Network Working Group
Internet Draft
Category: Standard Track
Expires: December 2014

Simon Delord (editor), Telstra
Raymond Key (editor)
Frederic Jounay, Orange CH
Yuji Kamite, NTT Communications
Zhihua Liu, China Telecom
Manuel Paul, Deutsche Telekom
Ruediger Kunze, Deutsche Telekom
Mach Chen, Huawei
Lizhong Jin

June 28, 2014

# Extension to LDP-VPLS for Ethernet Broadcast and Multicast
## draft-ietf-l2vpn-ldp-vpls-broadcast-exten-08



Abstract

This document proposes a simple extension to LDP-VPLS to improve bandwidth efficiency for Ethernet broadcast/multicast traffic within a carrier's network. It makes use of unidirectional point-to-multipoint PseudoWires to minimise payload frame duplication on physical links.

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].

Table of Contents



## 1. Introduction

This document proposes a simple extension to LDP-VPLS [RFC4762] to improve bandwidth efficiency for Ethernet broadcast/multicast traffic within a carrier's network. This bandwidth improvement is achieved by adding to the existing full-mesh of bidirectional point-to-point PseudoWires (P2P PWs) unidirectional point-to-multipoint PseudoWires (P2MP PWs) between selected PEs within a VPLS instance. With P2MP PWs, the ingress PE is not responsible for replicating the payload frame on each P2P PW towards the egress PE, instead the network elements along the physical path participate in the replication process. The replication is done by the underlying point-to-multi-point label switched path. This proposal allows for a large number of P2MP PWs to be carried through a single MPLS P2MP tunnel, thus, it is never necessary to maintain state in the network core for individual P2MP PWs.

## 2. Problem Statement & Motivation

### 2.1. Requirements for Broadcast/Multicast Support in VPLS

[RFC5501] provides an in-depth discussion on broadcast/multicast related requirements for VPLS. It highlights two specific issues:

- issue A: replication to non-member site.
- issue B: replication of PWs on shared physical path.

#### 2.1.1. Issue A: replication to non-member site

The current standard VPLS is a L2VPN service agnostic to customer's Layer 3 traffic, hence does not maintain any information about IP multicast group membership.  Although a Layer 3 IP multicast packet is encapsulated in a Layer 2 Ethernet multicast frame, the current standard VPLS treats Ethernet multicast frame in exactly the same way as Ethernet broadcast frame. There is therefore an issue that multicast traffic is sent to sites with no members.  Since the upstream PE does not maintain downstream membership information, it simply floods frames to all downstream PEs, and the downstream PEs forward them to directly connected CEs; however, those CEs might not be the members of any multicast group.

There are therefore two elements to Issue A:

- the PE to CE section (e.g. the AC), where a CE will receive unintended traffic.

- the PE to PE section within a VPLS instance, where a PE will receive multicast traffic even when it has no CE being member of any multicast group.

To address the PE to CE part, a PE might have to maintain multicast group information for CEs that are not kept in the existing VPLS solutions.

To address the PE to PE part and limit the flooding scope across the backbone, a PE needs to discover multicast group information from other remote PEs.

Both elements will present scalability concerns about state resources (memory, CPU, etc.) and their maintenance complexity.

Finally, if Layer-3 information is checked for transport,  the following  [RFC4665] requirement "a L2VPN service SHOULD be agnostic to customer's Layer3 traffic" can no longer be met.

### 2.1.2. Issue B: replication of PWs on shared physical path

Issue B on the other hand can still be improved without making use of any Layer-3-related information.

Issue B may still be considered acceptable when:

- Ethernet broadcast/multicast traffic volume is low; and
- The number of replications on each outgoing physical interface for a VPLS instance is small (e.g. not many PEs per VPLS instance).

However, with more broadcast/multicast applications (e.g. broadcast TV), Ethernet broadcast/multicast traffic volume may increase to a significant level. Assuming HDTV requires 10Mbps per channel, a bundle of 100 channels will require 1Gbps.

Furthermore, as MPLS networks expand from the core towards aggregation/access, more PEs may participate in a single VPLS instance. The number of replications on each outgoing physical interface for a VPLS instance is likely to increase.

## 2.2. Motivation

Based on the previous section, it may still be desirable for some carriers to look at improving issue B without having to look at Layer 3 information (Issue A).

One reason for this is that sometimes there is no L3 data to snoop. Another reason may be that some carriers may not be allowed to look above the L2 header, for example there may be regulatory issues with inspecting the customer payload. Also, some carriers may not want to do L3 snooping as Operations will naturally become more complicated if the number of managed objects (e.g. multicast groups) increases.

Another important point is that some carriers may want a manual and granular optimisation process that allows optimizations to certain services or areas but does not impact the rest of the network. For example, the bandwidth improvement process may only be required at specific locations in the network where bandwidth-intensive multicast broadcast Ethernet flows exist. It would also be beneficial if the optimisation process were incrementally deployable, so that the optimisation can still be leveraged even if there are portions of the network that are not able to support the features required by the optimisation process. A potential case would be a VPLS instance composed of both PEs supporting the proposed protocol extension and PEs not supporting it, the enhancement is then achieved between the compliant PEs only.

Finally, some carriers may also prefer a deterministic process to an entirely automated path selection algorithm that is network driven. [RFC5501] gives several reasons on why this may be the case:

- Accounting for various operator policies where the logical multicast topology within a carrier's network does not change dynamically in conjunction with a customer's multicast routing.

- Operations will naturally become more complicated if topology changes occur more frequently.

- Troubleshooting will tend to be difficult if a solution supports frequent dynamic membership changes with optimized transport within the carrier's network.

[RFC7117] is a solution that looks at solving both Issue A and Issue B. However, [RFC7117] proposes that, even for carriers who currently use [RFC4762] without auto-discovery mechanisms, BGP be introduced. This may also present operational challenges and complexities for some carriers, or this feature may simply not be supported on some of the network elements deployed.

### 2.3. Scope of the proposed solution

This draft therefore explores whether there is a way to improve Layer 2 Ethernet broadcast/multicast bandwidth simply and predictably with:

- Minimal extension to [RFC4762] and without the need to add BGP (e.g. no auto-discovery)
- Minimal impact to existing [RFC4762] deployed networks
- Operator driven optimisation (i.e. the operator decides where and how the bandwidth improvement should occur) to minimise the number of states and the potential operational complexities associated with dynamic changes within a carrier's core network.

## 3. Terminology

This document uses terminology described in [RFC4762] and [P2MP-PW-REQ].

## 4. Relevant IETF technologies for the proposed solution

The proposed solution relies on [RFC4762] existing mechanisms and complements them with extensions (P2MP LSPs and P2MP PWs) already standardised ([RFC4875]) or currently under development by the IETF ([RFC6388] and [P2MP-PW-LDP]).

### 4.1. P2MP LSPs

Similarly to what is defined in [RFC4762] where P2P PWs are multiplexed onto P2P LSPs, before the operator can start deploying P2MP PWs, an appropriate underlying layer made of P2MP LPSs needs to be configured (section 3.2 of [P2MP-PW-REQ]).

P2MP LSPs are used to minimise packet replication on specific physical links and to allow P routers in an MPLS domain to be transparent to services (e.g. a P Router will join the P2MP PSN tunnel operation but will have no knowledge of the P2MP PWs, same as [RFC4762]).

The mapping of the P2MP LSP over the physical topology is a key component of the bandwidth enhancement exercise and the operator needs to carefully consider where and how these P2MP LSPs should be deployed (see Appendix A for an example of a possible deployment).

Once configured, it is then possible to aggregate P2MP PWs over a particular P2MP LSP (similar to [RFC4762]).

### 4.2. P2MP PWs

P2MP PWs can be configured statically (e.g. by the operator) or via LDP on top of the P2MP LSPs. This configuration is done on a per PE per VPLS instance basis.

In a P2MP PW, the operator decides to connect one Root PE to at least two Leaf PEs (section 3.1 of [P2MP-PW-REQ]).

The Root PE is the headend of the P2MP PW (where a big Ethernet multicast/broadcast talker is connected - see example in Appendix A).

The Leaf PEs are the endpoints of the P2MP PW (they constitute the receivers where the broadcast/multicast traffic needs to be distributed to).

A Root PE may map more than one P2MP PW to a specific VPLS instance. In this case, the Root PE MUST NOT associate a leaf PE to more than one P2MP PW for a specific VPLS instance (this is to avoid a Leaf PE to receive duplicate copies of the same Ethernet frame from different P2MP PWs).

P2MP PWs are defined in [P2MP-PW-REQ] and one solution using LDP as the signalling mechanism between PEs is defined in [P2MP-PW-LDP].

## 5. Proposed extension to [RFC4762]

This section updates [RFC4762] by describing the extra rules to be applied within a VPLS when unidirectional P2MP PWs are added to the existing full-mesh of P2P PWs.

### 5.1. VPLS Reference Model

Figure 1 shows a topological model (not the physical topology) of a VPLS between four PEs with an arbitrary set of ACs attached to each VSI.

```
                    +---------+            +---------+
                    |   PE1   |            |   PE2   |
                    |  +---+  |            |  +---+  |
                    |  |   |  |            |  |   +-------AC4---
                    |  | V |  |            |  | V |  |
                    |  |   |  |            |  |   |  |
         ----AC1----+--+   |  |  Ethernet  |  |   +--+----AC5---
                    |  | S +--+-----PW-----+--+ S |  |
                    |  |   |  |            |  |   |  |
         ----AC2----+--+   |  |            |  |   |  |
                    |  | I |  |            |  | I |  |
                    |  |   |  | Ethernet   |  |   |  |
         ----AC3----+--+   |  |   PW +-----+--+   |  |
                    |  +-+-+  |     /      |  +-+-+  |
                    |    |  \ |    /       |    |    |
                    +----+---\+   /        +----+----+
                         |    \  /              |
                 Ethernet|     \/               |Ethernet
                       PW|     /\               |PW
                         |    /  \Ethernet      |
                    +----+---/+   \PW      +----+----+
                    |PE3 |  / |    \       |PE4 |    |
                    |  +-+-+  |     \      |  +-+-+  |
         ----AC6----+--+   |  |      \     |  |   +--+----AC8----
                    |  | V |  |       +----+--+ V |  |
                    |  |   |  |            |  |   |  |
         ----AC7----+--+   |  |  Ethernet  |  |   +--+----AC9----
                    |  | S +--+-----PW-----+--+ S |  |
                    |  |   |  |            |  |   |  |
                    |  |   |  |            |  |   |  |
                    |  | I |  |            |  | I |  |
                    |  |   |  |            |  |   |  |
                    |  |   |  |            |  |   |  |
                    |  +---+  |            |  +---+  |
                    |         |            |         |
                    +---------+            +---------+
```

Figure 1: Reference Diagram for VPLS

Figure 2 shows the proposed extensions to VPLS for Ethernet broadcast and multicast.  On top of the topology presented in Figure 2, two P2MP PWs have been added to the existing set of P2P PWs.

```
           +---------+            +---------+
           |   PE1   |            |   PE2   |
           |  +---+  |            |  +---+  |
           |  |   |  |            |  |   +-------AC4---
           |  | V |  |  Ethernet  |  | V |  |
           |  |   |  |    P2MP    |  |   |  |
----AC1----+--+   |  |     PW1    |  |   +--+----AC5---
           |  | S +--+->--+-->----+--+ S |  |
           |  |   |  |    |       |  |   |  |
----AC2----+--+   |  |    |       |  |   |  |
           |  | I |  |    |   +->-+--+ I |  |
           |  |   |  |    |   |   |  |   |  |
----AC3----+--+   |  |    |   |   |  |   |  |
           |  +---+  |    |   |   |  +---+  |
           |         |    |   |   |         |
           +---------+    |   |   +---------+
                          |   |
                          |   |
                          |   |
                          |   |
           +---------+    |   |   +---------+
           |PE3      |    |   |   |PE4      |
           |  +---+  |    |   |   |  +---+  |
----AC6----+--+   |  |    +->-----+--+   +--+----AC8----
           |  | V |  |        |   |  | V |  |
           |  |   |  |        |   |  |   |  |
----AC7----+--+   |  |        |   |  |   +--+----AC9----
           |  | S +--+->------+->-+--+ S |  |
           |  |   |  |   Ethernet |  |   |  |
           |  |   |  |     P2MP   |  |   |  |
           |  | I |  |     PW2    |  | I |  |
           |  |   |  |            |  |   |  |
           |  |   |  |            |  |   |  |
           |  +---+  |            |  +---+  |
           |         |            |         |
           +---------+            +---------+
```

Figure 2: Proposed Extensions to VPLS for Ethernet broadcast/multicast

P2MP PW1 is composed of PE1 as the root PE and PE2 and PE4 as leaf PEs.

P2MP PW2 is composed of PE3 as the root PE and PE2 and PE4 as leaf

PEs.

Note that for sake of clarity, Figure 2 does not show the full-mesh of P2P PWs presented in Figure 1.

Also note that the solution does not require that P2MP PWs be used on all PEs in the VPLS. For example, PE3 is not leaf of P2MP PW1 and PE1 is not leaf of P2MP PW2 and therefore, there is only a P2P PW between PE1 and PE3, and a P2P PW between PE2 and PE4.

## 5.2. Choosing PEs for a specific VPLS to be connected by a P2MP PW

This updates section 4.3 of [RFC4762].

VPLS is a full-mesh of P2P PWs and optionally a number of unidirectional P2MP PWs. At the difference of P2P PWs, not all PEs in a VPLS instance need to be connected via P2MP PWs.

For each P2MP PW on this VPLS instance:

- The operator selects one PE as the Root of the P2MP PW.

- The operator also selects two or more PEs belonging to the same VPLS instance to be Leafs of the P2MP PW

- Because there is already a full-mesh of bidirectional P2P PWs between all PEs, the P2MP PW is unidirectional only (e.g. from the Root PE to all the Leaf PEs connected to it).

- The operator also needs to make sure that there is an active P2MP LSP setup between the Root PE and the Leaf PEs:

- If there is already an active P2MP LSP setup between the Root PE and the Leaf PEs, then procedures described in 5.3 can be followed.

- If there is no P2MP LSP between the Root PE and the Leaf PEs then the operator needs to create first a P2MP LSP in order for procedures in 5.3 to be followed. Procedures to setup a P2MP LSP will vary based on the technology used and are described in [RFC6388] and [RFC4875].

## 5.3. Create and associate the P2MP PW to a specific VPLS Instance

This updates section 4.3 of [RFC4762].

Once that the endpoints of the P2MP PW have been selected and that there is an active P2MP LSP between them, the operator can then create and associate the P2MP PW to a specific VPLS instance. This activity can be done statically or via LDP [P2MP-PW-LDP].

Because P2MP PWs are used to demultiplex encapsulated Ethernet frames from multiple VPLS instances that are aggregated over the same P2MP transport LSP, it is necessary that a Leaf PE can associate unambiguously a P2MP PW aggregated within a P2MP LSP to both a specific VPLS instance and a Root PE.

In the static case, the operator is responsible for configuring all the required information on all PEs belonging to the P2MP PW.

In the LDP case, the P2MP PW is initiated by the Root PE by sending a P2MP PW LDP Label Mapping Message to each of the Leaf PEs.

This label mapping contains, the VPLS instance the P2MP PW is associated to, the P2MP LSP used to transport the P2MP PW and the P2MP PW MPLS Label.

The P2MP PW MPLS Label is upstream assigned and allocated according to the rules in [RFC5331].

The root PE imposes the upstream-assigned label on the outbound packets sent over the P2MP-PW and using this label a Leaf PE can identify the inbound packets arriving over the P2MP PW.

Detailed LDP message formats and P2MP PW setup procedures are described in [P2MP-PW-LDP].

### 5.4. Mapping more than one P2MP PW to a specific VPLS Instance on a specific Root PE

The proposed solution allows for a Root PE to map more than one P2MP PW to a specific VPLS instance (see example in Appendix A).

However in this case, the Root PE MUST NOT associate a leaf PE to more than one P2MP PW for a specific VPLS instance (this is to avoid a Leaf PE to receive duplicate copies of the same Ethernet frame from different P2MP PWs).

### 5.5. Flooding and Forwarding

This section updates section 4.1. of [RFC4762].

A root PE MUST NOT flood frames simultaneously over P2MP PW and P2P PW toward the same leaf PE.

For the flooding of an Ethernet broadcast/multicast frame over PWs to remote PEs participating in the VPLS:

- If there is P2MP PW towards a remote PE, the P2P PW associated with this remote PE will not be used. One copy of the frame will be forwarded on the P2MP PW for all the remote PEs associated with it.
- If there is no P2MP PW towards a remote PE, the P2P PW associated with this remote PE is used.

It should be noted that local policy on the Root PE at the operator's operational request can override any decision to flood and forward traffic over a P2MP PW for a VPLS instance. In that case, normal flooding procedures over P2P PWs described in 4.1 of [RFC4762] apply.

### 5.5.1. Flooding and Forwarding for Ethernet unknown unicast

In traditional Ethernet switched networks unknown unicast frames are handled the same way as broadcast and multicast Ethernet traffic (e.g. flooding). Similarly, current VPLS standards also handle unknown unicast traffic by flooding it across all P2P PWs.

The main purpose of this document is to address Ethernet broadcast and multicast traffic. For Ethernet unknown unicast frames there are two possibilities:

(1) forward the unknown unicast traffic on the P2MP PW, same as for Ethernet broadcast and multicast.
(2) keep the existing mechanism of [RFC4762] and flood over the mesh of P2P PWs.

A Root PE SHOULD support both the above as configurable options per VPLS instance. For a particular VPLS instance where frame reordering is considered unacceptable, option (2) is recommended.

### 5.6. Address Learning

This section updates section 4.2. of [RFC4762].

A Leaf PE MUST support the ability to perform MAC address learning for packets received on a P2MP PW.

When a Leaf PE receives an Ethernet frame on a P2MP PW it:
- First determines the VSI associated to the P2MP PW
- Then determines the Root PE of the P2MP PW
- Then determines the P2P PW associated with that Root PE
- Finally, creates a forwarding state in the VPLS instance for the P2P PW associated with the Root PE with a destination MAC address being the same as the source MAC address being learned.

### 5.7. Loop Free Topology

This updates section 4.4. of [RFC4762]

Paragraph 2 "must not forward from one PW to another" is applicable to P2MP PW & P2P PW.

### 5.8. Hierarchical VPLS

This refers to section 10 and 11 in [RFC4762]. No change is required.

### 5.9 P2MP PW Status

In case of a P2MP PW status change to not operational as per [P2MP-PW-LDP], then this should be treated as if this P2MP PW does not exist.

## 6. Local PE Implementation

This section is OPTIONAL.

As described in section 2.1.1, a PE receiving an IP multicast frame, will forward it to all ACs, including those with no member of the specific IP multicast group attached.

Unnecessary traffic consumes bandwidth on the access link and may become a concern from the customer perspective. In some cases, it may also be a security concern as the multicast frame may be forwarded to an endpoint other than the intended destinations.

Consequently, the use of some L3 related supplementary information in order to improve bandwidth consumption on the AC may be considered. Enabling L3 snooping on an AC basis only has an impact on the PE where the AC belongs, it does not impact the number of P2MP PW/LSPs used within the carrier's network and the state resources or the maintenance complexity associated with it.

Alternatives to L3 snooping such as static configuration of multicast Ethernet addresses & ports / interfaces for example are also possible.

## 7. Security Considerations

This document does not introduce any new security issues beyond those already described in [RFC4762].

## 8. IANA Considerations

This document does not require IANA assignment.

## 9. Acknowledgments

The authors would like to thank Marc Lasserre, Cesar Garrido, Lucy Yong, Neil Harrison, Samer Salam, Nocolai Leymann, Florin Balus, Olen Stokes, Ali Sajassi, Phil Bedard, Thomas Beckhaus, Linda Dunbar, Max Ng, Asaf Henig, Mark Whalley and Cary Wright for their valuable comments.

## 10. References

### 10.1. Normative References

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

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

[P2MP-PW-LDP] Sivabalan, et. al, "Signaling Root-Initiated Point-to-Multipoint Pseudowires using LDP", draft-ietf-pwe3-p2mp-pw-04, work in Progress, March 2012.

[RFC6388] Wijnands, et. al, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, November 2011.

[RFC4875] R. Aggarwal, Ed., D. Papadimitriou, Ed., S. Yasukawa, Ed., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

[RFC5331] R. Aggarwal, Y. Rekhter, E. Rosen, "MPLS Upstream Label Assignment and Context Specific Label Space", RFC 5331, August 2008.

### 10.2. Informative References

[RFC5501] Kamite, et al., "Requirements for Multicast Support in Virtual Private LAN Services", RFC 5501, March 2009.

[P2MP-PW-REQ] Jounay, et. al, "Requirements and Framework for Point-to-Multipoint Pseudowires over MPLS PSNs", draft-ietf-pwe3-p2mp-pw-requirements-10, work in progress, June 2014.

[RFC7117] Raggarwa, Kamite & Fang, "Multicast in VPLS", RFC 7117, February 2014.

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

## Author's Addresses

Simon Delord (editor)
Telstra
EMail: simon.delord@gmail.com

Raymond Key (editor)
EMail: raymond.key@ieee.org

Frederic Jounay
Orange CH
4 rue caudray 1020 Renens
Switzerland
EMail: frederic.jounay@orange.ch

Yuji Kamite
NTT Communications Corporation
Granpark Tower
3-4-1 Shibaura, Minato-ku
Tokyo  108-8118
Japan
EMail: y.kamite@ntt.com

Zhihua Liu
China Telecom
109 Zhongshan Ave.
Guangzhou 510630
China
EMail: zhliu@gsta.com

Manuel Paul
Deutsche Telekom
Winterfeldtstr. 21-27
10781 Berlin, Germany
EMail: manuel.paul@telekom.de

Ruediger Kunze
Deutsche Telekom
Winterfeldtstr. 21-27
10781 Berlin, Germany
EMail: ruediger.kunze@telekom.de

Mach(Guoyi) Chen
Huawei Technology Co., Ltd.
Q14 Huawei Campus, No. 156 Beiqing Road
Hai-dian District, Beijing 100095
China
EMail: mach.chen@huawei.com

Lizhong Jin
EMail: lizho.jin@gmail.com

## Appendix A. One example for broadcast video delivery

This section describes one deployment scenario in relation to broadcast video delivery and how the proposed solution would work.

One requirement of the model is that the application needs unicast data exchange (IP unicast transfer or control messages etc.) as a background environment. MAC-learning (and therefore VPLS) is effective to support it.

### A.1. Broadcast Video Delivery Topology

Figure 3 presents the physical topology of one broadcast video deployment.

```
                              |     |     |
                              AC    AC    AC
                              |     |     |
                            +---+ +---+ +---+
                            |PE3| |PE4| |PE5|
                            +---+ +---+ +---+
                               \    |    /
                                \   |   /                 +----+
                                 \  |  /              +---|PE9 |--AC--
                                  \ | /              /    +----+
   Ethernet        +---+          +---+        +---+/     +----+
   Broadcast->-AC--|PE1|----------|P1 |--------|P3 |------|PE10|--AC--
   Source          +---+\        /+---+        +---+\     +----+
                         \      /   |            |   \    +----+
                          \    /    |            |    +---|PE11|--AC--
                           \  /     |            |        +----+
                            \/      |            |
                            /\      |            |
                           /  \     |            |        +----+
                          /    \    |            |    +---|PE12|--AC--
                         /      \   |            |   /    +----+
   Ethernet        +---+/        \+---+        +---+/     +----+
   Broadcast->-AC--|PE2|----------|P2 |--------|P4 |------|PE13|--AC--
   Source          +---+          +---+        +---+\     +----+
                                  / | \              \    +----+
                                 /  |  \              +---|PE14|--AC--
                                /   |   \                 +----+
                               /    |    \
                            +---+ +---+ +---+
                            |PE6| |PE7| |PE8|
                            +---+ +---+ +---+
                              |     |     |
                              AC    AC    AC
```

```
                              |     |     |
```

Figure 3 - Physical Topology for Broadcast video

Figure 3 is split in three logical components:

- The Core network composed of P1, P2, P3 & P4. These 4 network elements are P routers connected in a ring.

- The Data Centers. These are a few large PoPs with high resiliency that hold the video content. PE1 & PE2 are located in one Data Center and are dual-homed to the core network. An Ethernet broadcast source is connected to each PE in the Data Center.

- The Aggregation network. The Aggregation network is responsible for aggregating last mile technology towards endusers (direct fiber, GPON, DSL, etc.). PE3, PE4, until PE14 are VPLS PE routers in an aggregation PoP and single-homed to the Core network.

There are two different video distribution services organised as follows:

- PE1 is connected to PE3, PE4, ... PE14 via VPLS instance-1.
- One Ethernet broadcast source is connected to PE1 into VPLS instance-1.
- PE2 is connected to PE3, PE4  ...PE14 via VPLS instance-2.
- One Ethernet broadcast source is connected to PE2 into VPLS instance-2.

## A.2. Impact of Physical Topologies on Ethernet Broadcast/multicast replication

Following the standard VPLS flooding mechanism, each time PE1 receives one broadcast frame from the Ethernet broadcast source on VPLS-1, PE1 will replicate 12 times the incoming frame.

Similarly, each time PE2 receives one broadcast frame from the Ethernet broadcast source on VPLS-2, PE2 will replicate 12 times the incoming frame.

### A.3. Proposed enhancement of Ethernet broadcast/multicast

The proposed enhancements are done in three steps:

- create P2MP LSPs for the infrastructure. These P2MP LSPs are used to carry one or more P2MP PWs.

- create unidirectional P2MP PWs by selectively choosing PEs where the optimisation should occur.

- forward Ethernet broadcast/multicast frames onto the P2MP PWs where these P2MP PWs have been created.

It is up to the network operator to decide how the distribution of the loading on physical link should occur.

Two different examples are presented below.

#### A.3.1. One possible enhancement scenario

##### A.3.1.1. Initial Deployment

In this scenario, the operator decides to create the following P2MP LSPs:

- PE1->PE3-5 via P1 called LSP1
- PE1->PE6-8 via P2 called LSP2
- PE1->PE9-11 via P3 called LSP3

- PE2->PE3-5 via P1 called LSP4
- PE2->PE6-8 via P2 called LSP5
- PE2->PE9-11 via P3 called LSP6

The operator then creates the following P2MP PWs:

- PE1->PE3-5 via P2MP PW1 over LSP1
- PE1->PE6-8 via P2MP PW2 over LSP2
- PE1->PE9-11 via P2MP PW3 over LSP3

- PE2->PE3-5 via P2MP PW4 over LSP4
- PE2->PE6-8 via P2MP PW5 over LSP5
- PE2->PE9-11 via P2MP PW6 over LSP6

There is no P2MP PWs between PE1 and PE12, PE13 and PE14.
There is no P2MP PWs between PE2 and PE12, PE13 and PE14.

There are several reasons why a P2MP PW may not be available on this part of the network (e.g. PE12, PE13 and PE14), for example:

- the hardware/software may not allow the support of the required features (P2MP LSPs and/or P2MP PWs).
- the operator does not need to improve multicast/broadcast services there (e.g. no specific bandwidth issue).
- the operator is currently under a migration phase where only part of the network is migrated at a time.

In this case, when PE1 receives one broadcast frame from the Ethernet broadcast source on VPLS-1:

- PE1 sends one copy of the broadcast frame onto P2MP PW1
- PE1 sends one copy of the broadcast frame onto P2MP PW2
- PE1 sends one copy of the broadcast frame onto P2MP PW3
- PE1 sends one copy onto the P2P PW towards PE12
- PE1 sends one copy onto the P2P PW towards PE13
- PE1 sends one copy onto the P2P PW towards PE14

PE1 only replicates 6 copies now (this is an improvement from 12 copies if only using P2P PWs).

#### A.3.1.2 Multiple P2MP PWs

Let's assume now that a new broadcast service is targeted at covering endusers geographically connected to PE9, PE10 and PE11.

For example, this could be a wholesale service, where another carrier with limited footprint for the region covered by PE9, PE10 and PE11 is seeking access for deploying its own broadcast application.

Based on the proposal in this document, and assuming that the application also needs unicast data exchange, if the new broadcast source is connected to PE1, it is then possible to:

- Create a new VPLS instance on PE1, PE9, PE10 and PE11 and a full-mesh of P2P PWs between all 4 PEs.

- Build a new P2MP PW, called P2MP PW7 between PE1, PE9, PE10 & PE11 that uses the existing P2MP LSP - LSP3.

This proposal allows for both P2MP PW3 and P2MP PW7 to be carried through a single MPLS P2MP tunnel, thus, removing the need to maintain state in the network core for individual P2MP PWs. The P routers in the core only need to be aware of the P2MP LSPs.

### A.3.2. Another possible enhancement scenario

In this scenario, the operator decides to create the following two P2MP LSPs:

- PE1-> PE3-14 via LSP1:
    - P1 as a branch towards PE3, PE4, PE5, P2 and P3
    - P2 as a branch towards PE6, PE7, PE8 and P4
    - P3 as a branch towards PE9, PE10 and PE11
    - P4 as a branch towards PE12, PE13 and PE14

- PE2-> PE3-14 via LSP2:
    - P2 as a branch towards PE6, PE7, PE8, P1 and P4
    - P1 as a branch towards PE3, PE4, PE5 and P3
    - P3 as a branch towards PE9, PE10 and PE11
    - P4 as a branch towards PE12, PE13 and PE14

The operator then creates the following P2MP PWs:

- PE1-> PE3-14 via P2MP PW1 over LSP1
- PE2-> PE3-14 via P2MP PW2 over LSP2

This case improves the P2P PW scenario as PE1 only replicates a single copy of the broadcast frame received from the Ethernet broadcast source.