Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 20, 2013

Zhang
Giuliano
Juniper Networks
Pacella
Verizon

February 16, 2013

Global Table Multicast with BGP-MVPN Procedures
draft-zzhang-mboned-mvpn-global-table-mcast-00.txt

Abstract

This document describes a way to implement Global Table Multicast, aka Internet Multicast, using BGP encodings and procedures for MVPN as specified in [RFC6514].

No protocol modification/extension is required. This is purely for informational and clarifying purposes only.

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 http://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 August 20, 2013.

Table of Contents

1. Introduction

[RFC6513] and [RFC6514] specify procedures and encodings to implement Multicast for L3VPNs (MVPN). [RFC6513] specifies general concepts and procedures that apply to PIM-based and/or BGP-based C-Multicast State Signaling (referred to PIM-MVPN and BGP-MVPN respectively), and [RFC6514] specifies BGP procedures and encodings used by both PIM-MVPN and BGP-MVPN.

While [RFC6513] and [RFC6514] assume the context of VPN, they can be used to implement Global (vs. VRF) Table Multicast as well, without any protocol modification/extension, even though the RFCs do not explicitly mention it.

Consider a provider network where the "core" part of it uses MPLS P2MP LSPs or Ingress Replication over either P2P LSPs (with RSVP-TE) or MP2P LSPs (with LDP) instead of PIM to carry multicast traffic among the border routers (BRs) of the core. Those BRs run PIM on interfaces connected to other routers outside the core.

This document describes how Global Table Multicast can be implemented using BGP-MVPN procedures. We start with a simple reference scenario below, and also discuss one slightly different scenario and another special one.

With Global Table Multicast implemented by BGP-MVPN procedures, all the features/characteristics of BGP-MVPN apply, including scaling, aggregation, flexible choice of provider tunnels, support for PIM-DM/ASM/SSM/Bidir as PE-CE multicast protocol, BSR and AUTO-RP as RP-to-group mapping protocols, etc.

2. 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].

3. Operation

In the simplest reference scenario, the BRs advertise to each other RPF routes to multicast sources via iBGP (with or without RRs in the middle) with Next Hop set to themselves. The routes could have been learned from other non-BRs via eBGP or IGP.

Conceptually and functionally, those BRs are just like MVPN PEs: connections to other routers outside the core can be treated as PE-CE interfaces and MVPN procedures can run among the PEs (i.e., BRs) for Discovery, Tunnel Binding, and Multicast State Signaling.

With that, using BGP-MVPN procedures for Global Table Multicast is straightforward and requires almost no further clarification. However, some popular practices are described below.

By default, RD 0:0 is used when advertising A-D routes for Global Table Multicast, though an implementation MAY support the configuration and use of a different RD value.

Similarly, when constructing the C-multicast Import RT as specified in Section 7 of [RFC6514], it is RECOMMENDED that the Local Administrator field is set to 0, though an implementation MAY use any value that can uniquely associate it to the global routing table (vs. a VRF).

3.1. IBGP session between BRs and non-BRs

In the simple reference scenario above, it is assumed that the BRs learn RPF routes from non-BRs via eBGP or IGP. The assumption is to illustrate the analogy to a true VPN environment. In another deployment scenario, the BRs could have learned the RPF routes over those iBGP sessions to non-BRs. If the BRs act as RRs and reflect the RPF routes to other BRs with polices to attach VRF Route Import Extended Community and Source AS Extended Community, BGP-MVPN procedures can still be used as described earlier. Even if the BRs do not act as RRs, the scenario could be considered analogous to what [RFC6368] describes. As long as BRs re-advertise those RPF routes with the above mentioned communities, BGP-MVPN procedures can be used as described earlier. Note that they do not even need to use the push/pop procedures in [RFC6368] - the only requirement is for the BRs to re-advertise the routes learned over iBGP sessions from non-BRs to other BRs over iBGP sessions.

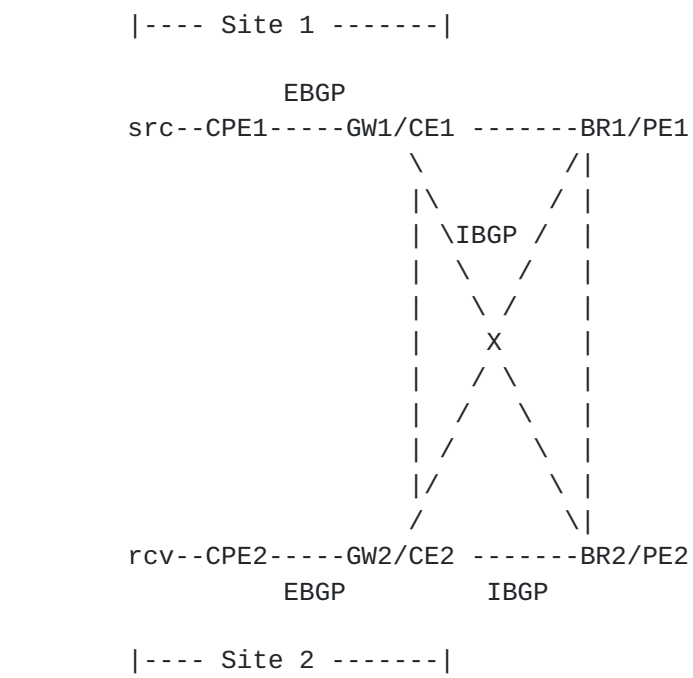
3.2. Non-BGP RPF Routes or BGP RPF routes not originated by the BRs

With true MVPN, the PEs advertise the RPF routes themselves as VPN-IP routes, and attach a VRF Route Import Extended Community that has the

C-multicast Import RT value for the VRF associated with the routes. The VRF Route Import Extended Community is extracted by egress PEs and attached to their C-Multicast Routes as Route Target Extended Community to control the distribution to and importation by relevant ingress PEs.

With Global Table Multicast, in both the simple reference scenario and the above mentioned variance, the BRs do (re-)advertise the RPF routes as required for BGP-MVPN. However, in other situations, it is possible that the RPF routes are not advertised by the BRs via BGP at all, hence they may not carry the VRF Route Import Extended Community.

Consider the following example:



There is a full-mesh of IBGP sessions among provider routers GW1/BR1/ BR2/GW2. EBGP sessions run between CPE1/GW1 and between CPE2/GW2. Border routers BR1/BR2 run BGP-MVPN procedures for Global Table Multicast. GW1 learns of BGP route to the src from CPE1 and advertises it to BRx/GW2.

Because GW1 does not run MVPN, BR2's route to the src (learned from GW1 instead of BR1) does not have the VRF Route Import Extended Community. Therefore, it would not be able to correctly attach a Route Target Extended Community corresponding to BR1 in its C-Multicast Routes.

To handle that situation, BR2 performs the following recursively. Note that the route in the following procedure is either the RPF route for the source itself, or the route to the Next Hop of the BGP route in the previous recursion.

- If the route is a BGP route with a VRF Route Import Extended Community, that VRF Route Import Extended Community is used.

- If the route is a BGP route without a VRF Route Import Extended Community, get the route to the Next Hop and recurse.

- If the route is an IGP route with a RSVP-TE LSP as next hop, and the LSP endpoint is a BR that advertises an Intra-AS I-PMSI A-D route (BR1 in the above example), a VRF Route Import Extended Community is constructed as BR_addr:0 to be associated with the RPF route, where the BR_addr is the Originating Router's IP Address of the Intra-AS I-PMSI A-D route.

If the above procedure does not produce a usable VRF Route Import Extended Community, then the RPF route is considered a local route (vs. a remote route that is associated with a remote BR). Note that the special process is necessary only if the BRs (that run MVPN procedures) do not advertise the RPF routes via BGP and include VRF Route Import Extended Community in such routes.

Constructing the VRF Route Import as BR_addr:0 by an egress BR in the above special situation explains why it is RECOMMENDED that the Local Administrator is set to 0 when an ingress BR constructs its C-Multicast Import RT - the zero value is a special value agreed on apriori by all (vs. a local value that is normally picked by the ingress router and signaled via the VRF Route Import Extended Community).

4. Security Considerations

This document raises no new security issues. Security considerations for the base protocol are covered in [RFC6514].

5. IANA Considerations

This document has no IANA considerations.

This section should be removed by the RFC Editor prior to final publication.

6. Acknowledgements

The authors would like to thank Rahul Aggarwal and Yakov Rehkter for their comments and suggestions.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC6513] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", RFC 6513, February 2012.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, February 2012.

7.2. Informative References

[RFC6368] Marques, P., Raszuk, R., Patel, K., Kumaki, K., and T. Yamagata, "Internal BGP as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 6368, September 2011.

Authors' Addresses

Jeffrey Zhang
Juniper Networks
10 Technology Park Dr.
Westford, MA 01886
US

Email: zzhang@juniper.net

Lenny Giuliano
Juniper Networks
2251 Corporate Park Drive
Herndon, VA 20171
US

Email: lenny@juniper.net

Dante J. Pacella
Verizon Communications
22001 Loudoun County Parkway
Ashburn, VA 20147

Email: dante.j.pacella@verizonbusiness.com