

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 30, 2015

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 27, 2014

Dual-Homing Coordination for MPLS Transport Profile (MPLS-TP)
Pseudowires
draft-cheng-pwe3-mpls-tp-dual-homing-coordination-00

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-cheng-pwe3-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].

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, 2015.

Copyright Notice

Copyright (c) 2014 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	2
2. Overview of the Proposed Solution	3
3. Protocol Extensions for MPLS-TP PW Dual-Homing Protection . .	4
3.1. Information Exchange Between Dual-Homing PEs	4
3.2. Protection Procedures	7
4. IANA Considerations	10
5. Security Considerations	10
6. References	10
6.1. Normative References	10
6.2. Informative References	11
Authors' Addresses	12

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 Attachment Circuit (AC) or the endpoint nodes.

In some scenarios such as mobile backhauling, the MPLS PWs are provisioned with dual-homing topology, in which at least the CE node

in 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.cheng-pwe3-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

The linear protection mechanisms for MPLS-TP network are defined in [RFC6378], [RFC7271] and [RFC7324]. When such mechanisms are applied to PW linear protection, both the working PW and the protection PW terminate 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 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 1) respectively to protect the failure occurs in the dual-homing PEs and the connected ACs. As specified in [I-D.cheng-pwe3-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.

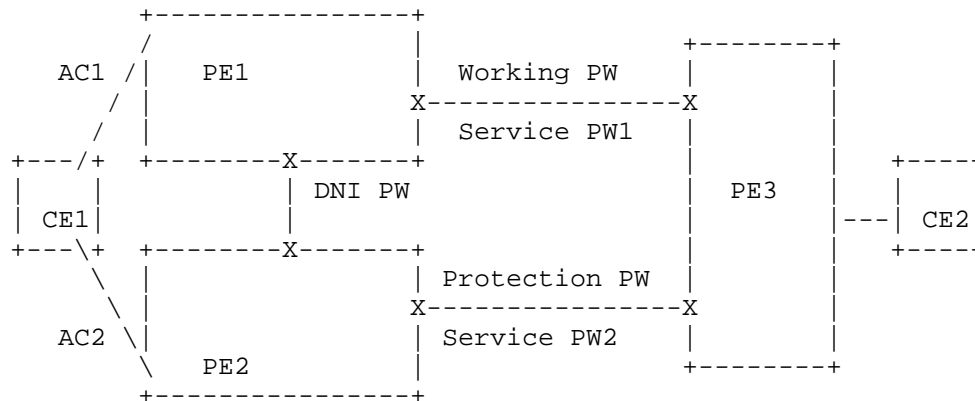


Figure 1. Dual-homing Protection with DNI-PW

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.cheng-pwe3-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:

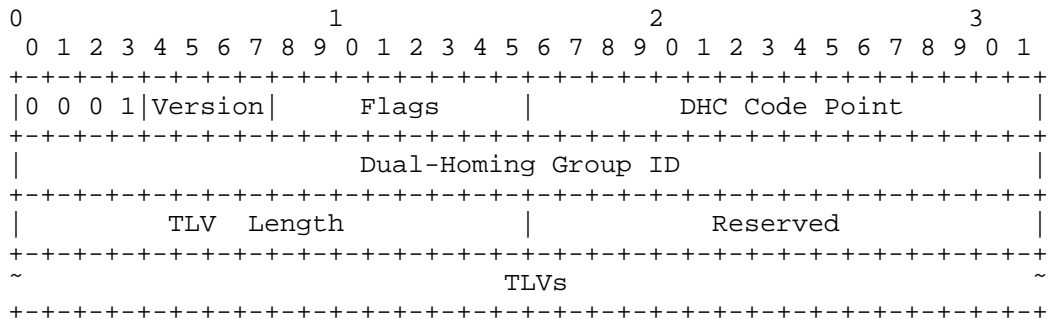


Figure 2. MPLS-TP Dual-Homing Coordination Message

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, 2 TLVs are defined in 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.

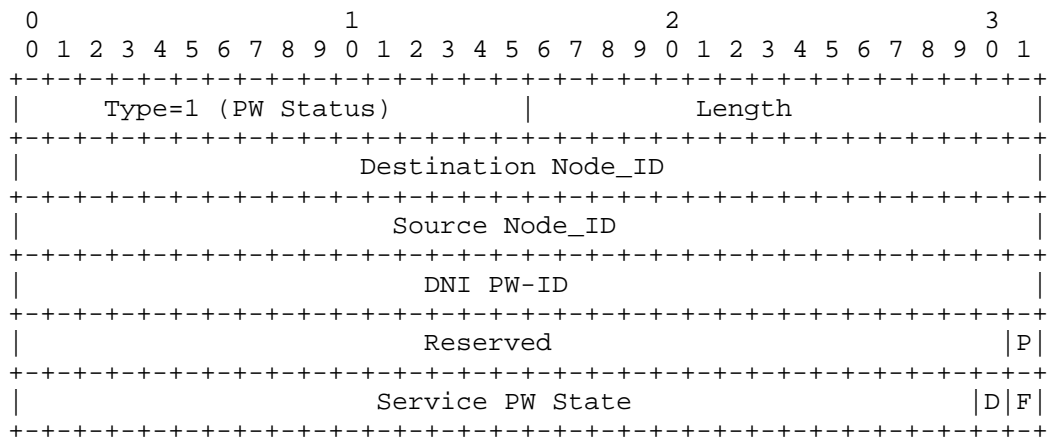


Figure 3. PW Status TLV

- The Destination Node_ID is the 32-bit Node_ID of the receiver PE.
- The Source Node_ID is the 32-bit Node_ID of the sending PE.
- The DNI PW-ID field contains the 32-bit 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 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: 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: 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.

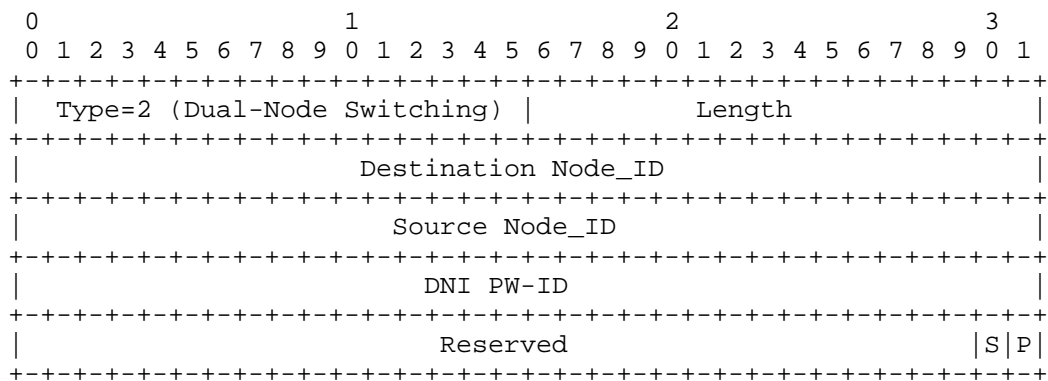


Figure 4. Dual-node Switching TLV

- 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 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.

The MPLS-TP DHC message is exchanged periodically between the dual-homing PEs. Whenever there is a change in the status of service PW on one dual-homing PE, it MUST be sent to the other dual-homing PE immediately using the PW status TLV in the DHC message. The Dual-Node Switching TLV is carried in the DHC message when a switchover request is issued by the protection PE according to the dual-homing forwarding state machine.

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.

On the 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 to the single-homing PE.

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

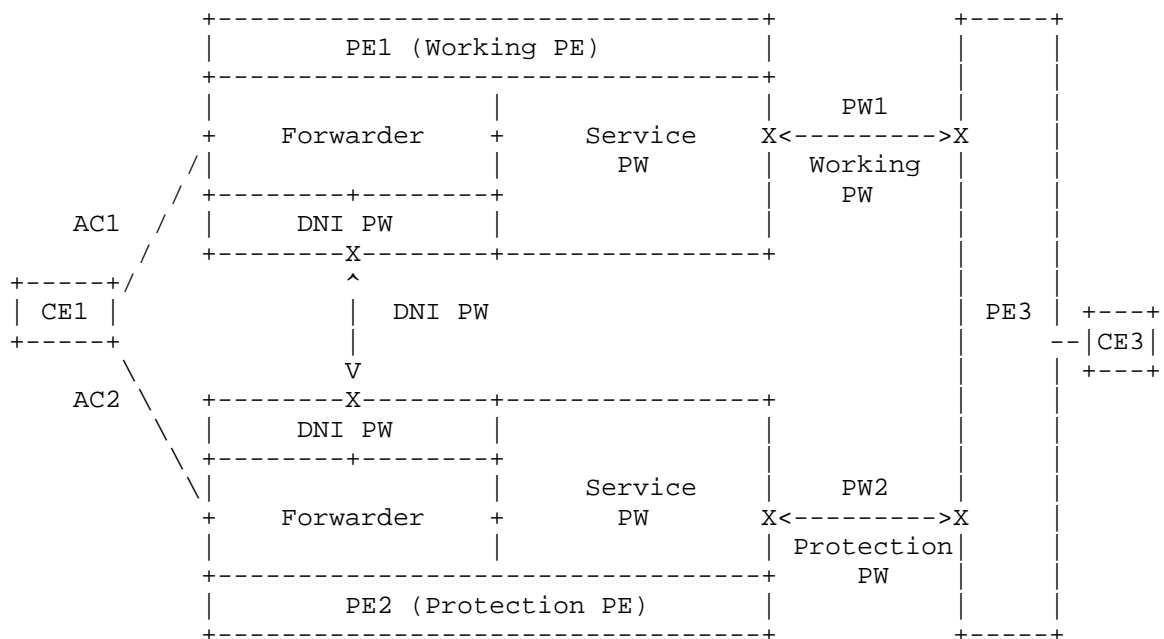


Figure 5. Components of PW dual-homing protection

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 status of service PW is determined by the OAM mechanism between the dual-homing PE and the remote PE.

DNI PW is the PW established 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 status of DNI PW is determined by OAM mechanism running between the dual-homing PEs. 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.

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

In order to perform dual-homing PW local protection, the service PW status and protection coordination requests need to be exchanged between the dual-homing PEs using the DHC message defined above.

Whenever there is a change in the status of service PW on the dual-homing PE, it MUST be sent to the other dual-homing PE in the same dual-homing group immediately using the PW status TLV in the DHC message. After the exchange of PW status, both the dual-homing PEs

could obtain the status of the working and protection service PWs. The status of DNI PW is determined by the OAM mechanisms between the dual-homing PEs, and the status of AC is determined by the AC redundancy mechanism. The protection PE SHOULD make the switchover decision according to the status of the connected AC, service PW and DNI PW, and SHOULD send the switchover request to the working PE using the Dual-node switching TLV in the DHC message. 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

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 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 state.

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 status of the working and protection PW. According to the forwarding state machine in Table 1, PE1 starts to forward traffic between the working PW and the DNI PW, while 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 the traffic is still forwarded using the working PW, PW switchover is not needed.

If some failure occurs in the PSN network which causes PW1 down, the working PE (PE1) or the remote PE (PE3) can detect the failure using MPLS-TP OAM mechanism. If PE1 detects the failure, it MUST inform PE2 the status of the working PW using the PW Status TLV in MPLS-TP DHC message. According to the forwarding state machine in Table 1, PE2 SHOULD set the connection between PW2 and the DNI PW, and PE1 SHOULD set the connection between the DNI PW and AC1. For switchover

coordination, PE2 MUST send a DHC message to PE1 with the S bit in the Dual-node switching TLV set, and send an appropriate protection coordination message on the protection PW (PW2) to PE3 for the remote side switchover from PW1 to PW2. Upon receipt of Dual-node switching TLV in the DHC message, PE1 MUST switch the traffic onto the connection between DNI PW and AC1. If PE3 detects the failure in PW1, it would send a protection coordination message on the protection PW (PW2) to inform PE2 to switchover to the protection PW. And PE2 MUST 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.

If some failure causes the working PE (PE1) down, both the remote PE (PE3) and the protection PE (PE2) would detect the failure using MPLS-TP OAM mechanisms. The status of AC1 changes to standby, and the state of AC2 changes to active according to AC redundancy mechanism. PE3 would send a protection coordination message on the protection path to inform its peer node (PE2) to switchover to the protection PW. 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. 2 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.cheng-pwe3-mpls-tp-dual-homing-protection]
Cheng, W., Wang, L., Li, H., Liu, K., Davari, S., and J. Dong, "Dual-Homing Protection for MPLS Transport Profile (MPLS-TP) Pseudowires", draft-cheng-pwe3-mpls-tp-dual-homing-protection-00 (work in progress), July 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", RFC 5586, June 2009.
- [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.
- [RFC7271] Ryoo, J., Gray, E., 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, June 2014.
- [RFC7324] Osborne, E., "Updates to MPLS Transport Profile Linear Protection", RFC 7324, July 2014.

6.2. 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-01 (work in progress), July 2014.
- [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

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

Email: alessandro.dalessandro@telecomitalia.it