

Network Working Group  
Internet Draft  
Category: Standards Track  
Expires: January 2011

F. Jounay  
J.L. Le Roux  
P. Niger  
France Telecom Orange

L.Jin  
ZTE

Y. Kamite  
NTT Communications

July 12, 2010

**LDP Extensions for Leaf-initiated Point-to-Multipoint Pseudowire  
draft-jounay-pwe3-leaf-initiated-p2mp-pw-02.txt**

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on January, 2011.

Abstract

This document provides a solution to extend LDP signaling to set up and maintain Point-to-Multipoint MultiSegment Pseudowire (P2MP MS-PW). The P2MP PW described in this draft is constructed by multiple unidirectional PW segments. Such an extension of existing point to point Pseudowire is made necessary by new applications. The document only deals with the leaf-initiated P2MP PW setup and maintenance. The

processing for setting up a P2MP MS-PW is based on the same as MLDP for P2MP LSP setup.

## 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

<a href="#">1. Terminology.....</a>	<a href="#">3</a>
<a href="#">2. Preliminary Notes.....</a>	<a href="#">3</a>
<a href="#">3. Introduction.....</a>	<a href="#">3</a>
<a href="#">4. P2MP MS-PW Reference Model.....</a>	<a href="#">3</a>
<a href="#">5. MPLS-PW to MPLS-PW Replication.....</a>	<a href="#">5</a>
<a href="#">6. Overview of the P2MP MS-PW Setup.....</a>	<a href="#">5</a>
<a href="#">7. LDP S-PE TLV.....</a>	<a href="#">5</a>
<a href="#">8. Configuration.....</a>	<a href="#">6</a>
<a href="#">9. Capability Negotiation Procedure.....</a>	<a href="#">6</a>
<a href="#">10. Signaling for P2MP MS-PW with dynamic S-PE routing.....</a>	<a href="#">6</a>
<a href="#">10.1. Label Map.....</a>	<a href="#">7</a>
<a href="#">10.1.1. Determining one's 'upstream PE'.....</a>	<a href="#">7</a>
<a href="#">10.1.2. Egress T-PE Operation.....</a>	<a href="#">7</a>
<a href="#">10.1.3. Branch S-PE Operation.....</a>	<a href="#">7</a>
<a href="#">10.1.4. Ingress T-PE Operation.....</a>	<a href="#">7</a>
<a href="#">10.2. Label Withdraw.....</a>	<a href="#">8</a>
<a href="#">10.2.1. Egress T-PE Operation.....</a>	<a href="#">8</a>
<a href="#">10.2.2. Branch S-PE Operation.....</a>	<a href="#">8</a>
<a href="#">10.2.3. Ingress T-PE Operation.....</a>	<a href="#">8</a>
<a href="#">10.2.4. Upstream PE Change.....</a>	<a href="#">8</a>
<a href="#">11. Security Considerations.....</a>	<a href="#">8</a>
<a href="#">12. IANA Considerations.....</a>	<a href="#">9</a>
<a href="#">13. Acknowledgments.....</a>	<a href="#">9</a>
<a href="#">14. References.....</a>	<a href="#">9</a>
<a href="#">14.1. Normative References.....</a>	<a href="#">9</a>
<a href="#">14.2. Informative References.....</a>	<a href="#">9</a>
<a href="#">Copyright Notice.....</a>	<a href="#">10</a>
<a href="#">Authors's Addresses..</a>	<a href="#">11</a>
<a href="#">Intellectual Property and Copyright Statements.....</a>	<a href="#">12</a>



## **1. Terminology**

This document uses acronyms and terminologies defined in [[RFC5036](#)], [[RFC3985](#)], [P2MP PW REQ] and [[RFC5254](#)].

## **2. Preliminary Notes**

The current version of the document does not cover:

- The mechanism for the leaves to discover the P2MP PW FEC identifying the P2MP MS-PW, out of the scope of this document.
- The P2MP PW Upstream Label Assignment required when the underlying layer is a P2MP LSP. This document describes LDP extensions for setting up P2MP MS-PW where the PW segments are supported over P2P PSN tunnels.

## **3. Introduction**

[P2MP PW REQ] describes a set of requirements for setting up a P2MP PW setup. In the MS-PW architecture, the underlying layer which supports a PW segment belonging to the PW tree may be either a unidirectional P2P or a P2MP PSN tunnel.

Note that a P2MP PW is optionally bidirectional [P2MP PW REQ]. This version of the document does not cover the return path from leaf to root, this point will be addressed in a next version.

This document describes LDP extensions for setting up P2MP MS-PW where the PW segments are supported over P2P PSN tunnels.

For that purpose the P2MP PW FEC element is reused from [ROOT INIT P2MP PW] to encode MS-PW parameters. The procedures for setting up a P2MP MS-PW are very similar with LDP mechanisms for setting P2MP LSP [[MLDP](#)], where hops are here S-PEs and T-PEs. Therefore a leaf can join the tree by sending a Label Map associated to this FEC towards the root.

## **4. P2MP MS-PW Reference Model**

Figure 1 describes the P2MP MS-PW reference model which is derived from [P2MP PW REQ] to support P2MP emulated services.



Note that a P2MP MS-PW may obviously transit through more than one S-PE along its path.



## 5. MPLS-PW to MPLS-PW Replication

Referencing Figure 1, PDUs are replicated to the Pseudowire segments at the PW label level. Hence the data plane does not need any special knowledge of the specific PW type. A simple standard MPLS label swap operation is sufficient to connect the PW segments. However when pushing a new PSN label the TTL SHOULD be set to 255, or some other locally configured fixed value. This process can be repeated as many times as necessary, the only limitation to the number of S-PEs traversed is imposed by the TTL field of the PW MPLS Label. The setting of the TTL of the PW MPLS label is a matter of local policy on the originating PE, but SHOULD be set to 255 except OAM packet.

## 6. Overview of the P2MP MS-PW Setup

The P2MP MS-PW setup relies on the use of the P2MP PW FEC Element defined also in [ROOT INIT P2MP PW]. The solution aims at setting up a unidirectional P2MP MS-PW.

The principle proposed here relies on a leaf-initiated P2MP MS-PW setup. In the proposed approach the leaf is assumed to know the P2MP PW FEC which contains the source AII address and the VPN ID (AGI). The procedure used for the P2MP PW FEC discovery by the leaves is out of scope of this document.

The document describes the solution to setup the P2MP MS-PW in the case the PW segments are supported individually over a P2P PSN tunnel. After a negotiation procedure between Ingress T-PE/S-PE and S-PE/Egress T-PEs to verify their P2MP PW FEC capability, the Egress T-PE sends a Label Map to its upstream PE selected to reach the SAII specified in the P2MP PW FEC. At turn the S-PE carries on the signalling procedure by sending if required a new Label Map towards its next hop to reach the source SAII. In the MS-PW architecture, the hop consists either in a branch S-PE or the Ingress T-PE. Each PE receiving the P2MP FEC installs a forwarding state to map traffic into the P2MP MS-PW.

The definition of PW Type, C bit, PW Info Length, AGI, SAII, and Optional Parameters are same as defined in [ROOT INIT P2MP PW].

## 7. LDP S-PE TLV

The S-PE TLV defined in [SEGMENT MS-PW] can also be applied for P2MP MS-PW. PW loop detection will be performed by S-PE which is same as [SEGMENT MS-PW], by using Sub-TLV of Local IP address of S-PE. When egress PE sends notification to ingress PE to indicate PW status, each S-PE on the path to ingress PE SHOULD append S-PE TLV with local IP address to LDP notification message, which allows the Root T-PE to build the P2MP MS-PW topology.





## 8. Configuration

After configuring on each T-PE the attached AIIIs, it is assumed that all the PEs (Ingress/Egress T-PEs and all S-PEs) maintain an AII PW routing table which gives for each AII as entry the "next hop" to reach that AII. This AII routing table can be filled manually or updated dynamically by means of some extended routing protocol like proposed in [DYN MS-PW]. The construction of the table is out of scope of the present document.

Each PE relies on its AII PW routing table to select the next hop PE (S-PE or T-PE) to reach a given AII.

The target-LDP session between T-PE and S-PE, or two S-PEs should be configured automatically or manually. The P2MP MS-PW signaling message should be transmitted over this target-LDP session.

## 9. Capability Negotiation Procedure

For the dynamic LDP protocol, the capability negotiation the solution MUST follow the PW Status Capability advertisement mechanism described in [ROOT INIT P2MP PW].

The PEs belonging to a given P2MP MS-PW MUST support the P2MP PW FEC Element used by LDP to setup the P2MP MS-PW.

## 10. Signaling for P2MP MS-PW with dynamic S-PE routing

The following defines the rules for the processing and propagation of the P2MP FEC Element for the Leaf-initiated P2MP MS-PW setup. The following notation is derived from [MLDP] and is used in the processing rules:

1. P2MP PW FEC Element <X, Y>: a FEC Element with SAII X, AGI Y.
2. P2MP PW Label Map <X, Y, L>: a Label Map message with a FEC TLV with a single P2MP PW FEC Element <X, Y> and Label TLV with label L.
3. P2MP PW Label Withdraw <X, Y, L>: a Label Withdraw message with a FEC TLV with a single P2MP PW FEC Element <X, Y> and Label TLV with label L.
4. P2MP MS-PW <X, Y> (or simply <X, Y>): a P2MP MS-PW with SAII X and AGI Y.
5. The notation  $L' \rightarrow \{ \langle I1, L1 \rangle \langle I2, L2 \rangle \dots, \langle In, Ln \rangle \}$  on branch S-PE S means that on receiving a packet with label L', S makes n copies



of the packet. For copy  $i$  of the packet,  $S$  swaps  $L'$  with  $L_i$  and sends it out over interface  $I_i$ .

The procedures below are organized by the role which the PE plays in the P2MP MS-PW. T-PE  $Z$  knows that it is an Egress T-PE by a discovery process which is outside the scope of this document. A T-PE is defined as an Egress T-PE if one or several leaf AIIs are configured. During the course of protocol operation, the Ingress T-PE recognizes its role because it owns the SAI of the PW tree.

### **10.1. Label Map**

The following lists procedures for generating and processing P2MP Label Map messages for PEs participating in a P2MP MS-PW.

For the approach described here we use downstream assigned labels.

#### **10.1.1. Determining one's 'upstream PE'**

For the case of P2MP MS-PW with dynamic S-PE routing, a PE  $Z$  that is part of P2MP MS-PW  $\langle X, Y \rangle$  determines the T-LDP peer  $U$  which lies on the best path from  $Z$  to the SAI. The path selection is achieved by means of looking up the AII PW routing table.  $U$  is  $Z$ 's "Upstream PE" for  $\langle X, Y \rangle$ .

#### **10.1.2. Egress T-PE Operation**

An Egress T-PE  $Z$  of P2MP MS-PW  $\langle X, Y \rangle$  determines its upstream PE  $U$  for  $\langle X, Y \rangle$ , allocates a label  $L$ , and sends a P2MP PW Label Map  $\langle X, Y, L \rangle$  to  $U$ .

#### **10.1.3. Branch S-PE Operation**

Suppose a branch S-PE  $Z$  receives a P2MP PW Label Map  $\langle X, Y, L \rangle$  from LDP peer  $T$ .  $Z$  checks whether it already has state for  $\langle X, Y \rangle$ . If not,  $Z$  allocates a label  $L'$ , and installs state to swap  $L'$  with  $L$  over interface  $I$  associated with peer  $T$ .  $Z$  also determines its upstream PE  $U$  for  $\langle X, Y \rangle$  and sends a P2MP PW Label Map  $\langle X, Y, L' \rangle$  to  $U$ .

If  $Z$  already has state for  $\langle X, Y \rangle$ , then  $Z$  does not send a Label Map message for P2MP MS-PW  $\langle X, Y \rangle$ . All that  $Z$  needs to do in this case is to update its forwarding state. Assuming its old forwarding state was  $L' \rightarrow \{ \langle I_1, L_1 \rangle \langle I_2, L_2 \rangle \dots, \langle I_n, L_n \rangle \}$ , its new forwarding state becomes  $L' \rightarrow \{ \langle I_1, L_1 \rangle \langle I_2, L_2 \rangle \dots, \langle I_n, L_n \rangle, \langle I, L \rangle \}$ .

#### **10.1.4. Ingress T-PE Operation**

Suppose the Ingress T-PE  $Z$  receives a P2MP Label Map  $\langle X, Y, L \rangle$  from

peer T. Z checks whether it already has forwarding state for  $\langle X, Y \rangle$ .

Jounay et al.

Expires January, 2011

[Page 7]

If not, Z creates forwarding state to push label L onto the traffic that Z wants to forward over the P2MP MS-PW.

If Z already has forwarding state for <X, Y>, then Z adds "push label L, send over interface I" to the nexthop, where I is the interface associated with peer T.

## **10.2. Label Withdraw**

The following lists procedures for generating and processing P2MP PW Label Withdraw messages for PEs that participate in a P2MP MS-PW.

### **10.2.1. Egress T-PE Operation**

If an Egress T-PE Z discovers that it has no longer leaves AII belonging to the P2MP MS-PW, it SHOULD send a P2MP PW Label Withdraw <X, Y, L> to its upstream PE U for <X, Y>, where L is the label it had previously advertised to U for <X, Y>.

### **10.2.2. Branch S-PE Operation**

If a branch S-PE Z receives a P2MP PW Label Withdraw message <X, Y, L> from a node W, it deletes label L from its forwarding state, and sends a P2MP PW Label Release message with label L to W.

If deleting L from Z's forwarding state for P2MP MS-PW <X, Y> results in no state remaining for <X, Y>, then Z propagates the P2MP PW Label Withdraw for <X, Y>, to its upstream T, by sending a P2MP PW Label Withdraw <X, Y, L1> where L1 is the label Z had previously advertised to T for <X, Y>.

### **10.2.3. Ingress T-PE Operation**

The procedure when the Ingress T-PE of a P2MP MS-PW receives a P2MP PW Label Withdraw message are the same as for branch S-PE, except that it would not propagate the P2MP PW Label Withdraw upstream (as it has no upstream).

### **10.2.4. Upstream PE Change**

If, for a given PE Z participating in a P2MP MS-PW <X, Y>, the upstream PE changes, say from U to U', then Z MUST update its forwarding state by deleting the state for label L, allocating a new label, L', for <X,Y>, and installing the forwarding state for L'. In addition Z MUST send a P2MP PW Label Map <X, Y, L'> to U' and send a P2MP PW Label Withdraw <X, Y, L> to U.

## **11. Security Considerations**

This section will be added in a future version.



## **12. IANA Considerations**

This draft does not define any new protocol element, and hence does not require any IANA action.

## **13. Acknowledgments**

## **14. References**

### **14.1. Normative References**

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), March 1997.
- [RFC5036] Andersson, L., Doolan, P., Feldman, N., Fredette, A., Thomas, B., "LDP Specification", October 2007.
- [RFC3985] Bryant, S., Pate, P. "PWE3 Architecture", March 2005
- [RFC5254] Bitar, N., Bocci, M., and Martini, L., "Requirements for inter domain Pseudo-Wires", October 2008

### **14.2. Informative References**

- [P2MP PW REQ] Jounay, F., Niger, P., Kamite, Y., Martini, L., Heron, G., Wang, L., Delord, .S, "Use Cases and signaling requirements for Point-to-Multipoint PW", Internet Draft, [draft-ietf-pwe3-p2mp-pw-requirements-02.txt](#), January 2010
- [DYN MS-PW] Balus, F., Bocci, M., Martini, L, "Dynamic Placement of Multi Segment Pseudo Wires", Internet Draft, [draft-ietf-pwe3-dynamic-ms-pw-10.txt](#), October 2009
- [SEGMENT MS-PW] Martini, L, &all, "Segmented Pseudowire", Internet Draft, [draft-ietf-pwe3-segmented-pw-15.txt](#), June 2010
- [MLDP] Minei, I., Kompella, K., Thomas, B., Wijnands, I. "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", Internet Draft, [draft-ietf-mpls-ldp-p2mp-10](#), July 2010





[ROOT INIT P2MP PW]

Martini, L., Jounay, &all , "Signaling  
Root-Initiated Point-to-Multipoint  
Pseudowires using LDP", [draft-martini-  
pwe3-p2mp-pw-01](#), October, 2009

#### Author's Addresses

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

Jean-Louis Le Roux  
France Telecom  
2, avenue Pierre-Marzin  
22307 Lannion Cedex  
FRANCE  
Email: jeanlouis.leroux@orange-ftgroup.com

Philippe Niger  
France Telecom  
2, avenue Pierre-Marzin  
22307 Lannion Cedex  
FRANCE  
Email: philippe.niger@orange-ftgroup.com

Yuji Kamite  
NTT Communications Corporation  
Tokyo Opera City Tower  
3-20-2 Nishi Shinjuku, Shinjuku-ku  
Tokyo 163-1421  
Japan  
Email: y.kamite@ntt.com

Lizhong Jin  
ZTE Corporation  
889, Bibo Road  
Shanghai, 201203, China  
Email: lizhong.jin@zte.com.cn

#### Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the  
document authors. All rights reserved.



This document is subject to [BCP 78](http://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.