Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 10, 2021

H. Bidgoli, Ed.
Nokia
F. Xu
Verizon
J. Kotalwar
Nokia
I. Wijnands
M. Mishra
Cisco System
Z. Zhang
Juniper Networks
July 09, 2020

PIM Signaling Through BIER Core
draft-ietf-bier-pim-signaling-09

Abstract

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.

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 January 10, 2021.

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 performs BIER encapsulation). Each BFIR must have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plane 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 have a valid BFR-id assigned. In this draft BIER will be used for forwarding and tunneling of control plane 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

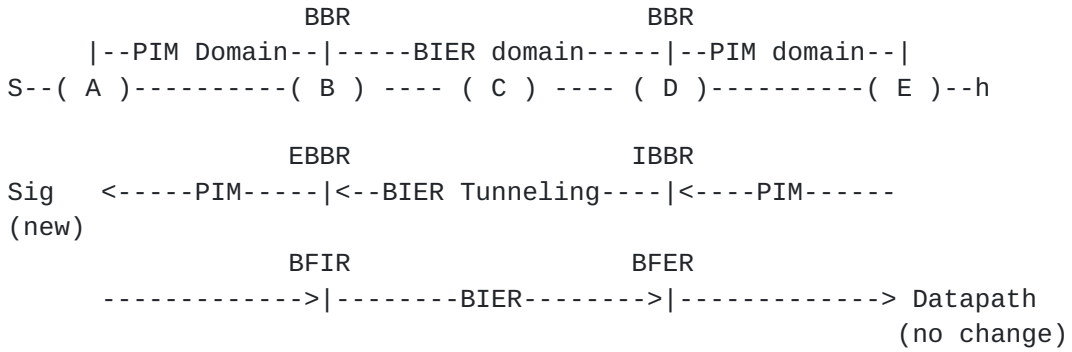

Figure 1: BIER boundry router

As per figure 1, the procedures of PIM signaling is done at the BIER boundary router. The BIER boundary routers (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 whether the join or prune is meant for a source that is reachable through the BIER domain. As an example, this source is located in 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, the IBBR will include a new PIM Join Attribute in the Join/prune message as per [RFC5384]. Two new "BIER IBBR" attributes are defined 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 BIER signaling message, a new PIM Join Attribute [RFC5384] is used. EBBR can use this attribute to build its multicast states, as described in EBBR procedure section. This new PIM join Attribute is added to PIM signaling message on the IBBR. Its format is as follow:

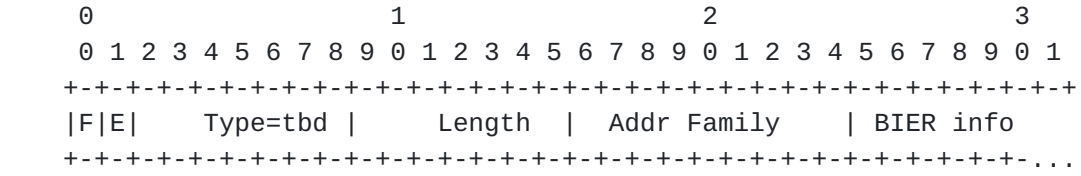

Figure 2: PIM Join Attribute

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 as per below figure

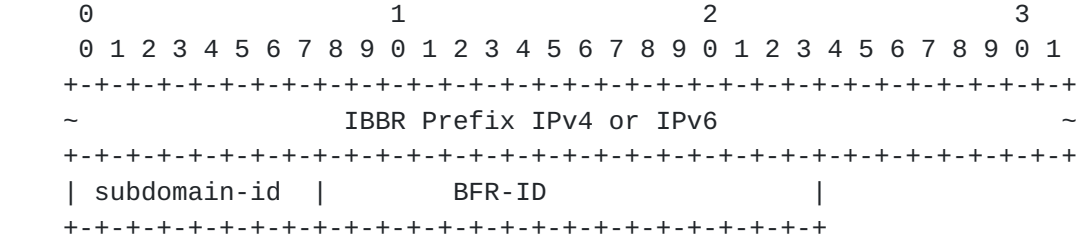

Figure 3: PIM Join Attribute detail

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.

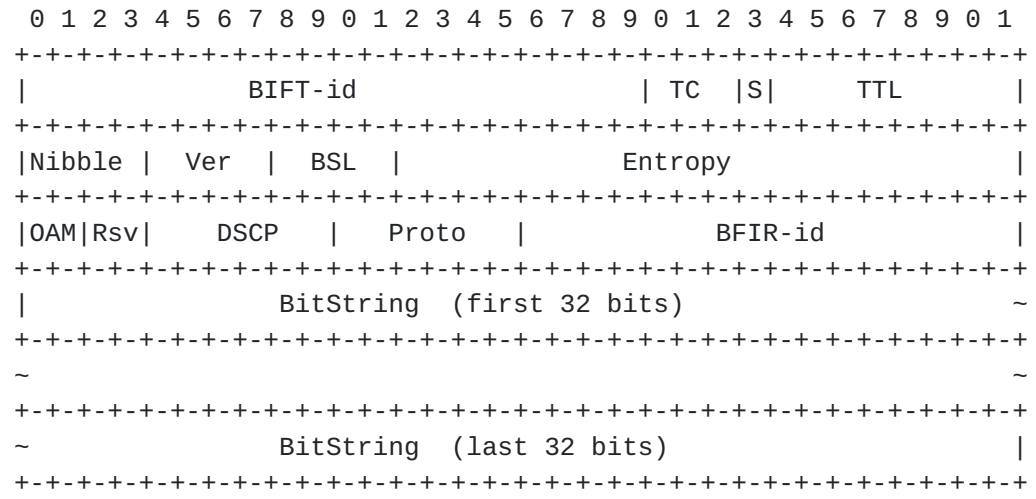

Figure 4: BIER header

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 on par with RFC 8279. No BIER router will examine the BIER packet encapsulating the PIM signaling packet. As such there is no multicast state built 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

EBBR will remove the BIER header and determine this is a signaling 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)

The EBBR will build a forwarding table for the arriving (S,G) using the obtained BFIR-id and the Sub-Domain information from BIER Header and/or the PIM join Attributes added to the PIM Signaling packet. In short it tracks all IBBRs interested in this (S,G). This is explained in section 4.1.

The multicast state on EBBR will contain PIM domain incoming interfaces, according to PIM specification and outgoing interfaces based on the above build forwarding table.

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 leaves joining RP is same as above. It should be noted that for ASM, the EBBRs are 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 determined 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 RFC 7761.

9. Acknowledgments

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

10. References

10.1. Normative References

[RFC8279] "Wijnands, IJ., Rosen, E., Dolganow, A., Przygienda, T. and S. Aldrin, "Multicast using Bit Index Explicit Replication"", October 2016.

10.2. 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.

[RFC8401] "Ginsberg, L., Przygienda, T., Aldrin, S., and Z. Zhang, "BIER Support via ISIS"", June 2018.

[RFC8444] "Psenak, P., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPF Extensions for Bit Index Explicit Replication"", June 2018.

Appendix A.

This appendix 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 consists 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:

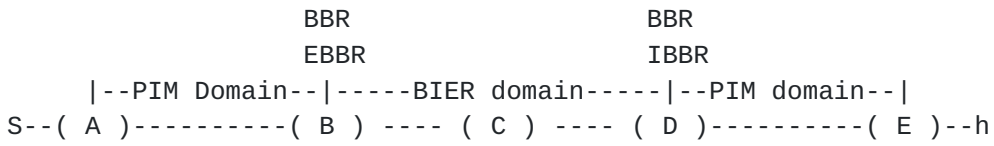

Figure 2

Figure 5: Static Route

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

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.

Authors' Addresses

Hooman Bidgoli (editor)
Nokia
Ottawa
Canada

Email: hooman.bidgoli@nokia.com

Fengman Xu
Verizon
Richardson
US

Email: fengman.xu@verizon.com

Jayant Kotalwar
Nokia
Montain View
US

Email: jayant.kotalwar@nokia.com

IJsbrand Wijnands
Cisco System
Diegem
Belgium

Email: ice@cisco.com

Mankamana Mishra
Cisco System
Milpitas
USA

Email: mankamis@cisco.com

Zhaohui Zhang
Juniper Networks
Boston
USA

Email: zzhang@juniper.com