Layer 2 and 3 Virtual Private Networks: Taxonomy, Technology, and Standardization Efforts

Paul Knight, Nortel Networks
Chris Lewis, Cisco Systems

ABSTRACT

Virtual private network services are often classified by the OSI layer at which the VPN service provider's systems interchange VPN reachability information with customer sites. Layer 2 and 3 VPN services are currently being designed and deployed, even as the related standards are being developed. This article describes the wide range of emerging L2 and L3 VPN architectures and technical solutions or approaches, and discusses the status of standards work. Some specific L2VPN and L3VPN technologies described here include virtual private LAN service, transparent LAN service, BGP/MPLS-based VPNs (RFC 2547bis), virtual router, and IPSec VPN approaches. We discuss recent and continuing standards efforts in the IETF l2vpn and l3vpn working groups, and related work in the Pseudo-Wire Emulation Edge-to-Edge working group, as well as in some other standards fora, and describe some mechanisms that provide membership, reachability, topology, security, and management functions.

INTRODUCTION AND NOTES ON TERMINOLOGY

This article provides a summary of some of the technologies and standards activities emerging in the layer 2 (L2) and layer 3 (L3) virtual private network (VPN) arenas. We provide a detailed look at the more recently emerging L2VPN areas, and take a more broadly descriptive approach to the more established L3VPN technologies.

There are many possible ways to engineer VPN services, and a variety of design approaches have been proposed or developed. Two Internet Engineering Task Force (IETF) working groups are chartered to help reduce the profusion of approaches and develop standards for VPN systems. These two working groups focus on L2VPNs and L3VPNs, and are named correspondingly (l2vpn, l3vpn). Much of this article describes the current work emerging from these working groups.

In this article, layer 3 is essentially synonymous with the Internet Protocol (IP); thus, an L3VPN is an IP VPN.

Within the context of VPNs, we use the following terms, which are defined in detail in the companion introductory article of this VPN standards series:

- Customer edge (CE) and provider edge (PE): VPN devices managed by the VPN customer and service provider, respectively
- User: An individual human using a computer
- Site: A collection of users in a local network
- VPN customer: An enterprise or organization controlling multiple sites

The Internet drafts cited in this article are categorized in Table 1 and referenced within the article as [Table 1-*n*], *n* being the number in the "Ref. #" column.

LAYER TWO VPNS

L2VPNs have existed for over 30 years. Currently, L2 services based on frame relay and asynchronous transfer mode (ATM) dominate the data revenues of the largest service providers. However, many providers are in the position of having built extensive dedicated L2 networks, separate IP-based networks, and a separate optical network; sometimes multiple public and private versions of each. This has a direct cost in terms of hardware and management. Some providers have concluded that having fewer multipurpose networks will be more cost effective than operating many special-purpose networks.

This thinking has led to interest in methods to converge L2 network infrastructures across a single infrastructure. There is a high level of interest in methods of delivering L2 services over a L3 IP network. Standards development work and product development by network equipment vendors are both proceeding rapidly.

0163-6804/04/$20.00 © 2004 IEEE

Selected VPN-related IETF Internet drafts				
VPN area	**VPN focus**	**Ref. #**	**Internet draft title**	**File name (as of April 2004)***
Layer 2	all	1	Framework for Layer 2 Virtual Private Networks (L2VPNs)	draft-ietf-l2vpn-l2-framework-04.txt
		2	Service Requirements for Layer 2 Provider Provisioned Virtual Private Networks	draft-ietf-l2vpn-requirements-01.txt
	IPLS	3	IP-Only LAN Service (IPLS)	draft-ietf-l2vpn-ipls-00.txt
	VPWS	4	Encapsulation Methods for Transport of ATM over IP and MPLS Networks	draft-ietf-pwe3-atm-encap-04.txt
		5	Frame Relay over Pseudo-Wires	draft-ietf-pwe3-frame-relay-02.txt
		6	Encapsulation Methods for Transport of Ethernet Frames over IP/MPLS Networks	draft-ietf-pwe3-ethernet-encap-05.txt
		7	Encapsulation Methods for Transport of PPP/HDLC over IP and MPLS Networks	draft-ietf-pwe3-ppp-hdlc-encap-mpls-02.txt
		8	Pseudowire Setup and Maintenance Using LDP	draft-ietf-pwe3-control-protocol-06.txt
		9	Layer 2 VPNs over Tunnels	draft-kompella-l2vpn-l2vpn-00.txt
		10	Provisioning Models and Endpoint Identifiers in L2VPN Signaling	draft-ietf-l2vpn-signaling-01.txt
		11	BGP-Based Auto-Discovery for L2VPNs	draft-hlmu-l2vpn-bgp-discovery-00.txt
		12	Using RADIUS for PE-Based VPN Discovery	draft-ietf-l2vpn-radius-pe-discovery-00.txt
	VPLS	13	Virtual Private LAN Services over MPLS	draft-ietf-ppvpn-vpls-ldp-01.txt
		14	Virtual Private LAN Service	draft-ietf-l2vpn-vpls-bgp-01.txt
	OAM	15	Pseudo Wire (PW) OAM Message Mapping	draft-nadeau-pwe3-oam-msg-map-04.txt
		16	Pseudo Wire (PW) Virtual Circuit Connection Verification (VCCV)	draft-ietf-pwe3-vccv-02.txt
		17	OAM Procedures for VPWS Interworking	draft-aissaoui-l2vpn-vpws-iw-oam-00.txt
Layer 3	all	18	Service Requirements for Layer 3 Provider Provisioned Virtual Private Networks	draft-ietf-l3vpn-requirements-00.txt
		19	A Framework for Layer 3 Provider Provisioned Virtual Private Networks	draft-ietf-l3vpn-framework-00.txt
	2547 bis	20	BGP/MPLS IP VPNs	draft-ietf-l3vpn-rfc2547bis-01.txt
		21	OSPF as the PE/CE Protocol in BGP/MPLS IP VPNs	draft-ietf-l3vpn-ospf-2547-01.txt
		22	Use of PE-PE IPsec in RFC 2547 VPNs	draft-ietf-l3vpn-ipsec-2547-02.txt
		23	Use of PE-PE GRE or IP in RFC 2547 VPNs	draft-ietf-l3vpn-gre-ip-2547-01.txt
		24	Applicability Statement for BGP/MPLS IP VPNs	draft-ietf-l3vpn-as2547-03.txt
	VR	25	Network-Based IP VPN Architecture Using Virtual Routers	draft-ietf-l3vpn-vpn-vr-01.txt
		26	Applicability Statement for Virtual-Router-Based Layer 3 PPVPN Approaches	draft-ietf-l3vpn-as-vr-01.txt
	IPSEC	27	An Architecture for Provider Provisioned CE-Based Virtual Private Networks Using IPSec	draft-ietf-l3vpn-ce-based-02.txt
	OAM	28	Framework for L3VPN Operations and Management	draft-ietf-l3vpn-mgt-fwk-01.txt
		29	Using BGP as an Auto-Discovery Mechanism for Provider-Provisioned VPNs	draft-ietf-l3vpn-bgpvpn-auto-01.txt
		30	CE-to-CE Member Verification for Layer 3 VPNs	draft-ietf-l3vpn-l3vpn-auth-00.txt
IPSEC	IPSEC	31	Internet Key Exchange (IKEv2) Protocol	draft-ietf-ipsec-ikev2-13.txt
		32	Security Architecture for the Internet Protocol	draft-ietf-ipsec-rfc2401bis-01.txt
		33	Use of IPsec Transport Mode for Dynamic Routing	draft-touch-ipsec-vpn-07.txt
		34	A Method to Provide Dynamic Routing in IPsec VPNs	draft-knight-ppvpn-ipsec-dynroute-03.txt

* For reference on the IETF Web site, please prepend "http://www.ietf.org/internet-drafts/" to the given file name. Note that the final two-digit version number is subject to frequent change; it is incremented as the Internet draft is revised. Also note that an Internet draft may become an IETF RFC document, and will then be assigned a different file name, although usually retaining the original title.

Table 1. *Selected VPN-related IETF Internet drafts.*

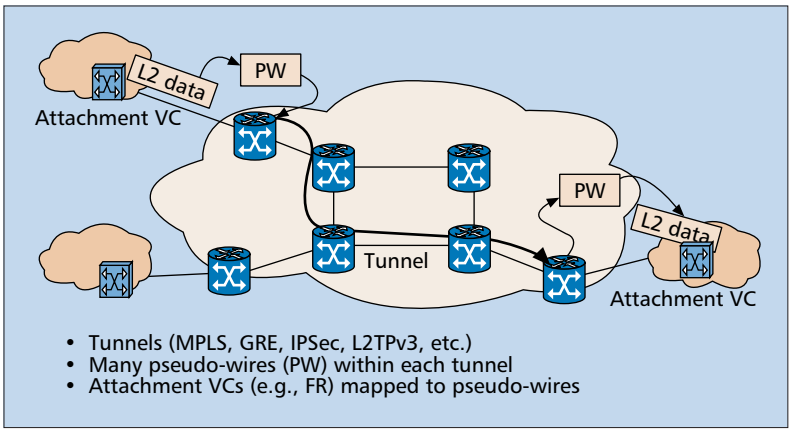

Figure 1. *L2VPN architecture*

There are essentially two service models delivered within the umbrella term of L2VPNs. The first is the point-to-point model offered by technologies like frame relay and ATM, based on distinct virtual circuits between two L3 routing endpoints. Ethernet point-to-point service is included. The second is a multipoint service, typically offered over Ethernet, that provides media access control (MAC) address learning and packet replication to make the WAN service offered by the service provider appear as if it is a LAN switch connected to each site. In terms of current deployed ports and provider revenue, the first application dominates the market.

The primary application of L2VPN technology today has been the creation of hub and spoke, full mesh, or partial mesh topologies built from point-to-point connections to interconnect VPN sites. This differs from the L3VPN model that offers a point-to-cloud type service, where each site's CE connection to the provider network simply sees an adjacent IP router as a gateway to all destinations.

The remainder of this section examines the taxonomy of the emerging L2 over packet services, their high-level architecture, and supporting standards work.

L2 VPN Architecture

The L2VPN framework is fully defined in [Table 1-1, 1-2]; however it may be summarized with reference to Fig. 1. L2 traffic within a specific attachment virtual circuit — frame relay, ATM, or Ethernet virtual LAN (VLAN) — is carried across a packet (IP or MPLS) core network in pseudo-wires (packet-based emulations of leased line, frame relay, ATM, or Ethernet physical wires), which are themselves transported within tunnels. This approach helps the service scale to support very large numbers of customers, each with many sites, since the routing devices within the packet network only need to know about tunnels between edge devices, not all the individual L2 connections that exist between the two edge devices. In overview, the L2 payload is encapsulated with a pseudo-wire header, which is further encapsulated within a tunnel header. The pseudo-wire header acts as the demultiplexer field at the tunnel termination.

L2VPN-over-Packet Taxonomy

Figure 2 defines the terms used when categorizing L2VPN-over-packet technologies. The same set of services can be delivered over an IP network using either IP-based or multiprotocol label switching (MPLS) tunneling. The two categories of L2VPN are the point-to-point virtual private wire service (VPWS) and multipoint virtual private LAN service (VPLS). VPWS provides emulation of L2 connections for frame relay data link channel identifiers (DLCIs), ATM permanent virtual circuits (PVCs), leased line formats, and Ethernet.

In the case of Ethernet VPWS, the service can be either Ethernet relay service (ERS) or Ethernet wire service (EWS). ERS takes into account Ethernet virtual LAN (VLAN) numbering and offers a frame-relay-like service to the end user, with the VLAN number replacing the function of the frame relay DLCI. EWS delivers port-based service, such that anything transmitted via the port is transported over a pseudowire, and no specific treatment of VLAN numbers occurs.

The other major category is the virtual private LAN service (VPLS), which is a multipoint Ethernet service. This is still based on pseudowire creation between PE devices, but adds MAC address learning, MAC-based forwarding and packet replication to the basic VPWS service to make the provider service operate like a LAN switch, from the perspective of the end user.

The final L2VPN type is the IP-only LAN service (IPLS), which is a subset of the generalized VPLS case, defined in [Table 1-3]. In IPLS, frames are forwarded based on their MAC destination addresses. However, the maintenance of the MAC forwarding tables is done via signaling, rather than via the MAC address learning in the data plane procedures of IEEE 802.1D. Further, Address Resolution Protocol (ARP) messages are proxied, rather than carried transparently. A limitation of IPLS is that it requires that CE devices be routers rather than Ethernet switches, which are supported by the VPLS architecture.

VPWS Technologies

A VPWS VPN is a collection of L2 circuits or pseudo-wires. Creating the pseudo-wires to form a L2VPN requires defining an encapsulation for pseudo-wire and tunnel transport, a control plane for session management and error notification, and potentially some auto-discovery capabilities. For deployment, it needs some provisioning and operation, administration, and maintenance (OAM) capabilities for service management. Interworking and OAM are discussed in a later section.

The encapsulation methods for each of the VPWS layer 2 transports (ATM, frame relay, Ethernet, and ppp-hdlc) are specified in [Table 1-4–7]. The content of tunnel and pseudo-wire encapsulation headers needs to be signaled across the packet network, along with circuit status information. Signaling may use point-to-point or broadcast mechanisms. The point-to-point approach [Table 1-8] requires

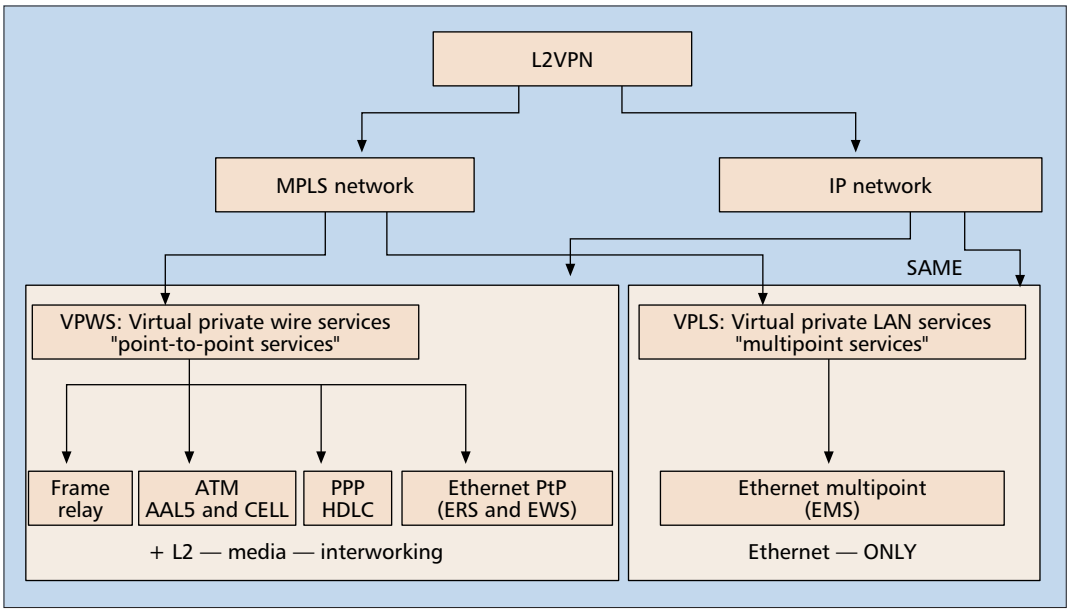

Figure 2. *L2 VPN service taxonomy: VPWS and VPLS.*

direct control sessions between each pair of PEs to support the exchange of control data. The broadcast approach [Table 1-9] leverages the existing Border Gateway Protocol (BGP) infrastructure within provider networks, which enables a single control channel to a BGP route reflector to be used to send a single copy of the information to all other PE devices.

It is technically possible for both approaches to work, and the debate on which is most efficient and best to deploy continues. The central issue is the debate about whether data that needs to be signaled as part of a VPWS VPN needs to go to all other members of that VPN. If the VPWS topology is a full mesh, there are good arguments for sending signaling data to all VPN members. If it is not a full mesh, the requirement for broadcasting signal elements to all VPN peers is less compelling. Currently fewer than 20 percent of deployed L2 VPNs utilize a full mesh topology in provider networks.

In terms of provisioning new packet-based L2VPN transports, there are also two approaches. The first replicates existing practice, using manual or management-system-based processes. The second is an auto-provisioning approach that is defined using either a broadcast signaling mechanism [Table 1-9] or a point-to-point signaling mechanism [Table 1-10].

Individual provisioning of pseudo-wires requires provisioning of the attachment circuits (ACs) at each PE/CE interface, a pseudo-wire specifying a virtual circuit identifier (VC-ID) and the remote PE. The VC-ID must be unique between the pair of PEs. The auto-provisioning model is based on the "colored pools" concept, meaning that each PE will have separate pools of circuit identifiers for each VPN (i.e., different VPNs are associated with different "colors"). This mechanism also requires the provisioning of ACs, but each AC is placed in a pool with a specific color. Through auto-discovery, each PE learns about all other PEs with pools with the same color.

Consider PE1, with a green pool of attachment circuits, and PE2, also with a green pool of ACs. When PE1 discovers PE2, PE1 removes an AC from its green pool. PE1 signals a pseudo-wire creation to PE2, specifying the green pool as the target, and binds the removed AC to the pseudo-wire. PE2 removes an AC from its green pool and binds it to the newly created pseudo-wire. The result is a full mesh of pseudo-wires between all the ACs of the same VPN, in this case the green VPN.

Proposals for using BGP as the auto-discovery mechanism are [Table 1-9, 1-11]. RADIUS is also being considered for this function [Table 1-12].

VPLS Technologies

Like VPWS, VPLS is built on L2 pseudo-wires. However, source MAC address learning, flooding of unknown, broadcast, and unicast frames, and forwarding based on MAC addresses are also defined, so the VPLS service appears to operate like a LAN switch to the end user. This is illustrated in Fig. 3.

The key point for VPLS service is that the pseudowires connecting VPN sites must use a full mesh so that split horizon can be used to prevent loops. This obviates the need for implementing the spanning tree protocol within the IP/MPLS core. The split horizon rule is used to stop loops, by preventing a PE from forwarding data received from one pseudo-wire within a VPLS to any other pseudo-wire; it is only forwarded to attachment circuits. Clearly, for this to function a full mesh of pseudo-wires per VPLS is required.

Work defining VPLS standards originated in the IETF, spread to the IEEE and then to the Metro Ethernet Forum (MEF). This work has now been organized along the following lines: The MEF covers service definition and user-to-network interface (UNI) aspects of the service, the IEEE deals with bridging aspects of the ser-

The key point for VPLS service is that the pseudowires connecting VPN sites must use a full mesh, so that split-horizon can be used to prevent loops. This obviates the need for implementing the spanning tree protocol within the IP/MPLS core.

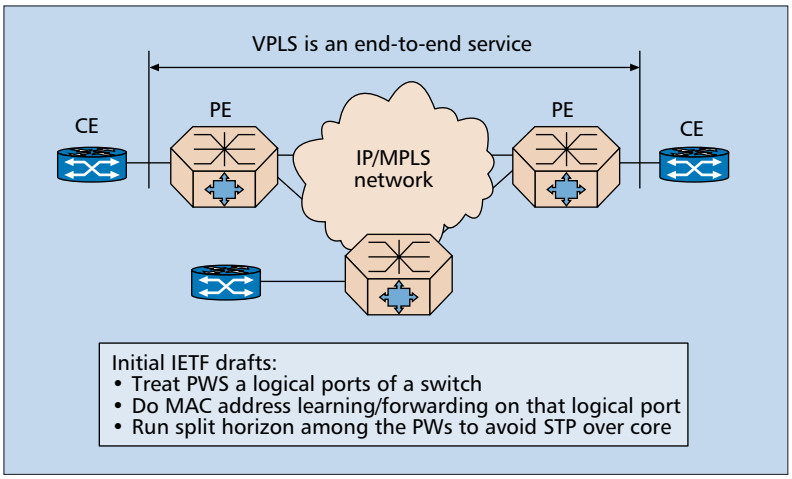

Figure 3. *Simplified VPLS architecture.*

vice and the network, and the IETF deals with provisioning of the service over the IP/MPLS portion of the network.

As in the VPWS case, there are proposals for VPLS implementation that are based on point-to-point signaling mechanisms [Table 1-13] and BGP-based broadcast mechanisms [Table 1-14].

[Table 1-13] also specifies the use of an access network based on tagging customer Ethernet frames with an IEEE 802.1q tag, to help with scalability of the service and cost reduction by bringing in lower-cost Ethernet switches as aggregation devices. This concept is illustrated in Fig. 4.

In this figure the more expensive MPLS-capable PE devices do not always extend to the edge of the provider's network, where customer sites are located. Where Ethernet aggregators exist, 802.1q tagged customer frames are given an additional 802.1q outer tag unique to that customer within the network. This process is called q-in-q tagging. There are inherent scalability issues with q-in-q tagging, as there are only a maximum possible 4096 tag values within the 802.1q header. Limiting the q-in-q domain to the access rather than the entire network and separating the q-in-q domains with an MPLS network are methods to alleviate this issue.

Interworking and OAM

The initial implementations of L2VPNs required the attachment circuits at both ends of the pseudo-wire to be the same L2 technology. This closely matched the service offerings of existing L2 networks. However, particularly with the introduction of metropolitan Ethernet services, the need to interwork between Ethernet and other L2 technologies has arisen. A typical example is a corporation with headquarters and main offices serviced by a MAN offering an L2 Ethernet service, while only frame relay service is available for other VPN sites. For this case the network needs to provide interworking between the Ethernet and frame relay connections at L2. Interworking requirements are discussed in detail in [Table 1-2].

L2VPN OAM standards work [Table 1-15, 1-16] is directed toward how to operationally verify these services once deployed on a provider network. Virtual circuit connection verification (VCCV) is a key technology for this requirement, as it supports connection verification applications for pseudo-wire VCs regardless of the underlying MPLS or IP tunnel technology. [Table 1-17] discusses OAM for VPWS in service interworking scenarios.

Current IETF Work Items for L2VPNs

Current IETF work items focus on management functions to support widescale deployment and operational support of these services. The L2VPN charter provides specific milestones and goals, and links to many of the documents cited here.

L3 VPN Architectures

As described in the L3VPN service requirements [Table 1-18] and framework [Table 1-19], L3VPNs can be either network- or CE-based

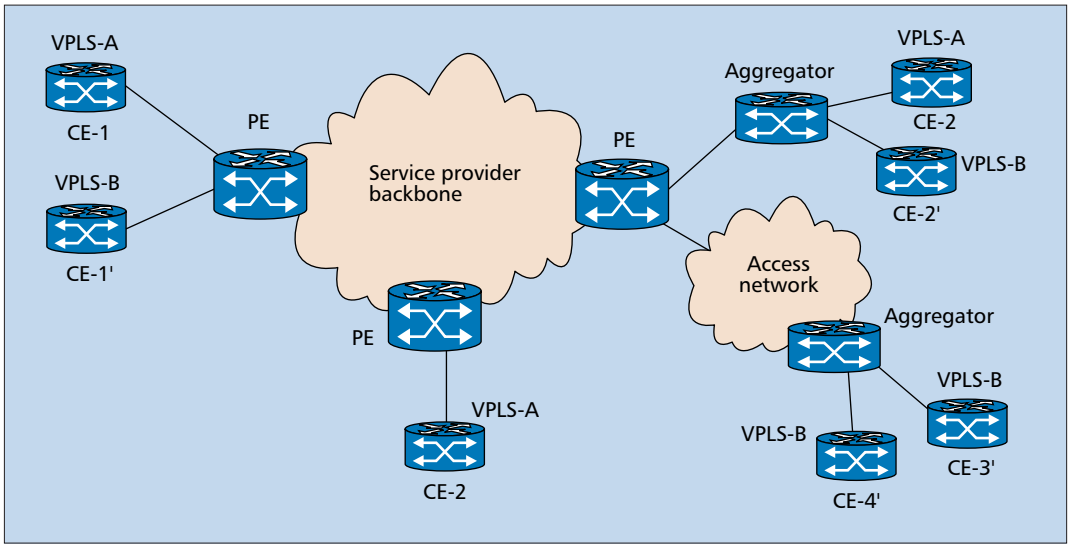

Figure 4. *VPLS system with distributed-PE.*

services. In network-based L3VPNs, a VPN service provider must configure its PE devices to establish the VPN connections for each customer site. The customer's CE device only needs to be able to perform ordinary routing functions. In contrast, CE-based L3VPNs can be constructed by customers without any specific involvement of a service provider beyond the provision of Internet access. However, the CE devices must provide all of the capabilities needed for the VPN.

From the IETF's l3vpn working group, three approaches have emerged as standards-track proposals. These are BGP/MPLS IP VPNs (known as RFC 2547bis, since it extends RFC 2547) [Table 1-20], virtual router IP VPNs [Table 1-25], and CE-based IPSec VPNs [Table 1-27].

Figure 5 depicts several common elements of the two network-based L3VPN approaches (RFC 2547bis and VR):

- Both approaches provide methods for storing and maintaining separation among the potentially overlapping address spaces of multiple VPN customers. For the VR approach, this is the routing table of the individual VR configured for each VPN; RFC 2547bis specifies a VPN routing and forwarding (VRF) table for each VPN. Generically, these mechanisms are virtual forwarding instances (VFIs).
- Both provide similar mechanisms for learning the reachability information (the IP networks or subnets) from each site. Each can use standard IP routing protocols or static configuration.
- Both address the requirement to distribute customer site reachability (or routes) across the VPN, and facilitate VPN membership discovery: how individual PEs or VRs discover their counterparts supporting other sites within a VPN.
- Both support the construction of tunnels between the PEs or VRs supporting the customer sites in a VPN. Tunnels simplify the requirement to maintain separation of data traffic among the various VPNs in an SP's network. The VR approach may use any kind of tunneling mechanism supported by the SP's network, including ATM or frame relay circuits, while RFC 2547bis only defines the use of MPLS- or IP-based tunneling.

Both the RFC 2547bis and VR L3VPN approaches are able to make use of BGP-based auto-discovery [Table1-29] to provide VPN membership capabilities. BGP-based auto-discovery is a mechanism that makes it possible for a PE router (or VR) handling connections to a particular VPN to discover all the other PE routers (or VRs) that are connected to the same VPN. For BGP auto-discovery to work, each PE/VR uses BGP to advertise that it is connected to a specific VPN. All other PE/VRs in the network, including the ones supporting that VPN, receive these BGP advertisements.

RFC2547bis L3VPNs

RFC 2547bis [Table 1-20] describes how an SP with an MPLS-enabled IP backbone can supply L3VPNs for customers. RFC 2547bis describes the interaction between CE routers and PE routers as peers. The CEs send routes from the customer site to the PEs, typically using a routing protocol. The use of BGP [Table 1-20] and OSPF [Table 1-21] are defined, as well as RIP and static configuration.

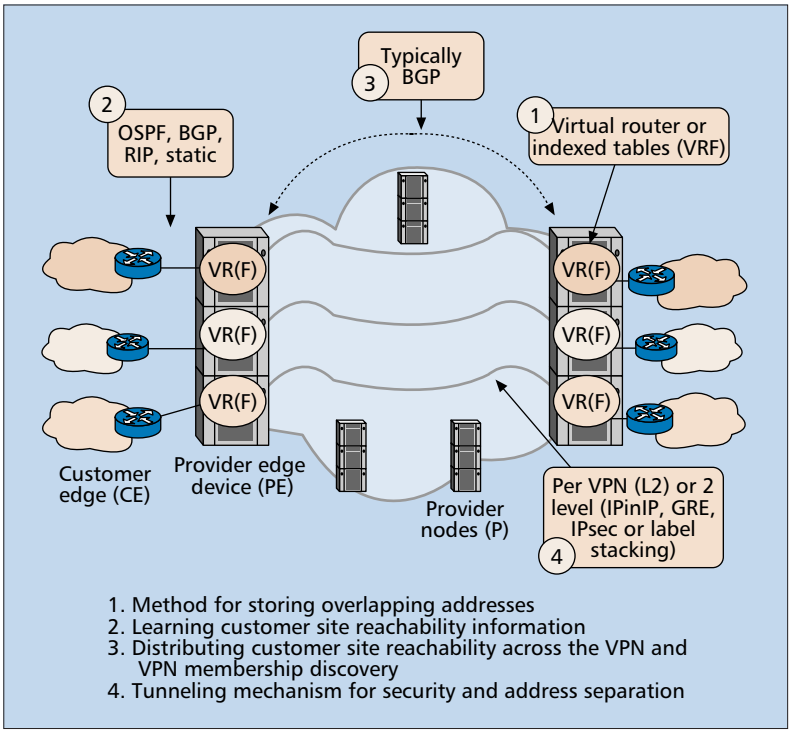

Figure 5. *Common components of network-based L3VPNs.*

In RFC 2547bis, the SP must use BGP in the core network. BGP interchanges the routes of each VPN among all the PE routers in that VPN. The RFC 2547bis approach specifies the use of MPLS labels to identify and separate the traffic of different VPNs. This ensures that all the routes from various VPNs are able to maintain separation, even in cases where multiple VPNs may share the same IP addresses.

The PE router delivers the routes from all of the other CE routers in that VPN to each connected CE. A key concept is that a CE router which belongs to a VPN does not establish any kind of routing relationship directly with the other CE routers in that VPN.

RFC 2547bis depends on MPLS mechanisms in the SP's network, and defines how MPLS labels and BGP, acting as a kind of distributed database, provide the mechanisms to deliver VPN services. Every VPN route is associated with an MPLS label, and when a VPN route is propagated through the SP's network by BGP, the MPLS label is propagated along with the route.

The MPLS label is used as a short tag (i.e., shorter than an IP destination address) that is part of an outer wrapper (or encapsulation) for each data packet, and allows it to follow a predetermined path through the SP's network with minimal processing (and thus minimal delay). A customer VPN data packet being sent through an SP's backbone is encapsulated using the MPLS label assigned to its destination address, within the context of that VPN. At this point, it is referred to as an MPLS packet.

RFC 2547bis establishes tunnels across the

With any tunnel type, one extremely important result of the use of tunnels and the MPLS encapsulation is that the backbone routers do not have to be burdened with any knowledge of the routes of the individual VPNs.

SP backbone between PEs within a VPN, and the MPLS packets travel through these tunnels to the proper destination PE. RFC 2547bis does not necessarily require MPLS tunnels across the SP backbone between PEs. The tunnels can use IPSec [Table 1-22]. Other IP-based or Generic Routing Encapsulation (GRE)-based encapsulation is also defined [Table 1-23]. With any tunnel type, one extremely important result of the use of tunnels and MPLS encapsulation is that the backbone routers need not be burdened with any knowledge of the routes of the individual VPNs.

Virtual Router L3 VPNs

The second major architecture being standardized for network-based L3VPNs is named for its key component, the virtual router. A VR provides all the capabilities of an individual physical router, but is actually an emulation of a router, instantiated as needed on an SP's PE device. Multiple VRs can exist on a single PE device. Several PE designs implement the VR concept with different internal architectures, but all provide VR-based L3VPN services with a well defined set of attributes, as described in [Table 1-25].

As an emulation of a physical router, a VR provides the same kinds of mechanisms, with related management tools, as a physical router. It provides for configuration of logical interfaces and routing relationships among the interfaces, along with a means of mapping the logical interfaces to the PE's physical interfaces. The VRs represent different VPNS, so they must maintain independent routing tables, and they are logically isolated so that the operation of a VR will not affect other VRs on that PE. Since the VRs are independent, a VPN's address space can overlap that of another VPN.

From the viewpoint of the customer's CE, the VR appears as a neighboring router within the customer's network. The CE interchanges routing (reachability) information with the VR. The CE sends data traffic addressed to other VPN sites via the VR. The VR connected to a site's CE is responsible for interchanging routing information with all the other VRs in that VPN.

RFC2547bis and VR Differences

A major distinction between the VR and RFC 2547bis approaches is the way they handle membership and reachability processes. Membership refers to the process by which each PE or VR learns about the other PE/VRs which are in its particular VPN. Reachability is the function used by a VPN model to inform the PE/VRs and their attached CEs about all the IP networks or subnetworks that exist in the VPN, and how to reach them.

To disseminate reachability information, the VR L3VPN model emulates individual routers communicating directly with the other routers in their own VPN, using standard routing protocols. It does this by constructing tunnels between the VRs in the VPN, using any tunneling capabilities supplied by the SP's network, including frame relay or ATM circuits; IPSec, IP-in-IP, or GRE tunnels; or MPLS paths. Each VR-based VPN can use its routing protocol of choice across the SP backbone, independent of other VPNs or the SP's own backbone routing protocol. The customer's site routing instances can be preserved across the SP backbone in most VR VPN implementations, subject to the capabilities of the tunneling mechanism employed.

To provide membership functions, the VR VPN model can use a variety of methods, including configuration via a management system, BGP-based auto-discovery, a VPN member directory server system, or the same mechanism used by RFC 2547bis. Each VR within a VPN is configured with a common VPN identifier, used as part of initiating the membership process.

In contrast to the VR model's methods of handling membership and reachability, the RFC 2547bis model "piggybacks" both reachability and membership information onto the BGP routing protocol used within the SP's core network. It uses extensions to the base BGP protocol, which were initially designed to carry non-IP routing information, to transport this information, essentially using the BGP routing protocol running across the SP's backbone network as an immense distributed database. Since all the reachability information is piggybacked onto the backbone BGP, the RFC 2547bis model must terminate the customer's VPN network layer at the PE connected to each site.

For detailed point-by-point comparisons of the L3VPN approaches, Applicability Statements [Table 1-24, 1-26] are being developed for each.

IPsec-based L3VPNs

The third major type of L3VPN approach is quite distinct from the VR and RFC2547bis models. It interconnects CE devices at the various VPN sites via the Internet, exchanging IP packets among the sites. In order to provide security for the VPN traffic as it travels through the Internet, almost all CE-based VPNs use IPSec. The Internet can provide L3 (IP) connectivity among customer sites without additional effort or expense by an SP to establish interconnections among locations. Thus, a CE-based IPSec VPN can be built and managed by an enterprise, or provided by an SP. [Table 1-27] describes an architecture that can be used by an SP to enable it to scale to serve large numbers of VPN customer sites using the CE-based IPSec model.

Although the basic IPSec standards in the IETF have provided the framework for the development of a large number of interoperable IPSec implementations, even these foundational standards are not finalized. The Internet Key Exchange (IKE), which provides the methodology for securely establishing and exchanging cryptographic keys, is being revised as IKEv2 [Table 1-31]. The IPSec Architecture (RFC 2401) is being rewritten, with the current revision known as RFC 2401bis [Table 1-32], and updates to the Encapsulating Security Protocol (ESP) and authentication header (AH) are underway as well.

Several areas of importance to IPSec VPNs in particular (as opposed to individual IPSec

connections) have not yet been fully specified by the IETF in an interoperable manner. These include:

Tunnel mode vs. transport mode: Although the current RFC 2401 specifies that only tunnel mode may be used between IPSec gateways (CEs), it has been argued that transport mode applied to IP-in-IP encapsulated data provides equivalent security and better properties for supporting dynamic routing. RFC 2401bis [Table 1-32] now allows transport mode between gateways, but there is still significant debate over how the two modes interact with dynamic routing protocols. [Table 1-33] is a major contribution supporting the use of transport mode, and [Table 1-34] adds additional support, but standards-track work addressing the IPSec VPN tunnel/transport mode issue is not yet evident.

Digital certificates and IPSec: The use of digital certificates tied to a public key infrastructure (PKI) offers enticing security benefits for IPSec VPNs, but the myriad options for configuring certificates, along with the limited facilities for automated negotiation within IKE, have led to few interoperable implementations of certificates in IPSec VPN environments. A new IETF working group, pki4ipsec, is addressing some of these issues for both IKE and IKEv2.

Dynamic routing: There are no clear standards for transporting and applying routing protocols between IPSec gateways, and the interaction of dynamic routing with the firewall-like functionalities of IPSec, particularly with tunnel mode, is likewise underspecified.

VPN-wide IPSec key establishment: IPSec is specified as a pairwise relationship, involving an elaborate and computationally intensive courtship to establish each security association and the keys used in it. Methodologies for securely managing the distribution and usage of VPN-wide keying material could improve the responsiveness of IPSec VPNs for time-sensitive but infrequent site-to-site connections, particularly for voice over IP applications.

L3 OAM

A number of operations and management issues common to all L3VPN approaches are discussed within a comprehensive framework in [Table 1-28]. Additional issues have been addressed in other documents. Auto-discovery of VPN sites, a significant labor-saving mechanism in VPN management, is presented in [Table 1-29], while a method of verifying VPN membership, for instance, in case of accidental cross-connection of VPN sites by an SP, is described in [Table 1-30].

ACKNOWLEDGMENTS

The authors gratefully acknowledge the helpful review and comments of Hamid Ould-Brahim.

BIOGRAPHIES

PAUL KNIGHT (paul.knight@nortelnetworks.com) is chair of the IETF pki4ipsec working group, and author or editor of several current IETF Internet drafts in the VPN, routing, and network security areas. He is a standards advisor with Nortel Networks, and has 20 years' experience as a network designer and network security consultant. He contributes to Web services and related standardization efforts in the Open Mobile Alliance, Broadband Content Delivery Forum, and other organizations.

CHRIS LEWIS (chrlewis@cisco.com) is a consulting systems engineer with Cisco Systems Inc. He has a B.Sc. in electrical engineering from Brunel University, Uxbridge, United Kingdom, and an M.B.A. from City University Business School, London, United Kingdom. He has over 15 years' experience with networks at financial institutions, market data providers, and now at Cisco Systems. He is the author of two networking books published by McGraw Hill and was previously a contributing editor at *Network Computing Magazine*. His areas of focus now include VPNs, MPLS technologies, and IP QoS.

Methodologies for securely managing the distribution and usage of VPN-wide keying material could improve the responsiveness of IPsec VPNs for time-sensitive but infrequent site-to-site connections, particularly for Voice over IP applications.