Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 12, 2012

Y. Kamite
NTT Communications
F. Jounay
France Telecom
B. Niven-Jenkins
Velocix
D. Brungard
AT&T
L. Jin
ZTE

July 11, 2011

Framework and Requirements for Virtual Private Multicast Service (VPMS)

draft-ietf-l2vpn-vpms-frmwk-requirements-04.txt

Abstract

This document provides a framework and service level requirements for Virtual Private Multicast Service (VPMS). VPMS is defined as a Layer 2 VPN service that provides point-to-multipoint connectivity for a variety of Layer 2 link layers across an IP or MPLS-enabled PSN. This document outlines architectural service models of VPMS and states generic and high level requirements. This is intended to aid in developing protocols and mechanisms to support VPMS.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 12, 2012.

Table of Contents

1. Introduction

1.1. Problem Statement

[RFC4664] describes different types of Provider Provisioned Layer 2 VPNs (L2 PPVPNs, or L2VPNs). Some of them are widely deployed today, such as Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS). A VPWS is a VPN service that supplies a Layer 2 (L2) point-to-point service. A VPLS is an L2 service that emulates Ethernet LAN service across a Wide Area Network (WAN).

For some use cases described hereafter, there are P2MP (point-to-multipoint) type services for Layer 2 traffic. However, there is no straightforward way to realize them based on the existing L2VPN specifications.

In a VPWS, a SP can set up point-to-point connectivity per a pair of CEs but it is not possible to replicate traffic for point-to-multipoint services in the SP's network side. A SP could build multiple Pseudowires (PWs) independently and have the CEs replicate traffic over them, but this is not only inconvenient for the customer, it's a waste of bandwidth resources.

In a VPLS, SPs can natively offer multipoint connectivity across their backbone. Although it is seemingly applicable for point-to-multipoint service as well, there remains extra complexity for SPs to filter unnecessary traffic between irrelevant sites (i.e., from a receiver PE to another receiver PE) because VPLS provides multipoint-to-multipoint connectivity between CEs. Moreover, VPLS's MAC-based learning/forwarding operation is unnecessary for some scenarios particularly if customers only need simple unidirectional point-to-multipoint service, or if they require non- Ethernet Layer 2 connectivity.

Consequently, there is a real need for a solution that natively provides point-to-multipoint service in L2VPN.

1.2. Scope of This Document

VPMS is defined as a Layer 2 service that provides point-to-multipoint connectivity for a variety of Layer2 link layers across an IP or MPLS-enabled PSN. VPMS is categorized as a class of provider-provisioned Layer 2 Virtual Private Networks (L2VPN).

This document introduces a new service framework, reference model and functional requirements for VPMS by extending the existing framework [RFC4664] and requirements [RFC4665] for L2VPNs.

The technical specifications are outside the scope of this document. There is no intent to specify solution-specific details.

This document provides requirements from both the Service Provider's and the Customer's point of view.

2. Conventions used in this document

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] .

3. Terminology

The content of this document makes use of the terminology defined in [RFC4026]. For readability purposes, we list some of the terms here in addition to some specific terms used in this document.

3.1. Acronyms

P2P: Point-to-Point

P2MP: Point-to-Multipoint

PW: Pseudowire

VPMS: Virtual Private Multicast Service

PE/CE: Provider/Customer Edge

P: Provider Router

AC: Attachment Circuit

PSN: Packet Switched Network

SP: Service Provider

VPMS instance: A service entity manageable in a VPMS that provides isolated service reachability domain to each CE. It corresponds to a so-called "VPN" as a specific set of sites that allows communication.

P2MP connection: A logical entity between PE/ACs in a given VPMS instance that transfers unidirectional traffic transparently from one local ingress AC to one or more remote egress ACs.

4. Use Cases

4.1. Ethernet Use Case

4.1.1. Ethernet-based multicast stream

For multicast traffic delivery, there is a requirement to deliver a unidirectional P2MP service in addition to the existing P2P service. The demand is growing to provide private (P2MP native Ethernet) services, for various applications such as IP- based delivery of TV broadcasting, content delivery networks, etc. Moreover, many digital audio/video devices (e.g., MPEG-TS, HD-SDI) that support Ethernet interfaces are becoming available, which will make Ethernet P2MP service more common. Also there are some applications that naturally suited to static transport of VPMS. For example, MPEG-TS/IP/Ethernet in DVB-H is typically static broadcast without any signaling in the upstream direction. VPMS could be a possible solution to provide these kinds of networking connectivity over PSNs.

Currently VPLS [RFC4761][RFC4762] is able to give P2MP-type replication for Ethernet traffic. Native VPLS already supports this capability via a full mesh of PWs, and an extension to optimize replication is also proposed [I-D.ietf-l2vpn-vpls-mcast] as an additional feature. However, VPLS by nature requires MAC-based learning and forwarding, which might not be needed in some cases by particular users. Generally, video distribution applications use unidirectional P2MP traffic, but may not always require any added complexity of MAC address management. In addition, VPLS is a service that essentially provides any-to-any connectivity between all CEs in a L2VPN as it emulates a LAN service. However, if only P2MP connectivity is required, the traffic between leaves is not allowed. It might require extra efforts to guarantee this behavior in VPLS. And in some P2MP scenarios there no traffic from leafs to root. In these cases, VPMS is a service that provides much simpler operation.

Note that VPMS provides single coverage of receiver membership; that is, there is no distinct differentiation for multiple multicast groups. All traffic from a particular Attachment Circuit (AC) will be forwarded toward the same remote receivers, even if the destination MAC address is changed. Basically in VPMS, destination MAC addresses are not used for forwarding, which is significantly different from VPLS. If MAC-based forwarding is preferred (i.e., multicast/unicast differentiation of MAC address), VPLS should be

chosen rather than VPMS.

4.1.2. Ethernet-based time/frequency synchronization

Nowadays there exist several solutions to provide synchronization for time and/or frequency reference by the packet-based technology of Ethernet. For example, PTPv2 (Precision Time Protocol version 2) is a time-transfer protocol defined in the IEEE 1588-2008 standard. It provides precise synchronization of packet-based networks (e.g., Ethernet). It adopts two-way time transfer approach for synchronization. Time transfer protocol may be operated in multicast or unicast mode in both directions, and it is mapped over the Ethernet/IP/UDP protocol stack.

Moreover, PTPv2 telecom profile is now discussed in ITU-T that defines a set of capabilities and extensions required to support telecommunication applications. It aims at providing frequency distribution with higher level of accuracy. It allows unicast mode or the mix of unicast/multicast modes for the transmission of the PTP messages.

In this aspect, VPMS might be considered as a potential packet-based infrastructure to deliver multicast messages in PTPv2 with efficient forwarding. Note, however, in PTPv2 telecom profile, multicast transport may not always be supported in all the parts of a telecom network because multicast might sometimes generate additional PDV (packet delay variation) compared to unicast. Therefore, VPMS use case and the corresponding solution for this purpose will need more study in the future (e.g., PDV issue to be checked).

4.2. ATM-based Use Case

4.2.1. ATM-based multicast stream

A use case of ATM-based service in VPMS could be to offer the capability for service providers to support IP multicast wholesale services over ATM in case the wholesale customer relies on ATM infrastructure. The P2MP support alleviates the constraint in terms of replication for ATM to support IP multicast services.

Another use case of VPMS for ATM is for audio/video stream applications. Today many digital TV broadcasting networks adopt ATM-based distribution systems with point-to-multipoint PVPs/PVCs. The transport network supports replicating ATM cells in transit nodes to efficiently deliver programs to multiple terminals. For migrating such ATM-based networks onto IP/MPLS-based networks, VPMS is considered to be a candidate solution.

4.3. TDM-based Use Case

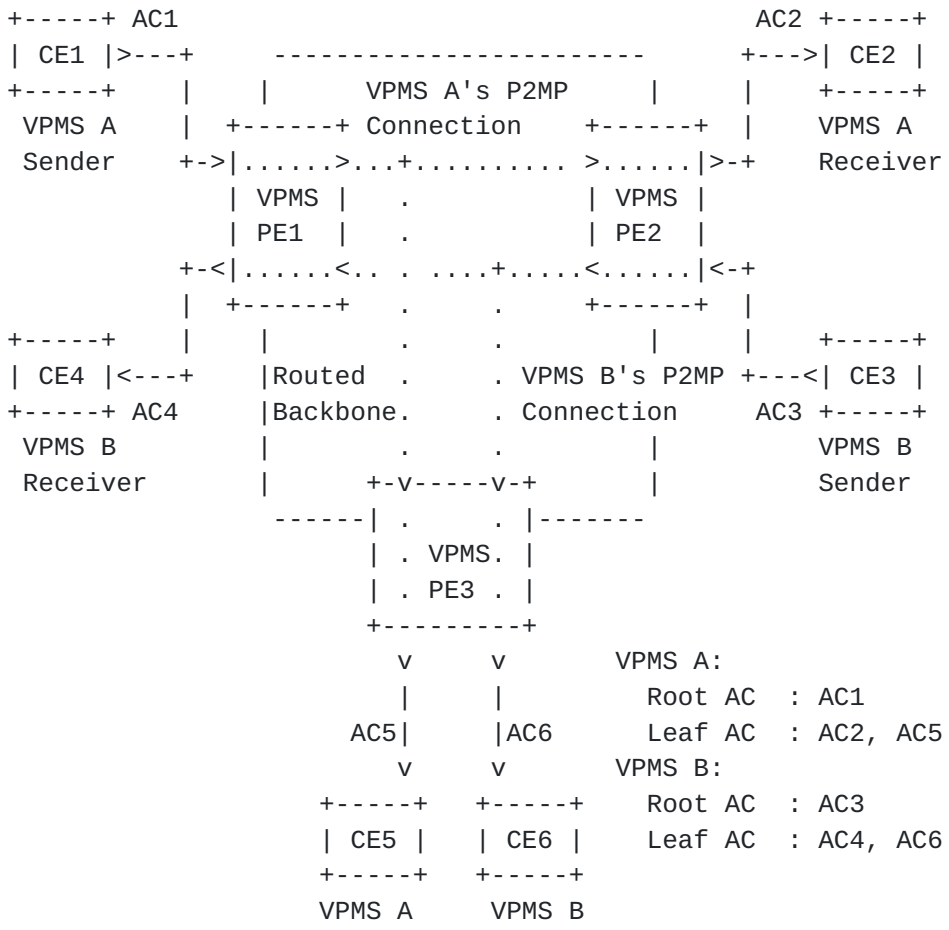
4.3.1. TDM-based multicast stream

Today the existing VPWS already supports TDM emulation services (SAToP, CESoPSN or TDMoIP). It is a Layer 1 service, not Layer 2 service; however, a common architecture is being used since they are all packet-based emulations over a SP's network. VPMS is also considered to be a solution for such TDM applications that require point-to-multipoint topology.

One use case is TDM-based multicast stream delivery, like video delivery. That is, data duplication is simply provided by Layer 1, without using upper layer's multicast protocols.

5. Reference Model

The VPMS reference model is shown in Figure 1.



Receiver Receiver

Figure 1: Reference Model for VPMS

A VPMS instance is defined as a service entity manageable in the VPMS architecture. A single VPMS instance provides an isolated service reachability domain to each CE, so it corresponds to a so-called "VPN" as it allows communication among a specific set of sites. A single VPMS instance provides a unique point-to-multipoint L2VPN service. In Figure 1, there are two VPMS instances shown, VPMS A and VPMS B. In principle, there is no traffic exchange allowed between these different instances, so they are treated as different VPNs.

In a VPMS, a single CE-PE link connection is used for transmitting frames for delivery to multiple remote CEs, with point-to-multipoint duplication. The SP's network (PE as well as P) has a role to replicate frames so that the sender's CE does not need to send multiple frames to individual receivers.

Like VPWS, an Attachment Circuit (AC) is provided to accommodate CEs in a VPMS. In a VPMS, an AC attached to a VPMS MUST be configured as "root" (sender) or "leaf" (receiver) not both. Any AC is associated with the role of either sending side (Tx) or receiving side (Rx) from the view of the CE. These will be named the root (sender) AC and leaf (receiver) AC respectively. Unless reverse traffic is optionally supported, a root AC does not transmit traffic back to a CE at upstream side, likewise a leaf AC does not receive traffic from a CE at downstream side. In Figure 1, AC1 and AC3 are configured as root ACs while AC2, AC4, AC5 and AC6 are configured as leaf ACs. In VPMS A, CE1 could send traffic via AC1, and CE2 and CE5 could receive the traffic.

A CE which is locally connected to a root AC is called a root (sender) CE. Also a CE which is locally connected to a leaf AC is called a leaf (receiver) CE. However, such CEs's roles will not be managed directly in VPMS because the configured AC's role (root or leaf) will automatically determine them.

Similarly, a PE which locally accommodates a root AC is called a root (sender) PE. A PE which locally accommodates a leaf AC is called a leaf (receiver) PE.

Root AC and leaf ACs are typically located at separate PEs as shown in Figure 1, but it is also allowed that a single PE locally has both a root AC and one or more leaf ACs.

Basically there is a one-to-one mapping between an attachment circuit and a VPMS instance. For example, all traffic from CE1 to PE1

(through AC1) is mapped to VPMS A (to CE2 and CE5).

In a VPMS, P2MP tree-shaped reachability is given from a single root AC to several leaf ACs. This will be named a "P2MP connection" in this VPMS framework. P2MP connection is a logical entity between PE/ACs in a given VPMS instance that transfers unidirectional traffic transparently from one local ingress AC to one or more remote egress ACs.

Similar to other L2VPN mechanisms, the VPMS architecture is based on PWs which may be using through IP or MPLS-enabled PSN tunnels over a routed backbone. That is, every P2MP connection can be instantiated by PW technology that supports P2MP traffic optimization (i.e., P2MP PW. See section 7.2.). P2MP traffic optimization will provide the benefit of traffic replication for high bandwidth efficiency. That is, the sender CE has only to transmit one stream towards the PE and it does not have to replicate traffic.

Regarding end-to-end traffic topology between the ACs, a single VPMS instance (i.e., one VPN) may correspond to a single P2MP connection. In Figure 1, VPMS A (one instance) has one P2MP connection (from AC1 to AC2 and AC5). However, there is also a case that a single VPMS consists of two or more P2MP connections grouped, which is typically used for multi-source or redundancy. The details are given in section 6.4 and 6.5.

VPMS can support various Layer 2 protocol services such as Ethernet, ATM, etc.

6. Customer Requirements

6.1. Service Topology

6.1.1. Point-to-Multipoint Traffic Support

A solution MUST support unidirectional point-to-multipoint traffic from a sender CE to multiple receiver CEs. A root CE can send traffic to one or more leaf CEs. Leaf CEs include not only the CEs which are located at remote sites, but also the local CEs which are connected to the same root PE (i.e., when a root AC and some of leaf ACs are co-located). If there is only one receiver in the instance, it is considered equivalent to unidirectional point-to-point traffic.

6.1.2. Reverse Traffic Support

There are cases where a reverse traffic flow is needed. A root CE may need to receive traffic from leaf CEs. There are some usage

scenarios for this, such as stream monitoring through a loopback mechanism, control channels which need feedback communication etc. A possible way to accomplish this is to provide different VPMS instances for reverse traffic, i.e. a root CE is a receiver of another VPMS instance. However, provisioning different VPNs for a particular customer would make its management task more complex. It is desired to have an alternative solution for supporting reverse traffic flow. This section provides additional requirements for this optional capability.

Therefore, a VPMS solution MAY support reverse traffic from a leaf AC to a root AC. Each reverse path is basically given in a P2P (unicast) manner. In other words, each leaf of the P2MP tree can individually send back traffic to the root. For this purpose a VPMS instance MAY have more than one reverse P2P connections as network entity; However, such network entities MUST have a common indetifier that enables themselves to be managed together in the same VPN. Thus any PWs used for such connections are expected to be assigned a common VPMS instance ID (i.e., VPN ID).

Note, a VPMS does not assume any-to-any multipoint reachability. Therefore, in principle, every leaf AC does not need to exchange traffic directly with other leaf ACs even if reverse traffic is supported.

Figure 3 illustrates this kind of scenario, where CE1 is configured as a sender in VPMS A. AC1 is a root AC, and AC2/AC3/AC4 are leaf ACs. P2MP connection is given for traffic in the forward direction. Unidirectional P2P connection is also provided for traffic in the reverse direction, from AC4 to AC1. Other reverse P2P connections can be provided similarly. (from AC2 to AC1 / from AC3 to AC1).

In this case, PEs need to deal with two types of traffic, locally-attached CE's sending (Tx) and receiving (Rx) flows. In Figure 3, they are both passing through the same physical PE-CE link (AC1 and AC4 respectively). But it is an implementation matter if Tx and Rx traffic are conveyed on the same physical link or separate links. It is also possible that a root PE multiplexes two ore more reverse traffic from different leaves and transmits it to an upstream CE over the same local physical link.

Note, in most implementations of VPWS today, every AC is always considered bidirectional. In contrast, in VPMS, every AC can be chosen unidirectional (if it is a totally unidirectional service), or bidirectional (if reverse traffic is supported).

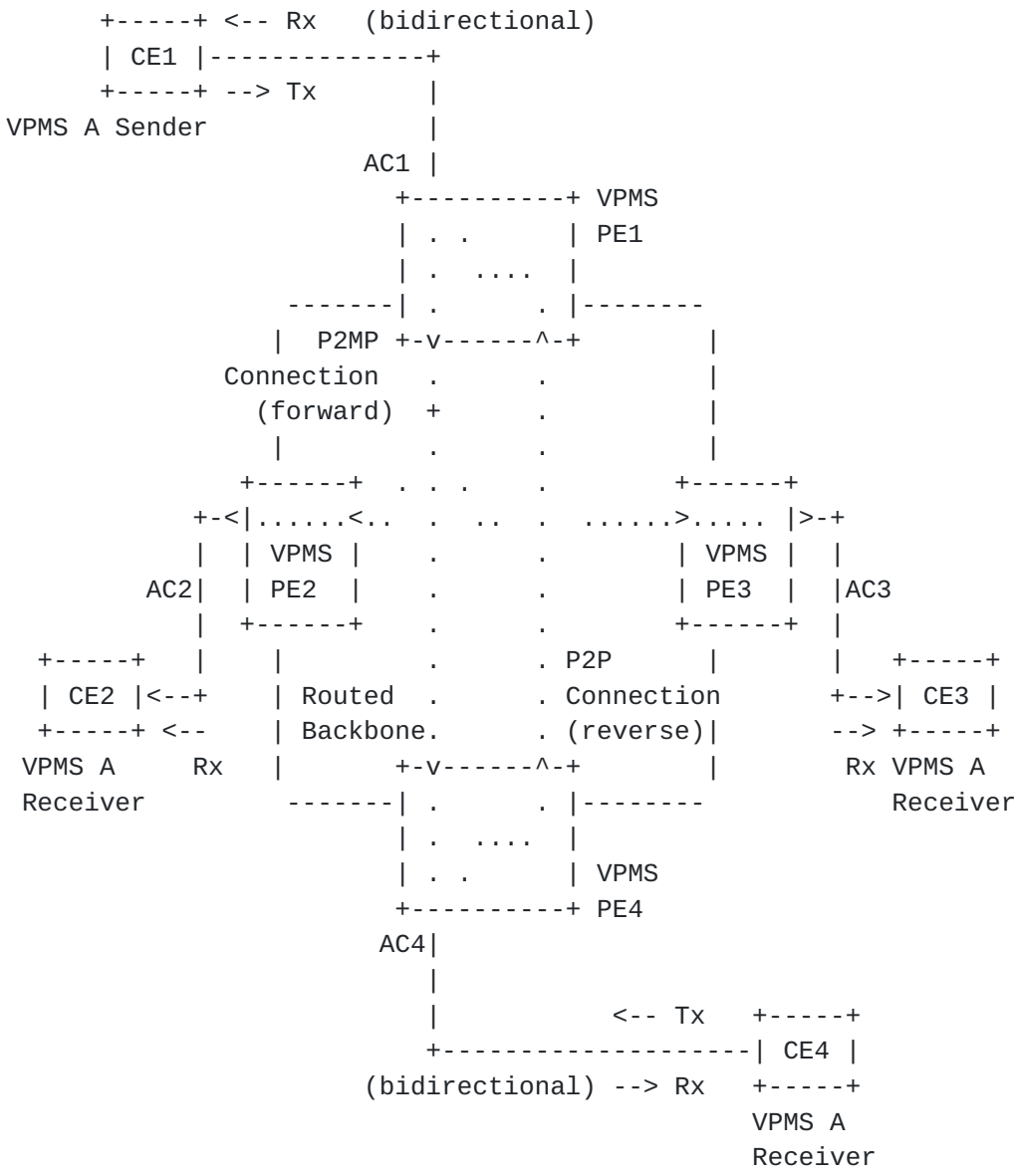


Figure 3: Reverse traffic support

6.2. Transparency

A solution is intended to provide Layer-2 traffic transparency. Transparency SHOULD be supported per VPMS instance basis. In other words, Layer-2 traffic can be transparently transported from a local CE to remote CEs in a given instance. Note, however, if service delimiting fields (VLAN Id in Ethernet, VPI/VCI in ATM, DLCI in FR etc.) are assigned by the SP, the Layer-2 traffic is not necessarily transparent. It will depend on the SP's choice if they assign it to each AC. Hence, it could be that some of the leaf CEs are receiving traffic that has different delimiting fields than the traffic for the

other leaf CEs. Hence, it could be that some of receiver CEs are getting traffic with different delimiting fields than the other receiver CEs.

The VPMS solution SHOULD NOT require any special packet processing by the end users (CEs).

6.3. Quality of Service (QoS)

A customer may require that the VPMS service provide guaranteed QoS. In particular, for real time applications which are considered common in point-to-multipoint delivery, delay and loss sensitive traffic MUST be supported. The solution SHOULD provide native QoS techniques for service class differentiation, such as IEEE 802.1p CoS for Ethernet.

For bandwidth committed services (e.g., ATM CBR), a solution SHOULD guarantee end-to-end bandwidth. It MAY provide flow admission control mechanisms to achieve that.

6.4. Multi-Source Support

A VPMS solution SHOULD support multiple sources in each VPN. That is, two or more sender CEs can exist in the same VPMS instance. Each sender CE SHOULD be able to be connected to physically separate root PEs, which will facilitate flexible topology design.

Additionally, traffic from sender CEs MAY have a common single coverage to receiver CEs, i.e., data from any sources can reach the same group of receivers. For example, in TV provisioning scenario, customers may need to receive two or more TV channels from different sources. In figure 3, suppose there are 7 hosts, CE1 to CE7 which belong to a given VPMS instance A. CE1, CE2 and CE3 are common sources of this VPN, and are attached to different PEs respectively. Traffic from each source can reach the same group of receivers, CE4 to CE7.

This kind of scenario will require multiple P2MP connections in a single VPMS instance (i.e., VPN). Each P2MP connection's root might be a completely different CE/AC, but leaf CE/ACs are overlapped. Since this is basically the same requirement as dual-homed scenario, see section 6.5.1 for details.

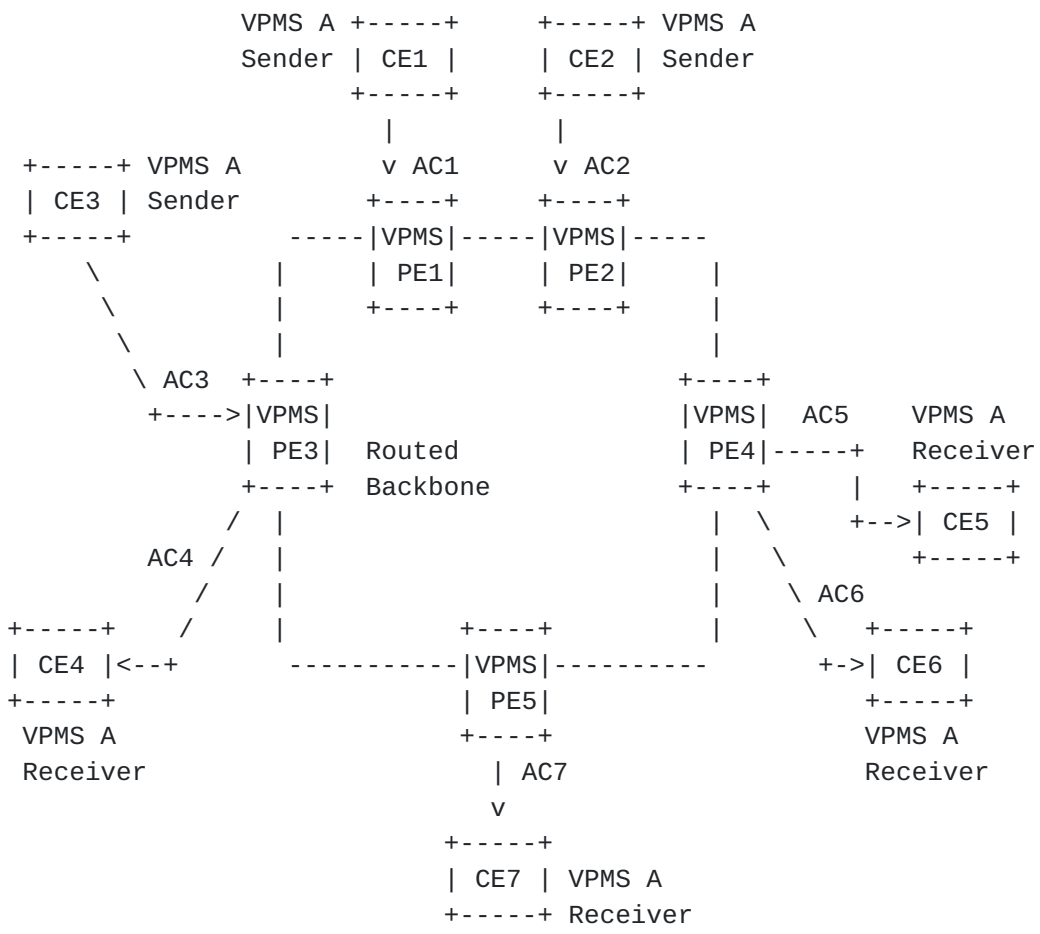


Figure 3: Multi-Source support

6.5. High Availability

A solution MUST provide protection and restoration mechanism for end-to-end services to ensure high availability.

There are multiple parts of the connection that can support protection and restoration: (1) CE to PE, (2) between PEs (3) inside core (PSN backbone). It is expected that (3) is fulfilled by existing PSN protection mechanisms (e.g., RSVP-TE FRR). Following subsections covers the requirements for redundancy support for (1) and (2).

6.5.1. Dual-homed Access Support

A solution MUST allow dual-homed redundant access from a CE to multiple PEs. This if beneficial to support reliable data delivery for customers. Figure 4 provides an example of this access topology.

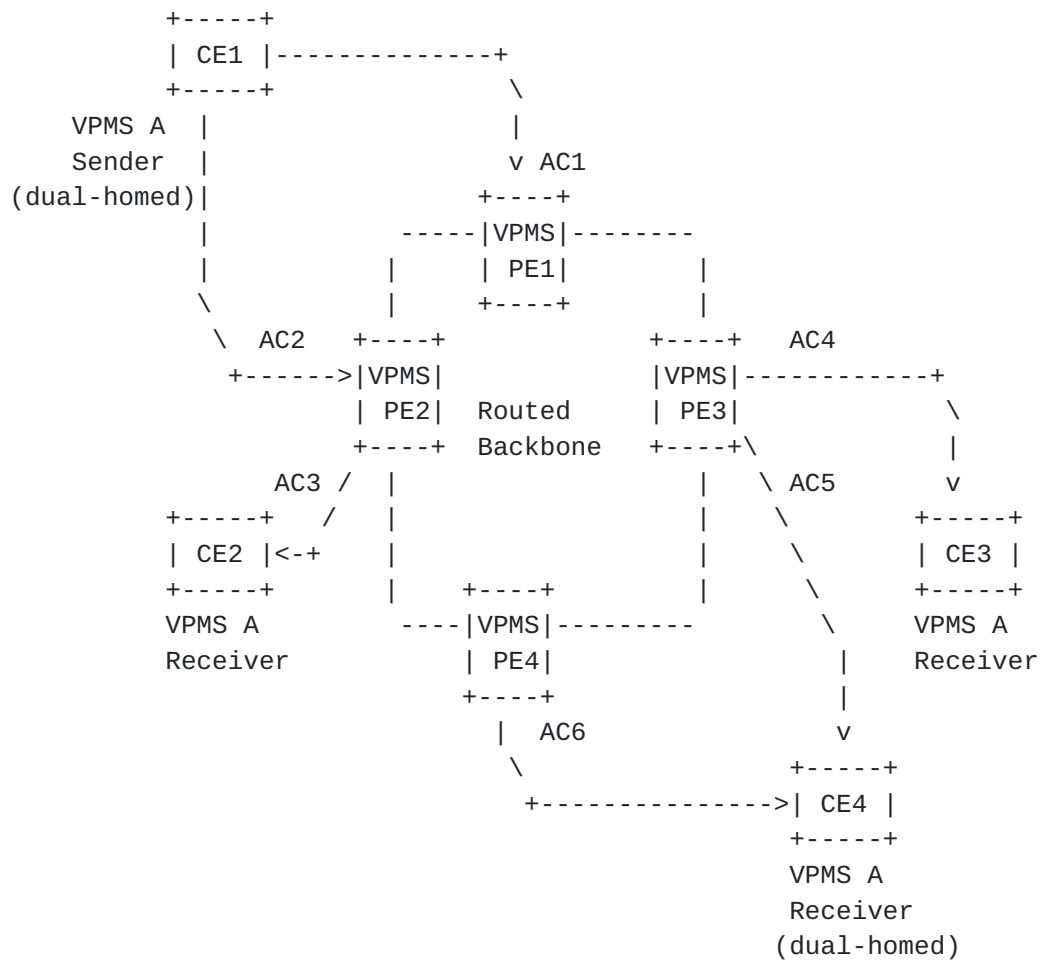


Figure 4: Dual-homing support

A solution SHOULD provide a protection mechanism between the redundant PEs to which a CE is dual-homed. This is because when an ingress root PE node fails whole traffic delivery will fail unless a backup root PE is provided, even in case of dual-homed access. Similarly, if an egress leaf PE node fails, traffic toward that CE is never received unless a backup leaf PE is provided.

In some cases, the data source is required to be highly reliable since it is often deployed as a centralized server that provides traffic to many receivers. Therefore, there is an additional requirement specifically about redundancy of root-side: each VPMS instance SHOULD be able to have multiple P2MP connections whose roots are located at separate root ACs. Those root ACs can be located at physically separate root PEs, whereas those trees will share common leaf ACs. This means that each P2MP connection has a single root AC, but several P2MP connections can be managed together inside a common VPN.

For example, in Figure 5, traffic from root AC1 and AC2 both reach receivers CE3 and CE4 while AC1, AC2, AC3 and AC4 all are associated with a single VPMS instance. This topology is reliable since there are redundant root PE/ACs. At the egress side, PE3 and PE4 select traffic from either root, PE1 or PE2. In this figure, each leaf PE has one leaf AC only (AC3 attached to PE3, and AC4 attached to PE4). Therefore, PEs will need to support PW protection and restoration mechanism so that two redundant P2MP connections are given among common ACs.

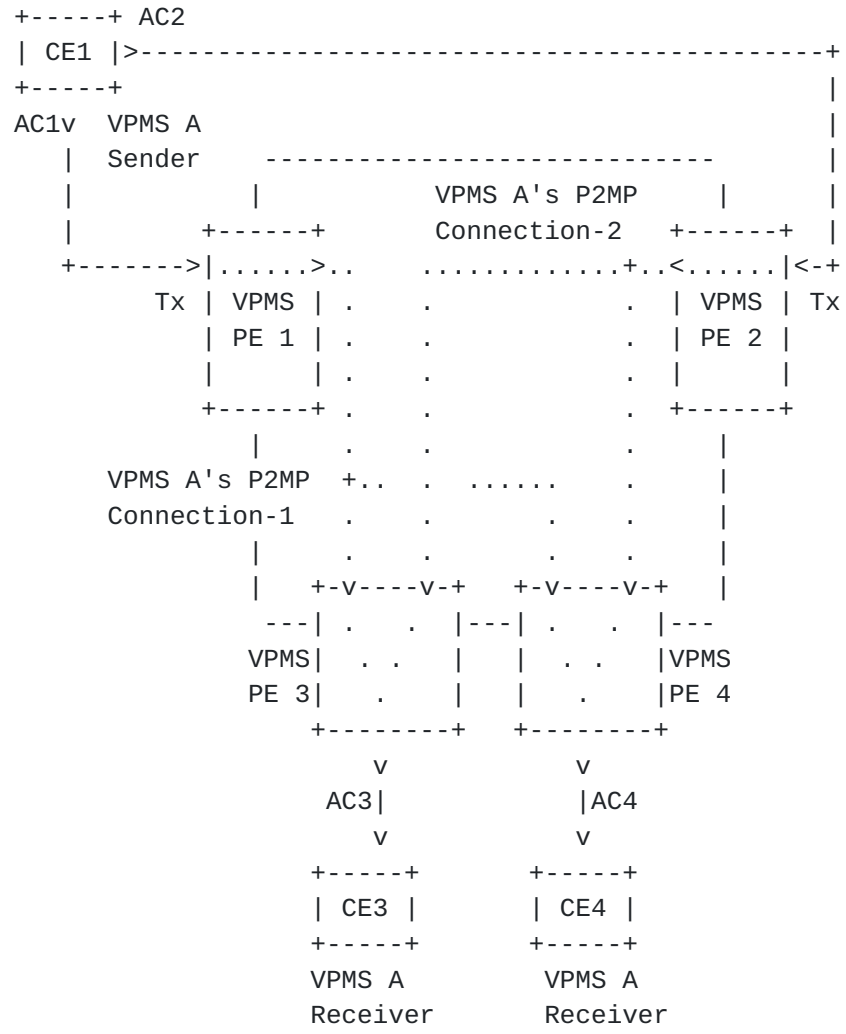


Figure 5: Multiple P2MP connections in Dual-homed Sender

6.5.2. Single/Dual Traffic Support in Dual-homed Access

When dual-homed access to root PEs is provided, a solution MAY allow a sender CE to transmit just a single copy of the traffic to either one of the two root (ingress) PEs or to transmit a copy of the traffic to both the PEs simultaneously. The latter scenario consumes more resource of CE-PE link than the single traffic scenario, but it is usually applicable when a source device has only a simple forwarding capability without any switchover functionality. In such a dual traffic case, the backup root (ingress) PE SHOULD be able to filter the incoming unnecessary traffic while the other root PE is active if it is needed by SP. In either case, single traffic or dual traffic, the switchover mechanism between root (ingress) PEs will be necessary to handle traffic appropriately in case of failure.

In the case of dual-homed access to leaf PEs, a solution MAY allow a receiver CE to receive a single copy of the traffic from either one of the two leaf (egress) PEs, or receive a copy of the traffic from both PEs simultaneously. The dual traffic approach is applicable if CE has fast switchover capability as a receiver by selecting either one of incoming traffic, but note that additional traffic resources are always consumed at PE-CE link of backup side. Specifically in the single traffic case, it might be needed to support switchover mechanism between egress PEs in case of failure.

6.6. Security

The basic security requirements from the view of customers are raised in section 6.5 of [RFC4665]. It also applies to VPMS.

In addition, a VPMS solution MAY have the mechanisms to activate the appropriate filtering capabilities (for example, MAC/VLAN filtering etc.), and it MAY be added with the control mechanism between particular sender/receiver sites inside a VPMS instance. For example, in Figure 1, filtering can be added such that traffic from CE1 to CE4/CE5 is allowed but traffic from CE1 to CE6 is filtered.

6.7. Reordering Prevention

A solution SHOULD prevent Layer-2 frame reordering when delivering customer traffic under normal conditions.

6.8. Failure reporting

A solution MAY provide information to the customer about failures. For example, if there is a loss of connectivity toward some of the receiver CEs, it is reported to the sender CE.

7. Service Provider Network Requirements

7.1. Scalability

A VPMS solution MUST be designed to scale well with an increase in the number of any of the following metrics:

- the number of PEs (per VPMS instance and total in a SP network)
- the number of VPMS instances (per PE and total)
- the number of root ACs / sender CEs (per PE, VPMS instance and total)
- the number of leaf ACs / receiver CEs (per PE, VPMS instance and total)

A VPMS solution SHALL document its scalability characteristics in quantitative terms. A solution SHOULD quantify the amount of state that a PE and a P device has to support.

The scalability characteristics SHOULD include:

- the processing resources required by the control plane in managing PWs (neighborhood or session maintenance messages, keepalives, timers, etc.)
- the processing resources required by the control plane in managing PSN tunnels
- the memory resources needed for the control plane
- other particular elements inherent to each solution that impact scalability

7.2. Pseudo Wire Signaling and PSN Tunneling

A VPMS solution SHOULD provide an efficient replication that can contribute to optimizing the bandwidth usage required in a SP's network. For supporting efficient replication, it is expected to take advantage of PW and PSN mechanisms that are capable of P2MP traffic.

Regarding PW mechanism, [I-D.ietf-pwe3-p2mp-pw-requirements] introduces P2MP PW concept and its requirements, showing two basic approaches of providing replication. One is SS (Single Segment)-PW model that provides replication by PSN tunnel such as P2MP LSP (i.e., by outer label layer), and the other is MS (Multi Segment)-PW model that provides replication by multiple interconnected PWs (i.e., by inner label layer). In either case, end-to-end P2MP topology (i.e., P2MP connection) in VPMS is common from the view of ACs. Requirements as a provider service specified in this document will be commonly applied regardless of P2MP PW's signaling model.

It is out of scope of this document how to extend and use PW mechanisms to realize P2MP connections. For example, it is under the scope of the solution work how to support forward/reverse traffic e.g., by a single PW signaling, coupling multiple PWs, or other ways.

This document does not raise any specific requirements for particular PSN tunneling schemes (point-to-point, point-to-multipoint and multipoint-to-multipoint) that are applied to VPMS. The actual type of PSN tunnel used in VPMS will be dependent on individual deployment scenarios (e.g., which PSN protocol is available in the core and how much of the network resources that operators will want to optimize).

7.3. Auto-discovery

A solution SHOULD support auto-discovery methods that dynamically allow VPMS related information to be discovered by the PEs to minimize the amount of configuration the SP must perform.

All of the requirements on discovery described in Section 7.3 of [RFC4665] SHOULD be satisfied in VPMS as well.

Auto-discovery will help operators' initial configuration of adding a new VPN (i.e., VPMS instance), adding/deleting new sender/receiver, and so on.

The candidate information treated in auto-discovery will be as follows:

- Information to indentify the location of each PE, e.g., PE router ID / IP address
- Information to identify the VPMS instance, that is, to identify a VPN
- Information to identify the type of ACs (root AC or leaf AC)
- Information to identify the P2MP connection that binds ACs
- Information to show if reverse traffic support is optionally desired
- SP-related information (AS number, etc. for an inter-provider case)

Following is an example scenario about adding a new leaf PE: suppose there are three PEs in an existing VPMS, PE1, PE2, PE3. PE1 is a root PE and has a AC1. PE2 and PE3 are leaves and have AC2 and AC3. every PE has the association among the information described above. Now a new PE4 having an AC4 is provisioned in the existing VPMS instance and this AC is configured as leaf. This information will be automatically discovered by the other existing remote PEs (i.e., ingress and egress PEs in the same VPMS instance). Once the ingress PE1 discovers this new PE/AC, it can automatically add AC4 as the new

leaf of P2MP connection topology according to P2MP PW signaling mechanism. The ingress PE1 will graft a new leaf (PE4) to the already existing P2MP connection which is now created from AC1 to AC2/AC3/AC4. This operation does not require any new configuration at the existing PEs.

Another example is about adding a new root PE: suppose there are one root PE (PE1/AC1) and three leaf PEs (PE2/AC2, PE3/AC3 and PE4/AC4). There is an existing P2MP connection from AC1 to AC2/AC3/AC4. Now the operator adds a new root PE/AC (PE5/AC5) for some reasons (e.g., multiple source sites, dual-homed access, root PE redundancy etc.). Then, auto-discovery mechanism advertises this information to all other members PE1/PE2/PE3/PE4, and a new P2MP connection from AC5 to AC2/AC3/AC4 is created by PW signaling.

Note that VPMS instance is created when one root AC and at least one leaf AC are added. In principle VPMS requires such minimum provisioning. Hence in dual-homing case of sender, only backup root PE can be dynamically added/deleted to/from VPMS without destroying the VPN.

7.4. Activation and Deactivation

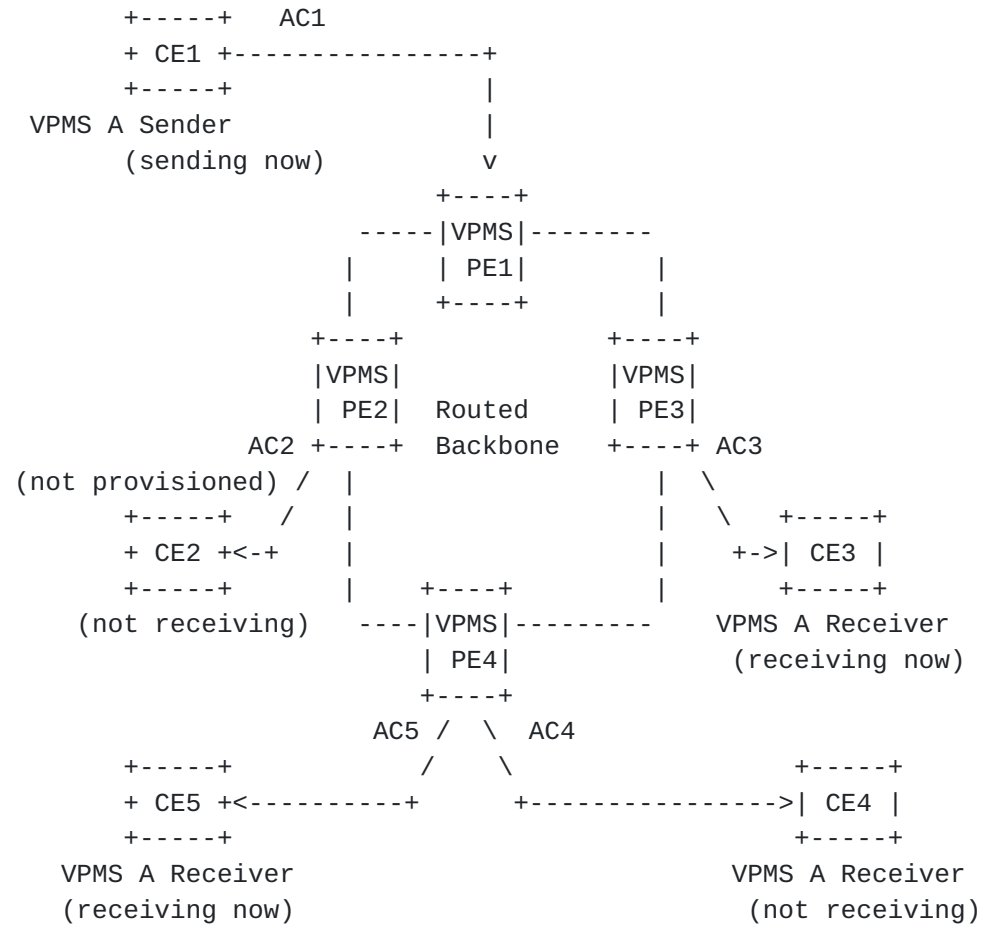
A solution SHOULD provide a way to activate/deactivate the administrative status of each AC. After initial provisioning, an SP might change connectivity configuration between particular CEs inside a single VPMS instance for operational reasons. This feature will be beneficial to help such a scenario.

For example, in Figure 6, AC1, AC3, AC4 and AC5 are initially provisioned for VPMS A. AC2 is not provisioned for any VPMSes. In VPMS A, CE1 is a sender and CE3, CE4 and CE5 are receivers. Traffic will usually flow from CE1 to all receivers, CE3, CE4 and CE5. However, for maintenance operation, application's request (e.g., stream program has changed) or some other reasons, AC4 needs to be set as administratively deactivated. Then it becomes necessary to turn off traffic from PE4 to CE4. This operation must be appropriately distinguished from failure cases.

When deactivating a particular site, backbone PSN/PW resources (e.g., admission control of PSN tunnel) MAY be released for that particular direction in order to provide that bandwidth to other services. In Figure 6, AC3 is now administratively activated and receiving traffic. However, if AC3 comes to be administratively deactivated, and if RSVP-TE (including P2P and/or P2MP) is used for backbone PSN, then TE reserved resources from PE1 to PE3 may be released.

In addition, a solution SHOULD allow single-sided activation

operation at a root (ingress) PE. In some scenarios, operators prefer centralized operation. This is often considered natural for one-way digital audio/video distribution applications: SPs often want to complete their service delivery by a single operation at one source PE, not by multiple operations at many leaf (egress) PEs. Figure 6 illustrates this scenario, where a SP only has to do single-sided operation at PE1 (source) to administratively activate/ deactivate various connections from AC1 to AC3, AC4 and/or AC5. It is not needed to perform operations on PE3 and PE4 directly.



AC1: Administratively activated
AC2: No VPMS provisioned
AC3: Administratively activated
AC4: Administratively deactivated
AC5: Administratively activated

Figure 6: Site activation and deactivation

7.5. Inter-AS Support

A solution SHOULD support inter-AS scenarios, where there is more than one provider providing a common VPMS instance and VPN. More specifically, it is necessary to consider the case where some of the PEs that compose one VPMS belong to several different ASes.

7.6. Co-existence with Existing L2VPNs

A solution MUST co-exist with the existing L2VPNs (e.g., VPWS, VPLS) across the same SP's network. A solution MUST NOT impede the operation of auto-discovery and signaling mechanism that are already supported by the PEs for those existing L2VPNs.

7.7. Operation, Administration and Maintenance

7.7.1. Fault Management

7.7.1.1. Fault Detection

A solution MUST provide tools that detect reachability failure and traffic looping of data transport in a VPMS instance. If multiple root ACs are supported (i.e., multiple P2MP connections are grouped together into a single VPMS instance), such tools MUST be able to perform distinguishing each P2MP connection.

7.7.1.2. Fault Notification

A solution MUST provide fault notification and trouble tracking mechanisms. (e.g. SNMP-trap and syslog that notify fault to remote NMS.)

In VPMS one point of failure at upstream often affects a number of downstream PEs and ACs that might raise a notification message. Hence notification messages MAY be summarized or compressed for operators' ease of management.

In case of receiver-side failure (leaf PE or its AC), this fault status SHOULD be able to be monitored at root PE. This will help an operator to monitor each leaf PE/AC in a centralized manner; that is, a root PE can collect leaf-side information. How this status is transferred depends on a solution.

In contrast, in case of sender-side failure (root PE or its AC), this fault status SHOULD also be able to be monitored at leaf PEs. This will help an operator to troubleshoot at leaf PEs (i.e., distinguish local AC's failure from remote root AC's failure easily).

In any case of failure at SP's network, fault information MAY be notified to the customer. Specifically, such fault MAY trigger generating customer OAM message toward CEs (e.g., AIS) and/or shutting down leaf ACs.

7.7.1.3. Fault Isolation

A solution MUST provide diagnostic/troubleshooting tools for data transport in a VPMS instance.

7.7.2. Testing

A solution MUST provide a mechanism for testing each data connectivity and verifying the associated information in a VPMS instance. The connectivity is between a root and all leaf ACs (i.e., each P2MP connection can be tested).

Operators will run testing before and after service activation. Testing mechanism SHOULD support end-to-end testing of the data path used by customer's data. End-to-end testing will have CE-to-CE path test and PE-to-PE path test. A solution MUST support PE-to-PE path test and MAY support CE-to-CE path test. In either case the minimum data path unit for each VPMS is unidirectional, hence if loopback testing is supported, additional consideration about reverse-path might also be needed (see section 6.1.2).

If there are multiple P2MP connections for redundancy (active/backup tree) in a common VPMS (like in Figure 4), testing mechanism MUST be able to check the connectivity over not only working P2MP connection but also protecting connection(s). This testing MUST be able to be performed from a root PE. It MAY also be able to be performed from a sender CE.

7.7.3. Performance Management

A solution MUST offer mechanisms to monitor traffic performance parameters and statistics of data traffic in VPMS.

A solution MUST provide access to:

- Traffic statistics (total traffic forwarded, incoming, outgoing, dropped, etc., by period of time)

A solution SHOULD provide access to:

- Performance information related to traffic usage, e.g., one-way delay, one-way jitter, one-way loss, delay variations (the difference of various one-way delay from a particular root PE to multiple leaf PEs) etc.

All or part of this information SHOULD be made available through standardized SNMP MIB Modules (Management Information Base).

It is expected that such information can be used for SLA monitoring between sender and receiver, to give the SP a clear picture of current service providing to the customer.

7.8. Security

Section 7.6. of [RFC4665] describes common Layer-2 VPN security requirements from service provider aspect, which also applies to VPMS. (For example, an SP network MUST be protected against malformed or maliciously constructed customer traffic, etc.)

This subsection adds VPMS-specific consideration and requirements.

In VPMS, all traffic is transported with multicast duplication in terms of end-to-end perspective, regardless of customer's individual protocol. A PE never processes CE's multicast control protocol (e.g., PIM, IGMP, MLD as Layer-3). Hence, in PE and P, basically the security threat from malicious customer's C-plane protocol is small.

In VPMS, there is security threat from malicious customers' D-plane traffic. A PE might receive a high volume of data from a CE. If there is no safeguard on PE, it will cause excessive replication in the SP network. Therefore, a VPMS solution SHOULD support traffic policing to limit the unwanted data traffic. Such a policing mechanism MUST be configurable per VPN basis, not the total of various VPNs to isolate malicious customer's traffic from others.

8. Security Considerations

The security requirements common to customers and service providers are raised in Section 5.5. of [RFC4665], which are fundamental for all Layer-2 VPN services. VPMS is a variant of Layer-2 VPN, and that statement also applies to VPMS.

Moreover, in this document, security requirements from the view of customers are shown in Section 6.5. Security requirements from the view of providers are shown in Section 7.8. They explain security considerations that are specific to VPMS.

9. IANA Considerations

This document has no actions for IANA.

10. Acknowledgments

Many thanks to Ichiro Fukuda, Kazuhiro Fujihara, Ukyo Yamaguchi and Kensuke Shindome for their ideas and feedback in documentation.

The authors gratefully acknowledge the valuable review and comments provided by Greg Mirsky and Yuji Tochio.

11. References

11.1. Normative References

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

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, March 2005.

11.2. Informative References

[I-D.ietf-l2vpn-vpls-mcast] Aggarwal, R., Kamite, Y., and L. Fang, "Multicast in VPLS", draft-ietf-l2vpn-vpls-mcast-09 (work in progress), July 2011.

[I-D.ietf-pwe3-p2mp-pw-requirements] Heron, G., Wang, L., Aggarwal, R., Vigoureux, M., Bocci, M., Jin, L., JOUNAY, F., Niger, P., Kamite, Y., DeLord, S., and L. Martini, "Requirements and Framework for Point-to-Multipoint Pseudowire", draft-ietf-pwe3-p2mp-pw-requirements-04 (work in progress), July 2011.

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

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

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling",

RFC 4761, January 2007.

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

Authors' Addresses

Yuji Kamite
NTT Communications Corporation
Granpark Tower
3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Email: y.kamite@ntt.com

Frederic Jounay
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
France

Email: frederic.jounay@orange-ftgroup.com

Ben Niven-Jenkins
Velocix
326 Cambridge Science Park
Milton Road, Cambridge
CB4 0WG
UK

Email: ben@niven-jenkins.co.uk

Deborah Brungard
AT&T
Rm. D1-3C22, 200 S. Laurel Ave.
Middletown, NJ, 07748
USA

Email: dbrungard@att.com

Lizhong Jin
ZTE Corporation
889, Bibo Road
Shanghai, 201203
China

Email: lizhong.jin@zte.com.cn