      Bidirectional Forwarding Detection (BFD) on Multi-chassis Link
                  Aggregation Group (MC-LAG) Interfaces
                     draft-mtm-rtgwg-bfd-mc-lag-01

Abstract

   This document describes the use of Bidirectional Forwarding Detection
   for Multi-chassis Link Aggregation Group to provide faster than Link
   Aggregation Control Protocol convergence. This specification
   enhances RFC 7130 "Bidirectional Forwarding Detection (BFD) on Link
   Aggregation Group (LAG) Interfaces".

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at https://datatracker.ietf.org/drafts/current/.

   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."

   This Internet-Draft will expire on May 22, 2021.

Table of Contents

1. Introduction

   The [RFC7130] defines the use of Bidirectional Forwarding Detection
   (BFD) on Link Aggregation Group (LAG) interfaces. A multi-chassis
   LAG (MC-LAG) is a type of LAG [IEEE.802.1AX.2008] with member links
   terminated on separate chassis. [IEEE.802.1AX.2008] does not specify
   MC-LAG but doesn't preclude it either. Link Aggregation Control
   Protocol (LACP), also defined in [IEEE.802.1AX.2008], can work with
   MC-LAG but, as in the LAG case, the fastest link failure detection
   interval is only in a range of single-digit seconds. This document
   defines how the mechanism defined to work on LAG interfaces [RFC7130]
   can be adapted to the MC-LAG case to enable sub-second detection of
   member link failure.

1.1. Conventions used in this document

1.1.1. Acronyms

   BFD: Bidirectional Forwarding Detection

   LAG: Link Aggregation Group

   LACP: Link Aggregation Control Protocol

   MC-LAG: Multi-chassis Link Aggregation Group

MPLS: Multi-Protocol Label Switching

1.1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in BCP
14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

2. Problem Statement

[RFC7130] does not specify the selection of the destination IP
address for the BFD control packet. The only requirement related to
the selection is in Section 2.1, stating that the use of the address
family across all member links of the given LAG MUST be consistent
across all the links. Thus it is implied that the same unicast IP
address will be used on all member links of the LAG as the use of
different destination addresses would defeat the purpose of [RFC7130]
transforming the case into a set of single-hop BFD sessions
[RFC5881]. But a single unicast IP address may not work in the MC-
LAG case as the member links are terminated on the separate chassis.
This document proposes overcoming this problem if using IP or Multi-
Protocol Label Switching (MPLS) data plane encapsulation.

3. BFD on MC-LAG with IP-only Data Plane

As described in [RFC7130], a micro-BFD session on the LAG interfaces
may use IPv4 or IPv6 address family. In some cases, two sessions,
one with IPv4 and one with IPv6 addresses, may run concurrently.
This document doesn't change any of these but specifies the selection
of the destination IP address in the MC-LAG use case:

o if IPv4 address family is used for the micro-BFD session, then an
   address from the link-local multicast address 224.0.0.0/24 range
   SHOULD be used as the destination IP address. The subnet
   broadcast address MAY be used as the destination IP address as
   well;

o if the address family used is IPv6, then the IPv6 All Routers
   address with the link scope, as defined in [RFC4291], FF02::2/128
   MUST be used as the destination IP address.

4. BFD on MC-LAG with IP/MPLS Data Plane

IP/UDP is the most natural encapsulation format for the case of
micro-BFD on MC-LAG over IP/MPLS data plane as displayed in Figure 1.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                     MPLS Label Stack                          ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-------
~                                                              ~   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+   IP
|                   Destination IP address                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ Header
|                     Source IP address                        |   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-------
|                        UDP header                            |
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                     BFD Control Packet                       ~
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
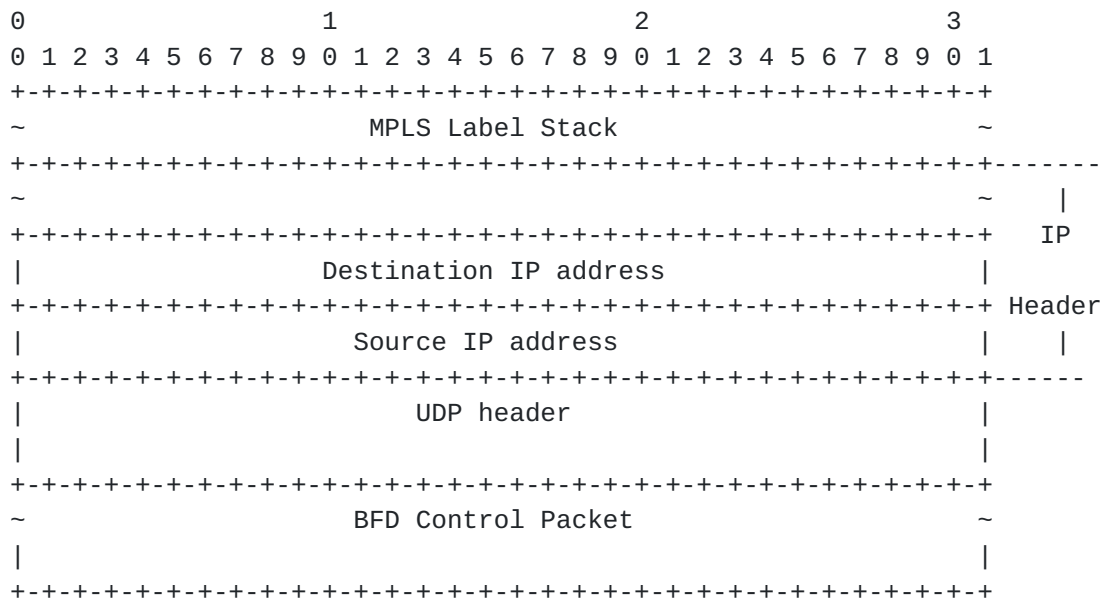
        Figure 1: BFD on MC-LAG member link on IPv4/MPLS data plane

   An IP and UDP headers immediately follow an MPLS label stack. The
   destination IP address MUST be set to the loopback address
   127.0.0.1/32 for IPv4 [RFC1812], or the loopback address ::1/128 for
   IPv6 [RFC4291]. TTL or Hop Limit field value MUST be set to 255,
   according to [RFC5881].

5. IANA Considerations

   This document makes no requests for IANA allocations. This section
   may be deleted by RFC Editor.

6. Security Considerations

   This document does not introduce new security concerns but inherits
   all security considerations discussed in [RFC5881] and [RFC7130].

7. Acknowledgements

   TBD

8. References

8.1. Normative References

   [IEEE.802.1AX.2008]
              "IEEE Standard for Local and metropolitan area networks -
              Link Aggregation", IEEE 802.1-AX, November 2008.

   [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>.

   [RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection
              (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881,
              DOI 10.17487/RFC5881, June 2010,
              <https://www.rfc-editor.org/info/rfc5881>.

   [RFC7130] Bhatia, M., Ed., Chen, M., Ed., Boutros, S., Ed.,
              Binderberger, M., Ed., and J. Haas, Ed., "Bidirectional
              Forwarding Detection (BFD) on Link Aggregation Group (LAG)
              Interfaces", RFC 7130, DOI 10.17487/RFC7130, February
              2014, <https://www.rfc-editor.org/info/rfc7130>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative

   [RFC1812] Baker, F., Ed., "Requirements for IP Version 4 Routers",
              RFC 1812, DOI 10.17487/RFC1812, June 1995,
              <https://www.rfc-editor.org/info/rfc1812>.

   [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing
              Architecture", RFC 4291, DOI 10.17487/RFC4291, February
              2006, <https://www.rfc-editor.org/info/rfc4291>.

Authors' Addresses

   Greg Mirsky
   ZTE Corp.

   Email: gregimirsky@gmail.com

   Jeff Tantsura
   Apstra

   Email: jefftant.ietf@gmail.com

   Gyan Mishra
   Verizon Inc.

   Email: gyan.s.mishra@verizon.com