Workgroup: Network Work group
Internet-Draft: draft-ietf-bier-ping-10
Published: 19 May 2023
Intended Status: Standards Track
Expires: 20 November 2023
Authors: N. Kumar, Cisco Systems, Inc.
C. Pignataro, Cisco Systems, Inc.
N. Akiya, Big Switch Networks
L. Zheng, Individual Contributor
M. Chen, Huawei Technologies
G. Mirsky, Ericsson

BIER Ping and Trace

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "BIER domain" without requiring intermediate routers to maintain any multicast-related per- flow state. BIER also does not require any explicit tree-building protocol for its operation. A multicast data packet enters a BIER domain at a "Bit-Forwarding Ingress Router" (BFIR), and leaves the BIER domain at one or more "Bit-Forwarding Egress Routers" (BFERs). The BFIR router adds a BIER header to the packet. The BIER header contains a bit-string in which each bit represents exactly one BFER to forward the packet to. The set of BFERs to which the multicast packet needs to be forwarded is expressed by setting the bits that correspond to those routers in the BIER header.

This document describes the mechanism and basic BIER OAM packet format that can be used to perform failure detection and isolation on the BIER data plane.

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 20 November 2023.

Table of Contents

1. Introduction
2. Conventions used in this document
  2.1. Terminology
  2.2. Requirements Language
3. BIER OAM
  3.1. BIER OAM message format
  3.2. Return Code
  3.3. BIER OAM TLV
    3.3.1. Original SI-BitString TLV
    3.3.2. Target SI-BitString TLV
    3.3.3. Incoming SI-BitString TLV
    3.3.4. Downstream Mapping TLV
    3.3.5. Responder BFER TLV
    3.3.6. Responder BFR TLV
    3.3.7. Upstream Interface TLV
  3.4. Multipath Entropy Data Encoding
4. Procedures
  4.1. BIER OAM Processing
  4.2. Per BFER ECMP Discovery
  4.3. Sending BIER Echo Request
  4.4. Receiving BIER Echo Request
  4.5. Sending Echo Reply
  4.6. Receiving Echo Reply
5. IANA Considerations
  5.1. UDP Port Number
  5.2. BIER OAM Parameters
  5.3. BIER OAM Message Type
  5.4. BIER Echo Request/Echo Reply Parameters
    5.4.1. BIER Echo Request/Echo Reply Message Types
    5.4.2. BIER Echo Reply Modes
    5.4.3. BIER Echo Return Codes
  5.5. TLVs

6. Security Considerations
7. Acknowledgement
8. References
  8.1. Normative References
  8.2. Informative References
Authors' Addresses

1. Introduction

[RFC8279] introduces and explains BIER architecture that provides optimal multicast forwarding through a "BIER domain" without requiring intermediate routers to maintain any multicast-related per- flow state. BIER also does not require any explicit tree-building protocol for its operation. A multicast data packet enters a BIER domain at a "Bit-Forwarding Ingress Router" (BFIR), and leaves the BIER domain at one or more "Bit-Forwarding Egress Routers" (BFERs). The BFIR router adds a BIER header to the packet. The BIER header contains a bit-string in which each bit represents exactly one BFER to forward the packet to. The set of BFERs to which the multicast packet needs to be forwarded is expressed by setting the bits that correspond to those routers in the BIER header.

Operations, Administration, and Maintenance (OAM) mechanisms are expected to support the detection of network failures. After the detection, operators localize and characterize the network defect. A query-based tool, e.g., ICMP [RFC0792] and LSP Ping [RFC8029], [RFC6425], is broadly used to detect and localize a network defect. Additionally, this mechanism can be used to check the consistency between the data and control planes. This document describes the mechanism and basic BIER OAM packet format that can be used to perform failure detection and isolation on the BIER data plane without any dependency on other layers, like the IP layer.

2. Conventions used in this document

2.1. Terminology

BFER - Bit-Forwarding Egress Router

BFIR - Bit-Forwarding Ingress Router

BFR - Bit-Forwarding Router

BIER - Bit Index Explicit Replication

DDMAP - Downstream Detailed Mapping TLV

ECMP - Equal Cost Multi-Path

OAM - Operation, Administration, and Maintenance

SI - Set Identifier

QTF - Querier Timestamp Format

RTF - Responder Timestamp Format

NTP - Network Time Protocol

MTU - Maximum Transmission Unit

DA - Downstream Address

DIA - Downstream Interface Address

DoS - Denial-of-Service

PTP - Precision Time Protocol

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. BIER OAM

BIER OAM is defined to stay within the BIER layer by directly following the BIER header without mandating the need for an IP header. [RFC8296] defines a 4-bit field as "Proto" to identify the payload following the BIER header. When the payload is BIER OAM, the "Proto" field will be set to 5 as defined in [RFC8296]

3.1. BIER OAM message format

The BIER OAM packet header format that follows the BIER header is displayed in Figure 1.

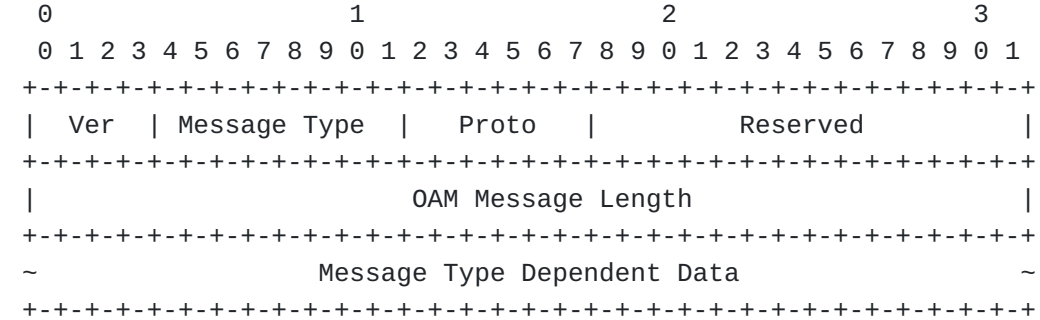

Figure 1: BIER OAM Header

Ver - a four-bit field that indicates the current version of the BIER OAM header. The current value is 1. The version number is to be incremented whenever a change is made that affects the ability of an implementation to parse or process the BIER OAM header correctly. For example, if syntactic or semantic changes are made to any of the fixed fields.

Message Type - a six-bit field that identifies OAM protocol. Values defined in this document are as follows:

Value	Description
1	Echo Request
2	Echo Reply

Proto - a six-bit field. This field is used to define if there is any data packet immediately following the OAM payload, which may be used by a hybrid OAM [RFC7799]. This field MUST be set to 0 if there is no data packet following the OAM payload. Value is one from the IANA registry "BIER Next Protocol Identifiers" [RFC8296].

Reserved - a fourteen-bit field. The value MUST be zeroed on transmission and ignored on receipt.

OAM Message Length - a two-octet field that reflects the length of the OAM message, including the header and the Messsage Type Dependent Data.

Message Type Dependant Data - a variable-length field that includes the OAM message identified by the value of the Message Type filed.

The Echo Request/Reply header format displayed in Figure 2

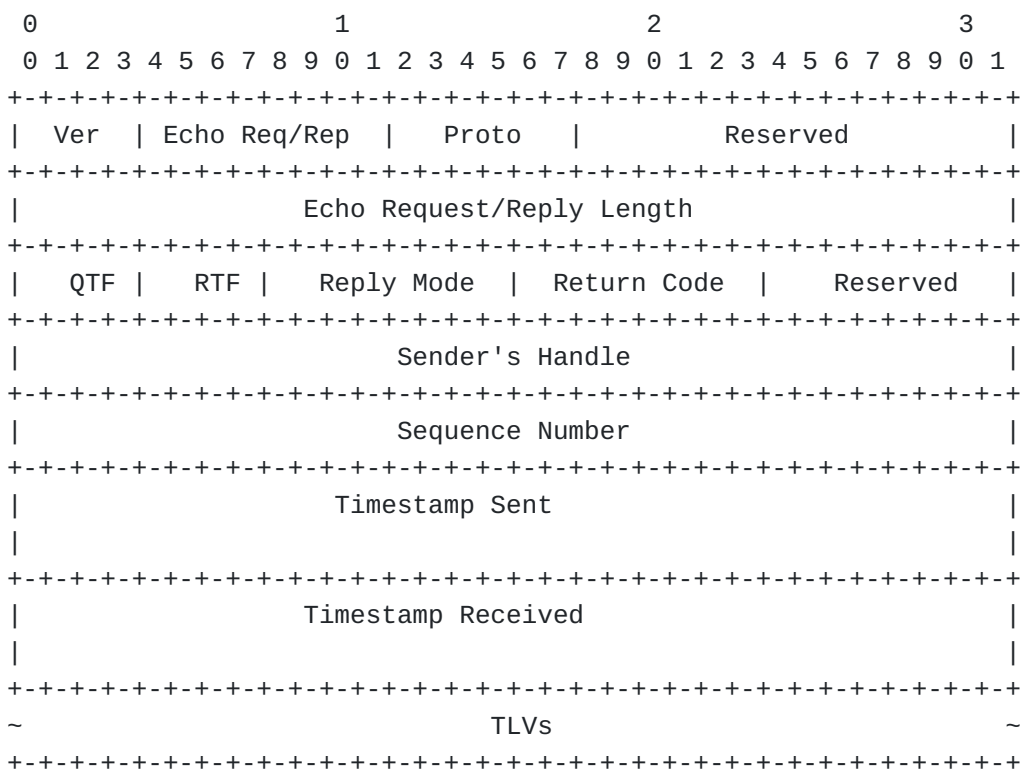

Figure 2: BIER Echo Request/Reply Format

Proto field MUST be set to 0 for Echo Request/Reply header.

QTF (Querier Timestamp Format) - a four-bit field. When the field is set to 2, the Timestamp Sent field is (in seconds and microseconds, according to the Initiator's clock) in NTP format [RFC5905]. When the value of the QTF field is 3, the Timestamp Sent's format is the IEEE 1588-2008 (1588v2) Precision Time Protocol (PTP) [IEEE.1588.2008] format. Any other value MAY be considered as a sanity check failure.

RTF (Responder Timestamp Format) - a four-bit field. When the field is set to 2, the Timestamp Received field is (in seconds and microseconds, according to the Initiator's clock) in NTP format [RFC5905]. When filed's value is 3, the format of the Timestamp Received is as defined in IEEE 1588-2008 (1588v2) Precision Time Protocol [IEEE.1588.2008]. Any other value MAY be considered as a sanity check failure.

The sender of the BIER Echo Request might receive the BIER Echo Reply with RTF different from the Sender's QTF, Thus, the Sender MUST be able to interpret both timestamp formats, i.e., NTP [RFC5905] and PTP [IEEE.1588.2008].

The Reply Mode - a one-octet field. The value MUST be set to one of the following values:

Value	Description
1	Do not Reply
2	Reply via IPv4/IPv6 UDP packet
3	Reply via BIER packet

When Reply Mode is set to 1, the receiver will not send any reply. This mode can be used for unidirectional path validation. When the Reply Mode is set to 2, the Responder Bit-Forwarding Router (BFR) encapsulates the Echo reply payload with the IP/UDP header. When the Initiator intends to validate the return BIER path, the Reply Mode will be set to 3 so that the Responder BFR will encapsulate the Echo Reply with the BIER header.

Return Code - a one-octet field. The value MUST be set to zero if the Type is "BIER Echo Request". The value of the Return Code filed MUST be set to one of the values defined in Section 3.2, if the Type is "BIER Echo Reply".

Reserved - a one-octet field. The Reserved field MUST be zeroed on transmit and MUST be ignored on receipt.

Sender's Handle - a four-octet field. The Sender's Handle is filled by the Initiator, and returned unchanged by responder BFR. This value can be used for matching the replies to the request.

Sequence Number - a four-octet field. The value of the field is assigned by the Initiator and can be used to detect any missed replies.

Timestamp - each field (Sent and Received) is an eight-octet field. The Timestamp Sent is the time when the Echo Request is sent. The Timestamp Received in Echo Reply is the time (accordingly to responding BFR clock) that the corresponding Echo Request was received. The format depends on the QTF/RTF value.

TLVs - Carries the TLVs as defined in Section 3.3.

3.2. Return Code

The responder uses the Return Code field to reply with a validity check or other error message to Initiator. It does not carry any meaning in Echo Request and MUST be set to zero. The Return Code can be one of the following:

Value	Value Meaning
0	No return code
1	Malformed Echo Request received
2	One or more of the TLVs is not supported
3	Replying BFR is the only BFER in header BitString
4	Replying BFR is one of the BFERs in header BitString
5	Packet-Forward-Success
6	Invalid Multipath Info Request
8	No matching entry in the forwarding table
9	Set-Identifier Mismatch
10	DDMAP Mismatch

"No return code" will be used by Initiator in the Echo Request. This value MUST NOT be used in Echo Reply.

"Malformed Echo Request received" will be used by any BFR if the received Echo Request packet is not properly formatted.

When a receiver does not support any TLV included in the Echo Request, the Return code will be set to "One or more of the TLVs is not supported" carrying the respective TLVs.

When the received header BitString in the Echo Request packet contains only its Bit-ID, "Replying BFR is the only BFER in header BitString" is set in the reply. This value implies that the receiver is BFER, and the packet is not forwarded to any more neighbors.

When the received header BitString in the Echo Request packet contains its Bit-ID in addition to other Bit-IDs, "Replying BFR is one of the BFERs in header BitString" is set in the reply. This value implies that the responder is a BFER and the packet is further forwarded to one or more neighbors.

Any transit BFR will send the Echo Reply with "Packet-Forward-Success", if the TLV in the received Echo Request is understood and the forwarding table has forwarding entries for the BitString. This behavior is demonstrated by a transit BFR during traceroute mode.

When the Echo Request is received with multipath info for more than one BFER, the Return Code is set to "Invalid Multipath Info Request".

If the BitString cannot be matched in the local forwarding table, the BFR will use "No matching entry in the forwarding table" in the reply.

If the BIER-MPLS label in the received Echo Request is not the one assigned for SI in Original SI-BitString TLV, "Set-Identifier Mismatch" is set in order to report the mismatch.

If the BitString in Header-H does not match the BitString in Egress BitString Sub-TLV of DDMAP TLV, a responding BFR will use "DDMAP Mismatch" to report the problem.

3.3. BIER OAM TLV

This section defines various TLVs that can be used in BIER OAM packet. The TLVs (Type-Length-Value tuples) have the following format:

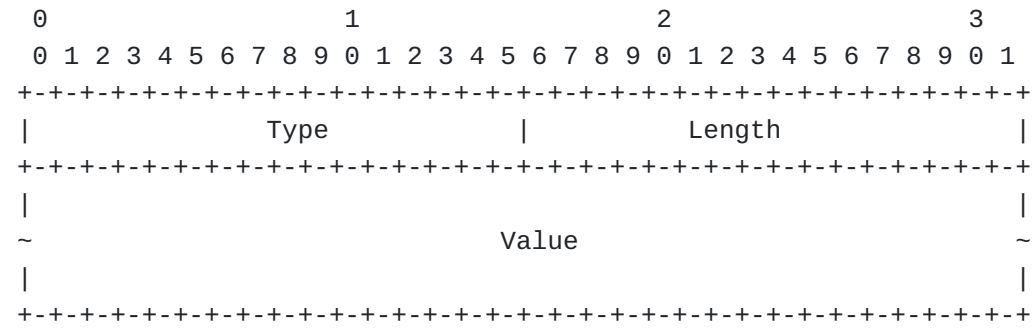

Figure 3: Type-Length-Value Format Used in BIER Echo Request/Reply

TLV Types are defined below; Length is the length of the Value field in octets. The Value field depends on the TLV Type.

3.3.1. Original SI-BitString TLV

The Original SI-BitString TLV carries the set of BFERs and carries the same BitString that the Initiator includes in the BIER header. This TLV has the following format:

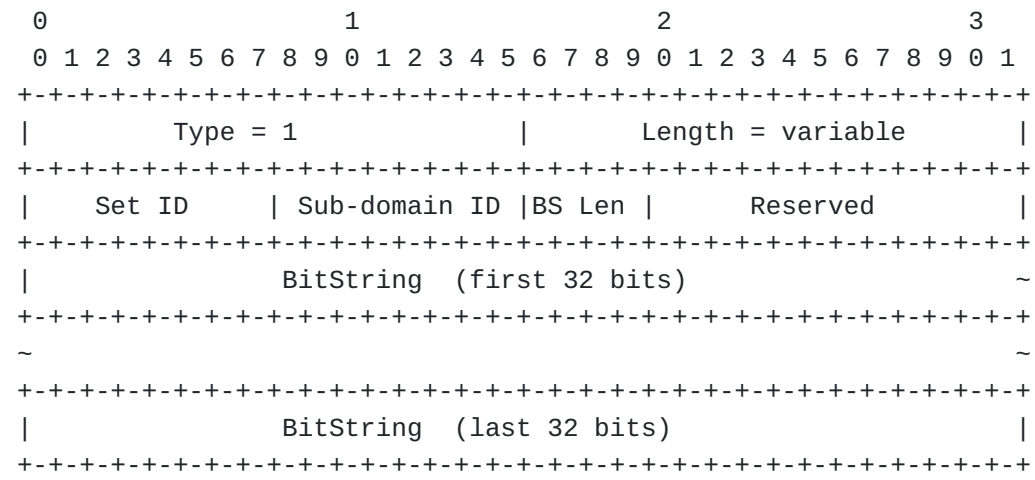

Figure 4: The Format of the Original SI-BitString TLV

Set ID - a one-octet field that is set to the value of the Set Identifier to which the BitString belongs. This value is derived as defined in [RFC8279].

Sub-domain ID - a one-octet field that is set to the Sub-domain value to which BFER in BitString belongs.

BS Len - a four-bit field that is set based on the length of BitString as defined in [RFC8296] reflected in four-octet words.

Reserved - a twelve-bit field. Its value MUST be zeroed on transmission and ignored on receipt.

BitString - a variable length field. The BitString field carries the set of BFR-IDs that Initiator will include in the BIER header. This TLV MUST be included by the Initiator in the Echo Request packet.

Any Initiator MUST include this TLV in the Echo Request packet.

3.3.2. Target SI-BitString TLV

The Target SI-BitString TLV carries the set of BFERs from which the Initiator expects the reply. This TLV has the following format:

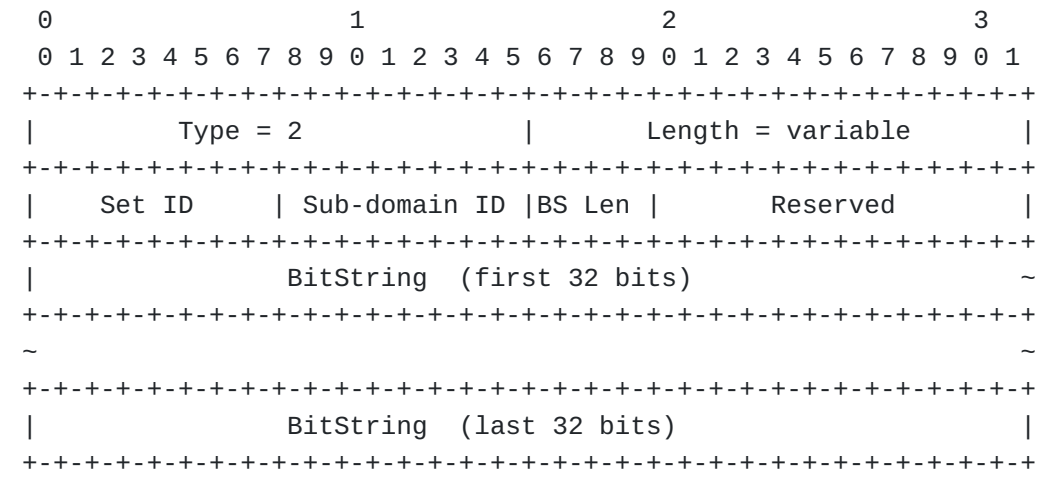

Figure 5: The Format of the Target SI-BitString TLV

Set ID field is set to the Set Identifier to which the BitString belongs.This value is derived as defined in [RFC8279].

Sub-domain ID is set to the Sub-domain value to which BFER in BitString belongs.

BS Len is set based on the length of BitString as defined in [RFC8296]

Reserved - the value MUST be zeroed on transmission and ignored on receipt.

The BitString field carries the set of BFR-IDs of BFER(s) that Initiator expects a response. The BitString in this TLV may be different from the BitString in the BIER header and allows control of the BFER responding to the Echo Request. This TLV MUST be included by Initiator in the BIER OAM packet if the Downstream Mapping TLV (Section 3.3.4) is included.

3.3.3. Incoming SI-BitString TLV

The Incoming SI-BitString TLV will be included by Responder BFR in Reply message and copies the BitString from the BIER header of incoming Echo Request message. This TLV has the following format:

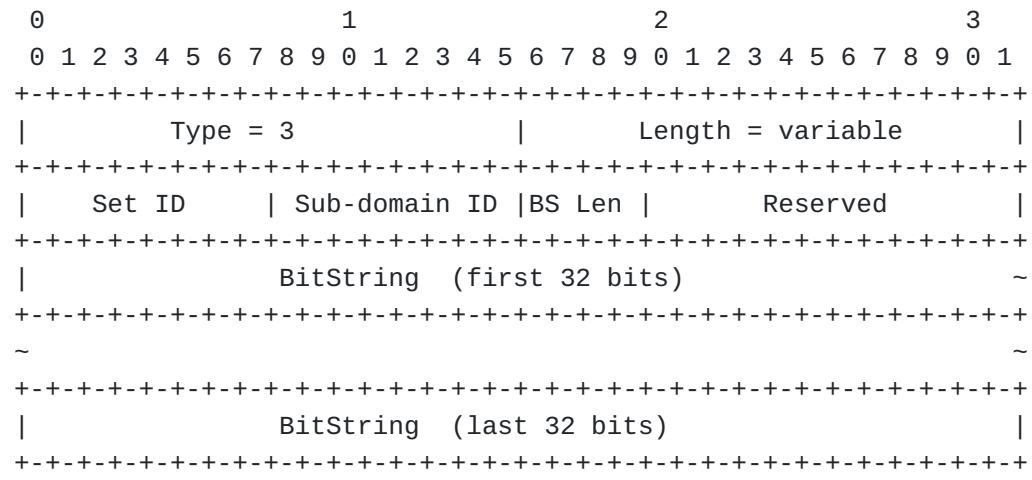

Figure 6: The Format of the Incoming SI-BitString TLV

Set ID field is set to the Set Identifier to which the BitString belongs. This value is derived as defined in [RFC8279]

Sub-domain ID is set to the Sub-domain value to which BFER in BitString belongs.

BS Len is set based on the length of BitString as defined in [RFC8296].

Reserved - the value MUST be zeroed on transmission and ignored on receipt.

The BitString field copies the BitString from the BIER header of the incoming Echo Request. A Responder BFR SHOULD include this TLV in Echo Reply if the Echo Request is received with the I flag set in Downstream Mapping TLV.

An Initiator MUST NOT include this TLV in Echo Request.

3.3.4. Downstream Mapping TLV

This TLV has the following format:

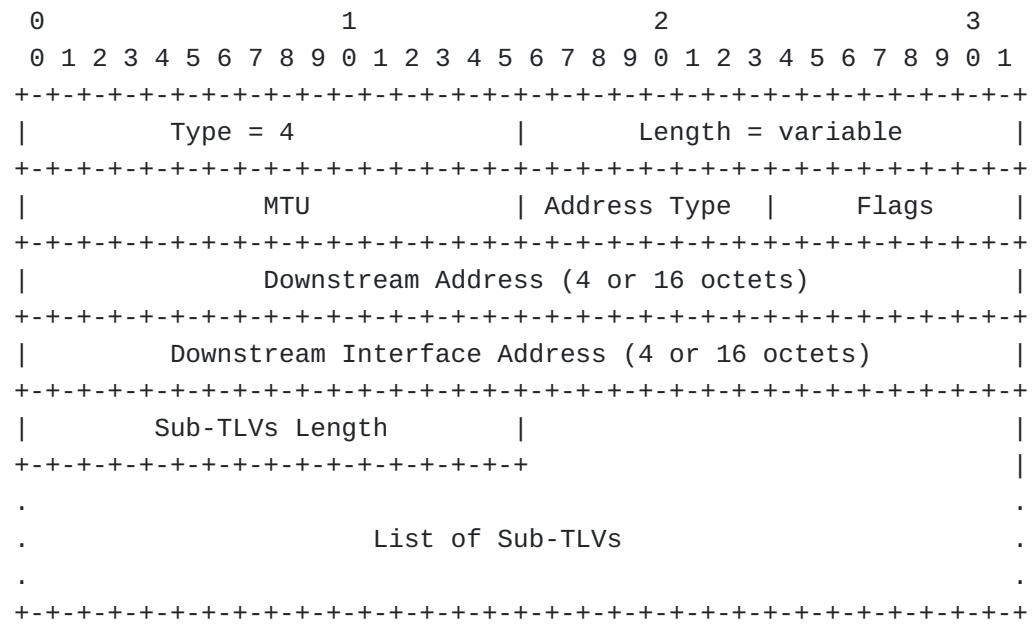

Figure 7: The Format of the Downstream Mapping TLV

MTU A two-octet field. The value is the Maximum Transmission Unit (MTU) value of the egress interface.

Address Type A one-octet field. The Address Type indicates the address type and length of the IP address for the downstream interface. The Address type is set to one of the following values:

Type	Addr. Type	DA Length	DIA Length
1	IPv4 Numbered	4	4
2	IPv4 Unnumbered	4	4
3	IPv6 Numbered	16	16
4	IPv6 Unnumbered	16	4

DA Length
Downstream Address (DA) field Length

DIA Length
Downstream Interface Address (DIA) field Length

Flags
The Flags field has the following format:

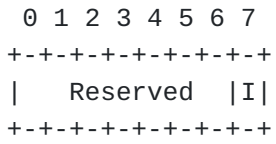

Figure 8: The Flags Field Format

Reserved - a seven-bit field. Its value MUST be zeroed on transmission and ignored on receipt.

I - a one-bit field. When I flag is set, the Responding BFR MUST include the Incoming SI- BitString TLV in Echo Reply message.

Downstream Address and Downstream Interface Address
each field is either four-octet or sixteen-octet, depending on the value of Address Type field.

If the Address Type is 1, the Downstream Address MUST be set to IPV4 BFR-Prefix of downstream BFR and Downstream Interface Address is set to the downstream interface address.

If the Address Type is 2, the Downstream Address MUST be set to IPV4 BFR-Prefix of downstream BFR and Downstream Interface Address is set to the index assigned by upstream BFR to the interface.

If the Address Type is 3, the Downstream Address MUST be set to IPV6 BFR-Prefix of downstream BFR and Downstream Interface Address is set to the downstream interface address.

If the Address Type is 4, the Downstream Address MUST be set to IPv6 BFR-Prefix of downstream BFR and Downstream Interface Address is set to the index assigned by upstream BFR to the interface.

3.3.4.1. Downstream Detailed Mapping Sub-TLVs

This section defines the optional Sub-TLVs that can be included in Downstream Mapping TLV.

Sub-TLV Type	Value
1	Multipath Entropy Data
2	Egress BitString

3.3.4.1.1. Multipath Entropy Data

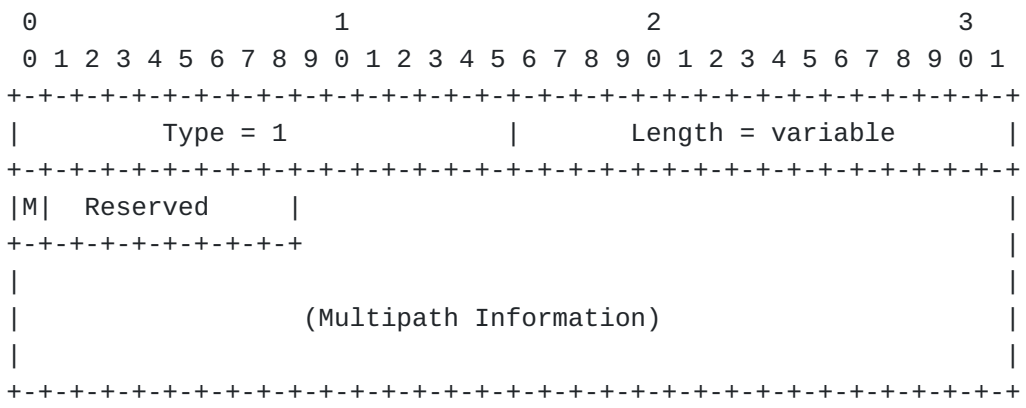

Figure 9: The Format of the Multipath Data Blob

M Flag This flag is set to 0 if all packets will be forwarded out through the interface defined in the Downstream Mapping TLV. When set to 1, Multipath Information will be defined by the Bit masked Entropy data.

Reserved The value MUST be zeroed on transmission and ignored on receipt.

Multipath Information Entropy Data is encoded as defined in Section 3.4.

3.3.4.1.2. Egress BitString Sub-TLV

Responder BFR MAY include this Sub-TLV with the rewritten BitString in the outgoing interface as defined in Section 6.1 of [RFC8279].

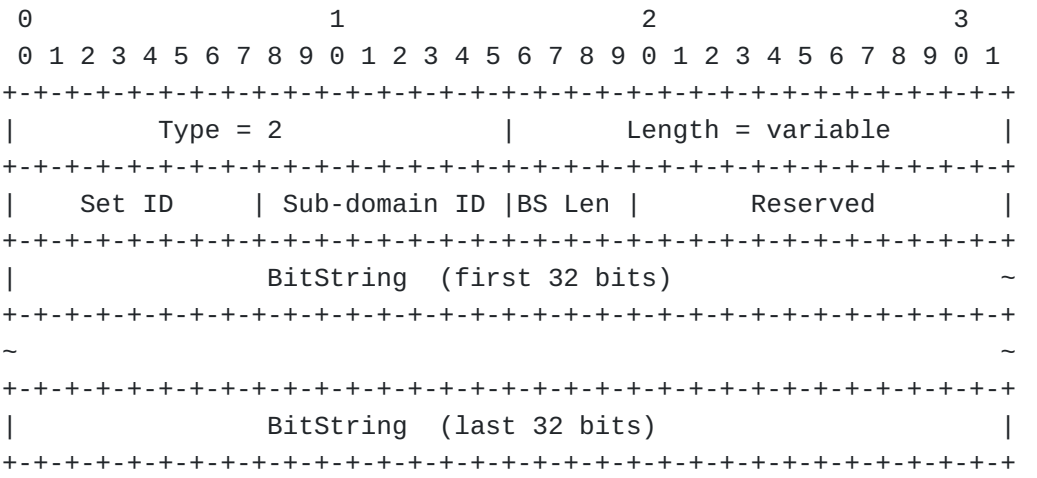

Figure 10: The Egress BitString Sub-TLV Format

Set ID field is set to the Set Identifier to which the BitString belongs. This value is derived as defined in [RFC8279].

Sub-domain ID is set to the Sub-domain value to which BFER in BitString belongs.

BS Len is set based on the length of BitString as defined in [RFC8296].

Reserved - the value MUST be zeroed on transmission and ignored on receipt.

The BitString field copies the rewritten BitString in the outgoing interface as defined in Section 6.1 of [RFC8279].

3.3.5. Responder BFER TLV

The BFER replying to the request MAY include the Responder BFER TLV. This TLV identifies the originator of BIER Echo Reply. This TLV has the following format:

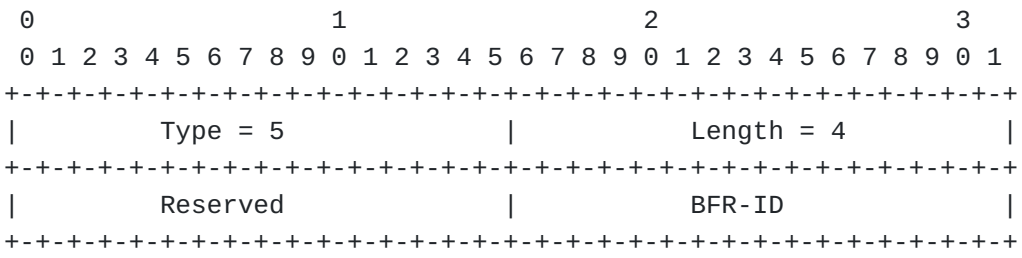

Figure 11: The Responder BFER TLV Format

Length A two-octet field. The value MUST be set to four.

Reserved A two-octet field. The value MUST be zeroed on transmission and ignored on receipt.

BFR-ID A two-octet field. The BFR-ID field carries the BFR-ID of the replying BFER. This TLV MAY be included by the Responding BFER in the BIER Echo Reply packet.

3.3.6. Responder BFR TLV

Any transit BFR replying to the request MAY include the Responder BFR TLV. This is used to identify the replying BFR without BFR-ID. This TLV has the following format:

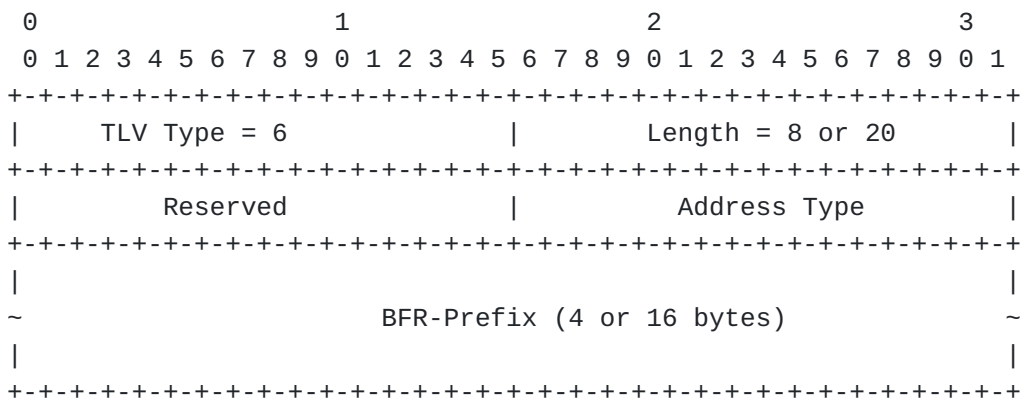

Figure 12: The Responder BFD TLV Format

Length The Length field, depending on the Address Type value - 8 or 20.

Reserved A two-octet field. The value MUST be zeroed on transmission and ignored on receipt.

Address Type A two-octet field. Set to 1 for IPv4 or 2 for IPv6.

BFR-Prefix This field carries the local BFR-Prefix of the replying BFR. This TLV MAY be included by Responding BFR in BIER Echo Reply packet.

3.3.7. Upstream Interface TLV

The BFR replying to the request MUST include the Upstream Interface TLV. This TLV identifies the incoming interface and the BIER-MPLS label in the incoming Echo Request. This TLV has the following format:

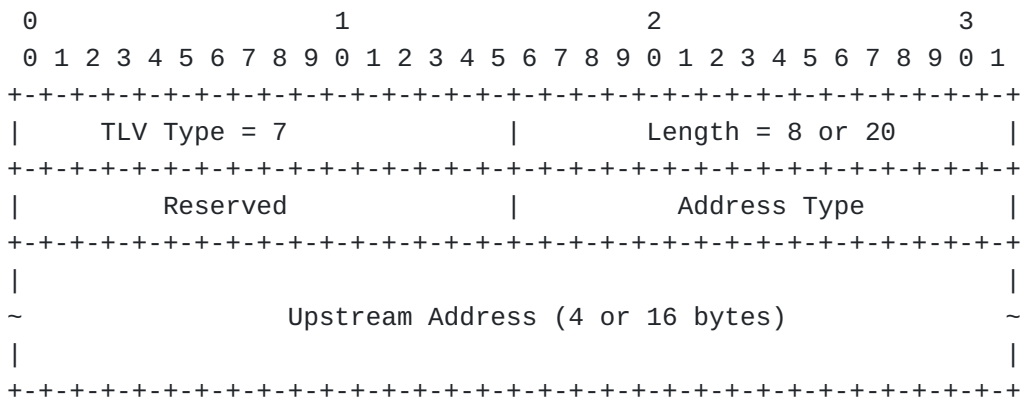

Figure 13: The Upstream Interface TLV Format

Length The Length field, depending on the Address Type value - 8 or 20.

Reserved
A two-octet field. The value MUST be zeroed on transmission and ignored on receipt.

Address Type A two-octet field. Set to 1 for IPv4 numbered, 2 for IPv4 Unnumbered, 3 for IPv6 numbered, or 4 for IPv6 Unnumbered.

Upstream Address
As defined in Section 3.3.4

3.4. Multipath Entropy Data Encoding

The size of the Entropy field in the BIER header is 20 bits, as defined in Section 2 of [RFC8296]. This encoding is the same as the Multipath Type 9 encoding, defined in Section 3.4.1.1.1 of [RFC8029].

4. Procedures

This section describes aspects of Ping and traceroute operations.

4.1. BIER OAM Processing

A BIER OAM packet MUST be sent to the BIER control plane for OAM processing if one of the following conditions is true:

*The receiving BFR is a BFER.

*TTL of BIER-MPLS Label (Section 2.1.1.1 [RFC8296]) expired.

The use of the Router Alert label to be deprecated as proposed in [I-D.ietf-mpls-lspping-norao].

4.2. Per BFER ECMP Discovery

As defined in [RFC8279], BIER follows the unicast forwarding path and allows load balancing over ECMP paths between BFIR and BFER. BIER OAM is expected to support ECMP path discovery between a BFIR and a given BFER and MUST support path validation and failure detection of any particular ECMP path between BFIR and BFER.

[RFC8296] proposes the BIER header with the Entropy field that can be leveraged to exercise all ECMP paths. The Initiator/BFIR will use a traceroute message to query each hop about the Entropy information for each of the downstream paths. To avoid complexity, it is suggested that the ECMP query is performed per BFER by carrying the required information in the BIER OAM message.

The Initiator MUST include Multipath Entropy Data Sub-TLV in Downstream Mapping TLV. It MUST also include the BFER in BitString TLV to which the Multipath query is performed.

Any transit BFR will reply with Bit-masked Entropy for each downstream path as defined in [RFC8029]

4.3. Sending BIER Echo Request

The Initiator MUST set the Message Type as 1 and Return Code as 0. The Proto field in the OAM packet MUST be set to 0. The choice of the Sender's Handle and Sequence Number is a local matter to the Initiator and SHOULD increment the Sequence Number by 1 for every subsequent Echo Request. The QTF field is set to Initiator's local timestamp format, and the TimeStamp Sent field is set to the time that the Echo Request is sent.

The Initiator MUST include Original SI-BitString TLV. The Initiator MUST NOT include more than one Original SI-BitString TLV. The Initiator infers the Set Identifier value and Sub-domain ID value from the respective BitString that will be included in the BIER header of the packet and includes the values in "SI" and Sub-Domain ID fields, respectively.

In Ping mode, the Initiator MAY include Target SI-BitString TLV to control the responding BFER(s) by listing all the BFERs from which the Initiator expects a response. In the traceroute mode, the Initiator MAY include Target SI-BitString TLV to control the path trace towards any specific BFER or set of BFERs. The Initiator on receiving a reply with the Return code "Replying BFR is the only BFER in the header BitString" or "Replying router is one of the BFERs in header BitString" SHOULD unset the respective BFR-id from Target SI-BitString for any subsequent Echo Request.

The Initiator MAY include Downstream Mapping TLV (Section 3.3.4) in the Echo Request to query additional information from transit BFRs and BFERs. In case of ECMP discovery, the Initiator MUST include the Multipath Entropy Data Sub-TLV and SHOULD set the Target SI-BitString TLV carrying a specific BFER ID.

The Initiator MUST encapsulate the OAM packet with the BIER header and MUST set the Proto as 5 and further encapsulates with BIER-MPLS label. In ping mode, the BIER-MPLS Label TTL MUST be set to 255. In traceroute mode, the BIER-MPLS Label TTL is set successively, starting from 1 and MUST stop sending the Echo Request if it receives a reply with Return code as "Replying router is the only BFER in BIER header BitString" from all BFER listed in Target SI-BitString TLV.

4.4. Receiving BIER Echo Request

Sending a BIER OAM Echo Request to control plane for payload processing is triggered as mentioned in Section 4.1.

Any BFR on receiving Echo Request MUST perform the basic sanity check. If the BFR cannot parse the OAM Dependent data payload completely because the value in the OAM Message Length field is incorrect, BFR MUST send Echo Reply with Return Code set to "Malformed Echo Request received" if the OAM Message Length is incorrect. If the packet sanity check is fine, it SHOULD initiate the below set of variables:

Reply-Flag
This flag is initially set to 1.

Interface-I
The incoming interface on which the Echo Request was received. This MAY be used to validate the Downstream Detailed Mapping TLV (DDMAP) info and populate the Upstream Interface TLV.

BIER-Label-L
The BIER-MPLS Label received as the top label of the received Echo Request. This MAY be used to validate if the packet is traversing the desired Set Identifier and sub-domain path.

Header-H
The BIER header of the received Echo Request. It can be used to validate the DDMAP info and to populate the Incoming SI-BitString TLV. Also, it can be used to perform entropy calculation considering a different field in the header and replying with Multipath Entropy Data Sub-TLV.

BFR MUST initialize the Best-return-code variable to the null value.

BFR will populate the Interface-I with the identifier of the interface over which the Echo Request is received, the top label in the MPLS stack of the received Echo Request to BIER-Label-L, and the BIER header to BIER-Header. If the received Echo Request carries Target SI-BitString TLV, a BFR SHOULD run the boolean AND operation between BitString in Header-H and BitString in Target SI-BitString TLV. If the resulting BitString is all-zero, reset Reply-Flag=0 and go to Section 4.5. Else:

*If the BIER-Label-L does not correspond to the local label assigned for {sub-domain, BitStringLen, SI} in Original SI-BitString TLV, Set the Best-return-code to "Set-Identifier Mismatch" and Go to Section 4.5.

This step allows the detection of a synchronization problem in the upstream BFR between BIER-Label and {sub-domain, BitStringLen, SI} that might cause an unintended packet leak between sub-domains.

- Set the Best-return-code to "One or more of the TLVsis not supported" if any of the TLVs in the Echo Request message is not supported. Go to Section 4.5.

- If the BitString in Header-H does not match the BitString in Egress BitString Sub-TLV of DDMAP TLV, set the Best-return-code to "DDMAP Mismatch" and go to Section 4.5. When there are more than one DDMAP TLV in the received Request packet, the Downstream Address and Downstream Interface Address should be matched with Interface-I to identify the right DDMAP TLV and then perform the BitString match.

This step allows the detection of a deviation between the BIER control plane and the BIER forwarding plane in the upstream node that may result in a forwarding loop or packet duplication.

- Set the Best-return-code to "Invalid Multipath Info Request", when the DDMAP TLV carries Multipath Entropy Data Sub-TLV, and if the Target SI-BitString TLV in the received Echo Request carries more than 1 BFER id. Go to Section 4.5. Else, list the ECMP downstream neighbors to reach BFR-id in Target SI-BitString TLV, calculate the Entropy considering the BitString in Header-H and Multipath Entropy Data Sub-TLV from received Echo Request. Store the Data for each Downstream interface in a temporary variable. Set the Best-return-code to 5 (Packet-Forward-Success) and goto Section 4.5

This step instructs the node to calculate the Entropy Data for each downstream interface to reach the BFER in Target SI-BitString TLV by considering the incoming BitString and Entropy Data.

- Set the Best-return-code to "Replying router is the only BFER in BIER header BitString", and go to Section 4.5 if the responder is BFER and there are no more bits in the BIER header BitString left for forwarding.

- Set the Best-return-code to "Replying router is one of the BFERs in BIER header BitString", and include Downstream Mapping TLV if the responder is BFER and there are more bits in BitString left for forwarding. Also, include the Multipath information as defined in Section 4.2 if the received Echo Request carries Multipath Entropy Data Sub-TLV. Go to Section 4.5.

- Set the Best-return-code to "No matching entry in the forwarding table", if the forwarding lookup, defined in Section 6.5 of

[RFC8279] does not match any entry for the received BitString in BIER header.

This step allows the detection of the missing BFR-id in the node's BIER forwarding table. It is difficult to detect the absence of the BFR-id if the Request includes more than one BFR-ids in the BitString and so may need to include the BFER-id that is not responding to detect such failure.

*Set the Best-return-code to "Packet-Forward-Success", and include Downstream Mapping TLV. Go to Section 4.5.

4.5. Sending Echo Reply

If Reply-Flag=0, BFR MUST release the variables and MUST NOT send any response to the Initiator. If Reply-Flag=1, proceed as below:

The Responder BFR SHOULD include the BitString from Header-H to Incoming SI-BitString TLV and include the Set ID, Sub-domain ID and BS Len that corresponds to BIER-Label-L. Responder BFR SHOULD include the Upstream Interface TLV and populate the address from Interface-I.

When the Best-return-code is "Replying BFR is one of the BFERs in header BitString", it MUST include Responder BFER TLV.

If the received Echo Request had DDMAP with Multipath Entropy Data Sub-TLV, Responder BFR MUST include DDMAP as defined in Section 3.3.4 for each outgoing interface over which the packet will be replicated and include the respective Multipath Entropy Data Sub-TLV. For each outgoing interface, the respective Egress BitString MUST be included in DDMAP TLV.

If the received Echo Request had DDMAP without Multipath Entropy Data Sub-TLV, Responder BFR MUST include DDMAP as defined in Section 3.3.4 for each outgoing interface over which the packet will be replicated. For each outgoing interface, respective Egress BitString MUST be included in DDMAP TLV.

When the Best-return-code is "Replying BFR is the only BFER in header BitString", it MUST include Responder BFER TLV.

The Responder MUST set the Message Type as 2 and Return Code as Best- return-code. The Proto field MUST be set to 0.

The Echo Reply can be sent as BIER-encapsulated, or IP/UDP encapsulated, depending on the Reply Mode in the received Echo Request. When the Reply Mode in the received Echo Request is set to 3, Responder appends the BIER header listing the BitString with BFIR ID (from Header- H), sets the Proto to 5, and sets the BFIR as 0.

When the Reply Mode in the received Echo Request is set to 2, Responder encapsulates with the IP/UDP header. The UDP destination port MUST be set to TBD1 (Section 5.1), and the source port MAY be set to TBD1 or other value selected from the Dynamic range of port numbers. The source IP is any local address of the responder, and the destination IP is derived from the BFIR-id of the BIER header [RFC8296] of the received Echo Request.

4.6. Receiving Echo Reply

The Initiator, upon receiving the Echo Reply, will use the Sender's Handle to match with Echo Request sent. If no match is found, the Initiator MUST ignore the Echo Reply.

If receiving Echo Reply has Downstream Mapping, the Initiator SHOULD copy the same to subsequent Echo Request(s).

If one of the Echo Reply is received with Return Code as "Replying BFR is one of the BFERs in header BitString", it SHOULD reset the BFR- id of the responder from Target SI-BisString TLV in subsequent Echo Request. This step helps avoid any BFR that is both BFER and transit BFR to respond with Echo Reply continuously.

5. IANA Considerations

The terms used in the IANA Considerations below are intended to be consistent with [RFC8126].

5.1. UDP Port Number

This document requests a UDP port TBD1 to be allocated by IANA for BIER Echo.

5.2. BIER OAM Parameters

IANA is requested to create and maintain the "BIER OAM Parameters" registry containing the sub-registries listed below.

5.3. BIER OAM Message Type

IANA is requested to create in the BIER OAM Parameters registry a sub-registry as follows:

Sub-registry Name: BIER OAM Message Type.

Assignment Policy:

2-31 IETF Review

32-62 First Come, First Served

Reference: [this document]

Value	Description	Reference
0	Reserved	This document
1	BIER Echo Request/Echo Reply	This document
2 - 31	Unassigned	This document
32-62	Unassigned	This document
63	Reserved	This document

Table 1: BIER OAM Message Type

5.4. BIER Echo Request/Echo Reply Parameters

IANA is requested to create in the BIER OAM registry the sub-registry BIER Echo Request/Echo Reply Parameters.

5.4.1. BIER Echo Request/Echo Reply Message Types

IANA is requested to create in the BIER Echo Request/Echo Reply Parameters the BIER Echo Request/Echo Reply Message Types sub-registry as follows:

Sub-registry Name: BIER Echo Request/Echo Reply Message Types

Assignment Policy:

3 - 175 IETF Review

176 - 239 First Come, First Served

Reference: [this document]

Value	Description	Reference
0	Reserved	This document
1	BIER Echo Request	This document
2	BIER Echo Reply	This document
3 - 175	Unassigned	This document
176 - 239	Unassigned	This document
240 - 251	Experimental	This document
252 - 254	Private Use	This document
255	Reserved	This document

Table 2: BIER Echo Request/Echo Reply Message Types

5.4.2. BIER Echo Reply Modes

IANA is requested to create in the BIER Echo Request/Echo Reply Parameters registry the new sub-registry as follows:

Sub-registry Name: BIER Echo Reply Mode

Assignment Policy:

8 - 175 IETF Review

176 - 239 First Come, First Served

Reference: [this document]

Value	Description	Reference
0	Reserved	This document
1	Do Not Reply	This document
2	Reply via an IPv4/IPv6 UDP Packet	This document
3	Reply via BIER packet	This document
4 - 175	Unassigned	This document
176 - 239	Unassigned	This document
240 - 251	Experiemntal	This document
252 - 254	Private Use	This document
255	Reserved	This document

Table 3: SFC Echo Reply Modes

5.4.3. BIER Echo Return Codes

IANA is requested to create in the BIER Echo Request/Echo Reply Parameters registry the new sub-registry as follows:

Sub-registry Name: BIER Echo Return Codes

Assignment Policy:

9 - 191 IETF Review

192 - 251 First Come, First Served

Reference: [this document]

Value	Description	Reference
0	No Return Code	This document
1	Malformed Echo Request received	This document
2	One or more of the TLVs is not supported	This document
3		This document

Value	Description	Reference
	Replying BFR is the only BFER in header BitString	
4	Replying BFR is one of the BFERs in header BitString	This document
5	Packet-Forward-Success	This document
6	Invalid Multipath Info Request	This document
7	Unassigned	This document
8	No matching entry in the forwarding table	This document
9	Set-Identifier Mismatch	This document
10	DDMAP Mismatch	This document
11 - 191	Unassigned	This document
192-251	Unassigned	This document
252-254	Private Use	This document
255	Reserved	This document

Table 4: BIER Echo Return Codes

5.5. TLVs

IANA is requested to create in the BIER OAM registry a sub-registry for the Type field of top-level TLVs as well as for any associated sub-TLVs. Note that the meaning of a sub-TLV is scoped by the TLV. The number of spaces for the sub-TLVs of various TLVs is independent.

The valid range for TLVs and sub-TLVs is 0-65535. Assignments in the ranges 0-16383 and 32768-49161 are made via Standards Action as defined in [RFC8126]; assignments in the ranges 16384-31743 and 49162-64511 are made via "Specification Required"; values in the ranges 31744-32767 and 64512-65535 are for Private Use and MUST NOT be allocated.

If a TLV or sub-TLV has a Type that falls in the range of Private Use, the Length MUST be at least 4, and the first four octets MUST be an SMI Private Enterprise Number that identifies the user, in network octet order. The rest of the Value field is private to the user.

The TLVs and Sub-TLVs defined in this document are not limited to Echo Request or Reply message types are applicable to other message types. The TLVs and Sub-TLVs requested by this document for the IANA consideration are the following:

Type	Sub-Type	Value Field
1		Original SI-BitString
2		Target SI-BitString
3		Incoming SI-BitString
4		Downstream Mapping
4	1	Multipath Entropy Data
4	2	Egress BitString
5		Responder BFER
6		Responder BFR
7		Upstream Interface

6. Security Considerations

The security consideration for BIER Ping is similar to ICMP [RFC0792] or LSP Ping [RFC8029], [RFC6425]. As with ICMP or LSP Ping, BFR can be exposed to Denial-of-Service (DoS) attacks, and it is RECOMMENDED to regulate the BIER Ping packet flow to the control plane. A rate limiter SHOULD be applied to avoid any attack. Although using BIER Echo Request in a DoS amplification attack is theoretically possible, spoofing BFIR ID in the BIER Header presents itself as a serious challenge. As a result, this threat is not a big concern.

As with ICMP or LSP Ping, a traceroute can be used to obtain network information. It is RECOMMENDED that the implementation checks the integrity of BFIR of the Echo messages against any locally secured list before processing the message further.

7. Acknowledgement

The authors would like to thank Antoni Przygienda, Eric Rosen, Faisal Iqbal, Jeffrey (Zhaohui) Zhang, and Shell Nakash for their review and comments.

The authors would like to thank Mankamana Mishra for his thorough review and comments.

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>.

[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>.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, DOI 10.17487/RFC5905, June 2010, <https://www.rfc-editor.org/info/rfc5905>.

[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>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <https://www.rfc-editor.org/info/rfc8279>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <https://www.rfc-editor.org/info/rfc8296>.

[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>.

[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>.

[IEEE.1588.2008] "Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", IEEE Standard 1588, March 2008.

8.2. Informative References

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <https://www.rfc-editor.org/info/rfc792>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[I-D.ietf-mpls-lspping-norao]
Kompella, K., Bonica, R., and G. Mirsky, "Deprecating the Use of Router Alert in LSP Ping", Work in Progress, Internet-Draft, draft-ietf-mpls-lspping-norao-01, 11 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-lspping-norao-01>.

Authors' Addresses

Nagendra Kumar
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709
US

Email: naikumar@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709-4987
US

Email: cpignata@cisco.com

Nobo Akiya
Big Switch Networks
Japan

Email: nobo.akiya.dev@gmail.com

Lianshu Zheng
Individual Contributor
China

Email: veronique_cheng@hotmail.com

Mach Chen
Huawei Technologies

Email: mach.chen@huawei.com

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com