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

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 23, 2018.

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:	Adjacencies:
BIFT-id (<SD:BSL:SI>) : BitPosition	<empty> or one or more per entry
BIFT-type: BIER-TE	
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.

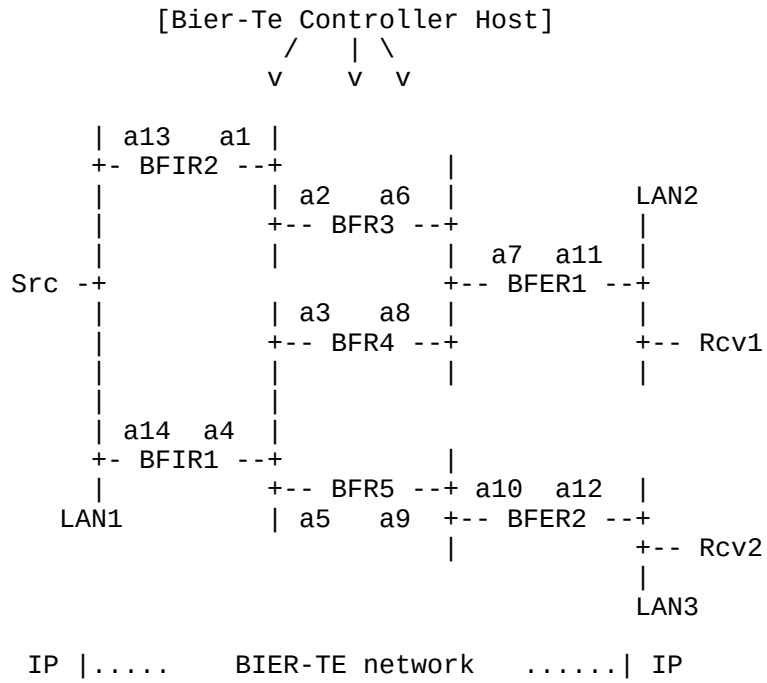

Figure 1 Forwarding Example

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