PWE3
Internet-Draft
Intended status: Standards Track
Expires: November 14, 2009

T. Nadeau, Ed.
BT
C. Pignataro, Ed.
Cisco Systems, Inc.

May 13, 2009

Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV) draft-ietf-pwe3-vccv-bfd-04

Status of this Memo

Copyright Notice

Abstract

This document describes new Connectivity Verification (CV) types using Bidirectional Forwarding Detection (BFD) with Virtual Circuit Connectivity Verification (VCCV). VCCV 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.

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].

The reader is expected to be familiar with the terminology and abbreviations defined in [RFC5085].

2. Introduction

This document describes new Connectivity Verification (CV) types using Bidirectional Forwarding Detection (BFD) with Virtual Circuit Connectivity Verification (VCCV). VCCV [RFC5085] provides a control channel that is associated with a Pseudowire (PW), as well as the corresponding operations and management functions such as connectivity/fault verification to be used over that control channel.

BFD [I-D.ietf-bfd-base] is used over the VCCV control channel primarily as a pseudowire fault detection mechanism, for detecting dataplane failures. Some BFD CV Types can additionally carry fault status 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, Frame-Relay or Ethernet. The specific details of such status interworking are out of the scope of this document, and are only noted here to illustrate the utility of BFD over VCCV for such purposes. Those details can be found in [I-D.ietf-pwe3-oam-msg-map].

The new BFD CV Types are PW Demultiplexer-agnostic, and hence applicable for both MPLS and L2TPv3 Pseudowire Demultiplexers. This document concerns itself with the BFD VCCV operation over Single-Segment Pseudowires (SS-PW). This specification describes procedures only for BFD asynchronous mode.

3. Bidirectional Forwarding Detection Connectivity Verification

VCCV can support several Connectivity Verification (CV) types. This section defines new CV Types for use when BFD is used as the VCCV payload.

Four CV Types are defined for BFD. Table 1 summarizes the BFD CV Types, grouping them by encapsulation (i.e., with vs. without IP/UDP headers) and by functionality (i.e., fault detection only vs. fault detection and status signaling).

	Fault Detection Only	Fault Detection and Status Signaling
BFD, IP/UDP encapsulation (with IP/UDP Headers)	0x04	0x08
BFD, PW-ACH encapsulation (without IP/UDP Headers)	0x10	0x20

Table 1: Bitmask Values for BFD CV Types

3.1. BFD CV Type Operation

When heart-beat indication is necessary for one or more PWs, the Bidirectional Forwarding Detection (BFD) [I-D.ietf-bfd-base] provides a means of continuous monitoring of the PW data path and, in some operational modes, propagation of PW receive and transmit defect state indications.

In order to use BFD, both ends of the PW connection need to agree on the BFD CV Type to use:

- For statically provisioned pseudowires, both ends need to be statically configured to use the same BFD CV Type (in addition to be statically configured for VCCV with the same CC Type).

- For dynamically established pseudowires, both ends of the PW must have signaled the existence of a control channel and the ability to run BFD on it (see Section 3.3 and Section 4).

Once a node has selected a valid BFD CV Type to use (either statically provisioned or selected dynamically after the node has both signaled and received signaling from its peer of these capabilities), it begins sending BFD control packets.

The BFD control packets are sent on the VCCV control channel. The use of the VCCV control channel provides the context required to bind and bootstrap the BFD session, since discriminator values are not exchanged; the pseudowire demultiplexer field (e.g., MPLS PW Label or L2TPv3 Session ID) provides the context to demultiplex the first BFD control packet, and thus single-hop BFD initialization procedures are followed (see Section 3 of [I-D.ietf-bfd-v4v6-1hop] and Section 6 of [I-D.ietf-bfd-generic]). A single BFD session exists per-pseudowire. Both PW endpoints take the Active role sending initial BFD Control packets with a "Your Discriminator" field of zero, and BFD Control

packets received with a "Your Discriminator" field of zero are associated to the BFD session bound to the PW. BFD MUST be run in asynchronous mode (see [I-D.ietf-bfd-base]).

The operation of BFD VCCV for PWs is therefore symmetrical. Both endpoints of the bidirectional pseudowire MUST send BFD messages on the VCCV control channel.

The details of the BFD state machine are as per Section 6.2 of [I-D.ietf-bfd-base]. The following scenario exemplifies the operation: When the downstream PE (D-PE) does not receive BFD control messages from its upstream peer PE (U-PE) during a certain number of transmission intervals (a number provisioned by the operator as "Detect Mult" or detection time multiplier [I-D.ietf-bfd-base]), D-PE declares that the PW in its receive direction is down. In other words, D-PE enters the "PW receive defect" state for this PW. After this calculated Detection Time (see Section 6.8.4 of [I-D.ietf-bfd-base]), D-PE declares the session Down, and signals this to the remote end via the State (Sta) with Diagnostic code 1 (Control Detection Time Expired). In turn, U-PE declares the PW is down in its transmit direction, setting the State to Down with Diagnostic code 3 (Neighbor signaled session down) in its control messages to D-PE. U-PE enters the "PW transmit defect" state for this PW. How it further processes this error condition, and potentially conveys this status to the attachment circuits is out of the scope of this specification, and is instead defined in [I-D.ietf-pwe3-oam-msg-map].

3.2. BFD Encapsulation

The VCCV message comprises a BFD Control packet [I-D.ietf-bfd-base] encapsulated as specified by the CV Type. There are two ways in which a BFD connectivity verification packet may be encapsulated over the VCCV control channel. This document defines four BFD CV Types (see Section 3), which can be grouped into two pairs of BFD CV Types from an encapsulation point of view. See Table 1 in Section 3 that summarizes the BFD CV Types.

o IP/UDP BFD Encapsulation (BFD with IP/UDP Headers)

  In the first method, the VCCV encapsulation of BFD includes the IP/UDP headers as defined in Section 4 of [I-D.ietf-bfd-v4v6-1hop]. BFD Control packets are therefore transmitted in UDP with destination port 3784 and source port within the rage 49152 through 65535. The IP Protocol Number and UDP Port numbers discriminate among the possible VCCV payloads (i.e., differentiate among ICMP Ping and LSP Ping defined in [RFC5085] and BFD).

The IP version (IPv4 or IPv6) MUST match the IP version used for signaling for dynamically established pseudowires, or MUST be configured for statically provisioned pseudowires. The source IP address is an address of the sender. The destination IP address is a (randomly chosen) IPv4 address from the range 127/8 or IPv6 address from the range 0:0:0:0:0:FFFF:127.0.0.0/104. The rationale is explained in Section 2.1 of [RFC4379]. The Time to Live/Hop Limit and Generalized TTL Security Mechanism (GTSM) procedures from Section 5 of [I-D.ietf-bfd-v4v6-1hop] apply to this encapsulation, and hence the TTL/Hop Limit is set to 255.

In this encapsulation, the BFD CV Type used in signaling (if used) is either 0x04 or 0x08.

o PW-ACH BFD Encapsulation (BFD without IP/UDP Headers)

In the second method, a BFD Control packet (format defined in Section 4 of [I-D.ietf-bfd-base]) is encapsulated directly in the VCCV control channel (see Sections 6 and 8 of [I-D.ietf-bfd-generic]) and the IP/UDP headers are omitted from the BFD encapsulation. Therefore, to utilize this encapsulation, a pseudowire MUST use the PW Associated Channel Header (PW-ACH) Control Word format for its Control Word (CW) or L2-Specific Sublayer (L2SS, used in L2TPv3).

In this encapsulation, a "raw" BFD Control packet (i.e., a BFD Control packet as defined in Section 4.1 of [I-D.ietf-bfd-base] without IP/UDP Headers) follows directly the PW-ACH. The PW-ACH Channel Type indicates that the Associated Channel carries "raw" BFD. The PW Associated Channel (PWAC) is defined in Section 5 of [RFC4385], and its Channel Type field is used to discriminate the VCCV payload types.

The usage of the PW-ACH on different VCCV CC Types is specified for CC Type 1, Type 2 and Type 3 respectively in Sections 5.1.1, 5.1.2 and 5.1.3 of [RFC5085], and in all cases requires the use of a CW (see Section 7 of [RFC4385]). When VCCV carries PW-ACH-encapsulated BFD (i.e., "raw" BFD), the PW-ACH (Pseudowire CW's or L2SS') Channel Type MUST be set to 0x0007 to indicate "BFD Control, PW-ACH-encapsulated" (i.e., BFD Without IP/UDP Headers, see Section 5.2). This is to allow the identification of the encased BFD payload when demultiplexing the VCCV control channel.

In this encapsulation, the BFD CV Type employed in signaling (if used) is either 0x10 or 0x20.

In summary, for the IP/UDP encapsulation of BFD (BFD with IP/UDP headers), if a PW Associated Channel Header is used, the Channel Type

MUST indicate either IPv4 (0x0021) or IPv6 (0x0057). For the PW-ACH encapsulation of BFD (BFD without IP/UDP headers), the PW Associated Channel Header MUST be used and the Channel Type MUST indicate BFD Control packet (0x0007).

3.3. CV Types for BFD

The CV Type is defined as a bitmask field used to indicate the specific CV Type or Types (i.e., none, one or more) of VCCV packets that may be sent on the VCCV control channel. The CV Types shown in the table below augment those already defined in [RFC5085]. Their values shown in parenthesis represent the numerical value corresponding to the actual bit being set in the CV Type bitfield.

BFD CV Types:

The defined values for the different BFD CV Types for MPLS and L2TPv3 PWs are:

Bit (Value)	Description
Bit 2 (0x04)	BFD IP/UDP-encapsulated, for PW Fault Detection only
Bit 3 (0x08)	BFD IP/UDP-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling
Bit 4 (0x10)	BFD PW-ACH-encapsulated, for PW Fault Detection only
Bit 5 (0x20)	BFD PW-ACH-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling

It should be noted that four BFD CV Types have been defined by combining two types of encapsulation with two types of functionality, see Table 1 in in Section 3.

Given the bidirectional nature of BFD, before selecting a given BFD CV Type capability to be used in dynamically established pseudowires, there MUST be common CV Types in the VCCV 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 remaining BFD CV Types).

The following list enumerates rules, restrictions and clarifications on the usage of BFD CV Types:

1. BFD CV Types used for fault detection and status signaling (i.e., CV Types 0x08 and 0x20) SHOULD NOT be used when a control protocol such as LDP [RFC4447] or L2TPV3 [RFC3931] is available that can signal the AC/PW status to the remote endpoint of the PW. More details can be found in [I-D.ietf-pwe3-oam-msg-map].

2. BFD CV Types used for fault detection only (i.e., CV Types 0x04 and 0x10) can be used whether a protocol that can signal AC/PW status is available or not. This includes both statically provisioned and dynamically signaled pseudowires.

   2.1. In this case, BFD is used exclusively to detect faults on the PW; if it is desired to convey AC/PW fault status, some means other than BFD are to be used. Examples include using LDP status messages when using MPLS as a transport (see Section 5.4 of [RFC4447]), and the Circuit Status AVP in an L2TPv3 SLI message for L2TPv3 (see Section 5.4.5 of [RFC3931]).

3. Pseudowires that do not use a CW or L2SS using the PW Associated Channel Header MUST NOT use the BFD CV Types 0x10 or 0x20 (i.e., PW-ACH encapsulation of BFD, without IP/UDP headers).

   3.1. PWs that use a PW-ACH include CC Type 1 (for both MPLS and L2TPv3 as defined in Sections 5.1.1 and 6.1 of [RFC5085]), and MPLS CC Types 2 and 3 when using a Control Word (as specified in Sections 5.1.2 and 5.1.3 of [RFC5085]). This restriction stems from the fact that the encapsulation uses the Channel Type in the PW-ACH.

   3.2. PWs that do not use a PW-ACH can use the VCCV BFD encapsulation with IP/UDP headers, as the only VCCV BFD encapsulation supported. Using the IP/UDP encapsulated BFD CV Types allows for the concurrent use of other VCCV CV Types that uses an encapsulation with IP headers (e.g., ICMP Ping or LSP Ping defined in [RFC5085]).

4. Only a single BFD CV Type can be selected and used. All BFD CV Types are mutually exclusive. After selecting a BFD CV Type, a node MUST NOT use any of the other three BFD CV Types.

5. Once a PE has chosen a single BFD CV Type to use, it MUST continue using it until such time as the PW is re-signaled. In order to change the negotiated and selected BFD CV Type, the PW must be torn-down and re-established.

4. Capability Selection

The precedence rules for selection of various CC and CV Types is clearly outlined in Section 7 of [RFC5085]. This section augments these rules when the BFD CV Types defined herein are supported. The selection of a specific BFD CV Type to use out of the four available CV Types defined is tied to multiple factors, as described in

Section 3.3. 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 multiple BFD CV Types are advertised, and after applying the rules in Section 3.3, the set that both ends of the pseudowire have in common is determined. If the two ends have more than one BFD CV Type in common, the following list of BFD CV Types is considered in the order lowest list number CV Type to highest list number CV Type, and the CV Type with the lowest list number is used:

1. 0x20 - BFD PW-ACH-encapsulated (without IP/UDP headers), for PW Fault Detection and AC/PW Fault Status Signaling

2. 0x10 - BFD PW-ACH-encapsulated (without IP/UDP headers), for PW Fault Detection only

3. 0x08 - BFD IP/UDP-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling

4. 0x04 - BFD IP/UDP-encapsulated, for PW Fault Detection only

This precedence order prioritizes superset of functionality and simplicity of encapsulation.

5. IANA Considerations

5.1. MPLS CV Types for the VCCV Interface Parameters Sub-TLV

The VCCV Interface Parameters Sub-TLV codepoint is defined in [RFC4446], and the VCCV CV Types registry is defined in [RFC5085]. This section lists the new BFD CV Types.

IANA is requested to augment the "VCCV Connectivity Verification Types" registry in the Pseudo Wires Name Spaces, reachable from [IANA.pwe3-parameters]. These are bitfield values. CV Type values 0x04 0x08, 0x10 and 0x20 are specified in Section 3 of this document.

MPLS Connectivity Verification (CV) Types:

Bit (Value)	Description
Bit 2 (0x04)	BFD IP/UDP-encapsulated, for PW Fault Detection only
Bit 3 (0x08)	BFD IP/UDP-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling
Bit 4 (0x10)	BFD PW-ACH-encapsulated, for PW Fault Detection only
Bit 5 (0x20)	BFD PW-ACH-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling

5.2. PW Associated Channel Type

The PW Associated Channel Types used by VCCV rely on previously allocated numbers from the Pseudowire Associated Channel Types Registry [RFC4385] in the Pseudo Wires Name Spaces reachable from [IANA.pwe3-parameters].

IANA is requested to reserve a new Pseudowire Associated Channel Type value as follows:

Value (in hex)	Protocol Name	Reference
0x0007	BFD Control, PW-ACH encapsulation (without IP/UDP Headers)	[This document]

5.3. L2TPv3 CV Types for the VCCV Capability AVP

This section lists the new BFD CV Types to be added to the existing "VCCV Capability AVP" registry in the L2TP name spaces. The Layer Two Tunneling Protocol "L2TP" Name Spaces are reachable from [IANA.l2tp-parameters].

IANA is requested to reserve the following L2TPv3 Connectivity Verification (CV) Types in the VCCV Capability AVP Values registry.

VCCV Capability AVP (Attribute Type AVP-TBD) Values

L2TPv3 Connectivity Verification (CV) Types:

Bit (Value)	Description
Bit 2 (0x04)	BFD IP/UDP-encapsulated, for PW Fault Detection only
Bit 3 (0x08)	BFD IP/UDP-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling
Bit 4 (0x10)	BFD PW-ACH-encapsulated, for PW Fault Detection only
Bit 5 (0x20)	BFD PW-ACH-encapsulated, for PW Fault Detection and AC/PW Fault Status Signaling

6. Congestion Considerations

The congestion considerations that apply to [RFC5085] apply to this mode of operation as well. There are no additional congestion considerations.

7. Security Considerations

Routers that implement the additional CV Types defined herein are subject to the same security considerations as defined in [RFC5085], [I-D.ietf-bfd-base], and [I-D.ietf-bfd-v4v6-1hop]. This specification does not raise any additional security issues beyond these. The IP/UDP-encapsulated BFD makes use of the TTL/Hop Limit procedures described in Section 5 of [I-D.ietf-bfd-v4v6-1hop], including the use of the Generalized TTL Security Mechanism (GTSM) as a security mechanism.

8. Acknowledgements

This work forks from a previous revision of the PWE3 WG document that resulted in [RFC5085], to which a number of people contributed, including Rahul Aggarwal, Peter B. Busschbach, Yuichi Ikejiri, Kenji Kumaki, Luca Martini, Monique Morrow, George Swallow, and others.

Mustapha Aissaoui, Sam Aldrin, Stewart Bryant, Peter B. Busschbach, Annamaria Fulignoli, Vishwas Manral, Luca Martini, Dave McDysan, Ben Niven-Jenkins, Pankil Shah, Yaakov Stein, and George Swallow provided useful feedback and valuable comments and suggestions improving newer versions of this document.

9. References

9.1. Normative References

[I-D.ietf-bfd-base]
Katz, D. and D. Ward, "Bidirectional Forwarding Detection", draft-ietf-bfd-base-09 (work in progress), February 2009.

[I-D.ietf-bfd-generic]
Katz, D. and D. Ward, "Generic Application of BFD", draft-ietf-bfd-generic-05 (work in progress), February 2009.

[I-D.ietf-bfd-v4v6-1hop]
Katz, D. and D. Ward, "BFD for IPv4 and IPv6 (Single Hop)", draft-ietf-bfd-v4v6-1hop-09 (work in progress), February 2009.

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

[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.

[RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

9.2. Informative References

[I-D.ietf-pwe3-oam-msg-map]
Martini, L. and M. Morrow, "Pseudo Wire (PW) OAM Message Mapping", draft-ietf-pwe3-oam-msg-map-10 (work in progress), April 2009.

[IANA.l2tp-parameters]
Internet Assigned Numbers Authority, "Layer Two Tunneling Protocol "L2TP"", May 2009, <http://www.iana.org/assignments/l2tp-parameters>.

[IANA.pwe3-parameters]
Internet Assigned Numbers Authority, "Pseudowire Name Spaces (PWE3)", April 2009, <http://www.iana.org/assignments/pwe3-parameters>.

[RFC3931] Lau, J., Townsley, M., and I. Goyret, "Layer Two Tunneling

Protocol - Version 3 (L2TPv3)", RFC 3931, March 2005.

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

Authors' Addresses

Thomas D. Nadeau (editor)
BT
BT Centre
81 Newgate Street
London, EC1A 7AJ
United Kingdom

Email: tom.nadeau@bt.com

Carlos Pignataro (editor)
Cisco Systems, Inc.
7200 Kit Creek Road
PO Box 14987
Research Triangle Park, NC 27709
USA

Email: cpignata@cisco.com