

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: June 27, 2014

Mingui Zhang
Huafeng Wen
Huawei
Jie Hu
China Telecom
December 24, 2013

STP Application of ICCP
draft-ietf-pwe3-iccp-stp-00.txt

Abstract

Inter-Chassis Communication Protocol (ICCP) supports the inter-chassis redundancy mechanism which achieves high network availability.

In this document, the PEs in a Redundant Group (RG) running ICCP are used to offer multi-homed connectivity to Spanning Tree Protocol (STP) networks. The ICCP TLVs for the STP application are defined, therefore PEs from the RG can make use of these TLVs to synchronize the state and configuration data of the STP network.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/1id-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2013 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
1.1.	Conventions used in this document	3
1.2.	Terminology	3
2.	The Use Case Scenario	3
2.1.	STP as an Application of ICCP	4
2.2.	A Private Solution: The BPDU Tunneling Model	4
3.	Spanning Tree Protocol Application TLVs	6
3.1.	STP Connect TLV	6
3.2.	STP Disconnect TLV	7
3.2.1.	STP Disconnect Cause TLV	8
3.3.	STP Config TLVs	8
3.3.1.	STP System Config	8
3.3.2.	STP Topology Changed Instances	9
3.3.3.	STP CIST Root Time	10
3.3.4.	STP MSTI Root Time	11
3.3.5.	STP Region Name	12
3.3.6.	STP Revision Level	12
3.3.7.	STP Instance Priority	13
3.3.8.	STP Configuration Digest	14
3.4.	STP Synchronization Request TLV	14
3.5.	STP Synchronization Data TLV	16
4.	Security Considerations	17
5.	IANA Considerations	17
	Acknowledgements	17
6.	References	17
6.1.	Normative References	17
6.2.	Informative References	17
	Author's Addresses	19

1. Introduction

Inter-Chassis Communication Protocol (ICCP) specifies a multi-chassis redundant mechanism, which enables PEs located in multi-chassis to act as a single Redundant Group (RG).

This document introduces Spanning Tree Protocol (STP) as a new application of ICCP. When a bridge network running STP is connected to a RG, the RG members should pretend to be a single root bridge to participate the operations of the STP. STP relevant information need be exchanged and synchronized among the RG members. ICCP TLVs for the Spanning Tree Protocol application are specified for this purpose.

1.1. Conventions used in this document

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](#)].

1.2. Terminology

STP: Spanning Tree Protocol
MSTP: Multiple Spanning Tree Protocol
DSLAM: Digital Subscriber Line Access Multiplexer
MST: Multiple Spanning Trees
CIST: Common and Internal Spanning Tree
MSTI: Multiple Spanning Tree Instance
BPDU: Bridge Protocol Data Unit

In this document, unless otherwise explicitly noted, when the term STP is used, it also covers MSTP.

2. The Use Case Scenario

In customers' broadband networks, bridged DSLAMs are usually geographically dispersed. It is a common case these DSLAMs are connected to carriers' L2VPN network at multiple points for the sake of reliability. Requirements from customers for this use case are listed as follows.

- o These DSLAMs are running STP.
- o These DSLAMs are not geographically close to each other. Multiple DSLAMs are connected to the carrier network at different locations.
- o When one connection to the carrier network fails, customers wish a connection in another location can continue to work after the re-

convergence of the STP rather than compromising the whole STP network. The failure of the connection may be due to the failure of the PE, the AC or even the CE (a DSLAM) itself.

- o Customers want to balance the load among those connections, therefore all those connections need be active.

In order to meet these requirements, the 'ICCP-STP' model is proposed in [Section 2.1](#). It introduces STP as a new application of ICCP. A private implementation is depicted in [Section 2.2](#) to serve as a contrast. Its issues are analyzed.

[2.1. STP as an Application of ICCP](#)

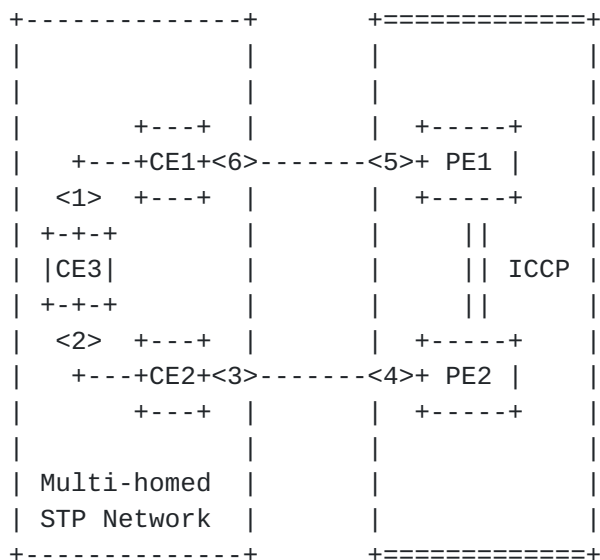


Figure 2.1: A STP network is multi-homed to an RG running ICCP.

Figure 2.1 shows an example topology of this model. With ICCP, the whole RG will be virtualized to be a single bridge. The RG pretends that the ports connected to the STP network (port <4>, <5>) are from the same bridge. All these ports emit configuration BPDU with the highest root priority to trigger the construction of the spanning tree. In this way, the STP will always broken a loop within the multi-homed STP network.

Each RG member has its BridgeIdentifier (the MAC address). The least significant one is elected as the BridgeIdentifier of the 'virtualized root bridge'.

[2.2. A Private Solution: The BPDU Tunneling Model](#)

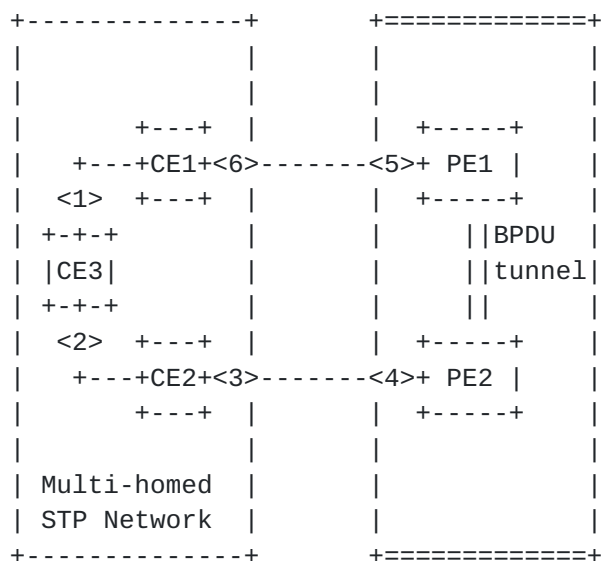


Figure 2.2: The BPDUs Tunneling Model

Figure 2.2 shows an example network of the 'BPDUs tunneling' model. Two PEs tunnel BPDUs of the STP network over PW. The OAM designed in [RFC 7023](#) can be adopted for the interworking between MPLS and Ethernet.

In this model, the ports connected to the STP network at the PEs' side are non-bridge ports (e.g., port <4>, <5>). The tunnel between PE1 and PE2 is a transparent tunnel of BPDUs. For CE1 and CE2, they regard that there is a direct link between them. Issues of this model are listed as follows.

- o Assume port <1> was blocked according to the STP calculation. Now, suppose link CE3-CE2 fails, port <2> is blocked while port <1> is unblocked. Since the BPDUs is just tunneled, PE2 is unaware of this change. It may continue to send traffic to CE3 via CE2 where a black-hole happens.

In order to handle this issue, PEs have to snoop the Topology Change (TC) message of the STP network so the tunnel is not "transparent" to BPDUs anymore. When the TC event is sensed, these PEs should withdraw MAC addresses of those instances affected by the TC event across the carrier's network.

- o When port <4> fails, CE1 is unaware. Suppose CE2 is the root bridge, port <6> has to wait for 3 STP HELLO Intervals (3*2s) and 2 Forwarding Delays (2*15s). The STP network takes at least 36 seconds to complete the convergence. The convergence process of the STP network is greatly slowed down. This actually changes the behavior of customers' STP networks.

In order to speed up the convergence, CE1 has to sense the state of the link between CE2 and PE2. The interworking PEs and CEs have to cope with various combinations of failures. There is no standard solution yet.

- o [RFC 7023](#) is scoped to only single segment PWS [[RFC6310](#)][[RFC 7023](#)]. When the STP network is attached to more than two PEs, a full mesh PWS have to be set up. It's not clear yet what behaviors these PEs should have.

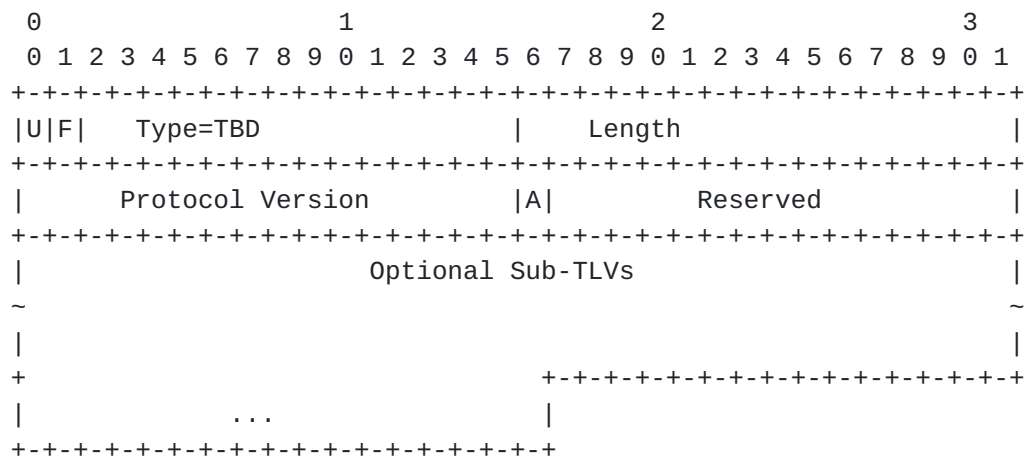
Compared with this 'BPDU tunneling' model, the advantage point of the "ICCP-STP" model is that PEs "participate" in the STP calculation, therefore we need not design the complex interworking mechanism.

3. Spanning Tree Protocol Application TLVs

This section discusses the ICCP TLVs for the Spanning Tree Protocol application.

3.1. STP Connect TLV

This TLV is included in the RG Connect message to signal the establishment of STP application connection.



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Connect TLV"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Protocol Version

The version of this particular protocol for the purposes of ICCP. This is set to 0x0001.

- A bit

Acknowledgement Bit. Set to 1 if the sender has received a STP Connect TLV from the recipient. Otherwise, set to 0.

- Reserved

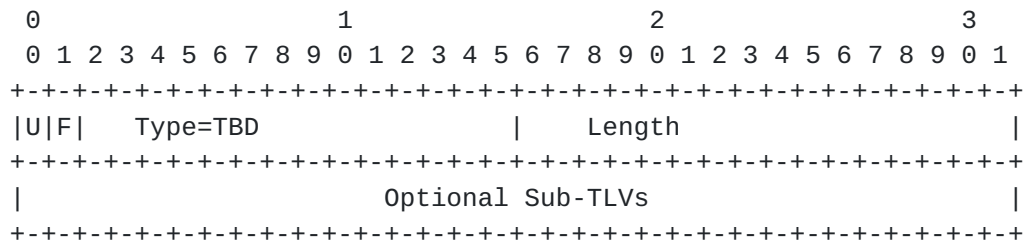
Reserved for future use.

- Optional Sub-TLVs

There are no optional Sub-TLVs defined for this version of the protocol.

3.2. STP Disconnect TLV

This TLV is used in an RG Disconnect Message to indicate that the connection for the STP application is to be terminated.



- U and F Bits

Both are set to 0.

- Type

```
set to TBD for "STP Disconnect TLV"
```

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Optional Sub-TLVs

The only optional Sub-TLV defined for this version of the protocol is the "STP Disconnect Cause" TLV defined next:

3.2.1. STP Disconnect Cause TLV

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F|   Type=TBD                               |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Disconnect Cause String          |
~                                                                    ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Disconnect Cause TLV"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Disconnect Cause String

Variable length string specifying the reason for the disconnect. Used for network management.

3.3. STP Config TLVs

The STP Config TLVs are sent in the RG Application Data message. When a STP Config TLV is received by a peering RB member, it SHOULD synchronize the configuration information contained in the TLV. TLVs specified from [section 3.3.1](#) through [section 3.3.9](#) contains such kind of configuration information.

3.3.1. STP System Config

This TLV announces the local node's STP System Parameters to the RG peers.


```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F|   Type=TBD                               |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     InstanceID List                    |
~                                                                           ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Topology Changed Instances"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- InstanceID List

The list of the instances whose topology is changed as indicated by the Topology Change Notification (TCN) Messages as specified in [\[802.1q\]](#) [section 13.14](#).

3.3.3. STP CIST Root Time

This TLV is used to report the Value of CIST Root Time to other members in the RG.

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F|   Type=TBD                               |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   MaxAge   |   MessageAge   |   FwdDelay   |   HelloTime   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| RemainingHops |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP CIST Root Time"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- MaxAge

The Maximum Age of this TLV.

- MessageAge

The actual age of this TLV.

- FwdDelay

The delay before the port enters the forwarding status.

- HelloTime

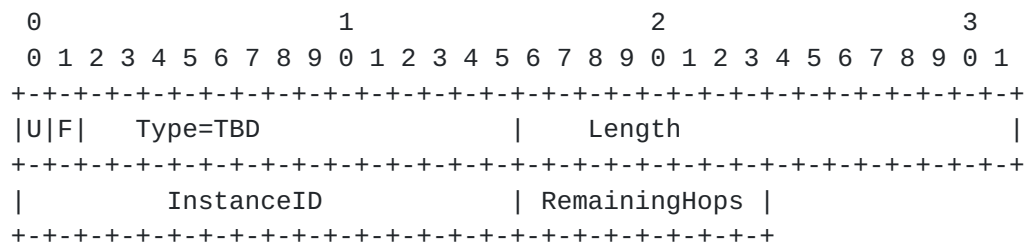
The interval between two continuous configuration BPDUs.

- RemainingHops

The remaining hops of this TLV

3.3.4. STP MSTI Root Time

This TLV is used to report the Value of MSTI Root Time to other members in the RG.



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP MSTI Root Time"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- InstanceID

The instance identification number of the MSTI.

- remainingHops

The remaining hops of this TLV

3.3.5. STP Region Name

This TLV is used to report the Value of Region Name to other members in the RG.

```

      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
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|U|F|  Type=TBD                |  Length                |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|                                Region Name              |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Region Name"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Region Name

The Name of the MST Region.

3.3.6. STP Revision Level

This TLV is used to report the Value of Revision Level to other members in the RG.


```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F|   Type=TBD                               |   Length                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Revision Level                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Revision Level"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Revision Level

The Revision Level as specified in [[802.1q](#)] [section 3.21](#);

3.3.7. STP Instance Priority

This TLV is used to report the Value of Instance Priority to other members in the RG.

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F|   Type=TBD                               |   Length                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Pri |   InstanceID                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Instance Priority"

- Length

- Pri

- InstanceID

The instance identification number of the MSTI.

3.3.8. STP Configuration Digest

[illegible]

- U and F Bits

- Type

```
set to TBD for "STP Configuration Digest"
```

- Length

Length of the STP Configuration Digest which is 16 octets.

- Configuration Digest

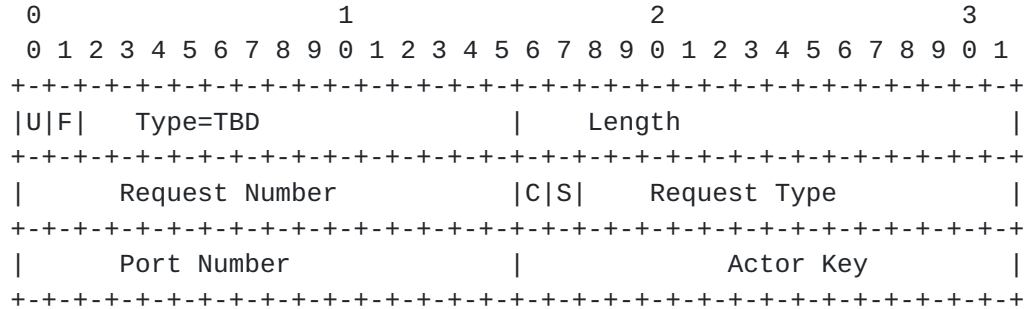
As specified in [\[802.1q\] section 13.7](#).

3.4.4. STP Synchronization Request TLV

The STP Synchronization Request TLV is used in the RG Application Data message. This TLV is used by a device to request from its peer to re-transmit configuration or operational state. The following information can be requested:

- system configuration and/or state
- configuration and/or state for a specific port

The format of the TLV is as follows:



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Synchronization Request TLV"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Request Number

2 octets. Unsigned integer uniquely identifying the request. Used to match the request with a response. The value of 0 is reserved for unsolicited synchronization, and MUST NOT be used in the STP Synchronization Request TLV.

- C Bit

Set to 1 if request is for configuration data. Otherwise, set to 0.

- S Bit

Set to 1 if request is for running state data. Otherwise, set to 0.

- Request Type

14-bits specifying the request type, encoded as follows:

0x00 Request System Data
0x01 Request Port Data
0x3FFF Request All Data

- Port Number

2 octets. When Request Type field is set to 'Request Port Data', this field encodes the STP Port Number for the requested port. When the value of this field is 0, it denotes that all ports, whose STP Key is specified in the "Actor Key" field, are being requested.

- Actor Key

2 octets. STP Actor key for the corresponding port. When the value of this field is 0 (and the Port Number field is 0 as well), it denotes that information for all ports in the system is being requested.

3.5. STP Synchronization Data TLV

The STP Synchronization Data TLV is used in the RG Application Data message. A pair of these TLVs is used by a device to delimit a set of TLVs that are being transmitted in response to an STP Synchronization Request TLV. The delimiting TLVs signal the start and end of the synchronization data, and associate the response with its corresponding request via the 'Request Number' field.

The STP Synchronization Data TLVs are also used for unsolicited advertisements of complete STP configuration and operational state data. The 'Request Number' field MUST be set to 0 in this case.

This TLV has the following format:

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Synchronization Data TLV"

- Length

Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Request Number

2 octets. Unsigned integer identifying the Request Number from the "STP Synchronization Request TLV" which solicited this synchronization data response.

- Flags

2 octets, response flags encoded as follows:

0x00 Synchronization Data Start

0x01 Synchronization Data End

4. Security Considerations

This document raises no new security issues.

5. IANA Considerations

The types used by the application TLVs defined in [Section 3](#) should be assigned.

Acknowledgements

The authors would like to thank the comments and suggestions from Gregory Mirsky.

6. References

[6.1. Normative References](#)

[ICCP] L. Martini, S. Salam, et al, "Inter-Chassis Communication Protocol for L2VPN PE Redundancy", [draft-ietf-pwe3-iccp-12.txt](#), work in progress.

[RFC7023] D. Mohan, Ed., N. Bitar, Ed., A. Sajassi, Ed., S. DeLord, P. Niger, R. Qiu., "MPLS and Ethernet Operations, Administration, and Maintenance (OAM) Interworking", [RFC 7023](#), October 2013.

[6.2. Informative References](#)

[RFC6310] Aissaoui, M., Busschbach, P., Martini, L., Morrow, M., Nadeau, T., and Y(J). Stein, "Pseudowire (PW) Operations, Administration, and Maintenance (OAM) Message Mapping", [RFC 6310](#), July 2011.

[802.1q] "IEEE Standard for Local and Metropolitan Area Networks---Virtual Bridged Local Area Networks.". IEEE Std 802.1 Q-2005, May 19, 2006.

Author's Addresses

Mingui Zhang
Huawei

Email: zhangmingui@huawei.com

Huafeng Wen
Huawei

Email: wenhuafeng@huawei.com

Jie Hu
China Telecom

Email: hujie@ctbri.com.cn

