Internet Engineering Task Force (IETF)

Request for Comments: 8185 Category: Standards Track

ISSN: 2070-1721

W. Cheng
L. Wang
H. Li
China Mobile
J. Dong
Huawei Technologies
A. D'Alessandro
Telecom Italia
June 2017

Dual-Homing Coordination for MPLS Transport Profile (MPLS-TP) Pseudowires Protection

Abstract

In some scenarios, MPLS Transport Profile (MPLS-TP) pseudowires (PWs) (RFC 5921) may be statically configured when a dynamic control plane is not available. A fast protection mechanism for MPLS-TP PWs is needed to protect against the failure of an Attachment Circuit (AC), the failure of a Provider Edge (PE), or a failure in the Packet Switched Network (PSN). The framework and typical scenarios of dual-homing PW local protection are described in RFC 8184. This document proposes a dual-homing coordination mechanism for MPLS-TP PWs that is used for state exchange and switchover coordination between the dual-homing PEs for dual-homing PW local protection.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc8185.

Copyright 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
2.	Requirements Language	3
3.	Overview of the Proposed Solution	4
4.	Protocol Extensions for Dual-Homing MPLS-TP PW Protection	5
4.	.1. Information Exchange Between Dual-Homing PEs	5
4 .	.2. Protection Procedures	9
5.	IANA Considerations	13
6.	Security Considerations	13
7.	References	14
7.	.1. Normative References	14
7.	.2. Informative References	15
Cont	tributors	16
Auth	hors' Addresses	17

1. Introduction

[RFC6372], [RFC6378], and [RFC7771] describe the framework and mechanism of MPLS Transport Profile (MPLS-TP) linear protection, which can provide protection for the MPLS Label Switched Path (LSP) and pseudowires (PWs) between the edge nodes. These mechanisms cannot protect against the failure of the Attachment Circuit (AC) or the edge nodes. [RFC6718] and [RFC6870] specify the PW redundancy framework and mechanism for protecting the AC or edge node against failure by adding one or more edge nodes, but it requires PW switchover in case of an AC failure; also, PW redundancy relies on Packet Switched Network (PSN) protection mechanisms to protect against the failure of PW.

In some scenarios such as mobile backhauling, the MPLS PWs are provisioned with dual-homing topology in which at least the Customer Edge (CE) node on one side is dual-homed to two Provider Edge (PE) nodes. If a failure occurs in the primary AC, operators usually prefer to perform local switchover in the dual-homing PE side and keep the working pseudowire unchanged, if possible. This is to avoid massive PW switchover in the mobile backhaul network due to AC failure in the mobile core site; such massive PW switchover may in turn lead to congestion caused by migrating traffic away from the preferred paths of network planners. Similarly, as multiple PWs share the physical AC in the mobile core site, it is preferable to keep using the working AC when one working PW fails in the PSN to potentially avoid unnecessary switchover for other PWs. To meet the above requirements, a fast dual-homing PW protection mechanism is needed to protect against failure in the AC, the PE node, and the PSN.

[RFC8184] describes a framework and several scenarios of dual-homing PW local protection. This document proposes a dual-homing coordination mechanism for static MPLS-TP PWs; the mechanism is used for information exchange and switchover coordination between the dual-homing PEs for the dual-homing PW local protection. The proposed mechanism has been implemented and deployed in several mobile backhaul networks that use static MPLS-TP PWs for the backhauling of mobile traffic from the radio access sites to the core site.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Overview of the Proposed Solution

Linear protection mechanisms for the MPLS-TP network are defined in [RFC6378], [RFC7271], and [RFC7324]. When such mechanisms are applied to PW linear protection [RFC7771], both the working PW and the protection PW are terminated on the same PE node. In order to provide dual-homing protection for MPLS-TP PWs, some additional mechanisms are needed.

In MPLS-TP PW dual-homing protection, the linear protection mechanism (as defined in [RFC6378], [RFC7271], and [RFC7324]) on the single-homing PE (e.g., PE3 in Figure 1) is not changed, while on the dual-homing side, the working PW and protection PW are terminated on two dual-homing PEs (e.g., PE1 and PE2 in Figure 1), respectively, to protect against a failure occurring in a PE or a connected AC. As described in [RFC8184], a dedicated Dual-Node Interconnection (DNI) PW is used between the two dual-homing PE nodes to forward the traffic. In order to utilize the linear protection mechanism [RFC7771] in the dual-homing PEs scenario, coordination between the dual-homing PE nodes is needed so that the dual-homing PEs can switch the connection between the AC, the service PW, and the DNI-PW properly in a coordinated fashion by the forwarder.

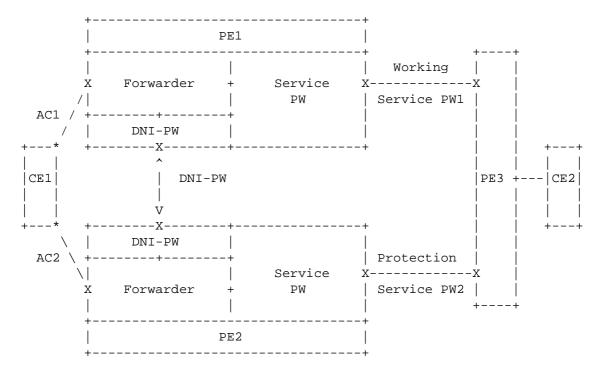


Figure 1: Dual-Homing Protection with DNI-PW

4. Protocol Extensions for Dual-Homing MPLS-TP PW Protection

In dual-homing MPLS-TP PW local protection, the forwarding states of the dual-homing PEs are determined by the forwarding state machine in Table 1.

Service PW	+ AC	+ DNI-PW	++ Forwarding Behavior
Active	Active	Up	Service PW <-> AC
Active	Standby	Up	Service PW <-> DNI-PW
Standby	Active	Up	DNI-PW <-> AC
Standby	Standby	Up	Drop all packets
Active	Active	Down	Service PW <-> AC
Active	Standby	Down	Drop all packets
Standby	Active	Down	Drop all packets
Standby	Standby	Down	Drop all packets

Table 1: Dual-Homing PE Forwarding State Machine

In order to achieve dual-homing MPLS-TP PW protection, coordination between the dual-homing PE nodes is needed to exchange the PW status and protection coordination requests.

4.1. Information Exchange Between Dual-Homing PEs

The coordination information will be sent on the DNI-PW over the Generic Associated Channel (G-ACh) as described in [RFC5586]. A new G-ACh channel type is defined for the dual-homing coordination between the dual-homing PEs of MPLS-TP PWs. This channel type can be used for the exchange of different types of information between the dual-homing PEs. This document uses this channel type for the exchange of PW status and switchover coordination between the dualhoming PEs. Other potential usages of this channel type are for further study and are out of the scope of this document.

The MPLS-TP Dual-Homing Coordination (DHC) message is sent on the DNI-PW between the dual-homing PEs. The format of the MPLS-TP DHC message is shown below:

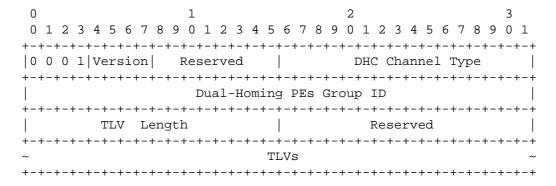


Figure 2: MPLS-TP Dual-Homing Coordination Message

The first 4 octets is the common G-ACh header as specified in [RFC5586]. The DHC Channel Type is the G-ACh channel type code point assigned by IANA (0x0009).

The Dual-Homing Group ID is a 4-octet unsigned integer to identify the dual-homing group to which the dual-homing PEs belong. It MUST be the same at both PEs in the same group.

The TLV Length field specifies the total length in octets of the subsequent TLVs.

In this document, two TLVs are defined in the MPLS-TP Dual-Homing Coordination message for dual-homing MPLS-TP PW protection:

Type	Description	Length
1	PW Status	20 bytes
2	Dual-Node Switching	16 bytes

The PW Status TLV is used by a dual-homing PE to report its service PW status to the other dual-homing PE in the same dual-homing group.

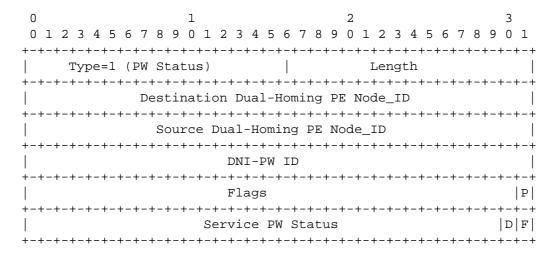


Figure 3: PW Status TLV

The Length field specifies the length in octets of the value field of the TLV.

The Destination Dual-Homing PE Node_ID is the 32-bit identifier of the receiver PE [RFC6370], which supports both IPv4 and IPv6 environments. Usually it is the same as the Label Switching Router ID (LSR ID) of the receiver PE.

The Source Dual-Homing PE Node_ID is the 32-bit identifier of the sending PE [RFC6370], which supports both IPv4 and IPv6 environments. Usually it is the same as the LSR ID of the sending PE.

The DNI-PW ID field contains the 32-bit PW ID [RFC8077] of the DNI-

The Flags field contains 32-bit flags, in which:

- o The P (Protection) bit indicates whether the Source Dual-Homing PE is the working PE (P=0) or the protection PE (P=1).
- o Other bits are reserved for future use, which MUST be set to 0 on transmission and MUST be ignored upon receipt.

The Service PW Status field indicates the status of the service PW between the sending PE and the remote PE. Currently, two bits are defined in the Service PW Status field:

o F bit: If set, it indicates Signal Fail (SF) [RFC6378] on the service PW. It can be either a local request generated by the PE itself or a remote request received from the remote PE.

- o D bit: If set, it indicates Signal Degrade (SD) [RFC6378] on the service PW. It can be either a local request or a remote request received from the remote PE.
- o Other bits are reserved for future use, which MUST be set to 0 on transmission and MUST be ignored upon receipt.

The Dual-Node Switching TLV is used by one dual-homing PE to send protection state coordination to the other PE in the same dual-homing group.

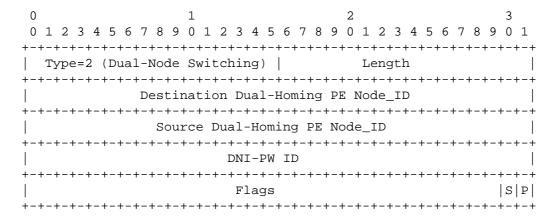


Figure 4: Dual-Node Switching TLV

The Length field specifies the length in octets of the value field of the TLV.

The Destination Dual-Homing PE Node_ID is the 32-bit identifier of the receiver PE [RFC6370]. Usually it is the same as the LSR ID of the receiver PE.

The Source Dual-Homing PE Node_ID is the 32-bit identifier of the sending PE [RFC6370]. Usually it is the same as the LSR ID of the sending PE.

The DNI-PW ID field contains the 32-bit PW-ID [RFC8077] of the DNI-

The Flags field contains 32-bit flags, in which:

o The P (Protection) bit indicates whether the Source Dual-Homing PE is the working PE (P=0) or the protection PE (P=1).

- o The S (PW Switching) bit indicates which service PW is used for forwarding traffic. It is set to 0 when traffic will be transported on the working PW, and it is set to 1 if traffic will be transported on the protection PW. The value of the S bit is determined by the protection coordination mechanism between the dual-homing PEs and the remote PE.
- o Other bits are reserved for future use, which MUST be set to 0 on transmission and MUST be ignored upon receipt.

When a change of service PW status is detected by one of the dual-homing PEs, it MUST be reflected in the PW Status TLV and sent to the other dual-homing PE as quickly as possible to allow for fast protection switching using three consecutive DHC messages. This set of three messages allows for fast protection switching even if one or two of these packets are lost or corrupted. After the transmission of the three rapid messages, the dual-homing PE MUST send the most recently transmitted service PW status periodically to the other dual-homing PE on a continual basis using the DHC message.

When one dual-homing PE determines that the active service PW needs to be switched from the working PW to the protection PW, it MUST send the Dual-Node Switching TLV to the other dual-homing PE as quickly as possible to allow for fast protection switching using three consecutive DHC messages. After the transmission of the three messages, the protection PW would become the active service PW, and the dual-homing PE MUST send the most recently transmitted Dual-Node Switching TLV periodically to the other dual-homing PE on a continual basis using the DHC message.

It is RECOMMENDED that the default interval of the first three rapid DHC messages be 3.3 ms, similar to [RFC6378], and the default interval of the subsequent messages is 1 second. Both the default interval of the three consecutive messages as well as the default interval of the periodic messages SHALL be configurable by the operator.

4.2. Protection Procedures

The dual-homing MPLS-TP PW protection mechanism can be deployed with the existing AC redundancy mechanisms. On the PSN side, a PSN tunnel protection mechanism is not required, as the dual-homing PW protection can also protect if a failure occurs in the PSN.

This section uses the one-side dual-homing scenario as an example to describe the dual-homing PW protection procedures; the procedures for a two-side dual-homing scenario would be similar.

On the dual-homing PE side, the role of working and protection PE are set by the management system or local configuration. The service PW connecting to the working PE is the working PW, and the service PW connecting to the protection PE is called the protection PW.

On the single-homing PE side, it treats the working PW and protection PW as if they terminate on the same remote PE node, thus normal MPLS-TP protection coordination procedures still apply on the single-homing PE.

The forwarding behavior of the dual-homing PEs is determined by the components shown in the figure below:

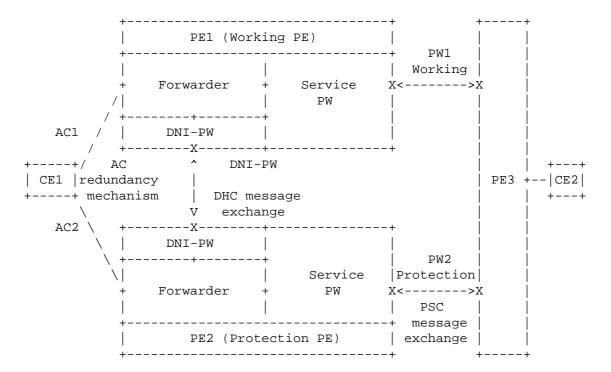


Figure 5: Components of One-Side Dual-Homing PW Protection

In Figure 5, for each dual-homing PE, the service PW is the PW used to carry service between the dual-homing PE and the remote PE. The state of the service PW is determined by the Operation, Administration, and Maintenance (OAM) mechanisms between the dual-homing PEs and the remote PE.

The DNI-PW is provisioned between the two dual-homing PE nodes. It is used to bridge traffic when a failure occurs in the PSN or in the ACs. The state of the DNI-PW is determined by the OAM mechanism between the dual-homing PEs. Since the DNI-PW is used to carry both

the DHC messages and the service traffic during protection switching, it is important to ensure the robustness of the DNI-PW. In order to avoid the DNI-PW failure due to the failure of a particular link, it is RECOMMENDED that multiple diverse links be deployed between the dual-homing PEs and the underlying Label Switched Path (LSP) protection mechanism SHOULD be enabled.

The AC is the link that connects a dual-homing PE to the dual-homed CE. The status of AC is determined by the existing AC redundancy mechanisms; this is out of the scope of this document.

In order to perform dual-homing PW local protection, the service PW status and Dual-Node Switching coordination requests are exchanged between the dual-homing PEs using the DHC message defined in Section 4.1.

Whenever a change of service PW status is detected by a dual-homing PE, it MUST be reflected in the PW Status TLV and sent to the other dual-homing PE immediately using the three consecutive DHC messages. After the transmission of the three rapid messages, the dual-homing PE MUST send the most recently transmitted service PW status periodically to the other dual-homing PE on a continual basis using the DHC message. This way, both dual-homing PEs have the status of the working and protection PW consistently.

When there is a switchover request either generated locally or received on the protection PW from the remote PE, based on the status of the working and protection service PW along with the local and remote request of the protection coordination between the dual-homing PEs and the remote PE, the active/standby state of the service PW can be determined by the dual-homing PEs. As the remote protection coordination request is transmitted over the protection path, in this case the active/standby status of the service PW is determined by the protection PE in the dual-homing group.

If it is determined on one dual-homing PE that switchover of the service PW is needed, this dual-homing PE MUST set the S bit in the Dual-Node Switching TLV and send it to the other dual-homing PE immediately using the three consecutive DHC messages. With the exchange of service PW status and the switching request, both dual-homing PEs are consistent on the active/standby forwarding status of the working and protection service PWs. The status of the DNI-PW is determined by PW OAM mechanism as defined in [RFC5085], and the status of ACs is determined by existing AC redundancy mechanisms: both are out of the scope of this document. The forwarding behavior on the dual-homing PE nodes is determined by the forwarding state machine as shown in Table 1.

Using the topology in Figure 5 as an example, in normal state, the working PW (PW1) is in active state, the protection PW (PW2) is in standby state, the DNI-PW is up, and AC1 is in active state according to the AC redundancy mechanism. According to the forwarding state machine in Table 1, traffic will be forwarded through the working PW (PW1) and the primary AC (AC1). No traffic will go through the protection PE (PE2) or the DNI-PW, as both the protection PW (PW2) and the AC connecting to PE2 are in standby state.

If a failure occurs in AC1, the state of AC2 changes to active according to the AC redundancy mechanism, while there is no change in the state of the working and protection PWs. According to the forwarding state machine in Table 1, PE1 starts to forward traffic between the working PW and the DNI-PW, and PE2 starts to forward traffic between AC2 and the DNI-PW. It should be noted that in this case only AC switchover takes place; in the PSN, traffic is still forwarded using the working PW.

If a failure in the PSN brings PW1 down, the failure can be detected by PE1 or PE3 using existing OAM mechanisms. If PE1 detects the failure of PW1, it MUST inform PE2 of the state of the working PW using the PW Status TLV in the DHC messages and change the forwarding status of PW1 to standby. On receipt of the DHC message, PE2 SHOULD change the forwarding status of PW2 to active. Then, according to the forwarding state machine in Table 1, PE1 SHOULD set up the connection between the DNI-PW and AC1, and PE2 SHOULD set up the connection between PW2 and the DNI-PW. According to the linear protection mechanism [RFC6378], PE2 also sends an appropriate protection coordination message [RFC6378] over the protection PW (PW2) to PE3 for the remote side to switchover from PW1 to PW2. If PE3 detects the failure of PW1, according to the linear protection mechanism [RFC6378], it sends a protection coordination message on the protection PW (PW2) to inform PE2 of the failure on the working PW. Upon receipt of the message, PE2 SHOULD change the forwarding status of PW2 to active and set up the connection according to the forwarding state machine in Table 1. PE2 SHOULD send a DHC message to PE1 with the S bit set in the Dual-Node Switching TLV to coordinate the switchover on PE1 and PE2. This is useful for a unidirectional failure that cannot be detected by PE1.

If a failure brings the working PE (PE1) down, the failure can be detected by both PE2 and PE3 using existing OAM mechanisms. Both PE2 and PE3 SHOULD change the forwarding status of PW2 to active and send a protection coordination message [RFC6378] on the protection PW (PW2) to inform the remote side to switchover. According to the existing AC redundancy mechanisms, the status of AC1 changes to

standby and the state of AC2 changes to active. According to the forwarding state machine in Table 1, PE2 starts to forward traffic between the PW2 and AC2.

5. IANA Considerations

IANA has assigned a new channel type for the "MPLS-TP Dual-Homing Coordination Message" from the "MPLS Generalized Associated Channel (G-ACh) Types (including Pseudowire Associated Channel Types)" subregistry within the "Generic Associated Channel (G-ACh) Parameters" registry.

Value	Description	Reference
0×0009	MPLS-TP Dual-Homing Coordination message	RFC 8185

IANA has created a new subregistry called "MPLS-TP DHC TLVs" within the "Generic Associated Channel (G-ACh) Parameters" registry. The registry has the following fields and initial allocations:

Type	Description	Length	Reference
0x0000	Reserved		
0x0001	PW Status	20 Bytes	RFC 8185
0×0002	Dual-Node Switching	16 Bytes	RFC 8185

The allocation policy for this registry is IETF Review, as specified in [RFC8126].

6. Security Considerations

MPLS-TP is a subset of MPLS and so builds upon many of the aspects of the MPLS security model. Please refer to [RFC5920] for generic MPLS security issues and methods for securing traffic privacy and integrity.

The DHC message defined in this document contains control information. If it is injected or modified by an attacker, the dual-homing PEs might not agree on which PE should be used to deliver the CE traffic, and this could be used as a denial-of-service attack against the CE. It is important that the DHC message be used within a trusted MPLS-TP network domain as described in [RFC6941].

The DHC message is carried in the G-ACh [RFC5586], so it is dependent on the security of the G-ACh itself. The G-ACh is a generalization of the Associated Channel defined in [RFC4385]. Thus, this document relies on the security mechanisms provided for the Associated Channel as described in those two documents.

As described in the Security Considerations section of [RFC6378], the G-ACh is essentially connection oriented, so injection or modification of control messages requires the subversion of a transit node. Such subversion is generally considered hard in connection-oriented MPLS networks and impossible to protect against at the protocol level. Management-level techniques are more appropriate. The procedures and protocol extensions defined in this document do not affect the security model of MPLS-TP linear protection as defined in [RFC6378].

Uniqueness of the identifiers defined in this document is guaranteed by the assigner (e.g., the operator). Failure by an assigner to use unique values within the specified scoping for any of the identifiers defined herein could result in operational problems. Please refer to [RFC6370] for more details about the uniqueness of the identifiers.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
 Requirement Levels", BCP 14, RFC 2119,
 DOI 10.17487/RFC2119, March 1997,
 http://www.rfc-editor.org/info/rfc2119.

- [RFC6370] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport
 Profile (MPLS-TP) Identifiers", RFC 6370,
 DOI 10.17487/RFC6370, September 2011,
 http://www.rfc-editor.org/info/rfc6370.
- [RFC6378] Weingarten, Y., Ed., Bryant, S., Osborne, E., Sprecher,
 N., and A. Fulignoli, Ed., "MPLS Transport Profile (MPLS-TP) Linear Protection", RFC 6378, DOI 10.17487/RFC6378,
 October 2011, http://www.rfc-editor.org/info/rfc6378>.

- [RFC7271] Ryoo, J., Ed., Gray, E., Ed., van Helvoort, H.,
 D'Alessandro, A., Cheung, T., and E. Osborne, "MPLS
 Transport Profile (MPLS-TP) Linear Protection to Match the
 Operational Expectations of Synchronous Digital Hierarchy,
 Optical Transport Network, and Ethernet Transport Network
 Operators", RFC 7271, DOI 10.17487/RFC7271, June 2014,
 http://www.rfc-editor.org/info/rfc7271.
- [RFC8077] Martini, L., Ed. and G. Heron, Ed., "Pseudowire Setup and
 Maintenance Using the Label Distribution Protocol (LDP)",
 STD 84, RFC 8077, DOI 10.17487/RFC8077, February 2017,
 http://www.rfc-editor.org/info/rfc8077.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
 May 2017, http://www.rfc-editor.org/info/rfc8174.

7.2. Informative References

- [RFC6718] Muley, P., Aissaoui, M., and M. Bocci, "Pseudowire Redundancy", RFC 6718, DOI 10.17487/RFC6718, August 2012, http://www.rfc-editor.org/info/rfc6718.
- [RFC6870] Muley, P., Ed. and M. Aissaoui, Ed., "Pseudowire Preferential Forwarding Status Bit", RFC 6870, DOI 10.17487/RFC6870, February 2013, http://www.rfc-editor.org/info/rfc6870.

- [RFC6941] Fang, L., Ed., Niven-Jenkins, B., Ed., Mansfield, S., Ed.,
 and R. Graveman, Ed., "MPLS Transport Profile (MPLS-TP)
 Security Framework", RFC 6941, DOI 10.17487/RFC6941, April
 2013, http://www.rfc-editor.org/info/rfc6941.
- [RFC7771] Malis, A., Ed., Andersson, L., van Helvoort, H., Shin, J.,
 Wang, L., and A. D'Alessandro, "Switching Provider Edge
 (S-PE) Protection for MPLS and MPLS Transport Profile
 (MPLS-TP) Static Multi-Segment Pseudowires", RFC 7771,
 DOI 10.17487/RFC7771, January 2016,
 http://www.rfc-editor.org/info/rfc7771.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, http://www.rfc-editor.org/info/rfc8126.
- [RFC8184] Cheng, W., Wang, L., Li, H., Davari, S., and J. Dong,
 "Dual-Homing Protection for MPLS and the MPLS Transport
 Profile (MPLS-TP) Pseudowires", RFC 8184,
 DOI 10.17487/RFC8184, June 2017.

Contributors

The following individuals substantially contributed to the content of this document:

Kai Liu Huawei Technologies

Email: alex.liukai@huawei.com

Shahram Davari Broadcom Corporation Email: davari@broadcom.com

Authors' Addresses

Weigiang Cheng China Mobile No.32 Xuanwumen West Street Beijing 100053 China

Email: chengweiqiang@chinamobile.com

Lei Wang China Mobile No.32 Xuanwumen West Street Beijing 100053 China

Email: Wangleiyj@chinamobile.com

Han Li China Mobile No.32 Xuanwumen West Street Beijing 100053 China

Email: Lihan@chinamobile.com

Jie Dong Huawei Technologies Huawei Campus, No. 156 Beiging Rd. Beijing 100095 China

Email: jie.dong@huawei.com

Alessandro D'Alessandro Telecom Italia via Reiss Romoli, 274 Torino 10148 Italy

Email: alessandro.dalessandro@telecomitalia.it