BIER Workgroup H. Bidgoli, Ed.
Internet Draft A. Dolganow
Intended status: Standard Track Nokia
Fengman Xu
Verizon

Expires: December 23, 2017 June 21, 2017

PIM Tunneling Through BIER Core
draft-hfa-bier-pim-tunneling-00

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides multicast forwarding through a "BIER domain" without requiring intermediate routers to maintain multicast related per-flow state. Neither does BIER require an explicit tree-building protocol for its operation. A multicast data packet enters a BIER domain at a "Bit-Forwarding Ingress Router" (BFIR), and leaves the BIER domain at one or more "Bit-Forwarding Egress Routers" (BFERs). The BFIR router adds a BIER header to the packet. Such header contains a bit-string in which each bit represents exactly one BFER to forward the packet to. The set of BFERs to which the multicast packet needs to be forwarded is expressed by the according set of bits switched on in BIER packet header.

This document describes the procedure needed for PIM to be tunneled through a BIER core. Allowing access CEs or PEs to run traditional PIM multicast services including draft rosen multicast MVPN through a core of BIER.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

Most Service Providers understand the benefit of BIER and would like a Core network that supports scalable multicast solution by removing the multicast states and deploying BIER.

That said greenfield deployment of BIER might not be possible for providers that deploy MVPN technology, or have more than 256 PEs in their network. Consider the following:

1. Most service providers deploy MVPN technology for multicast today. Their network structure typically is a Core network, which connects edge networks containing PEs. These PEs run MVPN services like draft rosen multicast vpns. In a typical tier one provider the number of PEs is well beyond 1K of routers. As the edge network expands the scale of multicast states in the core could test the routers limits. As such it is attractive to create a stateless BIER core which can transport MVPN technology. By deploying BIER in the core of the network the bottle necks for multicast states is removed. By pushing the multicast states to the edge provider (P) routers, a more manageable multicast state can be achieved.

2. Deploying greenfield BIER services for most providers could be a challenge. It might be attractive to deploy bier in multiple phases. Starting from the core of the network to remove the massive multicast state generated by traditional MVPN services is an ideal evolution path. By tunneling traditional MVPN technology through a BIER core an scalable and manageable network can be created.

3. Most vendors support 256 bits in the BIER header. Identifying only 256 PEs or CEs via a single BIER packet. Scaling beyond 256 PEs or CEs "might" require packet duplation depending on network topology. This packet duplication will be done via BIER Set Index (SI) which is explained in draft-ietf-bier-architecture. In the cases that packet duplication can’t be avoid it might be desirable to segment the network to traditional MVPN technology at the access and BIER in the core. By moving the Bier in the core all core routers could be presented via the 256 bits in the BIER header.

In all above cases it might be attractive to be able to tunnel traditional MVPN services over a BIER core.

This draft explains the procedure to tunnel PIM through a BIER core, as such enable tunneling of traditional MVPN services like draft-rosen multicast vpns through a core of BIER.

The procedures of PIM tunneling should be used at the BIER edge routers. The BIER edge routers (BER) are connected to legacy PIM routers on one side and BIER routers on the other side. PIM routers continue to send PIM state messages to the BER but the BER does not propagate PIM packets natively into the BIER sub-domain. Instead it will tunnel the PIM through BIER network.

In this draft the BFIR and BFER are the BER from multicast traffic point of view and not PIM signaling. That said the PIM BFIR (P-BFIR) and PIM (P-BFER) are BER from PIM signaling point of view.

As such a P-BFIR would be a BFER of the multicast traffic and a P-BFER would be a BFIR of the multicast traffic.

2. Conventions used in this document

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].

2.1. Definitions

Some of the terminology specified in [I-D.draft-ietf-bier-architecture-05] is replicated here and extended by necessary definitions:

BIER:

Bit Index Explicit Replication (The overall architecture of forwarding multicast using a Bit Position).

BFR:

Bit Forwarding Router (A router that participates in Bit Index Multipoint Forwarding). A BFR is identified by a unique BFR-prefix in a BIER domain.

BFIR:

Bit Forwarding Ingress Router (The ingress border router that inserts the BM into the packet). Each BFIR must have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plain packet (i.e. PIM) and forwarding dataplain packets. BFIR is term used for dataplain packet forwarding.

BFER:

Bit Forwarding Egress Router. A router that participates in Bit Index Forwarding as leaf. Each BFER must be a BFR. Each BFER must have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plain packet (i.e. PIM) and forwarding dataplain packets. BFIR is term used for dataplain packet forwarding.

P-BFIR:

PIM-Bit Forwarding ingress router. Ingress boundary router between PIM domain and BIER domain. It tunnels PIM packet through a BIER domain toward the source.

P-BFER:

PIM-Bit Forwarding egress router. Egress boundary router between BIER domain and PIM domain. It decapsulates PIM packet from a BIER tunnel and forwards it to the PIM domain.

BRT:

BIER RPF Table, is built on the P-BFER. It tracks which P-BFIR is interested in a group. It is used to map the group to the P-BFIR BIER prefix.

BFT:

Bit Forwarding Tree used to reach all BFERs in a domain.

BIFT:

Bit Index Forwarding Table.

BIER sub-domain:

A further distinction within a BIER domain identified by its unique sub-domain identifier. A BIER sub-domain can support multiple BitString Lengths.

BFR-id:

An optional, unique identifier for a BFR within a BIER sub-domain.

3. PIM Tunneling Through BIER domain

Suppose BIER sub-domain is to be an IGP area or instance. The BIER edge routers (BER) can be ABRs that are connected to edge network via IGP or BGP, or they can be any provider (P) router that is selected to act as the BER in that BIER sub-domain.

Each BER is configured as per BIER requirements explained in draft-ietf-bier-architecture.

The BERs receive PIM joins from the downstream routers because they are on the path toward the source. Additionally, on these BERs all interfaces which are PIM enabled are configured to tunnel PIM over BIER.

3.1. PIM-BFIR procedure

When PIM joins for a certain (S,G) arrives on a BER, in this case the P-BFIR (BFIR from PIM signaling point of view). This router first find the route to the source. The route to the source is assumed to be an IGP route. The BER tries to resolve the source (S), in the process it of resolving the source the SPF calculation can return the P-BFER that is in the path to the source.

The procedure to find the P-BFER (BFER from PIM signaling point of view) can be via 2 mechanism and is beyond the scope of this draft.

1.The P-BFER can be an ABR or ASBR router which is summarizing the route to the source, and as such is the source of this route.

2.The P-BFER can be resolved via SPF calculation and finding the first BFIR in the path the source.

The P-BFIR will become the BFER for multicast traffic point of view. This P-BFIR will track all the PIM interfaces that are interested in the (S,G) and create multicast states for all PIM routers attach to it. This BFER route will have incoming interface (RPF) as BIER "tunnel" interface and outgoing interface as the interface on which PIM Join was received. If there is another PIM Join for the same multicast (S,G) entry on some other interface, that interface gets added in the outgoing interface list.

The P-BFIR after discovering the P-BFER and its BFR-ID (flooded via IGP BIER extension) will construct the BIER header via the BIFT. The PIM packet is encapsulated in the BIER header and transported through BIER domain to P-BFER.

3.1.1. BIER packet construction at PIM BFIR

The BIER header will be encoded with the BFR-id of the P-BFER(with appropriate bit set in the bitstring) and PIM Join is then encapsulated in the packet.

```
        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|Rsv|    DSCP   |   Proto   |            BFIR-id            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                BitString  (first 32 bits)                     ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                                                               ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                BitString  (last 32 bits)                      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

BIERHeader.Proto = PIM

BIERHeader.BitString= Bit corresponding to the BFR-ID of the P-BFER

BIERHeader.BFIR-id = BFR-Id of the BER originating the encapsulated PIM packet, i.e. the P-BFIR.

Rest of the values in the BIER header are determined based on the network (MPLS/non-MPLS), capabilities (BSL), and network configuration.

3.2. Tunneling PIM through the BIER domain procedure

Throughout the BIER domain the BIER forwarding procedure is in par with draft-ietf-bier-architecture. No BIER router will examine the tunnel PIM packet. As such there is no multicast state build in the BIER domain.

The packet will be forwarded through the BIER domain until it reaches the BER with matching BFR-ID as in the BIERHeader.Bitstring. This BER (P-BFER) will examine the packet and know that it is a PIM packet from BIERHeader.Proto field and farther processing is needed.

3.3. PIM-BFER procedure

After receiving the BIER packet and processing the PIM payload encapsulated in BIER packet the P-BFER will remove the BIER header from PIM packet and lookup the route to the source, if the source is in PIM domain, it forwards the PIM packet toward the source.

With same token the P-BFER creates a multicast state with incoming interface as same interface that PIM join packet was forwarded and outgoing interfaces of BIER-Header.BFIR-id.

The P-BFER will also build a BIER reverse path forwarding table (BRT) table, using the BIERHeader.BFIR-id and the Group specified in the arriving PIM packet (S,G). BRT will be used by BFIR for datapath forwarding.

The router keeps track of all BFIR-ids interested in the Group specified in the (S,G) and updates the multicast state and populates the BRT accordingly.

It should be noted this router can also receive and forward PIM packet from other routers in the PIM domain and update the muticast state accordingly.

At this point the end-to-end multicast traffic flow setup is complete.

4. Datapath Forwarding

4.1. BIER reverse path forwarding table

The BIER RPF table (BRT) is needed on the BFIR so the multicast traffic can find the P-BFIR ID and append the correct BIT index to the BIER header for the multicast traffic before forwarding to BIER domain.

This table is built on the P-BFER buy using info from PIM packet and its corresponding BIER header. The PIM packet can provide the specific Group (G) address, meanwhile its corresponding BIER header can provide the originating P-BFIR ID. The P-BFIR is the last BIER router in the BIER domain or the BFER from datapath point of view.

These two pieces of information will be used to build BRT.

It should be noted that a single group can be associated with multiple P-BFIR IDs, as an example multiple MVPN leaf routers behind BIER domain are interested in the same group. These LEAF PEs are reachable via different P-BFIRs.

When the correct P-BFIR(s) (BFER(s)) are found in this table their P-BFIR ID can be used to do a lookup in BIFT and appropriate BIT indexes appended to the BIER header before forwarding the packet to BIER domain.

As an example in the network below:

```
     (BFIR)                      (BFER)
S--(P-BFER)---BIER-DOMAIN---(P-BFIR)--A (join (S,G1),(S,G2))
                                |-----C (join (S,G1),(S,G3))
                                |-----E (join (S,G3))
```

The BRT is as follow:

Group	P-BFIR
(G1)	C,A
(G2)	A
(G3)	E,C

And the BIFT is as follow:

BFR-id	BFR-NBR
(SI:Bitstring)	
1 (0:0001)	C
3 (0:0100)	E
4 (0:1000)	A

As such a multicast dataplain packet arriving with destination G1 will have the BITs (0:1001) and a packet arriving with destination of G3 will have the BITs of (0:0101)

4.2. Datapath traffic flow

When the multicast data traffic arrives on the BFIR (P-BFER) the router will find the destination IP of the traffic (i.e. group address) in the BRT. The BFIR then finds all the P-BFIR (BFER) that are interested in this group from the BRT table. The router then constructs the BIERHeader.BitString with all the BFIR interested in the group and will forward the packet to the BIER domain. The BFER(s) will accept the packets and remove the BIER header and forward the multicast packet as per pre-build multicast state for (G) and its outgoing interfaces.

5. PIM-ASM behavior

In case of PIM ASM the procedure for LEAFs joining RP or the source is same as above. The unicast (source registration) traffic from source to RP will be flooded throughout the BIER domain as regular unicast traffic without BIER involvement.

6. Draft Rosen multicast vpn behavior

Over the years draft rosen mvpn has evolved with many different type of signaling.

As an example, with AD and C-Multicast signaling of PIM or C-Multicast of PIM and AD of BGP.

The above mechanism works with draft rosen mvpn as long the C-multicast signaling is done via PIM. The provider PIM for MVPN can be forwarded from Root and LEAF PE with above explained mechanism.

The multicast traffic has to be forwarded via GRE tunnel. That said the AD signaling can be done via MP-BGP.

Future drafts will address the NG-MVPN and MPLS tunneling.

7. Security Considerations

TBD

7.1. Normative References

[BIER_ARCH] Wijnands, IJ., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", internet-draft draft-ietf-bier-architecture-05, October 2016.

[DRAFT ROSEN] E. Rosen, Y. Cai, I. Wijnands, draft-rosen-vpn.mcast-15, June 2010 (RFC 6037)

7.2. Informative References

[BGP_BIER_EXTENSIONS] Xu, X., Chen, M., Patel, K., Wijnands, I., and A. Przygienda, "BGP Extensions for BIER", internet-draft draft-ietf-bier-idr-extensions-02.txt, June 2017.

[BIER-OAM] Kumar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", internet-draft draft-ietf-bier-ping-01.txt, January 2017.

[BIER_MVPN] Rosen, E., Ed., Sivakumar, M., Wijnands, IJ., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using Bier", internet-draft draft-ietf-bier-mvpn-05, January 2017.

[ISIS_BIER_EXTENSIONS] Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER Support via ISIS", internet-draft draft-ietf-bier-isis-extensions-04.txt, March 2017.

[OSPF_BIER_EXTENSIONS] Psenak, P., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for Bit Index Explicit Replication", internet-draft draft-ietf-ospf-bier-extensions-05.txt, March 2017.

8. Acknowledgments <Add any acknowledgements>

Authors’ Addresses

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

Email: hooman.bidgoli@nokia.com

Fengman Xu
Verizon
400 International PKWY
Richardson, Tx 75081
US

Email: fengman.xu@verizon.com

Andrew Dolganow
Nokia
750D Chai Chee Rd
06-06, Viva Business Park
Singapore 469004

Email: Andrew.dolganow@nokia.com