

Internet Engineering Task Force (IETF)
Request for Comments: 7727
Category: Standards Track
ISSN: 2070-1721

M. Zhang
H. Wen
Huawei
J. Hu
China Telecom
January 2016

Spanning Tree Protocol (STP) Application
of the Inter-Chassis Communication Protocol (ICCP)

Abstract

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

In this document, Provider Edge (PE) devices in a Redundancy Group (RG) running ICCP are used to offer multihomed 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.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc7727>.

Copyright Notice

Copyright (c) 2016 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	4
1.1. Conventions Used in This Document	4
1.2. Terminology	4
2. Use Case	5
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 Sub-TLV	8
3.3. STP Configuration TLVs	8
3.3.1. STP System Config	9
3.3.2. STP Region Name	10
3.3.3. STP Revision Level	10
3.3.4. STP Instance Priority	11
3.3.5. STP Configuration Digest	12
3.4. STP State TLVs	12
3.4.1. STP Topology Changed Instances	12
3.4.2. STP CIST Root Time Parameters	14
3.4.3. STP MSTI Root Time Parameter	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	19
4.1.2. Local PE Isolation	19
4.2. ICCP STP Application Procedures	19
4.2.1. Initial Setup	19
4.2.2. Configuration Synchronization	20
4.2.3. State Synchronization	21
4.2.4. Failure and Recovery	22
5. Security Considerations	22
6. IANA Considerations	23
7. References	23
7.1. Normative References	23
7.2. Informative References	24
Acknowledgements	24
Authors' Addresses	25

1. Introduction

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

With the Spanning Tree Protocol (STP), a spanning tree will be formed over connected bridges by blocking some links between these bridges so that forwarding loops are avoided. This document introduces support of STP as a new application of ICCP. When a bridged STP network is connected to an RG, this STP application of ICCP enables the RG members to act as a single root bridge participating in the operations of STP.

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 Provider is 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 3.27)
MSTI: Multiple Spanning Tree Instance ([802.1q], Section 3.138)
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 to be active.
- o When one connection to the carrier network fails, customers require a connection in another location to continue to work after the reconvergence 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 attachment circuit (AC), or even the Customer Edge (CE) device itself.

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

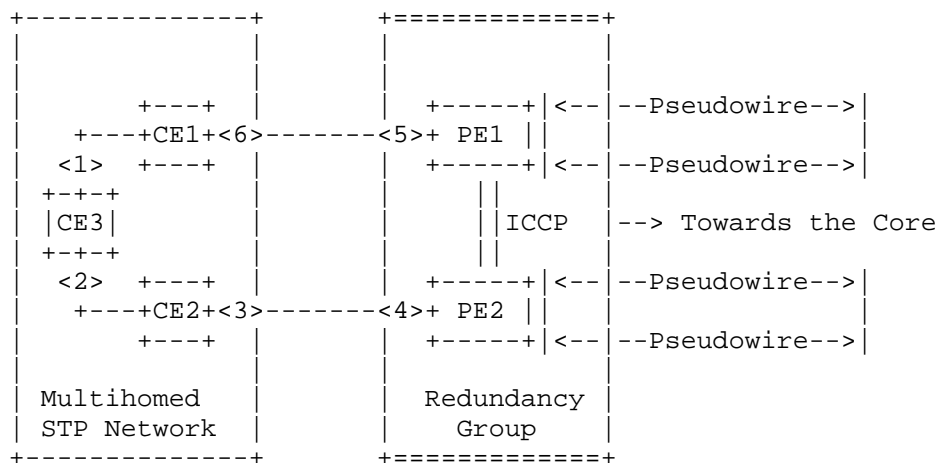


Figure 1: A STP network is multihomed to an RG running ICCP

Figure 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 (ports <4> and <5>) are for the same root bridge. All these ports send the configuration BPDU with the highest root priority to trigger the

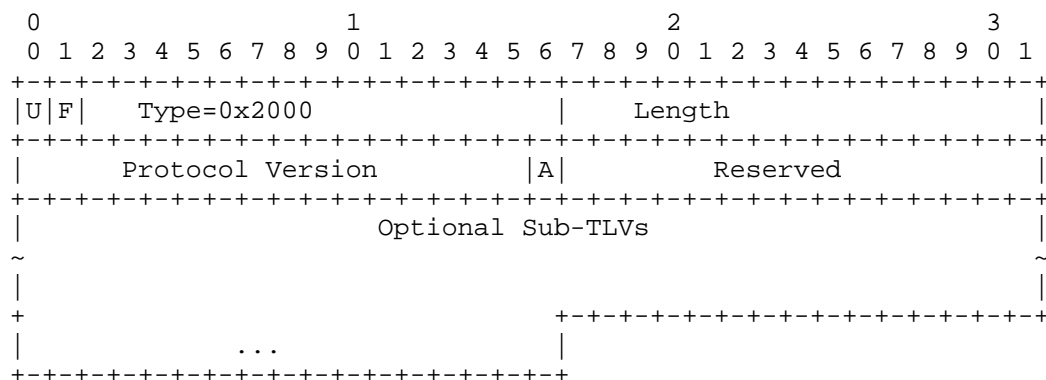
construction of the spanning tree. The link between the peering PEs is not visible to the bridge domains of the STP network. In this way, the STP will always break a possible loop within the multihomed STP network by breaking the whole network into separate islands so that each is attached to one PE. 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 multihomed 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 pseudowires (PWs) is established between PEs, and the "split horizon" rule 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 an ICCP STP application connection.



- U=F=0

- Type

Set to 0x2000 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 ICCP STP 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

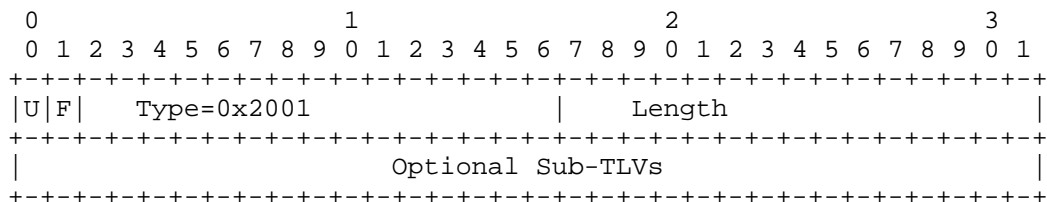
Reserved for future use. These bits MUST be sent as 0 and ignored on receipt.

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



- U=F=0

- Type

Set to 0x2001 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=0x200C   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+
|                                     Disconnect Cause String
|                                     |
~                                     ~
+-----+-----+-----+-----+-----+-----+-----+

```

- U=F=0

- Type

Set to 0x200C for "STP Disconnect Cause TLV"

- Length

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

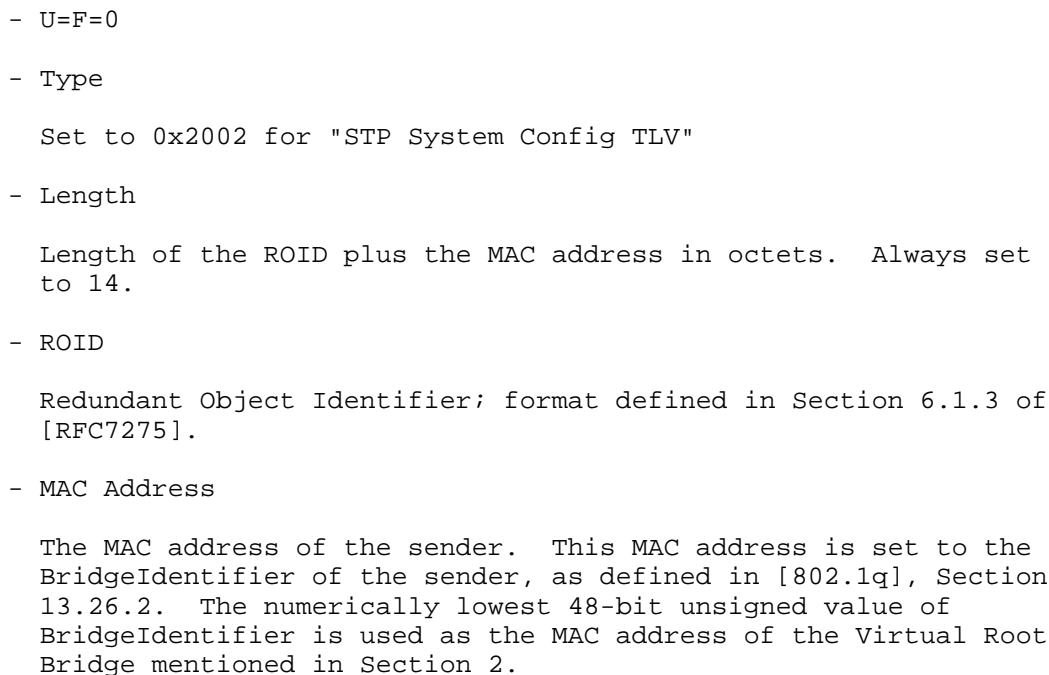
- Disconnect Cause String

Variable-length string specifying the reason for the disconnect, encoded in UTF-8 [RFC3629] format. Used for operational purposes.

3.3. STP Configuration TLVs

The STP Configuration TLVs are sent in the RG Application Data Message. When an STP Config TLV is received by a peer RG member, the member MUST synchronize with the configuration information contained in the TLV. TLVs specified in Sections 3.3.1 to 3.3.5 define specific configuration information.

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



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=0x2003                               |   Length   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Region Name    |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U=F=0

- Type

Set to 0x2003 for "STP Region Name TLV"

- 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 3.142.

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=0x2004                               |   Length   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Revision Level  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U=F=0

- Type

Set to 0x2004 for "STP Revision Level TLV".

- 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 13.8, item c.

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=0x2005   |   Length   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Pri |   InstanceID   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U=F=0

- Type

Set to 0x2005 for "STP Instance Priority TLV"

- 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-bit 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=0x2006   |      Length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Configuration Digest                                     |
~                                                                 ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U=F=0

- Type

Set to 0x2006 for "STP Configuration Digest TLV"

- Length

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

- Configuration Digest

As specified in [802.1q], Section 13.8, item d.

3.4. STP State TLVs

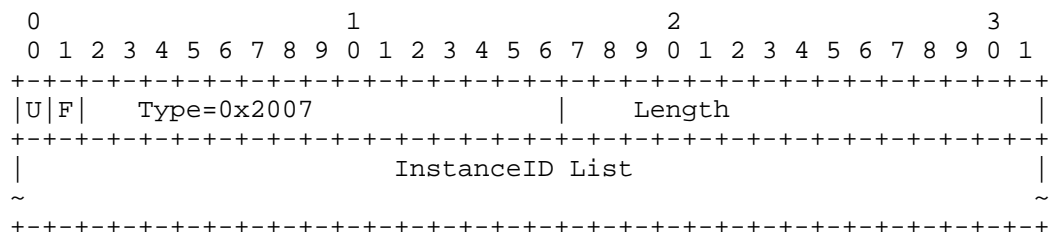
The STP State TLVs are sent in the RG Application Data Message. They are used by a PE device to report its STP status to other members in the RG. Such TLVs are specified in the following subsections.

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 Topology Change Notification (TCN) messages and generates this list. The receiving RG member MUST 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, and flushing out MAC addresses relevant to the instances listed in this TLV.

If the PE device supports MAC Address Withdrawal (see Section 6.2 of [RFC4762]), it SHOULD send a Label Distribution Protocol (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 receipt of the STP Topology Changed Instances TLV.



- U=F=0

- Type

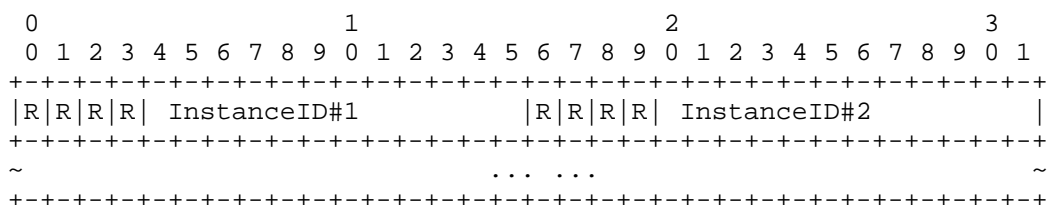
Set to 0x2007 for "STP Topology Changed Instances TLV"

- 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 the CIST or MSTIs whose topologies have changed as indicated by the TCN messages as specified in [802.1q], Section 13.14. The list is formatted by padding each InstanceID value to the 16-bit boundary as follows, where the bits in the "R" fields MUST be sent as 0 and ignored on receipt.



3.4.2. STP CIST Root Time Parameters

This TLV is used to report the Value of CIST Root Time Parameters ([802.1q], Section 13.26.7) 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 13.26.7.

```

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=0x2008                               |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   MaxAge                               |   MessageAge   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   FwdDelay                             |   HelloTime    |
+-----+-----+-----+-----+-----+-----+-----+-----+
| RemainingHops |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

- U=F=0

- Type

Set to 0x2008 for "STP CIST Root Time TLV"

- 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 13.26.7).

- 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 13.26.7).

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 13.26.7, 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=0x2009           |   Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Pri |   InstanceID           | RemainingHops |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- U=F=0

- Type

Set to 0x2009 for "STP MSTI Root Time TLV"

- 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-bit 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 14.4.1, item 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 that its peer retransmit 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=0x200A   |           Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Request Number   |C|S|   Request Type   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           InstanceID List           |
~                                     ~
+-----+-----+-----+-----+-----+-----+-----+-----+

```

- U=F=0

- Type

Set to 0x200A 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 it 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 the 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 the 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 they 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 MUST be divided into multiple sets, delimited by multiple pairs of STP Synchronization Data TLVs, and filled into multiple fragments. With the Request Number, lost fragments can be identified and re-requested.

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=0x200B   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Request Number   |   Reserved   |S|
+-----+-----+-----+-----+-----+-----+-----+

```

- U=F=0
- Type
 - set to 0x200B 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" that initiated this synchronization data response.
- Reserved
 - Reserved bits for future use. These MUST be sent as 0 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 the ICCP STP application should follow those procedures, with the changes presented in this section.

4.1. Common AC Procedures

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

4.1.1. Remote PE Node Failure or Isolation

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

4.1.2. Local PE Isolation

When a PE device detects that it has been isolated from the core network, then it needs to ensure that its AC redundancy mechanism will change the status of all active ACs to standby. The AC redundancy application SHOULD then send an RG Application Data Message in order to trigger failover to another active PE device 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 device is isolated from the MPLS core network.

4.2. ICCP STP Application Procedures

This section defines the procedures of the ICCP STP application that are applicable for Ethernet ACs.

4.2.1. Initial Setup

When an RG is configured on a system that supports the ICCP STP application, such systems MUST send an RG Connect Message with an STP Connect TLV to each PE device that is a member of the RG. The sending PE device 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 device MUST set the A bit to 0. If a PE device 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 device advertise its STP configuration and operational state in an unsolicited manner. A PE device SHOULD follow the order below when advertising its STP state upon initial application connection setup:

- Advertise the STP System Config TLV
- Advertise remaining Configuration TLVs
- Advertise State TLVs

The update of the information contained in the State TLVs depends on that in the Configuration TLVs. By sending the TLVs in the above order, the two peers may begin to update STP state as early as possible in the middle of exchanging these TLVs.

A PE device 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 an STP Connect TLV that has a differing Protocol Version, it MUST follow the procedures outlined in the Section 4.4.1 ("Application Versioning") of [RFC7275].

After the ICCP STP application connection has been established, every PE device 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 supports ICCP STP application MUST synchronize the configuration with other RG members. This is achieved via the use of STP Configuration 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 device 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 device receives any STP State TLV that it had not learned of before via an appropriate STP Configuration TLV, then the PE device MUST request synchronization of the configuration and state from its peer. If during such synchronization a PE device receives a State TLV that it has not learned before, then the PE device MUST send a NAK TLV for that particular TLV. The PE device 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 the 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 device MAY request its peer to retransmit previously advertised state. This is useful in case the PE device is recovering from a soft failure and attempting to relearn state. To request such retransmissions, a PE device MUST send a set of one or more STP Synchronization Request TLVs.

A PE device 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 RG 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 device has sent out a synchronization request for a particular PE device, it SHOULD silently ignore all TLVs that are from that node, are received prior to the synchronization response, and 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 device 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 device receives an advertisement of a Node ID value that is different from the value previously advertised, then the PE device MUST purge all state data previously received from that peer prior to the last synchronization.

4.2.4. Failure and Recovery

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

If the isolated PE device 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 device 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 a 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, a careful attack on a channel (such as the DoS attacks as described in [RFC7275]) can break STP within the customer network. Implementations need to provide mechanisms to mitigate these types of attacks. For example, the port between the PE device and the malicious CE device may be blocked.

6. IANA Considerations

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

IANA has made 13 allocations from this registry as shown below. IANA has allocated the codepoints from the range marked for assignment by IETF Review (0x2000-0x2FFF) [RFC5226]. Each assignment references this document.

Parameter	Type	Description
0x2000		STP Connect TLV
0x2001		STP Disconnect TLV
0x2002		STP System Config TLV
0x2003		STP Region Name TLV
0x2004		STP Revision Level TLV
0x2005		STP Instance Priority TLV
0x2006		STP Configuration Digest TLV
0x2007		STP Topology Changed Instances TLV
0x2008		STP CIST Root Time TLV
0x2009		STP MSTI Root Time TLV
0x200A		STP Synchronization Request TLV
0x200B		STP Synchronization Data TLV
0x200C		STP Disconnect Cause TLV

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, DOI 10.17487/RFC3629, November 2003, <<http://www.rfc-editor.org/info/rfc3629>>.
- [RFC4762] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <<http://www.rfc-editor.org/info/rfc4762>>.

- [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, DOI 10.17487/RFC7275, June 2014, <<http://www.rfc-editor.org/info/rfc7275>>.
- [802.1q] IEEE, "IEEE Standard for Local and Metropolitan Area Networks -- Bridges and Bridged Networks", IEEE Std 802.1Q-2014, DOI 10.1109/IEEESTD.2014.6991462, 2014.
- [802.1d1998] IEEE, "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, DOI 10.1109/IEEESTD.1998.95619, 1998.
- [802.1d2004] IEEE, "IEEE Standard for Local and metropolitan area networks -- Media Access Control (MAC) Bridges", IEEE Std 802.1D-2004, DOI 10.1109/ieeestd.2004.94569, 2004.

7.2. Informative References

- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <<http://www.rfc-editor.org/info/rfc5226>>.

Acknowledgements

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

Authors' Addresses

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

Email: zhangmingui@huawei.com

Huafeng Wen
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Email: wenhuafeng@huawei.com

Jie Hu
China Telecom
Beijing Information Science & Technology Innovation Park
Beiqijia Town Changping District,
Beijing 102209
China

Email: hujie@ctbri.com.cn

