BESS
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2020

Z. Zhang
W. Lin
Juniper Networks
J. Rabadan
Nokia

July 8, 2019

Multicast in L3VPNs Signaled by EVPN SAFI
draft-zzhang-bess-mcast-in-evpn-signaled-l3vpn-00

Abstract

[ietf-bess-evpn-prefix-advertisement] specifies an EVPN SAFI Type-5 route that can be used to signal L3VPNs. This document specifies procedures for multicast in such an L3VPN.

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 January 9, 2020.

Copyright Notice

Table of Contents

1. Terminology

It is expected that audience is familiar with EVPN and MVPN concepts and terminologies. For convenience, the following terms are briefly explained.

- PMSI: P-Multicast Service Interface - a conceptual interface for a PE to send customer multicast traffic to all or some PEs in the same VPN.

- I-PMSI: Inclusive PMSI - to all PEs in the same VPN.

- S-PMSI: Selective PMSI - to some of the PEs in the same VPN.

- Leaf A-D routes: For explicit leaf tracking purpose. Triggered by S-PMSI A-D routes and targeted at triggering route's originator.

- IMET A-D route: Inclusive Multicast Ethernet Tag A-D route. The EVPN equivalent of MVPN Intra-AS I-PMSI A-D route.

- SMET A-D route: Selective Multicast Ethernet Tag A-D route. The EVPN equivalent of MVPN Leaf A-D route but unsolicited and untargeted.

2. Introduction

Traditionally, an L3VPN is signaled with BGP "MPLS-labeled VPN address" SAFI and uses MPLS as provider tunnel as specified in [RFC4364>]. Multicast support in such an L3VPN is specified in [RFC6513] and [RFC6514].

[ietf-bess-evpn-prefix-advertisement] specifies another way of signaling L3VPN via EVPN SAFI Type-5 routes for two reasons:

- VXLAN tunnels can be used, either for deployment scenarios where MPLS is not desired or for the purpose of better ECMP hashing.

- In an environment where EVPN is already needed for L2VPN, an operator may prefer just using an additional EVPN route type to signal L3VPN routes, instead of using another SAFI. This is especially the case when L3VPN is used to provide inter-DC connection.

[ietf-bess-evpn-prefix-advertisement] does not define procedures for multicast. This document provides three options for different deployment scenarios.

2.1. Optimized Inter-Subnet Multicast for EVPN

If all multicast senders and receivers are in an EVPN domain (including both intra-DC and inter-DC cases), the Optimized Inter-Subnet Multicast (OISM) procedures defined in [ietf-bess-evpn-irb-mcast] is the best and preferred option. The advantages are that no new procedures are needed and Any Source Multicast (ASM) does not need PIM Rendezvous Point (RP) procedures.

This does require that, if not all BDs are presented on every PE, then a Supplemental Bridge Domain (SBD) needs to be configured on every PE. Since the "Interface-less IP-VRF-to-IP-VRF Model" defined in Section 4.4.1 of [ietf-bess-evpn-prefix-advertisement] does not use SBD, for multicast purpose it is better to move away from that model.

Additionally, in case of inter-DC, the SBD needs be stretched across DCs even if regular BDs are not stretched. If the number of PEs in all DCs becomes very large, segmentation procedures defined in [ietf-bess-evpn-bum-procedures-update] and further enhanced in [zzhang-bess-mvpn-evpn-cmcast-enhancements] can be used. Alternatively, MVPN procedures defined in [RFC6514] can be used/adapted for an L3VPN signaled by EVPN Type-5 routes, as described in the following two sections.

2.2. Using [RFC6514] Procedures

If the OISM procedure cannot be used for any of the following situations that use L3VPN signaled by EVPN Type-5 routes:

- There are senders/receivers not on a BD of an EVPN domain and OISM cannot extend to connect them.
- Stretching SBD across a DCI is not desired as described in the previous section.
- It's a pure L3VPN scenario, where EVPN does not add any value.

MVPN procedures defined in [RFC6514] can be used as is as long as:

- The MVPN procedures treat EVPN Type-5 routes the same as routes signaled with "MPLS-labeled VPN address" when it comes to UMH selection.
- The EVPN Type-5 routes to C-RP or C-src carry the VRF Route Import Extended Community and Source AS Extended Extended Community.

In other words, the only difference is that the routes used for UMH selecion now includes those signaled via EVPN Type-5 routes, and they MUST carry the two ECs mentioned above. The rest of [RFC6514] procedures are unchanged.

The EVPN Type-2 signaled IP routes may be used as well, though from MVPN point of view, they're no different from "local" routes associated with IRB interfaces.

2.3. Adapted [RFC6514] Procedures

Notice that, an operator may have chosen to use EVPN Type-5 routes to signal L3VPN because they wanted to avoid signaling another BGP SAFI. Using [RFC6514] procedures as described in the previous section defeats that purpose because a new MCAST-VPN SAFI has to be used.

That can be resolved by adapting the [RFC6514] procedures with EVPN SAFI, as described below.

RFC6514 uses 7 route types and only the Source Active route does not already have a corresponding EVPN route type:

MVPN		EVPN	
Type	Name	Type	Name
1	Intra-AS I-PMSI	3	IMET
2	Inter-AS I-PMSI	9	Per-Region I-PMSI
3	S-PMSI	10	S-PMSI
4	Leaf	11	Leaf
5	Source Active	TBD	Source Active (added in this spec)
6	(*,G) C-Multicast	6	SMET
7	(S,G) C-Multicast	7	SMET

As pointed out in [zzhang-bess-mvpn-evpn-cmcast-enhancements], the MVPN Type-6/7 C-multicast routes don't have leaf tracking semantics while EVPN SMET route has built-in leaf tracking semantics. Both have pros and cons depending on the situation. This document will specify when SMET routes used for MVPN do or do not need leaf tracking semantics and the corresponding procedures.

Also as pointed out in [zzhang-bess-mvpn-evpn-cmcast-enhancements], the MVPN Type-6/7 C-multicast routes are targeted while EVPN SMET routes are not. This document specifies that the EVPN SMET routes used for MVPN purpose will be targeted, except in a special case as mentioned in [zzhang-bess-mvpn-evpn-cmcast-enhancements].

With this, the MEG (MVPN/EVPN Gateway) [ietf-bess-evpn-irb-mcast] follows the adaped MVPN procedures as specified in this document instead of the [RFC6514] procedures on MVPN side.

Detailed procedures are specified in the following section.

3. Specifications

Details to be added.

4. Security Considerations

This document does not introduce new security risks. Whatever security aspects that are applicable to [RFC7432], [RFC6513], [RFC6514] and [ietf-bess-evpn-prefix-advertisement] apply here.

5. References

5.1. Normative References

[I-D.ietf-bess-evpn-irb-mcast]
Lin, W., Zhang, Z., Drake, J., Rosen, E., Rabadan, J., and A. Sajassi, "EVPN Optimized Inter-Subnet Multicast (OISM) Forwarding", draft-ietf-bess-evpn-irb-mcast-02 (work in progress), January 2019.

[I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in EVPN", draft-ietf-bess-evpn-prefix-advertisement-11 (work in progress), May 2018.

[I-D.zzhang-bess-mvpn-evpn-cmcast-enhancements]
Zhang, Z., Kebler, R., Lin, W., and E. Rosen, "MVPN/EVPN C-Multicast Routes Enhancements", draft-zzhang-bess-mvpn-evpn-cmcast-enhancements-01 (work in progress), March 2019.

[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>.

[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>.

5.2. Informative References

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[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>.

[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>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

Wen Lin
Juniper Networks

EMail: wlin@juniper.net

Jorge Rabadan
Nokia

EMail: jorge.rabadan@nokia.com