BESS
Internet-Draft
Updates: 6514 (if approved)
Intended status: Standards Track
Expires: November 25, 2021

Z. Zhang
L. Giuliano
Juniper Networks

May 24, 2021

MVPN and MSDP SA Interoperation
draft-ietf-bess-mvpn-msdp-sa-interoperation-08

Abstract

This document specifies the procedures for interoperation between Multicast Virtual Private Network (MVPN) Source Active routes and customer Multicast Source Discovery Protocol (MSDP) Source Active routes, which is useful for MVPN provider networks offering services to customers with an existing MSDP infrastructure. Without the procedures described in this document, VPN-specific MSDP sessions are required among the PEs that are customer MSDP peers. This document updates RFC6514.

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 November 25, 2021.

Table of Contents

1. Terminologies

Familiarity with MVPN [RFC6513] [RFC6514] and MSDP [RFC3618] protocols and procedures is assumed. Some terminologies are listed below for convenience.

- ASM: Any source multicast.
- SPT: Source-specific Shortest-path Tree.
- RPT: Rendezvous Point Tree.
- C-S: A multicast source address, identifying a multicast source located at a VPN customer site.
- C-G: A multicast group address used by a VPN customer.
- C-RP: A multicast Rendezvous Point for a VPN customer.

- C-Multicast: Multicast for a VPN customer.

- EC: Extended Community.

- GTM: Global Table Multicast, i.e., multicast in the default or global routing table vs. VRF table.

2. Introduction

Section "14. Supporting PIM-SM without Inter-Site Shared C-Trees" of [RFC6514] specifies the procedures for MVPN PEs to discover (C-S,C-G) via MVPN Source Active A-D routes and then send Source Tree Join (C-S,C-G) C-multicast routes towards the ingress PEs, to establish SPTs for customer ASM flows for which they have downstream receivers. (C-*,C-G) C-multicast routes are not sent among the PEs so inter-site shared C-Trees are not used and the method is generally referred to as "spt-only" mode.

With this mode, the MVPN Source Active routes are functionally similar to MSDP Source-Active messages. For a VPN, one or more of the PEs, say PE1, either acts as a C-RP and learns of (C-S,C-G) via PIM Register messages, or has MSDP sessions with some MSDP peers and learn (C-S,C-G) via MSDP SA messages. In either case, PE1 will then originate MVPN SA routes for other PEs to learn the (C-S,C-G).

[RFC6514] only specifies that a PE receiving the MVPN SA routes, say PE2, will advertise Source Tree Join (C-S,C-G) C-multicast routes if it has corresponding (C-*,C-G) state learnt from its CE. PE2 may also have MSDP sessions for the VPN with other C-RPs at its site, but [RFC6514] does not specify that PE2 advertises MSDP SA messages to those MSDP peers for the (C-S,C-G) that it learns via MVPN SA routes. PE2 would need to have an MSDP session with PE1 (that advertised the MVPN SA messages) to learn the sources via MSDP SA messages, for it to advertise the MSDP SA to its local peers. To make things worse, unless blocked by policy control, PE2 would in turn advertise MVPN SA routes because of those MSDP SA messages that it receives from PE1, which are redundant and unnecessary. Also notice that the PE1-PE2 MSDP session is VPN-specific (i.e., only for a single VPN), while the BGP sessions over which the MVPN routes are advertised are not.

If a PE does advertise MSDP SA messages based on received MVPN SA routes, the VPN-specific MSDP sessions with other PEs are no longer needed. Additionally, this MVPN/MSDP SA interoperation has the following inherent benefits for a BGP based solution.

- MSDP SA refreshes are replaced with BGP hard state.

- Route Reflectors can be used instead of having peer-to-peer sessions.

- VPN Extranet [RFC2764] mechanisms can be used to propagate (C-S,C-G) information across VPNs with flexible policy control.

While MSDP Source Active routes contain the source, group and RP addresses of a given multicast flow, MVPN Source Active routes only contain the source and group. MSDP requires the RP address information in order to perform MSDP peer-RPF. Therefore, this document describes how to convey the RP address information into the MVPN Source Active route using an Extended Community so this information can be shared with an existing MSDP infrastructure.

The procedures apply to Global Table Multicast (GTM) [RFC7716] as well.

2.1. MVPN RPT-SPT Mode

For comparison, another method of supporting customer ASM is generally referred to as "rpt-spt" mode. Section "13. Switching from a Shared C-Tree to a Source C-Tree" of [RFC6514] specifies the MVPN SA procedures for that mode, but those SA routes are a replacement for PIM-ASM assert and (s,g,rpt) prune mechanisms, not for source discovery purposes. MVPN/MSDP SA interoperation for the "rpt-spt" mode is outside the scope of this document. In the rest of the document, the "spt-only" mode is assumed.

3. Specification

The MVPN PEs that act as customer RPs or have one or more MSDP sessions in a VPN (or the global table in case of GTM) are treated as an MSDP mesh group for that VPN (or the global table). In the rest of the document, it is referred to as the PE mesh group. This PE mesh group MUST NOT include other MSDP speakers, and is integrated into the rest of MSDP infrastructure for the VPN (or the global table) following normal MSDP rules and practices.

When an MVPN PE advertises an MVPN SA route following procedures in [RFC6514] for the "spt-only" mode, it MUST attach an "MVPN SA RP-address Extended Community". This is a Transitive IPv4-Address-Specific Extended Community. The Local Administrative field is set to zero and the Global Administrative field is set to an RP address determined as the following:

- If the (C-S,C-G) is learnt as result of PIM Register mechanism, the local RP address for the C-G is used.

- If the (C-S,C-G) is learnt as result of incoming MSDP SA messages, the RP address in the selected MSDP SA message is used.

In addition to procedures in [RFC6514], an MVPN PE may be provisioned to generate MSDP SA messages from received MVPN SA routes, with or without local policy control. If a received MVPN SA route triggers an MSDP SA message, the MVPN SA route is treated as if a corresponding MSDP SA message was received from within the PE mesh group and normal MSDP procedure is followed (e.g. an MSDP SA message is advertised to other MSDP peers outside the PE mesh group). The (S,G) information comes from the (C-S,C-G) encoding in the MVPN SA NLRI and the RP address comes from the "MVPN SA RP-address EC" mentioned above. If the received MVPN SA route does not have the EC (this could be from a legacy PE that does not have the capability to attach the EC), the local RP address for the C-G is used. In that case, it is possible that the RP inserted into the MSDP SA message for the C-G is actually the MSDP peer to which the generated MSDP message is advertised, causing the peer to discard it due to RPF failure. To get around that problem the peer SHOULD use local policy to accept the MSDP SA message.

An MVPN PE MAY treat only the best MVPN SA route selected by the BGP route selection process (instead of all MVPN SA routes) for a given (C-S,C-G) as a received MSDP SA message (and advertise the corresponding MSDP message). In that case, if the selected best MVPN SA route does not have the "MVPN SA RP-address EC" but another route for the same (C-S, C-G) does, then the next best route with the EC SHOULD be chosen. As a result, when/if the best MVPN SA route with the EC changes, a new MSDP SA message is advertised if the RP address determined according to the newly selected MVPN SA route is different from before. The MSDP SA state associated with the previously advertised MSDP SA message with the older RP address will be timed out.

4. Security Considerations

RFC6514 specifies the procedure for a PE to generate an MVPN SA upon discovering a (C-S,C-G) flow (e.g. via a received MSDP SA message) in a VPN. This document extends this capability in the reverse direction - upon receiving an MVPN SA route in a VPN generate a corresponding MSDP SA and advertise it to MSDP peers in the same VPN. As such, the capabilities specified in this document introduce no additional security considerations beyond those already specified in RFC6514 and RFC3618. Moreover, the capabilities specified in this document actually eliminate the control message amplification that exists today where VPN-specific MSDP sessions are required among the PEs that are customer MSDP peers, which lead to redundant messages (MSDP SAs and MVPN SAs) being carried in parallel between PEs.

5. IANA Considerations

This document introduces a new Transitive IPv4 Address Specific Extended Community "MVPN SA RP-address Extended Community". IANA has registered subcode 0x20 in the Transitive IPv4-Address-Specific Extended Community Sub-Types registry for this EC.

6. Acknowledgements

The authors thank Eric Rosen and Vinod Kumar for their review, comments, questions and suggestions for this document. The authors also thank Yajun Liu for her review and comments.

7. References

7.1. 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>.

[RFC3618] Fenner, B., Ed. and D. Meyer, Ed., "Multicast Source Discovery Protocol (MSDP)", RFC 3618, DOI 10.17487/RFC3618, October 2003, <https://www.rfc-editor.org/info/rfc3618>.

[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>.

7.2. Informative References

[RFC2764] Gleeson, B., Lin, A., Heinanen, J., Armitage, G., and A. Malis, "A Framework for IP Based Virtual Private Networks", RFC 2764, DOI 10.17487/RFC2764, February 2000, <https://www.rfc-editor.org/info/rfc2764>.

[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>.

[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>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

Lenny Giuliano
Juniper Networks

EMail: lenny@juniper.net