

Network Working Group
Internet-Draft
Updates: 6514 (if approved)
Intended status: Standards Track
Expires: 24 May 2024

F. Duan
J. Xie
S. Chen
Huawei Technologies
21 November 2023

BGP MVPN in IPv6 Infrastructure Networks: Problems and Solution
Approaches
draft-duan-bess-mvpn-ipv6-infras-05

Abstract

MVPN deployment faces some problems while used in provider's IPv6 infrastructure networks. This document describes these problems, and corresponding solutions.

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 24 May 2024.

Copyright Notice

Copyright (c) 2023 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. Terminology	2
3. Problems and Solutions	2
3.1. Problems	2
3.2. Modification of C-Multicast Route NLRI	4
3.3. Route Reflection Control	6
4. Security Considerations	6
5. IANA Considerations	6
6. Acknowledgements	6
7. Normative References	6
Authors' Addresses	7

1. Introduction

BGP MVPN procedure is defined in [RFC6514]. As a mature MVPN technology, it has been accepted by most operators and vendors. In [RFC6515], BGP MVPN is updated for IPv6 infrastructure networks. However, the deployment of BGP MVPN in IPv6 network still faces some problems. This document describes these problems and corresponding solutions.

2. Terminology

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

3. Problems and Solutions

3.1. Problems

In [RFC6514] and [RFC6515], the following issues are critical for IPv6 infrastructure scenarios while a non-segmented inter-AS P-tunnel is being established between root PE and its leaf PEs, of which the reason is that the Source AS field (4 octets) of C-multicast route cannot hold a IPv6 address (16 octets).

1. In order to distinguish the C-multicast routes for a specific multicast c-flow (C-S, C-G) sent to different root PEs, section 11.1.3 of [RFC6514] proposed that "To support non-segmented inter-AS tunnels, the Source AS field in the C-multicast route is set to value of the Originating Router's IP Address field of the found Intra-AS I-PMSI A-D route". However, in NLRI of C-multicast route, Source AS field is 4 octets in length, while the Originating Router's IP Address field of Intra-AS I-PMSI A-D route is 16 octets length in provider's IPv6 networks. The 4-octet Source AS field cannot hold the Originating Router's IP Address in IPv6 network.
2. In order to control the propagation of C-multicast routes between different ASes, section 11.2 of [RFC6514] proposed that "Instead of matching the RD and Source AS carried in the C-multicast route against the RD and Source AS of an Inter-AS I-PMSI A-D route, the ASBR should match it against the RD and the Originating Router's IP Address of the Intra-AS I-PMSI A-D routes". However, Source AS field in NLRI of C-multicast route cannot be translated to the Originating Router's IP Address of the Intra-AS I-PMSI A-D routes in provider's IPv6 networks, because of the mismatch of their field length.

In the process of evolution to IPv6, IPv4 and IPv6 infrastructure addresses will coexist in the provider's network. The following figure is an example of BGP MVPN evolution to IPv6.

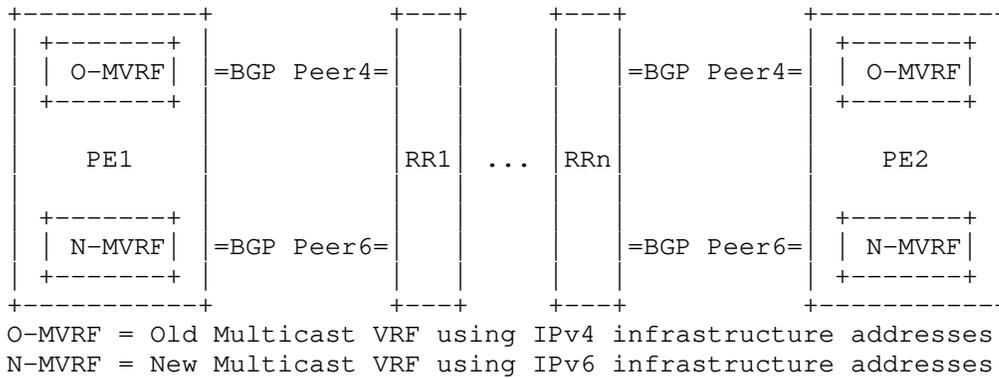


Figure 1: BGP MVPN Evolution to IPv6 Infrastructure

During the evolution process, IPv4 and IPv6 parallel BGP sessions are established between Provider Edge routers and route reflectors (RR). If the BGP MVPN routes are sent to all IPv4 and IPv6 BGP peers without any control, the number of the PATHs of these routes will be doubled with each reflection while BGP ADD-PATH [RFC7911] procedure is enabled on route reflectors.

3.2. Modification of C-Multicast Route NLRI

The solution to distinguish the C-multicast routes sent to different root PE is related to the way to distinguish UMH routes for a specific multicast source (C-S) sent from different root PE, which the later is not a problem of IPv6 infrastructure specific. In [RFC6514], it recommended that the RDs of root PE of a same MVPN were configured distinctly to perform selective forwarding selection, which was broken by GTM procedures defined in [RFC7716] because the UMH routes sent from different root PE through BGP SAFI 1 or SAFI 2 lack RD informations. There are also some MVPN deployment cases that the RDs of root PE may be configured with a same value for provisioning reasons. According to above description, whether the RDs of PE of a MVPN are same or not are two different deployment cases. This document addresses the C-multicast routes distinguishing issue for both cases. How to distinguish UMH routes in the cases of root PE with same RD is out of the scope of this document, because it is not IPv6 infrastructure specific.

To support non-segmented inter-AS tunnels in IPv6 infrastructure network, the C-multicast route NLRI is redefined as follow:

```

+-----+
|      RD      (8 octets)      |
+-----+
| Root Distinguisher (4 octets) |
+-----+
| Multicast Source Length (1 octet) |
+-----+
| Multicast Source (variable)      |
+-----+
| Multicast Group Length (1 octet) |
+-----+
| Multicast Group (variable)      |
+-----+

```

In the above figure, the Root Distinguisher field replaces the Source As field defined in [RFC6514]. When constructing a C-multicast route, leaf PE follows the following specification:

1. For the cases of IPv4 infrastructure or Intra-AS P-tunnel establishment in IPv6 infrastructure, the Root Distinguisher field MUST be treated as Source AS field and section 11.1.3 of [RFC6514] MUST be fully followed.
2. For non-segmented Inter-AS P-tunnel establishment in IPv6 infrastructure scenarios, if the RDs of ingress PEs are distinct (which can be detected from UMH routes), the Root Distinguisher field MUST be filled with the number of ingress AS.
3. For non-segmented Inter-AS P-tunnel establishment in IPv6 infrastructure scenarios, if the RDs of ingress PEs are same, a 4-octet distinct value MUST be assigned by leaf PE for each root PE. For example, each leaf PE uses a same well-known / configured hash algorithm to transform the IPv6 root IP to 4-octet distinct value for each ingress PE, or a provisioning method is used to globally assign different 4-octet IDs for each ingress PE. The Root Distinguisher field in C-multicast NLRI is filled with this value and a distinct C-multicast route will be sent to individual upstream root PE.

The solution to control the propagation of C-multicast route between different ASes is to use the IPv6 address included in IPv6 VRF Route Import Extended Community instead of Source AS field of C-multicast NLRI while locating Intra-AS AD route of the corresponding root PE the C-multicast sent to on ASBRs. This document recommends that the Local Administrator field of IPv6 VRF Route Import Extended Community is set to a non-zero value by root PEs even in GTM scenarios, of which the value is local assigned distinctly by root PE for both each MVPN and GTM instance. Accordingly, the IPv6 root address of a C-multicast route can be extracted from the only IPv6 VRF Route Import Extended Community carrying a non-zero Local Administrator field.

When receiving a C-multicast route from E-BGP neighbors, the ASBR checks whether an IPv6 VRF Route Import Extended Community with a non-zero Local Administrator field is included in this route and takes following actions:

1. If the IPv6 VRF Route Import Extended Community does not exist in the C-multicast route, the ASBR will treat the Root Distinguisher field as Source AS field and follows the description in section 11.2 of [RFC6514].
2. If the IPv6 VRF Route Import Extended Community does exist in the C-multicast route, the ASBR will match the IPv6 address carried in this extended community and the RD in C-multicast route NLRI against the Originating Router's IP Address and the RD of the

Intra-AS I-PMSI A-D routes. If the corresponding Intra-AS I-PMSI A-D route exists, the ASBR will propagate the C-multicast route in its local AS.

3.3. Route Reflection Control

To reduce BGP MVPN routes in Parallel IPv4 and IPv6 BGP sessions scenario, the following actions should be taken by sender PEs:

1. For Intra-AS I-PMSI A-D Route, S-PMSI A-D Route and Leaf A-D Route, if the Originating Router's IP Address field in the route is filled with an IPv6 address, it will be sent to the IPv6 BGP neighbors; otherwise, it will be sent to the IPv4 BGP neighbors.
2. For Inter-AS I-PMSI A-D Route and Source Active A-D Route, it is sent to both IPv6 BGP neighbors and IPv4 BGP neighbors.
3. For C-multicast route, if the IPv6 VRF Route Import Extended Community exists in the route, it will be sent to the IPv6 BGP neighbors; otherwise, it will be sent to the IPv4 BGP neighbors.

In the route reflectors, the part of routes which are received from IPv6 BGP neighbors will be reflected to other IPv6 BGP neighbors and the other part of routes which are received from IPv4 BGP neighbors will be reflected to other IPv4 BGP neighbors.

4. Security Considerations

This document introduces no new security considerations beyond those already specified in [RFC6514] and [RFC6515].

5. IANA Considerations

This document contains no actions for IANA.

6. Acknowledgements

Your name here

7. 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>>.

- [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>>.
- [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>>.
- [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>>.
- [RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", RFC 7911, DOI 10.17487/RFC7911, July 2016, <<https://www.rfc-editor.org/info/rfc7911>>.
- [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

Fanghong Duan
Huawei Technologies
Email: duanfanghong@huawei.com

Jingrong Xie
Huawei Technologies
Email: xiejingrong@huawei.com

Siyu Chen
Huawei Technologies
Email: chensiyu27@huawei.com