BIER Workgroup
Internet Draft
Intended status: Standard Track

H. Bidgoli
J. Kotalwar
Nokia
Z.Zhang
Juniper Networks
Eddie Leyton
Vrizon
Mankamana Mishra
I. Wijanands
Cisco System, Inc.

Expires: May 6, 2020

November 3, 2019

M-LDP Signaling Through BIER Core
draft-hb-bier-mldp-signaling-over-bier-01

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides multicast forwarding through a "BIER domain" without requiring intermediate routers to maintain multicast related per-flow state. Neither does BIER require an 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. Such 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 the according set of bits switched on in BIER packet header.

This document describes the procedure needed for mLDP tunnels to be signaled over and stitched through a BIER core, allowing LDP routers to run traditional Multipoint LDP services through a BIER core.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Table of Contents

1. Introduction

Some operators that are using mLDP P2MP LSPs for their multicast transport would like to deploy BIER technology in some segment of their network. This draft explains a method to signal mLDP services and stitch the mLDP datapath labels through a BIER domain, with minimal disruption and operational impact to the mLDP domain.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2.1. Definitions

Some of the terminology specified in [I-D.draft-ietf-bier-architecture-05] is replicated here and extended by necessary definitions:

BIER:

Bit Index Explicit Replication (The overall architecture of forwarding multicast using a Bit Position).

BFR:

Bit Forwarding Router (A router that participates in Bit Index Multipoint Forwarding). A BFR is identified by a unique BFR-prefix in a BIER domain.

BFIR:

Bit Forwarding Ingress Router (The ingress border router that inserts the Bit Map into the packet). Each BFIR must have a valid BFR-id assigned. BFIR is term used for dataplain packet forwarding.

BFER:

Bit Forwarding Egress Router. A router that participates in Bit Index Forwarding as leaf. Each BFER must be a BFR. Each BFER must have a valid BFR-id assigned. BFIR is term used for dataplain packet forwarding.

BBR:

BIER Boundary router. The router between the LDP domain and BIER domain.

IBBR:

Ingress BIER Boundary Router. The ingress router from signaling point of view. It maintains mLDP adjacency toward the LDP domain and determines if the mLDP FEC needs to be signaled across the BIER domain via targeted ldp.

EBBR:

Egress BIER Boundary Router. The egress router in BIER domain from signaling point of view. It terminates the targeted ldp signaling through BIER domain. It also keeps track of all IBBRs that are part of this p2mp tree

BFT:

Bit Forwarding Tree used to reach all BFERs in a domain.

BIFT:

Bit Index Forwarding Table.

BIER sub-domain:

A further distinction within a BIER domain identified by its unique sub-domain identifier. A BIER sub-domain can support multiple BitString Lengths.

BFR-id:

An optional, unique identifier for a BFR within a BIER sub-domain.

3. mLDP Signaling Through BIER domain

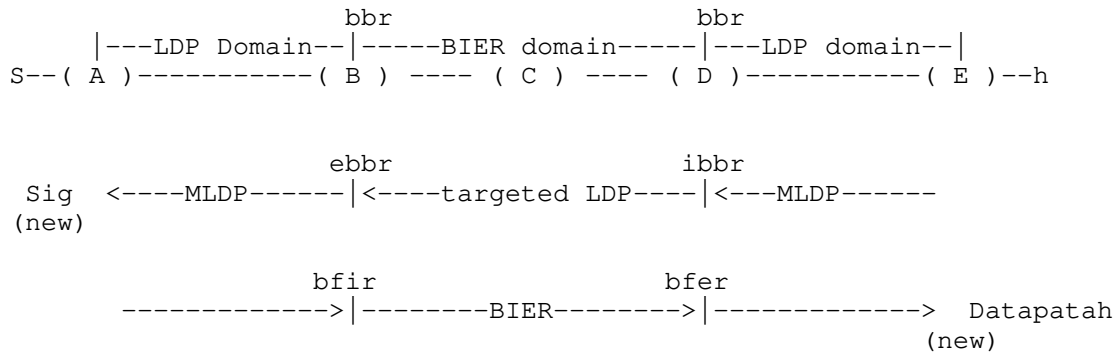


Figure 1: bier boundary router

As per figure 1, point-to-multipoint and multipoint-to-multipoint LSPs established via mLDP [RFC6388] can be signaled through a bier domain via targeted LDP sessions. This procedure is explained in [RFC7060] (Using LDP Multipoint Extension on Targeted LDP Sessions).

This documents provides some details and defines some needed procedures.

3.1. Ingress BBR procedure

The Ingress BBR (IBBR) is connected to the mLDP on one side and a bier domain on the other side. To connect the LDP domains via BIER domain IBBR needs to establish a targeted LDP session with EBBR closest to the root of the P2mp or mp2mp LSP. To do so IBBR will follow procedures in [RFC7060] in particular the section "6. targeted mLDP with Multicast Tunneling".

The target LDP session can be established manually via configuration or via automated mechanism.

3.1.1. Automatic tLDP session creation

A tLDP session can be generated automatically from every IBBR to EBBR. As an example when a mLDP FEC arrives on the IBBR, it can automatically start a tLDP Session with the EBBR. In this case both IBBR and EBBR should be in auto-discovery mode and react to the arriving FEC or tLDP Signaling packets (i.e. targeted hellos, keep-alives etc...).

The Root node address in the mLDP FEC can be used to find the EBBR. To identify the EBBR same procedures as [RFC7060} section 2.1 can be

used or the procedures as explained in the [draft-ietf-bier-pim-signaling] appendix A. After fining the IBBR the tLDP session can be initiated from the IBBR to EBBR.

3.1.1. ECMP Method on IBBR

If IBBR finds multiple equal cost EBBRs on the path to the Root, it can use a vendor specific algorithm to choose between the EBBRs. These algorithms are beyond the scope of this draft. As an example the IBBR can use the smallest EBBR IP address to establish its mLDP signaling to.

3.2. Egress BBR procedure method

The Egress BBR (EBBR) is connected to the mLDP domain which the root of the P2MP or MP2MP LSP resides on. The EBBR should accept the tLDP session generated form IBBR. It should assign a unique "upstream assigned label" for each arriving FEC generated by IBBRs.

The EBBR should follow the [RFC7060] procedures with following modifications:

- The label assigned by EBBR cannot be Implicit Null. This is to ensure that identity of each p2mp and/or mp2mp tunnel in BIER domain is uniquely distinguished.

- The label can be assigned from a domain-wide Common Block (DCB) [I-D.zzhang-bess-mvpn-evpn-aggregation-label], as well as upstream assigned.

- The Interface ID TLV [RFC6389]includes a new BIER sub-domain sub-tlv (type TBD)

The EBBR will also generate a new label and FEC toward the ROOT on the mLDP domain. The EBBR Should stitch this generate label with the "upstream assigned label" to complete the p2MP or MP2MP LSP. This stitch point should be stored on the datapath (ILM) table for packet forwarding.

With same token the EBBR should track all the arriving FECs and the IBBRs that are generating these FECs. EBBR will use this information to build the bier header for each set of common FEC arriving from the IBBRs.

3.2.1. IBBR procedure upon arriving upstream assigned label

Upon receiving the "upstream assigned label", IBBR should create its own stitching instruction between the "upstream assigned label" and

the down stream label that was signaled to it. IBBR should download these instructions to the datapath.

4. Datapath Forwarding

4.1. Datapath traffic flow

On BFIR when the MPLS label for P2MP/MP2MP LSP arrives a lookup in ILM table is done and the label is swapped with tLDP upstream assigned label. The BFIR will note all the BFERs that are interested in specific p2mp/mp2mp LSP (as per section 3.2). BFIR will put the corresponding BIER header with bit index set for all IBBRs interested in this P2MP LSP. BFIR will set the BIERHeader.Proto = MPLS and will forward the BIER packet into BIER domain.

In the BIER domain normal BIER forwarding procedure will be done, as per [RFC 8279]

The IBBRs will receive the BIER packet, will look at the protocol of BIER header (MPLS). BFER will remove the BIER header and will do a lookup in the ILM table for the upstream assigned label and perform its corresponding action.

It should be noted that these procedures are valid if BFIR is the ILER and/or BFER is the ELER as per [RFC 7060]

5. Recursive FEC

The above procedures also will work with a mLDP recursive FEC. The root used to determine the EBBR is the outer root of the FEC. The entire recursive FEC needs to be preserve when it is forwarded via tLDP and the label request.

6. IANA Considerations

This document contains no actions for IANA.

7. Security Considerations

TBD

8. References

8.1. Normative References

[BIER_ARCH] Wijnands, IJ., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication",

internet-draft draft-ietf-bier-architecture-08, October 2016.

8.2. Informative References

[BIER_MVPN] Rosen, E., Ed., Sivakumar, M., Wijnands, IJ., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using Bier", internet-draft draft-ietf-bier-mvpn-08, January 2017.

[ISIS_BIER_EXTENSIONS] Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER Support via ISIS", internet-draft draft-ietf-bier-isis-extensions-06.txt, March 2017.

[OSPF_BIER_EXTENSIONS] Psenak, P., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for Bit Index Explicit Replication", internet-draft draft-ietf-ospf-bier-extensions-09.txt, March 2017.

7. Acknowledgments Authors would like to acknowledge Jingrong Xie for his comments and help on this draft.

Authors' Addresses

Hooman Bidgoli (editor)
Nokia
600 March Rd.
Ottawa, Ontario K2K 2E6
Canada

Email: hooman.bidgoli@nokia.com

Jayant Kotalwar
Nokia
380 N Bernardo Ave,
Mountain View, CA 94043
US

Email: jayant.kotalwar@nokia.com

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

IJsbrand Wijnands
Cisco Systems

EMail: ice@cisco.com

Eddie Leyton
Vrizon

Email: Edward.leyton@verizonwireless.com

Mankamana Mishra
Cisco System
821 alder drive
Milpitas California
USA

Email: mankamis@cisco.com

BIER WG
Internet-Draft
Intended status: Standards Track
Expires: January 11, 2021

Quan Xiong
Greg Mirsky
ZTE Corporation
Fangwei Hu
Individual
Chang Liu
China Unicom

July 10, 2020

BIER BFD
draft-hu-bier-bfd-07.txt

Abstract

Point to multipoint (P2MP) BFD is designed to verify multipoint connectivity. This document specifies the application of P2MP BFD in BIER network.

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 11, 2021.

Copyright Notice

Table of Contents

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] provides the forwarding of multicast data packets through a multicast domain. It does so without requiring any explicit tree-building protocol and without requiring intermediate nodes to maintain any per-flow state.

[RFC8562] defines a method of using Bidirectional Forwarding Detection (BFD) to monitor and detect unicast failures between the sender (head) and one or more receivers (tails) in multipoint or multicast networks. [RFC8563] describes active tail extensions to the BFD protocol for multipoint networks.

This document describes the procedures for using such mode of BFD protocol to monitor connectivity between a multipoint sender, Bit-Forwarding Ingress Router (BFIR), and a set of one or more multipoint receivers, Bit-Forwarding Egress Routers (BFERs). The BIER BFD only supports the unidirectional multicast. This document defines the use

of P2MP BFD as per [RFC8562], and active tail as per [RFC8563] for BIER-specific domain.

2. Conventions used in this document

2.1. Terminology

This document uses the acronyms defined in [RFC8279] along with the following:

BFD: Bidirectional Forwarding Detection.

OAM: Operations, Administration, and Maintenance.

P2MP: Point to Multi-Point.

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 BFD Encapsulation

BIER BFD encapsulation uses the BIER OAM packet format defined in [I-D.ietf-bier-ping]. The value of the Message Type field MUST be set to BIER BFD (TBD1 by IANA). BFD Control Packet, defined in Section 4 [RFC5880] immediately follows the BIER OAM header. The operation of Multipoint BFD with the BFD Control Packet is described in [RFC8562].

4. BIER BFD Session Bootstrapping

As defined in [RFC8562], BIER BFD session MAY be established to monitor the state of the multipoint path. The BIER BFD session could be created for each multipoint path and the set of BFERs over which the BFIR is requested to run BIER BFD. The BFIR MUST advertise the multipoint path and the value of My Discriminator associated with the path to the set of BFERs. Bootstrapping a BIER BFD session MAY use BIER OAM message Section 4.1 or the control plane Section 4.2.

The BIER BFD bootstrapping MUST be repeated when the value of this discriminator being changed.

4.1. BIER OAM Bootstrapping

The BIER OAM could be used for bootstrapping the BIER BFD session. The BFIR sends the BIER OAM Echo request message carrying a BFD discriminator TLV which immediately follows the Target SI-Bitstring TLV (section 3.3.2 [I-D.ietf-bier-ping]).

The Target SI-Bitstring TLV MUST be used to carry the set of BFER information (including Sub-domain-id, Set ID, BS Len, Bitstring) for the purpose of the session establishment.

The BFD discriminator TLV is a new TLV for BIER OAM TLV with the type (TBD2 by IANA) and the length of 4. The value contains the 4-byte local discriminator generated by BFIR for this session. This discriminator MUST subsequently be used as the My Discriminator field in the BIER BFD session packets sent by BFIR. The format is as follows.

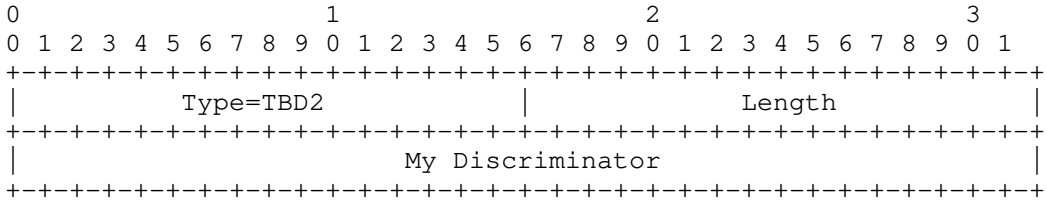


Figure 1: BFD discriminator TLV

4.2. IGP protocol Bootstrapping

An alternative option to bootstrap the BIER BFD is to advertise the BFD information in the control plane. This document defines a new BIER BFD Sub-sub-TLV carried in IS-IS and OSPF protocol.

The BFIR generates the My Discriminator value for each multicast flow and advertises it to the expecting BFERs which is indicated by the Bitstring which is carried in BIER BFD sub-sub-TLV. The corresponding BFERs SHOULD store the My Discriminator value for packet Demultiplexing.

4.2.1. IS-IS extension for BIER BFD

The new BIER BFD Sub-sub-TLV is carried within the BIER Info sub-TLV defined in [RFC8401]. The format is as follows.

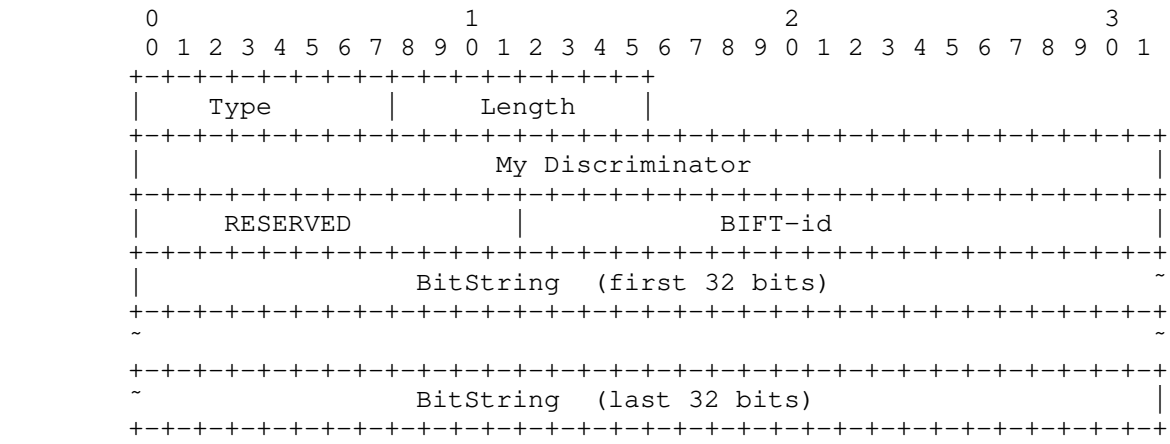


Figure 2: BIER BFD Sub-sub-TLV for IS-IS extension

Type: TBD3 by IANA.

Length: Length of the BIER BFD Sub-sub-TLV for IS-IS extension, in bytes.

My Discriminator: A unique, nonzero discriminator value generated by BFIR for each multipoint path.

The BitString field carries the set of BFR-IDs of BFER(s) that the BFIR expects to establish the BIER BFD session.

The BIFT-id represents a particular Bit Index Forwarding Table (BIFT) as per [RFC8279].

4.2.2. OSPF extension for BIER BFD

The new BIER BFD Sub-TLV is a sub-TLV of the BIER Sub-TLV defined in [RFC8444]. The format is as follows.

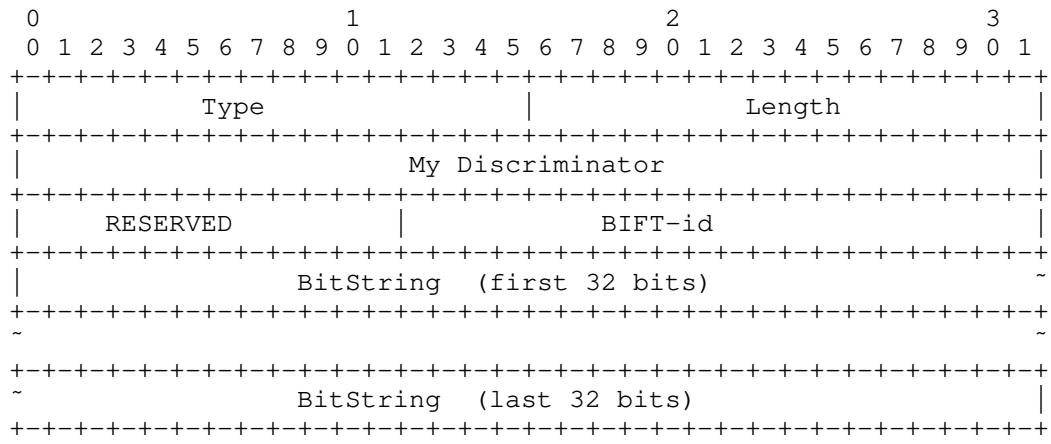


Figure 3: BIER BFD Sub-TLV for OSPF extension

Type: TBD4 by IANA.

Length: Length of the BIER BFD Sub-TLV for OSPF extension, in bytes.

Other fields in BIER BFD Sub-TLV is the same with section 4.2.1.

5. Discriminators and Packet Demultiplexing

As defined in [RFC8562], the BFIR sends BFD Control packets over the multipoint path via the BIER BFD session with My Discriminator set to the value assigned by the BFIR and the value of the Your Discriminator set to zero. The set of BFERs MUST demultiplex BFD packets based on a combination of the source address, My Discriminator value. The source address is BFIR-id and BIER MPLS Label (MPLS network) or BFIR-id and BIFT-id (Non-MPLS network) for BIER BFD. My Discriminator value is advertised in BIER BFD bootstrapping using one of the options described in Section 4.

6. Active Tail in BIER BFD

[RFC8563] defined an extension for Multipoint BFD, which allows the head to discover the state of a multicast distribution tree for any sub-set of tails. For BIER BFD in the active tail mode, the BFIR may learn the state and connectivity of the BFERs through allowing the BFERs to notify the BFIR. As per [RFC8563] provides detailed information on how the BFIR can use multipoint Poll sequence message or a combination of multicast and unicast Poll sequence messages to determine the state of the multicast tree. Also, [RFC8563] describes that a BFER can transmit an unsolicited unicast Poll sequence message

to the BFIR (note that a unicast message must be sent over a path which is disjoint from the multicast distribution tree).

6.1. Unsolicited Head Notification Mode

[I-D.mirsky-mpls-p2mp-bfd] provides detailed information on using the unsolicited notification method for P2MP MPLS LSP which is also applicable to BIER.

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

Upon detecting the failure, a BFER sends a BFD control packet with the following settings:

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

To improve the likelihood of notifying the BFIR of the failure, the BFER SHOULD transmit three BFD Control packets defined above in short succession.

A BFIR that has received the BFD Control packet, as described above, sends the unicast IP/UDP encapsulated BFD control packet with the Final (F) bit set to the BFER.

7. Security Considerations

For BIER OAM packet processing security considerations, see [I-D.ietf-bier-ping].

For general multipoint BFD security considerations, see [RFC8562].

No additional security issues are raised in this document beyond those that exist in the referenced BFD documents.

8. Acknowledgements

The authors would like to thank the comments and suggestions from Sandy Zhang, Jeffrey (Zhaohui) Zhang, Donald Eastlake 3rd.

9. IANA Considerations

9.1. BIER OAM Message Type

IANA is requested to assign a new type from the BIER OAM Message Type registry as follows:

Value	Description	Reference
TBD1	BIER BFD	[this document]

Table 1

9.2. BFD Discriminator TLV

IANA is requested to assign a new type from the BIER OAM TLV registry as follows:

Value	Description	Reference
TBD2	BFD discriminator TLV	[this document]

Table 2

9.3. BIER BFD Sub-sub-TLV

IANA is requested to assign a new BIER BFD Sub-sub-TLV within the BIER Info sub-TLV registry defined in [RFC8401] as follows:

Value	Description	Reference
TBD3	BIER BFD Sub-sub-TLV	[this document]

Table 3

9.4. BIER BFD Sub-TLV

IANA is requested to assign a new BIER BFD Sub-TLV from the BIER Sub-TLV registry defined in [RFC8444] as follows:

Value	Description	Reference
TBD4	BIER BFD Sub-TLV	[this document]

Table 4

10. References

10.1. Normative References

[I-D.ietf-bier-ping]
Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

[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>.

[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>.

[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>.

[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>.

[RFC8401] Ginsberg, L., Ed., Przygienda, T., Aldrin, S., and Z. Zhang, "Bit Index Explicit Replication (BIER) Support via IS-IS", RFC 8401, DOI 10.17487/RFC8401, June 2018, <https://www.rfc-editor.org/info/rfc8401>.

[RFC8444] Psenak, P., Ed., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, J., and S. Aldrin, "OSPFv2 Extensions for Bit Index Explicit Replication (BIER)", RFC 8444, DOI 10.17487/RFC8444, November 2018, <https://www.rfc-editor.org/info/rfc8444>.

[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>.

10.2. Informative References

[I-D.mirsky-mpls-p2mp-bfd] Mirsky, G., "BFD for Multipoint Networks over Point-to-Multi-Point MPLS LSP", draft-mirsky-mpls-p2mp-bfd-10 (work in progress), April 2020.

Authors' Addresses

Quan Xiong
ZTE Corporation
No.6 Huashi Park Rd
Wuhan, Hubei 430223
China

Phone: +86 27 83531060
Email: xiong.quan@zte.com.cn

Greg Mirsky
ZTE Corporation
USA

Email: gregimirsky@gmail.com

Fangwei Hu
Individual

Email: hufwei@gmail.com

Chang Liu
China Unicom
No.9 Shouti Nanlu
Beijing 100048
China

Phone: +86-010-68799999-7294
Email: liuc131@chinaunicom.cn

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 1, 2021

M. McBride
Futurewei
J. Xie
X. Geng
S. Dhanaraj
Huawei
R. Asati
Cisco
Y. Zhu
China Telecom
G. Mishra
Verizon Inc.
Z. Zhang
Juniper

September 28, 2020

BIER IPv6 Requirements
draft-ietf-bier-ipv6-requirements-09

Abstract

There have been several proposed solutions with BIER being used in IPv6. But there hasn't been a document which describes the problem and lists the requirements. The goal of this document is to describe the general BIER IPv6 encapsulation problem and detail solution requirements, thereby assisting the working group in the development of acceptable solutions.

Table of Contents

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides optimal multicast forwarding, without requiring intermediate routers to maintain per-flow state, through the use of a multicast-specific BIER header. [RFC8296] defines two types of BIER encapsulation: one is BIER MPLS encapsulation for MPLS environments, the other is non-MPLS BIER encapsulation to run without MPLS. This document describes non-MPLS BIER encapsulation in IPv6 environments. We explain the requirements of transporting multicast flow overlay payload through an IPv6 network underlay using BIER. The solutions

may use IPv6 forwarding plane and may include IPv6 encapsulation and/or generic IPv6 tunnelling.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

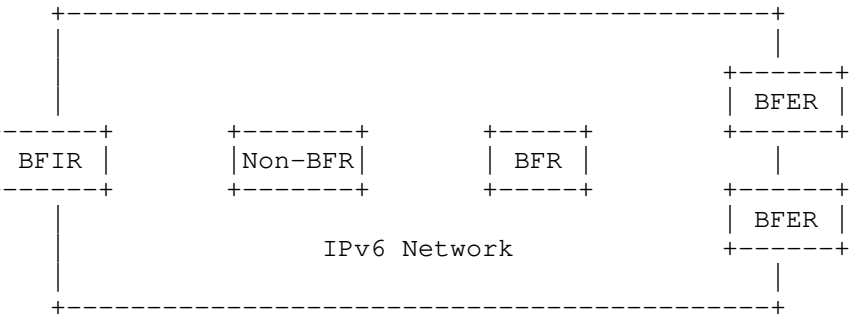
1.2. Terminology

- BIER: Bit Index Explicit Replication. Provides optimal multicast forwarding through adding a BIER header and removing state in intermediate routers.

2. Problem Statement

The problem is how to transport multicast packets, with non-MPLS BIER encapsulation, in an IPv6 environment. We need to determine where to put the BIER header in this IPv6 environment. With IPv6 encapsulation being increasingly used for unicast services, such as VPN or L2VPN, it may be desirable to have IPv6 encapsulation also used in BIER deployments for multicast services such as MVPN. It may also be desirable to not use IPv6 encapsulation except when IPv6 tunneling (native or GRE/UDP-like) is used to transport BIER packets over BIER-incapable routers.

Below is a simple scenario that needs BIER IPv6-based forwarding:



This scenario depicts the need to replicate BIER packets from a BFIR to BFERs across an IPv6 Service Provider core. Inside the IPv6 network, the BIER header is used to direct the packet from one BFR to the next BFRs, and either a IPv6 header or an L2/tunnel header is used to provide reachability between BFRs. The IPv6 environment may include a variety of link types, may be entirely IPv6, or may be dual stack. There may be cases where not all routers are BFR capable in

the IPv6 environment but still want to deploy BIER. Regardless of the environment, the problem is to deploy BIER, with non-MPLS BIER encapsulation, in an IPv6 network.

3. Requirements

There are several suggested requirements for BIER IPv6 solutions.

In this document, the requirements are divided into two levels: Mandatory and Optional. The requirement levels are determined based on the following factors:

If the requirement is required for a feature that is likely to be a potential deployment, the requirement level will be considered mandatory.

If the impact of not implementing the requirement may block BIER from been deployed, the requirement level will be considered mandatory.

3.1. Mandatory Requirements

Considering that these mandatory requirements are all well-known to the working group, and practical in normal deployment, they will be listed without a detailed description.

3.1.1. Support various L2 link types

The solution should support various kinds of L2 data link types.

3.1.2. Support BIER architecture

The solution must support the BIER architecture.

Supporting different multicast flow overlays, multiple sub-domains, multi-topologies, multiple sets, multiple Bit String Lengths, and deterministic ECMP are considered essential functions of BIER and need to be supported.

3.1.3. Support deployment with Non-BFR routers

The solution must support deployments with BIER-incapable routers. This is beneficial to the deployment of BIER, especially in early deployments when some routers do not support BIER forwarding but support IPv6 forwarding.

3.1.4. Support OAM

BIER OAM tools like [I-D.ietf-bier-ping] and [I-D.ietf-bier-pmmm-oam] should be supported, either directly using existing methods, or by specifying a new method for the same functionality. They are likely to be needed in normal BIER deployment for diagnostics.

3.2. Optional Requirements

The requirements in this section are listed as optional, and each requirement is explained with a detailed scenario. Note that fragmentation and IPSEC ESP are not BIER functions, they are provided by the upper IP layer.

3.2.1. Support Fragmentation

There are some cases where the Fragmentation/Assembly function is needed for BIER to work in an IPv6 network.

For example, a customer IPv6 multicast packet may be 1280 bytes and is required to be transported through an IPv6 network using BIER. Every link of the IPv6 network is no less than the requisite 1280 bytes [RFC8200], but the size of the payload that can be encapsulated in BIER (BIER-MTU) is less than 1280 bytes. In this case, it is not the appropriate action for a BFIR to drop the packet and advertise an MTU to the source [RFC8296]. Instead, some transport mechanism needs to provide the fragmentation and assembly function.

3.2.2. Support IPSEC ESP

There are some cases where the IPSEC ESP function may be needed to transport c-multicast packets through an IPv6 network with confidentiality using BIER technology.

A service provider may want to provide additional security SLA to its customer to ensure that the unencrypted c-multicast packet is not altered in the service provider's network. In this case, if the BIER technology is preferred for the multicast service, BIER with IPSEC ESP support may be a candidate solution. On the other hand, the traffic protection may be better provided via IPSEC or MACSEC at multicast flow overlay over and beyond the BIER domain.

4. IANA Considerations

Some BIER IPv6 encapsulation proposals do not require any action from IANA while other proposals require new IPv6 Option codepoints from IPv6 sub-registries, new "Next header" values, or require new IP

Protocol codes. This document, however, does not require anything from IANA.

5. Security Considerations

There are no security issues introduced by this draft.

6. Acknowledgement

Thanks to Eric Rosen for his listed set of initial requirements on the BIER WG mailing list.

7. Normative References

[I-D.ietf-bier-ping]
Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

[I-D.ietf-bier-pmmm-oam]
Mirsky, G., Zheng, L., Chen, M., and G. Fioccola, "Performance Measurement (PM) with Marking Method in Bit Index Explicit Replication (BIER) Layer", draft-ietf-bier-pmmm-oam-08 (work in progress), May 2020.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[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>.

Authors' Addresses

Mike McBride
Futurewei

Email: michael.mcbride@futurewei.com

Jingrong Xie
Huawei

Email: xiejingrong@huawei.com

Xuesong Geng
Huawei

Email: gengxuesong@huawei.com

Senthil Dhanaraj
Huawei

Email: senthil.dhanaraj@huawei.com

Rajiv Asati
Cisco

Email: rajiva@cisco.com

Yongqing Zhu
China Telecom

Email: zhuyq8@chinatelecom.cn

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Zhaohui Zhang
Juniper

Email: zzhang@juniper.net

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 20, 2021

H. Bidgoli, Ed.
Nokia
F. Xu
Verizon
J. Kotalwar
Nokia
I. Wijnands
M. Mishra
Cisco System
Z. Zhang
Juniper Networks
November 16, 2020

PIM Signaling Through BIER Core
draft-ietf-bier-pim-signaling-11

Abstract

Consider large networks deploying traditional PIM multicast service. Typically, each portion of these large networks have their own mandates and requirements.

It might be desirable to deploy BIER technology in some part of these networks to replace traditional PIM services. In such cases downstream PIM states need to be signaled over BIER Domain toward the source.

This draft explains the procedure to signal PIM joins and prunes through a BIER Domain, as such enable provisioning of traditional PIM services through a BIER Domain.

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 May 20, 2021.

Table of Contents

1. Introduction

Consider large networks deploying traditional PIM multicast service. Typically, each portion of these large networks have their own mandates and requirements.

It might be desirable to deploy BIER technology in some part of these networks to replace traditional PIM services. In such cases downstream PIM states need to be signaled over BIER Domain toward the source.

This draft explains the procedure to signal PIM joins and prunes through a BIER Domain, as such enable provisioning of traditional PIM services through a BIER Domain.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as describe in [RFC2119].

2.1. Definitions

Some of the terminology specified in [RFC8279] is replicated here and extended by necessary definitions:

BIER:

Bit Index Explicit Replication (The overall architecture of forwarding multicast using a Bit Position).

BFR:

Bit Forwarding Router (A router that participates in Bit Index Multipoint Forwarding).A BFR is identified by a unique BFR-prefix in a BIER domain.

BFIR:

Bit Forwarding Ingress Router (The ingress border router that performs BIER encapsulation). Each BFIR must have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plane packet (i.e. PIM) and forwarding dataplane packets. BFIR is the term used for dataplane packet forwarding.

BFER:

Bit Forwarding Egress Router. A router that participates in Bit Index Forwarding as leaf. Each BFER must have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plane packet (i.e. PIM) and forwarding dataplane packets. BFIR is the term used for dataplane packet forwarding.

BBR:

BIER Boundary router. A router between the PIM domain and BIER domain. Maintains PIM adjacency for all routers attached to it on the PIM domain and terminates the PIM adjacency toward the BIER domain.

IBBR:

Ingress BIER Boundary Router. An ingress router from signaling point of view. It maintains PIM adjacency toward the PIM domain and determines if PIM joins and prunes arriving from PIM domain need to be signaled across the BIER domain. If so it terminates the PIM adjacency toward the BIER domain and signals the PIM joins/prunes through the BIER core.

EBBR:

Egress BIER Boundary Router. An egress router in BIER domain from signaling point of view. It terminates the BIER packet and forwards the signaled joins and prunes into PIM Domain.

BFT:

Bit Forwarding Tree used to reach all BFERs in a domain.

BIFT:

Bit Index Forwarding Table.

BIER sub-domain:

A further distinction within a BIER domain identified by its unique sub-domain identifier. A BIER sub-domain can support multiple BitString Lengths.

BFR-id:

An optional, unique identifier for a BFR within a BIER sub-domain.

3. PIM Signaling Through BIER domain

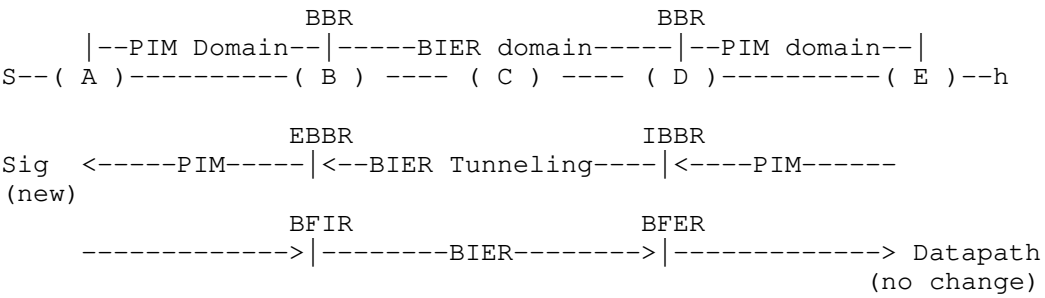


Figure 1: BIER boundary router

As per figure 1, the procedures of PIM signaling is done at the BIER boundary router. The BIER boundary routers (BBR) are connected to PIM capable routers toward the PIM domain and BIER routers toward the BIER domain. PIM routers in PIM domain continue to send PIM state messages to the BBR. The BBR will create PIM adjacency between all the PIM routers attached to it on the PIM domain. That said the BBR does not propagate all PIM packets natively into the BIER domain. Instead when it determines that the PIM join or prune messages needs to be signaled through the BIER domain it will tunnel the PIM packet through the BIER network. This tunneling is only done for signaling purposes and not for creating a PIM adjacency between the two disjoint PIM domains through the BIER domain.

The terminology ingress BBR (IBBR) and egress BBR (EBBR) are relative from signaling point of view.

The ingress BBR will determine if an arriving PIM join or prune needs to be signaled across the BIER domain. While the egress BBR will determine if the arriving BIER packet is a signaling packet and if so it will generate a PIM join/prune packet toward its attached PIM domain.

The BFER and BFIR are BBR from datapath point of view. It should be noted the new procedures in this draft are only applicable to signaling and there are no changes from datapath point of view.

3.1. Ingress BBR procedure

IBBR will create PIM adjacency to all PIM routers attached to it toward the PIM domain.

When a PIM join or prune for certain (S,G) arrives, the IBBR first determines whether the join or prune is meant for a source that is

reachable through the BIER domain. As an example, this source is located in a disjoint PIM domain that is reachable through the BIER domain. If so the IBBR will try to resolve the source via an EBBR closest to the source.

The procedure to find the EBBR (BFIR from datapath point of view) can be via many mechanisms explained in more detail in upcoming section.

After discovering the EBBR and its BFR-ID, the IBBR will include a new PIM Join Attribute in the join/prune message as per [RFC5384]. Two new "BIER IBBR" attributes are defined and explained in upcoming section. The PIM Join Attribute is used on EBBR to obtain necessary BIER information to build its multicast states. In addition the IBBR will change the PIM signaling packet source IP address to its BIER prefix address (standard PIM procedure). It will also keep the destination address as the well known multicast IP address. It then will construct the BIER header. The signaling packet, in this case the PIM join/prune packet, is encapsulated in the BIER header and transported through BIER domain to EBBR.

The IBBR will track all the PIM interfaces on the attached PIM domain which are interested in a certain (S,G). It creates multicast states for arriving (S,G)s from PIM domain, with incoming interface as BIER "tunnel" interface and outgoing interface as the PIM domain interface(s) on which PIM Join(s) were received on.

3.1.1. Determining EBBR on IBBR

As it was explained in the previous section, IBBR needs to determine the EBBR closest to the source. This is needed to encode the BIER header BitString field to forward the signaling packet through the BIER domain.

It should be noted, the PIM domains can be either part of the same IGP area as BIER domain(single area) or are stitched to the BIER domain via an ABR or ASBR routers. As such on IBBR, there can be many different procedures to determine the EBBR. Some examples of these procedures have been provided in Appendix A.

3.1.2. Considering ECMP in EBBR selection

If the lookup for source results into multiple EBBRs, then the EBBR selection algorithm should ensure that all signaling for a particular (C-S, C-G) is forwarded to a single EBBR. How this selection is done is vendor specific and beyond this draft. As an example it can be round robin per (C-S, C-G) or lowest EBBR IP for all (C-S, C-G)s.

3.1.3. PIM Signaling packet construction at IBBR

To ensure all necessary BIER information needed by EBBR is present in the BIER signaling message, a new PIM Join Attribute [RFC5384] is used. EBBR can use this attribute to build its multicast states, as described in EBBR procedure section. This new PIM join Attribute is added to PIM signaling message on the IBBR. Its format is as follow:

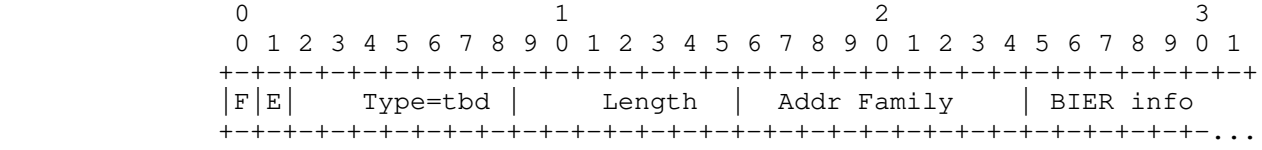


Figure 2: PIM Join Attribute

F bit: The Transitive bit. Specifies whether this attribute is transitive or non-transitive. MUST be set to zero. This attribute is ALWAYS non-transitive.

E bit: End-of-Attributes bit. Specifies whether this attribute is the last. Set to zero if there are more attributes. Set to 1 if this is the last attribute.

Type: TBD assign by IANA.

Length: The length in octets of the attribute value. MUST be set to the length in octets of the BIER info +1 octet to account for the Address Family field. For IPv4 AF Length = 7+1 For IPv6 AF Length = 19+1.

Addr Family: Signaled PIM Join/Prune address family as defined in [RFC7761].

BIER Info: IBBR Prefix (IPv4 or IPv6), SD, bfr-id as per below figure

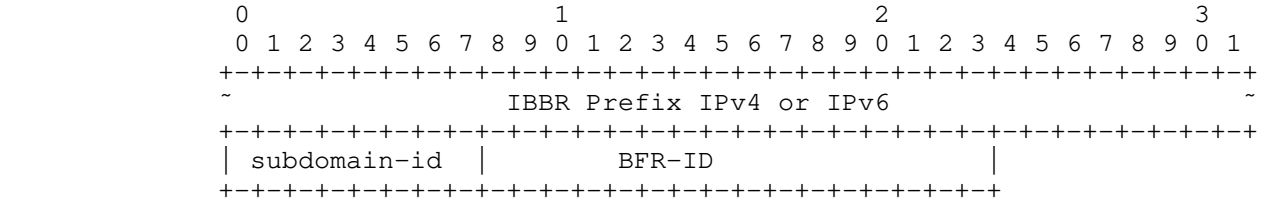


Figure 3: PIM Join Attribute detail

3.1.3.1. BIER packet construction at IBBR

The BIER header will be encoded with the BFR-id of the IBBR(with appropriate bit set in the BitString) and the PIM signaling packet is then encapsulated in the packet.

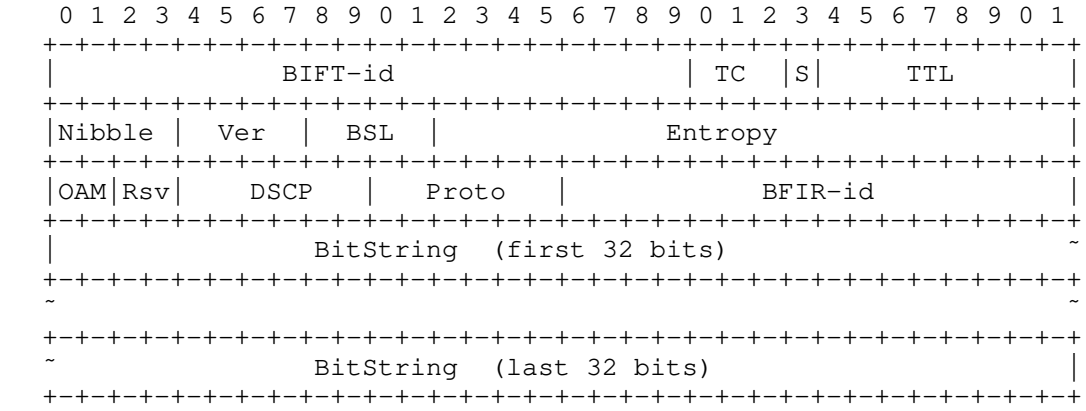


Figure 4: BIER header

BIERHeader.Proto = IPv4 or IPv6

BIERHeader.BitString= Bit corresponding to the BFR-ID of the EBBR

BIERHeader.BFIR-id = BFR-Id of the BBR originating the encapsulated PIM packet, i.e. the IBBR.

Rest of the values in the BIER header are determined based on the network (MPLS/non-MPLS), capabilities (BSL), and network configuration.

3.2. Signaling PIM through the BIER domain procedure

Throughout the BIER domain the BIER forwarding procedure is on par with [RFC8279]. No BIER router will examine the BIER packet encapsulating the PIM signaling packet. As such there is no multicast state built in the BIER domain.

The packet will be forwarded through the BIER domain until it reaches the BER with matching BFR-ID as in the BIERHeader.Bitstring. EBBR will remove the BIER header and examine the PIM IPv4 or IPv6 signaling packet further as per EBBR Procedure section.

3.3. EBBR procedure

EBBR will remove the BIER header and determine this is a signaling packet. The Received PIM join/prune Signaling packet is processed as if it were received from neighbors on a virtual interface, (i.e. as if the pim adjacency was present, regardless of the fact that there is no adjacency).

The EBBR will build a forwarding table for the arriving (S,G) using the obtained BFIR-id and the Sub-Domain information from BIER Header and/or the PIM join Attributes added to the PIM Signaling packet. In short it tracks all IBBRs interested in this (S,G). This is explained in section 4.1.

The multicast state on EBBR will contain PIM domain incoming interfaces, according to PIM specification and outgoing interfaces based on the above procedure to build the forwarding table.

It should be noted EBBR will maintain PIM adjacency toward the PIM domain and all PIM routers which are connected to it. At this point the end-to-end multicast traffic flow setup is complete.

4. Datapath Forwarding

4.1. BFIR tracking of (S,G)

For a specific Source and Group, BFIR (EBBR)should track all the interested BFERs (IBBRs) via PIM signaling messages arriving from the BIER Domain. BFIR builds its (s,g) forwarding state with incoming interface (IIF) as the Reverse Path Forwarding (RPF) interface (in attached PIM domain) towards the source. The outgoing interfaces are the tracked BFERs in the Bier Sub Domain.

4.2. Datapath traffic flow

When the multicast data traffic arrives on the BFIR (EBBR) the router will find all the interested BFERs for that specific (S,G). The router then constructs the BIERHeader.BitString with all the BFER interested in the group and will forward the packet to the BIER domain. The BFER(s) will accept the packets and remove the BIER header and forward the multicast packet as per pre-built multicast state for (S,G) and its outgoing interfaces.

5. PIM-SM behavior

The procedures described in this document can work with Any-Source Mutlicast (ASM) as long as static Rendevous Point (RP) or embedded RP

for IPv6 is used. Future drafts would cover Bootstrap Router (BSR) and more complicated SM discovery mechanisms.

It should be noted that this draft only signals PIM Joins and Prunes through the BIER domain and not any other PIM message types including PIM Hellos or Asserts. As such functionality related to these other type of massages will not be possible through a BIER domain with this draft and future drafts might cover these scenarios. As an example DR selection should be done in the PIM domain or if the PIM routers attached to IBBRs are performing DR selection there needs to be a dedicated PIM interface between these routers.

In case of PIM ASM Static RP or embedded RP for IPv6 the procedure for leaves joining RP is the same as above. It should be noted that for ASM, the EBBRs are determined with respect to the RP instead of the source.

6. Applicability to MVPN

With just minor changes, the above procedures apply to MVPN as well, with BFIR/BFER/EBBR/IBBR being VPN PEs. All the PIM related procedures, and the determination of EBBR happens in the context of a VRF, following procedures for PIM-MVPN.

When a PIM packet arrives from PIM domain attached to the VRF (IBBR), and it is determined that the source is reachable via the VRF through the BIER domain, a PIM signaling message is sent via BIER to the EBBR. In this case usually the PE terminating the PIM-MVPN is the EBBR. A label is imposed before the BIER header is imposed, and the "proto" field in the BIER header is set to 1 (for "MPLS packet with downstream-assigned label at top of stack"). The label is advertised by the EBBR/BFIR to associate incoming packets to its correct VRF. In many scenarios a label is already bound to the VRF loopback address on the EBBR/BFIR and it can be used.

When a multicast data packet is sent via BIER by an EBBR/BFIR, a label is imposed before the BIER packet is imposed, and the "proto" field in the BIER header is set to 1 (for "MPLS packet with downstream-assigned label at top of stack"). The label is assigned to the VPN consistently on all VRFs [draft-zzhang-bess-mvpn-evpn-aggregation-label-01].

If the more complicated label allocation scheme is needed for the data packets as specified in [draft-zzhang-bess-mvpn-evpn-aggregation-label-01], then additional PMSI signaling is needed as specified in [RFC6513].

To support per-area subdomain in this case, the ABRs would need to become VPN PEs and maintain per-VPN state so it is unlikely practical.

7. IANA Considerations

In the "PIM Join Attribute Types" registry, IANA to assign a new value [TBD] to the BIER Info Vector.

8. Security Considerations

The procedures of this document do not, in themselves, provide privacy, integrity, or authentication for the control plane or the data plane. For a discussion of the security considerations regarding the use of BIER, please see [RFC8279] and [RFC8296]. Security considerations regarding PIM protocol is based on [RFC7761].

9. Acknowledgments

The authors would like to thank Eric Rosen, Stig Venaas for thier reviews and comments.

10. References

10.1. Normative References

[RFC4607] "H. Holbrook, B. Cain, "Source-Specific multicast for IP"", October 2016.

[RFC8279] "Wijnands, IJ., Rosen, E., Dolganow, A., Przygienda, T. and S. Aldrin, "Multicast using Bit Index Explicit Replication"", October 2016.

10.2. Informative References

[draft-zzhang-bess-mvpn-evpn-aggregation-label-01] "Z. Zhang, E. Rosen, W. Lin, Z. Li, I.Wijnands, "MVPN/EVPN Tunnel Aggregation with Common labels"", April 2018.

[RFC2119] "S. Brandner, "Key words for use in RFCs to Indicate Requirement Levels"", March 1997.

[RFC5384] "A. Boers, I. Wijnands, E. Rosen, "PIM Join Attribute Format"", November 2008.

[RFC6513] "E. Rosen, R. Aggarwal, "Multicast in MPLS/BGP IP VPNs"", November 2008.

[RFC6826] "IJ. Wijnands, T. Echert, N. Leymann, M. Napierala, "Multipoint LDP In-Band Singnaling for PtP MPtMP LSP"", January 2013.

[RFC7761] "B.Fenner, M.Handley, H. Holbrook, I. Kouvelas, R. Parekh, Z.Zhang "PIM Sparse Mode"", March 2016.

[RFC8296] "IJ. Wijnands, E. Rosen, A. Dolganow, J. Yantsura, S. Aldrin, I. Meilik, "Encapsulation for BIER"", January 2018.

[RFC8401] "Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER Support via ISIS"", June 2018.

[RFC8444] "Psenak, P., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for Bit Index Explicit Replication"", June 2018.

[RFC8556] "Rosen, E., Ed., Sivakumar, M., Wijnands, IJ., Aldrin, S.,Dolganow, A., and T. Przygienda, "Multicast VPN Using BIER"", March 2018.

Appendix A.

This appendix provides some examples and routing procedures that can be used to determine the EBBR on IBBR. It should be noted, the PIM domains can be either part of the same IGP area as BIER domain(single area) or are stitched to the BIER domain via an ABR or ASBR routers. As such on IBBR, there can be many different procedures to determine the EBBR. Not all procedures are listed below.

A.1. SPF

On IBBR SPF procedures can be used to find the EBBR closest to the source.

Assuming the BIER domain consists of all BIER forwarding routers, SPF calculation can identify the router advertising the prefix for the source. A post process can find the EBBR by walking from the advertising router back to the IBBR in the reverse direction of shortest path tree branch until the first BFR is encountered.

A.2. Indirect next-hop

Alternatively, the route to the source could have an indirect next-hop that identifies the EBBR. These methods are explained in the following sections.

A.2.1. Static Route

On IBBR there can be a static route configured for the source, with source next-hop set as EBBR BIER prefix.

A.2.2. Interior Border Gateway Protocol (iBGP)

Consider the following topology:

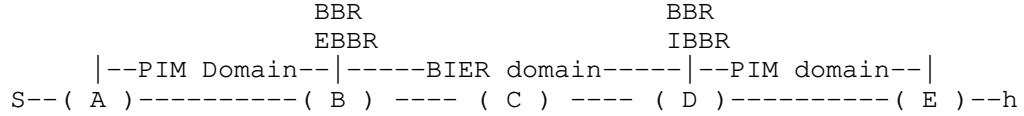


Figure 5: Static Route

Suppose BGP is enable between EBBR (B) and IBBR (D) and the PIM Domain routes are redistributed to the BIER domain via BGP. This would include the Multicast Source IP address (S), which resides in the PIM Domain. In such case BGP should use the same loopback interface as its next-hop as the BBR prefix. This will ensure that all PIM domain routes, including the Multicast Source IP address (S) are resolve via BBR's BIER prefix id as their next-hop. When the host (h) triggers a PIM join message to IBBR (D), IBBR tries to resolve (S). It resolves (S) via BGP installed route and realizes its next-hop is EBBR (B). IBBR will use this next-hop (B) to find its corresponding BIER bit index.

This procedure is inline with [RFC6826] mLDP in-band signaling section

A.3. Inter-area support

If each area has its own BIER sub-domain, the above procedure for post-SPF could identify one of the ABRs and the EBBR. If a sub-domain spans multiple areas, then additional procedures as described in A.2 is needed.

A.3.1. Inter-area Route summarization

In a multi-area topology, a BIER sub-domain can span a single area. Suppose this single area is constructed entirely of BIER capable routers and the ABRs are the BIER Boundary Routers attaching the BIER sub-domain in this area to PIM domains in adjacent areas. These BBRs can summarize the PIM domain routes via summary routes, as an example for OSPF, a type 3 summary LSAs can be used to advertise summary routes from a PIM domain area to the BIER area. In such scenarios the IBBR can be configured to look up the Source via IGP database and

use the summary routes and its Advertising Router field to resolve the EBBR. The IBBR needs to ensure that the IGP summary route is generated by a BFR. This can be achieved by ensuring that BIER Sub-TLV exists for this route. If multiple BBRs (ABRs) have generated the same summary route the lowest Advertising Router IP can be selected or a vendor specific hashing algorithm can select the summary route from one of the BBRs.

Authors' Addresses

Hooman Bidgoli (editor)
Nokia
Ottawa
Canada

Email: hooman.bidgoli@nokia.com

Fengman Xu
Verizon
Richardson
US

Email: fengman.xu@verizon.com

Jayant Kotalwar
Nokia
Montain View
US

Email: jayant.kotalwar@nokia.com

IJsbrand Wijnands
Cisco System
Diegem
Belgium

Email: ice@cisco.com

Mankamana Mishra
Cisco System
Milpitas
USA

Email: mankamis@cisco.com

Zhaohui Zhang
Juniper Networks
Boston
USA

Email: zzhang@juniper.com

Network Working Group
Internet-Draft
Updates: 8296 (if approved)
Intended status: Standards Track
Expires: January 14, 2021

J. Xie
Huawei Technologies
L. Geng
China Mobile
M. McBride
Futurewei
R. Asati
Cisco
S. Dhanaraj
Huawei
Y. Zhu
China Telecom
Z. Qin
China Unicom
M. Shin
LG Uplus
G. Mishra
Verizon Inc.
X. Geng
Huawei

July 13, 2020

Encapsulation for BIER in Non-MPLS IPv6 Networks
draft-xie-bier-ipv6-encapsulation-08

Abstract

This document proposes a BIER IPv6 (BIERv6) encapsulation for Non-MPLS IPv6 Networks using the IPv6 Destination Option extension header. This document updates RFC 8296.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] and [RFC8174].

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/.

Copyright Notice

Table of Contents

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides optimal multicast forwarding without requiring intermediate routers to maintain any per-flow state by using a multicast-specific BIER header.

[RFC8296] defines a common BIER Header format for MPLS and Non-MPLS networks. It has defined two types of encapsulation methods using the common BIER Header, (1) BIER encapsulation in MPLS networks, here-in after referred as MPLS BIER Header in this document and (2) BIER encapsulation in Non-MPLS networks, here-in after referred as Non-MPLS BIER Header in this document. [RFC8296] also assigned Ethertype=0xAB37 for Non-MPLS BIER Header packets to be directly carried over the Ethernet links.

This document proposes a BIER IPv6 encapsulation for Non-MPLS IPv6 Networks, defining a method to carry the standard Non-MPLS BIER header (as defined in [RFC8296]) in the native IPv6 header. A new IPv6 Option type - BIER Option is defined to encode the standard Non-MPLS BIER header and this newly defined BIER Option is carried under the Destination Options header of the native IPv6 Header [RFC8200].

The relationship of this document to BIER core standards is listed in Appendix A.

The relevant extensions to BIER Control-plane Standards are listed in Appendix B.

2. Terminology

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References.

The following new terms are used throughout this document:

- BIERv6 - Bit indexed explicit replication using IPv6 data plane.

- BIERv6 Domain - A limited-domain using BIERv6 encapsulation as specified in this document for transporting customer multicast packets from one router to multiple destination routers. It is usually managed by a single administrative entity, e.g., a service-provider. It could be a single AS network or a large-scale network that includes multiple ASes. BIER Domain is also used for the same meaning as BIERv6 domain in this document.

- BIERv6 Option - An Option type carried in IPv6 Destination Options Header (DO header, DOH) which includes the standard Non-MPLS BIER Header. It is in type-length-value (TLV) format. The value portion of the BIERv6 Option TLV, or the BIERv6 Option Data, is in the format of the standard Non-MPLS BIER header. BIER option is also used for the same meaning as BIERv6 option in this document.

- BIERv6 Header - An IPv6 Header with BIER Option.

- BIERv6 Packet - An IPv6 packet with BIERv6 Header. An IP/IPv6/ Ethernet multicast packet is encapsulated with an outside BIERv6 header and transformed to a BIERv6 packet on the ingress PE (BFIR). BIERv6 packet is transported by the transit routers (BFRs) through a BIERv6 domain towards egress PEs(BFERs). BIERv6 packet is decapsulated by the BFERs, with the original IP/IPv6/ Ethernet multicast packet being obtained and forwarded towards the multicast receivers .

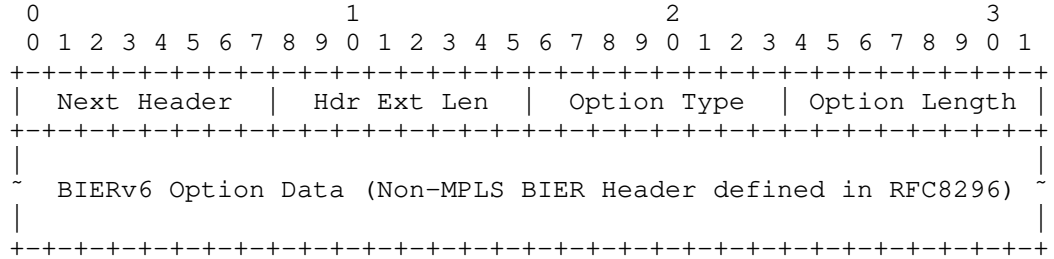
3. BIER IPv6 Encapsulation

3.1. BIER Option in IPv6 Destination Options Header

Destination Options Header and the Options that can be carried under this extension header is defined in [RFC8200]. This document defines a new Option type - BIER Option, to encode the Non-MPLS BIER header. As specified in Section 4.2 [RFC8200], the BIER Option follows type-length-value (TLV) encoding format and the standard Non-MPLS BIER header [RFC8296] is encoded in the value portion of the BIER Option TLV.

This BIER Option MUST be carried only inside the IPv6 Destination Options header and MUST NOT be carried under the Hop-by-Hop Options header.

The BIER Option is encoded in type-length-value (TLV) format as follows:



Next Header 8-bit selector. Identifies the type of header immediately following the Destination Options header.

Hdr Ext Len 8-bit unsigned integer. Length of the Destination Options header in 8-octet units, not including the first 8 octets.

Option Type To be allocated by IANA. See section 6.

Option Length 8-bit unsigned integer. Length of the option, in octets, excluding the Option Type and Option Length fields.

BIERv6 Option Data The BIERv6 Option Data contains the Non-MPLS BIER Header defined in RFC8296. Fields in the Non-MPLS BIER Header MUST be encoded as below.

BIFT-id: The BIFT-id is a domain-wide unique value in Non-MPLS IPv6 encapsulation. See Section 2.2 of RFC 8296.

TC: SHOULD be set to binary value 000 upon transmission and MUST be ignored upon. See Section 2.2 of RFC 8296.

S bit: SHOULD be set to 1 upon transmission, and MUST be ignored upon reception. See Section 2.2 of RFC 8296.

TTL: MUST be set to a value larger than 0 upon encapsulation, and SHOULD decrease by 1 by a BFR when forwarding a BIERv6 packet to a BFR adjacency. If the incoming TTL is 0, the packet is considered to be "expired". See Section 2.1.1.2 of RFC 8296.

Nibble: SHOULD be set to 0000 upon transmission, and MUST be ignored upon reception. See Section 2.2 of RFC 8296.

Ver: MUST be set to 0 upon transmission, and MUST be discarded when it is not 0 upon reception. See Section 2.2 of RFC 8296.

BSL: See Section 2.1.2 of RFC 8296.

Entropy: See Section 2.1.2 of RFC 8296.

OAM: See Section 2.1.2 of RFC 8296.

Rsv: See Section 2.1.2 of RFC 8296.

DSCP: SHOULD be set to binary value 000000 upon transmission and MUST be ignored upon reception. In BIERv6 encapsulation, uses Traffic Class field of IPv6 header instead.

Proto: SHOULD be set to 0 upon transmission and be ignored upon reception. In BIERv6 encapsulation, the functionality of this 6-bit Proto field is replaced by the Next Header field in Destination Options header or the last IPv6 extension header to indicate the type of the payload. This updates section 2.1.2 of [RFC8296] about Proto definition. Next Header value in BIERv6 encapsulation for common usage includes:

Value 4 for IPv4 packet as BIERv6 payload.

Value 41 for IPv6 packet as BIERv6 payload.

Value 143 for Ethernet packet as BIERv6 payload.

Multicast VPN (MVPN) service is considered as part of the BIER layering mode defined in [RFC8279], and should be supported by BIERv6 encapsulation. [I-D.xie-bier-ipv6-mvpn] illustrates how MVPN is supported in BIERv6 encapsulation without using this Proto field.

BIER-PING [I-D.ietf-bier-ping] is considered a useful function of the BIER architecture, and should be supported by BIERv6 encapsulation. How BIER-PING is supported in BIERv6 encapsulation without using this Proto field is outside the scope of this document.

BFIR-id: See Section 2.1.2 of RFC 8296.

BitString: See Section 2.1.2 of RFC 8296.

3.2. Destination Address in BIERv6 Encapsulation

When a BIERv6 packet is replicated to a next hop BFR, an unicast address of the next hop BFR is used as the destination address of the BIERv6 packet. Considerations of using unicast (or multicast) address is listed in Appendix C.

The unicast address used in BIERv6 packet targeting a BFR SHOULD be advertised as part of the BIER IPv6 Encapsulation. When a BFR advertises the BIER information with BIERv6 encapsulation capability, an IPv6 unicast address of this BFR MUST be selected specifically for BIERv6 packet forwarding. Locally this "BIER Specific" IPv6 address is initialized in FIB with a flag of "BIER specific handling", represented as End.BIER function.

If a BFR belongs to more than one sub-domain, it may (though it need not) have a different End.BIER in each sub-domain. If different End.BIER is used for each sub-domain, implementation SHOULD support

verifying the DA of a BIERv6 packet is the End.BIER address bound by the sub-domain of the packet.

For security deployment of BIERv6, the End.BIER address(es) is required to be allocated from an IPv6 address block, and the IPv6 address block is used for domain boundary security policy. See section 5.1 of this document for such security policy. Such kind of security policy using IPv6 address block follows the paradigm settled by the [RFC8754] section 5.

Deployment of BIERv6 in SRv6 network is allowed. In this case, the BIERv6 domain is the same as SRv6 domain, and the End.BIER address is allocated from the locator of SRv6.

To better understand the configuration mode of End.BIER address in BIERv6, [I-D.geng-bier-bierv6-yang] could be referenced.

For the convenience of such co-existence of BIERv6 and SRv6, the indication of End.BIER or "BIER specific handling" in FIB shares the same space as SRv6 Endpoints Behaviors defined in [I-D.ietf-spring-srv6-network-programming].

The following is an example pseudo-code of the End.BIER function:

```
 1. IF NH = 60 and HopLimit > 0                                  ;;Ref1
 2.   IF (OptType1 = BIER) and (OptLength1 = HdrExtLen*8 + 4) ;;Ref2
 3.     Lookup the BIER Header inside the BIER option TLV.
 4.     Forward via the matched entry.
 5.   ELSE                                                       ;;Ref3
 6.     Drop the packet and end the process.
 7. ELSE IF NH=ICMPv6 or (NH=60 and Dest_NH=ICMPv6)             ;;Ref4
 8.   Send to CPU.
 9. ELSE                                                        ;;Ref5
10.   Drop the packet.
```

Ref1: Destination options header follows the IPv6 header directly and HopLimit is bigger than zero.

Ref2: The first TLV is BIER type and is the only TLV present in Destination options header.

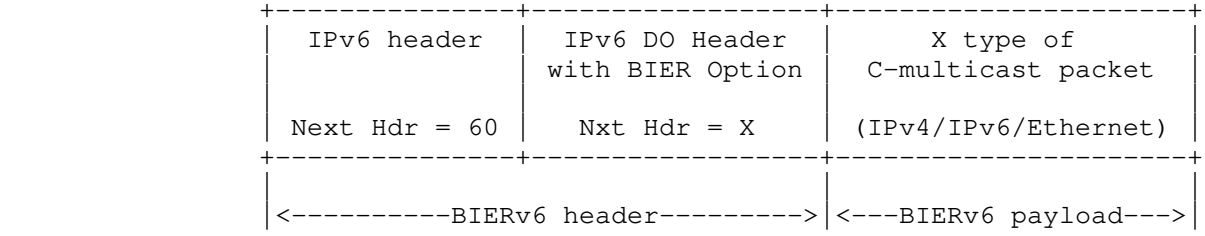
Ref3/Ref5: Undesired packet is droped because the destination address is the BIER specific IPv6 address (End.BIER function).

Ref4: An ICMPv6 packet using End.BIER as destination address.

3.3. BIERv6 Packet Format

As a multicast packet enters the BIER domain in a Non-MPLS IPv6 network, the multicast packet will be encapsulated with BIERv6 Header by the Ingress BFR (BFIR).

Typically a BIERv6 header would contain the Destination Options Header as the only Extensions Header besides IPv6 Header, as depicted in the below figure.



Format of the multicast packet with BIERv6 encapsulation carrying other extension headers along with Destination Options extension header is required to follow general recommendations of [RFC8200] and examples in other RFCs. [RFC6275] introduces how the order should be when other extension headers carries along with Home address option in a destination options header. Similar to this example, this document requires the Destination Options Header carrying the BIER option MUST be placed as follows:

- After the routing header, if that header is present
- Before the Fragment Header, if that header is present
- Before the AH Header or ESP Header, if either one of those headers is present

Source Address field in the IPv6 header MUST be a routable IPv6 unicast address of the BFIR in any case.

BFIR encodes the BIERv6 header in the above mentioned encapsulation format and forwards the BIERv6 packet to the nexthop BFR following the local BIFT table.

BFRs in the IPv6 network, processes and replicates the packets towards the BFERs using the local BIFT table. The BitString field in the BIERv6 Option Data may be changed by the BFRs as they replicate the packet. BFRs MUST follow the procedures defined in section 3.1 as they modify the other fields in the BIERv6 Option Data. The source address in the IPv6 header MUST NOT be modified by the BFRs.

4. BIERv6 Packet Processing

When a multicast packet enters the BIER domain, the Ingress BFR (BFIR) encapsulates the multicast packet with a BIERv6 Header, transforming it to a BIERv6 packet. The BIERv6 header includes an IPv6 header and a BIERv6 Option in IPv6 Destination Options Header. Source Address field in the IPv6 header MUST be set to a routable IPv6 unicast address of the BFIR. Destination Address field in the IPv6 header is set to the End.BIER address of the next-hop BFR the BIERv6 packet replicating to, no matter next-hop BFR is directly connected (one-hop) or not directly connected (multi-hop).

Upon receiving an BIERv6 packet, the BFR processes the IPv6 header first. This is the general procedure of IPv6.

If the IPv6 Destination address is an End.BIER IPv6 unicast address of this BFR, a 'BIER Specific Handling' indication will be obtained by the preceding Unicast DA lookup (FIB lookup). The BIER option, if exists, will be checked to decide which neighbor(s) to replicate the BIERv6 packet to.

It is a local behavior to handle the combination of extension headers, options and the BIER option(s) in destination options header when a 'BIER Specific Handling' indication is got by the preceding FIB lookup. Early deployment of BIERv6 may require there is only one BIER option TLV in the destination options header followed the IPv6 header. How other extension headers or more BIER option TLVs in a BIERv6 packet is handled is outside the scope of this document.

A packet having a 'BIER Specific Handling' indication but not having a BIER option is supposed to be a wrong packet or an ICMPv6 packet, and the process can be referred to the example in section 3.2.

A packet not having a 'BIER Specific Handling' indication but having a BIER option SHOULD be processed normally as unicast forwarding procedures, which may be a behavior of drop, or send to CPU, or other behaviors in existing implementations.

The Destination Address field in the IPv6 Header MUST change to the nexthop BFR's End.BIER Unicast address in BIERv6.

The Hop Limit field of IPv6 header MUST decrease by 1 when sending packets to a BFR neighbor, while the TTL in the BIER header MUST be unchanged on a Non-BIER router, or decrease by 1 on a BFR.

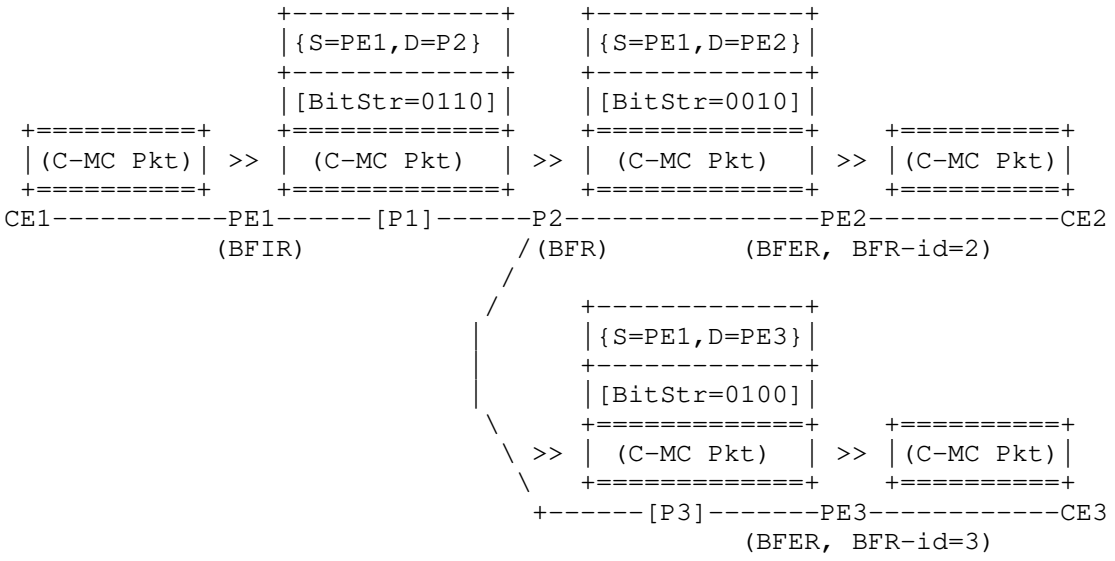
The BitString in the BIER header in the Destination Options Header may change when sending packets to a neighbor. Such change of BitString MUST be aligned with the procedure defined in RFC8279.

Because of the requirement to change the content of the option when forwarding BIERv6 packet, the BIER option type should have chg flag 1 per section 4.2 of RFC8200.

The procedures applies normally if a bit corresponding to the self bfr-id is set in the BitString field of the BIERv6 Option Data of the BIERv6 packet. The node is considered to be an Egress BFR (BFER) in this case. The BFER removes the BIERv6 header, including the IPv6 header and the Destination Options header, and copies the packet to the multicast flow overlay. The egress VRF of a packet may be determined by a further lookup on the IPv6 source address instead of the upstream-assigned MPLS Label as described in [RFC8556].

The Fragment Header, AH Header or ESP Header, if exists after the BIER options header, can be processed on BFER only as part of the multicast flow overlay process.

The following diagram shows the whole progression of the multicast packet as it enters the BIERv6 domain on PE1, and leaves the BIERv6 domain on PE2 and PE3.



{S=PE1,D=PE2}: Source address and Destination address in IPv6 header.
[BitStr=0110]: BitString value in IPv6 DO Header.
(C-MC Pkt): Customer MultiCast packet.

o PE1 is Provider Edge router, acting as BFIR.

- o P2 is Provider Core router, acting as BFR.
- o P1 and P3 are IPv6 routers, acting as Non-BFR.
- o PE2 and PE3 are Provider Edge routers, acting as BFER.
- o CE1 and CE2 are Customer Edge routers.

5. Security Considerations

BIER IPv6 encapsulation provides a new encapsulation based on IPv6 and BIER to transport multicast data packet in a BIER domain. The BIER domain can be a single IGP area, an anonymous system (AS) with multiple IGP areas, or multiple anonymous systems (ASes) operated by a network operator. A single BIER Sub-domain may be deployed through the whole BIER Domain, as illustrated in [I-D.geng-bier-ipv6-inter-domain].

This section reviews security considerations related to the BIER IPv6 encapsulation, based on security considerations of [RFC8279], [RFC8296], and other documents related to IPv6 extension.

It is expected that all nodes in a BIER IPv6 domain are managed by the same administrative entity. BIER-encapsulated packets should generally not be accepted from untrusted interfaces or tunnels. For example, an operator may wish to have a policy of accepting BIER-encapsulated packets only from interfaces to trusted routers, and not from customer-facing interfaces. See section 5.1 for normal Intra domain deployment.

For applications that require a BFR to accept a BIER-encapsulated packet from an interface to a system that is not controlled by the network operator, the security considerations of [RFC8296] apply.

BIER IPv6 encapsulation may cause ICMP packet sent to BFIR and cause security problems. See section 5.2 for ICMP related problems.

This document introduces a new option used in IPv6 Destination Options Header, together with the special-use IPv6 address called End.BIER in IPv6 destination address for BIER IPv6 forwarding. However the option newly introduced may be wrongly used with normal IPv6 destination address. See section 5.3 for problems introduced by the new IPv6 option with normal IPv6 destination address.

If the multicast data packet of a BIERv6 packet is altered by an intermediate router, contents of the multicast data packet will be damaged. BIER IPv6 encapsulation provides the ability of IPsec to

ensure the confidentiality or integrity for multicast data packet. See section 5.4 for this security problem.

If the BIERv6 encapsulation of a particular packet specifies a BitString (together with SI) other than the one intended by the BFIR, the packet is likely to be misdelivered. Some modifications of the BIER encapsulation, e.g., setting every bit in the BitString, may result in denial-of-service (DoS) attacks. This kind of DoS attack is a challenge not only in BIERv6 but also in BIER as specified in [RFC8279] and [RFC8296], as the BitString is required to change on BFR per the BIER forwarding procedures. This document does not provide new mechanisms to improve this kind of weakness.

A BIER router accepts and uses the End.BIER IPv6 address to construct BIFT only when the IPv6 address is configured explicitly, or received from a router via control-plane protocols. The received information is validated using existing authentication and security mechanisms of the control-plane protocols. BIER IPv6 encapsulation does not define any additional security mechanism in existing control-plane protocols, and it inherits any security considerations that apply to the control-plane protocols.

5.1. Intra Domain Deployment

Generally nodes outside the BIER Domain are not trusted: they cannot directly use the End.BIER of the domain. This is enforced by two levels of access control lists:

1. Any packet entering the BIER Domain and destined to an End.BIER IPv6 Address within the BIER Domain is dropped. This may be realized with the following logic. Other methods with equivalent outcome are considered compliant:

* allocate all the End.BIER IPv6 Address from a block S/s

* configure each external interface of each edge node of the domain with an inbound infrastructure access list (IACL) which drops any incoming packet with a destination address in S/s

* Failure to implement this method of ingress filtering exposes the BIER Domain to BIER attacks as described and referenced in [RFC8296].

2. The distributed protection in #1 is complemented with per node protection, dropping packets to End.BIER IPv6 Address from source addresses outside the BIER Domain. This may be realized with the following logic. Other methods with equivalent outcome are considered compliant:

* assign all interface addresses from prefix A/a

* assign all the IPv6 addresses used as source address of BIER IPv6 packets from a block B/b

* at node k, all End.BIER IPv6 addresses local to k are assigned from prefix Sk/sk

* configure each internal interface of each BIER node k in the BIER Domain with an inbound IACL which drops any incoming packet with a destination address in Sk/sk if the source address is not in A/a or B/b.

For simplicity of deployment, a configuration of IACL effective for all interfaces can be provided by a router. Such IACL can be referred to as global IACL(GIACL) .Each BIER node k then simply configs a GIACL which drops any incoming packet with a destination address in Sk/sk if the source address is not in A/a or B/b for the intra-domain deployment mode.

5.2. ICMP Error Processing

The BIERv6 BFR does not send ICMP error messages to the source address of a BIERv6 packet, there is still chance that Non-BFR routers send ICMP error messages to source nodes within the BIER Domain.

A large number of ICMP may be elicited and sent to a BFIR router, in case when a BIERv6 packet is filled with wrong Hop Limit, either error or malfeasance. A rate-limiting of ICMP packet should be implemented on each BFR.

The ingress node can take note of the fact that it is getting, in response to BIER IPv6 packet, one or more ICMP error packets. By default, the reception of such a packets MUST be countered and logged. However, it is possible for such log entries to be "false positives" that generate a lot of "noise" in the log; therefore, implementations SHOULD have a knob to disable this logging.

5.3. Security caused by BIER option

This document introduces a new option used in IPv6 Destination Options Header. An IPv6 packet with a normal IPv6 address of a router (e.g. loopback IPv6 address of the router) as destination address will possibly carry a BIER option.

For a router incapable of BIERv6, such BIERv6 packet will not be processed by the procedure described in this document, but be

processed as normal IPv6 packet with unknown option, and the existing security considerations for handling IPv6 options apply. Possible way of handling IPv6 packets with BIER option may be send to CPU for slow path processing, with rate-limiting, or be discarded according to the local policy.

For a router capable of BIERv6, such BIERv6 packet MUST NOT be forwarded, but should be processed as a normal IPv6 packet with unknown option, or additionally and optionally be countered and logged if the router is capable of doing so.

5.4. Applicability of IPsec

IPsec [RFC4301] uses two protocols to provide traffic security services -- Authentication Header (AH) [RFC4302] and Encapsulating Security Payload (ESP) [RFC4303]. Each protocol supports two modes of use: transport mode and tunnel mode. IPsec support both unicast and multicast. IPsec implementations MUST support ESP and MAY support AH.

This document assume IPsec working in tunnel mode with inner IPv4 or IPv6 multicast packet encapsulated in outer BIERv6 header and IPsec header(s).

IPsec used with BIER IPv6 encapsulation to ensure that a BIER payload is not altered while in transit between BFIR and BFERs. If a BFR in between BFIR and BFERs is compromised, there is no way to prevent the compromised BFR from making illegitimate modifications to the BIER payload or to prevent it from misforwarding or misdelivering the BIER-encapsulated packet, but the BFERs will detect the illegitimate modifications to the BIER Payload (or the inner multicast data packet). This could provide cryptographic integrity protection for multicast data transport. This capability of IPsec comes from the design that, the destination options header carrying the BIER header is located before the AH or ESP and the BFR routers in between BFIR and BFERs can process the BIER header without aware of AH or ESP.

For ESP, the Integrity Check Value (ICV) is computed over the ESP header, Payload, and ESP trailer fields. It doesn’t require the IP or extension header for ICV calculating, and thus the change of DA and BIER option data does not affect the function of ESP.

For AH, the Integrity Check Value (ICV) is computed over the IP or extension header fields before the AH header, the AH header, and the Payload. The IPv6 DA is immutable for unicast traffic in AH, and the change of DA in BIER IPv6 forwarding for multicast traffic is incompatible to this rule. How AH is extended to support multicast

traffic transporting through BIER IPv6 encapsulation is outside the scope of this document.

The detailed control-plane for BIER IPv6 encapsulation IPsec function is outside the scope of the document. Internet Key Exchange Protocol Version 2 (IKEv2) [RFC7296] and Group Security Association (GSA) [RFC5374] can be referred to for further studying.

6. IANA Considerations

6.1. BIER Option Type

Allocation is expected from IANA for a BIER Option Type codepoint from the "Destination Options and Hop-by-Hop Options" sub-registry of the "Internet Protocol Version 6 (IPv6) Parameters" registry. The value 0x70 is suggested.

Hex Value	act	chg	rest	Description	Reference
0x70	01	1	10000	BIER Option	This draft

6.2. End.BIER Function

Allocation is expected from IANA for an End.BIER function codepoint from the "SRv6 Endpoint Behaviors" sub-registry. The value 60 is suggested.

Value	Hex	Endpoint function	Reference
TBD	TBD	End.BIER	This draft

7. Acknowledgements

The authors would like to thank Stig Venaas for his valuable comments. Thanks IJsbrand Wijnands, Greg Shepherd, Tony Przygienda, Toerless Eckert, Jeffrey Zhang, Pascal Thubert for the helpful comments to improve this document.

Thanks Aijun Wang for comments about BIER OAM function in BIER IPv6 encapsulation.

Thanks Mach Chen for review and suggestions about BIER-PING function in BIER IPv6 encapsulation.

8. Contributors

Gang Yan

Huawei Technologies

China

Email: yangang@huawei.com

Yang(Yolanda) Xia

Huawei Technologies

China

Email: yolanda.xia@huawei.com

9. References

9.1. 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>.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI 10.17487/RFC4301, December 2005, <https://www.rfc-editor.org/info/rfc4301>.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302, DOI 10.17487/RFC4302, December 2005, <https://www.rfc-editor.org/info/rfc4302>.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, DOI 10.17487/RFC4303, December 2005, <https://www.rfc-editor.org/info/rfc4303>.

[RFC5374] Weis, B., Gross, G., and D. Ignjatic, "Multicast Extensions to the Security Architecture for the Internet Protocol", RFC 5374, DOI 10.17487/RFC5374, November 2008, <https://www.rfc-editor.org/info/rfc5374>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <https://www.rfc-editor.org/info/rfc6275>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <https://www.rfc-editor.org/info/rfc7296>.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[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>.

[RFC8401] Ginsberg, L., Ed., Przygienda, T., Aldrin, S., and Z. Zhang, "Bit Index Explicit Replication (BIER) Support via IS-IS", RFC 8401, DOI 10.17487/RFC8401, June 2018, <https://www.rfc-editor.org/info/rfc8401>.

[RFC8556] Rosen, E., Ed., Sivakumar, M., Przygienda, T., Aldrin, S., and A. Dolganow, "Multicast VPN Using Bit Index Explicit Replication (BIER)", RFC 8556, DOI 10.17487/RFC8556, April 2019, <https://www.rfc-editor.org/info/rfc8556>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

9.2. Informative References

[I-D.geng-bier-bierv6-yang]
Geng, X., Qin, Z., and F. Zheng, "YANG Data Model for Bierv6", draft-geng-bier-bierv6-yang-00 (work in progress), June 2020.

[I-D.geng-bier-ipv6-inter-domain]
Geng, L., Xie, J., McBride, M., and G. Yan, "Inter-Domain Multicast Deployment using BIERv6", draft-geng-bier-ipv6-inter-domain-01 (work in progress), January 2020.

[I-D.ietf-bier-ipv6-requirements]
McBride, M., Xie, J., Dhanaraj, S., Asati, R., Zhu, Y., and G. Mishra, "BIER IPv6 Requirements", draft-ietf-bier-ipv6-requirements-05 (work in progress), July 2020.

[I-D.ietf-bier-ping]
Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-16 (work in progress), June 2020.

[I-D.xie-bier-ipv6-isis-extension]
Xie, J., Wang, A., Yan, G., and S. Dhanaraj, "BIER IPv6 Encapsulation (BIERv6) Support via IS-IS", draft-xie-bier-ipv6-isis-extension-01 (work in progress), January 2020.

[I-D.xie-bier-ipv6-mvpn]
Xie, J., McBride, M., Dhanaraj, S., and L. Geng, "Use of BIER IPv6 Encapsulation (BIERv6) for Multicast VPN in IPv6 networks", draft-xie-bier-ipv6-mvpn-02 (work in progress), January 2020.

Appendix A. Relationship to BIER Core Standards

The BIER architecture [RFC8279] is inherited in this BIERv6 proposal, and the layering mode of BIER architecture is fully supported with some necessary extension to the data plane as well as the control plane standards.

The focus of this document is BIERv6 data plane, including the BIERv6 encapsulation and packet forwarding procedures. The common BIER header encoding [RFC8296] is maximum reused in this BIERv6 proposal.

To better understand the overall BIER IPv6 problem space and requirements, refer to [I-D.ietf-bier-ipv6-requirements].

Appendix B. Extensions to BIER Control-plane Standards

The relevant control-plane documents that have done or still to be done are listed below.

- Based on [RFC8401], IS-IS extension is defined in [I-D.xie-bier-ipv6-isis-extension] for intra-AS BIERv6 information advertisement and BIRT/BIFT building.
- OSPFv3 extension for intra-AS BIERv6 information advertisement and BIRT/BIFT building is to be defined.
- Based on this BIERv6 encapsulation, [I-D.geng-bier-ipv6-inter-domain] illustrates how inter-AS BIRT/ BIFT are built and how inter-AS multicast deployment is supported.
- BGP extension for inter-AS BIERv6 information advertisement and BIRT/BIFT building is to be defined.
- Based on [RFC8556], BGP-MVPN using BIERv6 encapsulation is defined in [I-D.xie-bier-ipv6-mvpn] for multicast service deployment.

Appendix C. Considerations of Using Unicast Address

BIER is generally a hop-by-hop and one-to-many architecture, and thus the IPv6 Destination Address (DA) being a Multicast Address is a way one may think of as an approach for both the two paradigms in BIERv6 encapsulation.

However using a unicast address has the following benefits:

1. Replicating a BIERv6 packet over a non-BIER capable router.
2. Fast rerouting a BIERv6 packet using a unicast by-pass tunnel.
3. Forwarding a BIERv6 packet to one of the many BFR neighbors connected on a LAN without imposing new requirements of snooping on switches.
4. Replicating a BIERv6 packet through an anonymous system(AS) to BFERs in other ASes, as illustrated in [I-D.geng-bier-ipv6-inter-domain].

Some of the above scenarios are assumed part of BIER architecture as described in [RFC8279], and some of them are the scalability aspects for inter-AS stateless multicast this document intends to support. This document intends to fulfil all these requirements (categorized

as multi-hop replication), and proposes to use unicast address for both one-hop replication and multi-hop replication.

Authors' Addresses

Jingrong Xie
Huawei Technologies

Email: xiejingrong@huawei.com

Liang Geng
China Mobile
Beijing 10053

Email: gengliang@chinamobile.com

Mike McBride
Futurewei

Email: mmcbride7@gmail.com

Rajiv Asati
Cisco

Email: rajiva@cisco.com

Senthil Dhanaraj
Huawei

Email: senthil.dhanaraj@huawei.com

Yongqing Zhu
China Telecom

Email: zhuyq8@chinatelecom.cn

Zhuangzhuang Qin
China Unicom

Email: qinzhuangzhuang@chinaunicom.cn

MooChang Shin
LG Uplus

Email: himzzang@lguplus.co.kr

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Xuesong Geng
Huawei

Email: gengxuesong@huawei.com

BIER Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 14, 2021

X. Min
Z. Zhang
ZTE Corp.
Y. Liu
China Mobile
N. Nainar
C. Pignataro
Cisco Systems, Inc.

July 13, 2020

Bit Index Explicit Replication (BIER) Encapsulation for In-situ OAM (IOAM) Data
draft-xzlnp-bier-ioam-00

Abstract

In-situ Operations, Administration, and Maintenance (IOAM) collects operational and telemetry information while the packet traverses a particular network domain. Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. The BIER header contains a bit-string in which each bit represents exactly one egress router to forward the packet to. This document outlines the requirements to carry IOAM data in BIER header and specifies how IOAM data fields are encapsulated in BIER header.

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 14, 2021.

Table of Contents

1. Introduction

In-situ Operations, Administration, and Maintenance (IOAM) collects operational and telemetry information while the packet traverses a particular network domain. [I-D.ietf-ippm-ioam-data] defines different IOAM data fields used to record various telemetry data from the transit nodes. The term "in-situ" refers to the fact that the IOAM data fields are added to the data packets rather than being sent within packets specifically dedicated to OAM.

Bit Index Explicit Replication (BIER), as defined in [RFC8279], is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building

protocol. The BIER header, as defined in [RFC8296], contains a bit-string in which each bit represents exactly one egress router to forward the packet to.

This document outlines the requirements to carry IOAM data in BIER header and specifies how IOAM data fields are encapsulated in BIER header.

2. Conventions Used in This Document

2.1. 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.

2.2. Abbreviations

Abbreviations used in this document:

BFER: Bit Forwarding Egress Router

BFIR: Bit Forwarding Ingress Router

BIER: Bit Index Explicit Replication

GRE: Generic Routing Encapsulation

IOAM: In-situ Operations, Administration, and Maintenance

OAM: Operations, Administration, and Maintenance

3. Requirements to carry IOAM data in BIER header

[I-D.ietf-bier-use-cases] lists many use cases for BIER. There are many multicast flows in one network. Some of the flows are sensitive for packet loss, delay and other factors, such as live video, real-time meeting. The network administrator wants to know the real-time statistics for these flows, such as delay, sequence, the I/O interface, and the usage of buffer, and so on.

So a method need to be used for measuring the packet real-time transportation guarantee. OAM function defined in [I-D.ietf-bier-pmmm-oam] can be used for packet loss and delay detection. This document attempts to provide a way to achieve on-path telemetry information collection through in-situ OAM.

4. IOAM data fields encapsulation in BIER header

The BIER header is defined in [RFC8279]. The BIER OAM header that follows BIER header is defined in [I-D.ietf-bier-ping]. IOAM-Data-Fields can either be carried in BIER using a new type of OAM message which follows the BIER OAM header (referred to as option 1), or be carried in BIER using a new next protocol header which immediately follows the BIER header (referred to as option 2). In this document, option 2 is selected and the reason is discussed in Section 5.1. An IOAM header is added containing the different IOAM-Data-Fields defined in [I-D.ietf-ippm-ioam-data].

In an administrative domain where IOAM is used, insertion of the IOAM header in BIER is enabled at the BFIRs, which also serve as IOAM encapsulating nodes by means of configuration, deletion of the IOAM header in BIER is enabled at the BFERs, which also serve as IOAM decapsulating nodes by means of configuration.

The Encapsulation format for IOAM over BIER is defined as follows:

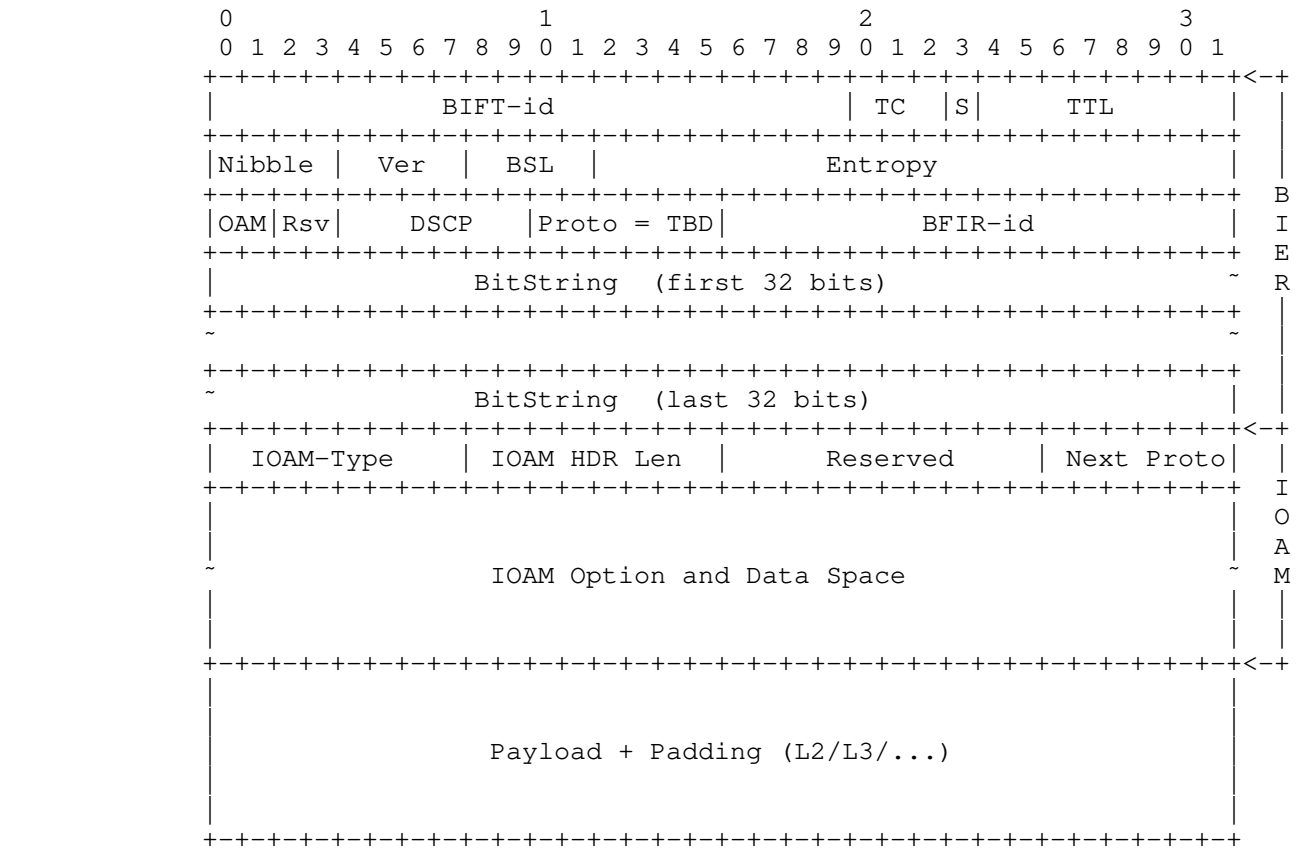


Figure 1: IOAM Encapsulation Format within BIER

The BIER header and fields are defined in [RFC8296]. Within the BIER header, a 6-bit field as "Proto" (Next Protocol) is used to identify the type of the payload immediately following the BIER header, The "Proto" value is set to TBD when the IOAM header is present.

The IOAM related fields in BIER are defined as follows:

IOAM-Type: 8-bit field defining the IOAM Option Type, as defined in Section 7.2 of [I-D.ietf-ippm-ioam-data].

IOAM HDR Len: 8-bit unsigned integer. Length of the IOAM header in 4-octet units.

Reserved: 10-bit reserved field MUST be set to zero upon transmission and ignored upon receipt.

Next Proto: 6-bit unsigned integer that identifies the type of payload immediately following this IOAM option. The semantics of this field are identical to the "Proto" field in [RFC8296].

IOAM Option and Data Space: IOAM option header and data is present as specified by the IOAM-Type field, and is defined in Section 4 of [I-D.ietf-ippm-ioam-data].

Multiple IOAM-Option-Types MAY be included within the BIER encapsulation. For example, if a BIER encapsulation contains two IOAM-Option-Types preceding a data payload, the Next Proto field of the first IOAM option will contain the value of TBD, while the Next Proto field of the second IOAM option will contain the "BIER Next Protocol" number indicating the type of the data payload. Each IOAM Option-Type MUST occur at most once within the same BIER encapsulation header.

5. Considerations

This section summarizes a set of considerations on the overall approach taken for IOAM data encapsulation in BIER, as well as deployment considerations.

5.1. Discussion of the encapsulation approach

Both the options described in section 4 are supposed to be feasible, nevertheless this document needs to select one as standardized encapsulation for IOAM over BIER. Considering the fact that the encapsulation format option 2 using a new next protocol header is more concise than option 1 using a new type of OAM message, and many other transport protocols, e.g. GRE, use a new next protocol header to encapsulate IOAM data, the encapsulation format option 2 is selected as the standardized one.

5.2. IOAM and the use of the BIER OAM bits

[RFC8296] defines a two-bits long field, referred to as OAM. [I-D.ietf-bier-pmmm-oam] describes how to use the two-bits OAM field for alternate marking performance measurement method. The BIER IOAM header and the BIER two-bits OAM field are orthogonal and can co-exist in the same packet header, i.e. a BIER packet with IOAM data can set the OAM field or not, and a BIER packet with OAM field set can also carry IOAM data or not.

6. Security Considerations

This document does not raise any additional security issues beyond those of the specifications referred to in the list of normative references.

7. IANA Considerations

In the "BIER Next Protocol Identifiers" registry defined in [RFC8296], a new Next Protocol Value for IOAM is requested from IANA as follows:

BIER Next Protocol Identifier	Description	Semantics Definition	Reference
TBD	In-situ OAM (IOAM)	Section 4	This Document

Table 1: New BIER Next Protocol Identifier for IOAM

8. Acknowledgements

The authors would like to acknowledge Greg Mirsky for his thorough review and very helpful comments.

9. References

9.1. Normative References

[I-D.ietf-bier-ping]
Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., remy@barefootnetworks.com, r., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-09 (work in progress), March 2020.

[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>.

[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>.

9.2. Informative References

[I-D.ietf-bier-pmmm-oam] Mirsky, G., Zheng, L., Chen, M., and G. Fioccola, "Performance Measurement (PM) with Marking Method in Bit Index Explicit Replication (BIER) Layer", draft-ietf-bier-pmmm-oam-08 (work in progress), May 2020.

[I-D.ietf-bier-use-cases] Nainar, N., Asati, R., Chen, M., Xu, X., Dolganow, A., Przygienda, T., Gulko, A., Robinson, D., Arya, V., and C. Bestler, "BIER Use Cases", draft-ietf-bier-use-cases-11 (work in progress), March 2020.

Authors' Addresses

Xiao Min
ZTE Corp.
Nanjing
China

Email: xiao.min2@zte.com.cn

Zheng(Sandy) Zhang
ZTE Corp.
Nanjing
China

Email: zhang.zheng@zte.com.cn

Yisong Liu
China Mobile
Beijing
China

Email: liuyisong@chinamobile.com

Nagendra Kumar Nainar
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States

Email: naikumar@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States

Email: cpignata@cisco.com

BIER WG
Internet-Draft
Intended status: Informational
Expires: January 14, 2021

Z. Zhang
G. Mirsky
Q. Xiong
ZTE Corporation
Y. Liu
China Mobile

July 13, 2020

BIER Source Protection
draft-zhang-bier-source-protection-01

Abstract

This document describes the multicast source protection functions in Bit Index Explicit Replication BIER domain.

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 14, 2021.

Copyright Notice

Table of Contents

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides 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).

Source Protection is not specific to the BIER environment. Source Protection means that to protect the multicast source node, two or more ingress routers, in BIER environment, ingress routers are BFIRs, may be used to connect the multicast source node. Generally, only one ingress router (BFIR) is selected to forward flows from multicast source node to egress routers, in BIER environment, egress routers are BFERs. The selected BFIR may be chosen based on local policies, BFERs that receiving the same multicast flow may elect to use the same or different BFIR. The BFIR and the path in use are referred to as working while all alternative available BFIRs and paths that can be used to receive the same multicast flow are referred to as protection.

For a BFER, when either the working BFIR or the working path fails, the BFER can select one of protecting BFIRs to get the multicast flow. The shorter the detection time is, the faster the flow recovers.

This document discusses the functions that can be used in failure detection for multicast source protection.

2. The Source Protection analysis

According to BIER architecture [RFC8279], BIER overlay protocols, which include MVPN [RFC8556], MLD [I-D.ietf-bier-mld], PIM [I-D.ietf-bier-pim-signaling], etc., are used to exchange the multicast flow information, so the BFER selects the UMH (Upstream Multicast Hop) BFIR as the ingress router, the BFIR collects the BFERs which want to receive the multicast flow. BIER transport is used to deliver the multicast packet to the destination BFERs. To ensure that the source flow protection is uninterrupted, the detection of a defect is at the BIER transport layer. The switchover is performed at the BIER overlay layer. When BFIR failure is detected, BIER overlay advertisement for BFIR re-select may be triggered.

As [I-D.ietf-bess-mvpn-fast-failover] describes, the root standby functions defined in section 4.2 can also be used in the BIER environment. About Cold Standby, Warm Standby, Hot Standby, more details will be added in future version.

The most important items in source detection are failure detection and switchover. Note that the failure detection includes node and working path failure monitoring. Similarly, BFIR switching and BFER switching are included in the switchover scenario.

The selected BFIR is referred to as Selected BFIR (S-BFIR) and the backup BFIR - as Backup BFIR (B-BFIR). For simplicity, only one B-BFIR is considered in the following case.

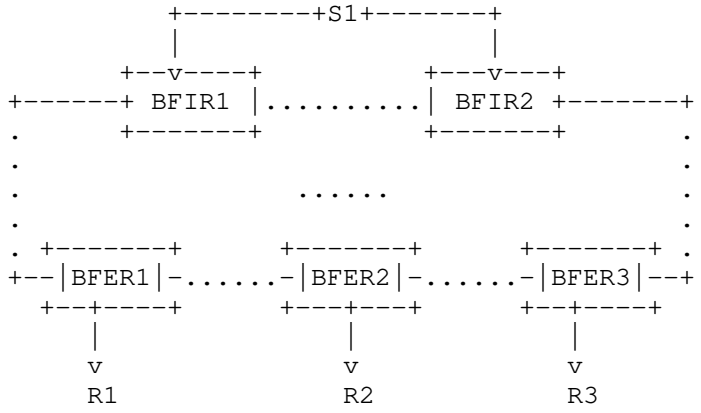


Figure 1: An Example of the Source Protection in BIER

In Figure 1, a multicast source S1 is connected to BFIR1 and BFIR2. In some deployments, only BFIR1 advertises S1 flow information to

BFERs by BIER overlay protocols, such as BGP(MVPN), MLD, PIM, etc. All the BFERs which want to receive the S1 flow will select BFIR1 as the S-BFIR, BFIR2 is the B-BFIR. In some other deployments, BFIR1 and BFIR2 both advertise S1 flows to BFERs by BIER overlay protocols, and some BFERs may select BFIR1 as S-BFIR, other BFERs may select BFIR2 as S-BFIR, BFIR1 and BFIR2 in charge of different BFERs, and they are complementary B-BFIR for the BFERs. We do not distinguish these two cases strictly.

2.1. Node failure monitoring

For example, if S1 connects BFIR1 and BFIR2 by a shared media, BFIR1 is the selected BFIR for multicast flow forwarding that comes from S1, BFIR2 can monitor BFIR1 node failing by BFD session [RFC5880] built on the shared media. Also, ping methods, include IPv4 ping [RFC0792] (when IPv4 prefix is used), LSP-Ping [RFC8029] (when mpls forwarding plane is built), IPv6 ping [RFC4443] (when IPv6 prefix is used), BIER ping [I-D.ietf-bier-ping] can also be used. In case there is no shared media among S1, BFIR1 and BFIR2, BFIR2 may monitor BFIR1 by BFD session or any type of ping methods across the BIER domain, in case there is no direct connection between BFIR1 and BFIR2, multiple hops will be traversed.

2.2. Monitoring of the Working Path for a Failure

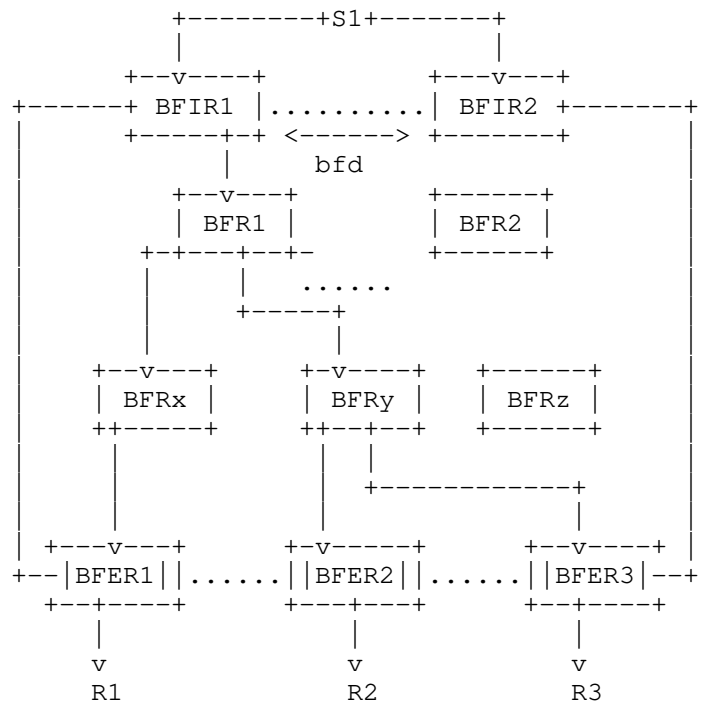


Figure 2: An Example of the Monitoring of the Working Path

In the case of a node failure detection, the path between B-BFIR and S-BFIR may not be the same as the path traversed by the data flow. For example, in Figure 2, the path from BFIR1 (S-BFIR) to all the BFERs is different from the path between BFIR1 and BFIR2 (B-BFIR). If the path between BFIR2 and BFIR1 is broken, BFIR2 will detect the failure and consider that BFIR1 is down. As a result, BFIR2 will take on the role of S-BFIR. But the path from BFIR1 to all of the BFERs may be working well and is not affected by the defect between BFIR1 and BFIR2. In this situation, the B-BFIR switches to S-BFIR unnecessarily, and potential packet loss will occur.

There are two options to monitor the working multicast distribution tree in BIER:

- S-BFIR monitors all the BFERs;

- each BFER monitors the S-BFIR.

In BIER transport environment, the monitor should be based on BIER forwarding.

When S-BFIR monitors all the related BFERs, multiple BFD sessions may be built between S-BFIR and each BFER. BIER ping [I-D.ietf-bier-ping] can also be used. Once S-BFIR finds that one BFER is lost by BFD session timeout or ping fail, S-BFIR should notify B-BFIR to take charge of flow forwarding for the lost BFER.

When BFER monitors the S-BFIR, multiple BFD sessions may be built between S-BFIR and each BFER. Or BIER ping [I-D.ietf-bier-ping] can also be used. Once BFER finds that the S-BFIR is lost, the BFER should select the B-BFIR as S-BFIR as soon as possible, and the BFER should advertisement the UMH selection to B-BFIR as soon as possible.

When MVPN is used as BIER overlay protocol, BFD discriminator defined section 3.1.6 in [I-D.ietf-bess-mvpn-fast-failover] can also be used to build the multipoint BFD among BFIR and BFERs.

BIER BFD [I-D.hu-bier-bfd] can be used to reduce the number of BFD sessions between S-BFIR and each of BFERs. If BIER BFD is used, only one multipoint BFD session will be built among S-BFIR and all the BFERs.

3. BFD and Ping

BFD and Ping can all be used in failure detection, but there are differences between them. A network administrator can select the appropriate mechanism according to the actual network.

3.1. BIER Ping

[I-D.ietf-bier-ping] describes the mechanism and basic BIER OAM packet format that can be used to perform failure detection and isolation on BIER data plane without any dependency on other layers like the IP layer.

In the example of Figure 1, BFER can monitor the status of BFIR and the path status between BFER and BFIR. BFER1 sends the BIER Ping packet to BFIR1. If BFER1 does not receive responses from BFIR1 in an expected period of time (may be multiplied average round-trip time), BFER1 will treat BFIR1 as a failed UMH, and BFER1 will select BFIR2 as the UMH and signal to BFIR2 to get multicast flow.

In this example, BFER1, BFER2, and BFER3 send BIER ping packet to BFIR1 separately. The timeout period MAY be set to different values depending on the local performance requirement on each BFER.

In the general case of more complex BIER topology, it cannot be guaranteed that the path used from BFIR1 to BFER1 is the same as in the reverse direction, i.e., from BFER1 to BFIR1. If that is not

guaranteed and the paths are not co-routed, then this method may produce false results, both false negative and false positive. The former is when ping fails while the multicast path and flow are OK. The latter is when the multicast path has a defect, but ping works. Thus, to improve the consistency of this method of detecting a failure in multicast flow transport, the path that the echo request from BFER1 traverses to BFIR1 must be co-routed with the path that the monitored multicast flow traverses through the BIER domain from BFIR1 to BFER1.

3.2. BIER BFD

[I-D.hu-bier-bfd] describes the application of P2MP BFD in a BIER network. And it describes the procedures for using such mode of BFD protocol to verify multipoint or multicast connectivity between a sender (BFIR) and one or more receivers (BFERs).

In the same example, BFIR1 sends the BIER Echo request packet to BFERs to bootstrap a p2mp BFD session. After BFER1, BFER2 and BFER3 receive the Echo request packet with BFD Discriminator and the Target SI-Bitstring TLVs, BFERs creates the BFD session of type MultipointTail [RFC8562] to monitor the status of BFIR1 and the working path. If BFERs have not received BFD packet from BFER1 for the Detection Time [RFC8562], BFER1 will treat BFIR1 as a failed UMH, and signal to BFIR2 to get the multicast flow.

The timeout period on each BFER MAY be set to a different value depending on the local performance requirement on each BFER. BFER monitors BFIR separately and selects its UMH independently from selections reached by other BFERs.

4. Security Considerations

Security considerations discussed in [RFC8279], [RFC8562], [I-D.ietf-bier-ping], [I-D.ietf-bess-mvpn-fast-failover] and [I-D.hu-bier-bfd] apply to this document.

5. Informative References

[I-D.hu-bier-bfd]
Xiong, Q., Mirsky, G., hu, f., and C. Liu, "BIER BFD", draft-hu-bier-bfd-07 (work in progress), July 2020.

[I-D.ietf-bess-mvpn-fast-failover]
Morin, T., Kebler, R., and G. Mirsky, "Multicast VPN fast upstream failover", draft-ietf-bess-mvpn-fast-failover-10 (work in progress), February 2020.

[I-D.ietf-bier-mld]
Pfister, P., Wijnands, I., Venaas, S., Wang, C., Zhang, Z., and M. Stenberg, "BIER Ingress Multicast Flow Overlay using Multicast Listener Discovery Protocols", draft-ietf-bier-mld-04 (work in progress), March 2020.

[I-D.ietf-bier-pim-signaling]
Bidgoli, H., Xu, F., Kotalwar, J., Wijnands, I., mishra, m., and Z. Zhang, "PIM Signaling Through BIER Core", draft-ietf-bier-pim-signaling-09 (work in progress), July 2020.

[I-D.ietf-bier-ping]
Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

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

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[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>.

[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>.

[RFC8556] Rosen, E., Ed., Sivakumar, M., Przygienda, T., Aldrin, S., and A. Dolganow, "Multicast VPN Using Bit Index Explicit Replication (BIER)", RFC 8556, DOI 10.17487/RFC8556, April 2019, <https://www.rfc-editor.org/info/rfc8556>.

[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>.

Authors' Addresses

Zheng Zhang
ZTE Corporation

Email: zzhang_ietf@hotmail.com

Greg Mirsky
ZTE Corporation

Email: gregimirsky@gmail.com

Quan Xiong
ZTE Corporation

Email: xiong.quan@zte.com.cn

Yisong Liu
China Mobile

Email: liuyisong@chinamobile.com