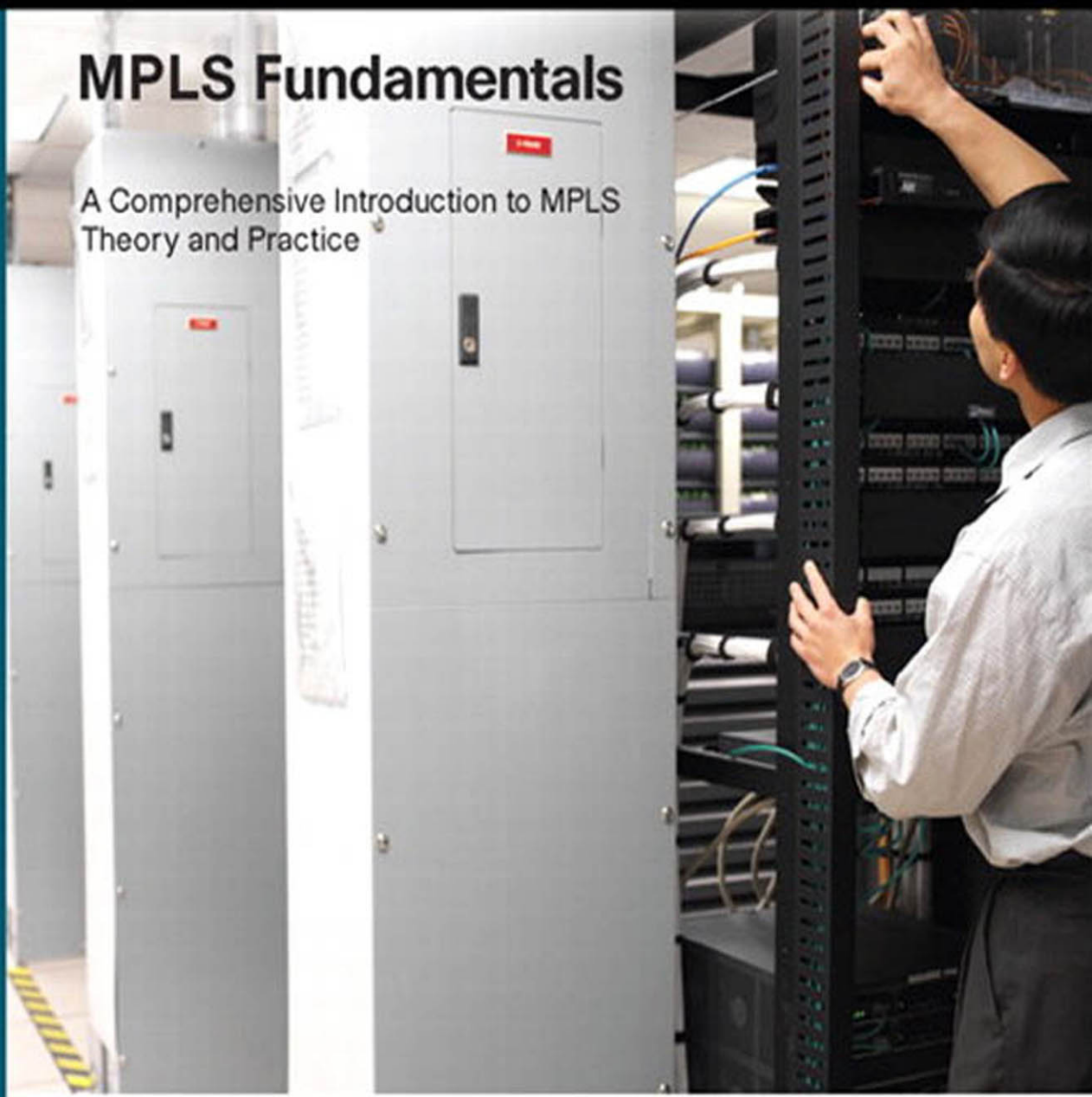


MPLS Fundamentals

A Comprehensive Introduction to MPLS Theory and Practice

ciscopress.com

Luc De Ghein, CCIE No.® 1897

MPLS Fundamentals

Luc De Ghein, CCIE
No. 1897

Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

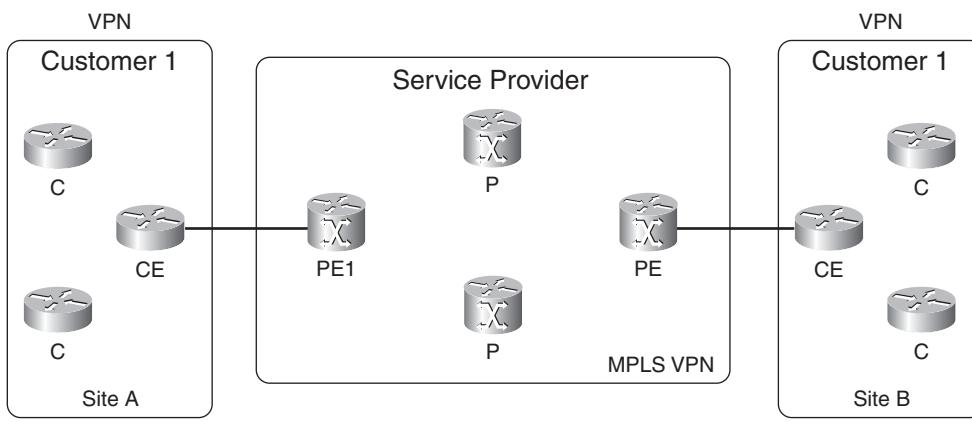
VPN Models

VPNs did exist before the arrival of MPLS. Most popular were Frame Relay or ATM technologies, providing VPN service at Layer 2. The provider had a Frame Relay or ATM backbone and supplied Layer 2 connectivity to the customer routers. This was commonly referred to as the *overlay model*. The service provider might have actually owned or managed the edge routers that were connected to the customer network. The point is that the routers were physically at the customer premises. Refer to the section "Peer-to-Peer VPN Model Versus Overlay VPN Model" in Chapter 1 for more information on this. Peer-to-peer VPN networks existed, but they were not popular. The main reason is that they were not easy to deploy and maintain because they needed distribute lists, IP packet filters, or GRE tunnels. As explained in Chapter 1, MPLS VPN is an example of a highly scalable peer-to-peer VPN model.

MPLS VPN Model

It is important to become familiar with the terminology concerning MPLS VPN. Look at Figure 7-1 for a schematic overview of the MPLS VPN model. A service provider is providing the common public infrastructure that customers use.

Figure 7-1 *MPLS VPN Schematic Overview*



A PE router is a provider edge (PE) router. It has a direct connection with the customer edge (CE) router at Layer 3. A provider (P) router is a router without the direct connection to the routers of the customer. In the MPLS VPN implementation, both P and PE routers run MPLS. This means that they must be able to distribute labels between them and forward labeled packets.

A CE router has a direct Layer 3 connection with the PE router. A customer (C) router is a router without a direct connection with the PE router. A CE router does not need to run MPLS.

Because the CE and PE routers interact at Layer 3, they must run a routing protocol (or static routing) between them. The CE router has only one peer outside of its own site: the PE router. If the CE router is multihomed, it can peer with multiple PE routers. The CE router does not peer with any of the CE routers from the other sites across the service provider network, as with the overlay model. The name *peer-to-peer model* is derived from the fact that the CE and PE form a peer at Layer 3.

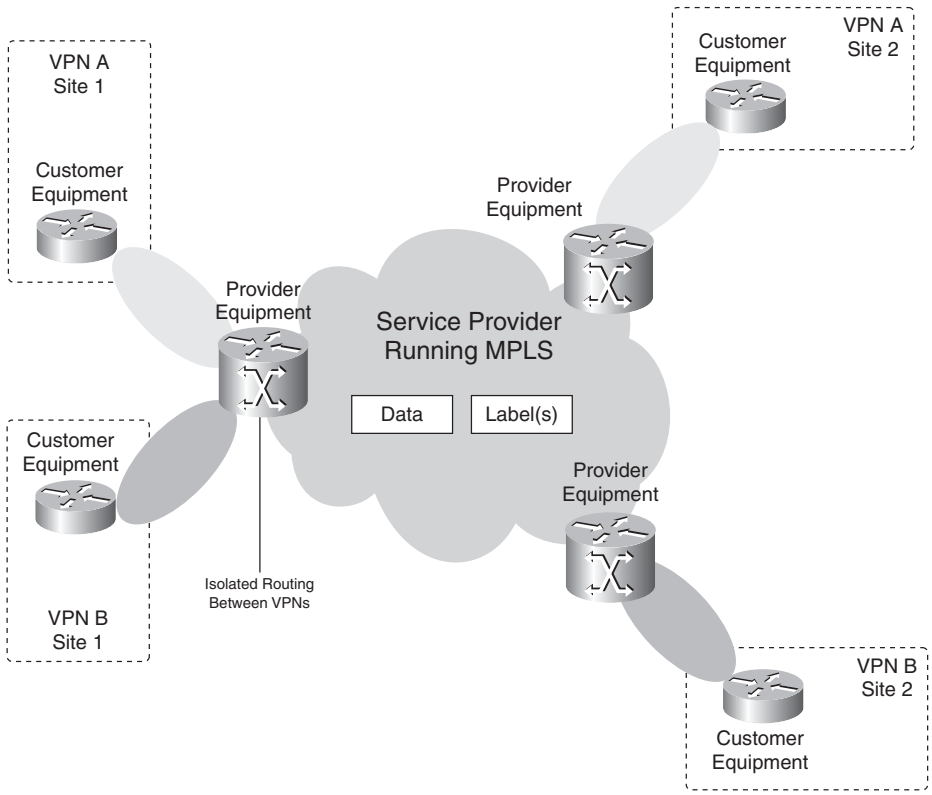
The *P* in VPN stands for private. As such, the customers of the service provider are allowed to have their own IP addressing scheme. This means that they can use registered IP addresses but also private IP addresses (see RFC 1918) or even IP addresses that are also used by other customers who are connecting to the same service provider (referred to as *overlapping IP addressing*). If the packets were to be forwarded as IP packets within the service provider network, this would cause problems, because the P routers would be confused. If the private and overlapping IP addressing scheme is not allowed, then every customer must be using a unique address range. In that case, the packets can be forwarded by looking up the destination IP address on every router in the service provider network. This means that all P and PE routers must have the complete routing table of every customer. This would be a large routing table. The only routing protocol that is capable of carrying a large number of routes is Border Gateway Protocol (BGP). This would mean that all P and PE routers would have to run internal BGP (iBGP) among them. However, this is not a VPN scheme, because it is not private to the customers.

Another solution is that every P and PE router has a private routing table for each customer. Several processes of one routing protocol (one process per VPN) could be running on all the routers to distribute the VPN routes. Running one routing process per VPN on every P router is not very scalable. Each time one VPN is added to the network, a new routing process must be added to every P router. Furthermore, if an IP packet enters a P router, how does the P router determine which VPN the packet belongs to to figure out which private routing table to use to forward the packet? If the packet is an IP packet, this is not possible. You could add one extra field to the IP packet indicating which VPN the IP packet belongs to. The P routers could then forward the IP packets by looking at this extra field and at the destination IP address. Again, all P routers would have to be aware of this extra field.

A scalable solution would be to have the P routers completely unaware of the VPNs. Then the P routers would not be burdened with having routing information for VPN routes. Can you achieve this by using MPLS? The answer is yes. The customer IP packets are labeled in the service provider network to achieve a private VPN for each customer. Furthermore, the P routers no longer need to have the routing table of the customers by using two MPLS labels. Therefore, BGP is not needed on the P routers. See the section "BGP-Free Core" in Chapter 1 for an explanation on this. The VPN routes are only known on the PE routers. As such, the VPN knowledge is present only on the edge routers of the MPLS VPN network, which makes the MPLS VPN solution scalable.

Figure 7-2 shows the MPLS VPN model: label switching packets in the service provider network and PE routers that are VPN aware.

Figure 7-2 *MPLS VPN Model*



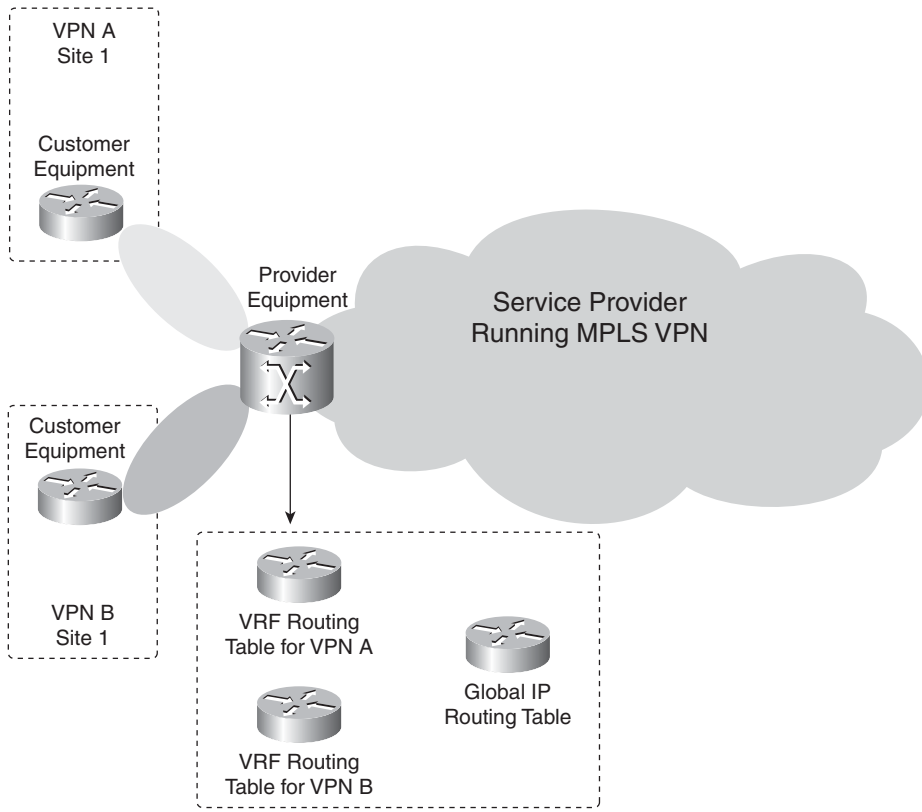
Architectural Overview of MPLS VPN

To achieve MPLS VPN, you need some basic building blocks on the PE routers. These building blocks are the following: VRF, route distinguisher (RD), route targets (RT), route propagation through MP-BGP, and forwarding of labeled packets.

Virtual Routing Forwarding

A virtual routing/forwarding (VRF) is a VPN routing and forwarding instance. It is the name for the combination of the VPN routing table, the VRF Cisco Express Forwarding (CEF) table, and the associated IP routing protocols on the PE router. A PE router has a VRF instance for each attached VPN. Look at Figure 7-3 to see that a PE router holds the global IP routing table, but also a VRF routing table per VPN connected to the PE.

Figure 7-3 *VRFs on a PE Router*



Because the routing should be separate and private for each customer (VPN) on a PE router, each VPN should have its own routing table. This private routing table is called the VRF routing table. The interface on the PE router toward the CE router can belong to only one VRF. As such, all IP packets received on the VRF interface are unambiguously identified as belonging to that VRF. Because there is a separate routing table per VPN, there is a separate CEF table per VPN to forward these packets on the PE router. This is the VRF CEF table. As with the global routing table and the global CEF table, the VRF CEF table is derived from the VRF routing table.

You create the VRF on the PE router with the **ip vrf** command. You use the **ip vrf forwarding** command to assign PE-CE interfaces on the PE router to a VRF. You can assign an interface to only one VRF, but you can assign several interfaces to the same VRF. The PE router then automatically creates a VRF routing table and CEF table. The VRF routing table does not differ

from a regular routing table in Cisco IOS other than that it is used for a set of VPN sites only and is completely separated from all other routing tables. The routing table as you have known it to this point will now be referred to as the *global* or the *default routing table*. Look at Example 7-1, where the VRF configured is the VRF *cust-one*.

Example 7-1 *Configuring a VRF*

```
!
ip vrf cust-one
 rd 1:1
 route-target export 1:1
 route-target import 1:1
!
interface Serial5/1
 ip vrf forwarding cust-one
 ip address 10.10.4.1 255.255.255.0
!

sydney#show ip route vrf cust-one

Routing Table: cust-one
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

     10.0.0.0/8 is variably subnetted, 9 subnets, 2 masks
B       10.10.2.0/24 [200/0] via 10.200.254.2, 00:31:04
C       10.10.4.0/24 is directly connected, Serial5/1
C       10.10.4.2/32 is directly connected, Serial5/1
B       10.10.100.1/32 [200/1] via 10.200.254.2, 00:31:04
B       10.10.100.3/32 [20/0] via 10.10.4.2, 00:13:29

sydney#show ip cef vrf cust-one
Prefix               Next Hop             Interface
0.0.0.0/0            no route
0.0.0.0/32           receive
10.10.2.0/24         10.200.214.1         POS0/1/0
10.10.4.0/24         attached             Serial5/1
10.10.4.0/32         receive
10.10.4.1/32         receive
10.10.4.2/32         attached             Serial5/1
10.10.4.255/32       receive
10.10.100.1/32       10.200.214.1         POS0/1/0
```

Example 7-1 *Configuring a VRF (Continued)*

```
10.10.100.3/32        10.10.4.2           Serial5/1
224.0.0.0/4           drop
224.0.0.0/24          receive
255.255.255.255/32    receive
```

The VRF routing table *cust-one* has prefixes in it that are populated by dynamic routing protocols and static routing, just like the global routing table. The concept of metrics, distance, next hop, and so on does not change. Because the VRF instance is associated with interfaces, only IP packets that are entering the PE router via those VRF interfaces are forwarded according to that VRF CEF table.

NOTE In Cisco IOS, CEF is the only switching method supported for forwarding IP packets from the VRF interface. As such, CEF must be enabled globally on all PE routers and all VRF interfaces.

RD

The VPN prefixes are propagated across the MPLS VPN network by Multiprotocol BGP (MP-BGP). The problem is that when BGP carries these IPv4 prefixes across the service provider network, they must be unique. If the customers had overlapping IP addressing, the routing would be wrong. To solve this problem, the concept of RDs was conceived to make IPv4 prefixes unique. The basic idea is that each prefix from each customer receives a unique identifier (the RD) to distinguish the same prefix from different customers. A prefix derived from the combination of the IPv4 prefix and the RD is called a vpnv4 prefix. MP-BGP needs to carry these vpnv4 prefixes between the PE routers.

NOTE MP-BGP is explained in the section "BGP" later in this chapter.

An RD is a 64-bit field used to make the VRF prefixes unique when MP-BGP carries them. The RD does not indicate which VRF the prefix belongs to. The function of the RD is not that of a VPN identifier, because some more complex VPN scenarios might require more than one RD per VPN. Each VRF instance on the PE router must have one RD assigned to it. This 64-bit value can have two formats: *ASN:nn* or *IP-address:nn*, where *nn* represents a number. The most commonly used format is *ASN:nn*, where *ASN* stands for autonomous system number. Usually, the service provider uses *ASN:nn*, where *ASN* is the autonomous system number that the Internet Assigned Numbers Authority (IANA) assigns to the service provider and nn is the number that the service provider uniquely assigns to the VRF. The RD does not impose semantics; it is just used to uniquely identify the VPN routes. This is needed because the IPv4 routes from one customer might be overlapping with the IPv4 routes from another. The combination of the RD with the IPv4 prefix provides a vpnv4 prefix, of which the address is 96 bits long. The mask is 32 bits long, just as it is for an IPv4