Network Working Group   H. Chen
Internet-Draft   M. McBride
Intended status: Standards Track   Futurewei
Expires: 28 August 2022   A. Wang
                                                        China Telecom
                                                           G. Mishra
                                                         Verizon Inc.
                                                             L. Liu
                                                            Fujitsu
                                                             X. Liu
                                                       Volta Networks
                                                    24 February 2022

BIER-TE for Broadcast Link
draft-chen-bier-te-lan-03

Abstract

   This document describes extensions to "Bit Index Explicit Replication
   Traffic Engineering" (BIER-TE) for supporting LANs (i.e., broadcast
   links). For a multicast packet with an explicit point-to-multipoint
   (P2MP) path traversing LANs, the packet is replicated and forwarded
   statelessly along the path.

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 28 August 2022.

Copyright Notice

Table of Contents

1. Introduction

   [I-D.ietf-bier-te-arch] introduces Bit Index Explicit Replication
   (BIER) Traffic/Tree Engineering (BIER-TE). It is an architecture for
   per-packet stateless explicit point to multipoint (P2MP) multicast
   path/tree. A Bit-Forwarding Router (BFR) in a BIER-TE domain has a
   BIER-TE Bit Index Forwarding Table (BIFT). A BIER-TE BIFT on a BFR
   comprises a forwarding entry for a BitPosition (BP) assigned to each

of the adjacencies of the BFR. If the BP represents a forward
connected adjacency, the forwarding entry for the BP forwards the
multicast packet with the BP to the directly connected BFR neighbor
of the adjacency. If the BP represents a BFER (i.e., egress node) or
say a local decap adjacency, the forwarding entry for the BP

decapsulates the multicast packet with the BP and passes a copy of
the payload of the packet to the packet's NextProto within the BFR.

In [I-D.ietf-bier-te-arch], for a LAN, the adjacency to each
neighboring BFR on the LAN is given a unique BitPosition. The
adjacency of this BitPosition is a forward connected adjacency
towards the BFR and this BitPosition is populated into the BIFT of
all the other BFRs on that LAN. This solution for a LAN does not
work in some cases.

For a packet with an explicit point-to-multipoint (P2MP) path, if the
path traverses some BFRs/nodes on a LAN, each of these BFRs/nodes on
the LAN may receive duplicated packets. Thus some of the egress
nodes will receive duplicated packets.

This document proposes a solution for LANs to resolve this issue.
For a packet with an explicit P2MP path traversing LANs (i.e.,
broadcast links), the packet is replicated and forwarded statelessly
along the path. Each of the egress nodes of the path will not
receive any duplicated packet.

1.1. Terminology

BIER: Bit Index Explicit Replication.

BIER-TE: BIER Traffic Engineering.

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.

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.

IGP: Interior Gateway Protocol.

LSDB: Link State DataBase.

OSPF: Open Shortest Path First.

IS-IS: Intermediate System to Intermediate System.

2. Example Application of Current BIER-TE with LAN

   This section illustrates an example application of the current BIER-
   TE defined in [I-D.ietf-bier-te-arch] to the BIER-TE topology with
   LAN in Figure 1.

2.1. Example BIER-TE Topology with LAN

   An example BIER-TE topology with a LAN for a BIER-TE domain is shown
   in Figure 1. It has 9 nodes/BFRs A, B, C, D, E, F, G, H and K.
   Nodes/BFRs D, F, E, H, A and K are BFERs and have local decap
   adjacency BitPositions (BPs for short) 1, 2, 3, 4, 5 and 6
   respectively. For simplicity, these BPs are represented by
   (SI:BitString), where SI = 0 and BitString is of 8 bits. BPs 1, 2,
   3, 4, 5 and 6 are represented by 1 (0:00000001), 2 (0:00000010), 3
   (0:00000100), 4 (0:00001000), 5 (0:00010000) and 6 (0:00100000)
   respectively.

                                 6
                               (K)
                                /14'
                               /
                     4'   /13' | 4
                  /----------(G)-------+   (H)

```
                     /                   16'    |_____/
                    /                    |          17'
                  /3'                _____|
     5    1'        2'    /            6'    /15'    |        18'
     ( A )-----------( B )-------------( C )        +--------( D )
      \10'             7'\    5'         \11'       |          1
        \                 \               \       LAN
         \                 \               \
          \         9'    \8'               \12'
           \-------------( E )             ( F )
                          3                  2
```
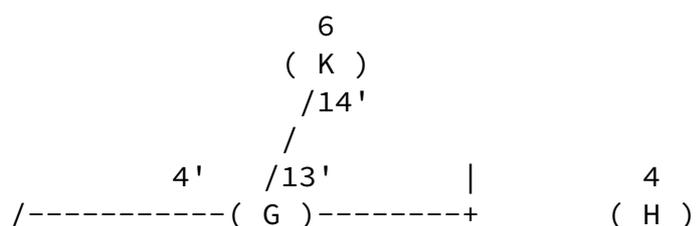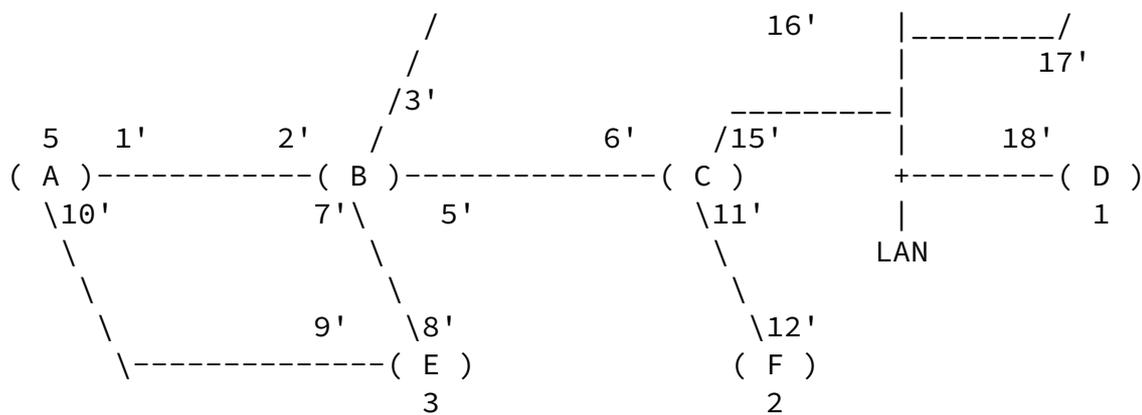
Figure 1: Example BIER-TE Topology with BP to BFR on LAN

The BitPositions for the forward connected adjacencies are represented by i', where i is from 1 to 18. In one option, they are encoded as (n+i), where n is a power of 2 such as 32768. For simplicity, these BitPositions are represented by (SI:BitString), where SI = (6 + (i-1)/8) and BitString is of 8 bits. BitPositions i'

(i from 1 to 18) are represented by 1'(6:00000001), 2'(6:00000010), 3'(6:00000100), 4'(6:00001000), 5'(6:00010000), 6'(6:00100000), 7'(6:01000000), 8'(6:10000000), 9'(7:00000001), 10'(7:00000010), . . . , 16'(7:10000000), 17'(8:00000001), 18'(8:00000010).

For a link between two nodes X and Y, there are two BitPositions for two forward connected adjacencies. These two forward connected adjacency BitPositions are assigned on nodes X and Y respectively. The BitPosition assigned on X is the forward connected adjacency of Y. The BitPosition assigned on Y is the forward connected adjacency of X.

For example, for the link between nodes B and C in the figure, two forward connected adjacency BitPositions 5' and 6' are assigned to two ends of the link. BitPosition 5' is assigned on node B to B's end of the link. It is the forward connected adjacency of node C. BitPosition 6' is assigned on node C to C's end of the link. It is the forward connected adjacency of node B.

For a LAN (i.e., broadcast link) connecting nodes X1, X2, ..., Xm, there are m BitPositions for m forward connected adjacencies. These m forward connected adjacency BitPositions are assigned on nodes X1,

X2, ..., Xm respectively.

    For the LAN connecting 4 nodes C, G, H and D in the figure, 4 forward
    connected adjacency BitPositions 15', 16', 17' and 18' are assigned
    to C, G, H and D respectively.

2.2. BIER-TE BIFT on BFR

    Every BFR in a BIER-TE domain/topology has a BIER-TE BIFT. This
    section shows the BIER-TE BIFT on every BFR/node of the BIER-TE
    topology with LAN in Figure 1.

    For the BIER-TE topology in Figure 1, each of 9 nodes/BFRs A, B, C,
    D, E, F, G, H and K has its BIER-TE BIFT for the topology. The BIFT
    on a BFR comprises a forwarding entry for each of the adjacencies of
    the BFR.

    The BIER-TE BIFT on BFR A (i.e., node A) is shown in Figure 2. There
    are three adjacencies of A. The 1st one is the forward connected
    adjacency from A to B (represented by BP 2'); the 2nd is the forward
    connected adjacency from A to E (represented by BP 9'); the 3rd is
    the local decap adjacency (represented by BP 5) for BFER (i.e.,
    egress) A. The BIFT on A has three forwarding entries.

```
          +----------------+--------------+------------+
          |  Adjacency BP  |    Action    |  BFR-NBR   |
          |  (SI:BitString)|              |  (Next Hop)|
          +================+==============+============+
          | 2'(6:00000010) | fw-connected |     B      |
          +----------------+--------------+------------+
          | 9'(7:00000001) | fw-connected |     E      |
          +----------------+--------------+------------+
          | 5 (0:00010000) | local-decap  |            |
          +----------------+--------------+------------+
```

                   Figure 2: BIER-TE BIFT on BFR A

    The 1st forwarding entry in the BIFT is for BitPosition 2', which is
    the forward connected adjacency from A to B. For a multicast packet

with BitPosition 2', which indicates that the P2MP path in the packet
traverses the adjacency from A to B, the forwarding entry forwards
the packet to B along the link from A to B.

The 2nd forwarding entry in the BIFT is for BitPosition 9', which is
the forward connected adjacency from A to E. For a multicast packet
with BitPosition 9', which indicates that the P2MP path in the packet
traverses the adjacency from A to E, the forwarding entry forwards
the packet to E along the link from A to E.

The 3rd forwarding entry in the BIFT locally decapsulates a multicast
packet with BitPosition 5 and passes a copy of the payload of the
packet to the packet's NextProto. It is for BitPosition 5, which is
the local decap adjacency for BFER (i.e., egress) A. For a multicast
packet with BitPosition 5, which indicates that the P2MP path in the
packet has node A as one of its destinations (i.e., egress nodes),
the forwarding entry decapsulates the packet and passes a copy of the
payload of the packet to the packet's NextProto within node A.

The BIER-TE BIFT on BFR B (i.e., node B) is shown in Figure 3. There
are four forward connected adjacencies of B. They are the forward
connected adjacencies from B to A (represented by BP 1'), B to G
(represented by BP 4'), B to C (represented by BP 6') and B to E
(represented by BP 8') respectively. The BIFT on B has four
forwarding entries for these adjacencies.

```
            +----------------+--------------+------------+
            |  Adjacency BP  |    Action    |  BFR-NBR   |
            |  (SI:BitString)|              |  (Next Hop)|
            +================+==============+============+
            | 1'(6:00000001) | fw-connected |     A      |
            +----------------+--------------+------------+
            | 4'(6:00001000) | fw-connected |     G      |
            +----------------+--------------+------------+
```

```
             | 6'(6:00100000) | fw-connected |     C      |
             +----------------+--------------+-----------+
             | 8'(6:10000000) | fw-connected |     E      |
             +----------------+--------------+-----------+
```

                    Figure 3: BIER-TE BIFT on BFR B

   The 1st forwarding entry in the BIFT is for BitPosition 1', which is
   the forward connected adjacency from B to A. For a multicast packet
   with BitPosition 1', which indicates that the P2MP path in the packet
   traverses the adjacency from B to A, the forwarding entry forwards
   the packet to A along the link from B to A.

   The 2nd forwarding entry in the BIFT is for BitPosition 4', which is
   the forward connected adjacency from B to G. For a multicast packet
   with BitPosition 4', which indicates that the P2MP path in the packet
   traverses the adjacency from B to G, the forwarding entry forwards
   the packet to G along the link from B to G.

   The 3rd forwarding entry in the BIFT is for BitPosition 6', which is
   the forward connected adjacency from B to C. For a multicast packet
   with BitPosition 6', which indicates that the P2MP path in the packet
   traverses the adjacency from B to C, the forwarding entry forwards
   the packet to C along the link from B to C.

   The 4-th forwarding entry in the BIFT is for BitPosition 8', which is
   the forward connected adjacency from B to E. For a multicast packet
   with BitPosition 8', which indicates that the P2MP path in the packet
   traverses the adjacency from B to E, the forwarding entry forwards
   the packet to E along the link from B to E.

   The BIER-TE BIFT on BFR C (i.e., node C) is shown in Figure 4. There
   are five forward connected adjacencies of C. They are the forward
   connected adjacencies from C to B (represented by BP 5'), C to F
   (represented by BP 12'), C to G (represented by BP 14'), C to H
   (represented by BP 15') and C to D (represented by BP 16')
   respectively. The BIFT on C has five forwarding entries for these
   adjacencies.

```
          +----------------+--------------+------------+
```

```
+----------------+-------------+-----------+
| Adjacency BP   |   Action    |  BFR-NBR  |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
| 5'(6:00010000) | fw-connected |    B     |
+----------------+-------------+-----------+
| 12'(7:00001000)| fw-connected |    F     |
+----------------+-------------+-----------+
| 16'(7:10000000)| fw-connected |    G     |
+----------------+-------------+-----------+
| 17'(8:00000001)| fw-connected |    H     |
+----------------+-------------+-----------+
| 18'(8:00000010)| fw-connected |    D     |
+----------------+-------------+-----------+
```

               Figure 4: BIER-TE BIFT on BFR C

The BIER-TE BIFT on BFR D (i.e., node D) is shown in Figure 5. There
are four adjacencies of D. Three of them are the forward connected
adjacencies from D to C (represented by BP 13'), D to G (represented
by BP 14') and D to H (represented by BP 15') respectively; the other
is the local decap adjacency (represented by BP 1) for BFER (i.e.,
egress) D. The BIFT on D has four forwarding entries for these
adjacencies.

```
+----------------+-------------+-----------+
| Adjacency BP   |   Action    |  BFR-NBR  |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
| 15'(7:01000000)| fw-connected |    C     |
+----------------+-------------+-----------+
| 16'(7:10000000)| fw-connected |    G     |
+----------------+-------------+-----------+
| 17'(8:00000001)| fw-connected |    H     |
+----------------+-------------+-----------+
| 1 (0:00000001) | local-decap |          |
+----------------+-------------+-----------+
```

               Figure 5: BIER-TE BIFT on BFR D

The BIER-TE BIFT on BFR E (i.e., node E) is shown in Figure 6. There
are three adjacencies of E. Two of them are the forward connected
adjacencies from E to B (represented by BP 7') and E to A
(represented by BP 10') respectively; the other is the local decap
adjacency (represented by BP 3) for BFER (i.e., egress) E. The BIFT
on E has three forwarding entries for these adjacencies.

```
+-----------------+-------------+-----------+
|   Adjacency BP  |    Action   |  BFR-NBR  |
|   (SI:BitString)|             | (Next Hop)|
+=================+=============+===========+
|  7'(6:01000000) | fw-connected|     B     |
+-----------------+-------------+-----------+
| 10'(7:00000010) | fw-connected|     A     |
+-----------------+-------------+-----------+
|  3 (0:00000100) | local-decap |           |
+-----------------+-------------+-----------+
```

                  Figure 6: BIER-TE BIFT on BFR E

   The BIER-TE BIFT on BFR F (i.e., node F) is shown in Figure 7. There
   are two adjacencies of F. The 1st one is the forward connected
   adjacencies from F to C (represented by BP 11'); the 2nd is the local
   decap adjacency (represented by BP 2) for BFER (i.e., egress) F. The
   BIFT on F has two forwarding entries for these adjacencies.

```
+-----------------+-------------+-----------+
|   Adjacency BP  |    Action   |  BFR-NBR  |
|   (SI:BitString)|             | (Next Hop)|
+=================+=============+===========+
| 11'(7:00000100) | fw-connected|     C     |
+-----------------+-------------+-----------+
|  2 (0:00000010) | local-decap |           |
+-----------------+-------------+-----------+
```

                  Figure 7: BIER-TE BIFT on BFR F

   The BIER-TE BIFT on BFR G (i.e., node G) is shown in Figure 8. There
   are four forward connected adjacencies of G. They are the
   adjacencies from G to B (represented by BP 3'), G to C (represented
   by BP 13'), G to H (represented by BP 15') and G to D (represented by
   BP 16') respectively. The BIFT on G has four forwarding entries for
   these adjacencies.

```
+----------------+-------------+-----------+
| Adjacency BP   |   Action    |  BFR-NBR  |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
|  3'(6:00000100)| fw-connected|     B     |
+----------------+-------------+-----------+
| 14'(7:00100000)| fw-connected|     K     |
+----------------+-------------+-----------+
| 15'(7:01000000)| fw-connected|     C     |
+----------------+-------------+-----------+
| 17'(8:00000001)| fw-connected|     H     |
+----------------+-------------+-----------+
| 18'(8:00000010)| fw-connected|     D     |
+----------------+-------------+-----------+
```

Figure 8: BIER-TE BIFT on BFR G

The BIER-TE BIFT on BFR H (i.e., node H) is shown in Figure 9. There
are four adjacencies of H. Three of them are the forward connected
adjacencies from H to C (represented by BP 13'), H to G (represented
by BP 14') and H to D (represented by BP 16') respectively; the other
is the local decap adjacency (represented by BP 4) for BFER (i.e.,
egress) H. The BIFT on H has four forwarding entries for these
adjacencies.

```
+----------------+-------------+-----------+
| Adjacency BP   |   Action    |  BFR-NBR  |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
| 15'(7:01000000)| fw-connected|     C     |
+----------------+-------------+-----------+
| 16'(7:10000000)| fw-connected|     G     |
+----------------+-------------+-----------+
| 18'(8:00000010)| fw-connected|     D     |
+----------------+-------------+-----------+
|  4 (0:00001000)| local-decap |           |
+----------------+-------------+-----------+
```

Figure 9: BIER-TE BIFT on BFR H

2.3. Example P2MP Path with LAN

   This section presents the issue about receiving duplicated packets at
   BFER H for any explicit P2MP path/tree from BFIR A to BFERs K, H and
   F in Figure 1 with minimum height of the tree, which is 3 (hops).
   Any path will go through the LAN in order to reach BFER H.

   There are only following explicit P2MP paths from A to K, H and D.

```
           K                                    K
          /14'          H                      /14'
         /           |___/17'                 /                    H
        G---------|                          G            |____/17'
       /4'          LAN                      /4'         _|
    2'  /       6'                        2'  /      6'  / |
  A--------B---------C                   A--------B---------C   LAN
                      \                                       \
                       \12'                                    \12'
                        F                                       F

Path 1 from A to K, H and F          Path 2 from A to K, H and F
```
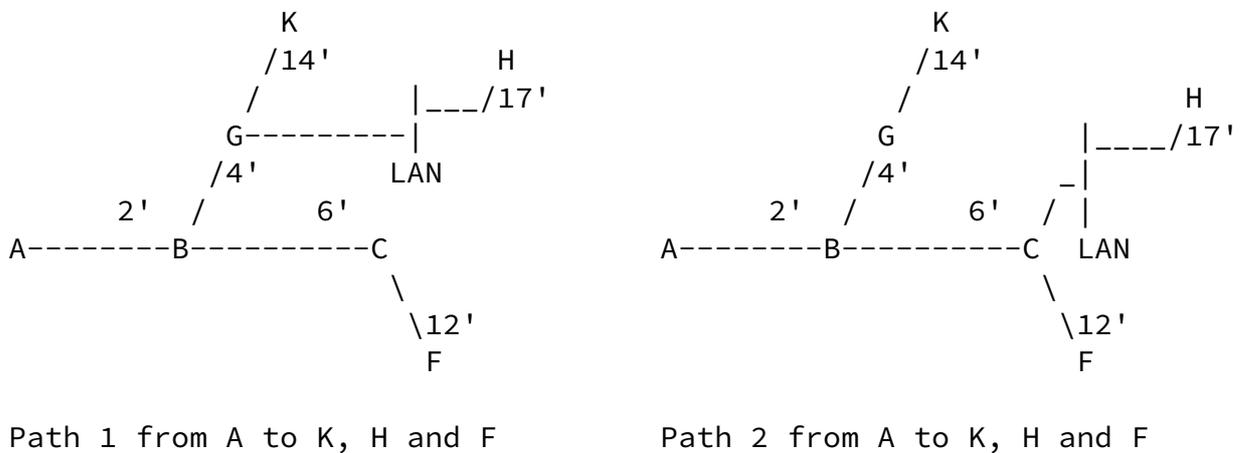
     Figure 10: All explicit paths with height 3 from A to K,H and F

   Path 1 and path 2 are represented by {2', 4', 6', 12', 14', 17', 2,
   4, 6}. Path 1 traverses the link/adjacency from A to B (indicated by
   BP 2'), the link/adjacency from B to G (indicated by BP 4'), the
   link/adjacency from B to C (indicated by BP 6'), the link/adjacency
   from G to K (indicated by BP 14'), the link/adjacency from G to H
   (indicated by BP 17') [the link/adjacency from C to H (also indicated
   by BP 17') for Path 2], and the link/adjacency from C to F (indicated
   by BP 12'). Path 1 is represented by {2', 4', 6', 12', 14', 17', 2,
   4, 6}. Path 2 has the same representation. The packet at A has this
   representation.

   For the packet with the P2MP path, A forwards the packet to B
   according to the forwarding entry for BP 2' in its BIFT.

After receiving the packet from A, B forwards the packet to G and C
according to the forwarding entries for BPs 4' and 6' in B's BIFT
respectively. The packet received by G has path {12', 14', 17', 2,
4, 6}. The packet received by C has path {12', 14', 17', 2, 4, 6}.

After receiving the packet from B, G sends a copy of the packet to K
according to the forwarding entry for BP 14' in G's BIFT, and another
copy of the packet to H according to the forwarding entry for BP 17'
in G's BIFT.

After receiving the packet from B, C copies and sends the packet to H
and F according to the forwarding entries for BPs 17' and 12' in C's
BIFT respectively.

Egress node H of the P2MP path receives the duplicated packets. One
packet is from G, and the same copy is from C.

The solution proposed for LANs in this document resolve this issue.
For a packet with an explicit P2MP path traversing LANs (i.e.,
broadcast links), the packet is replicated and forwarded statelessly
along the path. Each of the egress nodes of the path will not
receive any duplicated packet.

3. Improved BIER-TE with LAN

3.1. New BP Assignments for LAN

For all the nodes/BFRs attached to a LAN (i.e., broadcast link), it
is assumed that they are connected a pseudo node. In one
implementation, the pseudo node is the Designated Router (DR) of the
LAN in OSPF or the Designated Intermediate System (DIS) of the LAN in
IS-IS.

For the connection between the pseudo node and each of the nodes/BFRs
attached to a LAN, two BPs are assigned to it. One is for the
adjacency from the BFR to the pseudo node, the other is for the
adjacency from the pseudo node to the BFR.

The adjacency from a BFR to the pseudo node is called a LAN
adjacency. The adjacency from the pseudo node to a BFR is a forward
connected adjacency.

For example, suppose that the pseudo node for the LAN in Figure 1 is
Px. The BP assignments for the LAN (i.e., connections between Px and
BFRs C, G, H and D) are illustrated in Figure 11.

```
                                     6
                                   ( K )
                                   /14'
                                   /
                         4'     /13'          |             4
                       /----------( G )---------+         ( H )
                      /                17'      |  _____/
                     /                          18'| /20'     19'
                   /3'                  _____ Px
     5   1'        2'     /        6'     /15'   16' |22'   21'
   ( A )-----------( B )-------------( C )         +--------( D )
     \10'          7'\     5'        \11'                    1
      \               \                \
       \               \                \
        \           9'   \8'             \12'
         \-------------( E )            ( F )
                         3                2
```

Chen, et al.   Expires 28 August 2022   [Page 12]
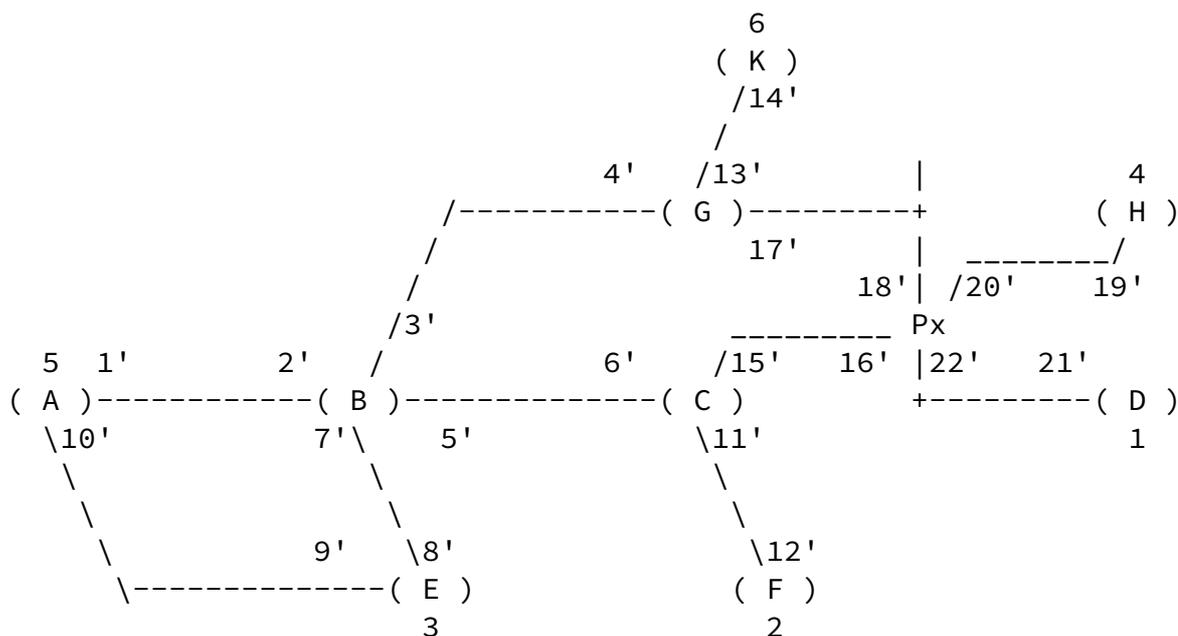

            Figure 11: Example BIER-TE Topology with BPs for LAN

   The connection/adjacency from Px to C is assigned BP 15', and the
   connection/adjacency from C to Px is assigned BP 16'.

   The connection/adjacency from Px to G is assigned BP 17', and the
   connection/adjacency from G to Px is assigned BP 18'.

   The connection/adjacency from Px to H is assigned BP 19', and the
   connection/adjacency from H to Px is assigned BP 20'.

   The connection/adjacency from Px to D is assigned BP 21', and the
   connection/adjacency from D to Px is assigned BP 22'.

   In an alternative, all the nodes/BFRs attached to a LAN are assumed
   fully connected each other (i.e., they are fully meshed). For a
   connection between any two BFRs on the LAN, two forward connected
   adjacencies are assigned to the two ends of the connection.

For example, there are four BFRs C, G, H and D attached to the LAN in Figure 1. There are six connections among these four BFRs. They are connections between C and G, C and H, C and D, G and H, G and D, H and D. Twelve BPs are needed for these six connections.

In general, for n BFRs attached to a LAN, there are n*(n-1)/2 connections among these n BFRs and n*(n-1) BPs are needed for these connections. This may not be scalable. But for this alternative, the BIER-TE BIFT on a BFR needs not to be changed except for considering the full mesh connections among the BFRs attached to a LAN.

3.2. Improved BIER-TE BIFT on BFR

Each BFR in a BIER-TE domain has a BIER-TE BIFT. For a BFR not attached to any LAN, the BIER-TE BIFT on the BFR is the same as before. For a BFR attached to a LAN, its BIER-TE BIFT is changed for considering the LAN.

For example, BFRs C, G, H and D are attached to a LAN in Figure 1. The BIER-TE BIFT on each of these four BFRs is changed for the new BP assignments for the LAN in Figure 11.

For a BFR attached to a LAN, suppose that the pseudo node for the LAN is Px. The improved BIER-TE BIFT on the BFR comprises a forwarding entry for the LAN adjacency from the BFR to Px and a secondary BIFT for Px. The secondary BIFT for Px on the BFR contains a forwarding entry for each of the forward connected adjacencies from Px to the BFRs attached to the LAN except for the adjacency from Px to the BFR.

For example, the improved BIER-TE BIFT on BFR C is illustrated in Figure 12. It comprises the forwarding entry for the LAN adjacency from C to Px (indicated by BP 16') and the secondary BIFT for Px on BFR C. The secondary BIFT contains three forwarding entries for three forward connected adjacencies from Px to G (indicated by 17'), Px to H (indicated by 19') and Px to D (indicated by 21') respectively.

```
        +-----------------+--------------+------------+
        | Adjacency BP    |    Action    |  BFR-NBR   |
```

```
| (SI:BitString) |                |  (Next Hop)  |
+================+================+==============+
| 17'(8:00000001)| fw-connected   |      G       |
+----------------+----------------+--------------+
| 19'(8:00000100)| fw-connected   |      H       |
+----------------+----------------+--------------+
| 21'(8:00010000)| fw-connected   |      D       |
+----------------+----------------+--------------+
          Secondary BIFT for Px on BFR C


+----------------+--------------+------------+
|  Adjacency BP  |    Action    |  BFR-NBR   |
| (SI:BitString) |              | (Next Hop) |
+================+==============+============+
|  5'(6:00010000)| fw-connected |     B      |
+----------------+--------------+------------+
| 12'(7:00001000)| fw-connected |     F      |
+----------------+--------------+------------+
| 16'(7:10000000)|lan-connected |     Px     |
+----------------+--------------+------------+
```

Figure 12: Improved BIER-TE BIFT on BFR C

The improved BIER-TE BIFT on BFR G is illustrated in Figure 13. It
comprises the forwarding entry for the LAN adjacency from G to Px
(indicated by BP 18') and the secondary BIFT for Px on BFR G. The
secondary BIFT contains three forwarding entries for three forward
connected adjacencies from Px to C (indicated by 15'), Px to H
(indicated by 19') and Px to D (indicated by 21') respectively.

```
+------------------+--------------+------------+
|  Adjacency BP    |    Action    |  BFR-NBR   |
| (SI:BitString)   |              | (Next Hop) |
+==================+==============+============+
```

```
| 15'(7:01000000) | fw-connected |     C      |
+----------------+-------------+-----------+
| 19'(8:00000100) | fw-connected |     H      |
+----------------+-------------+-----------+
| 21'(8:00010000) | fw-connected |     D      |
+----------------+-------------+-----------+
          Secondary BIFT for Px on BFR G


+----------------+-------------+-----------+
|  Adjacency BP  |   Action    |  BFR-NBR  |
|  (SI:BitString) |             | (Next Hop) |
+================+=============+===========+
|  3'(6:00000100) | fw-connected |     B      |
+----------------+-------------+-----------+
| 14'(7:00100000) | fw-connected |     K      |
+----------------+-------------+-----------+
| 18'(8:00000010) |lan-connected |     Px     |
+----------------+-------------+-----------+
```

Figure 13: Improved BIER-TE BIFT on BFR G

The improved BIER-TE BIFT on BFR H is illustrated in Figure 14. It
comprises the forwarding entry for the LAN adjacency from H to Px
(indicated by BP 20') and the secondary BIFT for Px on BFR H. The
secondary BIFT contains three forwarding entries for three forward
connected adjacencies from Px to C (indicated by 15'), Px to G
(indicated by 17') and Px to D (indicated by 21') respectively.

```
+-----------------+-------------+-----------+
| Adjacency BP    |   Action    | BFR-NBR   |
| (SI:BitString)  |             | (Next Hop)|
+=================+=============+===========+
| 15'(7:01000000) | fw-connected|     C     |
+-----------------+-------------+-----------+
| 17'(8:00000001) | fw-connected|     G     |
+-----------------+-------------+-----------+
| 21'(8:00010000) | fw-connected|     D     |
+-----------------+-------------+-----------+
          Secondary BIFT for Px on BFR H

+-----------------+-------------+-----------+
| Adjacency BP    |   Action    | BFR-NBR   |
| (SI:BitString)  |             | (Next Hop)|
+=================+=============+===========+
|  4 (0:00001000) | local-decap |           |
+-----------------+-------------+-----------+
| 20'(8:00001000) |lan-connected|    Px     |
+-----------------+-------------+-----------+
```

Figure 14: Improved BIER-TE BIFT on BFR H

The improved BIER-TE BIFT on BFR D is illustrated in Figure 15. It
comprises the forwarding entry for the LAN adjacency from D to Px
(indicated by BP 22') and the secondary BIFT for Px on BFR D. The
secondary BIFT contains three forwarding entries for three forward
connected adjacencies from Px to C (indicated by 15'), Px to G
(indicated by 17') and Px to H (indicated by 19') respectively.

```
+----------------+-------------+-----------+
| Adjacency BP   |   Action    | BFR-NBR   |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
| 15'(7:01000000)| fw-connected|     C     |
+----------------+-------------+-----------+
| 17'(8:00000001)| fw-connected|     G     |
+----------------+-------------+-----------+
| 19'(8:00000100)| fw-connected|     H     |
+----------------+-------------+-----------+
        Secondary BIFT for Px on BFR D

+----------------+-------------+-----------+
| Adjacency BP   |   Action    | BFR-NBR   |
| (SI:BitString) |             | (Next Hop)|
+================+=============+===========+
|  1 (0:00000001)| local-decap |           |
+----------------+-------------+-----------+
| 22'(8:00000100)|lan-connected|    Px     |
+----------------+-------------+-----------+
```

Figure 15: Improved BIER-TE BIFT on BFR D

3.3. Updated Forwarding Procedure

The forwarding procedure defined in [I-D.ietf-bier-te-arch] is
updated/enhanced for using an improved BIER-TE BIFT to support BIER-
TE with LAN.

The updated procedure is described in Figure 16. For a multicast
packet containing the BitString encoding an explicit P2MP path, if
the BP in the BitString is for a LAN adjacency to pseudo node Px for
the LAN, the updated forwarding procedure on a BFR sends the packet
towards Px's next hop nodes on the P2MP path encoded in the packet.

The procedure on a BFR "sends" (i.e., works as sending) the packet
with the BP for the LAN adjacency to Px according to the forwarding
entry for the BP in the improved BIER-TE BIFT on the BFR. And then
it acts on Px to "send" (i.e., works as sending) the packet to each
of the Px's next hop nodes that are on the P2MP path using the

secondary BIFT for Px.

It obtains the secondary BIFT for Px on the BFR, clears all the BPs
for the adjacencies of the BFR including the adjacency from the BFR
to Px, copies and sends the packet to each of the Px's next hop nodes
on the P2MP path using the secondary BIFT for Px.

For each Px's next hop node on the P2MP path, which is represented by
BP j in the packet's BitString, it gets the forwarding entry for BP j
from the secondary BIFT for Px, copies the packet, updates the copy's
BitString by clearing all the BPs for Px's adjacencies, and sends the
updated copy to the next hop node according to the forwarding entry.

```
   Packet = the packet received by BFR;
   FOR each BP k (from the rightmost in Packet's BitString) {
      IF BP k is local decap adjacency (or say BP of BFER) {
         Copies Packet, sends the copy to the multicast
         flow overlay and clears bit k in Packet's BitString
      } ELSE IF BP k is forward connected adjacency of the BFR {
         Finds the forwarding entry in the BIER-TE BIFT using BP k,
         Copies Packet, updates the copy's BitString by
         clearing all the BPs for the adjacencies of the BFR,
         and sends the updated copy to BFR-NBR
      } ELSE IF BP k is LAN adjacency to Px {
         Obtains the secondary BIFT for Px,
         Clears all the BPs for the adjacencies of the BFR,
         FOR each BP j (from the rightmost in Packet's BitString) {
            IF BP j is Px's forward connected adjacency {
               Gets the forwarding entry for BP j in the
               secondary BIFT for Px,
               Copies Packet, updates the copy's BitString by
               clearing all the BPs for Px's adjacencies,
               and sends the updated copy to BFR-NBR
            }
         }
      }
   }
```

                Figure 16: Updated Forwarding Procedure

4. Example Application of Improved BIER-TE

   This section illustrates an example application of improved BIER-TE
   to Figure 1. It shows the forwarding behaviors along an explicit
   P2MP path in Figure 11 going through the LAN in the figure.

   The new BP assignments for the LAN in Figure 1 is shown in Figure 11.
   The improved BIER-TE BIFT on each of the BFRs attached to the LAN is
   given in Section 3.2.

   The explicit P2MP path traverses the link/adjacency from A to B
   (indicated by BP 2'), the link/adjacency from B to G (indicated by BP
   4') and the link/adjacency from B to C (indicated by BP 6'), the
   link/adjacency from G to K (indicated by BP 14'), the link/adjacency

   from G to Px (indicated by BP 18'), the link/adjacency from C to F
   (indicated by BP 12'), and the link/adjacency from Px to H (indicated
   by BP 19'). This path is represented by {2', 4', 6', 12', 14', 18',
   19', 2, 4, 6}. The packet at A has this path.

   For the packet with the P2MP path, A forwards the packet to B
   according to the forwarding entry for BP 2' in its BIFT.

   After receiving the packet from A, B forwards the packet to G and C
   according to the forwarding entries for BPs 4' and 6' in B's BIFT
   respectively. The packet received by G has path {12', 14', 18', 19',
   2, 4, 6}. The packet received by C has path {12', 14', 18', 19', 2,
   4, 6}.

   After receiving the packet from B, G sends a copy of the packet to K
   according to the forwarding entry for BP 14' in G's improved BIER-TE
   BIFT and "sends" another copy of the packet to Px according to the
   forwarding entry for BP 18' in G's improved BIER-TE BIFT. After
   receiving the packet from G, which has path {12', 19', 2, 4, 6}, Px
   "sends" the packet to H according to the forwarding entry for BP 19'
   in the secondary BIFT for Px (a part of G's improved BIER-TE BIFT).

   After receiving the packet from G, which has path {12', 19', 2, 4,
   6}, K decapsulates the packet and passes a copy of the payload of the
   packet to the packet's NextProto within node K according to the
   forwarding entry for BP 6 in K's BIFT.

After receiving the packet from G, which has path {12', 2, 4, 6}, H
decapsulates the packet and passes a copy of the payload of the
packet to the packet's NextProto within node H according to the
forwarding entry for BP 4 in H's improved BIER-TE BIFT.

After receiving the packet from B, which has path {12', 14', 18',
19', 2, 4, 6}, C sends the packet to F according to the forwarding
entry for BP 12' in C's improved BIER-TE BIFT.

After receiving the packet from C, which has path {14', 18', 19', 2,
4, 6}, F decapsulates the packet and passes a copy of the payload of
the packet to the packet's NextProto within node F according to the
forwarding entry for BP 2 in F's BIER-TE BIFT.

Egress node H of the P2MP path does not receive any duplicated
packet.

5. Security Considerations

TBD.

6. IANA Considerations

No requirements for IANA.

7. Acknowledgements

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

8. References

8.1. Normative References

[I-D.ietf-bier-te-arch]
          Eckert, T., Menth, M., and G. Cauchie, "Tree Engineering
          for Bit Index Explicit Replication (BIER-TE)", Work in
          Progress, Internet-Draft, draft-ietf-bier-te-arch-12, 28
          January 2022, <https://www.ietf.org/archive/id/draft-ietf-
          bier-te-arch-12.txt>.

   [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>.

8.2. Informative References

   [I-D.eckert-bier-te-frr]
             Eckert, T., Cauchie, G., Braun, W., and M. Menth,
             "Protection Methods for BIER-TE", Work in Progress,
             Internet-Draft, draft-eckert-bier-te-frr-03, 5 March 2018,
             <https://www.ietf.org/archive/id/draft-eckert-bier-te-frr-
             03.txt>.

   [I-D.ietf-rtgwg-segment-routing-ti-lfa]
             Litkowski, S., Bashandy, A., Filsfils, C., Francois, P.,
             Decraene, B., and D. Voyer, "Topology Independent Fast
             Reroute using Segment Routing", Work in Progress,

             Internet-Draft, draft-ietf-rtgwg-segment-routing-ti-lfa-
             08, 21 January 2022, <https://www.ietf.org/archive/id/
             draft-ietf-rtgwg-segment-routing-ti-lfa-08.txt>.

   [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", Work in Progress, Internet-Draft, draft-ietf-spring-segment-protection-sr-te-paths-02, 21 January 2022, <https://www.ietf.org/archive/id/draft-ietf-spring-segment-protection-sr-te-paths-02.txt>.

[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,
United States of America
Email: Huaimo.chen@futurewei.com

Mike McBride
Futurewei
Email: michael.mcbride@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
United States of America
Phone: 301 502-1347
Email: gyan.s.mishra@verizon.com

Lei Liu
Fujitsu
United States of America
Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks
McLean, VA
United States of America
Email: xufeng.liu.ietf@gmail.com