

Internet Engineering Task Force (IETF)
Request for Comments: 9612
Category: Experimental
ISSN: 2070-1721

G. Mirsky
Ericsson
J. Tantsura
NVIDIA
I. Varlashkin
Google
M. Chen
Huawei
July 2024

Bidirectional Forwarding Detection (BFD) Reverse Path for MPLS Label Switched Paths (LSPs)

Abstract

Bidirectional Forwarding Detection (BFD) is expected to be able to monitor a wide variety of encapsulations of paths between systems. When a BFD session monitors an explicitly routed unidirectional path, there may be a need to direct the egress BFD peer to use a specific path for the reverse direction of the BFD session. This document describes an extension to the MPLS Label Switched Path (LSP) echo request that allows a BFD system to request that the remote BFD peer transmit BFD control packets over the specified LSP.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for examination, experimental implementation, and evaluation.

This document defines an Experimental Protocol for the Internet community. 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). Not all documents approved by the IESG are candidates for any level of Internet Standard; see 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/rfc9612>.

Copyright Notice

Copyright (c) 2024 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction
 - 1.1. Conventions Used in This document

1.1.1.	Terminology
1.1.2.	Requirements Language
2.	Problem Statement
3.	Control of the BFD Reverse Path
3.1.	BFD Reverse Path TLV
3.2.	Return Codes
3.3.	Failure Characterization
4.	Use Case Scenario
5.	Operational Considerations
6.	IANA Considerations
6.1.	BFD Reverse Path TLV
6.2.	Return Codes
7.	Security Considerations
8.	Normative References
	Acknowledgments
	Authors' Addresses

1. Introduction

[RFC5880], [RFC5881], and [RFC5883] established the Bidirectional Forwarding Detection (BFD) protocol for IP networks. [RFC5884] and [RFC7726] set rules for using BFD Asynchronous mode over MPLS Label Switched Paths (LSPs), while not defining means to control the path that an egress BFD system uses to send BFD control packets towards the ingress BFD system.

When BFD is used to detect defects of the traffic-engineered LSP, the path of the BFD control packets transmitted by the egress BFD system toward the ingress may be disjoint from the monitored LSP in the forward direction. The fact that BFD control packets are not guaranteed to follow the same links and nodes in both forward and reverse directions may be one of the factors contributing to false positive defect notifications (i.e., false alarms) at the ingress BFD peer. Ensuring that both directions of the BFD session use co-routed paths may, in some environments, improve the determinism of the failure detection and localization.

This document defines the BFD Reverse Path TLV as an extension to LSP ping [RFC8029] and proposes that it be used to instruct the egress BFD system to use an explicit path for its BFD control packets associated with a particular BFD session. IANA has registered this TLV in the "TLVs" registry defined by [RFC8029] (see Section 6.1). As a special case, forward and reverse directions of the BFD session can form a bidirectional co-routed associated channel.

The LSP ping extension described in this document was developed and implemented as a result of an operational experiment. The lessons learned from the operational experiment enabled the use of this extension between systems conforming to this specification. Further implementation is encouraged to better understand the operational impact of the mechanism described in the document.

1.1. Conventions Used in This document

1.1.1. Terminology

BFD: Bidirectional Forwarding Detection

FEC: Forwarding Equivalence Class

LSP: Label Switched Path

LSR: Label Switching Router

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

When BFD is used to monitor an explicitly routed unidirectional path (e.g., MPLS-TE LSP), BFD control packets in the forward direction would be in-band using the mechanism defined in [RFC5884]. However, the reverse direction of the BFD session would follow the shortest path route, which could be completely or partially disjoint from the forward path. This creates the potential for the failure of a disjoint resource on the reverse path to trigger a BFD failure detection, even though the forward path is unaffected.

If the reverse path is congruent with the forward path, the potential for such false positives is greatly reduced. For this purpose, this specification provides a means for the egress BFD peer to be instructed to use a specific path for BFD control packets.

3. Control of the BFD Reverse Path

To bootstrap a BFD session over an MPLS LSP, LSP ping [RFC8029] MUST be used with the BFD Discriminator TLV [RFC5884]. This document defines a new TLV, the BFD Reverse Path TLV, that can be used to carry information about the reverse path for the BFD session that is specified by the value in the BFD Discriminator TLV. The BFD Reverse Path TLV MAY contain zero or more sub-TLVs.

3.1. BFD Reverse Path TLV

The BFD Reverse Path TLV is an optional TLV within the LSP ping [RFC8029]. However, if used, the BFD Discriminator TLV MUST be included in an echo request message as well. If the BFD Discriminator TLV is not present when the BFD Reverse Path TLV is included, then it MUST be treated as a malformed echo request, as described in [RFC8029].

The BFD Reverse Path TLV carries information about the path onto which the egress BFD peer of the BFD session referenced by the BFD Discriminator TLV MUST transmit BFD control packets. The format of the BFD Reverse Path TLV is presented in Figure 1.

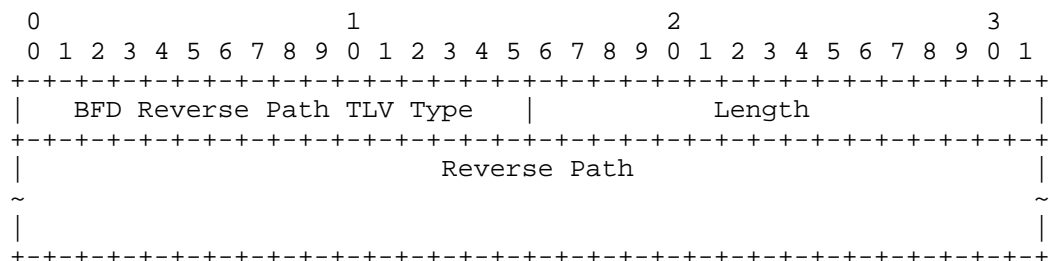


Figure 1: BFD Reverse Path TLV

BFD Reverse Path TLV Type:

This two-octet field has a value of 16384 (see Section 6).

Length:

This two-octet field defines the length in octets of the Reverse Path field.

Reverse Path:

This field contains zero or more sub-TLVs. Only non-multicast

Target FEC Stack sub-TLVs (already defined or to be defined in the future) for TLV Types 1, 16, and 21 in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry are permitted to be used in this field. Other sub-TLVs MUST NOT be used. (This implies that multicast Target FEC Stack sub-TLVs, e.g., the Multicast P2MP LDP FEC Stack sub-TLV and the Multicast MP2MP LDP FEC Stack sub-TLV, are not permitted in the Reverse Path field.)

If the egress LSR finds a multicast Target FEC Stack sub-TLV, it MUST send an echo reply with the received BFD Reverse Path TLV and BFD Discriminator TLV and set the Return Code to 192 ("Inappropriate Target FEC Stack sub-TLV present") (see Section 3.2). The BFD Reverse Path TLV includes zero or more sub-TLVs. However, the number of sub-TLVs in the Reverse Path field MUST be limited. The default limit is 128 sub-TLV entries, but an implementation MAY be able to control that limit. An empty BFD Reverse Path TLV (i.e., a BFD Reverse Path TLV with no sub-TLVs) is used to withdraw any previously set reverse path for the BFD session identified in the BFD Discriminator TLV. If no sub-TLVs are found in the BFD Reverse Path TLV, the egress BFD peer MUST revert to using the decision based on local policy, i.e., routing over an IP network, as described in Section 7 of [RFC5884].

If the egress peer LSR cannot find the path specified in the BFD Reverse Path TLV, it MUST send an echo reply with the received BFD Discriminator TLV and BFD Reverse Path TLV and set the Return Code to 193 ("Failed to establish the BFD session. The specified reverse path was not found.") (see Section 3.2). If an implementation provides additional configuration options, these can control actions at the egress BFD peer, including when the path specified in the BFD Reverse Path TLV cannot be found. For example, if the egress peer LSR cannot find the path specified in the BFD Reverse Path TLV, it MAY establish the BFD session over an IP network, as defined in [RFC5884]. Note that the Return Code required by the "MUST" clause in this paragraph does not preclude the session from being established over a different path as discussed in the "MAY" clause.

The BFD Reverse Path TLV MAY be used in the process of bootstrapping the BFD session as described in Section 6 of [RFC5884]. A system that supports this specification MUST support using the BFD Reverse Path TLV after the BFD session has been established. If a system that supports this specification receives an LSP ping with the BFD Discriminator TLV and no BFD Reverse Path TLV even though the reverse path for the specified BFD session was established according to the previously received BFD Reverse Path TLV, the egress BFD peer MUST transition to transmitting periodic BFD Control messages as described in Section 7 of [RFC5884]. If a BFD system that received an LSP ping with the BFD Reverse Path TLV does not support this specification, it will result in an echo response with the Return Code set to 2 ("One or more of the TLVs was not understood"), as described in Section 3 of [RFC8029].

3.2. Return Codes

This document defines the following Return Codes for the MPLS LSP echo reply:

"Inappropriate Target FEC Stack sub-TLV present" (192):

When a multicast Target FEC Stack sub-TLV is found in the received echo request, the egress BFD peer sends an echo reply with the Return Code set to 192 ("Inappropriate Target FEC Stack sub-TLV present") to the ingress BFD peer, as described in Section 3.1.

"Failed to establish the BFD session. The specified reverse path was not found." (193):

When a specified reverse path is unavailable, the egress BFD peer sends an echo reply with the Return Code set to 193 ("Failed to establish the BFD session. The specified reverse path was not found.") to the ingress BFD peer, as described in Section 3.1.

3.3. Failure Characterization

A failure detected by a BFD session that uses the BFD Reverse Path TLV could be due to a change in the FEC used in the BFD Reverse Path TLV. Upon detection of the network failure, the ingress BFD peer MUST transmit the LSP ping echo request with the Reply Path TLV [RFC7110] to verify whether the FEC is still valid. If the failure was caused by a change in the FEC used for the reverse direction of the BFD session, the ingress BFD peer MUST redirect the reverse path of the BFD session using another FEC in the BFD Reverse Path TLV and notify an operator.

4. Use Case Scenario

In the network presented in Figure 2, ingress LSR peer A monitors two tunnels to egress LSR peer H: A-B-C-D-G-H and A-B-E-F-G-H. To bootstrap a BFD session to monitor the first tunnel, ingress LSR peer A includes a BFD Discriminator TLV with a Discriminator value (e.g., foobar-1) [RFC7726]. Ingress LSR peer A includes a BFD Reverse Path TLV referencing the H-G-D-C-B-A tunnel to control the path from the egress LSR. To bootstrap a BFD session to monitor the second tunnel, ingress LSR peer A includes a BFD Discriminator TLV with a different Discriminator value (e.g., foobar-2) and a BFD Reverse Path TLV that references the H-G-F-E-B-A tunnel.

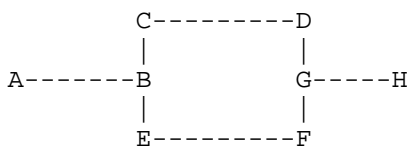


Figure 2: Use Case for BFD Reverse Path TLV

If an operator needs egress LSR peer H to monitor a path to ingress LSR peer A, e.g., the H-G-D-C-B-A tunnel, then by looking up the list of known reverse paths, it MAY find and use the existing BFD session.

5. Operational Considerations

When an explicit path is set as either Static or RSVP-TE LSP, corresponding sub-TLVs (defined in [RFC7110]) MAY be used to identify the explicit reverse path for the BFD session. If a particular set of sub-TLVs composes the Reply Path TLV [RFC7110] and does not increase the length of the Maximum Transmission Unit for the given LSP, that set can be safely used in the BFD Reverse Path TLV. If any of the sub-TLVs defined in [RFC7110] are used in the BFD Reverse Path TLV, then the periodic verification of the control plane against the data plane, as recommended in Section 4 of [RFC5884], MUST use the Reply Path TLV, as per [RFC7110], with that sub-TLV. By using the LSP ping with the Reply Path TLV, an operator monitors whether the reverse LSP is mapped to the same FEC as the BFD session at the egress BFD node. Selection and control of the rate of the LSP ping with the Reply Path TLV follows the recommendation in [RFC5884]:

| 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. An implementation MAY provide configuration
| options to control the rate of generation of the periodic LSP Ping
| Echo request messages.

Suppose an operator planned a network maintenance activity that

possibly affects the FEC used in the BFD Reverse Path TLV. In that case, the operator can avoid unnecessary disruption by using the LSP ping with a new FEC in the BFD Reverse Path TLV. But in some scenarios, proactive measures cannot be taken because the frequency of LSP ping messages is lower than the defect detection time provided by the BFD session. As a result, a change in the reverse-path FEC will first be detected as the BFD session's failure. An operator will be notified as described in Section 3.3.

6. IANA Considerations

6.1. BFD Reverse Path TLV

IANA has assigned the following value for the BFD Reverse Path TLV from the 16384-31739 range in the "TLVs" subregistry within the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry.

Type	TLV Name	Reference	Sub-TLV Registry
16384	BFD Reverse Path	RFC 9612	Only non-multicast sub-TLVs (already defined or to be defined in the future) in the "Sub-TLVs for TLV Types 1, 16, and 21" registry at https://www.iana.org/assignments/mpls-lsp-ping-parameters/mpls-lsp-ping-parameters.xml#sub-tlv-1-16-21 are permitted to be used in this field. Other sub-TLVs MUST NOT be used.

Table 1: New BFD Reverse Path TLV

6.2. Return Codes

IANA has assigned the following Return Code values from the 192-247 range in the "Return Codes" subregistry within the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry.

Value	Meaning	Reference
192	Inappropriate Target FEC Stack sub-TLV present	RFC 9612
193	Failed to establish the BFD session. The specified reverse path was not found.	RFC 9612

Table 2: New Return Codes

7. Security Considerations

Security considerations discussed in [RFC5880], [RFC5884], [RFC7726], [RFC8029], and [RFC7110] apply to this document.

The BFD Reverse Path TLV may be exploited as an attack vector by inflating the number of included sub-TLVs. The number of sub-TLVs MUST be limited to mitigate that threat. The default limit for the number of sub-TLVs is set to 128 (see Section 3.1). An implementation MAY use a mechanism to control that limit.

8. 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>>.
- [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>>.
- [RFC7110] Chen, M., Cao, W., Ning, S., Jounay, F., and S. Delord, "Return Path Specified Label Switched Path (LSP) Ping", RFC 7110, DOI 10.17487/RFC7110, January 2014, <<https://www.rfc-editor.org/info/rfc7110>>.
- [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>>.

Acknowledgments

The authors greatly appreciate the thorough reviews and helpful comments from Eric Gray and Carlos Pignataro. The authors much appreciate the help of Qian Xin, who provided information about the implementation of this specification.

Authors' Addresses

Greg Mirsky
Ericsson
Email: gregimirsky@gmail.com

Jeff Tantsura
NVIDIA
Email: jefftant.ietf@gmail.com

Ilya Varlashkin
Google
Email: imv@google.com

Mach(Guoyi) Chen
Huawei
Email: mach.chen@huawei.com