



Tutorial

L3VPNs

by Galina Pildush

Introduction to VPNs

[Eyewitness Testimony!](#)

Overview of VPNs

[Customer-Provisioned VPNs](#)

[Provider-Provisioned VPNs](#)

[MPLS and VPNs](#)

[Private Address Space Overlap: An Extranet Challenge](#)

[IETF Work](#)

VPN Topology Support

[Virtual Private Lines](#)

[Hub and Spoke Topologies](#)

[Fully Meshed Topologies](#)

RFC 2547bis Terminology That Has Become Common to VPNs

[Site versus Customer Premises](#)

[CE Routers](#)

[PE Routers](#)

[P Routers](#)

[Semipackets?](#)

L3 PPVPN Architectures

[VPN-IPv4 Address Structure](#)

[VRFs](#)

[Associating IP Packets with VRFs](#)

[Route Targets](#)

[Route Targets Formally Defined](#)

[Export and Import Targets](#)

[Why Are the Targets Called "Export" or "Import"?](#)

[Route Origin](#)

RFC 2547bis Operation

[BGP Signaling Plane](#)

[CE-PE Routing Information Exchange](#)

[PE-PE Routing Information Exchange](#)

[LSP Establishment](#)

[Customer Data Forwarding Plane](#)

[Provider Data Forwarding Plane](#)

RFC 2547bis Network Design

[Membership](#)

[PE Locations](#)

[Defining VRFs](#)

[Special Handling within RFC 2547bis](#)

[RFC 2547bis Scalability](#)

[Select Special Features of Cisco's L3VPN Implementation](#)

RFC 2547bis Basic Configuration

[Configuring MPLS within the Provider's Network](#)

[Defining VPNs](#)

[Configuring CE-PE Routing Information Exchange](#)

[Configuring PE-PE Routing Information Exchange](#)

Cisco IOS L3VPN Configuration Enhancements

[Configuring Complex Topologies](#)

[Recommended Troubleshooting Steps and Commands](#)

[Configuration Example](#)

[Conclusion](#)

[References](#)

Introduction to VPNs

The idea of multiple customers using a shared carrier infrastructure goes back to the beginnings of telephony; through X.25, Frame Relay, and ATM; and now to the concept of virtual private networks (VPNs). According to the IETF Provider-Provisioned VPN group, VPNs specifically run over IP or over designated sub-IP transports such as MPLS. While a Layer 2 VPN (L2VPN) may present exactly the same interface to the end user as Frame Relay, what distinguishes a true VPN from classic FR is that the L2VPN runs over IP or MPLS, not time-division multiplexed (TDM) transport.

Eyewitness Testimony!

Howard Berkowitz attended the VPN Birds of a Feather (BOF) meeting at the Orlando IETF meeting in 1988. Under the IETF rules, a BOF can meet twice before it must come up with a charter for a Working Group, or the WG cannot form. The BOF couldn't come up with a charter in Orlando, principally because telecommunications carriers did not want to exclude Frame Relay and ATM services they already offered from the definition of a VPN. They were concerned -- not without reason -- that if their services were not included, competitors would sell "new technology" against their working services.

It took several years before the IETF was able to form a consensus definition of VPNs. In Howard's opinion, several things had to happen before this was politically feasible. MPLS had to become well established as a "sub-IP" technology that still could use packets; the concept of the L2VPN had to be introduced as something telephone companies could offer with a FR or ATM interface; and it had to become a consensus that while TDM-based FR and ATM are solid technologies, they still had limitations and had basically reached their limits of major enhancement.

This Tutorial will familiarize you with the concept of provider-provisioned VPNs (PPVPNs) and their variations. Most of our emphasis will be on Cisco-supported variations, although there will be some references to various IETF proposals that Cisco does not yet support. By comparing some of these proposals to the Cisco approach, you can gain another level of insight into the Cisco implementation.

Overview of VPNs

In principle, a VPN is a private network that has been constructed over a shared public IP or sub-IP infrastructure. It is called *virtual* because (1) it does not require separate dedicated circuits between various locations and (2) it is based on the logical as opposed to physical separation of the facilities. It is called *private* because users of the network can maintain their own addressing and routing schemes, fully independent of and transparent to other customers.

The applicability of VPNs is enormous. Networks can join together various offices, customers and suppliers, or agents and corporate infrastructures. Figure 1 illustrates an example of such interconnectivity.

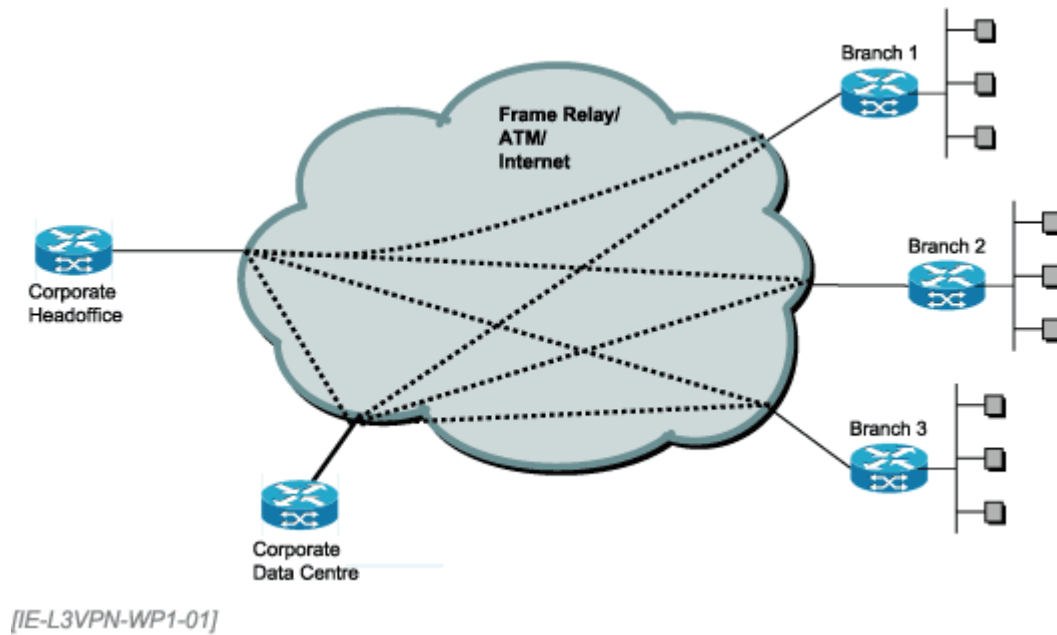


Figure 1. VPNs and Their Role

VPNs based on Frame Relay and ATM have proven to the market that customers can achieve connectivity using relatively secure bandwidth-guaranteed and reliable networks at a reasonable price.

The limitations of these networks lie in their ability to scale. Since the majority of the VPNs were based on PVC-built clouds, adding a site to a fully meshed architecture was and still is a major ordeal, very labor-intensive and error-prone. Just imagine adding a 101st site to the customer's VPN. This would require reprovisioning all the existing 100 sites!

The Internet has become a global connection transport at a reasonable price to corporate and private users. Today multiple corporations can extend their services using this public infrastructure at a reasonable cost for their own offices' interconnections as well as for their customers. VPNs based on the Internet offer corporations the same ability to interconnect as Frame Relay and ATM, coupled with the dynamics of virtual link setup and ease of management. Internet service providers (ISPs) can now offer not only traditional Internet service, but also value-added VPN services, thus generating additional revenue.

Customer-Provisioned VPNs

Customer-provisioned VPNs (CPVPNs) rely on customer equipment and provisioning for VPN management. Examples of CPVPNs include the well-known Layer 2 Tunneling Protocol (L2TP), IPSec, and Point-to-Point Tunneling Protocol (PPTP) models. L2TP is defined in [RFC 2661] and uses UDP for its transport. PPTP, on the other hand, uses TCP to transport PPP. IPSec [RFC 2401] uses authentication and encryption to tunnel the private IP traffic over an IP backbone. Although IPSec provides very strong security, the management requirements and intersite routing responsibilities are burdensome to customers. The provider receives the IP packets from customers and treats them like regular IP packets.

Provider-Provisioned VPNs

With provider-provisioned VPNs (PPVPNs), the provider's equipment is involved in VPN creation and management.

Some examples of PPVPNs include L3VPNs and L2VPNs. The focus of this Tutorial is L3VPNs. There are two basic architectures for VPNs; one based on virtual routers and the other on BGP-MPLS interaction. Cisco introduced the latter type in [RFC 2547].

Cisco does not support a VR model at this time; for details of VR, see [Knight 2003]. In general, the more telephony-oriented vendors such as Lucent and Nortel tend to prefer VR models. Juniper supports both the VR and the RFC 2547 models on various platforms.

The prime differences between L3VPNs and L2VPNs are as follows:

- L3VPN implementation requires providers to participate in customer's Layer 3 routing, while L2VPN implementation does not require this. Let's be clear about what we mean by *participation*.

RFC 2547 PEs don't pass routing updates, other than passively between the true customer routers (CEs). In VRs, however, the provider VR instance (PE) in the PE interacts with the routing protocol. The interaction is CE-PE-CE rather than CE-CE.

In other words, although RFC 2547 PE "knows" about customer routes, the presence of the RFC 2547 router is invisible to the customer. In VR schemes, the VR instance is visible to customer routers.

- L3VPNs require provider routers to manage and accept customers' routes, while L2VPN implementations result in transparency of customer routes to the provider.

The latest version championed by Cisco has evolved in the IETF, and is now a draft sometimes referred to as RFC 2547bis. The whole challenge of IP VPNs is to provide the techniques for customer routes to be exchanged between various sites without creating a routing conflict between various customers.

IETF Work

Currently there are no standards-track RFCs available for the L3VPNs. You can find the current drafts at <http://www.ietf.org/html.charters/ppvpn-charter.html>. The PPVPN Working Group within the IETF Sub-IP area is responsible for PPVPN standardization. New L3VPN and L2VPN Working Groups in the same subarea are responsible for being sure the PPVPN protocols are interoperable and architecturally consistent.

VPN Topology Support

L3VPN subscribers include two types of customers: those who are using Frame Relay and ATM

MPLS and VPNs

VPNs are mysterious because the meaning of the term keeps changing. The concept of VPNs has been around for a long time, even before the Internet took over the world. Carriers offered VPN services to enterprise customers as a replacement for dedicated trunk networks in voice communications. In the 1990s, Frame Relay and ATM services marched in, offering VPN data networks and replacing some of the dedicated leased lines.

Now that IP internetworking has spread all over the world, it makes sense to extend the VPN service concept into IP. A VPN based on IP is a bit of a challenge because all legacy VPN networks are connection-oriented and allow some form of service reservation. MPLS is the tool that allows IP-based VPNs with traffic engineering (TE) to exist. When TE is not a requirement but security is important, the VPN can be built over IPsec [Rosen 2003b]. Yet another alternative to MPLS between the VPN-aware provider routers, when neither TE nor security is required, is GRE tunnels or regular IP [Rekhter 2003].

Private Address Space Overlap: An Extranet Challenge

It is a well-known fact that many users abuse private network address 10/8. The term "abuse" refers to the fact that 10/8 is used by companies regardless of whether or not they require a Class A IP address. It is simply taken for granted that should there be a requirement to use a private address, it must be 10/8. Well, imagine interconnecting various customers who use 10/8, through the provider's IP cloud, without performing NAT. Quite a challenge, isn't it?

networks, and those who need to interconnect their new sites via an L3VPN using IP. When analyzing the interconnectivity architecture for the customers that wish to migrate from legacy VPNs to IP-based VPNs, the majority have had either fully meshed or hub-and-spoke topology structures, which totally depend on traffic flow and costs.

Let's look more closely at these two architectures to understand the impact on L3VPN implementations.

Virtual Private Lines

Provider VPNs often are as simple as a private line replacement, although the newer technologies allow more complex topologies.

Hub and Spoke Topologies

Hub-and-spoke topologies have been applied in the legacy forms of VPNs, such as Frame Relay and ATM, for reasons of cost. Typically carriers charge Frame Relay/ATM customers by the number of PVCs. To reduce the service cost, many customers implement hub-and-spoke topologies, choosing the central point as a hub (for example, the head office or a data center) and other locations as spokes. This, of course, would require any interspoke traffic to flow through the central node -- the hub.

Figure 2 illustrates the typical hub-and-spoke topology.

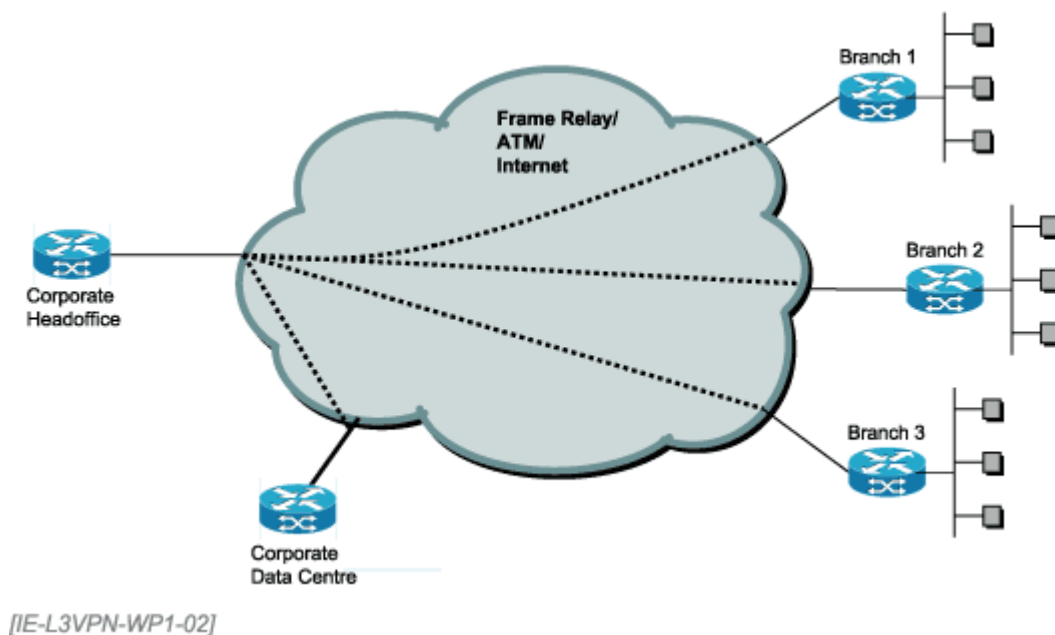


Figure 2. Hub-and-Spoke VPN Example

L3VPNs provide support for the hub-and-spoke topologies, should there be such a requirement.

Fully Meshed Topologies

Fully meshed topologies over Frame Relay and ATM have been rare. The main reason for this is the amount of manual labor involved. Both Frame Relay and ATM implementations typically were based on PVCs. Fully meshed topologies for N nodes require

$$N \times (N - 1) / 2$$

PVCs. Since PVCs have to be set up manually, any changes to the network required many coordination and configuration changes. This resulted in a limited number of fully meshed networks. An example of the fully meshed topology is illustrated in Figure 3.

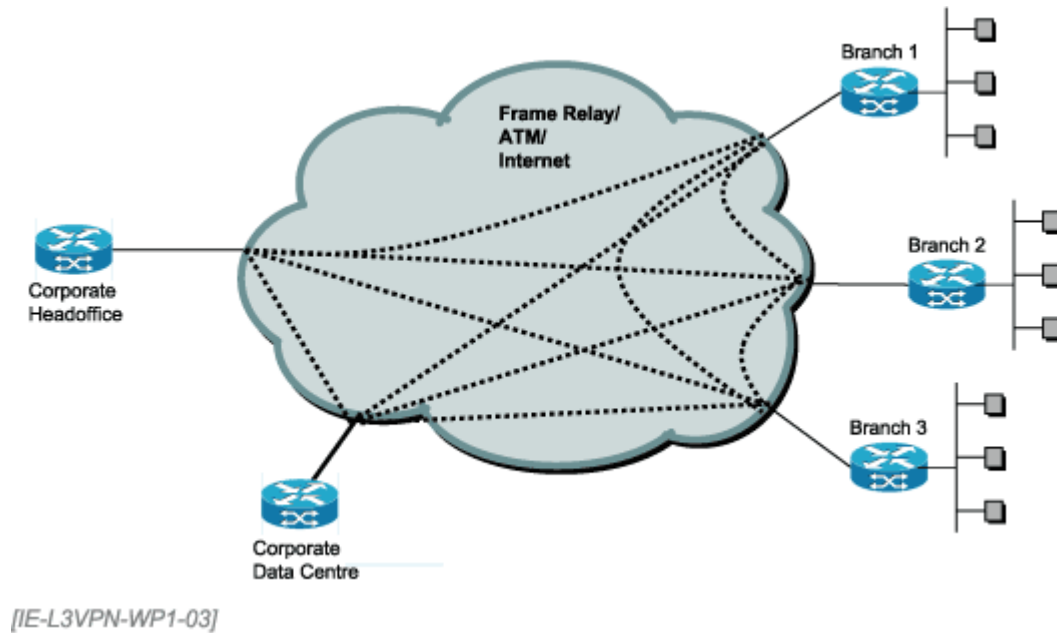


Figure 3. Fully Meshed VPN Example

L3VPNs can deliver fully meshed topologies without manual hassle because all the LSPs are set up dynamically.

RFC 2547bis Terminology That Has Become Common to VPNs

Prior to examining the operation of L3VPNs, you need to understand the terminology used, and the purpose of each of the components of the RFC 2547bis architecture. Using Figure 4, let's examine RFC 2547bis components.

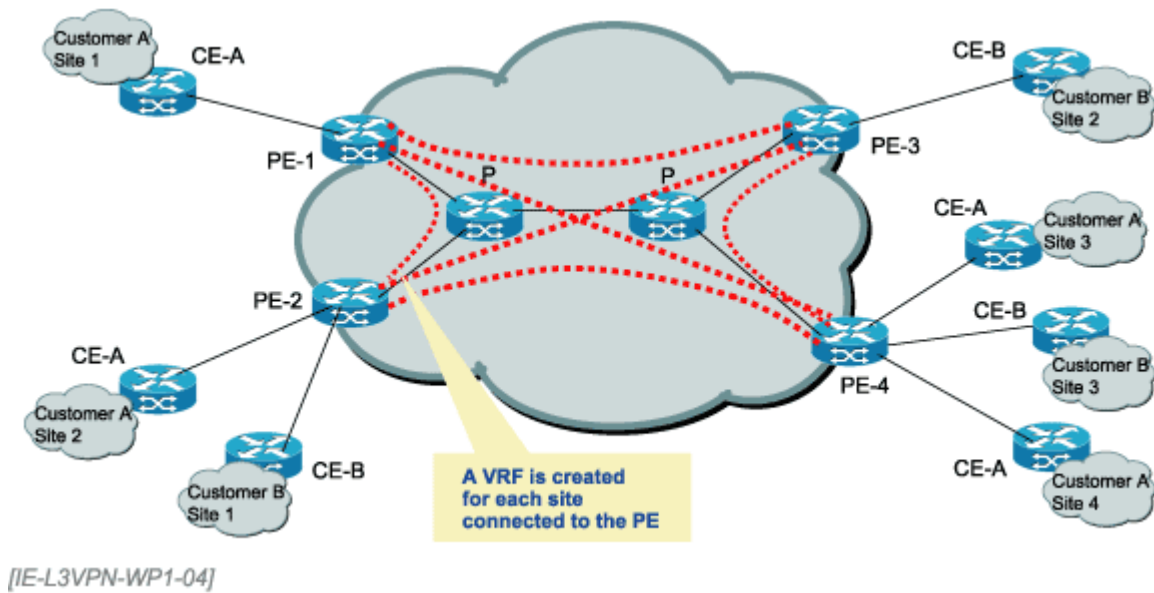


Figure 4. L3VPN Components

L3VPN components are as follows:

- CE -- Customer edge router
- PE -- Provider edge router
- P -- Provider router

CE Routers

Customer edge (CE) routers are the routers located at the customer premises. These routers interface with the PE routers and use some form of routing exchange, either static or dynamic, including such protocols as RIP, IGRP, EIGRP, OSPF, IS-IS, and BGP. The interconnection between a CE and a PE router can happen over any Layer 2 link.

Figure 4 illustrates seven CE routers, four of which interconnect Customer A (CE-A) and three of which interconnect Customer B (CE-B).

PE Routers

Provider edge (PE) routers are the routers located at the provider site and interfacing with the CE routers. These routers are aware of the customer's various VPNs and the network prefixes for each. Each of the PEs maintains a set of separate routing tables, fully independent of each other. Each of the routing tables belongs to a specific site. This separation allows duplicate addresses among various VPN customers and eliminates routing ambiguity.

Figure 4 indicates that there are four PE routers. In some

Site versus Customer Premises

Note that *customer premises* is not necessarily synonymous with *site*. In RFC 2547bis, a site can be a logical construct. According to [Rosen 2003a]:

"...[A] particular site may be divided by the customer into several 'virtual sites.' The SP may designate a particular set of VRFs to be used for routing packets from that site, and may allow the customer to set some characteristic of the packet, which is then used for choosing a particular VRF from the set.

"For example, each virtual site might be realized as a VLAN. The SP and the customer could agree that on packets arriving from a particular CE, certain VLAN values would be used to identify certain VRFs. Of course, packets from that CE would be discarded by the PE if they carry VLAN tag values that are not in the agreed upon set. Another way to accomplish this is to use IP source addresses. In this case PE uses the IP source address in a packet received from the CE, along with the interface over which the packet is received, to assign the packet to a particular VRF. Again, the customer would only be able to select from among the particular set of VRFs

cases a PE router can handle only one set of customers, but in the other instances a PE router is connected to multiple sites of various customers.

P Routers

Provider (P) routers are the backbone routers within a provider's network. PE routers interconnect via P routers, unless a PE router is directly connected to another PE router. Although both PE and P routers belong to a service provider's network, P routers are unaware of particular

VPNs, or even of the fact that VPNs actually exist. P routers do not carry VPN customer routes, nor do they participate in VPN control. P routers are only involved in the VPN forwarding plane, where they are responsible for sending semipackets and performing MPLS label swap and pop operations. This fact is a key element in L3VPNs, allowing them to scale better. Only PE routers are VPN-aware. Furthermore, no single PE router is usually required to hold all VPN customer state information, although it might in a small provider network

that the customer is allowed to use.

"If it is desired to have a particular host be in multiple virtual sites, then that host must determine, for each packet, which virtual site the packet is associated with. It can do this, e.g., by sending packets from different virtual sites on different VLANs, or out different network interfaces."

Semipackets?

I refer to packets encapsulated in the MPLS header as *semipackets*. L3VPNs require MPLS LSPs to be preset from PE to PE so that the payload information can be shipped from one PE to another for various customers, like ships in the night -- one customer is fully unaware of the other.

L3 PPVPN Architectures

As opposed to Frame Relay and ATM, an L3VPN operates on packets and shows the behavior of a routed, not a switched, system. There are two basic models, depending on whether or not the provider VPN interacts directly with the customer's routing protocols.

If the L3PPVPN does interact with the customer routing to support multiple customers, the PEs must support multiple *virtual routers* that appear to be part of the customer's routing system. Virtual routers are relatively easier for the router designer to implement, because they are one to a customer network.

The harder-to-build RFC 2547 model uses a more complex relationship between routing tables and site-specific forwarding tables. Let's examine the address structure used by the RFC 2547bis model to help us understand how the model achieves the support of overlapping address spaces for various customers' VPNs.

VPN-IPv4 Address Structure

The RFC 2547bis model uses VPN-IPv4 addresses, the format of which makes unambiguous any overlapping address spaces for various customers.

Figure 5 illustrates the VPN-IPv4 address. VPN addresses use an MP-BGP subaddress family identifier (SAFI 128). They use the same family identifier as the regular IPv4 routes.

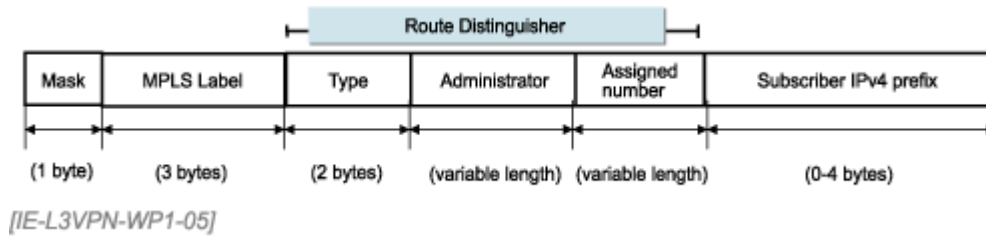


Figure 5. VPN-IPv4 Address

The VPN-IPv4 address consists of a 3-byte MPLS label (sometimes called a VRF label for reasons identified in the next section), a route distinguisher (RD), and a subscriber IPv4 prefix. The RD disambiguates routes, as it must be unique within the provider's space. Figure 6 identifies the RD fields in more detail.

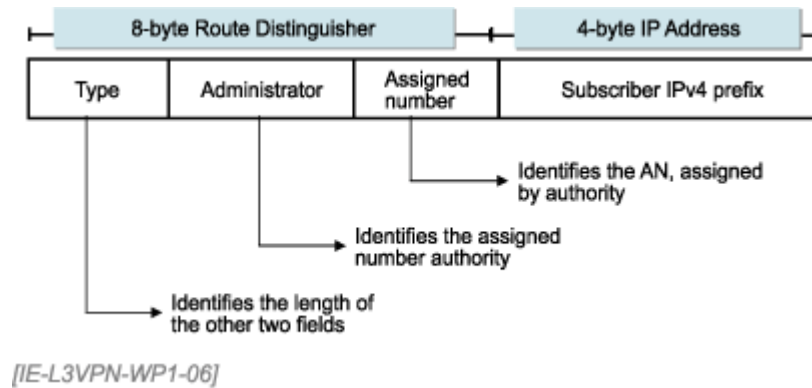


Figure 6. Route Distinguisher Structure

Cisco supports two RD formats, each of which uses a different approach to ensure uniqueness. A third method, not yet generally available from Cisco, uses the now-experimental 4-byte extended AS numbers. [Vohra 2003]

1. Type 0, where the length of the Administrator field is 2 bytes, and the Assigned Number field length is 4 bytes. The Administrator field depicts the provider's AS number, and the AN uses a 4-byte value unique within the AS. Please note that the Administrator field should contain a registered AS number from an appropriate routing registry such as ARIN or RIPE-NCC. This will guarantee uniqueness of the Administrator field, and it is then the responsibility of the administrator to guarantee uniqueness within that space.
2. Type 1, where the Administrator field length is 4 bytes and the Assigned Number field is 2 bytes. The Administrator field uses a 4-byte field in the format of an AS number. This often might be a number in the RFC 1918 private address range, which is convenient for enterprises. For example, 1055:55:155.1.1.0/24 or 5.5.5.5:80:155.1.1.0/24 will use the valid RD values.

VRFs

The RFC 2547bis model requires site-specific *VPN routing and forwarding tables* (VRFs) to exist in every PE router that is attached to that specific site. Figure 4 illustrates one PE router having two VRF tables - one for each attached VPN site. An RFC 2547bis router has at least two VRFs, one "default forwarding table" for non-VPN routes, and one or more such tables for VPNs.

Each VRF has a set of route targets (see below) of routes it will accept. You can think of these as analogous to a general BGP import/acceptance policy (e.g., accept all routes labeled with community

1:200 and reject all with 4:300).

The purpose of the VRF is to store all the routes learned from that attached site, as well as routes learned through the signaling plane of the L3VPN, MP-BGP from remote PEs. Within PEs, VPN policies are the deciding factor on which routes become part of which VRFs, based on the VPN-IPv4 route attribute called a route target (RT).

Please note that, in Cisco's implementation, a VRF consists of an IP routing table, a derived Cisco express forwarding (CEF) table, and a set of interfaces that use this forwarding table.

The RFC 2547bis draft [Rosen 2003a] allows a single CE to interconnect to more than one PE router. Should you wish to allow the traffic flow from that CE through any PE, then all your PE routers' VRFs must contain the same routes. You could, on the other hand, restrict the interconnectivity from a CE through some of the immediately attached PEs, while allowing the traffic flow through other PEs. In that case the immediately attached PEs' VRFs would be different from each other.

Furthermore, you could have a design where a single CE is associated with a set of VRFs within a PE. This is useful when you wish to divide a single VPN into sub-VPNs, imposing some sort of restrictions for inter-sub-VPN connectivity. For simplicity's sake, however, this Tutorial assumes that a single circuit, either physical or logical, is associated with a single VRF.

Associating IP Packets with VRFs

When a PE router receives a packet from a CE device, it must determine the attachment circuit over which the packet arrived, as this determines in turn the VRF (or set of VRFs) that can be used for forwarding that packet. In general, to determine the attachment circuit over which a packet arrived, a PE router takes note of the physical or logical interface over which the packet arrived, and possibly also takes note of some aspect of the packet's Layer 2 header. For example, if a packet's ingress attachment circuit is a frame relay VC, the identity of the attachment circuit can be determined from the physical frame relay interface over which the packet arrived, together with the DLCI field in the packet's frame relay header.

Although the PE's conclusion that a particular packet arrived on a particular Attachment Circuit may be partially determined by the packet's Layer 2 header, it must be impossible for a customer, by writing the header fields, to fool the SP into thinking that a packet which was received over one attachment circuit really arrived over a different one. In the example above, although the attachment circuit is determined partially by inspection of the DLCI field in the frame relay header, this field cannot be set freely by the customer. Rather, it must be set to a value specified by the SP, or else the packet cannot arrive at the PE router. [Rosen 2003a]

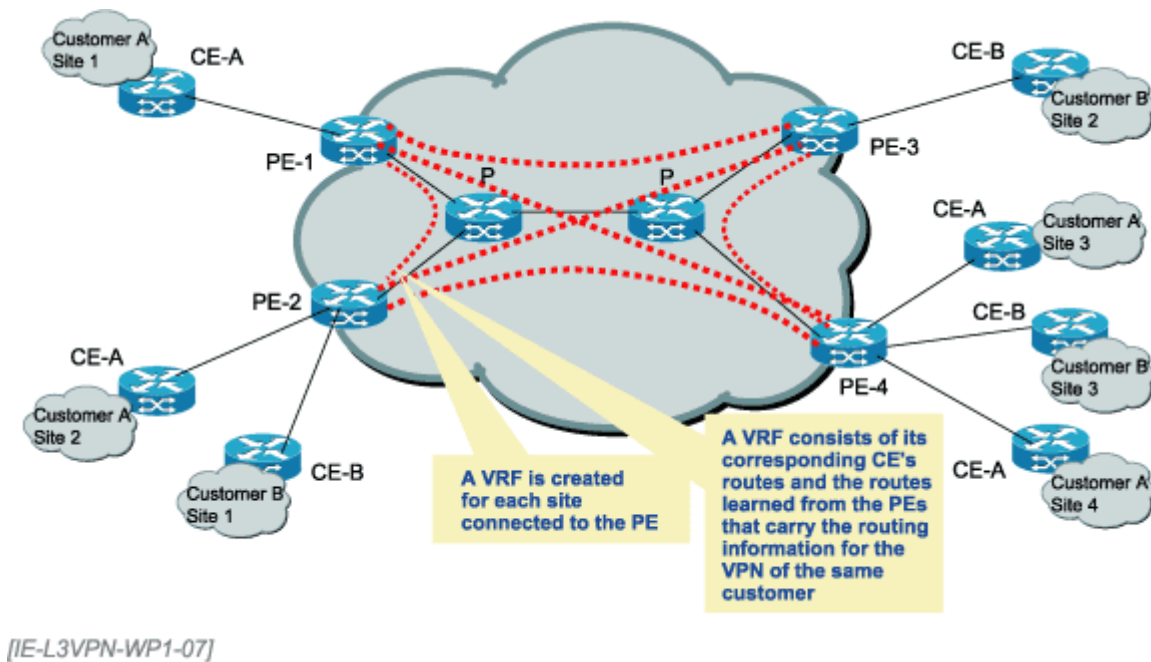


Figure 7. VRF Routes

Let's examine Figure 7, where Customer A is interconnected to PE-1, PE-2, and PE-4, while Customer B is interconnected to PE-2, PE-3, and PE-4. Customer A and Customer B cannot see each other's routes. PE routers interconnected to CE-A and CE-B contain VRFs for each customer, which are independent of each other. Should a PE have both CE-A and CE-B attached to it, it will have two VRFs. These PEs are PE-2 and PE-4. PE-1 and PE-2, on the other hand, contain only a single VRF, reflecting CE-A's and CE-B's routes, respectively.

The VRFs are populated using the dynamic or static routing protocols between a CE and its corresponding PE. Furthermore, each PE must send the learned routes to all other PEs. Using our example, PE-2 has two VRFs -- one for Customer A (CE-A) and another for Customer B (CE-B). It is PE-2's duty to send these routes to all other PE routers, (PE-1, PE-3, and PE-4), as well as to receive the routes sent from all other PEs, should those routes belong to one of the VPNs that the PE-2 is connected to. This means that PE-2 receives the route advertisement from PE-1, PE-3, and PE-4. These updates are done using the MP-BGP, which is addressed later in this Tutorial. PE-1 sends the content of its VRF to PE-2, which will be populated into VRF for customer A. PE-3 sends the content of its VRF to PE-2, which will be populated into the VRF for the customer B. Finally, PE-4 sends the context of its two VRFs to PE-2, populating both VRFs.

The idea of populating VRFs with proper routes is really a simple one, when you apply certain rules, or policies, allowing only specific routes to be part of specific VRFs. A PE must be able to distinguish routes for Customer A versus routes for Customer B. RFC 2547bis dictates the use of extended communities to recognize routes belonging to different customers. Based on those communities, you could write policies that will allow routes with a community X to be Customer A routes, and routes with a community Y to be Customer B routes.

Route Targets

Route targets are carried in BGP extended community attributes. Think about BGP communities in general and ask yourself what is the purpose of a BGP community. The BGP community allows us to tag routes in one place (say router A) and then recognize them somewhere else (say router B) and, as the result of such recognition, do something with those routes. There are well-known BGP communities, such as no-export and no-advertise, which routers are preprogrammed to recognize and deal with. You can also define your own communities. In that case you will have to write your own configurations that

will allow the router to recognize those communities and perhaps, as a result of such recognition, change a BGP attribute (for example, a local preference).

A route target community is an extended community type, as defined in [Sangli 2002]. It is used to recognize one customer's routes versus those of another customer. When an IPv4 route is learned statically or dynamically from the CE, the PE creates the VPN-IPv4 route and associates this route with the route target that is specific for that customer. Using our example in Figure 7, routes advertised from CE-A to PE-2 get tagged with a predefined community that will allow those routes to be recognized by other PE routers. Once PE-1 and PE-4 receive those routes, based on the route target community they will know that the routes belong to the VRFs associated with the Customer A. PE-3, however, will reject the route, because the VRF for Customer A does not exist there.

Route Targets Formally Defined

A route target community attribute can be thought of as identifying a set of sites (though it would be more precise to think of it as identifying a set of VRFs). Associating a particular route target community attribute with a route allows that route to be placed in the VRFs that are used for routing traffic received from the corresponding sites. [Rosen 2003a]

Export and Import Targets

Let's go back to the example illustrated in Figure 7. When the PE-2 router receives the routes from CE-A and CE-B, it places them into the corresponding VRFs. Next, prior to PE-2 advertising those routes to other PEs, the routes must be tagged with the export target community. PE-2, on the other hand, receives the VPN-IPv4 routes from PE-1, PE-3, and PE-4. Those were tagged with their own export target communities by PE-1, PE-3, and PE-4 before they were exported from those routers. When PE-2 receives these routes, it will use the import target community to identify the relative routes for specific VRF tables. Although the export target and the import target can be different from each other, they must be the same if a PE is to install the route heard from another PE into its VRF. This means that the export target sent by PE-4 to PE-2 must match the import target at PE-2 in order for PE-2 to install the route heard from PE-4 into the corresponding VRF.

Why Are the Targets Called "Export" or "Import"?

You can view the export target as the tag that is attached to the routes that get exported from a PE router into BGP toward other PEs. The import target gets imported with the advertised route into the router via BGP from other PE routers.

Since the format of the regular BGP community attribute allows only a 2-byte numbering space, BGP extended communities are used to define route targets. These are structured similarly to the RDs defined earlier in this Tutorial.

It is important to note that a VPN-IPv4 route can have only one RD, but it can have multiple route targets. The PEs can be configured to associate all the routes going to a specific CE with a specified route target. Or, the PEs could associate only certain routes of the specific CE with one route target, and other routes with another route target value. It is important for service providers and customers to agree on whether the customer is allowed to tag its routes with the route targets. First of all, in order for this flexibility to exist, the CE-PE routing protocol must be BGP. Next, if the customer tags the routes with the route targets, then the corresponding PE of the service provider must filter out the routes that are not allowed to propagate through the VPN cloud.

Route Origin

If the route target community associates a customer route belonging to a specific VRF, the route origin community associates a route with the site that originates the advertisement.

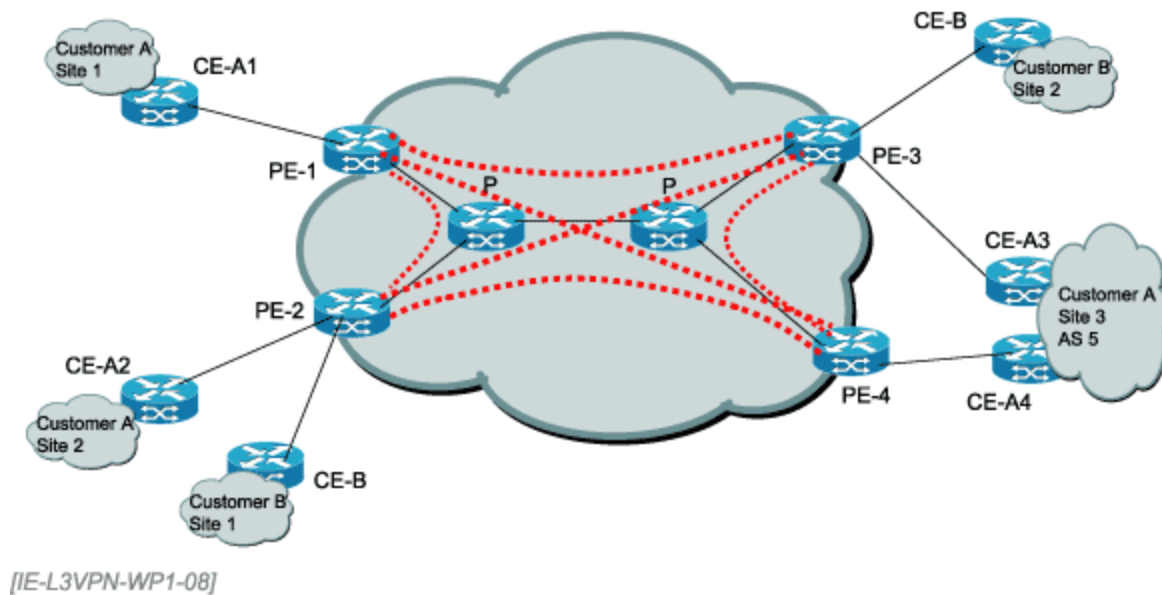


Figure 8. Route Origin Example

In Figure 8, Customer A site 3 is attached to two different PEs: PE-3 and PE-4. When PE-3, attached to CE-A3, receives the routes from PE-4 with the route origin community set as CE-A4, PE-3 will reject the importing of these routes into the VRF associated with CE-A4. This ensures more efficient VPN functionality and even prevents routing loops in some scenarios. Should these routes not be filtered, CE-A3 would send the traffic to PE-3 (or CE-A4 to PE-4), as EBGP routes are preferred to IGP in IOS. So, in a way, route origin allows you to implement policy routing.

You can view the use of route origin as an extension of the split horizon rule -- do not re-advertise the routes through an interface if you learned them through that interface.

RFC 2547bis Operation

This section reviews L3VPN operation. L3VPNs use MPLS as an underlying technology to create transparency to the P-level routers. L3VPN operation is divided into two planes:

1. The signaling plane (sometimes referred to as control flow), which is responsible for customer routes routing exchange between the PEs carried into the CE sites
2. The forwarding plane (sometimes referred to as data flow), which is responsible for data forwarding between the CE sites across the provider's network

If MPLS uses RSVP or LDP as the signaling protocol, L3VPNs use MP-iBGP as their signaling protocol.

BGP Signaling Plane

The L3VPN signaling plane uses MP-iBGP, whose responsibility is to exchange the CE routes between the PEs so that customer data can be forwarded. The signaling exchange between the CE and the PE can be any dynamic protocol (such as RIP, OSPF, IS-IS, BGP, etc.), or even static routing. Once the PE routers obtain the customers' routes, tacked independently in various PEs' VRFs, MP-iBGP conveys those routes to all other PEs, where they are put into the VRFs corresponding to the appropriate customers. Also, since the P routers have to be unaware of any VPNs for scalability reasons, MPLS has to run between the PEs to carry the payload data. MPLS, of course, requires its own signaling protocol - either RSVP or LDP.

CE-PE Routing Information Exchange

There is nothing new on the CE-PE routing exchange front. It is regular routing information exchange, should you choose the dynamic method. You can use any routing protocol you desire -- RIP, OSPF, IS-IS, or BGP. This routing exchange is localized between the CE and the PE. You can have two different sites for the same customer use different routing protocols between CE and PE. That is, CE-PE routing is fully localized. In addition, you can use static routing. Protocol selection depends solidly on the agreements between the service provider and the customer.

Should the CE-PE routing protocol be BGP, then the service provider may allow the customer to attach route targets to the routes before they get advertised to the attached PE. The customer and the service provider must agree on the route target values prior to use. Should this implementation method be chosen, the customer can specify, of course within the agreed limits, various routes and their corresponding route targets in real time. In this implementation the service provider must filter out the routes that contain improper route targets, which the customer is not allowed to use, based on certain agreements.

PE-PE Routing Information Exchange

PE-PE routing exchange is done with the help of MP-iBGP. MP-iBGP is very similar to regular IPv4 iBGP. That is, an iBGP session must be established between two BGP speakers before they can exchange any routing information. However, one critical extension of MP-BGP is the fact that it can handle VPN-IPv4 addresses (among other SAFI types), which are used for RFC 2547bis.

In summary, a PE router tags all the individual CE routes with an extended community that is unique for that customer. Then, once those routes are injected into the MP-iBGP, all other PEs receive them. Now a receiving PE needs to distinguish which routes belong to which VRF or the customer. Using the tagged extended communities, the PEs can distinguish the routes for various customers and inject those routes into the corresponding VRF.

Let's examine the PE-PE routing exchange closer, viewing Figure 9 using Customer A as our example.

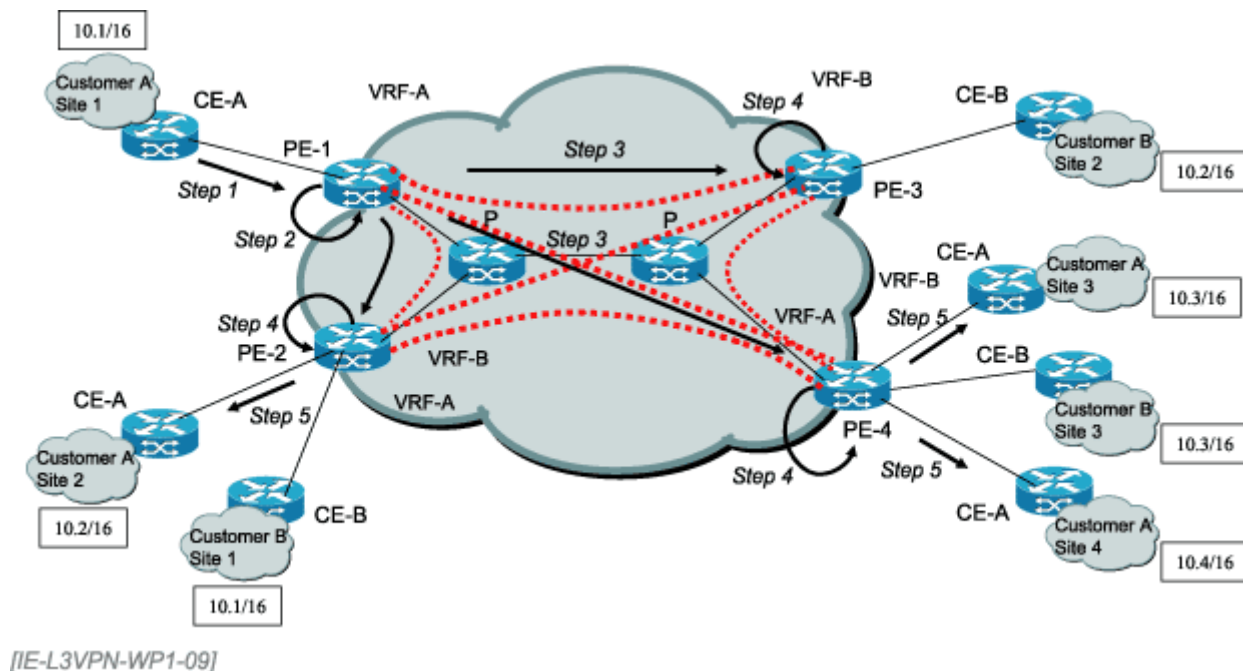


Figure 9. PE-PE Routing Information Exchange

Step 1. The site 1 CE-A device advertises the 10.1/16 routes to PE-1, using a dynamic routing protocol. PE-1 puts these routes into Customer's A VRF. Note that although there are no other customers attached to PE-1, VRFs are still used to allow PE-1 to add more customers later.

Step 2. PE-1 tags the CE-A routes with the extended VPN route target community associated with the Customer A. The route target community value of 10200:32 is set from the export list of route targets associated with Customer's A VRF.

Step 3. PE-1 generates an MP-iBGP update message containing the route 10.1/16, which was learned from CE-A site 1. This route has a VPN-IPv4 format. PE-1 sends this route to all its MP-iBGP peers configured on PE-1. Should there be other routes learned from CE-A site 1, they will be advertised also. Note that although the PE-1 and PE-2 routers do not share the same customer, you might still want to have the MP-iBGP sessions formed between them for future expansion. Assuming this is the case in our example, PE-1 sends the update to all PE routers (PE-2, PE-3, and PE-4).

Step 4. The remote PEs (PE-2, PE-3, and PE-4) receive the VPN route advertisement. These PEs use their import route targets to select which routes belong to which VRF. Cisco derives a CEF forwarding table from the VRF routing table. A separate set of routing and CEF tables is maintained for each VRF. These tables prevent information from being forwarded outside a specific VPN. In addition, these tables prevent packets that are outside of a VPN from being forwarded to a router within a VPN.

Step 5. The remote PEs forward the prefixes learned from PE-1 to their local CEs, if and only if those CEs are part of the same customer. The prefixes are forwarded via a CE-PE routing protocol, whatever it might be.

LSP Establishment

Based on the routing information stored in the VRF and CEF tables, packets are forwarded to their appropriate destinations using MPLS. MPLS is necessary for data forwarding. Think about it! How else can a packet with a private or a duplicate destination address traverse the provider's backbone? It can, but only in the hidden form. This is what MPLS tunnels will be doing -- hiding the payload packets.

According to [Rosen 2003a], "A PE router associates a label to each customer prefix learned from a CE router. Next, a PE router includes the label in the network reachability information for each prefix that it advertises to other PE routers. The PE may distribute the exact set of routes that appears in the VRF, or it may perform summarization and distribute aggregates of those routes, or it may do some of one and some of the other."

You can use either RSVP or LDP for LSP establishment. The P routers must be aware of the MPLS tunneling structure, but not the payloads running through it.

[Rosen 2003a] states, "Suppose that a PE has assigned label L to route R, and has distributed this label mapping via BGP. If R is an aggregate of a set of routes in the VRF, the PE will know that packets from the backbone which arrive with this label must have their destination addresses looked up in a VRF. When the PE looks up the label in its Label Information Base, it learns which VRF must be used. On the other hand, if R is not an aggregate, then when the PE looks up the label, it learns the egress attachment circuit, as well as the encapsulation header for the packet. In this case, no lookup in the VRF is done.

"We would expect that the most common case would be the case where the route is **not** an aggregate. The case where it is an aggregate can be very useful though if the VRF contains a large number of host routes (e.g., as in dial-in), or if the VRF has an associated LAN interface (where there is a different outgoing Layer 2 header for each system on the LAN, but a route is not distributed for each such system)."

See [Rosen 2003a] for a description of the potential choices for the router designer in assigning labels to routes.

Customer Data Forwarding Plane

Now that all the routes have been successfully exchanged, let's trace the payload packet flow between the two CEs. Assume that Customer A site 1 needs to communicate with site 4, as illustrated in Figure 10.

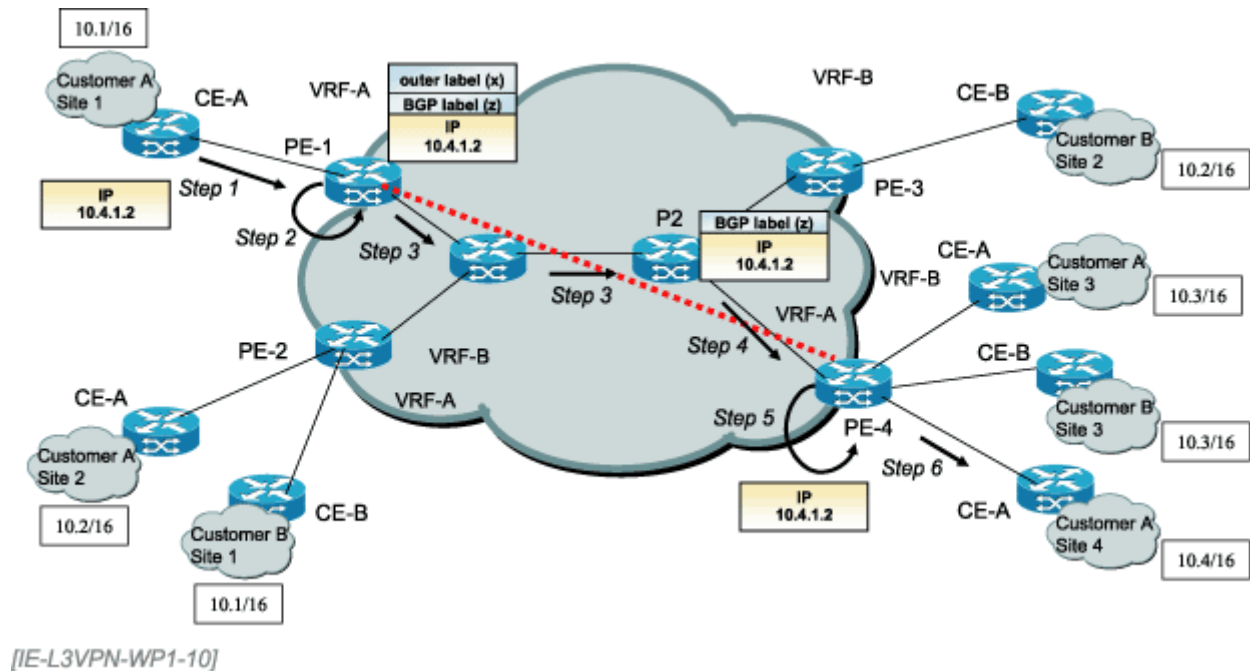


Figure 10. Payload Flow

Step 1. CE-A performs the longest match lookup on the packet addressed to 10.4/16. This lookup results in CE-A forwarding the packet to the IP address associated with the PE-1 router's VRF interface.

Step 2. The PE-1 router binds a label to the Customer A prefix learned from the PE-4 router. This results in the payload being encapsulated into inner and outer labels. This action is sometimes referred to as a double label push. The inner label is used by PE-4 to identify which CE that specific packet belongs to. The outer label is used to direct the packet to the correct PE router.

Step 3. When the packet leaves the PE-1 router, it simply traverses the LSP, which is specifically set up for PE-1 to PE-4 communication when carrying customer traffic. The P routers are now LSR routers and are responsible for swapping the outer MPLS label, x. The inner label, on the other hand, remains untouched throughout the entire LSP journey.

Step 4. When the packet reaches the penultimate hop router (PHR), which in our example is P2, the PHR performs the outer label pop, deencapsulating the outer MPLS label and leaving the inner label only. Then it forwards the packet to the LSP's egress point with only the VRF label.

Step 5. The PE-4 router uses the received VRF label to identify the specific VRF interface that must be used in forwarding the packet further to its destination. Upon the completion of the successful identification, the PE-4 router pops the inner label, leaving the regular IP packet intact.

Step 6. The PE-4 router sends the IPv4 packet to its outbound interface, pointing toward Customer A site 4.

Provider Data Forwarding Plane

It is important to note that the provider's core could already be running MPLS for other types of traffic. This could include providing services to various ISPs, where the core is traffic engineered in such a way that one set of routes is used for one set of ISPs, while another set of routes is used for other types of ISPs. Should that be the case, RFC 2547bis traffic can be viewed as just an additional type that is carried over the same core. Furthermore, should the provider deploy L2VPNs, utilizing the Kompella implementation method, it uses the same signaling plane as the RFC 2547bis [Kompella 2003]. Cisco does not support the Kompella implementation, although they do support Martini. Juniper supports both the Kompella and Martini implementations.

RFC 2547 payloads are just one more (labeled) type within the provider's MPLS cloud, or even in interconnected provider MPLS clouds. A motivation for 2547 is the idea that the provider is already likely to have a MPLS backbone, and RFC 2547 is one more service that layers onto it.

RFC 2547bis Network Design

Membership

You need to begin by defining the customer VPNs with which you will initially configure the overall system. At the most basic, this is a matter of identifying the customer sites in each VPN, and any requirements the VPN will have to communicate with other VPNs or with the public Internet.

For the moment, we'll defer special features. For basic VPN service, assume that the PE will know the addresses involved (i.e., of the CE) through static configuration.

PE Locations

Obviously, you need to know the identity and location of PE routers, as well as their connectivity.

Defining VRFs

To define a VRF, you need to know the specific PE for which you are defining, the sites that connect to it, and the VPNs that connect to that site. See Table 2 for basic configuration.

Special Handling within RFC 2547bis

RFC 2547bis Scalability

A key aspect of the RFC 2547bis model is that, while the set of PE routers can carry all VPN states for the SP's network, it would be extremely rare for a single PE to have to carry all states. You can extend this idea to route reflection and have multiple route reflectors that are responsible for various portions of the total VPN customer base. Your reasons for choosing one set of route reflectors versus another could depend solidly on the geographical location of the PEs and the most economical way of grouping them into a route reflector cluster.

One big advantage of route reflector deployment is the fact that you will minimize the number of MP-iBGP sessions in the SP network. As Howard Berkowitz states in his paper, "Realities Underlying Virtual Private Networks":

"Each PE can also attach to a BGP route reflector. If you do use route reflectors, you must be careful that router-specific information still propagates throughout the cluster."

[Chen 2003] defines the use of Outbound Route Filter (BGP-ORF) for BGP route reflectors. The use of ORF can improve efficiencies by allowing a route reflector to reflect only those routes about which a

particular route reflector client PE router cares. The PE router sends the BGP peer a list of route targets in which it is interested. The BGP peer applies this route target list as an outbound filter in such a way that the route reflector sends to the PE router only the routes that match at least one of its configured route targets. Deployment of ORF results in fewer BGP updates and less protocol traffic.

Cisco introduced ORF functionality in 12.0(11)ST. The function is a subset of the IETF proposals on ORF: it supports ORF for prefix-lists, but not for AS paths [Patel 2003] or communities. The latter two are the subject of different IETF proposals. See [Cisco 2003] for configuration details.

It is very important to emphasize that, if deployed, route reflectors are used only by the signaling plane of RFC 2547bis. The data forwarding plane does not involve the route reflectors at all. Regardless of whether or not the route reflectors are used for MP-iBGP exchange information, the payload packets traverse via the LSPs preset between the PEs.

Select Special Features of Cisco's L3VPN Implementation

Cisco's implementation of RFC 2547bis includes:

- **VRF Lite.** With VRF Lite, the ML series is considered as either a PE or a CE extension. It is considered a PE extension since it allows for VRF without the existence of MP-iBGP. It is considered a CE extension since this CE can have multiple VRFs and can serve many customers with one CE platform. Under VRF Lite, an ML-series CE can have multiple physical and/or logical interfaces with PEs for difference customers. It holds VRFs locally and does not distribute the VRFs to its connected PEs. It uses VRF information to send traffic to the appropriate interfaces when it receives the traffic from the ISP's PE routers.
- **Multi-VRF CE.** This feature was introduced in Cisco IOS release 12.2(4)T. It allows some of the functionality of a PE to be added to a multi-VRF CE router. This feature allows a CE to maintain separate VRF tables, which could extend the privacy and security of an MPLS VPN all the way to the branch level of the customer. The conversation between the CE and the PE is pure IP, without MPLS. Multi-VRF CE routers use VRF interfaces to form an adjacency similar to that of a VLAN on the customer side. Each VRF on the multi-VRF CE router is mapped to a VRF on the PE router.
- **VRF selection based on source IP address.** This feature was introduced in Cisco IOS release 12.0(22)S. It allows a specified interface on a PE router to route packets to different VPNs based on the source IP address of the packet. Once the packet is selected into the correct VRF table, it is forwarded as any other packet would be -- based on its destination address, using the VPN.

RFC 2547bis Basic Configuration

This section focuses on steps to be performed when implementing L3VPNs on the Cisco gear. First, it examines all the necessary steps that need to be performed, then illustrates the IOS L3VPN enhancements, and finally presents the configuration example.

The necessary steps consist of the following:

- MPLS tunnel predefinition
- MP-iBGP configuration between the PEs
- Configuring MPLS within the provider's network
- Defining VPNs

- Configuring CE-PE routing information exchange
- Configuring PE-PE routing information exchange

Configuring MPLS within the Provider's Network

In order for the payload packets of the customers' networks to be able to traverse the provider's networks, MPLS LSPs must be preset between the PEs. The commands necessary to achieve that are listed in Table 1.

Table 1. MPLS Setup within the Provider's Core

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# ip cef distributed	Enables CEF for MPLS forwarding. The key word distributed is required for 7500 series router to have a better performance.	Mandatory
2	Router(config)# tag-switching advertise-tags	Controls distribution of locally assigned (incoming) tags via the Tag Distribution Protocol (TDP).	Mandatory
3	Router(config-if)# tag-switching ip or Router(config-if)# mpls ip	Enables MPLS forwarding of IPv4 packets along normally routed paths for a particular interface. This must be configured on all the involved interfaces. The mpls ip version is the newer method.	Mandatory

Defining VPNs

VPN routing instances configuration must take place in all the PEs involved. You must perform the steps identified in the Table 2.

Table 2. Configuring VPN Definition

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# ip vrf vrf-name	Enters the VRF configuration mode, defines the VPN routing instance, and assigns the VRF name.	Mandatory
2	Router(config-vrf)# rd	Creates routing and forwarding	Mandatory

	<i>route-distinguisher</i>	table. <i>route-distinguisher</i> is an 8-byte value that is added to an IPv4 prefix to create a VPN IPv4 prefix.	
3	Router(config-vrf)# route-target {import export both} route-target-ext-community	Creates a list of import and/or export route target communities for the specified VRF. The key word import means that routing information will be imported from the target VPN extended community. The key word export means that routing information will be exported to the target VPN extended community. The key word both means that routing information will be exported and imported from/to the target VPN extended community.	Mandatory
4	Router(config-vrf)# import map route-map	Associates the specified route map with the VRF. If you use this command, you must define the specified route-map as well.	Optional
5	Router(config-vrf)# export map route-map	Associates the specified export route map with the VRF. If you use this command, you must define the specified route-map as well.	Optional
6	Router(config-if)# ip vrf forwarding vrf-name	Associates a VRF with an interface or a subinterface, facing the CE. The default for an interface is the global routing table. Please note one caveat -- executing this command on an interface removes the IP address. To rectify this, you need to reconfigure the interface IP address.	Mandatory

Configuring CE-PE Routing Information Exchange

Configuration steps for the CE-PE routing information exchange totally depend on the type of a routing protocol that is used between CE and PE, if any. You can also use static routing between the CE and the PE. Please note that, should you use a dynamic routing protocol, it must match the CE routing protocol for successful exchange of routing information.

Table 3 presents the steps necessary for a PE to communicate with a CE using BGP. Table 4 presents the same steps should RIP be used as the dynamic routing protocol between CE and PE. Table 5 presents the static route syntax definition between CE and PE. Please do not forget that the same steps must be performed in the corresponding CE. Note as well that you can also use OSPF or IS-IS between CE and PE.

Table 3. Configuring CE-PE Routing Information Exchange Deploying BGP

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# router	Configures EBGp with the CE	Mandatory

	<code>bgp autonomous-system</code>	end, where autonomous-system is the service provider's AS number	
2	Router(config-router)# <code>address-family ipv4 [unicast]</code> <code>vrf vrf-name</code>	Defines BGP PE-CE session for vrf-name	Mandatory
3	Router(config-router-af)# <code>neighbor {ip-address peer-group-name}</code> <code>remote-as number</code>	Specifies the CE's IP address, identifying it to the local AS	Mandatory
4	Router(config-router-af)# <code>neighbor ip-address activate</code>	Activates the advertisement of the IPv4 address family	Optional

Table 4. Configuring CE-PE Routing Information Exchange Deploying RIP

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# <code>router rip</code>	Enables RIP	Mandatory
2	Router(config-router)# <code>address-family ipv4 [unicast]</code> <code>vrf vrf-name</code>	Defines RIP parameters for the PE to CE routing sessions	Mandatory
3	Router(config-router-af)# <code>network prefix</code>	Enables RIP between CE and PE	Mandatory

Table 5. Configuring CE-PE Routing Information Exchange Deploying Static Routing

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# <code>ip route vrf vrf-name</code>	Defines static route parameters for every PE to CE to establish static routes for a VRF	Mandatory
2	Router(config-router)# <code>address-family ipv4 [unicast]</code> <code>vrf vrf-name</code>	Enters address family submode for configuring PE-PE MP-IBGP routing protocol	Mandatory
3	Router(config-router-af)# <code>redistribute static</code>	Redistributes VRF static routes into VRF BGP table	Mandatory
4	Router(config-router-af)# <code>redistribute static connected</code>	Redistributes directly connected networks into the VRF BGP table	Mandatory

Configuring PE-PE Routing Information Exchange

PE-PE routing information exchange occurs with the help of MP-IBGP. The steps necessary to achieve this are listed in Table 6.

Table 6. Configuring PE-PE Routing Information Exchange

Step	Command Syntax	Description	Necessity of Command
1	Router(config)# router bgp autonomous-system	Enters the iBGP routing process, where the autonomous-system is the AS number of the service provider.	Mandatory
2	Router(config-router)# neighbor {ip-address peer-group-name} remote-as number	Specifies another PE's IP address to form an iBGP session. Please note that the number is equal to autonomous-system specified in step 1.	Mandatory
3	Router(config-router)# neighbor ip-address activate	Activates the advertisement of the IPv4 address	Mandatory
4	Router(config-router)# address-family vpnv4 [unicast multicast]	Defines MP-iBGP parameters for VPN IPv4 NLRI exchange.	Mandatory
5	Router(config-router-af)# neighbor ip-address remote-as	Defines MP-iBGP session to exchange VPN IPv4 NRIs.	Mandatory
6	Router(config-router-af)# neighbor ip-address activate	Activates the advertisement of the IPv4 address family.	Mandatory

Cisco IOS L3VPN Configuration Enhancements

Cisco IOS L3VPN configuration enhancements, available as of IOS Release 12.0(7)T, provide increased BGP functionality, which enables service providers to better manage and control traffic flow within a VPN. These enhancements enable you to achieve the following:

- Implement hub-and-spoke VPN topologies
- Enable faster convergence for BGP VRF routes
- Limit the number of BGP VRF routes, resulting in better scalability
- Distribute BGP OSPF routing information

Table 7 lists the L3VPN enhancements and the corresponding IOS commands.

Table 7. L3VPN IOS Enhancements in Release 12.0(7)T

Command	Description	Enhancement Benefit
<code>bgp scan-time [import] scanner-interval</code>	Allows configuration of scanning intervals of BGP routers to decrease import processing time of routing information. The scanner-interval valid values range from 5	Decreases import processing time of VPN IPv4 routing information, resulting in faster convergence.

	to 60 seconds, with the default value being equal to 60 seconds.	
<code>maximum routes limit {warn-threshold warn-only}</code>	Allows limiting the number of routes in a VRF to prevent a PE router from importing too many routes. limit value can be from 1 to 4,294,967,295.	Prevents a PE router from importing too many routes into the VRF routing table. Also allows enforcement of the maximum number of members that can join a VPN from a particular site.
<code>neighbor allowas-in number</code>	Enables configuration of PE routers to allow CE routers to re-advertise all prefixes that contain duplicate autonomous system numbers (ASNs) to neighboring PE routers. Valid values of number range from 1 to 10.	Allows L3VPNs to have hub-and-spoke topologies, where a CE router can readvertise all prefixes containing duplicate AS numbers to neighboring PE routers.
<code>neighbor ip-address as-override</code>	Enables configuration of a PE router to reuse the same ASN on all sites within an MPLS VPN by overriding private ASNs. The ip-address specifies the router's IP address to override with the AS number provided.	Allows customers to configure BGP routes with the same AS number in multiple geographically dispersed sites, resulting in better scalability.

Configuring Complex Topologies

Import and export route targets allow flexible VPN designs. You could create fully meshed, partially meshed, or even hub-and-spoke topologies. In the fully meshed case, you would use the same route target for a single VPN in all the routers. A PE router advertising the route would leave with the export route target. A PE router receiving the route would use the same route target to import the route into the VRF for that VPN (so this route target would be the import route target). When you examine the other extreme topology, hub-and-spoke, you need to use two route targets, *hub* and *spoke*. The PE attached to the hub CE would use the hub route target when exporting routes to the PEs attached to the spoke CEs, and would use the spoke route targets when importing routes from the PEs attached to the spoke CEs. The PEs attached to the spoke CE would use the hub route target when importing routes from the PE attached to the hub CE, and would use the spoke route targets when exporting routes to the PEs attached to the hub CE.

The hub-and-spoke topologies predominantly were used for VPNs over Frame Relay and/or ATM. This is due to the cost of VCs. It is quite natural for a customer migrating from a legacy L2VPN to the L3VPN to proceed with the topology reflecting the current scenario, and then slowly migrate to the fully meshed environment.

Let's examine the hub-and-spoke topology more closely, analyzing the signaling flow between the spoke locations and the data forwarding between the spoke locations. Remember that the idea behind the hub-and-spoke topology is that the spoke CE routers cannot send traffic directly to each other. The traffic flow must go through the hub CE.

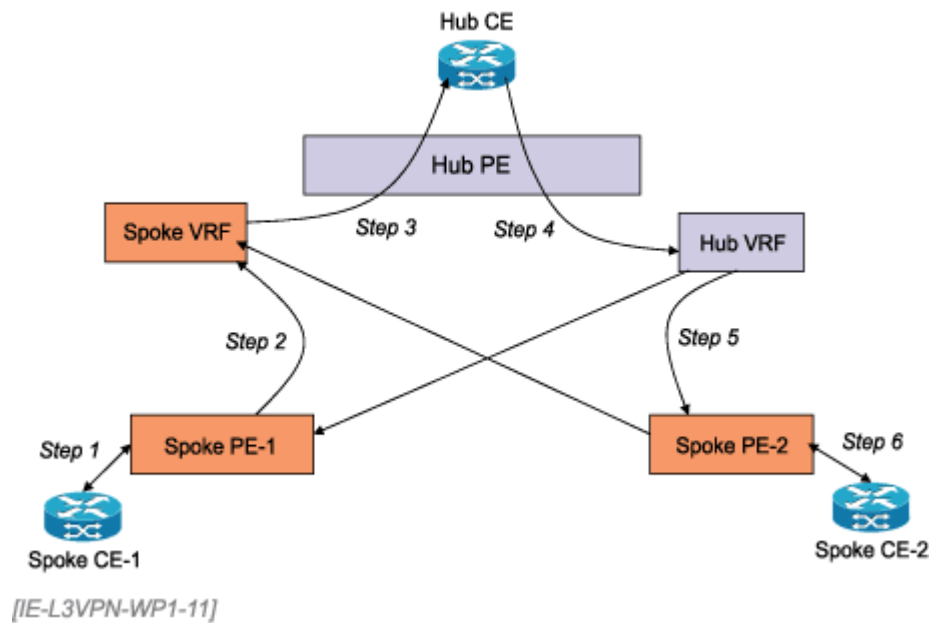


Figure 11. Hub-and-Spoke Topology Signaling Flow

Let's examine Figure 11 for the signaling flow between the spokes. The details of the Signaling flow are as follows:

Step 1. Spoke CE-1 advertises its routes to spoke PE-1.

Step 2. Spoke PE-1 advertises those routes to the spoke instance on the hub PE using the spoke route target.

Step 3. The hub PE sends the routes in the spoke instance to the hub CE router using a separate logical interface belonging to the VRF between the hub CE and the hub PE.

Step 4. The hub CE router re-advertises the routes (or it could generate the aggregate routes for all spoke CE sites). It sends the advertisement to the hub PE router into the hub instances using its own logical interface belonging to that VRF.

Step 5. The hub PE router advertises the routes from the hub instance to the spoke sites using the hub route target.

Step 6. The spoke sites match the routes with the hub route target and install those routes into their VRFs. Then the spoke PE sends those routes to the attached spoke CE router, in this case CE-2.

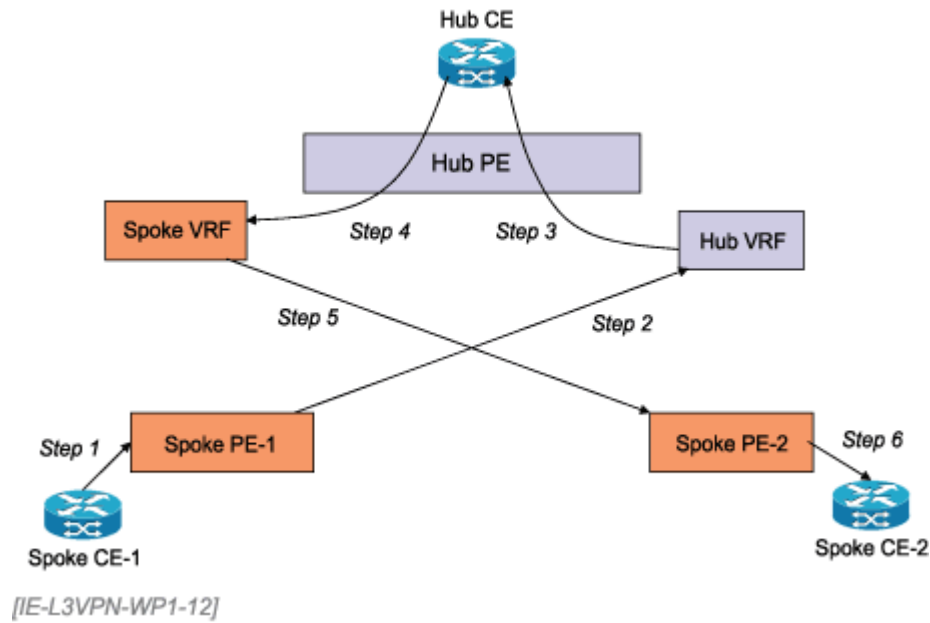


Figure 12. Hub-and-Spoke Topology Data Flow

Figure 12 illustrates the payload packet flow between the two spoke locations. The details are as follows.

Step 1. CE-1 sends the payload packet with the destination address of the CE-2.

Step 2. The PE-1 learns all the routes about site 2 through the hub instance of the PE router. Therefore it will send all the packets to the hub PE router.

Step 3. When the hub PE receives the packets, using the hub forwarding instance, it will forward the packet out of the logical interface between the hub PE and the hub CE to the hub CE.

Step 4. Since the hub CE learned about site 2 routes from the hub PE router's spoke instance, the hub CE router turns the packet around and sends it to the hub PE router VRF via the logical interface dedicated for that VRF.

Step 5. The spoke instance in the hub PE router forwards the packet to the spoke PE-2 router.

Step 6. The spoke PE-2 router forwards the packet to the CE-2.

Recommended Troubleshooting Steps and Commands

You need to apply a layered approach when troubleshooting any problem. Typical network troubleshooting methodologies require adherence to the layers of the OSI reference model. The context of L3VPN troubleshooting methodology uses not only the OSI reference model layers but also L3VPN layers. The focus of this Tutorial is to identify L3VPN layers that must be unpeeled while performing the troubleshooting.

Following are the layers of the L3VPN that must be checked when troubleshooting:

1. Signaling plane
 - o CE-PE routing exchange at one side of the VPN

- PE-PE MP-IBGP session and routing exchange
- PE-CE routing exchange at the opposite side of the VPN

2. Data forwarding plane

- LSPs must be established
- Packet forwarding between CE-PE at both ends of the VPN
- Packet forwarding between PEs.

Cisco IOS commands available to you for proper VPN operation verification are listed in Table 8.

Table 8. Cisco IOS VPN Operation Verification Commands

Command	Description
Router# show ip vrf	Displays the set of defined VRFs and interfaces
Router# show ip vrf [{brief detail interfaces}] vrf-name	Displays information about defined VRFs, in brief or detail format, and associated interfaces
Router# show ip route vrf vrf-name	Displays the IP routing table for a VRF vrf-name
Router# show ip protocols vrf vrf-name	Displays the routing protocol information for a VRF vrf-name
Router# show ip cef vrf vrf-name	Displays the CEF forwarding table associated with a VRF vrf-name
Router# show ip interface interface-number	Displays the VRF table associated with interface interface-number
Router# show ip bgp vpnv4 all [tags]	Displays information about all BGP sessions
Router# show tag-switching forwarding vrf vrf-name [prefix mask/length] [detail]	Displays label forwarding entries corresponding to VRF vrf-name routes advertised by this router

Configuration Example

Let's look at the configuration example of L3VPN using Cisco IOS. Figure 13 illustrates the topology for the example. You have two Customers, A and B. Both customer's sites are attached to PE-1 and PE-2. The ISP is running ISIS IGP inside of its AS.

The major configuration steps are:

1. PE-1 and PE-2 configurations
 - Configure route targets for Customers A and B.
 - Create VRFs for Customers A and B.

- o Configure the interfaces pointing towards P routers for MPLS switching.
- o Configure ISIS so that the interfaces point towards P routers.
- o Configure BGP between PE-1 and PE-2 routers, pointing to the Loopback interfaces.
- o Configure MP-BGP between PE-1 and PE-2 routers.

2. P configurations

- o Configure all the interfaces for MPLS switching.
- o Configure ISIS, including all the interfaces involved in ISIS routing.

The layout is shown in Figure 13.

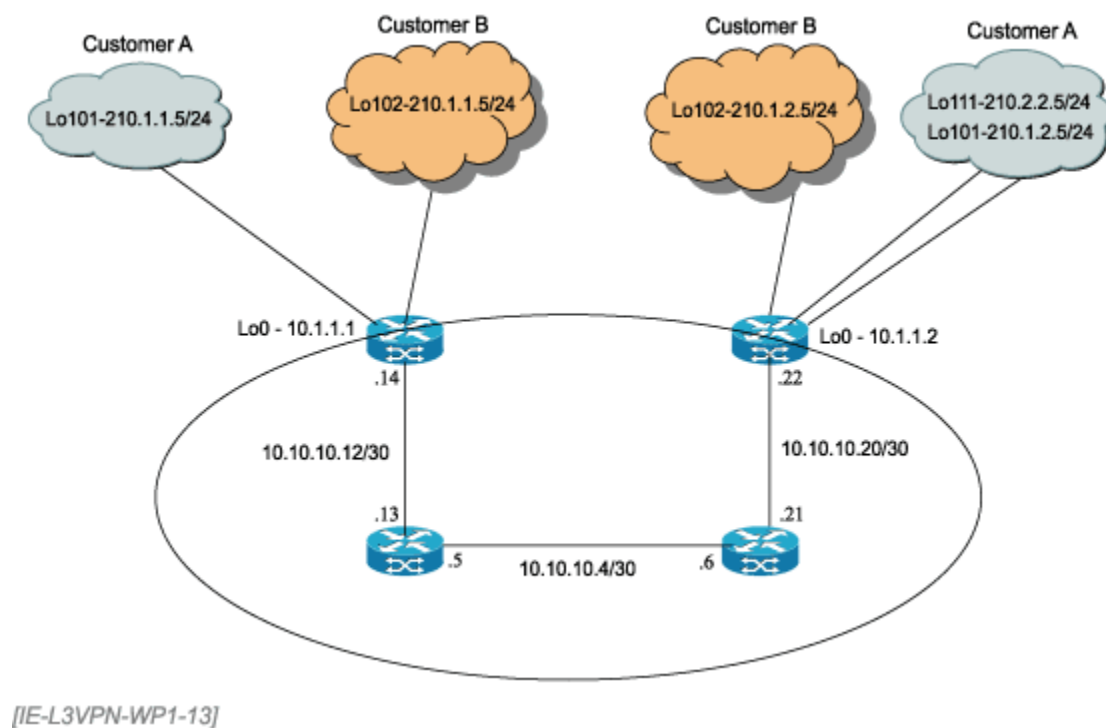


Figure 13. L3VPN Configuration Example

The detailed final configuration is shown in the L3VPN Lab Exercise that accompanies this Tutorial.

Conclusion

This Tutorial addresses L3VPN implementation using RFC 2547bis. L3VPNs are a new (or is it old?) method of building corporate private networks by leveraging the Internet infrastructure that is in place already and that continues to expand. Although L2VPNs have existed for a while utilizing the legacy Frame Relay and ATM technologies, L3VPNs promise enterprise interconnectivity at reduced operating costs. This is due to the fact that the service providers are collapsing their infrastructure into common technology -- IP, which results in reduced operating costs within a service provider space.

Although current IETF working groups and drafts present L2VPNs and L3VPNs, the choice of one versus

another will depend solidly on customer requirements. Should a customer wish to have a service provider participate in its Layer 3 routing, L3VPNs are the solution. Should a customer want the provider not to participate in the routing of its private IP traffic, L2VPNs are the solution.

When determining a suitable Layer 3 solution for a customer, several considerations must be taken into account:

1. The current VPN topology that the customer is deploying.
2. The future VPN topology -- hub-and-spoke versus fully meshed. It is most likely that a customer would want to migrate to fully meshed topology, enabling direct interconnectivity between all sites.
3. The number of customer sites.

Although a site may be in multiple VPNs, it is not necessarily the case that the route to a given system at that site should be the same in all the VPNs. Suppose, for example, we have an intranet consisting of sites A, B, and C, and an extranet consisting of A, B, C, and the "foreign" site D. Suppose that at site A there is a server, and we want clients from B, C, or D to be able to use that server. Suppose also that at site B there is a firewall. We want all the traffic from site D to the server to pass through the firewall, so that traffic from the extranet can be access-controlled. However, we don't want traffic from C to pass through the firewall on the way to the server, since this is intranet traffic.

It is possible to set up two routes to the server. One route, used by sites B and C, takes the traffic directly to site A. The second route, used by site D, takes the traffic instead to the firewall at site B. If the firewall allows the traffic to pass, it then appears to be traffic coming from site B, and follows the route to site A.

The RDs are structured so that every service provider can administer its own "numbering space" (i.e., can make its own assignments of RDs) without conflicting with the RD assignments made by any other service provider. An RD consists of three fields: a 2-byte type field, an administrator field, and an assigned number field.

The RD can also be used to create multiple different routes to the very same system. We have already discussed a situation in which the route to a particular server should be different for intranet traffic than for extranet traffic. This can be achieved by creating two different VPN-IPv4 routes that have the same IPv4 part, but different RDs. This allows BGP to install multiple different routes to the same system, and allows policy to be used to decide which packets use which route.

RDs are given this structure in order to ensure that an SP that provides VPN backbone service can always create a unique RD when it needs to do so. However, the structure is not meaningful to BGP; when BGP compares two such address prefixes, it ignores the structure entirely.

References

[Augustyn 2002] V. Augustyn et al. "Requirements for Virtual Private LAN Services (VPLS)" draft-ietf-ppvpn-vpls-requirements-01.txt

[Berkowitz 2002] Berkowitz, H. *Building Service Provider Networks*. New York: John Wiley & Sons, 2002.

[Callon 2003] R. Callon, M. Suzuki. "A Framework for Layer 3 Provider Provisioned Virtual Private Networks" draft-ietf-ppvpn-framework-08.txt March 2003.

[Carugi 2003] M. Carugi, D. McDysan. draft-ietf-ppvpn-requirements-06.txt, "Service requirements for

Layer 3 Provider Provisioned Virtual Private Networks" April 2003.

[Chen 2003] E. Chen, Y. Rekhter. "Cooperative Route Filtering Capability for BGP-4" draft-ietf-idr-route-filter-08.txt.

[Cisco 2003] Cisco Systems. "BGP Prefix-Based Outbound Route Filtering."
<http://www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s14/fsbgporf.htm>.

[Knight 2003] P. Knight et al. "Network based IP VPN Architecture using Virtual Routers" draft-ietf-ppvnpn-vpn-vr-04.txt. May 2003.

[Kompella 2003] K. Kompella, M. Leelanivas, Q. Vohra, "Layer 2 VPNs Over Tunnels" draft-kompella-ppvnpn-l2vpn-03.txt.

[Ould-Brahim 2003] H. Ould-Brahim et al. "Using BGP as an Auto-Discovery Mechanism for Provider-provisioned VPNs" draft-ietf-ppvnpn-bgpvpn-auto-05.txt. May 2003.

[Patel 2003] K. Patel, S. Hares, " Aspath Based Outbound Route Filter for BGP-4" draft-ietf-idr-aspath-orf-04.txt.

[Rekhter 2003] Y. Rekhter, E. Rosen. "Use of PE-PE GRE or IP in RFC2547 VPNs" draft-ietf-ppvnpn-gre-ip-2547-02.txt. July 2003.

[RFC 2401] "Security Architecture for the Internet Protocol." S. Kent, R. Atkinson November 1998.

[RFC 2547] "BGP/MPLS VPNs." E. Rosen, Y. Rekhter. March 1999.

[RFC 2661] Layer Two Tunneling Protocol L2TP." W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn, B. Palter. August 1999.

[RFC 2858] " Multiprotocol Extensions for BGP-4." T. Bates et al. June 2000.

[Rosen 2003a] E. Rosen, Y. Rekhter. "BGP/MPLS IP VPNs" draft-ietf-ppvnpn-rfc2547bis-04.txt May 2003.

[Rosen 2003b] E. Rosen et al. " Use of PE-PE IPsec in RFC2547 VPNs" draft-ietf-ppvnpn-ipsec-2547-03.txt February 2003.

[Sangli 2002] S. Sangli et al. "BGP Extended Communities Attribute" draft-ietf-idr-bgp-ext-communities-05.txt. May 2002.

[Vohra 2003] Q. Vohra, E. Chen. " BGP support for four-octet AS number space" draft-ietf-idr-as4bytes-06.txt. June 2003.

*[IE-L3VPN-WP1-F03]
[2003-07-31-04]*