Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

WQ. Cheng
L. Wang
H. Li
China Mobile
K. Liu
Huawei Technologies Co., Ltd.
S. Davari
Broadcom Corporation

October 21, 2013

MPLS-TP PWE3 dual-homed protection (MPDP) draft-cheng-mpls-tp-pwe3-dual-homed-protection-00

Abstract

This document presents the requirements for Dual-homed protection in the MPLS-TP networks and defines a protocol that can protect the failure of an attachment circuit (AC) or the failure of a provider edge (PE) node or the failure of a pseudowire (PW) in the packet-switched network (PSN).

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 24, 2014.

Table of Contents

1. Introduction

The linear protection and Ring protection mechanisms for MPLS-TP is described in RFC 6378, RFC 6974 and other IETF drafts. These mechanisms work within the PSN and provide fast recovery when link failure or P node failure occurs. However, they are unable to protect against the failure of a PE node or the failure of an attachment circuit.

The PW redundancy solution which is defined by RFC 6718 requires separate mechanisms to recover the PE and AC link failure from the PSN failure.

The operators need an end-to-end network's survivability for guaranteed services, so the protection mechanisms for AC,PE and failure within PSN are all needed. In order to meet the requirement, multiple layers and across nested recovery domains protection should be deployed. It raised the following issues:

o longer recovery time, because hold-off time should be set to avoid race scenarios, which makes the switching time longer.

o lower bandwidth efficiency because of multi-layer protections.

o extra configuration makes the operation and maintenance more complicated.

o if RFC 6718 is used, the AC link failure will result in protection switching performed within PSN, that means the failure of an AC are propagated to the remote PEs on the other side of the network.

In order to improve on RFC 6718, dual-homed protection mechanism should meet the following requirements

O Using a single layer protection for PSN,AC and PE failure

PWE3 Dual-Homed protection needs to recover PE failures, tunnel failures within PSN and AC link failures through a single layer protection mechanism so that the multi-layer protection can be avoided.

O Independent failure recovery

The principle of independent failure recovery is that the protection switching is solely performed within the network domain where the failure takes place. For instance, When a failure in a an AC happens, there is no need to inform the remote PE about the failure and there is no need to change the PW and path in the PSN.

O To deploy dual-homed network protection, as far as protocols which PE previously support, such as linear protection protocol, can be reused, upgrading remote PEs should be avoided.

According to RFC5654 "2.5.6. Topology-Specific Recovery Mechanisms", "MPLS-TP MAY support recovery mechanisms that are optimized for specific network topologies. These mechanisms MUST be interoperable with the mechanisms defined for arbitrary topology (mesh) networks to enable the protection of end-to-end transport paths ",this document presents a single layer dual-homed protection to meet those requirements.

2. Application scenarios of Dual-Homed protection

The application scenarios of Dual-homed protection can be classified into a One-side Dual-Homing topology and a Two-side Dual-homing topology.

2.1. One-side Dual-Homing topology

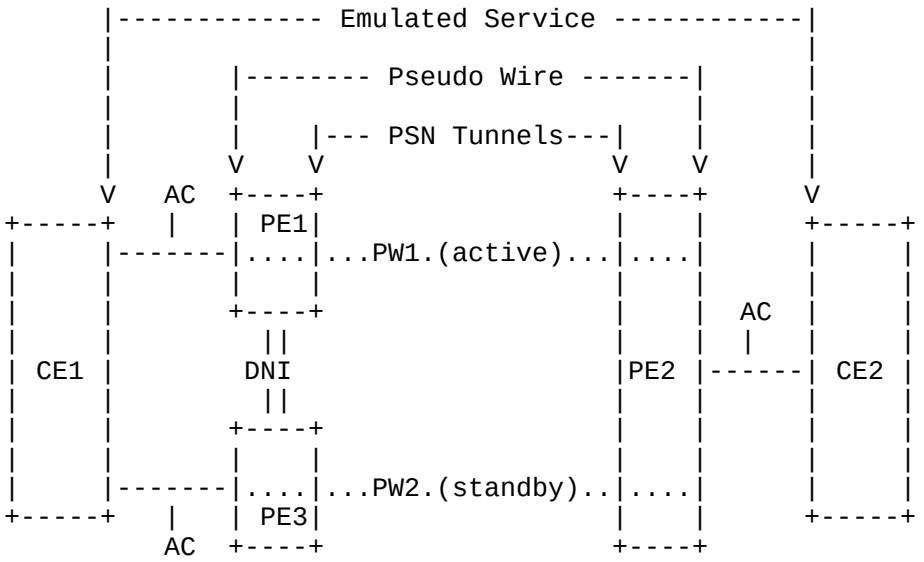
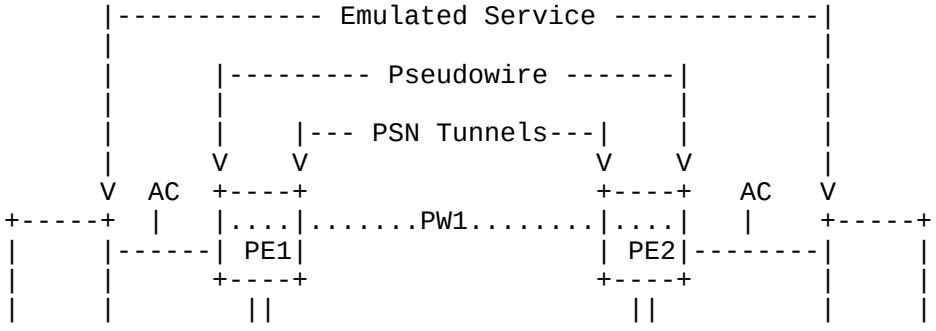


Figure 1 One-side PW Dual-Homing protection

Figure 1 illustrates the network scenario of one-side CE dual-homing topology. The dual-homing gateways for CE1 are PE1 and PE3, while PE2 is the single-homing for CE2. This scheme protects the node failures of PE1 and PE3 and the link failures between PE1 and CE1/PE3 and CE1. This scheme can be used in back hauling application scenarios. For example, nodeB serves for CE2 while RNC serves for CE1. PE2 works as an access layer MPLS-TP device while PE1 and PE3 works as a a pair of core layer MPLS-TP devices.

2.2. Two-side Dual-Homing topology



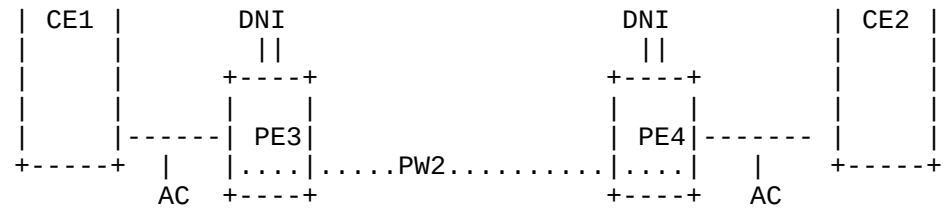


Figure 2 dual-side PW Dual-Homing protection

Figure 2 illustrates the network scenario of two-side CE dual-homing. The dual- homing nodes are PE1 and PE3 for CE1, and PE2 and PE4 for CE2, respectively. This scenario protects the PE1/PE3 and PE2/PE4 nodes failures. It also protects the links failure between PE1 and CE1, PE3 and CE1, PE2 and CE2, and PE4 and CE2. Meanwhile, dual-homing protection needs to handle the recovery of PSN Tunnel failure as well. As for broadband services provider, this scenario is mainly used in services for important business customers. Here, CE1 and CE2 can be regarded as service access points.

3. PWE3 dual-homed protection mechanism

In a PWE3 dual-homed protection mechanism, Multi-chassis PW protection is used between PEs, Multi-chassis LAG is used between dual-homed PE node group.

3.1. Multi-chassis PW protection

RFC6738(MPLS Transport Profile (MPLS-TP) Linear Protection) and ITU-T G.8131 have defined linear protection mechanism for MPLS-TP network. The PEs of working PW are the same as its protecting PW and the protection switching mechanism is running on each PE.

Dual-homed PW protection mechanism keeps the same protection switching mechanism with linear protection in the Remote PE (PE3 in figure 3) and the working PW and protecting PW are terminated in Dual-homed PEs (PE1 and PE2 in figure 3) respectively in order to protect Dual-homed PEs.The protection switching mechanism will be detailed in the following chapters.

Dual nodes interconnection (DNI) PW is set up between two dual-homed PE nodes, and it is used to bridge traffic when failure occurs.

Messages between two dual-homed node include channel status notify message and protection group status notify message. The format of these messages is TBD.

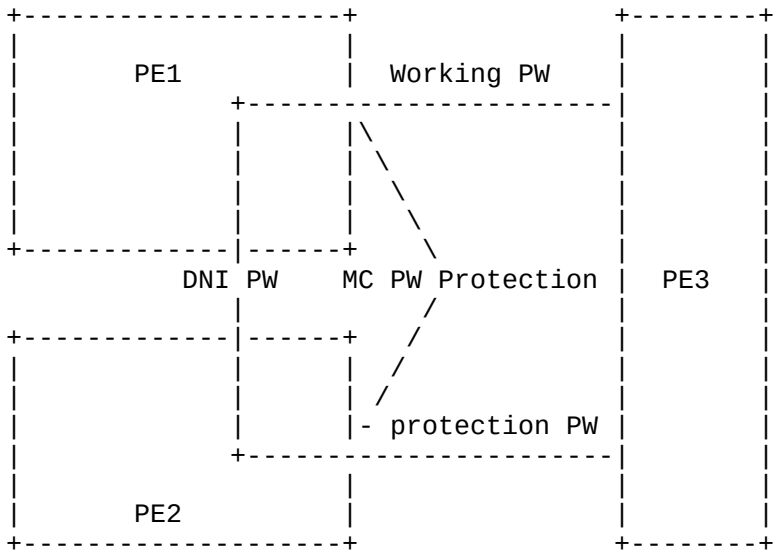


Figure 3 MC PW linear protection mechanism

The failure scenarios are listed as follows:

O One direction Failure of Working PW (PE1 to PE3):

PE3 detects failure and sends PSC or APS message in protection PW. When PE2 receives the failure information, it will exchange a switching message with PE3. At last, PE2 and PE3 will switch the traffic to the protection PW. PE2 will periodically send PE1 MC PW protection group status messages,and then PE1 will execute the switching according to the status of the MC PW protection.

O One direction Failure of Working PW (PE3 to PE1):

PE1 will detect the failure and send PE2 a MC PW protection message to notify work PW failure through DNI PW. PE2 will execute the switch with PE3 based on PSC or APS. PE2 will periodically send PE1 MC PW protection group status messages, and PE1 will execute switching according to the status of the MC PW protection.

O Bi-direction Failure of Working PW (Between PE3 and PE1):

Both of PE1 and PE3 will detect link fault respectively. PE3 executes switching to protection PW based on APS or PSC. PE1 sends PE2 a MC PW protection message to notify working PW failure through DNI PW,and then PE2 will execute switching to protection PW. PE2

will periodically send PE1 MC PW protection group status messages, and PE1 will execute switching according to the status of the MC PW protection.

O Working PE failure:

PE3 will detect failure, and send PSC or APS message in the protection PW. After PE2 exchanges the switching message with PE3, PE2 and PE3 will switch traffic to the protection PW.

3.2. Multi-chassis LAG

LAG(Link Aggregation Group) and LACP(Link Aggregation control protocol) is defined in IEEE802.1ax. LAG is used to expand bandwidth and protect link failure.

DRNI (Distributed Resilient Network Interconnect) and DRCP (Distributed Relay Control Protocol) is defined in IEEE P802.1AX-REV-D2.2. DRNI expands the concept of Link Aggregation so that, at either one or both ends of a Link Aggregation Group, the single Aggregation System is replaced by a Portal, composed from one to three Portal Systems.

In the document, DRNI is used to protect Ethernet link failure between CE and PE and dual-homed node Failure on the CE side as shown in figure 4.

O Working PE failure: Detailed signaling between PE1, PE2, CE is TBD

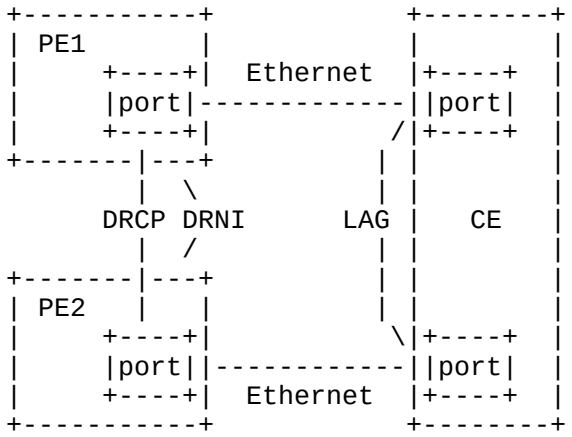


Figure 4 DRNI mechanism

4. Three point-switch collaboration

In dual-homed PE nodes, the protection mechanism in PSN network(MC PW protection) and protection mechanism between PE and CE(DRNI) should cooperate to ensure an appropriate switch operation.

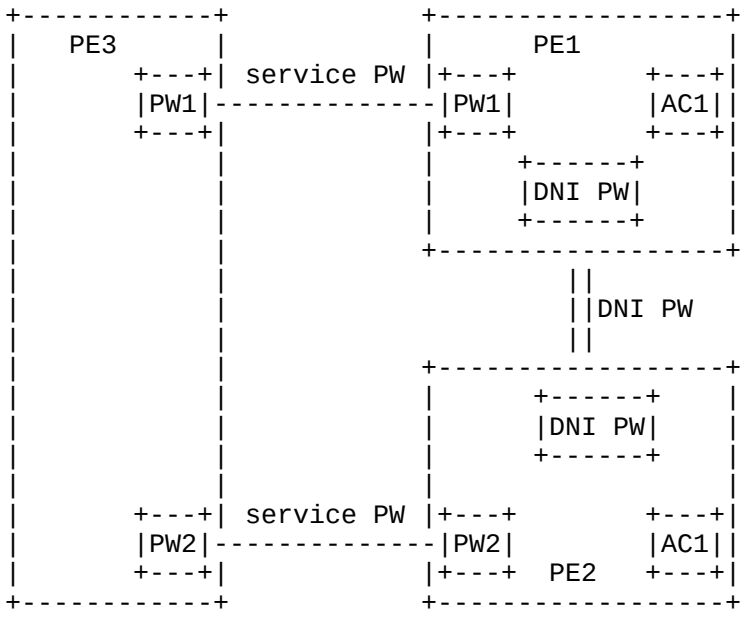


Figure 6 three point status machine

Service PW is the PW which carry service between dual-homed PE and remote PE. Service PW status is decided by MC PW protection mechanism.

DNI PW is the bridge PW between two dual-homed PE nodes. It is used to bridge traffic when PSN tunnel or AC failure occurs.

AC status is the status of AC port between dual-homed PW and CE, being either active or standby. It is decided by DRNI mechanism.

srv PW	AC	PW fwd	DNI fwd	AC fwd
Active	Active	to AC	to PW	to PW

Active	Standby	to DNI	to PW	to PW
Standby	Active	to AC	to AC	to DNI
Standby	Standby	to DNI	to AC	to DNI

Figure 7 three point status machine in dual-homed nodes

The principle of three point status machine in dual-homed nodes:

O If AC status is active, establish connection from service PW to AC. If AC status is standby, establish connection from service PW to DNI PW.

O If service PW is active, establish connection from AC to service PW. If service PW status is standby, establish connection from AC to DNI PW.

O If service PW is active, establish connection from DNI PW to service PW. If service PW status is standby, establish connection from DNI PW to AC.

5. Formal Syntax

None

6. Security Considerations

None

7. Acknowledgments

None

8. Author’s Addresses

None

9. References

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

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.

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

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

[RFC6974] Weingarten, Y., Bryant, S., Ceccarelli, D., Caviglia, D., Fondelli, F., Corsi, M., Wu, B., and X. Dai, "Applicability of MPLS Transport Profile for Ring Topologies", RFC 6974, July 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 Co., Ltd.
Huawei base, Bantian, Longgang District
Shenzhen 518129
China

Email: alex.liukai@huawei.com

Shahram Davari
Broadcom Corporation
3151 Zanker Road
San Jose, CA 95134-1933

Email: davari@broadcom.com