BIER WG   Fangwei Hu
Internet-Draft   Greg Mirsky
Intended status: Standards Track   Quan Xiong
Expires: April 14, 2019   ZTE Corporation
                                                               Chang Liu
                                                            China Unicom
                                                            Oct 11, 2018

BIER BFD
draft-hu-bier-bfd-02.txt

Abstract

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

Table of Contents

1. Introduction

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

[I-D.ietf-bfd-multipoint] defines a method of using Bidirectional Detection(BFD) to monitor and detect unicast failures between the sender (head) and one or more receivers (tails) in multipoint or multicast networks.

This document describes the procedures for using such mode of BFD protocol to verify multipoint or multicast connectivity between a multipoint sender (the "head", Bit-Forwarding Ingress Routers(BFIRs)) and a set of one or more multipoint receivers (the"tails", Bit-Forwarding Egress Routers(BFERs)). The BIER BFD only supports the unidirectional multicast. This document defines the use of BFD, as defined in [I-D.ietf-bfd-multipoint], for BIER domain. Use of BFD for multipoint networks active tail [I-D.ietf-bfd-multipoint-active-tail] is for further study.

2. Conventions used in this document

2.1. Terminology

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

BFD: Bidirectional Forwarding Detection.

OAM: Operations, Administration, and Maintenance.

P2MP: Point to Multi-Point.

2.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. BIER BFD Encapsulation

BIER BFD encapsulation uses the BIER OAM packet format defined in [I-D.ietf-bier-ping]. The value of the Msg Type field MUST be set to BIER BFD (TBD by IANA). BFD Control packet, defined in Section 4 [RFC5880] immediately follows the BIER OAM header.

4. Bootstrapping BIER BFD

The BIER OAM ping could be used for BIER BFD bootstrap. The multipoint header sends the BIER OAM packet with Target SI-Bitstring TLV (section 3.3.2 of [I-D.ietf-bier-ping]) carrying the set of BFER information (Sub-domain-id, Set ID, BS Len, Bitstring) to the multipoint tails to bootstrap the BIER BFD sessions.

5. Discriminators and Packet Demultiplexing

The tail(BFER) demultiplexes incoming BFD packets based on a combination of the source address and My discriminator as specified in [I-D.ietf-bfd-multipoint]. The source address is BFIR-id and BIER MPLS Label(MPLS network) or BFIR-id and BIFT-id(Non-MPLS network)for BIER BFD.

6. Security Considerations

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

For general multipoint BFD security considerations, see [I-D.ietf-bfd-multipoint].

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

7. Acknowledgements

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

8. IANA Considerations

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

Value	Description	Reference
TBD	BIER BFD	[this document]

Table 1

9. References

9.1. Normative References

[I-D.ietf-bfd-multipoint]
Katz, D., Ward, D., Networks, J., and G. Mirsky, "BFD for Multipoint Networks", draft-ietf-bfd-multipoint-18 (work in progress), June 2018.

[I-D.ietf-bier-ping]
Kumar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-03 (work in progress), January 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>.

[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>.

[RFC6213] Hopps, C. and L. Ginsberg, "IS-IS BFD-Enabled TLV", RFC 6213, DOI 10.17487/RFC6213, April 2011, <https://www.rfc-editor.org/info/rfc6213>.

[RFC6328] Eastlake 3rd, D., "IANA Considerations for Network Layer Protocol Identifiers", BCP 164, RFC 6328, DOI 10.17487/RFC6328, July 2011, <https://www.rfc-editor.org/info/rfc6328>.

[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>.

9.2. Informative References

[I-D.ietf-bfd-multipoint-active-tail] Katz, D., Ward, D., Networks, J., and G. Mirsky, "BFD Multipoint Active Tails.", draft-ietf-bfd-multipoint-active-tail-09 (work in progress), June 2018.

[ISO9577] ISO/IEC TR 9577:1999,, "International Organization for Standardization "Information technology - Telecommunications and Information exchange between systems - Protocol identification in the network layer"", 1999.

Authors' Addresses

Fangwei Hu
ZTE Corporation
No.889 Bibo Rd
Shanghai 201203
China

Phone: +86 21 68896273
Email: hu.fangwei@zte.com.cn

Greg Mirsky
ZTE Corporation
USA

Email: gregimirsky@gmail.com

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

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

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

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

INTERNET-DRAFT
Intended Status: Standards Track
Expires: April 26, 2018
R. Huang
N. Xia
N. Wei
Huawei
October 27, 2017

Encapsulation and Extension for BIER-TE
draft-huang-bier-te-encapsulation-extension-00

Abstract

This document proposes to extend the BIER packet format and some BIER-TE forwarding rules specified in BIER traffic engineering architecture.

Status of this Memo

Copyright and License Notice

Table of Contents

1 Introduction

[I-D.eckert-bier-te-arch] specifies BIER-TE: Traffic Engineering for Bit Index Explicit Replication (BIER). It shares part of the architecture with basic BIER as described in [I-D.ietf-bier-architecture], but uses every BitPosition of the BitString of a BIER-TE packet indicates one or more adjacencies instead of a BFER as in BIER.

BIER-TE proposes to share the packet format with BIER. Since it consumes much more BitPositions than BIER, it has scalability issue. For example, the maximum BitString length (BSL) that one BIER-TE packet can carry is 256, which means that one BIER-TE packet cannot pass over 256 numbered adjacencies. This is not a problem in BIER as for BIER all the BitPositions are either BFIRs or BFERs.

To alleviate this issue, one direct way is to allow one packet can travel over more than one Set Indentifier (SI) area. Based on it, this document proposes an encapsulation to solve this issue by extending the BIER packet format specified in [I-D.ietf-bier-mpls-encapsulation] and some BIER-TE forwarding rules in [I-D.eckert-bier-te-arch].

1.1 Terminology

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

2. BIER-TE Extension

2.1. Set Identifier

As described in [I-D.ietf-bier-architecture], Set Identifier (SI) is used to indicate the set of BFERs that one BIER packet can reach. In this document, SI is the segment area index. The number of adjacencies assigned BitPosition inside one segment area is not larger than the value of BSL.

2.2. Packet Travel Rule

As described in [I-D.eckert-bier-te-arch], packets that need to be sent to BFER in different SI require different BIER packets. If a packet travel from one BFIR to the BFERs with different SIs, the path for that packet can only be scheduled for those adjacencies belonging to the same SI carried by the packet, or some adjacencies may be assigned with multiple BitPosition as described in [I-D.xiong-bier-te-forwarding].

In this document, a packet is allowed to travel to multiple areas with different SIs. To do that, multiple bitstrings belonging to different SIs may be carried in the packet header. Considering the overhead of the BIER-TE header, the total length of all the bitstrings that a packet can carry is the maximum BSL 4096. For example, if the BSL is 256, then a packet can pass over at most 16 segment areas. If the topology of the network is well planned, this design is sufficient for use.

If all of the BitPosition in one of the Bitstrings are set to 0, it means that the packet will not travel to this area any more. The BFR could remove that BitString when forwarding the packet to the adjacencies.

2.3. The Bit Index Forwarding Table (BIFT)

The BIFT is used as described in [I-D.eckert-bier-te-arch], which is indexed by SI:BitPosition.

3. BIER-TE Encapsulation

3.1 Header

The BIER-TE encapsulation is illustrated as following. It reuses the format defined in [I-D.ietf-bier-mpls-encapsulation].

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              BIFT-id                  | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Nibble |  Ver  |  BSL  |              Entropy                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|OAM|B|E|    DSCP    |   Proto   |            BFIR-id           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                BitString  (first 32 bits)                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               BitString  (last 32 bits)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Extension Field (variable)                      |
```

B: This 1-bit field identifies a BIER head(0) or a BIER-TE head(1).

E: This 1-bit field identifies weather there is a extension field following the head.

3.2. Header Extension

The header extension is illustrated as following.

```
 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           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 The information of BitPostions                |
~                  in one specific segment area                 ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            ......                             |
```

This header extension uses TSV structure. The information of the BitPositions in one specific segment area (SI) is only carried in one TSV.

Type: 16 bits It defines the type of the extension. In this document, the type of TBD is specified.

Length: 16 bits It indicates the total octets of the information of BitPostions in one specific segment area.

Information of BitPostions in one specific segment area: This contains the information of the BitString in the specific segment area other than the previous header. In this document, the details of this field doesn't specified. It could be several BitStrings belonging to different segment areas populated together or It could be another BIER header nested in this extension. In the latter case, it will be easier for BFRs to pop the unused header.

4. BIER-TE Forwarding Example

Here, the same example from [I-D.eckert-bier-te-arch] is used as following. Assume the BSL is 8 (The BSL of 8 is used only in this example). pXX-Y indicate the BitPosition-SI assigned by the BIER-TE controller host to adjacencies in the BIER-TE topology.

```
               [Bier-Te Controller Host]
                      /    |
                     v     v  v
```

```
           |p5-2 p1-1 |
           +- BFIR2 --+          |
           |          | p2-1 p3-1|           LAN2
           |          +-- BFR3 --+           |
           |          |          | p4-1 p6-1 |
      Src -+                     +-- BFER1 --+
           |          |p7-1  p5-1|           |
           |          +-- BFR4 --+           +-- Rcv1
           |          |          |           |
           |          |
           |p6-2  p8-1|
           +- BFIR1 --+          |
           |          +-- BFR5 --+ p3-2 p4-2 |
         LAN1         |p1-2 p2-2 +-- BFER2 --+
                                 |           +-- Rcv2
                                             |
                                             LAN3

          IP  |..... BIER-TE network......| IP
```

Traffic needs to flow from BFIR2 towards Rcv1, Rcv2. The controller determines it wants to pass across the following paths:

```
                 -> BFER1 ---------------> Rcv1
    BFIR2 -> BFR3
                 -> BFR4 -> BFR5 -> BFER2 -> Rcv2
```

The BitString is set up in BFIR2 with 2 sets of BitStrings: S1:(p2, p4, p5, p6); S2:(p1, p3, p4). BFIR2 forwards based on that BitString.

BFR4 has the following BIFT:
p8-1: forward_connected(BFIR1) P1-2: forward_connected(BFR5)

BFR5 sees the sets of BitStrings: S1: (0...0); S2:(p3, p4). It pops the BitString of S1 and forward the packet out to BFER2.

Other forwarding rules are similar to those specified in [I-D.eckert-bier-te-arch].

5 Security Considerations

TBD

6 IANA Considerations

TBD.

7 References

7.1 Normative References

[KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[I-D.eckert-bier-te-arch] Eckert, T., Cauchie, G., Braun, W., and M. Menth, "Traffic Engineering for Bit Index Explicit Replication BIER-TE", draft-eckert-bier-te-arch-05 (work in progress), June 2017.

[I-D.ietf-bier-architecture] Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-bier-mpls-encapsulation] Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-10 (work in progress), October 2017.

7.2 Informative References

[I-D.zcxh-bier-te-forwarding] Zhu, Y., Chen, H., Xiong, Q., and F. Hu, "BIER-TE Forwarding", draft-zcxh-bier-te-forwarding-00 (work in progress), October 2017.

Authors' Addresses

Rachel Huang
Huawei
101 Software Avenue, Yuhua District
Nanjing 210012
China

EMail: rachel.huang@huawei.com

Nu Xia
Huawei
101 Software Avenue, Yuhua District
Nanjing 210012
China

EMail: xianu@huawei.com

Naiwen Wei
Huawei
China

EMail: weinaiwen@huawei.com

Network Working Group D. Purkayastha
Internet-Draft A. Rahman
Intended status: Informational D. Trossen
Expires: April 30, 2018 InterDigital Communications, LLC
October 27, 2017

Multicast HTTP using BIER
draft-purkayastha-multicast-http-using-bier-00

Abstract

HTTP Level multicast, using BIER, is described as a use case in [I-D.ietf-bier-use-cases]. In order to enable the use case, the document describes additional functions in the ingress and egress nodes to the BIER network. These functions are assumed to be part of the BIER multicast overlay.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

BIER Use Cases document [I-D.ietf-bier-use-cases] describes an "HTTP Level Multicast" scenario, where HTTP-level clients benefit from the dynamic multicast group formation enabled by BIER. HTTP multicast means aggregating individual HTTP Responses (e.g. for the same segment of a viral video) and mapping it onto the BIER multicast overlay. For this, the server NAP (network attachment point), creates a list of outstanding client NAP requests to the same HTTP request URI. When a response is available, BIER forwarding information is retrieved and used to send the HTTP response.

In this draft, we introduce the requirements for a BIER multicast overlay realizing this use case. It also describes the necessary functions that form the BIER multicast overlay and the operations that enable the desired "HTTP Level Multicast" behavior. We describe a list of protocol changes needed for the realization of the individual operations.

2. Requirements

A realization for the "HTTP multicast" use case may have the following requirements:

- o MUST support multiple FQDN-based service endpoints to exist in the overlay

- o MUST send FQDN-based service requests at the network level to a suitable FQDN-based service endpoint via policy-based selection of appropriate path information

- o MUST allow for multicast delivery of HTTP response to same HTTP request URI o MUST provide direct path mobility, where the path between the egress and ingress NAPs can be determined as being optimal (e.g., shortest path or direct path to a selected instance), is needed to avoid the use of anchor points and further reduce service-level latency

3. HTTP Multicast Overlay Components

Let us formulate the architecture of the BIER multicast overlay for the scenario outlined in [I-D.ietf-bier-use-cases]. This overlay is shown in Figure 1 below.

The multicast overlay is formed by the BFIR and BFER of the BIER layer and the additional NAP and PCE elements shown in the figure. When connecting to a standard IP routed peering network, a special NAP is utilized, shown as the border GW in the figure.

```
   +---------+   +----------+
   |         |   |          |
   +IP only  +---+ NAP      +---------|
   |reciever |   |          |         |
   |UE       |   +----------+         |
   +---------+                        |
                                +----------+  +----------+
                                |          |  |          |
                                |  BFER    |--|  BFR     |------|
                                |          |  |          |      |
                                +----------+  +----------+      |
                                                            +--------+
                                                  |-------| BFER   |
                                +---------+   +---|--+      +---|----+
                                |         |---| BFR  |          |
                                |  BFIR   |   +------+      +---------+
                                |         |                 | NAP     |
                                +---------+                 +---------+
   +---------+   +----------+       |                          |
   |         |   |          |       |                          |
   +IP only  +---+ NAP      +-------|                  +-----------------+
   |sender UE|   |          |                          | IP only sender  |
   +---------+   +----------+                          | and reciever    |
                                                       |     UE          |
                                                       +-----------------+
```

Figure 1: BIER Multicast Overlay for HTTP Multicast Use case

4. HTTP Multicast Overlay Operations

As shown in Figure 1, the multicast overlay includes a function called PCE (Path Computation Element function), which is responsible for selecting the correct multicast end point and possibly realizing path policy enforcement. The result of the selection is a BIER path identifier, which is delivered to the NAP upon initial path computation request (i.e., when sending a request to or response for a specific URL for the first time). The path identifier is utilized for any future request for a given URL-based request. All service end points indicate availability to the PCE through a registration procedure, the PCE will instruct all NAPs to invalidate previous path identifiers to the specific URL. This may result in an initial path computation request at the next service request forwarding. Through this, the newly registered service endpoint might be utilized if the policy-governed path computation selects said service instance.

In the architecture of Figure 1, an HTTP request is sent by an IP-based device towards the FQDN of the server defined in the HTTP request.

At the client facing NAP, the HTTP request is terminated at the HTTP level at a local HTTP proxy. We assume termination on the client side at Layer 3 and above protocols, such as TCP. Server NAP at the egress, terminates any transport protocol on the outgoing (server) side. These terminating functions are assumed to be part of the client/server NAP.

If no local BIER forwarding information exists to the server (NAP), a path computation entity (PCE) is consulted, which calculates a unicast path from the BFIR to which the client NAP is connect to the BFER to which the server NAP is connected. The PCE provides the forwarding information to the client NAP, which in turn caches the result.

Ultimately, the HTTP request is forwarded by the client NAP towards the server-facing NAP via the local BFIR. We assume a (TCP-friendly) transport protocol being used for the transmission between client and server NAP while not mandating the use of TCP for this transmission.

Upon arrival of an HTTP request at the server NAP, the server NAP proxy forwards the HTTP request as a well-formed HTTP request locally to the server.

If no BIER forwarding information exists for the reverse direction towards the requesting client NAP, this information is requested from the PCE, similar to the operation in forward direction.

Upon arrival of any further client NAP request at the server NAP to an HTTP request whose response is still outstanding, the client NAP is added to an internal request table. Optionally, the request is suppressed from being sent to the server.

Upon arrival of an HTTP response at the server NAP, the server NAP consults its internal request table for any outstanding HTTP requests to the same request. The server NAP retrieves the stored BIER forwarding information for the reverse direction for all outstanding HTTP requests and determines the path information to all client NAPs through a binary OR over all BIER forwarding identifiers with the same SI field. This newly formed joint BIER multicast response identifier is used to send the HTTP response across the network.

5. Required Protocol Changes

For the operations outlined in the previous section, we foresee the following protocol changes may be required:

- NAP-to-NAP protocol for HTTP: Map HTTP to BIER message exchange between client and server NAPs

- NAP-PCE protocol: Used for path computation and delivery of BIER routing information as well as path updates

- Overlay transport protocol: Used for transport-level exchange over BIER layer

- Registration protocol: Used to register FQDN service endpoints

- Content certificate distribution protocol: Used for HTTPS support

There is a similar ongoing work in SFC WG, which handles HTTP redirection [I-D.purkayastha-sfc-service-indirection]. The lower layers for the NAPs and PCE infrastructure is similar between the two approaches. Does the WG see value in supporting the requirements for BIER to enable HTTP Multicast Use case as defined in [I-D.ietf-bier-use-cases]? It also raises a relevant question, where shall the protocol work be done?

6. IANA Considerations

This document requests no IANA actions.

7. Security Considerations

TBD.

8. Informative References

[I-D.ietf-bier-use-cases]
Kumar, N., Asati, R., Chen, M., Xu, X., Dolganow, A., Przygienda, T., arkadiy.gulko@thomsonreuters.com, a., Robinson, D., Arya, V., and C. Bestler, "BIER Use Cases", draft-ietf-bier-use-cases-05 (work in progress), July 2017.

[I-D.purkayastha-sfc-service-indirection]
Purkayastha, D., Rahman, A., and D. Trossen, "USE CASE FOR HANDLING DYNAMIC CHAINING AND SERVICE INDIRECTION", draft-purkayastha-sfc-service-indirection-00 (work in progress), July 2017.

Authors' Addresses

Debashish Purkayastha
InterDigital Communications, LLC
Conshohocken
USA

Email: Debashish.Purkayastha@InterDigital.com

Akbar Rahman
InterDigital Communications, LLC
Montreal
Canada

Email: Akbar.Rahman@InterDigital.com

Dirk Trossen
InterDigital Communications, LLC
64 Great Eastern Street, 1st Floor
London EC2A 3QR
United Kingdom

Email: Dirk.Trossen@InterDigital.com
URI: http://www.InterDigital.com/

BIER Working Group   IJ. Wijnands
Internet-Draft   Cisco Systems
Intended status: Informational   X. Xu
Expires: August 30, 2018   Alibaba Group
                                                              H. Bidgoli
                                                                   Nokia
                                                       February 26, 2018

An Optional Encoding of the BIFT-id Field in the non-MPLS BIER Encapsulation
draft-wijnandsxu-bier-non-mpls-bift-encoding-01.txt

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. The Multicast packet is encapsulated using a BIER Header and transported through an MPLS or non-MPLS network. When MPLS is used as the transport, the Bit Indexed Forwarding Table (BIFT) is identified by a MPLS Label. When non-MPLS transport is used, the BIFT is identified by a 20bit value. This document describes one way of encoding the 20bit value.

Table of Contents

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. The Multicast packet is encapsulated [RFC8296] using a BIER Header and transported through an MPLS or non-MPLS network. When MPLS is used as the transport, the Bit Indexed Forwarding Table (BIFT) is identified by a MPLS Label. When non-MPLS transport is used, the BIFT is identified by a 20bit value. This document describes one way of encoding the 20bit value, based on the Sub-Domain (SD), Set Identifier (SI) and BitStringLength (BSL) values.

The BIER architecture requires that a BFR has a BIFT for every combination of <SD, SI, BSL> that is being used. When processing a BIER packet, the correct BIFT is inferred from the BIFT-id field of the encapsulation. When the non-MPLS encapsulation is used in a given BIER domain, it may be desirable for the a BIFT-id to be unique in that domain. This document describes an OPTIONAL method that can be used to form domain-wide unique BIFT-ids based on the <SD, SI, BSL> triples. If in the future the BIER architecture is extented with an additonal BIFT argument, this encoding does not generate domain-wide unique idenifiers anymore.

This encoding, if used, is only for the convenience of the network adminstrators. When forwarding a BIER packet, the BIFT-id is used as an opaque 20-bit value that identifies a BIFT; the forwarding procedures do not parse the 20-bit value, they just use it as a lookup key.

2. Terminology and Definitions

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References. For convenience, some of the more frequently used terms appear below.

BIER:
   Bit Indexed Explicit Replication.

BIFT-id:
   Bit Indexed Forwarding Table Identifier.

3. Specification of Requirements

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].

4. The Bit Index Forwarding Table

In MPLS networks a BIER label is allocated for each Bit Index Forwarding Table (BIFT) from the platform specific, downstream label database ([RFC8296]). This label is associated with a particular combination of BIER Sub-Domain (SD), Set Identifier (SI) and BitStringLength (BSL). In order for the network to know which MPLS label represents a particular combination of <SD, SI, BS>, this mapping has to be advertised through the network. This is currently done through an IGP or BGP. In MPLS networks this is not a drawback as the MPLS label has to be advertised anyway.

When the non-MPLS encoding is chosen, there is no need to advertise the BIFT-id to <SD, SI, BSL> mapping if the BIFT-id is domain-wide unique. For this reason we're defining an encoding that MAY be used by operators to compute the domain-wide unique BIFT-id values from the SD, SI and BSL. Although the BIFT-id is not expected to change, it may change when the BSL mismatch procedures [RFC8279] section 6.10.2 are applied.

5. The Non-MPLS Static BIFT Encoding

Find below the first 32 bits of the BIER header, encoding the SD, SI and BSL into the 20 bit BIFT-id field.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  BSL  |       SD      |       SI      |  TC |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

|-------- 20 bit BIFT-id Field ---------|
```

Figure 1

BSL: This 4-bit field encodes the length in bits of the BitString. These are the same values as documented in [RFC8296].

SD: This is a 8-bit field that encodes the Sub-Domain as described in [RFC8279].

SI: This is a 8-bit field that encodes the Set-ID as described in [RFC8279].

TC: This is a 3-bit field set to 000 (following [RFC8296]).

S: This is a 1-bit field set to 1 (following [RFC8296]).

TTL: See [RFC8296].

6. Security Considerations

This document does not introduce any new security considerations other than already discussed in [RFC8279].

7. IANA Considerations

There is no IANA consideration.

8. Acknowledgments

The authors like to thank the following people for their comments and contributions to this document; Eric Rosen, Neale Ranns, Jeffrey Zhang.

9. 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>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <https://www.rfc-editor.org/info/rfc8279>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <https://www.rfc-editor.org/info/rfc8296>.

Authors' Addresses

IJsbrand Wijnands
Cisco Systems
De Kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Xiaohu Xu
Alibaba Group

Email: xiaohu.xxh@alibaba-inc.com

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

Email: hooman.bidgoli@nokia.com

Network Working Group J. Xie
Internet-Draft Huawei
Intended status: Standards Track October 27, 2017
Expires: April 30, 2018

Multicast VPN Using MPLS P2MP and BIER
draft-xie-bier-mvpn-mpls-p2mp-00

Abstract

The Multicast Virtual Private Network (MVPN) specifications defines P-tunnels for carrying multicast traffic across the backbone. A variety of P-tunnel types are supported. Bit Index Explicit Replication (BIER) is a new architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state. The purpose of the current document is to specify one way of carrying multicast traffic over an SP MPLS backbone network using compatible method and encapsulation of BIER. It uses a pre-build P2MP as the BIER topology, and uses mLDP/RSVP-TE protocol extension to build BIER-related underlay routing and forwarding information in-band when building the P2MP topology.

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].

Status of This Memo

Table of Contents

1. Introduction

[RFC6513] and [RFC6514] specify the protocols and procedures that a Service Provider (SP) can use to provide Multicast Virtual Private Network (MVPN) service to its customers. Multicast tunnels are created through an SP's backbone network; these are known as "P-tunnels". The P-tunnels are used for carrying multicast traffic across the backbone. The MVPN specifications allow the use of several different kinds of P-tunnel technology. In an MPLS network, such P-tunnel can be mLDP P2MP or RSVP-TE P2MP.

Bit Index Explicit Replication (BIER) [I-D.ietf-bier-architecture] is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state.

BIER architecture requires routers participating in BIER to exchange BIER related information within a given domain. BIER architecture permits IGP/BGP or any other routing protocols to perform distribution of such information. Such routing protocols are defined as Underlay protocols.

In an MPLS network, [I-D.ietf-bier-mpls-encapsulation] define a BIER Header within it an initial 4 octets MPLS-Label, to encapsulate Multicast packet and transport through the MPLS network.

The purpose of the current document is to specify one way of carrying multicast traffic over an SP MPLS backbone network, using compatible method and encapsulation described in the above BIER documents. It uses a pre-build P2MP as the BIER topology, and uses mLDP/RSVP-TE protocol extension to build BIER-related underlay routing and forwarding information in-band when building the P2MP topology.

2. Terminology

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References. For convenience, some of the more frequently used terms and new terms list below.

- o LSP: Label Switch Path
- o LSR: Label Switching Router
- o P2MP: Point to Multi-point
- o P-tunnel: A multicast tunnel through the network of one or more SPs. P-tunnels are used to transport MVPN multicast data.
- o PMSI: Provider Multicast Service Interface
- o x-PMSI A-D route: a route that is either an I-PMSI A-D route or an S-PMSI A-D route.
- o PTA: PMSI Tunnel attribute. A type of BGP attribute known as the PMSI Tunnel attribute.
- o P2MP LSP based BIER: BIER using P2MP LSP as topology

3. Use of the PTA in MVPN Routes

3.1. Overview

According to [I-D.ietf-bier-architecture], the P2MP LSP based BIER is a REAL BIER, which using a P2MP LSP as the underlay topology. The P2MP LSP is not only a LSP, but also a topology as the BIER underlay. The P2MP LSP based BIER is P-tunnel, which is used for bearing multicast flows. Every flow can think as binding to an independent tunnel, which is constructed by the BitString in the BIER header of every packet of the flow. Multicast flows are transported in SPMSI-only mode, on P2MP LSP based BIER tunnels, and never directly on P2MP LSP tunnel.

3.2. Use of the PTA in x-PMSI A-D Routes

As defined in [RFC6514], the PMSI Tunnel attribute (PTA) carried by an x-PMSI A-D route identifies the P-tunnel that is used to instantiate a particular PMSI. If a PMSI is to be instantiated by P2MP LSP based BIER, the PTA is constructed by a BFIR, which is also a Ingress LSR. This document defines the following Tunnel Types:

+ TBD - RSVP-TE P2MP LSP based BIER

+ TBD - mLDP P2MP LSP based BIER

Allocation is expected from IANA for two new tunnel type codepoints from the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry. These codepoints will be used to indicate that the PMSIs is instantiated by MLDP or RSVP-TE extension with support of BIER.

When the Tunnel Type is set to RSVP-TE P2MP LSP based BIER, the Tunnel Identifier include two parts, as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  BSL  | Tunnel Number |            Must Be Zero               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       P2MP ID                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  MUST be zero                 |       Tunnel Range Base       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Extended Tunnel ID                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: PTA of RSVP-TE P2MP LSP based BIER

BSL: A 4 bits field. The values allowed in this field are specified in section 2 of [I-D.ietf-bier-mpls-encapsulation].

Tunnel Number: A 1 octet field encoding the Number of the Tunnel range. It MUST be greater than 0.

<Extended Tunnel ID, Reserved, Tunnel Range Base, P2MP ID>: A ID as carried in the RSVP-TE P2MP LSP SESSION Object defined in [RFC4875].

The "Tunnel Range" is the set of P2MP LSPs beginning with the Tunnel Range base and ending with ((Tunnel Range base)+(Tunnel Number)- 1). A unique Tunnel Range is allocated for the BSL and a Sub-domain-ID implicited by the P2MP.

The size of the Tunnel Range is determined by the number of Set Identifiers (SI) (section 1 of [I-D.ietf-bier-architecture]) that are used in the topology of the P2MP-LSP. Each SI maps to a single Tunnel in the Tunnel Range. The first Tunnel is for SI=0, the second Tunnel is for SI=1, etc.

When the Tunnel Type is set to mLDP P2MP LSP based BIER, the Tunnel Identifier include two parts, as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  BSL  | Tunnel Number |             Must Be Zero              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|P2MP Type(0x06)|        Address Family         | Address Length|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                       Root Node Address                       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Opaque Length(0x0007)      | OV Type(0x01) |OV Len(High 8b)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| (Low 8b)(0x04)|  Tunnel Range Base(High 24b)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   (Low 8b)    |
+-+-+-+-+-+-+-+-+
```

Figure 2: PTA of MLDP P2MP LSP based BIER

BSL: A 4 bits field. The values allowed in this field are specified in section 2 of [I-D.ietf-bier-mpls-encapsulation].

Tunnel Number: A 1 octet field encoding the Number of the Tunnel range. It MUST be greater than 0.

<Type=0x06, AF, AL, RootNodeAddr, Opqeue Length=0x0007, OV Type=0x01, OV Len=0x04, Tunnel Range Base>: A P2MP Forwarding Equivalence Class (FEC) Element, with a Generic LSP Identifier TLV as the opaque value element, defined in [RFC6388].

The "Tunnel Range" is the set of P2MP LSPs beginning with the Tunnel Range base and ending with ((Tunnel Range base)+(Tunnel Number)- 1). A unique Tunnel Range is allocated for the BSL and a Sub-domain-ID implicited by the P2MP.

The size of the Tunnel Range is determined by the number of Set Identifiers (SI) (section 1 of [I-D.ietf-bier-architecture]) that are used in the topology of the P2MP-LSP. Each SI maps to a single Tunnel in the Tunnel Range. The first Tunnel is for SI=0, the second Tunnel is for SI=1, etc.

When the Tunnel Type is any of the above two, The "MPLS label" field OPTIONAL contain an upstream-assigned non-zero MPLS label. It is assigned by the router (a BFIR) that constructs the PTA. Absence of an MPLS Label is indicated by setting the MPLS Label field to zero.

When the Tunnel Type is any of the above two, two of the flags, LIR and LIR-pF, in the PTA "Flags" field are meaningful. Details about the use of these flags can be found in [RFC6513], [I-D.ietf-bess-mvpn-expl-track] and [I-D.ietf-bier-mvpn]].

3.3. Use of the PTA in Leaf A-D routes

Before an egress PE can receive a (C-S,C-G) flow from a given ingress PE via RSVP-TE/MLDP P2MP LSP based BIER, the egress PE must have received one of the following x-PMSI A-D routes from the ingress PE:

- o A "more specific" x-PMSI A-D route, (C-S,C-G) S-PMSI A-D route.
- o A "less specific" x-PMSI A-D route, (C-*,C-*), (C-*,C-G), or (C-S,C-*) S-PMSI A-D route.

In which, the PTA tunnel Type is "RSVP-TE P2MP LSP based BIER" or "MLDP P2MP LSP based BIER".

The rules for determining which x-PMSI A-D route is the match for reception are given in [RFC6625]. If such a route is found, we refer to it as the "matching x-PMSI A-D route." If no matching x-PMSI A-D route for (C-S,C-G) is found, the egress PE cannot receive the (C-S,C-G) flow from the ingress PE via RSVP-TE/MLDP P2MP LSP based BIER until such time as a matching route is received.

When an egress PE determines that it needs to receive a (C-S,C-G) flow from a particular ingress PE via RSVP-TE/MLDP P2MP LSP based BIER, it originates a Leaf A-D route. Construction of the Leaf A-D route generally follows the procedures specified in [RFC6514], or optionally, the procedures specified in [I-D.ietf-bess-mvpn-expl-track]. However, when RSVP-TE/MLDP P2MP LSP based BIER is being used, the Leaf A-D route MUST carry a PTA that is constructed as follows:

1. The tunnel type MUST be set to RSVP-TE/MLDP P2MP LSP based BIER, corresponding to the PTA of the matching x-PMSI A-D route.

2. The MPLS label field SHOULD be set to zero.

3. The BFR-Prefix field of the Tunnel Identifier field MUST be set to the egress PE's IP-Address. This IP-Address is the same as the Originating Router's IP Addr field of the NLRI of the Leaf A-D route.

When an ingress PE receives such a Leaf A-D route, it learns the BFR-Prefix of the egress PE from the PTA. The ingress PE does not make any use the value of the PTA's MPLS label field.

Failure to properly construct the PTA cannot always be detected by the protocol, and will cause improper delivery of the data packets.

4. P2MP LSP based BIER Forwarding Procedures

The MVPN application plays the role of the "multicast flow overlay" as described in [I-D.ietf-bier-architecture].

This section specifies some OPTIONAL rules for forwarding a BIER-encapsulated data packet within a P2MP topology underlay.

These rules will produce the same results as the procedures in [I-D.ietf-bier-architecture], on condition that the underlay topology is a P2MP.

4.1. Overview

As [I-D.ietf-bier-architecture] describes:

1. BIER support using the default topology of the unicast IGP as the routing underlay. To quote from [I-D.ietf-bier-architecture]: "By default, each sub-domain uses the default topology of the unicast IGP as the routing underlay."

2. BIER also support using other topologies as the routing underlay, including a tree topology. To quote from [I-D.ietf-bier-architecture]: "Alternatively, one could deploy a routing underlay that creates a multicast-specific tree of some sort. Then BIER could be used to forward multicast data packets along the multicast-specific tree, while unicast packets follow the 'ordinary' OSPF best path."

This document specifies one OPTIONAL Forwarding Procedure of BIER encapsulation packet, on the condition that the BIER underlay topology is P2MP LSP, as describes in the above sections. Comparing to the Forwarding Procedure, which is described in [I-D.ietf-bier-architecture], and which is on a underlay of unicast IGP topology, there is some simplification:

1. Not need to Edit the BitString when forwarding packet to Neighbor, for the underlay P2MP topology is already loop-free.

2. Not need to use Entropy in the BIER Header, for current P2MP topology is already ECMP-eliminate.

The optional BIER forwarding procedure is, on the basis of P2MP forwarding procedure according to the BIER-MPLS label, and use the BitString to prune the undesired P2MP downstream.

The enhancement to the P2MP forwarding is to add a Forwarding BitMask to existing NHLFE defined in [RFC3031], for checking with the BitString in a packet, to determin whether the packet is to be forwarded or pruned. If the checking result by AND'ing a packet's BitString with the F-BM of the NHLFE (i.e., Packet->BitString &= F-BM) is non-zero, then forward the packet to the next-hop indicated by the NHLFE entry, and the Label is switched to the proper one in the NHLFE. If the result is zero, then do not forward the packet to the next-hop indicated by the NHLFE entry.

4.2. Building P2MP LSP based BIER forwarding state

When RSVP-TE/MLDP P2MP LSP based BIER are used, then it is not nessary to use IGP or BGP to build the BIER routing table and forwarding table. Instead, the BIER layer information is carried by MLDP or RSVP-TE, and when MLDP or RSVP-TE build P2MP LSP, it build the BIER forwarding state in-band.

The procedure for building RSVP-TE/MLDP P2MP LSP based BIER forwarding state using mLDP or RSVP-TE is outside the scope of this document.

4.3. Live-Live protection

As described above, loop and redundancy, ECMP and Entropy, are all not supported in current P2MP LSP underlay. There will be extra P2MP LSP convergence, after IGP convergence, in the case of link or node failure.

On the other hand, Multicast has special Service Level Aggrement(SLA), especially when multicast service is compressed or uncompressed video. Accordingly, there are some multicast-specific methods of protection, such as Live-Live. [RFC7431] defines a method of detecting failure locally by comparing the packets received from live-live paths. [I-D.ietf-bess-mvpn-fast-failover] defines a Live-Live method for protecting Multicast in MVPN.

This document specifies one OPTIONAL extension to enhance Live-Live protection, re-using the Entropy field of BIER header as a Sequence number of multicast packet, on the condition that the field is not used for ECMP, such as in the P2MP LSP topology described above.

This is an optional function of BIER Layer. If this function is enabled, every BFR of the domain is required to support, which means:

1. The BFIR (and Ingress LSR) will push a sequence-number in the Entropy field, per-flow per-packet.

2. The middle BFR will ignore the Entropy field, and not do the selection of multi-tables.

3. The BFER (and Egress LSR) will do packet check from live-live paths, and do forward packet with zero packet loss, on a per-flow basis.

5. Provisioning Considerations

P2MP LSP based BIER use concepts of both RSVP-TE/MLDP and BIER. Some provisioning considerations list below:

Sub-domain:

In P2MP LSP based BIER, every P2MP LSP is a specific BIER underlay topology, and an implicit Sub-domain. RSVP-TE/MLDP build the BIER information of the implicit sub-domain when build P2MP LSP. MVPN get the implicit sub-domain when specified with which RSVP-TE/MLDP P2MP LSP.

In the following conditions, there may be requirements to configure an explicit sub-domain ID for P2MP LSP based BIER:

1. P2MP LSP based BIER, use the native procedure of forwarding described in [I-D.ietf-bier-architecture], which require Consistent Per-Sub-domain BIFT.

2. P2MP LSP based BIER is shared by multiple VPNs, and an explicit sub-domain ID is configured as anchor for using by these VPNs.

When explicitly configing a sub-domain ID for P2MP LSP based BIER, the ID should be great than 255. For the [0-255] has been defined to use by IGP, BGP and MVPN, as specified in [I-D.ietf-bier-ospf-bier-extensions], [I-D.ietf-bier-isis-extensions], [I-D.ietf-bier-idr-extensions] and [I-D.ietf-bier-mvpn].

BFR-prefix:

In P2MP LSP based BIER, every BFR is also a LSR. So the BFR-prefix in the sub-domain is by default identified by LSR-id. Additionally, When BFR/LSR is also a MVPN PE, BFR-prefix is also the same as Originating Router's IP Address of x-PMSI A-D route or Leaf A-D route.

BFR-id:

When using protocols like RSVP-TE, which initializes P2MP LSP from a specific Ingress Node, BFR-id which is unique in P2MP LSP scope, can be auto-provisioned by Ingress Node, or conventionally configure on every Egress Nodes.

BSL and BIER-MPLS Label Block Size:

In P2MP LSP based BIER, Every P2MP LSP or implicit sub-domain requires a single BSL, and a specific BIER-MPLS Label block size for this BSL.

VPN-Label:

In P2MP LSP based BIER, a P2MP LSP based BIER 'P-tunnel' can be shared by multiple VPNs or a single VPN. When a P2MP LSP based BIER being shared by multiple VPNs, an Upstream-assigned VPN-Label is required. It can be auto-provisioned or manual configured by the BFIR or Ingress LSR.

6. IANA Considerations

Allocation is expected from IANA for two new tunnel type codepoints for "RSVP-TE P2MP LSP based BIER" and "MLDP P2MP LSP based BIER" from the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry.

7. Security Considerations

This document does not introduce any new security considerations other than already discussed in [I-D.ietf-bier-architecture].

8. Acknowledgements

TBD

9. References

9.1. Normative References

[I-D.ietf-bess-mvpn-expl-track]
Dolganow, A., Kotalwar, J., Rosen, E., and Z. Zhang, "Explicit Tracking with Wild Card Routes in Multicast VPN", draft-ietf-bess-mvpn-expl-track-03 (work in progress), September 2017.

[I-D.ietf-bess-mvpn-fast-failover]
Morin, T. and R. Kebler, "Multicast VPN fast upstream failover", draft-ietf-bess-mvpn-fast-failover-02 (work in progress), March 2017.

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-bier-idr-extensions]
Xu, X., Chen, M., Patel, K., Wijnands, I., and T. Przygienda, "BGP Extensions for BIER", draft-ietf-bier-idr-extensions-03 (work in progress), August 2017.

[I-D.ietf-bier-isis-extensions]
Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER support via ISIS", draft-ietf-bier-isis-extensions-06 (work in progress), October 2017.

[I-D.ietf-bier-mpls-encapsulation]
Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-11 (work in progress), October 2017.

[I-D.ietf-bier-mvpn]
Rosen, E., Sivakumar, M., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using BIER", draft-ietf-bier-mvpn-08 (work in progress), October 2017.

[I-D.ietf-bier-ospf-bier-extensions]
Psenak, P., Kumar, N., Wijnands, I., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for BIER", draft-ietf-bier-ospf-bier-extensions-09 (work in progress), October 2017.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S. Yasukawa, Ed., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, DOI 10.17487/RFC4875, May 2007, <https://www.rfc-editor.org/info/rfc4875>.

[RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November 2011, <https://www.rfc-editor.org/info/rfc6388>.

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <https://www.rfc-editor.org/info/rfc6513>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

[RFC6625] Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R. Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes", RFC 6625, DOI 10.17487/RFC6625, May 2012, <https://www.rfc-editor.org/info/rfc6625>.

[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>.

9.2. Informative 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>.

Author's Address

Jingrong Xie
Huawei
Q15 Huawei Campus, No.156 Beiqing Rd.
Beijing 100095
China

Email: xiejingrong@huawei.com

BIER WG Quan. Xiong
Internet-Draft Fangwei. Hu
Intended status: Standards Track Zheng. Zhang
Expires: January 19, 2018 ZTE Corporation
July 18, 2017

BIER-TE Encapsulation
draft-xiong-bier-te-encapsulation-00.txt

Abstract

Traffic Engineering for Bit Index Explicit Replication (BIER-TE) shares part of architecture , definition and packet format with Bit Index Explicit Replication (BIER) according to the introduce in [I-D.eckert-bier-te-arch]. BIER-TE supports the traffic engineering by explicit hop-by-hop forwarding and loose hop forwarding of packets.[I-D.ietf-bier-mpls-encapsulation] specifies a BIER encapsulation that BIER header contains a bitstring in which each bit represents exactly one egress router in the domain.

This document proposes a set of extensions to BIER encapsulation for BIER-TE. The extensions define the BIER-TE header which contains serveral bitstrings and each bit in each bitstring represents one or more adjacencies in BIER-TE domain. The encapsulation can be used both in an MPLS network and a non-MPLS network.

Copyright Notice

Table of Contents

1. Introduction

Traffic Engineering for Bit Index Explicit Replication (BIER-TE) shares part of architecture, definition and packet format with Bit Index Explicit Replication (BIER) according to the introductions in [I-D.eckert-bier-te-arch]. But BIER-TE supports the traffic engineering by explicit hop-by-hop forwarding and loose hop forwarding of packets. The BIER-TE controller host determines and assigns the BitPositions to the adjacencies which explicit paths passing through.

[I-D.ietf-bier-mpls-encapsulation] specifies a BIER encapsulation that BIER header contains a bitstring in which each bit represents exactly one egress router in the domain. But in BIER-TE every BitPosition of the BitString of a BIER-TE packet indicates one or more adjacencies instead of an egress router as in BIER. That MUST be a huge number of adjacencies from BFIR to all BFERs and the BitString in BIER encapsulation is related to SD,BSL and SI combination. For these distinct SD,BSL and SI combinations, there must be more than one BitStrings and the BFR must make many copies of multicast data packet. Even more, BitPositions of all adjacencies passing through BFIR to a BFER MAY be carried in different BitStrings and within the different packets.

Based on the discussion above, this document proposes a set of extensions to BIER encapsulation for BIER-TE. The extensions define the BIER-TE header which contains one or more bitstrings and each bit in each bitstring represents one or more adjacencies in BIER-TE domain.The encapsulation can be used both in an MPLS network and a non-MPLS network.

2. 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].

2.1. Terminology

The terminology is defined as [I-D.ietf-bier-architecture], [I-D.eckert-bier-te-arch] and [I-D.ietf-bier-mpls-encapsulation].

3. BIER-TE Encapsulation

The BIER-TE header is shown in Figure 1. It extends the BIER encapsulation and adds one or more BitString Sub-TLVs.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              BIFT-id                  | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Nibble |  Ver  |  BSL  |              Entropy                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|OAM|B|Rsv|    DSCP     |   Proto     |            BFIR-id        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                BitString  (first 32 bits)                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                BitString  (last 32 bits)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             BitString  Sub-TLVs (variable)                    |
+-                                                             -+
|                                                               |
          Figure 1:BIER-TE Header Format
```

B bit : 1bit, indicates BIER-TE packet when it is set. As the [I-D.eckert-bier-te-arch] mentioned, when a BFR receives a packet, it needs to interpret the BitString of a BIER-TE packet differently from a BIER packet and it is necessary to distinguish BIER from BIER-TE packets.

BitString Sub-TLV: identifies BitString related information and each BitString Sub-TLV corresponds to a particular combination of SD, BSL, SI and bitstring.The format details see section 3.1.

The definition of other fields is the same with [I-D.ietf-bier-mpls-encapsulation].

3.1. BitString Sub-TLV

This document proposes BitString Sub-TLV for BIER-TE header. The TLV is optional. The format of the new sub-TLV is shown in Figure 2 and 3.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   SubDomainId |       SI          |  BSL  |       Rsv         |N|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      BitString  (variable)                    |
+-                                                             -+
|                                                               |
```

Figure 2: BitString Sub-TLV Format 1

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               BIFT-id                 |       Rsv           |N|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  BitString  (variable)                        |
+-                                                             -+
|                                                               |
```

Figure 3: BitString Sub-TLV Format 2

The fields in the format are:

SD : 8bits, indicates the Sub-DomainId of the BitString in the range [0,255].

SI : 10bits,indicates the Set Identification ofthe BitString in the range [0,1023].

BSL : 4bits,indicates the length in bits of the BitString. If k is the length of the BitString, the value of this field is log2(k)-5. The values are supported as follows:

1: 64 bits

2: 128 bits

3: 256 bits

4: 512 bits

5: 1024 bits

6: 2048 bits

7: 4096 bits

N : 1bit, indicates that there are one or more BitString Sub-TLVs immediately preceding the TLV when it is set.

BitString: variable, identifies the collection of the adjacencies from BFIR to all BFERs and each BitString is related to SD,BSL and SI combination or BIFT-id of the packet.

BIFT-id: 22bits,The BIFT-id represents a particular Bit Index Forwarding Table (BIFT); see Section 6.4 of [I-D.ietf-bier-architecture]. Each BIFT corresponds to a particular combination of SD,BSL and SI.

4. Processing Rules with the BIER-TE Encapsulation

As defined in [I-D.eckert-bier-te-arch], the BIER-TE operations consists of four layers: the "Multicast Flow Overlay", the "BIER-TE Controller Host", the "Routing Underlay" and the "BIER-TE forwarding layer".The BIER-TE Multicast flow processing with BIER-TE encapsulation is as follows:

1. The BIER-TE Controller assigns the BitPositions for adjacencies based on the operator policy and populates the BitPositions to the BIFT of each BFR as mentioned in [I-D.eckert-bier-te-arch].

2. The Multicast Flow Overlay determines the BFIR and a set of BFERs and sends this information to the BIER-TE controller.

3. The BIER-TE controller calculates the explicit paths based on algorithms from BFIR to all BFERs.

4. The BIER-TE controller gets all adjacencies which the paths passing through and determines the list of bitstrings based on the SD,BSL and SI combination and BitPositions/adjacencies assignments. Each bit/BitPosition in each bitstring represents one or more adjacencies in BIER-TE domain. The BitPositions of the adjacencies that have the same SD,BSL and SI combination can be encoded into the same BitString. It then pushes the BitStrings into the BFIR.

5. When a BFIR receives a multicast packet from outside the BIER-TE domain, the BFIR carries out the following procedure:

a. The BFIR makes a copy of the multicast data packet and encapsulates the copy in a BIER-TE header as this document proposes(see Section 3). The BitStrings which received from the BIER-TE controller are mapped to the field of BitString Sub-TLVs.

b. The BFIR checks the BIER-TE header and get the BitString Sub-TLVs infomation. Then traverses the Sub-TLVs and related local BIFT which has the same SD,BSL and SI combination. The packet may then be transmitted to adjacencies/neighboring BFRs and applies to that copy with the forwarding procedure of [I-D.eckert-bier-te-arch].

5. Security Considerations

TBD.

6. IANA Considerations

TBD.

7. Acknowledgements

TBD.

8. Normative References

[I-D.eckert-bier-te-arch]
Eckert, T., Cauchie, G., Braun, W., and M. Menth, "Traffic Engineering for Bit Index Explicit Replication BIER-TE", draft-eckert-bier-te-arch-05 (work in progress), June 2017.

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-07 (work in progress), June 2017.

[I-D.ietf-bier-mpls-encapsulation]
Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-07 (work in progress), June 2017.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

Authors' Addresses

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

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

Fangwei Hu
ZTE Corporation
No.889 Bibo Rd
Shanghai 201203
China

Phone: +86 21 68896273
Email: hu.fangwei@zte.com.cn

Zheng Zhang
ZTE Corporation
No.50 Software Avenue,Yuhuatai District
Nanjing, Jiangsu Province 210012
China

Phone: +86 025 88016576
Email: zhang.zheng@zte.com.cn

BIER WG
Internet-Draft
Intended status: Standards Track
Expires: March 23, 2018

Yongqing. Zhu
Huanan. Chen
China Telecom
Quan. Xiong
Fangwei. Hu
ZTE Corporation
September 19, 2017

BIER-TE Forwarding
draft-zcxh-bier-te-forwarding-00.txt

Abstract

Traffic Engineering for Bit Index Explicit Replication (BIER-TE) shares part of architecture, definition and packet format with Bit Index Explicit Replication (BIER) according to the introduction in [I-D.eckert-bier-te-arch]. But BIER-TE supports the traffic engineering by explicit hop-by-hop forwarding and loose hop forwarding of packets.

This document proposes a set of extensions to realize the BIER-TE forwarding including the assignment of BitPositions to adjacencies and the configuration of Bit Index Forwarding Table (BIFT).

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Traffic Engineering for Bit Index Explicit Replication (BIER-TE) shares part of architecture, definition and packet format with Bit Index Explicit Replication (BIER) according to the introduction in [I-D.eckert-bier-te-arch]. But BIER-TE supports the traffic engineering by explicit hop-by-hop forwarding and loose hop forwarding of packets.

[I-D.ietf-bier-mpls-encapsulation] specifies a BIER encapsulation that BIER header contains a bitstring in which each bit represents one egress router in the domain. But in BIER-TE, every BitPosition of the BitString of a BIER-TE packet indicates one or more adjacencies which BFRs will transit packets passing though. BFRs recognizes BitStrings or packets for every Sub-Domain-ID(SD), BitStringLength(BSL) and Set Identification(SI) combination.

[I-D.eckert-bier-te-arch] discussed the process of the BIER-TE forwarding including Bit Index Forwarding Table (BIFT) and forwarding example. The BIER-TE controller host determines and assigns the BitPositions to the adjacencies which explicit paths can be built through them and then pushes the BitPositions/adjacencies to the BIFT which indexed by SI:BitPosition. The BIFT is configured to the routers known as "Bit-Forwarding Router" (BFR) which should be able to send packets to adjacencies connecting to other BFRs.

1.1. Motivation

As defined in [I-D.ietf-bier-architecture], a multicast data packet enters a domain at a "Bit-Forwarding Ingress Router" (BFIR), and leaves at one or more "Bit-Forwarding Egress Routers" (BFERs). For a multicast forwarding, the controller host needs to assign lots of BitPositions and use multiple SI and BSL within the same sub-domain. The distinct SD, BSL and SI combinations MUST be mapped to more than one BitStrings and carried in different packets.

As discussed in [I-D.eckert-bier-te-arch], the BIER-TE controller host tracks the BFR topology of the BIER-TE domain and determines the BitPositions and related BIFTs.Different with BIER, the BIFT related to the BitPositions which associated with a particular SD, BSL and SI combination need to be built throughout the whole network in BIER-TE.

The BFRs need to forward the packets based on the BitString and BIFT with a SD, BSL and SI combination.The BitPositions of these adjacencies passing through BFIR to each BFER must be assigned in the same SD, BSL and SI combination to ensure the multicast flow be forwarded to the BFER within the same packet. The assignment of BitPositions and the configuration of BIFT should be taken to considerations in detail.

1.2. Operation Overview

Based on the discussion above, this document proposes a set of extensions to realize the BIER-TE forwarding including the assignment of BitPositions to adjacencies and the configuration of BIFT. The main point is that the assignment of BitPositions and the configuration of the BIFT MUST be accomplished based on the explicit paths of multicast flow and be completed after the BFIR and BFERs are configured. The controller host SHOULD take charge of the management about multicast flow information.

The controller host dosen't need to track the topology to determine what adjacencies require BitPositions. The controller host MAY compute the explicit paths from BFIR to each BFER first and then assign the BitPositions including SD,BSL and SI combination to the adjacencies which the paths passing though respectively based on the policy. The assignment results need to meet the requirement that the BitPositions of the adjacencies from BFIR to each BFER could belong to a SD, BSL and SI combination. And then the controller pushes those BitPositions/adjacencies to the BIFT of the BFRs. The configuration of BIFT is not completed in BFR topology but incremental configuration based on the requirement of multicast flows.

1.3. 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].

2. BIER-TE Forwarding

This document proposes a general mechanism to extend the process in [I-D.eckert-bier-te-arch]. The operation for BIER-TE forwarding is as follows.

The controller host which representing the control plane of BIER-TE discovers the network topology information.

When the multicast flows needs to be forwarded in the BIER-TE network, the controller host tracks the multicast flow overlay to determine which multicast flow needs to be sent by a BFIR to which BFERs.

And based on the topology, the controller host needs to calculate the explicit paths from BFIR to every BFER across the BIER-TE domain according to the algorithms which is outside the scope of this document.

The BIER-TE controller host assigns the BitPositions including SD, BSL and SI combination to the adjacencies according to the assignment method and policy as discussed on the session 3.1. The BIFT for related BFRs which the explicit paths passing through SHOULD be populated by the controller host and configured once the BitPositions assigned with the detail in session 3.2 .

Finally, the BIER-TE controller host calculates the BitStrings according to the explicit paths and its related explicit SD, BSL and SI combination and pushes them into the BFIR as discussed in [I-D.eckert-bier-te-arch].

Once the BIFTs and BitStrings was programmed into the data plane of BFRs by the BIER-TE controller host, they can be used to forward packets according to the rules specified in the BIER-TE forwarding Procedures defined in [I-D.eckert-bier-te-arch].

2.1. The Assignment of BitPositions to Adjacencies

The BIER-TE controller host assigns the BitPositions including SD, BSL and SI combination to the adjacencies based on the explicit paths passing though. One or more BitPositions MAY be assigned to an adjacency with the different SD, BSL and SI combination. The assignment need to meet the requirement that the BitPositions of the adjacencies from BFIR to each BFER could belong to a SD, BSL and SI combination.

This document proposes a merthord for the assignment of BitPosition to adjacencies and defines two types of the assignment policy of BitPositions as following.

If the multicast flow needs to be sent from a BFIR to M BFERs along M explicit paths, the controller host MAY assign BitPositions for all adjacencies of M explicit paths with the K sets of SD, BSL and SI combinations which K M according to the assignment policy.

EXCLUSIVE-TYPE: Each multicast flow MAY use one or more SD, BSL and SI combination exclusively.

SHARING-TYPE: More than one multicast flows MAY share the same SD, BSL and SI combination. If the adjacencies of a path have been assigned to the same SI except some adjacencies which have not been assigned ever, the controller host SHOULD assign BP for these no-assigned adjacencies the same SI with the others. The premise is the index of the SI is enough for the assignment.

OPTIONALY, the policy of assignment MAY be configured by customers based on the requirement outside of the document.

2.2. The configuration of BIFT

The BIFT is populated by the BIER-TE control plane and exists in all BFRs as defined in [I-D.eckert-bier-te-arch]. This document proposes an extension to BIFT as the table 1 shown.

BIFT-id represents a particular BIFT and corresponds to a particular combination of SD, BSL, and SI. The value of BITF-id MUST be assigned by BIER-TE controller host and unique throughout the BIER-TE domain. The BIFT-id can be used in BIER encapsulation as discussed in [I-D.ietf-bier-mpls-encapsulation]. BIFT-type indicates the type of BIFT including BIER and BIER-TE.

Table 1 Extension of BIFT

Index: BIFT-id (<SD:BSL:SI>) : BitPosition BIFT-type: BIER-TE	Adjacencies: <empty> or one or more per entry
BIFT-id:1	forward_connected(interface,neighbor,DNR)
BIFT-id:2	forward_connected(interface,neighbor,DNR) forward_connected(interface,neighbor,DNR)
BIFT-id:3	local_decap([VRF])
BIFT-id:4	forward_routed([VRF,]l3-neighbor)
BIFT-id:5	<empty>
BIFT-id:6	ECMP({adjacency1,...adjacencyN}, seed)
...	
ID:BitStringLength	...

The BIFT in one sub-domain of a BFR is a table indexed by BIFT-id:BitPosition which populated by the controller host and configured once the BitPositions assigned. One or more BitPositions in table MAY correspond to the same adjacency. The configuration of BIFT is not completed before the service deployment but incremental configuration based on the requirement of multicast flows. The difference is that the configuration of BIFT is not to replace the table in BFR but update and add the BIFT-id:BitPosition items into BIFT.

When links or nodes fail or recover in the topology or service is deleted by customers, the related items need to be removed from BIFT with little effect on other BIFT items of other flows.

3. BIER-TE Forwarding Example

Step by step example of basic BIER-TE forwarding and using the network defined in [I-D.eckert-bier-te-arch]. The extension process for BIER-TE forwarding is shown as follows.

```
                          [Bier-Te Controller Host]
                                  /   | \
                                 v    v  v

                      | a13   a1 |
                      +- BFIR2 --+          |
                      |          | a2   a6  |           LAN2
                      |          +-- BFR3 --+           |
                      |          |          |  a7  a11  |
                 Src -+                     +-- BFER1 --+
                      |          | a3   a8  |           |
                      |          +-- BFR4 --+           +-- Rcv1
                      |          |          |           |
                      |          |
                      | a14  a4  |
                      +- BFIR1 --+          |
                      |          +-- BFR5 --+ a10  a12  |
                    LAN1         | a5   a9  +-- BFER2 --+
                                            |           +-- Rcv2
                                                        |
                                                        LAN3

                  IP |.....    BIER-TE network   ......| IP
```

Figure 1 Forwarding Example

aXX indicate the Adjacencies number assigned by the BIER-TE controller host according to the BIER-TE topology.

1, The BIER-TE controller discovers the network topology information and assigns the Adjacencies number for the adjacencies as the figure shown.

2, The BIER-TE controller tracks the multicast flow overlay to determine what multicast flow needs to be sent by which BFIR to which BFERs. Example is from BFIR2 to BFER1 and BFER2.

3, The BIER-TE controller computes the two optimal paths from BFIR2 to BFER1 and from BFIR2 to BFER2 respectively. The result is shown as follows.

a.BFIR2->BFER1:SRC->a13->a2->a7->a11->RCV

b.BFIR2->BFER2:SRC->a13->a2->a8->a5->a10->a12->RCV

4, The BIER-TE controller host assigns BP for the path from BFIR2 to BFER1 with the assignment method and the policy type is EXCLUSIVE-

TYPE. SD =0, SI=0, BSL=4, BIFT-id = 0, the BIFT-id:BitPosition is as follows:

a13->0:1

a2->0:2

a7->0:3

a11->0:4

5, The BIER-TE controller host assigns BP for the path from BFIR2 to BFER2. SD =0, SI=1, BSL=8, BIFT-id = 1, the BIFT-id:BitPosition is as follows:

a13->1:1

a2->1:2

a8->1:3

a5->1:4

a10->1:5

a12->1:6

6, Based on the assignment, the BIER-TE controller populates the according BIFTs and forwards it to the BFRs as the following shown.

BIFT BFIR2:

0:1: local_decap()

1:1: local_decap()

0:2: forward_connected(BFR3)

1:2: forward_connected(BFR3)

BIFT BFR3:

0:3: forward_connected(BFER1)

1:3: forward_connected(BFR4)

BIFT BFER1:

0:4: local_decap()

1:3: forward_connected(BFR4)

BIFT BFIR1:

1:4: forward_connected(BFR5)

BIFT BFR4:

0:3: forward_connected(BFER1)

1:4: forward_connected(BFR5)

BIFT BFR5:

1:5: forward_connected(BFER2)

BIFT BFER2:

1:6: local_decap()

7, The BitString is splited into two sub-BitStrings according to the BIFT-id by the BIER-TE controller.Examples for SI:Bitstring is 0:1111 and 1:00111111.

4. Security Considerations

TBD.

5. IANA Considerations

TBD.

6. Acknowledgements

TBD.

7. Normative References

[I-D.eckert-bier-te-arch]
Eckert, T., Cauchie, G., Braun, W., and M. Menth, "Traffic Engineering for Bit Index Explicit Replication BIER-TE", draft-eckert-bier-te-arch-05 (work in progress), June 2017.

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-bier-mpls-encapsulation]
Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-08 (work in progress), September 2017.

[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>.

Authors' Addresses

Yongqing Zhu
China Telecom
109 West Zhongshan Ave
Guangzhou, Guangdong 510630
China

Phone: +86 20 38639581
Email: zhuyq@gsta.com

Huanan Chen
China Telecom
109 West Zhongshan Ave
Guangzhou, Guangdong 510630
China

Phone: +86 20 38639346
Email: chenhuanan@gsta.com

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

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

Fangwei Hu
ZTE Corporation
No.889 Bibo Rd
Shanghai 201203
China

Phone: +86 21 68896273
Email: hu.fangwei@zte.com.cn

BIER Z. Zhang
Internet-Draft ZTE Corporation
Intended status: Standards Track A. Przygienda
Expires: April 29, 2018 Juniper Networks, Inc.
October 26, 2017

BIER in IPv6
draft-zhang-bier-bierin6-00

Abstract

BIER is a new architecture for the forwarding of multicast data packets. This document defines native IPv6 encapsulation for BIER hop-by-hop forwarding or BIERin6 for short.

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.

Table of Contents

1. Introduction

BIER [I-D.ietf-bier-architecture] is a new architecture for the forwarding of multicast data packets. It provides optimal forwarding through a "multicast domain' and it does not necessarily precondition construction of a multicast distribution tree, nor does it require intermediate nodes to maintain any per-flow state.

[I-D.ietf-bier-mpls-encapsulation] defines the BIER encapsulation format in MPLS and non-MPLS environment that rely on MPLS labels or special ethernet type support. In pure IPv6 environments a BIER packet could be forwarded by means of simple IPv6 hop-by-hop processing only, without any new hardware support. Ultimate hardware support is obviously possible but the encapsulation is especially interesting for environments like [RFC7368] where high throughput multicast forwarding performance is not decisive and could be initially done in slow-path or end host (assuming last-hop being IPv6 encapsulated).

This document defines native BIER IPv6 [RFC2460] encapsulation we call BIERv6. This encapsulation is aligned with the format defined in [I-D.ietf-bier-mpls-encapsulation] for a non-mpls version.

This document uses terminology defined in [I-D.ietf-bier-architecture].

2. IPv6 Header

The BIER packet itself is the payload of an IPv6 frame. The destination address field in IPv6 packet MAY be the neighbor's link-local or one of the loopback interface addresses of that neighbor. The destination address SHOULD be the BFR-prefix advertised by IGP/BGP extensions for BIER. TTL value of 1 MUST be used on the IPv6 packet.

The source address field in IPv6 packet MAY be the loopback interface address of the sending BFIR. The address SHOULD be the BFR-prefix advertised in IGP/BGP extension.

A new next-protocol type in IPv6 Next header field of TBD indicates the following BIER packet.

The Flow-ID in the IPv6 packet SHOULD be copied from the entropy field in the BIER encapsulation.

3. BIER Header

S bit in BIER header has no significance in this environment. It should be set to 1 upon transmission, but it MUST be ignored upon reception.

TC bits in BIER header have no significance in this environment since the IPv6 packet TC takes precedence on processing. It should be set to zero upon transmission, but it MUST be ignored upon reception.

The BIFT-id is used to indicate the combination of <SD, SI, BSL> normally; it should be set to the value advertised by the next-hop BFR through e.g. IGP [I-D.ietf-bier-ospf-bier-extensions], [I-D.zhang-bier-babel-extensions] or BGP [I-D.ietf-bier-idr-extensions] extension for BIER.

The remaining fields defined in BIER header MUST assume the same values and be afforded same treatement as specified in [I-D.ietf-bier-mpls-encapsulation].

4. IANA Considerations

IANA is requested to set up a new type of "Next header" registry value for BIERv6 in the "Assigned Internet Protocol Numbers" registry.

5. Security Considerations

General IPv6 and BIER security considerations apply.

6. References

6.1. Normative References

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-bier-idr-extensions]
Xu, X., Chen, M., Patel, K., Wijnands, I., and T. Przygienda, "BGP Extensions for BIER", draft-ietf-bier-idr-extensions-03 (work in progress), August 2017.

[I-D.ietf-bier-mpls-encapsulation]
Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-10 (work in progress), October 2017.

[I-D.ietf-bier-ospf-bier-extensions]
Psenak, P., Kumar, N., Wijnands, I., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for BIER", draft-ietf-bier-ospf-bier-extensions-09 (work in progress), October 2017.

[I-D.zhang-bier-babel-extensions]
Zhang, Z. and T. Przygienda, "BIER in BABEL", draft-zhang-bier-babel-extensions-01 (work in progress), June 2017.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

6.2. Informative References

[RFC7368] Chown, T., Ed., Arkko, J., Brandt, A., Troan, O., and J. Weil, "IPv6 Home Networking Architecture Principles", RFC 7368, DOI 10.17487/RFC7368, October 2014, <https://www.rfc-editor.org/info/rfc7368>.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation

EMail: zzhang_ietf@hotmail.com

Tony Przygienda
Juniper Networks, Inc.

EMail: prz@juniper.net

BIER WG Zheng. Zhang
Internet-Draft BenChong. Xu
Intended status: Standards Track ZTE Corporation
Expires: April 3, 2018 September 30, 2017

BIER Flooding Mechanism
draft-zhang-bier-flooding-00.txt

Abstract

This document introduces a method to flood BIER information in hybrid network to build BIER forwarding plane.

Status of This Memo

Copyright Notice

Table of Contents

1. Problem Statement

Some networks have been deployed widely are hybrid networks. There are different dynamic routing protocols running in the hybrid networks. Multicast services can also be provided in these kinds of networks because of the protocol independent feature of PIM.

BIER [I-D.ietf-bier-architecture] provides a new architecture for the forwarding of multicast data packets. It does not require a protocol for explicitly building multicast distribution trees, nor does it require intermediate nodes to maintain any per-flow state. [I-D.ietf-bier-isis-extensions] and [I-D.ietf-bier-ospf-bier-extensions] are good at establishing BIER forwarding plane in network which uses OSPF or IS-IS as BIER underlay protocol.

```
         |                |
.........|................|.........
.  ISIS  |                |        .
.        |                |        .
.       MR1--------------MR2       .
.        |  .          .  |        .
.        |    .      .    |        .
.        |      .  .      |        .
.        |       ..       |        .
.        |     .    .     |        .
.        |   .        .   |        .
.        | .            . |        .
.        MR3              MR4      .
.       /  \             /  \      .
....................................
       /    \           /    \
      /      \         /      \
     /        \       /        \
...................  ................
.    B1--------B2 .  .B3--------B4  .
.                 .  .              .
.   Rm            .  . Rx           .
.       ...       .  .     ...      .
.            Rn   .  .         Ry   .
.                 .  .              .
.         OSPF    .  .       OSPF   .
...................  ................
```

Figure 1: An typical hybrid network

In the mentioned networks, there are more than one dynamic routing protocols running in the networks. For example in figure 1, this is a partial typical network in actually deployment. Two different dynamic routing protocols and are used in the network. Sometimes static configured routes also are used in some parts of the network. In order to deploy BIER multicast, we can divide the network into several BIER domains. Obviously the efficiency slows down due to multiple encapsulating/ decapsulating executions.

2. Solution

The Bootstrap Router mechanism (BSR) [RFC5059] is a commonly used mechanism for distributing dynamic Group to RP mappings in PIM. It is responsible for flooding information about such mappings throughout a PIM domain, so that all routers in the domain can have the same information. [I-D.ietf-pim-source-discovery-bsr] defines a mechanism that can flood any kind of information throughout a PIM domain. This document borrows the idea from the two drafts, introduces a mechanism to flood BIER node’s information throughout a

BIER domain to build BIER forwarding plane. Nodes can use unicast forwarding table directly to establish BIER forwarding plane.

The validation processing of PFM messages is the same as the definition in [I-D.ietf-pim-source-discovery-bsr] section 3.2.

BIER node originates BIER information TLV and optional associated sub-TLVs in PFM message. The PFM messages are flooded by throughout the BIER domain. BFR gets routing information from the unicast forwarding table directly, and computes BIER forwarding table. Then BIER forwarding plane is established.

2.1. Scheduled Update

Because PIM advertisement is scheduled, the node's BIER information is refreshed periodically. In case one node's BIER information changes or expires, the other nodes recompute the BIER forwarding table. The holdtime in the BIER information TLV is used to make the item expired.

2.2. Triggered Update

If the BIER node's configuration changes, such as BFR-id, the node should send update PFM messages immediately. Then other nodes can recompute the new BIER forwarding table.

3. Message Format

3.1. PFM message

New TLVs are defined in PFM message to flood node's BIER information, such as BFR-id, BFR-prefix and so on. The new TLVs align exactly with the definition and restrictions in [I-D.ietf-bier-isis-extensions] and [I-D.ietf-bier-ospf-bier-extensions].

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|PIM Ver| Type  |N|  Reserved   |           Checksum            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Originator Address (Encoded-Unicast format)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type 1              |          Length 1             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value 1                            |
|                               .                               |
|                               .                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               .                               |
|                               .                               |
|           Type n              |          Length n             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value n                            |
|                               .                               |
|                               .                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: PFM message format

The format of PFM message is defined in [I-D.ietf-pim-source-discovery-bsr].

Originator Address: The router's address that originate the message. The address SHOULD be the same with the node's BFR-prefix.

The other fields is the same as definition in [I-D.ietf-pim-source-discovery-bsr].

3.2. BIER information TLV

A new type of TLV is defined in PFM message. This new type TLV is named by BIER information TLV. Two types of sub-TLV are associated with it. There is no optional BIER tree type sub-TLV in PFM message because of the independence of routing protocol.

```
 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 = 1            |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Reserved    | subdomain-id  |            BFR-id             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Holdtime         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              BFR-prefix (Encoded-Unicast format)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: BIER information TLV

- o Type: The value of type should be assigned by IANA.

- o Length: The total length of the BIER information TLV except for the first two fields.

- o Reserved: Must be 0 on transmission, ignored on reception. May be used in future version. 1 octets.

- o Subdomain-id: Unique value identifying the BIER sub-domain. 1 octet.

- o BFR-id: The value of BFR-id defined in [BIER-arch], 2 octets. 0 is invalid value. If the value of this field is set to 0, the whole TLV MUST be ignored and not forwarded.

- o Holdtime: The life cycle of the BIER information. The default value is 60s.

- o BFR-prefix: The BFR-prefix of the node in this sub-domain. The format for this address is given in the Encoded-Unicast address in [RFC7761].

A node may belong to several BIER sub-domains, so it is possible that there are multiple BIER information TLVs in the PFM message.

3.3. BIER MPLS Encapsulation sub-TLV

In case the nodes in the network support MPLS forwarding, BIER MPLS encapsulation sub-TLV can be advertised for a specific bitstring length for a certain (MT,subdomain). This sub-TLV may appear multiple times within single BIER information TLV. The format and restriction is the same as the definition in [I-D.ietf-bier-isis-extensions] and [I-D.ietf-bier-ospf-bier-extensions].

The type value of this sub-TLV should be assigned by IANA. The suggestion value is 1.

3.4. Optional BIER sub-domain BSL conversion sub-TLV

The format and restriction is the same as the definition in [I-D.ietf-bier-isis-extensions] and [I-D.ietf-bier-ospf-bier-extensions]. The type value of this sub-TLV should be assigned by IANA. The suggestion value is 2.

4. Security Considerations

The security considerations are mainly similar to what is documented in [I-D.ietf-pim-source-discovery-bsr].

5. IANA Considerations

This document requires the assignment of a new PFM TLV type for the BIER information Flooding Mechanism. IANA is also requested to create two sub-TLV types for BIER MPLS encapsulation sub-TLV and BIER sub-domain BSL conversion sub-TLV.

6. Normative References

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-bier-isis-extensions]
Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER support via ISIS", draft-ietf-bier-isis-extensions-05 (work in progress), July 2017.

[I-D.ietf-bier-ospf-bier-extensions]
Psenak, P., Kumar, N., Wijnands, I., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for BIER", draft-ietf-bier-ospf-bier-extensions-07 (work in progress), July 2017.

[I-D.ietf-pim-source-discovery-bsr]
Wijnands, I., Venaas, S., Brig, M., and A. Jonasson, "PIM flooding mechanism and source discovery", draft-ietf-pim-source-discovery-bsr-06 (work in progress), March 2017.

[RFC5059] Bhaskar, N., Gall, A., Lingard, J., and S. Venaas, "Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)", RFC 5059, DOI 10.17487/RFC5059, January 2008, <https://www.rfc-editor.org/info/rfc5059>.

[RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <https://www.rfc-editor.org/info/rfc7761>.

Authors' Addresses

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

Email: zhang.zheng@zte.com.cn

BenChong Xu
ZTE Corporation
No. 68 Zijinghua Road, Yuhuatai Distinct
Nanjing
China

Email: xu.benchong@zte.com.cn