Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 19, 2014

W. Cheng
L. Wang
H. Li
China Mobile
K. Liu
Huawei Technologies
S. Davari
Broadcom Corporation
J. Dong
Huawei Technologies

July 01, 2014

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

draft-cheng-pwe3-mpls-tp-dual-homing-protection-00

Abstract

In some scenarios, the MPLS Trasport 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). This document proposes a dual-homing protection mechanism for MPLS-TP PWs, which can provide fast protection for comprehensive failure scenarios including the failure of AC, the PE node or the PSN network.

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

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 19, 2014.

Copyright Notice

Table of Contents

1. Introduction

[RFC6372] and [RFC6378] describe the framework and mechanism of MPLS-TP Linear protection, which can provide protection for the MPLS LSP or PW between the edge nodes. Such mechanism does not protect the failure of the Attachement Circuit (AC) or the edge node. [RFC6718] [RFC6870] describe the framework and mechanism for PW redundancy to provide protection for AC or PE node failure. The PW redundancy mechanism is based on the signaling of Label Distribution Protocol (LDP), which is applicable to MPLS PWs or MPLS-TP PWs with a dynamic control plane. [I-D.ietf-pwe3-endpoint-fast-protection] describes a fast local repair mechanism for PW egress endpoint failures, which is based on PW redundancy, upstream label assignment and context specific label switching. Such mechanism is applicable to PWs with a dynamic control plane.

In some scenarios such as mobile backhauling, the MPLS-TP PWs are provisioned through either static configuration or management plane, where a dynamic control plane is not available. A fast protection mechanism is needed for these MPLS-TP PWs to protect against the failure of AC, the PE node or the PSN network.

In addition, if at least one side CE node is dual-homed to two PEs, and a fault occurs in the primary AC, operators usually prefer to perform switchover only in the AC side and keeps using the working pseudowire if possible. The purpose is to avoid massive PW switchover caused by AC failure in mobile backhaul networks and to achieve efficient and balanced link bandwidth utilization in the PSN network.

This document proposes a dual-homing protection mechanism for static MPLS-TP PWs, which can provide fast protection for comprehensive failure scenarios including failure of AC, the PE node or the PSN network, and meet the requirement of avoiding PW switchover when possible. The mechanism defined in this document is complementary to the existing protection mechanisms.

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 sites.

2. MPLS-TP PW Dual-Homing Protection Scenarios

The following sections describe the typical topology and application scenarios of MPLS-TP PW dual-homing protection. The scenarios can be classified into two categories: one-side dual-homing and two-side dual-homing.

2.1. One-side Dual-Homing Protection

```
           |<--------------- Emulated Service --------------->|
           |                                                  |
           |          |<------- Pseudo Wire ------>|          |
           |          |                            |          |
           |          |    |<-- PSN Tunnels-->|    |          |
           |          V    V                  V    V          |
           V    AC    +----+                  +----+          V
     +-----+    |     | PE1|                  |    |          +-----+
     |     |----------|........PW1.(working).......|          |     |
     |     |          |    |                  |    |          |     |
     |     |          +----+                  |    |     AC   |     |
     |     |            ||                    |    |     |    |     |
     | CE1 |           DNI PW                 |PE3 |----------| CE2 |
     |     |            ||                    |    |          |     |
     |     |          +----+                  |    |          |     |
     |     |          |    |                  |    |          |     |
     |     |----------|......PW2.(protection)......|          |     |
     +-----+    |     | PE2|                  |    |          +-----+
                AC    +----+                  +----+
                Figure 1. One-side PW dual-homing protection
```

Figure 1 illustrates the network scenario of one-side CE dual-homing protection. CE1 is dual-homed to PE1 and PE2, while CE2 is single-homed to PE3. This topology protects the node failures of PE1 and PE2 and the AC link failures between CE1 and PE1, PE2. This scheme can be used in mobile backhauling application scenarios. For example, NodeB serves as CE2 while RNC serves as CE1. PE3 works as an access side MPLS-TP device while PE1 and PE2 works as a core side MPLS-TP device.

2.2. Two-side Dual-Homing Protection

```
           |<---------------- Emulated Service -------------->|
           |                                                  |
           |          |<-------- Pseudowire ------>|          |
           |          |                            |          |
           |          |    |<-- PSN Tunnels-->|    |          |
           |          V    V                  V    V          |
           V    AC    +----+                  +----+     AC   V
     +-----+    |     |....|.......PW1........|....|     |    +-----+
     |     |----------| PE1|                  | PE3|----------|     |
     |     |          +----+                  +----+          |     |
     |     |            ||                      ||            |     |
     | CE1 |          DNI PW                  DNI PW          | CE2 |
     |     |            ||                      ||            |     |
     |     |          +----+                  +----+          |     |
     |     |          |    |                  |    |          |     |
     |     |----------| PE2|                  | PE4|--------- |     |
     +-----+    |     |....|.......PW2........|....|     |    +-----+
                AC    +----+                  +----+     AC
```

Figure 2. Two-side PW dual-homing protection

Figure 2 illustrates the network scenario of two-side CE dual-homing. CE1 is dual-homed to PE1 and PE2, and CE2 is dual-homed to PE3 and PE4. This topology can protect the node failure of the dual-homing PEs on both sides, and also protects the AC link failures between the CEs and their dual-homing PEs. Meanwhile, dual-homing PW protection can protect the failure occured in the PSN network. This scenario is mainly used for services providing for important business customers. In this case, CE1 and CE2 can be regarded as service access points.

3. Overview of the Proposed Solution

The linear protection mechanisms for MPLS-TP network are defined in [RFC6378] [I-D.ietf-mpls-tp-psc-itu]. When such mechanisms are applied to PW linear protection, both the working PW and the protection PW need to terminate on the same PE nodes. This section extends these mechanisms to provide dual-homing protection for MPLS-TP PWs.

With MPLS-TP PW dual-homing protection mechanism, the linear protection mechanisms on the Single-homing PE (e.g. PE3 in figure 3) are 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 3) respectively, to provide protection for the dual-homing PEs and the connected ACs. A dedicated Dual-Node Interconnection (DNI) PW is established between the two dual-homing PE nodes, which is used to bridge the PW traffic when failure happens in the working PW or the primary AC. In order to make the linear protection mechanism work under the dual-homing PEs scenario, some coordination between the dual-homing PE nodes is necessary. The protection switching mechanism is detailed in following sections.

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

Figure 3. Dual-homing Proctection with DNI PW

The following failure scenarios can be protected by MPLS-TP PW dual-homing protection:

o Failure of the working PW (PW1)

o Failure of the working AC (AC1)

o Failure of the working PE (PE1)

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

In order to achieve the MPLS-TP PW dual-homing protection, coordination is needed between the dual-homing PE nodes to communicate the service PW status and protection coordination requests. The forwarding state of dual-homing PEs is also specified.

4.1. Information Exchange Between Dual-Homing PEs

A new Channel Type is defined for the coordination between the dual-homing PEs of MPLS-TP PWs. Such channel type can be used for the exchange of different kinds of information. This document uses this channel type for the PW status exchange and protection coordination between the dual-homing PEs. Other use cases 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 an 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 Group ID                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         TLV  Length           |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              TLVs                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       Figure 4. 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.

2 TLVs are defined in MPLS-TP Dual-Homing Coordination message for MPLS-TP PW dual-homing 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.

```
   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 Node_ID                       |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                       Source Node_ID                          |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                         DNI PW-ID                             |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                         Reserved                            |P|
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                      Service PW Request                   |D|F|
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                    Figure 5. PW Status TLV
```

- The Destination Node_ID is the 32-bit Node_ID of the receiving PE.

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

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

- The P (Protection) bit indicates whether the message is sent by the working PE (P=0) or by the protection PE (P=1).

- The Service PW Request field indicates the protection request generated on the Service PW between the sending PE and the remote PE. Two bits are defined in the Service PW Request field:

o F bit: Indicates Signal Fail (SF) request 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: Indicates Signal Degrade (SD) request is 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 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 Node_ID                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Source Node_ID                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          DNI PW-ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Reserved                         |S|P|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6. Dual-node Switching TLV

- The Destination Node_ID is the 32-bit Node_ID of the receiving PE.

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

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

- The P (Protection) bit indicates whether the message is sent by the working PE (P=0) or by the protection PE (P=1). With the mechanism described in this document, only the protection PE could send out the DHC message with the Dual-node Switching TLV.

- The S (PW Switching) bit indicates which service PW is used for transporting user traffic. It is set to 0 when traffic is 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.

4.2. Protection Procedures

MPLS-TP PW dual-homing protection mechanism can work with the existing AC side redundancy mechanisms, e.g. MC-LAG. On PSN network side, PSN tunnel protection mechanism is not required, as the dual-homing PW protection can also protect the failure happened in the PSN network.

For the single-homing PE, it just treats the working PW and protection PW as if they terminate on the same remote PE node, thus normal protection coordination mechanisms still apply to the single-homing PE.

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

```
              +------------+             +------------------+
              |   PE3      |             |      PE1         |
              |       +---+| service PW1 |+---+        +---+|
              |       |PW1|--------------|PW1|        |AC1|\
              |       +---+| working PW  |+---+        +---+|\
              |            |             |     +-------+    | \
              |            |             |     |DNI PW|     |  \  AC1
              |            |             |     +-------+    |   \(active)
              |            |             +--------||--------+    \
 +-----+  AC3 |            |                      ||            +--\---+
 | CE2 |------|            |                      || DNI PW     |  CE1 |
 +-----|      |            |                      ||            +--/---+
              |            |             +--------||--------+    /
              |            |             |     +-------+    |   /
              |            |             |     |DNI PW|     |  / AC2
              |            |             |     +-------+    | /(standby)
              |            |             |                  |/
              |       +---+| service PW2 |+---+        +---+/
              |       |PW2|--------------|PW2|        |AC2||
              |       +---+|protection PW|+---+  PE2   +---+|
              +------------+             +------------------+
```

Figure 7. Components of PW Dual-Homing protection

In figure 7, for a dual-homing PE, Service PW is the PW which carries service between dual-homing PE and the remote PE. The status of

Service PW is determined by the OAM and protection switching coordination mechanisms between the dual-homing PEs and the remote PE.

DNI PW is the PW between the two dual-homing PE nodes. It is used to bridge traffic when failure occurs in PSN network or in the AC side. The status of DNI PW is determined by OAM protocol running between the dual-homing PEs.

AC is the link which connects the dual-homing PEs to the dual-homed CE. The AC status is determined by MC-LAG or other AC redundancy mechanisms.

The PW status and protection coordination requests are exchanged between the dual-homing PEs using the DHC message defined in section 4.1.

After the exchange of PW status information using the MPLS-TP DHC Message, both dual-homing PEs obtain the status of working and protection service PWs, the AC and the DNI PW. The forwarding behavoir of dual-homing PE nodes are determined by the forwarding state machine 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

In normal state, the working PW is in active state, and the primary AC is active state, according to Table 1, PW traffic will be forwarded between the working service PW and the primary AC (AC1). No traffic will go through the protection PE or the DNI PW, as both the protection service PW and the AC connecting to the protection PE are in standby state.

If AC1 goes down due to some fault, the AC side redundancy mechanism would switchover to the backup AC (AC2), the state of AC2 changes to active. There is no change in the status of working and protection PW. According to Table 1, PE1 starts to forward traffic between the working PW and the DNI PW, while PE2 starts to forward traffic between DNI PW and AC2. Note that in this case only AC switchover takes place, in PSN network traffic keeps transporting on the working PW and PW switchover is no needed.

If the working PW is down due to some fault in the PSN network, both the remote PE (PE3) and the working PE (PE1) would detect the failure using MPLS-TP OAM mechanisms, then PE3 would send a normal protection coordination message on the protection path to inform its peer node (PE2) to switchover to the protection PW. PE1 would also inform PE2 the working PW status (down) using the MPLS-TP DHC message. Then according to Table 1, PE2 starts to forward traffic between the protection PW and the DNI PW, and PE1 starts to forward traffic between the DNI PW and AC1.

If the working PE (PE1) goes down, both the remote PE(PE3) and the protection PE(PE2) would detect the failure using MPLS-TP OAM mechanisms, the state of AC1 would change to down, and the state of AC2 will change to active according to AC side redundancy mechansim. Then PE3 would send a normal protection coordination message on the protection path to inform its peer node (PE2) to switchover to the protection PW. Then according to table 1, PE2 starts to forward traffic between the protection PW and AC2.

5. IANA Considerations

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

This document creates a new registry called "MPLS-TP DHC TLVs" registry. 2 new TLVs are defined in this document:

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

6. 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.

7. References

7.1. Normative References

[I-D.ietf-mpls-tp-psc-itu]
Ryoo, J., Gray, E., Helvoort, H., D'Alessandro, A., Cheung, T., and E. Osborne, "MPLS Transport Profile (MPLS-TP) Linear Protection to Match the Operational Expectations of SDH, OTN and Ethernet Transport Network Operators", draft-ietf-mpls-tp-psc-itu-04 (work in progress), March 2014.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC6372] Sprecher, N. and A. Farrel, "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, September 2011.

[RFC6378] Weingarten, Y., Bryant, S., Osborne, E., Sprecher, N., and A. Fulignoli, "MPLS Transport Profile (MPLS-TP) Linear Protection", RFC 6378, October 2011.

7.2. Informative References

[I-D.ietf-pwe3-endpoint-fast-protection]
Shen, Y., Aggarwal, R., Henderickx, W., and Y. Jiang, "PW Endpoint Fast Failure Protection", draft-ietf-pwe3-endpoint-fast-protection-00 (work in progress), December 2013.

[RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

[RFC6718] Muley, P., Aissaoui, M., and M. Bocci, "Pseudowire Redundancy", RFC 6718, August 2012.

[RFC6870] Muley, P. and M. Aissaoui, "Pseudowire Preferential Forwarding Status Bit", RFC 6870, February 2013.

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