BIER Workgroup   H. Bidgoli, Ed.
Internet Draft   J. Kotalwar
Intended status: Standard Track   Nokia
                                                        Fengman Xu
                                                           Verizon
                                                      IJ. Wijnand
                                                  Mankamana Mishra
                                                Cisco Systems, Inc.
                                                         Z. Zhang
                                                  Juniper Networks
                                                      A. Dolganow
                                                      independent

Expires: May 3, 2020   October 31, 2019

                    PIM Signaling Through BIER Core
                    draft-ietf-bier-pim-signaling-08

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 Joins and Prunes
   to be signaled through a BIER core. Allowing PIM routers to run
   traditional PIM multicast services through a BIER core.

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), its areas, and its working groups. Note that
   other groups may also distribute working documents as Internet-

Copyright Notice

Table of Contents

1. Introduction

   Consider large networks deploying traditional PIM multicast service.
   Typically, each portion of these large networks have their own
   mandates and requirements.

   It might be desirable to deploy BIER technology in some part of these
   networks to replace traditional PIM services. In such cases
   downstream PIM states need to be signaled over BIER Domain toward the
   source.

   This draft explains the procedure to signal PIM joins and prunes
   through a BIER Domain, as such enable provisioning of traditional PIM
   services through a BIER Domain.

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. rfc8279] 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 dataplane packets. BFIR is term used for dataplane 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 dataplane packets. BFIR is term used for dataplane packet forwarding.

BBR:

BIER Boundary router. A router between the PIM domain and BIER domain. Maintains PIM adjacency for all routers attached to it on the PIM domain and terminates the PIM adjacency toward the BIER domain.

IBBR:

Ingress BIER Boundary Router. An ingress router from signaling point of view. It maintains PIM adjacency toward the PIM domain and determines if PIM joins and prunes arriving from PIM domain need to be signaled across the BIER domain. If so it terminates the PIM adjacency toward the BIER domain and signals the PIM joins/prunes through the BIER core.

EBBR:

Egress BIER Boundary Router. An egress router in BIER domain from signaling point of view. It terminates the BIER packet and forwards the signaled joins and prunes into PIM Domain.

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 Signaling Through BIER domain

```
                    BBR                     BBR
     |--PIM Domain--|-----BIER domain-----|--PIM domain--|
  S--( A )----------( B ) ---- ( C ) ---- ( D )----------( E )--h


                    EBBR                    IBBR
  Sig    <-----PIM-----|<--BIER Tunneling----|<----PIM------
  (new)


                   bfir                    bfer
       ------------->|--------BIER-------->|-------------> Datapath
                                                          (no change)
```
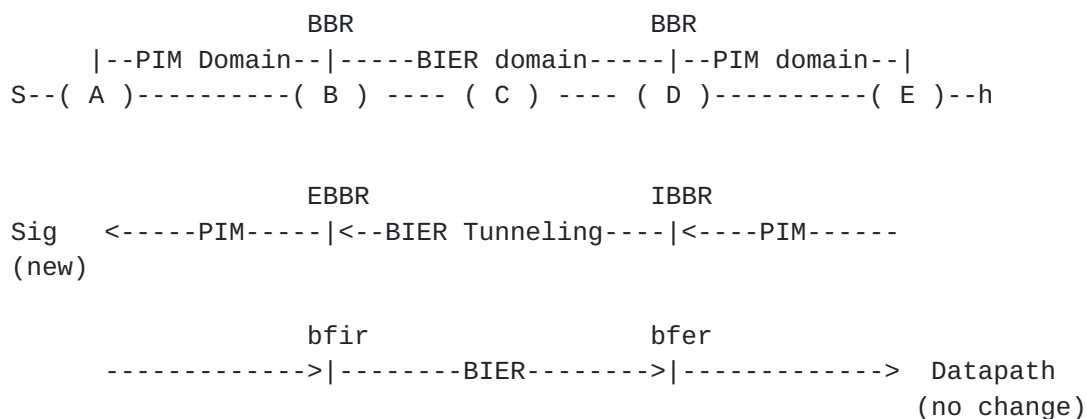
                    Figure 1: BIER boundary router

   As per figure 1, the procedures of PIM signaling is done at the BIER
   boundary router. The BIER boundary router (BBR) are connected to PIM
   capable routers toward the PIM domain and BIER routers toward the
   BIER domain. PIM routers in PIM domain continue to send PIM state
   messages to the BBR. The BBR will create PIM adjacency between all
   the PIM routers attached to it on the PIM domain. That said the BBR
   does not propagate all PIM packets natively into the BIER domain.
   Instead when it determines that the PIM join or prune messages needs
   to be signaled through the BIER domain it will tunnel the PIM packet
   through the BIER network. This tunneling is only done for signaling
   purposes and not for creating a PIM adjacency between the two
   disjoint PIM domains through the BIER domain.

The terminology ingress BBR (IBBR) and egress BBR (EBBR) are relative from signaling point of view.

The ingress BBR will determine if an arriving PIM join or prune needs to be signaled across the BIER domain. While the egress BBR will determine if the arriving BIER packet is a signaling packet and if so it will generate a PIM signaling packet toward its attached PIM domain.

The BFER and BFIR are BBR from datapath point of view. It should be noted the new procedures in this draft are only applicable to signaling and there are no changes from datapath point of view.

3.1. Ingress BBR procedure

IBBR will create PIM adjacency to all PIM routers attached to it toward the PIM domain.

When a PIM join or prune for certain (S,G) arrives, the IBBR first determines weather the join or prune is meant for a source that is reachable through the BIER domain. As an example, this source is located on a disjoint PIM domain that is reachable through the BIER domain. If so the IBBR will try to resolve the source via an EBBR closest to the source.

The procedure to find the EBBR (BFIR from datapath point of view) can be via many mechanisms explained in more detail in upcoming section.

After discovering the EBBR and its BFR-ID (flooded via IGP BIER extension), the IBBR will construct a PIM Join Attribute encoded as TLVs into the Source Address field of the PIM Join Message as per [RFC5384] and include it in PIM signaling message. Two new "BIER IBBR" attributes are define and explained in upcoming section. The PIM Join Attribute is used on EBBR to obtain necessary bier information to build its multicast states. In addition the IBBR will change the PIM signaling packet source IP address to its BIER prefix address (standard PIM procedure). It will also keep the destination address as the well known multicast IP address. It then will construct the BIER header. The signaling packet, in this case the PIM join/prune packet, is encapsulated in the BIER header and transported through BIER domain to EBBR.

The IBBR will track all the PIM interfaces on the attached PIM domain which are interested in a certain (S,G). It creates multicast states for arriving (S,G)s from PIM domain, with incoming interface as BIER

"tunnel" interface and outgoing interface as the PIM domain
interface(s) on which PIM Join(s) were received on.

3.1.1. Determining EBBR on IBBR

As it was explained in previous section, IBBR needs to determine the
EBBR closest to the source. This is needed to encode the BIER header
BitString field to forward the signaling packet through the BIER
domain.

It should be noted, the PIM domains can be either part of the same
IGP area as BIER domain(single area) or are stitched to the BIER
domain via an ABR or ASBR routers. As such on IBBR, there can be many
different procedures to determine the EBBR. Some examples of these
procedures have been provided in Appendix A.

3.1.2. Considering ECMP in EBBR selection

If the lookup for source results into multiple EBBRs, then the EBBR
selection algorithm should ensure that all signaling for a particular
(C-S, C-G) is forward to a single EBBR. How the this selection is
done is vendor specific and beyond this draft. As an example it can
be round robin per (C-S, C-G) or smallest EBBR IP for all (C-S, C-
G)s.

3.1.3. Pim Signaling packet construction at IBBR

To ensure all necessary bier information needed by EBBR is present in
the PIM signaling message, PIM Join Attribute [RFC5384] is used. EBBR
can use this attribute to build its multicast states, as described in
EBBR procedure section. The PIM Join Attribute format is as follow:

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |F|E|  Type=tbd |    Length     | Addr Family  |  BIER info
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-...
```

F bit: The Transitive bit. Specifies whether this attribute is
         transitive or non-transitive. MUST be set to zero. This
         attribute is ALWAYS non-transitive.

E bit: End-of-Attributes bit. Specifies whether this attribute is
         the last. Set to zero if there are more attributes. Set to 1
         if this is the last attribute.

Type: TBD assign by IANA

Length: The length in octets of the attribute value. MUST be set
        to the length in octets of the BIER info +1 octet to account
        for the Address Family field. For Ipv4 AF Length = 7+1 For
        Ipv6 AF Length = 19+1

Addr Family: Signaled PIM Join/Prune address family as defined in
        [RFC7761]

BIER Info: IBBR Prefix (ipv4 or ipv6), SD, bfr-id

3.1.3.1 BIER packet construction at IBBR

The BIER header will be encoded with the BFR-id of the IBBR(with
appropriate bit set in the bitstring) and the PIM signaling packet 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 = IPv4 or IPv6

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

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

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

3.2. Signaling PIM through the BIER domain procedure

Throughout the BIER domain the BIER forwarding procedure is in par

with [RFC 8279](#). No BIER router will examine the BIER packet
encapsulating the PIM signaling 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. EBBR
will remove the BIER header and examine the PIM IPv4 or IPv6
signaling packet farther as per EBBR Procedure section.

3.3. EBBR procedure

After receiving the BIER packet and determining this packet is a
signaling packet, EBBR will remove the BIER header from PIM packet.
The Received PIM join/prune Signaling packet is processed as if it
were received from neighbors on a virtual interface, (i.e. as if the
pim adjacency was presents, regardless of the fact there is no
adjacency)

With same token the EBBR creates a multicast state with incoming
interface as same interface that PIM join packet was forwarded and
outgoing interfaces of BIER tunnel to BFER identified with BFIR-id
and its corresponding Sub-Domain obtained from the BIER header or via
PIM Join Attributes added to the PIM signaling packet by the IBBR.

The EBBR will also build a forwarding table for the arriving (S,G)
using the obtained BFIR-id and the Sub-Domain information, in short
it tracks all IBBRs interested in this (S,G). This is explained in
[section 4.1](#).

It should be noted EBBR will maintain PIM adjacency toward the PIM
domain and all PIM routers which are connected to it.

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

4. Datapath Forwarding

4.1. BFIR tracking of (S,G)

For a specific Source and Group, BFIR (EBBR)should track all the
interested BFERs (IBBRs) via arriving PIM signaling from BIER Domain.
BFIR builds its (s,g)forwarding state with incoming interface (IIF)
as the RPF interface (in PIM domain) towards the source and one of
the outgoing interfaces as for sending to tracked BFERs in the SD.

4.2. Datapath traffic flow

When the multicast data traffic arrives on the BFIR (EBBR) the router
will find all the interested BFERs for that specific (S,G). The
router then constructs the BIERHeader.BitString with all the BFER
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 (S,G) and its outgoing interfaces.

5. PIM-SM behavior

The procedures described in this document can work with ASM as long
as static RP or embedded RP for IPv6 is used. Future drafts would
cover BSR and more complicated SM discovery mechanisms.

It should be noted that this draft only signals PIM Joins and Prunes
through the BIER domain and not any other PIM message types including
PIM Hellos or Asserts. As such functionality related to these other
type of massages will not be possible through a BIER domain with this
draft and future drafts might cover these scenarios. As an example DR
selection should be done in the PIM domain or if the PIM routers
attached to IBBRs are performing DR selection there needs to be a
dedicated PIM interface between these routers.

In case of PIM ASM Static RP or embedded RP for IPv6 the procedure
for LEAFs joining RP is same as above. It should be noted that for
ASM, the EBBR is determined with respect to the RP instead of the
source.

6. Applicability to MVPN

With just minor changes, the above procedures apply to MVPN as well,
with BFIR/BFER/EBBR/IBBR being VPN PEs.

All the PIM related procedures, and the determination of EBBR happens
in the context of a VRF, following procedures for PIM-MVPN.

When a PIM packet arrives from PIM domain attached to the vrf (IBBR),
and it is determine that the source is reachable via the vrf through
the BIER domain, a PIM signaling message is sent via BIER to the
EBBR. In this case usually the PE terminating the PIM-MVPN is the
EBBR. A label is imposed before the BIER header is imposed, and the
"proto" field in the BIER header is set to 1 (for "MPLS packet with
downstream-assigned label at top of stack"). The label is advertised
by the EBBR/BFIR to associate incoming packets to its correct VRF. In
many scenarios a label is already bound to the VRF loopback address
on the EBBR/BFIR and it can be used.

When a multicast data packet is sent via BIER by an EBBR/BFIR, a

label is imposed before the BIER packet is imposed, and the "proto" field in the BIER header is set to 1 (for "MPLS packet with downstream-assigned label at top of stack"). The label is assigned to the VPN consistently on all VRFs [draft-zzhang-bess-mvpn-evpn-aggregation-label].

If the more complicated label allocation scheme is needed for the data packets as specified in [draft-zzhang-bess-mvpn-evpn-aggregation-label], then additional PMSI signaling is needed as specified in [RFC6513].

To support per-area subdomain in this case, the ABRs would need to become VPN PEs and maintain per-VPN state so it is unlikely practical.

7. IANA Considerations

This document contains no actions for IANA.

8. Security Considerations

The procedures of this document do not, in themselves, provide privacy, integrity, or authentication for the control plane or the data plane. For a discussion of the security considerations regarding the use of BIER, please see RFC8279 and RFC8296. Security considerations regarding PIM protocol is based on RFC4601.

9. References

9.1. Normative References

[BIER_ARCH] Wijnands, IJ., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", rfc 8279, October 2016.

9.2. Informative References

[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-11, March 2018.

[ISIS_BIER_EXTENSIONS] Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER Support via ISIS", internet-draft draft-ietf-bier-isis-extensions-11.txt, June 2018.

[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-18.txt, June 2018.

10. Acknowledgments

The authors would like to thank Eric Rosen, Stig Venaas for thier
reviews and comments.

Appendix A

This section provides some examples and routing procedures that can
be used to determine the EBBR on IBBR. It should be noted, the PIM
domains can be either part of the same IGP area as BIER domain(single
area) or are stitched to the BIER domain via an ABR or ASBR routers.
As such on IBBR, there can be many different procedures to determine
the EBBR. Not all procedures are listed below.

A.1. SPF

On IBBR SPF procedures can be used to find the EBBR closest to the
source.

Assuming the BIER domain is consist of all BIER forwarding routers,
SPF calculation can identify the router advertising the prefix for
the source. A post process can find the EBBR by walking from the
advertising router back to the IBBR in the reverse direction of
shortest path tree branch until the first BFR is encountered.

A.2. Indirect next-hop

Alternatively, the route to the source could have an indirect next-
hop that identifies the EBBR. These methods are explained in the
following sections.

A.2.1. Static Route

On IBBR there can be a static route configured for the source, with
source next-hop set as EBBR BIER prefix.

A.2.2. Interior Border Gateway Protocol (iBGP)

Consider the following topology:

```
                   BBR                      BBR
                   EBBR                     IBBR
        |--PIM Domain--|-----bier domain-----|--PIM domain--|
    S--( A )----------( B ) ---- ( C ) ---- ( D )----------( E )--h
```
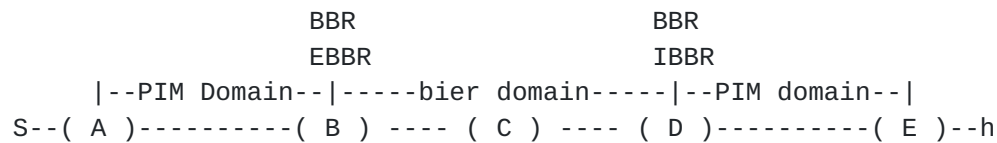
Figure 2

Suppose BGP is enable between EBBR (B) and IBBR (D) and the PIM
Domain routes are redistributed to the BIER domain via BGP. This
would include the Multicast Source IP address (S), which resides in
the PIM Domain. In such case BGP should use the same loopback
interface as its next-hop as the BBR prefix. This will ensure that
all PIM domain routes, including the Multicast Source IP address (S)
are resolve via BBR's bier prefix id as their next-hop. When the host
(h) triggers a PIM join message to IBBR (D), IBBR tries to resolve
(S). It resolves (S) via BGP installed route and realizes its next-
hop is EBBR (B). IBBR will use this next-hop (B) to find its
corresponding BIER bit index.

This procedure is inline with RFC6826 mLDP in-band signaling section
2.

A.3. Inter-area support

If each area has its own BIER sub-domain, the above procedure for
post-SPF could identify one of the ABRs and the EBBR. If a sub-domain
spans multiple areas, then additional procedures as described in A.2
is needed.

A.3.1. Inter-area Route summarization

In a multi-area topology, a BIER sub-domain can span a single area.
Suppose this single area is constructed entirely of Bier capable
routers and the ABRs are the BIER Boundary Routers attaching the BIER
sub-domain in this area to PIM domains in adjacent areas. These BBRs
can summarize the PIM domain routes via summary routes, as an example
for OSPF, a type 3 summary LSAs can be used to advertise summary
routes from a PIM domain area to the BIER area. In such scenarios the
IBBR can be configured to look up the Source via IGP database and use
the summary routes and its Advertising Router field to resolve the
EBBR. The IBBR needs to ensure that the IGP summary route is
generated by a BFR. This can be achieved by ensuring that bier sub-
tlv exists for this route. If multiple BBRs (ABRs) have generated the
same summary route the lowest Advertising Router IP can be selected
or a vendor specific hashing algorithm can select the summary route
from one of the BBRs.

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

Jayant Kotalwar
Nokia
380 N Bernardo Ave
Mountain View, CA 94043
US

Email: jayant.kotalwar@nokia.com

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

Email: Andrew.dolganow@nokia.com

IJsbrand Wijnands
Cisco System
De Kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Mankamana Mishra
Cisco System
821 alder drive
Milpitas California
USA

Email: mankamis@cisco.com

Zhaohui Zhang
Juniper Networks
USA

EMail: zzhang@juniper.net