

BIER
Internet-Draft
Intended status: Standards Track
Expires: 7 June 2025

Z. Zhang
Juniper Networks
4 December 2024

BIER Penultimate Hop Popping
draft-ietf-bier-php-16

Abstract

This document specifies a mechanism for Penultimate Hop Popping (PHP) in the Bit Index Explicit Replication (BIER) architecture. PHP enables the removal of the BIER header by the penultimate router, thereby reducing the processing burden on the final router in the delivery path. This extension to BIER enhances operational efficiency by optimizing packet forwarding in scenarios where the final hop's capabilities or requirements necessitate such handling. The document details the necessary extensions to the BIER encapsulation and forwarding processes to support PHP, providing guidance for implementation and deployment within BIER-enabled networks.

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 BCP 14 [RFC2119] [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 7 June 2025.

Copyright Notice

Copyright (c) 2024 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. Specifications	3
2.1. Signaling	4
2.2. BIRT/BIFT Calculation	5
3. Security Considerations	5
4. Operational Considerations	5
5. IANA Considerations	6
6. Acknowledgements	6
7. References	6
7.1. Normative References	6
7.2. Informative References	8
Author's Address	9

1. Introduction

The Bit Index Explicit Replication (BIER) architecture [RFC8279] consists of three layers: the "routing underlay", the "BIER layer", and the "multicast flow overlay". The multicast flow overlay is responsible for allowing BIER Forwarding Egress Routers (BFERs) to signal to BIER Forwarding Ingress Routers (BFIRs) their interest in receiving specific multicast flows, enabling BFIRs to encode the appropriate bitstring for forwarding by the BIER layer.

Multicast Virtual Private Network (MVPN) [RFC6513] [RFC6514] and Ethernet VPN (EVPN) [RFC7432] are two analogous overlays in which BGP Auto-Discovery routes for MVPN/EVPN are exchanged among all Provider Edge (PE) routers to signal which PEs should receive multicast traffic for all or certain flows. Typically, a consistent provider tunnel type is used for traffic delivery to all receiving PEs.

In a deployment scenario where MVPN/EVPN is in use and a sufficient number of provider routers support BIER, BIER can become the preferred provider tunnel type [RFC8556] [RFC9624]. However, some PEs may lack the capability to support BIER forwarding (they are referred to as BIER-incapable in this document). While it is possible for an ingress PE to send traffic to some PEs using one type of tunnel and to others using a different type, such a procedure can be complex and may result in suboptimal forwarding.

A potential solution to this issue is the use of Penultimate Hop Popping (PHP), whereby the upstream BIER Forwarding Router (BFR) pops the BIER header [RFC8296] and transmits the payload without BIER encapsulation. This transmission can occur either directly or indirectly through any type of tunnel to the PE. This mechanism is analogous to Multi-Protocol Label Switching (MPLS) PHP, except that the BIER header is removed.

The transition from an existing MVPN/EVPN deployment with conventional provider tunnels to a BIER-based solution, where some PEs are not BIER-capable, can be incremental. Initially, all PEs are upgraded to support BIER in the control plane, with those unable to perform BIER forwarding requesting PHP. Subsequently, BIER-capable ingress PEs can independently and incrementally switch to BIER transport.

While MVPN/EVPN is used as an example in the above discussion, BIER PHP is applicable to any scenario where the multicast flow overlay edge router does not support BIER, provided that the edge router does not need to identify the transmitting BFIR or participate in BIER Operations, Administration, and Maintenance (OAM) procedures.

This approach is effective when a BIER-incapable PE only needs to receive multicast traffic. However, if the PE also needs to send multicast traffic, it must perform Ingress Replication to a BIER-capable helper PE, which will then relay the packet to other PEs. The helper PE may be a Virtual Hub as defined in [RFC7024] for MVPN and [I-D.ietf-bess-evpn-virtual-hub] for EVPN, or an AR-Replicator as defined in [RFC9574] for EVPN.

2. Specifications

The BIER Penultimate Hop Popping (PHP) mechanism is designed specifically for scenarios where a multicast flow overlay router within a BIER domain does not support BIER forwarding, either completely or for specific BitStringLengths (BSL). In the latter case, PHP applies only to BIER packets with those particular BSLs. If the flow overlay router were capable of BIER forwarding, it would function as a BFER, and PHP would not be performed by the penultimate

hop.

The procedures outlined in this section are applicable only if, through means outside the scope of this document, it is established that all potential penultimate hop BFRs are capable of supporting PHP (i.e., able to remove the BIER header when forwarding to a requesting flow overlay router) and that the payload following the BIER header is one of the following:

- * MPLS packets with a downstream-assigned label at the top of the stack (i.e., the Proto field in the BIER header is set to 1). For instance, a label from a Domain-wide Common Block (DCB) as specified in [RFC9573].
- * IPv4/IPv6 multicast packets for which the Reverse Path Forwarding (RPF) check is disabled.

2.1. Signaling

In IS-IS signaling, a sub-TLV nested within another sub-TLV is referred to as a sub-sub-TLV (and further levels are possible, such as sub-sub-sub-TLV). In other signaling protocols, a sub-TLV nested within another sub-TLV is still referred to as a sub-TLV. For convenience, this document uses the term "sub-TLV" even when referring to a sub-sub-TLV in IS-IS, as there is no ambiguity in the terminology (e.g., MPLS Encapsulation).

A BIER-incapable router, when functioning as a multicast flow overlay router for BIER, MUST signal its BIER information as specified in [RFC8401], [RFC8444], [I-D.ietf-bier-ospfv3-extensions], or [I-D.ietf-bier-idr-extensions], including a PHP sub-TLV within the BIER sub-TLV (or TLV in the case of BGP) attached to the BIER-incapable router's BFR-prefix to request BIER PHP from other BFRs. The type of the sub-TLV or sub-sub-TLV is TBD, and the length is 0.

For MPLS encapsulation, the BIER-incapable multicast flow overlay router MAY omit the BIER MPLS Encapsulation sub-TLV, or it MUST set the Label field in the BIER MPLS Encapsulation sub-TLV to the Implicit Null Label [RFC3032].

In the case of MPLS encapsulation, if a BFER (which supports BIER but not a specific BSL) does not support a particular BSL, it MAY advertise a corresponding BIER MPLS Encapsulation sub-TLV with the Label field set to the Implicit Null Label to request PHP for that BSL. In this scenario, the PHP sub-TLV MUST NOT be included.

For non-MPLS encapsulation [I-D.ietf-bier-lsr-non-mpls-extensions], the BIER-incapable multicast flow overlay router MAY omit the BIER non-MPLS Encapsulation sub-TLV, or it MUST set the BIFT-id field in the BIER non-MPLS Encapsulation sub-TLV to 0.

Similarly, for non-MPLS encapsulation, if a BFER (which supports BIER but not a specific BSL) does not support a particular BSL, it MAY advertise a corresponding BIER non-MPLS Encapsulation sub-TLV but set the BIFT-id field to 0 to request PHP for that BSL. In this scenario, the PHP sub-TLV MUST NOT be included.

2.2. BIRT/BIFT Calculation

If a BFR adheres to Section 6.9 of [RFC8279] for handling BIER-incapable routers, it MUST treat a router as BIER-incapable for a specific BSL if the label in the corresponding MPLS Encapsulation sub-TLV advertised by the router is Implicit Null, or if the BIFT-id in the corresponding non-MPLS Encapsulation sub-TLV is 0. Additionally, the router MUST be treated as BIER-incapable for all BSLs if it advertises a PHP sub-TLV. Consequently, the router will not be utilized as a transit BFR for certain or all BSLs.

When the downstream neighbor (whether determined through IGP calculation or indicated in the BIER Nexthop sub-TLV in the case of BGP) for a BFR-prefix is the router that advertises the prefix with a PHP sub-TLV, an Implicit Null Label in its BIER MPLS Encapsulation sub-TLV, or a BIFT-id of 0 in its BIER non-MPLS Encapsulation sub-TLV, then, upon the creation or update of the corresponding BIRT or BIFT entry, the forwarding behavior MUST be that the BIER header is removed and the payload is forwarded to the downstream router without the BIER header, either directly or over any type of tunnel.

3. Security Considerations

This specification does not introduce additional security concerns beyond those already discussed in BIER architecture [RFC8279] and BIER signaling extensions to OSPF [RFC8444], IS-IS [RFC8401] and BGP [I-D.ietf-bier-idr-extensions].

4. Operational Considerations

BIER PHP can only be used when the conditions specified in Section 2 are met. The BIER OAM functionality is not available on the BIER-incapable flow overlay routers, but using PHP when the conditions are met is simpler than the alternative of using BIER to send to some whereas using non-BIER tunnels to send to other flow overlay routers.

[RFC8279] states that when the BIER header is decapsulated, by default its Time to Live (TTL) value is not inherited by the payload. That means in the PHP case, the payload's TTL will be decremented once more than in the non-PHP case.

5. IANA Considerations

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

Type	Name
----	----
TBD	BIER PHP Request

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

Type	Name
----	----
TBD	BIER PHP Request

This document requests a new sub-TLV type value from the OSPFv3 Extended LSA Sub-TLVs registry:

Type	Name	L2BM
----	----	----
TBD	BIER PHP Request	X

This document requests a new sub-TLV type value from the BGP BIER TLV sub-TLV Types registry requested in [I-D.ietf-bier-idr-extensions]:

Type	Name
----	----
TBD	BIER PHP Request

6. Acknowledgements

The author wants to thank Eric Rosen and Antonie Przygienda for their review, comments and suggestions. The author also wants to thank Senthil Dhanaraj for the suggestion of requesting PHP if a BFER does not support certain BSL.

7. References

7.1. Normative References

- [I-D.ietf-bier-idr-extensions]
Xu, X., Chen, M., Patel, K., Wijnands, I., Przygienda, T.,
and Z. J. Zhang, "BGP Extensions for BIER", Work in
Progress, Internet-Draft, draft-ietf-bier-idr-extensions-
15, 22 November 2024,
<<https://datatracker.ietf.org/doc/html/draft-ietf-bier-idr-extensions-15>>.
- [I-D.ietf-bier-lsr-non-mpls-extensions]
Dhanaraj, S., Yan, G., Wijnands, I., Psenak, P., Zhang, Z.
J., and J. Xie, "LSR Extensions for BIER non-MPLS
Encapsulation", Work in Progress, Internet-Draft, draft-
ietf-bier-lsr-non-mpls-extensions-03, 7 February 2024,
<<https://datatracker.ietf.org/doc/html/draft-ietf-bier-lsr-non-mpls-extensions-03>>.
- [I-D.ietf-bier-ospfv3-extensions]
Psenak, P., Nainar, N. K., and I. Wijnands, "OSPFv3
Extensions for BIER", Work in Progress, Internet-Draft,
draft-ietf-bier-ospfv3-extensions-07, 1 December 2022,
<<https://datatracker.ietf.org/doc/html/draft-ietf-bier-ospfv3-extensions-07>>.
- [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>>.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y.,
Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack
Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001,
<<https://www.rfc-editor.org/info/rfc3032>>.
- [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>>.
- [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>>.
- [RFC9573] Zhang, Z., Rosen, E., Lin, W., Li, Z., and IJ. Wijnands, "MVPN/EVPN Tunnel Aggregation with Common Labels", RFC 9573, DOI 10.17487/RFC9573, May 2024, <<https://www.rfc-editor.org/info/rfc9573>>.
- [RFC9624] Zhang, Z., Przygienda, T., Sajassi, A., and J. Rabadan, "EVPN Broadcast, Unknown Unicast, or Multicast (BUM) Using Bit Index Explicit Replication (BIER)", RFC 9624, DOI 10.17487/RFC9624, August 2024, <<https://www.rfc-editor.org/info/rfc9624>>.

7.2. Informative References

- [I-D.ietf-bess-evpn-virtual-hub] Patel, K., Sajassi, A., Drake, J., Zhang, Z. J., and W. Henderickx, "Virtual Hub-and-Spoke in BGP EVPNs", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-virtual-hub-00, 26 January 2020, <<https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-virtual-hub-00>>.
- [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>>.
- [RFC9574] Rabadan, J., Ed., Sathappan, S., Lin, W., Katiyar, M., and A. Sajassi, "Optimized Ingress Replication Solution for Ethernet VPNs (EVPNs)", RFC 9574, DOI 10.17487/RFC9574, May 2024, <<https://www.rfc-editor.org/info/rfc9574>>.

Author's Address

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net