

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2013

M. Chen
W. Cao
Huawei Technologies Co., Ltd
A. Takacs
Ericsson
P. Pan
Infinera
October 21, 2012

LDP extensions for Pseudowire Binding to LSP Tunnels
draft-caopwe3-mpls-tp-pw-over-bidir-lsp-07.txt

Abstract

Many transport services require that user traffic, in the forms of Pseudowires (PW), to be delivered on a single co-routed bidirectional LSP or two LSPs that share the same routes. In addition, the user traffic may traverse through multiple transport networks.

This document specifies an optional extension in LDP that enable the binding between PWs and the underlying LSPs.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 24, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. LDP Extensions	5
2.1. PSN Tunnel Binding TLV	5
2.1.1. PSN Tunnel Sub-TLV	7
3. Theory of Operation	8
4. PSN Binding Operation for SS-PW	9
5. PSN Binding Operation for MS-PW	12
6. Security Considerations	13
7. IANA Considerations	13
7.1. LDP TLV Types	13
7.1.1. PSN Tunnel Sub-TLVs	14
7.2. LDP Status Codes	14
8. Acknowledgements	14
9. References	14
9.1. Normative References	14
9.2. Informative References	14
Authors' Addresses	15

1. Introduction

Pseudo Wire (PW) Emulation Edge-to-Edge (PWE3) [RFC3985] is a mechanism to emulate layer 2 services, such as Ethernet p2p circuits. Such services are emulated between two Attachment Circuits (ACs) and the PW encapsulated layer 2 service payload is carried through Packet Switching Network (PSN) tunnels between Provider Edges (PEs). PWE3 typically uses Label Distribution Protocol (LDP) [RFC5036] or Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) [RFC3209] LSPs as PSN tunnels. The PEs select and bind the Pseudowires to PSN tunnels independently. Today, there is no protocol-based provisioning mechanism to associate PW's to PSN tunnels.

PW-to-PSN Tunnel binding has become increasingly common and important in many deployment scenarios. For instance, when connecting two remotely located sites, such as data centers, over the backbone, each site may deploy a high-performance router or switch to aggregate thousands of Ethernet VLAN flows. The aggregating routers and switches are interconnected via one or multiple MPLS/GMPLS LSP's, which may traverse through different routes or networks. Further, each Ethernet flow is offered to the customers as a bidirectional circuits with certain SLA attributes, such as bandwidth and latency. Hence, it's important for the operators to map the forwarding and reverse-direction traffic from an Ethernet circuit to the LSP's that are either bidirectional (e.g. GMPLS-initiated optical path) or co-routed.

The requirement for explicit control of PW-to-LSP mapping has been described in Section 5.3.2 ("Support for Explicit Control of PW-to-LSP Binding") of [RFC6373]. The following figure (Figure 1) provides the illustration.

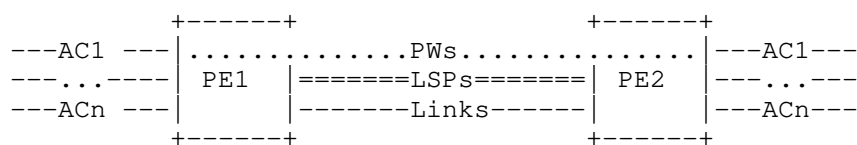


Figure 1: Explicit PW-to-LSP binding scenario

There are two PEs (PE1 and PE2) connected through multiple parallel links that may be on different fibers. Each link is managed and controlled as a bi-directional LSP. At each PE, there are a large number of bi-directional user flows from multiple Ethernet interfaces. Each user flow uses PW's to carry traffic on forwarding and reverse direction. The operators need to make sure that the user

flows (that is, the PW-pairs) to be carried on the same fiber (or, bidirectional LSP).

As mentioned above, there are a number of reasons behind this requirement. First, due to delay and latency constraints, traffic going over different fibers may require large amount of expensive buffer memory to compensate for the differential delay at the headend nodes. Further, the operators may apply different protection mechanisms on different parts of the network. As such, for optimal traffic management, traffic belongs to a particular user should traverse over the same fiber. That implies that both forwarding and reserve direction PW's that belong to the same user flow need to be mapped to the same co-routed bi-directional LSP or two LSPs with the same route.

Figure 2 illustrates a scenario where PW-LSP binding is not applied.

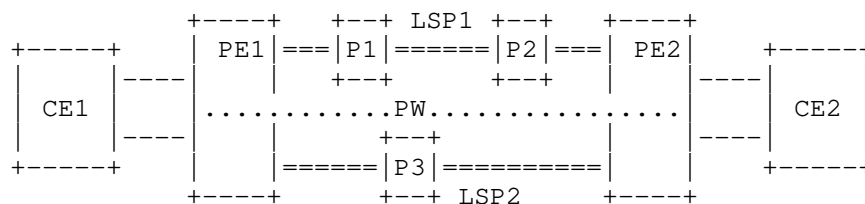


Figure 2: Inconsistent SS-PW to LSP binding scenario

LSP1 and LSP2 are two bidirectional connections on diverse paths. The operator is to deliver a bi-directional flow between PE1 and PE2. Using the existing mechanisms, it's possible that PE1 may select LSP1 (PE1-P1-P2-PE2) as the PSN tunnel for traffic from PE1 to PE2, while selecting LSP2 (PE1-P3-PE2) as the PSN tunnel for traffic from PE2 to PE1.

Consequently, the user traffic is delivered over two disjoint LSPs that may have very different service attributes in terms of latency and protection. This may not be acceptable as a reliable and effective transport service to the customers.

The similar problems may also exist in multi-segment PWs (MS-PWs), where user traffic on a particular PW may hop over different networks on forward and reverse directions.

One way to solve this problem is by introducing manual configuration at each PE to bind the PWs to the underlying PSN tunnels. However, this is prone to configuration errors and won't scale.

In this documentation, we will introduce an automatic solution by extending FEC 128/129 PW based on [RFC4447].

2. LDP Extensions

This document defines a new TLV, PSN Tunnel Binding TLV, to communicate tunnel/LSPs selection and binding requests between PEs. The TLV carries PW's binding profile and provides explicit or implicit information for the underlying PSN tunnel binding operation.

The binding TLV is optional, and MUST NOT affect the existing PW operation when not present in the messages.

The binding operation applies in both single-segment (SS) and multi-segment (MS) scenarios.

The extension supports two types of binding requests:

1. Strict binding: the requesting PE will choose and explicitly indicate the LSP information in the requests.
2. Congruent binding: a requesting PE will suggest an underlying LSP to a remote PE. On receive, the remote PE has the option to use the suggested LSP, or reply the information for an alternative.

In this document, the terminology of "tunnel" is identical to the "TE Tunnel" defined in Section 2.1 of [RFC3209], which is uniquely identified by a SESSION object that includes Tunnel end point address, Tunnel ID and Extended Tunnel ID. The terminology "LSP" is identical to the "LSP tunnel" defined in Section 2.1 of [RFC3209], which is uniquely identified by the SESSION object together with SENDER_TEMPLATE (or FILTER_SPEC) object that consists of LSP ID and Tunnel endpoint address.

2.1. PSN Tunnel Binding TLV

PSN Tunnel Binding TLV is an optional TLV and MUST be carried in the LDP Label Mapping message if PW to LSP binding is required. The format is as follows:

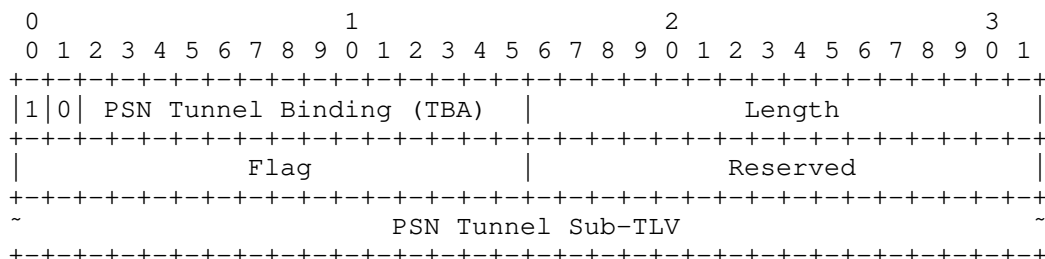
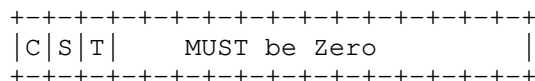


Figure 3: PSN Tunnel Binding TLV

The PSN Tunnel Binding TLV type is to be allocated by IANA

The Length field is 2 octets in length. It defines the length in octets of the entire TLV

The Flag field describes the binding requests, and has following format:



The flags are defined as the following:

C (Congruent path) bit: This informs the remote T-PE/S-PEs about the properties of the underlying LSPs. When set, the remote T-PE/S-PEs need to select LSPs with routes with the similar characteristics (that is, bidirectional or co-routed path). If there is no such tunnel available, the node may trigger the remote T-PE/S-PEs to establish a new LSP.

S (Strict) bit: This instructs the PEs with respect to the handling of the underlying LSPs. When set, the remote PE MUST use the tunnel/LSPs specified in the PSN Tunnel Sub-TLV as the PSN tunnel on the reverse direction of the PW, or the PW will fail to be established.

T (Tunnel Representation) bit: This indicates the format of the LSP tunnels. When the bit is set, the tunnel uses the tunnel information to identify itself, and the LSP Number fields in the PSN Tunnel sub-TLV (Section 2.1.1) MUST be set to zero. Otherwise, both tunnel and LSP information of the PSN tunnel are required. The default is set. The motivation for the T-bit is to support the MPLS protection operation where the LSP Number fields may be ignored.

C-bit and S-bit are mutually exclusive from each other, and cannot be set in the same message.

2.1.1. PSN Tunnel Sub-TLV

PSN Tunnel Sub-TLVs are designed for inclusion in the PSN Tunnel Binding TLV to specify the tunnel/LSPs to which a PW is required to bind.

Two sub-TLVs are defined: the IPv4 and IPv6 Tunnel sub-TLVs.

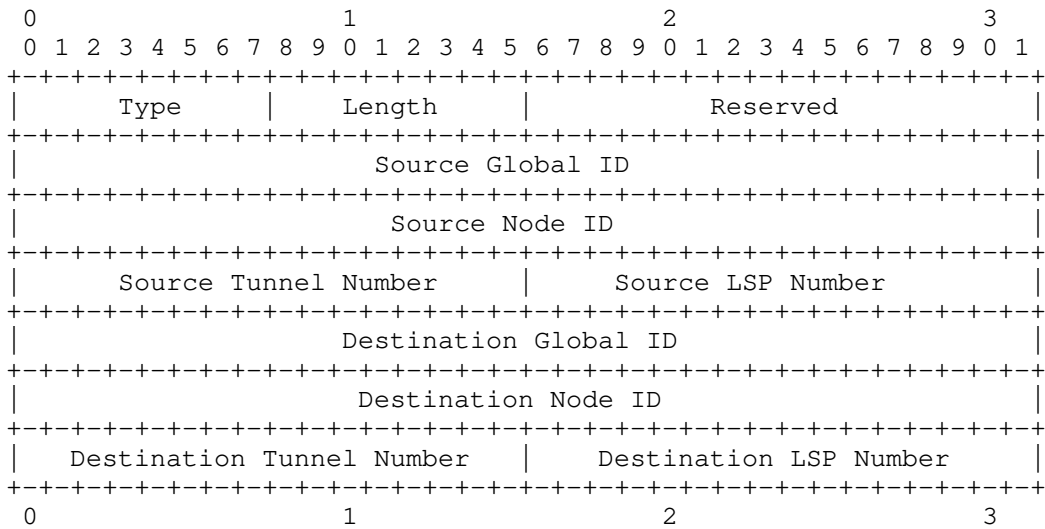


Figure 4: IPv4 PSN Tunnel sub-TLV format

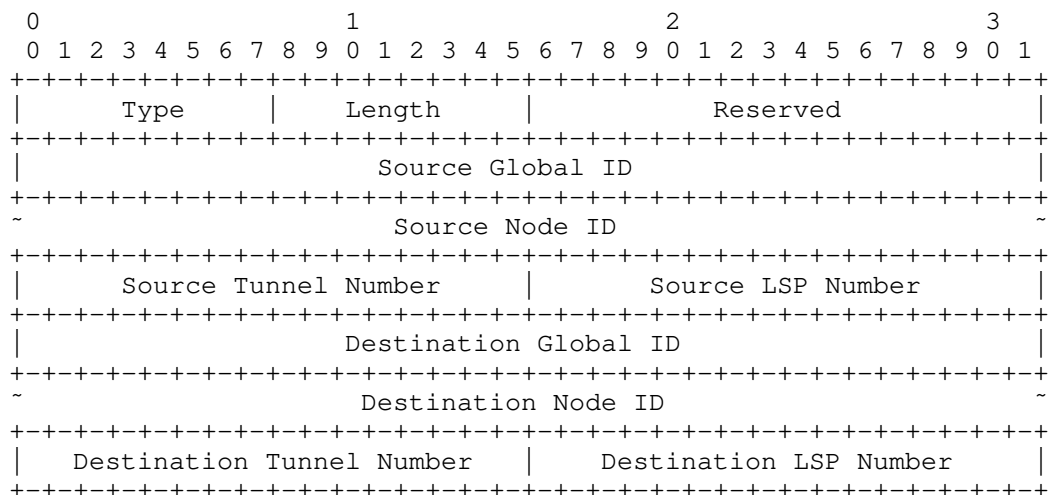


Figure 5: IPv6 PSN Tunnel sub-TLV format

The definition of Source and Destination Global/Node IDs and Tunnel/LSP Numbers are derived from [RFC6370]. This is to describe the underlying LSP's. Note that the LSP's in this notation is globally unique.

As defined in Section 4.6.1.2 and Section 4.6.2.2 of [RFC3209], the "Tunnel endpoint address" is mapped to Destination Node ID, and "Extended Tunnel ID" is mapped to Source Node ID. Both IDs can be IPv6 addresses.

A PSN Tunnel sub-TLV could be used to either identify a tunnel or a specific LSP. The T-bit in the Flag field defines the distinction as such that, when the T-bit is set, the Source/Destination LSP Number fields MUST be zero and ignored during processing. Otherwise, both Source/Destination LSP Number fields MUST have the actual LSP IDs of specific LSPs.

Each PSN Tunnel Binding TLV can only have one such sub-TLV.

3. Theory of Operation

During PW setup, the PEs may select desired forwarding tunnels/LSPs, and inform the remote T-PE/S-PEs about the desired reverse tunnels/LSPs.

Specifically, to set up a PW (or PW Segment), a PE may select a

candidate tunnel/LSP to act as the PSN tunnel. If none is available or satisfies the constraints, the PE will trigger and establish a new tunnel/LSP. The selected tunnel/LSP information is carried in the PSN Tunnel Binding TLV and sent with the Label Mapping message to the target PE.

Upon the reception of the Label Mapping message, the receiving PE will process the PSN Tunnel Binding TLV, determine whether it can accept the suggested tunnel/LSP or to find the reverse tunnel/LSP that meets the request, and respond with a Label Mapping message, which contains the corresponding PSN Tunnel Binding TLV.

It is possible that two PEs may request PSN binding to the same PW or PW segment over different tunnels/LSPs at the same time. There may cause collisions of tunnel/LSPs selection as both PEs assume the active role.

As defined in (Section 7.2.1, [RFC6073]), each PE may be generally categorized into active and passive roles:

1. Active PE: the PE which initiates the selection of the tunnel/LSPs and informs the remote PE;
2. Passive PE: the PE which obeys the active PE's suggestion.

In the remaining of this document, we will elaborate the operation for SS-PW and MS-PW:

1. SS-PW: In this scenario, both PE's for a particular PE may assume the active roles
2. MS-PW: One PE is active, while the other is passive. The PW's are setup using FEC 129

4. PSN Binding Operation for SS-PW

As illustrated in Figure-5, both PEs (say, PE1 and PE2) of a PW may independently initiate the setup. To perform PSN binding, the Label Mapping messages MUST carry a PSN Tunnel Binding TLV, and the PSN Tunnel sub-TLV MUST contains the desired tunnel/LSPs of the sender.

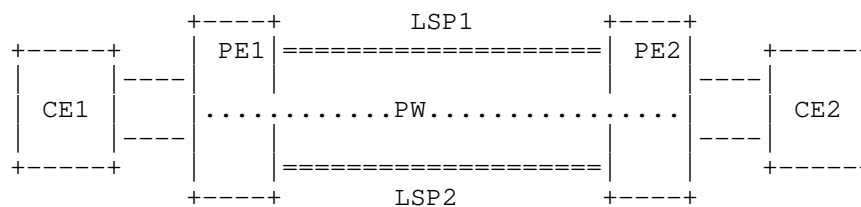


Figure 6: PSN binding operation in SS-PW environment

As outlined previously, there are two types of binding request: congruent and strict.

In strict binding, a PE (e.g., PE1) will mandate the other PE (e.g., PE2) to use a specified tunnel/LSP (e.g. LSP1) as the PSN tunnel on the reverse direction. In the PSN Tunnel Binding TLV, the S-bit MUST be set, the C-bit MUST be reset, and the Source and Destination IDs/Numbers MUST be filled.

On receive, if the S-bit is set, other than following the processing procedure defined in Section 5.3.3 of [RFC4447], the receiving PE (i.e. PE2) needs to determine whether to accept the indicated tunnel/LSP in PSN Tunnel Sub-TLV.

If the receiving PE (PE2) is also an active PE, and may have initiated the PSN binding requests to the other PE (PE1), if the received PSN tunnel/LSP is the same as it has been sent in the Label Mapping message by PE2, then the signaling has converged on a mutually agreed Tunnel/LSP. The binding operation is completed.

Otherwise, the receiving PE (PE2) MUST compare its own Node ID against the received Source Node ID. If it is numerically lower, the PE (PE2) will reply a Label Mapping message to complete the PW setup and confirm the binding request. The PSN Tunnel Binding TLV in the message MUST contain the same Source and Destination IDs/Numbers as in the received binding request, in the appropriate order. On the other hand, if the receiving PE (PE2) has a Node ID that is numerically higher than the Source Node ID carried in the PSN Tunnel Binding TLV, it MUST reply a Label Release message with status code set to "Reject to use the suggested tunnel/LSPs" and the received PSN Tunnel Binding TLV, and the PW will not be established.

To support congruent binding, the receiving PE can select the appropriated PSN tunnel/LSP for the reverse direction of the PW, so long as the forwarding and reverse PSNs share the same route.

Initially, a PE (PE1) sends a Label Mapping message to the remote PE (PE2) with the PSN Tunnel Binding TLV, with C-bit set, S-bit reset, and the appropriate Source and Destination IDs/Numbers. In case of

unidirectional LSPs, the PSN Tunnel Binding TLV may only contain the Source IDs/Numbers, the Destination IDs/Numbers are set to zero and left for PE2 to fill when responding the Label Mapping message.

On receive, since PE2 is also an active PE, and may have initiated the PSN binding requests to the other PE (PE1), if the received PSN tunnel/LSP has the same route as the one that has been sent in the Label Mapping message to PE1, then the signaling has converged. The binding operation is completed.

Otherwise, it needs to compare its own Node ID against the received Source Node ID. If it's numerically lower, PE2 needs to find/establish a tunnel/LSP that meets the congruent constraint, and reply a Label Mapping message with a PSN Binding TLV that contains the Source and Destination IDs/Numbers in the appropriate order. On the other hand, if the receiving PE (PE2) has a Node ID that is numerically higher than the Source Node ID carried in the PSN Tunnel Binding TLV, it MUST reply a Label Release message with status code set to "Reject to use the suggested tunnel/LSPs" and the received PSN Tunnel Binding TLV.

In both strict and congruent bindings, if T-bit is set, the LSP Number field MUST be set to zero. Otherwise, the field MUST contain the actual LSP number for the associated PSN LSP.

After a PW established, the operators may choose to move the PW's from the current tunnel/LSPs. Or, the underlying PSN is broken due to network failure. In this scenario, a new Label Mapping message MUST be sent to update the changes. Note that when T-bit is set, the working LSP broken will not trigger to update the changes if there are protection LSP's.

The message may carry a new PSN Tunnel Binding TLV, which contains the new Source and Destination Numbers/IDs. The handling of the new message should be identical to what has been described in this section.

However, if the new Label Binding message does not contain the PSN Tunnel Binding TLV, it declares the removal of any congruent/strict constraints. The PEs may not map the PW to the underlying PSN on purpose, the current independent PW to PSN binding will be used.

Further, as an implementation option, the PEs should not remove the traffic from an operational PW, until the completion of the underlying PSN tunnel/LSP changes.

5. PSN Binding Operation for MS-PW

MS-PW uses FEC 129 for PW setup. We refer the operation to Figure-6.

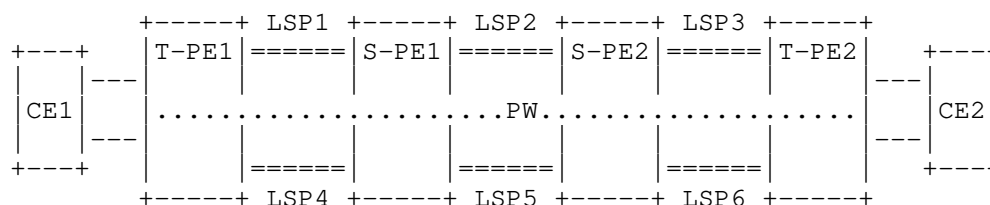


Figure 7: PSN binding operation in MS-PW environment

When an active PE (that is, T-PE1) starts to signal for a MS-PW, a PSN Tunnel Binding TLV MUST be carried in the Label Mapping message and sent to the adjacent S-PE (that is, S-PE1). The PSN Tunnel Binding TLV includes the PSN Tunnel sub-TLV that carries the desired tunnel/LSP of T-PE1's.

For strict binding, the initiating PE MUST set the S-bit, reset the C-bit and indicates the binding tunnel/LSP to the next-hop S-PE.

When S-PE1 receives the Label Mapping message, S-PE1 needs to determine if the signaling is for forward or reverse direction, as defined in Section 6.2.3 of [I-D.ietf-pwe3-dynamic-ms-pw].

If the Label Mapping message is for forward direction, and S-PE1 accepts the requested tunnel/LSPs from T-PE1, S-PE1 must save the tunnel/LSP information for reverse-direction processing later on. If the PSN binding request is not acceptable, S-PE1 MUST reply a Label Release Message to the upstream PE (T-PE1) with Status Code set to "Reject to use the suggested tunnel/LSPs".

Otherwise, S-PE1 relays the Label Mapping message to the next S-PE (that is, S-PE2), with the PSN Tunnel sub-TLV carrying the information of the new PSN tunnel/LSPs selected by S-PE1. S-PE2 and subsequent S-PEs will repeat the same operation until the Label Mapping message reaches to the remote T-PE (that is, T-PE2).

If T-PE2 agrees with the requested tunnel/LSPs, it will reply a Label Mapping message to initiate to the binding process on the reverse direction. The Label Mapping message contains the received PSN Tunnel Binding TLV for confirmation purposes.

When its upstream S-PE (S-PE2) receives the Label Mapping message, the S-PE relays the Label Mapping message to its upstream adjacent S-PE (S-PE1), with the previously saved PSN tunnel/LSP information in the PSN Tunnel sub-TLV. The same procedure will be applied on subsequent S-PEs, until the message reaches to T-PE1 to complete the PSN binding setup.

During the binding process, if any PE does not agree to the requested tunnel/LSPs, it can send a Label Release Message to its upstream adjacent PE with Status Code set to "Reject to use the suggested tunnel/LSPs".

For congruent binding, the initiating PE (T-PE1) MUST set the C-bit, reset the S-bit and indicates the suggested tunnel/LSP in PSN Tunnel sub-TLV to the next-hop S-PE (S-PE1).

During the MS-PW setup, the PEs have the option to ignore the suggested tunnel/LSP, and select another tunnel/LSP for the segment PW between itself and its upstream PE on reverse direction only if the tunnel/LSP is congruent with the forwarding one. Otherwise, the procedure is the same as the strict binding.

The tunnel/LSPs may change after a MS-PW being established. When a tunnel/LSP has changed, the PE that detects the change SHOULD select an alternative tunnel/LSP for temporary use while negotiating with other PEs following the procedure described in this section.

6. Security Considerations

The ability to control which LSP to carry traffic from a PW can be a potential security risk both for denial of service and traffic interception. It is RECOMMENDED that PEs do not accept the use of LSPs identified in the PSN Tunnel Binding TLV unless the LSP end points match the PW or PW segment end points. Furthermore, where security of the network is believed to be at risk, it is RECOMMENDED that PEs implement the LDP security mechanisms described in [RFC5036] and [RFC5920].

7. IANA Considerations

7.1. LDP TLV Types

This document defines new TLV [Section 2.1 of this document] for inclusion in LDP Label Mapping message. IANA is required to assign TLV type value to the new defined TLVs from LDP "TLV Type Name Space" registry.

7.1.1. PSN Tunnel Sub-TLVs

This document defines two sub-TLVs [Section 2.1.1 of this document] for PSN Tunnel Binding TLV. IANA is required to create a new registry ("PSN Tunnel Sub-TLV Name Space") for PSN Tunnel sub-TLVs and to assign Sub-TLV type values to the following sub-TLVs.

IPv4 PSN Tunnel sub-TLV - 0x01 (to be confirmed by IANA)

IPv6 PSN Tunnel sub-TLV - 0x02 (to be confirmed by IANA)

7.2. LDP Status Codes

This document defines a new LDP status codes, IANA is required to assigned status codes to these new defined codes from LDP "STATUS CODE NAME SPACE" registry.

"Reject to use the suggested tunnel/LSPs" - 0x0000003B (to be confirmed by IANA)

8. Acknowledgements

The authors would like to thank Adrian Farrel, Kamran Raza, Xinchun Guo, Mingming Zhu and Li Xue for their comments and help in preparing this document. Also this draft benefits from the discussions with Nabil Bitar, Paul Doolan, Frederic Journay, Andy Malis, Curtis Villamizar and Luca Martini.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.
- [RFC6370] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport Profile (MPLS-TP) Identifiers", RFC 6370, September 2011.

9.2. Informative References

- [I-D.ietf-pwe3-dynamic-ms-pw]
Martini, L., Bocci, M., and F. Balus, "Dynamic Placement

of Multi Segment Pseudowires",
draft-ietf-pwe3-dynamic-ms-pw-15 (work in progress),
June 2012.

- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
Tunnels", RFC 3209, December 2001.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching
(GMPLS) Signaling Functional Description", RFC 3471,
January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching
(GMPLS) Signaling Resource ReserVation Protocol-Traffic
Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-
Edge (PWE3) Architecture", RFC 3985, March 2005.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP
Specification", RFC 5036, October 2007.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS
Networks", RFC 5920, July 2010.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M.
Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.
- [RFC6373] Andersson, L., Berger, L., Fang, L., Bitar, N., and E.
Gray, "MPLS Transport Profile (MPLS-TP) Control Plane
Framework", RFC 6373, September 2011.

Authors' Addresses

Mach(Guoyi) Chen
Huawei Technologies Co., Ltd
Q14 Huawei Campus, No. 156 Beiqing Road, Hai-dian District
Beijing 100095
China

Email: mach@huawei.com

Wei Cao
Huawei Technologies Co., Ltd
Q14 Huawei Campus, No. 156 Beiqing Road, Hai-dian District
Beijing 100095
China

Email: wayne.caowei@huawei.com

Attila Takacs
Ericsson
Laborc u. 1.
Budapest 1037
Hungary

Email: attila.takacs@ericsson.com

Ping Pan
Infinera
169 West Java Drive, Sunnyvale, CA 94089
US

Email: ppan@infinera.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

J. Dong
M. Chen
Huawei Technologies
G. Mirsky
Ericsson
October 22, 2012

LDP Extensions for Lock Instruct and Loopback of Pseudowire in MPLS
Transport Profile
draft-dong-pwe3-mpls-tp-li-lb-02

Abstract

This document specifies extensions to the Label Distribution Protocol (LDP) to support provisioning of lock instruct and loopback mechanism for MPLS-TP Pseudowires.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. LDP Extensions	3
2.1. Extensions to MPLS-TP PW OAM Administration TLV	3
2.2. Extensions to PW Status TLV	4
3. Operations	4
3.1. Lock Instruct	4
3.2. Loopback	5
4. IANA Considerations	6
5. Security Considerations	6
6. References	6
6.1. Normative References	6
6.2. Informative References	7
Authors' Addresses	7

1. Introduction

The requirements for Lock Instruct (LI) and Loopback (LB) are specified in [RFC5860], and the framework of LI and LB is specified in [RFC6371]. [RFC6435] defines management plane based Lock Instruct (LI) and Loopback (LB) mechanisms, and an LI OAM message can be used for additional lock coordination between the MEPs. Management plane based LI and LB is suitable for scenarios where dynamic control plane is not available.

When a dynamic control plane is used for establishing MPLS-TP pseudowires (PWs), it's natural to use and extend the control plane protocol to provision LI and LB functions. Unlike other OAM mechanisms, LI and LB would modify the forwarding plane of a PW, thus without the involvement of control plane this may result in inconsistency between control plane and data plane. Besides, with control plane based mechanism, it does not need to rely on the TTL expiration to make the OAM requests reach particular MIP or MEP.

There are some existing control plane based OAM provisioning mechanisms for MPLS-TP. For example, [I-D.ietf-pwe3-oam-config] specifies the LDP extensions for the configuration of proactive OAM functions for MPLS-TP PWs when control plane is used.

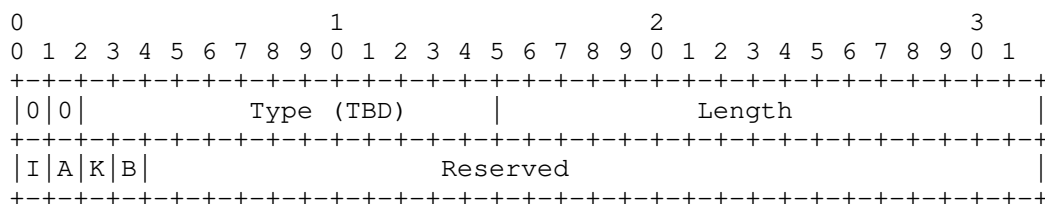
This document defines mechanisms similar to [I-D.ietf-pwe3-oam-config] to implement LI and LB functions for MPLS-TP PWs when MPLS-TP control plane is used. The mechanisms defined in this document are complementary to [RFC6435].

2. LDP Extensions

2.1. Extensions to MPLS-TP PW OAM Administration TLV

Two new flags (Lock bit and Loopback bit) are defined in MPLS-TP PW OAM Administration TLV [I-D.ietf-pwe3-oam-config].

Format of extended MPLS-TP PW OAM Administration TLV is as below:



Lock (K): When this bit is set, it indicates that the T-PE needs to

enable "Lock" function for this PW.

Loopback (B): When this bit is set, it indicates that the target node of this message SHOULD enable loopback function for this PW.

2.2. Extensions to PW Status TLV

Two new Status bits are defined in PW Status TLV:

Bit Mask	Description	
=====		
TBD1	Pseudowire in Lock Mode	[this document]
TBD2	Pseudowire in Loopback Mode	[this document]

3. Operations

The control plane based Lock Instruct and Loopback functions are applicable to both Single-Segment Pseudowire (SS-PW) [RFC3985] [RFC4447] and Multi-Segment Pseudowire (MS-PW) [RFC5659] [RFC6073].

3.1. Lock Instruct

When a PE/T-PE wants to put a PW into lock mode, it MUST send a Mapping message with the Lock (K) bit in the MPLS-TP PW OAM Administration TLV set.

For SS-PW, when the Mapping message arrives at the remote PE, the receiving PE SHOULD try to take the PW out of service. If the receiving PE locks the PW successfully, it SHOULD send a Notification message with PW status "Pseudowire in Lock Mode". Otherwise, it SHOULD send a Notification message with the LDP Status code set to "PW Lock Failure".

For MS-PW, when the Mapping message arrives at a downstream S-PE, the receiving S-PE SHOULD forward this Mapping message with the K bit unchanged towards the remote T-PE. When the Mapping message arrives at the remote T-PE, it SHOULD try to take the PW out of service. If the receiving T-PE locks the PW successfully, it SHOULD send a Notification message with PW status "Pseudowire in Lock Mode" to the upstream S-PE. Otherwise, it SHOULD send a Notification message with the LDP Status code set to "PW Lock Failure". On receipt of the Notification message, the S-PEs would know whether the MS-PW is in lock mode or not, and the S-PEs SHOULD forward the Notification message back to the Source T-PE.

When the PE/T-PE wants to take the PW out of the lock mode, it MUST send a Mapping message with the Lock (K) bit in the MPLS-TP PW OAM Administration TLV cleared. The receiving PE/T-PE SHOULD try to

unlock the PW. If the PW is unlocked successfully, the receiving PE/T-PE SHOULD send a Notification message with PW status bit "Pseudowire in Lock Mode" cleared. Otherwise, it SHOULD send a Notification message with the LDP Status code set to "PW Unlock Failure".

3.2. Loopback

When a PE/T-PE wants to put the remote PE/T-PE of a PW into loopback mode, it MUST send a Mapping message with both the Lock (K) bit and Loopback (B) bit in the MPLS-TP PW OAM Administration TLV set. When a T-PE wants to put a particular S-PE of the PW into loopback mode, it MUST send a Mapping message with both the Lock (K) bit and Loopback (B) bit set, and an Explicit Route Hop TLV (ER-Hop TLV) [I-D.ietf-pwe3-mspw-er] identifying the Target S-PE node MUST be carried in the Mapping message. The L flag in the ER-Hop TLV SHOULD be cleared. To ensure that the ER-Hop TLV identifies a single node as the Target S-PE, The PreLen field in the IPv4 prefix ER-Hop TLV SHOULD be set to 32, the PreLen field in the IPv6 prefix ER-Hop TLV SHOULD be set to 128, and the PreLen field in the L2 PW Address ER-Hop TLV SHOULD be set to 96. Information of the S-PE node can be collected using the SP-PE TLVs [RFC6073].

When the Mapping message arrives at the remote PE/T-PE, the receiving PE SHOULD try to put the PW in loopback mode. If the receiver node puts the PW into loopback mode successfully, it SHOULD send a Notification message with PW status "Pseudowire in Loopback Mode". Otherwise, it SHOULD send a Notification message with the LDP Status code set to "PW Enter Loopback Failure".

When a Mapping message with an ER-Hop TLV arrives an S-PE, the S-PE SHOULD check the ER-Hop TLV to see if it is the target S-PE of the message. If not, the S-PE SHOULD forward the message with the K and B bit unchanged to the next hop S-PE. When the Mapping message arrives at the target S-PE, the S-PE SHOULD parse the MPLS-TP PW OAM Administration TLV and try to put the PW into loopback mode. If the S-PE puts the PW into loopback mode successfully, it SHOULD send a Notification message with PW status set to "Pseudowire in Loopback Mode". An SP-PE TLV identifying the S-PE in loopback mode SHOULD also be carried in the Notification message. If the S-PE fails to put the PW into loopback mode, it SHOULD send a Notification message with the LDP Status code set to "PW Enter Loopback Failure". An SP-PE TLV identifying this S-PE SHOULD also be carried in the Notification message.

When the PE/T-PE wants to take the remote PE/T-PE out of the loopback mode, it MUST send a Mapping message with the Lock (K) bit set and Loopback (B) bit cleared. When the T-PE wants to take a particular

S-PE out of loopback mode, the message MUST also carry an ER-Hop TLV to identify the target S-PE. If the PW is taken out of loopback mode successfully, the receiving PE/T-PE/S-PE SHOULD send a Notification message with PW status bit "Pseudowire in Loopback Mode" cleared. Otherwise, it SHOULD send a Notification message with the LDP Status code set to "PW Exit Loopback Failure". For the S-PE case, An SP-PE TLV identifying this S-PE node SHOULD also be carried in the Notification message.

4. IANA Considerations

Two bits ("Lock" (K) and "Loopback" (B)) as defined in section 2.1 need to be allocated in the MPLS-TP PW OAM Administration TLV.

Two new PW Status bits as defined in section 2.2 need to be allocated in the "Pseudowire Status Codes" Registry.

Four new LDP status codes need to be assigned by the IANA in the LDP "STATUS CODE NAME SPACE":

Range/Value	E	Description
TBA	0	PW Lock Failure
TBA	0	PW Unlock Failure
TBA	0	PW Enter Loopback Failure
TBA	0	PW Exit Loopback Failure

5. Security Considerations

TBD

6. References

6.1. Normative References

- [I-D.ietf-pwe3-mspw-er]
Dutta, P., Bocci, M., and L. Martini, "Explicit Path Routing for Dynamic Multi-Segment Pseudowires", draft-ietf-pwe3-mspw-er-01 (work in progress), June 2012.
- [I-D.ietf-pwe3-oam-config]
Zhang, F., Bo, W., and E. Bellagamba, "Label Distribution Protocol Extensions for Proactive Operations, Administration and Maintenance Configuration of Dynamic MPLS Transport Profile PseudoWire", draft-ietf-pwe3-oam-config-01 (work in progress),

August 2012.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.
- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.
- [RFC5659] Bocci, M. and S. Bryant, "An Architecture for Multi-Segment Pseudowire Emulation Edge-to-Edge", RFC 5659, October 2009.
- [RFC5860] Vigoureux, M., Ward, D., and M. Betts, "Requirements for Operations, Administration, and Maintenance (OAM) in MPLS Transport Networks", RFC 5860, May 2010.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.
- [RFC6371] Busi, I. and D. Allan, "Operations, Administration, and Maintenance Framework for MPLS-Based Transport Networks", RFC 6371, September 2011.

6.2. Informative References

- [RFC6435] Boutros, S., Sivabalan, S., Aggarwal, R., Vigoureux, M., and X. Dai, "MPLS Transport Profile Lock Instruct and Loopback Functions", RFC 6435, November 2011.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Mach Chen
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: mach.chen@huawei.com

Greg Mirsky
Ericsson

Email: gregory.mirsky@ericsson.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

J. Dong
H. Wang
Huawei Technologies
October 22, 2012

Pseudowire Redundancy on S-PE
draft-dong-pwe3-redundancy-spe-03

Abstract

This document describes Multi-Segment Pseudowire (MS-PW) protection scenarios in which the pseudowire redundancy is provided on the Switching-PE (S-PE). Operations of the S-PEs which provide PW redundancy are specified. Signaling of the preferential forwarding status as defined in [I-D.ietf-pwe3-redundancy-bit] is reused.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. PW Redundancy on S-PE	3
3. S-PE Operations	4
4. VCCV Considerations	6
5. IANA Considerations	6
6. Security Considerations	7
7. Acknowledgements	7
8. References	7
8.1. Normative References	7
8.2. Informative References	7
Authors' Addresses	8

1. Introduction

[I-D.ietf-pwe3-redundancy] describes the framework and requirements for pseudowire (PW) redundancy, and [I-D.ietf-pwe3-redundancy-bit] specifies Pseudowire (PW) redundancy mechanism for scenarios where a set of redundant PWs is configured between provider edge (PE) nodes in single-segment pseudowire (SS-PW) [RFC3985] applications, or between terminating provider edge (T-PE) nodes in multi-segment pseudowire (MS-PW) [RFC5659] applications.

In some MS-PW scenarios, there are some benefits to provide PW redundancy on S-PEs, such as reducing the burden on the access T-PE nodes, and faster protection switching. This document describes some scenarios in which PW redundancy is provided on S-PEs, and specifies the operations of the S-PEs. Signaling of the preferential forwarding status as defined in [I-D.ietf-pwe3-redundancy-bit] is reused.

2. PW Redundancy on S-PE

In some MS-PW deployment scenarios, there are some benefits to provide PW redundancy on S-PEs. This section gives some examples of PW redundancy on S-PE.

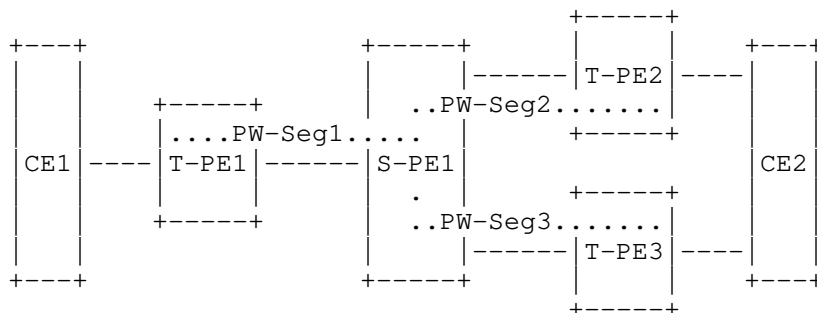


Figure 1. MS-PW Redundancy on S-PE

As illustrated in Figure 1, CE1 is connected to T-PE1 while CE2 is dual-homed to T-PE2 and T-PE3. T-PE1 is connected to S-PE1 only, and S-PE1 is connected to T-PE2 and T-PE3. The MS-PW is switched on S-PE1, and PW-Seg2 and PW-Seg3 provides resiliency on S-PE1 for failure of T-PE2 or T-PE3 or the connected ACs. PW-Seg2 is selected as primary PW segment, and PW-Seg3 is secondary PW segment.

MS-PW redundancy on S-PE is beneficial for the scenario in Figure 1 since T-PE1 as an access node may not be able to provide PW redundancy, especially when the PW-Seg1 between T-PE1 and S-PE1 is

statically configured. And with PW redundancy on S-PE, the number of PW segments needed between T-PE1 and S-PE1 is only half of the number of PW segments needed for end-to-end MS-PW redundancy. In addition, PW redundancy on S-PE could provide faster protection switching than end-to-end protection switching of MS-PW.

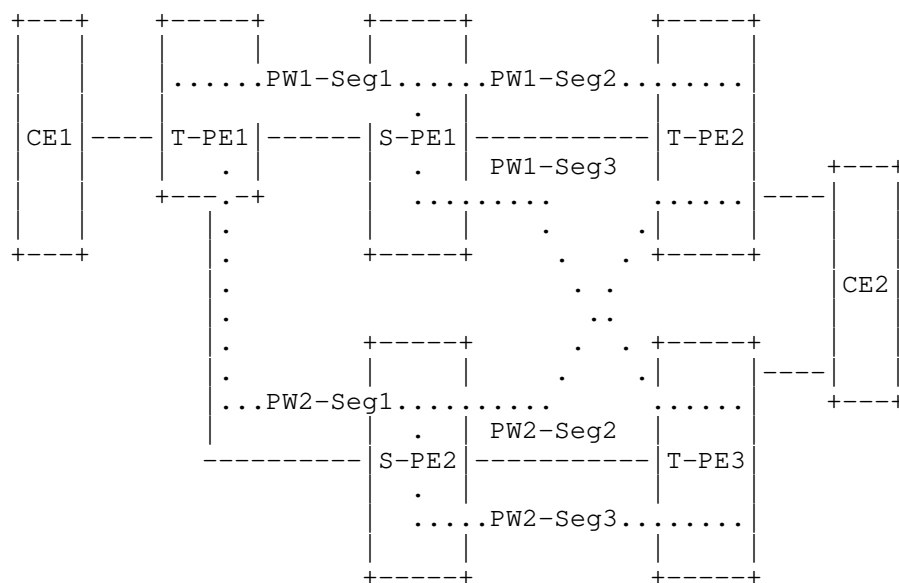


Figure 2. MS-PW Redundancy on S-PE with S-PE protection

As illustrated in Figure 2, CE1 is connected to T-PE1 while CE2 is dual-homed to T-PE2 and T-PE3. T-PE1 is connected to S-PE1 and S-PE2, and both S-PE1 and S-PE2 are connected to T-PE2 and T-PE3. There are two MS-PWs which are switched at S-PE1 and S-PE2 respectively to provide S-PE node protection. For MS-PW1, the S-PE1 provides resiliency using PW1-Seg2 and PW1-Seg3. For MS-PW2, the S-PE2 provides resiliency using PW2-Seg2 and PW2-Seg3. MS-PW1 is the primary PW and PW1-Seg2 is the primary PW segment.

MS-PW redundancy on S-PE is beneficial for the scenario in Figure 2 since it reduces the number of end-to-end MS-PWs required for both T-PE and S-PE protection. In addition, PW redundancy on S-PE could provide faster protection switching than end-to-end protection switching of MS-PW.

3. S-PE Operations

For an S-PE which provides PW redundancy, it is important to advertise proper preferential forwarding status to the PW segments on

both sides and perform protection switching according to the received status. This section specifies the operations of S-PEs on which PW redundancy is provisioned.

The S-PE SHOULD work as a Slave node for the single-connected side, and SHOULD work in Independent mode for the multi-connected side. The S-PE SHOULD pass the preferential forwarding status received from the single-connected side unchanged to the PW segments on the multi-connected side. The S-PE SHOULD advertise Standby status to the single-connected side if it receives Standby status from all the PW segments on the multi-connected side, and it SHOULD advertise Active status to the single-connected side if it receives Active status from any of the PW segments on the multi-connected side. For the single-connected side, the active PW segment is determined by the T-PE on this side, which works as the Master node. On the multi-connected side, the PW segment which has both local and remote Preferential Forwarding status as Active SHOULD be selected for traffic forwarding.

The Signaling of Preferential Forwarding bit defined in [I-D.ietf-pwe3-redundancy-bit] is reused in these scenarios.

For the scenario in Figure 1, assume the AC from CE2 to T-PE2 is active. In normal operation, S-PE1 would receive Active Preferential Forwarding status bit on the single-connected side from T-PE1, then it would advertise Active Preferential Forwarding status bit on both PW-Seg2 and PW-Seg3. T-PE2 and T-PE3 would advertise Active and Standby preferential status bit respectively to S-PE1, reflecting the forwarding state of the two ACs to CE2. By matching the local and remote Up/Down status and Preferential Forwarding status, PW-Seg2 would be used for traffic forwarding.

On failure of the AC between CE2 and T-PE2, the forwarding state of AC on T-PE3 is changed to Active. T-PE3 then advertises Active Preferential Status to S-PE1, and T-PE2 would advertise the Preferential Status bit of Standby to S-PE1. S-PE1 would perform the switchover according to the updated local and remote Preferential Forwarding status, and select PW-Seg3 for traffic forwarding. Since S-PE1 still connects to an Active PW segment on the multi-connected side, it will not advertise any change of the PW Preferential Forwarding status to T-PE1. T-PE1 would not be aware of the switchover on S-PE1.

For scenario of Figure 2, assume the AC from CE2 to T-PE2 is active. T-PE1 works in Master mode and it would advertise Active and Standby Preferential Forwarding status bit respectively to S-PE1 and S-PE2. According to the received Preferential Forwarding status bit, S-PE1 would advertise Active Preferential Forwarding status bit to both

T-PE2 and T-PE3, and S-PE2 would advertise Standby Preferential Forwarding status bit to both T-PE2 and T-PE3. T-PE2 would advertise Active Preferential Forwarding status bit to both S-PE1 and S-PE2, and T-PE3 would advertise Standby Preferential Forwarding status bit to both S-PE1 and S-PE2, reflecting the forwarding state of the two ACs to CE2. By matching the local and remote Up/Down Status and Preferential Forwarding status, PW1-Seg2 from S-PE1 to T-PE2 would be used for traffic forwarding. Since S-PE1 connects to the Active PW segment on the multi-connected side, it would advertise Active Preferential Forwarding status bit to T-PE1, and S-PE2 would advertise Standby Preferential Forwarding status bit to T-PE1 since it does not have any Active PW segment on the multi-connected side.

On failure of the AC between CE2 and T-PE2, the forwarding state of AC on T-PE3 is changed to Active. T-PE3 would then advertise Active Preferential Forwarding status bit to both S-PE1 and S-PE2, and T-PE2 would advertise Standby Preferential Forwarding status bit to both S-PE1 and S-PE2. S-PE1 would perform the switchover according to the updated local and remote Preferential Forwarding status, and select PW1-Seg3 for traffic forwarding. Since S-PE1 still has an Active PW segment on the multi-connected side, it would not advertise any change of the PW status to T-PE1. Thus T-PE1 would not be aware of the switchover on S-PE1.

If S-PE1 fails, T-PE1 would notice this through some detection mechanism and then advertise the Active Preferential Forwarding status bit to S-PE2, and PW2-Seg1 would be selected by T-PE1 for traffic forwarding. On receipt of the newly changed Preferential Forwarding status, S-PE2 would advertise the Active Preferential Forwarding status to both T-PE2 and T-PE3. T-PE2 and T-PE3 would also notice the failure of S-PE1 by some detection mechanism. Then by matching the local and remote Up/Down and Preferential Forwarding status, PW2-Seg2 would be selected for traffic forwarding.

4. VCCV Considerations

PW VCCV [RFC5085] CC type 1 "PW ACH" can be used with S-PE redundancy mechanism. VCCV CC type 2 "Router Alert Label" is not supported for MS-PW as specified in [RFC6073]. If VCCV CC type 3 "TTL Expiry" is to be used, the hop count from one T-PE to the remote T-PE needs to be obtained in advance. This can be achieved either by control plane SP-PE TLVs or through data plane tracing of the MS-PW.

5. IANA Considerations

This document makes no request of IANA.

6. Security Considerations

This document has the same security properties as in the PWE3 control protocol [RFC4447] and [I-D.ietf-pwe3-redundancy-bit].

7. Acknowledgements

The authors would like to thank Mach Chen, Lizhong Jin and Mustapha Aissaoui for their comments and discussions.

8. References

8.1. Normative References

- [I-D.ietf-pwe3-redundancy]
Muley, P., Aissaoui, M., and M. Bocci, "Pseudowire Redundancy", draft-ietf-pwe3-redundancy-09 (work in progress), June 2012.
- [I-D.ietf-pwe3-redundancy-bit]
Muley, P. and M. Aissaoui, "Pseudowire Preferential Forwarding Status Bit", draft-ietf-pwe3-redundancy-bit-08 (work in progress), September 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.
- [RFC5659] Bocci, M. and S. Bryant, "An Architecture for Multi-Segment Pseudowire Emulation Edge-to-Edge", RFC 5659, October 2009.

8.2. Informative References

- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.
- [RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Haibo Wang
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: rainsword.wang@huawei.com

PWE3
Internet-Draft
Intended status: Informational
Expires: January 25, 2015

YJ. Stein
RAD Data Communications
D. Black
EMC Corporation
B. Briscoe
BT
July 24, 2014

Pseudowire Congestion Considerations
draft-ietf-pwe3-congcons-02

Abstract

Pseudowires (PWs) have become a common mechanism for tunneling traffic, and may be found in unmanaged scenarios competing for network resources both with other PWs and with non-PW traffic, such as TCP/IP flows. It is thus worthwhile specifying under what conditions such competition is safe, i.e., the PW traffic does not significantly harm other traffic or contribute more than it should to congestion. We conclude that PWs transporting responsive traffic behave as desired without the need for additional mechanisms. For inelastic PWs (such as TDM PWs) we derive a bound under which such PWs consume no more network capacity than a TCP flow. We also propose employing a transport circuit breaker [I-D.fairhurst-tsvwg-circuit-breaker] that shuts down a TDM PW consistently surpassing this bound, as the emulated TDM service itself would be of insufficient quality.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 25, 2015.

Copyright Notice

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

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. PWs Comprising Elastic Flows	4
3. PWs Comprising Inelastic Flows	5
4. Security Considerations	16
5. IANA Considerations	16
6. Informative References	17
Appendix A. Loss Probabilities for TDM PWs	18
Appendix B. Effect of Packet Loss on Voice Quality for TDM PWs .	19
Authors' Addresses	22

1. Introduction

A pseudowire (PW) (see [RFC3985]) is a construct for tunneling a native service, such as Ethernet or TDM, over a Packet Switched Network (PSN), such as IPv4, IPv6, or MPLS. The PW packet encapsulates a unit of native service information by prepending the headers required for transport in the particular PSN (which must include a demultiplexer field to distinguish the different PWs) and preferably the 4 byte PWE3 control word.

PWs have no bandwidth reservation or control mechanisms, meaning that when multiple PWs are transported in parallel, and/or in parallel with other flows, there is no defined means for allocating resources for any particular PW, or for preventing negative impact of a particular PW on neighboring flows. Mechanisms for provisioning PWs in service provider networks are well understood and will not be discussed further here.

While PWs are most often placed in MPLS tunnels, there are several mechanisms that enable transporting PWs over an IP infrastructure. These include:

UDP/IP encapsulations defined for TDM PWs
([RFC4553][RFC5086][RFC5087]),
L2TPv3 based PWs,
MPLS PWs directly over IP according to RFC 4023 [RFC4023],
MPLS PWs over GRE over IP according to RFC 4023 [RFC4023].

Whenever PWs are transported over IP, they may compete for network resources with neighboring congestion-responsive flows (e.g., TCP flows). In this document we study the effect of PWs on such neighboring flows, and discover that the negative impact of PW traffic is generally no worse than that of congestion-responsive flows, ([RFC2914],[RFC5033]).

At first glance one may consider a PW transported over IP to be considered as a single flow, on a par with a single TCP flow. Were we to accept this tenet, we would require a PW to back off under congestion to consume no more bandwidth than a single TCP flow under such conditions (see [RFC5348]). However, since PWs may carry traffic from many users, it makes more sense to consider each PW to be equivalent to multiple TCP flows.

The following two sections consider PWs of two types.

Elastic Flows: Section 2 concludes that the response to congestion of a PW carrying elastic (e.g., TCP) flows is no different to the combined behaviour of the set of the same elastic flows were they not encapsulated within a PW.

Inelastic Flows: Section 3 considers the case of inelastic constant bit-rate (CBR) TDM PWs ([RFC4553][RFC5086][RFC5087]) competing with TCP flows. Such PWs require a preset amount of bandwidth, that may be lower or higher than that consumed by an otherwise unconstrained TCP flow under the same network conditions. In any case, such a PW is unable to respond to congestion in a TCP-like manner; on the other hand, the total bandwidth it consumes remains constant and does not increase to consume additional bandwidth as TCP rates back off. If the bandwidth consumed by a TDM PW is considered detrimental, the only available remedy is to completely shut down the PW, by using a transport circuit breaker mechanism. However, we will show that even before such an action is warranted, the PW will become unable to faithfully emulate the native TDM service; for example, when a TDM service is carrying voice grade telephony channels, the voice quality will degrade to below useful levels.

Thus, in both cases, pseudowires will not inflict significant harm on neighboring TCP flows, as in one case they respond adequately to congestion, and in the other they would be shut down due to being

unable to emulate the native service before harming neighboring flows.

2. PWs Comprising Elastic Flows

In this section we consider Ethernet PWs that primarily carry congestion-responsive traffic. We show that we automatically obtain the desired congestion avoidance behavior, and that additional mechanisms are not needed.

Let us assume that an Ethernet PW aggregating several TCP flows is flowing alongside several TCP/IP flows. Each Ethernet PW packet carries a single Ethernet frame that carries a single IP packet that carries a single TCP segment. Thus, if congestion is signaled by an intermediate router dropping a packet, a single end-user TCP/IP packet is dropped, whether or not that packet is encapsulated in the PW.

The result is that the individual TCP flows inside the PW experience the same drop probability as the non-PW TCP flows. Thus the behavior of a TCP sender (retransmitting the packet and appropriately reducing its sending rate) is the same for flows directly over IP and for flows inside the PW. In other words, individual TCP flows are neither rewarded nor penalized for being carried over the PW. An elastic PW does not behave as a single TCP flow, as it will consume the aggregated bandwidth of its component flows; yet if its component TCP flows backs off by some percentage, the bandwidth of the PW as a whole will be reduced by the very same percentage, purely due to the combined effect of its component flows.

This is, of course, precisely the desired behavior. Were individual TCP flows rewarded for being carried over a PW, this would create an incentive to create PWs for no operational reason. Were individual flows penalized, there would be a deterrence that could impede pseudowire deployment.

There have been proposals to add additional TCP-friendly mechanisms to PWs, for example by carrying PWs over DCCP. In light of the above arguments, it is clear that this would force the PW down to the bandwidth of a single flow, rather than N flows, and penalize the constituent TCP flows. In addition, the individual TCP flows would still back off due to their end points being oblivious to the fact that they are carried over a PW. This would further degrade the flow's throughput as compared to a non-PW-encapsulated flow, in contradiction to desirable behavior.

3. PWs Comprising Inelastic Flows

Inelastic PWs, such as TDM PWs ([RFC4553][RFC5086][RFC5087]), are potentially more problematic than the elastic PWs of the previous section. Being constant bit-rate (CBR), TDM PWs can not be made responsive to congestion. On the other hand, being CBR, they also do not attempt to capture additional bandwidth when neighboring TCP flows back off.

Since a TDM PW continuously consumes a constant amount of bandwidth, if the bandwidth occupied by a TDM PW endangers the network as a whole, the only recourse is to shut it down, denying service to all customers of the TDM native service. We can accomplish this by employing a transport circuit breaker, by which we mean an automatic mechanism for terminating a flow to prevent negative impact on other flows and on the stability of the network [I-D.fairhurst-tsvwg-circuit-breaker]. Note that a transport circuit breaker is intended as a protection mechanism of last resort, just as an electrical circuit breaker is only triggered when absolutely necessary. We should mention in passing that under certain conditions it may be possible to reduce the bandwidth consumption of a TDM PW. A prevalent case is that of a TDM native service that carries voice channels that may not all be active. Using the AAL2 mode of [RFC5087] (perhaps along with connection admission control) can enable bandwidth adaptation, at the expense of more sophisticated native service processing (NSP).

In the following we will show that for many cases of interest a TDM PW, treated as a single flow, will behave in a reasonable manner without any additional mechanisms. We will focus on structure-agnostic TDM PWs [RFC4553] although our analysis can be readily applied to structure-aware PWs (see Appendix A).

In order to quantitatively compare TDM PWs to TCP flows, we will compare the effect of TDM PW packets with that of TCP packets of the same packet size and sent at the same rate. This is potentially an overly pessimistic comparison, as TDM PW packets are frequently configured to be short in order to minimize latency, while TCP packets are free to be much larger.

There are two network parameters relevant to our discussion, namely the one-way delay D and the packet loss rate PLR. The one-way delay of a native TDM service consists of the physical time-of-flight plus 125 microseconds for each TDM switch traversed; and is thus very small as compared to typical PSN network-crossing latencies. Many protocols and applications running over TDM circuits thus expect extremely low delay, and thus in our comparisons we will only consider delays of a few milliseconds.

Regarding packet loss, the TDM PW RFCs specify behaviors upon detecting a lost packet. Structure-agnostic transport has no alternative to outputting an "all-ones" AIS pattern towards the TDM circuit, which, when long enough in duration, is recognized by the receiving TDM device as a fault indication (see Appendix A). International standards place stringent limits on the number of such faults tolerated. Calculations presented in the appendix show that only loss probabilities in the realm of fractions of a percent are relevant for structure-agnostic transport (see Appendix A). Structure-aware transport regenerates frame alignment signals thus hiding AIS indications resulting from infrequent packet loss. Furthermore, for TDM circuits carrying voice channels the use of packet loss concealment algorithms is possible (such algorithms have been previously described for TDM PWs). However, even structure-aware transport ceases to provide a useful service at about 2 percent loss probability. Hence, in our comparisons we will only consider PLRs of 1 or 2 percent.

RFC 5348 on TCP Friendly Rate Control (TFRC) [RFC5348] provides a simplified formula for TCP throughput as a function of delay and packet loss rate.

$$X = \frac{S}{R \left(\sqrt{2p/3} + 12 \sqrt{3p/8} p (1+32p^2) \right)}$$

where

X is average sending rate in Bytes per second,
 S is the segment (packet payload) size in Bytes,
 R is the round-trip time in seconds,
 p is the packet loss probability (i.e., PLR/100).

We can now compare the bandwidth consumed by TDM pseudowires with that of a TCP flow for given packet loss and delay. The results are depicted in the accompanying figures (available only in the PDF version of this document). In Figures 1 and 2 we see the conventional rate vs. packet loss plot for low-rate TDM (both T1 and E1) traffic, as well as TCP traffic with the same payload size (64 or 256 Bytes respectively). Since the TDM rates are constant (T1 and E1 having payload throughputs of 1.544 Mbps and 2.048 Mbps respectively), and the TDM service can only be faithfully emulated using SAToP up to a PLR of about half a percent, the T1 and E1 pseudowires occupy line segments on the graph. On the other hand, the TCP rate equation produces rate curves dependent on both delay and packet loss.

We see that in general for large packet sizes, short delays, and low packet loss rates, the TDM pseudowires consume much less bandwidth than TCP would under identical conditions. Only for small packets, long delays, and high packet loss ratios, do TDM PWs potentially consume more bandwidth, and even then only marginally. Similarly, in Figures 3 and 4 we repeat the exercise for higher rate E3 and T3 (rates 34.368 and 44.736 Mbps respectively) pseudowires, allowing delays and PLRs suitable appropriate for these signals. We see that the TDM pseudowires consume much less bandwidth than TCP, for all reasonable parameter combinations.

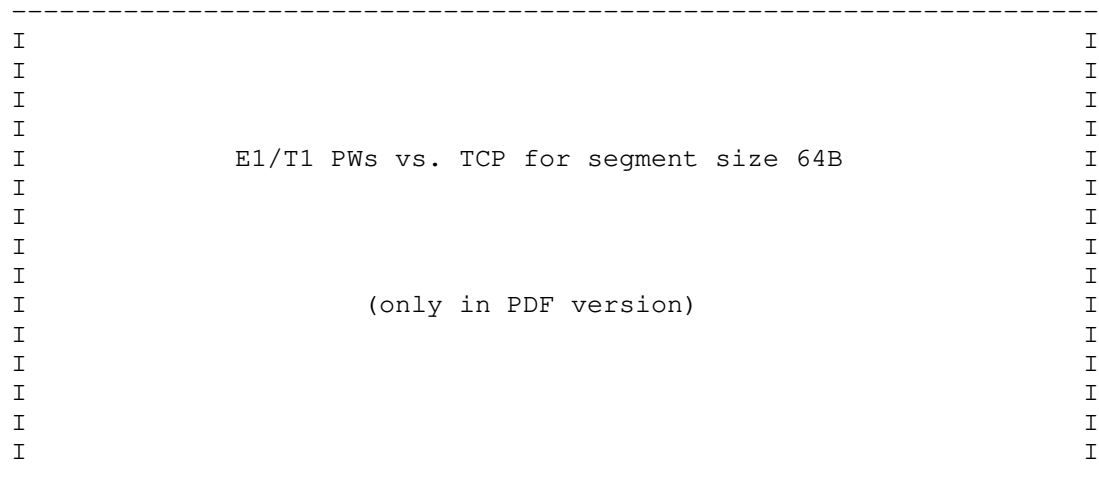


Figure 1 E1/T1 PWs vs. TCP for segment size 64B

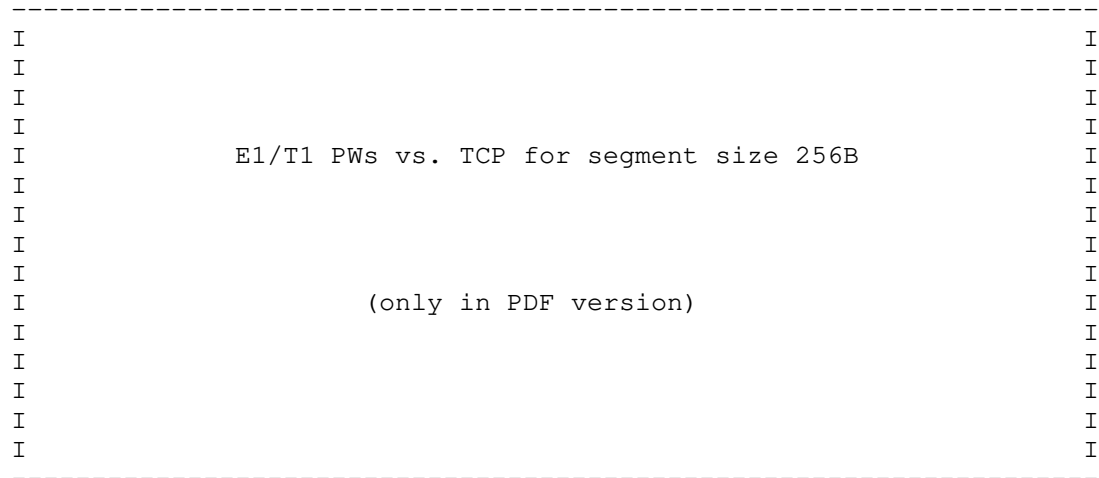


Figure 2 E1/T1 PWs vs. TCP for segment size 256B

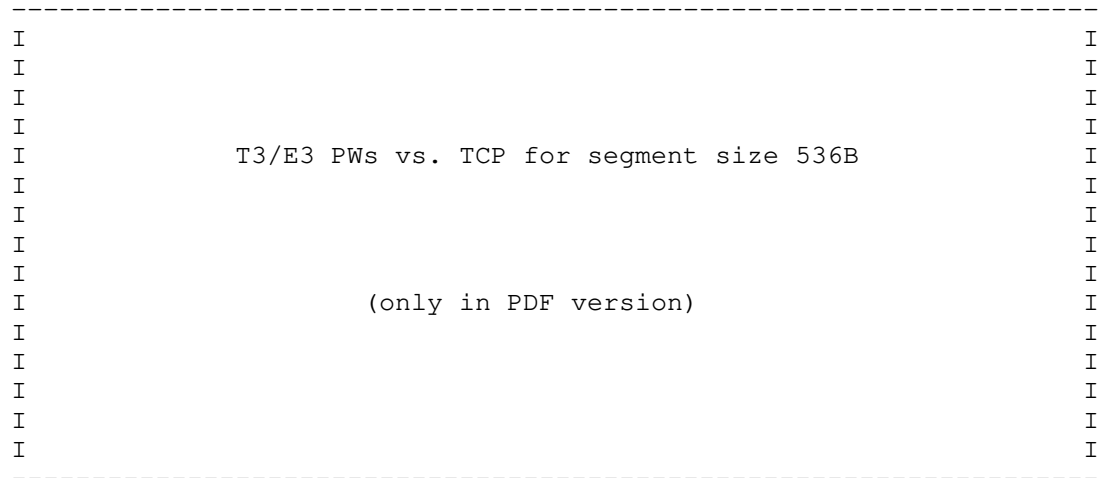


Figure 3 T3/E3 PWs vs. TCP for segment size 536B

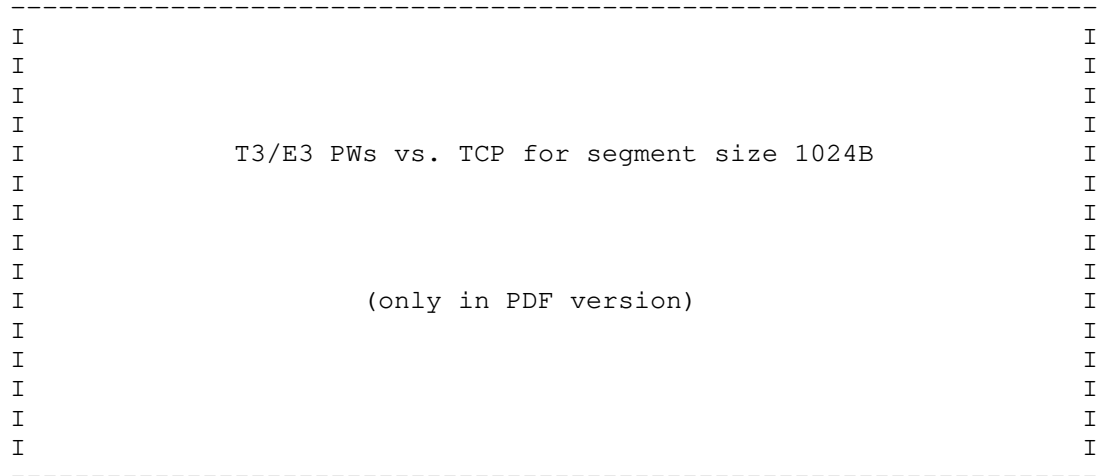


Figure 4 T3/E3 PWs vs. TCP for segment size 1024B

We can use the TCP rate equation to determine precise conditions under which a TDM PW consumes no more bandwidth than a TCP flow between the same endpoints would consume under identical conditions. Replacing the round-trip delay with twice the one-way delay D , setting the bandwidth to that of the TDM service BW , and the segment size to be the TDM fragment (taking into account the PWE3 control word), we obtain the following condition for a TDM PW.

$$D < \frac{4 S}{BW f(p)}$$

where

D is the one-way delay,
 S is the TDM segment size (packet excluding overhead) in Bytes,
 BW is TDM service bandwidth in bits per second,
 $f(p) = \sqrt{2p/3} + 12 \sqrt{3p/8} p (1+32p^2)$.

One may view this condition as defining an operating envelope for a TDM PW, as a TDM PW that occupies no more bandwidth than a TCP flow causes no more congestion than that TCP flow would. Under this condition it is safe to place the TDM PW along with congestion-responsive traffic such as TCP, without causing additional congestion. On the other hand, were the TDM PW to consume significantly more bandwidth a TCP flow, it could contribute disproportionately to congestion, and its mixture with congestion-responsive traffic might be inappropriate.

We derived this condition assuming steady-state conditions, and thus two caveats are in order. First, the condition does not specify how to treat a TDM PW that initially satisfies the condition, but is then faced with a deteriorating network environment. In such cases one additionally needs to analyze the reaction times of the responsive flows to congestion events. Second, the derivation assumed that the TDM PW was competing with long-lived TDM flows, because under this assumption it was straightforward to obtain a quantitative comparison with something widely considered to offer a safe response to congestion. Short-lived TCP flows may find themselves disadvantaged as compared to a long-lived TDM PW satisfying the condition.

We see in Figures 5 and 6 that TDM pseudowires carrying T1 or E1 native services satisfy the condition for all parameters of interest for large packet sizes (e.g., $S=512$ Bytes of TDM data). For the SAToP default of 256 Bytes, as long as the one-way delay is less than 10 milliseconds, the loss probability can exceed 0.3 or 0.6 percent. For packets containing 128 or 64 Bytes the constraints are more troublesome, but there are still parameter ranges where the TDM PW

consumes less than a TCP flow under similar conditions. Similarly, Figures 7 and 8 demonstrate that E3 and T3 native services with the SAToP default of 1024 Bytes of TDM per packet satisfy the condition for a broad spectrum of delays and PLRs.

Note that violating the condition for a short amount of time is not sufficient justification for shutting down the TDM PW. While TCP flows react within a round trip time, PW commissioning and decommissioning are time consuming processes that should only be undertaken when it becomes clear that the congestion is not transient.

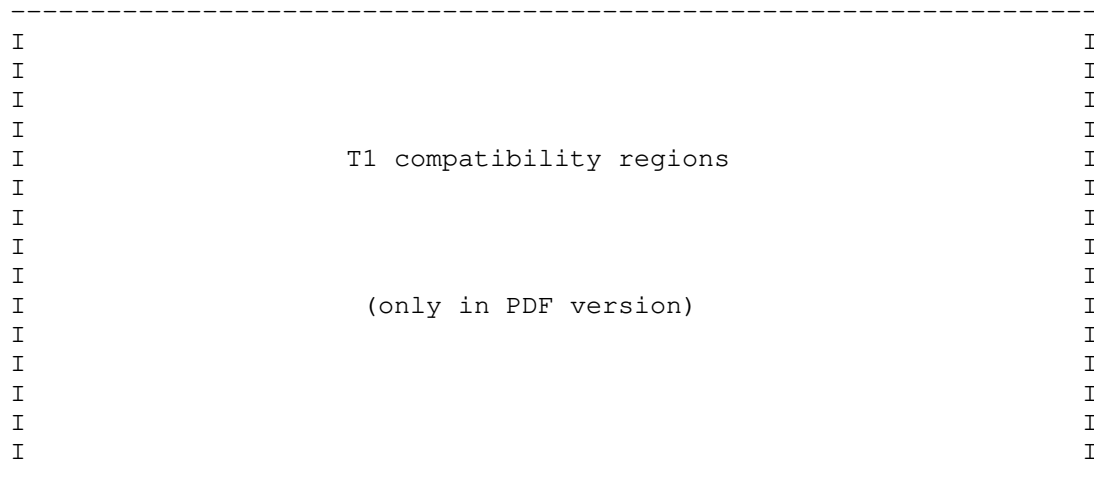


Figure 5 TCP Compatibility areas for T1 SAToP

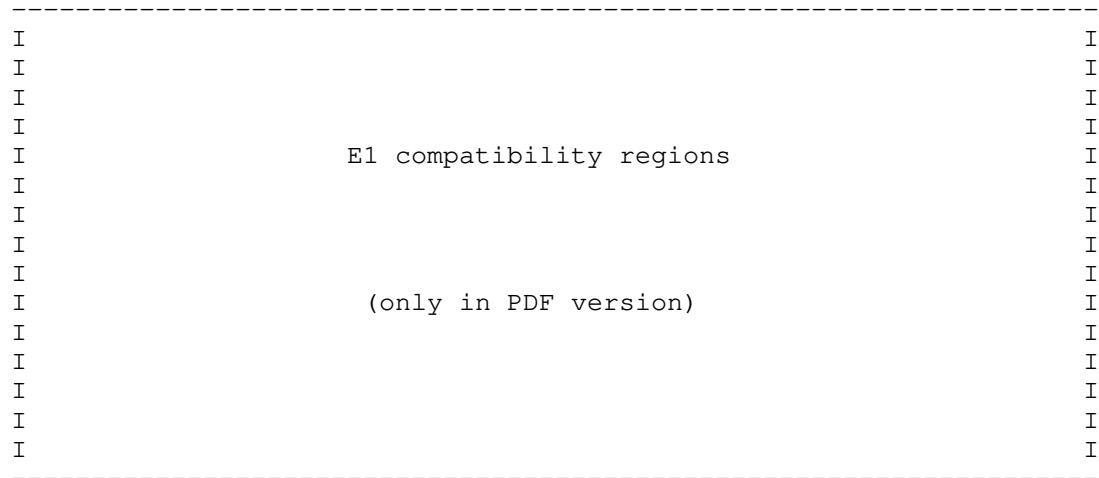


Figure 6 TCP Compatibility areas for E1 SAToP

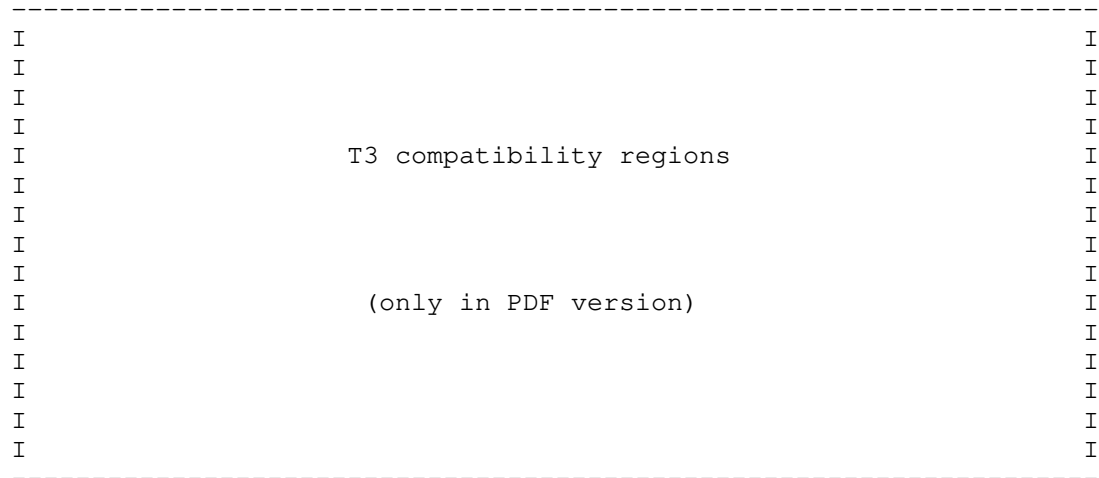


Figure 8 TCP Compatibility areas for T3 SAToP

4. Security Considerations

This document does not introduce any new congestion-specific mechanisms and thus does not introduce any new security considerations above those present for PWs in general.

5. IANA Considerations

This document requires no IANA actions.

6. Informative References

- [RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, September 2000.
- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.
- [RFC4023] Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, March 2005.
- [RFC4553] Vainshtein, A. and YJ. Stein, "Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)", RFC 4553, June 2006.
- [RFC5033] Floyd, S. and M. Allman, "Specifying New Congestion Control Algorithms", BCP 133, RFC 5033, August 2007.
- [RFC5086] Vainshtein, A., Sasson, I., Metz, E., Frost, T., and P. Pate, "Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)", RFC 5086, December 2007.
- [RFC5087] Stein, Y(J)., Shashoua, R., Insler, R., and M. Anavi, "Time Division Multiplexing over IP (TDMoIP)", RFC 5087, December 2007.
- [RFC5348] Floyd, S., Handley, M., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 5348, September 2008.
- [G775] International Telecommunications Union, "Loss of Signal (LOS), Alarm Indication Signal (AIS) and Remote Defect Indication (RDI) defect detection and clearance criteria for PDH signals", ITU Recommendation G.775, October 1998.
- [G826] International Telecommunications Union, "Error Performance Parameters and Objectives for International Constant Bit Rate Digital Paths at or above Primary Rate", ITU Recommendation G.826, December 2002.
- [P862] International Telecommunications Union, "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs", ITU Recommendation G.826, February 2001.

[I-D.stein-pwe3-tdm-packetloss]

Stein, Y(J). and I. Druker, "The Effect of Packet Loss on Voice Quality for TDM over Pseudowires", October 2003.

[I-D.fairhurst-tsvwg-circuit-breaker]

Fairhurst, G., "Network Transport Circuit Breakers", draft-fairhurst-tsvwg-circuit-breaker-01 (work in progress), May 2014.

Appendix A. Loss Probabilities for TDM PWs

ITU-T Recommendation G.826 [G826] specifies limits on the Errored Second Ratio (ESR) and the Severely Errored Second Ratio (SESR). For our purposes, we will simplify the definitions and understand an Errored Second (ES) to be a second of time during which a TDM bit error occurred or a defect indication was detected. A Severely Errored Second (SES) is an ES second during which the Bit Error Rate (BER) exceeded one in one thousand (10^{-3}). Note that if the error condition AIS was detected according to the criteria of ITU-T Recommendation G.775 [G826] a SES was considered to have occurred. The respective ratios are the fraction of ES or SES to the total number of seconds in the measurement interval.

For both E1 and T1 TDM circuits, G.826 allows ESR of 4% (0.04), and SESR of 1/5% (0.002). For E3 and T3 the ESR must be no more than 7.5% (0.075), while the SESR is unchanged.

Focusing on E1 circuits, the ESR of 4% translates, assuming the worst case of isolated exactly periodic packet loss, to a packet loss event no more than every 25 seconds. However, once a packet is lost, another packet lost in the same second doesn't change the ESR, although it may contribute to the ES becoming a SES. Assuming an integer number of TDM frames per PW packet, the number of packets per second is given by packets per second = $8000 / (\text{frames per packet})$, where prevalent cases are 1, 2, 4 and 8 frames per packet. Since for these cases there will be 8000, 4000, 2000, and 1000 packets per second, respectively, the maximum allowed packet loss probability is 0.0005%, 0.001%, 0.002%, and 0.004% respectively.

These extremely low allowed packet loss probabilities are only for the worst case scenario. In reality, when packet loss is above 0.001%, it is likely that loss bursts will occur. If the lost packets are sufficiently close together (we ignore the precise details here) then the permitted packet loss rate increases by the appropriate factor, without G.826 being cognizant of any change. Hence the worst-case analysis is expected to be extremely pessimistic for real networks. Next we will go to the opposite extreme and assume that all packet loss events are in periodic loss bursts. In

order to minimize the ESR we will assume that the burst lasts no more than one second, and so we can afford to lose no more than packet per second packets in each burst. As long as such one-second bursts do not exceed four percent of the time, we still maintain the allowable ESR. Hence the maximum permissible packet loss rate is 4%. Of course, this estimate is extremely optimistic, and furthermore does not take into consideration the SESR criteria.

As previously explained, a SES is declared whenever AIS is detected. There is a major difference between structure-aware and structure-agnostic transport in this regards. When a packet is lost SAToP outputs an "all-ones" pattern to the TDM circuit, which is interpreted as AIS according to G.775 [G775]. For E1 circuits, G.775 specifies for AIS to be detected when four consecutive TDM frames have no more than 2 alternations. This means that if a PW packet or consecutive packets containing at least four frames are lost, and four or more frames of "all-ones" output to the TDM circuit, a SES will be declared. Thus burst packet loss, or packets containing a large number of TDM frames, lead SAToP to cause high SESR, which is 20 times more restricted than ESR. On the other hand, since structure-aware transport regenerates the correct frame alignment pattern, even when the corresponding packet has been lost, packet loss will not cause declaration of SES. This is the main reason that SAToP is much more vulnerable to packet loss than the structure-aware methods.

For realistic networks, the maximum allowed packet loss for SAToP will be intermediate between the extremely pessimistic estimates and the extremely optimistic ones. In order to numerically gauge the situation, we have modeled the network as a four-state Markov model, (corresponding to a successfully received packet, a packet received within a loss burst, a packet lost within a burst, and a packet lost when not within a burst). This model is an extension of the widely used Gilbert model. We set the transition probabilities in order to roughly correspond to anecdotal evidence, namely low background isolated packet loss, and infrequent bursts wherein most packets are lost. Such simulation shows that up to 0.5% average packet loss may occur and the recovered TDM still conform to the G.826 ESR and SESR criteria.

Appendix B. Effect of Packet Loss on Voice Quality for TDM PWs

Packet loss in voice traffic can cause in gaps or artifacts that result in choppy, annoying or even unintelligible speech. The precise effect of packet loss on voice quality has been the subject of detailed study in the VoIP community, but VoIP results are not directly applicable to TDM PWs. This is because VoIP packets typically contain over 10 milliseconds of the speech signal, while

multichannel TDM packets may contain only a single sample, or perhaps a very small number of samples.

The effect of packet loss on TDM PWs has been previously reported [I-D.stein-pwe3-tdm-packetloss]. In that study it was assumed that each packet carried a single sample of each TDM timeslot (although the extension to multiple samples is relatively straightforward and does not drastically change the results). Four sample replacement algorithms were compared, differing in the value used to replace the lost sample:

1. replacing every lost sample by a preselected constant (e.g., zero or "AIS" insertion),
2. replacing a lost sample by the previous sample,
3. replacing a lost sample by linear interpolation between the previous and following samples,
4. replacing the lost sample by STatistically Enhanced INTERpolation (STEIN).

Only the first method is applicable to SAToPtransport, as structure awareness is required in order to identify the individual voice channels. For structure aware transport, the loss of a packet is typically identified by the receipt of the following packet, and thus the following sample is usually available. The last algorithm posits the LPC speech generation model and derives lost samples based on available samples both before and after each lost sample.

The four algorithms were compared in a controlled experiment in which speech data was selected from English and American English subsets of the ITU-T P.50 Appendix 1 corpus [P.50App1] and consisted of 16 speakers, eight male and eight female. Each speaker spoke either three or four sentences, for a total of between seven and 15 seconds. The selected files were filtered to telephony quality using modified IRS filtering and downsampled to 8 KHz. Packet loss of 0, 0.25, 0.5, 0.75, 1, 2, 3, 4 and 5 percent were simulated using a uniform random number generator (bursty packet loss was also simulated but is not reported here). For each file the four methods of lost sample replacement were applied and the Mean Opinion Score (MOS) was estimated using PESQ [P862]. Figure 5 depicts the PESQ-derived MOS for each of the four replacement methods for packet drop probabilities up to 5%.

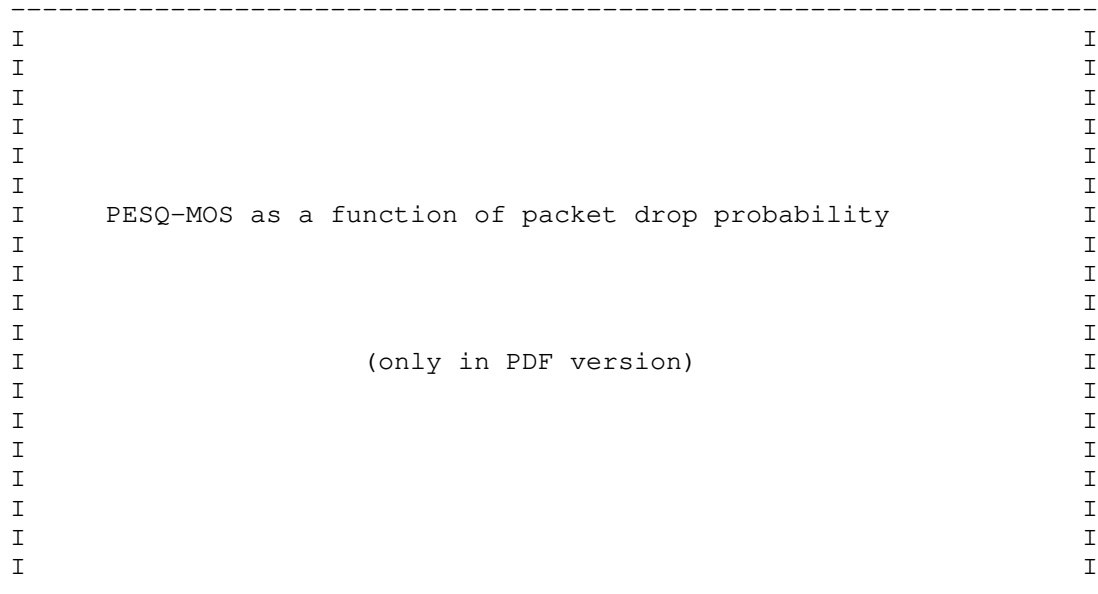


Figure 5 PESQ derived MOS as a function of packet drop probability

For all cases the MOS resulting from the use of zero insertion is less than that obtained by replacing with the previous sample, which in turn is less than that of linear interpolation, which is slightly less than that obtained by statistical interpolation.

Unlike the artifacts speech compression methods may produce when subject to buffer loss, packet loss here effectively produces additive white impulse noise. The subjective impression is that of static noise on AM radio stations or crackling on old phonograph records. For a given PESQ-derived MOS, this type of degradation is more acceptable to listeners than choppiness or tones common in VoIP.

If MOS>4 (full toll quality) is required, then the following packet drop probabilities are allowable:

```

zero insertion - 0.05 %
previous sample - 0.25 %
linear interpolation - 0.75 %
STEIN - 2 %

```

If MOS>3.75 (barely perceptible quality degradation) is acceptable, then the following packet drop probabilities are allowable:

zero insertion - 0.1 %
previous sample - 0.75 %
linear interpolation - 3 %
STEIN - 6.5 %

If MOS>3.5 (cell-phone quality) is tolerable, then the following packet drop probabilities are allowable:

zero insertion - 0.4 %
previous sample - 2 %
linear interpolation - 8 %
STEIN - 14 %

Authors' Addresses

Yaakov (Jonathan) Stein
RAD Data Communications
24 Raoul Wallenberg St., Bldg C
Tel Aviv 69719
ISRAEL

Phone: +972 (0)3 645-5389
Email: yaakov_s@rad.com

David L. Black
EMC Corporation
176 South St.
Hopkinton, MA 69719
USA

Phone: +1 (508) 293-7953
Email: david.black@emc.com

Bob Briscoe
BT
B54/77, Adastral Park
Martlesham Heath
Ipswich IP5 3RE
UK

Phone: +44 1473 645196
Email: bob.briscoe@bt.com
URI: <http://bobbbriscoe.net/>

PWE3
Internet-Draft
Updates: 5085 (if approved)
Intended status: Standards Track
Expires: March 6, 2015

T. Nadeau
lucidvision
L. Martini
S. Bryant
Cisco Systems
September 2, 2014

A Unified Control Channel for Pseudowires
draft-ietf-pwe3-vccv-for-gal-02

Abstract

This document describes a unified mode of operation for Virtual Circuit Connectivity Verification (VCCV), which provides a control channel that is associated with a pseudowire (PW). VCCV applies to all supported access circuit and transport types currently defined for PWs, as well as those being transported by the MPLS Transport Profile. This new mode is intended to augment those described in RFC5085. It describes new rules requiring this mode to be used as the default/mandatory mode of operation for VCCV. The older VCCV types will remain optional.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 6, 2015.

Copyright Notice

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

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Requirements Language and Terminology	2
2. Introduction	3
3. VCCV Control Channel When The Control Word is Used	5
4. VCCV Control Channel When The Control Word is Not Used	6
5. VCCV Capability Advertisement	7
6. Manageability Considerations	7
7. Security Considerations	7
8. IANA Considerations	7
8.1. VCCV Interface Parameters Sub-TLV	7
8.2. MPLS VCCV Control Channel (CC) Type 4	7
9. Acknowledgements	8
10. References	8
10.1. Normative References	8
10.2. Informative References	8
Authors' Addresses	9

1. Requirements Language and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

AC	Attachment Circuit [RFC3985].
AVP	Attribute Value Pair [RFC3931].
CC	Control Channel (used as CC Type).
CE	Customer Edge.
CV	Connectivity Verification (used as CV Type).
CW	Control Word [RFC3985].
L2SS	L2-Specific Sublayer [RFC3931].
LCCE	L2TP Control Connection Endpoint [RFC3931].

OAM	Operation and Maintenance.
PE	Provider Edge.
PSN	Packet Switched Network [RFC3985].
PW	Pseudowire [RFC3985].
PW-ACH	PW Associated Channel Header [RFC4385].
VCCV	Virtual Circuit Connectivity Verification [RFC5085].

2. Introduction

There is a need for fault detection and diagnostic mechanisms that can be used for end-to-end fault detection and diagnostics for a Pseudowire, as a means of determining the PW's true operational state. Operators have indicated in [RFC4377], and [RFC3916] that such a tool is required for PW operation and maintenance. To this end, the IETF's PWE3 Working Group defined the Virtual Circuit Connectivity Verification Protocol (VCCV) in [RFC5085]. Since then a number of interoperability issues have arisen with the protocol as it is defined.

Over time, a variety of VCCV options or "modes" have been created to support legacy hardware, these modes use of the CW in some cases, while in others the CW is not used. The difficulty of operating these different combinations of "modes" have been detailed in an implementation survey conducted by the PWE3 Working Group and documented in [RFC7079]. The implementation survey and the PWE3 Working Group have concluded that operators have difficulty deploying the VCCV OAM protocol due to the number of combinations and options for its use.

In addition to the implementation issues just described, the ITU-T and IETF have set out to enhance MPLS to make it suitable as an optical transport protocol. The requirements for this protocol are defined as the MPLS Transport Profile (MPLS-TP). The requirements for MPLS-TP can be found in [RFC5654]. In order to support VCCV when an MPLS-TP PSN is in use, the GAL-ACH had to be created [RFC5586]. This resulted in yet another mode of VCCV operation.

This document defines two modes of operation of VCCV: 1) with a control word or 2) without a control word, both with a ACH encapsulation making it possible to handle all of the other cases handled by the other modes of VCCV. The modes of operation defined in this document MUST be implemented.

Figure 1 depicts the architecture of a pseudowire as defined in [RFC3985]. It further depicts where the VCCV control channel resides within this architecture, which will be discussed in detail later in this document.

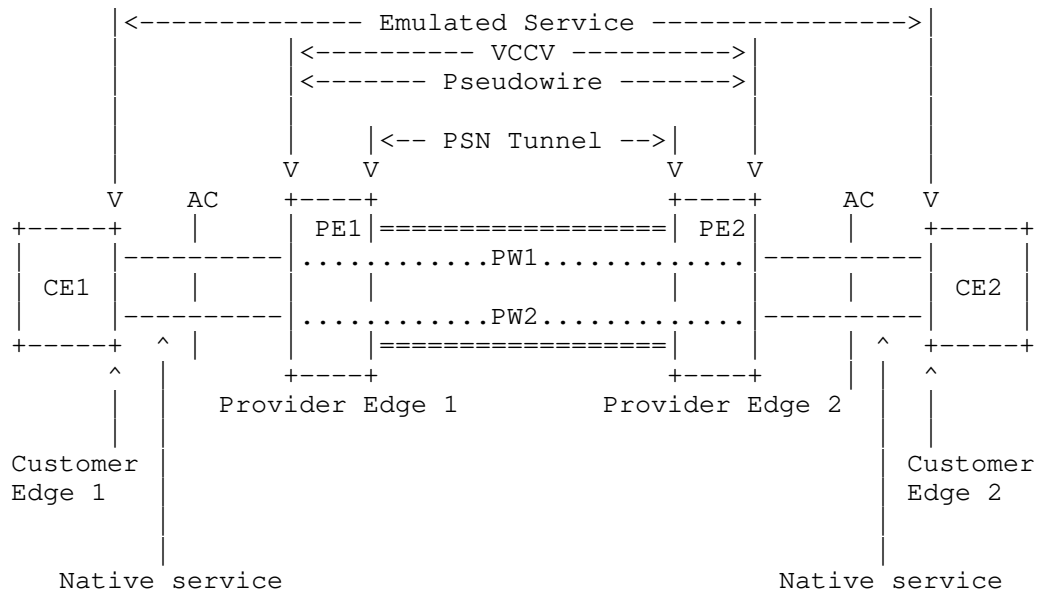


Figure 1: PWE3 VCCV Operation Reference Model

From Figure 1, Customer Edge (CE) routers CE1 and CE2 are attached to the emulated service via Attachment Circuits (AC), and to each of the Provider Edge (PE) routers (PE1 and PE2, respectively). An AC can be a Frame Relay Data Link Connection Identifier (DLCI), an ATM Virtual Path Identifier / Virtual Channel Identifier (VPI/VCI), an Ethernet port, or any other attachment type for which a PW is defined. The PE devices provide pseudowire emulation, enabling the CEs to communicate over the PSN. A pseudowire exists between these PEs traversing the provider network. VCCV provides several means of creating a control channel over the PW, between the PE routers that attach the PW.

Figure 2 depicts how the VCCV control channel is associated with the pseudowire protocol stack.

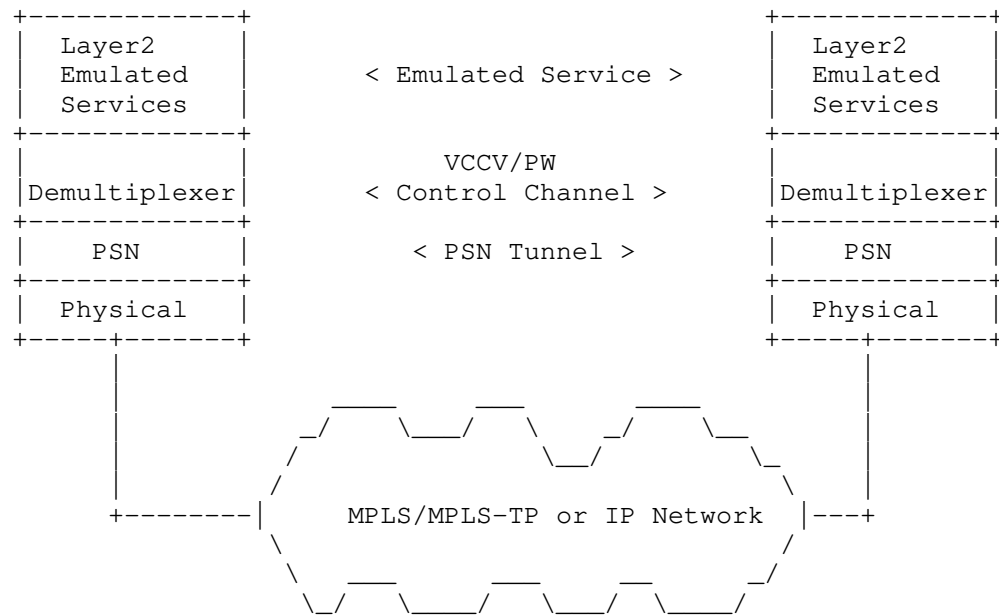


Figure 2: PWE3 Protocol Stack Reference Model including the VCCV Control Channel

VCCV messages are encapsulated using the PWE3 encapsulation as described in Section 3 and Section 4, so that they are handled and processed in the same manner (or in some cases, a similar manner) the PW PDUs for which they provide a control channel. These VCCV messages are exchanged only after the capability (the VCCV Control Channel and Connectivity Verification types) and the desire to exchange VCCV traffic has been advertised between the PEs (see Sections 5.3 and 6.3 of [RFC5085]), and VCCV type to use have been chosen.

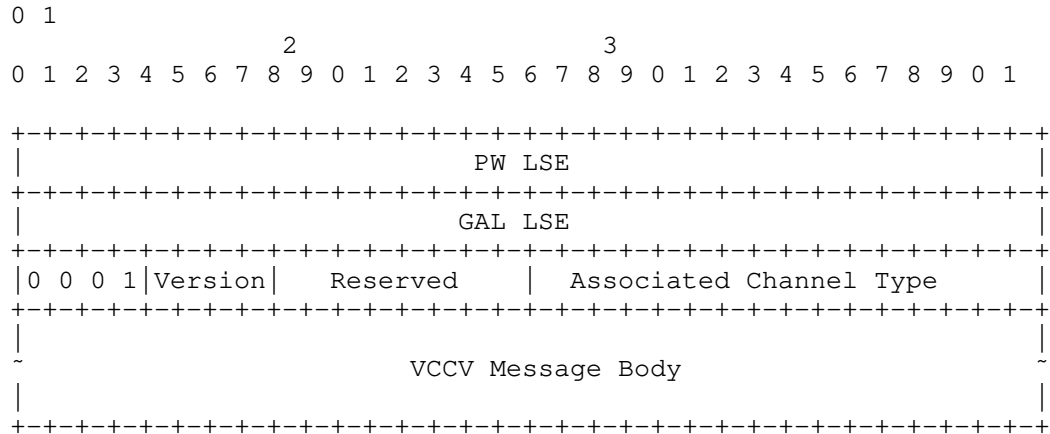
[EDITOR'S NOTE - Why are we talking about 6.3 which is L2TPv3 related in a text on GAL?]

3. VCCV Control Channel When The Control Word is Used

When the PWE3 Control Word is used to encapsulate pseudowire traffic, the rules described for encapsulating VCCV CC Type 1 as specified in section 9.5.1 of [RFC6073] and section 5.1.1 of [RFC5085] MUST be used. In this case the advertised CC Type is 1, and Associated Channel Types of 21, 07, or 57 are allowed.

4. VCCV Control Channel When The Control Word is Not Used

When the PWE3 Control Word is not used a new CC Type 4 is defined as follows:



EDITOR's note = when we wrote RFC3985 I seem to remember that TTL=1 was problematic do we want to specify TTL=1 in the text below?

EDITOR's note = not sure if it should be MUST or SHOULD in the text below.

When the PW is a single segment PW, the TTL field of the PW Label Stack Entry (LSE) SHOULD be set to 1. In the case of multi-segment pseudo-wires, the PW LSE TTL SHOULD be set to the value needed to reach the intended destination PE as described in [RFC6073].

The GAL LSE MUST contain the GAL reserved label as defined in [RFC5586].

As defined in [RFC4385] and [RFC4446] the first nibble of the next field is set to 0001b to indicate an ACH associated with a pseudowire instead of PW data. The Version and the Reserved fields MUST be set to 0, and the Channel Type is set to 0x0021 for IPv4, 0x0057 for IPv6 payloads [RFC5085] or 0x0007 for BFD payloads [RFC5885].

The Associated Channel Type defines how the "VCCV Message Body" field is to be interpreted by the receiver.

5. VCCV Capability Advertisement

The capability advertisement MUST match the c-bit setting that is advertised in the PW FEC element. If the c-bit is set, indicating the use of the control word, type 1 MUST be advertised and type 4 MUST NOT be advertised. If the c-bit is not set, indicating that the control word is not in use, type 4 MUST be advertised, and type 1 MUST NOT be advertised.

A PE supporting Type 4 MAY advertise other CC types as defined in [RFC5085]. If the remote PE also supports Type 4, then Type 4 MUST be used superseding the Capability Advertisement Selection rules of section 7 from [RFC5085]. If a remote PE does not support Type 4, then the rules from section 7 of [RFC5085] apply. If a CW is in use, then Type 4 is not applicable, and therefore the normal capability advertisement selection rules of section 7 from [RFC5085] apply.

6. Manageability Considerations

Editor's note - this is a placeholder - I am not sure if it is needed

7. Security Considerations

This document does not by itself raise any new security considerations beyond those described in [RFC5085].

8. IANA Considerations

8.1. VCCV Interface Parameters Sub-TLV

EDITOR'S NOTE ASFAICS this section can be deleted.

The VCCV Interface Parameters Sub-TLV code point is defined in [RFC4446]. IANA has created and will maintain registries for the CC Types and CV Types (bit masks in the VCCV Parameter ID). The CC Type and CV Type new registries (see Sections 8.1.1 and 8.1.2, respectively of [RFC5085]) have been created in the Pseudo Wires Name Spaces. The allocations must be done using the "IETF Review" policy defined in [RFC5226].

8.2. MPLS VCCV Control Channel (CC) Type 4

IANA is requested to assign a new bit from the MPLS VCCV Control Channel (CC) Types registry in the PWE3-parameters name space in order to identify VCCV type 4. It is recommended that Bit 3 be assigned to this purpose which would have a value of 0x08.

MPLS VCCV Control Channel (CC) Types

Bit (Value) =====	Description =====	Reference =====
Bit X (0x0Y)	Type 4	[This Specification]

9. Acknowledgements

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3931] Lau, J., Townsley, M., and I. Goyret, "Layer Two Tunneling Protocol - Version 3 (L2TPv3)", RFC 3931, March 2005.
- [RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, February 2006.
- [RFC4446] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", BCP 116, RFC 4446, April 2006.
- [RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.
- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", RFC 5586, June 2009.
- [RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.
- [RFC5885] Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, June 2010.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.

10.2. Informative References

- [RFC3916] Xiao, X., McPherson, D., and P. Pate, "Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)", RFC 3916, September 2004.

- [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.
- [RFC4377] Nadeau, T., Morrow, M., Swallow, G., Allan, D., and S. Matsushima, "Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks", RFC 4377, February 2006.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC7079] Del Regno, N. and A. Malis, "The Pseudowire (PW) and Virtual Circuit Connectivity Verification (VCCV) Implementation Survey Results", RFC 7079, November 2013.

Authors' Addresses

Thomas D. Nadeau
lucidvision

Email: tnadeau@lucidvision.com

Luca Martini
Cisco Systems

Email: lmartini@cisco.com

Stewart Bryant
Cisco Systems

Email: stbryant@cisco.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 06, 2014

N. Del Regno, Ed.
A. Malis, Ed.
Verizon Communications Inc
October 03, 2013

The Pseudowire (PW) & Virtual Circuit Connectivity Verification (VCCV)
Implementation Survey Results
draft-ietf-pwe3-vccv-impl-survey-results-03

Abstract

The IETF PWE3 Working Group has defined many encapsulations of various layer 1 and layer 2 service-specific PDUs and circuit data. In most of these encapsulations, use of the Pseudowire (PW) Control Word is required. However, there are several encapsulations for which the Control Word is optional, and this optionality has been seen in practice to possibly introduce interoperability concerns between multiple implementations of those encapsulations. This survey of the PW/VCCV user community was conducted to determine implementation trends and the possibility of always mandating the Control Word.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 06, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. PW/VCCV Survey Overview	4
1.2. PW/VCCV Survey Form	4
1.3. PW/VCCV Survey Highlights	6
2. Survey Results	6
2.1. Summary of Results	6
2.2. Respondents	6
2.3. Pseudowire Encapsulations Implemented	7
2.4. Number of Pseudowires Deployed	8
2.5. VCCV Control Channel In Use	9
2.6. VCCV Connectivity Verification Types In Use	12
2.7. Control Word Support for Encapsulations for which CW is Optional	14
2.8. Open Ended Question	15
3. Security Considerations	16
4. IANA Considerations	16
5. Acknowledgements	16
6. Appendix	16
6.1. Respondent 1	16
6.2. Respondent 2	17
6.3. Respondent 3	19
6.4. Respondent 4	20
6.5. Respondent 5	21
6.6. Respondent 6	22
6.7. Respondent 7	23
6.8. Respondent 8	24
6.9. Respondent 9	25
6.10. Respondent 10	27
6.11. Respondent 11	27
6.12. Respondent 12	29
6.13. Respondent 13	29
6.14. Respondent 14	31
6.15. Respondent 15	32
6.16. Respondent 16	33
6.17. Respondent 17	34
7. Informative References	36
Authors' Addresses	37

1. Introduction

Most pseudowire Emulation Edge-to-Edge (PWE3) encapsulations mandate the use of the Control Word (CW) to carry information essential to the emulation, to inhibit Equal-Cost Multipath (ECMP) behavior, and to discriminate Operations, Administration, and Maintenance (OAM) from Pseudowire (PW) packets. However, some encapsulations treat the Control Word as optional. As a result, implementations of the CW, for encapsulations for which it is optional, vary by equipment manufacturer, equipment model and service provider network. Similarly, Virtual Circuit Connectivity Verification (VCCV) supports three Control Channel (CC) types and multiple Connectivity Verification (CV) Types. This flexibility has led to reports of interoperability issues within deployed networks and associated drafts to attempt to remedy the situation.

The encapsulations and modes for which the Control Word is currently optional are:

- o Ethernet Tagged Mode [RFC4448]
- o Ethernet Raw Mode [RFC4448]
- o PPP [RFC4618]
- o HDLC [RFC4618]
- o Frame Relay Port Mode [RFC4618]
- o ATM (N:1 Cell Mode) [RFC4717]

Virtual Circuit Connectivity Verification (VCCV) [RFC5085] defines three Control Channel types for MPLS PW's: Type 1, using the pseudowire Control Word, Type 2, using the Router Alert (RA) Label, and Type 3, using TTL Expiration (e.g. MPLS PW Label with TTL == 1). While Type 2 (RA Label) is indicated as being "the preferred mode of VCCV operation when the Control Word is not present," RFC 5085 does not indicate a mandatory Control Channel to ensure interoperable implementations. The closest it comes to mandating a control channel is the requirement to support Type 1 (Control Word) whenever the control word is present. As such, the three options yield seven implementation permutations (assuming you have to support at least one Control Channel type to provide VCCV). Due to these permutations, interoperability challenges have been identified by several VCCV users.

In order to assess the best approach to address the observed interoperability issues, the PWE3 working group decided to solicit

feedback from the PW and VCCV user community regarding implementation. This document presents the survey questionnaire and the information returned by the user community who participated.

1.1. PW/VCCV Survey Overview

Per the direction of the PWE3 Working Group chairs, a survey was created to sample the nature of implementations of pseudowires, with specific emphasis on Control Word usage, and VCCV, with emphasis on Control Channel and Control Type usage. The survey consisted of a series of questions based on direction of the WG chairs and the survey opened to the public on November 4, 2010. The survey was conducted using the SurveyMonkey tool, <http://www.surveymonkey.com>. The survey ran from November 4, 2010 until February 25, 2011 and was repeatedly publicized on the PWE3 email list over that period.

The editors took precautions to ensure the validity of the sample and the data. Specifically, only responses with recognizable non-vendor company-affiliated email addresses were accepted. Unrecognizable or personal email addresses would have been contacted to determine their validity, but none were received. Only one response was received from each responding company. If multiple responses from a company had been received, they would have been contacted to determine whether the responses were duplicative or additive. This, however, did not occur.

1.2. PW/VCCV Survey Form

The PW/VCCV Implementation Survey requested the following information about user implementations (the lists of implementation choices were taken verbatim from the survey):

- Responding Organization. No provisions were made for anonymous responses, as all responses required a valid email address in order to validate the survey response. However, the results herein are reported anonymously, except for an alphabetic list of participating organizations in Section 2.2.

- Of the various encapsulations (and options therein) known at the time, including the WG draft for Fiber Channel, draft-ietf-pwe3-fc-encap (now [RFC6307]), which were implemented by the respondent. These included:

- o Ethernet Tagged Mode - RFC 4448
- o Ethernet Raw Mode - RFC 4448
- o SAToP - RFC 4553

- o PPP - RFC 4618
 - o HDLC - RFC 4618
 - o Frame Relay (Port Mode) - RFC 4619
 - o Frame Relay (1:1 Mode) - RFC 4619
 - o ATM (N:1 Mode) - RFC 4717
 - o ATM (1:1 Mode) - RFC 4717
 - o ATM (AAL5 SDU Mode) - RFC 4717
 - o ATM (AAL5 PDU Mode) - RFC 4717
 - o CEP - RFC 4842
 - o CESoPSN - RFC 5086
 - o TDMoIP - RFC 5087
 - o Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307]
- Approximately how many pseudowires of each type were deployed. Respondents could list a number, or for the sake of privacy, could just respond "In-Use" instead.
- For each encapsulation listed above, the respondent could indicate which Control Channel [RFC5085] was in use (see Section 1 for a discussion of these Control Channels). The options listed were:
- o Control Word (Type 1)
 - o Router Alert Label (Type 2)
 - o TTL Expiry (Type 3)
- For each encapsulation listed above, the respondent could indicate which Connectivity Verification types [RFC5085] were in use. The options were:
- o Internet Control Message Protocol (ICMP) Ping
 - o Label Switched Path (LSP) Ping
- For each encapsulation type for which the use of the Control Word is optional, the respondents could indicate the encapsulation for

which Control Word was supported by the equipment used and whether it was in use in the network. The encapsulations listed were:

- o Ethernet (Tagged Mode)
- o Ethernet (Raw Mode)
- o PPP
- o HDLC
- o Frame Relay (Port Mode)
- o ATM (N:1 Cell Mode)

- Finally, a freeform entry was provided for the respondent to provide feedback regarding PW and VCCV deployments, VCCV interoperability challenges, the survey or any network/vendor details they wished to share.

1.3. PW/VCCV Survey Highlights

There were seventeen responses to the survey that met the validity requirements in Section 1.1. The responding companies are listed below in Section 2.2.

2. Survey Results

2.1. Summary of Results

Prior to this survey, there was considerable speculation about whether the Control Word could always be mandated, with several proposals to do so. However, the survey showed that there was considerable deployment of PWs that did not use the the CW. The publication of this survey serves as a reminder of the extent of PWs without the CW in use, and hence a reminder that the CW-less modes cannot be deprecated in the near future.

2.2. Respondents

The following companies, listed here alphabetically as received in the survey responses, participated in the PW/VCCV Implementation Survey. Responses were only solicited from non-vendors (users and service providers), and no vendors responded (although if they had, their response would not have been included). The data provided has been aggregated. No specific company's response will be detailed herein.

- o AboveNet
- o AMS-IX
- o Bright House Networks
- o Cox Communications
- o Deutsche Telekom AG
- o Easynet Global Services
- o France Telecom Orange
- o Internet Solution
- o MTN South Africa
- o OJSC MegaFon
- o Superonline
- o Telecom New Zealand
- o Telstra Corporation
- o Time Warner Cable
- o Tinet
- o Verizon
- o Wipro Technologies

2.3. Pseudowire Encapsulations Implemented

The following question was asked: "In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented." Of all responses, the following list shows the percentage of responses for each encapsulation:

- o Ethernet Tagged Mode - RFC 4448 = 76.5%
- o Ethernet Raw Mode - RFC 4448 = 82.4%
- o SAToP - RFC 4553 = 11.8%
- o PPP - RFC 4618 = 11.8%

- o HDLC - RFC 4618 = 5.9%
- o Frame Relay (Port Mode) - RFC 4619 = 17.6%
- o Frame Relay (1:1 Mode) - RFC 4619 = 41.2%
- o ATM (N:1 Mode) - RFC 4717 = 5.9%
- o ATM (1:1 Mode) - RFC 4717 = 17.6%
- o ATM (AAL5 SDU Mode) - RFC 4717 = 5.9%
- o ATM (AAL5 PDU Mode) - RFC 4717 = 0.0%
- o CEP - RFC 4842 = 0.0%
- o CESoPSN - RFC 5086 = 11.8%
- o TDMoIP - RFC 5087 = 11.8%
- o Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307] = 5.9%

2.4. Number of Pseudowires Deployed

The following question was asked: "Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so." The following list shows the number of pseudowires in use for each encapsulation:

- o Ethernet Tagged Mode = 93,861
- o Ethernet Raw Mode = 94,231
- o SAToP - RFC 4553 = 20,050
- o PPP - RFC 4618 = 500
- o HDLC - RFC 4618 = 0
- o Frame Relay (Port Mode) - RFC 4619 = 5,002
- o Frame Relay (1:1 Mode) - RFC 4619 = 50,959
- o ATM (N:1 Mode) - RFC 4717 = 50,000
- o ATM (1:1 Mode) - RFC 4717 = 70,103

- o ATM (AAL5 SDU Mode) - RFC 4717 = 0
- o ATM (AAL5 PDU Mode) - RFC 4717 = 0
- o CEP - RFC 4842 = 0
- o CESoPSN - RFC 5086 = 21,600
- o TDMoIP - RFC 5087 = 20,000
- o Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307] = 0

In the above responses, on several occasions the response was in the form of "> XXXXX" where the response indicated a number greater than the one provided. Where applicable, the number itself was used in the sums above. For example, ">20K" and "20K+" yielded 20K.

Additionally, the following encapsulations were listed as "In-Use" with no quantity provided:

- o Ethernet Raw Mode: 2 Responses
- o ATM (AAL5 SDU Mode): 1 Response
- o TDMoIP: 1 Response

2.5. VCCV Control Channel In Use

The following instructions were given: "Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply." The numbers below indicate the number of responses. The responses were:

- o Ethernet Tagged Mode - RFC 4448
 - * Control Word (Type 1) = 7
 - * Router Alert Label (Type 2) = 3
 - * TTL Expiry (Type 3) = 3
- o Ethernet Raw Mode - RFC 4448
 - * Control Word (Type 1) = 8
 - * Router Alert Label (Type 2) = 4

- * TTL Expiry (Type 3) = 4
- o SAToP - RFC 4553
 - * Control Word (Type 1) = 1
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o PPP - RFC 4618
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o HDLC - RFC 4618
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o Frame Relay (Port Mode) - RFC 4619
 - * Control Word (Type 1) = 1
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o Frame Relay (1:1 Mode) - RFC 4619
 - * Control Word (Type 1) = 3
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 2
- o ATM (N:1 Mode) - RFC 4717
 - * Control Word (Type 1) = 1
 - * Router Alert Label (Type 2) = 0

- * TTL Expiry (Type 3) = 0
- o ATM (1:1 Mode) - RFC 4717
 - * Control Word (Type 1) = 1
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 1
- o ATM (AAL5 SDU Mode) - RFC 4717
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 1
 - * TTL Expiry (Type 3) = 0
- o ATM (AAL5 PDU Mode) - RFC 4717
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o CEP - RFC 4842
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0
- o CESoPSN - RFC 5086
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 1
- o TDMoIP - RFC 5087
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0

- * TTL Expiry (Type 3) = 0
- o Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307]
 - * Control Word (Type 1) = 0
 - * Router Alert Label (Type 2) = 0
 - * TTL Expiry (Type 3) = 0

2.6. VCCV Connectivity Verification Types In Use

The following instructions were given: "Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type." Note that BFD was not one of the choices. The responses were as follows:

- o Ethernet Tagged Mode - RFC 4448
 - * ICMP Ping = 5
 - * LSP Ping = 11
- o Ethernet Raw Mode - RFC 4448
 - * ICMP Ping = 6
 - * LSP Ping = 11
- o SAToP - RFC 4553
 - * ICMP Ping = 0
 - * LSP Ping = 2
- o PPP - RFC 4618
 - * ICMP Ping = 0
 - * LSP Ping = 0
- o HDLC - RFC 4618
 - * ICMP Ping = 0
 - * LSP Ping = 0
- o Frame Relay (Port Mode) - RFC 4619

- * ICMP Ping = 0
- * LSP Ping = 1
- o Frame Relay (1:1 Mode) - RFC 4619
 - * ICMP Ping = 2
 - * LSP Ping = 5
- o ATM (N:1 Mode) - RFC 4717
 - * ICMP Ping = 0
 - * LSP Ping = 1
- o ATM (1:1 Mode) - RFC 4717
 - * ICMP Ping = 0
 - * LSP Ping = 3
- o ATM (AAL5 SDU Mode) - RFC 4717
 - * ICMP Ping = 0
 - * LSP Ping = 1
- o ATM (AAL5 PDU Mode) - RFC 4717
 - * ICMP Ping = 0
 - * LSP Ping = 0
- o CEP - RFC 4842
 - * ICMP Ping = 0
 - * LSP Ping = 0
- o CESoPSN - RFC 5086
 - * ICMP Ping = 0
 - * LSP Ping = 1
- o TDMoIP - RFC 5087

- * ICMP Ping = 0
- * LSP Ping = 1
- o Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307]
 - * ICMP Ping = 0
 - * LSP Ping = 0

2.7. Control Word Support for Encapsulations for which CW is Optional

The following instructions were given: "Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional." The responses were:

- o Ethernet (Tagged Mode)
 - * Supported by Network/Equipment = 13
 - * Used in Network = 6
- o Ethernet (Raw Mode)
 - * Supported by Network/Equipment = 14
 - * Used in Network = 7
- o PPP
 - * Supported by Network/Equipment = 5
 - * Used in Network = 0
- o HDLC
 - * Supported by Network/Equipment = 4
 - * Used in Network = 0
- o Frame Relay (Port Mode)
 - * Supported by Network/Equipment = 3
 - * Used in Network = 1
- o ATM (N:1 Cell Mode)

* Supported by Network/Equipment = 5

* Used in Network = 1

2.8. Open Ended Question

Space was provided for user feedback. The following instructions were given: "Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share." Below are the responses, made anonymous. The responses are otherwise provided here verbatim.

1. BFD VCCV Control Channel is not indicated in the survey (may be required for PW redundancy purpose)
2. Using CV is not required at the moment
3. COMPANY has deployed several MPLS network elements, from multiple vendors. COMPANY is seeking a uniform implementation of VCCV Control Channel (CC) capabilities across its various vendor platforms. This will provide COMPANY with significant advantages in reduced operational overheads when handling cross-domain faults. Having a uniform VCCV feature implementation in COMPANY multi-vendor network leads to:
 - o Reduced operational cost and complexity
 - o Reduced OSS development to coordinate incompatible VCCV implementations.
 - o Increased end-end service availability when handling faults.In addition, currently some of COMPANY deployed VCCV traffic flows (on some vendor platforms) are not guaranteed to follow those of the customer's application traffic (a key operational requirement). As a result, the response from the circuit ping cannot faithfully reflect the status of the circuit. This leads to ambiguity regarding the operational status of our networks. An in-band method is highly preferred, with COMPANY having a clear preference for VCCV Circuit Ping using PWE Control Word. This preference is being pursued with each of COMPANY vendors.
4. PW VCCV is very useful tool for finding faults in each PW channel. Without this we can not find fault on a PW channel. PW VCCV using BFD is another better option. Interoperability challenges are with Ethernet OAM mechanism.
5. We are using L2PVPN ATOM like-to-like models - ATOMPLS - EoMPLS ATOMPLS : This service offered for transporting ATM cells over IP/MPLS core with Edge ATM CE devices including BPX, Ericsson Media Gateway etc. This is purely a Port mode with cell-packing configuration on it to have best performance. QoS marking is

done for getting LLQ treatment in the core for these MPLS encapsulated ATM packets. EoMPLS: This service offered for transporting 2G/3G traffic from network such as Node-B to RNC's over IP/MPLS backbone core network. QoS marking is done for getting guaranteed bandwidth treatment in the core for these MPLS encapsulated ATM packets. In addition to basic L2VPN service configuration, these traffic are routed via MPLS TE tunnels with dedicated path and bandwidth defined to avoid bandwidth related congestion.

6. EQUIPMENT MANUFACTURER does not provide options to configure VCCV control-channel and its sub options for LDP based L2Circuits. How can we achieve end-to-end management and fault detection of PW without VCCV in such cases?
7. I'm very interested in this work as we continue to experience interop challenges particularly with newer vendors to the space who are only implementing VCCV via control word. Vendors who have tailed their MPLS OAM set specifically to the cell backhaul space and mandatory CW have been known to fall into this space. That's all I've got.

3. Security Considerations

As this document is an informational report of the PW/VCCV User Implementation Survey results, no protocol security considerations are introduced.

4. IANA Considerations

This document has no actions for IANA.

5. Acknowledgements

We would like to thank the chairs of the PWE3 Working Group for their guidance and review of the Survey questions. We would also like to sincerely thank those listed in Section 2.2. who took the time and effort to participate.

6. Appendix

The detailed responses are included in this appendix. The respondent contact info has been removed.

6.1. Respondent 1

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 423

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.2. Respondent 2

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

SAToP - RFC 4553

CESoPSN - RFC 5086

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 5000

Ethernet Raw Mode - RFC 4448 - 1000

SAToP - RFC 4553 - 50

CESoPSN - RFC 5086 - 1600

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2), TTL Expiry (Type 3)

Ethernet Raw Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2), TTL Expiry (Type 3)

CESoPSN - RFC 5086: TTL Expiry (Type 3)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: ICMP Ping, LSP Ping

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

SAToP - RFC 4553: LSP Ping

CESoPSN - RFC 5086: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

I'm very interested in this work as we continue to experience interop challenges particularly with newer vendors to the space who are only implementing VCCV via control word. Vendors who have tailed their MPLS OAM set specifically to the cell backhaul space and mandatory CW have been known to fall into this space. That's all I've got.

6.3. Respondent 3

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

Frame Relay (Port Mode) - RFC 4619

Frame Relay (1:1 Mode) - RFC 4619

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 800

Ethernet Raw Mode - RFC 4448 - 50

Frame Relay (Port Mode) - RFC 4619 - 2

Frame Relay (1:1 Mode) - RFC 4619 - 2

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

No Response

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

No Response

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.4. Respondent 4

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 1000

Ethernet Raw Mode - RFC 4448 - 200

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

No Response

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

Ethernet Raw Mode - RFC 4448: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

EQUIPMENT MANUFACTURER does not provide options to configure VCCV control-channel and its sub options for LDP based L2Circuits. How can we achieve end-to-end management and fault detection of PW without VCCV in such cases?

6.5. Respondent 5

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

PPP - RFC 4618

Frame Relay (Port Mode) - RFC 4619

Frame Relay (1:1 Mode) - RFC 4619

Fiber Channel (Port Mode) - draft-ietf-pwe3-fc-encap [RFC6307]

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 4000

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2)

Ethernet Raw Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.6. Respondent 6

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 1000+

Ethernet Raw Mode - RFC 4448 - 500

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1)

Ethernet Raw Mode - RFC 4448: Control Word (Type 1)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: ICMP Ping, LSP Ping

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.7. Respondent 7

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Raw Mode - RFC 4448

ATM (1:1 Mode) - RFC 4717

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Raw Mode - RFC 4448 - 20

ATM (1:1 Mode) - RFC 4717 - 100

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

No Response

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Raw Mode - RFC 4448: LSP Ping

ATM (1:1 Mode) - RFC 4717: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode), PPP, HDLC, Frame Relay (Port Mode), ATM (N:1 Cell Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

We are using L2PVPN AToM like-to-like models - ATMoMPLS - EoMPLS
ATMoMPLS : This service offered for transporting ATM cells over IP/MPLS core with Edge ATM CE devices including BPX, Ericsson Media Gateway etc. This is purely a Port mode with cell-packing configuration on it to have best performance. QoS marking is done for getting LLQ treatment in the core for these MPLS encapsulated ATM packets. EoMPLS: This service offered for transporting 2G/3G traffic from network such as Node-B to RNC's over IP/MPLS backbone core network. QoS marking is done for getting guaranteed bandwidth treatment in the core for these MPLS encapsulated ATM packets. In addition to basic L2VPN service configuration, these traffic are routed via MPLS TE tunnels with dedicated path and bandwidth defined to avoid bandwidth related congestion.

6.8. Respondent 8

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Raw Mode - RFC 4448

ATM (AAL5 SDU Mode) - RFC 4717

TDMoIP - RFC 5087

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires

in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Raw Mode - RFC 4448 - In-Use

ATM (AAL5 SDU Mode) - RFC 4717 - In-Use

TDMoIP - RFC 5087 - In-Use

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Raw Mode - RFC 4448: Control Word (Type 1)

ATM (AAL5 SDU Mode) - RFC 4717: Router Alert Label (Type 2)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Raw Mode - RFC 4448: LSP Ping

ATM (AAL5 SDU Mode) - RFC 4717: LSP Ping

TDMoIP - RFC 5087: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Raw Mode), ATM (N:1 Cell Mode)

Used in Network: Ethernet (Raw Mode), ATM (N:1 Cell Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

PW VCCV is very useful tool for finding faults in each PW channel. Without this we can not find fault on a PW channel. PW VCCV using BFD is another better option. Interoperability challenges are with Ethernet OAM mechanism.

6.9. Respondent 9

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Frame Relay (1:1 Mode) - RFC 4619

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 19385

Frame Relay (1:1 Mode) - RFC 4619 - 15757

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Frame Relay (1:1 Mode) - RFC 4619: Control Word (Type 1)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Frame Relay (1:1 Mode) - RFC 4619: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode), PPP, HDLC, Frame Relay (Port Mode), ATM (N:1 Cell Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.10. Respondent 10

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Raw Mode - RFC 4448 - 325

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Raw Mode - RFC 4448: Control Word (Type 1)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: No Response

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.11. Respondent 11

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

PPP - RFC 4618 HDLC - RFC 4618

Frame Relay (1:1 Mode) - RFC 4619

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 2000

Ethernet Raw Mode - RFC 4448 - 100

PPP - RFC 4618 - 500

Frame Relay (1:1 Mode) - RFC 4619 - 200

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

No Response

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: ICMP Ping, LSP Ping

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

Frame Relay (1:1 Mode) - RFC 4619: ICMP Ping, LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode), PPP, HDLC

Used in Network: Ethernet (Tagged Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.12. Respondent 12

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Raw Mode - RFC 4448 - 50000

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Raw Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2), TTL Expiry (Type 3)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

No Response

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.13. Respondent 13

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

Frame Relay (1:1 Mode) - RFC 4619

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 3

Ethernet Raw Mode - RFC 4448 - 10-20

ATM (1:1 Mode) - RFC 4717 - 3

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1), TTL Expiry (Type 3)

Ethernet Raw Mode - RFC 4448: Control Word (Type 1), TTL Expiry (Type 3)

Frame Relay (1:1 Mode) - RFC 4619: Control Word (Type 1), TTL Expiry (Type 3)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: ICMP Ping, LSP Ping

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

Frame Relay (1:1 Mode) - RFC 4619: ICMP Ping, LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode), PPP, HDLC, Frame Relay (Port Mode), ATM (N:1 Cell Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode), Frame Relay (Port Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.14. Respondent 14

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 150

Ethernet Raw Mode - RFC 4448 - 100

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2)

Ethernet Raw Mode - RFC 4448: Control Word (Type 1), Router Alert Label (Type 2)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

Ethernet Raw Mode - RFC 4448: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode), PPP, HDLC, Frame Relay (Port Mode)

Used in Network: Ethernet (Tagged Mode), Ethernet (Raw Mode)

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

No Response

6.15. Respondent 15

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

Frame Relay (1:1 Mode) - RFC 4619

ATM (1:1 Mode) - RFC 4717

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 20,000

Ethernet Raw Mode - RFC 4448 - 1000

Frame Relay (1:1 Mode) - RFC 4619 - 30,000

ATM (1:1 Mode) - RFC 4717 - 20,000

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: TTL Expiry (Type 3)

Ethernet Raw Mode - RFC 4448: TTL Expiry (Type 3)

Frame Relay (1:1 Mode) - RFC 4619: TTL Expiry (Type 3)

ATM (1:1 Mode) - RFC 4717: TTL Expiry (Type 3)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

Ethernet Raw Mode - RFC 4448: LSP Ping

Frame Relay (1:1 Mode) - RFC 4619: LSP Ping

ATM (1:1 Mode) - RFC 4717: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: No Response

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

COMPANY has deployed several MPLS network elements, from multiple vendors. COMPANY is seeking a uniform implementation of VCCV Control Channel (CC) capabilities across its various vendor platforms. This will provide COMPANY with significant advantages in reduced operational overheads when handling cross-domain faults. Having a uniform VCCV feature implementation in COMPANY multi-vendor network leads to:

- o Reduced operational cost and complexity
- o Reduced OSS development to coordinate incompatible VCCV implementations.
- o Increased end-end service availability when handling faults.

In addition, currently some of COMPANY deployed VCCV traffic flows (on some vendor platforms) are not guaranteed to follow those of the customer's application traffic (a key operational requirement). As a result, the response from the circuit ping cannot faithfully reflect the status of the circuit. This leads to ambiguity regarding the operational status of our networks. An in-band method is highly preferred, with COMPANY having a clear preference for VCCV Circuit Ping using PWE Control Word. This preference is being pursued with each of COMPANY vendors.

6.16. Respondent 16

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

Ethernet Raw Mode - RFC 4448

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - 100

Ethernet Raw Mode - RFC 4448 - 100

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

No Response

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: ICMP Ping, LSP Ping

Ethernet Raw Mode - RFC 4448: ICMP Ping, LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: Ethernet (Tagged Mode), Ethernet (Raw Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

Using CV is not required at the moment

6.17. Respondent 17

2. In your network in general, across all products, please indicate which pseudowire encapsulations your company has implemented.

Ethernet Tagged Mode - RFC 4448

SAToP - RFC 4553

Frame Relay (Port Mode) - RFC 4619

Frame Relay (1:1 Mode) - RFC 4619

ATM (N:1 Mode) - RFC 4717

ATM (1:1 Mode) - RFC 4717

CESoPSN - RFC 5086

TDMoIP - RFC 5087

3. Approximately how many pseudowires are deployed of each encapsulation type. Note, this should be the number of pseudowires in service, carrying traffic, or pre-positioned to do so. ***Note, please indicate "In-Use" for any PW Encap Types which you are using but cannot provide a number.

Ethernet Tagged Mode - RFC 4448 - >40k

Ethernet Raw Mode - RFC 4448 - In-Use

SAToP - RFC 4553 - >20k

Frame Relay (Port Mode) - RFC 4619 - >5k

Frame Relay (1:1 Mode) - RFC 4619 - >5k

ATM (N:1 Mode) - RFC 4717 - >50k

ATM (1:1 Mode) - RFC 4717 - >50k

CESoPSN - RFC 5086 - >20k

TDMoIP - RFC 5087 - >20k

4. Please indicate which VCCV Control Channel is used for each encapsulation type. Understanding that users may have different networks with varying implementations, for your network in general, please select all which apply.

Ethernet Tagged Mode - RFC 4448: Control Word (Type 1)

SAToP - RFC 4553: Control Word (Type 1)

Frame Relay (Port Mode) - RFC 4619: Control Word (Type 1)

Frame Relay (1:1 Mode) - RFC 4619: Control Word (Type 1)

ATM (N:1 Mode) - RFC 4717: Control Word (Type 1)

ATM (1:1 Mode) - RFC 4717: Control Word (Type 1)

5. Please indicate which VCCV Connectivity Verification types are used in your networks for each encapsulation type.

Ethernet Tagged Mode - RFC 4448: LSP Ping

SAToP - RFC 4553: LSP Ping

Frame Relay (Port Mode) - RFC 4619: LSP Ping

Frame Relay (1:1 Mode) - RFC 4619: LSP Ping

ATM (N:1 Mode) - RFC 4717: LSP Ping

ATM (1:1 Mode) - RFC 4717: LSP Ping

6. Please indicate your network's support of and use of the Control Word for encapsulations for which the Control Word is optional.

Supported by Network/Equipment: ATM (N:1 Cell Mode)

Used in Network: No Response

7. Please use this space to provide any feedback regarding PW and VCCV deployments, VCCV interoperability challenges, this survey or any network/vendor details you wish to share.

BFD VCCV Control Channel is not indicated in the survey (may be required for PW redundancy purpose)

7. Informative References

- [RFC4448] Martini, L., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.
- [RFC4618] Martini, L., Rosen, E., Heron, G., and A. Malis, "Encapsulation Methods for Transport of PPP/High-Level Data Link Control (HDLC) over MPLS Networks", RFC 4618, September 2006.
- [RFC4717] Martini, L., Jayakumar, J., Bocci, M., El-Aawar, N., Brayley, J., and G. Koleyani, "Encapsulation Methods for Transport of Asynchronous Transfer Mode (ATM) over MPLS Networks", RFC 4717, December 2006.

- [RFC5085] Nadeau, T., Ed. and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", December 2007.
- [RFC6307] Black, D., Dunbar, L., Roth, M., and R. Solomon, "Encapsulation Methods for Transport of Fibre Channel Traffic over MPLS Networks", RFC 6307, April 2012.

Authors' Addresses

Christopher N. "Nick" Del Regno (editor)
Verizon Communications Inc
400 International Pkwy
Richardson, TX 75081
US

Email: nick.delregno@verizon.com

Andrew G. Malis (editor)
Verizon Communications Inc
60 Sylvan Road
Waltham, MA 02451
US

Email: andrew.g.malis@verizon.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 12, 2013

L. Jin
R. Chen
ZTE
S. Boutros
Cisco Systems
S. Kini
Ericsson
October 9, 2012

Static pseudowire configuration checking using Generic Associated
Channel (G-ACh) Advertisement Protocol
draft-jc-pwe3-static-config-check-01.txt

Abstract

This document defines a method to verify the configuration parameters of static pseudowires (PW). Since a static PW can be independently provisioned at each end of the PW there is a potential for a configuration parameter mismatch and this can result in the PW not being operational. This document introduces a configuration checking protocol to simplify the provisioning and ease trouble shooting.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 12, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Terminology	3
3. GAP Extensions	4
3.1. Static PW Application Message	4
3.2. PE Procedure for SS-PW	7
3.2.1. Sending PW application Element TLV	7
3.2.2. Receiving PW application Element TLV	7
3.2.3. PW Configuration Verification Process	8
3.2.4. Remote Label Advertisement	8
3.3. PE Procedure for MS-PW	8
4. Security Considerations	9
5. IANA Considerations	9
6. Acknowledgements	9
7. References	9
7.1. Normative references	9
7.2. Informative References	10
Authors' Addresses	10

1. Introduction

The manual configuration of static PW in MPLS and MPLS-TP network requires configuring different PW parameters at the two terminating PEs (Provider Edge). The PW parameters include PW-id, PW-Type, Control word setting, interface and VCCV parameters settings.

The PW provisioned parameters MUST be aligned, so as to make the PW operational. For dynamically signaled PW, the PW parameters are negotiated using the signaling protocol, and only when the PW parameters match at the terminating PE end points, the P2P (Point-to-Point) PW is made operational and can be used to forward data traffic.

In the absence of a signaling protocol, this draft defines a method to do static PW configuration verification, so as to ease the troubleshooting of end to end static PW provisioning in both MPLS and MPLS-TP networks. The mechanism to exchange the PW configuration parameters uses the Generic Associated Channel (G-ACH) Advertisement Protocol (GAP) defined in [I-D.ietf-mpls-gach-adv]. In this draft, the GAP functionality assumes that the PW's underlying PSN Tunnel with GAP enabled is operational.

In the following sections we will describe the extension to the GAP mechanism to do the PW configuration verification at the two terminating PEs for P2P PW. The P2MP (Point-to-Multipoint) PW configuration verification is for further study.

2. Terminology

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 [RFC2119].

This document uses some terms and acronyms as follows:

MPLS: Multi Protocol Label Switching.

OAM: MPLS Operations, Administration and Maintenance.

PE: Provide Edge Node.

T-PE: PW Terminating Provider Edge.

S-PE: PW Switching Provider Edge.

PW: PseudoWire.

TLV: Type, Length, and Value.

SS-PW: Single-segment PseudoWire

MS-PW: Multi-segment PseudoWire

3. GAP Extensions

3.1. Static PW Application Message

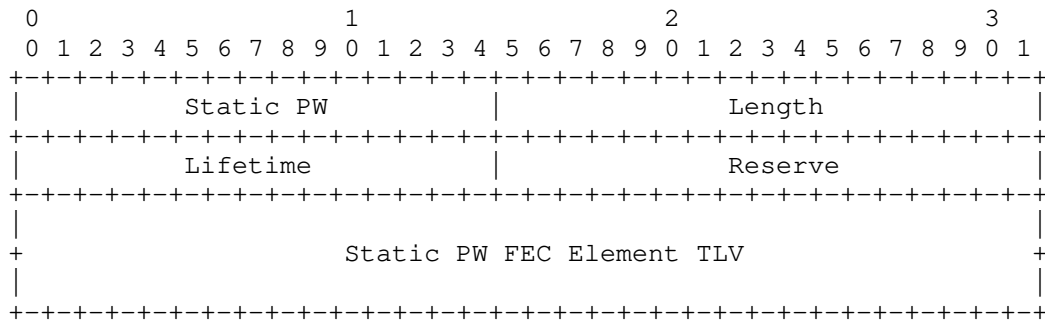


Figure 1

A new GAP application "Static PW" is defined in this draft. The Static PW Application ID is to be assigned by IANA, and suggested value is 0x0002.

Length: as per [I-D.ietf-mppls-gach-adv].

Lifetime: as per [I-D.ietf-mppls-gach-adv], and the default value is suggested to be 120 seconds.

Static PW FEC Element TLV for "Static PW" GAP application:

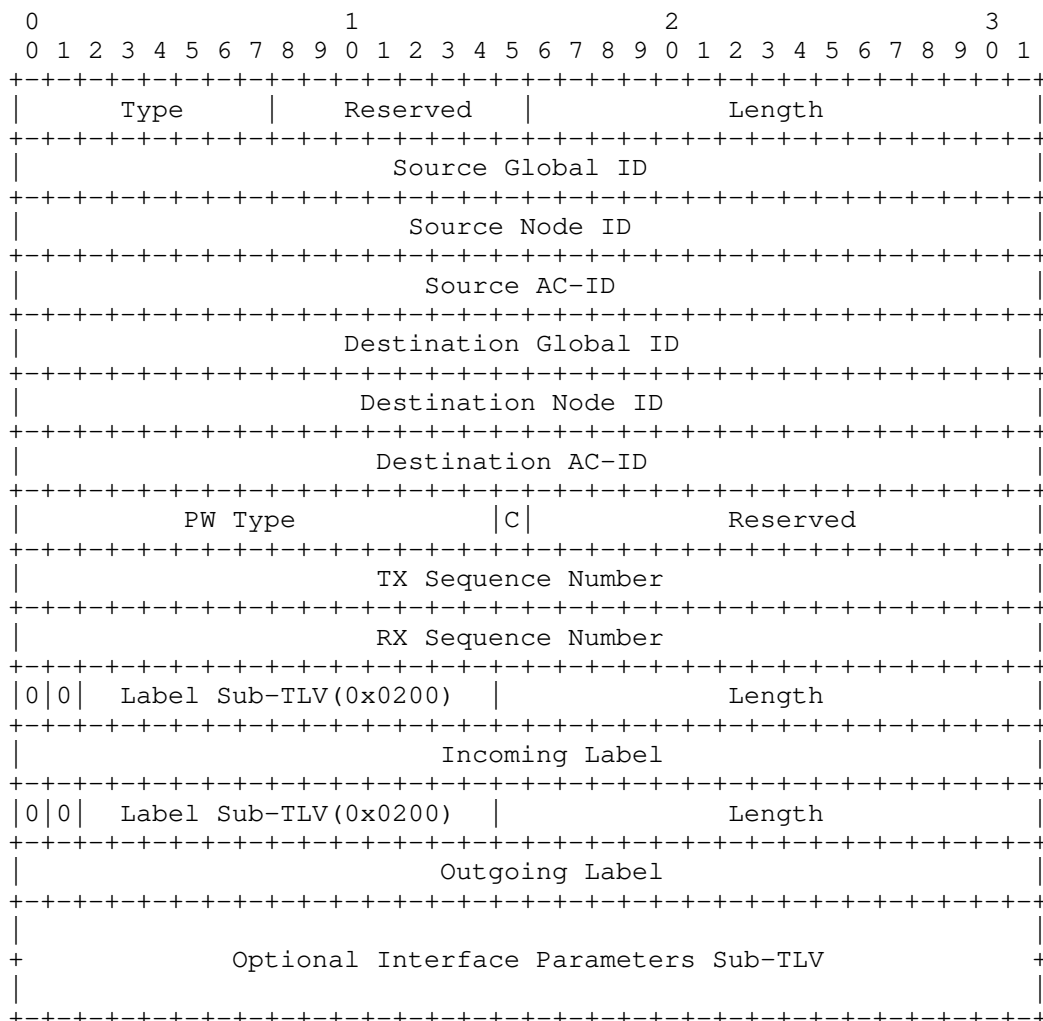


Figure 2

The Static PW FEC Element TLV type is to be assigned by IANA. The Length field specifies the length in octets of the Static PW FEC Element and all Optional Interface Parameters Sub-TLVs.

The Static PW FEC element TLV value MUST include the following:

- o The Global ID and Node ID fields MUST be set as per [RFC6370].

- o The AC-ID fields MUST be set as per [RFC5003].
- o PW-Type and control word bit (C) MUST be set as per [RFC4447].
- o TX Sequence Number: The transmitted message sequence number for the associated Static PW FEC Element TLV.
- o RX Sequence Number: The last received sequence number for the associated Static PW FEC Element TLV.
- o Two Generic Label TLVs as defined in [RFC5036] to encode static PW incoming and outgoing labels in the order shown above.
- o Optional Interface parameters Sub-TLV as defined in [RFC4447].

The GAP Suppress message defined in [I-D.ietf-mppls-gach-adv] only applies all TLVs for a given application. We define a new TLV, static PW suppress TLV, to suppress static PW FEC element transmission. Multiple static PW FEC element TLVs could be included in this TLV. The format would be as follows:

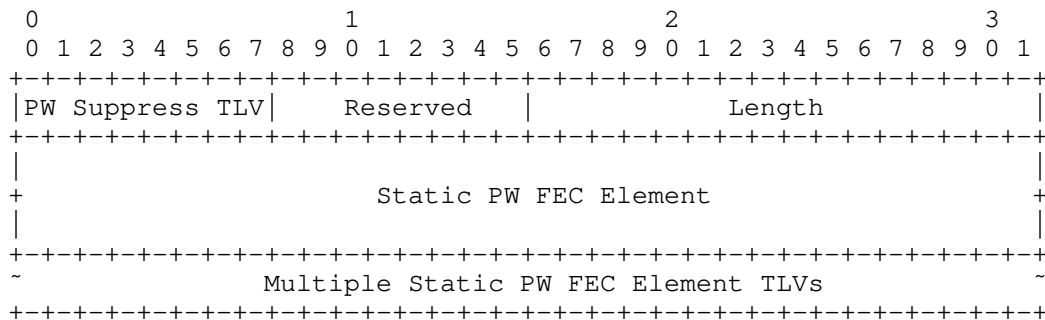


Figure 3

The type of static PW suppress TLV is to be assigned by IANA.

The static PW suppress TLV could be sent by a receiving PE to request a transmitting PE to stop sending GAP messages for the static PW FEC Element TLVs in the static PW suppress TLV.

The static PW application MUST follow all procedures defined in [I-D.ietf-mppls-gach-adv].

3.2. PE Procedure for SS-PW

The mechanism defined in this draft provides a verification tool for the P2P PW configuration information between two PEs. Upon the provisioning or re-provisioning of a PW at an endpoint PE, GAP messages carrying the static PW application TLV will be sent over the PW's corresponding PSN tunnel which the endpoints PEs of the P2P PW selects by local policy.

3.2.1. Sending PW application Element TLV

When a PW is configured at one endpoint PE, and the PW corresponding PSN Tunnel is operational and UP, the PE MUST send its local PW configuration information using the GAP over the PSN tunnel.

The transmitting PE MUST set the TX sequence number to a non-zero value in Static PW FEC Element TLV, and MUST increment the TX sequence number each time any local PW parameters change.

If the transmitting PE has previously received a GAP message with the static PW FEC Element, the transmitting PE MUST verify local PW parameters with the remote PE parameters as specified in section 4.2.3. The RX sequence number MUST be set to the previously received TX sequence number, otherwise set to zero.

3.2.2. Receiving PW application Element TLV

The receiving PE MUST update the remote PW parameters associated with a static PW FEC Element TLV, when the received TX sequence number in the GAP message is different from the last one received.

If the receiving PE has been provisioned locally with the PW parameters and has previously sent GAP message for the PW parameters, it MUST check if the RX sequence number in the received GAP message is equal to the TX sequence number it previously sent.

If the RX sequence number is equal, the receiving PE MUST send GAP message with static PW suppress TLV as a response to remote PE, and then verify local static PW parameters with the remote static PW FEC parameters as specified in section 3.2.3.

Otherwise, if the RX sequence number is not equal, the receiving PE MUST continue sending GAP message with static PW FEC element TLV, with the RX sequence number set to the last received TX sequence number from the remote PE.

If there is no local PW configuration associated with the static PW FEC Element TLV, the receiving PE MUST retain the remote static PW

FEC Element information.

Whenever PE receives the GAP message with static PW suppress TLV, it MUST stop sending GAP messages with the specified static PW FEC element TLVs included in the static suppress TLV.

The GAP message of static PW application SHOULD be sent at least three times within lifetime.

The mechanism described above applies as well for MS-PW.

3.2.3. PW Configuration Verification Process

Using source/destination Global-IDs, and source/destination node-ID and AC-IDs, to identify a locally provisioned static PW, once found, perform the following parameter verification checks:

1. Check the control word bit (C), and MUST do logical operation "AND". Only when both ends have the use of control word enabled, the result would be with control word presented on this PW.
2. Check PW type mismatch as defined in [RFC4447].
3. Check and negotiate interface parameters as defined in [RFC4447].
4. Check incoming and outgoing static PW labels. The local incoming label should be equal to remote outgoing label, and the local outgoing label should be equal to remote incoming label, otherwise checking failed.

3.2.4. Remote Label Advertisement

The mechanism described in this draft MAY also be used to communicate local static PW labels to allow for single side provisioning of labels. As such, only incoming label will be included in the GAP message and this label will be used by the remote PE as the output label for the PW.

3.3. PE Procedure for MS-PW

The mechanism described above for verifying the SS-PW configuration applies for MS-PW. As described in section 3.2, an S-PE MUST verify the incoming and outgoing static PW labels, however no other PW configuration parameters checking are needed at S-PE, since only the labels will be configured at S-PE.

An S-PE MUST pass through static PW application TLVs carried in GAP messages, from one PW segment's corresponding PSN tunnel to the other

PW segment's corresponding PSN tunnel.

4. Security Considerations

The mechanisms defined in this draft do not introduce any new threats more than what's described in [I-D.ietf-mpls-gach-adv].

5. IANA Considerations

IANA is requested to allocate a new "Static PW" Application ID in the "G-Ach Advertisement Protocol Applications" registry.

Application ID	Description	Reference
(TBD)	Static PW Application	(this draft)

This document requests that IANA create a new registry, "GAP Static PW Application: TLV objects", with fields and initial value as follows:

Type Name	Type ID	Reference
Static PW FEC Element	0	(this draft)
Static PW suppress TLV	1	(this draft)

The range of the Type ID field is 0 - 255.

The allocation policy for this registry is IETF Review.

6. Acknowledgements

The authors would like to thank Stewart Bryant, Dan Frost for their review and contributions.

7. References

7.1. Normative references

- [I-D.ietf-mpls-gach-adv]
Frost, D., Bryant, S., and M. Bocci, "MPLS Generic Associated Channel (G-ACh) Advertisement Protocol", draft-ietf-mpls-gach-adv-02 (work in progress), May 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

7.2. Informative References

- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.
- [RFC5003] Metz, C., Martini, L., Balus, F., and J. Sugimoto, "Attachment Individual Identifier (AII) Types for Aggregation", RFC 5003, September 2007.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", RFC 5036, October 2007.
- [RFC6370] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport Profile (MPLS-TP) Identifiers", RFC 6370, September 2011.

Authors' Addresses

Lizhong Jin
ZTE Corporation
889, Bibo Road
Shanghai, 201203, China

Email: lizhong.jin@zte.com.cn

Ran Chen
ZTE Corporation
No.19 East Huayuan Road
Beijing, 100191, China

Email: chen.ran@zte.com.cn

Sami Boutros
Cisco Systems, Inc.
3750 Cisco Way
San Jose, California 95134
USA

Email: sboutros@cisco.com

Sriganesh Kini
Ericsson
Ericsson
San Jose, CA 95134

Email: sriganesh.kini@ericsson.com

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: April 25, 2013

Mingui Zhang
Huafeng Wen
Huawei
October 22, 2012

STP Application of ICCP
draft-zhang-iccp-stp-01.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.

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/lid-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) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal

Provisions Relating to IETF Documents
(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Conventions used in this document	3
1.2. Terminology	3
2. The Use Case Scenario	3
2.1. Virtual Root Bridge	4
3. Spanning Tree Protocol Application TLVs	4
3.1. STP Connect TLV	4
3.2. STP Disconnect TLV	5
3.2.1. STP Disconnect Cause TLV	6
3.3. STP System Config TLV	7
3.3.1. MAC of the Root Bridge	7
3.3.2. STP Topology Changed Instances	8
3.3.3. STP CIST Root Time	8
3.3.4. STP MSTI Root Time	9
3.3.5. STP Region Name	10
3.3.6. STP Revision Level	11
3.3.7. STP Instance Priority	11
3.3.8. STP Configuration Digest	12
3.4. STP Synchronization Request TLV	13
3.5. STP Synchronization Data TLV	14
4. Security Considerations	15
5. IANA Considerations	15
6. References	15
6.1. Normative References	15
6.2. Informative References	15
Author's Addresses	17

1. Introduction

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

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

1.1. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. Terminology

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

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

2. The Use Case Scenario

It is a common case that an RG is connected to a bridge network where STP is running. For example, geographically dispersed DSLAMs of a Broadband Network may be connected by an RG. These DSLAMs constitute a typical STP network. For the sake of network resilience, it is reasonable to connect each RG member to this bridge network. The scenario in Figure 2.1 illustrates this kind of connection.

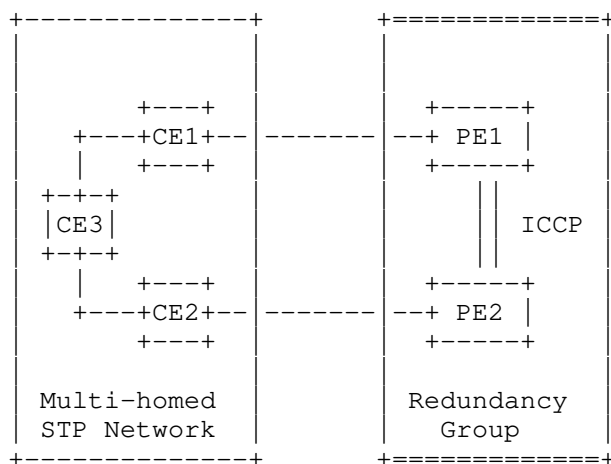


Figure 2.1: A STP network is multi-homed to an Redundant Group.

2.1. Virtual Root Bridge

With ICCP, the whole RG will be virtualized to be a single bridge. The RG pretends that the ports connected to the STP network are from the same bridge. All these ports emit configuration BPDU with the highest root priority to trigger the construction of the spanning tree. In this way, the STP will always broken a loop within the multi-homed STP network.

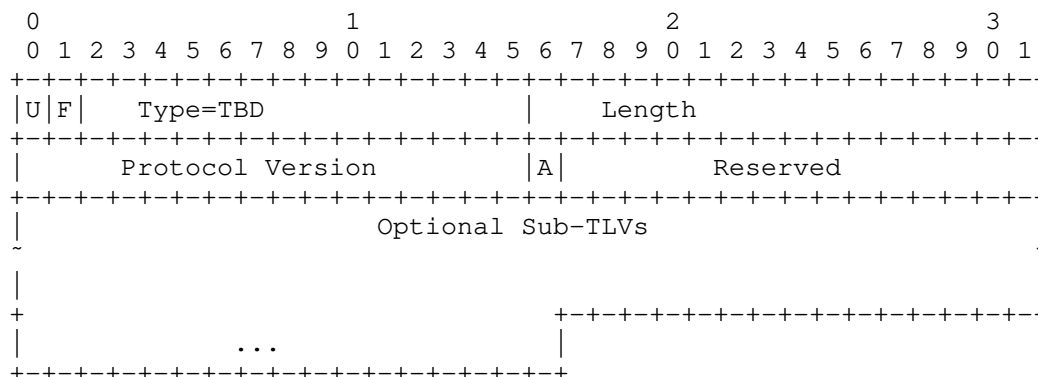
Each RG member has its BridgeIdentifier (the MAC address). The least significant one is elected as the BridgeIdentifier of the 'virtualized root bridge'.

3. Spanning Tree Protocol Application TLVs

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

3.1. STP Connect TLV

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Connect TLV"

- Length

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

- Protocol Version

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

- A bit

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

- Reserved

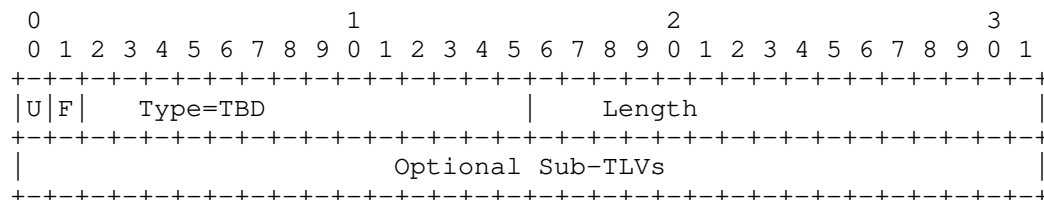
Reserved for future use.

- Optional Sub-TLVs

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

3.2. STP Disconnect TLV

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Disconnect TLV"

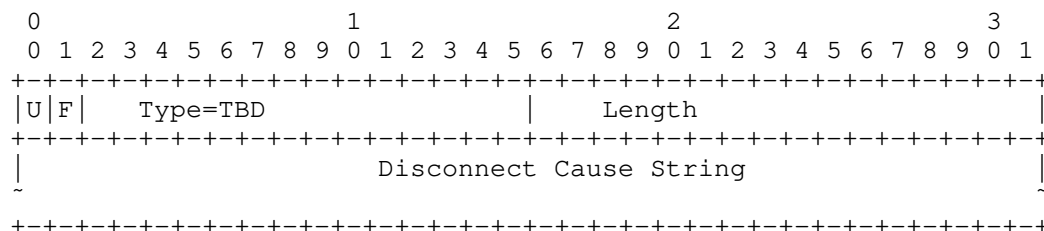
- Length

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

- Optional Sub-TLVs

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

3.2.1. STP Disconnect Cause TLV



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Disconnect Cause TLV"

- Length

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

- Disconnect Cause String

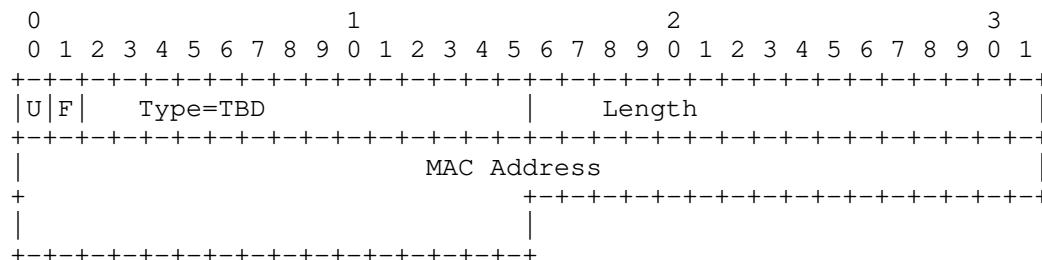
Variable length string specifying the reason for the disconnect. Used for network management.

3.3. STP System Config TLV

The STP System Config TLV is sent in the RG Application Data message. This TLV announces the local node's STP System Parameters to the RG peers. When a this TLV is received by a peering RB member, it SHOULD synchronize the configuration information contained in the TLV. TLVs specified from section 3.3.3 through section 3.3.8 contains such kind of configuration information.

3.3.1. MAC of the Root Bridge

This TLV is used to report the MAC address to be used as the MAC address of the root bridge to other members in the RG. In this document, this MAC address is set to the BridgeIdentifier of the sender, as defined in [802.1q] section 13.23.2. The RG member with the least significant unsigned BridgeIdentifier is elected as the root bridge.



- U and F Bits

Both are set to 0.

- Type

set to TBD for "MAC of the Root Bridge"

- Length

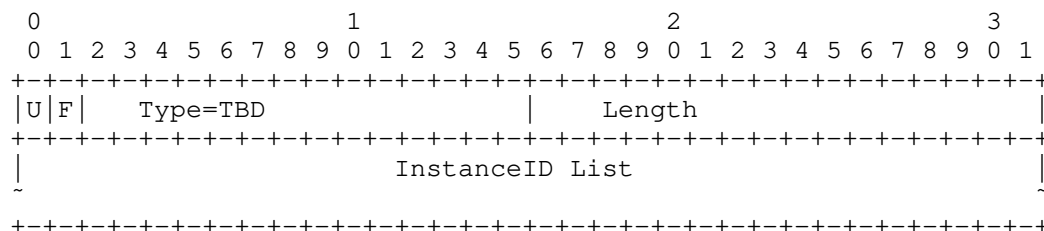
Length of the MAC address, which is 6 octets.

- MAC Address

The MAC address of the sender.

3.3.2. 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 to its port connected to the STP network, including the flush out of MAC addresses relevant to the instances listed in this TLV.



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Topology Changed Instances"

- Length

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

- InstanceID List

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

3.3.3. STP CIST Root Time

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

[illegible]

- U and F Bits

Both are set to 0.

- Type

```
set to TBD for "STP CIST Root Time"
```

- Length

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

- MaxAge

The Maximum Age of this TLV.

- MessageAge

The actual age of this TLV.

- FwdDelay

The delay before the port enters the forwarding status.

- HelloTime

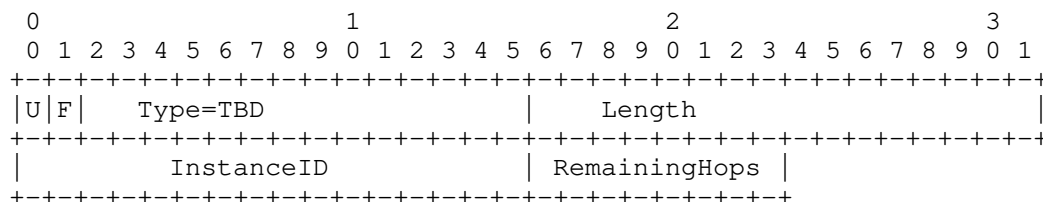
The interval between two continuous configuration BPDUs.

- RemainingHops

The remaining hops of this TLV

3.3.4. STP MSTI Root Time

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP MSTI Root Time"

- Length

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

- InstanceID

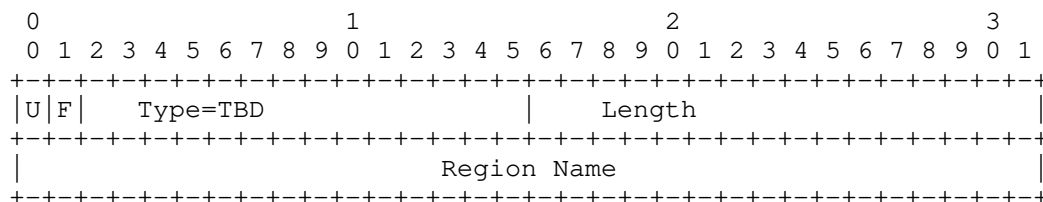
The instance identification number of the MSTI.

- remainingHops

The remaining hops of this TLV

3.3.5. STP Region Name

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Region Name"

- Length

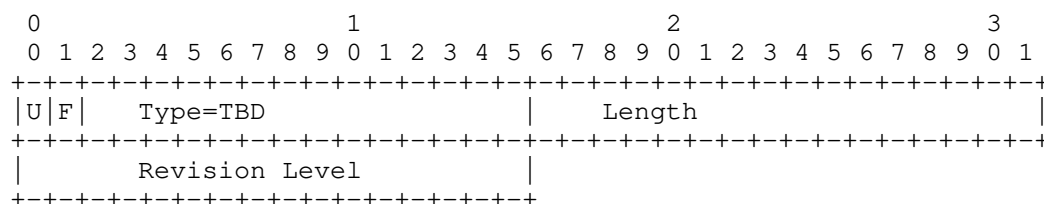
Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Region Name

The Name of the MST Region.

3.3.6. STP Revision Level

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Revision Level"

- Length

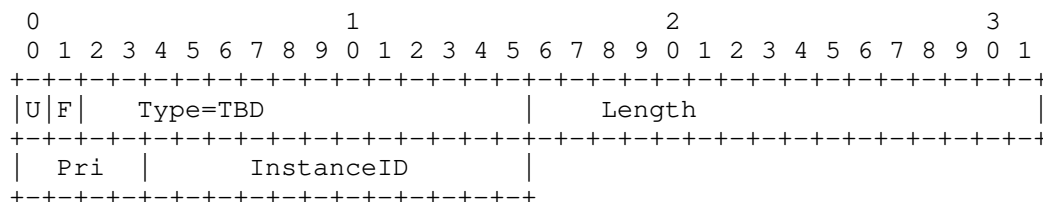
Length of the TLV in octets excluding the U-bit, F-bit, Type, and Length fields.

- Revision Level

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

3.3.7. STP Instance Priority

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



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Instance Priority"

- Length

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

- Pri

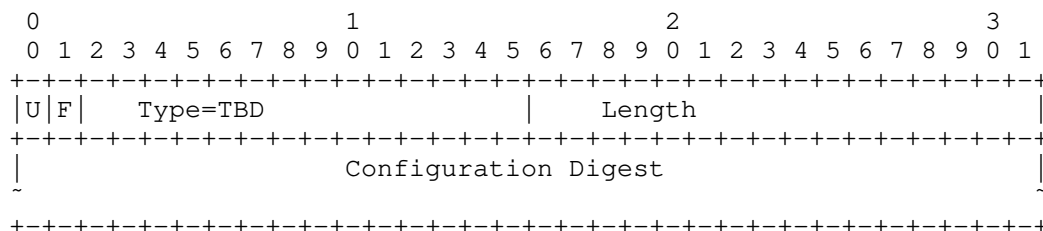
The Instance Priority

- InstanceID

The instance identification number of the MSTI.

3.3.8. STP Configuration Digest

This TLV is used to report the Value of STP VLAN Instance Mapping to other members in the RG.



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Configuration Digest"

- Length

Length of the STP Configuration Digest which is 16 octets.

- Configuration Digest

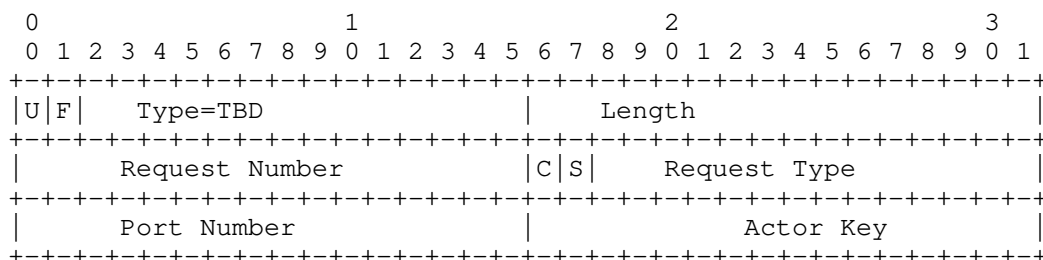
As specified in [802.1q] section 13.7.

3.4. STP Synchronization Request TLV

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

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

The format of the TLV is as follows:



- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Synchronization 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 uniquely identifying the request. Used to match the request with a response. The value of 0 is

reserved for unsolicited synchronization, and MUST NOT be used in the STP Synchronization Request TLV.

- C Bit

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

- S Bit

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

- Request Type

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

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

- Port Number

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

- Actor Key

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

3.5. STP Synchronization Data TLV

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

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

This TLV has the following format:

- U and F Bits

Both are set to 0.

- Type

set to TBD for "STP Synchronization Data TLV"

- Length

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

- Request Number

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

- Flags

2 octets, response flags encoded as follows:

0x00 Synchronization Data Start
0x01 Synchronization Data End

4. Security Considerations

This document raises no new security issues.

5. IANA Considerations

No new registry is requested to be assigned by IANA. RFC Editor: please remove this section before publication.

6. References

6.1. Normative References

[ICCP] L. Martini, S. Salam, et al, "AInter-Chassis Communication Protocol for L2VPN PE Redundancy", draft-ietf-pwe3-iccp-09.txt, work in progress.

6.2. Informative References

[802.1q] "IEEE Standard for Local and Metropolitan Area Networks---

Virtual Bridged Local Area Networks.". IEEE Std 802.1 Q-2005,
May 19, 2006.

Author's Addresses

Mingui Zhang
Huawei

Email: zhangmingui@huawei.com

Huafeng Wen
Huawei

wenhuafeng@huawei.com