Network Working Group
Internet Draft
Intended status: Standards Track
Expiration Date: September 2007

T. Nadeau (Ed)
C. Pignataro (Ed)
Cisco Systems, Inc.

R. Aggarwal (Ed)
Juniper Networks

March 2007

Pseudo Wire Virtual Circuit Connectivity Verification (VCCV)

draft-ietf-pwe3-vccv-13.txt

Status of this Memo

Abstract

This document describes Virtual Circuit Connection Verification (VCCV) which provides a control channel that is associated with a Pseudowire (PW), as well as the corresponding operations and management functions such as connectivity verification to be used over that control channel. VCCV applies to all supported access circuit and transport types

currently defined for PWs.

Table of Contents

1. Specification of requirements

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].

2. Introduction

As network operators deploy Pseudowire (PW) services, fault detection and diagnostic mechanisms particularly for the PSN portion of the network are pivotal. Specifically, the ability to provide end-to-end fault detection and diagnostics for an emulated PW service is critical for the network operator. Operators have indicated in [RFC4377] [RFC3916] that such a tool is required for PW deployments. This document describes procedures for a PSN-agnostic fault detection and diagnostics tool called Virtual Circuit Connection Verification (VCCV).

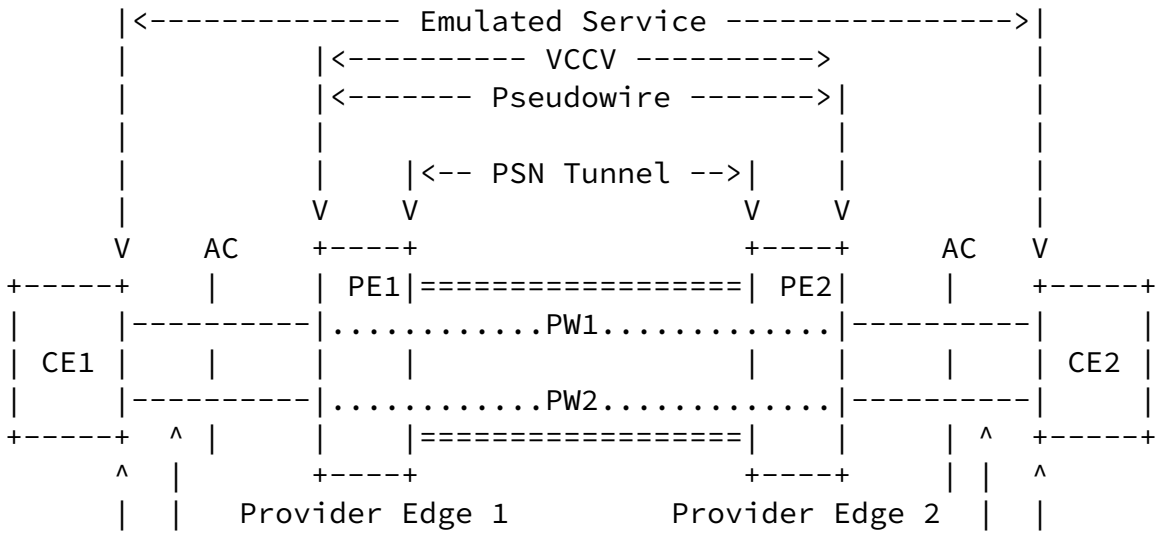


Figure 1: PWE3 VCCV Operation Reference Model

Figure 1 depicts the basic functionality of VCCV, and where it resides within the PWE3 VCCV Operation Reference Model [RFC3985]. Customer Edge (CE) routers CE1 and CE2 are attached to the emulated service via Access Circuits (ACs) to each of the Provider Edge (PE) Routers (PE1 and PE2). These PEs are in-turn, connected via a Pseudowire (PW) that traverses the provider network. VCCV provides several means of creating a control channel between PE routers that attach the PW.

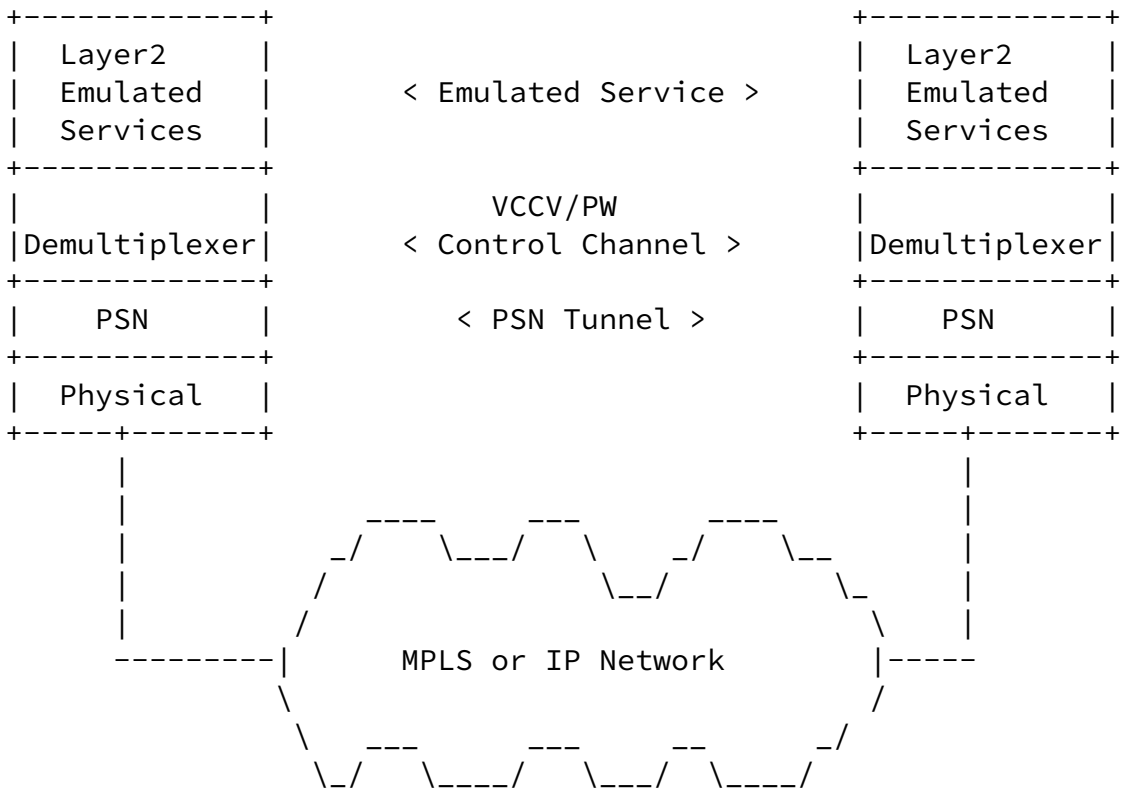


Figure 2: PWE3 Protocol Stack Reference Model
including the VCCV control channel.

Figure 2 depicts how the VCCV control channel is associated with the Pseudowire. Ping and other IP messages are encapsulated using the PWE3 encapsulation as described below in sections 5 and 6. These messages, referred to as VCCV messages, are exchanged only after the desire to exchange such traffic has been negotiated between the PEs (see Section 7.)

3. Overview of VCCV

VCCV defines a set of messages that are exchanged between PEs to verify connectivity of the Pseudowire. To make sure that VCCV packets follow the same path as the PW data flow, they SHOULD be encapsulated with the same PW demultiplexer and trasported over the same PSN tunnel. For example, if MPLS is the PSN in use, then the same label shim header (and label stack) MUST be incorporated. The only cases where this might not be possible is when out-of-band VCCV modes are used which require this encapsulation to be altered; however, these modes are NOT RECOMMENDED.

VCCV can be used both as a fault detection and/or a diagnostic

tool for Pseudowires. An operator can periodically invoke VCCV for proactive connectivity verification on an active Pseudowire, or on an ad hoc or as-needed as a means of manual connectivity verification. When invoking VCCV, the operator triggers a combination of one of its various Control Channel types (CC Types) and one of its various Connectivity Verification types (CV Types.) These include LSP-Ping, L2TPV3, or ICMP Ping [RFC0792] modes and are applicable depending on the underlying PSN.

Since a Pseudowire service is bi-directional, the reply MAY be sent in-band over the PW in the reverse direction. Responses MUST be encapsulated so that they follow the return path of the Pseudowire in this case. In-band responses MUST be attempted first. If an in-band test fails, the operator is advised to

then use a subsequent test using an out-of-band reply mode such as Reply Mode 4 from [RFC4379], which will return the result to the sender via an application level control channel to determine the fault's direction.

The control channel maintained with VCCV can carry fault detection status across a Pseudowire and convey this information between the endpoints of the Pseudowire. Furthermore, this information can then be translated into the native OAM status codes used by the native access technologies, such as ATM or Ethernet. The specific details of such status interworking is out of the scope of this document, and is only noted here to illustrate the utility of VCCV for such purposes. More complete details can be found in [OAM-MAP].

4. CC Types and CV Types

The VCCV Control Channel type (CC type) defines several possible types of control channel that VCCV can support. These control channels can in turn carry several types of protocols defined by the Connectivity Verification type (CV type). VCCV potentially supports multiple CV Types concurrently, but it only supports the use of a single CC Type. The specific type or types of VCCV packets that can be accepted and sent by a router are indicated during capability advertisement as described in sections 5.5 and 6.2. The various VCCV CV types supported MUST be used only when they apply to the context of the PW demultiplexor in use. For example, LSP Ping type should only be used when MPLS is utilized as the PSN.

Once a set of VCCV capabilities is received and advertised, a CC Type and CV Type(s) that match both the received and transmitted capabilites can be selected. That is, a PE router needs to only allow Types that are both received and advertised to be selected,

performing a logical AND between the received and transmitted bitflag fields. The specific CC Type and CV Type(s) are then chosen within the constraints and rules specified in Section 4.1.1, Section 4.1.2 and Section 7. Once a specific CC Type has been chosen (i.e., it matches both the transmitted and received VCCV CC capability), transmitted and replied to, this CC Type MUST be the only one used

until such time as the Pseudowire is re-signaled. In addition, based on these rules and the procedures defined in Section 5.2 of [RFC4447], the Pseudowire MUST be re-signaled if a different set of capabilities types is desired.

The CC and CV type indicator fields are defined as a bitmasks used to indicate the specific CC or CV type or types (i.e.: none, one or more) of control channel packets that may be sent on the VCCV control channel. These values represent the numerical value corresponding to the actual bit being set in the bitfield. The definition of each CC and CV Type is dependent on the context within which it is defined; please refer to the specific MPLS or L2TPv3 sections below.

Control Channel (CC) Types:

The defined values for CC Types are for MPLS PWs are:

```
Bit 0 (0x01) - Type 1: PWE3 control word with 0001b
               as first nibble as defined in [RFC4385].
Bit 1 (0x02) - Type 2: MPLS Router Alert Label.
Bit 2 (0x04) - Type 3: MPLS PW Demultiplexor Label
               TTL = 1 (Type 3).
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

The defined values for CC Types are for L2TPv3 PWs are:

```
Bit 0 (0x01) - L2-Specific Sublayer with V-bit set.
Bit 1 (0x02) - Reserved
Bit 2 (0x04) - Reserved
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

Connectivity Verification (CV) Types:

The defined values for CV Types are for MPLS PWs are:

```
Bit 0 (0x01) - ICMP Ping.
Bit 1 (0x02) - LSP Ping.
Bit 2 (0x04) - BFD for PW Fault Detection Only.
Bit 3 (0x08) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling.
Bit 4 (0x10) - BFD for PW Fault Detection Only. Carrying
               BFD payload without IP headers.
Bit 5 (0x20) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling. Carrying BFD payload
               without IP headers.
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

The defined values for CV Types are for L2TPv3 PWs are:

```
Bit 0 (0x01) - ICMP Ping.
Bit 1 (0x02) - Reserved
Bit 2 (0x04) - BFD for PW Fault Detection Only.
Bit 3 (0x08) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling.
Bit 4 (0x10) - BFD for PW Fault Detection Only. Carrying BFD
               payload without IP headers.
Bit 5 (0x20) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling. Carrying BFD payload without
               IP headers.
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

It should be noted that two pairs of CV Types have been defined when BFD is used. See Section 4.1.1 and 4.1.2.

If none of the types above are supported, the entire CC and CV Type Indicator fields SHOULD be transmitted as 0x00 (i.e.: all bits in the bitfield set to 0) to indicate this to the peer.

If no capability is signaled, then the peer MUST assume that the peer has no VCCV capability and follow the procedures specified in this document for this case.

4.1 Bidirectional Forwarding Detection

When heart-beat indication is necessary for one or more PWs, the Bidirectional Forwarding Detection (BFD) [BFD] provides a means of continuous monitoring of the PW data path and propagation of forward and reverse defect indications.

In order to use BFD, both ends of the PW connection must have signaled the existence of a common control channel and the ability to run BFD on it. Once a node has both signaled and received signaling from its peer of these capabilities and has chosen a single BFD CV Type as specified in Section 4.1.2, it MUST begin sending BFD control packets. The packets MUST be sent on the control channel. The use of the control channel provides the context required to bind and bootstrap the BFD session, thus single-hop BFD initialization procedures are followed [BFD], and BFD MUST be run in asynchronous mode [BFD].

When one of the PEs (PE2 from Figure 1) does not receive control messages from its peer PE (PE1 from Figure 1) during a certain number of transmission intervals (a number provisioned by the operator) PE2 declares that the PW in its receive direction is down. In other words, PE1 enters the "forward defect" state for this PW. PE1 then sends a message to PE2 with H=0 (i.e. "I do not hear you") and with Diagnostic code 1. In turn, PE2 declares the PW is down in its transmit direction and it uses Diagnostic code 3 in its control messages to PE1. PE2 enters the "reverse defect" state for this PW. How it further processes this error condition, and conveys this status the attachment circuits is out of the scope of this specification, and is instead defined in [OAM-MAP].

The VCCV message comprises a BFD packet [BFD] encapsulated as specified by the CV Type (see Section 4.1.1.)

4.1.1 BFD Encapsulation

VCCV defines two pairs of CV Types (see Section 4) which group two ways in which the BFD Connectivity Verification packets may be encapsulated. When the CV Type is either 0x04 or 0x08, the VCCV encapsulation includes the IP/UDP encapsulation as defined in Section 4 of [BFD-V4V6-1HOP]. However, when either CV Type 0x10 or 0x20 are employed, the IP/UDP headers are omitted. In this second group (i.e., cases using CV Type 0x10 or 0x20), the corresponding PW Associated Channel Header's or Layer-2 Specific Sublayer's Channel Type field MUST use the value of PW-ACT-TBD defined in Section 8.2 as a means of allowing the data plane to demultiplex the control channel and identify the encased BFD payload.

Additionally, only the CC Type 1 (PWE3 Control Word with 0001b as first nibble as defined in [RFC4385]) allows for the use of the BFD encapsulation without the IP/UDP headers (i.e., using CV Types 0x10 or 0x20) in conjunction with other CV Types that include an IP Header. That is, CV Types 0x10 or 0x20 MUST NOT be used along with other CV Types, unless the CC Type in used is Type 1 (PWE3 Control Word with 0001b as first nibble as defined in [RFC4385]). This

restriction stems from the fact that Type 1 is the only CC Type that contains a Protocol Identification (PID) field, the Associated Channel Type. If it is desired to concurrently have BFD along with a CV Type that includes an IP Header (e.g., LSP Ping), over a Control Channel utilizing CC Types 2 or 3, then only BFD encapsulations including IP/UDP headers (i.e., CV Types 0x04 or 0x08) can be used.

4.1.2 CV Types for BFD

As with other CV Types, and given the bidirectional nature of BFD, before selecting a given BFD CV Type capability to be used, there MUST be a match in the given CV Type capability advertised and received. That is, only BFD CV Types that were both advertised and received are available to be selected. Additionally, only one BFD CV Type can be used (selecting a BFD CV Type excludes all the rest BFD CV Types). The following list enumerates restrictions on the usage of BFD CV Types:

1. In the case of CV Type 0x08 or 0x20, the AC and PW status SHOULD be conveyed via BFD status codes as specified in [OAM-MAP].

2. The CV Types 0x08 and 0x20, however, SHOULD NOT be used when a control protocol such as LDP or L2TPV3 is available that can signal the AC/PW status to the remote endpoint of the PW [RFC4447].

3. In the case of type 0x04 or 0x10, BFD is used exclusively to detect faults on the PW and the status of those faults SHOULD be conveyed using some means other than BFD, such as using LDP status messages when using MPLS as a transport (see [RFC4447]), or the Circuit Status AVP in an L2TPv3 SLI message for L2TPv3 (see [RFC3931].)

4. Similarly, CV Types 0x04 and 0x10 SHOULD NOT be used when there is no control protocol available to signal the AC/PW status.

5. Only a single BFD CV Type can be seleced and used.

5. VCCV Control Channel for MPLS PSN

When MPLS is used to transport PW packets, VCCV packets are carried over the MPLS LSP as defined in this section. In order to apply IP monitoring tools a PWE3 PW, an operator may configure VCCV as a control channel for the PW between the PEs endpoints [RFC3985]. Packets sent across this channel from the source PE towards the destination PE either as in-band traffic with the PW's data, or out-of-band. In all cases, the control channel traffic MUST NOT be forwarded past the PE

endpoints towards the Customer Edge (CE) devices; instead, they must be intercepted at the PE endpoints for exception processing.

The capability of which control channel type (CC Type) to use is advertised by a PE to indicate which of the various control channel types are supported. Once the receiving PE has chosen a CC Type mode to use, it MUST continue using this mode until such time as the PW is re-signaled. Thus, if a new CC type is desired, the PW must be torn-down and re-established.

Ideally such a control channel would be completely inband. When a control word is present on the PW, it is possible to indicate the control channel by setting a bit in the control word header.

The following subsections define each of the currently defined VCCV Control Channel Types (CC Types).

5.1. Inband VCCV (Type 1)

The PW set-up protocol [RFC4447] determines whether a PW uses a control word. When a control word is used, it SHOULD have the following form for the purpose of indicating VCCV control channel messages. Note that for data, one uses the control word defined just above the MPLS payload [RFC4385].

The PW Associated Channel for VCCV control channel traffic is defined as follows in [RFC4385]:

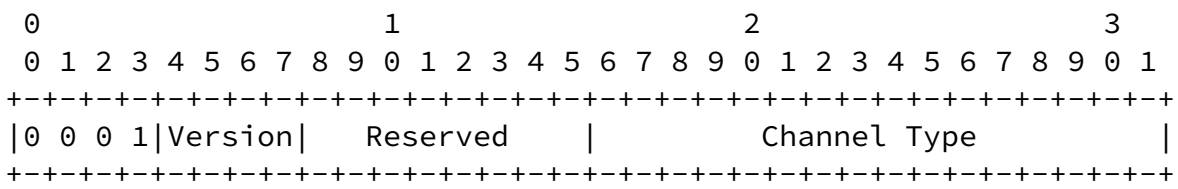


Figure 3: PW Associated Channel Header

The first nibble is set to 0001b to indicate a channel associated with a Pseudowire [RFC4385][RFC4446]. The Version and the Reserved fields are set to 0, and the Channel Type is set to 0x0021 for IPv4 and 0x0057 for IPv6 payloads. If the payload contains BFD without IP/UDP headers, it MUST use PW-ACT-TBD as the Channel Type (see Section 8.2.)

For example, the following is an example of how the ethernet ACH would be received [RFC4448] containing an LSP Ping payload corresponding to a choice of CC Type of 0x01 and a CV Type of 0x02:

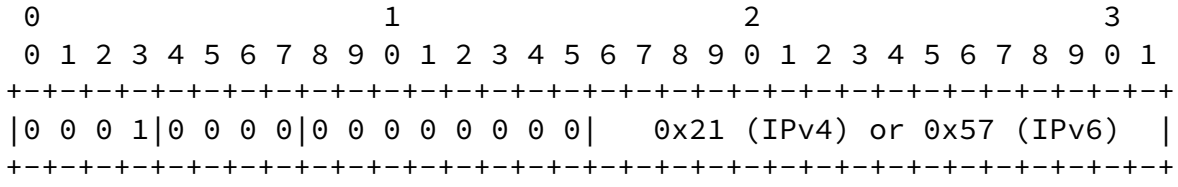


Figure 4: PW Associated Channel Header for VCCV

It should be noted that although some PW types are not required to carry the control word, this type of VCCV MUST only be used for those PW types that do employ the control word when it is in use. Additionally, this is the only CC Type mode that allows the concurrent usage of CV Types that are encapsulated with an IP Header (e.g., LSP Ping) along with other CV Types that lack an IP Header (e.g., BFD encapsulation as per CV Types 0x10 or 0x20.)

This is the preferred mode of VCCV operation when the control word is present.

5.2. Out-of-Band VCCV (Type 2)

A VCCV control channel can alternatively be created by using the MPLS router alert label [RFC3032] immediately above the PW label. It should be noted that this approach MAY result in a differnt equal cost multi-path (ECMP) hashing behavior than Pseudowire PDUs and thus result in the VCCV control channel traffic taking a path which differs from that of the actual data traffic under test.

This is the preferred mode of VCCV operation when the control word is not present.

5.3. TTL Expiry VCCV (Type 3)

The TTL of the PW label can be set to 1 to force the packet to be processed within the destination router's control plane. This is an inband control channel identification mechanism that is an alternate to Section 5.1.

To use this type, the control word MUST be used.

5.4 VCCV Connectivity Verification Types

5.4.1 MPLS LSP Ping

The LSP Ping header MUST be used in accordance with [RFC4379] and MUST also contain the target FEC Stack containing the sub-TLV of 8 for the "L2 VPN endpoint", 9 (deprecated) or 10 for "FEC 128 Pseudowire" or 11 for the "FEC 129 Pseudowire". The sub-TLV indicates the PW to be verified.

5.5 VCCV Capability Advertisement for MPLS PSN

To permit the indication of the type or types of PW control channel(s) and connectivity verification mode or modes over a particular PW, a VCCV parameter is defined below that is used as part of the PW establishment signaling. When a PE signals a PW and desires PW OAM for that PW, it MUST indicate this during PW establishment using the messages defined below. Specifically, for PE MUST include the VCCV interface parameter sub-TLV (0x0C) assigned in [RFC4446] in the PW setup message [RFC4447].

The decision of the type of VCCV control channel is left completely to the receiving control entity, although the set of choices is given by the sender in that it indicates the type or types of control channels and connectivity verification types that it can understand. The receiver SHOULD chose a single Control Channel type from the match between the choices sent and received, based on the capability advertisement selection specified below in Section 7, and it MUST continue to use this type for the duration of the life of the control channel. Changing Control Channel types after one has been established to be in use could potentially cause problems at the receiving end, and could also lead to interoperability issues, thus it is NOT RECOMMENDED.

When a PE sends a label mapping message for a PW, it uses the VCCV parameter to indicate the type of OAM control channels and connectivity verification type or types it is willing to receive and can send on that PW. The capablity of supporting a control channel or channels, and connectivity type or types used over that control channel or channels MUST be signaled before the remote PE may send VCCV messages, and then it can do so only on the control channel or channels, and using the connectivity verification type or types indicated.

If a PE receives VCCV messages prior to advertising capability for this message, it MUST discard these messages and not reply to them. In this case, the PE SHOULD increment an error counter and optionally issue a system and/or SNMP notification to indicate to the system administrator that this condition exists.

When LDP is used as the PW signaling protocol the requesting PE

indicates its configured VCCV capability or capabilities to the

remote PE by including the VCCV parameter with appropriate options indicating which control channel and connectivity verification types it supports in the VCCV interface parameter sub-TLV field of the PW ID FEC TLV (FEC 128) or in the interface parameter sub-TLV of the Genralized PW ID FEC TLV (FEC 129). The requesting PE MAY indicate that it supports multiple control channel options, and in doing so agrees to support any and all indicated types if transmitted to it, but MUST do so in accordance with the rules stipulated in Section 5.5.1 (VCCV Capability Advertisement Sub-TLV.)

Local policy may direct the PE to support certain OAM capability and to indicate it. The absence of the VCCV parameter indicates that no OAM functions are supported by the requesting PE, and thus the receiving PE MUST NOT send any VCCV control channel traffic to it. The reception of a VCCV parameter with no options set MUST be ignored as if one is not transmitted at all.

The receiving PE similarly indicates its supported control channel types in the label mapping message. These may or may not be the same as the ones that were sent to it. The sender should examine the set that is returned to understand which control channels it may establish with the remote peer, as specoified in Section 4 and Section 7. Similarly, it MUST NOT send control channel traffic to the remote PE for which the remote PE has not indicated it supports.

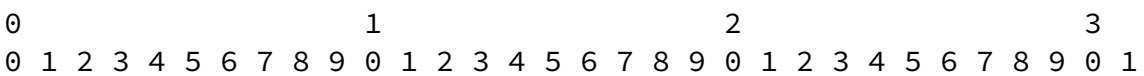
5.5.1 VCCV Capability Advertisement LDP Sub-TLV

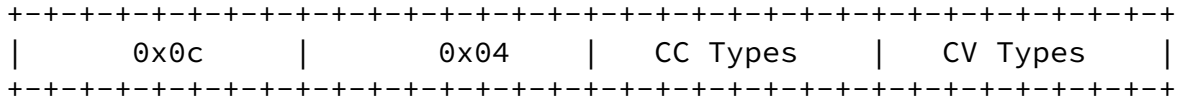
[RFC4447] defines an Interface Parameter Sub-TLV in the LDP PW ID FEC (FEC 128) and an Interface Parameters TLV in the LDP Generalized PW ID FEC (FEC 129) to signal different capabilities for specific PWs. An optional sub-TLV parameter is defined to indicate the capability of supporting none, one or more control channel and connectivity verification types for VCCV. This is the VCCV parameter field. If FEC 128 is used the VCCV parameter field is carried in the Interface Parameter sub-TLV. field If FEC 129 is used it is carried as an Interface Parameter sub-TLV in the Interface Parameters TLV.

The VCCV parameter ID is defined as follows in [RFC4446]:

Parameter ID	Length	Description
0x0c	4	VCCV

The format of the VCCV parameter field is as follows:





The Control Channel (CC Types) type field defines a bitmask used to indicate the type of control channel(s) (i.e.: none, one or more) that a router is capable of receiving control channel traffic on. If more than one control channel is specified, the router agrees to accept control traffic over either control channel; however, see the rules specified in Section 4 and Section 7 for more details. If none of the types are supported, a CC Type Indicator of 0x00 SHOULD be transmitted to indicate this to the peer. However, if no capability is signaled, then the PE MUST assume that its peer is incapable of receiving any of the VCCV CC Types and MUST NOT send any OAM control channel traffic to it. Note that the CC and CV types definitions are consistent regardless of the PW's transport or access circuit type. The CC and CV values are defined in Section 4.

6. VCCV Control Channel for L2TPv3/IP PSN

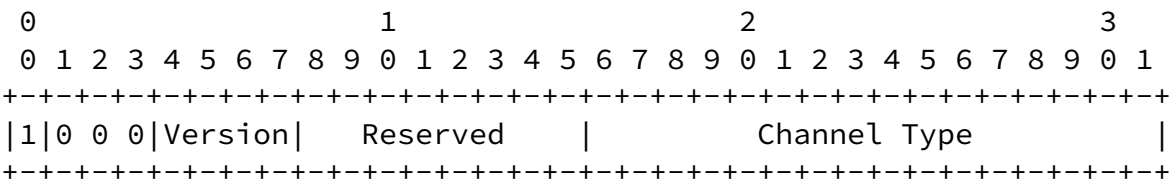
When L2TPv3 is used to setup a PW over an IP PSN, VCCV packets are carried over the L2TPv3 session as defined in this section. L2TPv3 provides a "Hello" keepalive mechanism for the L2TPv3 control plane that operates in-band over IP or UDP (see Section 4.4 of [RFC3931].) This built-in Hello facility provides dead peer and path detection only for the group of sessions associated with the L2TP Control Connection. VCCV, however, allows individual L2TP sessions to be tested. This provides a more granular mechanism which can be used to troubleshoot potential problems within the dataplane of L2TP endpoints themselves, or to provide additional connection status of individual Pseudowires.

The capability of which control channel type (CC Type) to use is advertised by a PE to indicate which of the potentially various control channel types are supported. Once the receiving PE has chosen a mode to use, it MUST continue using this mode until such time as the PW is re-signaled. Thus, if a new CC type is desired, the PW must be torn-down and re-established.

In order to carry VCCV messages within an L2TPv3 session data packet, the PW MUST be established such that an L2-Specific Sublayer (L2SS) that defines the V-bit is present. This document defines the V-bit for the Default L2-Specific Sublayer [RFC3931] and the ATM-Specific Sublayer [RFC4454] using the Bit 0 position (see Section 8.3.2 and

Section 8.3.3.) The L2-Specific Sublayer presence and type (either the Default or a PW-Specific L2SS) is signaled via the L2-Specific Sublayer AVP, Attribute Type 69, as defined in [RFC3931]. The V-bit

within the L2-Specific Sublayer is used to identify that a VCCV message follows, and when the V-bit is set the L2SS has the following format:



L2-Specific Sublayer Format when the V-bit (bit 0) is set

The VCCV messages are distinguished from user data by the V-bit. The V-bit is set to 1, indicating that a VCCV session message follows. The next three bits MUST be set to 0 when sending and ignored upon receipt. The remaining fields comprising 28 bits (i.e., Version, Reserved and Channel Type) follow the same definition, format and number registry from Section 5 of [RFC4385].

Depending on the CV Type in use, the Channel Type can indicate IPv4, IPv6 (see [RFC4385]) or BFD (see Section 8.2) as VCCV payload directly following the L2SS. For CV Types of 0x01, 0x04 and 0x08, the Channel Type can indicate IPv4 (0x21) or IPv6 (0x57); for CV Types of 0x10 and 0x20, the Channel Type indicates BFD Without IP/UDP Header (PW-ACT-TBD).

6.1. L2TPv3 VCCV Message

The VCCV message over L2TPv3 directly follows the L2-Specific Sublayer with the V-bit set. It could either contain an ICMP Echo packet as described in Section 6.1.1, or a BFD packet as described in Section 6.1.2.

6.1.1. L2TPv3 VCCV using ICMP Ping

When this connectivity verification mode is used, an ICMP Echo packet [RFC0792] achieves connectivity verification. The ICMP Ping packet directly follows the L2SS with the V-bit set. In the ICMP Echo request, the IP Header fields MUST have the following values: the destination IP address is set to the remote LCCE's IP address for the tunnel endpoint, the source IP address is set to the local LCCE's IP address for the tunnel endpoint, and the TTL is set to 1.

6.1.2. L2TPv3 VCCV using BFD

The L2TPv3 Session ID provides the context to demultiplex the first BFD control packet. See Section 4.1, Section 4.1.1 and Section 4.1.2 for additional details on BFD usage, BFD encapsulation and BFC CV Types.

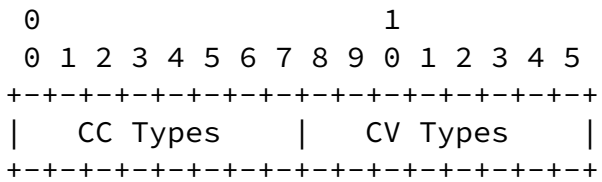
6.2. L2TPv3 VCCV Capability Indication

A new optional AVP is defined in Section 6.2.1 to indicate the VCCV capabilities during session establishment. An LCCE MUST signal its desire to use connectivity verification for a particular L2TPv3 session and its VCCV capabilities using the VCCV Capability AVP.

6.2.1. L2TPv3 VCCV Capability AVP

The "VCCV Capability AVP", Attribute type AVP-TBD, specifies the VCCV capabilities as a pair of bitflags for the Control Channel (CC) and Connectifity Verification (CV) Types. This AVP is exchanged during session establishment (in ICRQ, ICRP, OCRQ or OCRP messages). The value field has the following format:

VCCV Capability AVP (ICRQ, ICRP, OCRQ, OCRP)



CC Types:

The Control Channel (CC) Types field defines a bitmask used to indicate the type of control channel(s) that may be used to receive OAM traffic on for the given Session. The router agrees to accept VCCV traffic at any time over any of the signaled VCCV control channel types. CC Type values are defined in Section 4. Although there is only one value defined in this document, the CC Types field is included for forward compatibility should further CC Types need to be defined in the future.

A CC Type of 0x01 may only be requested when there is an L2-Specific Sublayer that defines the V-bit present. If a CC Type of 0x01 is requested without requesting an L2-Specific Sublayer AVP with an L2SS type that defines the V-bit, the session MUST be disconnected with a CDN message.

If no CC Type is supported, a CC Type Indicator of 0x00 SHOULD be sent.

CV Types:

The Connectifity Verification (CV) Types field defines a bitmask used to indicate the specific type or types (i.e.: none, one or more) of control packets that may be sent on the specified VCCV control channel. CV Type values are defined in Section 4.

If no VCCV Capability AVP is signaled, then the LCCE MUST assume that the peer is incapable of receiving VCCV and MUST NOT send any OAM control channel traffic to it.

All L2TP AVPs have an M (Mandatory) bit, H (Hidden) bit, Length, and Vendor ID. The Vendor ID for the VCCV Capability AVP MUST be 0, indicating that this is an IETF-defined AVP. This AVP MAY be hidden (the H bit MAY be 0 or 1). The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 8.

6.3. L2TPv3 VCCV Operation

An LCCE sends VCCV messages on an L2TPv3 signaled Pseudowire for fault detection and diagnostic of the L2TPv3 session. The VCCV message travels inband with the Session and follows the exact same path as the user data for the session, because the IP header and L2TPv3 Session header are identical. The egress LCCE of the L2TPv3 session intercepts and processes the VCCV message, and verifies the signaling and forwarding state of the Pseudowire on reception of the VCCV message. Any faults detected can be signaled in the VCCV response. It is to be noted that the VCCV mechanism for L2TPv3 is primarily targeted at verifying the Pseudowire forwarding and signaling state at the egress LCCE. It also helps when L2TPv3 Control Connection and Session paths are not identical.

An LCCE MUST NOT send VCCV packets on an L2TPv3 session unless it has received VCCV capability by means of the VCCV Capability AVP from the remote end. If an LCCE receives VCCV packets and its not VCCV capable or it has not sent VCCV capability indication to the remote end, it MUST discard these messages. It should also increment an error counter. In this case the LCCE MAY optionally issue a system and/or SNMP notification.

As specified in Section 4, CV Types sent and received need to match in order to be used. Specifically, and also because BFD is bidirectional in nature, when using BFD as the connectivity verification type, an LCCE must send VCCV packets on an L2TPv3 session only if it has signaled VCCV capability with a BFD CV Type to

the remote end and received VCCV capability with a matching BFD CV Type from the remote end.

7. Capability Advertisement Selection

When a PE receives a VCCV capability advertisement, the advertisement may potentially contain more than one CC or CV Type. Only matching capabilities can be selected. When multiple capabilities match, only one CC Type MUST be used, and the CV Type or CV Types to be used MUST follow the restrictions described in sections 4, 4.1.1 and 4.1.2.

In particular, once a valid CC Type is used by a PE (traffic sent using that encapsulation), the PE MUST NOT send any traffic down another CC Type control channel.

For cases were multiple CC Types are advertised, the following precedence rules apply when choosing the single CC Type to use:

0x01 - PWE3 control word with 0001b as first nibble
0x04 - MPLS PW Demultiplexor Label TTL = 1
0x02 - MPLS Router Alert Label

The selection of the BFD CV Type to use out of the four BFD CV Types defined in this document is tied to multiple factors: Given that BFD is bidirectional in nature, only CV Types that are both received and sent in VCCV capability signaling advertisement can be selected. When a control protocol that can signal the AC/PW status is not available, CV Types CV Types 0x04 and 0x10 SHOULD NOT be used. When a control protocol that can signal the AC/PW status (such as LDP [RFC4447] or L2TPv3 [RFC3931]) is available, CV Types 0x08 and 0x20 SHOULD NOT be used. All BFD CV Types are mutually exclusive with the rest, and selecting a BFD CV Type prevents the use of any of the other three BFD CV Types. Finally, only CC Type 0x01 (for both MPLS and L2TPv3) supports the concurrent use of BFD CV Types 0x10 or 0x20 along with another CV Type that uses an encapsulation with IP headers. Therefore, if it is desired to use a CV Type of 0x10 or 0x20 simultaneously with a CV Type that uses IP Headers, then CC Type 0x01 MUST be used.

8. IANA Considerations

8.1. VCCV Interface Parameters Sub-TLV

The VCCC Interface Parameters Sub-TLV codepoint is defined in [RFC4446]. IANA is requested to create and maintain registries for the CC Types and CV Types (bitmasks in the VCCV Parameter ID). The CC

Type and CV Type new registries (see Section 8.1.1 and 8.1.2 respectively) are to be created in the Pseudo Wires Name Spaces, reachable from <http://www.iana.org/assignments/pwe3-parameters>. The allocations must be done using the "IETF Consensus" policy defined in RFC2434.

8.1.1. Control Channel Types (CC Types)

IANA is requested to set up a registry of "VCCV Control Channel Types". These are 8 bitfield values. CC Type values 0x01, 0x02, and 0x04 are specified in Section 4 of this document. The remaining bitfield values (0x08, 0x10, 0x20, 0x40 and 0x80) are to be assigned by IANA using the "IETF Consensus" policy defined in [RFC2434]. A VCCV Control Channel Type description and a reference to an RFC approved by the IESG are required for any assignment from this registry.

```
Bit 0 (0x01) - Type 1: PWE3 control word with 0001b
               as first nibble as defined in [RFC4385].
Bit 1 (0x02) - Type 2: MPLS Router Alert Label.
Bit 2 (0x04) - Type 3: MPLS PW Demultiplexor Label
               TTL = 1 (Type 3).
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

8.1.2. Connectivity Verification Types (CV Types)

IANA is requested to set up a registry of "VCCV Control Verification Types". These are 8 bitfield values. CV Type values 0x01, 0x02, 0x04 0x08, 0x10 and 0x20 are specified in Section 4 of this document. The remaining bitfield values (0x40 and 0x80) are to be assigned by IANA using the "IETF Consensus" policy defined in [RFC2434]. A VCCV Control Verification Type description and a reference to an RFC approved by the IESG are required for any assignment from this registry.

```
Bit 0 (0x01) - ICMP Ping.
Bit 1 (0x02) - LSP Ping.
Bit 2 (0x04) - BFD for PW Fault Detection Only.
Bit 3 (0x08) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling.
Bit 4 (0x10) - BFD for PW Fault Detection Only. Carrying
               BFD payload without IP headers.
Bit 5 (0x20) - BFD for PW Fault Detection and AC/PW Fault
               Status Signaling. Carrying BFD payload
```

without IP headers.
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved

8.2 PW Associated Channel Type

The PW Associated Channel Types used by VCCV as defined above in sections 4.1, 4.2 and 4.3 rely on previously allocated numbers from the Pseudowire Associated Channel Types Registry [RFC4385] in the Pseudo Wires Name Spaces reachable from <http://www.iana.org/assignments/pwe3-parameters>. In particular, 0x21 (Internet Protocol version 4) MUST be used whenever an IPv4 payload follows the Pseudowire Associated Channel Header, or 0x57 MUST be used when an IPv6 payload follows the Pseudowire Associated Channel Header.

In cases where raw BFD follows the Pseudowire Associated Channel Header (i.e.: the IP/UDP encapsulation as specified in [BFD] will not be present), a new Pseudowire Associated Channel Types Registry [RFC4385] entry of PW-ACT-TBD is used. IANA is requested to reserve a new Channel Types value as follows:

Value (in hex)	Protocol Name	Reference
PW-ACT-TBD	BFD Without IP/UDP Header	[This document]

8.3. L2TPv3 Assignments

Sections 8.3.1 through 8.3.3 are registrations of new L2TP values for registries already managed by IANA. Section 8.3.4 requests a new registry to be added to the existing L2TP name spaces, and be maintained by IANA accordingly. The Layer Two Tunneling Protocol "L2TP" Name Spaces are reachable from <http://www.iana.org/assignments/l2tp-parameters>.

8.3.1. Control Message Attribute Value Pairs (AVPs)

An additiona AVP Attribute is specified in Section 6.2.1. It is required to be defined by IANA as described in Section 2.2 of [RFC3438].

Attribute Type	Description
AVP-TBD	VCCV Capability AVP

8.3.2. Default L2-Specific Sublayer bits

The Default L2-Specific Sublayer contains 8 bits in the low-order portion of the header. This document defines one reserved bits in the Default L2-Specific Sublayer in Section 6, which may be assigned by IETF Consensus [RFC2434]. It is required to be assigned by IANA.

```
Default L2-Specific Sublayer bits - per [RFC3931]
---------------------------------
Bit 0 - V (VCCV) bit
```

8.3.3. ATM-Specific Sublayer bits

The ATM-Specific Sublayer contains 8 bits in the low-order portion of the header. This document defines one reserved bits in the ATM-Specific Sublayer in Section 6, which may be assigned by IETF Consensus [RFC2434]. It is required to be assigned by IANA.

```
ATM-Specific Sublayer bits - per [RFC4454]
--------------------------
Bit 0 - V (VCCV) bit
```

8.3.4. VCCV Capability AVP Values

This is a new registry for IANA to maintain in the L2TP Name Spaces.

IANA is requested to maintain a registry for the CC Types and CV Types bitmasks in the VCCV Capability AVP, defined in Section 6.2.1. The allocations must be done using the "IETF Consensus" policy defined in [RFC2434]. A VCCV CC or CV Type description and a reference to an RFC approved by the IESG are required for any assignment from this registry.

IANA is requested to reserve the following bits in this registry:

```
VCCV Capability AVP (Attribute Type AVP-TBD) Values
--------------------------------------------------

Control Channel (CC) Types
```

Bit 0 (0x01) - L2-Specific Sublayer with V-bit set.
Bit 1 (0x02) - Reserved
Bit 2 (0x04) - Reserved
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved

Connectifity Verification (CV) Types

Bit 0 (0x01) - ICMP Ping
Bit 1 (0x02) - Reserved
Bit 2 (0x04) - BFD for PW Fault Detection Only.
Bit 3 (0x08) - BFD for PW Fault Detection and AC/PW Fault Status Signaling.
Bit 4 (0x10) - BFD for PW Fault Detection Only. Carrying BFD payload without IP headers.
Bit 5 (0x20) - BFD for PW Fault Detection and AC/PW Fault Status Signaling. Carrying BFD payload without IP headers.
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved

9. Security Considerations

Routers that implement the mechanism described herein are subject to to additional denial-of-service attacks as follows:

An intruder may impersonate an LDP peer in order to force a failure and reconnection of the TCP connection. Please see the Security Considerations section of [RFC3036] details.

An intruder could intercept or inject VCCV packets effectively providing false positives or false negatives.

An intruder could deliberately flood a peer router with VCCV messages to either obtain services without authorization or to

deny services to others.

A misconfigured or misbehaving device could inadvertantly flood a peer router with VCCV messages which could result in a denial of services. In particular, if a router is either implicitly or explicitly indicated that it cannot support one or all of the types of VCCV, but is sent those messages in sufficient quantity, could result in a denial of service.

All of attacks above which concern the L2TPv3 or LDP control planes may be countered by use of a control message authentication scheme between LDP or L2TPv3 peers, such as the MD5-based scheme outlined in [RFC3036] or the MD5 or SHA-1 based schemes described in [RFC3931] to provide mutual peer authentication and individual control message integrity and authenticity checking (see Section 8.1 of [RFC3931]). Implementation of IP source address filters may also aid in deterring

these types of attacks.

VCCV message throttling mechanisms should be employed, especially in distributed implementations which have a centralized control plane processor with various line cards attached by some data path. In these architectures VCCV messages may be processed on the central processor after being forwarded there by the receiving line card. In this case, the path between the line card and the control processor may become saturated if appropriate VCCV traffic throttling is not employed, which could lead to a denial of service. Such filtering is also useful for preventing the processing of unwanted VCCV messages, such as those which are sent on unwanted (and perhaps unadvertised) control channel types or VCCV types.

VCCV spoofing requires MPLS PW label spoofing and spoofing the PSN tunnel header. As far as the PW label is concerned the same considerations as specified in [RFC3031] apply. If the PSN is a MPLS tunnel, PSN tunnel label spoofing is also required. For L2TPv3, data packet spoofing considerations are outlined in Section 8.2 of [RFC3931]. While the L2TPv3 Session ID provides traffic separation, the optional Cookie provides additional protection to thwarts spoofing attacks. To maximize protection against a variety of dataplane attacks, a 64-bit cookie can be used. L2TPv3 can also be run over IPsec as detailed in Section 4.1.3 of [RFC3931].

10. Acknowledgements

The authors would like to thank Hari Rakotoranto, Michel Khouderchah, Bertrand Duvivier, Vanson Lim, Chris Metz, W. Mark Townsley, Eric Rosen, Dan Tappan, Danny McPherson, Luca Martini, Don O'Connor, Neil Harrison, Danny Prairie and Mustapha Aissaoui for their valuable comments and suggestions.

11. References

11.1. Normative References

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

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

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y.,

Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[RFC3036] Andersson, L., Doolan, P., Feldman, N., Fredette, A., and B. Thomas, "LDP Specification", RFC 3036, January 2001.

[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.

[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.

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[BFD] Katz, D. and D. Ward, "Bidirectional Forwarding Detection", draft-ietf-bfd-base-05 (work in progress), June 2006.

11.2. Informative References

[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.

[RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

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

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[OAM-MAP] Nadeau, T., "Pseudo Wire (PW) OAM Message Mapping", draft-ietf-pwe3-oam-msg-map-04 (work in progress),

March 2006.

[RFC3438] Townsley, W., "Layer Two Tunneling Protocol (L2TP) Internet Assigned Numbers Authority (IANA) Considerations Update", BCP 68, RFC 3438, December 2002.

[RFC4448] Martini, L., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS

Networks", RFC 4448, April 2006.

[RFC4454] Singh, S., Townsley, M., and C. Pignataro, "Asynchronous Transfer Mode (ATM) over Layer 2 Tunneling Protocol Version 3 (L2TPv3)", RFC 4454, May 2006.

[BFD-V4V6-1HOP] Katz, D. and D. Ward, "BFD for IPv4 and IPv6 (Single Hop)", draft-ietf-bfd-v4v6-1hop-05 (work in progress), June 2006.

12. Editors' Addresses

Thomas D. Nadeau
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719

Email: tnadeau@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7025 Kit Creek Road
PO Box 14987
Research Triangle Park, NC 27709

EMail: cpignata@cisco.com

Rahul Aggarwal
Juniper Networks
1194 North Mathilda Ave.
Sunnyvale, CA 94089

Email: rahul@juniper.net

13. Contributors' Addresses

George Swallow
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719

Email: swallow@cisco.com

Monique Morrow
Cisco Systems, Inc.
Glatt-com
CH-8301 Glattzentrum
Switzerland

Email: mmorrow@cisco.com

Yuichi Ikejiri
NTT Communication Corporation
1-1-6, Uchisaiwai-cho, Chiyoda-ku
Tokyo 100-8019
Shinjuku-ku, JAPAN

Email: y.ikejiri@ntt.com

Kenji Kumaki
KDDI Corporation
KDDI Bldg. 2-3-2,
Nishishinjuku,
Tokyo 163-8003,
JAPAN

E-mail: ke-kumaki@kddi.com

Peter B. Busschbach
Lucent Technologies
67 Whippany Road
Whippany, NJ, 07981

E-mail: busschbach@lucent.com

Vasile Radoaca
Nortel Networks
Billerica, MA, 01803

Email: vasile@nortelnetworks.com

14. Intellectual Property Statement

15. Full Copyright Statement