Internet Engineering Task Force (IETF)
Request for Comments: 9780
Updates: 8562
Category: Standards Track
ISSN: 2070-1721

G. Mirsky
Ericsson
G. Mishra
Verizon Inc.
D. Eastlake 3rd
Independent

May 2025

Bidirectional Forwarding Detection (BFD) for Multipoint Networks over Point-to-Multipoint MPLS Label Switched Paths (LSPs)

Abstract

This document describes procedures for using Bidirectional Forwarding Detection (BFD) for multipoint networks to detect data plane failures in point-to-multipoint MPLS Label Switched Paths (LSPs) and Segment Routing (SR) point-to-multipoint policies with an SR over MPLS (SR-MPLS) data plane.

Furthermore, this document updates RFC 8562 by recommending the use of an IPv6 address from the Dummy IPv6 Prefix address block 100:0:0:1::/64 and discouraging the use of an IPv4 loopback address mapped to IPv6.

In addition, this document describes the applicability of LSP Ping (as an in-band solution) and the control plane (as an out-of-band solution) to bootstrap a BFD session. The document also describes the behavior of the active tail for head notification.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9780.

Table of Contents

1. Introduction

[RFC8562] defines a method of using Bidirectional Forwarding Detection (BFD) [RFC5880] to monitor and detect 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 to which the bfd.SessionType is set on a BFD endpoint.

This document describes procedures for using such modes of the BFD protocol to detect data plane failures in point-to-multipoint (P2MP) MPLS Label Switched Paths (LSPs) and Segment Routing (SR) point-to-multipoint policies with an SR over MPLS (SR-MPLS) data plane.

The document also describes the applicability of LSP Ping (an in-band solution) and out-of-band solutions to bootstrap a BFD session in this environment.

Historically, an address in the IPv6-mapped IPv4 loopback address block ::ffff:127.0.0.1/128 was mandated, although functionally, an IPv6 address from that address block is not analogous to its IPv4 counterpart. Furthermore, using the loopback address as the destination address, even for an inner IP encapsulation of a tunneled packet, violates Section 2.5.3 of [RFC4291]. Hence, IANA has allocated 100:0:0:1::/64 as a new Dummy IPv6 Prefix (Section 7.1) for destination IPv6 addresses used for IP/UDP encapsulation of management, control, and OAM (Operations, Administration, and Maintenance) packets. A source-only IPv6 dummy address is used as the destination to generate an exception and a reply message to the request message received. This document starts the transition to using the IPv6 addresses from the Dummy IPv6 Prefix address block 100:0:0:1::/64 as the IPv6 destination address in the IP/UDP encapsulation of active OAM over the MPLS data plane. Thus, this document updates [RFC8562] by recommending the use of an IPv6 address from the Dummy IPv6 Prefix address block 100:0:0:1::/64 (Section 7.1) while acknowledging that an address from the ::ffff:127.0.0.1/128 address block might be used by existing implementations. This document discourages the use of an address in the IPv6-mapped IPv4 loopback address block.

This document also describes the behavior of the active tail for head notification.

2. Conventions Used in This Document

2.1. Terminology

ACH: Associated Channel Header

BFD: Bidirectional Forwarding Detection

GAL: G-ACh Label

G-ACh: Generic Associated Channel

LSP: Label Switched Path

LSR: Label Switching Router

MPLS: Multiprotocol Label Switching

P2MP: Point-to-Multipoint

PW: Pseudowire (PW)

SR: Segment Routing

SR-MPLS: SR over MPLS

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 Demand mode from the very start of a point-to-multipoint (P2MP) BFD session. Because the head doesn't receive any BFD Control packets from a tail, the head of the P2MP BFD session transmits all BFD Control packets with the value of the Your Discriminator field set to zero. As a result, a tail cannot demultiplex BFD sessions using Your Discriminator, as defined in [RFC5880]. To demultiplex BFD sessions, [RFC8562] requires that the tail use the source IP address, My Discriminator, and the identity of the multipoint tree from which the BFD Control packet was received. 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 5.7 of [RFC8562]. 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. This document updates Section 5.8 of [RFC8562] regarding the selection of the IPv6 destination address as follows:

* The sender of an MPLS echo request SHOULD use an address from the Dummy IPv6 Prefix address block 100:0:0:1::/64 (see Section 7.1).

* The sender of an MPLS echo request MAY select the IPv6 destination address from the ::ffff:7f00/104 address block.

Section 1.2 of [RFC6790] lists several advantages of generating the entropy value by an ingress Label Switching Router (LSR) compared to when a transit LSR infers entropy using the information in the MPLS label stack or payload. This specification further clarifies the

following if multiple alternative paths for the given P2MP LSP Forwarding Equivalence Class (FEC) exist:

* The MultipointHead SHOULD use the Entropy Label [RFC6790] used for LSP Ping [RFC8029] to exercise those particular alternative paths; or

* The MultipointHead MAY use the UDP port number to possibly exercise those particular alternate paths.

3.2. Non-IP Encapsulation of Multipoint BFD

In some environments, the overhead of extra IP/UDP encapsulations may be considered burdensome, which makes the use of more compact Generic Associated Channel (G-ACh) [RFC5586] encapsulation attractive. Also, the validation of the IP/UDP encapsulation of a BFD Control packet in a P2MP BFD session may fail because of a problem related to neither the MPLS label stack nor BFD. Avoiding unnecessary encapsulation of P2MP BFD over an 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 (shown in Figure 1) MUST use the G-ACh Label (GAL) [RFC5586] at the bottom of the label stack followed by an Associated Channel Header (ACH). If a BFD Control packet in PW-ACH encapsulation (without IP/UDP Headers) is 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 Multipoint BFD Session (0x0013) (see Section 7.2). To provide the identity of the MultipointHead for the particular multipoint BFD session, a Source Address TLV, as defined in Section 4.1 of [RFC7212], MUST immediately follow a BFD Control packet. The use of other TLVs is outside the scope of this document.

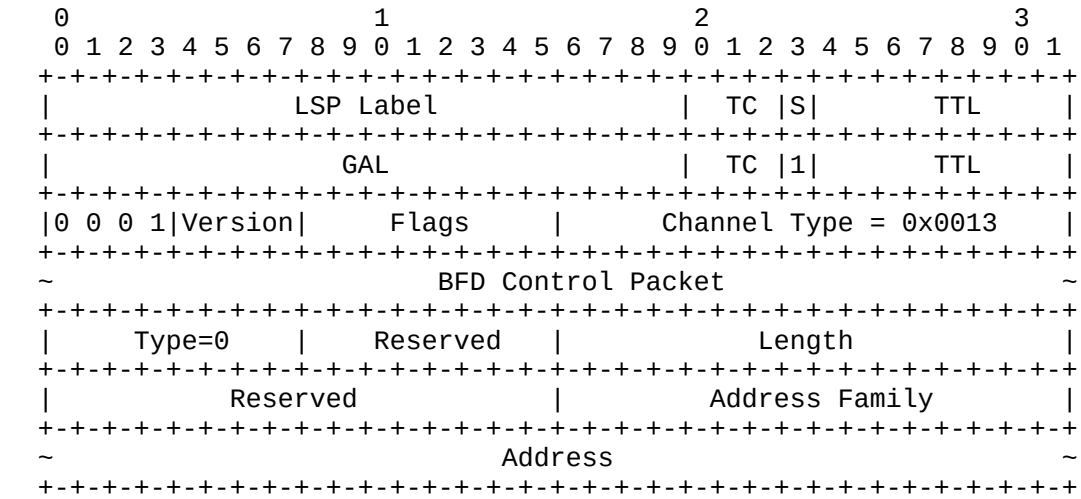

Figure 1: Non-IP Encapsulation for Multipoint BFD over a Multicast MPLS LSP

The fields in Figure 1 are interpreted as follows:

* The top three four-octet words are defined in [RFC5586].

* The BFD Control Packet field is defined in [RFC5880].

* All the remaining fields are defined in Section 4.1 of [RFC7212].

4. Bootstrapping Multipoint BFD

4.1. LSP Ping

LSP Ping is the part of the on-demand OAM toolset used 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 both egress and ingress endpoints.

LSP Ping, as defined in [RFC6425], MAY be used to bootstrap MultipointTail. If LSP Ping is used, it MUST include the Target FEC Stack TLV [RFC8029] and the BFD Discriminator TLV [RFC5884]. For the case of P2MP MPLS LSP, the Target FEC Stack TLV MUST use sub-TLVs defined in Section 3.1 of [RFC6425]. For the case of P2MP SR policy with an SR-MPLS data plane, an implementation of this specification MUST follow the procedures defined in [RFC8287]. Setting the value of the Reply Mode field to "Do not reply" [RFC8029] for the LSP Ping to bootstrap the MultipointTail of the P2MP BFD session is RECOMMENDED. Indeed, because BFD over a multipoint network uses BFD Demand mode, the MPLS echo reply from a tail has no useful information to convey to the head, unlike in the case of BFD over a P2P MPLS LSP [RFC5884]. A MultipointTail that receives an LSP Ping that includes the BFD Discriminator TLV MUST do the following:

* validate the LSP Ping;

* associate the received BFD Discriminator value with the P2MP LSP;

* create a P2MP BFD session and set bfd.SessionType = MultipointTail as described in [RFC8562]; and

* use the source IP address of the 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. Control Plane

The BFD Discriminator attribute MAY be used to bootstrap a multipoint BFD session on a tail, following the format and procedures given in Section 3.1.6 of [RFC9026].

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

[RFC8562] defines how BFD Demand mode can be used in multipoint networks. When applied in MPLS, the procedures specified in [RFC8562] allow an egress LSR to detect a failure in the part of the P2MP MPLS LSP from the ingress LSR to that egress LSR. The ingress LSR is not aware of the state of the P2MP LSP. [RFC8563], using mechanisms defined in [RFC8562], defines the behavior of an active tail. An active tail might notify the head of the detected failure and respond 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 of [RFC8563]) and is the simplest of the methods described in [RFC8563]. The use of this method in BFD over P2MP MPLS LSP is discussed in this document. Analysis of other methods for a head to learn of the state of the P2MP MPLS LSP is outside the scope of this document.

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

* On an ingress LSR:
  - bfd.SessionType is MultipointHead.
  - bfd.RequiredMinRxInterval is nonzero, allowing egress LSRs to send BFD Control packets.
* On an egress LSR:
  - bfd.SessionType is MultipointTail.
  - bfd.SilentTail is set to zero.

Section 5.2.1 of [RFC8563] notes that "the tail sends unsolicited BFD packets in response to the detection of a multipoint path failure" but does not provide specifics about the information in the packets or the frequency of transmissions. The procedure for an active tail with unsolicited notifications for P2MP MPLS LSP is defined below.

Upon detecting the failure of the P2MP MPLS LSP, an egress LSR sends a BFD Control packet with the following settings:

* The Poll (P) bit is set.
* The Status (Sta) field is set to the Down value.
* The Diagnostic (Diag) field is set to the 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.

The BFD Control packet MAY be encapsulated in IP/UDP with the destination IP address of the ingress LSR and the UDP destination port number set to 4784 per [RFC5883]. If non-IP encapsulation is used, then a BFD Control packet is encapsulated using PW-ACH encapsulation (without IP/UDP Headers) with Channel Type 0x0007 [RFC5885].

The BFD Control packets are transmitted at the rate of one per second until either 1) the egress LSA receives a control packet from the ingress LSR that is valid for this BFD session and has the Final (F) bit set or 2) the defect condition clears. However, to improve the likelihood of notifying the ingress LSR of the failure of the P2MP MPLS LSP, the egress LSR SHOULD initially transmit three BFD Control packets (as defined above) in short succession. The actual transmission of the periodic BFD Control packet MUST be jittered by up to 25% within one-second intervals. Thus, the interval MUST be reduced by a random value of 0 to 25%, to reduce the possibility of congestion on the ingress LSR’s data and control planes.

As described above, an ingress LSR that has received the BFD Control packet sends the unicast IP/UDP encapsulated BFD Control packet with the Final (F) bit set to the egress LSR. In some scenarios (e.g., when a P2MP LSP is broken close to its root and the number of egress LSRs is significantly large), the root might receive a large number of notifications. The notifications from leaves to the root will not use resources allocated for the monitored multicast flow and, as a result, will not congest that particular flow, although they may negatively affect other flows. However, the control plane of the ingress LSR might be congested by the BFD Control packets transmitted

by egress LSRs and the process of generating unicast BFD Control packets, as noted above. To mitigate that, a BFD implementation that supports this specification is RECOMMENDED to use a rate limiter of received BFD Control packets passed to the ingress LSR's control plane for processing.

6. Security Considerations

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

Also, BFD for P2MP MPLS LSPs MUST follow the requirements listed in Section 4.1 of [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. The operator MAY consider the size of the packet the MultipointHead transmits periodically as using IP/UDP encapsulation, which adds up to 28 octets (more than 50% of the BFD Control packet length) compared to G-ACh encapsulation.

7. IANA Considerations

7.1. IPv6 Special-Purpose Address

IANA has allocated the following in the "IANA IPv6 Special-Purpose Address Registry" [IANA-IPv6-REG]:

Address Block: 100:0:0:1::/64
Name: Dummy IPv6 Prefix
RFC: RFC 9780
Allocation Date: 2025-04
Termination Date: N/A
Source: True
Destination: False
Forwardable: False
Globally Reachable: False
Reserved-by-Protocol: False

7.2. MPLS Generalized Associated Channel (G-ACh) Type

IANA has allocated the following value in the "MPLS Generalized Associated Channel (G-ACh) Types" registry [IANA-G-ACh-TYPES].

Value	Description	Reference
0x0013	Multipoint BFD Session	RFC 9780

Table 1: Multipoint BFD Session G-ACh Type

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>.

[RFC5885] Nadeau, T., Ed. and C. Pignataro, Ed., "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, DOI 10.17487/RFC5885, June 2010, <https://www.rfc-editor.org/info/rfc5885>.

[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>.

[RFC8287] Kumar, N., Ed., Pignataro, C., Ed., Swallow, G., Akiya, N., Kini, S., and M. Chen, "Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes", RFC 8287, DOI 10.17487/RFC8287, December 2017, <https://www.rfc-editor.org/info/rfc8287>.

[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>.

[RFC9026] Morin, T., Ed., Kebler, R., Ed., and G. Mirsky, Ed., "Multicast VPN Fast Upstream Failover", RFC 9026, DOI 10.17487/RFC9026, April 2021, <https://www.rfc-editor.org/info/rfc9026>.

8.2. Informative References

[IANA-G-ACh-TYPES] IANA, "MPLS Generalized Associated Channel (G-ACh) Types", <https://www.iana.org/assignments/g-ach-parameters>.

[IANA-IPv6-REG] IANA, "IANA IPv6 Special-Purpose Address Registry", <https://www.iana.org/assignments/iana-ipv6-special-registry>.

[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>.

Acknowledgements

The authors sincerely appreciate the comments received from Andrew Malis, Italo Busi, and Shraddha Hegde. The authors also appreciate the thought-stimulating questions from Carlos Pignataro.

Authors' Addresses

Greg Mirsky
Ericsson
Email: gregimirsky@gmail.com

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com

Donald Eastlake 3rd
Independent
2386 Panoramic Circle
Apopka, FL 32703
United States of America
Email: d3e3e3@gmail.com