Path Autogeneration in BIER
draft-zhang-bier-path-autogeneration-01

Abstract

   [I-D.ietf-bier-architecture] BIER introduces a method for multicast
   flow forwarding, without storing states in every node along the
   multicast path. This document introduces a method of establishing
   multicast path automatically.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   [I-D.ietf-bier-architecture] BIER is a technology of multicast flow
   forwarding. [I-D.ietf-bier-mpls-encapsulation] introduces a way to
   use mpls in BIER forwarding. This document introduces a method of
   establishing multicast path automatically.

   Upstream-assigned label is used in this document. Associate with the
   label indicated the BFIR, the two labels compose the keywords for
   multicast flow forwarding. Every node along the multicast path used
   the two labels to forward multicast flow without label changing.
   Obviously, the nodes along the path establish the mpls forwarding
   table, when they receive the packet which include the label
   combination and the path specification.

   BFIR sends the first packet with label combination and path
   specification, every node along the path build the mpls forwarding
   table, and forward the packet to next node according to the path
   specification. If there is already one mpls forwarding item which
   has the same label combination, the node combines the next-hop in the
   packet to the existed forwarding item.

2. Packet Formats

   This document introduces a new TLV that is carried by the BIER
   packet, and this TLV is composed by the nodes in the multicast path.
   There is a flag in the BIER header indicates that a path TLV is be
   carried. One of the reserved flag may be used to signal this.
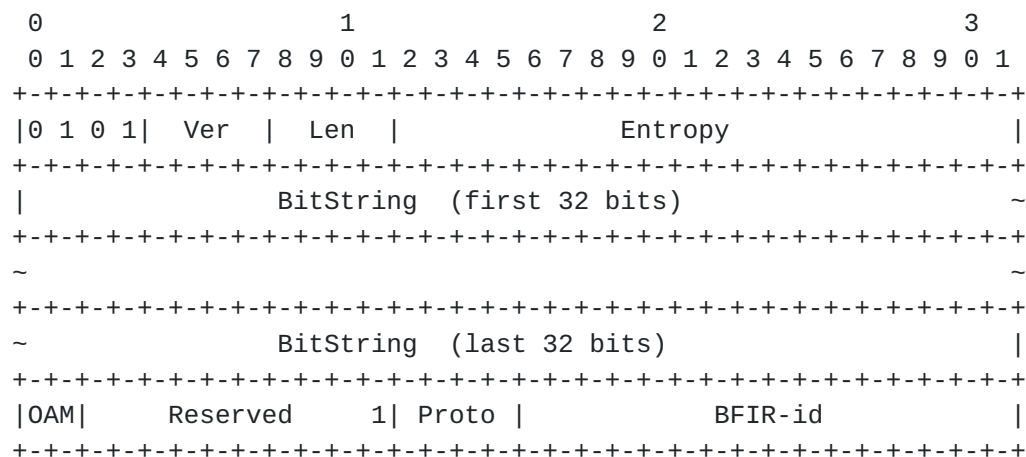
```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |0 1 0 1|  Ver  |  Len  |             Entropy                   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               BitString  (first 32 bits)                     ~
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 ~                                                              ~
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 ~               BitString  (last 32 bits)                      |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |OAM|     Reserved    1| Proto |           BFIR-id             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                     Figure 1 BIER Header
```

2.1. Node function

The path list is composed by nodes along the multicast path. The
multicast path is decided by BFIR in advance through PCE or other
calculations or configurations.

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          Type             |             Length               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                        BFIR label                            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                   Upstream-assigned label                    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          Id               |             Id                   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 ~                                                              ~
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 ~          Id               |           BFER-id                |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   Figure 2 Path Specification
```

o Type TBD, indicate that there is a path TLV.

o Length The length of the TLV.

o BFIR label The label of BFIR.

o Upstream-assigned label The label that is assigned by BFIR for the
   specific multicast flow.

o Id The node BFR-id or the link id along the multicast path. They
   are carried by the packets in order.

o BFER-id The BFR-id of BFER.

Like the ingress replication, BFIR repeats BIER packets several times
according to every BFER of the multicast flow. And then BFIR sends
the BIER packets with the two labels.

The withdraw packet is the same format with the path packet. But the
type should be another type, and the last node that is carried by the
withdraw packet is the canceled node.

2.2. TE function

Also, the method of path auto-generation can be used in BIER-TE
forwarding.

The format of TLV that is carried by BIER header is same as the
previous section. And the difference is that the BitString is BIER-
TE forwarding BitString. The packet is forwarding by the BIER-TE
Forwarding Table.

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |            Type               |            Length             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         BFIR label                           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     Upstream-assigned label                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |              BitString  (first 32 bits)                      ~
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   ~                                                              ~
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   ~              BitString  (last 32 bits)                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
               Figure 3 TE path specification
```
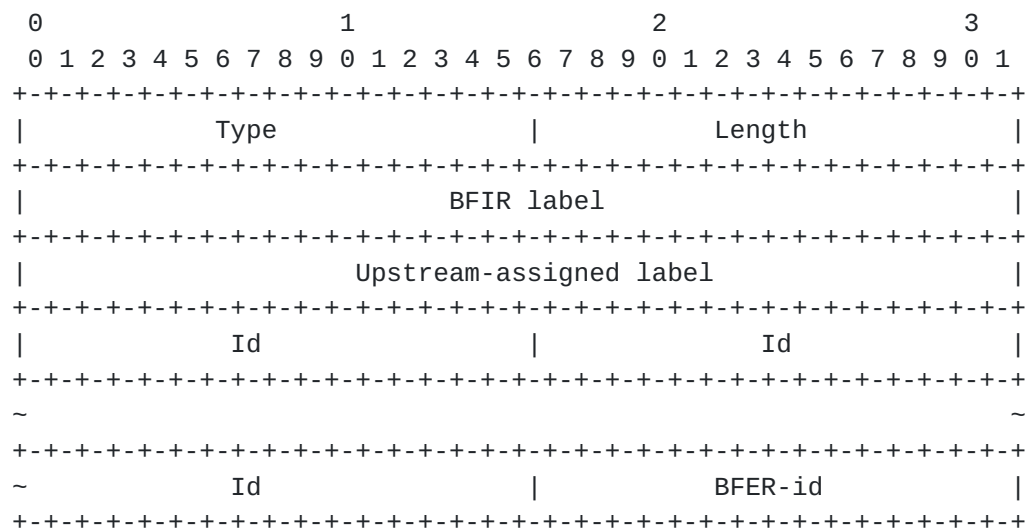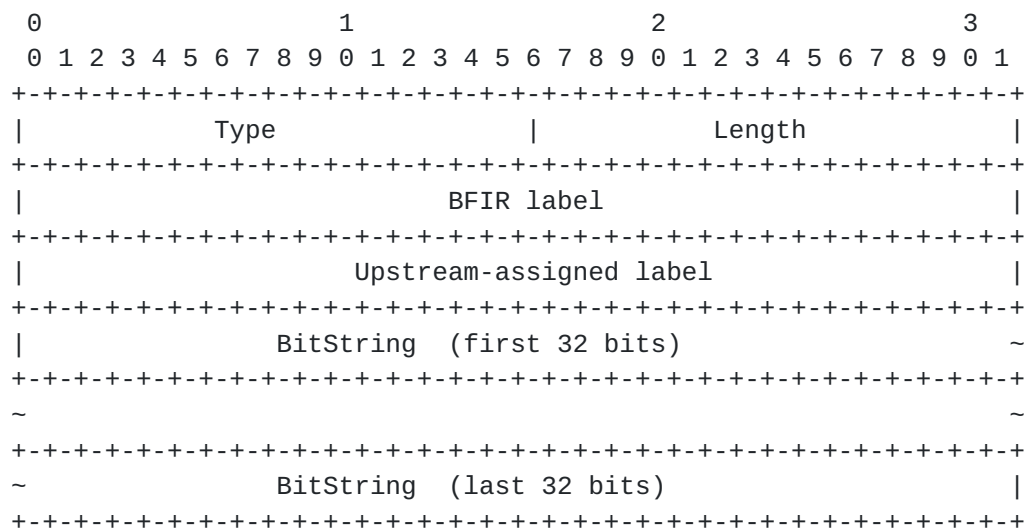
o Type TBD, indicate that there is a TE path TLV.

o Length The length of the TLV.

o BFIR label The label of BFIR.

o Upstream-assigned label The label that is assigned by BFIR for the
   specific multicast flow.

o BitString The adjacency of TE path, and it is the same as the
   BitString in the BIER header.

When the multicast flow travels the BIER-TE path by BIER-TE
forwarding, the two level labels also can be established in every
BFR. If some adjacency should be withdrawn from the path, the type
will be another type of TE path withdraw TLV.

3. Procedures

3.1. Node procedure

3.1.1. Sending the path packets

When BFIR gains the BFER information of a specific multicast flow,
BFIR encapsulates the receiving flow with the path list, and sends to
every BFER individually.Particularly, the encapsulated packet that is
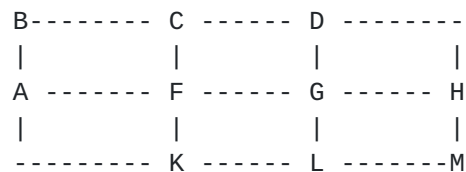sent by BFIR includes the label combination.

```
        B-------- C ------ D --------
        |         |        |         |
        A ------- F ------ G ------ H
        |         |        |         |
        --------- K ------ L -------M
               Figure4 An Example
```

For example, in figure 3, A is BFIR for a specific multicast flow.
And D, H, M is BFER. According the PCE calculation, the multicast
flow should be sent along these paths: A-F-G-D/H, A-K-L-M. So A
encapsulates the flow with the two label combination and makes three
copies, and then sends to D/H/M separately. In node A, there is one
mpls label forwarding item which the ingress label is the two labels
combination, and the next-hops are F and K with the same two labels
out. A encapsulates the formal multicast flow with the two labels
and forwards to next hops.

For the stability consideration, BFIR sends the path packets every
once in a while. The interval may be 30 minutes. And when new BFER
joined, BFIR sends the path packets to the new BFER immediately.

3.1.2. Withdraw Nodes

According to the PCE calculation, some existed nodes in the path may
be canceled. BFIR sends the withdraw path packet which the last node
is the canceled node.

For example in figure 1, if the bandwidth in the network is changed,
the multicast flow should be forwarded along these ways: A-B-C-D-H,

A-K-L-M. The specific multicast flow will not pass by the node F and
G anymore. Except A sending a new path packet along the path
A-B-C-D-H, A also sends two withdraw path packets to node F and G.

3.1.3. BFR and BFER

When BFR and BFER receive the path packets, they build an item in the
mpls forwarding table according to the label combination that is
carried by the packets. If there is an item already in the
forwarding table, the nodes should add the next-hop which is the next
node in the path packet.

When BFR and BFER receive the withdraw path packets which indicate
that they should be canceled, the BFR and BFER delete the mpls
forwarding item which line with the label combination. If the middle
node finds that the next node that along the path is the canceled
node, the middle node deletes the next-hop in the mpls forwarding
item. If there is no next-hop in the mpls forwarding item, the node
deletes the mpls item in the mpls forwarding table.

3.2. BIER-TE procedure

3.2.1. Sending the path packets

When BFIR gains the BFER information of a specific multicast flow,
and BFIR knows all the adjacencies that the flow should travel, BFIR
encapsulates the receiving flow with the BIER-TE adjacencies, and
sends to BIER domain. If there are several SI should be
encapsulated, the flow will be encapsulated by different BitString
several times.

```
          B-p1----p2-- C -p3----p4-- D -p5-----------
          |            |             |            |
          p6           p7            p8           p9
          |            |             |            |
           A-p10---p11- F -p12---p13- G -p14----p15- H
          |            |             |            |
          p16          p17           p18          p19
          |            |             |            |
           --------p20- K -p21---p22- L -p23----p24--M
                   Figure 5 BIER-TE example
```

For example, in figure 4, A is BFIR for a specific multicast flow.
And D, H, M is BFER. According the PCE calculation, the multicast
flow should be sent along these paths: A-F-G-D/H, A-K-L-M. The
BitPositions in BitString are p8, p10, p12, p14, p16, p21, p23. A
encapsulates the flow with the two label combination and the BIER-TE
BitString.

Like the node function, when the packet goes through the path of
BIER-TE, every BFR establishes the label forwarding item.

3.2.2. Withdraw BIER-TE adjacencies

If some adjacencies in the existed path should be withdraw, the
ingress node send the packet with withdraw type of TLV. The
BitString in BIER header is previous TE path adjacency, and the
BitString in the TLV is composed of withdraw adjacencies. When BFR
receives the packet, BFR will withdraw the associated adjacencies
itself.

4. Security Considerations

There should be some common security methods to guarantee the
validation of path packets.

5. IANA considerations

IANA is requested to allocate four types in TLVs of BIER path
packets. Two for node function and two for TE function.

6. Normative References

[I-D.eckert-bier-te-arch]
          Eckert, T. and G. Cauchie, "Traffic Enginering for Bit
          Index Explicit Replication BIER-TE", draft-eckert-bier-te-
          arch-02 (work in progress), October 2015.

[I-D.ietf-bier-architecture]
          Wijnands, I., Rosen, E., Dolganow, A., P, T., and S.
          Aldrin, "Multicast using Bit Index Explicit Replication",
          draft-ietf-bier-architecture-03 (work in progress),
          January 2016.

[I-D.ietf-bier-mpls-encapsulation]
          Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., and
          S. Aldrin, "Encapsulation for Bit Index Explicit
          Replication in MPLS Networks", draft-ietf-bier-mpls-
          encapsulation-03 (work in progress), February 2016.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: zhang.zheng@zte.com.cn

Bo Wu
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: wu.bo@zte.com.cn

Cui(Linda) Wang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: wang.cui1@zte.com.cn