

VPLS Technical Tutorial

One of the main reasons why Multi-Protocol Label Switching (MPLS), generally accepted as the de-facto convergence technology, is so attractive to service providers, is that MPLS facilitates the deployment and management of virtual private networks (VPNs). While MPLS-based layer 3 VPNs have been gaining considerable momentum in the industry for some time, many now acknowledge that MPLS-based layer 2 VPNs, particularly Virtual Private LAN Services (VPLSs), will become very important in a service provider's service offering. VPLS enables service providers to offer multipoint Ethernet Virtual LAN services to a large number of customers. This technical tutorial describes the basic operation of VPLS and the requirements on the different involved network elements.

Table of contents

1	Introduction
1	VPLS over MPLS: solution overview
3	How does VPLS work?
4	Creating the PWs
4	MAC learning and packet forwarding
6	H-VPLS
7	Inter-metro service
8	Conclusion
8	References
9	Abbreviations

Introduction

Virtual private networks (VPNs) have evolved considerably since their introduction in the early 1980s, when they were built using dedicated leased lines. Frame Relay, which was introduced in the 1990s, is today the predominant VPN offering worldwide.

After the introduction of Multi-Protocol Label Switching (MPLS) in the late 1990s, a number of new VPN types were defined. The service providers' acceptance of MPLS as the network convergence technology of choice led to considerable attention being paid to MPLS-based VPNs, which offer easy service delivery within service providers' networks as well as service delivery to the users.

The various types of MPLS-based VPNs can be classified in a number of ways. One straightforward way is to base the classification on the service being offered to the customer. Typically this is either a layer 2 [1,2] or a layer 3 point-to-point service or multipoint service. This results in the following interesting VPN types:

- Layer 3 multipoint VPNs or Internet Protocol (IP) VPNs; these are often referred to as Virtual Private Routed Networks (VPRNs).
- Layer 2 point-to-point VPNs, which basically consist of a collection of separate virtual leased lines (VLLs) or pseudo wires (PWs).
- Layer 2 multipoint VPNs, or VPLSs, as discussed in this paper.

MPLS-based IP VPNs, which were introduced some years ago by the early adopters, are currently enjoying healthy growth. The two strong points of this VPN service are its multipoint nature and its support for IP. VLLs, which were introduced more recently, offer a clear migration for traditional Frame Relay/Asynchronous Transfer Mode (FR/ATM) VPNs to the converged MPLS network without replacing the customer premises equipment and without affecting the customer's service experience.

Although VPLSs were only recently introduced, already numerous operators are offering them commercially. Like MPLS-based IP VPNs, VPLS is a multipoint service, but unlike IP VPNs it can transport non-IP traffic; it also leverages the well-known advantages of Ethernet. VPLS is also used within a service provider's network to aggregate services for delivery to residential and enterprise customers.

This paper focuses on the basics of VPLS and Hierarchical VPLS (H-VPLS) as they are described in standardization forums and widely supported by leading vendors; Alcatel was a founder of VPLS and H-VPLS. The company's innovations and solutions are the topic of a separate paper in the Alcatel Telecommunications Review [3].

VPLS over MPLS: solution overview

VPLS, also known as Transparent LAN Service (TLS) or E-LAN service, is a layer 2 multipoint VPN that allows multiple sites to be connected in a single bridged domain over a provider managed IP/MPLS network [4]. All customer sites in a VPLS instance (that is, a VPLS for a particular enterprise) appear to be on the same Local Area Network (LAN), regardless of their locations. VPLS uses an Ethernet interface with the customer, simplifying the LAN/WAN (Wide Area Network) boundary and allowing rapid and flexible service provisioning.

A VPLS-capable network consists of customer edges (CEs), provider edges (PEs) and a core MPLS network:

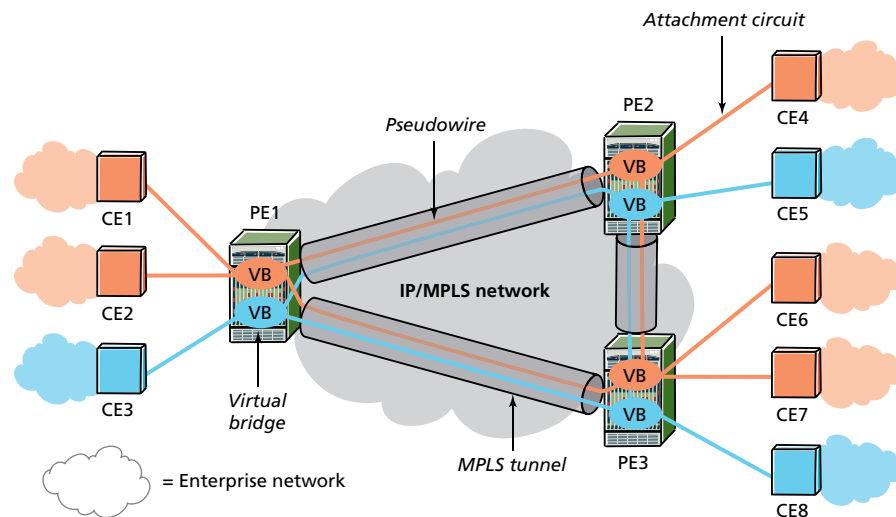
- The CE device is a router or switch located at the customer's premises; it can be owned and managed by the customer, or owned and managed by the service provider. It is connected to the PE through an Attachment Circuit (AC). In the case of VPLS, it is assumed that Ethernet is the interface between the CE and the PE.

- The PE device is where all the VPN intelligence resides, where the VPLS originates and terminates, and where all the necessary tunnels are set up to connect to all the other PEs. As VPLS is an Ethernet layer 2 service, the PE must be capable of Media Access Control (MAC) learning, bridging and replication on a per-VPLS basis.
- The IP/MPLS core network interconnects the PEs; it does not really participate in the VPN functionality. Traffic is simply switched based on the MPLS labels.

The basis of any multipoint VPN service (IP VPN or VPLS) is the full mesh of MPLS tunnels (Label Switched Paths [LSPs], also called outer tunnels) that are set up between all the PEs participating in the VPN service. The Label Distribution Protocol (LDP) is used to set up these tunnels; alternatively, the Resource Reservation Protocol – Traffic Engineering (RSVP-TE) or a combination of LDP and RSVP-TE can be used. Multipoint VPNs can be created on top of this full mesh, hiding the complexity of the VPN from the backbone routers.

For every VPLS instance, a full mesh of inner tunnels (PWs) is created between all the PEs that participate in the VPLS instance. An auto-discovery mechanism locates all the PEs participating in a given VPLS instance. This mechanism is not specified in the draft specification, so the service provider can either configure the PE with the identities of all the other PEs in a given VPLS, or can select its preferred auto-discovery mechanism, for example, the Remote Authentication Dial-In User Service (RADIUS).

Figure 1. VPLS reference model



Pseudo wire technology is standardized by the Internet Engineering Task Force (IETF) Pseudo Wire Emulation Edge to Edge (PWE3) Working Group [5]. PWs are historically also known as Martini tunnels, and the extensions to the LDP protocol to allow signaling of PWs are sometimes called Martini signaling.

A PW consists of a pair of point-to-point single hop unidirectional LSPs in opposite directions, each identified by a PW label, also called a virtual connection (VC) label. PW labels are exchanged between a pair of PEs using the targeted LDP signaling protocol. The VPLS identifier is exchanged with the labels, so that both PWs can be linked and be associated with a particular VPLS instance. Note that this exchange of PW labels has to take place between each pair of PEs participating in a given VPLS instance, and that the PW labels only have local significance between each pair of PEs.

The creation of PWs with a pair of LSPs enables a PE to participate in MAC learning: when the PE receives an Ethernet frame with an unknown source MAC address, the PE knows on which VC it was sent.

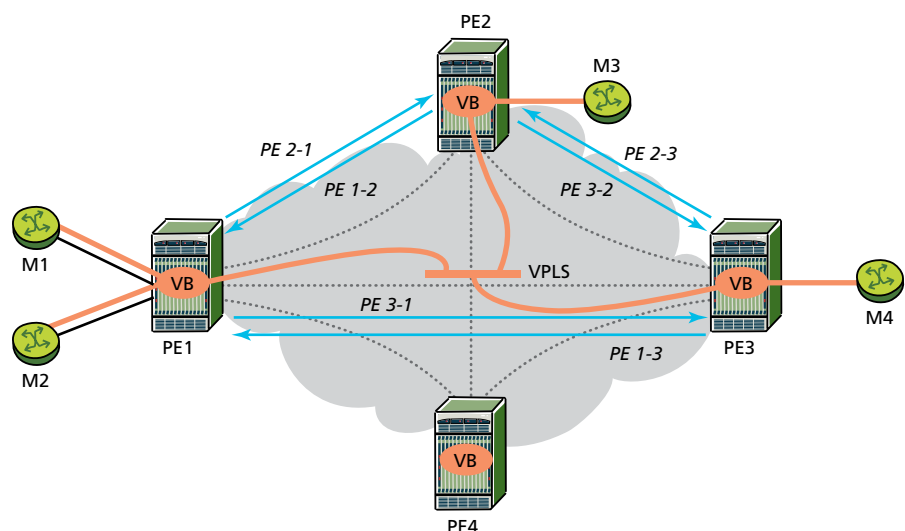
The PE routers must support all classical Ethernet features, such as MAC learning, packet replication and forwarding. They learn the source MAC addresses of the traffic arriving on their access and network ports. From a functional point of view, this means that the PEs must implement a bridge for each VPLS instance; this is often called a virtual bridge (VB), as shown in Figure 1. The VB functionality is realized in the PE through a Forwarding Information Base (FIB) for each VPLS instance; this FIB is populated with all the learned MAC addresses. All traffic is switched based on MAC addresses and forwarded between all participating PE routers using the LSP tunnels. Unknown packets (that is, the destination MAC address has not been learned) are replicated and forwarded on all LSPs to the PE routers participating in that service until the target station responds and the MAC address is learned by the PE routers associated with that service.

To prevent forwarding loops, the so-called split horizon rule is used. In the VPLS context, this rule basically implies that a PE must never send a packet on a PW if that packet has been received from a PW. This ensures that traffic cannot form a loop over the backbone network using PWs. The fact that there is always a full mesh of PWs between the PE devices ensures that every destination within the VPLS will be reached by a broadcast packet.

How does VPLS work?

It is assumed here that there is a full mesh of MPLS tunnels between the four PEs connected to the MPLS network. A VPLS instance identified by Service-identifier 101 (Svc-id 101) has to be created between PE1, PE2 and PE3; PE4 does not participate in the considered VPLS instance. Assume that this configuration was determined using an unspecified auto-discovery mechanism. M1, M2, M3 and M4 are end-stations at different customer sites and their ACs to their respective PE devices (see Figure 2) have been configured in the PEs to belong to a particular VPLS instance: Svc-id 101.

Figure 2. PW signaling



PE1 -> PE2: For SVC-ID 101 use VC label PE 2-1
 PE2 -> PE1: For SVC-ID 101 use VC label PE 1-2
 PE1 -> PE3: For SVC-ID 101 use VC label PE 3-1
 PE3 -> PE1: For SVC-ID 101 use VC label PE 1-3
 PE3 -> PE2: For SVC-ID 101 use VC label PE 2-3
 PE2 -> PE3: For SVC-ID 101 use VC label PE 3-2

Creating the PWs

Three PWs need to be created, each consisting of a pair of unidirectional LSPs or VCs. For VC-label signaling between PEs, each PE initiates a targeted LDP session to the peer PE and communicates to the peer PE which VC label to use when sending packets for the considered VPLS. The specific VPLS instance is identified in the signaling exchange using a service identifier (for example, Svc-id 101). In the example that follows, PE1 indicates to PE2: “If you have traffic to send to me for Svc-id 101, use VC label pe2-1 in the encapsulation of the packets”. Likewise, PE2 indicates to PE1: “If you have traffic to send to me for Svcid 101, use VC label pe1-2 in the encapsulation of the packets”. In this way, the first PW is created.

MAC learning and packet forwarding

Once the VPLS instance with Svc-id 101 has been created, the first packets can be sent and the MAC learning process starts. Assume M3 is sending a packet to PE2 destined for M1 (M3 and M1 are each identified by a unique MAC address), as shown in Figure 3:

- PE2 receives the packet and learns (from the source MAC address) that M3 can be reached on local port 1/1/2:0; it stores this information in the FIB for Svc-id 101.
- PE2 does not yet know the destination MAC address M1, so it floods the packet to PE1 with VC label pe2-1 (on the corresponding MPLS outer tunnel) and to PE3 with VC label pe2-3 (on the corresponding MPLS outer tunnel). The packet format is shown in Figure 4.
- PE1 learns from VC label pe2-1 that M3 is behind PE2; it stores this information in the FIB for Svc-id 101.
- PE3 learns from VC label pe2-3 that M3 is behind PE2; it stores this information in the FIB for Svc-id 101.
- PE1 strips off label pe2-1, does not know the destination M1 and floods the packet on ports 1/1/1:100 and 1/1/1:200; PE1 does not flood the packet to PE3 because of the split horizon rule.
- PE3 strips off label pe2-3, does not know the destination M1 and sends the packet on port 1/1/2:0; PE3 does not flood the packet to PE1 because of the split horizon rule.
- M1 receives the packet.

Figure 3. VPLS learning

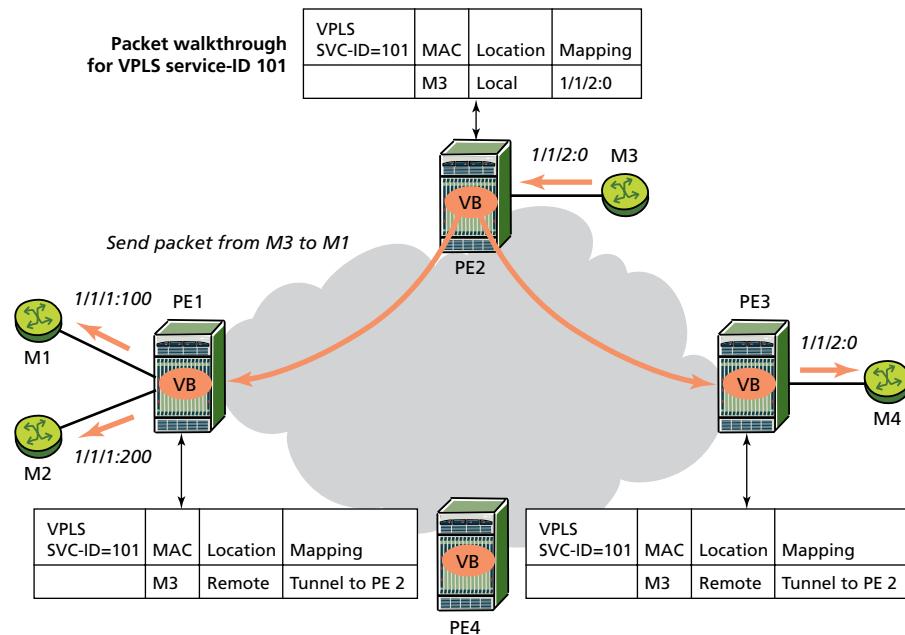
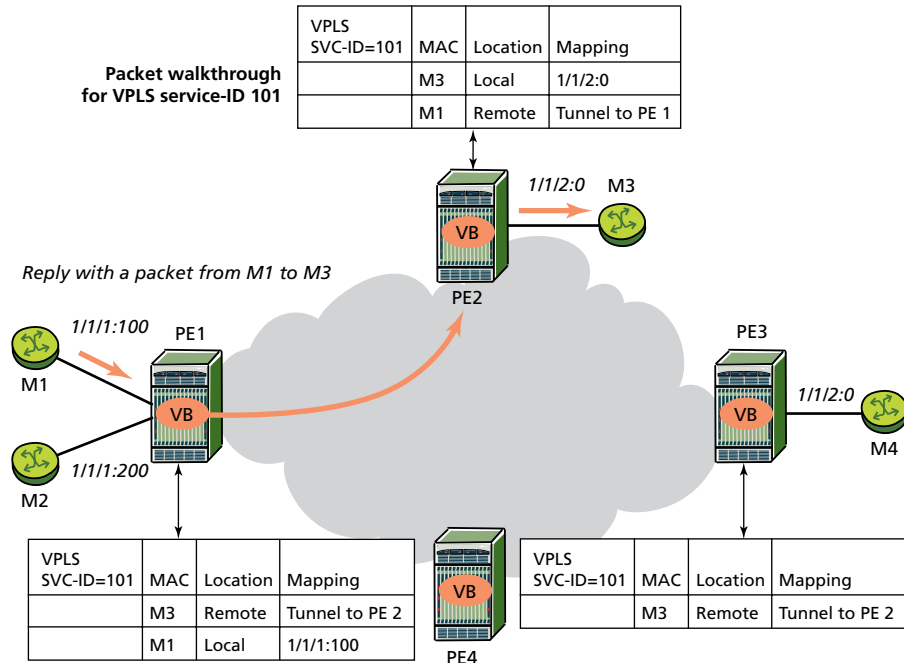


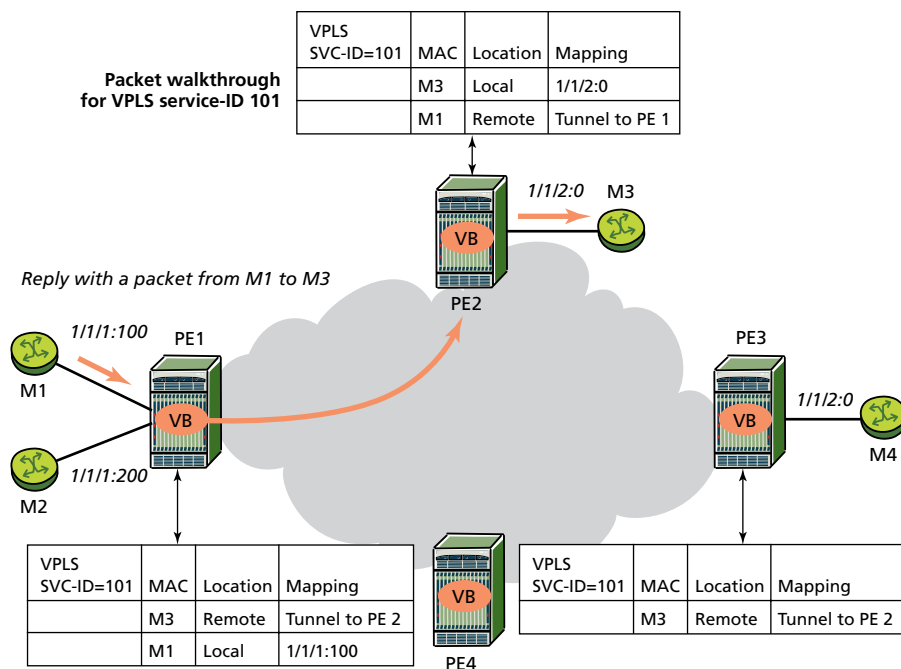
Figure 4. VPLS packet format



When M1 receives the packet from M3, it replies with a packet to M3 (see Figure 5):

- PE1 receives the packet from M1 and learns that M1 is on local port 1/1/1:100; it stores this information in the FIB for Svc-id 101.
- PE1 already knows that M3 can be reached through PE2 and therefore only sends the packet to PE2 using VC label pe1-2.
- PE2 receives the packet for M3; it knows that M3 is reachable on port 1/1/2:0.
- M3 receives the packet.

Figure 5. VPLS packet forwarding

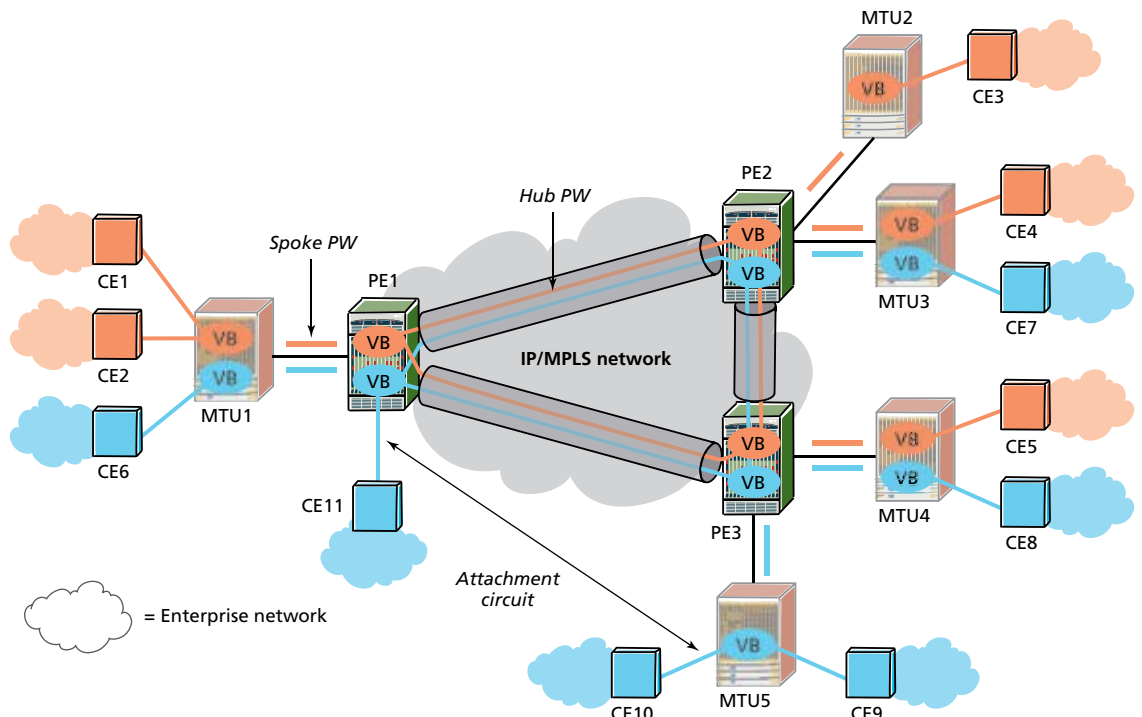


The H-VPLS architecture builds on the base VPLS solution and expands it to provide several scaling and operational advantages [4]. It is especially useful in large scale deployments with numerous PEs and/or multi-tenant units (MTUs).

Service providers deploy MTUs in multi-tenant buildings to serve the various enterprises in these buildings; each enterprise can potentially belong to a different VPLS VPN. Service providers then need to aggregate the MTU traffic towards the PE device in the central office or point of presence (PoP), as shown in CE IP/MPLS CE Figure 6. A traditional MTU is an Ethernet device that supports all layer 2 switching functions, including the normal bridging functions of learning and replication on all of its ports; it is typically dedicated to one enterprise. To share WAN resources more efficiently between customers, it is possible to extend the VPLS functionality to the MTUs. In this case, the MTUs act like PE devices, leading to a large number of MTUs participating in the VPLS. In a network with numerous PEs/MTUs, this would lead to scalability limitations in terms of the number of PWs to be maintained, packets to be PE1 replicated and MAC addresses to be 1/1/1:100 VB maintained.

The scaling advantages of H-VPLS are obtained by introducing hierarchy, thereby eliminating the need for a full mesh of LSPs and PWs between all participating devices. Hierarchy is achieved by augmenting the base VPLS core mesh of PE-to-PE PWs (referred to as hub PWs) with access PWs (called spoke PWs) to form a two-tier hierarchical VPLS model, as shown in Figure 6. Spoke PWs are created between the MTUs and the PE routers. H-VPLS offers the flexibility of utilizing different types of connection for the spoke PW implementation: either an IEEE 802.1Q tagged connection or an MPLS LSP with LDP signaling.

Figure 6. H-VPLS reference model



H-VPLS also offers several operational advantages by centralizing the major functions (for example, VPLS end-point auto-discovery, participating in a routed backbone, maintaining a full mesh of tunnel LSPs and multiple full meshes of PWs) in the PoP PE routers. This makes it possible to use lower-cost, low-maintenance MTU devices, thereby reducing the overall capital expenditure and operating expenses because typically there are an order of magnitude more MTU devices than PE routers. Another operational advantage offered by H-VPLS is centralized provisioning, with fewer elements to touch when turning up service for a customer. Adding a new MTU device requires some configuration of the local PE router but does not require any signaling to other PE routers or MTU devices, thereby greatly simplifying the provisioning process.

In H-VPLS, a CE is attached to an MTU through an attachment circuit. An AC from a specific customer is associated (by configuration) with a virtual bridge which is dedicated to that customer within the considered MTU (see Figure 6). An AC may be a physical or a virtual LAN (VLAN) tagged logical port. In the basic scenario, an MTU has one uplink to a PE. This uplink contains one spoke PW for each VPLS served by the MTU. The end-points of this spoke PW are an MTU and a PE. Spoke PWs can be implemented using LDP-signaled MPLS PWs if the MTU is MPLS enabled. Alternatively, they can be implemented using provider VLANs (P-VLAN) whereby every VLAN on the MTU-PE uplink of an Ethernet aggregation network identifies a spoke PW. In Figure 6, the uplink between MTU1 and PE1 carries two PWs because MTU1 has two VPLS customers attached. As the MTU has only one PW per VPLS, its operation is straightforward:

- Ethernet frames with known MAC addresses are switched accordingly within the VPLS.
- Frames with unknown or broadcast MAC addresses that are received from the PW are replicated and sent to all attached CE devices within the VPLS.
- Frames with unknown or broadcast MAC addresses that are received from a CE device are sent over the PW to the PE and to all other attached CE devices within the VPLS.
- Unknown MAC addresses are learned and aged1 within the VPLS both for frames coming from the PW and for frames coming from CE devices.

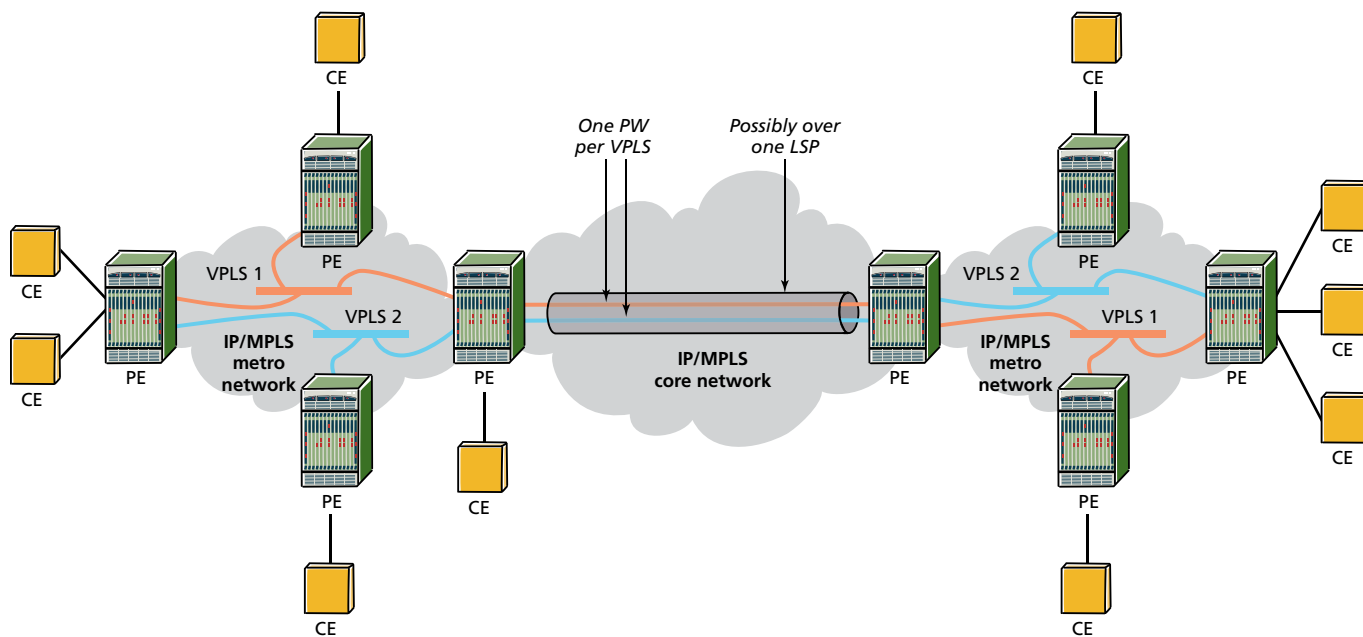
The PE device needs to implement one VB for each VPLS served by the PE-attached MTUs; the spoke PWs are seen as ACs from different customers. As such, a particular spoke PW is associated with the PE VB dedicated to the considered VPLS instance. In the core network, the PE has a full mesh of PWs to all other PEs that serve the VPLS (as in the normal VPLS scenario). These core PWs are called hub PWs. From a control plane level and data plane point of view, operation of the PE is the same as in the basic VPLS scenario.

Inter-metro service

H-VPLS enables VPLS to span multiple metro networks, as shown in Figure 7. A spoke connection is used to connect each VPLS between the two metros. In its simplest form, this could be a single tunnel LSP. A set of ingress and egress PW labels is exchanged between the border PE devices to create a PW for each VPLS instance to be transported over this LSP. The PE routers at each end treat this inter-metro PW as a virtual spoke connection for the VPLS in the same way as they treat PE-MTU connections. This architecture minimizes the signaling overhead and avoids a full mesh of VCs and LSPs between the two metro networks.

Once a MAC address has not been used for some time it is removed from the table; this is known as aging.

Figure 7. H-VPLS used as an inter-metro service



Conclusion

MPLS-based layer 2 services, such as VLL and VPLS, are being offered by service providers worldwide. Their success can be attributed to the fact that they use MPLS in the service provider's network combined with FR/ATM and Ethernet as handoff to the enterprise for VLL and Ethernet for VPLS.

MPLS-based layer 2 services offer enterprise customers exactly what they need for intersite connectivity: protocol transparency, scalable and granular bandwidth from 64 kbit/s to 1 Gbit/s, fast service activation and provisioning, and a simplified LAN/WAN boundary. VPLS also enables service providers to deliver a scalable VPN service offering that can be combined with Internet access on a consolidated IP/MPLS infrastructure, thereby reducing operating expenses. VPLS has already received widespread industry support from both vendors and service providers.

Alcatel-Lucent supports VPLS and H-VPLS in a broad range of products, including data and optical products, complemented by powerful network and services management.

References

- [1] IETF L2VPN working group: <http://www.ietf.org/html.charters/l2vpn-charter.html>.
- [3] J. Witters, G. van Kersen, J. De Clerq, S. Khandekar: "Keys to Successful VPLS Deployment", Alcatel Telecommunications Review, 4th Quarter 2004, pp 428-432.
- [4] M. Lasserre, V. Kompella: "Virtual Private LAN Services over MPLS", work in progress, <http://www.ietf.org/internet-drafts/draft-ietf-l2vpn-vpls-ldp03.txt>.
- [5] IETF PWE3 working group: <http://www.ietf.org/html.charters/pwe3-charter.html>.

Abbreviations

AC	attachment circuit
ATM	Asynchronous Transfer Mode
CE	customer edge
DA	destination address
DSL	Digital Subscriber Line
FCS	Frame Check Sequence
FIB	Forwarding Information Base
FR	Frame Relay
H-VPLS	Hierarchical Virtual Private LAN Service
IP	Internet Protocol
LAN	local area network
LDP	Label Distribution Protocol
LSP	Label Switched Path
MAC	Media Access Control
MPLS	Multi-Protocol Label Switching
MTU	multi-tenant unit
PE	provider edge
PoP	point of presence
PW	pseudo wire
PWE3	Pseudo-Wire Emulation Edge-to-Edge
RADIUS	Remote Authentication Dial-In User Service
RSVP-TE	Resource Reservation Protocol with Traffic Engineering Extensions
SA	source address
TLS	Transport Layer Security
VB	virtual bridge
VC	virtual connection
VLAN	virtual local area network
VLL	virtual leased line
VPLS	Virtual Private LAN Service
VPRN	Virtual Private Routed Network
VPN	virtual private network

www.alcatel-lucent.com Alcatel, Lucent, Alcatel-Lucent and the Alcatel-Lucent logo are trademarks of Alcatel-Lucent. All other trademarks are the property of their respective owners. The information presented is subject to change without notice. Alcatel-Lucent assumes no responsibility for inaccuracies contained herein. Copyright © 2009 Alcatel-Lucent. All rights reserved.
CPG2896090205 (03)

