INTERNET-DRAFT   Mingui Zhang
Intended Status: Proposed Standard   Huafeng Wen
Expires: February 5, 2016   Huawei
                                                                 Jie Hu
                                                          China Telecom
                                                         August 4, 2015

STP Application of ICCP
draft-ietf-pwe3-iccp-stp-04.txt

Abstract

Inter-Chassis Communication Protocol (ICCP) supports an inter-chassis redundancy mechanism which is used to support high network availability.

In this document, the PEs in a Redundancy Group (RG) running ICCP are used to offer multi-homed connectivity to Spanning Tree Protocol (STP) networks to improve availability of the STP networks. The ICCP TLVs and usage for the ICCP STP application are defined.

Table of Contents

1. Introduction

Inter-Chassis Communication Protocol (ICCP [RFC7275]) specifies a multi-chassis redundancy mechanism which enables PEs located in a multi-chassis arrangement to act as a single Redundancy Group (RG).

This document introduces support of Spanning Tree Protocol (STP) as a new application of ICCP. This STP application of ICCP supports when a bridged STP network is connected to a RG, the RG members act as a single root bridge participating in the operations of STP protocol. STP relevant information needs to be exchanged and synchronized among the RG members. New ICCP TLVs for the ICCP STP application are specified for this purpose.

From the point of view of the customer, the Service Providers is still providing a Virtual Private LAN Service (VPLS) [RFC4762].

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

ICCP: Inter-Chassis Communication Protocol
VPLS: Virtual Private LAN Service
STP: Spanning Tree Protocol
MSTP: Multiple Spanning Tree Protocol
MST: Multiple Spanning Trees
CIST: Common and Internal Spanning Tree ([802.1q] Section 3.4)
MSTI: Multiple Spanning Tree Instance ([802.1q] Section 3.22)
BPDU: Bridge Protocol Data Unit

In this document, unless otherwise explicitly noted, the term STP also covers MSTP.

2. Use Case

Customers widely use Ethernet as an access technology [RFC4762]. It's common that one customer's Local Area Network (LAN) has multiple bridges connected to a carrier's network at different locations for reliability purposes. Requirements for this use case are listed as follows.

o Customers desire to balance the load among their available connections to the carrier's network, therefore all the connections need be active.

o When one connection to the carrier network fails, customers require a connection in another location to 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 itself.

In order to meet these requirements, the 'ICCP-STP' model is proposed. It introduces STP as a new application of ICCP.

```
          +---------------+      +=============+
          |               |      |             |
          |               |      |             |
          |       +---+   |      |  +-----+|<--|--Pseudowire-->|
          |   +---+CE1+<6>-------<5>+ PE1 ||   |               |
          |  <1>  +---+   |      |  +-----+|<--|--Pseudowire-->|
          | +-+-+         |      |     ||      |
          | |CE3|         |      |     ||ICCP  |--> Towards the Core
          | +-+-+         |      |     ||      |
          |  <2>  +---+   |      |  +-----+|<--|--Pseudowire-->|
          |   +---+CE2+<3>-------<4>+ PE2 ||   |               |
          |       +---+   |      |  +-----+|<--|--Pseudowire-->|
          |               |      |             |
          | Multi-homed   |      |  Redundancy |
          | STP Network   |      |    Group    |
          +---------------+      +=============+
```

Figure 2.1: A STP network is multihomed to 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. Each RG member has its BridgeIdentifier (the MAC address). The numerically lowest one is used as the BridgeIdentifier of the 'virtualized root bridge'. The RG acts as if the ports connected to the STP network (port <4>, <5>) are for the same root bridge. All these ports send the configuration BPDU with the highest root priority to trigger the construction of the spanning tree. The link between the peering PEs is not visible to the bridge domains of the STP network. In this way, the STP will always break a possible loop within the multi-homed STP network by breaking the whole network into separate islands so that each is attached to one PE. That forces all PEs in the RG to be active. This is different from a generic VPLS [RFC4762] where the root bridge resides in the customer network and the multi-homed PEs act in the active-standby mode. Note that the specification of VPLS remains unchanged other than for this operation. For instance, a full-mesh of PWs is established between PEs, and split-horizon is still used to perform the loop-breaking through the core.

3. Spanning Tree Protocol Application TLVs

This section specifies the ICCP TLVs for the ICCP STP application. The Unknown TLV bit (U-bit) and the Forward unknown TLV bit (F-bit) of the following TLVs MUST be sent as cleared and processed on receipt as specified in [RFC7275].

3.1. STP Connect TLV

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

```
 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=TBA1               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Protocol Version         |A|         Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Optional Sub-TLVs                        |
~                                                               ~
|                                                               |
+                                +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             ...                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA1 (value to be assigned by IANA) 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 STP ICCP application protocol. This document defines version 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. These bits MUST be sent as zero and

ignored on receipt.

- 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 RG Disconnect Message to indicate that the connection for the ICCP STP application is to be terminated.

```
 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=TBA2               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Optional Sub-TLVs                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA2 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" sub-TLV, defined below:

3.2.1. STP Disconnect Cause sub-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=TBA13              |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Disconnect Cause String                  |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA13 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, to be used for operational purposes.

3.3. STP Config TLVs

The STP Config TLVs are sent in the RG Application Data message. When STP Config TLV is received by a peer RG member, it MUST synchronize the configuration information contained in the TLV. TLVs specified from Section 3.3.1 through Section 3.3.5 defines specific 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=TBA3               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              ROID                             |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         MAC Address                           |
+                               +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA3 for "STP System Config"

- Length

Length of the ROID plus the MAC address in octets. Always set to 14.

-ROID

Redundant Object Identifier, format defined in Section 6.1.3 of [RFC7275].

- MAC Address

The MAC address of the sender. This MAC address is set to the BridgeIdentifier of the sender, as defined in [802.1q] Section 13.23.2. The numerically lowest 48 bit unsigned value of BridgeIdentifier is used as the MAC address of the Virtual Root Bridge mentioned in Section 2.1.

3.3.2. STP Region Name

This TLV carries the value of Region Name.

```
 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=TBA4               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Region Name                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

set to TBA4 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 as specified in [802.1q] Section 3.18.

3.3.3. STP Revision Level

This TLV carries the value of Revision Level.

```
    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=TBA5               |    Length                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |       Revision Level          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  Set to TBA5 for "STP Revision Level".

- Length

  Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields. Always set to 2.

- Revision Level

  The Revision Level as specified in [802.1q] Section 13.7 bullet 3);

3.3.4. STP Instance Priority

This TLV carries 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=TBA6               |    Length                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  Pri  |      InstanceID       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA6 for "STP Instance Priority"

- Length

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

- Pri

  The Instance Priority. It is interpreted as unsigned integer with higher value indicating a higher priority.

- InstanceID

  The 12 bits Instance Identifier of the CIST or MSTI. This parameter takes a value in the range 1 through 4094 for MSTI as defined in [802.1q] Section 12.8.1.2.2 and takes value of 0 for CIST.

3.3.5. STP Configuration Digest

This TLV carries the value of STP VLAN Instance Mapping.

```
 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=TBA7             |    Length                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Configuration Digest                         |
~                                                                 ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA7 for "STP Configuration Digest"

- Length

  Length of the STP Configuration Digest. Always set to 16 octets.

- Configuration Digest

  As specified in [802.1q] Section 13.7 bullet 4).

3.4. STP State TLVs

The STP State TLVs are sent in the RG Application Data message. They are used by a PE to report its STP status to other members in the RG. Such TLVs are specified in the following subsections.

3.4.1. STP Topology Changed Instances

This TLV is used to report the Topology Changed Instances to other members of the RG. The sender monitors TCN messages and generates this list. The receiving RG member SHOULD initiate the Topology Change event, including sending BPDU with the Topology Change flag set to 1 out of the designated port(s) of the Topology Changed bridge domains of the STP network, flushing out of MAC addresses relevant to the instances listed in this TLV.

If the PE supports MAC Address Withdrawal (see Section 6.2 of [RFC4762]), it SHOULD send an LDP Address Withdraw Message with the list of MAC addresses towards the core over the corresponding LDP sessions. It is not necessary to send such a message to PEs of the same RG since the flushing of their MAC address tables should have been performed upon the receipt of "STP Topology Changed Instances" 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=TBA8               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       InstanceID List                         |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA8 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 InstanceIDs of CIST or MSTIs whose topologies have changed as indicated by the Topology Change Notification (TCN) Messages as specified in [802.1q] Section 13.14. The list is formatted by padding Instance ID value to 16 bit boundary as follows, where the bits in the "R" fields MUST be sent as zero and ignored on receipt.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|R|R| InstanceID#1          |R|R|R|R| InstanceID#2          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                             ... ...                           ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

3.4.2. STP CIST Root Time Parameters

This TLV is used to report the Value of CIST Root Time Parameters ([802.1q] Section 13.23.7) to other members of the RG. All time parameter values are in seconds with a granularity of 1. For ranges and default values of these parameter values, refer to [802.1d1998] Section 8.10.2 Table 8-3, [802.1d2004] Section 17.14 Table 17-1, and [802.1q] Section 13.23.7.

```
 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=TBA9               |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    MaxAge                     |   MessageAge                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    FwdDelay                   |   HelloTime                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| RemainingHops |
+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA9 for "STP CIST Root Time"

- Length

  Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields. Always set to 9.

- MaxAge

  The Max Age of the CIST. It is the maximum age of the information transmitted by the bridge when it is the Root Bridge ([802.1d2004] Section 17.13.8).

- MessageAge

The Message Age of the CIST (see [802.1q] Section 13.23.7).

- FwdDelay

  The Forward Delay of the CIST. It is the delay used by STP Bridges to transition Root and Designated Ports to Forwarding ([802.1d2004] Section 17.13.5).

- HelloTime

  The Hello Time of the CIST. It is the interval between periodic transmissions of Configuration Messages by Designated Ports ([802.1d2004] Section 17.13.6).

- RemainingHops

  The remainingHops of the CIST ([802.1q] Section 13.23.7) .

3.4.3. STP MSTI Root Time Parameter

This TLV is used to report the parameter value of MSTI Root Time to other members of the RG. As defined in [802.1q] Section 13.23.7, it is the value of remainingHops for the given MSTI.

```
 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=TBA10              |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Pri  |  InstanceID           | RemainingHops |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA10 for "STP MSTI Root Time"

- Length

  Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields. Always set to 3.

- Pri

  The Instance Priority. It is interpreted as an unsigned integer with higher value indicating a higher priority.

- InstanceID

  The 12 bits Instance IDentifier of the Multiple Spanning Tree Instance (MSTID). As defined in [802.1q] Section 12.8.1.2.2, this parameter takes a value in the range 1 through 4094.

- RemainingHops

  The remainingHops of the MSTI. It is encoded in the same way as in [802.1q] Section 14.6.1 bullet f).

3.5. 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:

- configuration and/or state of the STP system,
- configuration and/or state for a given list of instances.

The format of the TLV is as follows:

```
 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=TBA11              |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Request Number           |C|S|   Request Type            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        InstanceID List                        |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA11 for "STP Synchronization Request TLV"

- Length

  Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields. Always set to 4.

- Request Number

  2 octets. Unsigned integer uniquely identifying the request.

Used to match the request with a corresponding response. The value of 0 is reserved for unsolicited synchronization, and MUST NOT be used in the STP Synchronization Request TLV. As indicated in [RFC7275], given the use of TCP, there are no issues associated with the wrap-around of the Request Number.

- C-bit

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

- S-bit

  Set to 1 if the 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 data of the listed instances
0x3FFF	Request System Data and data of all instances

- InstanceID List

  The InstanceIDs of CIST or MSTIs, format specified in Section 3.4.1.

3.6. STP Synchronization Data TLV

The pair of STP Synchronization Data TLVs are used by sender 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. It's REQUIRED that each pair of STP Synchronization Data TLVs occur in the same fragment. When the total size of the TLVs to be transmitted exceeds the maximal size of a fragment, these TLVs SHOULD be divided into multiple sets, delimited by multiple pairs of STP Synchronization Data TLVs, and filled into multiple fragments.

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.

STP Synchronization Data TLV has the following format:

```
    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=TBA12              |    Length                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |     Request Number            |    Reserved                 |S|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- U=F=0

- Type

  set to TBA12 for "STP Synchronization Data TLV"

- Length

  Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields. Always set to 4.

- Request Number

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

- Reserved

  Reserved bits for future use. These MUST be sent as zero and ignored on receipt.

-S

  S = 0: Synchronization Data Start
  S = 1: Synchronization Data End

4. Operations

Operational procedures for AC redundancy applications have been specified in Section 9.2 of [RFC7275]. The operational procedures of ICCP STP application should follow these procedures except the changes presented in this section.

4.1. Common AC Procedures

For the generic procedures of AC redundancy applications defined in Section 9.2.1 of [RFC7275], the following changes are introduced.

4.1.1. Remote PE Node Failure or Isolation

When a local PE node detects that a remote PE that is a member of the same RG is no longer reachable (using the mechanisms described in Section 5 of [RFC7275]), the local PE checks if it has redundancy ACs for the affected services. In case of redundant ACs present, and if the local PE has the new highest bridge priority, the local PE node becomes the virtual root bridge for corresponding ACs.

4.1.2. Local PE Isolation

When a PE node detects that it has been isolated from the core network, then it should ensure that its AC redundancy mechanism will change the status of all active ACs to standby. The AC redundancy application SHOULD then send ICCP "Application Data" message in order to trigger failover to another active PE in the RG. Note that this works only in the case of dedicated interconnect (Sections 3.2.1 and 3.2.3), since ICCP will still have the path to the peer, even though the PE is isolated from the MPLS core network.

4.2. ICCP STP Application Procedures

This section defines the procedures that are specific to the ICCP STP application which are applicable for Ethernet ACs.

4.2.1. Initial Setup

When a RG is configured on a system that supports the ICCP STP application, such system MUST send an "RG Connect" message with "STP Connect TLV" to each PE that is member of the RG. The sending PE MUST set the A bit to 1 in that TLV if it has received a corresponding "STP Connect TLV" from its peer PE; otherwise, the sending PE MUST set the A bit to 0. If a PE receives an "STP Connect TLV" from its peer after sending its own TLV with the A bit set to 0, it MUST resend the TLV with the A bit set to 1. A system considers the ICCP STP application connection to be operational when it has both sent and received "STP Connect TLVs" with the A bit set to 1. When the ICCP STP application connection between a pair of PEs is operational, the two devices can start exchanging "RG Application Data" messages for the ICCP STP application. This involves having each PE advertise its STP configuration and operational state in an unsolicited manner. A PE SHOULD follow the following order when advertising its STP state upon initial application connection setup:

- Advertise system configuration TLV
- Advertise remaining configuration TLVs
- Advertise state TLVs

A PE MUST use a pair of "STP Synchronization Data TLVs" to delimit the entire set of TLVs that are being sent as part of this

unsolicited advertisement.

If a system receives an "RG Connect" message with "STP Connect TLV" that has a differing Protocol Version, it MUST follow the procedures outlined in the "Application Versioning" Section of [RFC7275].

After the ICCP STP application connection has been established, every PE MUST communicate its system level configuration to its peers via the use of "STP System Config TLV".

When the ICCP STP application is administratively disabled on the PE, or on the particular RG, the system MUST send an "RG Disconnect" message containing "STP Disconnect TLV".

4.2.2. Configuration Synchronization

A system that support ICCP STP application MUST synchronize the configuration with other RG members. This is achieved via the use of "STP Config TLVs". The PEs in the RG MUST all agree on the common MAC address to be associated with the virtual root bridge. It is possible to achieve this via consistent configuration on member PEs. However, in order to protect against possible misconfigurations, a virtual root bridge identifier MUST be set to the MAC address advertised by the PE with the numerically lowest BridgeIdentifier (i.e., the MAC address of the bridge) in the RG.

Furthermore, for a given ICCP STP application, an implementation MUST advertise the configuration prior to advertising its corresponding state. If a PE receives any STP State TLV that it had not learned of before via an appropriate STP Config TLV, then the PE MUST request synchronization of the configuration and state from its peer. If during such synchronization a PE receives a State TLV that it has not learned before, then the PE MUST send a NAK TLV for that particular TLV. The PE MUST NOT request resynchronization in this case.

4.2.3. State Synchronization

PEs within the RG need to synchronize their state for proper STP operation. This is achieved by having each system advertise its running state in STP State TLVs. Whenever any STP parameter either on CE or PE side is changed, the system MUST transmit an updated TLV for the affected STP instances. Moreover, when the administrative or operational state changes, the system MUST transmit an updated state TLV to its peers.

A PE MAY request its peer to retransmit previously advertised state. This is useful in case of the PE recovering from a soft failure and attempting to relearn state. To request such retransmissions, a PE

MUST send a set of one or more "STP Synchronization Request TLVs".

A PE MUST respond to a "STP Synchronization Request TLV" by sending the requested data in a set of one or more STP configuration or state TLVs delimited by a pair of "STP Synchronization Data TLVs".

Note that the response may span across multiple RG Application Data messages, for example when MTU limits are exceeded; however, the above ordering MUST be retained across messages, and only a single pair of Synchronization Data TLVs MUST be used to delimit the response across all Application Data Messages.

A PE device MAY readvertise its STP state in an unsolicited manner. This is done by sending the appropriate State TLVs delimited by a pair of "STP Synchronization Data TLVs" and using a 'Request Number' of 0.

While a PE has sent out a synchronization request for a particular PE node, it SHOULD silently ignore all TLVs from that node, that are received prior to the synchronization response and which carry the same type of information being requested. This saves the system from the burden of updating state that will ultimately be overwritten by the synchronization response. Note that TLVs pertaining to other systems should continue to be processed normally.

If a PE receives a synchronization request for an instance that doesn't exist or is not known to the PE, then it MUST trigger the unsolicited synchronization of all information by restarting the initialization.

If during the synchronization operation a PE receives an advertisement of a Node ID value which is different from the value previously advertised, then the PE MUST purge all state data previously received from that peer prior to the last synchronization.

4.2.4. Failure and Recovery

When a PE that is active for the ICCP STP application encounters a core isolation fault [RFC7275], it SHOULD attempt to fail-over to a peer PE which hosts the same RG. The default fail-over procedure is to have the failed PE bring down the link(s) towards the multi-homed STP network. This will cause the STP network to reconverge and to use the other links that are connected to the other PE(s) in the RG. Other procedures for triggering fail-over are possible, and are outside the scope of this document.

If the isolated PE is the one that has the numerically lowest BridgeIdentifier, PEs in the RG MUST synchronize STP configuration

and state TLVs and determine a new virtual root bridge as specified in Section 4.2.2.

Upon recovery from a previous fault, a PE SHOULD NOT reclaim the role of the virtual root for the STP network even if it has the numerically lowest BridgeIdentifier among the RG. This minimizes traffic disruption.

Whenever the virtual root bridge changes, the STP Topology Changed Instances TLV lists the instances that are affected by the change. These instances MUST undergo a STP reconvergence procedure when this TLV is received as defined in Section 3.4.1.

5. Security Considerations

This document specifies an application running on the channel provided by ICCP [RFC7275]. The security considerations on ICCP apply in this document as well.

For the ICCP STP application, an attack on channel (running in the provider's network) can break not only the ability to deliver traffic across the provider's network, but also the ability to route traffic within the customer's network. That is, careful attack on channel (such as the DOS attacks as described in [RFC7275]) can break STP within the customer network. Implementations SHOULD provide mechanisms to mitigate these types of attacks. For example, the port between the PE and the malicious CE may be blocked.

6. IANA Considerations

The IANA maintains a top-level registry called "Pseudowire Name Spaces (PWE3)". It has a sub-registry called "ICC RG Parameter Types".

IANA is requested to make 13 allocations from this registry as shown below. IANA is requested to allocate the codepoints in sequential block starting from the next available value in the range marked for assignment by IETF review (0x2000-0x2FFF) [RFC5226]. All assignments should reference this document.

Parameter Type	Description
TBA1	STP Connect TLV
TBA2	STP Disconnect TLV
TBA3	STP System Config TLV
TBA4	STP Region Name TLV
TBA5	STP Revision Level TLV
TBA6	STP Instance Priority TLV

TBA7	STP Configuration Digest TLV
TBA8	STP Topology Changed Instances TLV
TBA9	STP STP CIST Root Time TLV
TBA10	STP MSTI Root Time TLV
TBA11	STP Synchronization Request TLV
TBA12	STP Synchronization Data TLV
TBA13	STP Disconnect Cause TLV

Acknowledgements

Authors would like to thank the comments and suggestions from Ignas Bagdonas, Adrian Farrel, Andrew G. Malis, Gregory Mirsky and Alexander Vainshtein.

7. References

7.1. Normative References

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

[RFC4762] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.

[RFC7275] Martini, L., Salam, S., Sajassi, A., Bocci, M., Matsushima, S., and T. Nadeau, "Inter-Chassis Communication Protocol for Layer 2 Virtual Private Network (L2VPN) Provider Edge (PE) Redundancy", RFC 7275, June 2014.

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

[802.1d1998] "Information technology---Telecommunications and information exchange between systems---Local and metropolitan area networks---Common specifications--Part 3: Media Access Control (MAC) Bridges". ANSI/IEEE Std 802.1D, 1998 Edition.

[802.1d2004] "IEEE Standard for Local and metropolitan area networks--- Media Access Control (MAC) Bridges". IEEE Std 802.1 D-2004.

7.2. Informative References

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

Author's Addresses

Mingui Zhang
Huawei Technologies
No. 156 Beiqing Rd. Haidian District,
Beijing 100095
P.R. China

EMail: zhangmingui@huawei.com

Huafeng Wen
Huawei Technologies
101 Software Avenue,
Nanjing 210012
P.R. China

EMail: wenhuafeng@huawei.com

Jie Hu
China Telecom

EMail: hujie@ctbri.com.cn