Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2016

P. Mohapatra
Sproute Networks
A. Sreekantiah
K. Patel
B. Pithawala
Cisco Systems
A. Lo
Arista Networks

October 16, 2015

Automatic Route Target Filtering for legacy PEs
draft-ietf-idr-legacy-rtc-06

Abstract

This document describes a simple procedure that allows "legacy" BGP speakers to exchange route target membership information in BGP without using mechanisms specified in [RFC4684]. The intention of the proposed technique is to help in partial deployment scenarios and is not meant to replace [RFC4684].

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 April 18, 2016.

Copyright Notice

Copyright (c) 2015 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 (http://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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction 2
1.1. Requirements Language 3
2. Basic Idea 3
3. Detailed Operation 3
3.1. Legacy PE Behavior 3
3.2. RR Behavior 5
3.2.1. Generating Route Target Membership NLRIs for the legacy PE clients 6
4. ROUTE_FILTER Community 6
5. Deployment Considerations 7
6. Contributors 7
7. Acknowledgments 7
8. IANA Considerations 7
9. Security Considerations 7
10. References 7
10.1. Normative References 7
10.2. Informational References 8
Authors' Addresses 8

1. Introduction

[RFC4684] provides a powerful and general means for BGP speakers to exchange and propagate Route Target reachability information and constrain VPN route distribution to achieve high scale. However, it requires that all the BGP speakers in the network are upgraded to support this functionality. For example, in a network with route reflectors (RR), if one PE client in the cluster doesn't support constrained distribution, the cluster degenerates into storing and

processing all the VPN routes. The route reflectors need to request and store all the network routes since they do not receive route target membership information from the legacy PEs. The RR will also generate all those routes to the legacy PEs and the legacy PEs will end up filtering the routes and store the subset of VPN routes that are of interest.

This document specifies a mechanism for such legacy PE devices using existing configuration and toolset to provide similar benefits as [RFC4684]. At the same time, it is backward-compatible with the procedures defined in [RFC4684]. It also allows graceful upgrade of the legacy router to be [RFC4684] capable.

1.1. 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 RFC 2119 [RFC2119].

2. Basic Idea

The basic idea is to make use of VPN unicast route exchange from the legacy PEs to a new BGP speaker (e.g. an RR) to signal RT membership. The legacy PEs announce a set of "special" routes with mapped RTs to the RR along with a standard community (defined in this document). The presence of the community triggers the RR to extract the RTs and build RT membership information.

3. Detailed Operation

3.1. Legacy PE Behavior

The following simple steps are performed on the legacy PE device:

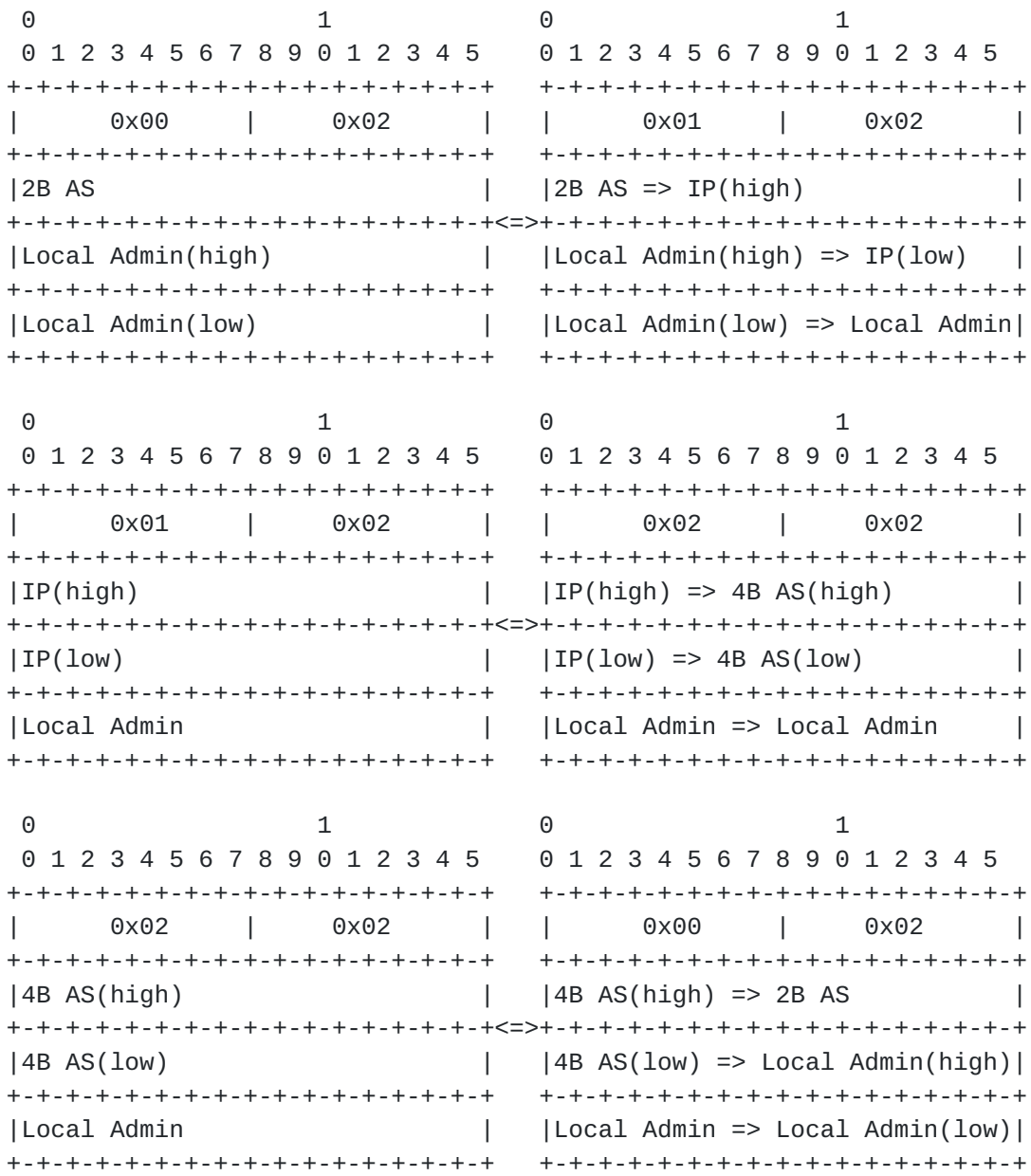
- Collect the "import route targets" of all the configured customer VRFs. Let's call this set 'IRTS'.

- Create a special "route-filter VRF" with a route distinguisher(RD) that's configured with the same value across the network for all legacy PE devices. Note: the equivalence of the RD value is for optimization - the operator may choose to use different values.

- Originate one or more routes in this VRF and attach a subset of 'IRTS' as "translated route-target extended communities" with each route so as to evenly distribute the RTs (and to make sure they can fit into one BGP UPDATE message). Collectively, the union of the "translated route-target extended communities" of all these routes is equal to the set 'IRTS'. The translated RTs are

attached as export route-targets for the routes originated in the route-filter VRF.

o The translation of the IRTs is necessary in order to refrain from importing "route-filter" VRF routes into VPN VRFs that would import the same route-targets. The translation of the IRTS is done as follows. For a given IRT, the equivalent translated RT (TRT) is constructed by means of swapping the value of the high-order octet of the Type field for the IRT (as defined in [RFC4360]).



As an example, if IRT R= 65500:12244(hex: 0x0002ffdc00002fd4), equivalent route-filter TRT: 255.220.0.0:12244(hex: 0x0102ffdc00002fd4). One shortcoming of the translation mechanism is a possible collision between IRTs and TRTs if the network has been configured with RTs of multiple higher order octet types (2-byte AS, IP address, and 4-byte AS). It is expected that such a configuration is rare in practice.

- o As an alternative to the translation of the IRTS, the subset of the 'IRTS' can be attached as-is (without swapping the type field as described earlier) as "export route-target extended communities" with each route so as to evenly distribute the RTs (and to make sure they can fit into one BGP UPDATE message). In this case, the IRT subsets can be attached in outbound policy to avoid the route-filter VRFs from being imported into VPN VRFs. Also in this case, the route-filter VRF routes must be tagged with a different special community (from that associated with the translated RTs) as described in Section 4 so that the receiving BGP speaker can distinguish the two cases.

- o The routes are marked with NO_ADVERTISE and NO_EXPORT well-known communities as well as the appropriate new community that's defined in this document Section 4. Note that there is no specific provision made to disallow configuration of subsequent route policies that can potentially alter the set of communities attached to "route-filter" VRF routes. The protocol behavior in such a case is undefined and the use of those policy statements is discouraged.

3.2. RR Behavior

Upon receiving the "route-filter" routes, the BGP speaker does its usual processing to store them in its local RIB. It recognizes them as route-filter routes based on the association of the new standard community as defined in this document. If required (as indicated by the community value), it translates the attached route-target extended communities (TRT) to equivalent import route-targets (IRT). Finally it creates the route-target filter list for each legacy client by collecting the entire set of route targets. From this point onwards, the behavior is similar to that defined in [RFC4684]. The RR does not propagate the routes further because of their association with NO_ADVERTISE community. Also the VPN EoR that is sent by the legacy PE should also be used as an indication that the legacy PE is done sending the route-filter information as per the procedures defined in [RFC4684] for implementing a EoR mechanism to signal the completion of initial RT membership exchange.

3.2.1. Generating Route Target Membership NLRIs for the legacy PE clients

The RR MAY also translate the received extended communities from legacy clients into route target membership NLRIs as if it had received those NLRIs from the client itself. This is useful for further propagation of the NLRIs to rest of the network to create RT membership flooding graph. When the route_filter routes are received with same RD (from all legacy PE speakers), processing of the paths to generate equivalent NLRIs becomes fairly easy.

4. ROUTE_FILTER Community

This memo defines four BGP communities that are attached to BGP UPDATE messages at the legacy PE devices and processed by the route reflectors as defined above. They are as follows:

Community	Meaning
ROUTE_FILTER_v4	RTs are attached as-is for VPNv4 route filtering
...	...
ROUTE_FILTER_v6	RTs are attached as-is for VPNv6 route filtering
...	...
ROUTE_FILTER_TRANSLATED_v4	Translated RTs are attached for VPNv4 route filtering
...	...
ROUTE_FILTER_TRANSLATED_v6	Translated RTs are attached for VPNv6 route filtering

In the absence of (or lack of support of) AF specific communities (ROUTE_FILTER_v6, ROUTE_FILTER_TRANSLATED_v6), the ROUTE_FILTER_v4 or ROUTE_FILTER_TRANSLATED_v4 MAY be treated by an implementation as a default VPN route-filter community to build a combination VPN filter for all VPN AFs (VPNv4, VPNv6) present on the RR. This is in accordance with the procedures in [RFC4684] to build combination route-filters for VPN AFs and AF specific route-filters defined in [I-D.keyur-bgp-af-specific-rt-constrain]. If this is the case, then subsequent receipt of any "route-filter" routes with AF specific communities (ROUTE_FILTER_v6, ROUTE_FILTER_TRANSLATED_v6) will override the default filters sent with ROUTE_FILTER_v4 or ROUTE_FILTER_TRANSLATED_v4 for the VPNv6 AFI when support for the AF specific communities exists.

5. Deployment Considerations

When both the legacy PE and the RR support extended community based Outbound Route Filtering as in [I-D.chen-bgp-ext-community-orf] this may be used as a alternate solution for the legacy PE to signal RT membership information, in order to realize the same benefits as [RFC4684]. Also extended community ORF can be used amongst the RRs in lieu of [RFC4684] to realize similar benefits.

6. Contributors

Significant contributions were made by Stephane Litkowski, Luis M Tomotaki and James Uttaro which the authors would like to acknowledge.

7. Acknowledgments

The authors would like to thank Rob Shakir for his review and comments.

8. IANA Considerations

IANA shall assign new code points from BGP first-come first-serve communities for the four communities as listed in Section 4.

9. Security Considerations

There are no additional security risks introduced by this design.

10. References

10.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <http://www.rfc-editor.org/info/rfc4360>.

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

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, DOI 10.17487/RFC4684, November 2006, <http://www.rfc-editor.org/info/rfc4684>.

10.2. Informational References

[I-D.chen-bgp-ext-community-orf]
Chen, E., Lo, A., and K. Patel, "Extended Community Based Outbound Route Filter for BGP-4", draft-chen-bgp-ext-community-orf-02 (work in progress), December 2011.

[I-D.keyur-bgp-af-specific-rt-constrain]
Patel, K., Raszuk, R., Djernaes, M., Dong, J., and M. Chen, "IPv6 AF Extensions for Route Target Distribution", draft-keyur-bgp-af-specific-rt-constrain-01 (work in progress), March 2011.

Authors' Addresses

Pradosh Mohapatra
Sproute Networks

Email: mpradosh@yahoo.com

Arjun Sreekantiah
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: asreekan@cisco.com

Keyur Patel
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: keyupate@cisco.com

Burjiz Pithawala
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: bpithaw@cisco.com

Alton Lo
Arista Networks
5470 Great America Parkway
Santa Clara, CA 95054
USA

Email: altonlo@aristanetworks.com