

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 06, 2013

M. Chen
W. Cao
Huawei Technologies Co., Ltd
A. Takacs
Ericsson
P. Pan
Infinera
June 04, 2013

LDP extensions for Pseudowire Binding to LSP Tunnels
draft-ietf-pwe3-mpls-tp-pw-over-bidir-lsp-01.txt

Abstract

Many transport services require that user traffic, in the forms of Pseudowires (PW), to be delivered on a single co-routed bidirectional LSP or two LSPs that share the same routes. In addition, the user traffic may traverse through multiple transport networks.

This document specifies an optional extension in LDP that enable the binding between PWs and the underlying LSPs.

Requirements Language

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](#)].

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on December 06, 2013.

Copyright Notice

Copyright (c) 2013 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	2
2.	LDP Extensions	4
2.1.	PSN Tunnel Binding TLV	5
2.1.1.	PSN Tunnel Sub-TLV	6
3.	Theory of Operation	8
4.	PSN Binding Operation for SS-PW	8
5.	PSN Binding Operation for MS-PW	11
6.	Security Considerations	12
7.	IANA Considerations	12
7.1.	LDP TLV Types	12
7.1.1.	PSN Tunnel Sub-TLVs	13
7.2.	LDP Status Codes	13
8.	Acknowledgements	13
9.	References	13
9.1.	Normative References	13
9.2.	Informative References	13
	Authors' Addresses	14

[1.](#) Introduction

Pseudowire (PW) Emulation Edge-to-Edge (PWE3) [[RFC3985](#)] is a mechanism to emulate layer 2 services, such as Ethernet Point-to-Point (P2P) circuits. Such services are emulated between two Attachment Circuits (ACs) and the PW encapsulated layer 2 service payload is carried through Packet Switching Network (PSN) tunnels between Provider Edges (PEs). PWE3 typically uses Label Distribution Protocol (LDP) [[RFC5036](#)] or Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) [[RFC3209](#)] LSPs as PSN tunnels. The PEs select and bind the Pseudowires to PSN tunnels independently. Today, there is no protocol-based provisioning mechanism to associate PWs to PSN tunnels.

PW-to-PSN Tunnel binding has become increasingly common and important in many deployment scenarios. For instance, when connecting two remotely located sites, such as data centers, over the backbone, each site may deploy a high-performance router or switch to aggregate thousands of Ethernet VLAN flows. The aggregating routers and switches are interconnected via one or multiple MPLS/GMPLS LSPs, which may traverse through different routes or networks. Further, each Ethernet flow is offered to the customers as a bidirectional circuits with certain SLA attributes, such as bandwidth and latency. Hence, it's important for the operators to map the forwarding and reverse-direction traffic from an Ethernet circuit to a single bidirectional co-routed LSP or two LSPs that share the same route.

The requirement for explicit control of PW-to-LSP mapping has been described in [Section 5.3.2](#) ("Support for Explicit Control of PW-to-LSP Binding") of [[RFC6373](#)]. The following figure (Figure 1) provides the illustration.

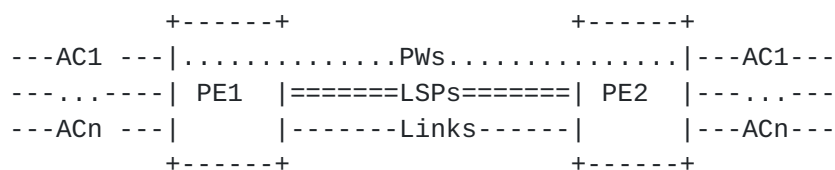


Figure 1: Explicit PW-to-LSP binding scenario

There are two PEs (PE1 and PE2) connected through multiple parallel links that may be on different fibers. Each link is managed and controlled as a bi-directional LSP. At each PE, there are a large number of bi-directional user flows from multiple Ethernet interfaces. Each user flow uses PWs to carry traffic on forwarding and reverse direction. The operators need to make sure that the user flows (that is, the PW-pairs) to be carried on the same fiber (or, bidirectional LSP).

As mentioned above, there are a number of reasons behind this requirement. First, due to delay and latency constraints, traffic going over different fibers may require large amount of expensive buffer memory to compensate for the differential delay at the headend nodes. Further, the operators may apply different protection mechanisms on different parts of the network. As such, for optimal traffic management, traffic belongs to a particular user should traverse over the same fiber. That implies that both forwarding and reserve direction PWs that belong to the same user flow need to be mapped to the same co-routed bi-directional LSP or two LSPs with the same route.

Figure 2 illustrates a scenario where PW-LSP binding is not applied.

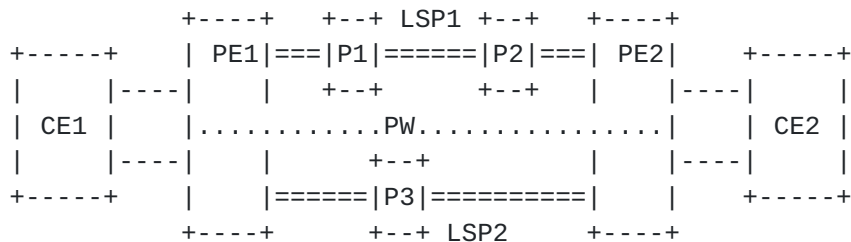


Figure 2: Inconsistent SS-PW to LSP binding scenario

LSP1 and LSP2 are two bidirectional connections on diverse paths. The operator is to deliver a bi-directional flow between PE1 and PE2. Using the existing mechanisms, it's possible that PE1 may select LSP1 (PE1-P1-P2-PE2) as the PSN tunnel for traffic from PE1 to PE2, while selecting LSP2 (PE1-P3-PE2) as the PSN tunnel for traffic from PE2 to PE1.

Consequently, the user traffic is delivered over two disjoint LSPs that may have very different service attributes in terms of latency and protection. This may not be acceptable as a reliable and effective transport service to the customers.

The similar problems may also exist in multi-segment PWs (MS-PWs), where user traffic on a particular PW may hop over different networks on forward and reverse directions.

One way to solve this problem is by introducing manual configuration at each PE to bind the PWs to the underlying PSN tunnels. However, this is prone to configuration errors and won't scale.

In this documentation, we will introduce an automatic solution by extending FEC 128/129 PW based on [\[RFC4447\]](#).

2. LDP Extensions

This document defines a new optional TLV, PSN Tunnel Binding TLV, to communicate tunnel/LSPs selection and binding requests between PEs. The TLV carries PW's binding profile and provides explicit or implicit information for the underlying PSN tunnel binding operation.

The binding operation applies in both single-segment (SS) and multi-segment (MS) scenarios.

The extension supports two types of binding requests:


```

+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|C|S|T|      MUST be Zero      |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```


The flags are defined as the following:

C (Congruent path) bit: This informs the remote T-PE/S-PEs about the properties of the underlying LSPs. When set, the remote T-PE/S-PEs need to select LSP with the same route as the PSN tunnel. If there is no such tunnel available, the node may trigger the remote T-PE/S-PEs to establish a new LSP.

S (Strict) bit: This instructs the PEs with respect to the handling of the underlying LSPs. When set, the remote PE MUST use the tunnel/LSP specified in the PSN Tunnel Sub-TLV as the PSN tunnel on the reverse direction of the PW, or the PW will fail to be established.

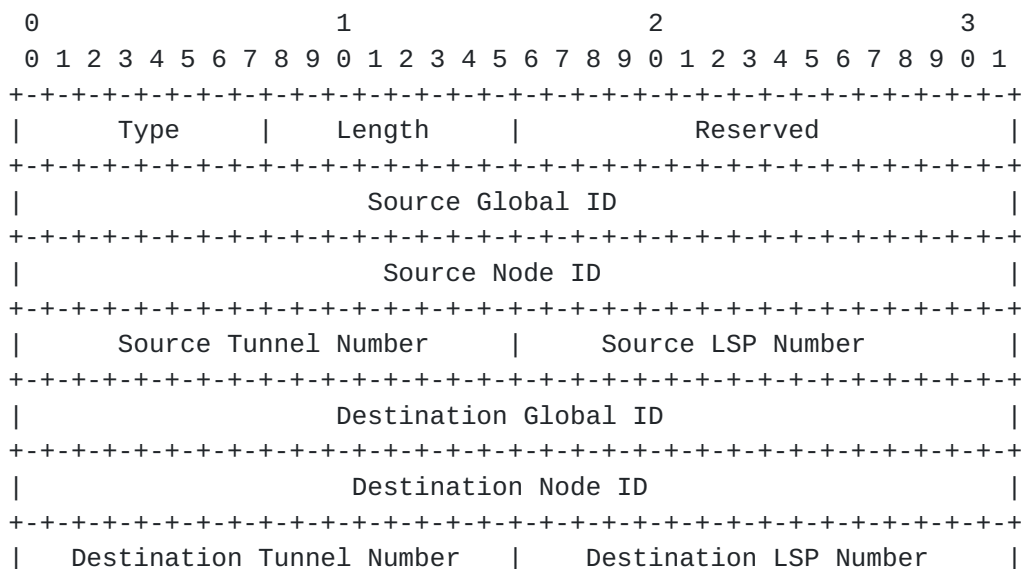
T (Tunnel Representation) bit: This indicates the format of the LSP tunnels. When the bit is set, the tunnel uses the tunnel information to identify itself, and the LSP Number fields in the PSN Tunnel sub-TLV ([Section 2.1.1](#)) MUST be set to zero. Otherwise, both tunnel and LSP information of the PSN tunnel are required. The default is set. The motivation for the T-bit is to support the MPLS protection operation where the LSP Number fields may be ignored.

C-bit and S-bit are mutually exclusive from each other, and cannot be set in the same message.

2.1.1. PSN Tunnel Sub-TLV

PSN Tunnel Sub-TLVs are designed for inclusion in the PSN Tunnel Binding TLV to specify the tunnel/LSPs to which a PW is required to bind.

Two sub-TLVs are defined: the IPv4 and IPv6 Tunnel sub-TLVs.



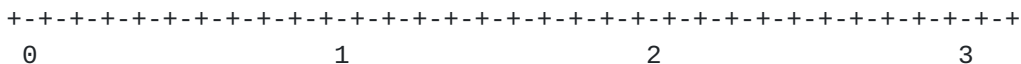


Figure 4: IPv4 PSN Tunnel sub-TLV format

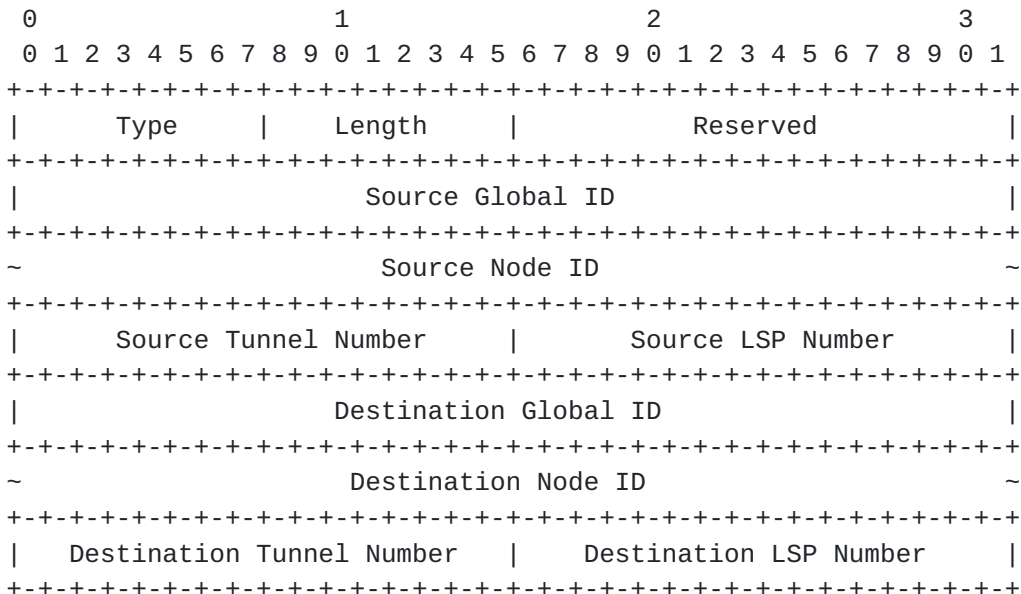


Figure 5: IPv6 PSN Tunnel sub-TLV format

The definition of Source and Destination Global/Node IDs and Tunnel/LSP Numbers are derived from [\[RFC6370\]](#). This is to describe the underlying LSPs. Note that the LSPs in this notation are globally unique.

As defined in [Section 4.6.1.2](#) and [Section 4.6.2.2 of \[RFC3209\]](#), the "Tunnel endpoint address" is mapped to Destination Node ID, and "Extended Tunnel ID" is mapped to Source Node ID. Both IDs can be IPv6 addresses.

A PSN Tunnel sub-TLV could be used to either identify a tunnel or a specific LSP. The T-bit in the Flag field defines the distinction as such that, when the T-bit is set, the Source/Destination LSP Number fields MUST be zero and ignored during processing. Otherwise, both Source/Destination LSP Number fields MUST have the actual LSP IDs of specific LSPs.

Each PSN Tunnel Binding TLV can only have one such sub-TLV.

3. Theory of Operation

During PW setup, the PEs may select desired forwarding tunnels/LSPs, and inform the remote T-PE/S-PEs about the desired reverse tunnels/LSPs.

Specifically, to set up a PW (or PW Segment), a PE may select a candidate tunnel/LSP to act as the PSN tunnel. If none is available or satisfies the constraints, the PE will trigger and establish a new tunnel/LSP. The selected tunnel/LSP information is carried in the PSN Tunnel Binding TLV and sent with the Label Mapping message to the target PE.

Upon the reception of the Label Mapping message, the receiving PE will process the PSN Tunnel Binding TLV, determine whether it can accept the suggested tunnel/LSP or to find the reverse tunnel/LSP that meets the request, and respond with a Label Mapping message, which contains the corresponding PSN Tunnel Binding TLV.

It is possible that two PEs may request PSN binding to the same PW or PW segment over different tunnels/LSPs at the same time. There may cause collisions of tunnel/LSPs selection as both PEs assume the active role.

As defined in ([Section 7.2.1](#), [[RFC6073](#)]), each PE may be generally categorized into active and passive roles:

1. Active PE: the PE which initiates the selection of the tunnel/LSPs and informs the remote PE;
2. Passive PE: the PE which obeys the active PE's suggestion.

In the remaining of this document, we will elaborate the operation for SS-PW and MS-PW:

1. SS-PW: In this scenario, both PEs for a particular PW may assume the active roles.
2. MS-PW: One PE is active, while the other is passive. The PWs are setup using FEC 129.

4. PSN Binding Operation for SS-PW

As illustrated in Figure-5, both PEs (say, PE1 and PE2) of a PW may independently initiate the setup. To perform PSN binding, the Label Mapping messages MUST carry a PSN Tunnel Binding TLV, and the PSN Tunnel sub-TLV MUST contains the desired tunnel/LSPs of the sender.

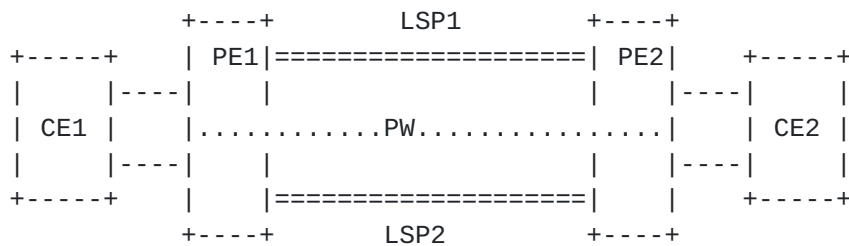


Figure 6: PSN binding operation in SS-PW environment

As outlined previously, there are two types of binding request: congruent and strict.

In strict binding, a PE (e.g., PE1) will mandate the other PE (e.g., PE2) to use a specified tunnel/LSP (e.g. LSP1) as the PSN tunnel on the reverse direction. In the PSN Tunnel Binding TLV, the S-bit MUST be set, the C-bit MUST be reset, and the Source and Destination IDs/Numbers MUST be filled.

On receive, if the S-bit is set, other than following the processing procedure defined in [Section 5.3.3 of \[RFC4447\]](#), the receiving PE (i.e. PE2) needs to determine whether to accept the indicated tunnel/LSP in PSN Tunnel Sub-TLV.

If the receiving PE (PE2) is also an active PE, and may have initiated the PSN binding requests to the other PE (PE1), if the received PSN tunnel/LSP is the same as it has been sent in the Label Mapping message by PE2, then the signaling has converged on a mutually agreed Tunnel/LSP. The binding operation is completed.

Otherwise, the receiving PE (PE2) MUST compare its own Node ID against the received Source Node ID. If it is numerically lower, the PE (PE2) will reply a Label Mapping message to complete the PW setup and confirm the binding request. The PSN Tunnel Binding TLV in the message MUST contain the same Source and Destination IDs/Numbers as in the received binding request, in the appropriate order. On the other hand, if the receiving PE (PE2) has a Node ID that is numerically higher than the Source Node ID carried in the PSN Tunnel Binding TLV, it MUST reply a Label Release message with status code set to "Reject to use the suggested tunnel/LSPs" and the received PSN Tunnel Binding TLV, and the PW will not be established.

To support congruent binding, the receiving PE can select the appropriated PSN tunnel/LSP for the reverse direction of the PW, so long as the forwarding and reverse PSNs share the same route.

Initially, a PE (PE1) sends a Label Mapping message to the remote PE (PE2) with the PSN Tunnel Binding TLV, with C-bit set, S-bit reset,

and the appropriate Source and Destination IDs/Numbers. In case of unidirectional LSPs, the PSN Tunnel Binding TLV may only contain the Source IDs/Numbers, the Destination IDs/Numbers are set to zero and left for PE2 to fill when responding the Label Mapping message.

On receive, since PE2 is also an active PE, and may have initiated the PSN binding requests to the other PE (PE1), if the received PSN tunnel/LSP has the same route as the one that has been sent in the Label Mapping message to PE1, then the signaling has converged. The binding operation is completed.

Otherwise, it needs to compare its own Node ID against the received Source Node ID. If it is numerically lower, PE2 needs to find/establish a tunnel/LSP that meets the congruent constraint, and reply a Label Mapping message with a PSN Binding TLV that contains the Source and Destination IDs/Numbers in the appropriate order. On the other hand, if the receiving PE (PE2) has a Node ID that is numerically higher than the Source Node ID carried in the PSN Tunnel Binding TLV, it MUST reply a Label Release message with status code set to "Reject to use the suggested tunnel/LSPs" and the received PSN Tunnel Binding TLV.

In both strict and congruent bindings, if T-bit is set, the LSP Number field MUST be set to zero. Otherwise, the field MUST contain the actual LSP number for the related PSN LSP.

After a PW established, the operators may choose to move the PWs from the current tunnel/LSPs to other tunnel/LSPs. Or, the underlying PSN tunnel is broken due to network failure. In this scenario, a new Label Mapping message MUST be sent to update the changes. Note that when T-bit is set, the working LSP broken will not trigger to update the changes if there are protection LSPs.

The message may carry a new PSN Tunnel Binding TLV, which contains the new Source and Destination Numbers/IDs. The handling of the new message should be identical to what has been described in this section.

However, if the new Label Binding message does not contain the PSN Tunnel Binding TLV, it declares the removal of any congruent/strict constraints. The PEs may not map the PW to the underlying PSN on purpose, the current independent PW to PSN binding will be used.

Further, as an implementation option, the PEs should not remove the traffic from an operational PW, until the completion of the underlying PSN tunnel/LSP changes.

5. PSN Binding Operation for MS-PW

MS-PW uses FEC 129 for PW setup. We refer the operation to Figure-6.

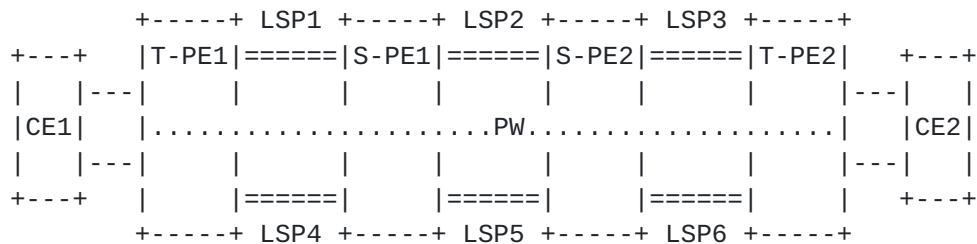


Figure 7: PSN binding operation in MS-PW environment

When an active PE (that is, T-PE1) starts to signal a MS-PW, a PSN Tunnel Binding TLV MUST be carried in the Label Mapping message and sent to the adjacent S-PE (that is, S-PE1). The PSN Tunnel Binding TLV includes the PSN Tunnel sub-TLV that carries the desired tunnel/LSP of T-PE1's.

For strict binding, the initiating PE MUST set the S-bit, reset the C-bit and indicate the binding tunnel/LSP to the next-hop S-PE.

When S-PE1 receives the Label Mapping message, S-PE1 needs to determine if the signaling is for forward or reverse direction, as defined in Section 6.2.3 of [[I-D.ietf-pwe3-dynamic-ms-pw](#)].

If the Label Mapping message is for forward direction, and S-PE1 accepts the requested tunnel/LSPs from T-PE1, S-PE1 must save the tunnel/LSP information for reverse-direction processing later on. If the PSN binding request is not acceptable, S-PE1 MUST reply a Label Release Message to the upstream PE (T-PE1) with Status Code set to "Reject to use the suggested tunnel/LSPs".

Otherwise, S-PE1 relays the Label Mapping message to the next S-PE (that is, S-PE2), with the PSN Tunnel sub-TLV carrying the information of the new PSN tunnel/LSPs selected by S-PE1. S-PE2 and subsequent S-PEs will repeat the same operation until the Label Mapping message reaches to the remote T-PE (that is, T-PE2).

If T-PE2 agrees with the requested tunnel/LSPs, it will reply a Label Mapping message to initiate to the binding process on the reverse direction. The Label Mapping message contains the received PSN Tunnel Binding TLV for confirmation purposes.

When its upstream S-PE (S-PE2) receives the Label Mapping message, the S-PE relays the Label Mapping message to its upstream adjacent S-PE (S-PE1), with the previously saved PSN tunnel/LSP information in the PSN Tunnel sub-TLV. The same procedure will be applied on subsequent S-PEs, until the message reaches to T-PE1 to complete the PSN binding setup.

During the binding process, if any PE does not agree to the requested tunnel/LSPs, it can send a Label Release Message to its upstream adjacent PE with Status Code set to "Reject to use the suggested tunnel/LSPs".

For congruent binding, the initiating PE (T-PE1) MUST set the C-bit, reset the S-bit and indicates the suggested tunnel/LSP in PSN Tunnel sub-TLV to the next-hop S-PE (S-PE1).

During the MS-PW setup, the PEs have the option to ignore the suggested tunnel/LSP, and select another tunnel/LSP for the segment PW between itself and its upstream PE on reverse direction only if the tunnel/LSP is congruent with the forwarding one. Otherwise, the procedure is the same as the strict binding.

The tunnel/LSPs may change after a MS-PW being established. When a tunnel/LSP has changed, the PE that detects the change SHOULD select an alternative tunnel/LSP for temporary use while negotiating with other PEs following the procedure described in this section.

6. Security Considerations

The ability to control which LSP to carry traffic from a PW can be a potential security risk both for denial of service and traffic interception. It is RECOMMENDED that PEs do not accept the use of LSPs identified in the PSN Tunnel Binding TLV unless the LSP end points match the PW or PW segment end points. Furthermore, where security of the network is believed to be at risk, it is RECOMMENDED that PEs implement the LDP security mechanisms described in [[RFC5036](#)] and [[RFC5920](#)].

7. IANA Considerations

7.1. LDP TLV Types

This document defines new TLV [[Section 2.1](#) of this document] for inclusion in LDP Label Mapping message. IANA is required to assign TLV type value to the new defined TLVs from LDP "TLV Type Name Space" registry.

7.1.1. PSN Tunnel Sub-TLVs

This document defines two sub-TLVs [[Section 2.1.1](#) of this document] for PSN Tunnel Binding TLV. IANA is required to create a new registry ("PSN Tunnel Sub-TLV Name Space") for PSN Tunnel sub-TLVs and to assign Sub-TLV type values to the following sub-TLVs.

IPv4 PSN Tunnel sub-TLV - 0x01 (to be confirmed by IANA)

IPv6 PSN Tunnel sub-TLV - 0x02 (to be confirmed by IANA)

7.2. LDP Status Codes

This document defines a new LDP status codes, IANA is required to assigned status codes to these new defined codes from LDP "STATUS CODE NAME SPACE" registry.

"Reject to use the suggested tunnel/LSPs" - 0x0000003B (to be confirmed by IANA)

8. Acknowledgements

The authors would like to thank Adrian Farrel, Kamran Raza, Xinchun Guo, Mingming Zhu and Li Xue for their comments and help in preparing this document. Also this draft benefits from the discussions with Nabil Bitar, Paul Doolan, Frederic Journay, Andy Malis, Curtis Villamizar and Luca Martini.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", [RFC 4447](#), April 2006.
- [RFC6370] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport Profile (MPLS-TP) Identifiers", [RFC 6370](#), September 2011.

9.2. Informative References

- [I-D.ietf-pwe3-dynamic-ms-pw]
Martini, L., Bocci, M., and F. Balus, "Dynamic Placement of Multi Segment Pseudowires", [draft-ietf-pwe3-dynamic-ms-pw-16](#) (work in progress), December 2012.

- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", [RFC 3471](#), January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#), January 2003.
- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", [RFC 3985](#), March 2005.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", [RFC 5036](#), October 2007.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", [RFC 5920](#), July 2010.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", [RFC 6073](#), January 2011.
- [RFC6373] Andersson, L., Berger, L., Fang, L., Bitar, N., and E. Gray, "MPLS Transport Profile (MPLS-TP) Control Plane Framework", [RFC 6373](#), September 2011.

Authors' Addresses

Mach(Guoyi) Chen
Huawei Technologies Co., Ltd
Q14 Huawei Campus, No. 156 Beiqing Road, Hai-dian District
Beijing 100095
China

Email: mach.chen@huawei.com

Wei Cao
Huawei Technologies Co., Ltd
Q14 Huawei Campus, No. 156 Beiqing Road, Hai-dian District
Beijing 100095
China

Email: wayne.caowei@huawei.com

Attila Takacs
Ericsson
Laborc u. 1.
Budapest 1037
Hungary

Email: attila.takacs@ericsson.com

Ping Pan
Infinera
169 West Java Drive, Sunnyvale, CA 94089
US

Email: ppan@infinera.com

