Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 26 January 2022

H. Bidgoli, Ed.
Nokia
F. Xu
Verizon
J. Kotalwar
Nokia
I. Wijnands
M. Mishra
Cisco System
Z. Zhang
Juniper Networks
25 July 2021

PIM Signaling Through BIER Core
draft-ietf-bier-pim-signaling-12

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 the BIER Domain toward the source.

This draft specifies the procedure to signal PIM join/prune messages through a BIER Domain, as such enabling the provisioning of traditional PIM services through a BIER Domain. These procedures are valid for forwarding PIM join/prune messages to the Source (SSM) or Rendezvous Point (ASM).

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 26 January 2022.

Table of Contents

1. Introduction

It might be desirable to simplify/upgrade some part of an existing network to BIER technology, removing any legacy multicast protocols like PIM. This simplification should be done with minimum interruption or disruption to the other parts of the network from singling, services and software upgrade point of view. To do so this draft is specifies procedures for signaling multicast join and prune messages over the BIER domain, this draft is not trying to create FULL PIM adjacency over a BIER domain between two PIM nodes. The PIM adjacency is terminated at BIER edge routers and only join/prune signaling messages are transported over the BIER network. It just so happened that this draft chose signaling messages to be in par with PIM join/prune messages. These signaling messages are forwarded upstream toward the BIER edge router on path to the Source or Rendezvous point. These signaling messages are encapsulated in a BIER header.

2. Conventions used in this document

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.

2.1. Definitions

An understanding of the BIER architecture [RFC8279] and the related terminology is expected. The following are some of the new definitions used in this draft.

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 signals join/prune messages across the BIER domain to EBBR as needed.

EBBR:

Egress BIER Boundary Router. An egress router in BIER domain from signaling point of view. It maintains PIM adjacency to all upstream PIM routers. It terminates the BIER signaling packets and creates necessary PIM join/prune messages into PIM Domain.

3. PIM Signaling Through BIER domain

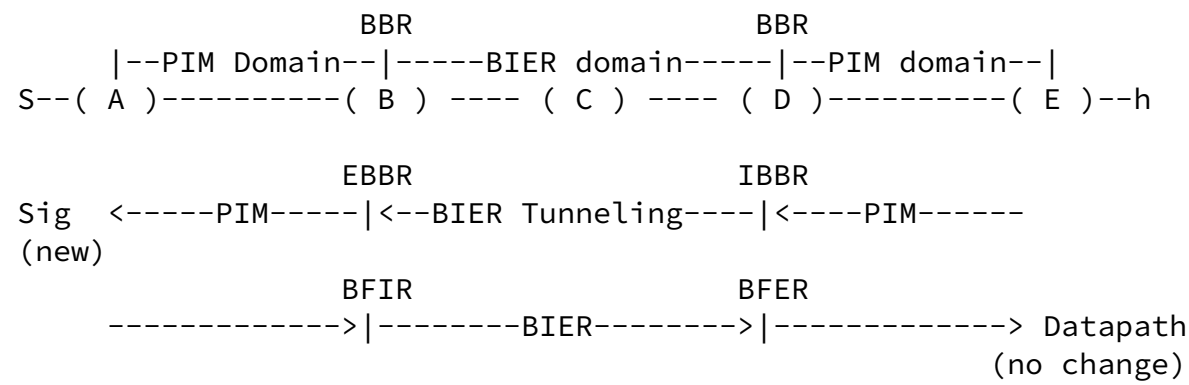


Figure 1: BIER boundary router

Figure 1 illustrates the operation of the BIER Boundary router (BBR). BBRs are connected to PIM routers in the PIM domain and BIER routers in 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. Each BBR determines if a BIER Signaling Join or Prune message needs to be transmitted through the BIER domain. This draft has chosen these BIER signaling messages to be PIM join/prune message, as such an implementation could chose to tunnel actual PIM join/prune messages

through 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) is relative only from a signaling point of view.

The egress BBR will determine if the arriving BIER packet is a signaling packet and if so it will generate a PIM join/prune packet toward its attached PIM domain.

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

The IBBR maintains a PIM adjacency [RFC7761] with any PIM router attached to it on the PIM domain.

When a PIM Join or Prune message is received, the IBBR determines whether the Source or RP is reachable through the BIER domain. The EBBR is the BFR through which the Source or RP is reachable. In PIM terms [RFC7761], the EBBR is the RPF Neighbor, and the RPF Interface is the BIER "tunnel" used to reach it. The mechanisms used to find the EBBR are outside the scope of this document and there can be many mechanism depending on if the source or RP are in same area or autonomous system (AS) or in different area or AS -- some examples are provided in Appendix A.

If the lookup for source or rendezvous point results into multiple EBBRs, different IBBRs could choose different EBBRs for the same flow. As long as a unique IBBR chooses a unique EBBR for the same flow. On downstream these EBBRs will send traffic to their corresponding IBBRs.

After discovering the EBBR and its BFR-id, the IBBR MUST use the BIER Information Vector (Section 3.1.1) which is a PIM Join Attribute type [RFC5384]. The EBBR uses this attribute to obtain the necessary BIER information to build its multicast state. The signaling packet, in this case a PIM Join/Prune message, is encapsulated in the BIER Header and forwarded through the BIER domain to the EBBR. The source

address of the PIM packets MUST be set to IBBR local BFR-prefix. The destination address MUST be set to ALL-PIM-ROUTERS [RFC7761].

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 join messages 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. New Pim Join Attribute, BIER Information Vector

The new PIM Join Attribute " BIER Information Vector" is defined as follow based on [RFC5384]

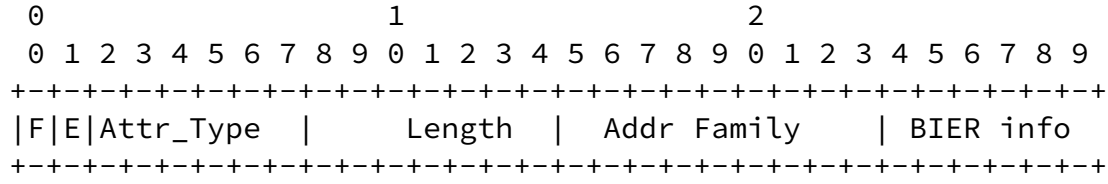


Figure 2: PIM Join Attribute

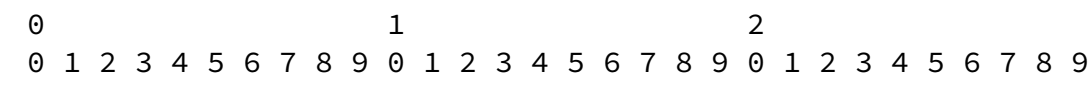
F bit: Transitive Attribute, as specified in [RFC5384]. MUST be set to zero as this attribute is always non-transitive. If EBBR receives this attribute type with the F bit set it must discard the Attribute.

E bit: End of Attributes, as specified in [RFC5384]

Attr_Type: TBD assign by IANA.

Length: Length of the value field, as specified in [RFC5384]. MUST be set to the length of the BIER Info field + 1. For IPv4 the length is 8, and 20 for IPv6. Incorrect length value compare to the Addr Family must be discarded.

Addr Family: PIM address family as specified in [RFC7761]. Unrecognized Address Family must be discarded.



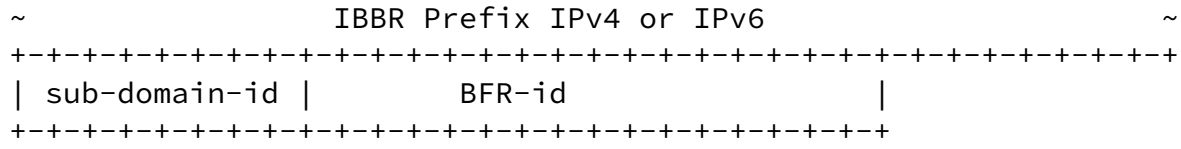


Figure 3: PIM Join Attribute detail

BIER Info: IBBR's BFR-prefix (IPv4 or IPv6), sub-domain-id, BFR-id

3.1.1.1. BIER packet construction at the 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.

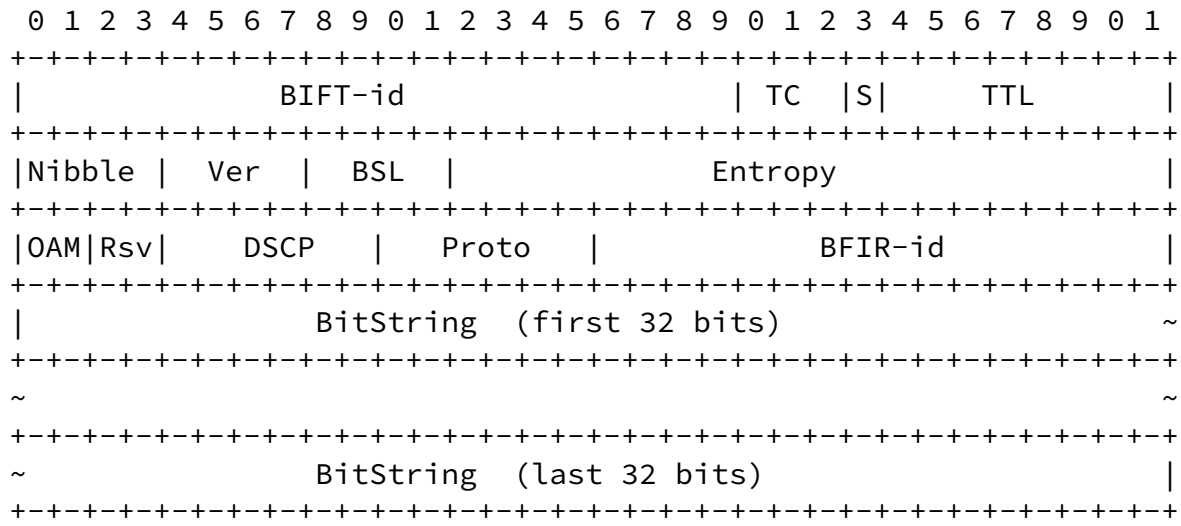


Figure 4: BIER header

BIERHeader.Proto = PIM Addrress Family

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

BIERHeader.BFIR-id = BFR-Id of the BBR originating the encapsulated signaling 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 according to [RFC8279]. No BIER router will examine the BIER the 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 EBBR indicated by the BIERHeader.Bitstring. Only this targeted EBBR router will remove the BIER header and examine the PIM IPv4 or IPv6 signaling packet further as per EBBR Procedure section.

3.3. EBBR procedure

EBBR removes the BIER Header and determine this is a signaling packet. The Received signaling packet, PIM join/prune message, is processed as if it were received from neighbors on a virtual interface, (i.e. as if the pim adjacency was present, regardless of the fact that 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 signaling packet. In short it tracks all IBBRs interested in this (S,G). For a specific Source and Group, EBBR SHOULD track all the interested IBBRs via signaling messages arriving from the BIER Domain. BFER builds its (s,g) forwarding state with incoming interface (IIF) as the Reverse Path Forwarding (RPF) interface (in attached PIM domain) towards the source or rendezvous point. The outgoing interfaces include a virtual interface that represent BIER forwarding to tracked IBBRs.

The EBBR maintains a PIM adjacency [RFC7761] with any PIM router attached to it on the PIM domain. At this point the end-to-end multicast traffic flow setup is complete.

4. Datapath Forwarding

4.1. Datapath traffic flow

When multicast data traffic arrives on the BFIR (EBBR) it forwards the traffic, through the BIER domain, to all interested IBBRs following the procedures specified in [RFC8279]. The BFER(s) (IBBR(s)) also follow the procedures in [RFC8279] and forward the multicast packet through its outgoing interface(s).

5. PIM-SM behavior

The procedures described in this document can be used with Any-Source Mutlicast (ASM) as long as a static Rendezvous Point (RP) or embedded RP for IPv6 is used[RFC3956].

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. The register messages are unicas encapsulatedt from the source to RP as such they are forwarded without these procedures.

In case of PIM ASM Static RP or embedded RP for IPv6 the procedure for leaves joining RP is the 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-01].

If the more complicated label allocation scheme is needed for the data packets as specified in [draft-zzhang-bess-mvpn-evpn-aggregation-label-01], 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

IANA is requested to assign a value (TBD) to the BIER Information Vector PIM Join Attribute from the PIM Join Attribute Types registry.

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]. The security consideration for [RFC7761] aslso apply.

9. Acknowledgments

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

10. References

10.1. Normative References

[RFC2119] "S. Brandner, "Key words for use in RFCs to Indicate Requirement Levels"", March 1997.

[RFC5384] "A. Boers, I. Wijnands, E. Rosen, "PIM Join Attribute Format"", November 2008.

[RFC7761] "B.Fenner, M.Handley, H. Holbrook, I. Kouvelas, R. Parekh, Z.Zhang "PIM Sparse Mode"", March 2016.

[RFC8174] "B. Leiba, "ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words"", May 2017.

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

[RFC8296] "IJ. Wijnands, E. Rosen, A. Dolganow, J. Yantsura, S. Aldrin, I. Meilik, "Encapsulation for BIER"", January 2018.

10.2. Informative References

[draft-zzhang-bess-mvpn-evpn-aggregation-label-01] "Z. Zhang, E. Rosen, W. Lin, Z. Li, I.Wijnands, "MVPN/EVPN Tunnel Aggregation with Common labels"", April 2018.

[RFC3956] "P.Savola, B. Haberman "Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address"".

[RFC6513] "E. Rosen, R. Aggarwal, "Multicast in MPLS/BGP IP VPNs"", November 2008.

Appendix A. Determining the EBBR

This appendix provides some examples of routing procedures that can be used to determine the EBBR at the IBBR.

A.1. Link-State Protocols

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

A static route to the source can be configured on the IBBR with the next-hop set as the EBBR's BFR-prefix.

A.2.2. Interior Border Gateway Protocol (iBGP)

Consider the following topology:

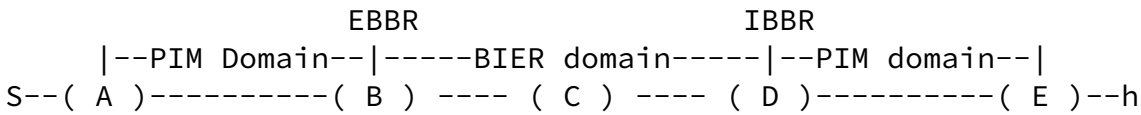


Figure 5: Static Route

Suppose BGP is enable between EBBR (B) and IBBR (D) and the PIM Domain routes are redistributed to the BIER domain via BGP, performing next-hop-self for these routes. This would include the Multicast Source IP address (S). In such case BGP should use the same next-hop as the EBBR BIER prefix. This will ensure that all PIM domain routes, including the Multicast Source IP address (S) are resolve via EBBR's BIER prefix address. 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).

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,
United States of America

Email: fengman.xu@verizon.com

Jayant Kotalwar
Nokia
Montain View,
United States of America

Email: jayant.kotalwar@nokia.com

IJsbrand Wijnands
Cisco System
Diegem
Belgium

Email: ice@cisco.com

Mankamana Mishra
Cisco System
Milpitas,
United States of America

Email: mankamis@cisco.com

Zhaohui Zhang
Juniper Networks

Boston,
United States of America

Email: zzhang@juniper.com