Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 5, 2012

J. Uttaro
AT&T
P. Mohapatra
D. Smith
Cisco Systems
R. Raszuk
NTT MCL Inc.
J. Scudder
Juniper Networks
October 3, 2011

BGP ACCEPT_OWN Community Attribute
draft-ietf-l3vpn-acceptown-community-04.txt

Abstract

Under certain conditions it is desirable for a BGP route reflector to be able to modify the Route Target list of a VPN route that is distributed by the route reflector, enabling the route reflector to control how a route originated within one VRF is imported into other VRFs. This technique works effectively as long as the VRF that exports the route is not on the same PE as the VRF(s) that import the route. However, due to the constraints of the BGP protocol, it does not work if the two are on the same PE. This document describes a modification to the BGP protocol allowing this technique to work when the VRFs are on the same PE, allowing the technique to be used in a standard manner throughout an autonomous system.

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 5, 2012.

Copyright Notice

Table of Contents

1. Introduction

In certain scenarios, a BGP speaker may maintain multiple "VPN Routing and Forwarding tables", or VRFs [RFC4364]. Under certain conditions, it is desirable for a route reflector to be able to modify the Route Target (RT) list of a VPN route that is distributed by the route reflector, enabling the route reflector to control how a route originated within one VRF is imported into other VRFs. Though it is possible to perform such policy control directly on the originator, it may be operationally cumbersome in an autonomous system with a large number of border routers having complex BGP policies.

The technique of the route reflector modifying the RT list works effectively as long as the VRF that exports the route is not on the same PE as the VRF(s) that import the route. However, due to the constraints of the BGP protocol, it does not work if the two are on the same PE. This is because per the BGP specification [RFC4271], a BGP speaker rejects prefix advertisements received that were originated by itself. In an autonomous system with route reflectors, the route reflector attaches the ORIGINATOR_ID attribute to the UPDATE messages so that if such prefix advertisements reach the originator, the originator can reject them by simply checking the ORIGINATOR_ID attribute. The BGP specification also mandates that a route should not be accepted from a peer when the NEXT_HOP attribute matches the receiver's own "IP address".

This document proposes a modification to BGP's behavior by defining a new community [RFC1997] value, in order to allow the technique of RT list modification by the route reflector to be used in a standard manner throughout an autonomous system, irrespective of whether the VRFs are on the same, or different PEs.

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. ACCEPT_OWN Community

This memo defines a new BGP community, ACCEPT_OWN, whose value as assigned by IANA is 0xFFFF0001. Processing of the ACCEPT_OWN community SHOULD be controlled by configuration. The functionality SHOULD default to being disabled, as further specified in Section 2.3.

2.1. Route Acceptance

A router MAY accept a route whose ORIGINATOR_ID or NEXT_HOP value matches that of the receiving speaker if all of the following are true:

- Processing of the ACCEPT_OWN community is enabled by configuration.
- The route in question carries the ACCEPT_OWN community.
- The route in question was originated from a source VRF on the router (as determined by inspecting the Route Distinguisher).
- The route in question is targeted to one or more destination VRFs on the router (as determined by inspecting the Route Target(s)).
- At least one destination VRF is different from the source VRF.

A route MUST never be accepted back into its source VRF, even if it carries one or more Route Targets (RTs) which match that VRF.

2.2. Propagating ACCEPT_OWN Between Address Families

The ACCEPT_OWN community controls propagation of routes which can be associated with a source VRF by inspection of their Route Distinguisher and with a target VRF by inspection of their Route Target list (for example VPN routes with a SAFI of 128). As such, it SHOULD NOT be attached to any routes which cannot be associated with a source VRF. This implies that when propagating routes into a VRF, the ACCEPT_OWN community should not be propagated. Likewise, if a route carrying the ACCEPT_OWN community is received in an address family which does not allow the source VRF to be looked up, the ACCEPT_OWN community MUST be discarded. An OPTIONAL message may be logged in this case.

2.3. Configuration Control

ACCEPT_OWN handling SHOULD be controlled by configuration, and SHOULD default to being disabled. When ACCEPT_OWN is disabled by configuration (either explicitly or by default), the router MUST NOT apply the special route acceptance rules detailed in Section 2.1. The router SHOULD still apply the propagation rules detailed in Section 2.2.

3. Decision Process

If a BGP speaker supports ACCEPT_OWN and is configured for the extensions defined in this document, the following step is inserted after the LOCAL_PREF comparison step in BGP decision process:

> When comparing a pair of routes for a BGP destination, the route attached with ACCEPT_OWN community is preferred over the route that does not have the community.

In all other respects, the decision process remains unchanged. This extra step MUST only be invoked during the best path selection process of VPN-IP routes [RFC4364] (i.e. it MUST NOT be invoked for the best path selection of "imported" IP routes in a VRF). The purpose of the extra step is to allow the paths advertised by the route reflector with ACCEPT_OWN community to be selected as best over other paths that the BGP speaker may have received, hence enabling the applications ACCEPT_OWN is designed for.

4. Deployment Considerations

The ACCEPT_OWN community as described in this document is useful within a single autonomous system which uses a single layer of route reflectors. Its use with hierarchical route reflectors would require further specification and is out of scope for this document. Likewise, its use across multiple autonomous systems is out of scope for this document.

5. Other Applications

This approach may also be relevant to other scenarios where a BGP speaker maintains multiple routing contexts using an approach different from that of [RFC4364], as long as the specific approach in use has the property that the BGP speaker originates and receives routes within a particular context. In such a case, "VRF" in this document should be understood to mean whatever construct provides a routing context in the specific technology under consideration. Likewise, "Route Distinguisher" should be understood to mean whatever construct allows a route's originator to associate that route with its source context, and "Route Target" should be understood to mean whatever construct allows a route to be targeted for import into a context other than its source.

6. Security Considerations

ACCEPT_OWN as described above permits a router's own route prefix to be advertised to a different VRF on that router. In this respect, such a route is similar to any other BGP route and shares the same set of security vulnerabilities and concerns. No new fundamental security issues are introduced by ACCEPT_OWN.

7. IANA Considerations

IANA has assigned the value 0xFFFF0001 from BGP well-known communities registry for ACCEPT_OWN community. No additional IANA action is required.

8. Acknowledgments

The authors would like to thank Yakov Rekhter, Jim Guichard, Clarence Filsfils, John Mullooly, Jeff Haas, Pranav Mehta, and Tamas Mondal for their valuable comments and suggestions. The decision process changes were suggested by Pranav Mehta to solve the remote extranet problem.

9. Normative References

[RFC1997] Chandrasekeran, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, August 1996.

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

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

Appendix A. Local Extranet Application (non-informative)

One of the applications for this behavior is auto-configuration of extranets within MPLS VPN networks. Consider the following topology:

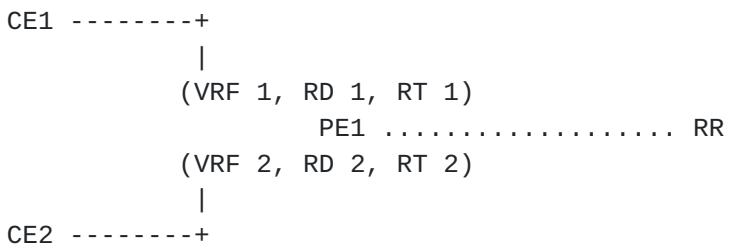

Figure 1: Extranet Application

Within the above topology, PE1 receives a prefix X from CE1. Prefix X is installed in VRF 1 and is advertised to the route reflector with route distinguisher (RD) 1 and route target (RT) 1 as configured on PE1. The requirement is to import prefix X into VRF 2 and advertise it to CE2 in support of extranet VPN connectivity between CE1/VRF1 and CE2/VRF2. Current BGP mechanisms for MPLS VPNs [RFC4364] require changing the import RT value and/or import policy for VRF 2 on PE1. This is operationally cumbersome in a network with a large number of border routers having complex BGP policies. Alternatively, using the new ACCEPT_OWN community value, the route reflector can simply re-advertise prefix X back to PE1 with RT 2 appended. In this way, PE1 will accept prefix X despite its ORIGINATOR_ID or NEXT_HOP value, import it into VRF 2 as a result of RT 2, and will then determine the correct adjacency rewrite within VRF 1 based on the RD value (1) and the prefix. Note that the RT 1 value originally attached to the route will simply be ignored since associated with the source VRF 1. The same operation needs also to happen in the reverse direction (VRF 1 learning a route from VRF 2) to achieve establishment of an extranet VPN strictly via the route reflector without changing the BGP policy of PE1 in any way. A router performing such an extranet application can accept a route with its own ORIGINATOR_ID or NEXT_HOP value only if the VRF in which the router originated the route is different than the VRF in which the router accepts the re-advertised route.

Authors' Addresses

James Uttaro
AT&T
200 S. Laurel Avenue
Middletown, NJ 07748
USA

Email: uttaro@att.com

Pradosh Mohapatra
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: pmohapat@cisco.com

David J. Smith
Cisco Systems
111 Wood Avenue South
Iselin, NJ 08830
USA

Email: djsmith@cisco.com

Robert Raszuk
NTT MCL Inc.
101 S Ellsworth Avenue Suite 350
San Mateo, CA 94401
US

Email: robert@raszuk.net

John Scudder
Juniper Networks
1194 N. Mathilda Ave
Sunnyvale, CA 94089
USA

Email: jgs@juniper.net