Network Working Group   A. Mishra
Internet-Draft   O3b Networks
Intended status: Standards Track   M. Jethanandani
Expires: May 25, 2018   November 21, 2017

                        BFD Performance Measurement
                     draft-am-bfd-performance-00

Abstract

   This document describes an extension to the Bidirectional Forwarding
   Detection (BFD) protocol to determine the optimal BFD transmit
   interval for links with high one-way delay.

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 determine the smallest BFD
   transmit interval that can be supported on the link. This is
   achieved by actively measuing the one-way delay for each BFD session
   and setting the BFD session intervals based on the measured delay.
   This allows the BFD session to adapt to the fastest rate feasible on
   the current active path.

2. Use Cases

   To ensure stability, the BFD interval is typically set to value
   greater than the one-way delay of the link. This value is currently
   manually tuned based on the largest one-way delay in the set of links
   over which the session can be established.

   The method described in this proposal is useful in networks where the
   network latency is high, or varies with time. Trans-oceanic links
   and connectivity over geo-synchronous satellites are typical examples
   of links where the latency is high and the difference in latency on
   primary and backup paths can be significant.

   Another use-case is connectivity using satellites in mid-earth orbit
   (MEO) or low-earth orbit (LEO). In these systems the one-way delay,
   while it is low (25msec to 150 msec), varies with time. This

variation, based on various factors, can be as high as 30 msec. With
mobile receivers, such as ships, the delay when using such
connectivity can be non-trivial to predict. This requires an
automated method to determine the optimal BFD interval to allow
fastest possible recovery in case of failure.

Many networks employ the use of diverse link types for redundancy
where each link has significantly different link characteristics.
For example, using geo-stationary orbit (GEO) satellite backup for
MEO/LEO connectivity, or using fibre backup for MEO connectivity.
The end-to-end BFD sessions for services running on top of the
diverse transport will benefit from adaptive BFD rate.

3. BFD Performance TLV

The functionality proposed for BFD performance measurement is
achieved by proposing a new BFD Performance TLV to the BFD control
frame. This TLV leverages the delay measurement method defined in
RFC 6374 [RFC6374]. As BFD Version 1 control frame does not have
unused flags, the BFD Performance TLV overloads the BFD
Authentication Flag and uses a new auth type BFDP-AUTH-TYPE (code-
point TBA). The BFD Performance TLV merges the MPLS delay
measurement message with the BFD authentication TLV (while removing
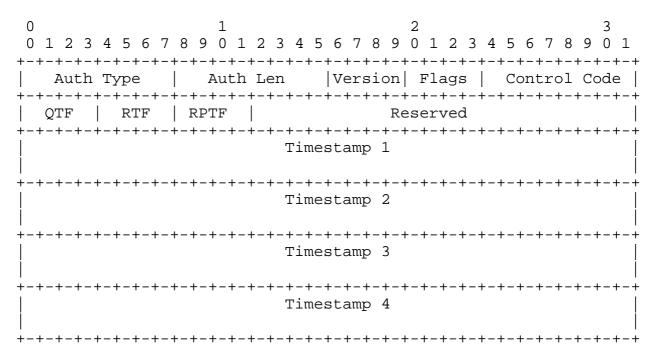fields that are not required for this application)

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   Auth Type    |    Auth Len     |Version| Flags |  Control Code |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  QTF  |  RTF  | RPTF  |                 Reserved                |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Timestamp 1                           |
   |                                                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Timestamp 2                           |
   |                                                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Timestamp 3                           |
   |                                                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Timestamp 4                           |
   |                                                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: BFD Performance TLV

where:

Auth Type: The Authentication Type, which in this case is BFDP-AUTH-TYPE (value to be assigned).

Auth Len: The length of the Authentication Section, in bytes.

Version: Currently set to 0.

Flags: As specified in Section 3.1 of RFC 6374 [RFC6374]. The T flag is set to 1.

Control Code: As specified in Section 3.1 of RFC 6374 [RFC6374].

QTF: Querier Timestamp Format. The format of the timestamp values written by the querier, as specified in Section 3.4 of RFC 6374 [RFC6374].

RTF: Responder Timestamp Format. The format of the timestamp values written by the responder, as specified in Section 3.4 of RFC 6374 [RFC6374].

RPTF: Responder's Preferred Timestamp Format. The timestamp format preferred by the responder, as specified in Section 3.4 of RFC 6374 [RFC6374].

Timestamp 1-4: Referring to Section 2.4 of RFC 6374 [RFC6374], when a query is sent from A, Timestamp 1 is set to T1 and the other timestamp fields are set to 0. When the query is received at B, Timestamp 2 is set to T2. At this point, B copies Timestamp 1 to Timestamp 3 and Timestamp 2 to Timestamp 4, and re-initializes Timestamp 1 and Timestamp 2 to 0. When B transmits the response, Timestamp 1 is set to T3. When the response is received at A, Timestamp 2 is set to T4. The actual formats of the timestamp fields written by A and B are indicated by the Querier Timestamp Format and Responder Timestamp Format fields respectively.

The mapping of timestamps to the Timestamp 1-4 fields is designed to ensure that transmit timestamps are always written at the same fixed offset in the packet, and likewise for receive timestamps. This property is important for hardware processing.

4. Theory of Operations

This delay measurement follows the method defined in Section 2.4 of RFC 6374 [RFC6374].

The message is classified using the BFD authentication method defined in RFC5880 [RFC5880].

Method for determining the optimal BFD interval for a link with
certain delay charateristics is implementation specific and beyond
the scope of this document.

5. IANA Requirements

Requesting new BFD Authentication Type for BFD Performance TLV.

6. Security Consideration

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

7. Normative References

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

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

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay
           Measurement for MPLS Networks", RFC 6374,
           DOI 10.17487/RFC6374, September 2011,
           <https://www.rfc-editor.org/info/rfc6374>.

Authors' Addresses

Ashesh Mishra
O3b Networks

Email: mishra.ashesh@gmail.com

Mahesh Jethanandani

Email: mjethanandani@gmail.com