Network Working Group
Internet Draft
Category: Standards Track
Expiration Date: September 2010

R. Aggarwal
Juniper Networks

F. Jounay
France Telecom

March 08, 2010

Point-to-Multipoint Pseudo-Wire Encapsulation

draft-raggarwa-pwe3-p2mp-pw-encaps-01.txt

Status of this Memo

Copyright and License Notice

Abstract

A Point-to-Multipoint (P2MP) Pseudowire (PW) is a mechanism that emulates the essential attributes of a P2MP Telecommunications service such as P2MP ATM over a Packet Switched Network (PSN).

This document describes the encapsulation and data plane procedures for a P2MP PW. These procedures are meant to be independent of the control plane used to signal a P2MP PW.

Table of Contents

1. Specification of requirements

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. Introduction

A Point-to-Multipoint (P2MP) Pseudowire (PW) is a mechanism that emulates the essential attributes of a P2MP Telecommunications service such as P2MP ATM over a Packet Switched Network (PSN). One of the applicabilities of a P2MP PW is to deliver a Layer 2 multicast service, that carries multicast frames (encoded using Layer 2 or IP mechanisms) from a multicast source to one or more multicast receivers.

The required functions of P2MP PWs include encapsulating service-

specific PDUs arriving at an ingress Attachment Circuit (AC), and carrying them across a tunnel to one or more egress ACs, managing their timing and order, and any other operations required to emulate the behavior and characteristics of the service as faithfully as possible.

P2MP PWs extend the PWE3 architecture [RFC3985] to offer a P2MP Telecommunications service. They follow the PWE3 architecture as described in [RFC3985] with modifications as outlined in this document. One notable difference between point-to-point (P2P) PWs as outlined in [RFC3985] and P2MP PWs is that the former emulate a bidirectional service whereas the latter emulate a unidirectional service.

This document describes the encapsulation and data plane procedures for a P2MP PW. These procedures are meant to be independent of the control plane used to signal a P2MP PW.

3. P2MP PW Semantics

A P2MP PW provides a mechanism for the root CE to send traffic to one or more leaf CEs over a PSN.

A root CE in a sender site sends traffic on one or more ACs to the root PE. The root PE delivers this traffic over a P2MP PW to one or more leaf PEs. Each leaf PE in turn delivers this traffic to one or more leaf CEs in a receiver site. A particular leaf CE MUST receive this traffic over a single AC.

A particular leaf CE may receive traffic from multiple sender CEs. Traffic from different sender CEs is received by a leaf PE over unique P2MP PWs. The leaf PE must use unique ACs to send traffic received over unique P2MP PW, to the same leaf CE or different leaf CEs.

4. P2MP PW Encapsulation

An architectural building block of P2MP PWs is that routers not directly connected to VPN customers should carry no VPN state, or at minimum this state should not grow linearly with the number of individual connections provisioned on the edge devices.

In order to achieve this traffic belonging to different P2MP PWs, which may be in different L2VPNs, may be carried over the same PSN tunnel. The PSN tunnels MUST be based on P2MP MPLS LSPs signaled

using either RSVP-TE [RFC4875] or P2MP LDP [mLDP]. Penultimate-hop-popping MUST be disabled. The egress PEs MUST NOT advertise IMPLICIT NULL or EXPLICIT NULL for P2MP PSN Tunnels that are used to carry P2MP PW traffic. This document uses the terms P-Multicast Tree and P2MP PSN Tunnel inter-changeably.

When multiple P2MP PWs are carried over the same P2MP MPLS PSN tunnel there is a need to identify at the leaf PE the P2MP PW the packet belongs to. This is done by using a P2MP PW demultiplexor that allows an egress PE to determine in the data plane, the P2MP PW for which the packet is intended. A MPLS label is used as the P2MP PW demultiplexor. The root PE MUST use this label as the bottom-most label while encapsulating a customer data packet in a P2MP PW. Each of the leaf PEs must be able to associate this inner label with the same P2MP PW and use it to demultimplex the traffic received over the P2MP PSN tunnel.

This document requires the use of an upstream assigned MPLS label [RFC5331] as the P2MP PW demultiplexor.

The MPLS upstream assigned label that is used as the P2MP PW demultiplexor MUST be assigned by the root PE i.e. the PE connected to the root CE. It MUST be distributed by the root PE to the leaf PEs i.e. the PEs connected to the receiver CEs via a control plane mechanism or via provisioning. The control plane mechanisms used to achieve this are outside the scope of this document.

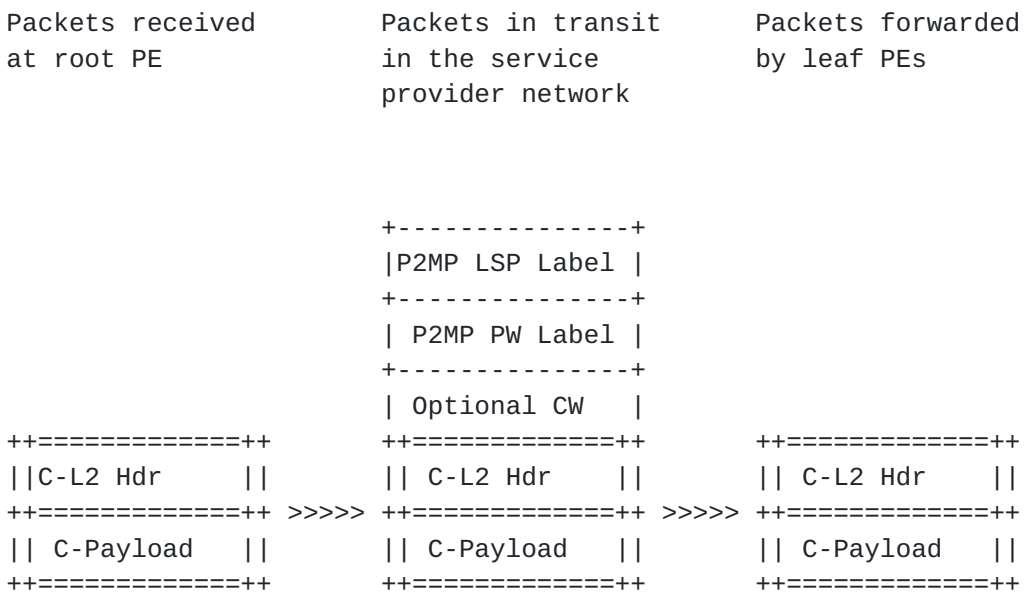
5. P2MP PW Encapsulation Type

The PW encapsulation types specified in [RFC4446] MUST be used for P2MP PWs.

6. Data Forwarding

6.1. MPLS Tree Encapsulation

The following diagram shows the progression of a L2 multicast packet as it enters and leaves the SP network when MPLS trees are being used. RSVP-TE P2MP LSPs are examples of such trees. The modification to the Layer 2 frame, by the root PE, depends on the Layer 2 encapsulation type. This document requires that the encapsulation methods used in transporting of Layer 2 frames over tunnels be the same as described in [RFC4448], [RFC4618], [RFC4619], and [RFC4717]. Note that these encapsulation methods may result in inserting a control-word in the P2MP PW encapsulation as shown in the figure below.



6.2. Demultiplexing P2MP PW Traffic

Demultiplexing P2MP PW traffic on an egress PE requires the receiving PE to determine the P2MP PW the packet belongs to.

6.2.1. One P-Multicast Tree - One P2MP PW Mapping

When a P-Multicast tree is mapped to only one P2MP PW, determining the tree on which the packet is received is sufficient to determine the P2MP PW on which the packet is received. The tree is determined based on the tree encapsulation. When MPLS encapsulation is used, eg: RSVP-TE P2MP LSPs, the outer MPLS label is used to determine the tree. Penultimate-hop-popping MUST be disabled on the MPLS LSP (RSVP-TE P2MP LSP or LDP P2MP LSP).

6.2.2. One P-Multicast Tree - Many P2MP PW Mapping

When traffic belonging to multiple P2MP PWs is carried over the same tree, the P2MP PW that the packet belongs to is identified by using an inner label. This label determines the P2MP PW for which the packet is intended. The root PE uses this label as the inner label while encapsulating a customer multicast data packet. Each of the leaf PEs must be able to associate this inner label with the same P2MP PW and use it to demultimplex the traffic received over the P2MP LSP.

This document requires the use of upstream label assignment by the root PE [RFC5331]. Hence the inner label is assigned by the root PE. When the egress PE receives a packet over a P2MP LSP, the outer encapsulation [in the case of MPLS P2MP LSPs, the outer MPLS label] specifies the label space to perform the inner label lookup. The same label space MUST be used by the egress PE for all P-Multicast trees that have the same root [RFC5331].

If the tree uses MPLS encapsulation, as in RSVP-TE P2MP LSPs, the outer MPLS label and the incoming interface provides the label space of the label beneath it. This assumes that penultimate-hop-popping is disabled. The egress PE MUST NOT advertise IMPLICIT NULL or EXPLICIT NULL for that tree. Once the label representing the tree is popped off the MPLS label stack, the next label is the demultiplexing information that allows the proper P2MP PW to be determined. This determines the set of egress CE ACs that the packet needs to be forwarded to. The egress PE then strips the inner MPLS label and sends the packet to this set of egress CEs.

6.3. Layer 2 MTU

This document requires that the Layer 2 MTU configured on all the access circuits connecting the CEs to PEs, for a given P2MP PW be the same. The P2MP PW signaling mechanisms must provide a means for ensuring this.

The MTU on the Layer 2 access links MUST be chosen such that the size of the L2 frames plus the P2MP PW header does not exceed the MTU of the SP network. Layer 2 frames that exceed the MTU after encapsulation MUST be dropped.

7. Security Considerations

TBD

8. IANA Considerations

TBD

9. Acknowledgments

Thanks to Yakov Rekhter and Kireeti Kompella for the discussions that lead to this document.

10. References

10.1. Normative References

[VPLS-MCAST] R. Aggarwal et. al., "Multicast in VPLS", draft-ietf-l2vpn-vpls-mcast, work in progress.

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

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

10.2. Informative References

[VPMS-REQ] Y. Kamite, F. Jounay, "Framework and Requirements for Virtual Private Multicast Service (VPMS)", draft-ietf-l2vpn-vpms-frmwk-requirements, work in progress

[RFC3985] S. Bryant et. al., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

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

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

[RFC4875] R. Aggarwal et. al, "Extensions to RSVP-TE for Point to Multipoint TE LSPs", draft-ietf-mpls-rsvp-te-p2mp-07.txt

[MLDP] I. Minei et. al, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", draft-ietf-mpls-ldp-p2mp-02.txt

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

[RFC4618] Martini, L., Rosen, E., Heron, G., and A. Malis,

"Encapsulation Methods for Transport of PPP/High-Level Data Link Control (HDLC) over MPLS Networks", RFC 4618, September 2006.

[RFC4619] Martini, L., Kawa, C., and A. Malis, "Encapsulation Methods for Transport of Frame Relay over Multiprotocol Label Switching (MPLS) Networks", RFC 4619, September 2006.

[RFC4717] Martini, L., Jayakumar, J., Bocci, M., El-Aawar, N., Brayley, J., and G. Koleyni, "Encapsulation Methods for Transport of Asynchronous Transfer Mode (ATM) over MPLS Networks", RFC 4717, December 2006.

11. Author's Address

Rahul Aggarwal
Juniper Networks
1194 North Mathilda Ave.
Sunnyvale, CA 94089
Email: rahul@juniper.net

Frederic Jounay
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
FRANCE
Email: frederic.jounay@orange-ftgroup.com