

INTERNET-DRAFT
Intended Status: Standard Track

Sami Boutros(Ed.)
VMware

Updates: RFC7385

Siva Sivabalan(Ed.)
Cisco Systems

Expires: May 16, 2018

November 12, 2017

Signaling Root-Initiated Point-to-Multipoint Pseudowire using LDP
draft-ietf-pals-p2mp-pw-04.txt

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. This document updates RFC7385 by re-assigning reserved value 0xFF to be the wildcard transport tunnel type.

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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2017 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.

Table of Contents

1	Introduction	3
1.1	Terminology	4
2.	Signaling P2MP PW	4
2.1	PW ingress to egress incompatibility issues	6
2.2	P2MP PW FEC	6
2.2.1	P2MP PW Upstream FEC Element	6
2.2.2	P2P PW Downstream FEC Element	10
2.3	Typed Wildcard FEC Format for new FEC	10
2.4	Group ID usage	11
2.5	Generic Label TLV	11
3.	LDP Capability Negotiation	12
4.	P2MP PW Status	13
5	Security Considerations	13
6	Acknowledgment	13
7	IANA Considerations	14
7.1.	FEC Type Name Space	14
7.2.	LDP TLV Type	14
7.3.	mLDP Opaque Value Element TLV Type	14
7.4.	Selective Tree Interface Parameter sub-TLV Type	14
7.5.	Wildcard PMSI tunnel type	15
8	References	15
8.1.	Normative References	15
8.2.	Informative References	16
	Contributors	16
	Authors' Addresses	17

As outlined in [RFC7338], even though the traffic flow from a Root-PE (R-PE) to Leaf-PE(s) (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.

The P2MP PW requires an MPLS LSP to carry the PW traffic, and the MPLS packets carrying the PW upstream label will be encapsulated according to the methods described in [RFC5332].

1.1 Terminology

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].

FEC: Forwarding Equivalence Class

LDP: Label Distribution Protocol

mLDP: Label Distribution Protocol for P2MP/MP2MP 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 of MS-PW

TE: Traffic Engineering

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

L-PE: Leaf-PE - egress PE.

2. Signaling P2MP PW

In order to advertise labels as well as exchange PW related LDP messages, PEs must establish LDP sessions among themselves. 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 PSN LSP set up between R-PE and L-PE(s). Note that the procedure to set up the P2MP PSN LSP is different depending on the signaling protocol used (RSVP-TE or mLDP).

In case of mLDP, a Leaf-PE can decide to join the P2MP LSP at any time. In the case of RSVP-TE, the P2MP 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 join, add or prune existing or new L-PE(s) from the tree. In any case, the PW setup can ignore these differences, and simply assume that the P2MP PSN LSP is available when needed.

P2MP PW signaling is initiated by the R-PE which sends a separate P2MP-PW LDP label mapping message to each of the the L-PE(s) belonging to that P2MP PW. This label mapping message will contain the following:

1. A FEC TLV containing P2MP PW Upstream FEC element that includes Transport LSP sub TLV.
2. An Interface Parameters TLV, as described in [RFC4447bis].
3. A PW Grouping TLV, as described in [RFC4447bis].
4. 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 PW 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 PW Downstream 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 MUST be 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 the associated P2MP PSN LSP is in place. If the associated P2MP PSN LSP is not in place, and its type is LDP P2MP LSP, the L-PE MUST attempt to join the P2MP LSP associated with the P2MP PW. If the associated P2MP PSN 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. If an L-PE fails to join the P2MP PSN LSP, that L-PE MUST not enable the PW, and MUST notify the user. In this case, a PW status message with status code of 0x00000008 (Local PSN-facing PW (ingress) Receive Fault) MUST also be sent to the R-PE.

2.1 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 with status code 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.

2.2 P2MP PW FEC

[RFC4447bis] specifies two types of LDP FEC elements called "Pwid FEC Element" and "Generalized Pwid FEC Element" used to signal P2P PWs. This document defines two new types of FEC elements called "P2MP PW Upstream FEC Element" and "P2P PW Downstream FEC Element". These FEC elements are associated with a mandatory upstream assigned label and an optional downstream assigned label respectively.

2.2.1 P2MP PW Upstream FEC Element

A new FEC type for the P2MP PW Upstream FEC Element is encoded as follows:

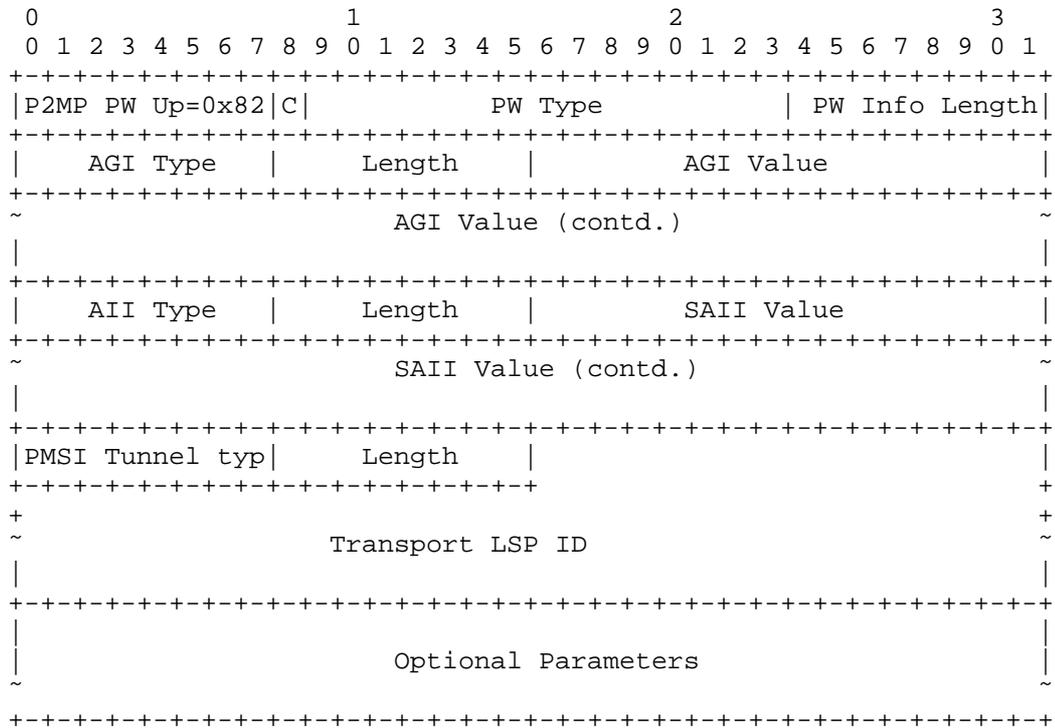


Figure 2: P2MP PW Upstream FEC Element

* P2MP PW Up:

8 bits representation for the P2MP PW Upstream FEC type.

* PW Type:

15 bits representation of PW type as specified in [RFC4447bis].

* 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, SAI, PMSI Tunnel info, 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, SAI, etc.) present, nor any

interface parameters TLVs. Alternatively, typed wildcard FEC described in section 3.3, can be used to achieve the same or to have better filtering.

* 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 [RFC4447bis].

* 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.

* PMSI Tunnel info

PMSI Tunnel info is the combination of PMSI Tunnel Type, Length and Transport LSP ID.

A P2MP PW MUST be associated with a transport LSP which can be established using RSVP-TE or mLDP.

* PMSI Tunnel Type:

The PMSI tunnel type is defined in [RFC6514].

When the type is set to mLDP P2MP LSP, the Tunnel Identifier is a P2MP FEC Element as defined in [RFC6388]. A new mLDP Opaque Value Element type for L2VPN-MCAST application as specified in the IANA considerations MUST be used.

* Transport LSP ID: This is the Tunnel Identifier which is defined in [RFC6514].

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 in which case P2MP PWs will be sharing the same R-PE and L-PE(s). An R-PE may also have many P2MP PWs with disjoint L-PE sets.

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.

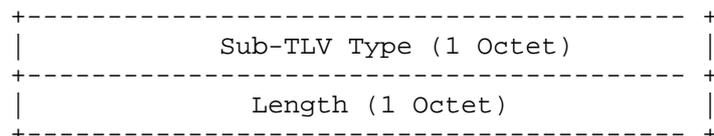
In the case of RSVP-TE P2MP LSP, only the R-PE initiates the signaling of P2MP LSP.

* 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 TLVs: Interface Parameters TLV, and Group ID TLV.

The Interface Parameters TLV and Group ID TLV specified in [RFC4447bis] can also be used in conjunction with P2MP PW FEC in a label message. For Group ID TLV, the sender and receiver of these TLVs should follow the same rules and procedures specified in [RFC4447bis]. For Interface Parameters TLV, the procedure differs from the one specified in [RFC4447bis] due to specifics of P2MP connectivity. When the interface parameters are signaled by a 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, as described in the IANA considerations and according to the format described in [RFC4447bis]. The value of the sub-TLV contains the source and the group for a given multicast tree as shown in Figure 3. Also, if a P2MP PW is associated with multiple selective trees, the corresponding label mapping message will carry more than one instance 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.



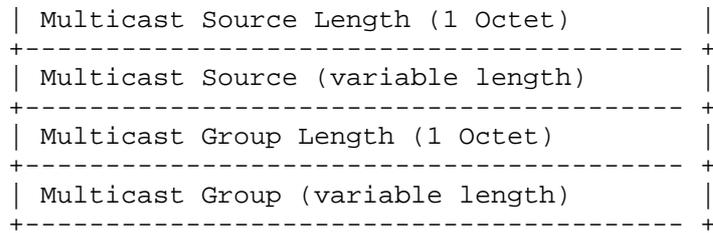


Figure 3: Selective Tree Interface Parameter Sub-TLV Value

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.

2.2.2 P2P PW Downstream FEC Element

The optional P2P PW Downstream FEC Element is encoded as follows:

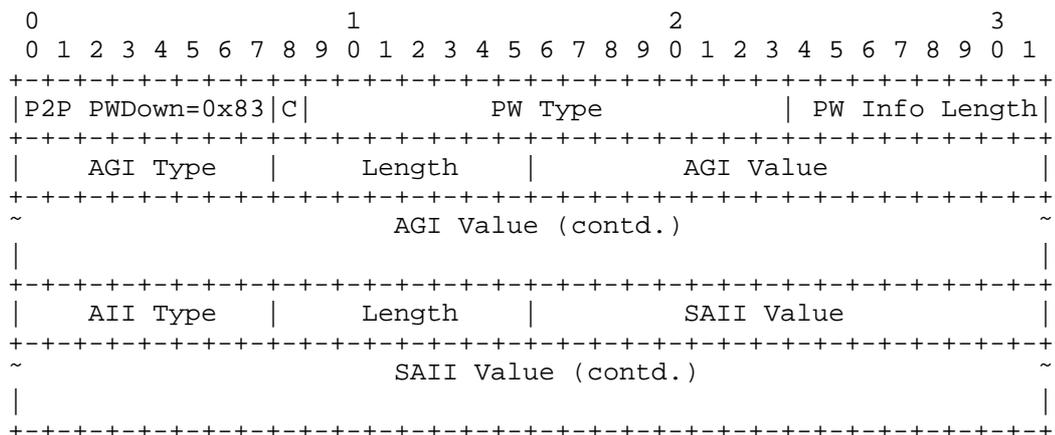


Figure 4: P2P PW Downstream FEC Element

The definition of the fields in the P2P PW Downstream FEC Element is the same as those of P2MP PW Upstream FEC Element shown in Figure 2.

2.3 Typed Wildcard FEC Format for new FEC

[RFC5918] defines the general notion of a "Typed Wildcard" FEC Element, and requires FEC designer to specify a typed wildcard FEC element for newly defined FEC element types. This document defines two new FEC elements, "P2MP PW Upstream" and "P2P PW Downstream" FEC

element, and hence requires us to define their Typed Wildcard format.

[RFC6667] defines Typed Wildcard FEC element format for other PW FEC Element types (PWid and Gen. PWid FEC Element) in section 2 as follows:

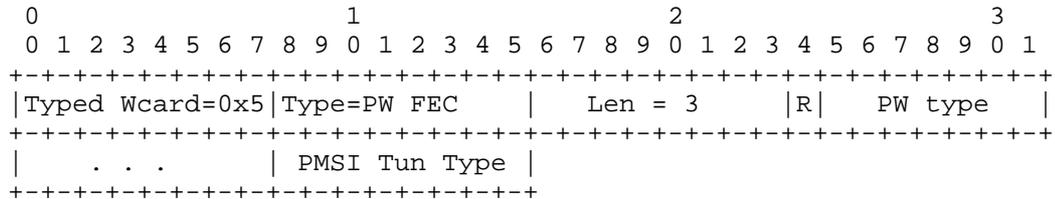


Figure 5: Typed Wildcard Format for P2MP PW FEC Elements

[RFC6667] specifies that "Type" field can be either "PWid" (0x80) or "Generalized PWid" (0x81) FEC element type. This document reuses the existing typed wildcard format as specified in [RFC6667] and illustrated in Figure 5 and extends the definition of "Type" field to also include "P2MP PW Upstream" and "P2P PW Downstream" FEC element types. This document adds an additional field "PMSI Tun Type". This document reserves PMSI tunnel Type 0xFF to mean "wildcard" transport tunnel type. The PMSI tunnel Type field only applies to Typed wildcard P2MP PW Upstream FEC and MUST be set to "wildcard" for "P2P PW Downstream FEC" typed wildcard element.

2.4 Group ID usage

The Grouping TLV as defined in [RFC4447bis] 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 [RFC4447bis].

2.5 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 [RFC4447bis]. For a given P2MP PW, a single upstream-assigned label is allocated by the R-PE, and is advertised to all L-PEs using the Generic Label TLV in label mapping message containing the P2MP PW Upstream FEC element.

The R-PE can 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 and P2P PW Downstream FEC in label mapping message.

3. 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] to advertise the LDP "P2MP PW 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 supports dynamic capability advertisement, then the P2MP PW Capability TLV MUST be included in the LDP Initialization message during the session establishment. An LSR having P2MP PW capability MUST recognize both P2MP PW Upstream FEC Element and P2P PW Downstream FEC Element in LDP label messages.

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

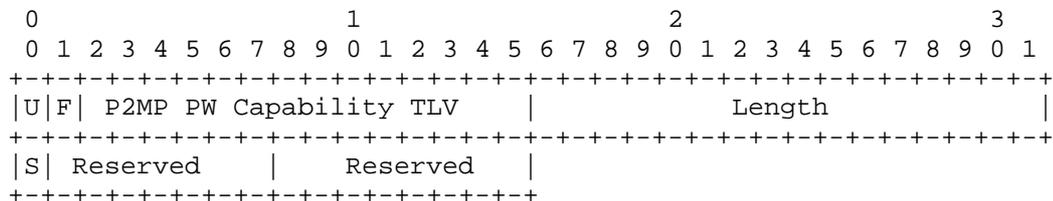


Figure 7: LDP P2MP PW Capability TLV

* U-bit:

SHOULD be 1 (ignore if not understood).

* F-bit:

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

* P2MP PW Capability 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.

* Length:

MUST be set to 2 (octet).

4. P2MP PW Status

In order to support the proposed mechanism, an LSR MUST be capable of handling PW status. As such, PW status negotiation procedure described in [RFC4447bis] is not applicable to P2MP PW. An LSR MUST NOT claim to be P2MP PW capable by sending a LDP P2MP PW Capability TLV if it is not also capable of handling PW status.

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 [RFC4447bis] to report the PW status. If no PW status notification is required, then no PW status notification is sent (for example if the P2MP PW is established and operational with a status code of Success (0x00000000), a PW status message is not necessary). A PW status message sent from an L-PE to R-PE MUST contain the P2P PW Downstream 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 PW Upstream 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 [RFC6425].

5 Security Considerations

In general the security measures described in [RFC4447bis] are adequate for this protocol. However the use of MD5 as the method of securing an LDP control plane is no longer considered adequately secure. Implementations should be written in such a way that they can migrate to a more secure cryptographic hash function when the next authentication method to be used in the LDP might not be simple hash based authentication algorithm.

6 Acknowledgment

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

7 IANA Considerations

7.1. FEC Type Name Space

This document uses two new FEC element types, number 0x82 and 0x83 are suggested for assignment from the registry "FEC Type Name Space" for the Label Distribution Protocol (LDP RFC5036):

Value	Hex	Name	Reference
-----	-----	-----	-----
130	0x82	P2MP PW Upstream FEC Element	RFCxxxx
131	0x83	P2P 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:
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 "LDP MP Opaque Value Element basic type" name space defined in [RFC6388].

The following value is suggested for assignment:

```

TLV type  Description
13        L2VPN-MCAST application TLV

```

Length: 4

Value: A 32-bit integer, unique in the context of the root, as identified by the root's address.

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" defined in [RFC4446].

The following value is suggested for assignment:

```

TLV type  Description

```

0x1b Selective Tree Interface Parameter.

7.5. WildCard PMSI tunnel type

This document requests that IANA modify the following entry in the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry within the "Border Gateway Protocol (BGP) Parameters" namespace previously assigned by RFC7385 as "reserved".

Value	Meaning	Reference
0xFF	wildcard transport tunnel type	[This document]

8 References

8.1. Normative References

- [RFC2119] Bradner, S, "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March, 1997.
- [RFC4447bis] "Pseudowire Setup and Maintenance using the Label Distribution Protocol", Martini, L., et al., draft-ietf-pals-rfc4447bis-05.txt, July 2016.
- [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", RFC 5331, August 2008.
- [RFC5332] T. Eckert, E. Rosen, Ed., R. Aggarwal, Y. Rekhter, "MPLS Multicast Encapsulations", RFC 5332, August 2008.
- [RFC6388] I. Minei, K. Kompella, I. Wijnands, B. Thomas, "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.
- [RFC6514] R. Aggarwal, E. Rosen, T. Morin, Y. Rekhter, "BGP Encodings

and Procedures for Multicast in MPLS/BGP IP VPNs", RFC6514, February 2012.

[RFC5561] B.Thomas, K.Raza, S.Aggarwal, R.Agarwal, JL. Le Roux, "LDP Capabilities", RFC 5561, July 2009.

[RFC5918] R. Asati, I. Minei, and B. Thomas, "LDP Typed Wildcard Forwarding Equivalence Class", RFC 5918, August 2010.

[RFC6667] K. Raza, S. Boutros, and C. Pignataro, "LDP Typed Wildcard FEC for Pwid and Generalized Pwid FEC Elements", RFC 6667, July 2012.

8.2. 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.

[RFC7338] F. Jounay, et. al, "Requirements for Point to Multipoint Pseudowire", RFC7338, September 2014.

[RFC6425] A. Farrel, S. Yasukawa, "Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS)- Extensions to LSP Ping", RFC6425, November 2011.

Contributors

The following co-authors have also contributed to this document:

Luca Martini
Cisco Systems, Inc.
Email: lmartini@cisco.com

Maciek Konstantynowicz
Cisco Systems, Inc.
e-mail: maciek@cisco.com

Gianni Del Vecchio
Swisscom
Email: Gianni.DelVecchio@swisscom.com

Thomas D. Nadeau
Brocade

Email: tnadeau@lucidvision.com

Frederic Jounay
Orange CH
Email: Frederic.Jounay@salt.ch

Philippe Niger
Orange CH
Email: philippe.niger@orange.com

Yuji Kamite
NTT Communications Corporation
Email: y.kamite@ntt.com

Lizhong Jin
Email: lizho.jin@gmail.com

Martin Vigoureux
Nokia
Email: martin.vigoureux@nokia.com

Laurent Ciavaglia
Nokia
Email: laurent.ciavaglia@nokia.com

Simon Delord
Telstra
E-mail: simon.delord@gmail.com

Kamran Raza
Cisco Systems
Email: skraza@cisco.com

Authors' Addresses

Sami Boutros
VMware Inc.
Email: sboutros@vmware.com

Siva Sivabalan
Cisco Systems, Inc.
Email: msiva@cisco.com