Internet Engineering Task Force (IETF)
Request for Comments: 6246
Category: Informational
ISSN: 2070-1721

A. Sajassi, Ed.
F. Brockners
Cisco Systems
D. Mohan, Ed.
Nortel
Y. Serbest
AT&T

June 2011

Virtual Private LAN Service (VPLS) Interoperability with Customer Edge (CE) Bridges

Abstract

One of the main motivations behind Virtual Private LAN Service (VPLS) is its ability to provide connectivity not only among customer routers and servers/hosts but also among customer IEEE bridges. VPLS is expected to deliver the same level of service that current enterprise users are accustomed to from their own enterprise bridged networks or their Ethernet Service Providers.

When customer edge (CE) devices are IEEE bridges, then there are certain issues and challenges that need to be accounted for in a VPLS network. The majority of these issues have been addressed in the IEEE 802.1ad standard for provider bridges and they can be leveraged for VPLS networks. This document extends the provider edge (PE) model described in RFC 4664 based on IEEE 802.1ad bridge module, and it illustrates a clear demarcation between the IEEE bridge module and IETF LAN emulation module. By doing so, it shows that the majority of interoperability issues with CE bridges can be delegated to the 802.1ad bridge module, thus removing the burden on the IETF LAN emulation module within a VPLS PE.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Virtual Private LAN Service (VPLS) is a LAN emulation service intended for providing connectivity between geographically dispersed customer sites across MANs/WANs (over MPLS/IP), as if they were connected using a LAN. One of the main motivations behind VPLS is its ability to provide connectivity not only among customer routers and servers/hosts but also among IEEE customer bridges. If only connectivity among customer IP routers/hosts is desired, then an IP-only LAN Service [IPLS] solution could be used. The strength of the VPLS solution is that it can provide connectivity to both bridge and non-bridge types of CE devices. VPLS is expected to deliver the same level of service that current enterprise users are accustomed to from their own enterprise bridged networks [802.1D] [802.1Q] today or the same level of service that they receive from their Ethernet Service Providers using IEEE 802.1ad-based networks [802.1ad] (or its predecessor, QinQ-based networks).

When CE devices are IEEE bridges, then there are certain issues and challenges that need to be accounted for in a VPLS network. The majority of these issues have been addressed in the IEEE 802.1ad standard for provider bridges and they can be leveraged for VPLS networks. This document extends the PE model described in [RFC4664] based on the IEEE 802.1ad bridge module and illustrates a clear demarcation between IEEE bridge module and IETF LAN emulation module. By doing so, it describes that the majority of interoperability issues with CE bridges can be delegated to the 802.1ad bridge module,

thus removing the burden on the IETF LAN emulation module within a VPLS PE. This document discusses these issues and, wherever possible, suggests areas to be explored in rectifying these issues. The detailed solution specification for these issues is outside of the scope of this document.

This document also discusses interoperability issues between VPLS and IEEE 802.1ad networks when the end-to-end service spans across both types of networks, as outlined in [RFC4762].

This document categorizes the CE-bridge issues into two groups: 1) mandatory and 2) optional. The issues in group (1) need to be addressed in order to ensure the proper operation of CE bridges. The issues in group (2) would provide additional operational improvement and efficiency and may not be required for interoperability with CE bridges. Sections 5 and 6 discuss these mandatory and optional issues, respectively.

1.1. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Ethernet Service Instance

Before starting the discussion of bridging issues, it is important to clarify the Ethernet Service definition. The term VPLS has different meanings in different contexts. In general, VPLS is used in the following contexts [RFC6136]: a) as an end-to-end bridged LAN service over one or more networks (one of which is an MPLS/IP network), b) as an MPLS/IP network supporting these bridged LAN services, and c) as (V)LAN emulation. For better clarity, we differentiate between its usage as network versus service by using the terms VPLS network and VPLS instance, respectively. Furthermore, we confine VPLS (both network and service) to only the portion of the end-to-end network that spans an MPLS/IP network. For an end-to-end service (among different sites of a given customer), we use the term "Ethernet Service Instance" or ESI.

We define the Ethernet Service Instance (ESI) as an association of two or more Attachment Circuits (ACs) over which an Ethernet service is offered to a given customer. An AC can be either a User-Network Interface (UNI) or a Network-Network Interface (NNI); furthermore, it can be an Ethernet interface or a VLAN, it can be an ATM or Frame Relay Virtual Circuit, or it can be a PPP/HDLC (PPP/High-Level Data

Link Control) interface. If an ESI is associated with more than two ACs, then it is a multipoint ESI. In this document, wherever the keyword ESI is used, it means multipoint ESI unless stated otherwise.

An ESI can correspond to a VPLS instance if its associated ACs are only connected to a VPLS network, or an ESI can correspond to a Service VLAN if its associated ACs are only connected to a Provider-Bridged network [802.1ad]. Furthermore, an ESI can be associated with both a VPLS instance and a Service VLAN when considering an end-to-end service that spans across both VPLS and Provider-Bridged networks. An ESI can span across different networks (e.g., IEEE 802.1ad and VPLS) belonging to the same or different administrative domains.

An ESI most often represents a customer or a specific service requested by a customer. Since traffic isolation among different customers (or their associated services) is of paramount importance in service provider networks, its realization shall be done such that it provides a separate Media Access Control (MAC) address domain and broadcast domain per ESI. A separate MAC address domain is provided by using a separate MAC forwarding table (e.g., Forwarding Information Base (FIB), also known as filtering database [802.1D]) per ESI (for both VPLS and IEEE 802.1ad networks). A separate broadcast domain is provided by using a full mesh of pseudowires per ESI over the IP/MPLS core in a VPLS network and/or a dedicated Service VLAN per ESI in an IEEE 802.1ad network.

3. VPLS-Capable PE Model with Bridge Module

[RFC4664] defines three models for VPLS-capable PE (VPLS-PE), based on the bridging functionality that needs to be supported by the PE. If the CE devices can be routers/hosts or IEEE bridges, the second model from [RFC4664] is the most suitable, and it is both adequate to provide the VPLS level of service and consistent with the IEEE standards for Provider Bridges [802.1ad]. We briefly describe the second model and then expand upon this model to show its sub-components based on the [802.1ad] Provider Bridge model.

As described in [RFC4664], the second model for VPLS-PE contains a single bridge module supporting all the VPLS instances on that PE , where each VPLS instance is represented by a unique VLAN inside that bridge module (also known as a Service VLAN or S-VLAN). The bridge module has a single "Emulated LAN" interface over which it communicates with all VPLS forwarders, and each VPLS instance is represented by a unique S-VLAN tag. Each VPLS instance can consist of a set of pseudowires, and its associated forwarder can correspond to a single VLAN as depicted in Figure 1 below. Thus, sometimes it is referred to as VLAN emulation.

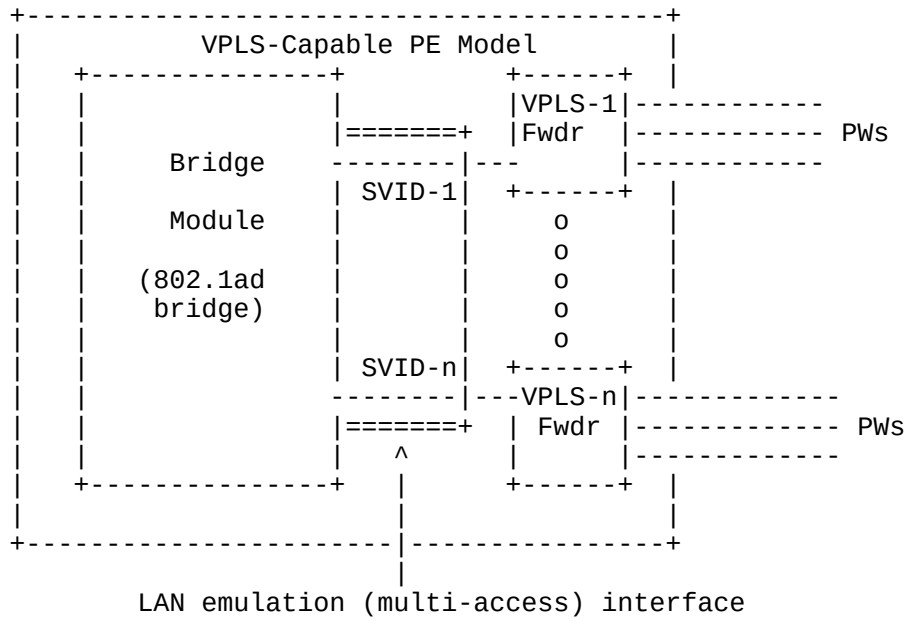

Figure 1. VPLS-Capable PE Model

Customer frames associated with a given ESI carry the S-VLAN ID for that ESI over the LAN emulation interface. The S-VLAN ID is stripped before transmitting the frames over the set of pseudowires (PWs) associated with that VPLS instance (assuming raw mode PWs are used as specified in [RFC4448]).

The bridge module can itself consist of one or two sub-components, depending on the functionality that it needs to perform. Figure 2 depicts the model for the bridge module based on [802.1ad].

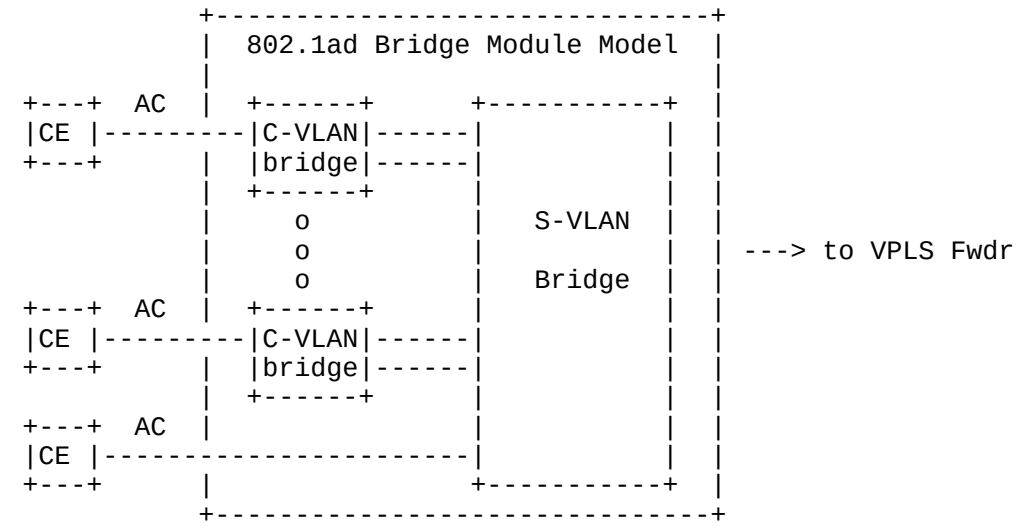

Figure 2. Model of the 802.1ad Bridge Module

The S-VLAN bridge component is always required and it is responsible for tagging customer frames with S-VLAN tags in the ingress direction (from customer UNIs) and removing S-VLAN tags in the egress direction (toward customer UNIs). It is also responsible for running the provider's bridge protocol -- such as Rapid Spanning Tree Protocol (RSTP), Multiple Spanning Tree Protocol (MSTP), Generic VLAN Registration Protocol (GVRP), GARP Multicast Registration Protocol (GMRP), etc. -- among provider bridges within a single administrative domain.

The customer VLAN (C-VLAN) bridge component is required when the customer Attachment Circuits are VLANs (aka C-VLANs). In such cases, the VPLS-capable PE needs to participate in some of the customer's bridging protocol such as RSTP and MSTP. Such participation is required because a C-VLAN at one site can be mapped into a different C-VLAN at a different site or, in case of asymmetric mapping, a customer Ethernet port at one site can be mapped into a C-VLAN (or group of C-VLANs) at a different site.

The C-VLAN bridge component does service selection and identification based on C-VLAN tags. Each frame from the customer device is assigned to a C-VLAN and presented at one or more internal port-based interfaces, each supporting a single service instance that the customer desires to carry that C-VLAN. Similarly, frames from the provider network are assigned to an internal interface or 'LAN' (e.g, between C-VLAN and S-VLAN components) on the basis of the S-VLAN tag. Since each internal interface supports a single service instance, the

S-VLAN tag can be, and is, removed at this interface by the S-VLAN bridge component. If multiple C-VLANs are supported by this service instance (e.g., via VLAN bundling or port-based service), then the frames will have already been tagged with C-VLAN tags. If a single C-VLAN is supported by this service instance (e.g., VLAN-based), then the frames will not have been tagged with a C-VLAN tag since C-VLAN can be derived from the S-VLAN (e.g., one-to-one mapping). The C-VLAN-aware bridge component applies a port VLAN ID (PVID) to untagged frames received on each internal 'LAN', allowing full control over the delivery of frames for each C-VLAN through the Customer UNI Port.

4. Mandatory Issues

4.1. Service Mapping

Different Ethernet AC types can be associated with a single Ethernet Service Instance (ESI). For example, an ESI can be associated with only physical Ethernet ports, VLANs, or a combination of the two (e.g., one end of the service could be associated with physical Ethernet ports and the other end could be associated with VLANs). In [RFC4762], unqualified and qualified learning are used to refer to port-based and VLAN-based operation, respectively. [RFC4762] does not describe the possible mappings between different types of Ethernet ACs (e.g., 802.1D, 802.1Q, or 802.1ad frames). In general, the mapping of a customer port or VLAN to a given service instance is a local function performed by the local PE, and the service provisioning shall accommodate it. In other words, there is no reason to restrict and limit an ESI to have only port-based ACs or to have only VLAN-based ACs. [802.1ad] allows for each customer AC (either a physical port, a VLAN, or a group of VLANs) to be mapped independently to an ESI that provides better service offerings to enterprise customers. For better and more flexible service offerings and for interoperability purposes between VPLS and 802.1ad networks, it is imperative that both networks offer the same capabilities in terms of customer ACs mapping to the customer service instance.

The following table lists possible mappings that can exist between customer ACs and their associated ESIs. As can be seen, there are several possible ways to perform such mappings. In the first scenario, it is assumed that an Ethernet physical port only carries untagged traffic and all traffic is mapped to the corresponding service instance or ESI. This is referred to as "port-based with untagged traffic". In the second scenario, it is assumed that an Ethernet physical port carries both tagged and untagged traffic and all that traffic is mapped to the corresponding service instance or ESI. This is referred to as "port-based with tagged and untagged traffic". In the third scenario, it is assumed that only a single

VLAN is mapped to the corresponding service instance or ESI. This is referred to as "VLAN-based". Finally, in the fourth scenario, it is assumed that a group of VLANs from the Ethernet physical interface is mapped to the corresponding service instance or ESI. This is referred to as "VLAN bundling".

Ethernet I/F & Associated Service Instance(s)				
	Port-based untagged	Port-based tagged & untagged	VLAN-based	VLAN bundling
Port-based untagged	Y	N	Y(Note-1)	N
Port-based tagged & untagged	N	Y	Y(Note-2)	Y
VLAN-based	Y(Note-1)	Y(Note-2)	Y	Y(Note-3)
VLAN Bundling	N	Y	Y(Note-3)	Y

Note-1: In this asymmetric mapping scenario, it is assumed that the CE device with "VLAN-based" AC is capable of supporting [802.1Q] frame format.

Note-2: In this asymmetric mapping scenario, it is assumed that the CE device with "VLAN-based" AC can support [802.1ad] frame format because it will receive Ethernet frames with two tags, where the outer tag is an S-VLAN and the inner tag is a C-VLAN received from "port-based" AC. One application example for such CE device is in a Broadband Remote Access Server (BRAS) for DSL aggregation over a Metro Ethernet network.

Note-3: In this asymmetric mapping scenario, it is assumed that the CE device with "VLAN-based" AC can support the [802.1ad] frame format because it will receive Ethernet frames with two tags, where the outer tag is an S-VLAN and the inner tag is a C-VLAN received from "VLAN bundling" AC.

If a PE uses an S-VLAN tag for a given ESI (either by adding an S-VLAN tag to customer traffic or by replacing a C-VLAN tag with a S-VLAN tag), then the frame format and EtherType for S-VLAN SHALL adhere to [802.1ad].

As mentioned before, the mapping function between the customer AC and its associated ESI is a local function; thus, when the AC is a single customer VLAN, it is possible to map different customer VLANs at different sites to a single ESI without coordination among those sites.

When a port-based mapping or a VLAN-bundling mapping is used, then the PE may use an additional S-VLAN tag to mark the customer traffic received over that AC as belonging to a given ESI. If the PE uses the additional S-VLAN tag, then in the opposite direction the PE SHALL strip the S-VLAN tag before sending the customer frames over the same AC. However, when VLAN-mapping mode is used at an AC and if the PE uses the S-VLAN tag locally, then if the Ethernet interface is a UNI, the tagged frames over this interface SHALL have a frame format based on [802.1Q]. In such a case, the PE SHALL translate the customer tag (C-VLAN) into the provider tag (S-VLAN) upon receiving a frame from the customer. In the opposite direction, the PE SHALL translate from provider frame format (802.1ad) back to customer frame format (802.1Q).

All the above asymmetric services can be supported via the PE model with the bridge module depicted in Figure 2 (based on [802.1ad]).

4.2. CE Bridge Protocol Handling

When a VPLS-capable PE is connected to a CE bridge, then -- depending on the type of Attachment Circuit -- different protocol handling may be required by the bridge module of the PE. [802.1ad] states that when a PE is connected to a CE bridge, then the service offered by the PE may appear to specific customer protocols running on the CE in one of the four ways:

a) Transparent to the operation of the protocol among CEs of different sites using the service provided, appearing as an individual LAN without bridges;

b) Discarding frames, acting as a non-participating barrier to the operation of the protocol;

c) Peering, with a local protocol entity at the point of provider ingress and egress, participating in and terminating the operation of the protocol; or

d) Participation in individual instances of customer protocols.

All the above CE bridge protocol handling can be supported via the PE model with the bridge module depicted in Figure 2 (based on [802.1ad]). For example, when an Attachment Circuit is port-based, then the bridge module of the PE can operate transparently with respect to the CE's RSTPs or MSTPs (and thus no C-VLAN component is required for that customer UNI). However, when an Attachment Circuit is VLAN-based (either VLAN-based or VLAN bundling), then the bridge module of the PE needs to peer with the RSTPs or MSTPs running on the CE (and thus the C-VLAN bridge component is required). In other words, when the AC is VLAN-based, then protocol peering between CE and PE devices may be needed. There are also protocols that require peering but are independent from the type of Attachment Circuit. An example of such protocol is the link aggregation protocol [802.1AX]; however, this is a media-dependent protocol as its name implies.

[802.1ad] reserves a block of 16 MAC addresses for the operation of C-VLAN and S-VLAN bridge components. Also, it shows which of these reserved MAC addresses are only for C-VLAN bridge components, which are only for S-VLAN bridge components, and which apply to both C-VLAN and S-VLAN components.

4.3. Partial Mesh of Pseudowires

A VPLS service depends on a full mesh of pseudowires, so a pseudowire failure reduces the underlying connectivity to a partial mesh, which can have adverse effects on the VPLS service. If the CE devices belonging to an ESI are routers running link state routing protocols that use LAN procedures over that ESI, then a partial mesh of PWs can result in "black holing" traffic among the selected set of routers. And if the CE devices belonging to an ESI are IEEE bridges, then a partial mesh of PWs can cause broadcast storms in the customer and provider networks. Furthermore, it can cause multiple copies of a single frame to be received by the CE and/or PE devices. Therefore, it is of paramount importance to be able to detect PW failure and to take corrective action to prevent creation of partial mesh of PWs.

When the PE model depicted in Figure 2 is used, then [802.1ag] procedures could be used for detection of partial mesh of PWs. [802.1ag] defines a set of procedures for fault detection, verification, isolation, and notification per ESI.

The fault detection mechanism of [802.1ag] can be used to perform connectivity check among PEs belonging to a given VPLS instance. It checks the integrity of a service instance end-to-end within an administrative domain, e.g., from one AC at one end of the network to another AC at the other end of the network. Therefore, its path

coverage includes the bridge module within a PE and it is not limited to just PWs. Furthermore, [802.1ag] operates transparently over the full mesh of PWs for a given service instance since it operates at the Ethernet level (and not at the PW level). It should be noted that since a PW consists of two unidirectional Label Switched Paths (LSPs), then one direction can fail independently of the other. Even in this case, the procedures of [802.1ag] can provide a consistent view of the full mesh to the participating PEs by relying on remote defect indication (RDI).

Another, less preferred, option is to define a procedure for detection of partial mesh; in this procedure, each PE keeps track of the status of its PW Endpoint Entities (EEs, e.g., VPLS forwarders) as well as the EEs reported by other PEs. Therefore, upon a PW failure, the PE that detects the failure not only takes notice locally but also notifies other PEs belonging to that service instance so that all the participant PEs have a consistent view of the PW mesh. Such a procedure is for the detection of partial mesh per service instance, and in turn it relies on additional procedure for PW failure detection such as Bidirectional Forward Detection (BFD) or Virtual Circuit Connectivity Verification (VCCV). Given that there can be tens (or even hundreds) of thousands of PWs in a PE, there can be scalability issues with such fault detection/notification procedures.

4.4. Multicast Traffic

VPLS follows a centralized model for multicast replication within an ESI. VPLS relies on ingress replication. The ingress PE replicates the multicast packet for each egress PE and sends it to the egress PE using point-to-point PW over a unicast tunnel. VPLS operates on an overlay topology formed by the full mesh of pseudo-wires. Thus, depending on the underlying topology, the same datagram can be sent multiple times down the same physical link. VPLS currently does not offer any mechanisms to restrict the distribution of multicast or broadcast traffic of an ESI throughout the network, which causes an additional burden on the ingress PE through unnecessary packet replication. This in turn causes additional load on the MPLS core network and additional processing at the receiving PE where extraneous multicast packets are discarded.

One possible approach to delivering multicast more efficiently over a VPLS network is to include the use of IGMP snooping in order to send the packet only to the PEs that have receivers for that traffic, rather than to all the PEs in the VPLS instance. If the customer bridge or its network has dual-home connectivity, then -- for proper operation of IGMP snooping -- the PE must generate a "General Query" over that customer’s UNIs upon receiving a customer topology change

notification as described in [RFC4541]. A "General Query" by the PE results the customer multicast MAC address(es) being properly registered at the PE when there are customer topology changes. It should be noted that IGMP snooping provides a solution for IP multicast packets and is not applicable to general multicast data.

Using the IGMP snooping as described, the ingress PE can select a subset of PWs for packet replication, thus avoiding sending multicast packets to the egress PEs that don't need them. However, the replication is still performed by the ingress PE. In order to avoid replication at the ingress PE, one may want to use multicast distribution trees (MDTs) in the provider core network; however, this brings some potential pitfalls. If the MDT is used for all multicast traffic of a given customer, then this results in customer multicast and unicast traffic being forwarded on different PWs and even on a different physical topology within the provider network. This is a serious issue for customer bridges because customer Bridge Protocol Data Units (BPDUs), which are multicast data, can take a different path through the network than the unicast data. Situations might arise where either unicast OR multicast connectivity is lost. If unicast connectivity is lost but multicast forwarding continues to work, the customer spanning tree would not take notice which results in loss of its unicast traffic. Similarly, if multicast connectivity is lost, but unicast is working, then the customer spanning tree will activate the blocked port, which may result in a loop within the customer network. Therefore, the MDT cannot be used for both customer multicast control and data traffic. If it is used, it should only be limited to customer data traffic. However, there can be a potential issue even when it is used for customer data traffic since the MDT doesn't fit the PE model described in Figure 1 (it operates independently from the full mesh of PWs that correspond to an S-VLAN). It is also not clear how connectivity fault management (CFM) procedures (802.1ag) used for the ESI integrity check (e.g., per service instance) can be applied to check the integrity of the customer multicast traffic over the provider MDT. Because of these potential issues, the specific applications of the provider MDT to customer multicast traffic shall be documented and its limitations be clearly specified.

5. Optional Issues

5.1. Customer Network Topology Changes

A single CE or a customer network can be connected to a provider network using more than one User-Network Interface (UNI). Furthermore, a single CE or a customer network can be connected to more than one provider network. [RFC4665] provides some examples of such customer network connectivity; they are depicted in Figure 3

below. Such network topologies are designed to protect against the failure or removal of network components from the customer network, and it is assumed that the customer leverages the spanning tree protocol to protect against these cases. Therefore, in such scenarios, it is important to flush customer MAC addresses in the provider network upon the customer topology change in order to avoid black-holing of customer frames.

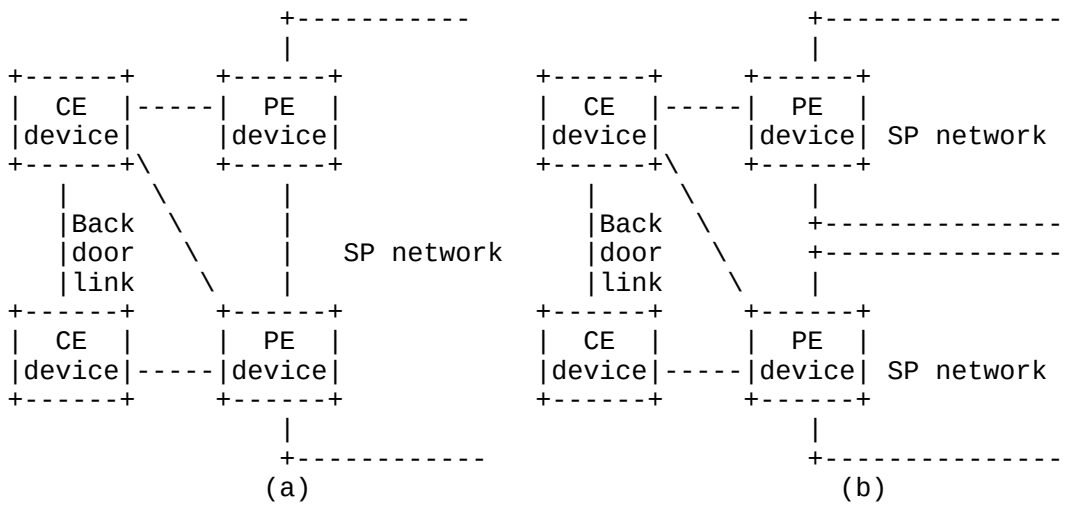

Figure 3. Combination of Dual-Homing and Backdoor Links for CE Devices

The customer networks use their own instances of the spanning tree protocol to configure and partition their active topology so that the provider connectivity doesn’t result in a data loop. Reconfiguration of a customer’s active topology can result in the apparent movement of customer end stations from the point of view of the PEs. There are two methods for addressing this issue based on the provider bridge model depicted in Figure 1. In the first method, the Topology Change Notification (TCN) message received from the CE device is translated into one or more out-of-band "MAC Address Withdrawal" messages as specified in [RFC4762]. In the second method, the TCN message received from the CE device is translated into one or more in-band "Flush" messages per [p802.1Qbe]. The second method is recommended because of ease of interoperability between the bridge and LAN emulation modules of the PE.

5.2. Redundancy

[RFC4762] talks about dual-homing of a given Multi-Tenant Unit switch (MTU-s) to two PEs over a provider MPLS access network to provide protection against link and node failure. For example, in case the primary PE fails or the connection to it fails, then the MTU-s uses the backup PWs to reroute the traffic to the backup PE. Furthermore, it discusses the provision of redundancy when a provider Ethernet access network is used and how any arbitrary access network topology (not just hub-and-spoke) can be supported using the provider's MSTP protocol. It also discusses how the provider MSTP for a given access network can be confined to that access network and operate independently from MSTP protocols running in other access networks.

In both types of redundancy mechanism (Ethernet and MPLS access networks), only one PE is active for a given VPLS instance at any time. In case of an Ethernet access network, core-facing PWs (for a VPLS instance) at the PE are blocked by the MSTP; whereas, in case of a MPLS access network, the access-facing PW is blocked at the MTU-s for a given VPLS instance.

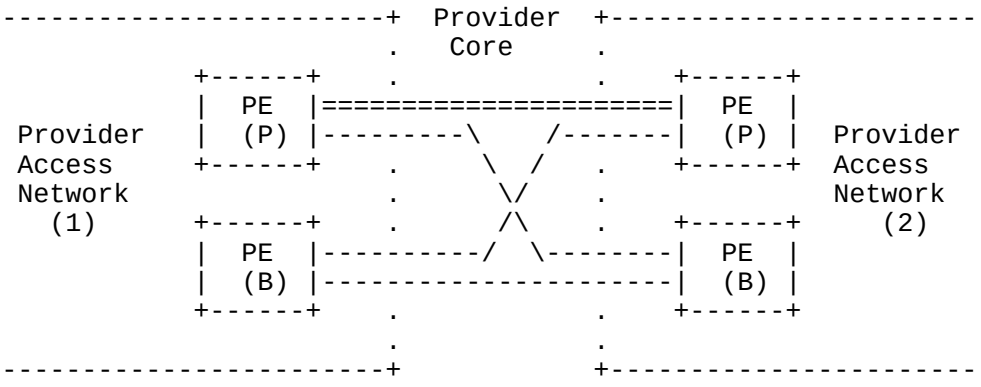

Figure 4. Bridge Module Model

Figure 4 shows two provider access networks each with two PEs that are connected via a full mesh of PWs for a given VPLS instance. As shown in the figure, only one PE in each access network serves as a Primary PE (P) for that VPLS instance and the other PE serves as the backup PE (B). In this figure, each primary PE has two active PWs originating from it. Therefore, when a multicast, broadcast, and unknown unicast frame arrives at the primary PE from the access network side, the PE replicates the frame over both PWs in the core even though it only needs to send the frame over a single PW (shown with "==" in Figure 4) to the primary PE on the other side. This is an unnecessary replication of the customer frames and consumes core-

network bandwidth (half of the frames get discarded at the receiving PE). This issue is aggravated when there are more than two PEs per provider access network -- e.g., if there are three PEs or four PEs per access network, then 67% or 75%, respectively, of core-network bandwidth for multicast, broadcast, and unknown unicast are respectively wasted.

Therefore, it is recommended to have a protocol among PEs that can disseminate the status of PWs (active or blocked) among themselves. Furthermore, it is recommended to have the protocol tied up with the redundancy mechanism such that (per VPLS instance) the status of active/backup PE gets reflected on the corresponding PWs emanating from that PE.

The above discussion was centered on the inefficiency regarding packet replication over MPLS core networks for current VPLS redundancy mechanism. Another important issue to consider is the interaction between customer and service provider redundancy mechanisms, especially when customer devices are IEEE bridges. If CEs are IEEE bridges, then they can run RSTPs or MSTPs. RSTP convergence and detection time is much faster than its predecessor (IEEE 802.1D STP, which is obsolete). Therefore, if the provider network offers a VPLS redundancy mechanism, then it should provide transparency to the customer’s network during a failure within its network, e.g., the failure detection and recovery time within the service provider network should be less than the one in the customer network. If this is not the case, then a failure within the provider network can result in unnecessary switch-over and temporary flooding/loop within the customer’s network that is dual-homed.

5.3. MAC Address Learning

When customer devices are routers, servers, or hosts, then the number of MAC addresses per customer sites is very limited (most often one MAC address per CE). However, when CEs are bridges, then there can be many customer MAC addresses (e.g., hundreds of MAC addresses) associated with each CE.

[802.1ad] has devised a mechanism to alleviate MAC address learning within provider Ethernet networks that can equally be applied to VPLS networks. This mechanism calls for disabling MAC address learning for an S-VLAN (or a service instance) within a provider bridge (or PE) when there is only one ingress and one egress port associated with that service instance on that PE. In such cases, there is no need to learn customer MAC addresses on that PE since the path through that PE for that service instance is fixed. For example, if a service instance is associated with four CEs at four different sites, then the maximum number of provider bridges (or PEs) that need

to participate in that customer MAC address learning is only three, regardless of how many PEs are in the path of that service instance. This mechanism can reduce the number of MAC addresses learned in a hierarchical VPLS (H-VPLS) with QinQ access configuration.

If the provider access network is of type Ethernet (e.g., IEEE 802.1ad-based network), then the MSTP can be used to partition the access network into several loop-free spanning tree topologies where Ethernet service instances (S-VLANs) are distributed among these tree topologies. Furthermore, GVRP can be used to limit the scope of each service instance to a subset of its associated tree topology (thus limiting the scope of customer MAC address learning to that sub-tree). Finally, the MAC address disabling mechanism (described above) can be applied to that sub-tree to further limit the number of nodes (PEs) on that sub-tree that need to learn customer MAC addresses for that service instance.

Furthermore, [802.1ah] provides the capability of encapsulating customers' MAC addresses within the provider MAC header. A MTU-s capable of this functionality can significantly reduce the number of MAC addresses learned within the provider network for H-VPLS with QinQ access, as well as H-VPLS with MPLS access.

6. Interoperability with 802.1ad Networks

[RFC4762] discusses H-VPLS provider-network topologies with both Ethernet [802.1ad] and MPLS access networks. Therefore, it is important to ensure seamless interoperability between these two types of networks.

Provider bridges as specified in [802.1ad] are intended to operate seamlessly with customer bridges and provide the required services. Therefore, if a PE is modeled based on Figures 1 and 2, which include a [802.1ad] bridge module, then it should operate seamlessly with Provider Bridges given that the issues discussed in this document have been taken into account.

7. Acknowledgments

The authors would like to thank Norm Finn and Samer Salam for their comments and valuable feedback.

8. Security Considerations

In addition to the security issues described in [RFC4762], the following considerations apply:

- When a CE that is a customer bridge is connected to the VPLS network, it may be desirable to secure the end-to-end communication between the customer bridge nodes across the VPLS network. This can be accomplished by running [802.1AE] MAC security between the C-VLAN components of the customer bridges. In this case, the VPLS PEs must ensure transparent delivery of the encryption/security protocol datagrams using the Bridge Group Address [802.1ad].

- When a CE that is a customer bridge is connected to the VPLS network, it may be desirable to secure the communication between the customer bridge and its directly connected PE. If the PE is modeled to include a [802.1ad] bridge module, then this can be achieved by running MAC security between the customer bridge and the S-VLAN component of the VPLS PE as described in Section 7.7.2 of [802.1AX].

- When an 802.1ad network is connected to a VPLS network, it is possible to secure the NNI between the two networks using the procedures of [802.1AE] and [802.1AX] between the S-VLAN components of the Provider Edge Bridge and the attached VPLS PE, as long as the PE is modeled to include an [802.1ad] bridge module.

9. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4762] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.

[802.1ad] IEEE 802.1ad-2005, "Amendment to IEEE 802.1Q-2005. IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks Revision-Amendment 4: Provider Bridges".

[802.1AE] IEEE 802.1AE-2006, "IEEE Standard for Local and Metropolitan Area Networks - Media Access Control (MAC) Security".

[802.1ag] IEEE 802.1ag-2007, "IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management".

[802.1ah] IEEE 802.1ah-2008, "IEEE Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks Amendment 7: Provider Backbone Bridges".

[802.1AX] IEEE 802.1AX-2008 "IEEE Standard for Local and Metropolitan Area Networks - Link Aggregation".

10. Informative References

[IPLS] Shah, H., Rosen, E., Le Faucheur, F., and G. Heron, "IP-Only LAN Service (IPLS)", Work in Progress, February 2010.

[RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.

[RFC4541] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, May 2006.

[RFC4664] Andersson, L., Ed., and E. Rosen, Ed., "Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, September 2006.

[RFC4665] Augustyn, W., Ed., and Y. Serbest, Ed., "Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks", RFC 4665, September 2006.

[RFC6136] Sajassi, A., Ed., and D. Mohan, Ed., "Layer 2 Virtual Private Network (L2VPN) Operations, Administration, and Maintenance (OAM) Requirements and Framework", RFC 6136, March 2011.

[802.1D] IEEE 802.1D-2004, "IEEE Standard for Local and Metropolitan Area Networks - Media access control (MAC) Bridges (Incorporates IEEE 802.1t-2001 and IEEE 802.1w)".

[802.1Q] IEEE Std. 802.1Q-2003 "Virtual Bridged Local Area Networks".

[p802.1Qbe] IEEE Draft Standard P802.1Qbe, "IEEE Draft Standard for Local and Metropolitan Area Networks -- Virtual Bridged Local Area Networks Amendment: Multiple I-SID Registration Protocol".

Authors' Addresses

Ali Sajassi (editor)
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
EMail: sajassi@cisco.com

Frank Brockners
Cisco Systems, Inc.
Hansaallee 249
40549 Duesseldorf
Germany
EMail: fbrockne@cisco.com

Dinesh Mohan (editor)
Nortel
Ottawa, ON K2K3E5
EMail: dinmohan@hotmail.com

Yetik Serbest
AT&T Labs
9505 Arboretum Blvd.
Austin, TX 78759
EMail: yetik_serbest@labs.att.com