Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 2, 2020

J. Xie
Huawei Technologies
M. McBride
Futurewei
S. Dhanaraj
Huawei Technologies
L. Geng
China Mobile

July 1, 2019

Use of BIER IPv6 Encapsulation (BIERv6) for Multicast VPN in IPv6 networks
draft-xie-bier-ipv6-mvpn-01

Abstract

This draft defines the procedures and messages for using Bit Index Explicit Replication (BIER) for Multicast VPN Services in IPv6 networks using the BIER IPv6 encapsulation. It provides a migration path for Multicast VPN service using BIER MPLS encapsulation in MPLS networks to multicast VPN service using BIER IPv6 encapsulation (BIERv6) in IPv6 networks.

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 2, 2020.

Table of Contents

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. BIERv6 refers to the deployment of BIER in IPv6 networks using the BIER IPv6 encapsulation format defined in [I-D.xie-bier-ipv6-encapsulation].

[I-D.ietf-spring-srv6-network-programming] introduces the Network programming concepts in SRv6 networks and explains how the 128-bit IPv6 address can be used as SRv6 SID in the format LOC:FUNCT, where LOC part of the SID is routable, while FUNCT part of the SID is an opaque identification of a local function bound to the SID. It has

also defined some well known standard functions like End.DT4 - Endpoint with decaps and IPv4 table lookup for L3VPN (equivalent to per-VRF VPN label).

[I-D.dawra-bess-srv6-services] defines the TLVs to associate a function like End.DT4 with the L3VPN Unicast routes advertised via BGP. It also details how the functions of End.DT4, End.DT6, End.DT46 (End.DTx) can be used to identify a L3VPN/EVPN instead of using a VPN Label in MPLS-VPN [RFC4364] of the received data packet and thereby realize the L3VPN Services in the SRv6 Networks. However, it covers unicast services exclusively.

This document describes a method to realize MVPN services using BIER as a P-tunnel in the IPv6 Networks (BIERv6 Networks). It defines a method to use an SRv6 Service SID, called Src.DTx in this document, as source address of an IPv6 header, to identify the MVPN instance at the Egress PE. The LOC part and FUNCT part of this SRv6 Service SID represent the context and the upstream-assigned VPN Label respectively in MVPN scenario's as defined in [RFC8556].

In particular, MVPN deployment in IPv6 networks relies on L3VPN deployment on IPv6 networks firstly, thus the c-multicast routing procedure like UMH Selection can be done. The L3VPN deployment in IPv6 networks can be referred to [I-D.dawra-bess-srv6-services].

GTM defined in [RFC7716] is also covered in this document, as GTM shares the same BGP-MVPN signaling, while providing an approach of Non-VPN multicast over a service provider core with various P-tunnel type. For the same reason of UMH selection, and the requirement of basic operation like ping (e.g, to the multicast source address), the Global IPv4/IPv6 over SRv6 Core as described in [I-D.dawra-bess-srv6-services] is also required.

2. Terminology

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References. Additionally the following terms are used through out the document.

- BIERv6 - BIER in IPv6 networks using the BIERv6 encapsulation format defined in [I-D.xie-bier-ipv6-encapsulation].
- SRv6 - Segment Routing instantiated on the IPv6 dataplane as defined in [I-D.ietf-spring-srv6-network-programming].
- SRv6 SID - SRv6 Segment Identifier as defined in [I-D.ietf-spring-srv6-network-programming].

- o End.DTx - Refers to the functions End.DT6, End.DT4, End.DT46 defined in [I-D.ietf-spring-srv6-network-programming].

- o Src.DTx - Refers to the functions Src.DT4, Src.DT6, Src.DT46 defind in this document.

- o SRv6 L3 Service - L3VPN/Global-L3 service in SRv6 networks defined in [I-D.dawra-bess-srv6-services], or MVPN/GTM service in BIERv6 networks defined in this document.

3. Use of PTA and Prefix-SID Attribute in x-PMSI A-D Routes

The BGP-MVPN I-PMSI A-D (Type 1) or S-PMSI A-D (Type 3) route (called x-PMSI A-D route in this document), advertised by Ingress PE carries the BIER (Type 11) PTA as specified in [RFC8556]. The BIER PTA carried in the x-PMSI A-D route is used for explicitly tracking the receiver-site PEs which are interested in a specific multicast flow. It includes three BIER-specific fields, Sub-domain-id, BFR-id, and BFR-prefix. For BIER P-tunnel using the BIERv6 encapsulation in IPv6 networks, the BFR-prefix field in the PTA MUST be set to the BFIR IPv6 prefix and the MPLS Label field in the PTA MUST set to 0. For MVPN over BIERv6, the Src.DTx IPv6 address of the BFIR is used to identify the VRF instead of an MPLS Label. The Src.DTx IPv6 Address (Src.DT6 or Src.DT4 or Src.DT46) MUST be carried within an SRv6 L3 Service TLV [I-D.dawra-bess-srv6-services] of BGP Prefix-SID attribute in the x-PMSI A-D route.

The Ingress PE encapsulates the c-multicast IP packet with BIERv6 header and the source address in the outer IPv6 header will be set to the Src.DTx IPv6 address advertised in the BGP-MVPN x-PMSI A-D routes. See section 3 of [I-D.xie-bier-ipv6-encapsulation] for the detailed packet format.

Egress PE (BFER) receiving the x-PMSI A-D routes with BIER PTA and SRv6 L3 Service TLV learns the Src.DTx IPv6 address and uses it to identify the VRF of the c-multicast packet.

When Egress PE receives a BIERv6 packet and the self bfr-id is set in the bit-string field of the BIERv6 header, it retrieves the Src.DTx IPv6 address from the source address of the IPv6 header to determine the VRF and the Address Family (AF) of the c-multicast data packet, and performs the MFIB lookup in the corresponding table.

4. MVPN over BIERv6 Core

[RFC8556] specifies the protocol and procedures to be followed by the Ingress and Egress PEs to use BIER as a P-tunnel for MVPN in MPLS networks. This section specifies the required changes and procedures

in addition to support BIER as a P-tunnel in IPv6 networks using BIERv6.

In a IPv6 service provider network, many of the IP address fields used in the BGP-MVPN routes are IPv6 address as specified in [RFC6515]. These are listed below.

- o "Originating Router's IP Address" in the NLRI of Type 1 or Type 3 BGP-MVPN route is an IPv6 address.
- o "Network Address of Next Hop" field in the MP_REACH_NLRI attribute is an IPv6 address.
- o Route Targets Extended Community (EC) used in C-multicast join (Type 6 or 7) route or Leaf A-D (Type 5) route is an IPv6 Address Specific Extended Community, where the Global Administrator field will be an IPv6 address identifies the Upstream PE or the UMH.
- o "VRF Route Import Extended Community (EC)" carried by unicast VPN-IPv4 or VPN-IPv6 routes as [RFC6515] specifies, or SAFI 1, 2, or 4 unicast routes, or MVPN (SAFI 5) Source-Active routes as [RFC7716] specifies.

On the Ingress PE (BFIR), the BGP-MVPN x-PMSI A-D route is constructed as per the procedures specified in [RFC8556] and with the following specifications.

- o MPLS Label field in the BIER PTA MUST be set to Zero.
- o BFR-prefix field in the BIER PTA MUST be set to the Ingress PEs (BFIR) IPv6 BFR-Prefix Address. It does not need to be the same as the other IPv6 address of the x-PMSI AD route.
- o Route MUST also carry an BGP Prefix SID attribute with an SRv6 L3 Service TLV carrying an Src.DTx IPv6 address uniquely identifying the MVPN instance.

If the MVPN is IPv4 MVPN, the Src.DTx can be either Src.DT4 or Src.DT46. If the MVPN is IPv6 MVPN, the Src.DTx can be either Src.DT6 or Src.DT46. The distribution of the x-PMSI A-D routes uses the Src.DTx according to the local configuration, and is independent to the use of End.DTx in VPN-IP unicast routes of this VPN. For example, one can use End.DT46 for VPNv4 and VPNv6 unicast routes, but use Src.DT4 for the MVPN routes for the same VPN. Another example, one can use End.DX for VPNv4 unicast routes, but use Src.DT46 for the MVPN routes for the same VPN.

BFIR MAY carry the BGP Prefix-SID attribute only in I-PMSI A-D route when I-PMSI A-D route is used, while other S-PMSI A-D routes do not carry the BGP Prefix-SID attribute.

BFIR MAY carry the BGP Prefix-SID attribute only in wildcard S-PMSI A-D routes when the "S-PMSI Only" mode as described in [RFC6625] is used, while other S-PMSI A-D routes do not carry the BGP Prefix-SID attribute.

On the Egress PE (BFER), the BGP-MVPN x-PMSI A-D route is processed as per the procedures specified in [RFC8556] and with the following specifications:

- The MPLS Label field in the BIER PTA of the BGP-MVPN x-PMSI A-D route MUST be ignored and MUST not be used for the identification of the VRF.

- The BGP-MVPN x-PMSI A-D route MUST be dropped if the BFR-prefix field in the BIER PTA is not an IPv6 address.

- The BGP-MVPN x-PMSI A-D route MUST be dropped if it does not carry a Src.DTx IPv6 address in the SRv6 L3 Service TLV in BGP Prefix SID attribute.

- Leaf A-D route originated by the Egress PE (BFER) MUST carry the BIER PTA with the BFR-prefix field set to the BFER IPv6 BFR-prefix.

Valid BGP-MVPN x-PMSI A-D route received by an Egress PE (BFER) is stored locally, and the Src.DTx IPv6 Address carried in the SRv6 L3 service TLV is used to identify the VRF of a c-multicast data packet. This may be populated into forwarding table only when there is c-multicast flow state with UMH of the specific BFIR this Src.DTx located in.

If more than one x-PMSI A-D routes belonging to the same VRF has different Src.DTx value, the processing is determined by the local policy of the BFER.

If more than one x-PMSI A-D routes belonging to different VRF has the same Src.DTx value, the BFER must log an error, and a BIERv6 packet with this Src.DTx as the IPv6 source address MUST be dropped.

The BGP Prefix-SID attribute (which may include the Src.DTx in SRv6 L3 Service TLV) MUST NOT be carried in Leaf A-D route upon sending, and MUST be ignored upon reception.

5. GTM over BIERv6 Core

As specified in [RFC7716], Global Table Multicast (GTM) uses the same Subsequent Address Family Identifier (SAFI) value, the same Network Layer Reachability Information (NLRI) format, and the same procedures of MVPN with only a few adaptions. It support for both IPv4 and IPv6 multicast flows over either an IPv4 or IPv6 SP infrastructure. GTM over BIERv6 core is obviously a case of IPv4/IPv6 multicast over an IPv6 SP infrastructure with BIERv6 data-plane.

The BIER (Type 11) PTA attribute and the BGP Prefix-SID attribute are carried in the x-PMSI A-D route in GTM cases. When the a BGP-MVPN x-PMSI A-D route is received by Egress PE, it is stored locally, and the Src.DTx IPv6 Address of the Ingress PE in the route is used to determine the VRF of a packet, which is the 'public' VRF in the case of GTM.

There are some other attributes listed below for GTM over a BIERv6 core:

- Route Distinguishers - the RD field of a BGP-MVPN route's NLRI MUST be set to zero (i.e., to 64 bits of zero) to represent a Non-VPN GTM. See section 2.2 of [RFC7716].

- Route Targets Extended Community (EC) - The RT EC carried by the BGP-MVPN C-multicast (Type 6 or 7) route or Leaf A-D (Type 4) route MUST be an IPv6-address-specific Extended Community (EC). The Global Administrator field identifies the Upstream PE or the UMH, and the Local Administrator field MUST always be set to zero in GTM case.

- VRF Route Import Extended Community (EC) - The VRF Route Import EC used in BIERv6 core MUST be an IPv6-address-specific EC if used, either used in UMH-eligible unicast routes having a SAFI of 1, 2, or 4, or used in the MVPN (SAFI of 5) Source Active A-D route.

GTM IPv4 multicast over an BIERv6 core may be considered an alternative to support IPv4 IPTV content delivery during transition to IPv6 period comparing to [RFC8114]. They both use IPv4-in-IPv6 encapsulation, while BIERv6 uses an additional BIER header within an IPv6 Extension header to support stateless core.

6. Data Plane

6.1. Encapsulation of Multicast Traffic

BIER IPv6 encapsulation (BIERv6) [I-D.xie-bier-ipv6-encapsulation] is used for forwarding the c-multicast traffic through an IPv6 core. The following diagram shows the progression of an MVPN c-multicast packet as it enters and leaves the intra-AS service-provider network.

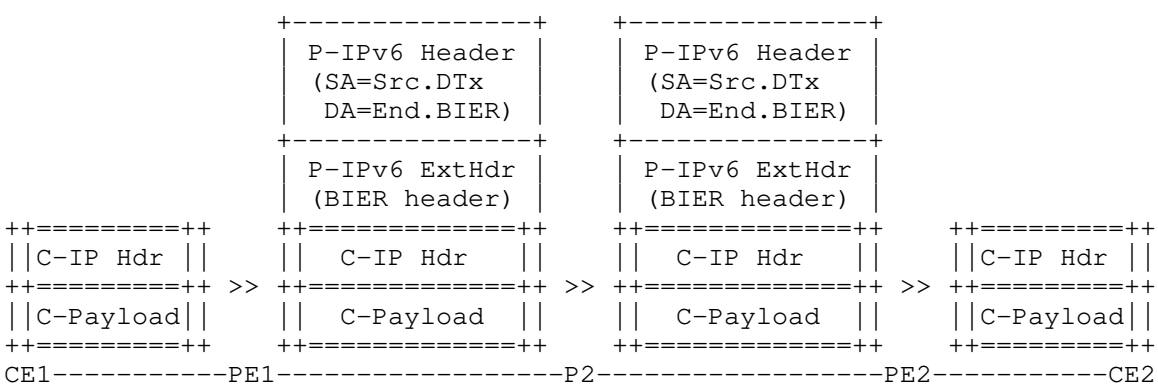

Figure 1: BIERv6 MVPN/GTM Intra-AS

In case of inter-AS scenario, BIERv6 packets may travel through unicast to a Boarder Router (BR), and then replicate in a single intra-AS BIERv6 domain. How such non-segmented BIERv6 scenario can be supported is outside the scope of this document.

How segmented MVPN, for example, between BIERv6 and BIERv6, or between BIERv6 and Ingress Replication(IR) in Non-MPLS IPv6 networks, is outside the scope of this document.

The Src.DTx SHOULD support as destination address of an ICMPv6 packet. The following is an example pseudo-code of the Src.DTx function as destination address:

```
1. IF Last_NH = ICMPv6   ;;Ref1
2.   Send to CPU.
3. ELSE
4.   Drop the packet.
```

Ref1: ICMPv6 packet using Src.DT4, Src.DT6 or Src.DT46 as destination address.

6.2. MTU

Each BFIR is expected to know the Maximum Transmission Unit (MTU) of the BIER domain. This may be known by provisioning, or by method specified in [draft-ietf-bier-mtud]. The section 3 of [RFC8296] applies.

6.3. TTL

The ingress PE (BFIR) should not copy the Time to Live (TTL) field from the payload IP header received from a CE router to the delivery IP header. Setting the TTL of the delivery IP header is determined by the local policy of the ingress PE (BFIR) router per section 3 of [RFC8296].

7. Security Considerations

The security considerations SEC-1, SEC-2, SEC-3 defined in [I-D.ietf-spring-srv6-network-programming] apply equally to this document.

8. IANA Considerations

Allocation is expected from IANA for the following Src.DTx functions codepoints from the "SRv6 Endpoint Behaviors" sub-registry.

Values 68, 69, 70 is suggested for Src.DT6, Src.DT4, Src.DT46 respectively.

Value	Hex	Endpoint function	Reference
TBD	TBD	Src.DT6	This draft
TBD	TBD	Src.DT4	This draft
TBD	TBD	Src.DT46	This draft

Src.DT6 Source address indicating decapsulation and IPv6 table lookup e.g. IPv6-MVPN (equivalent to per-VRF VPN label in RFC8556)
Src.DT4 Source address indicating decapsulation and IPv4 table lookup e.g. IPv4-MVPN (equivalent to per-VRF VPN label in RFC8556)
Src.DT46 Source address indicating decapsulation and IP table lookup e.g. IP-MVPN (equivalent to per-VRF VPN label)

9. Acknowledgements

TBD.

10. References

10.1. Normative References

[I-D.dawra-bess-srv6-services]
Dawra, G., Filsfils, C., Dukes, D., Brissette, P., Sethuram, S., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Decraene, B., Matsushima, S., and S. Zhuang, "SRv6 BGP based Overlay services", draft-dawra-bess-srv6-services-00 (work in progress), March 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-00 (work in progress), April 2019.

[I-D.xie-bier-ipv6-encapsulation]
Xie, J., Geng, L., McBride, M., Dhanaraj, S., Yan, G., and Y. Xia, "Encapsulation for BIER in Non-MPLS IPv6 Networks", draft-xie-bier-ipv6-encapsulation-01 (work in progress), June 2019.

[RFC6515] Aggarwal, R. and E. Rosen, "IPv4 and IPv6 Infrastructure Addresses in BGP Updates for Multicast VPN", RFC 6515, DOI 10.17487/RFC6515, February 2012, <https://www.rfc-editor.org/info/rfc6515>.

[RFC6625] Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R. Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes", RFC 6625, DOI 10.17487/RFC6625, May 2012, <https://www.rfc-editor.org/info/rfc6625>.

[RFC7716] Zhang, J., Giuliano, L., Rosen, E., Ed., Subramanian, K., and D. Pacella, "Global Table Multicast with BGP Multicast VPN (BGP-MVPN) Procedures", RFC 7716, DOI 10.17487/RFC7716, December 2015, <https://www.rfc-editor.org/info/rfc7716>.

[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>.

10.2. Informative 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>.

Authors' Addresses

Jingrong Xie
Huawei Technologies

Email: xiejingrong@huawei.com

Mike McBride
Futurewei

Email: mmcbride7@gmail.com

Senthil Dhanaraj
Huawei Technologies

Email: senthil.dhanaraj@huawei.com

Liang Geng
China Mobile

Email: gengliang@chinamobile.com