Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 4, 2016

A. Mishra
Ciena Corporation
M. Jethanandani
Cisco Systems
A. Saxena
Ciena Corporation
S. Pallagatti
Juniper Networks
M. Chen
Huawei
P. Fan
China Mobile

March 3, 2016

BFD Stability
draft-ashesh-bfd-stability-04.txt

Abstract

This document describes extensions to the Bidirectional Forwarding Detection (BFD) protocol to measure BFD stability. Specifically, it describes a mechanism for detection of BFD frame loss as well as local delay measurements for BFD transmitter and receiver.

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 RFC 2119 [RFC2119].

Table of Contents

1. Introduction

The Bidirectional Forwarding Detection (BFD) [RFC5880] protocol operates by transmitting and receiving control frames, generally at high frequency, over the datapath being monitored. In order to prevent significant data loss due to a datapath failure, the tolerance for lost or delayed frames in the Detection Time, as defined in BFD [RFC5880] is set to the smallest feasible value.

This document proposes a mechanism to detect delayed or lost frames in a BFD session in addition to the datapath fault detection mechanisms of BFD. Such a mechanism presents significant value to measure the stability of BFD sessions and provides data to the operators for the cause of a BFD failure.

This document does not propose BFD extension to measure data traffic loss or delay on a link or tunnel and the scope is limited to BFD frames.

2. Use Cases

Legacy BFD cannot detect any BFD frame delay or loss if delay or loss does not last for dead interval. This draft proposes a method to distinguish between a dropped and a delayed frame on the receiver. For example, if the receiver receives BFD CC frame k at time t but receives frame k+1 at time t+9.9ms for a 3.3ms BFD interval, the frame is delayed. However, if the receiver receives frame k+3 at time t+10ms, and never receives frame k+1 and/or k+2, then it has experienced a drop. Delays can be because of congestion in the network or because of delays in the BFD transmitter or receiver.

This proposal enables BFD engine to generate diagnostic information on the health of each BFD session that could be used to preempt a failure on a link that BFD was monitoring by allowing time for a corrective action to be taken.

In a faulty datapath scenario, operator can use BFD health information to trigger delay and loss measurement OAM protocol (Connectivity Fault Management (CFM) or Loss Measurement (LM)-Delay Measurement (DM)) to further isolate the issue.

3. BFD Null-Authentication TLV

The functionality proposed for BFD stability measurement is achieved by appending the Null-Authentication TLV (as defined in Optimizing BFD Authentication [I-D.ietf-bfd-optimizing-authentication]) to the BFD control frame that do not have authentication enabled.

4. Theory of Operations

This mechanism allows operator to measure the loss, transmitter delay and receiver delay of BFD CC frames.

When using MD5 or SHA authentication, BFD uses authentication TLV that carries the Sequence Number. However, if non-meticulous authentication is being used, or no authentication is in use, then the non-authenticated BFD frames MUST include NULL-Auth TLV.

4.1. Loss Measurement

Loss measurement counts the number of BFD control frames missed at the receiver during any Detection Time period. The loss is detected by comparing the Sequence Number field in the Auth TLV (NULL or otherwise) in successive BFD CC frames. The Sequence Number in each successive control frame generated on a BFD session by the transmitter is incremented by one.

The first BFD NULL-Auth TLV processed by the receiver that has a non-zero sequence number is used for bootstrapping the logic. Each successive frame after this is expected to have a Sequence Number that is one greater than the Sequence Number in the previous frame. When the Sequence Number wraps around it should start from 1 instead of 0.

4.2. Delay Measurement

Delay measurement can be done locally & independently on the transmitter & receiver. Hence it is out of the scope of this document. Following is an example of how the delay measurement can be achieved on both sides:

Transmitter Delay:

Delay measurements on the transmitter can be made by calculating the time difference between software BFD engine transmitting the frame and the time when the hardware puts the frame on the wire.

Receiver Delay:

Delay measurement can be made using the time difference between the time hardware received a BFD Frame and the time software BFD Engine processed the frame.

While a constant delay may not be indicator of instability, large transient delays can decrease the BFD session stability significantly. BFD MAY choose to inform the operator about any of the delays when the delay measurement crosses a particular threshold value.

5. IANA Requirements

N/A

6. Security Consideration

Other than concerns raised in BFD [RFC5880] there are no new concerns with this proposal.

7. Contributors

Manav Bhatia

8. Acknowledgements

Authors would like to thank Nobo Akiya, Jeffery Haas, Peng Fan, Dileep Singh, Basil Saji, Sagar Soni and Mallik Mudigonda who also contributed to this document.

9. Normative References

[I-D.ietf-bfd-optimizing-authentication]
Jethanandani, M., Mishra, A., Saxena, A., and M. Bhatia, "Optimizing BFD Authentication", draft-ietf-bfd-optimizing-authentication-01 (work in progress), February 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <http://www.rfc-editor.org/info/rfc5880>.

Authors' Addresses

Ashesh Mishra
Ciena Corporation
3939 North 1st Street
San Jose, CA 95134
USA

Email: mishra.ashesh@outlook.com
URI: www.ciena.com

Mahesh Jethanandani
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: mjethanandani@gmail.com
URI: www.cisco.com

Ankur Saxena
Ciena Corporation
3939 North 1st Street
San Jose, CA 95134
USA

Email: ankurpsaxena@gmail.com
URI: www.ciena.com

Santosh Pallagatti
Juniper Networks
Juniper Networks, Exora Business Park
Bangalore, Karnataka 560103
India

Email: santoshpk@juniper.net

Mach Chen
Huawei

Email: mach.chen@huawei.com

Peng Fan
China Mobile
32 Xuanwumen West Street
Beijing, Beijing
China

Email: fanp08@gmail.com