

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: April 25, 2015

Mingui Zhang
Huafeng Wen
Huawei
Jie Hu
China Telecom
October 22, 2014

STP Application of ICCP
draft-ietf-pwe3-iccp-stp-02.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. The operation logic of the application and the usage of these ICCP TLVs are specified.

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) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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	4
1.1. Conventions used in this document	4
1.2. Terminology	4
2. The Use Case Scenario	4
2.1. STP as an Application of ICCP	5
2.2. The BPDUs Tunneling Model	6
3. Spanning Tree Protocol Application TLVs	7
3.1. STP Connect TLV	7
3.2. STP Disconnect TLV	8
3.2.1. STP Disconnect Cause TLV	9
3.3. STP Config TLVs	9
3.3.1. STP System Config	10
3.3.2. STP Region Name	10
3.3.3. STP Revision Level	11
3.3.4. STP Instance Priority	12
3.3.5. STP Configuration Digest	12
3.4. STP State TLVs	13
3.4.1. STP Topology Changed Instances	13
3.4.2. STP CIST Root Time	14
3.4.3. STP MSTI Root Time	15
3.5. STP Synchronization Request TLV	16
3.6. STP Synchronization Data TLV	17
4. Operations	18
4.1. Common AC Procedures	18
4.1.1. Remote PE Node Failure or Isolation	18
4.1.2. Local PE Isolation	18
4.2. STP Application Procedures	18
4.2.1. Initial Setup	19
4.2.2. Configuration Synchronization	19
4.2.3. Status Synchronization	20
4.2.4. Failure and Recovery	21
5. Security Considerations	22

6.	IANA Considerations	22
	Acknowledgements	22
7.	References	22
7.1.	Normative References	22
7.2.	Informative References	23
	Author's Addresses	24

1. Introduction

Inter-Chassis Communication Protocol (ICCP [[RFC7257](#)]) 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 an RG, the RG members 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 STP application are specified for this purpose.

From the point of view of the customers, ISPs are still providing Virtual Private LAN Service (VPLS). Familiarity with [[RFC4762](#)] is assumed in this document.

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
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. ISPs provides these customers with VPLS. It's common that one customer broadband network has multiple DSLAMs connected to a carrier's network at different locations 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 to the carrier's network, 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. In [Section 2.2](#), a counter-part solution based on the BPDU tunneling technique is addressed as a contrast. Its issues are analyzed.

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

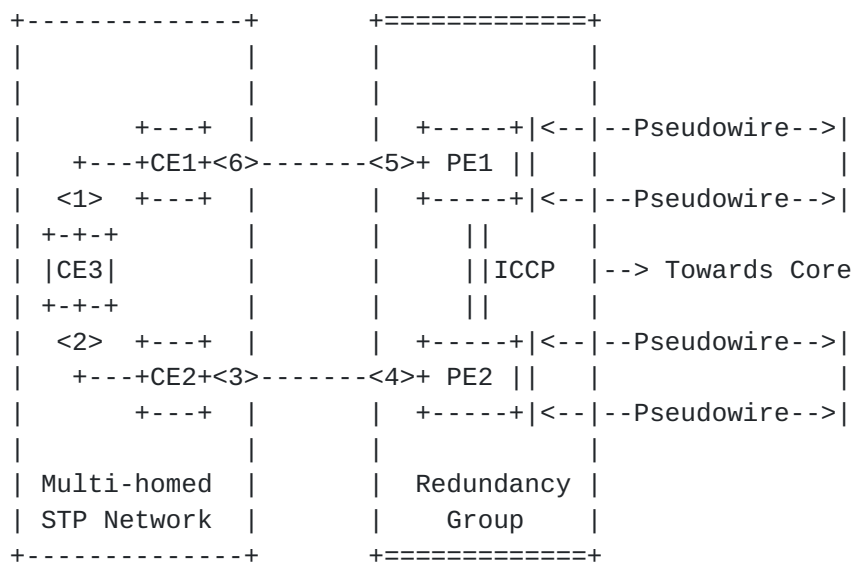


Figure 2.1: A STP network is multi-homed to an RG running ICCP.

Figure 2.1 shows an example topology of this model. With ICCP, the whole RG will be virtualized to be a single bridge. Each RG member has its BridgeIdentifier (the MAC address). The least significant one is used as the BridgeIdentifier of the 'virtualized root bridge'. The RG pretends that the ports connected to the STP network (port <4>, <5>) are from the same root bridge. All these ports emit 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 STP network. In this way, the STP will always break a loop within the multi-homed STP network. As if the whole network is broken into separate islands that each is attached to one PE. It forces all PEs in the RG to be active. This is different from a generic VPLS per [RFC4762] where the root bridge resides in the customer network and the multi-homed PEs usually act in the active-standby mode. Note that the specification of VPLS should remain unchanged other than this operation. For instance, a full-mesh of PWs is established between PEs, split-horizon is still used to perform the loop-free protection through the core, etc.

2.2. The BPDU Tunneling Model

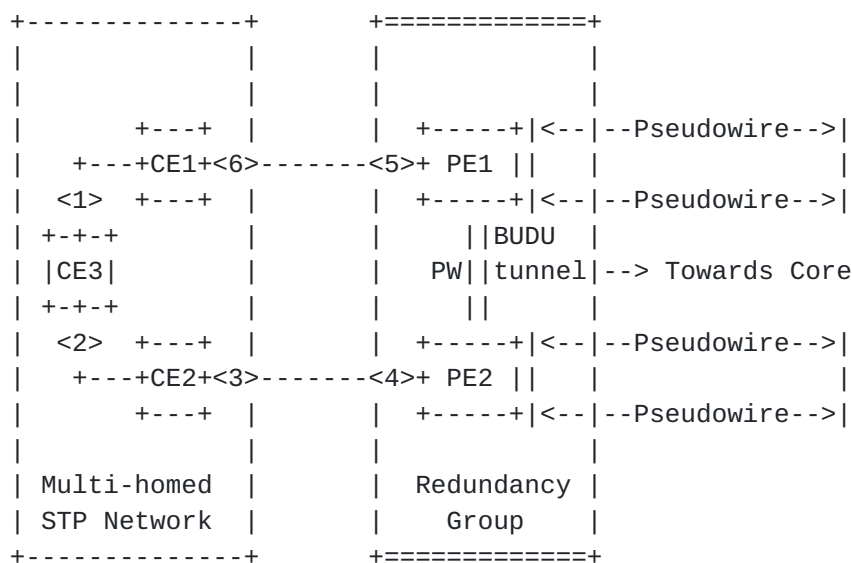


Figure 2.2: The BPDU Tunneling Model

Figure 2.2 shows an example network of the 'BPDU tunneling' model. Two PEs tunnel customer traffic and BPDUs of the STP network over a PW. The OAM designed in [RFC7023] 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. Consider the remote traffic from PEX to PE2 then to

CE3. PE2 continues to send traffic to CE3 via CE2 where a black-hole happens.

In order to handle this issue, PEs need 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.

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 specifies the ICCP TLVs for the STP application.

3.1. STP Connect TLV

This TLV is included in the RG Connect message to signal the establishment of 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=0x0040                               |   Length   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Protocol Version                       |A|   Reserved   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Optional Sub-TLVs                                     |
~                                                                 ~
|                                                                 |
+                                     +---+---+---+---+---+---+---+---+
|                                     ...                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```


- U and F Bits

Both are set to 0.

- Type

set to 0x0040 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.

```

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=0x0041           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Optional Sub-TLVs    |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

```
set to 0x0041 for "STP Disconnect TLV"
```

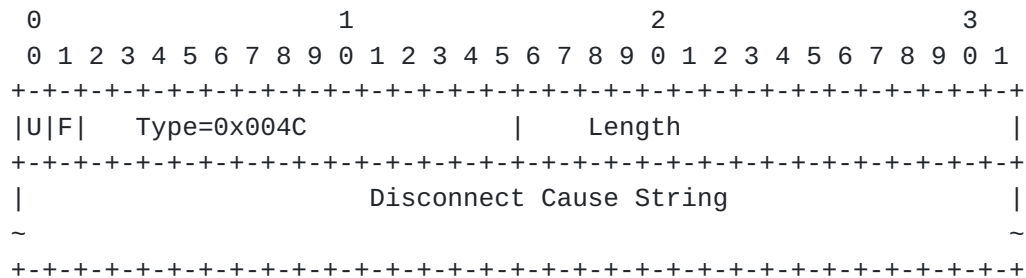
- Length

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



- U and F Bits

Both are set to 0.

- Type

set to 0x004C 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.

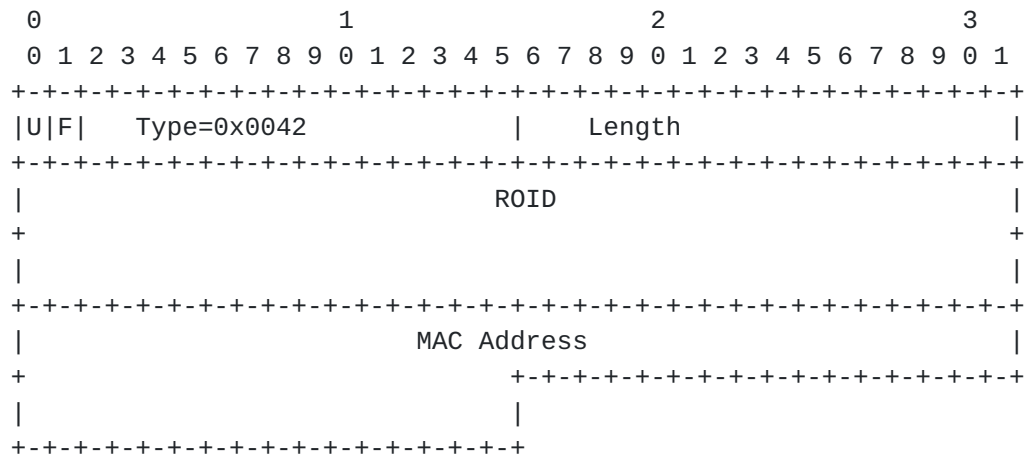
3.3. STP Config TLVs

The STP Config TLVs are sent in the RG Application Data message. When a STP Config TLV is received by a peering RG member, it MUST synchronize the configuration information contained in the TLV. TLVs

specified from [Section 3.3.1](#) through [Section 3.3.5](#) 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.



- U and F Bits

Both are set to 0.

- Type

set to 0x0042 for "STP System Config"

- Length

Length of the MAC address, which is 6 octets.

- ROID

As defined in [Section 6.1.3 of \[RFC7257\]](#).

- 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 the least significant unsigned 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 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=0x0046               |   Length               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Region Name                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0046 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.

3.3.3. 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=0x0047               |   Length               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Revision Level              |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0047 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](#);

3.3.4. 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=0x0048           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Pri |   InstanceID         |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0048 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

- InstanceID

The instance identification number of the MSTI.

3.3.5. STP Configuration Digest

This TLV is used to report the Value of STP VLAN Instance Mapping 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=0x0049                |  Length                |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|                                Configuration Digest        |
~                                                                    ~
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0049 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. 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 in the RG. The receiver RG member SHOULD enforce the Topology Change event, including sending a BPDU with the Topology Change flag set to 1 out of the designated port(s) connected to the STP network, the flushing out of MAC addresses relevant to the instances listed in this TLV, etc.

If the PE supports the MAC Address Withdrawal (see [Section 6.2 of \[RFC4762\]](#)), it SHOULD send an LDP Address Withdraw Message with the list of MAC addresses to be removed to all remote peering PEs over the corresponding LDP sessions. It's unnecessary to send such a message to PEs in the same RG since the flushing of their MAC address tables should have been performed upon the STP Topology Changed Instances TLV is received.


```

      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=0x0043           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     InstanceID List                                     |
~                                                                 ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0043 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](#).

3.4.2. 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=0x0044           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   MaxAge   |   MessageAge   |   FwdDelay   |   HelloTime   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| RemainingHops |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0044 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

3.4.3. STP MSTI Root Time

This TLV is used to report the Value of MSTI 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=0x0045          |  Length          |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|          InstanceID          | RemainingHops |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

- U and F Bits

Both are set to 0.

- Type

set to 0x0045 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

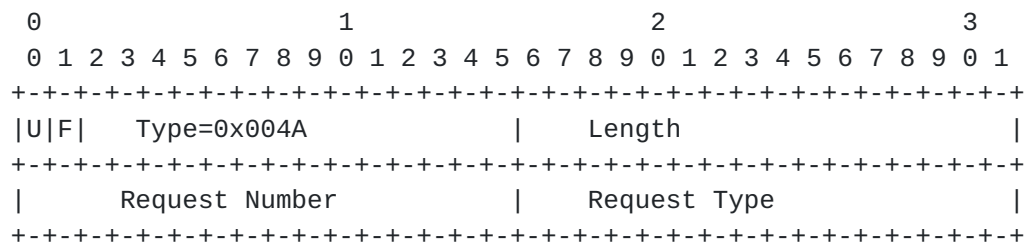
The remaining hops of this TLV

3.5. 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:

- 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 0x004A 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.

- Request Type

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

```
0x00    Request Configuration Data
0x01    Request State Data
0x3FFF  Request All Data
```

3.6. 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:

```

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=0x004B           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   Request Number           |   Flags           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

- U and F Bits

Both are set to 0.

- Type

set to 0x004B 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. Operations

Operation procedures for AC redundancy applications have been specified in [Section 9.2 of \[RFC7257\]](#). As an AC redundancy application, the operation procedures of the STP application should follow most of 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 \[RFC7257\]](#), the following changes are introduced.

4.1.1. Remote PE Node Failure or Isolation

When a 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 \[RFC7257\]](#)), the local PE determines if it has redundant ACs for the affected services. If the local PE has the highest priority (after the failed PE), then it becomes the virtual root bridge.

4.1.2. Local PE Isolation

When a PE node detects that it has been isolated from the core network (i.e., all core-facing interfaces/links are not operational), then it should ensure that its AC redundancy mechanism will change the status of any active ACs to standby. The AC redundancy application SHOULD then send ICCP "Application Data" messages 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 a path to the peer, even though the PE is isolated from the MPLS core network.

4.2. STP Application Procedures

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

4.2.1. Initial Setup

When an RG is configured on a system to support the STP application, the system MUST send an "RG Connect" message with "STP Connect TLV" to each PE that is member of the same RG. The sending PE MUST set the A bit to 1 in the said 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 the said TLV with the A bit set to 0, it MUST resend the TLV with the A bit set to 1. A system considers the STP application connection to be operational when it has sent and received "STP Connect TLVs" with the A bit set to 1. When the STP application connection between a pair of PEs is operational, the two devices can start exchanging "RG Application Data" messages for the STP application. This involves having each PE advertise its STP configuration and operational state in an unsolicited manner. A PE SHOULD subscribe to 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 [[RFC7257](#)].

After the 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 STP application is disabled on the device, or is unconfigured for the RG in question, the system MUST send an "RG Disconnect" message with "STP Disconnect TLV".

4.2.2. Configuration Synchronization

A system that support the STP application MUST synchronize the configuration with other RG members. This is achieved via the use of "STP Config TLVs". The PEs in an RG MUST all agree on the 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 misconfiguration, a system **MUST** use, for the virtual root bridge, the MAC address supplied by the PE with the numerically least significant BridgeIdentifier (i.e., the MAC address of the bridge) in the RG.

Furthermore, for a given STP application, an implementation **MUST** advertise the configuration prior to advertising its 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 respective peer. If during a synchronization, a PE receives a State TLV that it has not learned of before, then the PE **MUST** send a NAK TLV for the offending TLV. The PE **MUST NOT** request re-synchronization in this case.

4.2.3. Status Synchronization

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

A PE **MAY** request that its peer retransmit previously advertised state. This is useful for example when the PE is 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". The TLVs comprising the response **MUST** be ordered in the RG Application Data message(s) such that the Synchronization Response TLV with the "Synchronization Data Start" flag precedes the various other STP configuration or state TLVs encoding the requested data. These, in turn, **MUST** precede the Synchronization Data TLV with the "Synchronization Data End" flag.

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** re-advertise its STP state in an unsolicited manner. This is done by sending the appropriate Config and State TLVs

delimited by a pair of "STP Synchronization Data TLVs" and using a 'Request Number' of 0.

While a PE has a pending synchronization request for a system, it SHOULD silently ignore all TLVs for said system, 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, are to continue to be processed per normal in this case.

If a PE receives a synchronization request that doesn't exist or is not known to the PE, then it MUST trigger an unsolicited synchronization of all system information (i.e. replay the initialization sequence).

If a PE learns, as part of a synchronization operation from its peer, that the latter is advertising a Node ID value which is different from the value previously advertised, then the PE MUST purge all state data previously learnt from that peer prior to the last synchronization.

4.2.4. Failure and Recovery

When a PE that is active for a STP application encounters a core isolation fault [[RFC7257](#)], it SHOULD attempt to fail-over to a peer PE which hosts the same R0. The default fail-over procedure is to have the failed PE bring down the link(s) towards the multi-homed STP network (e.g. by bringing down the line-protocol). This will cause the STP network reconverges and uses 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 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](#) (i.e., the PE with the least significant BridgeIdentifier will become the virtual root bridge.).

Upon recovery from a previous fault, a PE MAY reclaim the role of the virtual root for the STP network if it has the lowest BridgeIdentifier among the RG. PEs in the RG MUST synchronize the STP configuration and state TLVs.

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. See security considerations of [\[RFC7257\]](#).

6. IANA Considerations

The ICC RG parameter types used by the application TLVs defined in [Section 3](#) of this document should be assigned. ICC RG parameter types are 14-bit values. Parameter Type values 1 through 0x003A are specified in [\[RFC7257\]](#). This document adds the specification of values 0x003B through 0x004C. Parameter Type values 0x004D through 0x1FFF are to be assigned by IANA, using the "Expert Review" policy defined in [\[RFC5226\]](#). Suggested parameter type space value allocations for the STP application are specified below:

Parameter Type	Description
-----	-----
0x003B-0x003F	Reserved
0x0040	STP Connect TLV
0x0041	STP Disconnect TLV
0x0042	STP System Config TLV
0x0043	STP Region Name TLV
0x0044	STP Revision Level TLV
0x0045	STP Instance Priority TLV
0x0046	STP Configuration Digest TLV
0x0047	STP Topology Changed Instances TLV
0x0048	STP STP CIST Root Time TLV
0x0049	STP MSTI Root Time TLV
0x004A	STP Synchronization Request TLV
0x004B	STP Synchronization Data TLV
0x004C	STP Disconnect Cause TLV

Acknowledgements

Authors would like to thank the comments and suggestions from 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.

[RFC7257] L. Martini, S. Salam, et al, "Inter-Chassis Communication Protocol for L2VPN PE Redundancy", [draft-ietf-pwe3-iccp-16.txt](#), in RFC Ed Queue.

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

7.2. Informative References

[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.

[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.

Author's Addresses

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

E-Mail: zhangmingui@huawei.com

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

E-Mail: wenhuafeng@huawei.com

Jie Hu
China Telecom

E-Mail: hujie@ctbri.com.cn

