

BIER
Internet-Draft
Intended status: Standards Track
Expires: May 7, 2020

Z. Zhang
A. Przygienda
Juniper Networks
A. Sajassi
Cisco Systems
J. Rabadan
Nokia
November 4, 2019

EVPN BUM Using BIER
draft-ietf-bier-evpn-02

Abstract

This document specifies protocols and procedures for forwarding broadcast, unknown unicast and multicast (BUM) traffic of Ethernet VPNs (EVPN) using Bit Index Explicit Replication (BIER).

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.

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 May 7, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Terminologies	3
2. Use of the PMSI Tunnel Attribute	4
2.1. Auxiliary Information	5
2.2. Explicit Tracking	6
2.2.1. Using IMET/SMET routes	6
2.2.2. Using S-PMSI/Leaf A-D Routes	6
2.3. MPLS Label in PTA	7
3. Multihoming Split Horizon	8
4. Data Plane	8
4.1. Encapsulation and Transmission	8
4.1.1. At a BFIR that is an Ingress PE	8
4.1.2. At a BFIR that is a P-tunnel Segmentation Point	10
4.2. Disposition	11
4.2.1. At a BFER that is an Egress PE	11
4.2.2. At a BFER that is a P-tunnel Segmentation Point	11
5. IANA Considerations	11
6. Security Considerations	11
7. Acknowledgements	12
8. References	12
8.1. Normative References	12
8.2. Informative References	13
Authors' Addresses	14

1. Introduction

[RFC7432] and [RFC8365] specify the protocols and procedures for Ethernet VPNs (EVPNs). For broadcast, unknown unicast and multicast (BUM) traffic, provider/underlay tunnels (referred to as P-tunnels) are used to carry the BUM traffic. Several kinds of tunnel technologies can be used, as specified in [RFC7432].

Bit Index Explicit Replication (BIER) ([RFC8279]) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. The

purpose of this document is to specify the protocols and procedures to transport EVPN BUM traffic using BIER.

The EVPN BUM procedures specified in [RFC7432] and extended in [I-D.ietf-bess-evpn-bum-procedure-updates], [I-D.ietf-bess-evpn-igmp-mld-proxy], and [I-D.zzhang-bess-mvpn-evpn-cmcast-enhancements] are much aligned with MVPN procedures. As such, this document is also very much aligned with [RFC8556]. For terseness, some background, terms and concepts are not repeated here. Additionally, some text is borrowed verbatim from [RFC8556].

1.1. Terminologies

- o BFR: Bit-Forwarding Router.
- o BFIR: Bit-Forwarding Ingress Router.
- o BFER: Bit-Forwarding Egress Router.
- o BFR-Prefix: An IP address that uniquely identifies a BFR and is routeable in a BIER domain.
- o C-S: A multicast source address, identifying a multicast source located at a VPN customer site.
- o C-G: A multicast group address used by a VPN customer.
- o C-flow: A customer multicast flow. Each C-flow is identified by the ordered pair (source address, group address), where each address is in the customer's address space. The identifier of a particular C-flow is usually written as (C-S,C-G). Sets of C-flows can be identified by the use of the "C-*" wildcard (see [RFC6625]), e.g., (C-*,C-G).
- o P-tunnel. A multicast tunnel through the network of one or more SPs. P-tunnels are used to transport MVPN multicast data
- o IMET Route: Inclusive Multicast Ethernet Tag Auto-Discovery route. Carried in BGP Update messages, these routes are used to advertise the "default" P-tunnel for a particular broadcast domain.
- o SMET Route: Selective Multicast Ethernet Tag Auto-Discovery route. Carried in BGP Update messages, these routes are used to advertise the C-flows that the advertising PE is interested in.
- o S-PMSI A-D route: Selective Provider Multicast Service Interface Auto-Discovery route. Carried in BGP Update messages, these

routes are used to advertise the fact that particular C-flows are bound to (i.e., are traveling through) particular P-tunnels.

- o PMSI Tunnel attribute (PTA). This BGP attribute carried is used to identify a particular P-tunnel. When C-flows of multiple VPNs are carried in a single P-tunnel, this attribute also carries the information needed to multiplex and demultiplex the C-flows.

2. Use of the PMSI Tunnel Attribute

[RFC7432] specifies that Inclusive Multicast Ethernet Tag (IMET) routes carry a PMSI Tunnel Attribute (PTA) to identify the particular P-tunnel to which one or more BUM flows are being assigned, the same as specified in [RFC6514] for MVPN. [RFC8556] specifies the encoding of PTA for use of BIER with MVPN. Much of that specification is reused for use of BIER with EVPN and much of the text below is borrowed verbatim from [RFC8556].

The PMSI Tunnel Attribute (PTA) contains the following fields:

- o "Tunnel Type". The same codepoint 0x0B that IANA has assigned for [RFC8556] for the new tunnel type "BIER" is used for EVPN as well.
- o "Tunnel Identifier". When the "tunnel type" field is "BIER", this field contains two subfields. The text below is exactly as in [RFC8556].
 - 1 The first subfield is a single octet, containing the sub-domain-id of the sub-domain to which the BFIR will assign the packets that it transmits on the PMSI identified by the NLRI of the IMET, S-PMSI A-D, or per-region I-PMSI A-D route that contains this PTA. How that sub-domain is chosen is outside the scope of this document.
 - 2 The second subfield is a two-octet field containing the BFR-id, in the sub-domain identified in the first subfield, of the router that is constructing the PTA.
 - 3 The third subfield is the BFR-Prefix (see [RFC8279]) of the originator of the route that is carrying this PTA. This will either be a /32 IPv4 address or a /128 IPv6 address. Whether the address is IPv4 or IPv6 can be inferred from the total length of the PMSI Tunnel attribute.

The BFR-prefix need not be the same IP address that is carried in any other field of the x-PMSI A-D route, even if the BFIR is the originating router of the x-PMSI A-D route.

- o "MPLS label". For EVPN-MPLS [RFC7432], this field contains an upstream-assigned MPLS label. It is assigned by the BFIR. Constraints on the way in which the originating router selects this label are discussed in Section 2.3. For EVPN-VXLAN/NVGRE/GENEVE [RFC8365], this field is a 24-bit VNI/VSID of global significance.
- o "Flags". When the tunnel type is BIER, two of the flags in the PTA Flags field are meaningful. Details about the use of these flags can be found in Section 2.2.
 - * "Leaf Info Required per Flow (LIR-pF)"
[I-D.ietf-bess-mvpn-expl-track]
 - * "Leaf Info Required Bit (LIR)"
- o "Auxiliary Information". This is optional, present if the total length of the PTA is larger than the sum of lengths of the fields before this one. It is in the form of a series of TLVs.

Note that if a PTA specifying "BIER" is attached to an IMET, S-PMSI A-D, or per-region I-PMSI A-D route, the route MUST NOT be distributed beyond the boundaries of a BIER domain. That is, any routers that receive the route must be in the same BIER domain as the originator of the route. If the originator is in more than one BIER domain, the route must be distributed only within the BIER domain in which the BFR-Prefix in the PTA uniquely identifies the originator. As with all MVPN routes, distribution of these routes is controlled by the provisioning of Route Targets.

2.1. Auxiliary Information

For the "Auxiliary Information", one TLV is defined in this document - Tunnel Encapsulation TLV. The value part of the TLV is a Tunnel TLV as defined in [I-D.ietf-idr-tunnel-encaps].

This MAY be used when VXLAN/NVGRE/GENEVE encapsulation with an IP header (and UDP header in case of VXLAN/GENVE) is the BIER payload. Normally that is not needed with BIER, except when BIER PHP [I-D.ietf-bier-php] is used and the encapsulation (after BIER header is popped) between the BIER Penultimate Hop and the egress PE does not have a way to indicate the next header is VXLAN/NVGRE/GENEVE. In

that case the full VXLAN/NVGRE/GENEVE encapsulation with an IP header MUST be used. The tunnel type (VXLAN/NVGRE/GENEVE), endpoint, and some tunnel specific information MAY be specified in the Tunnel TLV or MAY be provisioned on PEs. The tunnel endpoint MUST be an IP multicast address and the receiving egress PE MUST be set up to receive and process packets addressed to the address. The same multicast address can be used for all BDs, as the the inner VXLAN/NVGRE/GENEVE header will be used to identify BDs.

2.2. Explicit Tracking

When using BIER to transport an EVPN BUM data packet through a BIER domain, an ingress PE functions as a BFIR (see [RFC8279]). The BFIR must determine the set of BFERs to which the packet needs to be delivered. This can be done in either of two ways in the following two sections.

2.2.1. Using IMET/SMET routes

Both IMET and SMET (Selective Multicast Ethernet Tag [I-D.ietf-bess-evpn-igmp-mld-proxy]) routes provide explicit tracking functionality.

For an inclusive PMSI, the set of BFERs to deliver traffic to includes the originators of all IMET routes for a broadcast domain. For a selective PMSI, the set of BFERs to deliver traffic to includes the originators of corresponding SMET routes.

The SMET routes do not carry a PTA. When an ingress PE sends traffic on a selective tunnel using BIER, it uses the upstream assigned label that is advertised in its IMET route.

Only when selectively forwarding is for all flows without tunnel segmentation, SMET routes are used without the need for S-PMSI A-D routes. Otherwise, the procedures in the following section apply.

2.2.2. Using S-PMSI/Leaf A-D Routes

There are two cases where S-PMSI/Leaf A-D routes are used as discussed in the following two sections.

2.2.2.1. Selective Forwarding Only for Some Flows

With the SMET procedure, a PE advertises an SMET route for each (C-S,C-G) or (C-*,C-G) state that it learns on its ACs, and each SMET route is tracked by every PE in the same broadcast domain. It may be desired that SMET routes are not used to reduce the burden of explicit tracking.

In this case, most multicast traffic will follow the I-PMSI (advertised via IMET route) and only some flows follow S-PMSIs. To achieve that, S-PMSI/Leaf A-D routes can be used, as specified in [I-D.ietf-bess-evpn-bum-procedure-updates].

The rules specified in Section 2.2.1 and Section 2.2.2 of [RFC8556] apply.

2.2.2.2. Tunnel Segmentation

Another case where S-PMSI/Leaf A-D routes are necessary is tunnel segmentation, which is also specified in [I-D.ietf-bess-evpn-bum-procedure-updates], and further clarified in [I-D.zhang-bess-mvpn-evpn-cmcast-enhancements] for segmentation with SMET routes. This is only applicable to EVPN-MPLS.

The rules specified in Section 2.2.1 of [RFC8556] apply. Section 2.2.2 of [RFC8556] do not apply, because similar to MVPN, the LIR-pF flag cannot be used with segmentation.

2.2.2.3. Applicability of Additional MVPN Specifications

As with the MVPN case, Section "3. Use of the PMSI Tunnel Attribute in Leaf A-D routes" of [RFC8556] apply.

Notice that, [RFC8556] refers to procedures specified in [RFC6625] and [I-D.ietf-bess-mvpn-expl-track]. Those two documents were specified for MVPN but are actually applicable to IP multicast payload in EVPN as well.

2.3. MPLS Label in PTA

Rules in section 2.1 of [RFC8556] apply, EXCEPT the following three bullets (they do NOT apply to EVPN) in that section:

- o If the two routes do not have the same Address Family Identifier (AFI) value, then their respective PTAs MUST contain different MPLS label values. This ensures that when an egress PE receives a data packet with the given label, the egress PE can infer from the label whether the payload is an IPv4 packet or an IPv6 packet.
- o If the BFIR is an ingress PE supporting MVPN extranet ([RFC7900]) functionality, and if the two routes originate from different VRFs on this ingress PE, then the respective PTAs of the two routes MUST contain different MPLS label values.
- o If the BFIR is an ingress PE supporting the "Extranet Separation" feature of MVPN extranet (see Section 7.3 of [RFC7900]), and if

one of the routes carries the "Extranet Separation" extended community but the other does not, then the respective PTAs of the two routes MUST contain different MPLS label values.

3. Multihoming Split Horizon

For EVPN-MPLS, [RFC7432] specifies the use of ESI labels to identify the ES from which a BUM packet originates. A PE receiving that packet from the core side will not forward it to the same ES. The procedure works for both Ingress Replication (IR) and RSVP-TE/mLDP P2MP tunnels, using downstream- and upstream-assigned ESI labels respectively. For EVPN-VXLAN/NVGRE/GENEVE, [RFC8365] specifies local-bias procedures, with which a PE receiving a BUM packet from the core side knows from encapsulation the ingress PE so it does not forward the packet to any multihoming ESes that the ingress PE is on, because the ingress PE already forwarded the packet to those ESes, regardless of whether the ingress PE is a DF for those ESes.

With BIER, the local-bias procedure still applies for EVPN-VXLAN/NVGRE/GENEVE as the BFIR-id in the BIER header identifies the ingress PE. For EVPN-MPLS, ESI label procedures also still apply though two upstream assigned labels will be used (one for identifying the broadcast domain and one for identifying the ES) - the same as in the case of using a single P2MP tunnel for multiple broadcast domains. The BFIR-id in the BIER header identifies the ingress PE that assigned those two labels.

4. Data Plane

Similar to MVPN, the EVPN application plays the role of the "multicast flow overlay" as described in [RFC8279].

4.1. Encapsulation and Transmission

A BFIR could be either an ingress PE or a P-tunnel segmentation point. The procedures are slightly different as described below.

4.1.1. At a BFIR that is an Ingress PE

To transmit a BUM data packet, an ingress PE first determines the route matched for transmission and routes for tracking leaves according to the following rules.

1. If selective forwarding is not used, or it is not an IP Multicast packet after the ethernet header, the IMET route originated for the BD by the ingress PE is the route matched for transmission. Leaf tracking routes are all other received IMET routes for the BD.

2. Otherwise, if selective forwarding is used for all IP Multicast traffic based on SMET routes, the IMET route originated for the BD by the ingress PE is the route matched for transmission. Received SMET routes for the BD that best match the source and destination IP address are leaf tracking routes.
3. Otherwise, route matched for transmission is the S-PMSI A-D route originated by the ingress PE for the BD, that best matches the packet's source and destination IP address and has a PTA specifying a valid tunnel type that is not "no tunnel info". Leaf tracking routes are determined as following:
 - 1) If the match for transmission route carries a PTA that has the LIR flag set but does not have the LIR-pF flag set, the routes matched for tracking are Leaf A-D routes whose "route key" field is identical to the NLRI of the S-PMSI A-D route.
 - 2) If the match for transmission route carries a PTA that has the LIR-pF flag, the leaf tracking routes are Leaf A-D routes whose "route key" field is derived from the NLRI of the S-PMSI A-D route according to the procedures described in Section 5.2 of [I-D.ietf-bess-mvpn-expl-track].

Note that in both cases, SMET routes may be used in lieu of Leaf A-D routes, as a PE may omit the Leaf A-D route in response to an S-PMSI A-D route with LIR or LIR-pF bit set, if an SMET route with the corresponding Tag, Source and Group fields is already originated [I-D.ietf-bess-evpn-bum-procedure-updates]. In particular, in the second case above, even though the SMET route does not have a PTA attached, it is still considered as a Leaf A-D route in response to a wildcard S-PMSI A-D route with the LIR-pF bit set.

4. Otherwise, route matched for transmission and leaf tracking routes are determined as in rule 1.

If no route is matched for transmission, the packet is not forwarded onto a p-tunnel. If the tunnel that the ingress determines to use based on the route matched for transmission (and considering interworking with PEs that do not support certain tunnel types per procedures in [I-D.ietf-bess-evpn-igmp-mld-proxy]) requires leaf tracking (e.g. Ingress Replication, RSVP-TE P2MP tunnel, or BIER) but there are no leaf tracking routes, the packet will not be forwarded onto a p-tunnel either.

The following text assumes that BIER is the determined tunnel type. The ingress PE pushes an upstream assigned ESI label per [RFC7432] if the following conditions are all met:

- o The packet is received on a multihomed ES.
- o It's EVPN-MPLS.
- o ESI label procedure is used for split-horizon.

The MPLS label from the PTA of the route matched for transmission is then pushed onto the packet's label stack for EVPN-MPLS. For EVPN-VXLAN/NVGRE/GENEVE, a VXLAN/NVGRE/GENEVE header is prepended to the packet with the VNI/VSID set to the value in the PTA's label field, and then an IP/UDP header is prepended if needed (e.g. for PHP purpose).

Then the packet is encapsulated in a BIER header and forwarded, according to the procedures of [RFC8279] and [RFC8296]. See especially Section 4, "Imposing and Processing the BIER Encapsulation", of [RFC8296]. The "Proto" field in the BIER header is set to 2 in case of EVPN-MPLS, or a value to be assigned in case of EVPN-VXLAN/NVGRE/GENEVE (Section 5) when IP header is not used, or 4/6 if IP header is used for EVPN-VXLAN/NVGRE/GENEVE.

In order to create the proper BIER header for a given packet, the BFIR must know all the BFERs that need to receive that packet. This is determined from the set of leaf tracking routes.

4.1.2. At a BFIR that is a P-tunnel Segmentation Point

In this case, the encapsulation for upstream segment of the p-tunnel includes (among other things) a label that identifies the x-PMSI or IMET A-D route that is the match for reception on the upstream segment. The segmentation point re-advertised the route into one or more downstream regions. Each instance of the re-advertised route for a downstream region has a PTA that specify tunnel information that is the same as or different from that of the route for a different region. For any particular downstream region, the route matched for transmission is the re-advertised route, and the leaf tracking routes are determined as following if needed for the tunnel type:

- o If the route matched for transmission is an x-PMSI route, it must have the LIR flag set in its PTA and the leaf tracking routes are all the matching Leaf A-D and SMET routes received in the downstream region.
- o If the route matched for transmission is an IMET route, the leaf tracking routes are all the IMET routes for the same BD received in the downstream region.

If the downstream region uses BIER, the packet is forwarded as following: the upstream segmentation's encapsulation is removed and the above mentioned label is swapped to the upstream-assigned label in the PTA of the route matched for transmission, and then a BIER header is imposed as in Section 4.1.1.

4.2. Disposition

The same procedures in section 4.2 of [RFC8556] are followed for EVPN-MPLS, except some EVPN specifics discussed in the following two sub-sections in this document.

For EVPN-VXLAN/NVGRE/GENEVE, the only difference is that the payload is VXLAN/NVGRE/GENEVE (with or without an IP header) and the VNI/VSID field in the VXLAN/NVGRE/GENEVE header is used to determine the corresponding mac VRF or broadcast domain.

4.2.1. At a BFER that is an Egress PE

Once the corresponding mac VRF or broadcast domain is determined from the upstream assigned label or VNI/VSID, EVPN forwarding procedures per [RFC7432] or [RFC8365] are followed. In case of EVPN-MPLS, if there is an inner label in the label stack following the BIER header, that inner label is considered as the upstream assigned ESI label for split horizon purpose.

4.2.2. At a BFER that is a P-tunnel Segmentation Point

This is only applicable to EVPN-MPLS. The same procedures in Section 4.2.2 of [RFC8556] are followed, subject to multihoming procedures specified in [I-D.ietf-bess-evpn-bum-procedure-updates].

5. IANA Considerations

This document requests two assignments in "BIER Next Protocol Identifiers" registry, with the following two recommended values:

- o 7: Payload is VXLAN encapsulated (no IP/UDP header)
- o 8: Payload is NVGRE encapsulated (no IP header)
- o 9: Payload is GENEVE encapsulated (no IP/UDP header)

6. Security Considerations

To be updated.

7. Acknowledgements

The authors thank Eric Rosen for his review and suggestions. Additionally, much of the text is borrowed verbatim from [RFC8556].

8. References

8.1. Normative References

- [I-D.ietf-bess-evpn-bum-procedure-updates]
Zhang, Z., Lin, W., Rabadan, J., Patel, K., and A. Sajassi, "Updates on EVPN BUM Procedures", draft-ietf-bess-evpn-bum-procedure-updates-07 (work in progress), August 2019.
- [I-D.ietf-bess-evpn-igmp-mld-proxy]
Sajassi, A., Thoria, S., Patel, K., Drake, J., and W. Lin, "IGMP and MLD Proxy for EVPN", draft-ietf-bess-evpn-igmp-mld-proxy-04 (work in progress), September 2019.
- [I-D.ietf-bess-evpn-optimized-ir]
Rabadan, J., Sathappan, S., Lin, W., Katiyar, M., and A. Sajassi, "Optimized Ingress Replication solution for EVPN", draft-ietf-bess-evpn-optimized-ir-06 (work in progress), October 2018.
- [I-D.ietf-bess-mvpn-expl-track]
Dolganow, A., Kotalwar, J., Rosen, E., and Z. Zhang, "Explicit Tracking with Wild Card Routes in Multicast VPN", draft-ietf-bess-mvpn-expl-track-13 (work in progress), November 2018.
- [I-D.ietf-idr-tunnel-encaps]
Patel, K., Velde, G., and S. Ramachandra, "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-14 (work in progress), September 2019.
- [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>>.
- [RFC6625] Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R. Qiu, "Wildcard in Multicast VPN Auto-Discovery Routes", RFC 6625, DOI 10.17487/RFC6625, May 2012, <<https://www.rfc-editor.org/info/rfc6625>>.

- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <<https://www.rfc-editor.org/info/rfc7432>>.
- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.
- [RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/info/rfc8296>>.
- [RFC8317] Sajassi, A., Ed., Salam, S., Drake, J., Uttaro, J., Boutros, S., and J. Rabadan, "Ethernet-Tree (E-Tree) Support in Ethernet VPN (EVPN) and Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 8317, DOI 10.17487/RFC8317, January 2018, <<https://www.rfc-editor.org/info/rfc8317>>.
- [RFC8556] Rosen, E., Ed., Sivakumar, M., Przygienda, T., Aldrin, S., and A. Dolganow, "Multicast VPN Using Bit Index Explicit Replication (BIER)", RFC 8556, DOI 10.17487/RFC8556, April 2019, <<https://www.rfc-editor.org/info/rfc8556>>.

8.2. Informative References

- [I-D.boutros-bess-evpn-geneve]
Boutros, S., Sajassi, A., Drake, J., Rabadan, J., and S. Aldrin, "EVPN control plane for Geneve", draft-boutros-bess-evpn-geneve-04 (work in progress), March 2019.
- [I-D.ietf-bier-php]
Zhang, Z., "BIER Penultimate Hop Popping", draft-ietf-bier-php-03 (work in progress), October 2019.
- [I-D.keyupate-bess-evpn-virtual-hub]
Patel, K., Sajassi, A., Drake, J., Zhang, Z., and W. Henderickx, "Virtual Hub-and-Spoke in BGP EVPNs", draft-keyupate-bess-evpn-virtual-hub-02 (work in progress), September 2019.

- [I-D.zzhang-bess-mvpn-evpn-cmcast-enhancements]
Zhang, Z., Kebler, R., Lin, W., and E. Rosen, "MVPN/EVPN
C-Multicast Routes Enhancements", draft-zzhang-bess-mvpn-
evpn-cmcast-enhancements-01 (work in progress), March
2019.
- [RFC8365] Sajassi, A., Ed., Drake, J., Ed., Bitar, N., Shekhar, R.,
Uttaro, J., and W. Henderickx, "A Network Virtualization
Overlay Solution Using Ethernet VPN (EVPN)", RFC 8365,
DOI 10.17487/RFC8365, March 2018,
<<https://www.rfc-editor.org/info/rfc8365>>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

Antoni Przygienda
Juniper Networks

EMail: prz@juniper.net

Ali Sajassi
Cisco Systems

EMail: sajassi@cisco.com

Jorge Rabadan
Nokia

EMail: jorge.rabadan@nokia.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 1, 2021

M. McBride
Futurewei
J. Xie
X. Geng
S. Dhanaraj
Huawei
R. Asati
Cisco
Y. Zhu
China Telecom
G. Mishra
Verizon Inc.
Z. Zhang
Juniper
September 28, 2020

BIER IPv6 Requirements
draft-ietf-bier-ipv6-requirements-09

Abstract

There have been several proposed solutions with BIER being used in IPv6. But there hasn't been a document which describes the problem and lists the requirements. The goal of this document is to describe the general BIER IPv6 encapsulation problem and detail solution requirements, thereby assisting the working group in the development of acceptable solutions.

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 April 1, 2021.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Requirements Language	3
1.2. Terminology	3
2. Problem Statement	3
3. Requirements	4
3.1. Mandatory Requirements	4
3.1.1. Support various L2 link types	4
3.1.2. Support BIER architecture	4
3.1.3. Support deployment with Non-BFR routers	4
3.1.4. Support OAM	5
3.2. Optional Requirements	5
3.2.1. Support Fragmentation	5
3.2.2. Support IPSEC ESP	5
4. IANA Considerations	5
5. Security Considerations	6
6. Acknowledgement	6
7. Normative References	6
Authors' Addresses	7

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides optimal multicast forwarding, without requiring intermediate routers to maintain per-flow state, through the use of a multicast-specific BIER header. [RFC8296] defines two types of BIER encapsulation: one is BIER MPLS encapsulation for MPLS environments, the other is non-MPLS BIER encapsulation to run without MPLS. This document describes non-MPLS BIER encapsulation in IPv6 environments. We explain the requirements of transporting multicast flow overlay payload through an IPv6 network underlay using BIER. The solutions

may use IPv6 forwarding plane and may include IPv6 encapsulation and/or generic IPv6 tunnelling.

1.1. 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 RFC 2119 [RFC2119].

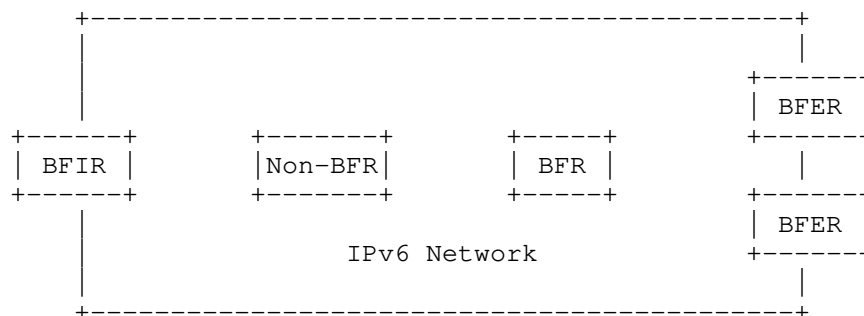
1.2. Terminology

- o BIER: Bit Index Explicit Replication. Provides optimal multicast forwarding through adding a BIER header and removing state in intermediate routers.

2. Problem Statement

The problem is how to transport multicast packets, with non-MPLS BIER encapsulation, in an IPv6 environment. We need to determine where to put the BIER header in this IPv6 environment. With IPv6 encapsulation being increasingly used for unicast services, such as VPN or L2VPN, it may be desirable to have IPv6 encapsulation also used in BIER deployments for multicast services such as MVPN. It may also be desirable to not use IPv6 encapsulation except when IPv6 tunneling (native or GRE/UDP-like) is used to transport BIER packets over BIER-incapable routers.

Below is a simple scenario that needs BIER IPv6-based forwarding:



This scenario depicts the need to replicate BIER packets from a BFIR to BFERs across an IPv6 Service Provider core. Inside the IPv6 network, the BIER header is used to direct the packet from one BFR to the next BFRs, and either a IPv6 header or an L2/tunnel header is used to provide reachability between BFRs. The IPv6 environment may include a variety of link types, may be entirely IPv6, or may be dual stack. There may be cases where not all routers are BFR capable in

the IPv6 environment but still want to deploy BIER. Regardless of the environment, the problem is to deploy BIER, with non-MPLS BIER encapsulation, in an IPv6 network.

3. Requirements

There are several suggested requirements for BIER IPv6 solutions.

In this document, the requirements are divided into two levels: Mandatory and Optional. The requirement levels are determined based on the following factors:

If the requirement is required for a feature that is likely to be a potential deployment, the requirement level will be considered mandatory.

If the impact of not implementing the requirement may block BIER from been deployed, the requirement level will be considered mandatory.

3.1. Mandatory Requirements

Considering that these mandatory requirements are all well-known to the working group, and practical in normal deployment, they will be listed without a detailed description.

3.1.1. Support various L2 link types

The solution should support various kinds of L2 data link types.

3.1.2. Support BIER architecture

The solution must support the BIER architecture.

Supporting different multicast flow overlays, multiple sub-domains, multi-topologies, multiple sets, multiple Bit String Lengths, and deterministic ECMP are considered essential functions of BIER and need to be supported.

3.1.3. Support deployment with Non-BFR routers

The solution must support deployments with BIER-incapable routers. This is beneficial to the deployment of BIER, especially in early deployments when some routers do not support BIER forwarding but support IPv6 forwarding.

3.1.4. Support OAM

BIER OAM tools like [I-D.ietf-bier-ping] and [I-D.ietf-bier-pmmm-oam] should be supported, either directly using existing methods, or by specifying a new method for the same functionality. They are likely to be needed in normal BIER deployment for diagnostics.

3.2. Optional Requirements

The requirements in this section are listed as optional, and each requirement is explained with a detailed scenario. Note that fragmentation and IPSEC ESP are not BIER functions, they are provided by the upper IP layer.

3.2.1. Support Fragmentation

There are some cases where the Fragmentation/Assembly function is needed for BIER to work in an IPv6 network.

For example, a customer IPv6 multicast packet may be 1280 bytes and is required to be transported through an IPv6 network using BIER. Every link of the IPv6 network is no less than the requisite 1280 bytes [RFC8200], but the size of the payload that can be encapsulated in BIER (BIER-MTU) is less than 1280 bytes. In this case, it is not the appropriate action for a BFIR to drop the packet and advertise an MTU to the source [RFC8296]. Instead, some transport mechanism needs to provide the fragmentation and assembly function.

3.2.2. Support IPSEC ESP

There are some cases where the IPSEC ESP function may be needed to transport c-multicast packets through an IPv6 network with confidentiality using BIER technology.

A service provider may want to provide additional security SLA to its customer to ensure that the unencrypted c-multicast packet is not altered in the service provider's network. In this case, if the BIER technology is preferred for the multicast service, BIER with IPSEC ESP support may be a candidate solution. On the other hand, the traffic protection may be better provided via IPSEC or MACSEC at multicast flow overlay over and beyond the BIER domain.

4. IANA Considerations

Some BIER IPv6 encapsulation proposals do not require any action from IANA while other proposals require new IPv6 Option codepoints from IPv6 sub-registries, new "Next header" values, or require new IP

Protocol codes. This document, however, does not require anything from IANA.

5. Security Considerations

There are no security issues introduced by this draft.

6. Acknowledgement

Thanks to Eric Rosen for his listed set of initial requirements on the BIER WG mailing list.

7. Normative References

[I-D.ietf-bier-ping]

Nainar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", draft-ietf-bier-ping-07 (work in progress), May 2020.

[I-D.ietf-bier-pmmm-oam]

Mirsky, G., Zheng, L., Chen, M., and G. Fioccola, "Performance Measurement (PM) with Marking Method in Bit Index Explicit Replication (BIER) Layer", draft-ietf-bier-pmmm-oam-08 (work in progress), May 2020.

[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>>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/info/rfc8296>>.

Authors' Addresses

Mike McBride
Futurewei

Email: michael.mcbride@futurewei.com

Jingrong Xie
Huawei

Email: xiejingrong@huawei.com

Xuesong Geng
Huawei

Email: gengxuesong@huawei.com

Senthil Dhanaraj
Huawei

Email: senthil.dhanaraj@huawei.com

Rajiv Asati
Cisco

Email: rajiva@cisco.com

Yongqing Zhu
China Telecom

Email: zhuyq8@chinatelecom.cn

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Zhaohui Zhang
Juniper

Email: zzhang@juniper.net

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: May 7, 2020

N. Kumar
IJ. Wijnands
M. Mishra
Cisco
November 4, 2019

Flex Algorithm for BIER
draft-nainar-bier-flex-algo-00

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "BIER domain" without requiring intermediate routers to run explicit tree-building protocol or to maintain multicast-related, per-flow state. IGP protocols are extended to carry BFR-Id and other encapsulation informations that are used by traditional path computing algorithm using link metric for a loop-free best path selection.

This document defines a constrained based path selection using IGP flexible Algorithm for BIER.

Requirements Language

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 RFC 2119 [RFC2119] RFC 8174 [RFC8174] when and only when, they appear in all capitals, as shown here.

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 May 7, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Acronyms and Terminology	3
1.1.1. Acronyms	3
1.1.2. Terminology	3
2. Flexible Algorithm	3
3. Constraint Forwarding Identifier	4
3.1. BFR ID Mapping for Flexible Algorithm	4
3.2. BIER-MPLS Label Mapping for Flexible Algorithm	5
4. IGP Extensions Flexible Algorithm	6
4.1. ISIS	6
4.2. OSPF	6
5. Security Considerations	6
6. IANA Considerations	6
7. Acknowledgements	6
8. References	6
8.1. Normative References	6
8.2. Informative References	7
Authors' Addresses	7

1. Introduction

[RFC8279] defines Bit Index Explicit Replication (BIER), an architecture that provides optimal multicast forwarding through a "BIER domain" without requiring intermediate routers to run explicit tree-building protocol or to maintain multicast-related, per-flow state. [RFC8401] and [RFC8444] defines the IGP protocols extensions to carry BFR-Id and other encapsulation informations that are used by traditional path computing algorithm using link metric for a loop-free best path selection.

The ability to compute constrained path using attributes beyond the basic link metric and steering the multicast traffic over such constrained path brings a lot of benefits such as efficient load distribution, path dis-jointness and resiliency. Bandwidth-aware, delay-sensitive or multi-planar are some of the examples for such constrained path selection. The path computation and traffic steering over flexible algorithm based constrained path requires advertising a set of Path constraints associated to each link and a unique dataplane based identifier to differentiate the data packets that needs to be steered over such computed constrained paths.

This document specifies the IGP protocol extensions and the mechanism to implement IGP Flexible Algorithm for BIER network.

1.1. Acronyms and Terminology

1.1.1. Acronyms

TBD

1.1.2. Terminology

This document uses the terminologies defined in [RFC8279], [RFC8296], and so the readers are expected to be familiar with the same.

2. Flexible Algorithm

Different types of constraints may be used to compute a path over the BIER network. Link performance and multi-plane are some of the common examples for such constraints. An example multiplane BIER network is shown in below figure 1.

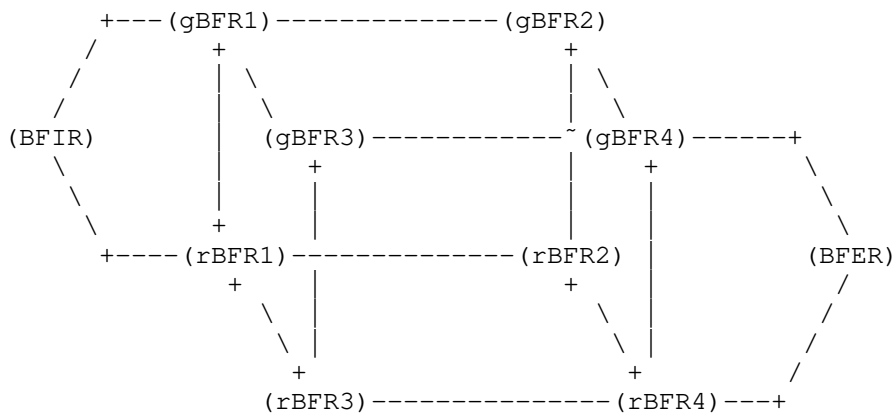


Figure 1. MultiPlane BIER Network

The above BIER network is enabled with "green" and "red" planes by assigning a contiguous set of BFRs to each plane. For example, gBFR1, gBFR2, gBFR3 and gBFR4 belongs to "green" plane while rBFR1, rBFR2, rBFR3 and rBFR4 belong to "red" plane. BFIR and BFER are enabled with both the planes.

Any BFR must have a mechanism to identify the set of constraints associated to each algorithm so that a loop free path can be computed. Any BFR must have a mechanism to map the data packet to the associated constrained path for loop free constrained forwarding.

3. Constraint Forwarding Identifier

This section explains different mechanism for identifying the constraints forwarding in the BIER encapsulated data packet.

3.1. BFR ID Mapping for Flexible Algorithm

For each Flexible Algorithm, a domain wide unique BFR-ID will be assigned with BFR-Prefix for each participating BFER within the BIER domain.

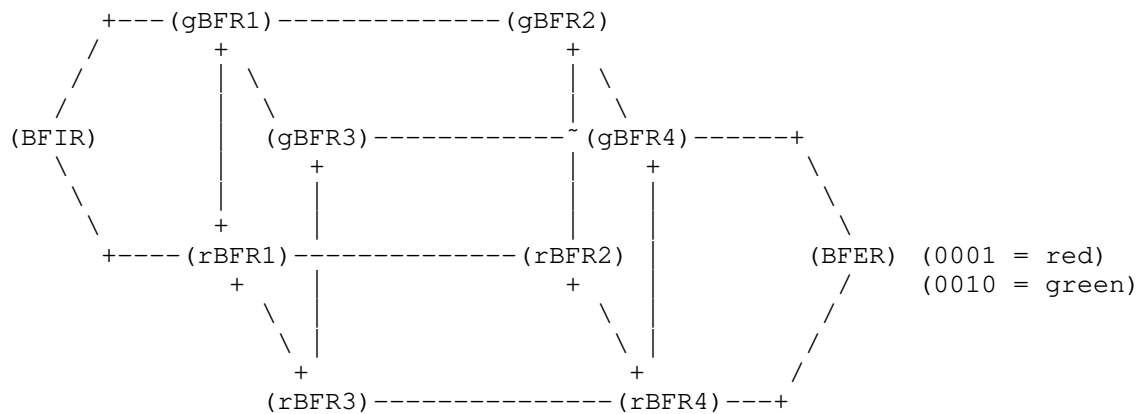


Figure 2. BFR-ID Mapping

Each BFER is assigned with domain wide unique BFR-ID for each Flexible Algorithm. In Figure 2, BFER assigns 0001 for "red" plane while using 0010 for "green" plane. Any BFR participating in one plane may not have the BFR-ID associated with other planes.

BFIR pushes the relevant BFR-ID to enforce the forwarding over any specific constraint path which can be influenced by a local policy.

3.2. BIER-MPLS Label Mapping for Flexible Algorithm

For each Flexible Algorithm, a locally unique BIER-MPLS label is assigned by each participating BFR within the BIER domain. In this option, each BFER is assigned with just one BFR-ID as mentioned in [RFC8279].

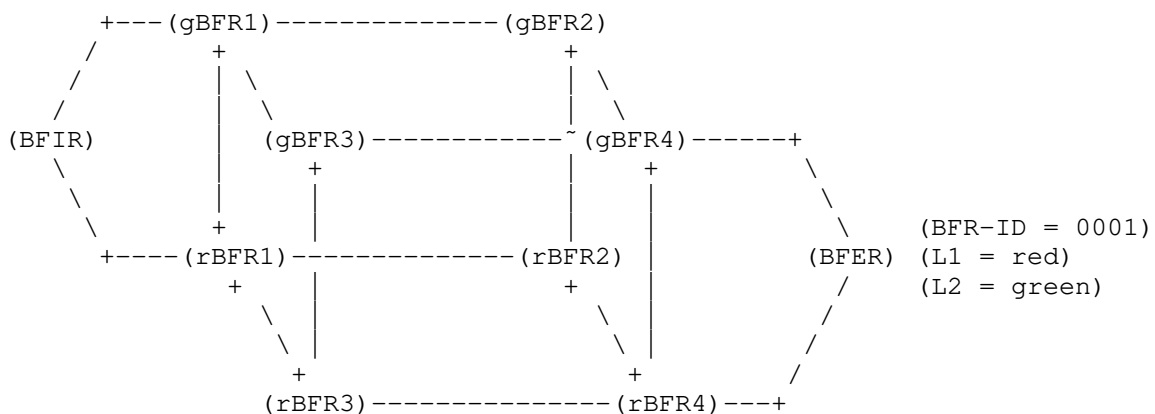


Figure 3. BIER-MPLS Label Mapping

Each BFR is assigned with locally unique BIER-MPLS for each Flexible Algorithm. The BIER-MPLS label along with the relevant constraints are advertised to other BFR using extensions defined in section x. In Figure 3, BFER is assigned with BFR-ID of 1 and advertise BIER-MPLS label L1 for "red" plane and L2 for "green" plane.

BFIR pushes the relevant BIER-MPLS advertised by the nexthop. Any BFR participating in both the plane will have the forwarding instruction for both the planes populated in different BIFT. The incoming BIER-MPLS label is used to identify the plane and the BIFT to perform the lookup and forwarding.

Additional details about non-MPLS BIER encapsulation will be included in later revisions.

4. IGP Extensions Flexible Algorithm

This section defines the IGP protocol extensions for BIER Flexible Algorithm.

4.1. ISIS

4.2. OSPF

5. Security Considerations

To be Updated.

6. IANA Considerations

TBD.

7. Acknowledgements

To be Updated.

8. References

8.1. Normative References

- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.
- [RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/info/rfc8296>>.
- [RFC8401] Ginsberg, L., Ed., Przygienda, T., Aldrin, S., and Z. Zhang, "Bit Index Explicit Replication (BIER) Support via IS-IS", RFC 8401, DOI 10.17487/RFC8401, June 2018, <<https://www.rfc-editor.org/info/rfc8401>>.
- [RFC8444] Psenak, P., Ed., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, J., and S. Aldrin, "OSPFv2 Extensions for Bit Index Explicit Replication (BIER)", RFC 8444, DOI 10.17487/RFC8444, November 2018, <<https://www.rfc-editor.org/info/rfc8444>>.
- [RFC8459] Dolson, D., Homma, S., Lopez, D., and M. Boucadair, "Hierarchical Service Function Chaining (hSFC)", RFC 8459, DOI 10.17487/RFC8459, September 2018, <<https://www.rfc-editor.org/info/rfc8459>>.

8.2. Informative References

- [I-D.ietf-lsr-flex-algo] Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-algo-04 (work in progress), September 2019.

Authors' Addresses

Nagendra Kumar
Cisco Systems, Inc.

Email: naikumar@cisco.com

Ijsbrand Wijnands
Cisco Systems, Inc.

Email: iwijnand@cisco.com

Mankamana Mishra
Cisco Systems, Inc.

Email: mankamis@cisco.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 25, 2019

S. Venaas
IJ. Wijnands
M. Mishra
Cisco Systems, Inc.
M. Sivakumar
Juniper Networks
October 22, 2018

PIM Flooding Mechanism and Source Discovery for BIER
draft-venaas-bier-pfm-sd-00

Abstract

PIM Flooding Mechanism and Source Discovery (PFM-SD) is a mechanism for source discovery within a PIM domain. PIM signaling over BIER has been defined, allowing for BIER to interoperate with PIM. This document defines PFM-SD over BIER, such that PFM-SD can be used by PIM in a PIM domain to discover sources that are reachable via BIER. Also, this document provides PFM-SD extensions to discover the BIER ingress router closest to the source. This can be used by BIER overlays, such as PIM signaling over BIER, to determine which router to signal.

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 April 25, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

(<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Conventions Used in This Document	3
2. PFM over BIER	3
3. PFM Ingress BIER Router TLV	3
4. Group Source Holdtime Metric TLV	4
5. BIER signaling enhancements	6
6. Security Considerations	6
7. IANA Considerations	6
8. References	6
8.1. Normative References	6
8.2. Informative References	7
Authors' Addresses	7

1. Introduction

PIM Flooding Mechanism (PFM) and Source Discovery (SD) [RFC8364] provides a generic flooding mechanism for distributing information throughout a PIM domain. In particular it allows for source discovery. There are various deployment scenarios where PIM and BIER need to co-exist. For instance, consider migration scenarios where a few routers in a PIM domain are upgraded to support BIER. In that case, one may use PIM Signaling Through BIER Core [I-D.ietf-bier-pim-signaling], allowing PIM to build trees passing through the BIER routers. This document defines PFM over BIER. This allows PFM to pass through the BIER routers, allowing PFM to be used in the PIM domain.

One challenge with PIM signaling over BIER [I-D.ietf-bier-pim-signaling] is to determine which BIER router is closest to the source. A number of options are discussed in that document. This document provides an alternative solution for discovering which BIER router to signal. It may also be used with other signaling mechanisms such as IGMP/MLD [I-D.ietf-bier-mld]. This is achieved by introducing two new PFM TLVs. When a BIER router forwards a PFM message into BIER, it adds a new TLV specifying the BIER sub-domain, its BFR-ID and its BIER prefix. Also, any Group Source Holdtime TLVs, defined in [RFC8364], are replaced with new TLVs that include the router's cost of reaching the sources.

Length: The length of the value in octets.

Sub-domain-id: The ID of the sub-domain that this PFM is forwarded into. The length is 1 octet.

Reserved: MUST be set to 0, and ignored when received. The length is 1 octet.

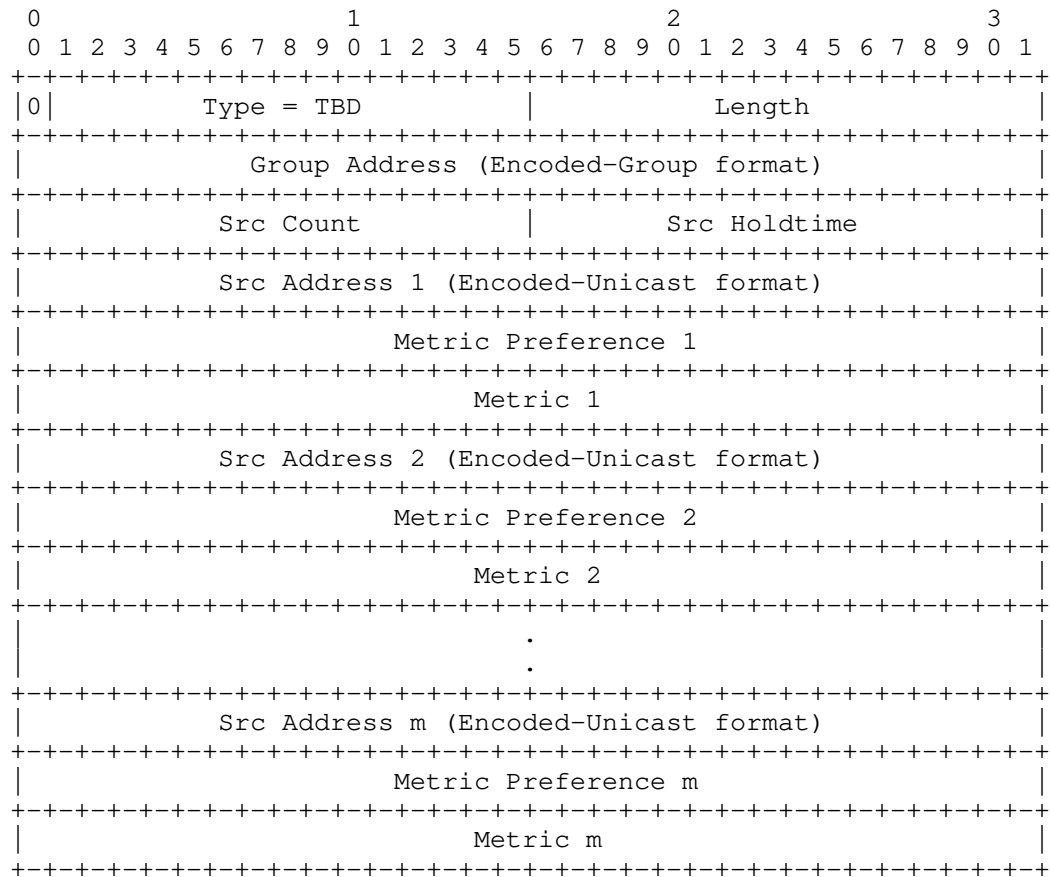
BFR-id: The BFR-id of the router that added this TLV in the sub-domain specified. The length is 2 octets.

BFR-prefix: The BFR-prefix of the router that added this TLV in the sub-domain specified. This length is 4 octets for IPv4 and 16 octets for IPv6.

4. Group Source Holdtime Metric TLV

When a router forwards a PFM message into a BIER domain, it should replace all Group Source Holdtime TLVs defined in [RFC8364] with the Group Source Holdtime Metric TLVs defined here. They are the same, except here we also add metric preference and metric. The metric preference and metric MUST be set to this router's metric and preference to reach the specified source. If the source is not reachable, the TLV MUST be omitted. This TLV is used together with the PFM Ingress BIER Router TLV is used to indicate the ingress router's cost of reaching the source.

When a router receives a message containing this TLV, it SHOULD store this information, but it MUST NOT forward these TLVs. If forwarding into another BIER domain, the metric preference and metric MUST be updated with this router's cost of reaching the source. If forwarding into a PIM domain, all the TLVs SHOULD be replaced with Group Source Holdtime TLVs as defined in [RFC8364]. The same information is used, except that the metric preference and metric are left out. One could potentially make use of the metric in a PIM domain as well, but it is not clear whether this is useful, and the PIM routers may not support this TLV.



0: The Transitive bit is set to 0.

Type: Type is TBD.

Length: The length of the value in octets.

Group Address: The group that sources are to be announced for. The format for this address is given in the Encoded-Group format in [RFC7761].

Src Count: The number of source addresses that are included.

Src Holdtime: The Holdtime (in seconds) for the included source(s).

Src Address: The source address for the corresponding group. The format for these addresses is given in the Encoded-Unicast address in [RFC7761].

Metric Preference: Preference value assigned to the unicast routing protocol that provided the route to the source.

Metric: The unicast routing table metric associated with the route used to reach the source. The metric is in units applicable to the unicast routing protocol used.

5. BIER signaling enhancements

A BIER border router SHOULD cache all the Group Source Holdtime Metric TLVs it receives, along with the respective PFM Ingress BIER Router TLV. This allows the router to determine which sources are active, and which BIER border router is closest to the source. The sub-domain ID, BFR-id and BFR-prefix in the TLV provide the necessary information for use by signaling mechanisms such as [I-D.ietf-bier-pim-signaling] to signal the preferred ingress router. It may also be used by [I-D.ietf-bier-mld]. IGMP/MLD reports would generally be sent to all BIER routers as it is not known which sources are active and which routers can reach them. But by using the enhancements in this document, a source-specific report can be sent to the router closest to the source. Also a group report might be set to the set of routers that are closest to the sources for that group. This reduces the amount of receiver state on the BIER routers, and also the amount of messages each routers needs to process.

6. Security Considerations

TBD

7. IANA Considerations

This document defines two new PFM TLVs that needs to be assigned from the "PIM Flooding Mechanism Message Types" registry.

8. References

8.1. Normative References

- [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>>.

- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <<https://www.rfc-editor.org/info/rfc7761>>.
- [RFC8364] Wijnands, IJ., Venaas, S., Brig, M., and A. Jonasson, "PIM Flooding Mechanism (PFM) and Source Discovery (SD)", RFC 8364, DOI 10.17487/RFC8364, March 2018, <<https://www.rfc-editor.org/info/rfc8364>>.

8.2. Informative References

- [I-D.ietf-bier-mld]
Pfister, P., Wijnands, I., Venaas, S., Wang, C., Zhang, Z., and M. Stenberg, "BIER Ingress Multicast Flow Overlay using Multicast Listener Discovery Protocols", draft-ietf-bier-mld-01 (work in progress), June 2018.
- [I-D.ietf-bier-pim-signaling]
Bidgoli, H., Dolganow, A., Kotalwar, J., Xu, F., mishra, m., and Z. Zhang, "PIM Signaling Through BIER Core", draft-ietf-bier-pim-signaling-04 (work in progress), October 2018.

Authors' Addresses

Stig Venaas
Cisco Systems, Inc.
Tasman Drive
San Jose CA 95134
USA

Email: stig@cisco.com

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Mankamana Mishra
Cisco Systems, Inc.
Tasman Drive
San Jose CA 95134
USA

Email: mankamis@cisco.com

Mahesh Sivakumar
Juniper Networks
1133 Innovation Way
Sunnyvale CA 94089
USA

Email: sivakumar.mahesh@gmail.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 7, 2020

M. Sivakumar
Juniper Networks
S. Venaas
Cisco Systems, Inc.
Z. Zhang
ZTE Corporation
March 6, 2020

IGMPv3/MLDv2 Message Extension
draft-venaas-pim-igmp-mld-extension-01

Abstract

IGMP and MLD protocols are extensible, but no extensions have been defined so far. This document provides a well-defined way of extending IGMP and MLD, including a new extension type to distinguish between different extensions.

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 September 7, 2020.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Conventions used in this document	2
3. Multicast Listener Query Extension	2
4. Version 2 Multicast Listener Report Extension	4
5. IGMP Membership Query Extension	4
6. IGMP Version 3 Membership Report Extension	5
7. Security Considerations	6
8. IANA Considerations	6
9. References	7
9.1. Normative References	7
9.2. Informative References	7
Authors' Addresses	7

1. Introduction

In this document, we describe a generic method to extend IGMPv3 [RFC3376] and MLDv2 [RFC3810] messages to accommodate information other than what is contained in the current message formats. This is done by introducing an extension-type field in the message formats to indicate the application for which the extension is done. This will be followed by the actual value of the extension.

The extension will be part of additional data as mentioned in [RFC3810] Section 5.1.12 (resp. [RFC3376] Section 4.1.10) for query messages and [RFC3810] Section 5.2.12 (resp. [RFC3376] Section 4.2.11) for report messages.

One such extension is being defined in [I-D.ietf-bier-mld]

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.

3. Multicast Listener Query Extension

The MLD query format with extension is shown below

0

1

2

3

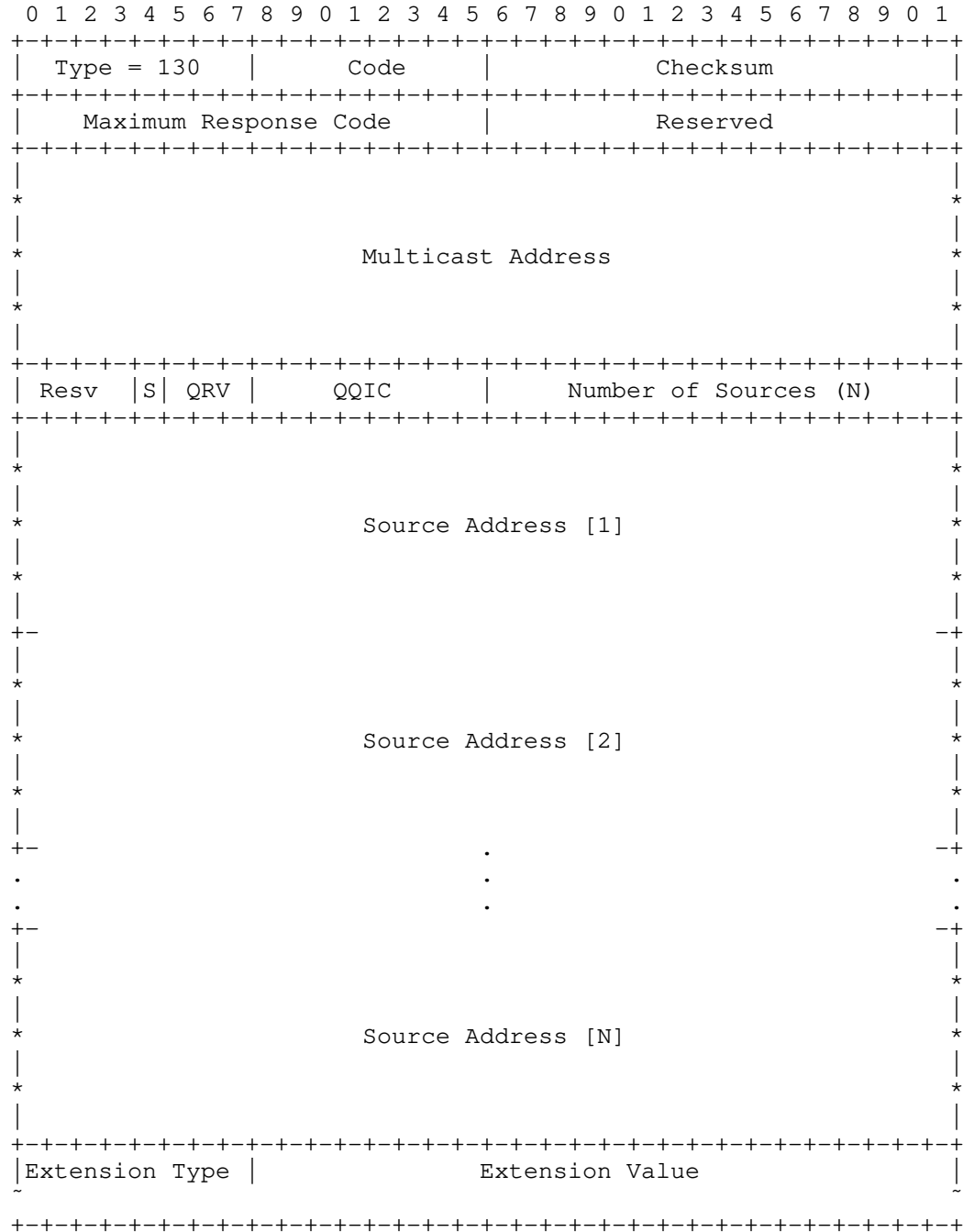


Figure 2: MLD Query Extension

4. Version 2 Multicast Listener Report Extension

The MLD report format with extension is shown below

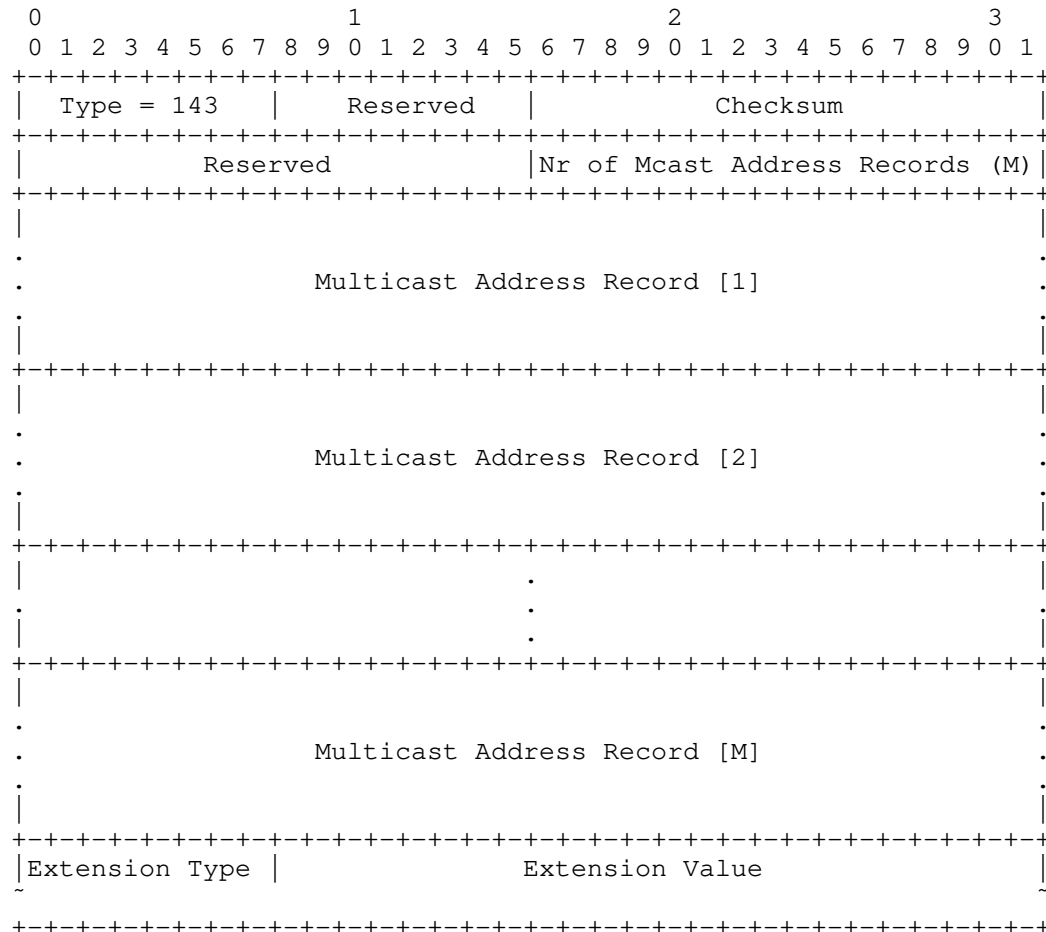


Figure 3: MLD Report Extension

5. IGMP Membership Query Extension

The IGMP query format with the extension is shown below

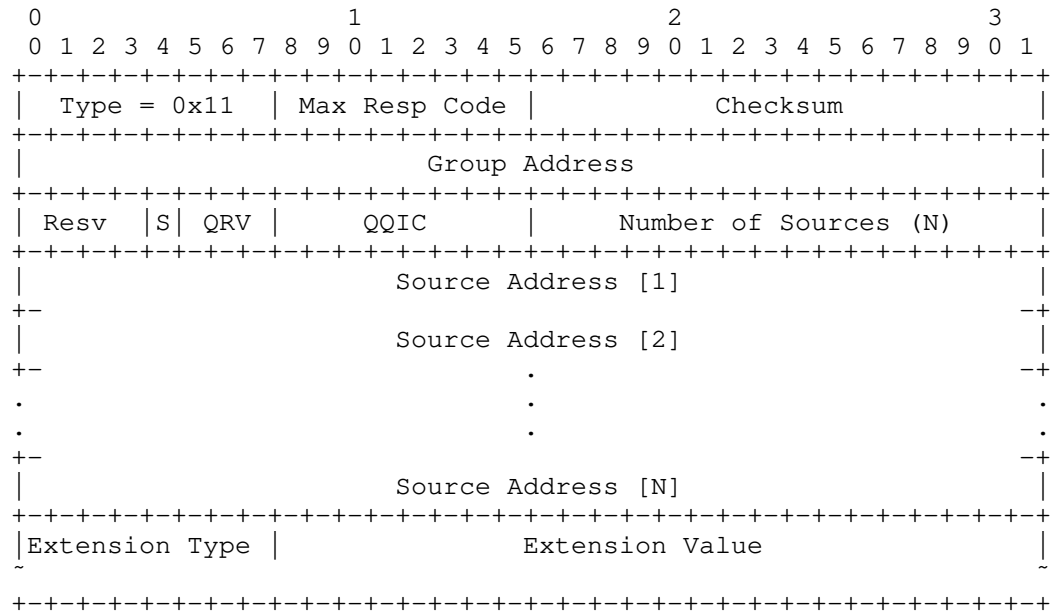


Figure 4: IGMP Query Extension

6. IGMP Version 3 Membership Report Extension

The IGMP report format with the extension is shown below

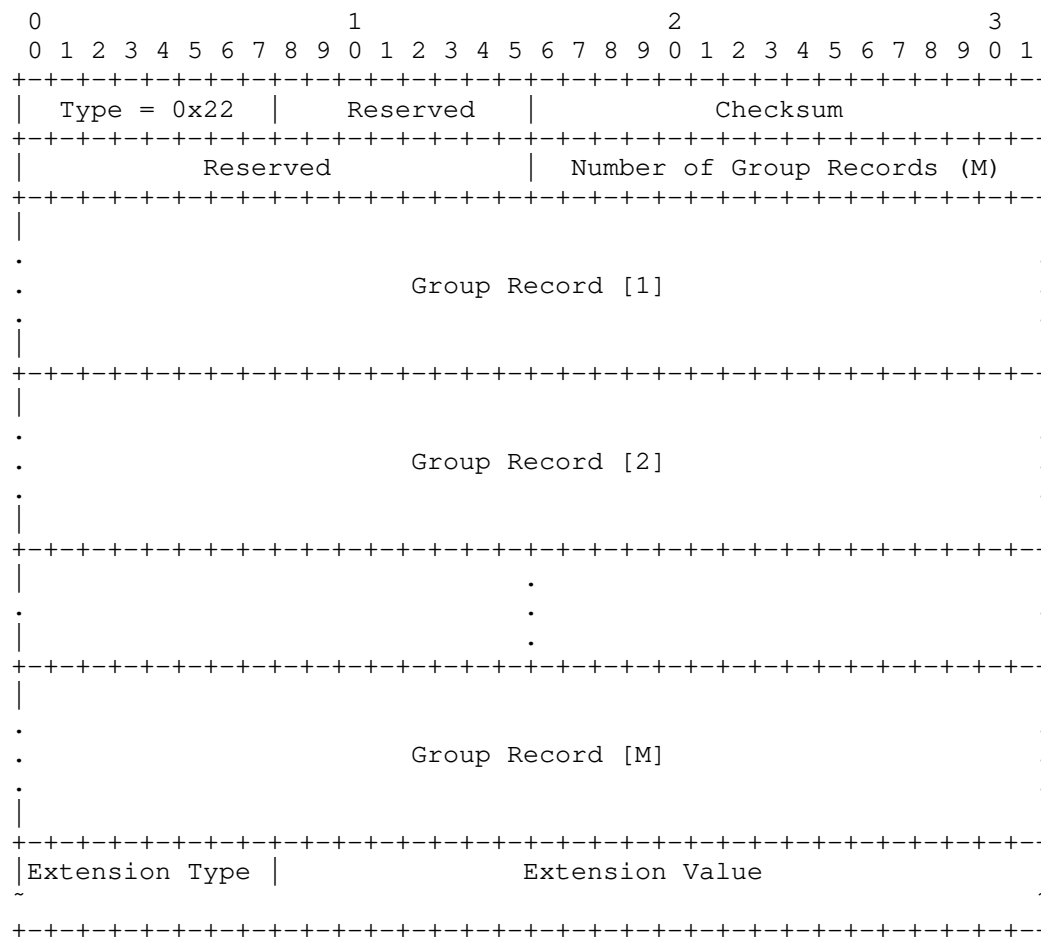


Figure 5: IGMP Report Extension

7. Security Considerations

This document extends MLD (resp. IGMP) message formats. As such, there is no impact on security or changes to the considerations in [RFC3810] and [RFC3376].

8. IANA Considerations

This document requests that IANA creates a new registry for IGMP/MLD extension-types.

9. References

9.1. Normative References

- [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>>.
- [RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, DOI 10.17487/RFC3376, October 2002, <<https://www.rfc-editor.org/info/rfc3376>>.
- [RFC3810] Vida, R., Ed. and L. Costa, Ed., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, DOI 10.17487/RFC3810, June 2004, <<https://www.rfc-editor.org/info/rfc3810>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

9.2. Informative References

- [I-D.ietf-bier-mld]
Pfister, P., Wijnands, I., Venaas, S., Wang, C., Zhang, Z., and M. Stenberg, "BIER Ingress Multicast Flow Overlay using Multicast Listener Discovery Protocols", draft-ietf-bier-mld-04 (work in progress), March 2020.

Authors' Addresses

Mahesh Sivakumar
Juniper Networks
64 Butler St
Milpitas CA 95035
USA

Email: sivakumar.mahesh@gmail.com

Stig Venaas
Cisco Systems, Inc.
Tasman Drive
San Jose CA 95134
USA

Email: stig@cisco.com

Zheng (Sandy) Zhang
ZTE Corporation
No. 50 Software Ave, Yuhuatai District
Nanjing 210000
China

Email: zhang.zheng@zte.com.cn

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 21, 2020

J. Xie
Huawei Technologies
L. Geng
China Mobile
M. McBride
Futurewei
R. Asati
Cisco
S. Dhanaraj
Huawei
July 20, 2019

Encapsulation for BIER in Non-MPLS IPv6 Networks
draft-xie-bier-ipv6-encapsulation-03

Abstract

This document proposes a BIER IPv6 (BIERv6) encapsulation for Non-MPLS IPv6 Networks using the IPv6 Destination Option extension header.

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] and [RFC8174].

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 21, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Terminology	3
3. BIER IPv6 Encapsulation	3
3.1. BIER Option in IPv6 Destination Options Header	3
3.2. Multicast and Unicast Destination Address	6
3.3. BIERv6 Packet Format	8
4. BIERv6 Packet Processing	9
5. Security Considerations	10
6. IANA Considerations	11
6.1. BIER Option Type	11
6.2. End.BIER Function	11
7. Acknowledgements	11
8. Contributors	11
9. References	12
9.1. Normative References	12
9.2. Informative References	12
Authors' Addresses	13

1. Introduction

Bit Index Explicit Replication (BIER) [RFC8279] is an architecture that provides optimal multicast forwarding without requiring intermediate routers to maintain any per-flow state by using a multicast-specific BIER header.

[RFC8296] defines a common BIER Header format for MPLS and Non-MPLS networks. It has defined two types of encapsulation methods using the common BIER Header, (1) BIER encapsulation in MPLS networks, here-in after referred as MPLS BIER Header in this document and (2) BIER encapsulation in Non-MPLS networks, here-in after referred as Non-MPLS BIER Header in this document. [RFC8296] also assigned

Ethertype=0xAB37 for Non-MPLS BIER Header packets to be directly carried over the Ethernet links.

This document proposes a BIER IPv6 encapsulation for Non-MPLS IPv6 Networks, defining a method to carry the standard Non-MPLS BIER header (as defined in [RFC8296]) in the native IPv6 header. A new IPv6 Option type - BIER Option is defined to encode the standard Non-MPLS BIER header and this newly defined BIER Option is carried under the Destination Options header of the native IPv6 Header [RFC8200].

This document details one of the proposed solutions for transporting BIER packets in an IPv6 network. To better understand the overall BIER IPv6 problem space, use cases and proposed solutions, refer to [I-D.ietf-bier-ipv6-requirements].

2. Terminology

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References.

The following new terms are used throughout this document:

- o BIERv6 - BIER IPv6.
- o BIER Option - An Option type carried in IPv6 Destination Options Header which includes the standard Non-MPLS BIER Header.
- o BIERv6 Header - An IPv6 Header with BIER Option.
- o BIERv6 Packet - An IPv6 packet with BIERv6 Header. Such an IPv6 packet typically carries the user multicast payload and is forwarded by BFRs in the BIERv6 network towards the multicast receivers.

3. BIER IPv6 Encapsulation

3.1. BIER Option in IPv6 Destination Options Header

Destination Options Header and the Options that can be carried under this extension header is defined in [RFC8200]. This document defines a new Option type - BIER Option, to encode the Non-MPLS BIER header. As specified in Section 4.2 [RFC8200], the BIER Option follows type-length-value (TLV) encoding format and the standard Non-MPLS BIER header [RFC8296] is encoded in the value portion of the BIER Option TLV.

TTL: MUST be set to 0 upon transmission, and MUST be ignored upon reception. The function of TTL is replaced by the Hop Limit field in IPv6 header.

Nibble: SHOULD be set to 0000 upon transmission, and MUST be ignored upon reception. See Section 2.2 of RFC 8296.

Ver: MUST be set to 0 upon transmission, and MUST be discarded when it is not 0 upon reception. See Section 2.2 of RFC 8296.

BSL: See Section 2.1.2 of RFC 8296.

Entropy: See Section 2.1.2 of RFC 8296.

OAM: See Section 2.1.2 of RFC 8296.

Rsv: See Section 2.1.2 of RFC 8296.

DSCP: SHOULD be set to binary value 000000 upon transmission and MUST be ignored upon reception. In IPv6 BIER encapsulation, uses highest 6-bit of Traffic Class field of IPv6 header to hold a Differentiated Services Codepoint [RFC2474].

Proto: SHOULD be set to 0 upon transmission and MUST be ignored upon reception. In IPv6 BIER encapsulation, the functionality of this 6-bit Proto field is replaced by the Next Header field in Destination Options header, which is the last IPv6 extension header, to indicate the BIER payload, which is also IPv6 payload.

For BIER Proto 1, indicating a Downstream-assigned MPLS payload, use Next Header value 137.

For BIER Proto 2, indicating an Upstream-assigned MPLS payload, there is no Next Header code currently. An upstream-assigned MPLS label within the context of special BFIR router, which in turn is represented by the BFIR-id and the Sub-domain indirectly indicated by the BIFT-id in a BIER-MPLS or BIER-ETH packet, can be replaced by an IPv6 source address in a BIER IPv6 encapsulation packet in a direct manner. In this case, use Next Header value 4 for IPv4 payload, or value 41 for IPv6 payload.

For BIER Proto 3, indicating an Ethernet payload, use Next Header value 97.

For BIER Proto 4, indicating an IPv4 payload, use Next Header value 4.

For BIER Proto 5, indicating a BIER-OAM payload, use Next Header value 58. How the BIER-PING is supported with BIER IPv6 encapsulation is outside the scope of this document.

For BIER Proto 6, indicating an IPv6 payload, use Next Header value 41.

BFIR-id: See Section 2.1.2 of RFC 8296.

BitString: See Section 2.1.2 of RFC 8296.

3.2. Multicast and Unicast Destination Address

BIER is generally a hop-by-hop and one-to-many architecture, and thus the IPv6 Destination Address (DA) being a Multicast Address is a way one may think of as an approach for both the two paradigms in BIERv6 encapsulation.

However using a unicast address has the following benefits:

1. Tunneling a BIERv6 packet over a non-BIER capable router.
2. Fast rerouting a BIERv6 packet using a unicast by-pass tunnel.
3. Forwarding a BIERv6 packet to one of the many BFR neighbors connected on a LAN.
4. Connecting BIER domains, for example Data Center domains, in an overlay manner.

Some of the above functions are assumed very basic requirements and part of BIER architecture as described in [RFC8279]. This document uses unicast address for both one-hop replication and multi-hop replication.

The unicast address used in BIERv6 packet targeting a BFR SHOULD be the IPv6 BFR-Prefix advertised from this BFR. When a BFR advertises the BIER information with BIERv6 encapsulation capability, the IPv6 BFR-prefix of this BFR MUST be selected specifically for BIERv6 packet forwarding. Locally this "BIER Specific" IPv6 address is initialized in FIB with a flag of "BIER specific handling", represented as End.BIER function. For convenience, the indication in FIB share the same space as SRv6 Endpoints Behaviors defined in [I-D.ietf-spring-srv6-network-programming]. Apart from this sharing

of code space, there is nothing dependent on SRv6. The co-existence of BIERv6 and SRv6 is outside the scope of this document.

BFR Prefix is used only in control plane in BIER MPLS encapsulation but not used in data plane. While in BIERv6, BFR prefix is used in both control plane and data plane. For the convenience of migration to BIERv6, it is RECOMMENDED to use an "exclusive" IPv6 address as BFR prefix when deploying BIER-MPLS in IPv6 networks. The "exclusive" IPv6 address should not be used for general purpose like BGP session establishment, but for "BIER specific" function. For Non-BIER IPv6 routers, the IPv6 address is a regular IPv6 prefix reachable through IPv6 unicast routing.

The following is an example of configuring a BIER specific IPv6 address and using this address as BFR prefix:

```
# Config a BIER specific IPv6 address with 128-bit mask on loopback0.
interface loopback0
  ipv6 address 2001:DB8::AB37 128 End.BIER

# Config the BIER-specific IPv6 address on loopback0 as BFR Prefix.
bier sub-domain 6 ipv6-underlay
bfr-prefix interface loopback0
```

The address used as "BIER specific" IPv6 address can be from inside the scope of an SRv6 Locator or outside the scope of the SRv6 Locator(s) since it is a host prefix (128-bit prefix-length prefix).

Each "BIER specific" address can be used in one or many sub-domains as BFR-prefix, such that it can be associated with one or many Multi-Topologies (MTs) or algorithms.

More than one "BIER specific" address are also allowed as different BFR-prefix of more than one sub-domain, as described in section 2 of [RFC8279].

The following is an example pseudo-code of the End.BIER function:

```
1. IF NH = 60 and HopLimit > 0 ;;;Ref1
2.   IF (OptType1 = BIER) and (OptLength1 = HdrExtLen*8 + 4) ;;;Ref2
3.     Lookup the BIER Header inside the BIER option TLV.
4.     Forward via the matched entry.
5.   ELSE ;;;Ref3
6.     Drop the packet and end the process.
7. ELSE IF NH=ICMPv6 or (NH=60 and Dest_NH=ICMPv6) ;;;Ref4
8.   Send to CPU.
9. ELSE ;;;Ref5
10.  Drop the packet.
```

Ref1: Destination options header follows the IPv6 header directly and HopLimit is bigger than zero.

Ref2: The first TLV is BIER type and is the only TLV present in Destination options header.

Ref3/Ref5: Undesired packet is dropped because the destination address is the BIER specific IPv6 address (End.BIER function).

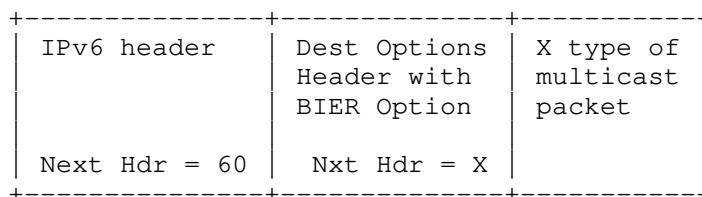
Ref4: An ICMPv6 packet using End.BIER as destination address.

3.3. BIERv6 Packet Format

As a multicast packet enters the BIER domain in a Non-MPLS IPv6 network, the multicast packet will be encapsulated with BIERv6 Header.

Typically a BIERv6 header would contain the Destination Options Header as the only Extensions Header besides IPv6 Header. However, it is allowed and possible for other extension headers to appear along with the Destination Options Header as long as the requirements listed in section 3.1 of this document is met.

Format of the multicast packet with BIERv6 encapsulation carrying only the Destination Options header is depicted in the below figure.



Format of the multicast packet with BIERv6 encapsulation carrying other extension headers along with Destination Options extension header is required to follow general recommendations of [RFC8200] and examples in other RFCs. [RFC6275] introduces how the order should be when other extension headers carries along with Home address option in a destination options header. Similar to this example, this document requires the Destination Options Header carrying the BIER option MUST be placed as follows:

- o After the routing header, if that header is present
- o Before the Fragment Header, if that header is present

- o Before the AH Header or ESP Header, if either one of those headers is present

Source Address field in the IPv6 header MUST be a routable IPv6 unicast address of the BFIR in any case.

BFIR encodes the Non-MPLS BIER header in the above mentioned encapsulation format and forwards the BIERv6 packet to the nexthop BFR following the local BIFT table.

BFRs in the IPv6 network, processes and replicates the packets towards the BFRs using the local BIFT table. The bit-string field in the Non-MPLS BIER header may be changed by the BFRs as they replicate the packet. BFRs MUST follow the procedures defined in section 3.1 as they modify the other fields in the Non-MPLS BIER header. The source address in the IPv6 header MUST NOT be modified by the BFRs.

4. BIERv6 Packet Processing

There is no BIER-specific processing, and all the 8 steps in section 6.5 of RFC8279 apply to BIERv6 packet processing. However, there are some IPv6-specific processing procedures due to the base and general procedures of IPv6.

On the overlay layer, when a multicast packet enters the BIER domain in a Non-MPLS IPv6 network, the Ingress BFR (BFIR) encapsulates the multicast packet with a BIERv6 Header, transforming it to a BIERv6 packet. The BIERv6 header includes an IPv6 header and IPv6 Destination Options Header within a standard Non-MPLS BIER header. Source Address field in the IPv6 header MUST be set to a routable IPv6 unicast address of the BFIR. Destination Address field in the IPv6 header is set to the BFR prefix of the next-hop BFR the BIERv6 packet replicating to, no matter next-hop BFR is directly connected (one-hop) or not directly connected (multi-hop).

On the BIER layer, upon receiving an BIERv6 packet, the BFR processes the IPv6 header first. This is the general procedure of IPv6.

If the IPv6 Destination address is an IPv6 BFR-Prefix unicast address of this BFR, a 'BIER Specific Handling' indication will be obtained by the preceding Unicast DA lookup (FIB lookup). The BIER option, if exists, will be checked to decide which neighbor(s) to replicate the BIERv6 packet to.

It is a local behavior to handle the combination of extension headers, options and the BIER option(s) in destination options header when a 'BIER Specific Handling' indication is got by the preceding

FIB lookup. Early deployment of BIERv6 may require there is only one BIER option TLV in the destination options header followed the IPv6 header. How other extension headers or more BIER option TLVs in a BIERv6 packet is handled is outside the scope of this document.

A packet having a 'BIER Specific Handling' indication but not having a BIER option is supposed to be a wrong packet or an ICMPv6 packet, and the process can be referred to the example in section 3.2.

A packet not having a 'BIER Specific Handling' indication but having a BIER option SHOULD be processed normally as unicast forwarding procedures, which may be a behavior of drop, or send to CPU, or other behaviors in existing implementations.

The Destination Address field in the IPv6 Header MUST change to the nexthop BFR's BFR Prefix if Unicast address is used in BIERv6.

The Hop Limit field of IPv6 header MUST decrease by 1 when sending packets to a BFR neighbor, while the TTL in the BIER header MUST be unchanged.

The BitString in the BIER header in the Destination Options Header may change when sending packets to a neighbor. Such change of BitString MUST be aligned with the procedure defined in RFC8279. Because of the requirement to change the content of the option when forwarding BIERv6 packet, the BIER option type should have chg flag 1 per section 4.2 of RFC8200.

The procedures applies normally if a bit corresponding to the self bfr-id is set in the bit-string field of the Non-MPLS BIER header of the BIERv6 packet. The node is considered to be an Egress BFR (BFER) in this case. The BFER removes the BIERv6 header, including the IPv6 header and the Destination Options header, and copies the packet to the multicast flow overlay. The egress VRF of a packet may be determined by a further lookup on the IPv6 source address instead of the upstream-assigned MPLS Label as described in [RFC8556].

The Fragment Header, AH Header or ESP Header, if exists after the BIER options header, can be processed on BFER only as part of the multicast flow overlay process.

5. Security Considerations

A BIERv6 packet with a special IPv6 Destination Address, the BFR prefix functioning as End.BIER, would be processed by BIER forwarding procedure only when the 'BIER Specific Handling' flag has been obtained ahead in the FIB lookup of the IPv6 header. Otherwise the packet with an IPv6 BIER Option will not be processed by the

procedure but be processed as normal IPv6 packet. A possible way of handling IPv6 packets with extension may be send to CPU for slow path processing.

6. IANA Considerations

6.1. BIER Option Type

Allocation is expected from IANA for a BIER Option Type codepoint from the "Destination Options and Hop-by-Hop Options" sub-registry of the "Internet Protocol Version 6 (IPv6) Parameters" registry. The value 0x70 is suggested.

Hex Value	act	chg	rest	Description	Reference
0x70	01	1	10000	BIER Option	This draft

6.2. End.BIER Function

Allocation is expected from IANA for an End.BIER function codepoint from the "SRv6 Endpoint Behaviors" sub-registry. The value 60 is suggested.

Value	Hex	Endpoint function	Reference
TBD	TBD	End.BIER	This draft

7. Acknowledgements

The authors would like to thank Stig Venaas for his valuable comments. Thanks IJsbrand Wijnands, Greg Shepherd, Tony Przygienda, Toerless Eckert, Jeffrey Zhang for the helpful comments to improve this document.

8. Contributors

Gang Yan
Huawei Technologies
China
Email: yangang@huawei.com

Yang(Yolanda) Xia
Huawei Technologies
China
Email: yolanda.xia@huawei.com

9. References

9.1. Normative References

- [RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <<https://www.rfc-editor.org/info/rfc6275>>.
- [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.
- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.
- [RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/info/rfc8296>>.
- [RFC8556] Rosen, E., Ed., Sivakumar, M., Przygienda, T., Aldrin, S., and A. Dolganow, "Multicast VPN Using Bit Index Explicit Replication (BIER)", RFC 8556, DOI 10.17487/RFC8556, April 2019, <<https://www.rfc-editor.org/info/rfc8556>>.

9.2. Informative References

- [I-D.ietf-bier-ipv6-requirements] McBride, M., Xie, J., Dhanaraj, S., and R. Asati, "BIER IPv6 Requirements", draft-ietf-bier-ipv6-requirements-01 (work in progress), July 2019.

- [I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J.,
daniel.voyer@bell.ca, d., Matsushima, S., and Z. Li, "SRv6
Network Programming", draft-ietf-spring-srv6-network-
programming-01 (work in progress), July 2019.
- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Authors' Addresses

Jingrong Xie
Huawei Technologies

Email: xiejingrong@huawei.com

Liang Geng
China Mobile
Beijing 10053

Email: gengliang@chinamobile.com

Mike McBride
Futurewei

Email: mmcbride7@gmail.com

Rajiv Asati
Cisco

Email: rajiva@cisco.com

Senthil Dhanaraj
Huawei

Email: senthil.dhanaraj@huawei.com

BIER
Internet-Draft
Intended status: Standards Track
Expires: March 26, 2020

Zheng. Zhang
ZTE Corporation
Bo. Wu
Individual
Zhaohui. Zhang
Juniper Networks
IJsbrand. Wijnands
Cisco Systems, Inc.
September 23, 2019

BIER Prefix Redistribute
draft-zwzw-bier-prefix-redistribute-03

Abstract

This document defines a BIER proxy function to interconnect different underlay routing protocol areas in a network. And a new BIER proxy range sub-TLV is also defined to convey BIER BFR-id information across the routing areas.

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.

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 26, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Problem statement	2
2. Proposal	4
3. Advertisement	6
3.1. BIER proxy range sub-TLV	6
4. Example	7
5. IANA Considerations	8
6. Security Considerations	8
7. Acknowledgements	8
8. Normative References	8
Authors' Addresses	9

1. Problem statement

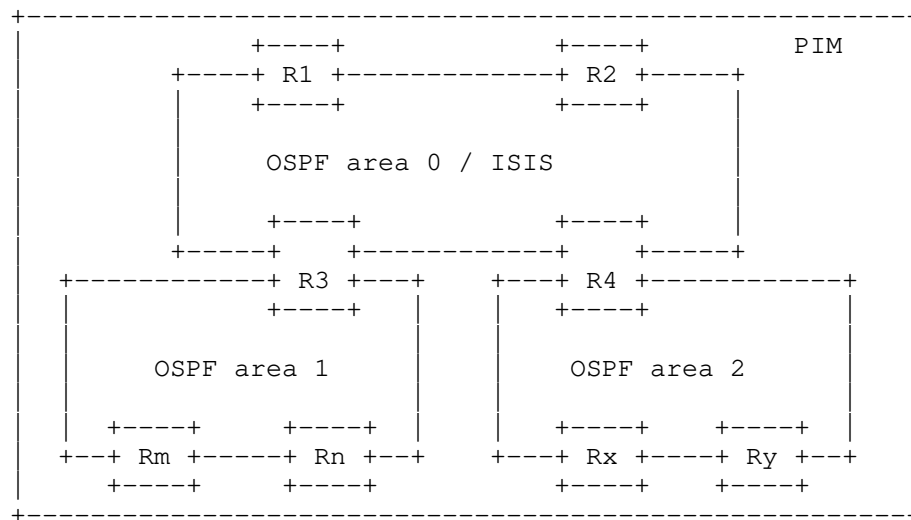


Figure 1

Figure 1 shows that there are three areas in a traditional network. In some deployment situations, different routing protocols may also be used in a network. There are just small number of routers in each area of the network. Currently, multicast services are provided in this hybrid network by using protocol independent feature of PIM.

BIER could be a candidate multicast protocol to replace PIM to reduce multicast states in the hybrid network. BIER [RFC8279] is a new architecture for the forwarding of multicast data packets. It does not require a protocol for explicitly building multicast distribution trees, nor does it require intermediate nodes to maintain any per-flow state. In order to build BIER forwarding plane, BIER key parameters must be flooded in one BIER domain such as BFR-prefix, BFR-id, subdomain-id, and so on. The routing protocols which are used to flood these BIER parameters are called BIER routing underlay. The associated routing protocol extensions are defined in documents such as [RFC8401], [RFC8444], [I-D.ietf-bier-idr-extensions], [I-D.ietf-bier-ospfv3-extensions], and so on.

Figure 2

Except the hybrid network, there is the situation that several areas formed by one same IGP protocol need to be merged into one BIER domain in existing network. The prefix redistribution method defined in this document can be used too.

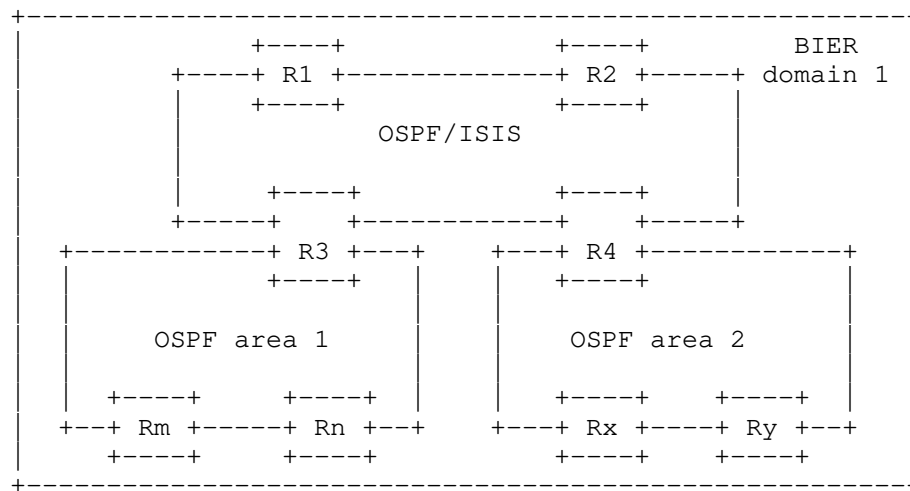


Figure 3

It is more efficient to deploy BIER by creating one BIER domain for the hybrid network to achieve forwarding benefit.

Since the limitation of the BIER routing protocol scope, BFR-id is confined to only one routing area. A BIER proxy function is introduced to transport BIER BFR-id information in a BIER domain across multiple routing protocol areas. So BIER forwarding tables can be built across multiple underlay routing protocols to replace encapsulation/decapsulation processing. In the current deployment, border router (ABR) has a similar role, ABR summarizes unicast routing information from one routing protocol area and sends it to another routing area by new routing protocol messages. So ABR can implement BIER proxy function to summarize BIER BFR-id information from one routing protocol area and sends it to another routing area.

In figure 3, R3 and R4 connect two areas which running different routing protocols, they can be used as BIER proxies to transport BIER information. For example, after R3 receives BFR-ids information from OSPF area 1 and sends it to ISIS routing area, the routers in ISIS routing area can generate BIER forwarding items toward the BFR-ids in OSPF area 1. Similarly, R3 receives BFR-ids information from ISIS area and sends it to OSPF area 1, the routers in OSPF area 1 can build BIER forwarding items toward the BFR-ids in ISIS area. R4 does the same function, the BIER forwarding plane is constructed accordingly.

3. Advertisement

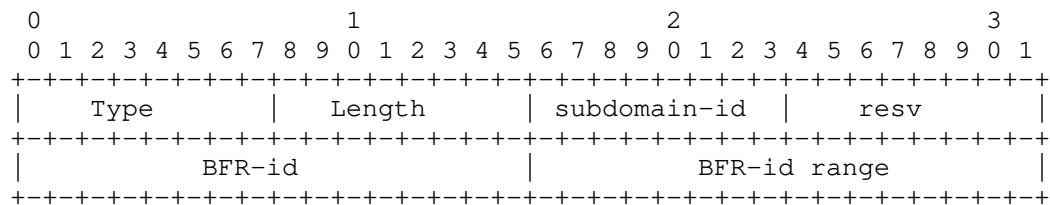
According to [RFC8279], each BFER needs to have a unique (in each sub-domain) BFR-id, and each BFR and BFER floods itself BIER info sub-TLV and associated sub-sub-TLVs in the BIER domain. To keep consistent with the definition in [RFC8444], [I-D.ietf-bier-ospfv3-extensions], and [RFC8401], BIER info sub-TLV defined in [RFC8401] and BIER sub-TLV defined in [RFC8444], [I-D.ietf-bier-ospfv3-extensions] is reused to convey the BFR-id information. OSPF extended Prefix Opaque LSA [RFC7684] and TLVs 235, 237 defined in [RFC5120], or TLVs 135 [RFC5305], or TLV 236 [RFC5308] are still used to carry the BFR-id / BFR-prefix information.

The key parameters got from the original routing protocol should be adapted to the format of next routing protocol, such as BFR-prefix, BFR-id, subdomain-id, and so on. Some parameters like BAR, MT-ID has local significance, So they should be set to same values with BIER proxy own advertisement when BIER proxy advertise them to the next routing area.

And as the two BIER info sub-sub-TLVs (sub-TLVs) including MPLS encapsulation and BSL conversion also have local significance. The information carried in these two sub-sub-TLV need not, but MAY, be advertised to next routing area.

3.1. BIER proxy range sub-TLV

In case unicast default route and aggregated / summarized routes are used in some routing areas and routers in next area can not see the specific BFR-prefix routes from original area, the prefix advertised should be set to default route or aggregated / summarized routes. Like in figure 3, in case R3/R4 does not advertise specific ISIS unicast routes to OSPF area and only advertises unicast default route or aggregated / summarized route to OSPF area 1/2, when R3/R4 advertises BIER info sub-TLV to OSPF area 1/2, R3 MUST advertise the prefix with default route or aggregated / summarized route. In that case, multiple BFR-ids will be mapped to one prefix. In order to advertise BFR-ids optimally, we define a new BIER proxy range sub-TLV to advertise the information of BFR-ids.



- o Type: TBD to indicate the BIER proxy range sub-TLV.
- o Length: variable.
- o Subdomain-id: The subdomain-id from original advertisement.
- o resv: The reserved field.
- o BFR-id: The first BFR-id from original advertisement.
- o BFR-id range: The range of BFR-ids with one subdomain-id.

The BIER proxy range sub-TLV is attached to the aggregated / summarized route prefix or default route prefix. The summarized / aggregated / default prefix may need multiple BIER proxy range sub-TLVs if the BFR-ids covered by the prefix are allocated from different ranges (even if they're from a single range but if some BFR-ids in the range map to some BIER prefixes that are covered by a different summarized / aggregated prefix, then that single large range needs to be broken into smaller ranges).

The BFR-ids associated with the summarized prefix can be advertised individually in the BIER range sub-TLV. Though BFR-id's range can increase advertisement efficiency, necessary configuration / policy should be provided to guide the range generation of BFR-ids. Otherwise unwanted amount of updates may occur when a BFR-id is removed from the range.

Because a summarized / default prefix covers many BIER prefixes, the mapping between a BIER prefix and its BFR-id is no longer conveyed in the routing underlay. As a result, the mapping must be provided by other means, e.g. in the multicast overlay.

4. Example

As in figure 3, R3 and R4 as BIER proxy, R3 as an example should advertise the BIER BFR-ids information from ISIS area to OSPF area 1 with the advertiser set to R3 itself, and advertise BIER info from OSPF area 1 to ISIS area as well. In case R3 and R4 generates specific BFR-prefix and BFR-ids from the original area to the next area, BIER info sub-TLV defined in [RFC8401] and BIER sub-TLV defined in [RFC8444], or [I-D.ietf-bier-ospfv3-extensions] is reused to convey the BFR-id information. All the routers generate BIER forwarding items to other area toward BIER proxy according to [RFC8279].

In case BIER proxy can not advertise specific BFR-prefix but aggregated / summarized / default prefix from the original area to

the next area, BIER proxy range sub-TLV is used to convey the information. Suppose that Rm is an ingress router, R1, R2, Rx and Ry is egress router, the BFR-ids of these egress router are 31, 55, 112, 157. The BFR prefixes of them are 10.1.1.5, 10.1.1.50, 203.1.1.10, 203.1.1.60. Suppose that summarized prefixes are advertised into OSPF area. The summarized prefixes are 10.1.1.0/24 and 203.1.1.0/24. All the routers in OSPF area 1 compute forwarding table for unicast / BIER according to the summarized prefixes, and they can get to these prefixes by routes toward proxy R3.

Rm encapsulate multicast flow with BIER header that with 31, 55, 112 and 157 bit set in the BIER header (Supposed that 256 BitStringLength is used). The routers in OSPF area 1 forward packet toward R3. R3 forwards packet according to the BFR-ids set in the BIER header normally. Later packet reaches R1, R2 and R4. Similarly, R4 forwards packet into OSPF area 2 normally. Finally packet reaches Rx and Ry.

5. IANA Considerations

IANA is requested to set up a new types of sub-TLV (TLV) registry value for BIER proxy range advertisement in OSPF, ISIS, BGP, etc.

6. Security Considerations

Implementations must assure that malformed TLV and Sub-TLV permutations do not result in errors which cause hard protocol failures.

7. Acknowledgements

The authors would like to thank Stig Venaas for his valuable comments and suggestions.

8. Normative References

[I-D.ietf-bier-idr-extensions]

Xu, X., Chen, M., Patel, K., Wijnands, I., and T. Przygienda, "BGP Extensions for BIER", draft-ietf-bier-idr-extensions-07 (work in progress), September 2019.

[I-D.ietf-bier-ospfv3-extensions]

Psenak, P., Kumar, N., and I. Wijnands, "OSPFv3 Extensions for BIER", draft-ietf-bier-ospfv3-extensions-00 (work in progress), May 2019.

- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <<https://www.rfc-editor.org/info/rfc5120>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <<https://www.rfc-editor.org/info/rfc5308>>.
- [RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <<https://www.rfc-editor.org/info/rfc7684>>.
- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.
- [RFC8401] Ginsberg, L., Ed., Przygienda, T., Aldrin, S., and Z. Zhang, "Bit Index Explicit Replication (BIER) Support via IS-IS", RFC 8401, DOI 10.17487/RFC8401, June 2018, <<https://www.rfc-editor.org/info/rfc8401>>.
- [RFC8444] Psenak, P., Ed., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., Zhang, J., and S. Aldrin, "OSPFv2 Extensions for Bit Index Explicit Replication (BIER)", RFC 8444, DOI 10.17487/RFC8444, November 2018, <<https://www.rfc-editor.org/info/rfc8444>>.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation

EMail: zzhang_ietf@hotmail.com

Bo Wu
Individual

EMail: w1973941761@163.com

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

IJsbrand Wijnands
Cisco Systems, Inc.

EMail: ice@cisco.com