CHAPTER 2

Deployment Architectures

As mentioned earlier, network virtualization extension across the WAN can be broadly classified into two categories:

- Inter-MAN/ large campus connectivity or WAN core
- Virtualized branch aggregations or WAN edge

WAN Core

After creating MPLS MAN islands, this is the next logical step when migrating to MPLS-based enterprise networks. The different options for interconnecting MPLS MANs are:

- MPLSoL2 service—If it is a legacy Layer 2 or Layer 2 VPN service from SP
- MPLSoGRE—If it is a Layer 3 VPN service from SP
- Carrier Supporting Carrier (CSC)

Typically, in a large enterprise, the WAN core consists of dedicated point-to-point, high-bandwidth links. We do not expect these to move to Layer 3-based (such as Layer 3 VPNs) connections because these are deemed critical links that require the fastest possible round-trip times and higher bandwidths —hence Layer 2circuits are preferred. Additionally, these are few in numbers and hence cost advantages of Layer 3 services are not necessarily applicable.

While all three options are discussed in this chapter, only MPLSoL2 Service is discussed in depth in WAN Core—MPLSoL2 Service since it is expected to be the most-widely deployed.

MPLSoL2 Service

This is the simplest deployment model for connecting MANs if the enterprise already has Layer 2 connectivity between them either via legacy WAN (FR/ATM/POS) or via Layer 2 VPN service (AToM) from a provider. The migration involves converting the edge devices into a P/PE and making it part of the MPLS MAN network.

The MANs are assumed to be already MPLS enabled and configured for enterprise-deployed VPNs. As shown in Figure 2-1, the WAN edge router used for interconnecting MANs plays the role of a P device. It is expected to label switch packets between the MANs across the SP network.

Figure 2-1 MPLS Over Legacy Layer 2 Service

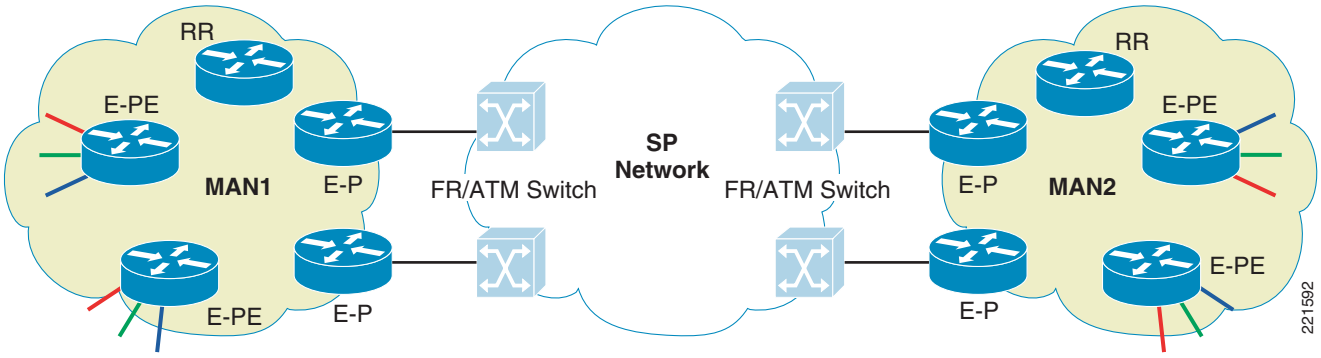
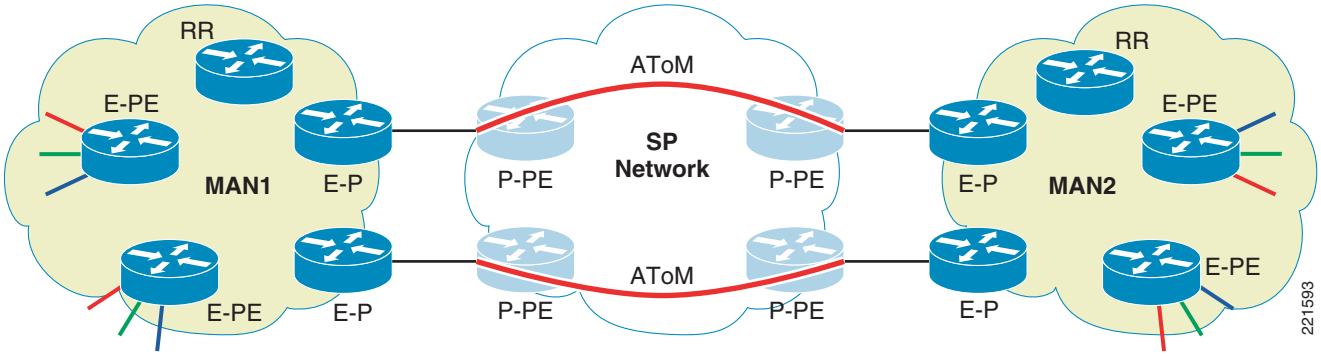


Figure 2-2 MPLS Over Layer 2 VPN Service



From a control plane perspective, the following are expected to be run over the Layer 2 link:

- IGP such as EIGRP or OSPF for MPLS device reachability (P/PE/RR)
- LDP for label distribution
- MP-iBGP for VPN route/label distribution

If these MAN islands/campuses are under different administrative control, then Inter-AS can be implemented. Typical Inter-AS models are:

- Back-to-back VRFs
- ASBR-to-ASBR with MP-eBGP
- ASBR-to-ASBR with multihop EBGP using Route Reflectors

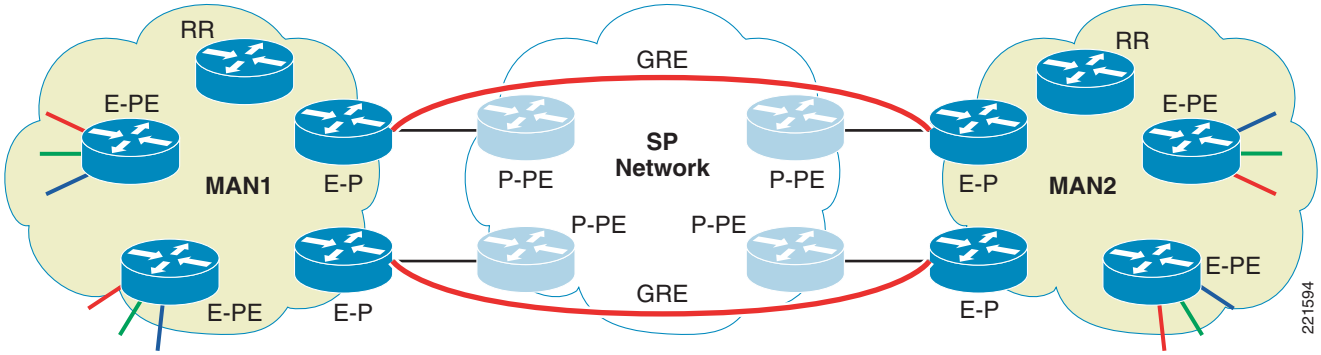
Apart from being a simple solution to deploy, it also offers wider platform options. All the platforms that support P roles should be deployable. All the features that would be deployed within a MPLS network (such as TE) can also be deployed across the WAN core.

MPLSoGRE

The implementation assumes that the enterprise has a Layer 3-based service such a Layer 3 VPNs from a provider interconnecting the MPLS MANs. The MANs may have multiple connections between them to provide load balancing and/or redundancy. It might be desirable to obtain the redundant connectivity services from multiple providers.

As shown in Figure 2-3, the WAN edge router used for interconnecting MANs plays the role of a P device even though it is a CE for the SP VPN service. It is expected to label switch packets between the MANs across the SP network.

Figure 2-3 MPLS over GRE

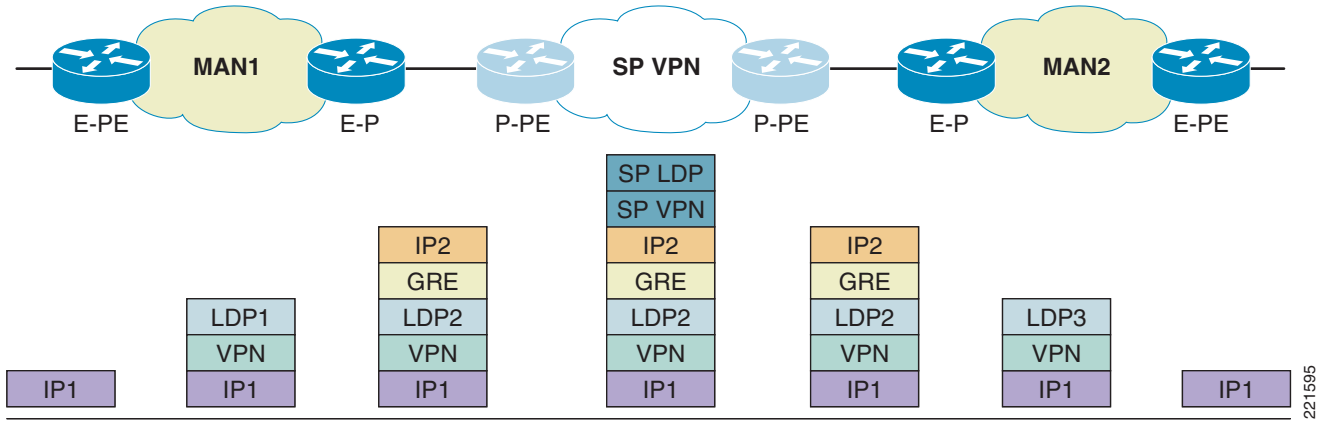


A point-to-point GRE tunnel is set up between each edge router pair (if full mesh is desired). From a control plane perspective, the following are expected to be run within the GRE tunnels:

- IGP such as EIGRP or OSPF for MPLS device reachability (P/PE/RR)
- LDP for label distribution
- MP-iBGP for VPN route/label distribution

Once the route/label distribution is done in the control plane, the enterprise edge device acts like a label switching router (LSR/P) where it treats the GRE interfaces as normal access interfaces. Figure 2-4 shows end-to-end packet flow between campuses across different MANs.

Figure 2-4 MPLS over GRE—Forwarding Plane



As can be seen from the headers, this adds a large amount of overhead to the MTU and hence is not the most desired option. In addition, platform support is also limited to the 7200 and ISRs. Thus none of the high-end platforms (7600, 12000) support it. 7600 supports MPLSoGRE in PE-PE mode only, but ideally we would prefer P-P setup for inter-MAN connectivity. Thus MPLSoGRE is not a preferred option for Inter-MAN connectivity.

Carrier Supporting Carrier (CSC)

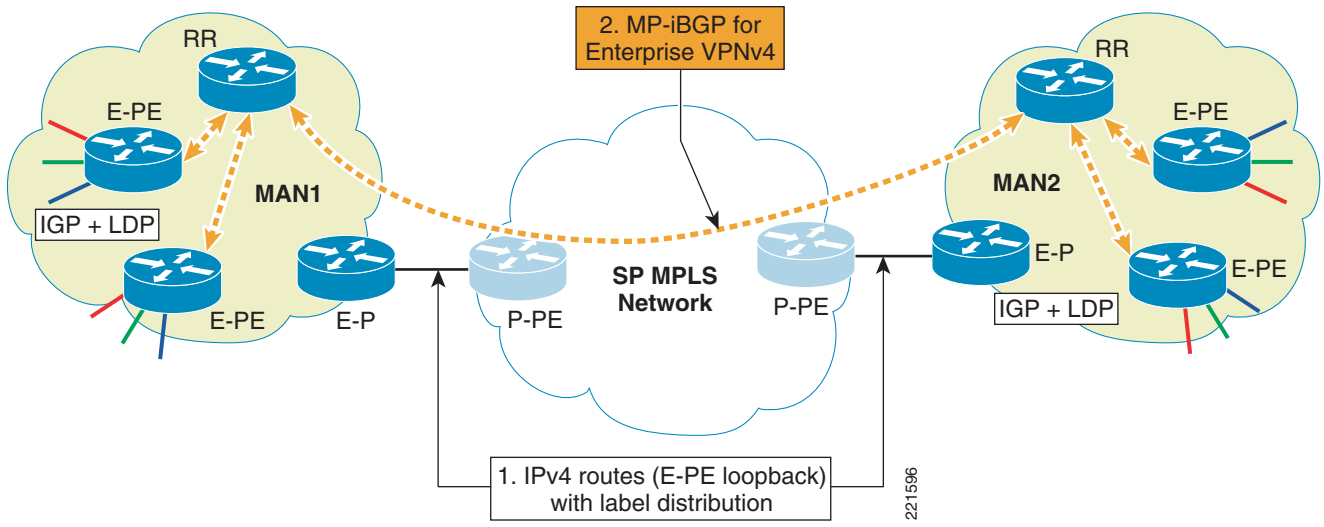
The Carrier Supporting Carrier was developed for MPLS enabled SPs to support other MPLS/VPN SPs. With the growth in enterprise network growth, CSC is a service that SPs can provide to large enterprises as well. For an enterprise, this involves getting a label transport service from a provider. The enterprise edge devices perform the P role in the enterprise MPLS network.

The advantages of a label transport service include the fact that there is no overlay such as GRE required. The SP provides any-to-any connectivity as well, so multiple dedicated links are not required between the MANs. Based on the incoming label, the SP network ensures that the packet is forwarded to the correct MAN location.

From a control plane perspective there are two major elements (Figure 2-5):

1. A IPv4 route and label exchange between the enterprise edge P and the Provider PE—the only routes that needed to be exchanged are for the PE and RR reachability (loopbacks). This can be achieved in two ways:
   a. Running a IGP with the Provider PE to advertise the loopback addresses and running LDP to advertise the labels associated with those loopback addresses.
   b. Alternatively, the SP may prefer to run EBGP+label option to advertise the loopback addresses along with their associated labels.
2. Enterprise running MP-iBGP between its RRs to exchange VPNv4+lable information for its own VPNs.

Figure 2-5 Carrier Supporting Carrier (CSC)

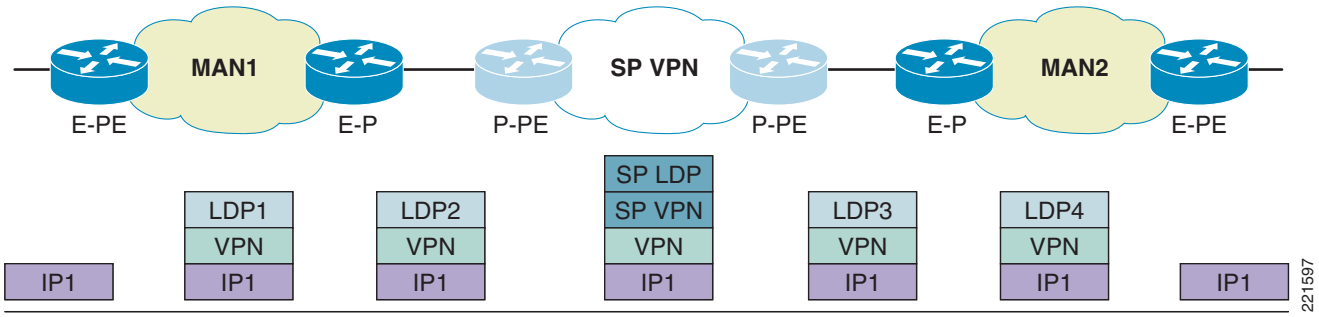


In the forwarding plane, as shown in Figure 2-6, the following occurs:

- The P router at the enterprise edge (E-P1) receives a labeled packet from the MAN and label switches it based on the label learned via the provider for E-P2.
- The provider PE (P-PE1) label switches the incoming top label with the label advertised via P-PE2. This is treated as the "VPN" label within the provider network.
- P-PE1 prepends the top label (SP LDP) for P-PE1 to P-PE2 reachability.
- P-PE2 receives the packet with SP LDP label popped (if PHP is enabled). It label switches the exposed label corresponding to the one advertised by E-P2.

- E-P2 label switches the top label and sends it across the MAN2 network to the appropriate PE.

Figure 2-6 Carrier Supporting Carrier—Forwarding Plane



Depending on the size of the deployments, the platforms can range from 12000, 7600, 7304, 7200, or 3800.

CSC does have two major draw backs:

- There is a very limited offering from the SPs to the enterprises—their CSC service is designed primarily for other SPs. Additionally, since the VPNs are now maintained by the enterprises, the SP essentially sells a single VPN service.
- Since it is essentially a Layer 3 service from a provider, the enterprise may wish to encrypt the traffic, which is not feasible as currently there are no mechanisms that allow for encryption of labelled packets.

WAN Edge

If enterprises have requirements for virtualization of the branches, then the following deployment models are technically available:

- Multi-VPN service from SP
- Carrier Supporting Carrier
- MPLSoL2 infrastructure
- Self-Deployed Multi-VRF with mGRE/DMVPN (DMVPN per VRF)
- MPLS VPN over DMVPN— 2547oDMVPN (Hub and Spoke only)
- MPLS VPN over IP using L2TPv3 (2547oL2TPv3)

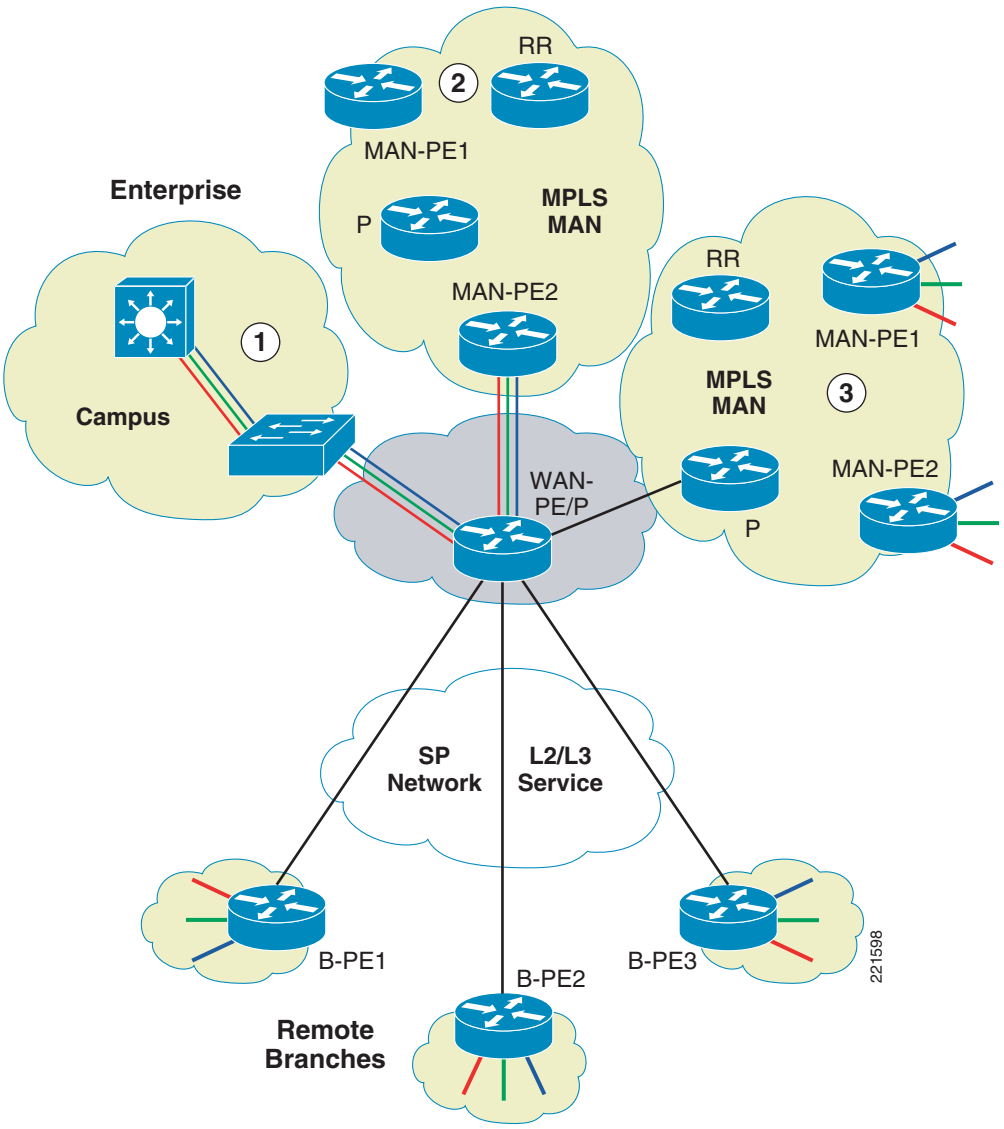
There are scenarios where virtualization may be required at the large branches only. The rest of the branches may not have any VRFs, but may have to be placed in there own VRF, default common VRF, or global table at the headend depending on enterprise requirements.

The deployment models should support integration of branches ranging from 10s to 1000s. The interface types on the headend are expected to range from the DS3 to GE. The speeds on the branches would typically be T1 or below.

WAN Edge Integration

The deployment models listed above allow connectivity between the segmented branches and the headend. The headend itself (as shown inFigure 2-7) may be connected to the rest of the core network in at least three different ways (discussed below). Not every model may support these integration options.

Figure 2-7 Integrating WAN Edge with MPLS MAN



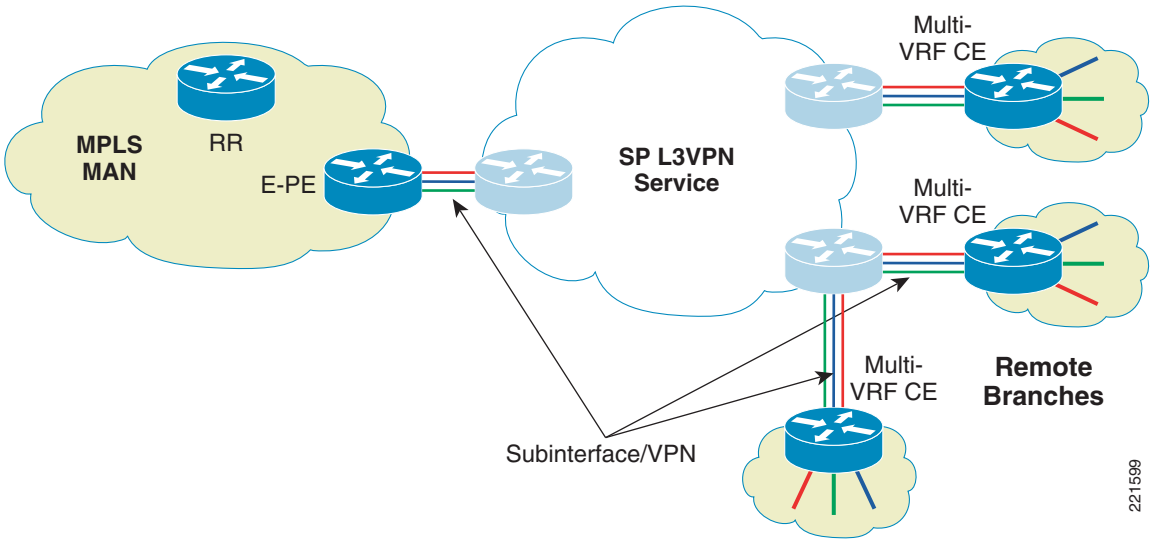
1. Direct connectivity with campus—In this scenario the WAN edge router is running in VRF-lite mode towards the campus where the VRFs from the branches are extended into the campus by using VLANs. The assumption here is that campus is virtualized using a combination of VLANs and VRF-lite as well. A separate routing instance is running per VRF within the campus that is extended to the branches via the WAN edge device. This can be deployed with any of the deployment models above.

2. Back-to-back PEs with MPLS MAN—The WAN edge router is running VRF-lite mode towards the core network in this scenario as well, but instead of connecting into a VRF-lite campus, it connects back-to-back with a MPLS MAN PE. This option is ideal for scenarios where the WAN Edge router may have P capabilities to extend the core MPLS to the branches or if it cannot be directly made into a core PE. This option can be deployed with any of the models above.
3. Direct Connectivity with MPLS MAN—This is the easiest integration option since the WAN edge router performs the P functionality and the branch PEs are integrated directly with the core MPLS network. This is ideal for MPLSoL2 and 2547oDMVPN deployments although its currently not supported fully in the later scenario.

Multi-VPN Service from Provider

A simple solution for enterprises to extend virtualization to the branches is to obtain multiple Layer 3 VPN services from a provider. As shown in Figure 2-8, the branch routers become Multi-VRF CEs and the headend may be a Multi-VRF CE or a PE (if it is directly connected to the MPLS MAN). This may be a desirable solution from a cost perspective if there are a small number of branches that have a requirement for virtualization and the number of VRFs is low.

Figure 2-8 Multi-VPN service from SP



In the control plane each of the CEs run a routing protocol such as OSPF, EIGRP or BGP with the SP PE on per VRF basis. Thus any design recommendations implemented while getting a single VPN service (in terms of routing, QoS, Multicast, etc.) would have to be followed for each of these VPN instances as well.

Carrier Supporting Carrier

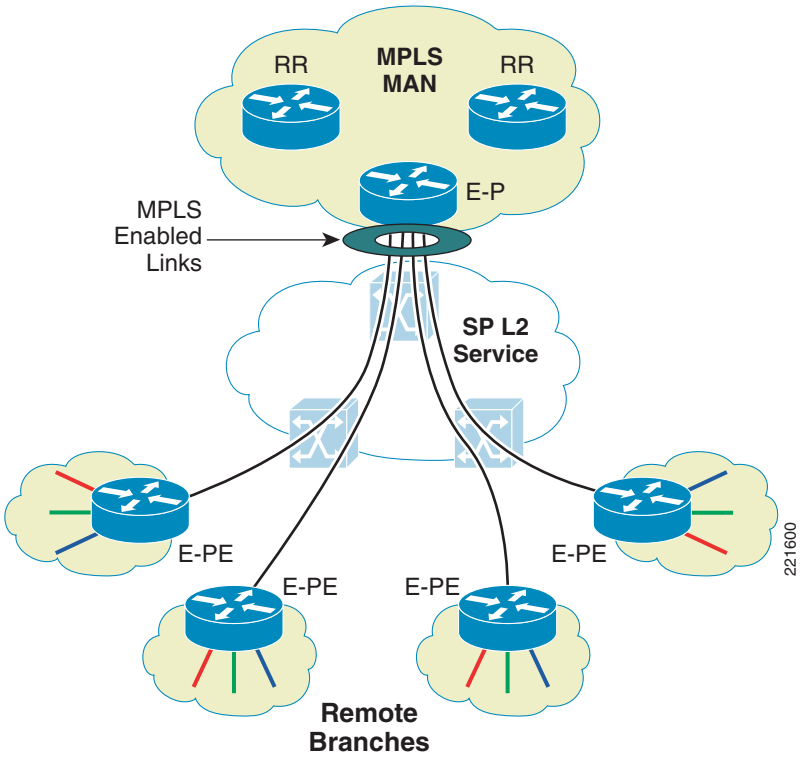
This is same labeled transport service that was discussed in Carrier Supporting Carrier (CSC). The requirement here is limited to branch routers such as ISRs.

MPLSoL2 Service

This model assumes that the enterprise has existing Layer 2 services for connecting branches and wants to enable MPLS over them. Since such Layer 2 connectivity is typically hub and spoke or partial mesh, the MPLS overlay also inherits the same connectivity characteristics. If spoke-to-spoke communication is required, it has to be handled via the hub.

The branch aggregation router is converted into a P role for the MPLS network and is expected to label switch packets as shown in Figure 2-9. The branch routers become PE routers with VRF interfaces facing the branch and MPLS-enabled interface facing the headend.

Figure 2-9 MPLSoL2 Service



In the control plane, each of the remote branch PEs would have a LDP session and a IGP session with the headend aggregator. They would also have MP-iBGP sessions with the route reflectors that would typically reside behind the headend aggregating device.

In cases where virtualization is not required at certain branches, then those branch routers do not need to have their WAN connection MPLS enabled. On the headend, depending on the connectivity model (point-to-point vs. multipoint) and interface flexibility, the enterprise has a few options:

- If using multipoint interface(s) at headend, then separate the MPLS-enabled and the non-MPLS connections into separate multipoint groups. Within the non-MPLS group, they may need to be further separated based on the VRF(s) into which they need to be placed.
- If using point-to-point interfaces, then each individual connection can be MPLS enabled or placed in a VRF.
- A combination of point-to-point and multipoint interfaces can be supported as well.

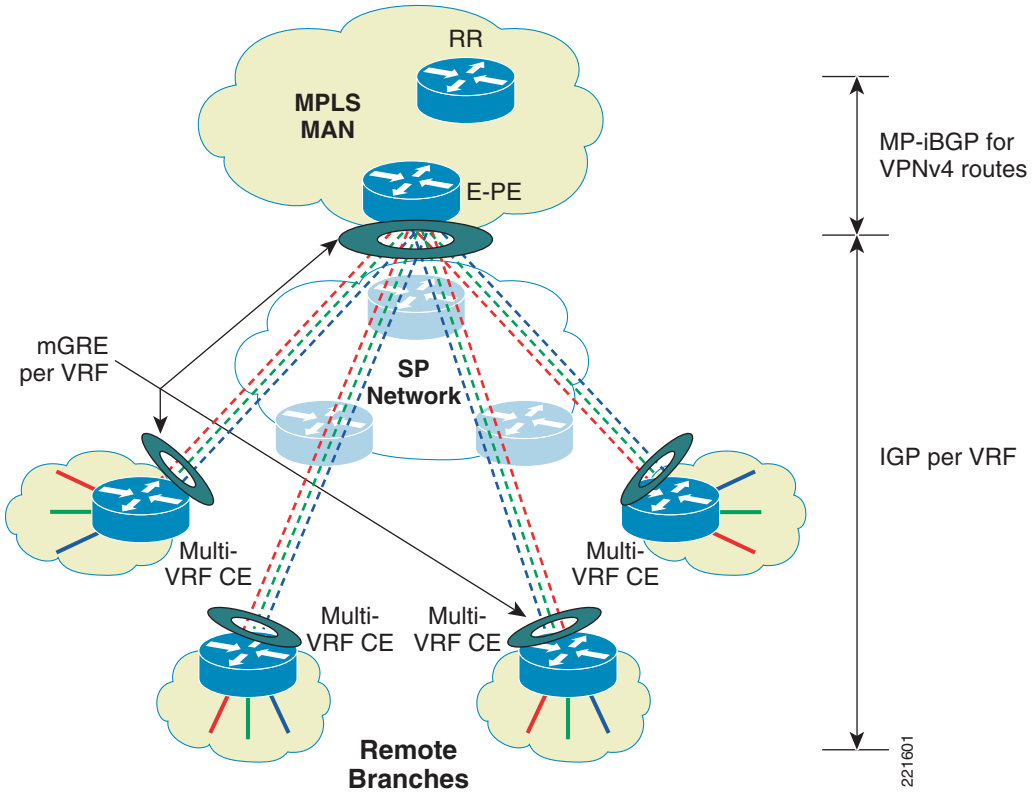
Note In each of the cases, the aggregation device have to perform both the roles—P as well as PE.

DMVPN per VRF

This model can be used over a Layer 2 or Layer 3 service from a provider. If it is a Layer 3 VPN service, then the enterprise purchases only a single VPN from the provider but overlays its own VPNs by using a combination of Multi-VRF and GRE. The headend has a mGRE tunnel per VRF, the branches have either GRE (if no spoke-to-spoke communication is required) or mGRE (if spoke-to-spoke communication is required) tunnels per VRF. DMVPN provides for spoke-to-spoke communication as well as encryption (although encryption is optional for our deployment model). By configuring mGRE on certain spokes, it provides them with the ability to create dynamic tunnels to other spokes (which should be configured with mGRE as well) on a per-VRF basis.

Most enterprises only have a partial mesh requirement—large sites need to be meshed together but the smaller sites are typically only hub and spoke. Thus the deployment is expected to be a combination of GRE and mGRE at the spokes (Figure 2-10).

Figure 2-10 DMVPN per VRF



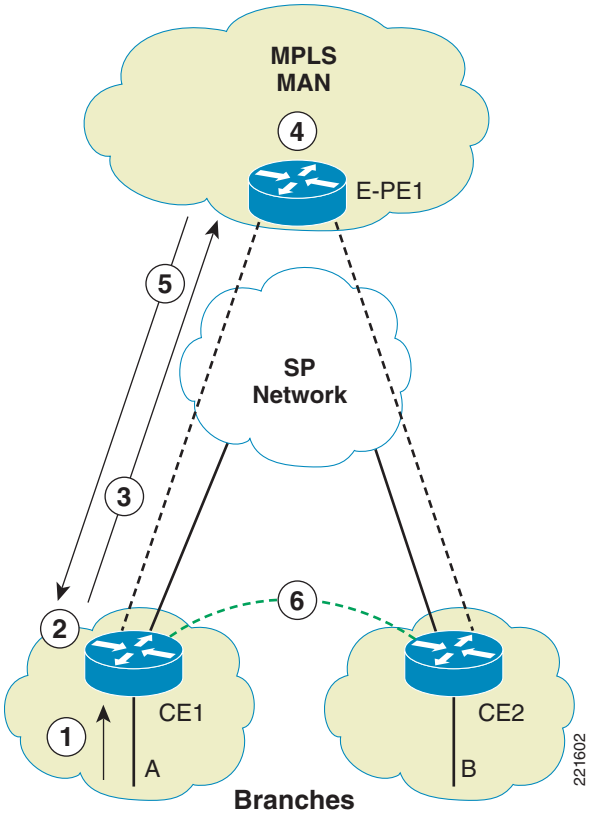
The hub device, while aggregating the branches, is also a PE for the MPLS MAN network. It has a IGP instance running within each VPN with each of the spokes. The IPv4 addresses learned from the spokes are converted to VPNv4 addresses before being advertised to the RRs using MP-iBGP. IGP might limit the scale of the deployment requiring the use of multiple hub routers. Scaling options include:

- Use of BGP as the hub and spoke protocol with dedicated RRs for this purpose

- Having multiple termination devices at the headend based—for example, one for each VRF if there are a low number of VRFs

DMVPN uses NHRP to keep track of the next-hop to physical address mapping. The hub is the NHRP server that maintains the mapping table on per VRF basis. In the example shown in Figure 2-11, once the GRE tunnel is established with the hub, both Branch1 (CE1) and Branch2 (CE2) register with the hub (E-PE1) using NHRP on per VRF basis. The hub learns about each of the branch VPN routes and advertises them back out to the other branches.

Figure 2-11 DMVPN per VRF—Spoke-to-Spoke Communication

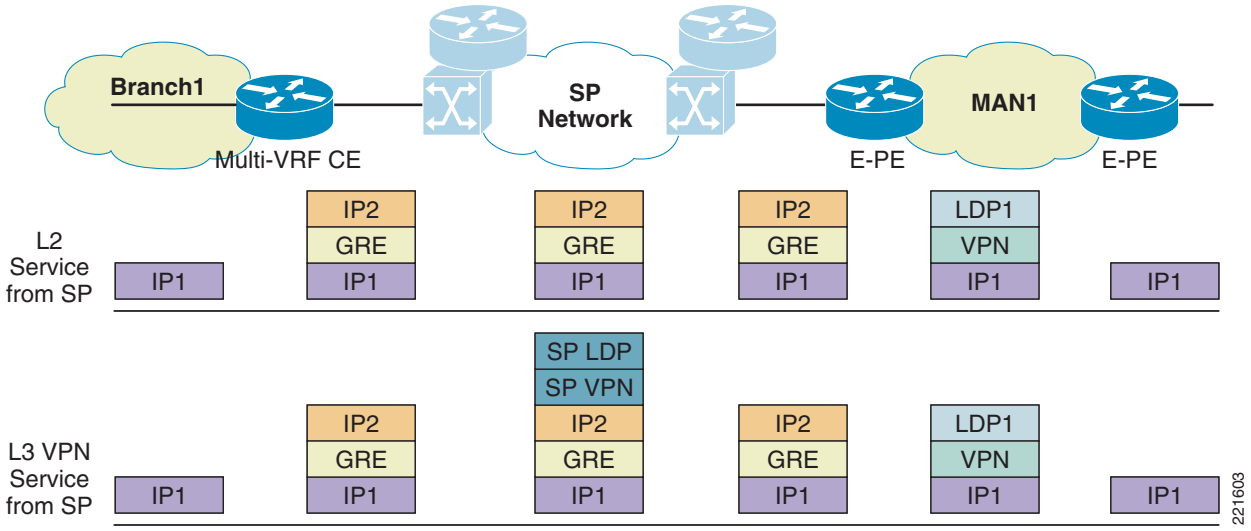


As shown in Figure 2-11, in the forwarding plane, the following sequence occurs on a per-VRF basis:

1. A sends packets to CE1 destined for B.
2. CE1 looks up its VRF table and finds B with next hop of CE2's GRE tunnel address.
3. CE1 sends a NHRP query to E-PE1 to resolve the next-hop address.
4. E-PE1 looks up the per-VRF NHRP database and associates CE2's tunnel address with its physical interface address.
5. E-PE1 sends the NHRP response back to CE1 with CE2's physical interface address.
6. CE1 sets up the dynamic GRE tunnel with CE2.

Figure 2-12 shows the end-to-end packet encapsulation when using this model. The packets from the remote branches are encapsulated in GRE on a per-VRF basis and forwarded across the SP network. The enterprise PE at the hub site decapsulates the GRE headers and performs label pushing based on the VRF in which the GRE interface is configured. The packets are then forwarded as normal MPLS VPN packets across the MAN (with the VPN and the LDP label).

Figure 2-12 DMVPN per VRF—Forwarding Plane



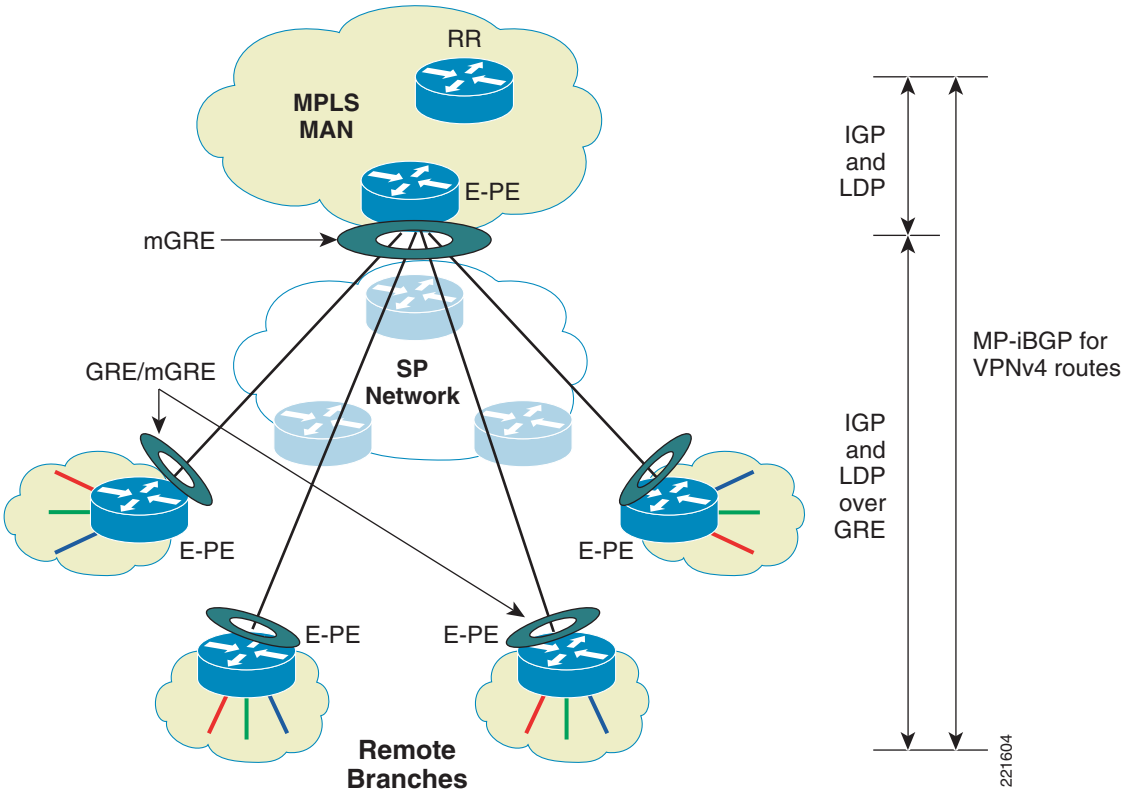
The hub site PE can be a 7600 or a 7200 depending on the scale or encryption requirements. The spokes can be aa ISR such as 2800 and 3800 depending on the branch requirements in terms of number of VRFs, throughput, and other non-transport related features.

MPLS VPN over DMVPN—2547oDMVPN (Hub & Spoke Only)

This model does not have some of the scale limitations of the Multi-VRF based solutions because the GRE tunnels are created outside the VRFs and hence a single tunnel can be shared for transporting many VRFs. The hub is configured with a single mGRE tunnel while spokes have a single GRE tunnel.

Note This is designed to be used for hub and spoke communication only and currently the dynamically created spoke-to-spoke tunnels are not supported.

Figure 2-13 2547oDMVPN (Hub & Spoke Only)



As shown in Figure 2-13, in the control plane the following protocols exist:

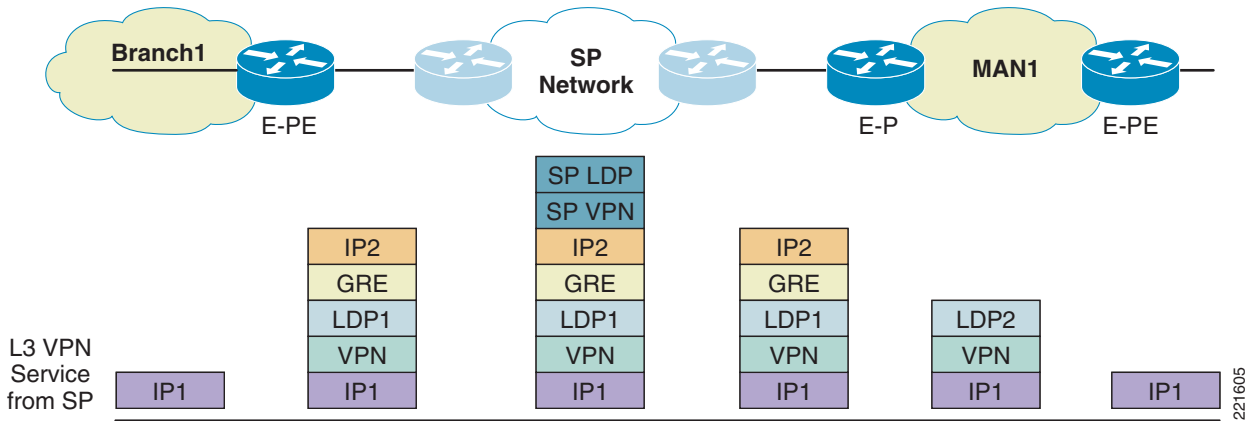
- Routing protocol of the provider to learn the Branch and headend's physical interface addresses (tunnel source address). Statics could be used as well if these are easily summarizable.
- GRE tunnel between the branch PE and the headend P.
- IGP running in the enterprise global space over the GRE tunnel to learn remote PE's and RR's loopback address.
- LDP session over the GRE tunnel with label allocation/advertisement for the GRE tunnel address by the branch router.
- MP-iBGP session with RR, where the branch router's BGP source address is the tunnel interface address—this forces the BGP next-hop lookup for the VPN route to be associated with the tunnel interface.

Additionally, IPsec can be used to encrypt the GRE tunnels; encryption happens after the GRE encapsulation.

Hub as a P Router

As shown in Figure 2-14, the branch router attaches the appropriate VPN label for the destination along with the LDP label advertised by the hub P for the destination next-hop address. It then encapsulates the labeled packet in a GRE tunnel with the hub P as the destination before sending it to the provider. Since in this example SP is providing Layer 3 VPN service, it further prepends its own VPN and LDP labels for transport within its network. The hub P receives a GRE encapsulated labeled packet. It decapsulated the tunnel headers before label switching it out to the appropriate outgoing interface in the MPLS MAN for the packet to reach the eventual PE destination.

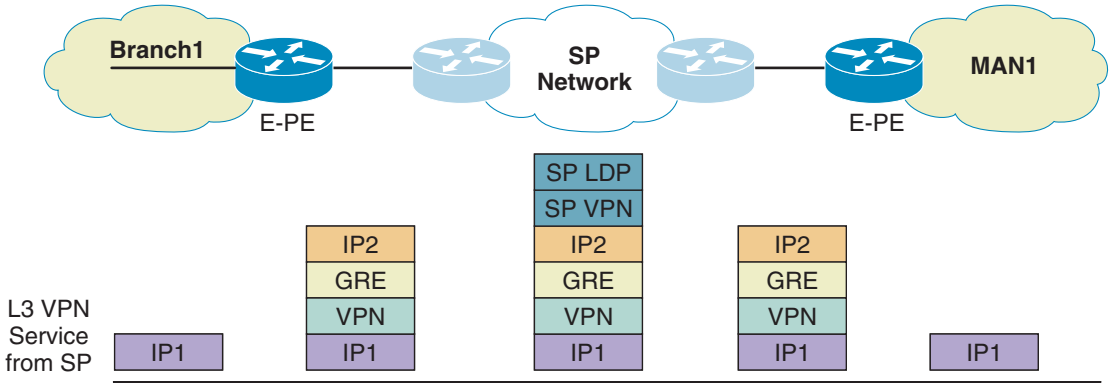
Figure 2-14 2547oDMVPN—Forwarding Plane with Hub as a P



Hub as a PE Router

As shown in Figure 2-15, the branch router attaches the appropriate VPN label for the destination advertised by the hub PE router. It then encapsulates the labeled packet in a GRE tunnel with the hub PE as the destination before sending it to the provider. Since in this example SP is providing Layer 3 VPN service, it further prepends its own VPN and LDP labels for transport within its network. The hub PE receives a GRE encapsulated labeled packet. It decapsulated the tunnel headers before forwarding it out to the appropriate outgoing interface based on the VPN label information and the VRF routing table.

Figure 2-15 2547oDMVPN—Forwarding Plane with Hub as a PE



MPLS VPN Over IP Using L2TPv3—2547oL2TPv3

This is currently not supported on the relevant platforms (7600, ISRs) and hence is not discussed, but is listed for the sake of completeness.

The rest of the guide focuses on providing design and implementation guidelines for the following deployment model for WAN Core:

- MPLSoL2 Service

The rest of the guide also focuses on providing design and implementation guidelines for the following deployment model for WAN Edge:

- MPLSoL2 Infrastructure
- DMVPN per VRF
- 2547oDMVPN