BFD Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 3, 2020

G. Mirsky
X. Min
ZTE Corp.

July 2, 2019

Extended Bidirectional Forwarding Detection
draft-mirmin-bfd-extended-01

Abstract

This document describes a mechanism to extend the capabilities of Bidirectional Forwarding Detection (BFD). These extensions enable BFD to measure performance metrics like packet loss and packet delay. Also, a method to perform lightweight on-demand authentication is defined in this specification.

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 3, 2020.

Copyright Notice

Table of Contents

1. Introduction

[RFC5880] provided the base specification of Bidirectional Detection (BFD) as the light-weight mechanism to monitor a path continuity between two systems and detect a failure in the data-plane. Since its introduction, BFD has been broadly deployed. There were several attempts to introduce new capabilities in the protocol, some more successful than others. One of the significant obstacles to extending BFD capabilities may be seen in the compact format of the BFD control message. This document introduces an Extended BFD control message and describes the use of the new format for new BFD capabilities.

2. Conventions used in this document

2.1. Terminology

BFD: Bidirectional Forwarding Detection

G-ACh Generic Associated Channel

HMAC Hashed Message Authentication Code

MTU Maximum Transmission Unit

PMTUD Path MTU Discovery

p2p: Point-to-Point

TLV Type, Length, Value

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. Extended BFD Control Message

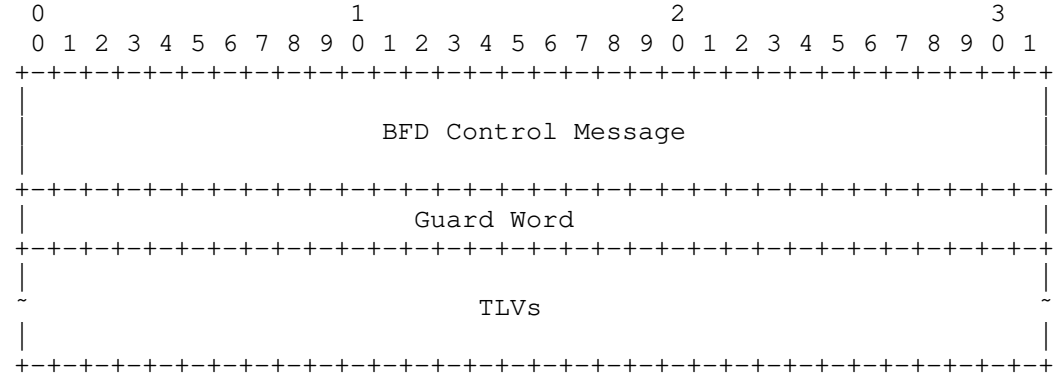

Figure 1: Extended BFD Control Message Format

where fields are defined as the following:

- BFD control message as defined [RFC5880].
- Guard word - four octets long field to identify the role of the BFD system - sender or responder.
- TLVs - variable-length field that contains commands and/or data encoded as type-length-value (TLV).

If an Extended BFD control message is encapsulated in IP/UDP, the value of the Total Length in the IP header includes the length of the Extended BFD control message while the value of the Length field of the BFD control message equals the value as defined in [RFC5880]. If an Extended BFD control message is to be used over Generic Associated

Channel (G-ACh), e.g., [RFC6428] new code point for G-ACh may be allocated.

Figure 2 displays the generic TLV format.

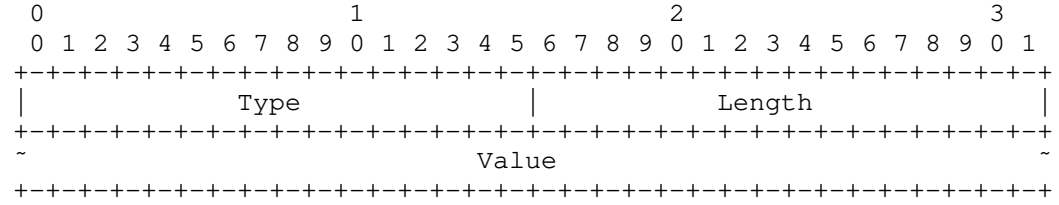

Figure 2: General Type-Length-Value Encoding

where fields are defined as the following:

- Type - two octets long field that defines the encoding of the Value field
- Length - two octets long field equals length on the Value field in octets.
- Value - depends on the Type.

TLVs may be included within other TLVs, in which case the former TLVs are referred to as sub-TLVs. Sub-TLVs have independent types.

3.1. Extended BFD Capability Negotiation

A BFD system also referred to as a node in this document, that supports Extended BFD first MUST discover whether other nodes in the given BFD session support the Extended BFD. The node MUST send Extended BFD control message initiating the Poll Sequence as defined in [RFC5880]. If the remote system fails to respond with the Extended BFD control message and the Final flag set, then the initiator node MUST conclude that the BFD peer does not support the use of the Extended BFD control messages.

The first Extended BFD control message initiating the Poll Sequence SHOULD include the Capability TLV that lists capabilities that may be used at some time during the lifetime of the BFD session. The format of the Capability TLV and the capabilities that use the Extended BFD control message are presented in Figure 3.

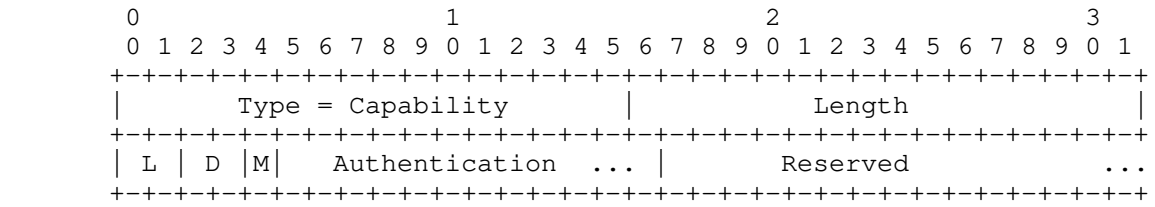

Figure 3: Capability TLV Format

where fields are defined as the following:

- Type - TBA1 allocated by IANA in Section 4

- Length - two octets long field equals length on the Capability field in octets. The value of the Length field MUST be a multiple of 4.

- Loss - two bits size field. The least significant of two bits is set if the node is capable of measuring packet loss using periodically transmitted Extended BFD control message. The most significant of two bits is set if the node is capable of measuring packet loss using the Poll Sequence with Extended BFD control message.

- Delay - two bits size field. The least significant of two bits is set if the node is capable of measuring packet delay using periodically transmitted Extended BFD control message. The most significant of two bits is set if the node is capable of measuring packet delay using the Poll Sequence with Extended BFD control message.

- MTU- one-bit size field. Set if the node is capable of using the Extended BFD control message in Path MTU Discovery (PMTUD). [Ed.note: Definition of the PMTUD using the Extended BFD control message is for further version.]

- (Lightweight) Authentication - variable-length field. The Authentication field is used by a BFD system to advertise its lightweight authentication capabilities. The format and the use of the Authentication field are defined in Section 3.5.1.

- Reserved - MUST be zeroed on transmission and ignored on receipt. The Reserved field is zero-padded to align the length of the Capability TLV to a 4-octet boundary.

The remote BFD node that supports this specification MUST respond to the Capability TLV with the Extended BFD control message that

includes the Capability TLV listing capabilities the responder supports. The responder MUST set the Final flag in the Extended BFD control message.

3.2. Padding TLV

Padding TLV MAY be used to generate Extended BFD control packets of the desired length.

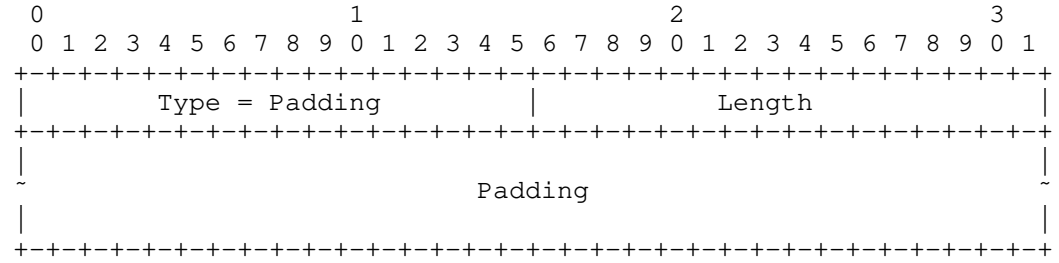

Figure 4: Padding TLV Format

where fields are defined as the following:

- Type - TBA1 allocated by IANA in Section 4
- Length - two octets long field equals length on the Padding field in octets.
- Padding - variable-length field. MUST be zeroed on transmit and ignored on receipt.

3.3. Diagnostic TLV

Diagnostic TLV MAY be used to characterize the result of the last requested operation.

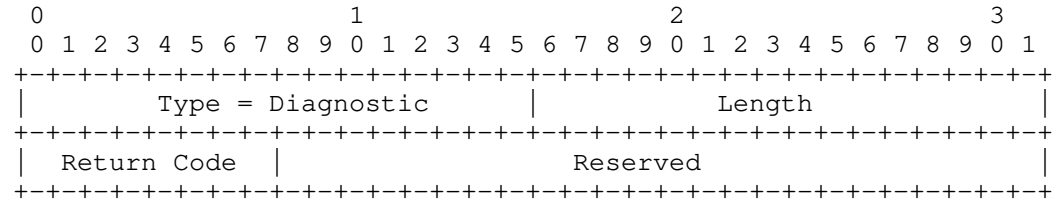

Figure 5: Diagnostic TLV Format

where fields are defined as the following:

- Type - TBA6 allocated by IANA in Section 4.

- Length - MUST be set to four.

- Return Code - eight bits-long field. The responding BFD system can set it to one of the values defined in Section 4.3.

- Reserved - 24 bits-long field. MUST be zeroed on transmit and ignored on receipt.

3.4. Performance Measurement with Extended BFD Control Message

Loss measurement, delay measurement, and loss/delay measurement messages can be used in the Extended BFD control message to support one-way and round-trip measurements. All the messages are encapsulated as TLVs with Type values allocated by IANA, Section 4.

The sender MAY use the Performance Metric TLV (presented in Figure 6) to measure performance metrics and obtain the measurement report from the receiver.

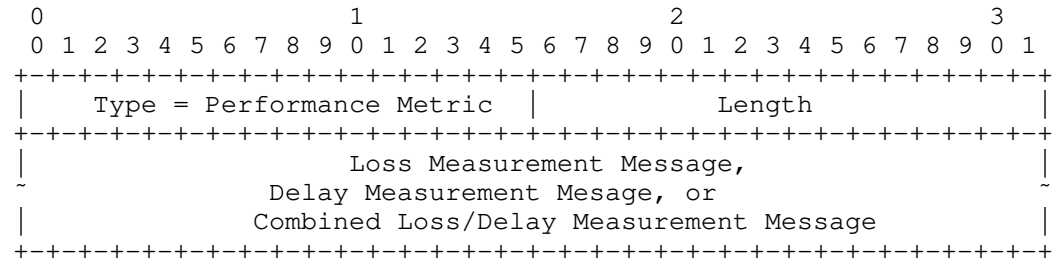

Figure 6: Performance Metric TLV Format

where fields are defined as the following:

- Type - TBA3 through TBA5 allocated by IANA in Section 4 as follows:

  * TBA3 - Loss Measurement Type;

  * TBA4 - Delay Measurement Type;

  * TBA5 - Combined Loss/Delay Measurement Type

- Length - two octets long field equals length on the Metric sub-TLVs field in octets. The value of the Length field MUST be a multiple of 4.

- Value - various performance metrics measured either directly or using synthetic methods accordingly using the messages defined in Sections 3.1 through 3.3 [RFC6374].

To perform one-way loss and/or delay measurement, the BFD node MAY periodically transmit the Extended BFD message with one of the TLVs listed above in Asynchronous mode. To perform synthetic loss measurement, the sender MUST monotonically increment the counter of transmitted test packets. Also, direct-mode loss measurement, as described in [RFC6374], is supported. Procedures to negotiate and manipulate transmission intervals defined in Sections 6.8.2 and 6.8.3 in [RFC5880] SHOULD be used to control the performance impact of using the Extended BFD for performance measurement in the particular BFD session.

To measure the round-trip loss and/or delay metrics the BFD node transmits the Extended BFD control message with the Performance Metric TLV with the Poll flag set. Before the transmission of the Extended BFD control message with the Performance Metric TLV, the receiver MUST clear the Poll flag and set the Final flag.

3.5. Lightweight Authentication

Using BFD without any security measures, for example, by exchanging BFD control packets without authentication, increases the risk of an attack, especially over multiple nodes. Thus, using BFD without security measures may cause false positive as well as false negative defect detection situations. In the former, an attacker may spoof BFD control packets pretending to be a remote peer and can thus view the BFD session operation even though the real path had failed. In the latter, the attacker may spoof altered BFD control message indicating that the BFD session is un-operational even though the path and the remote BFD peer operate normally.

BFD technology[RFC5880] includes optional authentication protection of BFD control packets to minimize the chances of attacks in a networking system. However, at least some of the supported authentication protocols do not provide sufficient protection in modern networks. Also, current BFD technology requires authentication of each and every BFD control packet. Such an authentication requirement can put a computational burden on networking devices, especially in the Asynchronous mode, at least because authenticating each BFD control packet can require substantial computing resources to support packet exchange at high rates.

This specification defines a lightweight on-demand mode of authentication for a BFD session. The lightweight authentication is

an optional mode that can be used when the BFD Authentication [RFC5880] is not in use (bfd.AuthType is zero). The mechanism includes negotiation (Section 3.5.1) and on-demand authentication (Section 3.5.2) phases. During the former, BFD peers advertise supported authentication capabilities and independently select the commonly supported mode of the authentication. In the authentication phase, each BFD system transmits, at certain events and periodically, authenticated BFD control packets in Poll Sequence.

3.5.1. Lightweight Authentication Mode Negotiation

Figure 7 displays the format of the Authentication field that is part of the Capability Encoding:

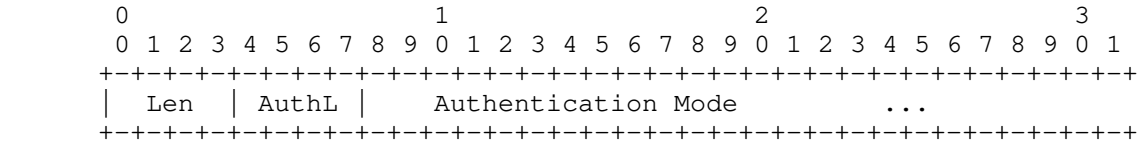

Figure 7: Lightweight Authentication Capability Field Format

where fields are defined as the following:

- Len (Length) - four-bits long field. The value of the Length field is equal to the length of the Authentication field, including the Length, in octets.

- AuthL (Authentication Length) - four bits size field. The value of the field is, in four octets long words, the longest authentication signature the BFD system is capable of supporting for any of the methods advertised in the AuthMode field.

- Authentication Mode - variable-length field. It is a bit-coded field that a BFD system uses to list modes of lightweight authentication it supports.

A BFD system uses Capability TLV, defined in Section 3.1, to discover the commonly supported mode of the Lightweight Authentication. The system sets the values in the Authentication field according to properly reflect its authentication capabilities. The BFD system transmits the Extended BFD control packet with Capability TLV as the first in a Poll Sequence. The remote BFD system that supports this specification receives the Extended BFD control packet with the advertised Lightweight Authentication modes and stores information locally. The system responds with the advertisement of its Lightweight Authentication capabilities in the Extended BFD control packet with the Final flag set. Each BFD system uses local and

received information about Lightweight Authentication capabilities to deduce the commonly supported modes and selects from that set the one that uses the strongest authentication with the longest signature. If the common set is empty, i.e., none of supported by one BFD system authentication method is supported by another, an implementation MUST reflect this in its operational state and SHOULD notify an operator.

3.5.2. Using Lightweight Authentication

After BFD peers select an authentication mode for using in Lightweight Authentication each BFD system MUST use it to authenticate each Extended BFD control packet transmitted as part of a Poll Sequence using Lightweight Authentication TLV presented in Figure 8.

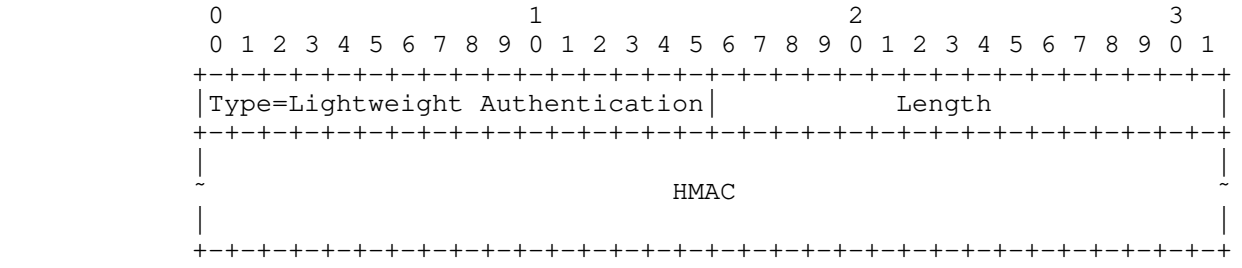

Figure 8: Lightweight Authentication TLV Format

where fields are defined as the following:

- Type - TBA8 allocated by IANA in Section 4
- Length - two octets long field equals length on the HMAC (Hashed Message Authentication Code) field in octets. The value of the Length field MUST be a multiple of 4.
- HMAC - the hash value calculated on the preceding Extended BFD control packet data.

The Lightweight Authentication TLV MUST be the last TLV in an Extended BFD control packet. Padding TLV (Section 3.2) MAY be used to align the length of the Extended BFD control packet, excluding the Lightweight Authentication TLV, at multiple of 16 boundary.

The BFD system that receives the Extended BFD control packet with the Lightweight Authentication TLV MUST first validate the .authentication by calculating the hash over the Extended BFD control packet. If the validation succeeds, the receiver MUST transmit the Extended BFD control packet with the Final flag set and the value of

the Return code field in Diagnostic TLV set to None value (Table 5). If the validation of the lightweight authentication fails, then the BFD system MUST transmit the Extended BFD control packet with the Final flag set and the value of the Return Code field of the Diagnostic TLV set to Lightweight Authentication failed value (Table 5). The BFD system MUST have a control policy that defines actions when the system receives the Lightweight Authentication failed return code.

4. IANA Considerations

4.1. Extended BFD Message Types

IANA is requested to create the Extended BFD Message Types registry. All code points in the range 1 through 32759 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC8126]. Code points in the range 32760 through 65279 in this registry shall be allocated according to the "First Come First Served" procedure as specified in [RFC8126]. Remaining code points are allocated according to Table 1:

Value	Description	Reference
0	Reserved	This document
1- 32767	Mandatory TLV, unassigned	IETF Review
32768 - 65279	Optional TLV, unassigned	First Come First Served
65280 - 65519	Experimental	This document
65520 - 65534	Private Use	This document
65535	Reserved	This document

Table 1: Extended BFD Type Registry

This document defines the following new values in Extended BFD Type registry:

Value	Description	Reference
TBA1	Padding	This document
TBA2	Capability	This document
TBA3	Loss Measurement	This document
TBA4	Delay Measurement	This document
TBA5	Combined Loss/Delay Measurement	This document
TBA6	Diagnostic	This document
TBA8	Lightweight Authentication	This document

Table 2: Extended BFD Types

4.2. Lightweight Authentication Modes

IANA is requested to create a Lightweight Authentication Modes registry. All code points in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC8126].

This document defines the following new values in the Lightweight Authentication Modes registry:

Bit Position	Value	Description	Reference
0	0x1	Keyed SHA-1	This document
1	0x2	Meticulous Keyed SHA-1	This document
2	0x4	SHA-256	This document

Table 3: Lightweight Authentication Modes

4.3. Return Codes

IANA is requested to create the Extended BFD Return Codes registry. All code points in the range 1 through 250 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC8126]. Remaining code points are allocated according to Table 4:

Value	Description	Reference
0	Reserved	This document
1- 250	Unassigned	IETF Review
251-253	Experimental	This document
254	Private Use	This document
255	Reserved	This document

Table 4: Extended BFD Return Codes Registry

This document defines the following new values in Extended BFD Return Codes registry:

Value	Description	Reference
0	None	This document
1	One or more TLVs was not understood	This document
2	Lightweight Authentication failed	This document

Table 5: Extended BFD Return Codes

5. Security Considerations

This document does not introduce new security aspects but inherits all security considerations from [RFC5880], [RFC6428], and [RFC6374].

6. 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>.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC6428] Allan, D., Ed., Swallow, G., Ed., and J. Drake, Ed., "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428, DOI 10.17487/RFC6428, November 2011, <https://www.rfc-editor.org/info/rfc6428>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

Appendix A. Acknowledgements

TBD

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Xiao Min
ZTE Corp.

Email: xiao.min2@zte.com.cn