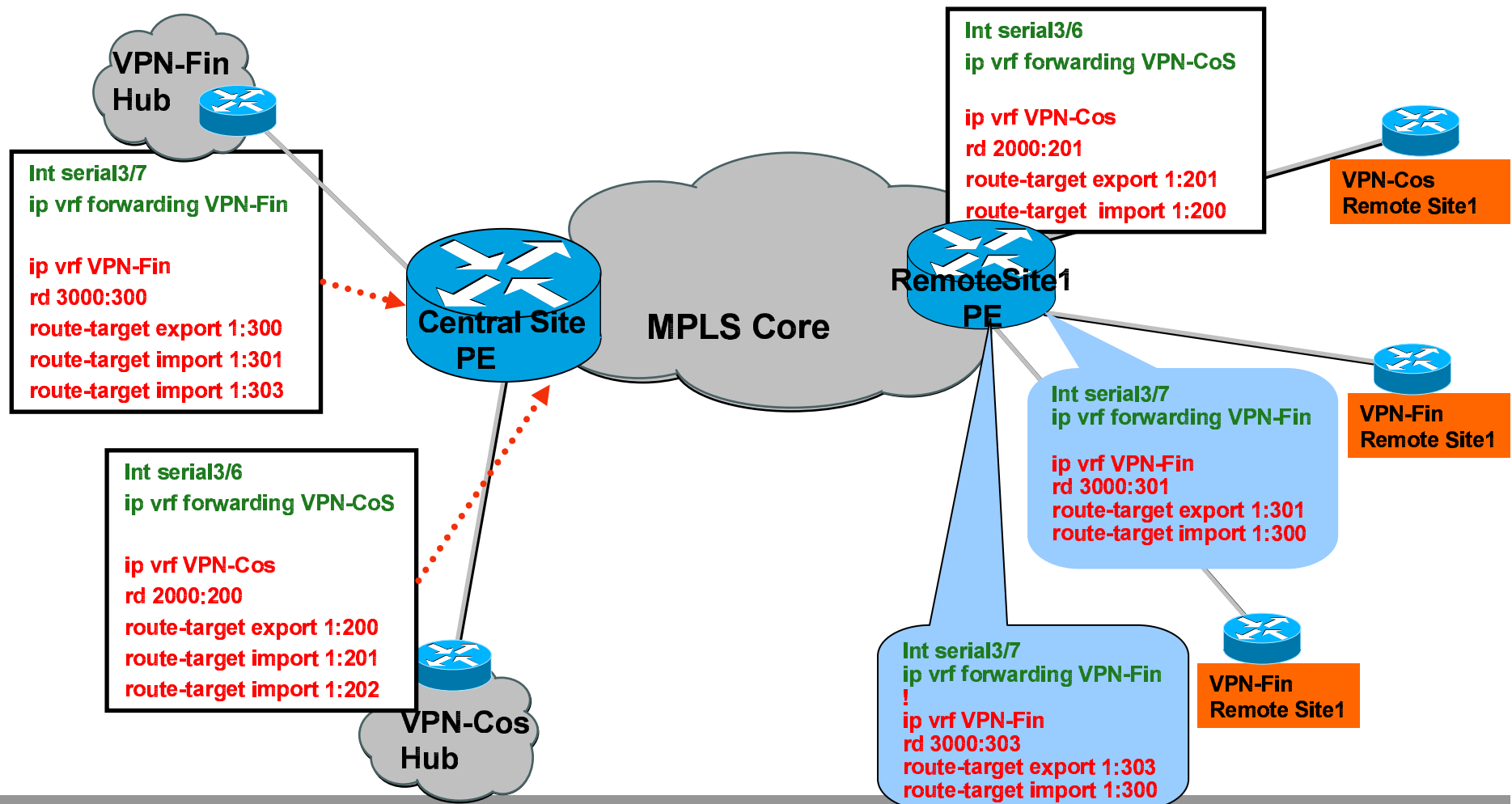
No redistribution in to BGP is required
since central site CE-PE routing
protocol is eBGP

```
Central_PE:
router bgp 100
no synchronization
neighbor 10.10.100.1 remote-as 100
neighbor 10.10.100.1 update-source loopback 0

neighbor 10.10.100.5 remote-as 100
neighbor 10.10.100.5 update-source loopback 0

neighbor 10.10.100.9 remote-as 100
neighbor 10.10.100.9 update-source loopback 0
no auto-summary
!
Central PE needs to MP-BGP peer with all the other PEs
In this MPLS Core topology.
!
address-family ipv4 vrf VPN-Aero
neighbor 10.10.100.30 remote-as 1000
neighbor 10.10.100.30 activate
exit-address-family
!
address-family ipv4 vrf VPN-Cos
neighbor 10.10.100.31 remote-as 2000
neighbor 10.10.100.31 activate
exit-address-family
!
address-family vpnv4 VPN-Fin
neighbor 10.10.100.32 remote-as 3000
neighbor 10.10.100.32 activate
exit-address-family
!
address-family vpnv4
neighbor 10.10.100.1 activate
neighbor 10.10.100.1 send-community extended
neighbor 10.10.100.5 activate
neighbor 10.10.100.5 send-community extended
neighbor 10.10.100.9 activate
neighbor 10.10.100.9 send-community extended
```

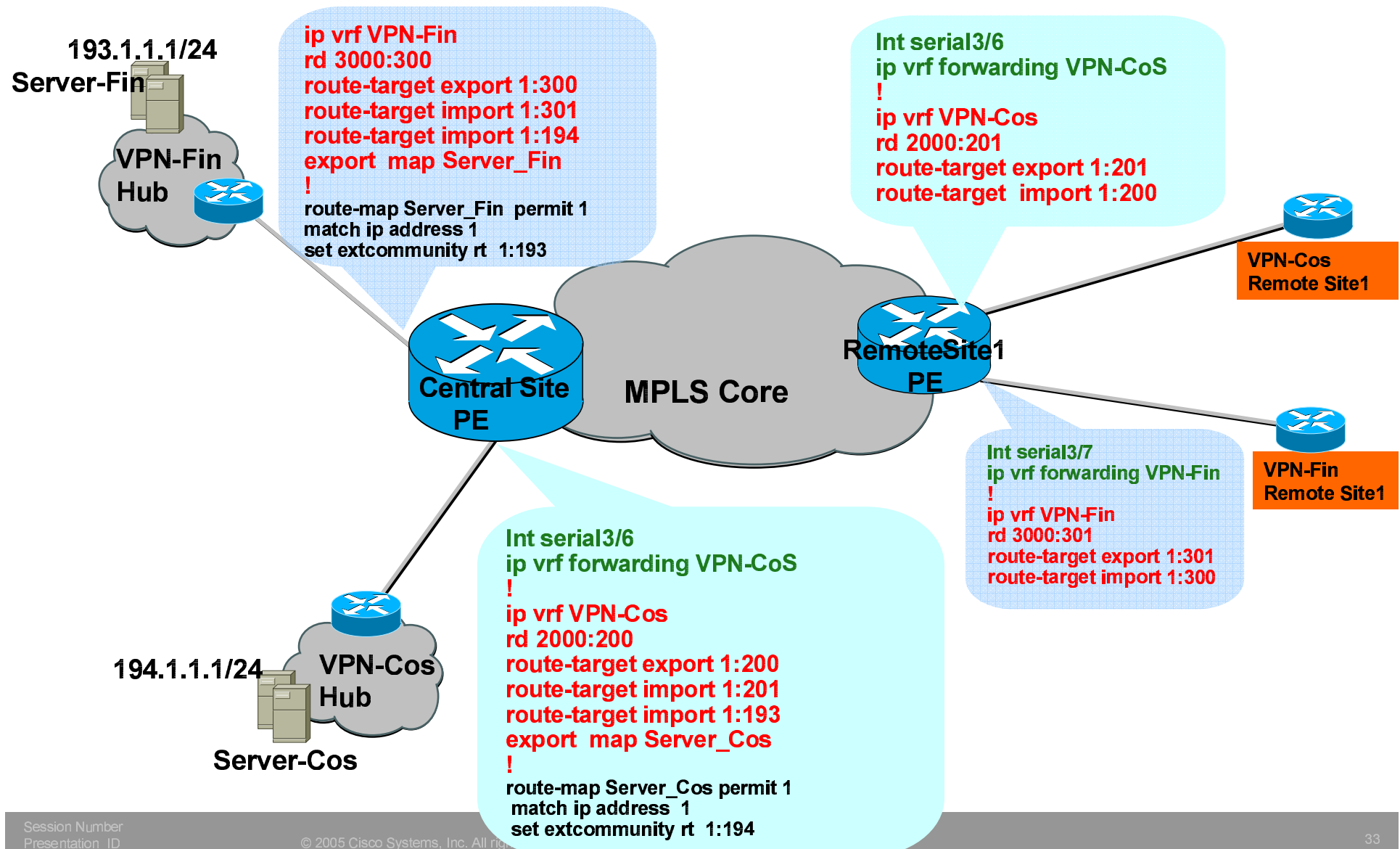
Basic Intranet Model – Hub/Spoke



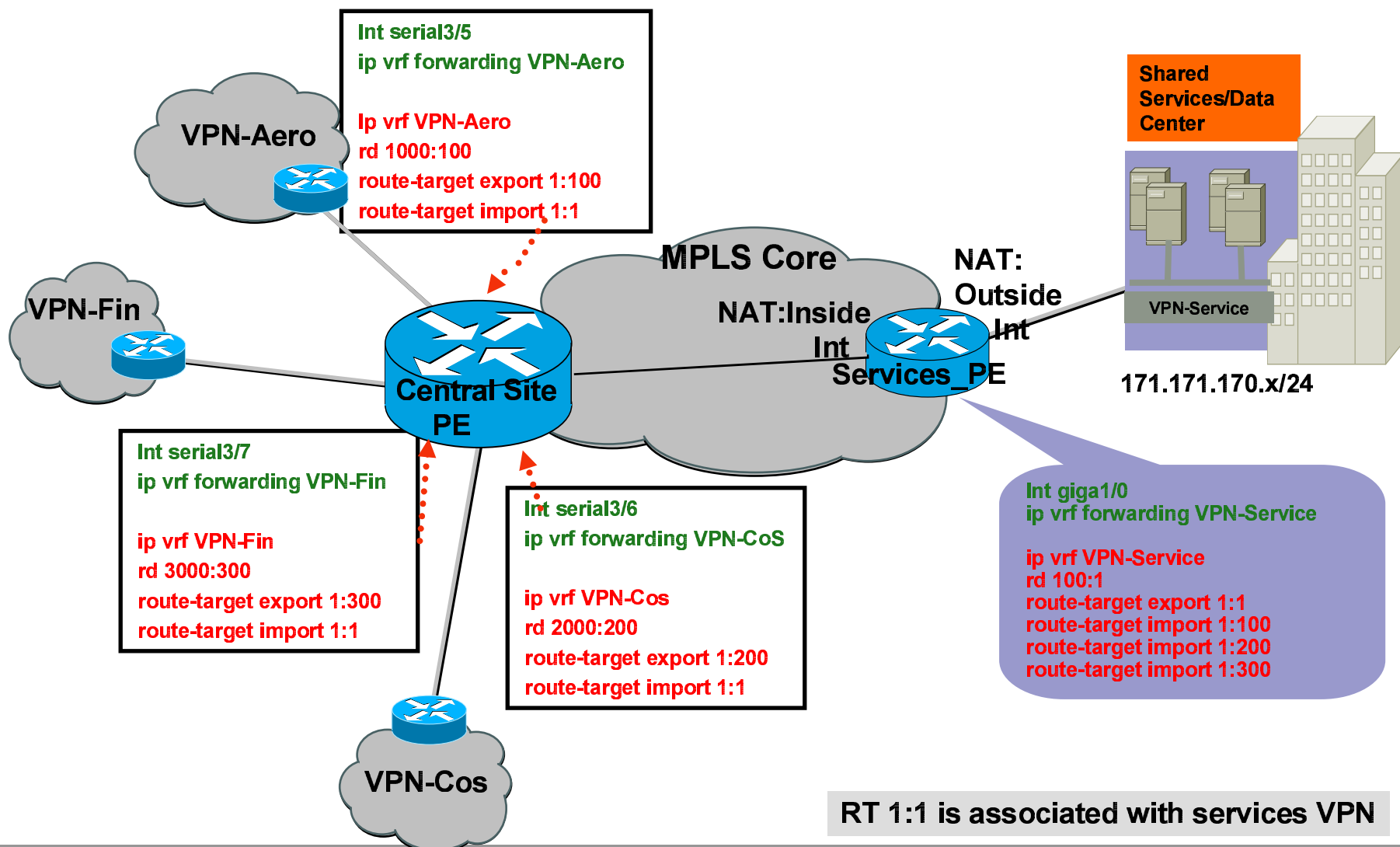
MPLS VPN Extranet Support

- **Extranet support is simply the import of routes from one VRF into another VRF which services a different VPN**
- **Controlled through the use of route target**
 - If we import the routes, we have access;**
 - Therefore the RT defines the service and not the site**
- **Various topologies are viable using this technique**
- **Example of how part of the different VPN networks can communicate**

Extranet Model – Hub/Spoke



Central/Shared Services



Internet Access

- **Default Route**

Use the Global Routing table to access the Internet Gateway with “Global” option

Need a static route for VPN NW pointing to the CE and also redistributing into BGP to advertise to the Internet

Site should be using public addresses or NATed at the CPEs

Dynamic Default Route

Static Default Route

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

Internet Access (Cont.)

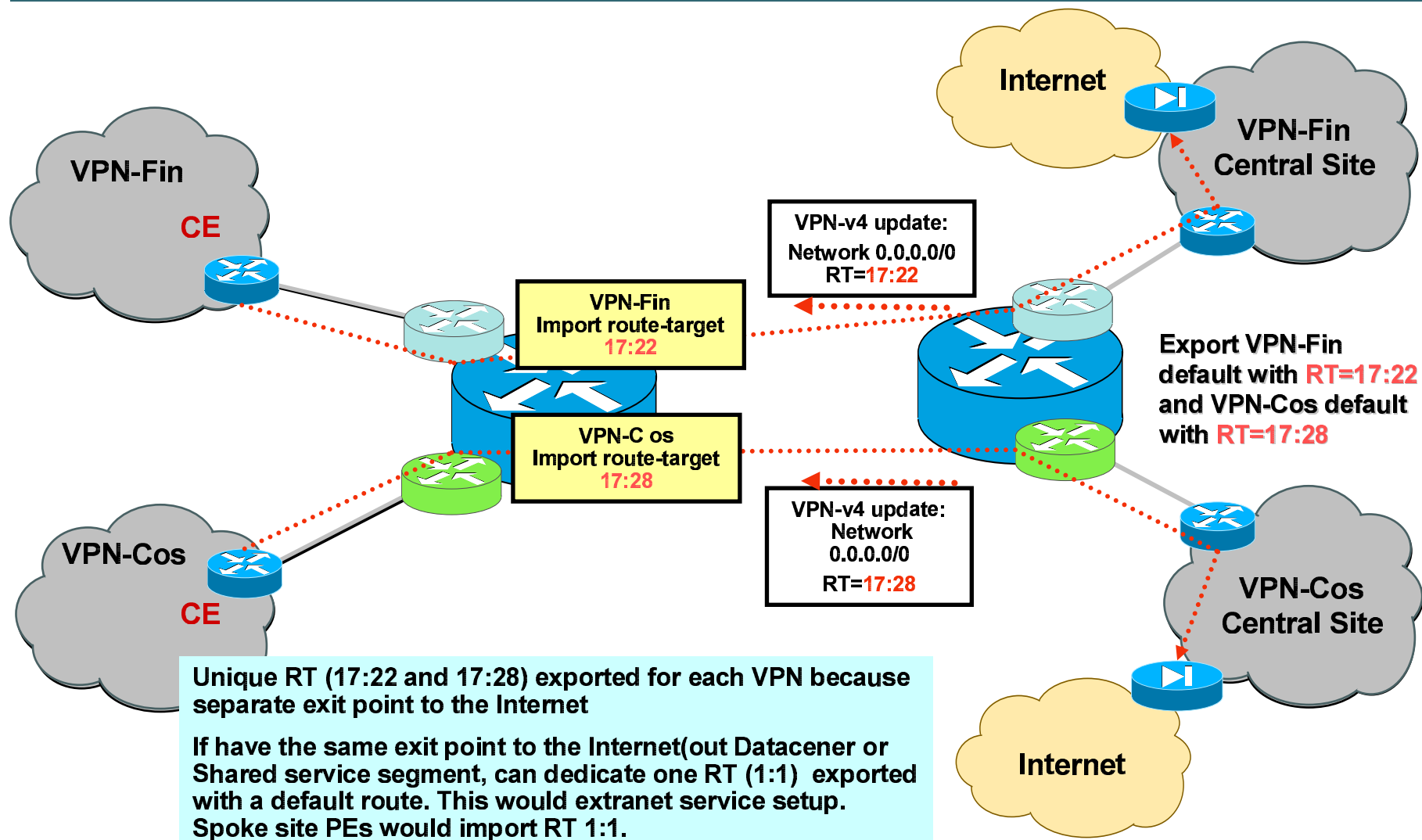
- **Internet access via a VRF that contains the default route to the Internet gateway**

Extranet with Internet VRF

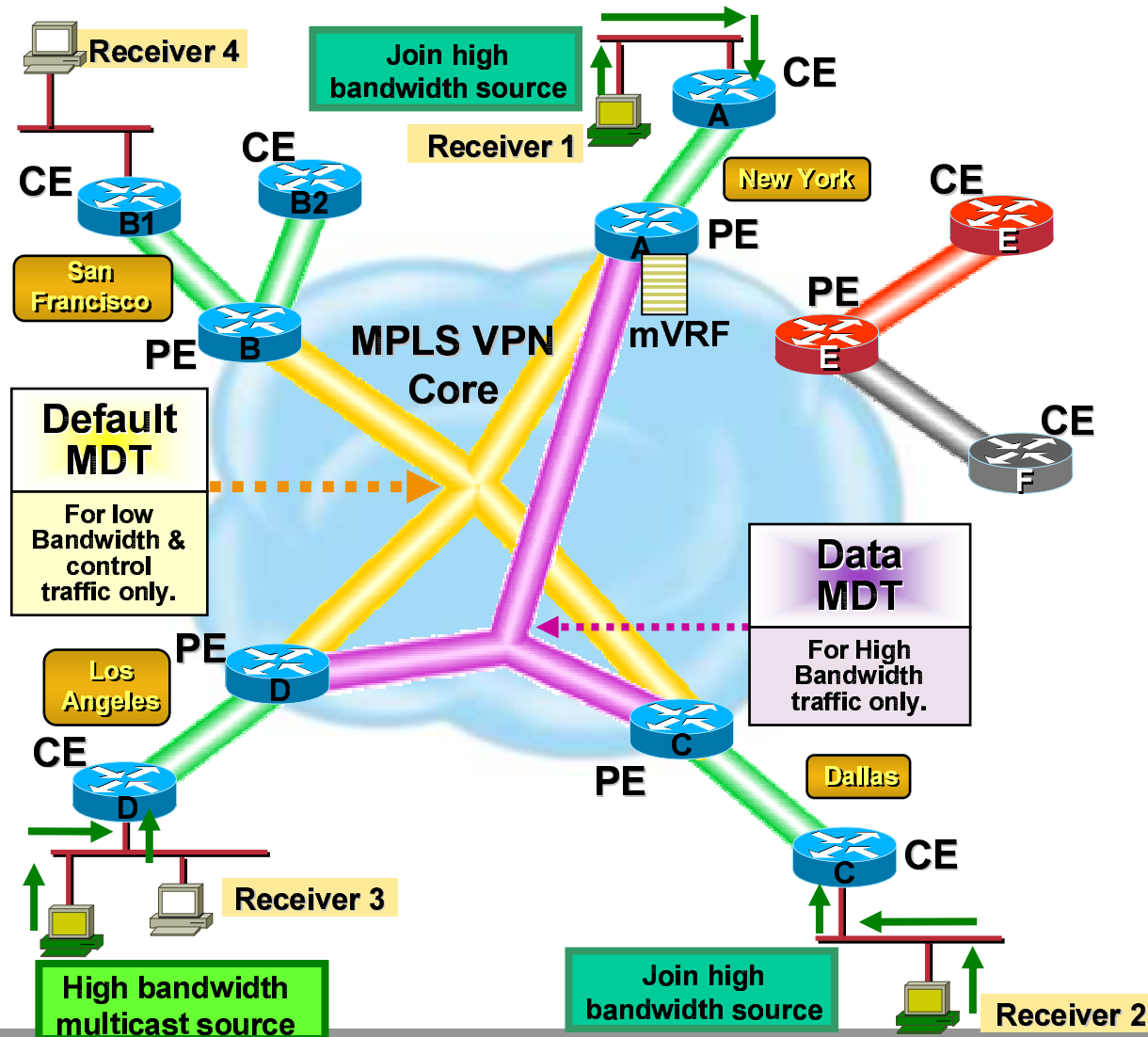
- **A VRF that contains full Internet table**

Scalability issues, security issues, should be avoided

Internet Access – Dynamic Default Route



mVPN Concept and Fundamentals



- Multicast over MPLS L3 VPNs:mVPNs or GRE tunnels
- Default MDT is (per PE) per VRF in the core, touches all PEs where the VRF resides & used for customer PIM control and data packets and always present
- Server sends traffic; interested receivers 1 & 2 join high bandwidth source
- PIM adjcencies formed between PE-CE devices (not CE-CE)
- PE are always a root to the MDT; PE is also a leaf to MDT rooted on remote PEs

Configuration example: SSM or Bi-Dir in the Backbone, PIM-SM in the VPN

```
!  
ip vrf VPN-Fin
```

```
rd 1:1  
route-target export 1:1  
route-target import 1:1  
mdt default 232.0.0.1
```

← group for default MDT

mdt data 232.5.0.0 0.0.0.255 threshold 100

← needed for data-mdts (kbps)

```
!  
ip multicast-routing  
ip multicast-routing vrf VPN-Fin
```

← enables multicast routing for this VRF

```
!  
interface FastEthernet1/0/0  
ip vrf forwarding VPN-Fin  
ip address 172.16.140.1 255.255.255.0  
ip pim sparse-dense-mode  
!  
interface GigabitEthernet4/0/0  
ip address 10.0.2.1 255.255.255.0  
ip router isis  
ip pim sparse-mode  
ip route-cache distributed  
tag-switching ip  
!
```

```
router isis  
net 49.1111.3333.3333.3333.00  
!
```

```
router bgp 1  
no synchronization  
neighbor 10.0.0.2 remote-as 1  
neighbor 10.0.0.2 update-source Loopback0  
!
```

address-family ipv4 vrf VPN-Fin

```
neighbor 172.16.140.2 remote-as 1001  
neighbor 172.16.140.2 activate  
exit-address-family  
!
```

address-family mdt vrf VPN-Fin

```
neighbor 172.16.140.2 remote-as 1001  
neighbor 172.16.140.2 activate  
exit-address-family  
!
```

```
address-family vpnv4  
neighbor 10.0.0.2 activate  
neighbor 10.0.0.2 send-community extended  
exit-address-family  
!
```

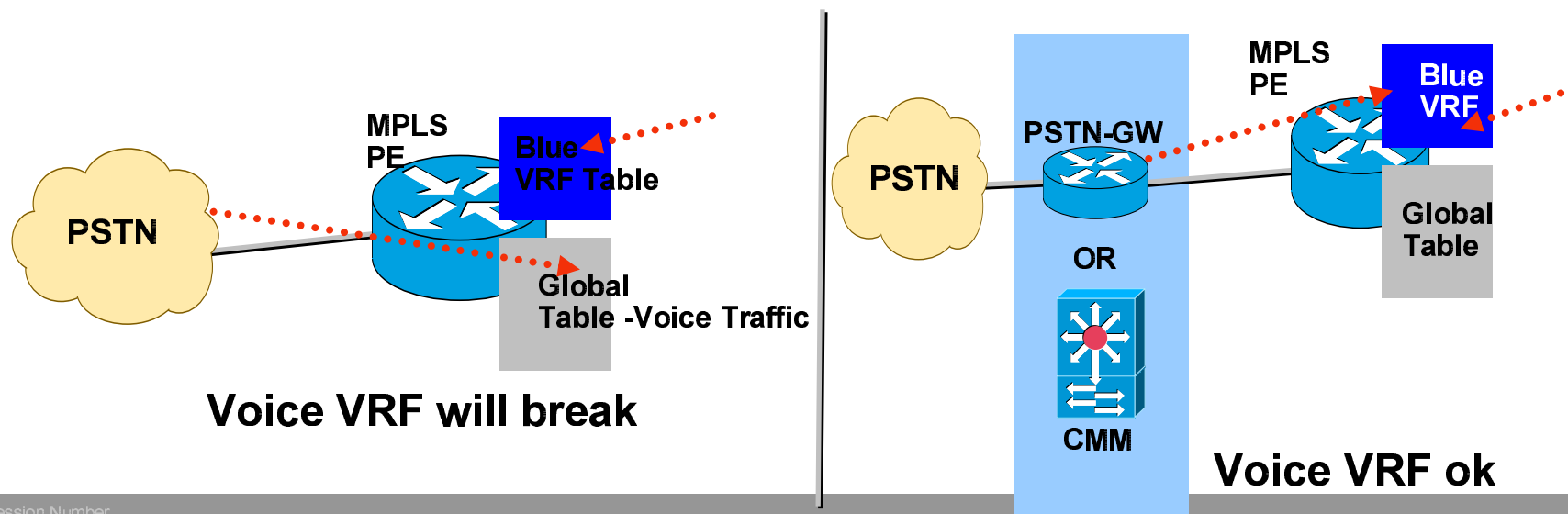
ip pim ssm default

Commands in red are the only commands needed to enable the VPN for multicast

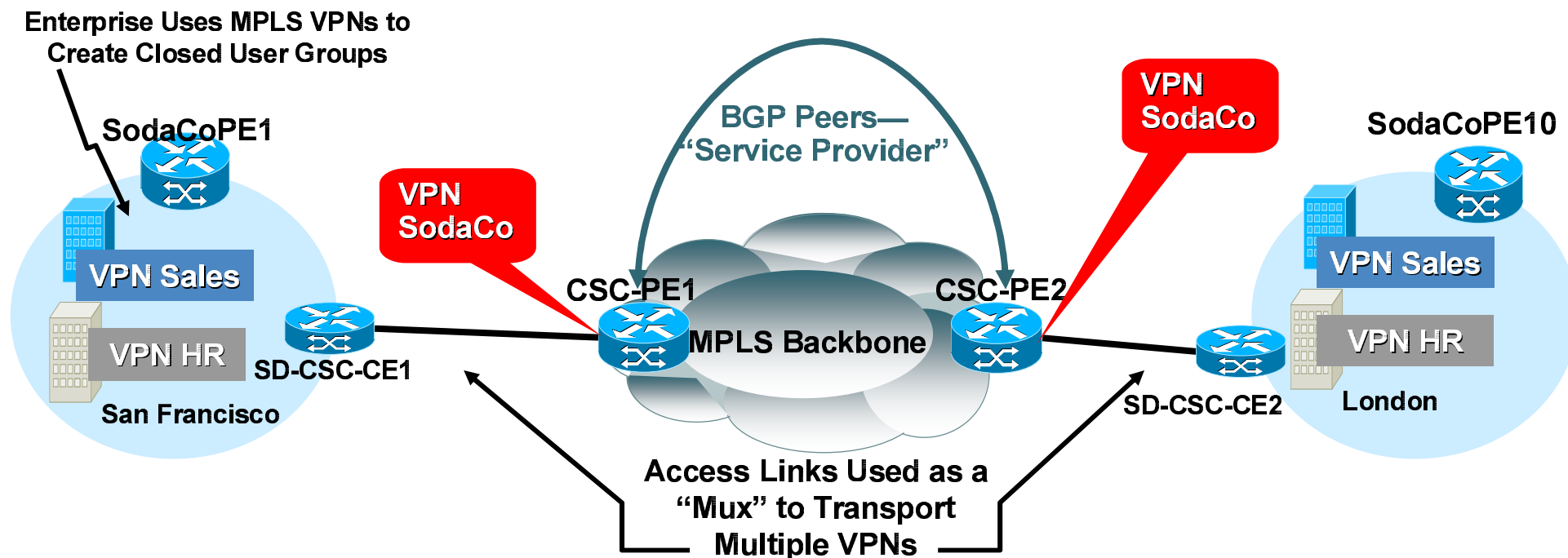
1. Enable native multicast in the core
2. Set PIM adjacencies between PE & CE
3. Create VRF, default and data MDTs

VoIP VPN Services

- Setup just like any other VRF
- Voice signaling protocols are not VRF aware yet
- Voice traffic needs to be packetized on a separate box from where VRF interface exists (MultiVRF CE & VPN PE)
Need to have a separate voice gateway from a PE when connecting to the PSTN network



What if Enterprise Customers with MPLS VPNs Needs MPLS VPN Services from a Service Provider?

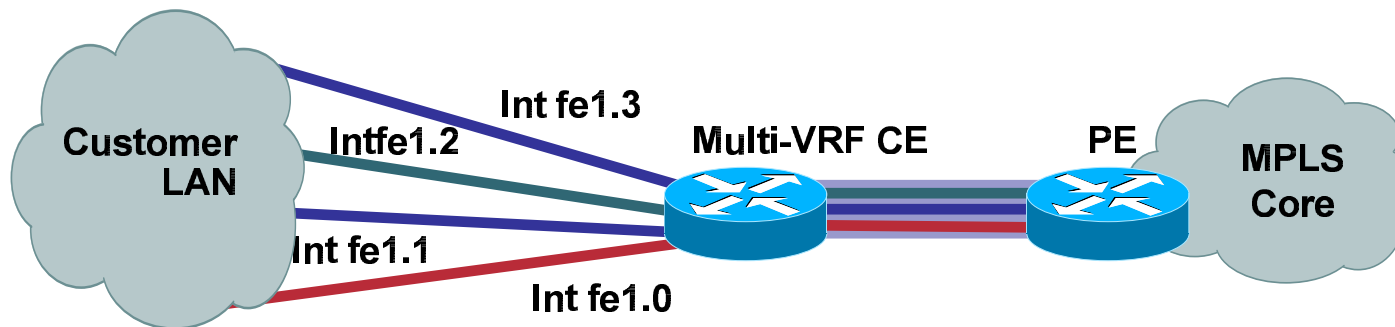


- PE1 and PE10 peer and exchange external VPNv4 routing
- CSC-PEs treat CSC-CE as a VPN sites. CSC-PEs exchange VPNv4 routes
- No MP-BGP on CSC-CEs just IGP+LDP or eBGP+IPv4 Labels
 - CSC-CE1 and CSC-CE2 don't carry any external VPNv4 routing
 - connect MPLS VPN networks and extend LSP path
- VPN Sales, HR, etc. get carried within VPN SodaCo in the provider core

Additional L3 VPN Options

- **Multi-VRF VPNs**
- **MPLS VPNs over IP**
 - MPLS VPNs over GRE**
 - Manual point to point**
 - Dynamic multipoint**
 - IPsec Encryption**
 - MPLS VPNs over L2tpV3**
 - Manual point to point**
 - Dynamic multipoint**

Multi-VRF: Extending VPN functionality to CPE w/o full MPLS



- Reduced # of edge devices per VPN (metro area, multi-tenant bldg, multi-VLAN support, etc.)
- No MPLS functionality on the CE, no CE-PE label exchange, no MP-BGP, no VRF ID or labels are created in control or forwarding plane
- Pair of ingress and egress interface is bound to a VRF
- Separate dedicated VRF instances + a global routing table
- Overlapping address space is supported
- Same routing protocol support as in normal VRF; Local inter-VRF routing is supported as well

Common Multi-VRF Applications

1. Provide Internet and VPN services using the same CE

**Traffic separation on VRF lite CE instead of on a PE;
public and private traffic are kept separate**

There is an option to put Internet table in a VRF on a CE

2. Multi-VRF for campus segmentation

3. Multi-VRF campus segmentation expansion over MAN or WAN

LAN/MAN/WAN core could be MPLS or IP

GRE/L2TPv3, point to point or dynamic multipoint

MPLS over Dynamic Multipoint GRE

- MPLS or GRE relies on ACLs
- Mixed tunneling environments are not easily supported – if other PEs cannot decapsulate MPLS over GRE then VPN traffic could be blackholed

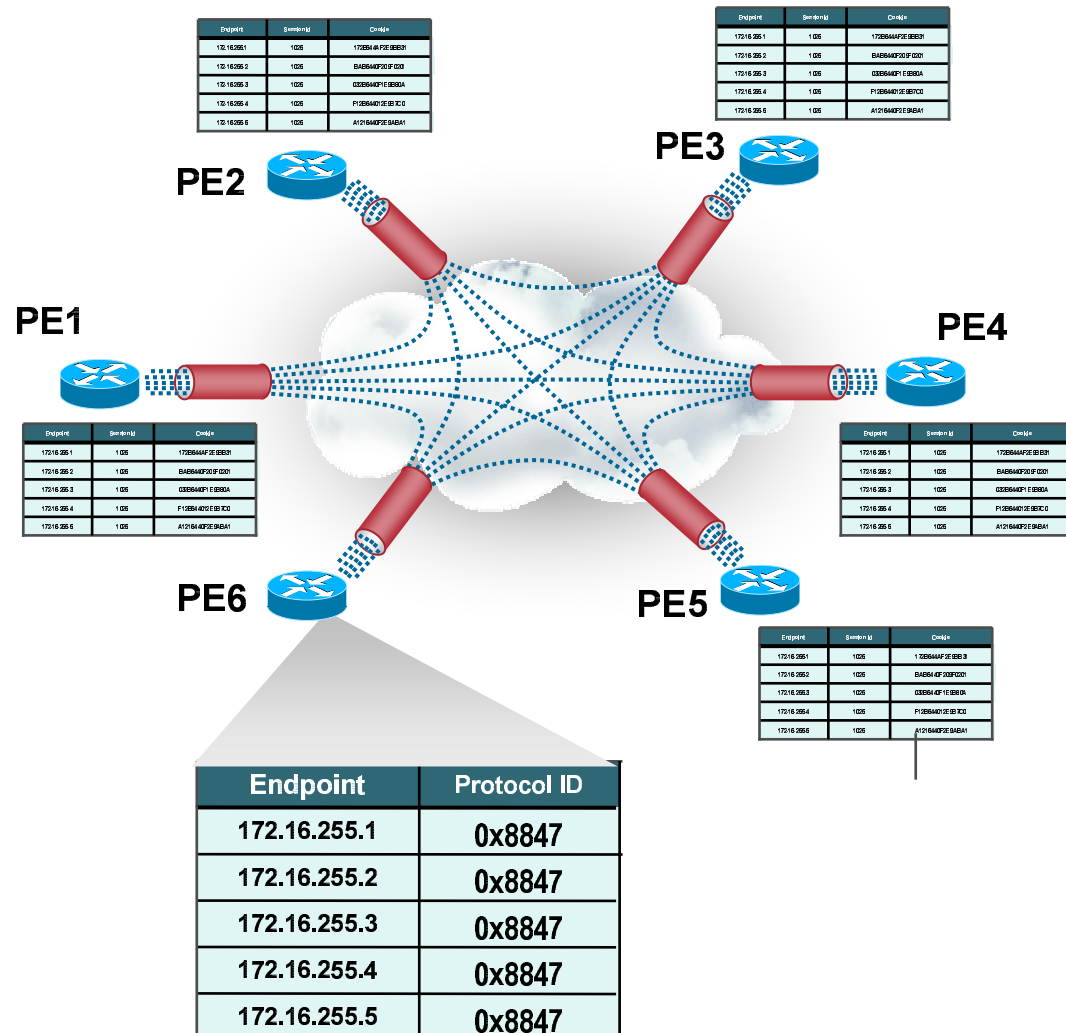
-Still Needed: A method for PEs to advertise if they are able to receive MPLS over IP traffic, and with what type of encapsulation

Bind tunnel end pts to tunnel capabilities

-MTU size

-Security: a hacker guess a correct 20bit MPLS label?

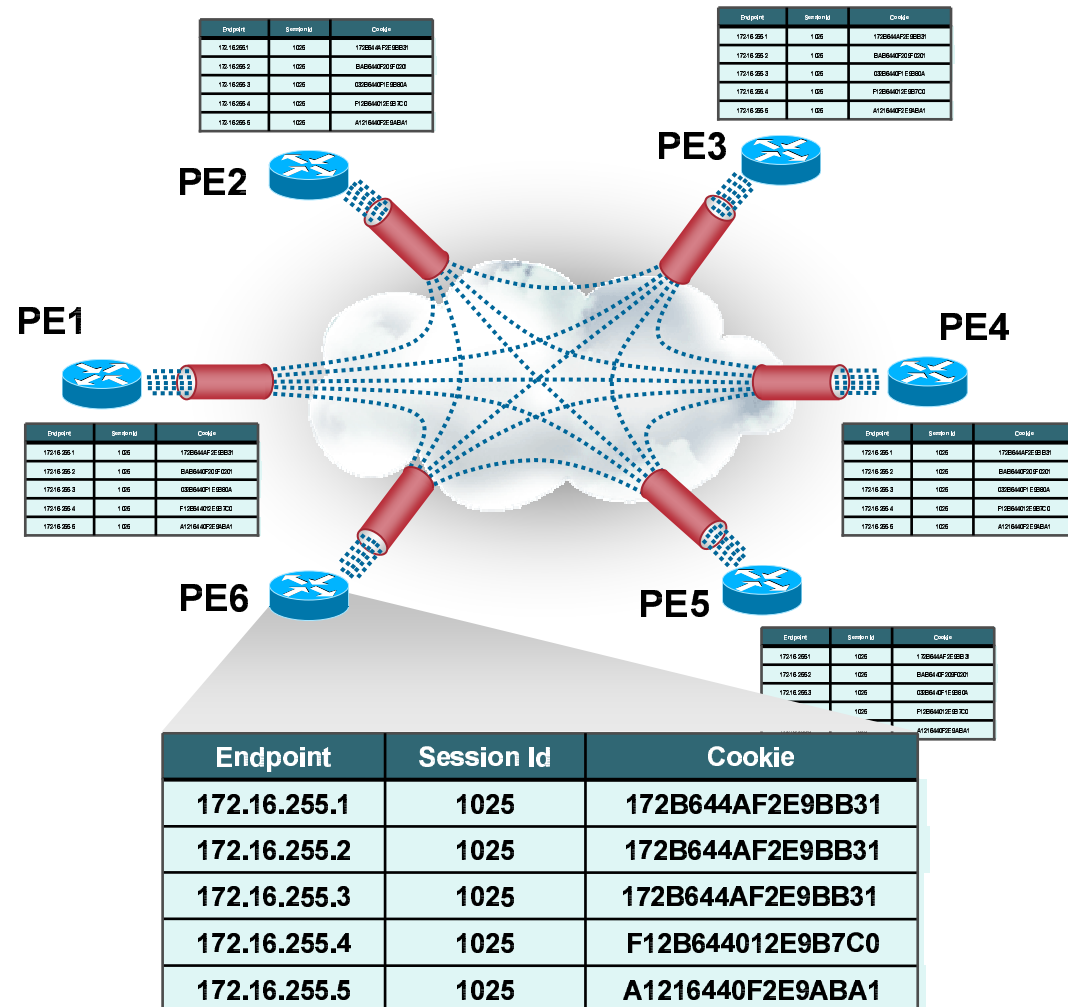
100 pps attack rate & **100** active VPN labels (routes) on a PE in **1 minute** and **45 seconds!**



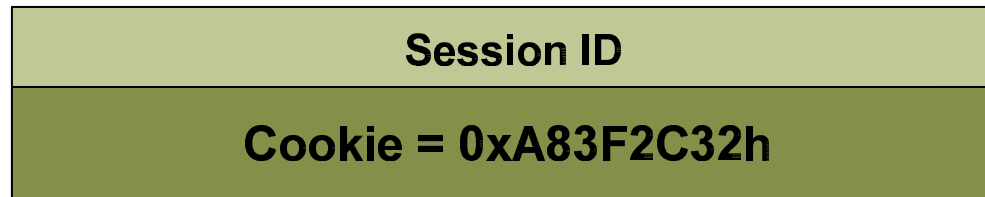
Tunneling Technologies - L2TPv3 w/BGP Tunnel SAFI, more robust solution

- One L2TPv3 Multipoint session is dynamically created on each PE for receiving traffic from other PE's (point to point L2TPv3 signaling is not used)
- BGP advertises tunnel capabilities via Tunnel SAFI - MPLS over L2TPv3 traffic only sent to PEs which can handle it
- Tunnel SAFI also includes per-PE Session ID and Cookie pair
- How quickly can a hacker guess a correct 64 bit L2tpv3 cookie?

@10Mbps attack rate – 6K years!



MPLS over L2TPv3 - L2TPv3 Packet Authentication Check with Cookie



- **64-bit value must match for each packet**
- **Not a 64-bit lookup! Just a very fast compare based on the Session ID lookup**
- **No encryption hardware needed**
- **Rather than checking an IP SA or DA, L2TPv3 “seeds” each packet with an unguessable value selected at random by each PE, and advertised to other PEs in the VPN via the BGP Tunnel SAFI**
- **Somewhat like an ACL, but simple to manage and virtually impossible for a hacker to guess**

VPN Services Over IP Tunnels

Review of Capabilities

	Static IP	Static GRE Overlay	Dynamic Multi-point GRE	L2TPv3 w/SAFI
Encapsulates MPLS over IP	Yes	Yes	Yes	Yes
Tested in a large active deployment	?	Yes	?	Yes
Avoids full mesh via scalable, dynamic, p2mp tunnels	No	No	Yes	Yes
Avoids blackholes by advertising tunnel capabilities	No	No	No	Yes
Encapsulation facilitates highspeed lookup and distributed processing assist	No	No	No	Yes
Simple, scalable, anti-spoofing protection built-in	No	No	No	Yes

MPLS over IP Tunnels - Summary of what can be done with this technology

- **Extending the reach of MPLS**

MPLS services (such as RFC 2547 VPNs) based on IP Tunnels can cross multiple providers (Inter-provider) or administrative domains (Inter-AS) to reach customers anywhere IP reaches

- **Migration to MPLS**

MPLS/MPLS where available, MPLS/IP where not

MPLS over IP Tunnels - Summary of what can be done with this technology

- **Operational Flexibility**

Some service providers do not yet have (or do not yet want) MPLS in their core networks, but still want to offer their customers MPLS-based services

- **Scaling MPLS VPN deployments**

IP route aggregation allows for scaling MPLS VPNs across a very large number of PEs without increasing the number of PE-PE LSPs and associated /32 routes advertised in an IGP

MPLS over IP Tunnels - Summary of Available Tunnelling Technologies

- **Static MPLS over GRE may be used to connect a small number of isolated nodes or disparate MPLS networks, but is not recommended for high scale deployments**
- **Dynamic Multipoint Tunneling available with GRE or L2TPv3 solves the manual provisioning problem with static GRE tunnels, but still can allow blackholes**
- **The BGP Tunnel SAFI prevents blackholes to routers which cannot decapsulate a given type of IP tunnel, allowing staged migration to MPLS**

MPLS over IP Tunnels - Summary of Available Tunnelling Technologies (Cont.)

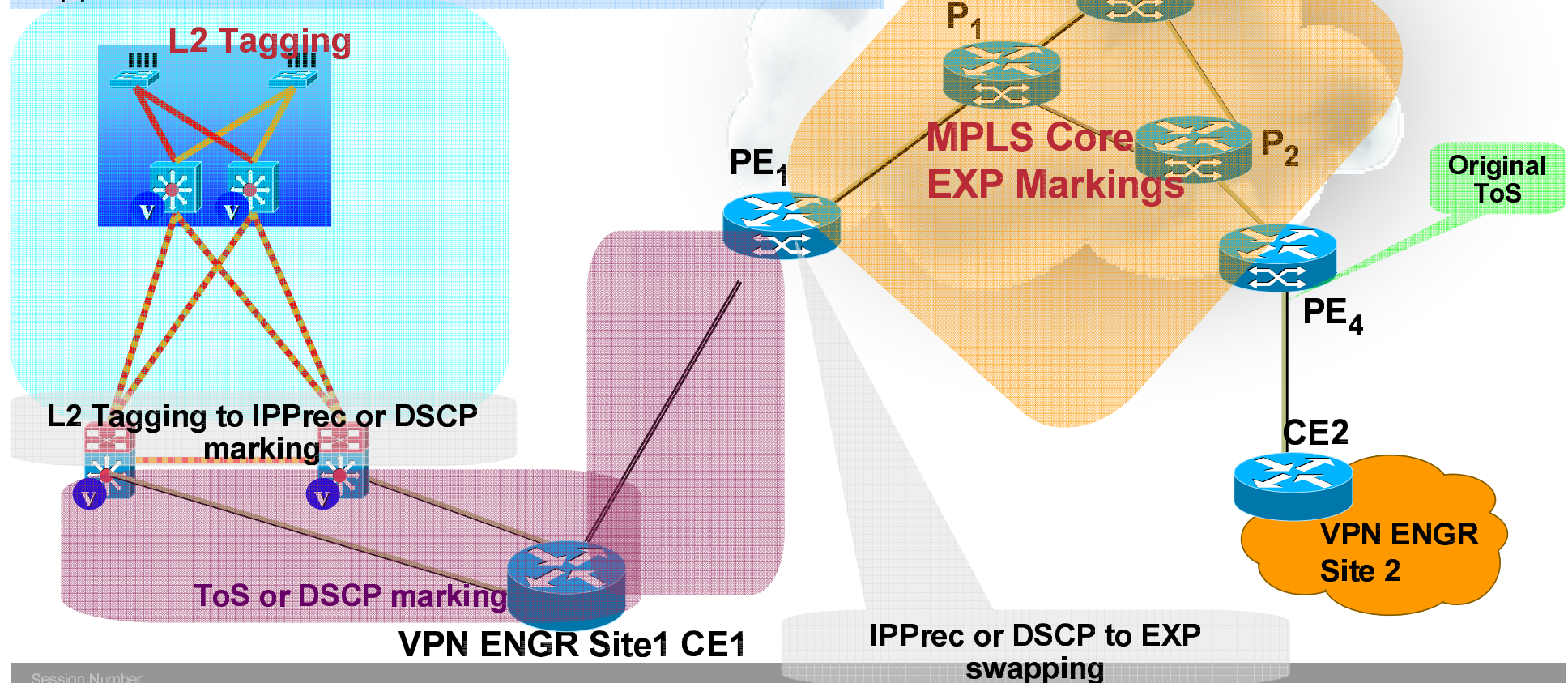
- **IPsec can provide strong security, but is expensive from an opex and capex perspective**
- **L2TPv3 includes lightweight yet strong anti-spoofing protection, with zero additional opex complexity over mGRE, and no reliance on ACLs**
- **Conclusion: MPLS over L2TPv3 w/BGP Tunnel SAFI is the most feature rich and proven MPLS over IP Tunnel offering among the choices available**

Agenda

- **MPLS Basics**
- **MPLS L3 VPNs**
- ***MPLS QoS***
- **Applying MPLS in the Enterprise**
- **Case Studies**

MPLS Quality of Service

L3 Access/Edge: IP S.D. Add, Ports, ToS, DSCP
AccessL2: L2-tag, MAC address
PE: ToS, DSCP, MPLS-Exp bits; Core: MPLS EXP bits.
If more than 8 CoS used, will need to consolidate classes (preferably) on CE egress int. as EXP only supports 8 classes.

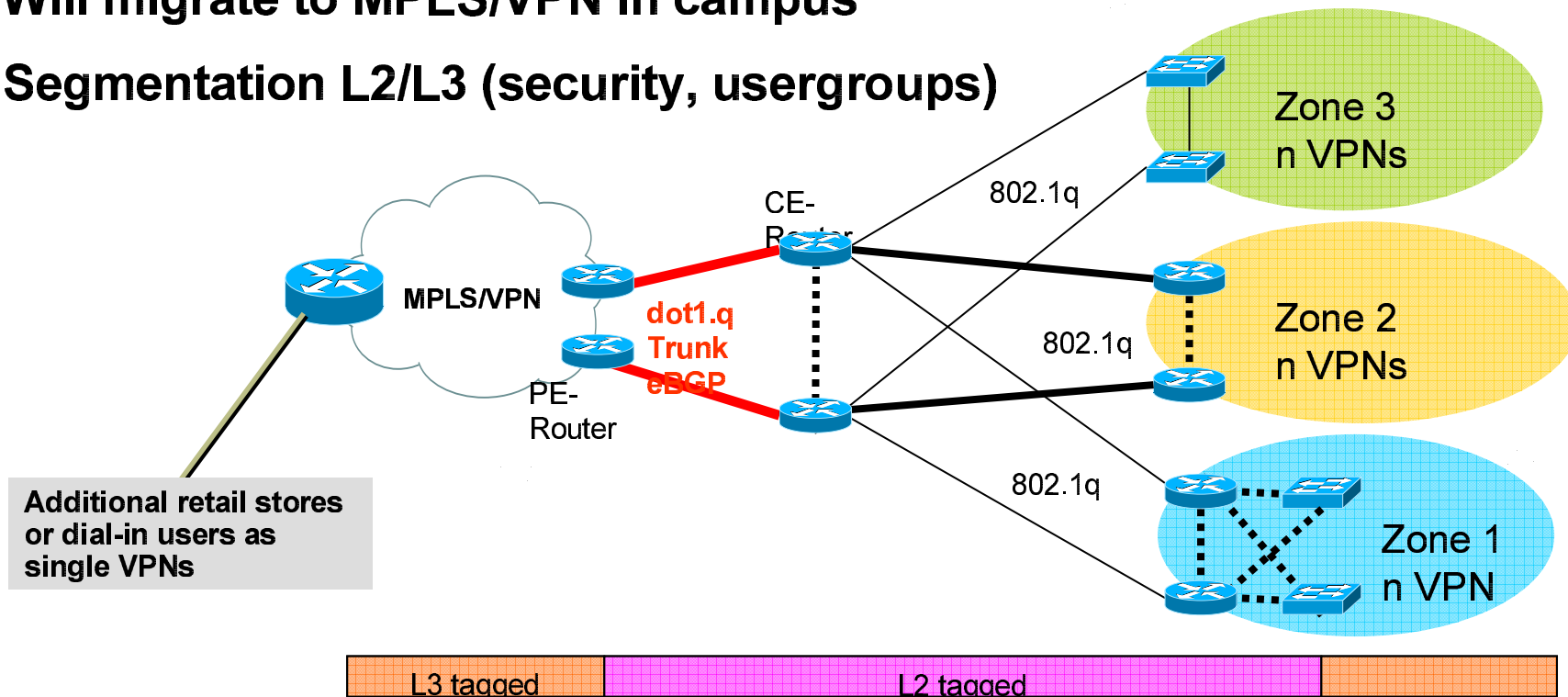


Agenda

- **MPLS Basics**
- **MPLS L3 VPNs**
- **MPLS QoS**
- ***Case Studies***

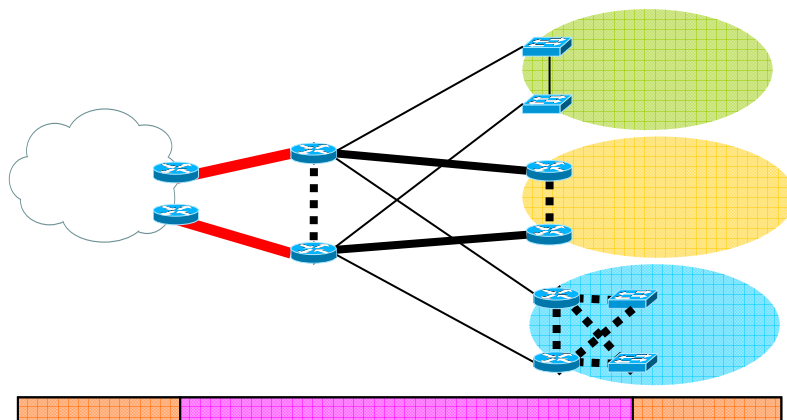
VRF-lite Customer - Retail

- Uses MPLS SP for WAN network
- Multi-VRF for PE-CE and CE traffic separation (VLAN mapping)
- Will migrate to MPLS/VPN in campus
- Segmentation L2/L3 (security, usergroups)



VRF-lite Customer - Retail

- Type of PE/P-Box: Cisco 7200
- P-Topology: P2P (6)
- Cisco IOS Software Release 12.3
- Core Routing: IS-IS, MPBGP
- Type of core links: OC3 POS
- QoS on Core Links: LLQ
- MPBGP RR: no



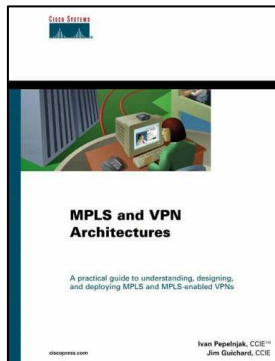
- Type of CE-Box: Cisco 3750M / SUP720
- CE-Topology: directly connected VLAN
- VRFlight#: < 10
- Routes per VRFlight#: < 1,000
- Cisco IOS Software Release 12.2/12.2SX
- Edge routing: directly connected
- PE-CE links: GE channel/trunks
- QoS on PE-CE Links: HW Prio
- CE additional functions: Marking
- Multicast: no

References

- **Today**

CE may write Community with customer identification

Literature

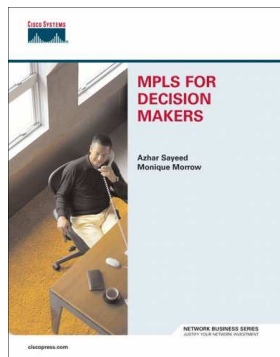
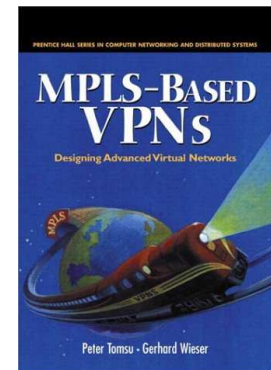


MPLS and VPN Architectures(Volume I & II)

**Ivan Pepelnjak
Jim Guichard**

MPLS –Based VPNs

**Peter Tomsu
Gerhard Wieser**



MPLS for Decisions Makers

**Azhar Sayeed
Monique Morrow**

Configuration example: SSM or Bi-Dir in the Backbone, PIM-SM in the VPN; MDT

```
ip vrf VPN-Fin
```

```
rd 1:1
route-target export 1:1
route-target import 1:1
mdt default 239.1.1.0
```

!If use 239 for default then don't need to use address-family for mdt

← group for default MDT

```
mdt data 232.5.0.0 0.0.0.255 threshold 100
```

← needed for data-mdts (kbps)

```
!
```

```
ip multicast-routing
```

```
ip multicast-routing vrf VPN-Fin
```

← enables multicast routing for this VRF

```
!
```

```
interface FastEthernet1/0/0
ip vrf forwarding VPN-Fin
ip address 172.16.140.1 255.255.255.0
ip pim sparse-dense-mode
```

```
!
```

```
interface GigabitEthernet4/0/0
ip address 10.0.2.1 255.255.255.0
ip router isis
ip pim sparse-mode
```

```
ip route-cache distributed
ip ipsec switching
```

```
router isis
net 49.1111.3333.3333.3333.00
!
```

```
router bgp 1
no synchronization
neighbor 10.0.0.2 remote-as 1
neighbor 10.0.0.2 update-source Loopback0
!
```

```
address-family ipv4 vrf VPN-Fin
```

```
neighbor 172.16.140.2 remote-as 1001
neighbor 172.16.140.2 activate
exit-address-family
!
```

```
address-family mdt vrf VPN-Fin
```

```
neighbor 172.16.140.2 remote-as 1001
neighbor 172.16.140.2 activate
exit-address-family
!
```

```
address-family vpnv4
neighbor 10.0.0.2 activate
neighbor 10.0.0.2 send-community extended
exit-address-family
!
```

```
ip pim ssm default
```

Commands in red are the only commands needed to enable the VPN for multicast.

1. Enable native multicast in the core
2. Set PIM adjacencies between PE & CE
3. Create VRF, default and data MDTs