Bidirectional Forwarding Detection (BFD) for Multi-point Networks and Protocol Independent Multicast - Sparse Mode (PIM-SM) Use Case
draft-ietf-pim-bfd-p2mp-use-case-02

Abstract

   This document discusses the use of Bidirectional Forwarding Detection
   (BFD) for multi-point networks to provide nodes that participate in
   Protocol Independent Multicast - Sparse Mode (PIM-SM) with the sub-
   second convergence. Optional extension to PIM-SM Hello, as specified
   in RFC 7761, to bootstrap point-to-multipoint BFD session. also
   defined in this document.

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 January 26, 2020.

Table of Contents

1. Introduction

   Faster convergence in the control plane, in general, is beneficial
   and allows minimizing periods of traffic blackholing, transient
   routing loops, and other scenarios that may negatively affect service
   data flow. That equally applies to unicast and multicast routing
   protocols.

   [RFC7761] is the current specification of the Protocol Independent
   Multicast - Sparse Mode (PIM-SM) for IPv4 and IPv6 networks.
   Confirming implementation of PIM-SM elects a Designated Router (DR)
   on each PIM-SM interface. When a group of PIM-SM nodes is connected
   to shared-media segment, e.g., Ethernet, the one elected as DR is to
   act on behalf of directly connected hosts in the context of the PIM-
   SM protocol. Failure of the DR impacts the quality of the multicast
   services it provides to directly connected hosts because the default
   failure detection interval for PIM-SM routers is 105 seconds.
   Introduction of Backup DR (BDR), proposed in
   [I-D.ietf-pim-dr-improvement], improves convergence time in the PIM-
   SM over shared-media segment but still depends on long failure
   detection interval.

   Bidirectional Forwarding Detection (BFD) [RFC5880] had been
   originally defined to detect failure of point-to-point (p2p) paths -
   single-hop [RFC5881], multihop [RFC5883]. [RFC8562] extends the BFD
   base specification [RFC5880] for multipoint and multicast networks,
   which precisely characterizes deployment scenarios for PIM-SM over
   LAN segment. This document demonstrates how point-to-multipoint
   (p2mp) BFD can enable faster detection of PIM-SM router failure and

thus minimize multicast service disruption. The document also
defines the extension to PIM-SM [RFC7761] and
[I-D.ietf-pim-dr-improvement] to bootstrap a PIM-SM router to join in
p2mp BFD session over shared-media link.

1.1. Conventions used in this document

1.1.1. Terminology

BFD: Bidirectional Forwarding Detection

BDR: Backup Designated Router

DR: Designated Router

p2mp: Pont-to-Multipoint

PIM-SM: Protocol Independent Multicast - Sparse Mode

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

[RFC7761] does not provide a method for fast, e.g., sub-second,
failure detection of a neighbor PIM-SM router. BFD already has many
implementations based on HW that are capable of supporting multiple
sub-second sessions concurrently.

3. Applicability of p2mp BFD

[RFC8562] may provide an efficient and scalable solution for the
fast-converging environment that demonstrates the head-tails
relationship. Each such group presents itself as p2mp BFD session
with its head being the root and other routers being tails of the
p2mp BFD session. Figure 1 displays the new optional BFD
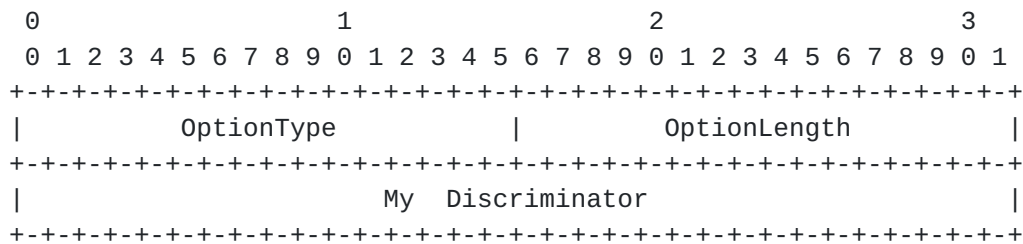Discriminator TLV to bootstrap tail of the p2mp BFD session.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            OptionType         |          OptionLength         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       My  Discriminator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   Figure 1: BFD Discriminator TLV to Bootstrap P2MP BFD session

where new fields are interpreted as:

   OptionType is a value (TBA1) assigned by IANA Section 4 that
   identifies the TLV as BFD Discriminator TLV;

   OptionLength value is always 4

   My Discriminator - My Discriminator value allocated by the root of
   the p2mp BFD session.

If PIM-SM routers that support this specification are configured to
use p2mp BFD for faster convergence, then the router to be monitored,
referred to as 'head', MUST create BFD session of type
MultipointHead, as defined in [RFC8562]. If the head doesn't support
[I-D.ietf-pim-dr-improvement], but, for example, uses procedures
defined in [I-D.mankamana-pim-bdr], then it MUST include BFD TLV in
its PIM-Hello message. If the head uses extensions defined in
[I-D.ietf-pim-dr-improvement], then DR MUST include BFD TLV in its
Hello message. The DR Address TLV also MUST be included in the Hello
message. For a BDR it is RECOMMENDED to include BFD TLV in its Hello
message. If BDR includes BFD TLV, then the BDR Address TLV also MUST
be present in the Hello message. Then the head MUST begin periodic
transmission of BFD control packets. Source IP address of the BFD
control packet MUST be the same as the source IP address of the PIM-
Hello with BFD TLV messages being transmitted by the head. The
values of My Discriminator in the BFD control packet and My
Discriminator field of the BFD TLV in PIM-Hello, transmitted by the
head MUST be the same. When a PIM-SM router is configured to monitor
the head by using p2p BFD, referred to through this document as
'tail', receives PIM-Hello packet with BFD TLV it MAY create p2mp BFD
session of type MultipointTail, as defined in [RFC8562].

Because p2mp BFD doesn't use the three-way handshake and the head
transmits BFD control packets with the value of Your Discriminator
field set to zero, [RFC8562] modified how a BFD system demultiplexes
received BFD control packet. The tail demultiplexes p2mp BFD test
session based on head's source IP address, the My Discriminator value
it learned from BFD Discriminator TLV and the identity of the

multipoint path that the BFD control packet was received from. The
Detection Time for p2mp BFD sessions is defined differently from the
definition provided in [RFC5880]. The Detection Time for each
MultipointTail session is calculated as the product of the last
received values of Desired Min TX Interval and Detect Mult. A tail
declares the BFD session down after the Detection Timer expires. If
the tail has detected MultipointHead failure, it MUST remove the
neighbor. If the failed head node was PIM-SM DR or BDR, the tail MAY
start DR Election process as specified in Section 4.3.2 [RFC7761] or
Section 4.1 [I-D.ietf-pim-dr-improvement] respectively.

If the head ceased to include BFD TLV in its PIM-Hello message, tails
MUST close the corresponding MultipointTail BFD session. Thus the
tail stops using BFD to monitor the head and reverts to the
procedures defined in [RFC7761] and [I-D.ietf-pim-dr-improvement].

3.1. Multipoint BFD Encapsulation

The MultipointHead of p2mp BFD session when transmitting BFD control
packet:

   MUST set TTL value to 1;

   SHOULD use group address ALL-PIM-ROUTERS ('224.0.0.13' for IPv4
   and 'ff02::d' for IPv6) as destination IP address

   MAY use network broadcast address for IPv4 or link-local all nodes
   multicast group for IPv6 as the destination IP address;

   MUST set destination UDP port value to 3784 when transmitting BFD
   control packets, as defined in [RFC8562].

4. IANA Considerations

IANA is requested to allocate a new OptionType value from PIM Hello
Options registry according to:

```
+-------------+---------------+------------------+---------------+
| Value Name  | Length Number | Name Protocol    | Reference     |
+-------------+---------------+------------------+---------------+
| TBA         | 4             | BFD Discriminator | This document |
+-------------+---------------+------------------+---------------+
```

                 Table 1: BFD Discriminator option type

5. Security Considerations

Security considerations discussed in [RFC7761], [RFC5880], and
[RFC8562], and [I-D.ietf-pim-dr-improvement] apply to this document.

An implementation that supports this specification SHOULD use a
mechanism to control the maximum number of BFD sessions that can be
active at the same time.

6. Acknowledgments

Authors cannot say enough to express their appreciation of comments
and suggestions we received from Stig Venaas.

7. Normative References

[I-D.ietf-pim-dr-improvement]
          Zhang, Z., hu, f., Xu, B., and m. mishra, "PIM DR
          Improvement", draft-ietf-pim-dr-improvement-07 (work in
          progress), January 2019.

[I-D.mankamana-pim-bdr]
          mishra, m., "PIM Backup Designated Router Procedure",
          draft-mankamana-pim-bdr-02 (work in progress), April 2019.

[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>.

[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>.

[RFC5883] Katz, D. and D. Ward, "Bidirectional Forwarding Detection
          (BFD) for Multihop Paths", RFC 5883, DOI 10.17487/RFC5883,
          June 2010, <https://www.rfc-editor.org/info/rfc5883>.

[RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I.,
          Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent
          Multicast - Sparse Mode (PIM-SM): Protocol Specification
          (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March
          2016, <https://www.rfc-editor.org/info/rfc7761>.

   [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>.

   [RFC8562] Katz, D., Ward, D., Pallagatti, S., Ed., and G. Mirsky,
              Ed., "Bidirectional Forwarding Detection (BFD) for
              Multipoint Networks", RFC 8562, DOI 10.17487/RFC8562,
              April 2019, <https://www.rfc-editor.org/info/rfc8562>.

Authors' Addresses

   Greg Mirsky
   ZTE Corp.

   Email: gregimirsky@gmail.com

   Ji Xiaoli
   ZTE Corporation
   No.50 Software Avenue, Yuhuatai District
   Nanjing
   China

   Email: ji.xiaoli@zte.com.cn