

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: April 30, 2012

M. Chen  
W. Cao  
Huawei Technologies Co., Ltd  
A. Takacs  
Ericsson  
P. Pan  
Infinera  
October 28, 2011

**LDP extensions for Explicit Pseudowire to transport LSP mapping**  
**draft-cao-pwe3-mppls-tp-pw-over-bidir-lsp-04.txt**

**Abstract**

A bidirectional Pseudowire (PW) service currently uses two unidirectional PWs each carried over a unidirectional LSP. Each end point of a PW or segment of multi-segment PW (MS-PW) independently selects the LSP to use to carry the PW for which it is the head end.

Some transport services may require that bidirectional PW traffic follows the same paths through the network in both directions. Therefore, PWs may be required to use LSP with the same paths. Co-routed bidirectional LSPs or unidirectional LSPs with the same route (links and nodes) allow this service to be provided.

This document specifies an optional extension to LDP that allows both ends of a PW (or segment of a MS-PW) to select and bind to the same co-routed bidirectional LSP or two unidirectional LSPs with the same route.

**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 April 30, 2012.

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



## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">4</a>
<a href="#">2.</a>	LDP Extensions . . . . .	<a href="#">5</a>
<a href="#">2.1.</a>	PSN Tunnel Binding TLV . . . . .	<a href="#">6</a>
<a href="#">2.1.1.</a>	PSN Tunnel Sub-TLV . . . . .	<a href="#">7</a>
<a href="#">3.</a>	Theory of Operation . . . . .	<a href="#">8</a>
<a href="#">4.</a>	PSN Binding Operation for SS-PW . . . . .	<a href="#">9</a>
<a href="#">5.</a>	PSN Binding Operation for MS-PW . . . . .	<a href="#">11</a>
<a href="#">6.</a>	Security Considerations . . . . .	<a href="#">13</a>
<a href="#">7.</a>	IANA Considerations . . . . .	<a href="#">13</a>
<a href="#">7.1.</a>	LDP TLV Types . . . . .	<a href="#">13</a>
<a href="#">7.1.1.</a>	PSN Tunnel Sub-TLVs . . . . .	<a href="#">13</a>
<a href="#">7.2.</a>	LDP Status Codes . . . . .	<a href="#">14</a>
<a href="#">8.</a>	Acknowledgements . . . . .	<a href="#">14</a>
<a href="#">9.</a>	References . . . . .	<a href="#">14</a>
<a href="#">9.1.</a>	Normative References . . . . .	<a href="#">14</a>
<a href="#">9.2.</a>	Informative References . . . . .	<a href="#">14</a>
	Authors' Addresses . . . . .	<a href="#">15</a>



## 1. Introduction

Pseudo Wire (PW) Emulation Edge-to-Edge (PWE3) [[RFC3985](#)] is a mechanism to emulate a number of layer 2 services, such as Asynchronous Transfer Mode (ATM), Frame Relay or Ethernet. 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). Today PWE3 generally uses two reverse unidirectional Label Distribution Protocol (LDP) [[RFC5036](#)] or Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) [[RFC3209](#)] LSPs as PSN tunnels, and each of the PEs selects and binds PSN tunnel independently. There is no protocol-based provision to explicitly associate a PW with a specific PSN tunnel.

For transport applications it has been identified that many transport services may require bidirectional traffic that follows congruent paths. When co-routed bidirectional LSPs [[RFC3471](#)][[RFC3473](#)] are used as PSN tunnels, this requirement can be fulfilled if both PEs of a specific/segment PW select and bind to the same co-routed bidirectional LSPs. In the case of unidirectional LSPs, LSPs with the same route need to be selected to support the PW. However, current mechanisms cannot guarantee appropriate mapping of PWs to underlying LSPs.

The lack of the control over LSP-PW binding may introduce service issues in operation, as shown in Figure 1.

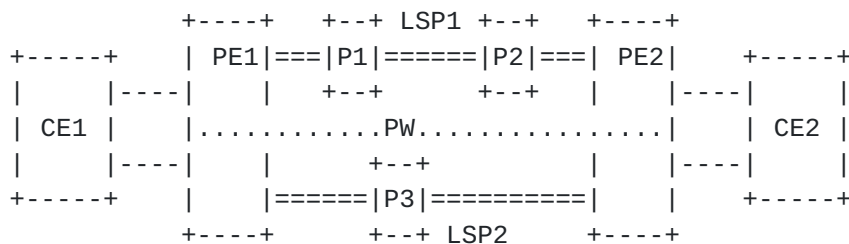


Figure 1: Inconsistent SS-PW to LSP binding scenario

There are two bidirectional LSPs: LSP1 and LSP2, along diverse paths. A bidirectional PW service is offered 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 the PE1->PE2 direction of the PW, while PE2 may select LSP2 (PE2-P3-PE2) as the PSN tunnel for the PE2->PE1 direction of the PW.

Consequently, the bidirectional PW service is delivered over two disjoint LSPs, which may have completely different service attributes



in terms of bandwidth and latency. If service offering requires consistent traffic behavior on forward and reverse direction, this may not be acceptable.

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 and the underlying PSN tunnels. However, this is prone to configuration errors and does not scale.

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

## 2. LDP Extensions

This document defines a new TLV, PSN Tunnel Binding TLV, to communicate tunnel/LSPs selection and binding requests between PEs at the bi-directional PW's setup time. The TLV carries PW's binding profile and provides both explicit and implicit information on the underlying PSN tunnels.

The binding TLV is optional, and MUST NOT affect the existing PW operation when not present in the messages.

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

Presently, the extension supports two types of binding requests:

1. Congruent binding: the requesting PE will ask the underlying LSPs to have the same route (across the same links and nodes). The response PE can select either co-routed bidirectional LSP or unidirectional LSP as the reverse PSN tunnel, as long as the selected LSP has the same route with the LSP the requesting PE selected.
2. Strict binding: the requesting PE will choose and explicitly indicate both forwarding and reverse LSP's in the requests.

In this document, the terminology of "tunnel" is identical to the "TE Tunnel" defined in [Section 2.1 of \[RFC3209\]](#), which is uniquely identified by a SESSION object that includes Tunnel end point address, Tunnel ID and Extended Tunnel ID. The terminology "LSP" is identical to the "LSP tunnel" defined in [Section 2.1 of \[RFC3209\]](#), which is uniquely identified by the SESSION object together with





SENDER\_TEMPLATE (or FILTER\_SPEC) object that consists of LSP ID and Tunnel end point address.

## 2.1. PSN Tunnel Binding TLV

PSN Tunnel Binding TLV is an optional TLV and MUST be carried in the LDP Label Mapping message if explicit PW to PSN tunnel binding is required. The format of this TLV is as follows:

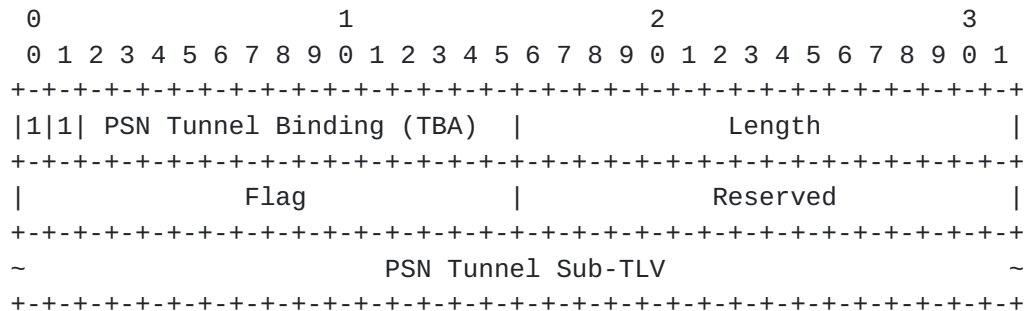
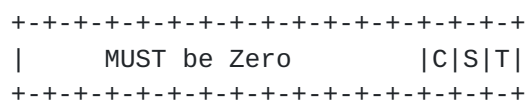


Figure 2: PSN Tunnel Binding TLV

The PSN Tunnel Binding TLV type is to be allocated by IANA.

The Length field is 2 octets in length. It defines the length in octets of the entire TLV.

The Flag field describes the binding requests, and has following format:



Three flags have been defined at the present time.

C (Congruent path) bit: This informs the remote T-PE/S-PEs about the properties of the underlying PSN tunnels. When set, the remote T-PE/S-PEs need to select tunnel/LSPs with the same route (e.g., the same co-routed bidirectional LSP as the requesting PE selected). If there is no satisfied tunnel, it may trigger the remote T-PE/S-PEs to establish a new tunnel.

S (Strict) bit: This instructs the PEs with respect to the handling of the underlying PSN tunnels. When set, the remote PE MUST use the tunnel/LSPs 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 PSN tunnels. When the bit is set, the PSN 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.

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.

In this document two sub-TLVs are defined: the IPv4/IPv6 Tunnel sub-TLVs. The format of the PSN Tunnel sub-TLVs is as follows:

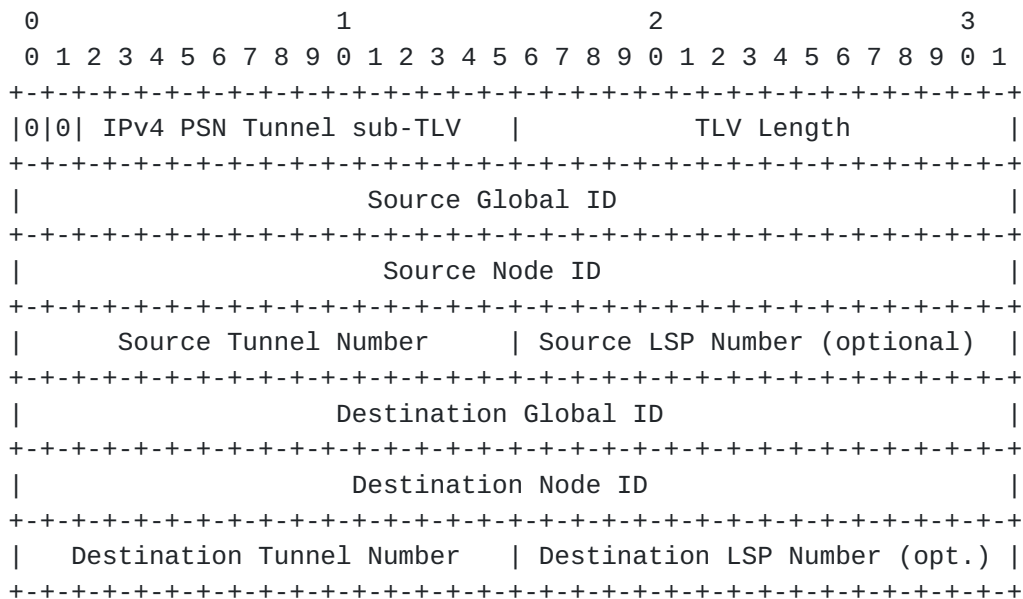


Figure 3: IPv4 PSN Tunnel sub-TLV format



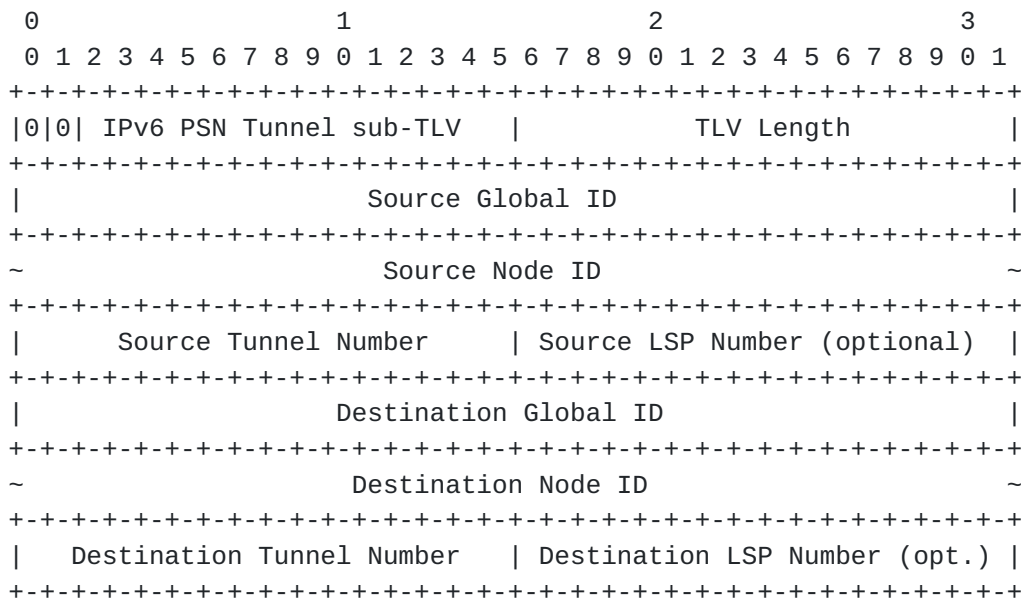


Figure 4: IPv6 PSN Tunnel sub-TLV format

The definition of Source and Destination Global/Node IDs and Tunnel/LSP Numbers are derived from [\[RFC6370\]](#). The notation is designed to describe co-routed or associated bi-directional LSPs, which is suitable in the context of the work here.

As defined in [Section 4.6.1.2](#) and [Section 4.6.2.2 of \[RFC3209\]](#), the "Tunnel end point 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 determines whether it stands for tunnel or LSP.

When the T-bit is set, it identifies a tunnel, and the Source/Destination LSP Number fields MUST be set to 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 no one available



or satisfies the constraints, the PE may trigger to 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 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 co-routed or bidirectional tunnels/LSPs at the same time. There may cause collisions of tunnel/LSPs selection as both PEs assume the active role.

The PEs can be generally categorized into two types:

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.

Segmented PW has defined the active/passive role election ([Section 7.2.1](#), [RFC6073]). This document will not define any new procedures.

In the remaining of this document, it will elaborate the operation in two situations:

1. SS-PW: In this scenario, both PEs of a PW assume 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.





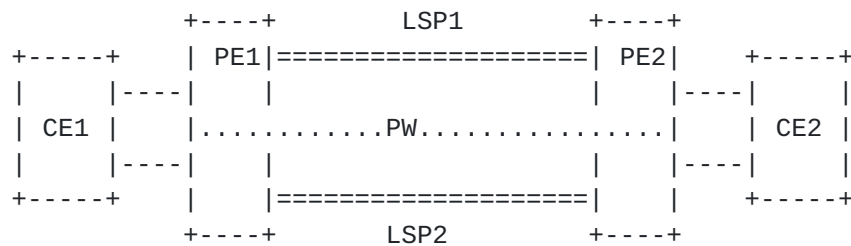


Figure 5: PSN binding operation in SS-PW environment

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), it **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.

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 have 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, it needs to compare its own Node ID against the received Source Node ID. If it's numerically lower, PE2 needs to find/establish a tunnel/LSP that meets the congruent constraint, and then 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 associated PSN LSP.

After a PW established, the operators may choose to switch the PW from the current tunnel/LSPs. Or, the underlying PSN is broken due to network failure. In this scenario, a new Label Mapping message MUST be sent to update the changes. Noting 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.



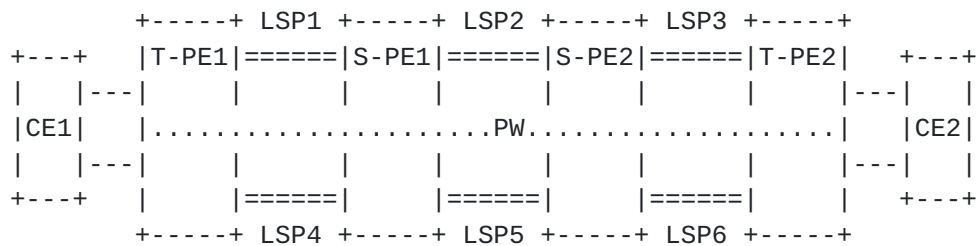


Figure 6: PSN binding operation in MS-PW environment

When the active PE (T-PE1) starts to signal for a MS-PW, a PSN Tunnel Binding TLV MUST be carried in the Label Mapping message and sent to the adjacent S-PE (say 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 (T-PE1) MUST set the S-bit, reset the C-bit and indicates the binding tunnel/LSP to the next-hop S-PE (S-PE1).

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 (S-PE2), with the PSN Tunnel sub-TLV carrying the information of the new PSN tunnel/LSPs selected by S-PE1 for the next PW segment. S-PE2 and subsequent S-PEs will repeat the same operation until the Label Mapping message reaches to the remote T-PE (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, 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-14](#) (work in progress), July 2011.
- [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.

#### Authors' Addresses

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

Email: mach@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  
Sri Mohana Satya Srinivas Singamsetty  
US

Email: ppan@infinera.com

