BIER Z. Zhang
Internet-Draft Juniper Networks
Intended status: Standards Track July 2, 2018
Expires: January 3, 2019

BIER Penultimate Hop Popping
draft-zzhang-bier-php-00

Abstract

Bit Index Explicit Replication (BIER) can be used as provider tunnel for MVPN/GTM [RFC6514] [RFC7716] or EVPN BUM [RFC7432]. It is possible that not all routers in the provider network support BIER and there are various methods to handle BIER incapable transit routers. However the MVPN/EVPN PEs are assumed to be BIER capable - they are BFIRs/BFERs. This document specifies a method to allow BIER incapable routers to act as MVPN/EVPN PEs with BIER as the transport, by having the upstream BFR (connected directly or indirectly via a tunnel) of a PE remove the BIER header and send the payload to the PE.

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

Table of Contents

1. Terminologies

Familiarity with BIER/MVPN/EVPN protocols and procedures is assumed. Some terminologies are listed below for convenience.

[To be added].

2. Introduction

The BIER architecture includes three layers: the "routing underlay", the "BIER layer", and the "multicast flow overlay". The multicast flow overlay is responsible for the BFERs to signal to BFIRs that they are interested in receiving certain multicast flows so that BFIRs can encode the correct bitstring for BIER forwarding by the BIER layer.

MVPN and EVPN are two similar overlays where BGP Auto-Discovery routes for MVPN/EVPN are exchanged among all PEs to signal which PEs need to receive multicast traffic for all or certain flows.

Typically the same provider tunnel type is used for traffic to reach all receiving PEs.

Consider an MVPN/EVPN deployment where enough P/PE routers are BIER capable for BIER to become the preferred the choice of provider tunnel. However, some PEs cannot be upgraded to support BIER forwarding. While there are ways to allow an ingress PE to send traffic to some PEs with one type of tunnel and send traffic to some other PEs with a different type of tunnel, the procedure becomes complicated and forwarding is not optimized.

One way to solve this problem is to use Penultimate Hop Popping (PHP) so that the upstream BFR can pop the BIER header and send the payload "natively" (note that the upstream BFR can be connected directly or indiretly via a tunnel to the PE). This is similar to MPLS PHP though it is the BIER header that is popped. In case of MPLS encapsulation, even the signaling is similar - a BIER incapable router signals as if it supported BIER, but to request PHP at the penultimate hop, it signals an Implicit Null label instead of a regular BIER label as the Label Range Base in its BIER MPLS Encapsulation sub-TLV.

In order for the PE to be able to correctly forward the packets resulting from the PHP, certain conditions must be met, as specified in Section 3.

While the above text uses MVPN/EVPN as example, BIER PHP is applicable to any scenario where the multicast flow overlay edge router does not support BIER.

This works well if a BIER incapable PE only needs to receive multicast traffic. If it needs to send multicast traffic as well, then it must Ingress Replicate to a BIER capable helper PE, who will in turn relay the packet to other PEs. The helper PE is either a Virtual Hub as specified in [RFC7024] for MVPN and [I-D.keyupate-bess-evpn-virtual-hub] for EVPN, or an AR-Replicator as specified in [I-D.ietf-bess-evpn-optimized-ir] for EVPN.

3. Specifications

The procedures in this section can be applied only if, by means outside the scope of this document, it is known that one of the following conditions is met.

- o The payload after BIER header is IPv4 or IPv6 (i.e., the Proto field in the BIER header is 4 or 6).

Notice that in this case the Destination Address in the IPv4/IPv6 header must be in the address space for the BIER layer.

o The payload after BIER header is MPLS packet with downstream-assigned label at top of stack (i.e., the Proto field in the BIER header is 2), For example, labels from a Domain-wide Common Block (DCB) are used as specified in [I-D.zzhang-bess-mvpn-evpn-aggregation-label].

For MPLS encapsulation, a BIER incapable router, if acting as a multicast flow overlay router, MUST signal its BIER information as specified in [RFC8401] or [I-D.ietf-bier-ospf-bier-extensions] or [I-D.ietf-bier-idr-extensions], with the Label Range Base in the BIER MPLS Encapsulation sub-TLV set to Implicit Null Label [RFC3032].

For non-MPLS encapsulation, a PHP sub-sub-TLV is included in the BIER sub-TLV attached to the BIER incapable router's BIER prefix to request BIER PHP from other BFRs. The sub-sub-TLV's type is TBD, and the length is 0. The PHP sub-sub-TLV MAY be used for MPLS encapsulation as well.

If a BFR follows section 6.9 of [RFC8279] to handle BIER incapable routers, it must treat a router as BIER incapable if the Label Range Base dvertised by the router is Implicit Null, or if the router advertises a PHP sub-sub-TLV, so that the router is not used as a transit BFR.

If the downstream neighbor for a BIER prefix is the one advertising the prefix with a PHP sub-sub-TLV or with an Implicit Null Label as the Label Range Base in its BIER MPLS Encapsulation sub-sub-TLV, then when the corresponding BIRT or BIFT entry is created/updated, the forwarding behavior MUST be that the BIER header is removed and the payload be sent to the downstream router without the BIER header, either directly or over a tunnel.

4. Security Considerations

To be added.

5. IANA Considerations

This document requests a new sub-sub-TLV type value from the "Sub-sub-TLVs for BIER Info Sub-TLV" registry in the "IS-IS TLV Codepoints" registry:

Type	Name
TBD	BIER PHP Request

This document also requests a new sub-TLV type value from the OSPFv2 Extended Prefix TLV Sub-TLV registry:

Type	Name
TBD	BIER PHP Request

6. Acknowledgements

The authors want to thank Eric Rosen and Antonie Przygienda for their review, comments and suggestions.

7. References

7.1. Normative References

[I-D.ietf-bess-evpn-optimized-ir]
Rabadan, J., Sathappan, S., Henderickx, W., Sajassi, A., Isaac, A., and M. Katiyar, "Optimized Ingress Replication solution for EVPN", draft-ietf-bess-evpn-optimized-ir-03 (work in progress), February 2018.

[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-05 (work in progress), March 2018.

[I-D.ietf-bier-ospf-bier-extensions]
Psenak, P., Kumar, N., Wijnands, I., Dolganow, A., Przygienda, T., Zhang, Z., and S. Aldrin, "OSPFv2 Extensions for BIER", draft-ietf-bier-ospf-bier-extensions-18 (work in progress), June 2018.

[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-00 (work in progress), March 2017.

[I-D.zzhang-bess-mvpn-evpn-aggregation-label]
Zhang, Z., Rosen, E., Lin, W., Li, Z., and I. Wijnands, "MVPN/EVPN Tunnel Aggregation with Common Labels", draft-zzhang-bess-mvpn-evpn-aggregation-label-01 (work in progress), April 2018.

[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>.

[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>.

7.2. Informative References

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <https://www.rfc-editor.org/info/rfc6513>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

[RFC7024] Jeng, H., Uttaro, J., Jalil, L., Decraene, B., Rekhter, Y., and R. Aggarwal, "Virtual Hub-and-Spoke in BGP/MPLS VPNs", RFC 7024, DOI 10.17487/RFC7024, October 2013, <https://www.rfc-editor.org/info/rfc7024>.

[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>.

Author's Address

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net