MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2021

G. Mirsky
ZTE Corp.
G. Mishra
Verizon Inc.

October 18, 2020

BFD for Multipoint Networks over Point-to-Multi-Point MPLS LSP
draft-mirsky-mpls-p2mp-bfd-11

Abstract

This document describes procedures for using Bidirectional Forwarding Detection (BFD) for multipoint networks to detect data plane failures in Multiprotocol Label Switching (MPLS) point-to-multipoint (p2mp) Label Switched Paths (LSPs) using active tails with unsolicited notifications mode. It also describes the applicability of LSP Ping, as in-band, and the control plane, as out-band, solutions to bootstrap a BFD session in this environment.

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 April 21, 2021.

Table of Contents

1. Introduction

[RFC8562] defines a method of using Bidirectional Detection (BFD) [RFC5880] to monitor and detect unicast failures between the sender (head) and one or more receivers (tails) in multipoint or multicast networks. [RFC8562] added two BFD session types - MultipointHead and MultipointTail. Throughout this document, MultipointHead and MultipointTail refer to the value bfd.SessionType is set on a BFD system. This document describes procedures for using such a mode of BFD protocol to detect data plane failures in Multiprotocol Label Switching (MPLS) point-to-multipoint (p2mp) Label Switched Paths (LSPs). The document also describes the applicability of out-band solutions to bootstrap a BFD session in this environment.

2. Conventions used in this document

2.1. Terminology

MPLS: Multiprotocol Label Switching

LSP: Label Switched Path

BFD: Bidirectional Forwarding Detection

p2mp: Point-to-Multipoint

FEC: Forwarding Equivalence Class

G-ACh: Generic Associated Channel

ACH: Associated Channel Header

GAL: G-ACh Label

2.2. 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.

3. Multipoint BFD Encapsulation

[RFC8562] uses BFD in the Demand mode from the very start of a point-to-multipoint (p2mp) BFD session. Because the head doesn't receive any BFD Control packet from a tail, the head of the p2mp BFD session transmits all BFD Control packets with the value of Your Discriminator field set to zero. As a result, a tail cannot demultiplex BFD sessions using Your Discriminator, as defined in [RFC5880]. [RFC8562] requires that to demultiplex BFD sessions, the tail uses the source IP address, My Discriminator, and the identity of the multipoint tree from which the BFD Control packet was received. The p2mp MPLS LSP label MAY provide the identification of the multipoint tree in case of an inclusive p-tree or upstream assigned label in case of aggregate p-tree. If the BFD Control packet is encapsulated in IP/UDP, then the source IP address MUST be used to demultiplex the received BFD Control packet as described in Section 3.1. The non-IP encapsulation case is described in Section 3.2.

3.1. IP Encapsulation of Multipoint BFD

[RFC8562] defines IP/UDP encapsulation for multipoint BFD over p2mp MPLS LSP:

UDP destination port MUST be set to 3784;

destination IP address MUST be set to the loopback address 127.0.0.1/32 for IPv4, or the loopback address ::1/128 for IPv6 [RFC4291].

This specification further clarifies that:

if multiple alternative paths for the given p2mp LSP Forwarding Equivalence Class (FEC) exist, the MultipointHead SHOULD use Entropy Label [RFC6790] used for LSP Ping [RFC8029] to exercise that particular alternative path;

or the MultipointHead MAY use the UDP port number as discovered by LSP Ping traceroute [RFC8029] as the source UDP port number to possibly exercise that particular alternate path.

3.2. Non-IP Encapsulation of Multipoint BFD

In some environments, the overhead of extra IP/UDP encapsulations may be considered as overburden, thus making the use of more compact G-ACh encapsulation attractive. Also, the validation of the IP/UDP encapsulation of BFD Control packet of p2mp BFD session may fail because of a problem not related to neither MPLS label stack nor to BFD. Avoiding unnecessary encapsulation of p2mp BFD over MPLS LSP improves the accuracy of the correlation of the detected failure and defect in MPLS LSP. Non-IP encapsulation for multipoint BFD over p2mp MPLS LSP MUST use Generic Associated Channel (G-ACh) Label (GAL) (see [RFC5586]) at the bottom of the label stack followed by Associated Channel Header (ACH). If BFD Control, PW-ACH encapsulation (without IP/UDP Headers) channel to be used in ACH, an implementation would not be able to verify the identity of the MultipointHead and, as a result, will not properly demultiplex BFD packets. Hence, a new channel type value is needed. The Channel Type field in ACH MUST be set to TBA1 value Section 6. To provide the identity of the MultipointHead for the particular multipoint BFD session, a Source Address TLV [RFC7212] MUST immediately follow a BFD Control message.

4. Bootstrapping Multipoint BFD

4.1. LSP Ping

LSP Ping is the part of the on-demand OAM toolset to detect and localize defects in the data plane and verify the control plane against the data plane by ensuring that the LSP is mapped to the same FEC, at the egress, as the ingress.

LSP Ping, as defined in [RFC6425], MAY be used to bootstrap MultipointTail. If the LSP Ping used, it MUST include the Target FEC

TLV and the BFD Discriminator TLV defined in [RFC5884]. The Target FEC TLV MUST use sub-TLVs defined in Section 3.1 [RFC6425]. It is RECOMMENDED setting the value of Reply Mode field to "Do not reply" [RFC8029] for the LSP Ping to bootstrap MultipointTail of the p2mp BFD session. Indeed, because BFD over a multipoint network is using BFD Demand mode, the LSP echo reply from a tail has no useful information to convey to the head, unlike in the case of the BFD over a p2p MPLS LSP [RFC5884]. A MultipointTail that receives the LSP Ping that includes the BFD Discriminator TLV:

- MUST validate the LSP Ping;
- MUST associate the received BFD Discriminator value with the p2mp LSP;
- MUST create a p2mp BFD session and set bfd.SessionType = MultipointTail as described in [RFC8562];
- MUST use the source IP address of LSP Ping, the value of BFD Discriminator from the BFD Discriminator TLV, and the identity of the p2mp LSP to properly demultiplex BFD sessions.

Besides bootstrapping a BFD session over a p2mp LSP, LSP Ping SHOULD be used to verify the control plane against the data plane periodically by checking that the p2mp LSP is mapped to the same FEC at the MultipointHead and all active MultipointTails. The rate of generation of these LSP Ping Echo request messages SHOULD be significantly less than the rate of generation of the BFD Control packets because LSP Ping requires more processing to validate the consistency between the data plane and the control plane. An implementation MAY provide configuration options to control the rate of generation of the periodic LSP Ping Echo request messages.

4.2. Operation of Multipoint BFD with Active Tail over P2MP MPLS LSP

[RFC8562] defined how the BFD Demand mode can be used in multipoint networks. When applied in MPLS, procedures specified in [RFC8562] allow an egress LSR to detect a failure of the part of the MPLS p2mp LSP from the ingress LSR. The ingress LSR is not aware of the state of the p2mp LSP. [RFC8563], using mechanisms defined in [RFC8562], defined an "active tail" behavior. An active tail might notify the head of the detected failure and responds to a poll sequence initiated by the head. The first method, referred to as Head Notification without Polling, is mentioned in Section 5.2.1 [RFC8563], is the simplest of all described in [RFC8563]. The use of this method in BFD over MPLS p2mp LSP is discussed in this document. Analysis of other methods of a head learning of the state of an MPLS p2mp LSP is outside the scope of this document.

As specified in [RFC8563] for the active tail mode, BFD variables MUST be as follows:

On an ingress LSR:

- bfd.SessionType is MultipointHead;
- bfd.RequiredMinRxInterval is set to nonzero allowing egress LSRs to send BFD Control packets.

On an egress LSR:

- bfd.SessionType is MultipointTail;
- bfd.SilentTail is set to zero.

In Section 5.2.1 [RFC8563] is noted that "the tail sends unsolicited BFD packets in response to the detection of a multipoint path failure" but without the specifics on the information in the packet and frequency of transmissions. This document defines the procedure of the active tail with unsolicited notifications for p2mp MPLS LSP as specified below.

Upon detecting the failure of the p2mp MPLS LSP, an egress LSR sends BFD Control packet with the following settings:

- the Poll (P) bit is set;
- the Status (Sta) field set to Down value;
- the Diagnostic (Diag) field set to Control Detection Time Expired value;
- the value of the Your Discriminator field is set to the value the egress LSR has been using to demultiplex that BFD multipoint session;
- BFD Control packet is encapsulated in IP/UDP with the destination IP address of the ingress LSR and the UDP destination port number set to 4784 per [RFC5883]
- these BFD Control packets are transmitted at the rate of one per second until either it receives the valid for this BFD session control packet with the Final (F) bit set from the ingress LSR or the defect condition clears.

To improve the likelihood of notifying the ingress LSR of the failure of the p2mp MPLS LSP, the egress LSR SHOULD transmit three BFD Control packets defined above in short succession.

An ingress LSR that has received the BFD Control packet, as described above, sends the unicast IP/UDP encapsulated BFD Control packet with the Final (F) bit set to the egress LSR.

4.3. Control Plane

BGP-BFD Attribute [I-D.ietf-bess-mvpn-fast-failover] MAY be used to bootstrap multipoint BFD session on a tail.

5. Security Considerations

This document does not introduce new security aspects but inherits all security considerations from [RFC5880], [RFC5884], [RFC7726], [RFC8562], [RFC8029], and [RFC6425].

Also, BFD for p2mp MPLS LSP MUST follow the requirements listed in section 4.1 [RFC4687] to avoid congestion in the control plane or the data plane caused by the rate of generating BFD Control packets. An operator SHOULD consider the amount of extra traffic generated by p2mp BFD when selecting the interval at which the MultipointHead will transmit BFD Control packets. Also, the operator MAY consider the size of the packet the MultipointHead transmits periodically as using IP/UDP encapsulation adds up to 28 octets, which is more than 50% of BFD Control packet length, comparing to G-ACh encapsulation.

6. IANA Considerations

IANA is requested to allocate value (TBA1) from its MPLS Generalized Associated Channel (G-ACh) Types registry.

Value	Description	Reference
TBA1	Multipoint BFD Session	This document

Table 1: Multipoint BFD Session G-ACh Type

7. Acknowledgements

The author sincerely appreciates the comments received from Andrew Malis and thought stimulating questions from Carlos Pignataro.

8. References

8.1. 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>.

[RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <https://www.rfc-editor.org/info/rfc5586>.

[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>.

[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>.

[RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, DOI 10.17487/RFC5884, June 2010, <https://www.rfc-editor.org/info/rfc5884>.

[RFC6425] Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011, <https://www.rfc-editor.org/info/rfc6425>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7212] Frost, D., Bryant, S., and M. Bocci, "MPLS Generic Associated Channel (G-ACh) Advertisement Protocol", RFC 7212, DOI 10.17487/RFC7212, June 2014, <https://www.rfc-editor.org/info/rfc7212>.

[RFC7726] Govindan, V., Rajaraman, K., Mirsky, G., Akiya, N., and S. Aldrin, "Clarifying Procedures for Establishing BFD Sessions for MPLS Label Switched Paths (LSPs)", RFC 7726, DOI 10.17487/RFC7726, January 2016, <https://www.rfc-editor.org/info/rfc7726>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[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>.

[RFC8563] Katz, D., Ward, D., Pallagatti, S., Ed., and G. Mirsky, Ed., "Bidirectional Forwarding Detection (BFD) Multipoint Active Tails", RFC 8563, DOI 10.17487/RFC8563, April 2019, <https://www.rfc-editor.org/info/rfc8563>.

8.2. Informative References

[I-D.ietf-bess-mvpn-fast-failover] Morin, T., Kebler, R., and G. Mirsky, "Multicast VPN Fast Upstream Failover", draft-ietf-bess-mvpn-fast-failover-11 (work in progress), October 2020.

[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>.

[RFC4687] Yasukawa, S., Farrel, A., King, D., and T. Nadeau, "Operations and Management (OAM) Requirements for Point-to-Multipoint MPLS Networks", RFC 4687, DOI 10.17487/RFC4687, September 2006, <https://www.rfc-editor.org/info/rfc4687>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com