MPLS Working Group Internet Draft Intended status: Standards Track Expires: February 2012 Dave Allan, Ed. Ericsson

George Swallow Ed. Cisco Systems, Inc

> John Drake Ed. Juniper

> > August 2011

Proactive Connectivity Verification, Continuity Check and Remote Defect indication for MPLS Transport Profile draft-ietf-mpls-tp-cc-cv-rdi-06

Abstract

Continuity Check, Proactive Connectivity Verification and Remote Defect Indication functionalities are required for MPLS-TP OAM.

Continuity Check monitors a label switched path for any loss-ofcontinuity defect. Connectivity Verification augments Continuity Check in order to provide confirmation that the desired source is connected to the desired sink. Remote defect indication enables an End Point to report, to its associated End Point, a fault or defect condition that it detects on a pseudo wire, label switched path or Section.

This document specifies specific extensions to BFD and methods for proactive Continuity Check, Continuity Verification, and Remote Defect Indication for MPLS-TP label switched paths, pseudo wires and Sections using Bidirectional Forwarding Detection as extended by this memo.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in <u>RFC2119</u> [1].

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of <u>BCP 78</u> and <u>BCP 79</u>.

Allan et al., Expires February 2012

[Page 1]

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on February 9th, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

<u>1</u> . Introduction <u>3</u>
<u>2</u> . Conventions used in this document <u>4</u>
<u>2.1</u> . Terminology <u>4</u>
<u>3</u> . MPLS-TP CC, proactive CV and RDI Mechanism using BFD <u>5</u>
<u>3.1</u> . Existing Capabilities <u>5</u>
<u>3.2</u> . CC, CV, and RDI Overview <u>6</u>
<u>3.3</u> . ACH code points for CC and proactive CV
<u>3.4</u> . MPLS-TP BFD CC Message format <u>7</u>
<u>3.5</u> . MPLS-TP BFD proactive CV Message format
<u>3.5.1</u> . Section MEP-ID <u>9</u>
<u>3.5.2</u> . LSP MEP-ID <u>10</u>

<u>3.5.3</u> . PW Endpoint MEP-ID <u>10</u>
3.6. BFD Session in MPLS-TP terminology11
<u>3.7</u> . BFD Profile for MPLS-TP <u>12</u>
<u>3.7.1</u> . Session initiation and Modification <u>13</u>
<u>3.7.2</u> . Defect entry criteria <u>13</u>
<u>3.7.3</u> . Defect entry consequent action <u>15</u>
<u>3.7.4</u> . Defect exit criteria <u>15</u>
<u>3.7.5</u> . State machines <u>15</u>
3.7.6. Configuration of MPLS-TP BFD sessions <u>18</u>
<u>3.7.7</u> . Discriminator values <u>18</u>
<u>4</u> . Configuration Considerations <u>18</u>
<u>5</u> . Acknowledgments <u>19</u>
<u>6</u> . IANA Considerations <u>19</u>
<u>7</u> . Security Considerations
<u>8</u> . References
<u>8.1</u> . Normative References <u>20</u>
<u>8.2</u> . Informative References <u>21</u>

1. Introduction

In traditional transport networks, circuits are provisioned on two or more switches. Service Providers (SP) need Operations, Administration and Maintenance (OAM) tools to detect mis-connectivity and loss-ofcontinuity of transport circuits. Both pseudo wires (PWs) and MPLS-TP label switched paths (LSPs) [12] emulating traditional transport circuits need to provide the same Continuity Check (CC), proactive Continuity Verification (CV), and Remote Defect Indication (RDI) capabilities as required in RFC 5860[3]. This document describes the use of Bidirectional Forwarding Detection (BFD)[4] for CC, proactive CV, and RDI of a PW, LSP or sub-path maintenance entity (SPME) between two Maintenance Entity Group End Points (MEPs).

As described in [13], CC and CV functions are used to detect loss-ofcontinuity (LOC), and unintended connectivity between two MEPs (e.g. mis-merging or mis-connectivity or unexpected MEP).

RDI is an indicator that is transmitted by a MEP to communicate to its peer MEP that a signal fail condition exists. RDI is only used for bidirectional LSPs and is associated with proactive CC & CV BFD control packet generation.

This document specifies the BFD extension and behavior to satisfy the CC, proactive CV monitoring and the RDI functional requirements for both co-routed and associated bi-directional LSPs. Supported encapsulations include generic alert label (GAL)/generic associated

channel (G-ACh), virtual circuit connectivity verification (VCCV) and UDP/IP. Procedures for uni-directional point-to-point (p2p) and point-to-multipoint (p2mp) LSPs are for further study.

This document utilizes identifiers for MPLS-TP systems as defined in [9]. Work is on-going in the ITU-T to define a globally-unique semantic for ITU Carrier Codes (ICCs), and future work may extend this document to utilize ICCs as identifiers for MPLS-TP systems.

The mechanisms specified in this document are restricted to BFD asynchronous mode.

2. Conventions used in this document

<u>2.1</u>. Terminology

- ACH: Associated Channel Header
- BFD: Bidirectional Forwarding Detection
- CC: Continuity Check
- CV: Connectivity Verification
- GAL: G-ACh Label
- G-ACh: Generic Associated Channel
- LDI: Link Down Indication
- LKI: Lock Instruct
- LKR: Lock Report
- LSP: Label Switched Path
- LSR: Label Switching Router
- ME: Maintenance Entity
- MEG: Maintenance Entity Group
- MEP: Maintenance Entity Group End Point
- MIP: Maintenance Entity Group Intermediate Point
- Allan et al., Expires February 2012

Internet-Draft draft-ietf-mpls-tp-cc-cv-rdi-06

August 2011

MPLS: Multi-Protocol Label Switching

MPLS-OAM: MPLS Operations, Administration and Maintenance

MPLS-TP: MPLS Transport Profile

MPLS-TP LSP: Uni-directional or Bidirectional Label Switched Path representing a circuit

MS-PW: Multi-Segment PseudoWire

NMS: Network Management System

OAM: Operations, Administration, and Maintenance [14]

PW: Pseudo Wire

PDU: Protocol Data Unit

P/F: Poll-Final

RDI: Remote Defect Indication

SPME: Sub-Path Maintenance Entity

TTL: Time To Live

TLV: Type Length Value

VCCV: Virtual Circuit Connectivity Verification

3. MPLS-TP CC, proactive CV and RDI Mechanism using BFD

This document describes procedures for achieve combined CC, CV and RDI functionality within a single MPLS-TP MEG using BFD. This augments the capabilities that can be provided for MPLS-TP LSPs using existing specified tools and procedures.

<u>3.1</u>. Existing Capabilities

A CC-only mode may be provided via protocols and procedures described in $\frac{\text{RFC 5885}[7]}{\text{Mith ACH channel 7}}$. These procedures may be applied to bi-directional LSPs (via the use of the GAL), as well as PWs.

Implementations may also interoperate with legacy equipment by implementing <u>RFC 5884[8]</u> for LSPs and <u>RFC 5085[10]</u> for PWs, in addition to the procedures documented in this memo. In accordance with <u>RFC 5586[2]</u>, when BFD control packets are encapsulated in an IP

Allan et al., Expires February 2012 [Page 5]

August 2011

header, the fields in the IP header are set as defined in <u>RFC</u> <u>5884[8]</u>. When IP encapsulation is used CV mis-connectivity defect detection can be performed by inferring a globally unique source on the basis of the combination of the source IP address and My Discriminator" fields.

3.2. CC, CV, and RDI Overview

The combined CC, CV, and RDI functionality for MPLS-TP is achieved by multiplexing CC and CV PDUs within a single BFD session. The CV PDUs are augmented with a source MEP ID TLV to permit mis-connectivity detection to be performed by sink MEPs.

The interleaving of PDUs is achieved via the use of distinct encapsulations and code points for generic associated channel (G-ACh) encapsulated BFD depending on whether the PDU format is CC or CV:

- CC format: defines a new code point in the Associated Channel Header (ACH) described in <u>RFC 5586[2]</u>.This format supports Continuity Check and RDI functionalities.
- o CV format: defines a new code point in the Associated Channel Header (ACH) described in <u>RFC 5586[2]</u>. The ACH with "MPLS-TP Proactive CV" code point indicates that the message is an MPLS-TP BFD proactive CV message, and information for CV processing is appended in the form of the source MEP ID TLV.

RDI is communicated via the BFD diagnostic field in BFD CC messages, and the diagnostic code field in CV messages MUST be ignored. It is not a distinct PDU. As per [4], a sink MEP SHOULD encode a diagnostic code of "1 - Control detection time expired" when the interval times detect multiplier have been exceeded. A sink MEP SHOULD encode a diagnostic code of "5 - Path Down" as a consequence of the sink MEP receiving LDI. A sink MEP MUST encode a diagnostic code of "XX -

Misconnectivity defect" (to be assigned by IANA) when CV PDU processing indicates a misconnectivity defect. A sink MEP that has started sending diagnostic code 5 SHOULD NOT change it to 1 when the detection timer expires.

Allan et al., Expires February 2012 [Page 6]

Internet-Draft draft-ietf-mpls-tp-cc-cv-rdi-06

3.3. ACH code points for CC and proactive CV

Figure 1 illustrates the G-ACh encoding for BFD CC-CV-RDI functionality.

0 2 1 3 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 | BFD CC/CV Code Point |0 0 0 1|Version| Flags

Figure 1: ACH Indication of MPLS-TP CC/CV/RDI

The first nibble (0001b) indicates the G-ACh as per <u>RFC 5586[2]</u>.

The version and the flags are set to 0 as specified in [2].

The code point is either

- BFD CC code point = XX1. [XX1 to be assigned by IANA from the PW Associated Channel Type registry.] or,

- BFD proactive CV code point = XX2. [XX2 to be assigned by IANA from the PW Associated Channel Type registry.]

CC and CV PDUs apply to all pertinent MPLS-TP structures, including PWs, MPLS LSPs (including SPMEs), and Sections.

CC and CV operation is simultaneously employed on a maintenance entity (ME) within a single BFD session. The expected usage is that normal operation is to send CC BFD protocol data units (PDUs) interleaved with a CV BFD PDU (augmented with a source MEP-ID and identified as requiring additional processing by the different ACh channel type). The insertion interval for CV PDUs is one per second. Detection of a loss-of-continuity defect is the detect multiplier (fixed at 3 for the CC code point) times the session periodicity. Mis-connectivity defects are detected in a maximum of one second. 3.4. MPLS-TP BFD CC Message format

The format of an MPLS-TP CC Message is shown below.

Allan et al., Expires February 2012

[Page 7]

0 2 1 3 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 0 0 0 1|Version| Flags | BFD CC Code point BFD Control Packet ~ ~ T Figure 2: MPLS-TP CC Message

As shown in figure 2, the MPLS-TP CC message consists of the BFD control packet as defined in [4] pre-pended by the G-ACh.

3.5. MPLS-TP BFD proactive CV Message format

The format of an MPLS-TP CV Message is shown below.

0 1 2 3 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 0 0 0 1|Version| Flags | BFD CV Code Point BFD Control Packet \sim ~ MEP Source ID TLV ~

Figure 3: MPLS-TP CV Message

As shown in figure 3, the MPLS-TP CV message consists of the BFD control packet as defined in [4] pre-pended by the ACH, and appended by a MEP source ID TLV.

A MEP Source ID TLV is encoded as a 2 octet field that specifies a Type, followed by a 2 octet Length Field, followed by a variable length Value field. A BFD session will only use one encoding of the Source ID TLV.

The length in the BFD control packet is as per [4]; the length of the MEP Source ID TLV is not included. There are 3 possible Source MEP

Allan et al., Expires February 2012 [Page 8]

TLVs (corresponding to the MEP-IDs defined in [9]) [type fields to be assigned by IANA]. The type fields are:

X1 - Section MEP-ID

X2 - LSP MEP-ID

X3 - PW MEP-ID

When GAL label is used, the time to live (TTL) field of the GAL MUST be set to at least 1, and the GAL MUST be the end of stack label (S=1) as per [<u>2</u>].

A node MUST NOT change the value in the MEP Source ID TLV.

When digest based authentication is used, the Source ID TLV MUST NOT be included in the digest

3.5.1. Section MEP-ID

The IP compatible MEP-IDs for MPLS-TP sections is the interface ID. The format of the Section MEP-ID TLV is:

0	1	2	3		
0123456789	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1		
+-	+ - + - + - + - + - + - + - + - + - + -	+ - + - + - + - + - + - + - + - + - + -	+-+-+		
Туре	Lengt	h			
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++-	+ - + - + - + - + - + - + - + - + - + -	+ - + - + - + - + - + - + - + - + - + -	+-+-+		
MPLS-TP Global_ID					
+-					
MPLS-TP Node Identifier					
+-					
MPLS-TP Interface Number					
+-					

Figure 4: Section MEP-ID TLV format

Where the type is of value 'X1' (to be assigned by IANA). The length is the length of the value fields. The MPLS-TP Global ID, Node Identifier and Interface Numbers are as per [9].

[Page 9] Allan et al., Expires February 2012

3.5.2. LSP MEP-ID

The fields for the LSP MEP-ID is as defined in [9]. This is applicable to both LSPs and SPMEs. This consists of 32-bit MPLS-TP Global ID, the 32-bit Node Identifier, followed by the 16-bit Tunnel_Num (that MUST be unique within the context of the Node Identifier), and the 16-bit LSP_NUM (that MUST be unique with the context of the Tunnel Num). The format of the TLV is:

Θ	1	2	3		
0123456789	0 1 2 3 4 5 6	78901234	15678901		
+-	-+-+-+-+-+-+-+	+ - + - + - + - + - + - + - + - + - + -	+ - + - + - + - + - + - + - +		
Туре	1	Length			
+-	-+-+-+-+-+-+-+	+ - + - + - + - + - + - + - + - + - + -	+ - + - + - + - + - + - + - +		
	MPLS-TP Glo	bal_ID	1		
+-					
	MPLS-TP Node]	Identifier	1		
+-					
Tunnel_M	Num	LSP_1	lum		
+-					

Figure 5: LSP MEP-ID TLV format

Where the type is of value 'X2' (to be assigned by IANA). The length is the length of the value fields. The MPLS-TP Global ID, Node Identifier, Tunnel Num and LSP_Num are as per [9].

3.5.3. PW Endpoint MEP-ID

The fields for the MPLS-TP PW Endpoint MEP-ID is as defined in [9]. The format of the TLV is:

0 1 2 3 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 Allan et al., Expires February 2012 [Page 10]

| Type | Length MPLS-TP Global_ID MPLS-TP Node Identifier AC ID AGI Type | AGI Length | AGI Value AGI Value (contd.)

Figure 6: PW Endpoint MEP-ID TLV format

Where the type is value 'X3' (to be assigned by IANA). The length is the length of the following data. The Global ID, Node Identifier and Attachment Circuit (AC)_ID are as per [9]. The Attachment Group Identifier (AGI) Type is as per $[\underline{6}]$, and the AGI length is the length of the AGI value field.

3.6. BFD Session in MPLS-TP terminology

A BFD session corresponds to a CC and proactive CV OAM instance in MPLS-TP terminology. A BFD session is enabled when the CC and proactive CV functionality is enabled on a configured Maintenance Entity (ME).

When the CC and proactive CV functionality is disabled on an ME, the BFD session transitions to the ADMIN DOWN State and the BFD session ends.

A new BFD session is initiated when the operator enables or reenables the CC and CV functionality.

All BFD state changes and P/F exchanges MUST be done using CC packets. P/F and session state information in CV packets MUST be ignored.

Allan et al., Expires February 2012 [Page 11]

3.7. BFD Profile for MPLS-TP

BFD operates in asynchronous mode utilizing the encapsulation defined in <u>section 3</u> for all sessions in a given MEG. For LSPs, SPMEs and sections this is GAL/G-ACh encapsulated BFD using the code points specified in <u>section 3.3</u>. For PWs, this is G-ACh or GAL/G-ACh encapsulated BFD using the code points specified in <u>section 3.3</u>. In this mode, the BFD Control packets are periodically sent at configurable time rate. This rate is a fixed value common for both directions of MEG for the lifetime of the MEG.

This document specifies bi-directional BFD for p2p transport LSPs; hence all BFD packets MUST be sent with the M bit clear.

There are two modes of operation for bi-directional LSPs. One in which the session state of both directions of the LSP is coordinated and one constructed from BFD sessions in such a way that the two directions operate independently but are still part of the same MEG. A single bi-directional BFD session is used for coordinated operation. Two independent BFD sessions are used for independent operation. It should be noted that independent operation treats session state and defect state as independent entities. For example an independent session can be in the UP state while receiving RDI indication. For a coordinated session, the session state will track the defect state.

In coordinated mode, an implementation SHOULD NOT reset bfd.RemoteDiscr until it is exiting the DOWN state.

In independent mode, an implementation MUST NOT reset bfd.RemoteDiscr upon transitioning to the DOWN state.

Overall operation is as specified in [4] and augmented for MPLS in [8]. Coordinated operation is as described in [4]. Independent operation requires clarification of two aspects of [4]. Independent operation is characterized by the setting of bfd.MinRxInterval to zero by the MEP that is typically the session originator (referred to as the source MEP), and there will be a session originator at either end of the bi-directional LSP. Each source MEP will have a corresponding sink MEP that has been configured to a Tx interval of zero.

This memo specifies a preferred interpretation of the base spec on how a MEP with a BFD transmit rate set to zero behaves. One interpretation is that no periodic messages on the reverse component of the bi-directional LSP originate with that MEP, it will only originate messages on a state change.

Allan et al., Expires February 2012 [Page 12]

draft-ietf-mpls-tp-cc-cv-rdi-06 August 2011 Internet-Draft

The first clarification is that when a state change occurs a MEP set to a transmit rate of zero sends BFD control messages with a one second period on the reverse component until such time that the state change is confirmed by the session peer. At this point the MEP set to a transmit rate of zero can resume quiescent behavior. This adds robustness to all state transitions in the RxInterval=0 case.

The second is that the originating MEP (the one with a non-zero bfd.TxInterval) will ignore a DOWN state received from a zero interval peer. This means that the zero interval peer will continue to send DOWN state messages that include the RDI diagnostic code as the state change is never confirmed. This adds robustness to the exchange of RDI indication on a uni-directional failure (for both session types DOWN with a diagnostic of either control detection period expired or neighbor signaled session down offering RDI functionality).

A further extension to the base specification is that there are additional OAM protocol exchanges that act as inputs to the BFD state machine; these are the Link Down Indication [5] and the Lock Instruct/Lock Report transactions; Lock Report interaction being optional.

3.7.1. Session initiation and Modification

Session initiation occurs starting from MinRx = 1 second, MinTx >= 1 second and the detect multiplier = 3. Once in the UP state, poll/final discipline is used to modify the periodicity of control message exchange from their default rates to the desired rates and set the detect multiplier to 3. Once the desired rate has been reached using the poll/final mechanism, implementations SHOULD NOT attempt further rate modification.

In the rare circumstance where an operator has a reason to further change session parameters, beyond the initial migration from default values; poll/final discipline can be used with the caveat that a peer implementation may consider a session change unacceptable and/or bring the BFD session down via the use of the ADMIN DOWN state.

3.7.2. Defect entry criteria

There are further defect criteria beyond those that are defined in [4] to consider given the possibility of mis-connectivity defects. The result is the criteria for a LSP direction to transition from the

Allan et al., Expires February 2012 [Page 13] defect free state to a defect state is a superset of that in the BFD base specification $[\underline{4}]$.

The following conditions cause a MEP to enter the defect state for CC PDUs (in no particular order):

- 1. BFD session times out (loss-of-continuity defect).
- 2. Receipt of a link down indication or lock report.

And the following will cause the MEP to enter the mis-connectivity defect state for CV operation (again not in any particular order):

- 1. BFD control packets are received with an unexpected encapsulation (mis-connectivity defect), these include: - receiving an IP encoded CC or CV BFD control packet on a LSP configured to use GAL/G-ACh, or vice versa (note there are other possibilities that can also alias as an OAM packet)
- 2. Receipt of an unexpected globally unique Source MEP identifier (Mis-connectivity defect). Note that as each encoding of the Source MEP ID TLV contains unique information (there is no mechanical translation possible between MEP ID formats), receipt of an unexpected source MEP ID type is the same as receiving an unexpected value.
- 3. Receipt of a session discriminator that is not in the local BFD database in the Your Discriminator field (mis-connectivity defect).
- 4. Receipt of a session discriminator that is in the local database but does not have the expected label (mis-connectivity defect).
- 5. IF BFD authentication is used, receipt of a message with incorrect authentication information (password, MD5 digest, or SHA1 hash).

The effective defect hierarchy (order of checking) is

- 1. Receiving nothing.
- 2. Receiving link down indication. E.g. a local link failure, an MPLS-TP LDI indication, or Lock Report.
- 3. Receiving from an incorrect source (determined by whatever means).

Allan et al., Expires February 2012 [Page 14]

- Receiving from a correct source (as near as can be determined), but with incorrect session information).
- 5. Receiving BFD control packets in all discernable ways correct.

<u>3.7.3</u>. Defect entry consequent action

Upon defect entry a sink MEP will assert signal fail into any client (sub-)layers. It will also communicate session DOWN to its session peer using CC messages.

The blocking of traffic as a consequent action MUST be driven only by a defect's consequent action as specified in [13] section 5.1.1.2.

When the defect is mis-connectivity, the section, LSP or PW termination will silently discard all non-OAM traffic received. The sink MEP will also send a defect indication back to the source MEP via the use of a diagnostic code of mis-connectivity defect to be assigned by IANA.

3.7.4. Defect exit criteria

3.7.4.1. Exit from a loss-of-continuity defect

For a coordinated session, exit from a loss-of-connectivity defect is as described in figure 7 which updates $[\underline{4}]$.

For an independent session, exit from a loss-of-connectivity defect occurs upon receipt of a well formed BFD control packet from the peer MEP as described in figures 8 and 9.

3.7.4.2. Exit from a mis-connectivity defect

Exit from a mis-connectivity defect state occurs when no CV messages with mis-connectivity defects have been received for a period of 3.5 seconds.

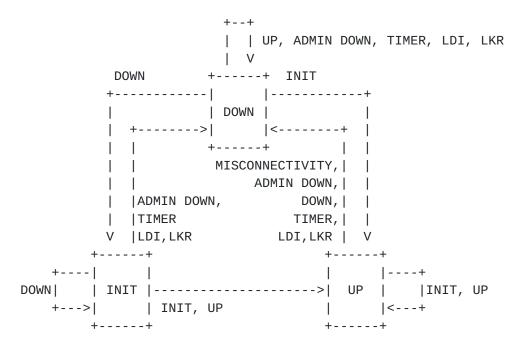
<u>3.7.5</u>. State machines

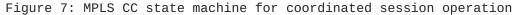
The following state machines update $[\underline{4}]$. They have been modified to include LDI and LKR as specified in $[\underline{5}]$ as inputs to the state machine and to clarify the behavior for independent mode. LKR is an optional input.

The coordinated session state machine has been augmented to indicate LDI and optionally LKR as inputs to the state machine. For a session

Allan et al.,	Expires February 2012	[Page 15]
---------------	-----------------------	-----------

that is in the UP state, receipt of LDI or optionally LKR will transition the session into the DOWN state.





For independent mode, there are two state machines. One for the source MEP (which requested bfd.MinRxInterval=0) and the sink MEP (which agreed to bfd.MinRxInterval=0).

The source MEP will not transition out of the UP state once initialized except in the case of a forced ADMIN DOWN. Hence LDI and optionally LKR do not enter into the state machine transition from the UP state, but do enter into the INIT and DOWN states.

> +--+ | UP, ADMIN DOWN, TIMER, LDI, LKR I V DOWN +----+ INIT

Allan et al., Expires February 2012 [Page 16]

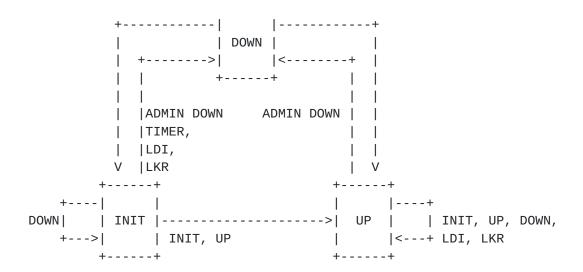


Figure 8: MPLS CC State machine for source MEP for independent session operation

The sink MEP state machine (for which the transmit interval has been set to zero) is modified to:

1) Permit direct transition from DOWN to UP once the session has been initialized. With the exception of via the ADMIN DOWN state, the source MEP will never transition from the UP state, hence in normal unidirectional fault scenarios will never transition to the INIT state.

			+	+					
					ADMIN	DOWN,	TIMER,	LDI,	LKR
			'	V					
D	DOWN	+.			+ INI	ET, UP			
+ -							+		
			DOW	N					
	+	>			<	+			
		+ -			+				

Allan et al., Expires February 2012 [Page 17]

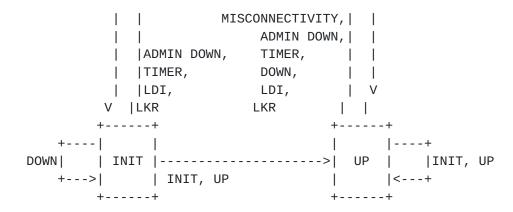


Figure 9: MPLS CC State machine for the sink MEP for independent session operation

3.7.6. Configuration of MPLS-TP BFD sessions

Configuration of MPLS-TP BFD session parameters and coordination of same between the source and sink MEPs is out of scope of this memo.

3.7.7. Discriminator values

In the BFD control packet the discriminator values have either local to the sink MEP or no significance (when not known).

My Discriminator field MUST be set to a nonzero value (it can be a fixed value), the transmitted Your Discriminator value MUST reflect back the received value of My Discriminator field or be set to 0 if that value is not known.

Per <u>RFC5884 Section 7</u> [8], a node MUST NOT change the value of the My Discriminator" field for an established BFD session.

4. Configuration Considerations

The following is an example set of configuration parameters for a BFD session:

Mode and Encapsulation RFC 5884 - BFD CC in UDP/IP/LSP

Allan et al.,	Expires February 2012	[Page 18]
---------------	-----------------------	-----------

RFC 5885 - BFD CC in G-ACh
RFC 5085 - UDP/IP in G-ACh
MPLS-TP - CC/CV in GAL/G-ACh or G-ACh

For MPLS-TP, the following additional parameters need to be configured:

- 1) Session mode, coordinated or independent
- 2) CC periodicity
- 3) The MEP ID for the MEPs at either end of the LSP
- Whether authentication is enabled (and if so, the associated parameters)

And the the discriminators used by each MEP, both bfd.LocalDiscr and bfd.RemoteDiscr can optionally be configured or locally assigned. Finally a detect multiplier of 3 is directly inferred from the code points.

5. Acknowledgments

Nitin Bahadur, Rahul Aggarwal, Tom Nadeau, Nurit Sprecher and Yaacov Weingarten also contributed to this document.

<u>6</u>. IANA Considerations

This draft requires the allocation of two channel types from the IANA "PW Associated Channel Type" registry in <u>RFC4446</u> [6].

XX1 MPLS-TP CC message

XX2 MPLS-TP CV message

This draft requires the creations of a source MEP-ID TLV registry. The parent registry will be the "PW Associated Channel Type" registry of <u>RFC4446</u> [6]. All code points within this registry shall be allocated according to the "Standards Action" procedures as specified in [11].

The initial values are:

X1 - Section MEP-ID

X2 - LSP MEP-ID

Allan et al., Expires February 2012

[Page 19]

X3 - PW MEP-ID

The items tracked in the registry will be the type, and the associated name and reference. The source MEP-ID TLV will require standards action registration procedures for additional values.

This memo requests a code point from the registry for BFD diagnostic codes [4]:

XX -

- mis-connectivity defect

7. Security Considerations

The use of CV improves network integrity by ensuring traffic is not "leaking" between LSPs.

Base BFD foresees an optional authentication section (see [4] <u>section 6.7</u>); that can be applied to this application. Although the source MEP-ID TLV is not included in the BFD authentication digest, there is a chain of trust such that the discriminator associated with the digest is also associated with the expected MEP-ID which will prevent impersonation of CV messages in this application.

This memo specifies the use of globally unique identifiers for MEP IDs. This provides absolutely authoritative detection of persistent leaking of traffic between LSPs. Non-uniqueness can result in undetected leaking in the scenario where two LSPs with common MEP IDs are misconnected. This would be considered to be undesirable but rare, it would also be difficult to exploit for malicious purposes as at a minimum, both a network end point, and a node that was a transit point for the target MEG would need to be compromised.

8. References

8.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, March 1997.
- [2] Bocci, M. et al., "MPLS Generic Associated Channel ", <u>RFC</u> <u>5586</u>, June 2009
- [3] Vigoureux, M., Betts, M. and D. Ward, "Requirements for Operations Administration and Maintenance in MPLS Transport Networks", <u>RFC5860</u>, May 2010

Allan et al., Expires February 2012 [Page 20]

- [4] Katz, D. and D. Ward, "Bidirectional Forwarding Detection", <u>RFC 5880</u>, June 2010
- Swallow, G. et al., "MPLS Fault Management OAM", draft-[5] ietf-mpls-tp-fault-04 (work in progress), April 2011
- [6] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", RFC 4446, April 2006
- Nadeau, T. et al. "Bidirectional Forwarding Detection [7] (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV) ", IETF <u>RFC 5885</u>, June 2010
- Aggarwal, R. et.al., "Bidirectional Forwarding Detection [8] (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, June 2010
- [9] Bocci, M. and G. Swallow, "MPLS-TP Identifiers", draftietf-mpls-tp-identifiers-06 (work in progress), June 2011
- [10] Nadeau, T, et al., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", <u>RFC 5085</u>, December 2007
- [11] Narten, T., Alvestrand, H., "Guidelines for Writing an IANA Considerations Section in RFCs", IETF RFC 5226, May 2008

8.2. Informative References

- [12] Bocci, M., et al., "A Framework for MPLS in Transport Networks", <u>RFC5921</u>, July 2010
- [13] Allan, D., and Busi, I. "MPLS-TP OAM Framework", draftietf-mpls-tp-oam-framework-11 (work in progress), February 2011
- [14] Andersson et. al., "OAM Terminology", IETF RFC 6291, June 2011

Authors' Addresses

Dave Allan Ericsson Email: david.i.allan@ericsson.com

John Drake Juniper Email: jdrake@juniper.net

George Swallow Cisco Systems, Inc. Email: swallow@cisco.com

Annamaria Fulignoli Ericsson Email: annamaria.fulignoli@ericsson.com

Sami Boutros Cisco Systems, Inc. Email: sboutros@cisco.com

Martin Vigoureux Alcatel-Lucent Email: martin.vigoureux@alcatel-lucent.com

Siva Sivabalan Cisco Systems, Inc. Email: msiva@cisco.com

David Ward Juniper Email: dward@juniper.net

Robert Rennison ECI Telecom Email: robert.rennison@ecitele.com