

VPLS Technology White Paper

Copyright © 2021 New H3C Technologies Co., Ltd. All rights reserved.

No part of this manual may be reproduced or transmitted in any form or by any means without prior written consent of New H3C Technologies Co., Ltd.

Except for the trademarks of New H3C Technologies Co., Ltd., any trademarks that may be mentioned in this document are the property of their respective owners.

This document provides generic technical information, some of which might not be applicable to your products.

The information in this document is subject to change without notice.

Contents

Overview	1
Background	1
Benefits	1
VPLS implementation	2
Concepts	2
VPLS network architecture	3
Establishment of a PW	3
Static PW	4
LDP PW	4
BGP PW	4
BGP auto-discovery LDP PW	5
MAC address management	6
MAC address learning	6
MAC address aging	7
MAC address withdrawal	7
Traffic forwarding and flooding	7
Unicast traffic forwarding and flooding	7
Multicast and broadcast traffic forwarding and flooding	7
VPLS packet encapsulation	7
Packet encapsulation on an AC	7
Packet encapsulation on a PW	8
VLAN tag processing during packet forwarding	8
VLAN tag processing when Ethernet access mode and Ethernet data encapsulation type are used	9
VLAN tag processing when Ethernet access mode and VLAN data encapsulation type are used	9
VLAN tag processing when VLAN access mode and Ethernet data encapsulation type are used	10
VLAN tag processing when VLAN access mode and VLAN data encapsulation type are used	10
Loop avoidance	11
PW redundancy	11
PW states	11
PW redundancy operation modes	11
Primary/backup switchover	11
H-VPLS implementation	12
H-VPLS access modes	12
Link redundancy in H-VPLS	14
Loop avoidance in H-VPLS	15
Flow label	15
Flow label capabilities	15
Packet forwarding based on flow labels	15
Restrictions	16
QinQ configuration and packet encapsulation on PWs	16
H-VPLS QinQ access	16
Comware implementation characteristics	17
H-VPLS networking characteristics	17
MAC address reclaiming	17
BFD detection and redundancy	17
Flexible networking modes	18
Hub-spoke networking	18
Local E-tree networking	18
Application scenarios	19
References	19

Overview

Background

With the globalization of economy, more and more enterprises are spreading across an increasingly wider area, and employees travel more frequently. All these appeal for services that can enable enterprises to interconnect their branches, so that employees can easily access the corporate networks from any place.

Originally, service providers met the previously mentioned requirements by providing leased lines. But this method has significant disadvantages. It is not applicable when there are a large amount of branches or the number of branches grows quickly. Besides, this method is relatively expensive and a network based on leased lines is hard to manage.

After Asynchronous Transfer Mode (ATM) and Frame Relay (FR) emerged, service providers turned to them to provide virtual circuits. Enterprises can establish their own Layer 3 networks for IP or IPX traffic based on the virtual circuits. However, the virtual links are point-to-point Layer 2 links and a network based on them is complex to configure and maintain, especially when a new site joins.

Later, with IP networks present almost everywhere around the world, service providers began to focus on how to provide enterprises with low-cost private networks over existing IP networks. The technology that was developed to answer this demand is MPLS VPN, which is easy to configure and supports flexible bandwidth setting.

MPLS VPNs fall into two categories: MPLS L3VPN and MPLS L2VPN. MPLS L3VPN requires that the service providers participate in the internal routing management on user networks. The original MPLS L2VPN technology Virtual Leased Line (VLL) provides point-to-point L2VPN services on public networks. The virtual links established by VLL function just as they were physical links connecting the sites directly, but only point-to-point exchange is supported in this environment.

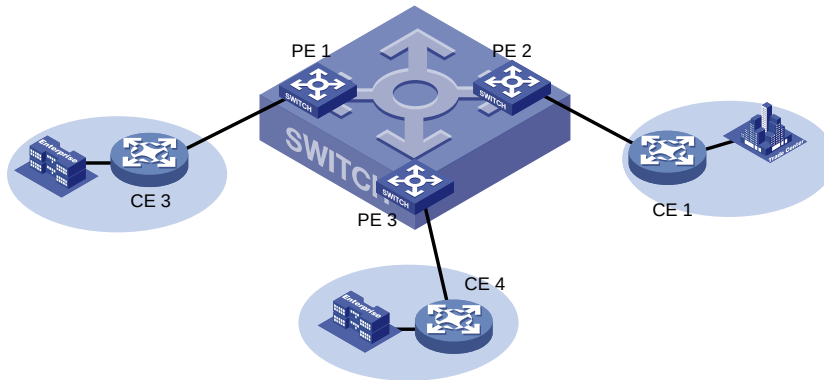
Virtual Private LAN Service (VPLS) is developed based on the traditional VLL solution. It supports multipoint-to-multipoint communication and proves to be a better solution for service providers.

Benefits

VPLS combines the advantages of Ethernet and MPLS and implements all the functions that traditional LANs provide. It can connect geographically dispersed Ethernets through the service provider IP/MPLS networks so that the Ethernets can work as a single LAN.

As shown in [Figure 1](#), a service provider simulates an Ethernet bridge on the MPLS backbone by using VPLS. The bridge forwards frames based on MAC address or MAC address and VLAN tag. In the simplest case, all sites connected to the PEs belong to a single VPLS instance, and each CE needs to communicate with the other CEs in the VPLS instance. For the CEs, the MPLS backbone functions just like an Ethernet bridge.

Figure 1 Ethernet bridge emulated by VPLS



VPLS provides these advantages:

- VPLS uses Ethernet interfaces at the user side, supporting quick and flexible service deployment at the border between the LAN and the WAN.
- With VPLS, users control and maintain routing policies on the user networks. This simplifies the management on the service provider network.
- All CEs of a VSI belong to a same subnet. This makes IP address planning much easier.
- The VPLS service is invisible to users. It is not involved in IP addressing and routing.

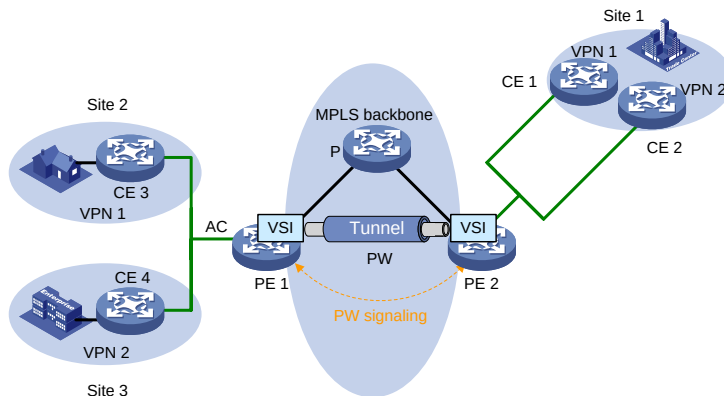
VPLS implementation

Concepts

- **CE**—Customer edge device that is directly connected with the service provider network.
- **PE**—Provider edge device that connects one or more CEs to the service provider network for VPN service access. A PE maps and forwards packets between private networks and public network tunnels. In H-VPLS networking scenarios, a PE can be a UPE or NPE.
- **UPE**—User facing provider edge device that functions as the user access convergence device.
- **NPE**—Network provider edge device that functions as the network core PE. An NPE resides at the edge of a VPLS network core domain and provides transparent VPLS transport services between core networks.
- **Service delimiter**—Identifier that a service provider adds before a user data packet to identify which VPN the packet belongs to. A service delimiter is local significant. A typical example of service delimiter is the outer tag in QinQ.
- **QinQ**—802.1Q in 802.1Q, a tunneling protocol based on 802.1Q. It offers a point-to-multipoint L2VPN service mechanism. QinQ allows packets to carry both the private network VLAN tags and the public network VLAN tags to travel across the service provider network. It provides a simpler Layer 2 VPN tunneling service.

VPLS network architecture

Figure 2 VPLS network architecture



As shown in Figure 2, a VPLS network consists of these primary components:

- **AC**—Attachment circuit that connects a user with the service provider, that is, a link between a CE and a PE. The ends of an AC can only be Ethernet interfaces.
- **PW**—Pseudowire, bidirectional virtual connection between VSIs on two PEs. A PW consists of two unidirectional LSPs.
- **Tunnel**—Direct channel between a local PE and the peer PE for transparent data transmission in between. A tunnel can be an MPLS tunnel or a GRE tunnel. It carries one or more PWs over an IP/MPLS backbone.
- **PW signaling**—Protocol that is the fundament of VPLS. It is used for creating and maintaining PWs and automatically discovering VSI peer PEs. Currently, there are two PW signaling protocols: LDP and BGP.
- **VPLS instance**—A customer network might contain multiple geographically dispersed sites. The service provider uses VPLS to connect all the sites to create a single Layer 2 VPN, which is referred to as a VPLS instance. Sites in different VPLS instances cannot communicate with each other at Layer 2.
- **VSI**—Virtual switch instance, an Ethernet bridge function entity of a VPLS instance on a PE. It forwards Layer 2 frames based on MAC address and VLAN tag.

Establishment of a PW

A PW is a communication tunnel on the public network. It can be established on an MPLS tunnel (a common LSP or a CR-LSP) or a GRE tunnel. For a PW to be established, you need to do the following configurations:

1. Establish an MPLS or GRE tunnel between the local end and the peer PE.
2. Determine the address of the peer PE. If the peer PE is in the same VSI as the local end, you can specify the address of the peer PE manually, or let the signaling protocol find the peer PE automatically.
3. Use the LDP or BGP signaling protocol to assign multiplex distinguishing flags (that is, PW labels) and advertise the assigned PW flags to the peer PE, establish unidirectional LSPs and further establish a PW. If a PW is established on an MPLS tunnel, a packet transported over the PW will contain two levels of labels. The inner label, called a PW label, identifies the PW to which the packet belongs so that the packet is forwarded to the correct CE; while the outer label, called the public network MPLS tunnel label, is for guaranteeing the correct transmission of the packet on the MPLS tunnel.

The following describes how to use each of the signaling protocols to establish a PW respectively.

Static PW

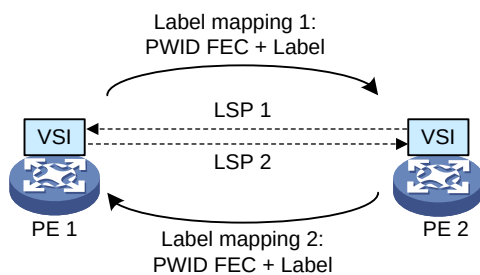
To establish a static PW, configure the peer PE address, and the incoming and outgoing PW labels for the PW on the two PEs.

Static PWs are easy to configure but cannot adapt to network changes. Therefore, static PWs are applicable only to small stable networks.

LDP PW

To create an LDP PW, specify the address of the remote PE, and use LDP to advertise the PW-label binding to the remote PE. After the two PEs receive the PW-label binding from each other, they establish an LDP PW.

Figure 3 LDP PW establishment



As shown in [Figure 3](#), a PW is established by using LDP in the following steps:

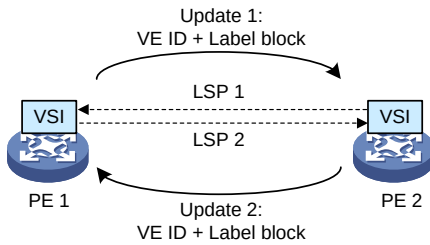
1. After being associated with a VSI, each PE uses LDP in downstream unsolicited (DU) mode to send a label mapping message to its peer PE unsolicitedly. The message contains the PWid FEC (FEC 128), the PW label bound with the FEC, and the interface settings such as the maximum transmission unit.
2. Upon receiving such a message, a PE determines whether it is associated with the PWID. If so, it accepts the label mapping message.
3. After a unidirectional LSP is established in each direction, the PW is formed. A PW can be considered as a virtual Ethernet interface of a VSI.

LDP VPLS is easy to implement. However, as LDP does not provide automatic VPLS member discovery mechanism, the PE peers need to be manually configured and every PE needs to be reconfigured whenever a new PE joins.

BGP PW

To create a BGP PW, configure BGP to advertise label block information to the remote PE. After the two PEs receive label block information from each other, they use the label block information to calculate the incoming and outgoing labels and create the BGP PW.

Figure 4 BGP PW establishment



As shown in [Figure 4](#), a PW is established by using BGP in the following steps:

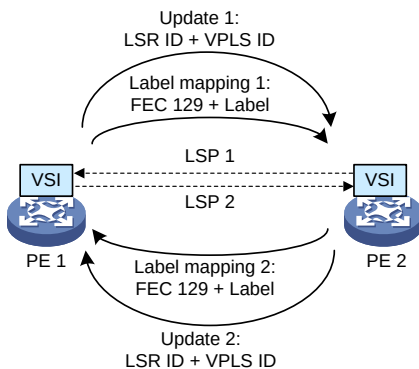
1. A PE advertises its VE ID and label block information to all peer PEs by a BGP update message. A VE ID is the unique identifier of a site connected with the PE in the VPN and is assigned by the service provider. A label block consists of a group of consecutive labels.
2. Upon receiving an update message, a PE figures out a unique label value based on its own VE ID information and the label block in the update message and uses the label value as the PW label. At the same time, the PE gets the PW label of the peer PE based on the VE ID in the message and its local label block.
3. After two PE peers receive update messages from each other and figure out the PW labels, a PW is established between the two PEs.

BGP VPLS implements automatic VPLS member discovery by VPN target configurations and requires no manual configuration when a PE joins or exits, thus featuring high scalability. However, the BGP protocol is complex in itself.

BGP auto-discovery LDP PW

To create a BGP auto-discovery LDP PW, configure BGP to automatically find the remote PE, and use LDP to advertise the PW-label binding to the remote PE. After the two PEs receive the PW-label binding from each other, they establish a BGP auto-discovery LDP PW.

Figure 5 BGP auto-discovery LDP PW establishment



As shown in [Figure 5](#), a BGP auto-discovery LDP PW is established in the following steps:

1. A PE advertises its LSR ID and VPLS ID to all peer PEs by using BGP update messages. After being associated with a VSI, each PE uses LDP in DU mode to send a label mapping message to its peer PE unsolicitedly.
2. The information advertised by BGP includes the ID (for example, LSR ID) and VPLS ID of the advertising PE.
3. The receiving PE compares the received VPLS ID with its own VPLS ID. If the two VPLS IDs are identical, the two PEs use LDP to establish a PW. If not, the PEs do not establish a PW.

4. After discovering a peer PE address available for PW establishment through BGP, the PE uses LDP in DU mode to send a label mapping message to the peer PE unsolicitedly. The FEC type in the LDP message is Generalized PWid FEC Element (FEC 129), which contains the VPLS ID, Source Attachment Individual Identifier (SAII), and Target Attachment Individual Identifier (TAII). The SAII is the LSR ID of the advertising PE. The TAI identifies the remote PE and is advertised by the remote PE. VPLS ID+SAII+TAII uniquely identifies a PW in a VPLS instance.
5. After the two PEs receive the PW-label binding from each other, they establish a BGP auto-discovery LDP PW.

BGP auto-discovery LDP PWs combine the advantages of LDP PWs and BGP PWs, realizing automatic discovery of VPLS members through BGP and label exchange between VPLS members through LDP.

MAC address management

For user networks, a VPLS network emulates an Ethernet bridge, which forwards packets based on MAC address or both MAC address and VLAN tag. Each PE associated with a particular VPLS service establishes a VSI for the VPLS instance, and each VSI maintains a MAC address table and supports packet flooding and forwarding, and MAC address learning and aging.

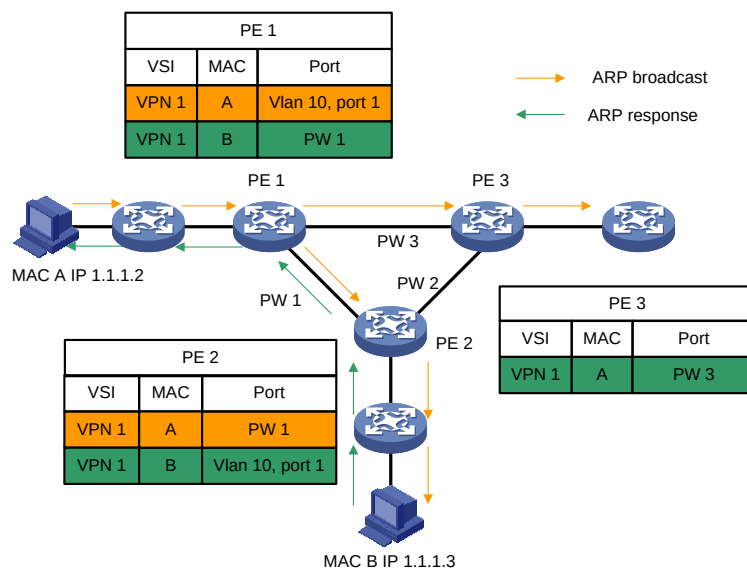
MAC address learning

VPLS provides reachability through source MAC learning. A PE maintains a MAC address table for each VSI.

As shown in Figure 6, a PE learns source MAC addresses in the following ways:

- Learning the source MAC addresses of directly connected sites.
If the source MAC address of a packet from a CE does not exist in the MAC address table, the PE learns the source MAC address of the AC connected to the CE.
- Learning the source MAC addresses of remote sites connected through PWs.
A VSI regards a PW as a logical Ethernet interface. If the source MAC address of a packet from a PW does not exist in the MAC address table, the PE learns the source MAC address on the PW of the VSI.

Figure 6 MAC address learning



MAC address aging

Learned MAC address entries that are no more in use need to be aged out by an aging mechanism. The aging mechanism functions based on the source MAC addresses of received packets. Whenever a PE receives a packet, it sets an activated or effective flag for the corresponding MAC address entry. If a MAC address entry does not have an activated or effective flag set in a specified period of time, it will be removed from the MAC address table.

MAC address withdrawal

When an AC or a PW goes down, the PE deletes MAC addresses on the AC or PW. Then it sends an LDP address withdrawal message to notify all other PEs in the VPLS instance to delete those MAC addresses.

Traffic forwarding and flooding

Unicast traffic forwarding and flooding

After a PE receives a unicast packet from an AC, the PE searches the MAC address table of the VSI bound to the AC for packet forwarding.

- If a match is found, the PE forwards the packet according to the matching entry.
 - If the outgoing interface in the entry is a PW, the PE inserts the PW label to the packet, and adds the public tunnel header to the packet. It then forwards the packet to the remote PE over the PW. If the PW is carried on an LSP or MPLS TE tunnel, each packet on the PW contains two labels. The inner label is the PW label, which identifies the PW and ensures that the packet is forwarded to the correct VSI. The outer label is the public LSP or MPLS TE tunnel label, which ensures that the packet is correctly forwarded to the remote PE.
 - If the outgoing interface in the entry is a local interface, the PE directly forwards the packet to the local interface.
- If no match is found, the PE floods the packet to all other ACs and PWs in the VSI.

After a PE receives a unicast packet from a PW, the PE searches the MAC address table of the VSI bound to the PW for packet forwarding.

- If a match is found, the PE forwards the packet through the outgoing interface in the matching entry.
- If no match is found, the PE floods the packet to all ACs in the VSI.

Multicast and broadcast traffic forwarding and flooding

After a PE receives a multicast or broadcast packet from an AC, the PE floods the packet to all other ACs and the PWs in the VSI bound to the AC.

After a PE receives a multicast or broadcast packet from a PW, the PE floods the packet to all ACs in the VSI bound to the PW.

VPLS packet encapsulation

Packet encapsulation on an AC

Two packet encapsulation types are available on an AC: VLAN access and Ethernet access.

- **VLAN access:** The Ethernet header of a packet sent from a CE to a PE or from a PE to a CE includes a VLAN tag, which is added in the header as a service delimiter for the service provider network to identify the user. The tag is called P-Tag.
- **Ethernet access:** The Ethernet header of a packet sent from a CE to a PE or from a PE to a CE does not contain any service delimiter. If the header contains a VLAN tag, it is the internal VLAN tag of the user and means nothing to the PE. Such an internal VLAN tag of a user is called U-Tag.

Packet encapsulation on a PW

A PW is uniquely identified by its PWID and PW encapsulation type. Two PE peers must advertise the same PWID and PW encapsulation type.

Two packet encapsulation types are available on a PW: Ethernet and VLAN.

- **Ethernet**—P-tag is not transferred on a PW.
 - For a packet from a CE:
 - If the packet contains a P-tag, the PE removes the P-tag, and adds a PW label and an outer tag into the packet before forwarding it.
 - If the packet contains no P-tag, the PE directly adds a PW label and an outer tag into the packet before forwarding it.
 - For a packet to a CE:
 - If the access mode of the AC is VLAN, the PE adds a P-tag into the packet before sending it to the CE.
 - If the access mode of the AC is Ethernet, the PE directly sends the packet to the CE.

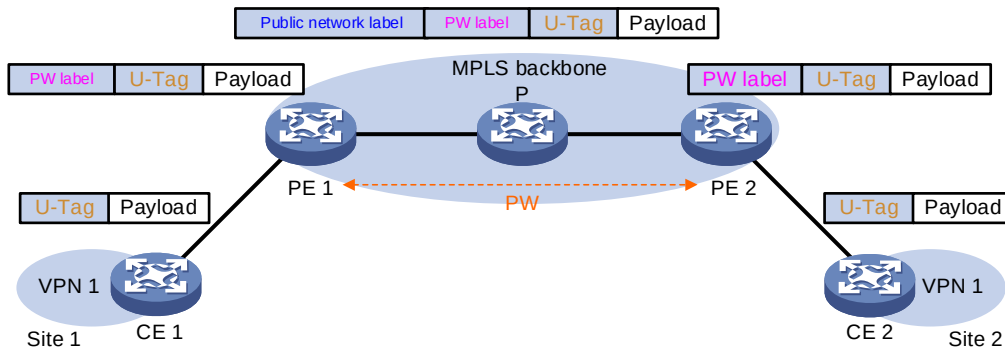
You cannot rewrite or remove existing tags.
- **VLAN**—Packets transmitted over a PW must carry a P-tag.
 - For a packet from a CE:
 - If the peer PE does not require the ingress to rewrite the P-tag, the PE keeps the P-tag unchanged for the packet, and then encapsulates the packet. If the packet contains no P-tag, the PE adds a null label (the label value is 0) into the packet, and then encapsulates the packet.
 - If the peer PE requires the ingress to rewrite the P-tag, the PE changes the P-tag to the expected VLAN tag (the tag value might be 0), and then adds a PW label and an outer tag into the packet. If the packet contains no P-tag, the PE adds a VLAN tag expected by the peer PE (the tag value might be 0), and then adds a PW label and an outer tag into the packet.
 - For a packet to a CE:
 - If the access mode of the AC is VLAN, the PE rewrites or retains the P-tag before forwarding the packet.
 - If the access mode of the AC is Ethernet, the PE removes the P-tag before forwarding the packet.

VLAN tag processing during packet forwarding

The VLAN tag processing mechanism varies by the access mode and data encapsulation type.

VLAN tag processing when Ethernet access mode and Ethernet data encapsulation type are used

Figure 7 VLAN tag processing when Ethernet access mode and Ethernet data encapsulation type are used

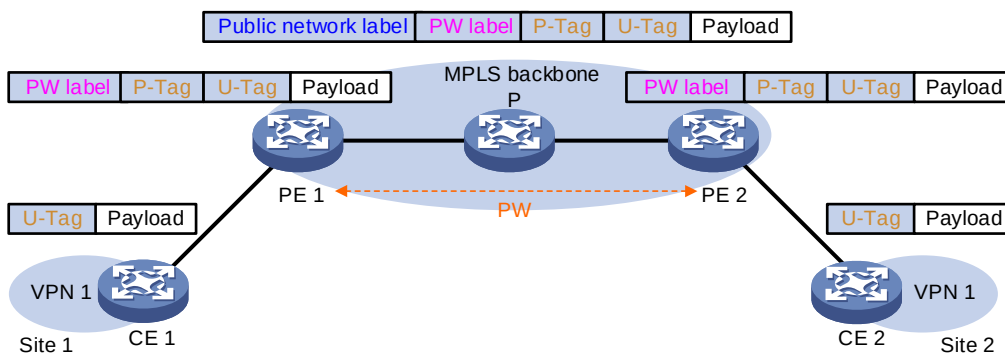


As shown in Figure 7, when Ethernet access is used on ACs and Ethernet mode is used on PWs, a packet is forwarded as follows:

1. CE 1 sends a packet carrying information about the VLAN to which the user belongs, that is, the U-Tag, to PE 1.
2. Upon receiving the packet, PE 1 selects a proper PW based on the destination MAC address or both the native VLAN of the user and the destination MAC address, and then adds the PW label of the PW into the packet.
3. To forward the packet across the public network through an MPLS tunnel, PE 1 adds the public network tunnel label into the packet.
4. When PE 2 receives the packet, it gets to know the VSI to which the packet belongs by the PW label and forwards the payload together with the U-Tag to CE 2.

VLAN tag processing when Ethernet access mode and VLAN data encapsulation type are used

Figure 8 VLAN tag processing when Ethernet access mode and VLAN data encapsulation type are used

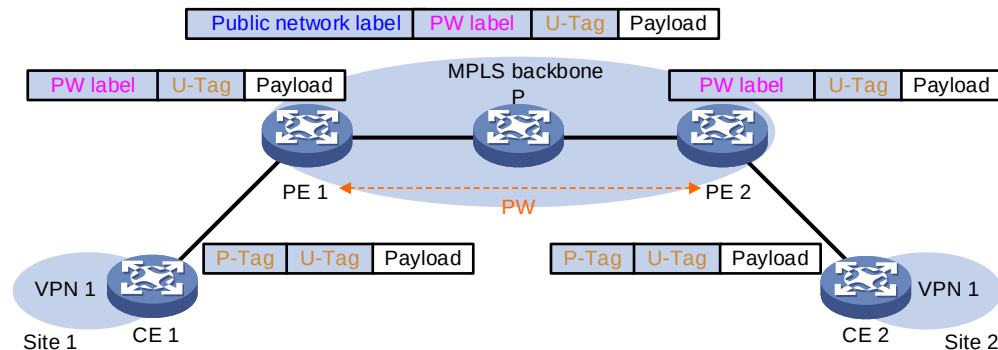


As shown in Figure 8, when Ethernet access is used on ACs and VLAN mode is used on PWs, the packet forwarding process is similar to that when Ethernet access is used on ACs and Ethernet mode is used on PWs. The only difference is that packets transferred on the PWs must carry P-Tags.

When PE 1 receives from CE 1 a packet without a P-Tag, it adds the VLAN tag that the peer PE expects or a null tag into the packet, and then pushes two levels of MPLS labels and forwards the packet out. When PE 2 receives the packet, it removes the two levels of MPLS labels and the P-Tag before forwarding the packet to CE 2.

VLAN tag processing when VLAN access mode and Ethernet data encapsulation type are used

Figure 9 VLAN tag processing when VLAN access mode and Ethernet data encapsulation type are used

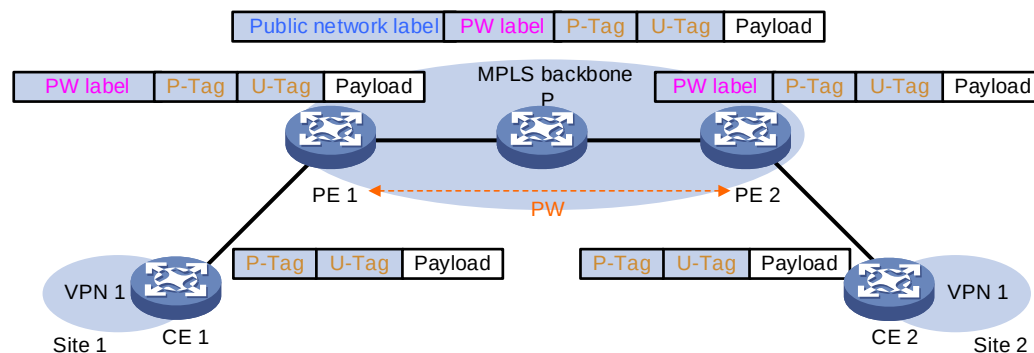


As shown in Figure 9, when VLAN access is used on ACs and Ethernet mode is used on PWs, a packet is forwarded as follows:

1. CE 1 sends a packet carrying the service delimiter P-Tag to PE 1.
2. Upon receiving the packet, PE 1 removes the P-Tag, adds two levels of MPLS labels into the packet, and then forwards the packet through a public network MPLS tunnel to PE 2.
3. When PE 2 receives the packet, it removes the two levels of MPLS labels, adds a P-Tag, and then forwards the packet to CE 2.

VLAN tag processing when VLAN access mode and VLAN data encapsulation type are used

Figure 10 VLAN tag processing when VLAN access mode and VLAN data encapsulation type are used



As shown in Figure 10, when VLAN access is used on ACs and VLAN mode is used on PWs, the packet forwarding process is similar to that when VLAN access is used on ACs and Ethernet mode is used on PWs. The only difference is that packets transferred on the PWs must carry P-Tags and that PEs receiving the packets may change or retain the P-Tags in the packets.

Loop avoidance

In general, Layer 2 networks use the STP to avoid loops. This is not applicable for VPLS networks because the users cannot sense the service provider network. Therefore, enabling STP in the private networks means nothing to the service provider network. In VPLS, PW full mesh and split horizon forwarding are used to avoid loops:

- PEs are logically fully meshed, that is, each PE must create for each VPLS forwarding instance a tree to all the other PEs of the instance.
- Each PE must support horizontal split forwarding to avoid loops. When receiving a packet from a PW, a PE does not forward the packet to the other PWs of the VSI. That is, any two PEs communicate through the PW directly connecting them, rather than a third PE. This is why PW full mesh is required for each VSI instance.

PW redundancy

PW redundancy provides redundant links between PEs so that the customer networks can communicate when the path over one PW fails. In the current software version, only static PWs and LDP PWs support PW redundancy.

PW states

A PW can be in either of the following states:

- **Active**—The PW is in active state and can forward packets.
- **Standby**—The PW is in standby state and cannot forward packets.

PW redundancy operation modes

LDP PWs support the independent and master/slave PW redundancy operation modes.

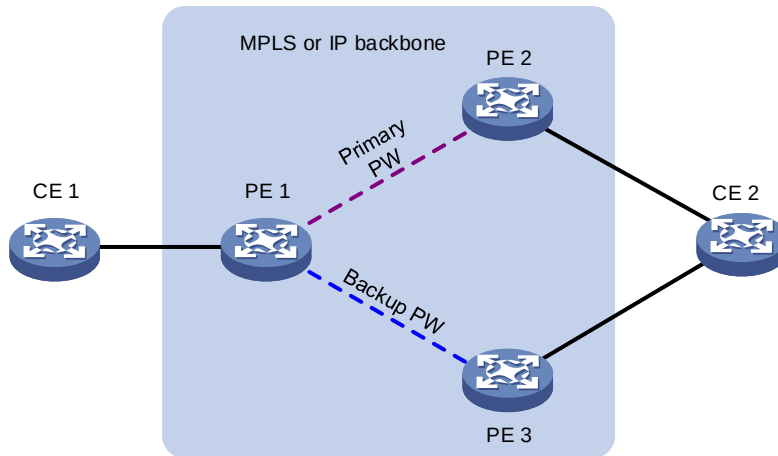
- **Independent mode**—The two PEs of a PW use LDP to advertise their respective PW active/standby state to each other. A PW can forward traffic only when it is up and active at both ends of the PW. In this mode, make sure both PEs of a PW use the independent PW redundancy operation mode.
- **Master/slave mode**—One PE of a PW operates as the master node and the other PE operates as the slave node. The master PE determines the PW active/standby state and then uses LDP to advertise the PW state to the slave PE. The slave PE uses the same PW state as the master PE based on the information received from the master PE. In this way, the master and slave PEs for the set of redundant PWs can use the same active PW to forward user packets.

A slave node does not advertise the PW active/standby state to the master node, and the master node ignores the PW active/standby state received from the slave nodes.

Primary/backup switchover

As shown in [Figure 11](#), PE 1 establishes two PWs (one primary and one backup). The CEs communicate through the primary PW. When the primary PW fails, PE 1 brings up the backup PW and forwards packets from CE 1 to CE 2 through the backup PW. When CE 2 receives the packets, it updates its MAC address table, so that packets from CE 2 to CE 1 also travel through the backup PW.

Figure 11 VPLS PW redundancy



VPLS determines whether the primary PW fails according to the LDP session status or the BFD result. The backup PW is used when one of the following conditions exists:

- The public tunnel of the primary PW is deleted, or BFD detects that the public tunnel has failed.
- The primary PW is deleted because the LDP session between PEs goes down, or BFD detects that the primary PW has failed.
- A manual PW switchover is performed.

For static PWs, the PE switches the traffic to the backup PW after the primary PW fails.

For LDP PWs, the local PE determines the PW state and then advertises the state information to the remote PE through the PW Preferential Forwarding bit in LDP messages.

- When the PW Preferential Forwarding bit is set, the PW is in standby state.
- When the PW Preferential Forwarding bit is not set, the PW is in active state.

The two PEs negotiate the PW for traffic forwarding based on the PW state and PW redundancy mode.

H-VPLS implementation

As described previously, VPLS requires that the PEs are fully meshed. Therefore, for a VPLS instance, this equation exists: the number of PWs = the number of PEs $\times$ (the number of PEs – 1) $\div$ 2. When the VPLS network is large, there will be a great amount of PWs. As a result, the PW signaling cost will be very high and the network will be difficult to manage and scale. H-VPLS is introduced to simplify network management and improve network scalability.

In H-VPLS, a PE can be a UPE or NPE. A UPE functions as a multi-tenant unit (MTU) and connects CEs to the service provider network. An NPE resides at the edge of a VPLS network core domain and provides transparent VPLS transport services between core networks. Only NPEs are required to be fully meshed and a UPE does not need to be connected with all NPEs. Using a hierarchical structure, H-VPLS reduces the number of PWs and eases the PW signaling burden.

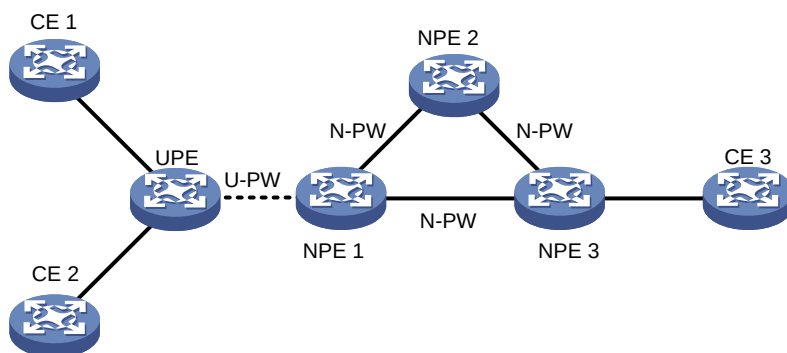
H-VPLS access modes

According to the connection modes between UPEs and NPEs, there are two H-VPLS access modes:

- LSP access
- QinQ access

LSP access

Figure 12 H-VPLS LSP access



As shown in Figure 12, UPE functions as the MTU and only establishes a U-PW with NPE 1; it does not establish any virtual link with any other peer. Note that for a U-PW to be established, you need to create a VSI instance and specify the peer on the NPE and UPE respectively and ensure that the two devices have the same PWID.

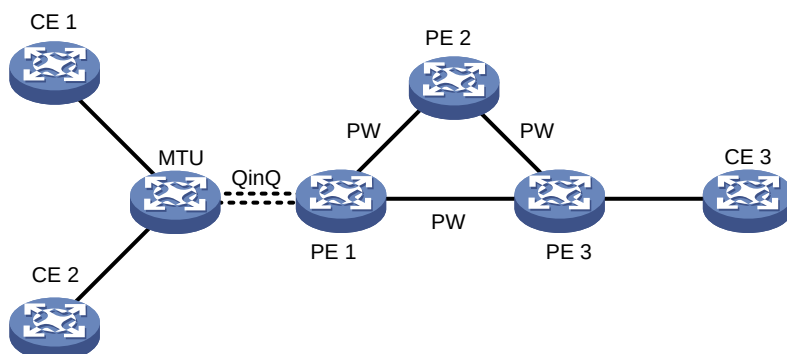
In the above scenario, a packet is forwarded as follows:

1. For each packet from CE 1 or CE 2, UPE adds the PW label of U-PW (that is, the PW label that NPE 1 assigns for PW multiplexing/demultiplexing) into the packet, and then forwards the packet to NPE 1.
2. Upon receiving the packet, NPE 1 figures out the VSI to which the packet belongs based on the PW label, adds the PW label of N-PW according to the destination MAC address of the packet, and then forwards the packet out.
3. When NPE 1 receives a packet from N-PW, it adds the PW label of U-PW into the packet and forwards the packet to UPE, which forwards the packet to the destination CE in turn.

For packets to be exchanged between CE 1 and CE 2, UPE can forward them directly without NPE 1 if it supports bridging. However, the situation is different for packets with unknown destination MAC addresses and broadcast packets. Upon receiving such a packet from a local CE, UPE not only broadcasts the packet to the other local CEs through the bridging function, but also forwards it through U-PW to NPE 1, which replicates the packet and sends a copy to each peer CE.

QinQ access

Figure 13 H-VPLS QinQ access



As shown in Figure 13, MTU is a standard bridging device and an Ethernet QinQ connection is established on between MTU and PE 1. Note that for the connection to be established, you need to enable QinQ for the interfaces connecting CEs on MTU and configure VLAN access mode on PE 1.

In the above scenario, when MTU receives a packet from CE 1 or CE 2, it adds the outer VLAN flag into the packet and forwards the packet to PE 1. Based on its VLAN access mode, PE 1 interprets the outer VLAN flag as the service provider VLAN flag, that is, the service delimiter assigned by the service provider for the user. By this delimiter, PE 1 maps the packet to the corresponding VSI instance and forwards (unicasts or broadcasts) the packet according to the VSI instance.

The following details the forwarding process:

1. With QinQ enabled on the interfaces connecting CEs, MTU adds a VLAN tag into each packet from the CEs and forwards the packet through the QinQ tunnel transparently to PE 1.
2. Upon receiving a packet from MTU, PE 1 determines the VSI to which the packet belongs by the VLAN tag in the packet, adds the PW label of the PW for the destination MAC address of the packet into the packet, and then forwards the packet out.
3. When PE 1 receives a packet from a PW, it determines the VSI to which the packet belongs by the PW label and labels the packet with the VLAN tag for the destination MAC address of the packet. Then, it forwards the packet through the QinQ tunnel to the MTU, which in turn forwards the packet to the destination CE.

For packets to be exchanged between CE 1 and CE 2, MTU can forward them directly without PE 1 because it supports bridging. However, the situation is different for packets with unknown destination MAC addresses and broadcast packets. Upon receiving such a packet from a local CE, MTU not only broadcasts it to the other local CEs through the bridging function, but also forwards it through the QinQ tunnel to PE 1, which replicates the packet and sends a copy to each peer CE.

Link redundancy in H-VPLS

If there is only a single link between a UPE and a NPE or between a MTU and a PE, the connectivity of all VPNs supported by the convergence device will be broken in case the link fails. Therefore, link redundancy is required for the two H-VPLS access modes. Normally, the convergence device uses only one link, the primary link, for access. When the primary link fails, it uses the backup link instead.

In H-VPLS LSP access mode, as LDP runs between a UPE and a NPE, the status of the primary PW can be determined by checking the status of the LDP session. In H-VPLS QinQ access mode, you need to configure STP between a MTU and its connected PE, so that a backup link is used instead when the primary link fails.

Figure 14 Link redundancy in H-VPLS LSP access mode

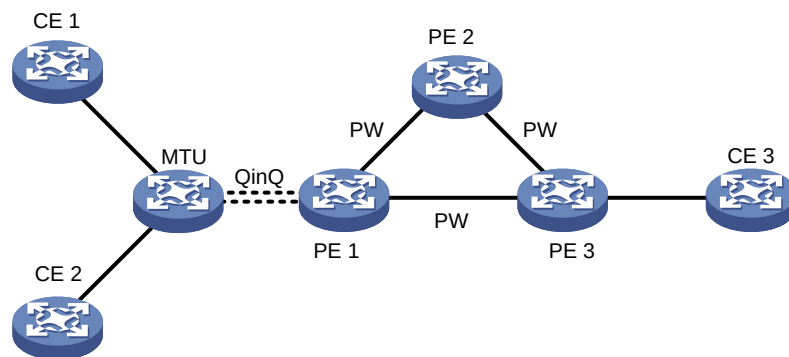
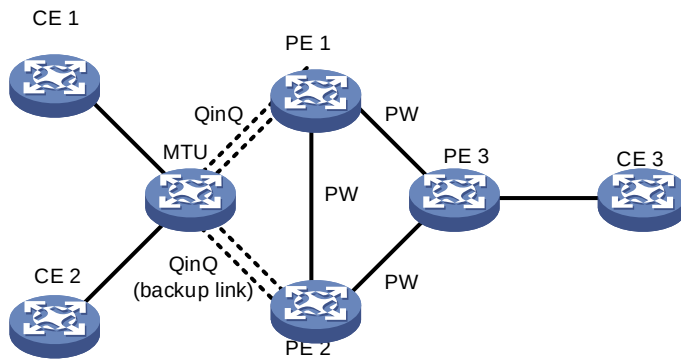


Figure 15 Link redundancy in H-VPLS QinQ access mode



Loop avoidance in H-VPLS

Loop avoidance in H-VPLS is different from that in common VPLS in the following aspects:

- Only NPEs need to be fully meshed and a UPE does not need to be connected with all NPEs.
- Upon receiving a packet from a PW connected with another NPE, a NPE does not forward the packet to the other PWs of the VSI that are connected with NPEs, but it forwards the packet to the PWs connected with UPEs.
- When a NPE receives a packet from a PW connected with a UPE, it forwards the packet to all PWs of the VSI that are connected with other NPEs.

Flow label

Packets carrying different types of traffic might be transmitted through the same PW and encapsulated with the same PW label. The P devices forward the traffic flows of a PW over the same path even if Equal Cost Multiple Paths (ECMPs) exist.

The L2VPN flow label feature can enable a P device to perform load sharing on packets based on the flow types.

Flow label capabilities

You can enable the flow label sending, receiving, or both sending and receiving capabilities on a PE.

- The sending capability enables a PE to send packets with flow labels. The PE adds a flow label before it adds a PW label to a packet during PW encapsulation.
- The receiving capability enables a PE to identify the flow label in a received packet and removes the flow label before forwarding the packet.

For two PE to successfully negotiate the flow label capabilities, make sure one end has the sending capability and the other end has the receiving capability.

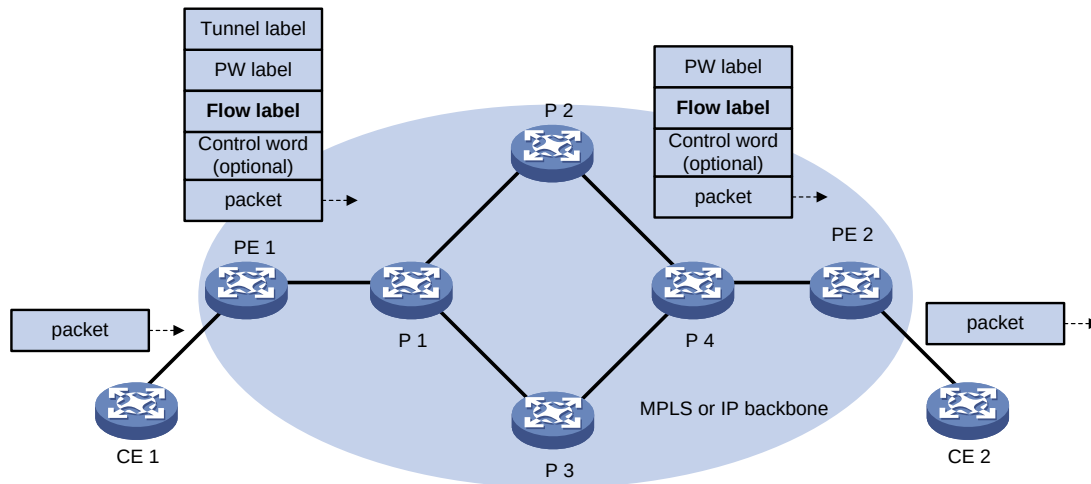
For static PWs, you must manually configure flow label capabilities for the local and remote PEs. For dynamic PWs, you can configure the local and remote PEs to negotiate the flow label capabilities or manually configure flow label capabilities for the PEs.

Packet forwarding based on flow labels

As shown in [Figure 16](#), after you configure this feature, the P and PE devices process packets as follows:

- When the ingress PE encapsulates a packet, it adds a flow label before it adds a PW label. The ingress PE adds different flow labels for packets of different traffic types.
- The P devices perform load sharing on packets based on the flow labels.
- The egress PE removes both the PW and flow labels from a packet before forwarding the packet.

Figure 16 Packet forwarding based on flow labels



Restrictions

QinQ configuration and packet encapsulation on PWs

Some products do not resolve the outer tags of received packets according to "[Packet encapsulation on an AC](#)". Instead, they process tags based on whether QinQ is enabled on the private network interfaces and the packet encapsulation types of PWs. When using these products, note that:

- When a VPLS service board processes VPLS traffic, it always treats outer tags as P-Tags (service delimiters). If QinQ is not enabled on the private network interfaces and PWs are working in Ethernet mode, the board directly removes the outer tag, even if the tag is a U-Tag. Therefore, if you do not want the U-tags to be removed, enable QinQ on the private network interfaces, so that the board removes only the tags that QinQ adds.
- The value of the Requested VLAN ID field that is advertised to the peer depends on the number of ACs supported by VPLS, that is, the number of interfaces that can be bound with a VSI. If a VSI is bound with a single AC, the value of the Requested VLAN ID field is the VLAN ID of the local private network interface; otherwise, it is 0. Therefore, if a VSI can be bound with more than one AC and the encapsulation type of the PW is tagged, the P-Tag of a packet on the PW is a null tag.

H-VPLS QinQ access

When configuring H-VPLS QinQ access mode, note that:

- On an MTU, you need to enable STP or MSTP on the Ethernet interfaces connected with NPEs and CEs, so that BPDU messages are exchanged between NPEs to avoid loops.
- On an NPE, ensure that MSTP is not enabled on the Ethernet interfaces connected with an MTU. Otherwise, BPDU messages will not be able to be transferred normally.

- To prevent BPDU messages from being transferred to other PEs in the VPLS network domain, you need to map BPDU messages to a VPLS instance that is different from that for user data packets.

Comware implementation characteristics

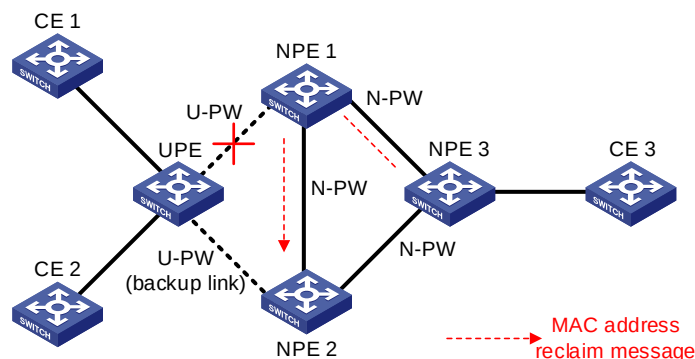
H-VPLS networking characteristics

MAC address reclaiming

A UPE belongs to two NPEs. When the PW in use (for example, the primary PW) fails, the UPE initiates a PW switchover to use the other PW. However, in a short period of time after the original PW fails, the NPEs of the other sites still send packets to the original PW, which the NPE cannot forward. To improve the convergence speed, a mechanism is introduced to inform other NPEs of the PW switchover as quickly, so that they remove their MAC entries for the VSI, initiate MAC address relearning, and reestablish MAC forwarding paths. This mechanism is implemented by LDP address reclaim messages.

In the implementation of Comware, the UPE sends MAC address reclaim messages. As shown in [Figure 17](#), the UPE sends a MAC address reclaim message to the NPE connected by the newly activated PW, which in turn forwards the message to other NPES.

Figure 17 MAC address reclaiming in Comware



An address reclaim message contains a MAC TLV. When a device receives an address reclaim message, it removes or relearns MAC addresses according to the parameters specified in the TLV. When the quantity of MAC addresses is large, a null MAC address list can be sent to improve convergence performance. When a NPE receives an address reclaim message containing a null MAC address list, it removes all MAC addresses of the specified VSI except the one learned from the PE sending the message. Comware supports only reclaiming MAC addresses by sending a null MAC address list.

BFD detection and redundancy

In an H-VPLS network, two PWs are established between a UPE and two NPEs for redundancy: one is the primary and the other is the secondary. When the primary PW fails, traffic is switched to the secondary PW for continual communication.

BFD is a detection mechanism used throughout the network for quick detection and monitoring of link connectivity. It sends detection packets at an interval that can be as short as 10 ms to detect the route reachability between a UPE and a NPE. When BFD detects on a UPE that a connection to a NPE fails, the UPE initiates a PW switchover and uses the backup PW to forward traffic, thus ensuring the continuity of communication.

Flexible networking modes

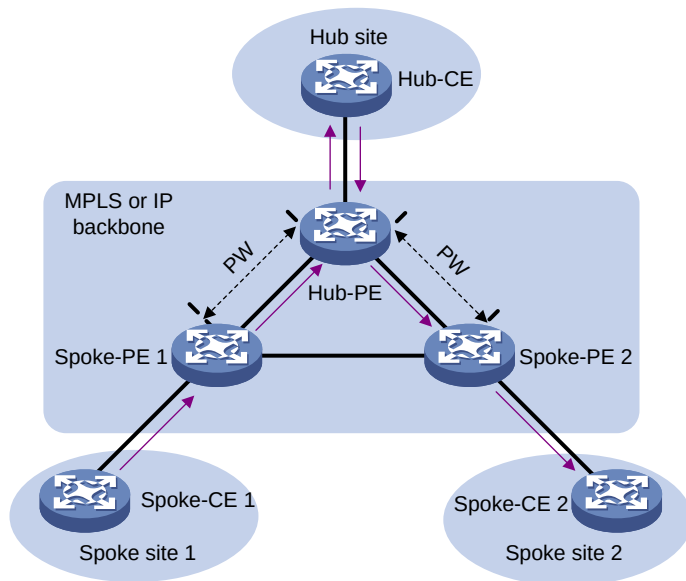
Except for traditional VPLS networks, Comware also supports hub-spoke networks and local E-tree networks.

Hub-spoke networking

The hub-spoke network model has one hub site and multiple spoke sites. The spoke sites cannot directly communicate with each other. Traffic between spoke sites must travel through the hub site, so the hub site can implement centralized traffic management.

Only static PWs and LDP PWs support the hub-spoke network model.

Figure 18 Hub-spoke model of VPLS



As shown in [Figure 18](#):

- The CE in the hub site is called a hub-CE.
- The PE connected to the hub site is called a hub-PE.
- A CE in a spoke site is called a spoke-CE.
- A PE connected to a spoke site is called a spoke-PE.
- A link (AC or PW) towards the hub site is called a hub link. You must specify one and only one hub link in a VSI.
- A link (AC or PW) towards a spoke site is called a spoke link.

In the hub-spoke VPLS network, a PE performs MAC address learning and forwarding as follows:

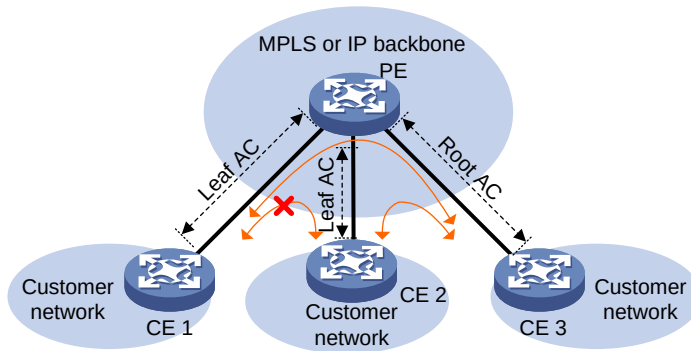
1. After receiving a packet from a spoke link, the PE performs MAC address learning and forwards the packet to the hub link.
2. After receiving a packet from the hub link, the PE does not learn the MAC address. Instead, it searches the MAC address table to forward the packet to a spoke link.

Local E-tree networking

E-Tree networks are applicable to P2MP services. In a local E-tree network, ACs are classified into root ACs and leaf ACs. To improve security and isolate ACs, enable E-Tree on the PEs. Then,

communication is allowed only between root links and between root and leaf links. Leafs cannot communicate with each other.

Figure 19 Local E-tree network

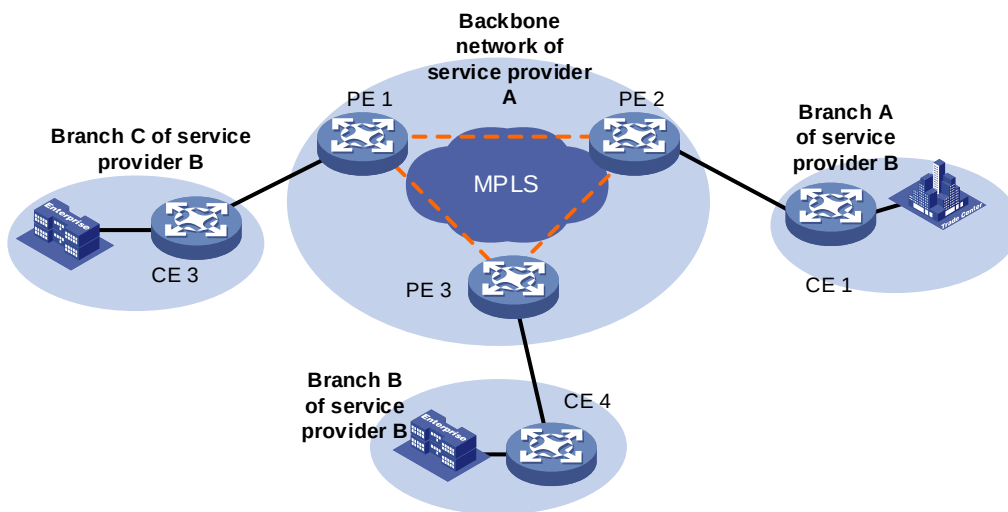


Application scenarios

VPLS applies to large enterprises that hold their own maintenance staff and feature large network scale, great quantity of routes, geographically dispersed branches, and higher demand for service quality.

As shown in Figure 20, service provider A holds a backbone that covers the whole country, and service provider B wants to rent the bandwidth of service provider A to connect its branches geographically dispersed in several cities. Service provider B has enough network management and maintenance capability. To provide high quality services and keep the privacy of routes, service provider B adopts VPLS for its network.

Figure 20 Typical VPLS network



References

- RFC 4447: *Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)*
- RFC 4761: *Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling*
- RFC 4762: *Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling*