PIM Working Group
Internet-Draft
Intended status: Standards Track
Expires: 12 June 2022

G. Mirsky
Ericsson
J. Xiaoli
ZTE Corporation

9 December 2021

Fast Failover in Protocol Independent Multicast - Sparse Mode (PIM-SM) Using Bidirectional Forwarding Detection (BFD) for Multipoint Networks draft-ietf-pim-bfd-p2mp-use-case-10

Abstract

This document specifies how Bidirectional Forwarding Detection for multipoint networks can provide sub-second failover for routers that participate in Protocol Independent Multicast - Sparse Mode (PIM-SM). An extension to the PIM Hello message used to bootstrap a point-to-multipoint BFD session is 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 12 June 2022.

Copyright Notice

Table of Contents

1. Introduction

Faster convergence in the control plane minimizes the periods of traffic blackholing, transient routing loops, and other situations that may negatively affect service data flow. Faster convergence in the control plane is beneficial 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. A conforming implementation of PIM-SM elects a Designated Router (DR) on each PIM-SM interface. When a group of PIM-SM nodes is connected to a shared media segment, e.g., Ethernet, the node elected as DR acts 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.

Bidirectional Forwarding Detection (BFD) [RFC5880] was originally defined to detect a failure of a point-to-point (p2p) path, single-hop [RFC5881] or multihop [RFC5883]. In some PIM-SM deployments, a

p2p BFD can be used to detect a failure and enable faster failover. [RFC8562] extends the BFD base specification [RFC5880] for multipoint and multicast networks, which matches the deployment scenarios for PIM-SM over a LAN segment. A BFD system in p2mp environment that transmits BFD Control messages using the BFD Demand mode [RFC5880] creates less BFD state than the Asynchronous mode. Point-to-multipoint (p2mp) BFD can enable faster detection of PIM-SM router failure compared to PIM-SM without BFD and thus minimize multicast service disruption. The monitored PIM-SM router acts as the head and other routers as tails of a p2mp BFD session. This document defines the monitoring of a PIM-SM router using p2mp BFD. This document also defines the extension to PIM-SM [RFC7761] to bootstrap a PIM-SM router to join in p2mp BFD session over shared media segment.

1.1. Conventions used in this document

1.1.1. Terminology

This document uses terminology defined in [RFC5880], [RFC8562], and [RFC7761]. Familiarity with these specifications and the terminology used is expected.

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. BFD Discriminator PIM Hello Option

Figure 1 displays the new optional BFD Discriminator PIM Hello option to bootstrap a tail of the p2mp BFD session.

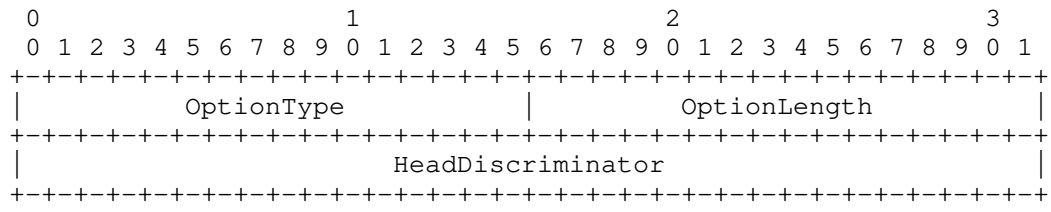

Figure 1: BFD Discriminator PIM Hello Option

where new fields are interpreted as:

OptionType: TBA.

OptionLength: MUST be set to 4.

HeadDiscriminator: the four-octet field MUST be included in the BFD Discriminator PIM-SM Hello option. The value MUST NOT be zero. It equals the value of My Discriminator ([RFC5880]) allocated by the head.

If the value of the OptionLength field is not equal to 4, the BFD Discriminator PIM Hello option is considered malformed, and the receiver MUST stop processing PIM Hello options. If the value of the HeadDiscriminator field equals zero, then the BFD Discriminator PIM Hello option MUST be considered invalid, and the receiver MUST ignore it. The receiver SHOULD log a notification regarding the malformed or invalid BFD Discriminator Hello option under the control of a throttling logging mechanism.

2.1. Using P2MP BFD in PIM Router Monitoring

If the head is no longer serving the function that prompted it to be monitored, then it MUST cease including the BFD Discriminator PIM Hello option in its PIM-Hello message, and it SHOULD shut down the BFD session following the procedures described in Section 5.9 [RFC8562].

The head MUST create a BFD session of type MultipointHead [RFC8562]. Note that any PIM-SM router, regardless of its role, MAY become a head of a p2mp BFD session. To control the volume of BFD control traffic on a shared media segment, an operator should carefully select PIM-SM routers configured as a head of a p2mp BFD session. The head MUST include the BFD Discriminator PIM Hello option in its PIM Hello messages.

A PIM-SM router that is configured to monitor the head by using p2mp BFD is referred to throughout this document as a "tail". When such a tail receives a PIM-Hello packet with the BFD Discriminator PIM Hello option, the tail MAY create a p2mp BFD session of type MultipointTail, as defined in [RFC8562].

The node that includes the BFD Discriminator PIM Hello option transmits BFD Control packets periodically. For the tail to correctly demultiplex BFD [RFC8562], the source address and My Discriminator of the BFD packets MUST be the same as the source address and the HeadDiscriminator, respectively, of the PIM Hello message. If that is not the case, the tail BFD node would not be able to monitor the state of the PIM-SM node, that is, the head of the p2mp BFD session, though the regular PIM-SM mechanisms remain fully operational.

If the tail detects a MultipointHead failure [RFC8562], it MUST delete the corresponding neighbor state and follow procedures defined in [RFC7761] for the DR and additional neighbor state deletion after the neighbor timeout expires.

If the head ceases to include the BFD Discriminator PIM Hello option in its PIM-Hello message, tails SHOULD close the corresponding MultipointTail BFD session without affecting the PIM state in any way. Thus, the tail stops using BFD to monitor the head and reverts to the procedures defined in [RFC7761].

2.2. P2MP BFD in PIM DR Load Balancing

[RFC8775] specifies the PIM Designated Router Load Balancing (DRLB) functionality. Any PIM router that advertises the DRLB-Cap Hello Option can become the head of a p2mp BFD session, as specified in Section 2.1. The head router administratively sets the bfd.SessionState to Up in the MultipointHead session [RFC8562] only if it is a Group Designated Router (GDR) Candidate, as specified in Sections 5.5 and 5.6 of [RFC8775]. If the router is no longer the GDR, then it MUST shut down following the procedures described in Section 5.9 [RFC8562]. For each GDR Candidate that includes BFD Discriminator option in its PIM Hello, the PIM DR MUST create a MultipointTail session [RFC8562]. PIM DR demultiplexes BFD sessions based on the value of the My Discriminator field and the source IP address. If PIM DR detects a failure of one of the sessions, it MUST remove that router from the GDR Candidate list and immediately transmit a new DRLB-List option.

2.3. Multipoint BFD Encapsulation

The MultipointHead of a p2mp BFD session when transmitting BFD Control packets:

MUST set TTL or Hop Limit value to 255 (Section 5 [RFC5881]). Similarly, all received BFD Control packets that are demultiplexed to the session MUST be discarded if the received TTL or Hop Limit is not equal to 255;

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

3. IANA Considerations

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

Value	Length	Name	Reference
TBA	4	BFD Discriminator Option	This document

Table 1: BFD Discriminator option type

4. Security Considerations

This document defines a way to accelerate detecting a failure that affects PIM functionality by using BFD. The operation of either protocol is not changed.

The security considerations discussed in [RFC7761], [RFC5880], [RFC5881], [RFC8562], and [RFC8775] apply to this document.

5. Acknowledgments

The authors cannot say enough to express their appreciation of the comments and suggestions we received from Stig Venaas. The authors greatly appreciate the comments and suggestions by Alvaro Retana that improved the clarity of this document.

6. References

6.1. 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>.

[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>.

[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>.

[RFC8775] Cai, Y., Ou, H., Vallepalli, S., Mishra, M., Venaas, S., and A. Green, "PIM Designated Router Load Balancing", RFC 8775, DOI 10.17487/RFC8775, April 2020, <https://www.rfc-editor.org/info/rfc8775>.

6.2. Informative References

[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>.

Authors' Addresses

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com

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

Email: ji.xiaoli@zte.com.cn