

Internet Engineering Task Force
Internet Draft
Expiration Date: September 2011
Intended status: Standards Track

Sami Boutros
Luca Martini
Siva Sivabalan
Cisco

Maciek Konstantynowicz
Juniper

Frederic Jounay
Philippe Niger
France Telecom

Gianni Del Vecchio
Swisscom

Thomas D. Nadeau
Huawei

Yuji Kamite
NTT Communications

Simon Delord
Telstra

Lizhong Jin
ZTE

Laurent Ciavaglia
Martin Vigoureux
Alcatel-Lucent

March 2, 2011

Signaling Root-Initiated Point-to-Multipoint Pseudowires using LDP

[draft-ietf-pwe3-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>.

Internet Draft draft-ietf-pwe3-p2mp-pw-02.txt

March 2, 2011

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

This Internet-Draft will expire on September 2, 2010

Abstract

This document specifies a mechanism to signal Point-to-Multipoint (P2MP) Pseudowires (PW) tree using LDP. Such a mechanism is suitable for any Layer 2 VPN service requiring P2MP connectivity over an IP or MPLS-enabled PSN. A P2MP PW established via the proposed mechanism is root initiated.

Table of Contents

1	Specification of Requirements	3
2	Introduction	3
3	Terminology	4
4	Signaling the P2MP PW	5
4.1	PW ingress to egress incompatibility issues	6
4.2	P2MP PW FEC Element	7
4.3	Group ID usage	9
4.4	Generic Label TLV	9
4.5	Transport LSP TLV	10
5	LDP Capability Negotiation	12
6	P2MP PW status	13
7	Security Considerations	13
8	IANA Considerations	13
8.1	FEC Type Name Space	13
8.2	LDP TLV TYPE	14
8.3	mLDP Opaque Value Element TLV TYPE	14
9	References	14
9.1	Normative References	14
9.2	Informative References	15
10	Author's Addresses	16

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) emulates the essential attributes of a unidirectional P2MP Telecommunications service such as P2MP ATM over PSN. A major difference between a Point-to-Point (P2P) PW outlined in [[RFC3985](#)] and a P2MP PW is that the former is intended for bidirectional service whereas the latter is intended for both unidirectional, or optionally bidirectional service. Requirements for P2MP PW are described in [[P2MP-PW-REQ](#)].

P2MP PW can be constructed as either Single Segment (P2MP SS-PW) or Multi Segment (P2MP MS-PW) Pseudowires as mentioned in [[P2MP-PW-REQ](#)]. P2MP MS-PW is outside the scope of this document. A reference model for P2MP PW is depicted in Figure 1 below. A transport LSP associated with a P2MP SS-PW SHOULD be a P2MP MPLS LSP (i.e., P2MP TE tunnel established via RSVP-TE [[RFC4875](#)] or P2MP LSP established via mLDP [[mLDP](#)]) spanning from the Root-PE to the Leaf-PE(s) of the P2MP SS-PW tree. For example, in Figure 1, PW1 can be associated with a P2MP TE tunnel or P2MP LSP setup using [[mLDP](#)] originating from PE1 and terminating at PE2 and PE3.

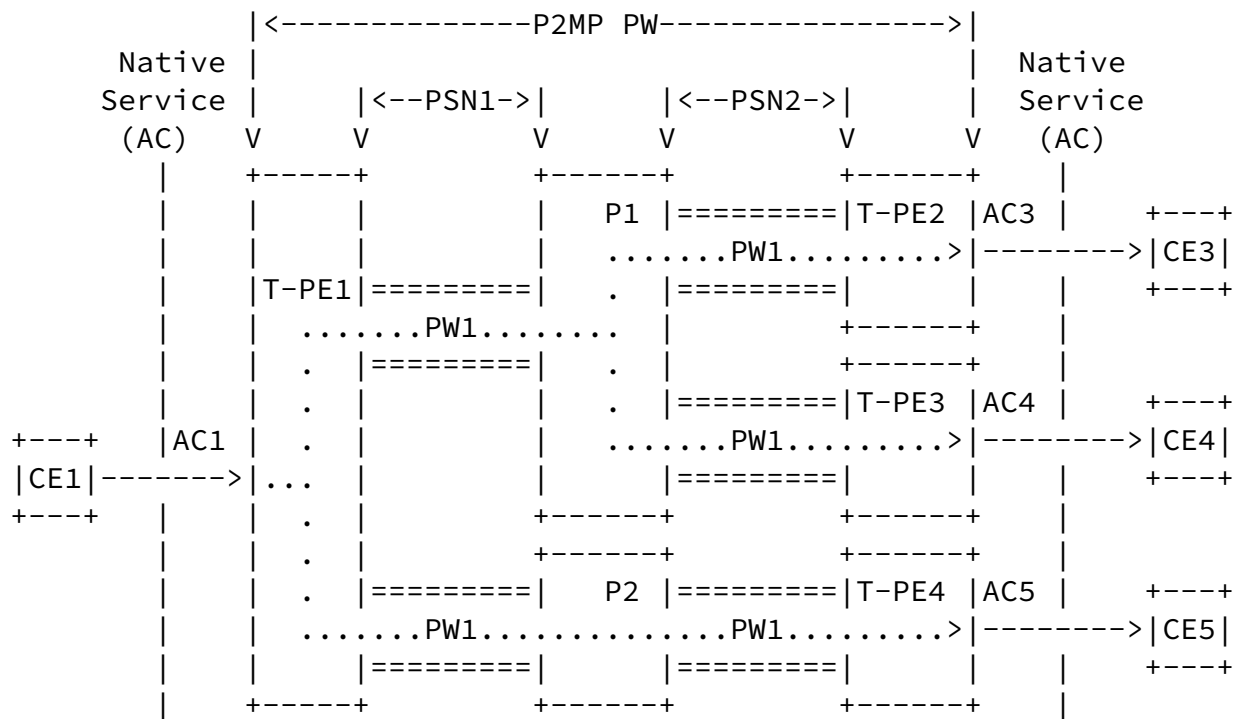


Figure 1: P2MP PW

Mechanisms for establishing P2P SS-PW using LDP are described in [RFC4447]. In this document, we specify a method to signal P2MP PW using LDP. In particular, we define new TLVs, parameters, and status codes to facilitate LDP to signal and maintain P2MP PWs.

Note that even though the traffic flow from a Root-PE to Leaf-PE(s)

is P2MP in nature, it may be desirable for any Leaf-PE to send unidirectional P2P traffic destined only to the Root-PE. The proposed mechanism takes such an option into consideration.

The P2MP PW requires an MPLS LSP to carry the PW traffic. the PW MPLS packet will be encapsulated according to the methods described in [\[RFC5332\]](#).

[3.](#) Terminology

FEC: Forwarding Equivalence Class

LDP: Label Distribution Protocol

mLDP: Label Distribution Protocol for P2MP LSP

LSP: Label Switching Path

MS-PW: Multi-Segment Pseudowire

P2P: Point to Point

P2MP: Point to Multipoint

PE: Provider Edge

PSN: Packet Switched Network

PW: Pseudowire

SS-PW: Single-Segment Pseudowire

S-PE: Switching Provider Edge Node of MS-PW

TE: Traffic Engineering

R-PE: Root-PE - ingress PE, PE initiating P2MP PW setup.

L-PE: Leaf-PE - egress PE.

4. Signaling the P2MP PW

In order to advertise labels as well as exchange PW related LDP messages, PEs must establish LDP sessions among themselves using the Extended Discovery Mechanisms. A PE discovers other PEs that are to be connected via P2MP PWs either via manual configuration or autodiscovery [[RFC6074](#)].

Root-PE and each Leaf-PE MUST be configured with the same FEC as defined in the following section.

P2MP PW requires that there is an active P2MP PSN LSP set up between Root-PE and Leaf-PE(s). Note that the procedure to set up the P2MP PSN LSP is different depending on the protocol used: RSVP-TE or mLDP.

In case of mLDP a Leaf-PE can decide to join the P2MP LSP at any time, while in case of RSVP-TE the P2MP LSP is set up by the Root-PE, generally at the initial service provisioning time. It should be noted that local policy can override any decision to join, add or prune existing or new Leaf-PE(s) from the tree.

In any case the PW setup can ignore these differences, and simply assume that the P2MP tunnel is available when needed.

The P2MP PW is initiated by the root (source) Provider Edge router

(R-PE), by simply sending an P2MP-PW LDP label mapping message to all the Leaf Provider Edge routers L-PEs. This label mapping message will contain the following:

- i. P2MP PW FEC element.
- ii. an Interface Parameters TLV, as described in [[RFC4447](#)] sec 5.3.2.1
- iii. a PW Grouping TLV, as described in [[RFC4447](#)] sec 5.3.2.2
- iv. a Transport LSP TLV.
- v. a label TLV for the upstream-assigned label R-PE to L-PE direction.
- vi. MAY contain a downstream-assigned label for the L-PE to R-PE direction.

The LDP liberal label retention mode is used, and per [[RFC5036](#)] requirement the label request message MUST also be supported.

The Upstream-assigned label is allocated according to the rules in [\[RFC5331\]](#).

When a Leaf-PE receives a PW Label Mapping Message, it MUST verify the associated P2MP transport LSP is in place.

If the associated transport P2MP LSP is not in place, and the transport LSP TLV type is LDP P2MP LSP, a Leaf-PE SHOULD attempt to join the P2MP transport associated with the P2MP PW.

If the associated transport P2MP LSP is not in place, and the transport LSP TLV type is RSVP-TE P2MP LSP, a Leaf-PE SHOULD await RSVP-TE P2MP LSP signaling from the Root-PE.

[4.1.](#) PW ingress to egress incompatibility issues

If a Root-PE signals a PW with a pw type, CW mode, or interface parameters that a particular Leaf-PE cannot accept, then the L-PE must simply not enable the PW, and notify the user. In this case a PW status message of 0x00000001 - Pseudowire Not Forwarding MUST also be sent to the R-PE.

Note that this procedure does not apply if the L-PE had not been provisioned with this particular P2MP PW. In this case according to the LDP liberal label retention rules, no action is taken.

[4.2.](#) P2MP PW FEC Element

[RFC4447] specifies two types of LDP FEC elements called "Pwid FEC Element" and "Generalized Pwid FEC Element" used to signal P2P PWs. We define a new type of FEC element called "P2MP PW FEC Element" whose type is 0x82 (Pending IANA Allocation) and is encoded as follows:

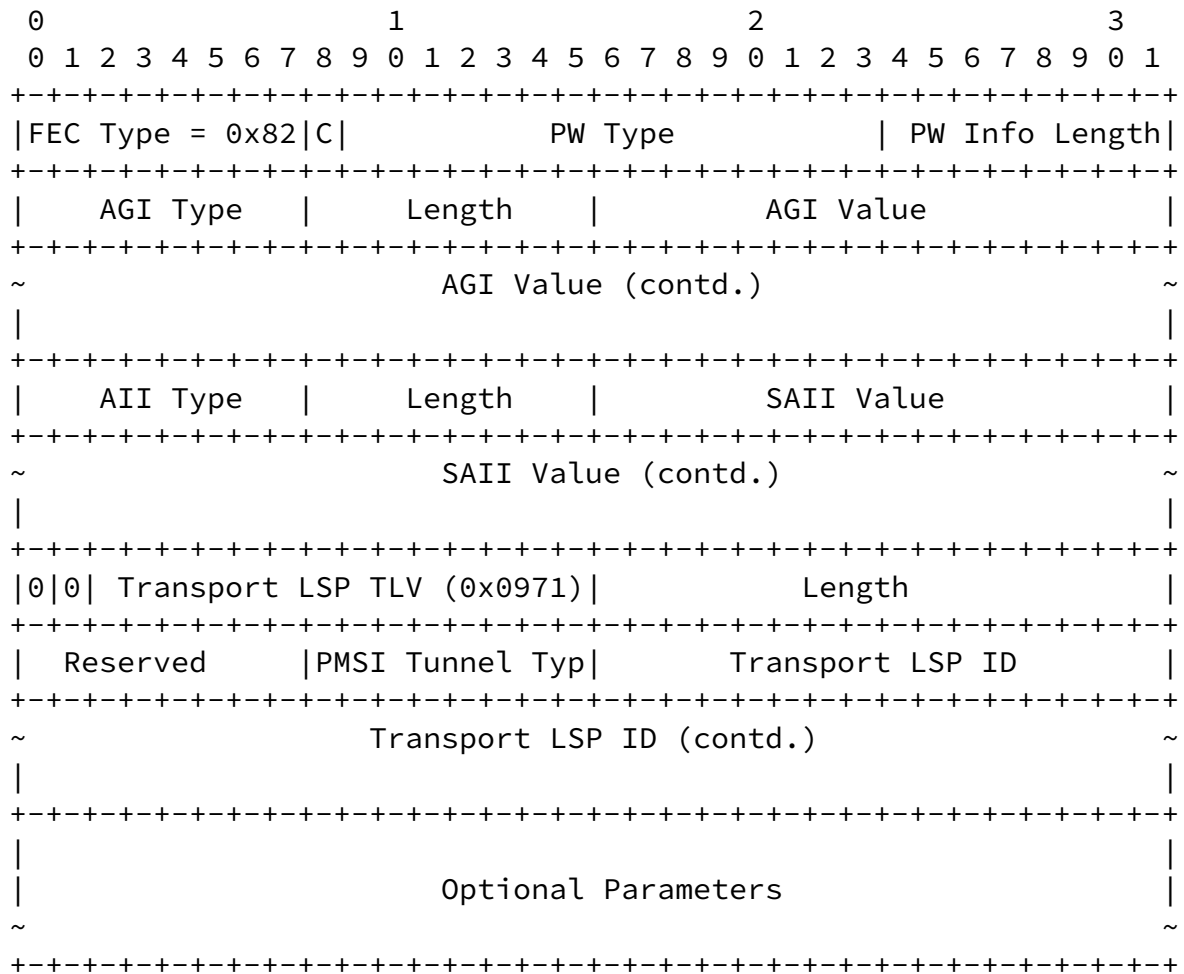


Figure 2: P2MP PW FEC Element

* PW Type:

15-bit representation of PW type, and the assigned values are assigned by IANA.

* C bit:

A value of 1 or 0 indicates whether control word is present or absent for the P2MP PW.

* PW Info Length:

Sum of the lengths of AGI, SAII and Optional Parameters field in octets. If this value is 0, then it references all PWs using the specified grouping ID. In this case, there are no other FEC element fields (AGI, SAII, etc.) present, nor any interface parameters TLVs.

★ AGI:

Attachment Group Identifier can be used to uniquely identify VPN or VPLS instance associated with the P2MP PW. This has the same format as that of the Generalized PWid FEC element [[RFC4447](#)].

★ SAII:

Source Attachment Individual Identifier is used to identify the root of the P2MP PW. The root is represented using AII type 2 format specified in [[RFC5003](#)]. Note that the SAII can be omitted by simply setting the length and type to zero.

P2MP PW is identified by the Source Attachment Identifier (SAI). If the AGI is non-null, the SAI is the combination of the SAII and the AGI, if the AGI is null, the SAI is the SAII.

★ Transport LSP TLV:

A P2MP PW MUST be associated with a transport LSP. The Transport LSP TLV contains the information required to identify the transport LSP. Note that the Transport LSP TLV MUST immediately follow the FEC, but is not part of the FEC, and SHOULD NOT be used in other messages where the FEC is used.

★ Optional Parameters:

The Optional Parameter field can contain some TLVs that are not part of the FEC, but are necessary for the operation of the PW. This document defines two such parameters: Interface Parameters TLV, and Group ID TLV.

The Interface Parameters TLV and Group ID TLV specified in [[RFC4447](#)] can also be used in conjunction with P2MP PW FEC. For Group ID TLV the sender and receiver of these TLVs should follow the same rules and procedures specified in [[RFC4447](#)]. For Interface Parameters TLV the procedure differs from the one specified in [[RFC4447](#)] due to specifics of P2MP connectivity. When the interface parameters are signaled by the Root-PE, the Leaf-PE must check if its configured value(s) is less than or equal to the threshold value provided by the

Root-PE (e.g. MTU size (Ethernet), max number of concatenated ATM cells, etc)). For other interface parameters like CEP/TDM Payload bytes (TDM), the value MUST match exactly the the one signaled by the Root-PE.

Note that since the LDP label mapping message is only sent by the R-PE to all the L-PEs it is not possible to negotiate any interface parameters.

[4.3.](#) Group ID usage

The Grouping TLV as defined in [[RFC4447](#)] contains a group ID capable of indicating an arbitrary group membership of a P2MP-PW. This group ID can be used in LDP "wild card" status, and withdraw label messages, as described in [[RFC4447](#)].

[4.4.](#) Generic Label TLV

For a given P2MP PW, a single upstream-assigned label is allocated by the Root-PE, and is advertised to all Leaf-PEs using the Generic Label TLV in the label mapping message containing the P2MP PW FEC element. The Root-PE imposes the upstream-assigned label on the outbound packets sent over the P2MP-PW, and using this label a Leaf-PE identifies the inbound packets arriving over the P2MP PW. Even though the P2MP PW is unidirectional, it may be possible for a Root-PE to receive traffic from any Leaf-PE using a unidirectional P2P PW in the reverse direction as outlined in [[P2MP-PW-REQ](#)]. For this purpose, the Root-PE can also allocate a unique downstream-assigned label for each Leaf-PE from which it is intended to receive P2P traffic. In other words, Label Mapping Message for a P2MP PW from a Root-PE to a Leaf-PE MUST carry a upstream-assigned label and MAY carry an OPTIONAL downstream-assigned label.

As in the case of P2P PW signaling, P2MP PW labels are carried within Generic Label TLV contained in LDP Label Mapping Message. A Generic Label TLV is formatted and processed as per the rules and procedures specified in [[RFC4447](#)]. But, as mentioned above, a Label Mapping Message for a P2MP PW can have up to two Generic Label TLVs; one for upstream-assigned label (always) and another for downstream-assigned label (optional). In the case of two Generic Label TLVs, the first TLV (from the beginning of the message) carries upstream-assigned label and the next generic label TLV carries the downstream-assigned label as shown below:

Internet Draft

[draft-ietf-pwe3-p2mp-pw-02.txt](#)

March 2, 2011

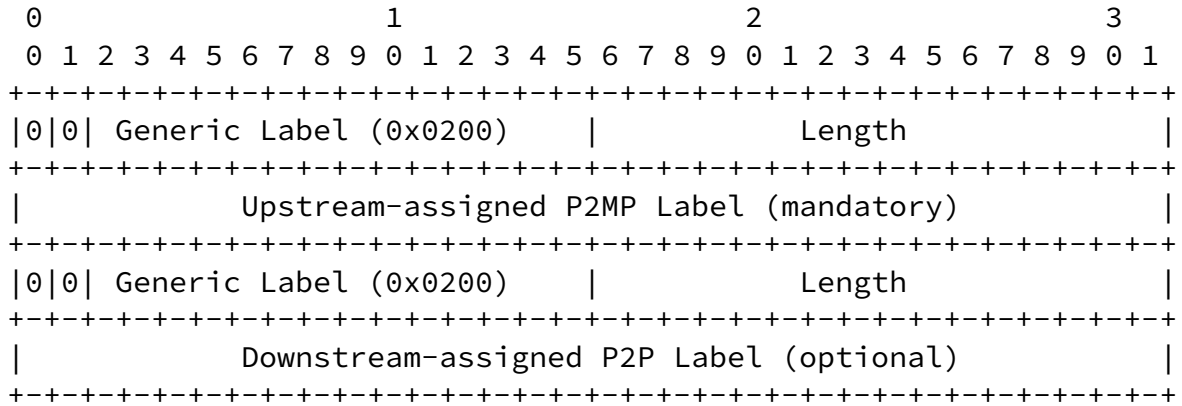
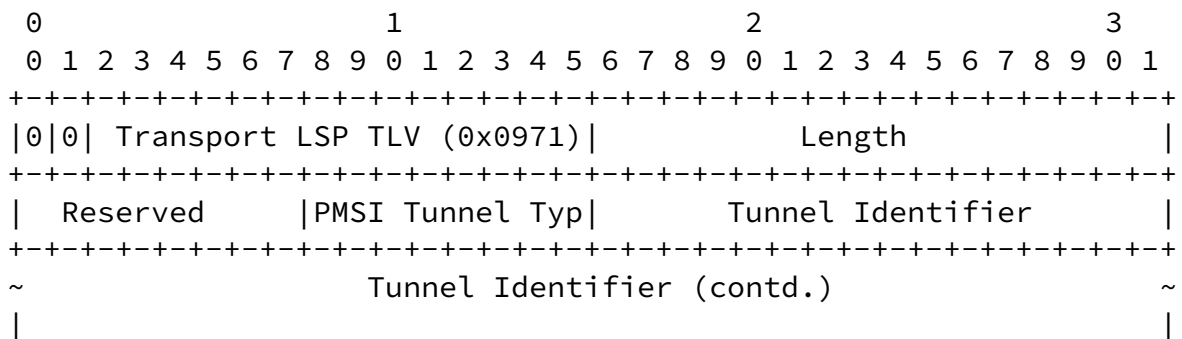


Figure 3: Generic Label TLVs in P2MP PW Label Mapping Message

Note that other type of TLVs may appear between the above generic label TLVs, however any other generic label TLV MUST NOT appear between the upstream-assigned P2MP Label TLV, and downstream-assigned P2P Label TLV.

[4.5.](#) Transport LSP TLV

A P2MP PW MUST be associated with a transport LSP which can be established using RSVP-TE or mLDP. Thus, a Label Mapping Message MUST contain the identity of the transport LSP. For this purpose, this specification introduces a new TLV called "Transport LSP TLV" which has the following format:



+ - + - + - + - + - + - + - + - + - + - + - + - + - + - + - + - + - +

Note: TLV number pending IANA allocation.

[Page 10]

[draft-ietf-pwe3-p2mp-pw-02.txt](#)

* Reserved Flags:

* PMSI Tunnel Type:

When the type is set to mLDP P2MP LSP, the Tunnel Identifier is a P2MP FEC Element as defined in [[mLDP](#)]. A new mLDP Opaque Value Element type for L2VPN-MCAST application needs to be allocated.

Editor Comment: The content of the Opaque Value Element TLV is a TBD.

* Tunnel Identifier:

Transport LSP TLV MUST be present only in the Label Mapping Message. An Root-PE sends Label Mapping Message as soon as the transport LSP ID associated with the P2MP PW is known (e.g., via configuration) regardless of the operational state of the transport LSP. Similarly, a Root-PE does not withdraw the labels when the corresponding transport LSP goes down. Furthermore, a Leaf-PE retains the P2MP PW labels regardless of the operational status of the transport LSP.

Note that a given transport LSP can be associated with more than one P2MP PW and all P2MP PWs will be sharing the same Root-PE and Leaf-PE(s).

In the case of LDP P2MP LSP, when a Leaf-PE receives the Label Mapping Message, it can initiate the process of joining the P2MP LSP tree associated with the P2MP PW.

In the case of RSVP-TE P2MP LSP, only the Root-PE initiates the signaling of P2MP LSP.

5. LDP Capability Negotiation

The capability of supporting P2MP PW must be advertised to all LDP peers. This is achieved by using the methods in [RFC5561] and advertising the P2MP PW LDP capability TLV. If an LDP peer supports the dynamic capability advertisement, this can be done by sending a new capability message with the S bit set for the P2MP PW capability TLV. If the peer does not support dynamic capability advertisement, then the P2MP PW TLV MUST be included in the LDP initialization procedures in the capability parameter [RFC5561].

In line with requirements listed in [RFC5561] the following TLV is defined to indicate the P2MP PW capability:

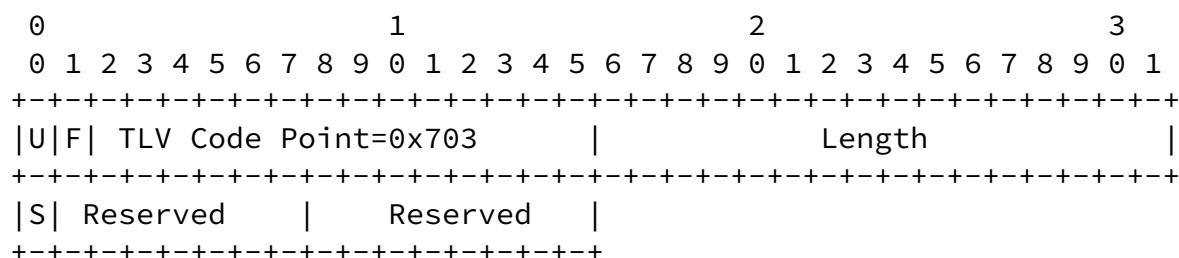


Figure 5: P2MP PW LDP Capability TLV

Note: TLV number pending IANA allocation.

* U-bit:

SHOULD be 1 (ignore if not understood).

* F-bit:

SHOULD be 0 (don't forward if not understood).

* TLV Code Point:

The TLV type, which identifies a specific capability. The P2MP PW capability code point is requested in the IANA allocation section below.

* S-bit:

The State Bit indicates whether the sender is advertising or withdrawing the P2MP PW capability. The State bit is used as follows:

- 1 - The TLV is advertising the capability specified by the TLV Code Point.

- 0 - The TLV is withdrawing the capability specified by the TLV Code Point.

6. P2MP PW status

In order to support the proposed mechanism, a node MUST be capable of handling PW status. As such, PW status negotiation procedure described in [RFC4447] is not applicable to P2MP PW.

Once a Leaf-PE successfully process a Label Mapping Message for a P2MP PW, it MUST send appropriate PW status according to the procedure specified [RFC4447] to notify the PW status. If there is no PW status notification required, then no PW status notification is sent. (for example if the P2MP PW is established and operational with a status of 0x00000000 no pw status message is necessary).

PW status message sent from any Leaf-PE to Root-PE contains P2MP PW

FEC to identify the PW. Finally, a Root-PE also sends PW status to Leaf-PE(s) to reflect its view of a P2MP PW state.

Connectivity status of the underlying P2MP LSP that P2MP PW is associated with, can be verified using LSP Ping and Traceroute procedures described in [[P2MP-LSP-PING](#)].

[7.](#) Security Considerations

The security measures described in [[RFC4447](#)] is adequate for the proposed mechanism.

[8.](#) IANA Considerations

[8.1.](#) FEC Type Name Space

This document uses a new FEC element types, number 0x82 will be requested as an allocation from the registry "FEC Type Name Space" for the Label Distribution Protocol (LDP [RFC5036](#)):

| Value | Hex | Name | Reference |
|-------|-------|---------------------|-----------|
| ----- | ----- | ----- | ----- |
| 130 | 0x82 | P2MP PW FEC Element | RFCxxxx |

[8.2.](#) LDP TLV TYPE

This document uses a new LDP TLV types, IANA already maintains a registry of name "TLV TYPE NAME SPACE" defined by [RFC5036](#). The following values are suggested for assignment:

| TLV type | Description |
|----------|------------------------|
| 0x0971 | Transport LSP TLV |
| 0x0703 | P2MP PW Capability TLV |

[8.3.](#) mLDP Opaque Value Element TLV TYPE

This document requires allocation of a new mLDP Opaque Value Element Type from the LDP MP Opaque Value Element type name space defined in [\[mLDP\]](#).

The following value is suggested for assignment:

| TLV type | Description |
|----------|-----------------------------|
| 0x3 | L2VPN-MCAST application TLV |

[9.](#) References

[9.1.](#) Normative References

- [RFC2119] Bradner. S, "Key words for use in RFCs to Indicate Requirement Levels", [RFC 2119](#), March, 1997.
- [RFC4447] "Transport of Layer 2 Frames Over MPLS", Martini, L., et al., [rfc4447](#) April 2006.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", [RFC 5036](#), October 2007.
- [RFC5003] C. Metz, L. Martini, F. Balus, J. Sugimoto, "Attachment Individual Identifier (AII) Types for Aggregation", [RFC5003](#), September 2007.
- [RFC5331] R. Aggarwal, Y. Rekhter, E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", [rfc5331](#), August 2008.
- [RFC5332] T. Eckert, E. Rosen, Ed., R. Aggarwal, Y. Rekhter, "MPLS Multicast Encapsulations", [rfc5332](#), August 2008.

- [mLDP] I. Minei, K. Kompella, I. Wijnands, B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", [draft-ietf-mpls-ldp-p2mp-06](#), Work In Progress, April 2009.

[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).", [rfc4875](#), May 2007.

[L3VPN-MCAST] R. Aggarwal, E. Rosen, T. Morin, Y. Rekhter,
"BGP Encodings and Procedures for Multicast in MPLS/BGP IP
VPNs", [draft-ietf-l3vpn-2547bis-mcast-bgp-08.txt](#),
Work in Progress, October 2009.

[RFC5561] B.Thomas, K.Raza, S.Aggarwal, R.Agarwal, JL. Le Roux,
"LDP Capabilities", [rfc5561](#), July 2009.

9.2. Informative References

[RFC3985] Stewart Bryant, et al., "PWE3 Architecture",
[RFC3985](#)

[RFC6074] E. Rosen,W. Luo,B. Davie,V. Radoaca "Provisioning,
Auto-Discovery, and Signaling in Layer 2 Virtual Private
Networks (L2VPNs)", [RFC6074](#), January 2011.

[P2MP-PW-REQ] F. Jounay, et. al, "Requirements for Point
to Multipoint Pseudowire",
[draft-ietf-pwe3-p2mp-pw-requirements-03.txt](#), Work in Progress,
August 2010.

[P2MP-LSP-PING] A. Farrel, S. Yasukawa, "Detecting Data Plane
Failures in Point-to-Multipoint Multiprotocol Label Switching
(MPLS) - Extensions to LSP Ping",
[draft-ietf-mpls-p2mp-lsp-ping-15.txt](#), Work In Progress,
January 2011.

10. Author's Addresses

Luca Martini
Cisco Systems, Inc.
9155 East Nichols Avenue, Suite 400
Englewood, CO, 80112
e-mail: lmartini@cisco.com

Sami Boutros
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
e-mail: sboutros@cisco.com

Siva Sivabalan
2000 Innovation Drive
Kanata, ONTARIO K2K 3E8
CANADA
e-mail: msiva@cisco.com

Maciek Konstantynowicz
Juniper Networks
UNITED KINGDOM
e-mail: maciek@juniper.net

Gianni Del Vecchio
Swisscom (Schweiz) AG
Zentweg 9
CH-3006 Bern
Switzerland
e-mail: Gianni.DelVecchio@swisscom.com

Thomas D. Nadeau
Huawei Technologies
2330 Central Expy
Santa Clara, CA 95050
USA
e-mail: thomas.nadeau@huawei.com

Internet Draft

[draft-ietf-pwe3-p2mp-pw-02.txt](#)

March 2, 2011

Frederic Jounay
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
FRANCE
Email: frederic.jounay@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
889 Bibo Road,
Shanghai, 201203
P.R.China
Email: lizhong.jin@zte.com.cn

Martin Vigoureux
Alcatel-Lucent
Route de Villejust
Nozay, 91620
France
Email: martin.vigoureux@alcatel-lucent.com

Laurent Ciavaglia
Alcatel-Lucent
Route de Villejust
Nozay, 91620
France
Email: Laurent.Ciavaglia@alcatel-lucent.com

Boutros, et al.

[Page 17]

Internet Draft [draft-ietf-pwe3-p2mp-pw-02.txt](#)

March 2, 2011

Simon Delord
Alcatel-Lucent

E-mail: simon.a.delord@team.telstra.com

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) 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.

Acknowledgments

The authors wish to acknowledge the contributions of Ali Sajassi.

Expiration Date: September 2011