

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 6, 2021

H. Chen
Futurewei
A. Wang
China Telecom
G. Mishra
Verizon Inc.
Y. Fan
Casa Systems
L. Liu
Fujitsu
X. Liu
Volta Networks
November 2, 2020

BIER Egress Protection
draft-chen-bier-egress-protect-00

Abstract

This document describes a mechanism for fast protection against the failure of an egress node of a "Bit Index Explicit Replication" (BIER) domain. It does not have any per-flow state in the core of the domain. For a multicast packet to an egress node of the domain, when the egress node fails, its upstream hop as a PLR sends the packet to the egress' backup node once the PLR detects the failure.

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] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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 6, 2021.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

- [1.](#) Introduction [2](#)
- [1.1.](#) Terminology [3](#)
- [2.](#) Overview of BIER Egress Protection [4](#)
- [3.](#) Protocol Extensions [5](#)
- [3.1.](#) Extensions to OSPF [5](#)
- [3.2.](#) Extensions to IS-IS [6](#)
- [4.](#) BIER Extensions [7](#)
- [4.1.](#) Egress Protection Bit Index Routing Tables [7](#)
- [4.2.](#) Egress Protection Bit Index Forwarding Tables [8](#)
- [4.3.](#) Updated Forwarding Procedure [8](#)
- [4.4.](#) Switching between EP and Normal Forwarding [9](#)
- [5.](#) Example Application of BIER Egress Protection [10](#)
- [5.1.](#) Example BIER Topology [10](#)
- [5.2.](#) BIRT and BIFT on a BFR [11](#)
- [5.3.](#) EP-BIRTs and EP-BIFTs on a BFR [12](#)
- [5.4.](#) Forwarding using EP-BIFT [14](#)
- [6.](#) Security Considerations [16](#)
- [7.](#) IANA Considerations [16](#)
- [8.](#) Acknowledgements [16](#)
- [9.](#) References [16](#)
- [9.1.](#) Normative References [16](#)
- [9.2.](#) Informative References [17](#)
- Authors' Addresses [18](#)

1. Introduction

[RFC8279] specifies "Bit Index Explicit Replication" (BIER). It provides optimal forwarding of multicast data packets through a "multicast/BIER domain". It does not require the use of a protocol

for explicitly building multicast distribution trees, and it does not require intermediate nodes to maintain any per-flow state.

This document describes a mechanism for fast protection against the failure of an egress node of a "Bit Index Explicit Replication" (BIER) domain, which is called BIER Egress Protection.

This BIER Egress Protection does not require intermediate nodes to maintain any per-flow state for fast protection against the failure of an egress node of the flow.

1.1. Terminology

BFR: Bit-Forwarding Router.

BFIR: Bit-Forwarding Ingress Router.

BFER: Bit-Forwarding Egress Router.

BFR-id: BFR Identifier. It is a number in the range [1,65535].

BFR-NBR: BFR Neighbor.

F-BM: Forwarding Bit Mask.

BFR-prefix: An IP address (either IPv4 or IPv6) of a BFR.

BIRT: Bit Index Routing Table. It is a table that maps from the BFR-id (in a particular sub-domain) of a BFER to the BFR-prefix of that BFER, and to the BFR-NBR on the path to that BFER.

BIFT: Bit Index Forwarding Table.

FRR: Fast Re-Route.

PLR: Point of Local Repair.

LFA: Loop-Free Alternate.

RLFA: Remote LFA.

DLFA: Remote LFA with Directed forwarding.

IGP: Interior Gateway Protocol.

LSDB: Link State DataBase.

SPF: Shortest Path First.

SPT: Shortest Path Tree.

OSPF: Open Shortest Path First.

IS-IS: Intermediate System to Intermediate System.

LSA: Link State Advertisement in OSPF.

LSP: Link State Protocol Data Unit (PDU) in IS-IS.

SPT-old(R): The SPT rooted at node R using LSDB before X fails (i.e., old LSDB).

SPT-new(R, X): The SPT rooted at node R using LSDB without X after X fails (i.e., new LSDB).

P-Space P(R,X): The set of nodes that are reachable from R without going through X. In other words, it is the set of nodes that are not downstream of X in SPT-old(R).

Extended P-Space P'(R,X): The set of nodes that are reachable from R or a neighbor of R, without going through X.

Q-Space Q(D,X): The set of nodes that do not use X to reach destination D using the old LSDB.

PQ node(R,X): A member of both the P-Space P(R, X) (or the extended P-Space P'(R, X)) and the Q-Space (D, X).

2. Overview of BIER Egress Protection

For fast protecting an egress node of a BIER domain, a backup egress node is configured on the egress node. After the configuration, the egress node distributes the information about the backup egress to its direct neighbors.

For clearly distinguishing between an egress node and a backup egress node, an egress node is called a primary egress node sometimes.

For a multicast packet to a primary egress node of the domain, when the primary egress node fails, its upstream hop as a point of local repair (PLR) sends the packet to the backup egress node configured to protect the primary egress node once the PLR detects the failure. The upstream hop of the primary egress node is its direct neighbor.

A Bit-Forwarding Router (BFR) in a BIER sub-domain builds and maintains an "Egress Protection Bit Index Routing Table" (EP-BIRT) for each of its BFR Neighbors (BFR-NBRs) that are egress nodes of the

domain to provide fast protection against the failure of an egress node. The BFR builds each EP-BIRT based on a BIRT defined in [RFC8279]. An "Egress Protection Bit Index Forwarding Table" (EP-BIFT) is derived from an EP-BIRT in a way that is similar to the way in which a BIFT is derived from a BIRT, which is defined in [RFC8279].

Once the BFR as a PLR detects the failure of its BFR-NBR X that is a primary egress node of the domain, for a multicast packet targeting to the primary egress node, the PLR uses the EP-BIFT for X to send the packet to the backup egress node configured to protect the primary egress node.

3. Protocol Extensions

This section defines extensions to OSPF and IS-IS for advertising the backup information (including the backup egress node for protecting a primary egress node) to its direct neighbors.

3.1. Extensions to OSPF

When a node P (as a primary egress node) has a backup egress node configured to protect against its failure, node P advertises the information about the backup egress node to its neighbors in its router information opaque LSA of LS type 9. The information is included in a backup egress node TLV. The format of the TLV is shown in Figure 1.

After each of the neighbors receives the backup egress node TLV from node P, it knows that node P as a primary egress node will be protected by the backup egress node in the TLV. Once detecting the failure of node P, it sends the packet targeting to node P towards the backup egress node.

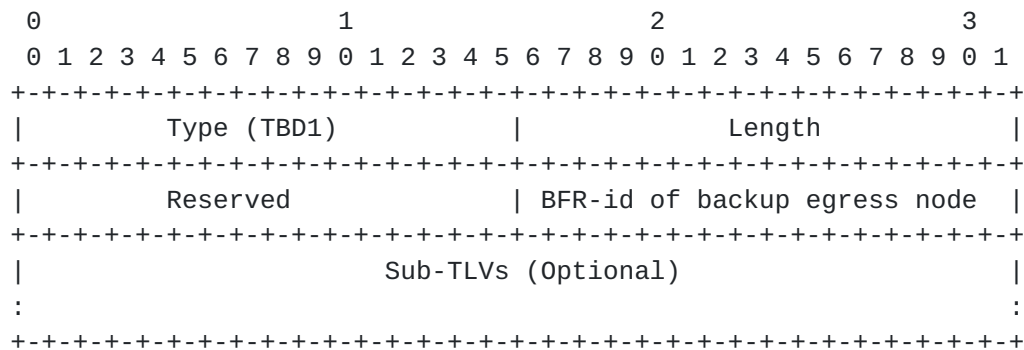


Figure 1: OSPF Backup Egress TLV

Type: 2 octets, its value (TBD1) is to be assigned by IANA.

Length: 2 octets, its value is 4 plus the length of the Sub-TLVs included. If no Sub-TLV is included, its value is 4.

Reserved: 2 octets, it MUST be set to zero when sending and be ignored while receiving.

BFR-id of backup egress node: 2 octets, its value is the BFR-id of the backup egress node configured to protect against the failure of the primary egress node (i.e., the node advertising this TLV).

Sub-TLVs (Optional): No Sub-TLV is defined now.

3.2. Extensions to IS-IS

For supporting fast protection against the failure of a primary egress node in a BIER domain, a new IS-IS TLV, called IS-IS backup egress node TLV, is defined. It contains the BFR-id of a backup egress node.

When a node P (as a primary egress node) has a backup egress node configured to protect against its failure, node P advertises the information about the backup egress node to its neighbors using a IS-IS backup egress node TLV.

This TLV may be advertised in IS-IS Hello (IIH) PDUs, LSPs, or in Circuit Scoped Link State PDUs (CS-LSP) [RFC7356]. The format of the TLV is shown in Figure 2.

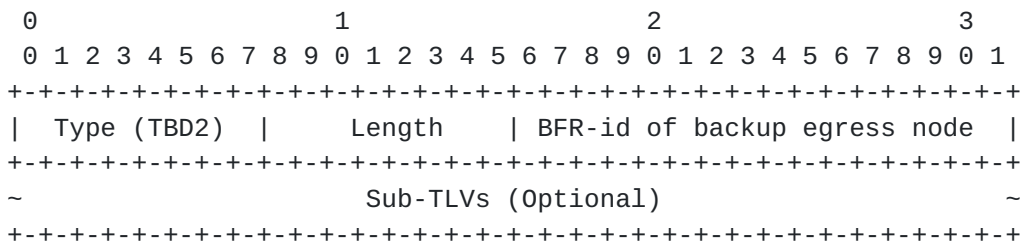


Figure 2: IS-IS Backup Egress TLV

Type: 1 octet, its value (TBD2) is to be assigned by IANA.

Length: 1 octet, its value is 2 plus the length of the Sub-TLVs included. If no Sub-TLV is included, its value is 2.

BFR-id of backup egress node: 2 octets, its value is the BFR-id of the backup egress node configured to protect against the failure of the primary egress node (i.e., the node advertising this TLV).

Sub-TLVs (Optional): No Sub-TLV is defined now.

4. BIER Extensions

4.1. Egress Protection Bit Index Routing Tables

If a BFR is a direct neighbor of an egress node in a BIER sub-domain, it builds and maintains a number of "Egress Protection Bit Index Routing Tables" (EP-BIRTs). There is an EP-BIRT for each of the BFR's neighbors that are egress nodes of the domain. The BFR builds each EP-BIRT based on its BIRT. Comparing to the BIRT, an EP-BIRT has a piece of new backup information for each BFER.

The new backup information for a BFER indicates if the BFER as an egress node is protected by the BFR. If so, the information further includes the backup egress node configured to protect the BFER.

In one implementation, the new backup information is represented by {EP, BE-BFER}. EP (short for Egress Protection) is a flag, indicating whether the BFER as an egress node is protected. EP = 1 means that the BFER is protected. EP = 0 means that the BFER is not protected. BE-BFER (short for Backup Egress BFER) is the BFER (i.e., BFER-id) of the backup egress node when EP = 1. BE-BFER is NULL (0) when EP = 0.

In the EP-BIRT for BFR-NBR X that is an egress node, the row having X as BFER and as its next hop BFR-NBR contains the new backup information {EP = 1, BE-BFER}, where BE-BFER is the BFER (i.e., BFER-id) of backup egress node for protecting the egress node. Each of the other rows in the EP-BIRT contains the new backup information {EP = 0, BE-BFER = NULL}.

When the egress node fails, for a multicast packet targeting to the primary egress node BFER (PE-BFER), the BFR sends the packet to the BE-BFER through using the route to the backup egress node. The BFR clears the bit for PE-BFER and adds the bit for BE-BFER in the packet's BitString first, and then forwards the packet according to the forwarding entry for BE-BFER.

The EP-BIRT for BFR-NBR X that is an egress node considers the failure of X. It has a route or say a next hop (i.e., BFR-NBR N on the path, where N is not X) to every BFER except for X.

The BFR may build the EP-BIRT for BFR-NBR X by copying its BIRT to the EP-BIRT and sets the new information for each BFER to empty such as {EP = 0, BE-BFER = NULL} first. And then it updates each of the rows in the EP-BIRT that has X as BFER or next hop BFR-NBR X.

For the BFR-id of a BFER in the EP-BIRT for egress node X, when the next hop BFR-NBR on the path to the BFER is X, the BFR checks whether

the BFER is X. If the BFER is not X, the BFR changes next hop BFR-NBR X to a backup next hop (BNH) when there is a BNH on a backup path to the BFER without going through X and the link from the BFR to X. If the BFER is X, the BFR adds the new backup information {EP = 1, BE-BFER} for the BFER as PE-BFER.

If there is not any BNH to a BFER to protect against the failure of X, the next hop BFR-NBR X to the BFER in the EP-BIRT for BFR-NBR X is changed to NULL. For a multicast packet having the BFER as one of its destinations, if the next hop BFR-NBR to the BFER is NULL, the BFR does not send the packet to the next hop BFR-NBR NULL but drops it when X fails.

Note: In another option, the next hop BFR-NBR X to the BFER in the EP-BIRT for BFR-NBR X keeps unchanged when there is not any BNH to the BFER to protect against the failure of X. In this case, for a multicast packet having the BFER as one of its destinations, the BFR sends the packet to X when X fails.

In one implementation, the BNH is the Loop-Free Node-Protecting Alternate defined in [[RFC5286](#)] to protect against the failure of X and link from the BFR to X. In another implementation, the BNH is the virtual Loop-Free Alternate (LFA), i.e., PQ node, defined in [[RFC7490](#)]. In a special case, a PQ node is a Loop-Free Node-Protecting Alternate defined in [[RFC5286](#)].

[4.2.](#) Egress Protection Bit Index Forwarding Tables

From each EP-BIRT on the BFR, an "Egress Protection Bit Index Forwarding Table" (EP-BIFT) is derived. In addition to having a route to a BFER in each row of the EP-BIFT which is the same as the EP-BIRT, it has a "Forwarding Bit Mask" (F-BM) in its each row. For the rows in the EP-BIRT that have the same SI and the same BFR-NBR and the same new backup information {EP, BE-BFER}, the F-BM for each of these rows in the EP-BIFT is the logical OR of the BitStrings of these rows.

This EP-BIFT is programmed into the data plane and is not used to forward any packet in normal operations. It is activated to forward a packet with a BIER header once the BFR detects the failure of BFR-NBR. The header contains SI, BitString, BitStringLength, and sub-domain.

[4.3.](#) Updated Forwarding Procedure

The forwarding procedure defined in [[RFC8279](#)] is updated/enhanced for an EP-BIFT to consider the egress protection (i.e., the new information {EP, BE-BFER} in the EP-BIFT). For a multicast packet

with the BitString indicating a BFER as one of its destinations, the updated forwarding procedure sends the packet towards the backup egress node of the BFER if the BFER is protected. It checks whether EP = 1 in the forwarding entry for the BFER. If EP = 1, the procedure clears the bit for the BFER as PE-BFER and adds the bit for BE-BFER in the packet's BitString first, and then forwards the packet using the row (i.e., forwarding entry) for BE-BFER.

The updated procedure is described in Figure 3. It is used with an EP-BIFT for BFR-NBR X as egress node on a BFR to forward multicast packets when X fails. It can also be used with a BIFT on the BFR to forward multicast packets in normal operations if the new backup information in each row of the BIFT is empty such as {EP = 0, BE-BFER = NULL}.

```

Packet = the packet received by BFR;
FOR each BFER k (from the rightmost in Packet's BitString) {
  IF BFER k is the BFR itself {
    copies Packet, sends the copy to the multicast
    flow overlay and clears bit k in Packet's BitString
  } ELSE {
    finds the row in the EP-BIFT for the sub-domain using
    Packet's SI and BitString as the key/index
    IF EP == 1 {
      clears bit k in Packet's BitString; //BFER k is PE-BFER
      adds bit j in Packet's BitString; //BFER j is BE-BFER
    } ELSE {
      IF BFR-NBR in the row is not NULL {
        Copies Packet, updates the copy's BitString by ANDing
        it with F-BM in the row, sends updated copy to BFR-NBR
      } // BFR-NBR == NULL, not sent Packet to BFR-NBR
      updates Packet's BitString by ANDing it with the INVERSE
      of the F-BM in the row
    }
  }
}

```

Figure 3: Updated Forwarding Procedure

4.4. Switching between EP and Normal Forwarding

The EP-BIFTs will be pre-computed and installed ready for activation when an egress node failure is detected. Once the BFR detects the failure of its BFR-NBR X as an egress, it activates the EP-BIFT for X to forward packets with BIER headers and de-activates its BIFT. After activation of the EP-BIFT, it remains in effect until it is no longer required.

In general, when the routing protocol has re-converged on the new topology taking into account the failure of X, the BIRT is re-computed using the updated LSDB and the BIFT is re-derived from the BIRT. Once the BIFT is installed ready for activation, it is activated to forward packets with BIER headers and the EP-BIFT for X is de-activated.

From the new topology, the BFR computes/re-computes the EP-BIRT for each BFR-NBR Y as an egress of the BFR and the EP-BIFT for Y is derived/re-derived from the EP-BIRT for Y. The EP-BIFT is installed/re-installed ready for activation when Y fails.

5. Example Application of BIER Egress Protection

This section illustrates an example application of BIER Egress Protection on a BFR in a BIER topology in Figure 4.

5.1. Example BIER Topology

An example BIER topology for a BIER sub-domain is shown in Figure 4. It has 8 nodes/BFRs A, B, C, D, E, F, G and H. Each of the links connecting these nodes/BFRs has a cost. The link cost of 1 is default and is not indicated in the figure. The link costs of other values such as 2 and 3 are indicated in the figure.

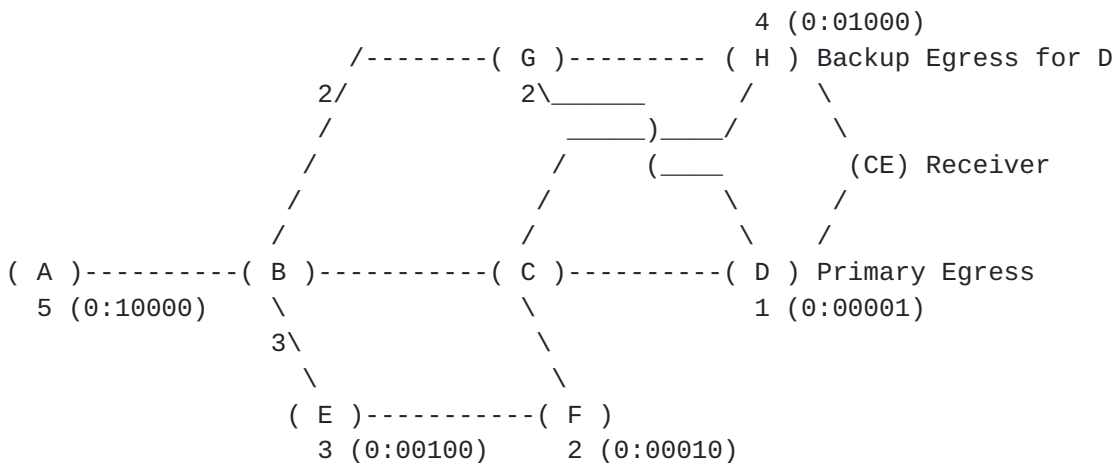


Figure 4: Example BIER Topology

Nodes/BFRs D, F, E, H and A are BFRs and have BFR-ids 1, 2, 3, 4, and 5 respectively. For simplicity, these BFR-ids are represented by (SI:BitString), where SI = 0 and BitString is of 5 bits. BFR-ids 1, 2, 3, 4, and 5 are represented by (0:00001), (0:00010), (0:00100), (0:01000) and (0:10000) respectively.

BFER H is configured to protect BFER D on BFR D. Suppose that this information is distributed to BFR D's neighbors BFR C and BFR G by IGP. BFR C and BFR G know that H is the backup egress to protect the primary egress D.

CE is a multicast traffic Receiver, which is dual homed to primary egress node D and backup egress node H for protecting primary egress D. During normal operations, there is no multicast traffic to CE from backup egress node H and CE receives the multicast traffic only from primary egress node D. There is no duplicated traffic to receiver CE. This is different from MoFRR in [RFC7431], where duplicated traffic is sent to a dual homed receiver. When primary egress node D fails, the multicast traffic is sent to CE from backup egress node H.

5.2. BIRT and BIFT on a BFR

Every BFR in a BIER sub-domain/topology builds and maintains a Bit Index Routing Table (BIRT). For the BIER topology in Figure 4, each of 8 nodes/BFRs A, B, C, D, E, F, G and H builds and maintains a BIRT using the LSDB for the topology.

The BIRT built on BFR C (i.e. node C) is shown in Figure 5.

```

+-----+-----+-----+
|   BFR-id   | BFR-Prefix | BFR-NBR  |
| (SI:BitString) | of Dest BFER | (Next Hop) |
+=====+=====+=====+
|  1 (0:00001) |    D       |    D     |
+-----+-----+-----+
|  2 (0:00010) |    F       |    F     |
+-----+-----+-----+
|  3 (0:00100) |    E       |    F     |
+-----+-----+-----+
|  4 (0:01000) |    H       |    H     |
+-----+-----+-----+
|  5 (0:10000) |    A       |    B     |
+-----+-----+-----+
    
```

Figure 5: Bit Index Routing Table on BFR C

The 1st row in the BIRT indicates that the next hop BFR-NBR on the shortest path to BFER D with BFR-id 1 is BFR D.

The 2nd row in the BIRT indicates that the next hop BFR-NBR on the shortest path to BFER F with BFR-id 2 is BFR F.

The 3rd row in the BIRT indicates that the next hop BFR-NBR on the shortest path to BFER E with BFR-id 3 is BFR F.

The 4-th row in the BIRT indicates that the next hop BFR-NBR on the shortest path to BFER H with BFR-id 4 is BFR H.

The 5-th row in the BIRT indicates that the next hop BFR-NBR on the shortest path to BFER A with BFR-id 5 is BFR B.

From this BIRT on BFR C, a Bit Index Forwarding Table (BIFT) is derived. This BIFT is shown in Figure 6.

The 2nd and 3-th rows in the BIRT have the same SI = 0 and next hop BFR-NBR = F. The F-BM for each of these two rows in the BIFT is the logical OR of the BitStrings of these rows, which is 00110 (00010 OR 00100 = 00110).

The F-BM for 1st row in the BIFT is 00001.

The F-BM for 4-th row in the BIFT is 01000.

The F-BM for 5-th row in the BIFT is 10000.

```

+-----+-----+-----+
|   BFR-id   | F-BM   | BFR-NBR   |
| (SI:BitString) |       | (Next Hop) |
+=====+=====+=====+
|  1 (0:00001) | 00001  |    D      |
+-----+-----+-----+
|  2 (0:00010) | 00110  |    F      |
+-----+-----+-----+
|  3 (0:00100) | 00110  |    F      |
+-----+-----+-----+
|  4 (0:01000) | 01000  |    H      |
+-----+-----+-----+
|  5 (0:10000) | 10000  |    B      |
+-----+-----+-----+
    
```

Figure 6: Bit Index Forwarding Table on BFR C

5.3. EP-BIRTs and EP-BIFTs on a BFR

Each of the BFRs that are neighbors of egress nodes (i.e., BFERs) in a BIER sub-domain/topology builds and maintains a number of Egress Protection Bit Index Routing Tables (EP-BIRTs).

For the BIER topology in Figure 4,

BFR B is the neighbor of BFERs A and E;
 BFR C is the neighbor of BFERs D, F and H;
 BFR E is the neighbor of BFER F;
 BFR F is the neighbor of BFER E;
 BFR G is the neighbor of BFERs D and H.

Each of 5 nodes/BFRs B, C, E, F and G builds and maintains a number of EP-BIRTs using the LSDB for the topology for its every BFR-NBR as an egress node.

For example, BFR C (i.e., node C) in the BIER topology builds and maintains three EP-BIRTs for its three BFR-NBRs (BFERs D, F and H) that are egress nodes respectively.

The EP-BIRT for BFER D built by BFR C based on the BIRT on BFR C (refer to Figure 5) is shown in Figure 7.

The BIRT is copied to the EP-BIRT for BFER D (i.e., the first three columns of the EP-BIRT). The new backup information (i.e., the 4-th column) for every row in the EP-BIRT is initialized to {EP = 0, BE-BFER = 0/NULL}.

```

+-----+-----+-----+-----+
| BFR-id | BFR-Prefix | BFR-NBR | {EP, BE-BFER} |
|(SI:BitString)| of Dest BFER |(Next Hop)| (Backup Info) |
+-----+-----+-----+-----+
| 1 (0:00001) | D | D/NULL | EP=1, BE-BFER=H |
+-----+-----+-----+-----+
| 2 (0:00010) | F | F | EP=0, BE-BFER=0 |
+-----+-----+-----+-----+
| 3 (0:00100) | E | F | EP=0, BE-BFER=0 |
+-----+-----+-----+-----+
| 4 (0:01000) | H | H | EP=0, BE-BFER=0 |
+-----+-----+-----+-----+
| 5 (0:10000) | A | B | EP=0, BE-BFER=0 |
+-----+-----+-----+-----+
    
```

Figure 7: EP-BIRT for BFER D on BFR C

In the EP-BIRT for BFER D, the row that has BFR-NBR == D is the 1st row. This row has the new backup information {EP = 1, BE-BFER = H}, which indicates that BFER D (i.e., primary egress node D) is protected by BFER H (i.e., backup egress node H). Each of the other rows has the new backup information {EP = 0, BE-BFER = 0/NULL}.

The 1st row in the EP-BIRT indicates that the next hop BFR-NBR on the path to BFER D with BFR-id 1 is NULL (changed to NULL from D). There is no backup next hop (BNH) to D when D fails.

The 2nd row in the EP-BIRT indicates that the next hop BFR-NBR on the path to BFER F with BFR-id 2 is BFR F.

The 3rd row in the EP-BIRT indicates that the next hop BFR-NBR on the path to BFER E with BFR-id 3 is BFR F.

The 4-th row in the EP-BIRT indicates that the next hop BFR-NBR on the path to BFER H with BFR-id 4 is BFR H.

The 5-th row in the EP-BIRT indicates that the next hop BFR-NBR on the path to BFER A with BFR-id 5 is BFR B.

From this EP-BIRT for BFER D on BFR C, an Egress Protection Bit Index Forwarding Table (EP-BIFT) is derived. This EP-BIFT for BFER D is shown in Figure 8.

The 2nd and 3rd rows in the EP-BIRT have the same SI = 0, the same next hop BFR-NBR = E and the same backup information {EP=0, BE-BFER=0}. The F-BM for each of these two rows in the EP-BIFT is the logical OR of the BitStrings of these rows, which is 00110 (00010 OR 00100 = 00110).

BFR-id (SI:BitString)	F-BM	BFR-NBR (Next Hop)	{EP, BE-BFER} (Backup Info)
1 (0:00001)	00001	NULL	EP=1, BE-BFER=H
2 (0:00010)	00110	F	EP=0, BE-BFER=0
3 (0:00100)	00110	F	EP=0, BE-BFER=0
4 (0:01000)	01000	H	EP=0, BE-BFER=0
5 (0:10000)	10000	B	EP=0, BE-BFER=0

Figure 8: EP-BIFT for BFER D on BFR C

The F-BM for 1st row in the EP-BIFT is 00001.

The F-BM for 4-th row in the EP-BIFT is 01000.

The F-BM for 5-th row in the EP-BIFT is 10000.

5.4. Forwarding using EP-BIFT

Suppose that there is a multicast traffic from BFR A as ingress/BFIR to egresses/BFERs D, F and E. For every packet of the traffic, after receiving it, BFR A adds a BIER header into the packet and sends the

packet with the BIER header to BFR B, which sends the packet BFR C. The BIER header contains (SI:BitString) = (0:00111) for egresses/BFRs D, F and E.

In normal operations, after receiving the packet from BFR B, BFR C copies, updates and sends the packet to BFR D and BFR F using the BIFT on BFR C according to the forwarding procedure defined in [\[RFC8279\]](#).

Once BFR C detects the failure of its BFR-NBR D, which is a BFER, after receiving the packet from BFR B, BFR C copies, updates and sends the packet using the EP-BIFT for BFER D on BFR C according to the updated forwarding procedure.

For the packet targeting to BFER D (i.e., primary egress node), BFR C sends it towards BFER H (i.e., backup egress node), which is configured to protect BFER D.

For example, once BFR C detects the failure of its BFR-NBR D, after receiving the packet from BFR B, BFR C copies, updates and sends the packet to BFR H and BFR F using the EP-BIFT for BFER D on BFR C.

The packet received by BFR C from BFR B contains (SI:BitString) = (0:00111). The rightmost one bit in BitString is bit 1. For BFER 1 (0:00001) (i.e., BFR D as BFER), BFR C gets the 1st row (i.e., forwarding entry) in the EP-BIFT for BFER D. EP = 1 in the row indicates that BFER D is protected against the failure of D. BFR C clears bit 1 in Packet's BitString and sets bit 4 (i.e., the bit for BE-BFER = H) in Packet's BitString to one. The BitString in Packet is 01110 now. This lets BFR C send Packet to BE-BFER H.

For the packet containing BitString = 01110, the rightmost one bit in BitString is bit 2. For BFER 2 (0:00010) (i.e., BFR F as BFER), BFR C gets the 2nd row (i.e., forwarding entry) in the EP-BIFT for BFER D. EP = 0 and the next hop BFR-NBR is F in the row. BFR C copies, updates and sends the packet to F.

The packet sent to F contains the updated BitString = 00110, which is 01110 & F-BM in the 2nd row = 01110 & 00110 = 00110.

After sending the packet to F, BFR C updates the original packet by ANDing its BitString with the INVERSE of the F-BM in the 2nd row. The updated BitString = 01000, which is 01110 & ~F-BM in the row = 01110 & 11001 = 01000.

For the packet containing BitString = 01000, the rightmost one bit in BitString is bit 4. For BFER 4 (0:01000) (i.e., BFR H as BFER), BFR C gets the 4-th row (i.e., forwarding entry) in the EP-BIFT for BFER

D. EP = 0 and the next hop BFR-NBR is H in the row. BFR C copies, updates and sends the packet to H. The packet sent to H contains BitString = 01000.

After sending the packet to H, BFR C updates the original packet by ANDing its BitString with the INVERSE of the F-BM in the 4-th row. The updated BitString = 00000, which is 01000 & ~F-BM in the row = 01000 & 10111 = 00000.

The updated packet has BitString without any one bit. BFR C finishes forwarding the packet to F and H (backup for D). BFR F will send the packet to E.

6. Security Considerations

TBD.

7. IANA Considerations

No requirements for IANA.

8. Acknowledgements

The authors would like to thank people for their comments to this work.

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>>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", [RFC 5226](#), DOI 10.17487/RFC5226, May 2008, <<https://www.rfc-editor.org/info/rfc5226>>.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", [RFC 5250](#), DOI 10.17487/RFC5250, July 2008, <<https://www.rfc-editor.org/info/rfc5250>>.
- [RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", [RFC 5286](#), DOI 10.17487/RFC5286, September 2008, <<https://www.rfc-editor.org/info/rfc5286>>.

- [RFC5714] Shand, M. and S. Bryant, "IP Fast Reroute Framework", [RFC 5714](#), DOI 10.17487/RFC5714, January 2010, <<https://www.rfc-editor.org/info/rfc5714>>.
- [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>>.
- [RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding Scope Link State PDUs (LSPs)", [RFC 7356](#), DOI 10.17487/RFC7356, September 2014, <<https://www.rfc-editor.org/info/rfc7356>>.
- [RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", [RFC 7490](#), DOI 10.17487/RFC7490, April 2015, <<https://www.rfc-editor.org/info/rfc7490>>.
- [RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", [RFC 7684](#), DOI 10.17487/RFC7684, November 2015, <<https://www.rfc-editor.org/info/rfc7684>>.
- [RFC7770] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", [RFC 7770](#), DOI 10.17487/RFC7770, February 2016, <<https://www.rfc-editor.org/info/rfc7770>>.
- [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>>.
- [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>>.

[9.2.](#) Informative References

- [I-D.ietf-rtgwg-segment-routing-ti-lfa]
Litkowski, S., Bashandy, A., Filsfils, C., Decraene, B.,
Francois, P., Voyer, D., Clad, F., and P. Camarillo,
"Topology Independent Fast Reroute using Segment Routing",
[draft-ietf-rtgwg-segment-routing-ti-lfa-04](#) (work in
progress), August 2020.
- [I-D.ietf-spring-segment-protection-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu,
"Segment Protection for SR-TE Paths", [draft-ietf-spring-
segment-protection-sr-te-paths-00](#) (work in progress),
September 2020.
- [RFC7431] Karan, A., Filsfils, C., Wijnands, IJ., Ed., and B.
Decraene, "Multicast-Only Fast Reroute", [RFC 7431](#),
DOI 10.17487/RFC7431, August 2015,
<<https://www.rfc-editor.org/info/rfc7431>>.
- [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>>.
- [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>>.

Authors' Addresses

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing, 102209
China

Email: wangaj3@chinatelecom.cn

Gyan S. Mishra
Verizon Inc.
13101 Columbia Pike
Silver Spring MD 20904
USA

Phone: 301 502-1347
Email: gyan.s.mishra@verizon.com

Yanhe Fan
Casa Systems
USA

Email: yfan@casa-systems.com

Lei Liu
Fujitsu

USA

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks

McLean, VA
USA

Email: xufeng.liu.ietf@gmail.com

