Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2016

W. Cheng
L. Wang
H. Li
China Mobile
K. Liu
Huawei Technologies
S. Davari
Broadcom Corporation
J. Dong
Huawei Technologies
A. D'Alessandro
Telecom Italia
October 19, 2015

Dual-Homing Coordination for MPLS Transport Profile (MPLS-TP) Pseudowires Protection
draft-ietf-pals-mpls-tp-dual-homing-coordination-01

Abstract

In some scenarios, the MPLS Transport Profile (MPLS-TP) Pseudowires (PWs) are provisioned through either static configuration or management plane, where a dynamic control plane is not available. A fast protection mechanism for MPLS-TP PWs is needed to protect against the failure of Attachment Circuit (AC), the failure of Provider Edge (PE) and also the failure in the Packet Switched Network (PSN). The framework and scenarios for dual-homing pseudowire (PW) local protection are described in [draft-ietf-pals-mpls-tp-dual-homing-protection]. This document proposes a dual-homing coordination mechanism for MPLS-TP PWs, which is used for state exchange and coordination between the dual-homing PEs for dual-homing PW local protection.

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].

Table of Contents

1. Introduction

[RFC6372], [RFC6378] and [I-D.ietf-pals-ms-pw-protection] describe the framework and mechanism of MPLS-TP Linear protection, which can provide protection for the MPLS LSP and PW between the edge nodes. These mechanisms does not protect the failure of the Attachment Circuit (AC) or the edge nodes. [RFC6718] and [RFC6870] specifies the PW redundancy framework and mechanism for protecting the AC or edge node failure by adding one or more edge nodes, but it requires

PW switchover in case of a AC failure, also PW redundancy relies on PSN protection mechanisms to protect 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 CE node on one side is dual-homed to two PEs. If a failure occurs in the primary AC, operators usually prefer to perform switchover only 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 the AC failure in the core site, and also could achieve efficient and balanced link bandwidth utilization. Similarly, it is preferable to keep using the working AC when one working PW fails in PSN network. A fast dual-homing PW protection mechanism is needed to protect the failure in AC, the PE node and the PSN network to meet the above requirements.

[I-D.ietf-pals-mpls-tp-dual-homing-protection] describes a framework and several scenarios for dual-homing pseudowire (PW) local protection. This document proposes a dual-homing coordination mechanism for static MPLS-TP PWs, which is used for information exchange and coordination between the dual-homing PEs for the dual-homing PW local protection. The proposed mechanism has been deployed in several mobile backhaul networks which use static MPLS-TP PWs for the backhauling of mobile traffic from the RF sites to the core site.

2. Overview of the Proposed Solution

Linear protection mechanisms for MPLS-TP network are defined in [RFC6378], [RFC7271] and [RFC7324]. When such mechanisms are applied to PW linear protection [I-D.ietf-pals-ms-pw-protection], both the working PW and the protection PW are terminated on the same PE nodes. 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 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 the failure occurs in the dual-homing PEs and the connected ACs. As specified in [I-D.ietf-pals-mpls-tp-dual-homing-protection], a dedicated Dual-Node Interconnection (DNI) PW is provisioned between the two dual-homing PE nodes, which is used to bridge the traffic between the dual-homing PEs when failure happens in the working PW or the primary AC. In order to make the linear protection mechanism work in the dual-homing PEs scenario, some coordination between the dual-homing PE nodes is needed, so that the dual-homing PEs can set the connection between AC, the service PW and the DNI-PW properly in a coordinated fashion.

```
               +----------------+
               /                |                +--------+
          AC1 /|   PE1          |  Working PW    |        |
             / |                X----------------X        |
            /  |                |  Service PW1   |        |
       +---/+  +--------X-------+                |        |   +----+
       |    |           | DNI PW                 |  PE3   |   |    |
       | CE1|           |                        |        |---| CE2|
       +---\+  +--------X-------+                |        |   +----+
            \  |                | Protection PW  |        |
             \ |                X----------------X        |
          AC2 \|                | Service PW2    |        |
               \   PE2          |                +--------+
               +----------------+
```

Figure 1. Dual-homing Protection with DNI-PW

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

In dual-homing MPLS-TP PW local protection, the forwarding state of the dual-homing PEs are determined by the forwarding state machine as defined in [I-D.ietf-pals-mpls-tp-dual-homing-protection]. In order to achieve the MPLS-TP PW dual-homing protection, coordination between the dual-homing PE nodes is needed to exchange the PW status and protection coordination requests.

3.1. Information Exchange Between Dual-Homing PEs

The coordination information will be sent over the G-ACh as described in [RFC5586]. A new G-ACh channel type is defined for the coordination between the dual-homing PEs of MPLS-TP PWs. This channel type can be used for the exchange of different kinds of information between the dual-homing PEs. This document uses this channel type for the PW status exchange and switchover coordination between the dual-homing PEs. Other potential usage 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 MPLS-TP DHC message is shown below:

```
    0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |0 0 0 1|Version|     Flags     |         DHC Code Point        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     Dual-Homing PEs Group ID                  |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |         TLV  Length           |           Reserved            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                              TLVs                             ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2. MPLS-TP Dual-Homing Coordination Message

The Dual-Homing Group ID is a 4-octet unsigned integer to identify the dual-homing PEs in the same dual-homing group.

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

Type	Description	Length
1	PW State	20 Bytes
2	Dual-Node Switching	16 Bytes

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

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |     Type=1 (PW Status)        |          Length               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |              Destination Dual-homing PE Node_ID               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                Source Dual-homing PE Node_ID                  |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                         DNI PW-ID                             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                         Flags                               |P|
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                      Service PW State                     |D|F|
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3. PW State TLV

- The Destination Dual-homing PE Node_ID is the 32-bit Identifier of the receiver PE.

- The Source Dual-homing PE Node_ID is the 32-bit identifier of the sending PE.

- The DNI PW-ID field contains the 32-bit PW-ID of the DNI PW.

- The Flag field contains 32 bit flags.

  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.

- The Service PW State field indicates the state of the Service PW between the sending PE and the remote PE. Currently two bits are defined in the Service PW Request field:

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

  o D bit: If set, it indicates Signal Degrade (SD) generated 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 and MUST be set to 0 on transmission and SHOULD be ignored upon receipt.

The Dual-Node Switching TLV is used by the protection dual-homing PE to send protection state coordination to the working dual-homing PE.

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  Type=2 (Dual-Node Switching) |          Length               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |              Destination Dual-homing PE Node_ID               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                Source Dual-homing PE Node_ID                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         DNI PW-ID                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                          Flags                            |S|P|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4. Dual-node Switching TLV

- The Destination Dual-homing PE Node_ID is the 32-bit Identifier of the receiver PE.

- The Source Dual-homing PE Node_ID is the 32-bit identifier of the sending PE.

- The DNI PW-ID field contains the 32-bit PW-ID of the DNI PW.

- The Flag field contains 32 bit flags.

- The P (Protection) bit indicates whether the Source Dual-homing PE is the working PE (P=0) or the protection PE (P=1). With the mechanism described in this document, only the protection PE can send DHC message with the Dual-node Switching TLV therefore it is always set to 1 when sent.

- 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 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 protection PE and the remote PE.

The MPLS-TP DHC message is exchanged periodically between the dual-homing PEs. Whenever a change of service PW state is detected by a dual-homing PE, it MUST be reflected in the PW State TLV and sent to the other dual-homing PE using a DHC message immediately. The Dual-Node Switching TLV is carried in the DHC message when a switchover request is issued by the protection PE according to the linear protection mechanism.

3.2. Protection Procedures

The dual-homing MPLS-TP PW protection mechanism can be deployed with the existing AC redundancy mechanisms, e.g. Multi-Chassis Link Aggregation Group (MC-LAG). On the PSN network side, PSN tunnel protection mechanism is not required, as the dual-homing PW protection can also protect the failure occurs in the PSN network.

This section takes one-side dual-homing scenario as example to describe the dual-homing PW protection procedures, the procedures for two-side dual-homing scenario would be similar.

On dual-homing PE side, the role of working and protection PE are set by NMS 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 protection PW.

On single-homing PE side, it just treats the working PW and protection PW as if they terminate on the same remote PE node, thus normal MPLS-TP protection coordination mechanisms 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:

```
             +----------------------------------+           +-----+
             |        PE1 (Working PE)          |           |     |
             +----------------------------------+           |     |
             |                 |                |    PW1    |     |
             +    Forwarder    +     Service    X<--------->X     |
            /|                 |       PW       |  Working  |     |
           / +--------+--------+                |    PW     |     |
     AC1  /  |     DNI PW      |                |           |     |
         /   +--------X--------+----------------+           |     |
 +-----+/             ^                                     |     |
 | CE1 |              |  DNI PW                             | PE3 | +---+
 +-----+              |                                     |     --|CE3|
        \             V                                     |     | +---+
     AC2 \   +--------X--------+----------------+           |     |
          \  |     DNI PW      |                |           |     |
           \ +--------+--------+                |           |     |
            \|                 |     Service    |    PW2    |     |
             +    Forwarder    +       PW       X<--------->X     |
             |                 |                | Protection|     |
             +----------------------------------+     PW    |     |
             |        PE2 (Protection PE)       |           |     |
             +----------------------------------+           +-----+
```

Figure 5. Components of one-side dual-homing PW protection

In figure 5, for a dual-homing PE, service PW is the PW used to carry service between the dual-homing PE and the remote PE. The state of service PW is determined by some mechanisms between the dual-homing PE and the remote PE (e.g. by OAM).

DNI PW is provisioned between the two dual-homing PE nodes. It is used to bridge traffic when failure occurs in the PSN network or in the ACs. The state of DNI PW is determined by some mechanisms between the dual-homing PEs (e.g. by OAM). Since DNI PW is used to carry both the coordination messages and service traffic, it is RECOMMENDED to provision multiple links between the dual-homing PEs and use some protection mechanism for the DNI PW.

AC is the link which connects the dual-homing PEs to the dual-homed CE. The status of AC is determined by some AC redundancy mechanisms (e.g. by MC-LAG).

In order to perform dual-homing PW local protection, the service PW state and Dual-node switching coordination requests are exchanged between the dual-homing PEs using the DHC message defined above.

Whenever a change of service PW state is detected by a dual-homing PE, it MUST be reflected in the PW State TLV and sent to the other dual-homing PE using a DHC message immediately. In case there is a switchover request received on the protection PW from the remote PE, the protection PE MUST set the switchover request in the Dual-Node Switching TLV and send it to the working PE using the DHC message. After the exchange of service PW state and switching request, both dual-homing PEs could determine the Active/Standby forwarding status of the working and protection service PWs. The status of DNI PW and the ACs are determined by other mechanisms out of the scope of this document. The forwarding behavior of the dual-homing PE nodes is determined by the forwarding state machine as shown in the following table:

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

Table 1. Dual-homing PE Forwarding State Machine

Take the topology in figure 5 as 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 AC side redundancy mechanism. According to 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 status.

If some 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 PSN network traffic is still forwarded using the working PW.

If some failure in the PSN network causes PW1 to go down, the failure can be detected by the working PE (PE1) or the remote PE (PE3) using some mechanisms (e.g. by OAM). If PE1 detects the failure of PW1, it

MUST inform PE2 the state of working PW using the PW State TLV in DHC message 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 linear protection mechanism, PE2 also sends an appropriate protection coordination message on the protection PW (PW2) to PE3 for the remote side switchover from PW1 to PW2. If PE3 detects the failure of PW1, according to linear protection mechanism, it sends a protection coordination message on the protection PW (PW2) to inform PE2 to switchover to the protection PW. Upon receipt of the message, PE2 SHOULD change the forwarding status of PW2 to active and set up the connection according to Table 1. PE2 SHOULD send a DHC message to PE1 with the S bit in the Dual-Node Switching TLV set to coordinate the switchover on PE1 and PE2. This can be useful for unidirectional failure which was not detected by PE1.

If some failure causes the working PE (PE1) to go down, the failure can be detected by both the protection PE(PE2) and the remote PE(PE3) using some mechanisms (e.g. OAM). PE2 changes the forwarding status of PW2 to active, and PE3 sends a protection coordination message on the protection path to inform PE2 to switchover to the protection PW. According to AC redundancy mechanism, 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 protection PW and AC2.

4. IANA Considerations

IANA needs to assign one new channel type for "MPLS-TP Dual-Homing Coordination message" from the "Pseudowire Associated Channel Types" registry.

This document creates a new registry called "MPLS-TP DHC TLVs" registry. Two new TLVs are defined in this document:

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

5. Security Considerations

Procedures and protocol extensions defined in this document do not affect the security model of MPLS-TP linear protection as defined in [RFC6378]. Please refer to [RFC5920] for MPLS security issues and generic methods for securing traffic privacy and integrity.

6. References

6.1. Normative References

[I-D.ietf-pals-mpls-tp-dual-homing-protection]
Cheng, W., Wang, L., Li, H., Liu, K., Davari, S., Dong, J., and A. D'Alessandro, "Dual-Homing Protection for MPLS and MPLS-TP Pseudowires", draft-ietf-pals-mpls-tp-dual-homing-protection-00 (work in progress), April 2015.

[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>.

[RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <http://www.rfc-editor.org/info/rfc5586>.

[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>.

[RFC7324] Osborne, E., "Updates to MPLS Transport Profile Linear Protection", RFC 7324, DOI 10.17487/RFC7324, July 2014, <http://www.rfc-editor.org/info/rfc7324>.

6.2. Informative References

[I-D.ietf-pals-ms-pw-protection]
Malis, A., Andersson, L., Helvoort, H., Shin, J., Wang, L., and A. D'Alessandro, "S-PE Protection for MPLS and MPLS-TP Static Multi-Segment Pseudowires", draft-ietf-pals-ms-pw-protection-03 (work in progress), October 2015.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <http://www.rfc-editor.org/info/rfc5920>.

[RFC6372] Sprecher, N., Ed. and A. Farrel, Ed., "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, DOI 10.17487/RFC6372, September 2011, <http://www.rfc-editor.org/info/rfc6372>.

[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>.

Authors' Addresses

Weiqiang 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

Kai Liu
Huawei Technologies
Huawei Base, Bantian, Longgang District
Shenzhen 518129
China

Email: alex.liukai@huawei.com

Shahram Davari
Broadcom Corporation
3151 Zanker Road
San Jose 95134-1933
United States

Email: davari@broadcom.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing 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