Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 12 January 2023

J. Dong
S. Zhuang
Huawei Technologies
G. Van de Velde
Nokia

11 July 2022

BGP Extended Community for Identifying the Target Nodes
draft-dong-idr-node-target-ext-comm-05

Abstract

BGP has been used to distribute different types of routing and policy information. In some cases, the information distributed may be only intended for one or a particular group of BGP nodes in the network. Currently BGP does not have a generic mechanism of designating the target nodes of the routing information. This document defines a new type of BGP Extended Community called "Node Target". The mechanism of using the Node Target Extended Community to steer BGP route distribution to particular BGP nodes is specified.

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 12 January 2023.

Copyright Notice

Table of Contents

1. Introduction

BGP [RFC4271] has been used to distribute different types of routing and policy information. In some cases, the information distributed may be only intended for one or a group of receiving BGP nodes in the network. One typical use case is the distribution of BGP Flow Spec [RFC8955] [RFC8956] rules only to a particular group of BGP nodes. Such a targeted distribution mechanism is considered useful as it can save the resources on nodes which do not need that information.

Currently BGP does not have a generic mechanism of designating the set of nodes to which the information is to be distributed. Route Target (RT) as defined in [RFC4364] was designed for the matching of VPN routes into the target VPN Routing and Forwarding tables (VRFs) on the PE nodes. [I-D.ietf-idr-segment-routing-te-policy] introduces the mechanism of steering the SR Policy information to the target head end node based on RT, it is only applicable to the SR Policy Address Family. Although it is possible to reuse RT to control the distribution of non-VPN information to one or a group of receiving nodes, such mechanism is not applicable when the information to be distributed is VPN-specific and is advertised with another set of RTs for the VRF matching, as the matching or any of the VPN RT in the BGP route would result in that route being imported to a local VRF, regardless of whether the receiving node is the target node or not. Thus a general mechanism which is independent from the control of VPN route to VRF import is needed.

Another possible approach is to configure, on each router, a community and the corresponding policies to match the community to determine whether to accept the received routes or not. Such mechanism relies on manual configuration thus is considered error-prone. It is preferable by some operators that an automatic approach can be provided, which would make the operation much easier.

This document defines a new type of BGP Extended Community called "Node Target". The mechanism of using the Node Target extended community to control the BGP route distribution only to particular BGP nodes is also specified.

1.1. 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.

2. Node Target Extended Communities

This section defines a new BGP Extended Community [RFC4360] called "Node Target Extended Community". It can be a transitive extended community with the high-order octet of the type set to 0x01, or a non-transitive extended community with the high-order octet type set to 0x41. The sub-type of the Node Target Extended Community is TBA.

The format of Node Target Extended Community is shown in Figure 1.

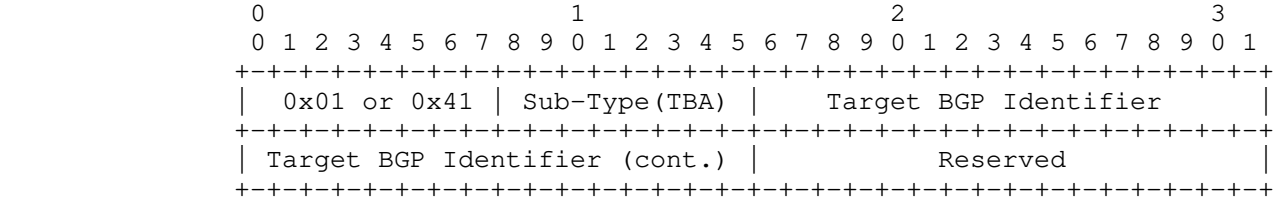

Figure 1. Node Target extended community

Where:

Target BGP Identifier (4 octets): The BGP Identifier of a target node. It is a 4-octet, unsigned, non-zero integer as defined in [RFC6286].

Reserved field (2 octets): Reserved for future use, MUST be set to zero on transmission and ignored on receipt.

One or more Node Target extended communities MAY be carried in an Update message to designate a group of target BGP nodes.

3. Procedures

In this section, the mechanism for intra-domain scenario is described, the mechanism for inter-domain scenario is for further study. The domain here refers to an administrative domain, which may consists of one or multiple ASes managed by a single operator.

When a network controller or BGP speaker plans to advertise some BGP routing or policy information only to one or a group of BGP nodes in the network, it MUST put the BGP Identifier of each target node into the Node Target extended communities, and attach the Node Target extended communities to the routes or policies to be advertised.

When a BGP speaker receives a BGP Update which contains one or more Node Target extended communities, it MUST check the target BGP Identifiers carried in the Node Target extended communities of the Update.

* If the target BGP Identifier in any of the Node Target extended community matches with the local BGP Identifier, this node is one of the target nodes of the Update, the information in the Update is eligible to be kept and installed on this node.

  - If this node is a Route Reflector, and in the Update there is one or more Node Target extended communities which contains non-local BGP Identifiers, information in the Update are eligible be reflected to its peers according to the rules defined in [RFC4456]. The default behavior for the RR in this case is to reflect the Update to all its peers without checking their BGP Identifiers. Depends on a configurable policy, the RR MAY further check the BGP Identifiers of its peers to determine the set of peers which are the target nodes of the Update, and only reflect the information in the Update to the matched BGP peers.

  - If this node is an Autonomous System Border Router (ASBR), and the BGP Identifiers of one or more of its EBGP peers match with the Node Target extended communities in the Update, information in the Update is eligible to be advertised to the matched EBGP peers.

* If the target BGP Identifier in any of the Node target extended community does not match with the local BGP Identifier, this node is not the target node of Update, the information in the Update is not eligible to be installed on this node.

- If this node is a Route Reflector, information in the Update is eligible to be reflected to its peers according to the rules defined in [RFC4456]. The default behavior for the RR in this case is to reflect the Update to all its peers without checking their BGP Identifiers. Depends on a configurable policy, the RR MAY check the BGP Identifiers of its peers to determine the set of peers which are the target nodes of the Update, and only reflect the information in the Update to the matched BGP peers.

4. Compatibility Considerations

The Node Target extended community introduced in this document can be deployed incrementally in the network. For BGP speakers which understand the Node Target extended community, it is used to determine whether the nodes are the target nodes of the Update. For BGP speakers which do not understand the Node Target extended community, it will be ignored and the information in the Update will be processed and advertised based on normal BGP procedure. Although this could ensure that the target nodes can always obtain the information needed, this may result in unnecessary state maintained on the legacy BGP nodes. If the information advertised is the Flow Spec rules, the legacy BGP speakers may install unnecessary Flowspec rules, this may have impact on traffic which matches such rules, thus may result in unexpected traffic steering or filtering behaviors on such nodes. This may be mitigated by setting appropriate routing policies on the legacy BGP nodes.

5. IANA Considerations

This document requests that IANA assigns one new sub-type for "Node Target Extended Community" from the "Transitive IPv4-Address-Specific Extended Community" registry of the "BGP Extended Communities" registry.

This document requests that IANA assigns the same sub-type for "Node Target Extended Community" from the "Non-Transitive IPv4-Address-Specific Extended Community" registry of the "BGP Extended Communities" registry.

6. Security Considerations

The mechanism defined in this document can limit the scope of the receiving nodes of BGP Updates, which make it possible for an attacker to do fine-grained targeting of malicious BGP Updates only to a restricted set of routers. This would potentially make it more difficult for a network administrator to discover an attack. This may be mitigated by filtering the Node Target Extended Communities at the administrative network boundaries.

7. Contributors

Haibo Wang
Email: rainsword.wang@huawei.com

8. Acknowledgements

The authors would like to thank Zhenbin Li, Ercin Torun, Jeff Haas, Robert Raszuk and John Scudder for the review and discussion of this document.

9. References

9.1. 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>.

[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, <https://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, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[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>.

9.2. Informative References

[I-D.ietf-idr-segment-routing-te-policy] Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-segment-routing-te-policy-18, 16 June 2022, <https://www.ietf.org/archive/id/draft-ietf-idr-segment-routing-te-policy-18.txt>.

[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>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <https://www.rfc-editor.org/info/rfc5575>.

[RFC6286] Chen, E. and J. Yuan, "Autonomous-System-Wide Unique BGP Identifier for BGP-4", RFC 6286, DOI 10.17487/RFC6286, June 2011, <https://www.rfc-editor.org/info/rfc6286>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

[RFC8956] Loibl, C., Ed., Raszuk, R., Ed., and S. Hares, Ed., "Dissemination of Flow Specification Rules for IPv6", RFC 8956, DOI 10.17487/RFC8956, December 2020, <https://www.rfc-editor.org/info/rfc8956>.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing
100095
China
Email: jie.dong@huawei.com

Shunwan Zhuang
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing
100095
China
Email: zhuangshunwan@huawei.com

Gunter Van de Velde
Nokia
Antwerp
Belgium
Email: gunter.van_de_velde@nokia.com