MPLS Working Group
Internet Draft
Intended status: Standards Track
Expires: November 2010

Dave Allan, Ed.
Ericsson

George Swallow Ed.
Cisco Systems, Inc

John Drake Ed.
Juniper

May 11, 2010

Proactive Connection Verification, Continuity Check and Remote Defect indication for MPLS Transport Profile
draft-asm-mpls-tp-bfd-cc-cv-04

Abstract

Continuity Check (CC), Proactive Connectivity Verification (CV) and Remote Defect Indication (RDI) functionalities required for are MPLS-TP OAM.

Continuity Check monitors the integrity of the continuity of the path for any loss of continuity defect. Connectivity verification monitors the integrity of the routing of the path between sink and source for any connectivity issues. RDI enables an End Point to report, to its associated End Point, a fault or defect condition that it detects on a PW, LSP or Section.

This document specifies methods for proactive CV, CC, and RDI for MPLS-TP Label Switched Path (LSP), PWs and Sections using Bidirectional Forwarding Detection (BFD).

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 RFC2119 [1].

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

In traditional transport networks, circuits are provisioned on two or more switches. Service Providers (SP) need OAM tools to detect mis-connectivity and loss of continuity of transport circuits. Both PWs and MPLS-TP LSPs [6] emulating traditional transport circuits need to provide the same CC and proactive CV capabilities as required in draft-ietf-mpls-tp-oam-requirements[3]. This document describes the use of BFD for CC, proactive CV, and RDI of a PW, LSP or PST between two Maintenance Entity Group End Points (MEPs).

As described in [8], Continuity Check (CC) and Proactive Connectivity Verification (CV) functions are used to detect loss of continuity (LOC), and unintended connectivity between two MEPs (e.g. mismerging or misconnection or unexpected MEP).

The Remote Defect Indication (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 connections and is associated with proactive CC & CV packet generation.

This document specifies the BFD extension and behavior to satisfy the CC, proactive CV monitoring and the RDI functional requirements for bi-directional paths. Procedures for uni-directional paths are for further study.

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

1.1. Authors

David Allan, John Drake, George Swallow, Annamaria Fulignoli, Sami Boutros, Siva Sivabalan, David Ward, Martin Vigoureux.

2. Conventions used in this document

2.1. Terminology

ACH: Associated Channel Header

BFD: Bidirectional Forwarding Detection

CV: Connection Verification

GAL: Generalized Alert Label

LSR: Label Switching Router

MEG: Maintenance Entity Group

MEP: Maintenance Entity Group End Point

MIP: Maintenance Entity Group Intermediate Point

MPLS-OAM: MPLS Operations, Administration and Maintenance

MPLS-TP: MPLS Transport Profile

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

MS-PW: Multi-Segment PseudoWire

NMS: Network Management System

PW: Pseudo Wire

RDI: Remote Defect Indication.

TTL: Time To Live

TLV: Type Length Value

2.2. Issues for discussion

1) Requirement for additional BFD diagnostic codes

   1. When periodicity of CV cannot be supported

   2. For mis-connectivity defect

2) Do we continue to separate CC and CV as separate functions, or collapse them into a single CC+CV behavior given CV is a superset of CC?

3) Is receipt of an unexpected discriminator really a problem?

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

This document proposes distinct encapsulations and code points for BFD depending on whether the mode of operation is CC or CV:

- o CC mode: defines a new code point in the Associated Channel Header (ACH) described in [2].In this mode Continuity Check and RDI functionalities are supported.

- o CV mode: defines a new code point in the Associated Channel Header (ACH) described in [2]. Under MPLS label stack, the ACH with "MPLS Proactive CV" code point indicates that the message is an MPLS BFD proactive CV and CC message.

- o RDI: is communicated via the BFD state field in BFD CC and CV messages. It is not a distinct PDU.

3.1. ACH code points for CC and proactive CV

```
 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     |0xHH   BFD CC/CV Code Point    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: ACH Indication of MPLS-TP Connection Verification

The first nibble (0001b) indicates the ACH.

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

The code point is either

- BFD CC code point = 0xHH. [HH to be assigned by IANA from the PW Associated Channel Type registry.] or,

- BFD proactive CV code point = 0xHH. [HH to be assigned by IANA from the PW Associated Channel Type registry.]

Both CC and CV modes apply to PWs, MPLS LSPs (including tandem connection monitoring), and Sections.

It's possible to run BFD in CC mode on some transport paths and BFD in CV mode on other transport paths. For a given Maintenance Entity Group (MEG) only one mode can be used. A MEP that is configured to support CC mode and receives CV BFD packets, or vice versa, MUST consider them as an unexpected packet, i.e. detect a mis-connectivity defect.

3.2. MPLS BFD CC Message format

The format of an MPLS CC Message format 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     |    0xHH BFD CC Code point     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                  BFD Control Packet                           ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: MPLS CC Message

3.3. MPLS BFD proactive CV Message format

The format of an MPLS CV Message format is shown below, ACH TLVs [5] MUST precede the BFD control packet.

```
 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     |    0xHH  BFD CV Code Point    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    ACH TLV Header                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~          Unique MEP-ID of source of the BFD packet            ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                  BFD Control Packet                           ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: MPLS CV Message

As shown in Figure 3, BFD Control packet as defined in [4] is transmitted as MPLS labeled packets along with ACH, ACH TLV Header defined in Section 3 of RFC 5586 and one ACH TLV object carrying the unique MEP Identifier of the source of the BFD packet defined in [7]

When GAL label is used, the TTL field of the GAL MUST be set to at least 1, and the GAL will be the end of stack label (S=1).

3.4. BFD Session in MPLS-TP terminology

A BFD session corresponds to a CC or a proactive CV OAM instance in MPLS-TP terminology.

A BFD session is enabled when the CC or proactive CV functionality is enabled on a configured Maintenance Entity (ME) or in the case of an associated bi-directional path, pair of Maintenance Entities.

On a Sink MEP, a BFD session can be in DOWN, INIT or UP state as detailed in [4].

When on a ME the CC or proactive CV functionality is disabled, 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 re-enables the CC or CV functionality on the same ME.

3.5. BFD Profile for MPLS-TP

BFD MUST operate in asynchronous mode. In this mode, the BFD Control packets are periodically sent at configurable time rate. This rate is typically a fixed value for the lifetime of the session. In the rare circumstance where an operator has a reason to change session parameters, poll/final discipline is used.

The transport profile is designed to operate independent of the control plane; hence the C bit SHOULD be set.

This document specifies bi-directional BFD for p2p transport paths, hence the M bit MUST be clear.

There are two modes of operation for bi-directional paths. One in which both directions of the path fate share and one constructed from BFD sessions in such a way that the two directions operate independently. A single bi-directional BFD session is used for fate sharing operation. Two independent BFD sessions are used for independent operation.

Fate sharing operation is as described in [4]. Independent operation requires clarification of two aspects of [4]. Independent operation is characterized by the setting of MinRxInterval to zero by the MEP that is typically the session originator, and there will be a session originator at either end of the bi-directional path.

The base spec is unclear on aspects of how a session with a BFD source set to zero interval behaves. One interpretation is that no periodic messages originate with that source, it will only originate messages on a state change.

The first clarification is that when a state change occurs a zero interval source send BFD control messages with a one second period until such time that the state change is confirmed by the session peer. At this point the zero interval source 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 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 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 control detection period expired offering RDI functionality).

The normal usage is that 1:1 protected paths must use fate sharing, and independent operation applies to 1+1 protected paths.

3.5.1. Session initiation

In all scenarios a BFD session starts with both ends in the DOWN state. DOWN state messages exchanged include the desired Tx and Rx rates for the session. If a node cannot support the Min Tx rate desired by a peer MEP it does not transition from down to the INIT state and sends a diagnostic code (TBD) indicating that the requested Tx rate cannot be supported.

Otherwise once a transition from DOWN to INIT has occurred, the session progresses as per [4].

3.5.2. Defect entry criteria

There are further defect criteria beyond that defined in [4] to consider given the possibility of mis-connectivity and mis-configuration defects. The result is the criteria for a path direction to transition from the defect free state to a defect state is a superset of that in the BFD base specification [4].

The following conditions case a MEP to enter the defect state:

1. BFD session times out (Loss of Continuity defect),
2. BFD control packets are received with an unexpected encapsulation (Mis-connectivity defect), these include
   - a PW receiving a packet with a GAL
   - an LSP receiving an IP header instead of a GAL

   (note there are other possibilities but these can also alias
3. Receipt of an unexpected globally unique Source MEP identifier (Mis-connectivity defect),
4. Receipt of an unexpected session discriminator (Mis-connectivity defect)
5. Receipt of an unexpected M bit (Session Mis-configuration defect)

The effective defect hierarchy (order of checking) is

1. Receiving nothing

2. Receiving from an incorrect source (determined by whatever means)

3. Receiving from a correct source (as near as can be determined), but with incorrect session information)

4. Receiving control packets in all discernable ways correct.

3.5.3. 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.

The blocking of traffic as consequent action MUST be driven only by a defect's consequent action as specified in draft-ietf-mpls-tp-oam-framework Error! Reference source not found. section 5.1.1.2. When the defect is mis-braching, the transport path termination will silently discard all non-oam traffic received.

3.5.4. Defect exit criteria

Exit from a Loss of continuity defect

For a fate sharing session exit from a loss of connectivity defect is as described in [4].

For an independent session, exit from a loss of connectivity defect occurs upon receipt of a well formed control packet from the peer MEP.

Exit from a session mis-configuration defect

[editors: for a future version of the document]

Exit from a mis-connectivity defect

The exit criteria for a mis-connectivity defect is determined by the maximum of the set of min Rx session time times the multiplier that have been received. A session can transition from DOWN to UP (independent mode) or DOWN to INIT (fate sharing mode) when both correctly formed control packets are being exchanged, and no mis-connected control packets have been received in the specified interval.

3.5.5. Configuration of MPLS-TP BFD sessions

[Editors note, for a future revision of the document]

3.5.6. Discriminator values

MPLS labels at peer MEPs are used to provide context for the received BFD packets.

In the BFD control packet the discriminator values have either local or no significance.

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.

4. Acknowledgments

To be added in a later version of this document

5. IANA Considerations

To be added in a later version of this document

6. Security Considerations

The security considerations for the authentication TLV need further study.

Base BFD foresees an optional authentication section (see [4] section 6.7); that can be extended also to the tool proposed in this document.

Authentication methods that require checksum calculation on the outgoing packet must extend the checksum also on the ME Identifier Section. This is possible but seems uncorrelated with the solution proposed in this document: it could be better to use the simple password authentication method.

7. References

7.1. Normative References

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

[2] Bocci, M. et al., " MPLS Generic Associated Channel ", RFC 5586 , June 2009

[3] Vigoureux, M., Betts, M. and D. Ward, "Requirements for OAM in MPLS Transport Networks", draft-ietf-mpls-tp-oam-requirements-06 (work in progress), March 2010

[4] Katz, D. and D. Ward, "Bidirectional Forwarding Detection", draft-ietf-bfd-base-11 (work in progress), February 2009

[5] Boutros, S. et al., "Definition of ACH TLV Structure", draft-ietf-mpls-tp-ach-tlv-02 (work in progress), March 2010

7.2. Informative References

[6] Bocci, M., et al., "A Framework for MPLS in Transport Networks", draft-ietf-mpls-tp-framework-12, (work in progress), May 2010

[7] Bocci, M. and G. Swallow, "MPLS-TP Identifiers", draft-swallow-mpls-tp-identifiers-02 (work in progress), March 2010

[8] Allan, D., Busi, I. and B. Niven-Jenkins, "MPLS-TP OAM Framework", draft-ietf-mpls-tp-oam-framework-06 (work in progress), April 2010

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