Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 17, 2020

L. Geng
China Mobile
J. Xie
Huawei Technologies
M. McBride
Futurewei
G. Yan
Huawei Technologies

January 14, 2020

Inter-Domain Multicast Deployment using BIERv6 draft-geng-bier-ipv6-inter-domain-01

Abstract

Bit Index Explicit Replication IPv6 encapsulation (BIERv6) introduces an approach to use IPv6 extension header to carry BIER header with IPv6 unicast address as destination address. It provides the ability to replicate a packet from one router to another router in a different domain as well as in the same domain. This document introduces the techniques for multicast deployment across multiple domains using BIERv6, and demonstrate how BIERv6 is beneficial for such deployment.

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 July 17, 2020.

Table of Contents

1. Introduction

Bit Index Explicit Replication [RFC8296] IPv6 encapsulation (BIERv6) described in [I-D.xie-bier-ipv6-encapsulation] introduces an approach to use IPv6 extension header to carry BIER header. One BIERv6 option, using IPv6 unicast address as destination address provides the ability to replicate a packet from one router to another router in a different domain as well as in the same domain. This document introduces the techniques for multicast deployment across multiple domains using BIERv6, and demonstrates how BIERv6 is beneficial for such deployment.

2. Terminology

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

3. Inter-domain Multicast Overview

It is common to deploy multicast services across multiple domains.

One typical scenario for this type of deployment is in a service-provider network for MVPN service as described in [I-D.ietf-bier-ipv6-requirements]. Service provider network tends to be very heterogeneous with full-mesh backbone network, and metro networks with fabric for dense area coverage or ring-shaped for sparse area coverage. The backbone network and metro networks are autonomous systems interconnected by border routers (BRs). Multicast-based delivery of video need to be set up from a source router on the backbone to each of the boundary routers of each metro network.

This scenario may have some variant. For example, multicast source router is a Top of Rack (TOR) switch in a service provider data center(SPDC) connected to backbone with data center gateway(s) (DC-GW), and multicast receiver is the home broadband subscribers connected to boundary routers (e.g. BNG) of each metro network. Operators may want to set up multicast-based delivery from TOR to BNGs seamlessly without segmentation or stitching on DC-GW(s) or BR(s).

It is described as hierachical multicast in this document.

Another typical scenario for inter-domain multicast deployment is in peering network as described in [RFC8313] to set up multicast-based delivery of content across inter-domain peering points.

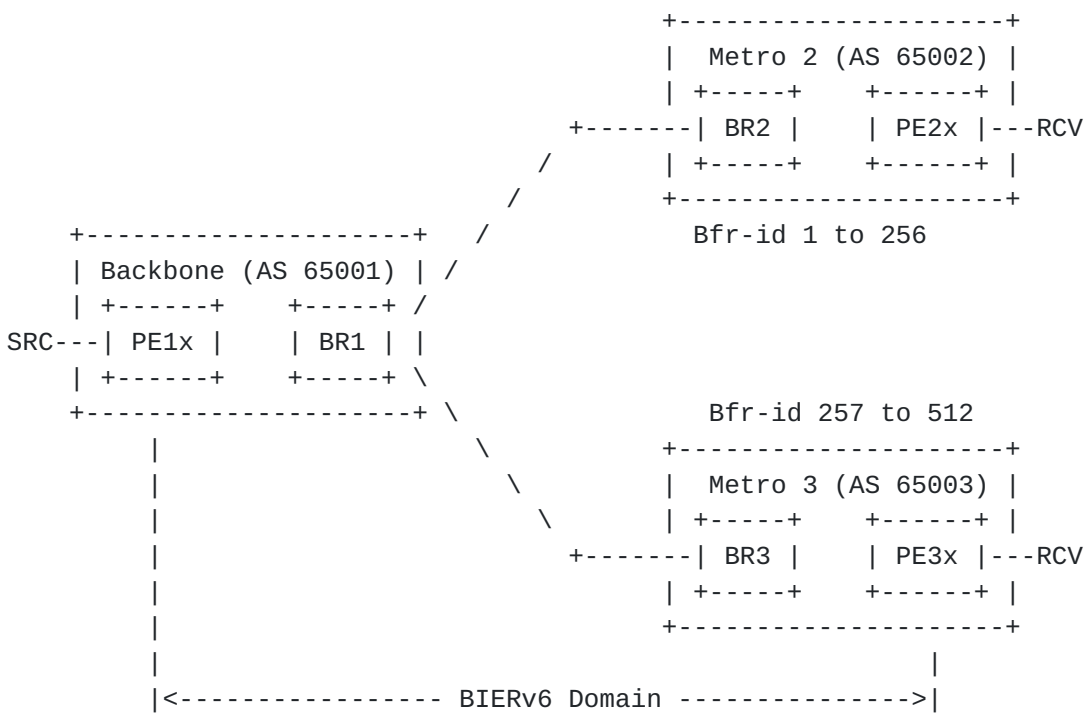
This scenario may have some variant. For example, interconnected content delivery networks (CDNs) (described in [RFC6770]) owned by Network Service Providers (NSPs) or Enterprise Service Providers may need to deliver multicast from one to others.

It is described as peering multicast in this document.

4. Inter-domain Multicast Deployment using BIERv6

4.1. Hierarchical Multicast

Following is an example of hierarchical deployment of multicast.



BR = Border Router
SRC = Multicast Source
RCV = Multicast Receiver

Figure 1: Inter-Domain Hierarchical Multicast

Multicast source is connected to PE1x, and multicast receivers are connected to PE2x and PE3x.

PE1x, PE2x, PE3x is located in Backbone (AS 65001), Metro 2 (AS 65002), and Metro 3 (AS 65003) respectively, and BR1, BR2, BR3 is boarder of these three domains. They belong to a single administrative domain.

IGP underlay for BIERv6 is deployed in Metro2, Metro3 respectively. The bfr-ids in Metro2 and Metro 3 should be divided rationally.

PE1x, PE2x, PE3x uses 2001:DB8::E1, 2001:DB8::E2, 2001:DB8::E3 as End.BIER IPv6 address respectively.

BR1, BR2, BR3 uses 2001:DB8::B1, 2001:DB8::B2, 2001:DB8::B3 as End.BIER IPv6 address respectively.

All of them use the Non-MPLS static BSL-SD-SI BIFT encoding method described in [I-D.ietf-bier-non-mpls-bift-encoding] as the auto-generation method.

On BR1, static configuration can be used to construct inter-domain BIERv6 forwarding table.

```
    bier sub-domain 6 ipv6-underlay
      bfr-prefix interface loopback0
      end-bier 2001:DB8::B1
      encapsulation ipv6 bsl 256 max-si 2
      static-bift
        nexthop end-bier 2001:DB8::B2 bfr-id 1 to 256
        nexthop end-bier 2001:DB8::B3 bfr-id 257 to 512
```

Accordingly, the following BIFTs will be constructed:

```
    BIFT correspond to SD<6>/BSL<256>/SI<0>
      (neighbor = 2001:DB8::B2, F-BM = ffff....ffff)
    BIFT correspond to SD<6>/BSL<256>/SI<1>
      (neighbor = 2001:DB8::B3, F-BM = ffff....ffff)
```

On PE1x, static configuration can be used to construct inter-domain BIERv6 forwarding table. Note that PE1x doesn't need to assign a valid BFR-id uniquely among many.

```
    bier sub-domain 6 ipv6-underlay
      bfr-prefix interface loopback0
      end-bier 2001:DB8::E1
      encapsulation ipv6 bsl 256 max-si 2
      static-bift
        nexthop end-bier 2001:DB8::B1 bfr-id 1 to 512
```

Accordingly, the following BIFTs will be constructed:

```
    BIFT correspond to SD<6>/BSL<256>/SI<0>
      (neighbor = 2001:DB8::B1, F-BM = ffff....ffff)
    BIFT correspond to SD<6>/BSL<256>/SI<1>
      (neighbor = 2001:DB8::B1, F-BM = ffff....ffff)
```

Use of BGP as inter-domain underlay protocol to advertise the BIER information from BR2 or BR2 to BR1, or from BR1 to PE1x is outside the scope of this document.

On each domain, two redundant border routers may be deployed, and anycase IPv6 address can be used on each pair of BRs as End.BIER IPv6 address.

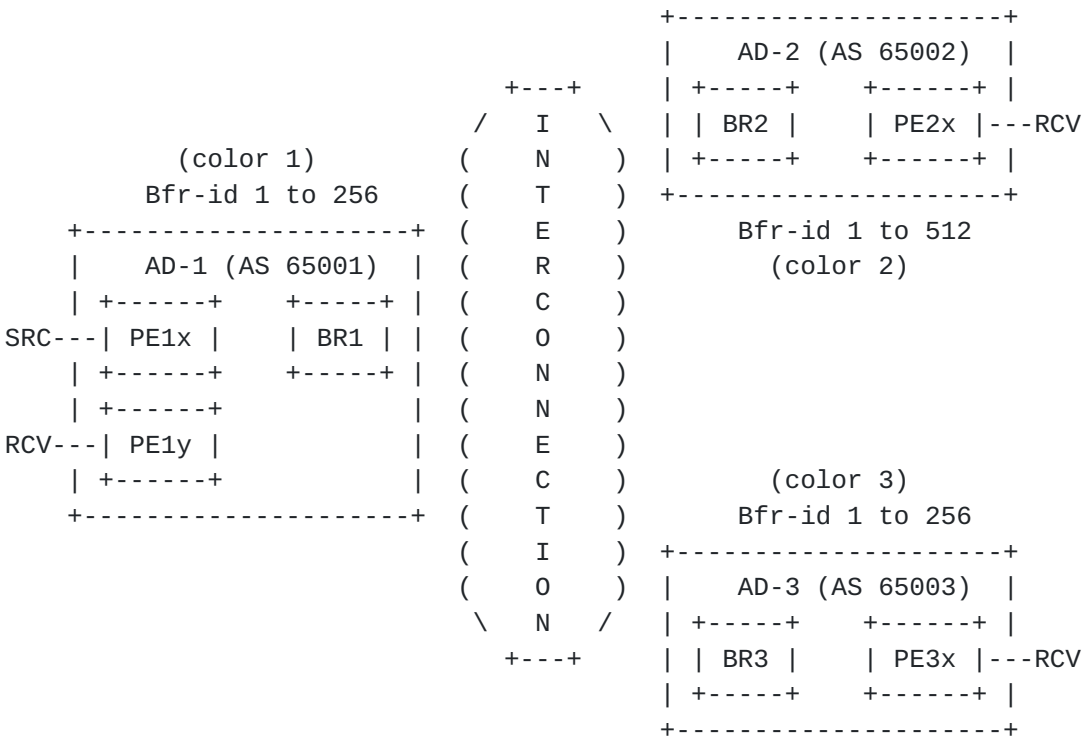
Inter-Domain BIER will converge normally when unicast converge and the BIFT will be reconstructed accordingly.

For multicast overlay layer, there are no extensions needed. MVPN is deployed on PE1x, PE2x and PE3x using sub-domain 6 and bsl 256 without segmentation on border router(s).

Note: Use of the IPv6 address configured on PE1 to identify an MVPN instance can eliminate the need for BFR-id configuration on PE1x, which otherwise has to be configured from the space of a sub-domain.

4.2. Peering Multicast

Following is an example of peering deployment of multicast.



AD = Administrative Domain (independent autonomous system)
BR = Border Router
SRC = Multicast Source
RCV = Multicast Receiver

Figure 2: Inter-Domain Peering Multicast

Each Administrative Domain AD-1, AD-2 or AD-3 is configured a unique color. Color 1, 2, 3 are used in this example.

For routing underlay layer, the ingress router uses IGP protocol (IS-IS as example in this document) for the domain it belongs to, and uses static configuration for the domain it doesn't belong to.

Below is an example of routing underlay configuration on PE1x. Note that PE1x doesn't need to assign a valid BFR-id per color.

```
    # PE1x routing underlay layer configuration
    bier sub-domain 6 ipv6-underlay
      bfr-prefix interface loopback0
      end-bier 2001:DB8::E1
      encapsulation ipv6 bsl 256 max-si 1
      color 1 protocol isis
      color 2 static-bift
        nexthop end-bier 2001:DB8::B2 bfr-id 1 to 512
      color 3 static-bift
        nexthop end-bier 2001:DB8::B3 bfr-id 1 to 256
```

The following lists the BIFT that will be constructed on PE1x:

```
   BIFT corresponding  to SD<6>/BSL<256>/SI<0> for color 1 ;;Ref1
   BIFT corresponding  to SD<6>/BSL<256>/SI<0> for color 2 ;;Ref2
   BIFT corresponding  to SD<6>/BSL<256>/SI<1> for color 2 ;;Ref3
   BIFT corresponding  to SD<6>/BSL<256>/SI<0> for color 3 ;;Ref4
```

Ref1: BIFT constructed using IGP.

Ref2: BIFT constructed using static configuration, with BR2 a multi-hop BFR neighbor of PE1x.

Ref3: BIFT constructed using static configuration, with BR2 a multi-hop BFR neighbor of PE1x.

Ref3: BIFT constructed using static configuration, with BR3 a multi-hop BFR neighbor of PE1x.

For multicast overlay layer, the color extended community defined in [RFC5512] is carried in Leaf A-D route together with the PTA attribute.

(1) PE in each domain gets the color it belongs to. This can be done by configuration on each PE in each domain.

(2) PE carries a color attribute in BGP-MVPN Leaf A-D route when advertising to Ingress PE as response to explicit-tracking initiated by the Ingress PE. This can be done by configuration on MVPN deployment. Refer to [I-D.xie-bier-ipv6-mvpn] for other attributes needed to be used.

(3) The Ingress PE gets the Leaf A-D route, learns the BFERs of a color (representing a domain) interested in a multicast flow, and

constructs the overlay forwarding table. Below is an example of the overlay forwarding table on PE1x:

```
  (VRF<X>, S<S1>, G<G1>)
    (Color<1>, SD<6>, BSL<256>, SI<0>, BitString<0001>) ;;Ref1
    (Color<2>, SD<6>, BSL<256>, SI<0>, BitString<0001>) ;;Ref2
    (Color<2>, SD<6>, BSL<256>, SI<1>, BitString<0001>) ;;Ref3
    (Color<3>, SD<6>, BSL<256>, SI<0>, BitString<0001>) ;;Ref4
```

Ref1: packet will be replicated according to the BitString<0001> and the BIFT constructed using the IGP for SD<6>/BSL<256>/SI<0> for color 1.

Ref2: packet will be replicated according to the BitString<0001> and the BIFT constructed using the static-bift configuration for SD<6>/BSL<256>/SI<0> for color 2.

Ref3: packet will be replicated according to the BitString<0001> and the BIFT constructed using the static-bift configuration for SD<6>/BSL<256>/SI<1> for color 2.

Ref3: packet will be replicated according to the BitString<0001> and the BIFT constructed using the static-bift configuration for SD<6>/BSL<256>/SI<1> for color 3.

Note: BFR-id configuration on PE1x is only necessary when PE1x will act as BFER, for example, there is multicast packet from PE2x to PE1x. The BFR-ids in color 1, 2, 3 is independent on each other.

5. Security Considerations

The procedures of this document do not, in themselves, provide privacy, integrity, or authentication for the control plane or the data plane.

6. IANA Considerations

No IANA Allocation is required in this document.

7. Acknowledgements

TBD.

8. References

8.1. Normative References

[I-D.ietf-bier-ipv6-requirements]
McBride, M., Xie, J., Dhanaraj, S., and R. Asati, "BIER IPv6 Requirements", draft-ietf-bier-ipv6-requirements-03 (work in progress), November 2019.

[I-D.ietf-bier-non-mpls-bift-encoding]
Wijnands, I., Xu, X., and H. Bidgoli, "An Optional Encoding of the BIFT-id Field in the non-MPLS BIER Encapsulation", draft-ietf-bier-non-mpls-bift-encoding-02 (work in progress), August 2019.

[I-D.xie-bier-ipv6-encapsulation]
Xie, J., Geng, L., McBride, M., Asati, R., and S. Dhanaraj, "Encapsulation for BIER in Non-MPLS IPv6 Networks", draft-xie-bier-ipv6-encapsulation-04 (work in progress), December 2019.

[I-D.xie-bier-ipv6-mvpn]
Xie, J., McBride, M., Dhanaraj, S., and L. Geng, "Use of BIER IPv6 Encapsulation (BIERv6) for Multicast VPN in IPv6 networks", draft-xie-bier-ipv6-mvpn-01 (work in progress), July 2019.

[RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", RFC 5512, DOI 10.17487/RFC5512, April 2009, <https://www.rfc-editor.org/info/rfc5512>.

[RFC6770] Bertrand, G., Ed., Stephan, E., Burbridge, T., Eardley, P., Ma, K., and G. Watson, "Use Cases for Content Delivery Network Interconnection", RFC 6770, DOI 10.17487/RFC6770, November 2012, <https://www.rfc-editor.org/info/rfc6770>.

[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>.

[RFC8313] Tarapore, P., Ed., Sayko, R., Shepherd, G., Eckert, T., Ed., and R. Krishnan, "Use of Multicast across Inter-domain Peering Points", BCP 213, RFC 8313, DOI 10.17487/RFC8313, January 2018, <https://www.rfc-editor.org/info/rfc8313>.

8.2. 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

Liang Geng
China Mobile
Beijing 10053

Email: gengliang@chinamobile.com

Jingrong Xie
Huawei Technologies

Email: xiejingrong@huawei.com

Mike McBride
Futurewei

Email: mmcbride7@gmail.com

Gang Yan
Huawei Technologies

Email: yangang@huawei.com