

Internet Engineering Task Force (IETF)
Request for Comments: 5885
Category: Standards Track
ISSN: 2070-1721

T. Nadeau, Ed.
BT
C. Pignataro, Ed.
Cisco Systems, Inc.
June 2010

Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)

Abstract

This document describes 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.

Status of This Memo

This is an Internet Standards Track document.

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

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc5885>.

Copyright Notice

Copyright (c) 2010 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
2. Specification of Requirements	3
3. Bidirectional Forwarding Detection Connectivity Verification	3
3.1. BFD CV Type Operation	4
3.2. BFD Encapsulation	5
3.3. CV Types for BFD	7
4. Capability Selection	9
5. IANA Considerations	10
5.1. MPLS CV Types for the VCCV Interface Parameters Sub-TLV	10
5.2. PW Associated Channel Type	10
5.3. L2TPv3 CV Types for the VCCV Capability AVP	11
6. Congestion Considerations	11
7. Security Considerations	12
8. Acknowledgements	12
9. References	12
9.1. Normative References	12
9.2. Informative References	13

1. Introduction

This document describes 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 [RFC5880] is used over the VCCV control channel primarily as a pseudowire fault detection mechanism, for detecting data-plane 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 Operations, Administration, and Maintenance (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 [OAM-MSG-MAP].

The new BFD CV Types are PW demultiplexer-agnostic, and hence applicable for both MPLS and Layer Two Tunneling Protocol version 3 (L2TPv3) pseudowire demultiplexers. This document concerns itself with the BFD VCCV operation over single-segment pseudowires (SS-PWs). This specification describes procedures only for BFD asynchronous mode.

2. 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].

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 versus without IP/UDP headers) and by functionality (i.e., fault detection only versus 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) [RFC5880] 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 being 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 Sections 3.3 and 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:

- o 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 [RFC5881] and Section 6 of [RFC5882]).

- o 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.
- o BFD MUST be run in asynchronous mode (see [RFC5880]).

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 [RFC5880]. 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 [RFC5880]), 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 [RFC5880]), 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 defined in [OAM-MSG-MAP].

3.2. BFD Encapsulation

The VCCV message comprises a BFD Control packet [RFC5880] 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, which 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 [RFC5881]. BFD Control packets are therefore transmitted in UDP with destination port 3784 and source port within the range 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 [RFC5881] apply to this encapsulation, and hence the TTL/Hop Limit is set to 255.

If the PW is established by signaling, then the BFD CV Type used for this encapsulation 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 [RFC5880]) is encapsulated directly in the VCCV control channel (see Sections 6 and 8 of [RFC5882]) 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 (see [RFC5586]) 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 [RFC5880] 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.

If the PW is established by signaling, then the BFD CV Type used for this encapsulation 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 parentheses 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 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 [OAM-MSG-MAP].
2. BFD CV Types used for fault detection only (i.e., CV Types 0x04 and 0x10) can be used whether or not a protocol that can signal AC/PW status is available. 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 Attribute Value Pair (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 use 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 when 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 of the lowest list number CV Type to the 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

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 has augmented the "VCCV Connectivity Verification (CV) Types" registry in the Pseudowire Name Spaces reachable from [IANA]. 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 Pseudowire Name Spaces reachable from [IANA].

IANA has reserved a new Pseudowire Associated Channel Type value as follows:

Registry:

Value	Description	TLV Follows	Reference
-----	-----	-----	-----
0x0007	BFD Control, PW-ACH encapsulation (without IP/UDP Headers)	No	[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].

IANA has reserved the following L2TPv3 Connectivity Verification (CV) Types in the VCCV Capability AVP Values registry.

VCCV Capability AVP (Attribute Type 96) 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. This section describes explicitly how they apply.

BFD as a VCCV application is required to provide details on congestion and bandwidth considerations. BFD provides with a desired minimum transmit interval and a required minimum receive interval, negotiates the transmission interval using these configurable fields, and has a packet of fixed size (setting the transmission rate). Therefore, it results in a configuration limited bandwidth utilization. As stated in [RFC5085], this is sufficient protection against congestion as long as BFD's configured maximum bit-rate is minimal compared to the bit-rate of the pseudowire the VCCV channel is associated with. If the pseudowire bit-rate can't be guaranteed to be minimal, like potentially for highly variable bit-rate and/or congestion responsive pseudowires, BFD will be required to operate using an adaptive congestion control mechanism (for example, including a throttled transmission rate on "congestion detected" situations, and a slow-start after shutdown due to congestion and until basic connectivity is verified).

Since the bandwidth utilized by BFD is configuration-limited, the VCCV channel **MUST NOT** be rate-limited below this maximum configurable bandwidth or BFD will not operate correctly. The VCCV channel could provide rate-limiting above the maximum BFD rate, to protect from a misbehaving BFD application, so that it does not conflict and can coexist. Additionally, the VCCV channel **SHOULD NOT** use any additional congestion control loop that would interfere or negatively interact with that of BFD. 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], [RFC5880], and [RFC5881]. 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 [RFC5881], 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

- [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.

- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection", RFC 5880, June 2010.
- [RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881, June 2010.
- [RFC5882] Katz, D. and D. Ward, "Generic Application of Bidirectional Forwarding Detection (BFD)", RFC 5882, June 2010.

9.2. Informative References

- [IANA] Internet Assigned Numbers Authority, "Protocol Registries", <<http://www.iana.org>>.
- [OAM-MSG-MAP] Aissaoui, M., Busschbach, P., Morrow, M., Martini, L., Stein, Y., Allan, D., and T. Nadeau, "Pseudowire (PW) OAM Message Mapping", Work in Progress, March 2010.
- [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.
- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", RFC 5586, June 2009.

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