Network Working Group
Internet Draft
Intended status: Standards Track
Expires: April 15, 2012

Siva Sivabalan (Ed.)
Sami Boutros (Ed.)
Luca Martini
Cisco Systems

Frederic Jounay
Philippe Niger
France Telecom

Maciek Konstantynowicz
Juniper

Thomas D. Nadeau
CA Technologies

Gianni Del Vecchio
Swisscom

Simon Delord
Telstra

Yuji Kamite
NTT Communications

Laurent Ciavaglia
Martin Vigoureux
Alcatel-Lucent

Lizhong Jin
ZTE

October 27, 2011

Signaling Root-Initiated Point-to-Multipoint Pseudowire using LDP
draft-ietf-pwe3-p2mp-pw-03.txt

Status of this Memo

Internet-Draft draft-ietf-pwe3-p2mp-pw-03.txt October 2011

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. 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 and bidirectional services. Requirements for P2MP PW are described in [P2MP-PW-REQ].

A P2MP PW can be constructed as either Single Segment (P2MP SS-PW) or Multi Segment (P2MP MS-PW) Pseudowire 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 (R-PE) to the Leaf-PE(s) (L-PEs) 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 T-PE1 and terminating at T-PE2 and T-PE3.

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.

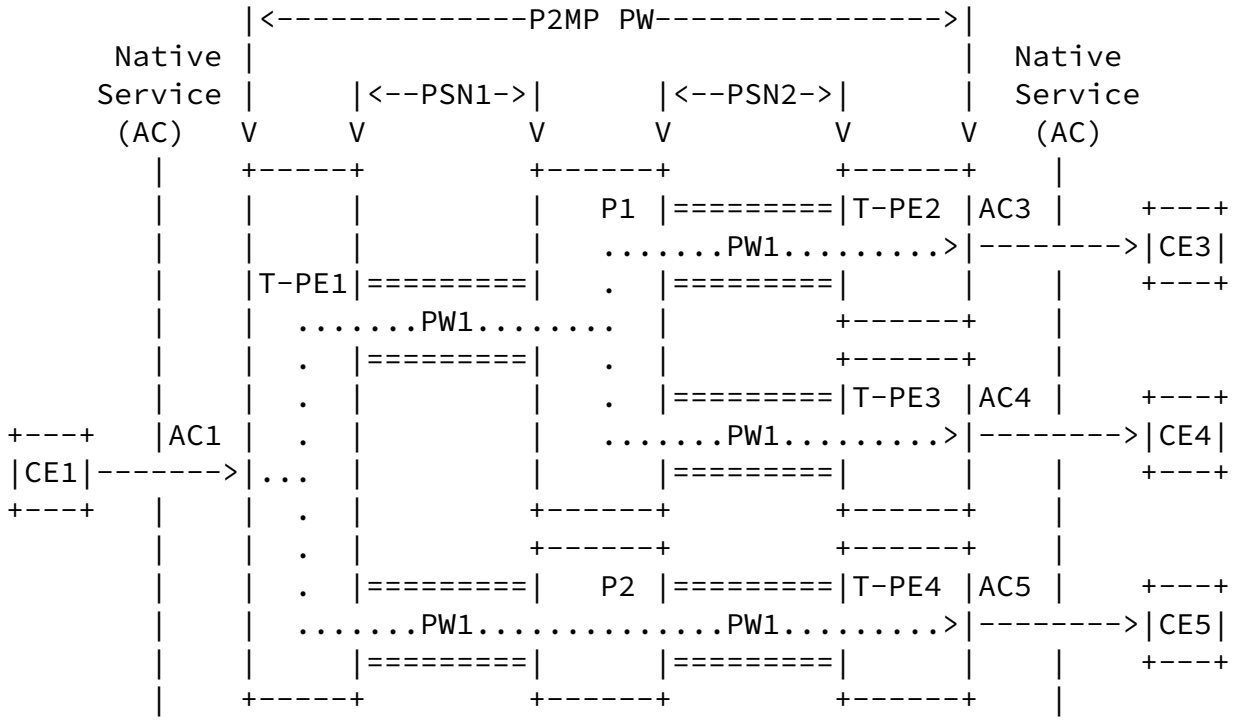

Figure 1: P2MP PW

As outlined in [P2MP-PW-REQ], even though the traffic flow from an R-PE to L-PEs is P2MP in nature, it may be desirable for any L-PE to send unidirectional P2P traffic destined only to the R-PE. The proposed mechanism takes such option into consideration.

A P2MP PW requires an MPLS LSP to carry the PW traffic, and the MPLS packets carried over the PW will be encapsulated according to the methods described in [RFC5332].

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 RFC-2119 [RFC2119].

2. 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.

3. Signaling 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].

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

P2MP PW requires that there is an active P2MP transport LSP set up between R-PE and L-PE(s). The procedure to set up the P2MP transport LSP is different depending on the protocol used (RSVP-TE or mLDP).

In case of mLDP, an L-PE can decide to join the P2MP LSP at any time, while in the case of RSVP-TE the P2MP transport LSP is set up by the R-PE, generally at the initial service provisioning time. It should be noted that local policy can override any decision to add or prune existing or new L-PE(s) to/from the tree. In any case, the PW setup can ignore these differences, and simply assume that the P2MP transport LSP is available when needed.

A P2MP PW signaling is initiated by the R-PE simply by sending a P2MP-PW LDP label mapping message to the L-PE(s) belonging to that P2MP PW. This label mapping message will contain the following:

1. A P2MP Upstream PW FEC element.
2. An Interface Parameters TLV, as described in [RFC4447].
3. A PW Grouping TLV as described in [RFC4447].
4. A Transport LSP TLV.
5. A label TLV for the upstream-assigned label used by R-PE for the traffic going from R-PE to L-PE(s).

The R-PE imposes the upstream-assigned label on the outbound packets sent over the P2MP-PW, and using this label an L-PE identifies the inbound packets arriving over the P2MP PW.

Additionally, the R-PE MAY send label mapping message(s) to one or more L-PE(s) to signal unidirectional P2P PW(s). The L-PE(s) can use such PW(s) to send traffic to the R-PE. This optional label mapping message will contain the following:

1. P2P Downstream PW FEC element.
2. A label TLV for the down-stream assigned label used by the corresponding L-PE to send traffic to the R-PE.

The LDP liberal label retention mode is used, and per requirements specified in [RFC5036], the label request message MUST also be supported.

The upstream-assigned label is allocated according to the rules in [RFC5331].

When an L-PE receives a PW Label Mapping Message, it MUST verify that the associated P2MP transport LSP is in place. If the associated P2MP transport LSP is not in place, and its type is LDP P2MP LSP, the L-PE SHOULD attempt to join the P2MP LSP. If the P2MP transport LSP is not in place, and its type is RSVP-TE P2MP LSP, the L-PE SHOULD wait till the P2MP transport LSP is signaled.

3.1. PW ingress to egress incompatibility issues

If an R-PE signals a PW with a pw type, CW mode, or interface parameters that a particular L-PE cannot accept, then the L-PE must 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.

3.2. P2MP PW FEC

[RFC4447] specifies two types of LDP FEC elements called "PWid FEC Element" and "Generalized PWid FEC Element" used to signal P2P PWs. We define two new types of FEC element called "P2MP Upstream PW FEC Element" and "P2P Downstream PW FEC Element". These FEC elements are associated with a mandatory upstream assigned label and an optional downstream assigned label respectively.

FEC type of the P2MP Upstream PW FEC Element is 0x82 (pending IANA allocation) and is encoded as follows:

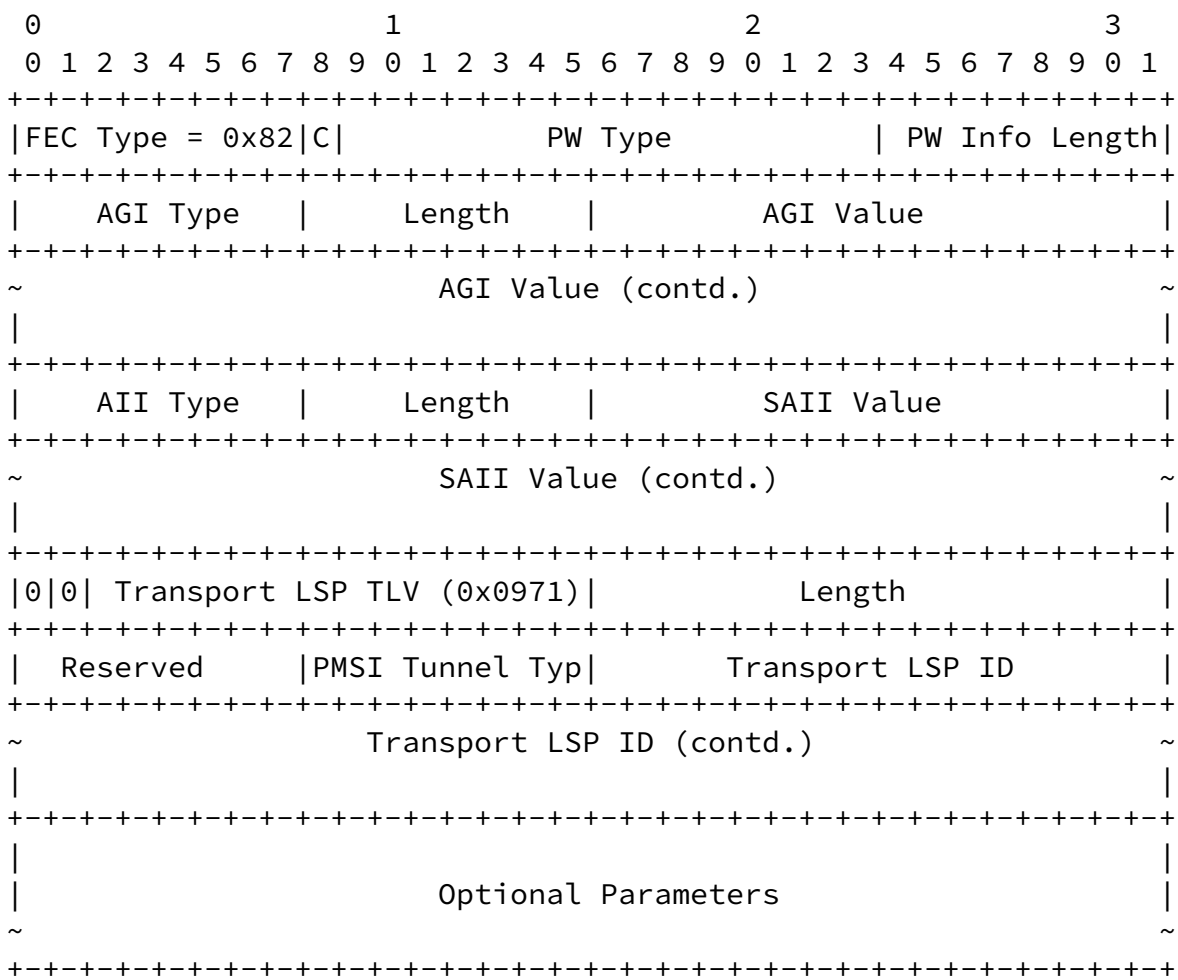

Figure 2: P2MP Upstream PW FEC Element

* 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 neither 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 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. 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 proposed mechanism uses two such TVLs: 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 an R-PE, each L-PE must check if its configured value(s) is less than or equal to the threshold value provided by the R-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 exactly match the values signaled by the R-PE.

Multicast traffic stream associated with a P2MP PW can be selective or inclusive. To support the former, this document defines a new optional Selective Tree Interface Parameter sub-TLV (type is pending IANA allocation) according to the format described in [RFC4447]. The value of the sub-TLV contains the source and the group for a given multicast tree as shown in Figure 3. This is similar to the way (S, G) is defined in [VPLS-MCAST]. Also, if a P2MP PW is associated with multiple selective trees, the corresponding label mapping message will carry more than one instances of this Sub-TLV. Furthermore, in the absence of this sub-TLV, the P2MP PW is associated with all multicast traffic stream originating from the root.

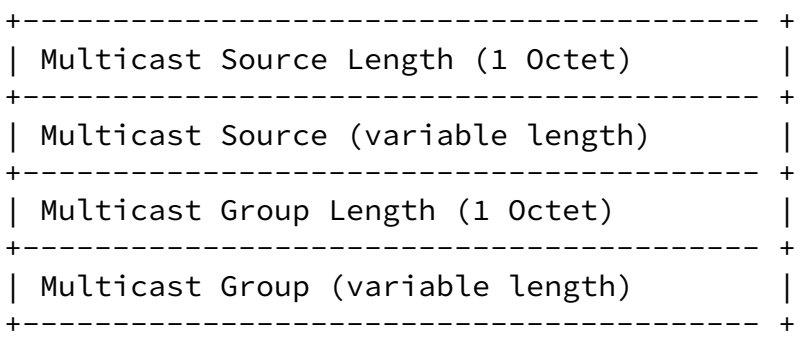

Figure 3: Selective Tree Interface Parameter Sub-TLV Value

The Multicast Source field contains the address of the multicast source. The Multicast Source field contains an IPv4 address or IPv6 address depending on whether the Multicast Source Length is 32 or 128. The Multicast Source Length can be set to 0 to indicate wildcard.

The Multicast Group field contains the address of the multicast group. The Multicast Group field contains an IPv4 address or IPv6 address depending on whether the Multicast Group Length is 32 or 128. The Multicast Group Length can be set to 0 to indicate wildcard.

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

The type of optional P2P Downstream PW FEC Element is 0x83 (pending IANA allocation), and is encoded as follows:

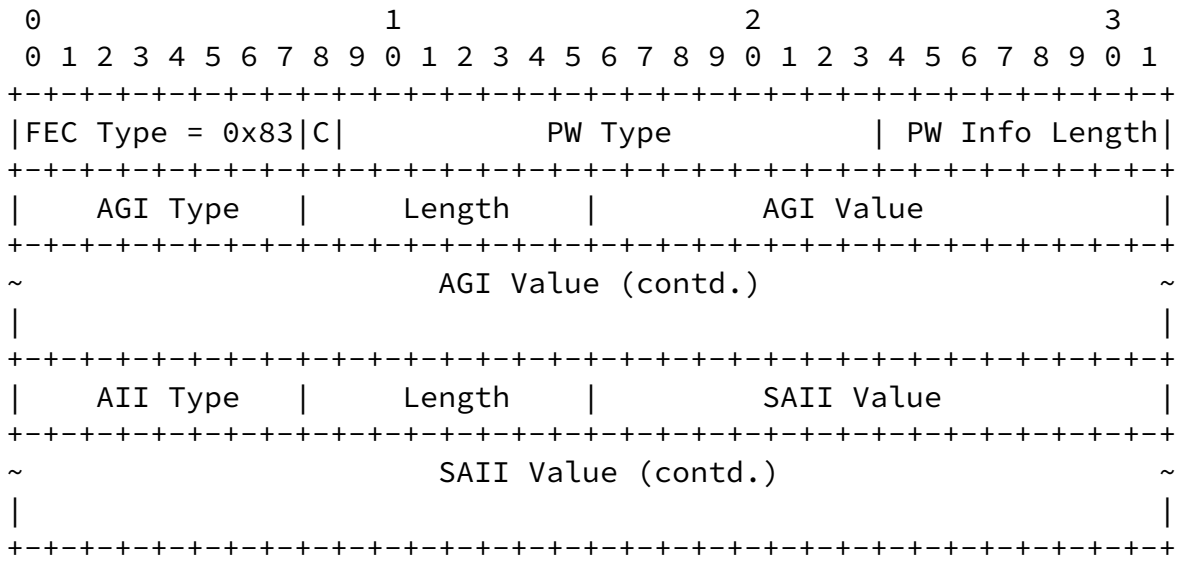

Figure 4: P2P Downstream PW FEC Element

The definition of the fields in the P2P Downstream PW FEC Element is the same as those of P2MP Upstream PW FEC Element.

3.3. 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].

3.4. Generic Label TLV

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]. For a given P2MP PW, a single upstream-assigned label is allocated by the R-PE, and is advertised to all the L-PEs using the Generic Label TLV in label mapping message containing the P2MP Upstream PW FEC element.

The R-PE MAY also allocate a unique label for each L-PE from which it intends to receive P2P traffic. Such a label is advertised to the L-PE using Generic Label TLV in label mapping message.

3.5. 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:

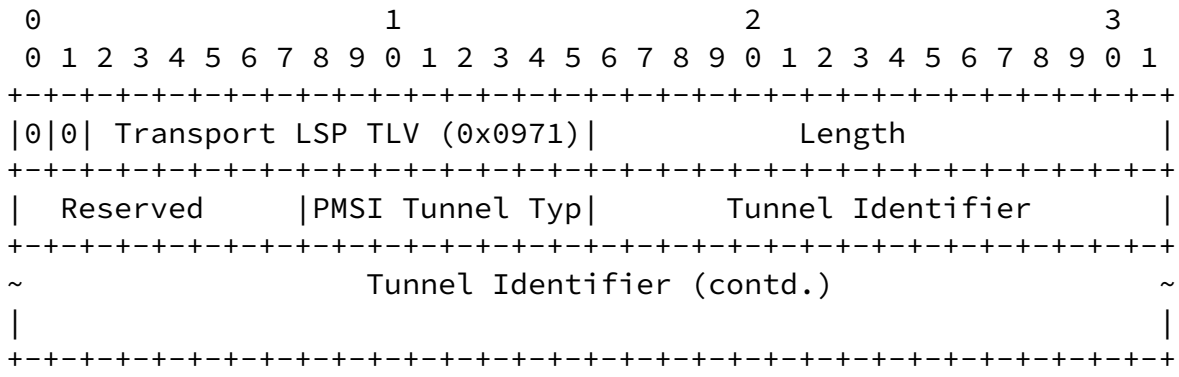

Figure 5: Transport LSP TLV

Note: TLV number pending IANA allocation.

* Reserved Flags:

  Reserved bits Must be set to 0 when transmitting the message, and ignored on receiving the message.

* PMSI Tunnel Type:

  The Transport LSP Type identifies the type of technology used to establish a transport LSP. The PMSI tunnel type is defined in [L3VPN-MCAST].

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:

The Tunnel containing the Transport LSP is identified by the Tunnel Identifier which is defined in [L3VPN-MCAST].

Transport LSP TLV MUST be present only in the Label Mapping Message. An R-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 that transport LSP. Similarly, an R-PE does not withdraw the labels when the corresponding transport LSP goes down. Furthermore, an L-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 PWs and all P2MP PWs will be sharing the same R-PE and L-PE(s).

In the case of LDP P2MP LSP, when an L-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 R-PE initiates the signaling of P2MP LSP.

4. 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]. An LSR having P2MP PW capability MUST recognize both P2MP Upstream FEC Element and P2P Downstream FEC Element in LDP Label Binding Message.

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

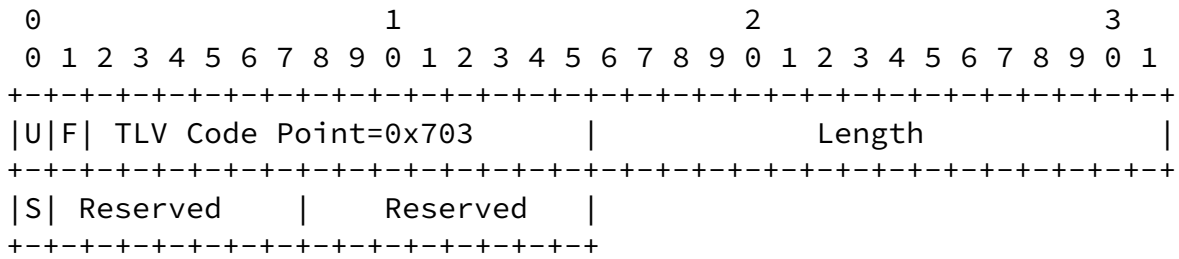

Figure 6: P2MP PW LDP Capability TLV

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 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.

5. 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 an L-PE successfully processes 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, pw status message is not necessary). PW status message sent from any L-PE to R-PE contains P2P Downstream PW FEC to identify the PW.

An R-PE also sends PW status to L-PE(s) to reflect its view of a P2MP PW state. Such PW status message contains P2MP Upstream PW FEC to identify the PW.

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].

6. Security Considerations

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

7. IANA Considerations

7.1. FEC Type Name Space

This document uses two new FEC element types, number 0x82 and 0x83 will be requested as an allocation from the registry "FEC Type Name Space" for the Label Distribution Protocol (RFC5036):

Value	Hex	Name	Reference
130	0x82	P2MP PW Upstream FEC Element	RFCxxxx
131	0x83	P2MP PW Downstream FEC Element	RFCxxxx

7.2. 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

7.3. 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

7.4. Selective Tree Interface Parameter sub-TLV Type

This document requires allocation of a sub-TLV from the registry "Pseudowire Interface Parameters Sub-TLV Type".

The following value is suggested for assignment:

TLV type	Description
0x0D	Selective Tree Interface Parameter.

8. Acknowledgment

Authors would like thank Kamran Raza, Andre Pelletier, and Parag Jain for their valuable suggestions.

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.

[VPLS-MCAST] R. Aggarwal, Y. Kamite, L. Fang, and Y. Rekhter, "Multicast in VPLS", draft-ietf-l2vpn-vpls-mcast-09.txt, Work In Progress, July 2011.

Author's Addresses

Siva Sivabalan
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario, K2K 3E8
Canada

Email: msiva@cisco.com

Sami Boutros

Cisco Systems, Inc.
3750 Cisco Way
San Jose, California 95134
USA
Email: sboutros@cisco.com

Luca Martini
Cisco Systems, Inc.
9155 East Nichols Avenue, Suite 400
Englewood, CO, 80112
United States
Email: lmartini@cisco.com

Maciek Konstantynowicz
Juniper Networks
UNITED KINGDOM
Email: maciek@juniper.net

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

Thomas D. Nadeau
CA Technologies
273 Corporate Drive
Portsmouth, NH 03801
USA
Email: thomas.nadeau@ca.com

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

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

Full Copyright Statement

Intellectual Property Statement

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the UETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of RFC 5378. No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under RFC 5378, shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.