BIER Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2019

D. Merling
M. Menth
University of Tuebingen

March 05, 2019

BIER Fast Reroute
draft-merling-bier-frr-00

Abstract

BIER is a scalable multicast overlay [RFC8279] that utilizes some routing underlay, e.g., IP, to build up its Bit Index Forwarding Tables (BIFTs). This document proposes a Fast Reroute Extension for BIER (BIER-FRR). In case of a link or node failure, the routing underlay may first utilize FRR techniques to restore connectivity and then its forwarding tables converge. After that, BIER can update its BIFTs, which requires time. BIER packets may not be delivered until the last procedure has finished. With BIER-FRR, a BIER Forwarding Router (BFR) can deliver BIER packets again after a link or node failures as soon as the connectivity within the routing underlay is restored and the BFR is informed about a next-hop (NH) that is unreachable on a lower layer. BIER-FRR provides a mode for link protection and node protection. For link protection, it tunnels traffic to the next-hop using the underlying routing. For node protection, it forwards BIER packets to their specific next-hop and next-next-hops using tunnels in the underlying routing after applying a suitable backup bitmask to the bitstring in the BIER header of each packet. This procedure prevents duplicates. If topology allows, BIER-FRR achieves full protection against any single component failure. For link protection, BIER-FRR requires only a minor change to the forwarding logic. For node protection, BIER-FRR also requires backup entries in the BIFT.

This document describes the concept and operating principles of BIER-FRR. It defines the necessary modifications to the BIER forwarding Procedure and the BIFT, and explains how backup entries are computed.

Copyright Notice

Table of Contents

1. Introduction

With BIER [RFC8279], Bit-Forwarding Routers (BFRs) forward packets based on a bitstring in the BIER header using the information in their Bit Index Forwarding Tables (BIFTs). In case of a persistent link or node failure, BIER traffic may not be delivered until the BIFT has been updated based on the re-converged routing underlay. The routing underlay restores connectivity more quickly than BIER, in particular if the routing underlay leverages fast reroute (FRR) mechanisms because then the forwarding ability is retained before forwarding tables of the routing underlay have converged. In this document we propose Fast Reroute Extension for BIER (BIER-FRR). It enables a BFR to quickly reroute BIER packets as soon as the underlying routing works again and it is informed about a next-hop (NH) that is unreachable on a lower layer.

We first explain the problem and distinguish link and node failures for that purpose. In case of a persistent link failure, a BFR cannot deliver BIER traffic until the NH in the BIFT is updated with an appropriate node. In case of a node failure, the entire multicast subtree behind the failed not is not reachable until the BIFT is updated. Thus, in either case, BIER's connectivity is restored only after the underlying routing has converged and the BIFTs have been updated. This may require substantial time during which BIER traffic is dropped. An exception are unreachable NHs to which BIER traffic is sent through tunnels in the routing underlay. They are reachable again as soon as the routing underlay works again.

BIER-FRR tackles this problems with two different modes that have different implementation complexity: link protection and node protection. In any case, a BFR with an unreachable NH needs to be informed about the failure and acts as a point of local repair (PLR). E.g., BFD mechanisms may be used [I-D.hu-bier-bfd] to detect failed NHs. With BIER-FRR link protection, a BFR tunnels affected BIER packets towards the NH using a tunnel in the underlying routing. Then, this traffic can be delivered as soon as the underlying routing works again. With BIER-FRR node protection, a BFR tunnels affected BIER packets to the NH and all relevant next-next-hops (NNHs) in the underlying routing after applying suitables backup bitmasks to the

bitstring in the BIER header. This procedure ensures that both the NH and all potential multicast subtrees receive the traffic if possible, and it prevents potential duplicates and loops on the BIER layer. Thus, BIER-FRR basically implements 1:1 protection. The latter is discussed in [I-D.xiong-bier-resilience] without proposing a specific mechanism.

This document describes the concept of BIER-FRR, protection properties, the computation of backup bitmasks, and gives a detailed BIER-FRR forwarding example.

2. Terminology

The following sections require the understanding of certain abbreviations and definitions that were defined within other documents, especially [RFC8279]. To facilitate the reading of this document, they are shortly explained in this section.

- BFR: Bit-Forwarding Router, Section 1 of [RFC8279]

  A device that acts as a BIER forwarding device in the BIER domain.

- BFIR: Bit-Forwarding Ingress Router, Section 1 of [RFC8279]

  Entry point to the BIER domain. A BFIR adds a BIER header to a multicast packet.

- BFER: Bit-Forwarding Egress Router, Section 1 of [RFC8279]

Exit point of the BIER domain. A BFER removes the BIER header of a multicast packet.

- NH: Next-hop

  The next downstream BFR to which a packet is forwarded.

- BIFT: Bit Index Forwarding Table, Section 6.4 of [RFC8279]

  A BFR uses its BIFT to determine the NH(s) of a BIER packet. The BIFT maps a F-BM to each NH of a BFR.

- F-BM: Forwarding bitmask Section 6.4 of [RFC8279]

  A F-BM is a bitmask that indicates which destinations are reached via the subtree of the corresponding NH. A F-BM is applied to the BIER header by bitwise AND'ing the F-BM with the bitstring in the BIER header. This prevents duplicates at BFERs.

- Routing underlay: Section 4.1 of [RFC8279]

  The routing underlay connects pairs of BFR. If a typical Interior Gateway Protocol (IGP) like OSPF is used, the multicast packets will be forwarded on shortest paths. Other routing underlays with different path layouts are possible. The routing underlay is used to determine the NH entries of the BIFT.

- BIER Layer: Section 4.2 of [RFC8279]

  Conceptually the BIER layer is placed above the routing underlay. The BIER layer can be understood as a transport layer for multicast packets. It consists of all necessary protocols and mechanisms to forward a BIER packet from a BFIR, over potentially multiple BFRs, to a set of BFERs.

- Multicast Overlay: Section 4.3 of [RFC8279]

  Conceptually the multicast overlay is placed above the BIER layer. It is used to maintain information about egress points of multicast groups. The multicast overlay passes those information to the BIER layer which then determines the corresponding BIER

headers.

- PLR: Point of Local Repair

  A node that cannot forward a packet due to an unreachable NH.

- BIER-FRR: Fast Reroute Extension for BIER (BIER-FRR)

  A mechanism to restore connectivity on the BIER layer as soon as BFRs are informed about non-reachable NHs and the underlying routing works again.

- BIER-FRR link protection

  A mode of BIER-FRR that can handle only link failures.

- BIER-FRR node protection

  A mode of BIER-FRR that can handle both link and node failures. It is more complex than BIER-FRR link protection.

- NNH: Next-next-hop

  Next downstream BFR after the NH.

2.1. 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 [RFC2119].

3. Problem Statement

We first consider the impact of link failures and then the one of node failures on the behavior of a BIER network without BIER-FRR.

3.1. Link Failures

The effect of a link failure depends on the technology used for encapsulation of BIER packets. We distinguish three cases. (1) BIER

packets are carried over some lower-layer technology with protection. (2) BIER packets are tunneled through the routing underlay. (3) BIER packets are carried over some lower-layer technology without protection.

3.1.1. BIER Encapsulation within a Lower-Layer Technology with Protection

MPLS is an example for a lower-layer technology with protection capabilities. In case of a link failure, first packets are lost, then the protection mechanism of the lower-layer technology quickly restores the link. From then on, packets are no longer lost.

3.1.2. BIER Encapsulation within the Routing Underlay

IP is an example for a routing underlay. The routing underlay is expected to deal with failures of lower- layer technologies. In case of a link failure, packets are lost. If the failure persists due to a lower-layer technology without protection, the routing underlay is informed about the failure. The routing underlay may leverage FRR techniques, e.g., Loop-Free Alternates (LFAs) [RFC5286], to quickly restore reachability so that packets are delivered again which are sent encapsulated the within routing underlay. From then on, also BIER packets encapsulated within the routing underlay are delivered again.

At the same time, routing reconvergence is triggered. When the routing has converged after some time, forwarding tables of the routing underlay are updated. Based on them, new NHs for BIFTs are computed and installed so that the PLR delivers BIER packets to a different NH than the one that is still unreachable via the lower-layer technology.

3.1.3. BIER Encapsulation within a Lower-Layer Technology without Protection

Ethernet is an example for a lower-layer technology without protection. In case of a link failure, the failure persists from the perspective of BIER and the routing underlay unless the failure is repaired. As a consequence, packets are lost. Then, the routing underlay acts as described above to restore rechability and finally

updates its forwarding tables. By that time, BIER packets encapsulated within the lower-layer technology are still dropped. Then, new NHs for BIFTs are computed based on the new forwarding tables of the routing underlay and are installed. From then on, BIER can deliver packets again over a different NH.

When BIER-FRR is used, BIER packets can be delivered again as soon as the BFR is informed about the unreachable NH and routing underlay works again.

3.2. Node Failures

The effect of node failures is more severe. First, the packets cannot be delivered to the failed node. This, however, cannot be repaired. Second, multicast distribution trees downstream of a failed NH cannot receive traffic as the failed NH replicate traffic towards relevant NNHs. This problem is solved neither by lower-layer technologies with link protection nor by BIER encapsulation within the routing underlay. Therefore, BIER packets sent to the failed NH are dropped until BIFTs are updated based on reconverged forwarding tables of the routing underlay. This may require quite some time.

When BIER-FRR node protection is used, BIER packets can be delivered along the affected multicast tree as soon as the BFR is informed about the unreachable NH and the routing underlay works again.

4. Fast Reroute Extension for BIER (BIER-FRR)

This section describes the concept of BIER-FRR. In case of a link or node failure, it reroutes BIER packets until the BIFTs are updated or the failure is repaired. BIER-FRR offers two modes: link protection and node protection. Their protection level is summarized.

4.1. BIER-FRR Link Protection

We introduce the mechanism and illustrate it by an example.

4.1.1. Mechanism

When a BFR is informed about an unreachable NH, it tunnels all BIER packets towards that NH through the routing underlay. As soon as the routing underlay works again, the BIER packets are delivered to the NH if the NH still works. Then, the NH forwards the BIER packets further along the multicast distribution tree.

4.1.2. Example

Figure 1 shows an example toplogy for the routing underlay and Figure 2 the multicast distribution tree for BFR 1 on the BIER Layer which is computed based on shortest paths.

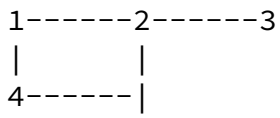

Figure 1: Example topology for the routing underlay.

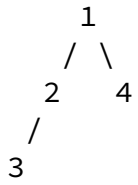

Figure 2: Multicast distribution tree for BFR 1 on the BIER Layer.

When BFR 1 sends a packet to BFR 3, the NH is BFR 2. If link 1<->2 fails, packets encapsulated within a lower-layer technology can no longer be delivered from BFR 1 to BFR 2. As soon as BFR 1 is informed that BFR 2 is no longer reachable, it encapsulates the BIER packets to BFR 3 within the routing underlay towards BFR 2. When the routing underlay has restored connectivity, the BIER packets are tunneled from BFR 1 via BFR 4 to BFR 2 which decapsulates them. Then BFR further forwards the BIER packets to BFR 3.

4.2. BIER-FRR Node Protection

We introduce the mechanism, illustrate it by an example, and explain how needed backup BIFT entries are computed.

4.2.1. Mechanism

When a BFR is informed about an unreachable NH, it tunnels all affected BIER packets to that NH if the NH receives a copy of the BIER packet, and to all NNHs over which copies of the BIER packet are

to be delivered. The latter are the relevant NNHs. Before tunneling the packets, their bitstrings are modified using backup F-BMs to avoid forwarding loops and duplicates.

The BIFT and its operation are explained in detail in Section 6.4 of [RFC8279]. We briefly revise them to facilitate further reading before introducing backup BIFT entries to support the solution presented above. Figure 3 illustrates the structure of the BIFT. The BIFT contains for each BFR-id a F-BM and the next hop (BFR-NBR). The BFR-id is mapped to a position in the bitstring of the BIER header; for this purpose, the bits within a bitstring are counted from right to left starting with 1. The F-BM is also a bitstring and indicates all BFRs that are reachable through the multicast distribution subtree via BFR-NBR. As a result, the F-BMs in a BIFT are either identical (same BFR-NBR) or disjoint with regard to activated bits. For transmission of BIER packets, the BIFT is used together with a copy of the bitstring of the BIER header. Processing is performed by the following loop. An entry from the BIFT is selected that holds a BFR-id which is set in the copy of the bitstring. Then, the F-BM of that entry is applied to the bitstring of the BIER packet and then the BIER packet is sent to the indicated BFR-NBR. The bits of the F-BM are cleared in the bitstring copy and the loop restarts. It ends when all bits in the bitstring copy are zero.

BFR-id	F-BM	BFR-NBR
1		

Figure 3: Structure of the BIFT according to [RFC8279].

BFR-id	F-BM	BFR-NBR
1	primary F-BM	primary NH
	backup F-BM	backup NH
...	...	...

Figure 4: Structure of a BIFT with primary and backup entries.

To support BIER-FRR node protection, backup BIFT entries for protected BFR-NBRs are added to the BIFT. That structure is illustrated in Figure 4. We call the normal BIFT entries primary

entries. Backup BIFT entries have the same structure as primary BIFT entries and are used for forwarding in the same way. The set of

active bits in a primary BIFT entry must equal the set of active bits in its corresponding backup entries to guarantee that all destinations in the multicast distribution subtree via BFR-NBR are protected.

If the BFR-NBR of a primary BIFT entry is reachable, the corresponding backup BIFT entries are ignored in the forwarding process. If the BFR-NBR of a primary BIFT entry is unreachable, the BIER packet is processed using the corresponding backup BIFT entries instead of the primary BIFT entry. BIER packets sent by a backup BIFT entry MUST be tunneled through the routing underlay to the backup BFR-NBR after application of the backup F-BM.

There are other options to organize the backup entries just as there are options for more scalable BIFT organization.

4.2.2. Example

Figure 5 shows an example toplogy for the routing underlay and Figure 6 the multicast distribution trees for BFR 1 and BFR 2 on the BIER Layer which are computed based on shortest paths.

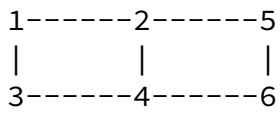

Figure 5: Example topology for the routing underlay.

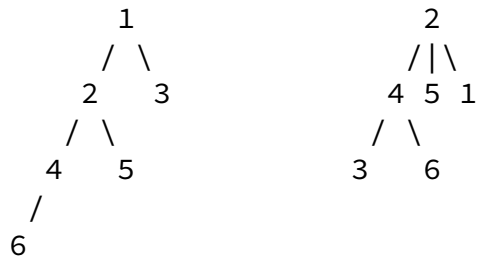

Figure 6: Multicast distribution trees for BFR 1 and BFR 2 on the BIER Layer.

Figure 7 gives the BIFT for BFR 1 with backup entries.

FRR-BIFT BFR 1		
BFR-id	F-BM	BFR-NBR
1	000001	-
	-	-
2	111010	2
	000010	2
3	000100	3
	000100	3
4	111010	2
	101000	4
5	111010	2
	010000	5
6	111010	2
	101000	4

Figure 7: BIFT of BFR 1 with backup entries.

When BFR 1 sends a BIER packet to BFR 6, the NH is BFR 2. If link 1<->2 fails, BIER packets encapsulated within a lower-layer technology can no longer be delivered from BFR 1 to BFR 2. As soon as BFR 1 is informed that BFR 2 is no longer reachable, it applies backup BIFT entries to forward affected BIER packets. That means, it modifies the bitstring of BIER packets towards BFR 6 with the

appropriate backup F-BM and sends them to backup NH BFR 4 after encapsulation within the routing underlay. Therefore, the packets are tunneled from BFR 1 via BFR 3 to BFR 4. BFR 4 decapsulates the packet and a copy of the BIER packet is delivered to BFR 6.

4.2.3. Computation of Backup BIFT Entries

We explain the computation and give an example.

4.2.3.1. Computation

BIER-FRR node protection ensures that a PLR can send BIER packets in case of an unreachable NH to all BFRs in the downstream multicast subtree of the unreachable NH. For this purpose, backup entries for these BFRs need to be provided in the BIFT of the PLR. We compute them differently for the NHs of PLRs and for all other BFRs which

belong to multicast subtrees starting with a NNH. This leads to two computation rules:

1. BIER packets for NHs are sent to the NHs (backup-NH = NH) via a tunnel and the backup F-BM must ensure that these BIER packets are not forwarded any further. That means, the backup F-BM contains only the BFR-id of the NH.

2. BIER packets for other BFRs are sent via a tunnel to the NNH in the multicast subtree they belong to. Also all other BFRs in the same multicast subtree should be reached with the same BIER packet. Therefore, the backup F-BM for a BFR contains the BFR-ids for all BFRs in its multicast subtree starting with the respective NNH. Thus, the corresponding backup F-BM can be computed by ANDing the PLR's F-BM for the NH and the NH's F-BM for the specific NNH.

4.2.3.2. Example

We consider the BIFT of BFR 1 in Figure 7.

Example for rule (1): The backup BIFT entry for BFR 2 has a F-BM that just contains BFR 2 (000010).

Example for rule (2): The backup BIFT entry for BFR 4 has a F-BM that

contains BFR 4 and BFR 6 (101000). It is computed ANDing the F-BM of BFR 1 for BFR 2 (111010) and the F-BM of BFR 2 for BFR 4 (101100). The latter has been derived from the multicast distribution tree of BFR 2 in Figure 6.

4.3. Protection Level

BIER-FRR is a protection scheme that reacts when a NH is no longer reachable. It is a local mechanism that does not require signaling or cooperation with other nodes, possibly except for the detection of locally unreachable NHs.

The protection ensures that BIER multicast traffic is forwarded to all destinations that are reachable over the routing underlay and that no duplicates occur. The protection is fast as it works as soon as the BFR is informed about a unreachable NH and the underlying routing works again after the failure occurred.

BIER-FRR link protection is able to protect single link failures within a network provided that the underlying routing can restore full connectivity. Multiple link failures within a network are not necessarily protected.

BIER-FRR node protection protects both single link and single node failures within a network provided that the underlying routing can restore full connectivity. Multiple link and node failures within a network are not necessarily protected.

The design of BIER-FRR guarantees loop-freeness on the BIER layer. Since the BIER packet is tunneled, the BIER is header is only used for forwarding if the tunneled packet arrives at the designated BFR. Loop-freeness on the routing underlay is out of the scope of this document.

5. Necessary Changes to the BIER Architecture

This section serves as an overview to list the necessary conceptual features or changes that are required for BIER-FRR.

5.1. Unicast Tunneling

Unicast tunnels to connect two not directly adjacent BFRs are already available. This feature is described in Section 6.9 of [RFC8279].

5.2. Detecting Unreachable N(N)Hs

A liveness component (e.g. BFD) has to be added to enable the detection of unreachable NHs. This feature has been proposed in [I-D.hu-bier-bfd].

5.3. BIFT with backup entries

The BIFT has to be extended with backup entries as described in Section XXX. When the regular BIER forwarding procedure yields an unreachable NH, the backup entry contains a backup F-BM for header modification and a NNH to which the BIER packet should be tunneled to.

6. Security Considerations

This memo does not extend the security considerations for BIER.

7. IANA Considerations

This document requests no action by IANA.

8. References

8.1. Normative References

[I-D.hu-bier-bfd]
hu, f., Mirsky, G., Xiong, Q., and C. Liu, "BIER BFD", draft-hu-bier-bfd-02 (work in progress), October 2018.

[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>.

[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>.

8.2. Informative References

[I-D.xiong-bier-resilience] Xiong, Q., hu, f., and G. Mirsky, "The Resilience for BIER", draft-xiong-bier-resilience-01 (work in progress), October 2018.

[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>.

Authors' Addresses

Daniel Merling
University of Tuebingen
Sand 13
Tuebingen 72076
Germany

Phone: +49 7071 29-70507
Email: daniel.merling@uni-tuebingen.de

Michael Menth
University of Tuebingen
Sand 13
Tuebingen 72076

Germany

Phone: +49 7071 29-70505
Email: menth@uni-tuebingen.de