

Network Working Group
Internet-Draft
Updates: 1997 (if approved)
Intended status: Standards Track
Expires: December 15, 2019

J. Borkenhagen
AT&T
R. Bush
IIJ & Arrcus
R. Bonica
Juniper Networks
S. Bayraktar
Cisco Systems
June 13, 2019

Well-Known Community Policy Behavior
draft-ietf-grow-wkc-behavior-08

Abstract

Well-Known BGP Communities are manipulated differently across various current implementations; resulting in difficulties for operators. Network operators should deploy consistent community handling across their networks while taking the inconsistent behaviors from the various BGP implementations into consideration.. This document recommends specific actions to limit future inconsistency, namely BGP implementors must not create further inconsistencies from this point forward.

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 December 15, 2019.

Copyright Notice

Copyright (c) 2019 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 (<https://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.

Table of Contents

1. Introduction	2
2. Manipulation of Communities by Policy	3
3. Community Manipulation Policy Differences	3
4. Documentation of Vendor Implementations	3
4.1. Note on an Inconsistency	4
5. Note for Those Writing RFCs for New Community-Like Attributes	5
6. Action Items	5
7. Security Considerations	5
8. IANA Considerations	5
9. Acknowledgments	6
10. Normative References	6
Authors' Addresses	6

1. Introduction

The BGP Communities Attribute was specified in [RFC1997] which introduced the concept of Well-Known Communities. In hindsight, [RFC1997] did not prescribe as fully as it should have how Well-Known Communities may be manipulated by policies applied by operators. Currently, implementations differ in this regard, and these differences can result in inconsistent behaviors that operators find difficult to identify and resolve.

This document describes the current behavioral differences in order to assist operators in generating consistent community-manipulation policies in a multi-vendor environment, and to prevent the introduction of additional divergence in implementations.

This document recommends specific actions to limit future inconsistency, namely BGP implementors MUST NOT create further inconsistencies from this point forward.

2. Manipulation of Communities by Policy

[RFC1997] says:

"A BGP speaker receiving a route with the COMMUNITIES path attribute may modify this attribute according to the local policy."

One basic operational need is to add or remove one or more communities to the set. The focus of this document is another common operational need, to replace all communities with a new set. To simplify this second case, most BGP policy implementations provide syntax to "set" community that operators use to mean "remove any/all communities present on the route, and apply this set of communities instead."

Some operators prefer to write explicit policy to delete unwanted communities rather than using "set;" i.e. using a "delete community *:*" and then "add community x:y ..." configuration statements in an attempt to replace all communities. The same community manipulation policy differences described in the following section exist in both "set" and "delete community *:*" syntax. For simplicity, the remainder of this document refers only to the "set" behaviors, which we refer to collectively as each implementation's "'set" directive.'

3. Community Manipulation Policy Differences

Vendor implementations differ in the treatment of certain Well-Known communities when modified using the syntax to "set" the community. Some replace all communities including the Well-Known ones with the new set, while others replace all non-Well-Known Communities but do not modify any Well-Known Communities that are present.

These differences result in what would appear to be identical policy configurations having very different results on different platforms.

4. Documentation of Vendor Implementations

In this section we document the syntax and observed behavior of the "set" directive in several popular BGP implementations to illustrate the severity of the problem operators face.

In Juniper Networks' Junos OS, "community set" removes all communities, Well-Known or otherwise.

In Cisco IOS XR, "set community" removes all communities except for the following:

Numeric	Common Name
0:0	internet
65535:0	graceful-shutdown
65535:1	accept-own rfc7611
65535:65281	NO_EXPORT
65535:65282	NO_ADVERTISE
65535:65283	NO_EXPORT_SUBCONFED (or local-AS)

Communities not removed by Cisco IOS XR

Table 1

Cisco IOS XR does allow Well-Known communities to be removed only by explicitly enumerating one at a time, not in the aggregate; for example, "delete community accept-own". Operators are advised to consult Cisco IOS XR documentation and/or Cisco support for full details.

On Extreme networks' Brocade NetIron: "set community X" removes all communities and sets X.

In Huawei's VRP product, "community set" removes all communities, Well-Known or otherwise.

In OpenBGPD, "set community" does not remove any communities, Well-Known or otherwise.

Nokia's SR OS has several directives that operate on communities. Its "set" directive is called using the "replace" keyword, replacing all communities, Well-Known or otherwise, with the specified communities.

4.1. Note on an Inconsistency

The IANA publishes a list of Well-Known Communities [IANA-WKC].

Cisco IOS XR's set of Well-Known communities that "set community" will not overwrite diverges from the IANA's list of Well-Known communities. Quite a few Well-Known communities from IANA's list do not receive special treatment in Cisco IOS XR, and at least one community on Cisco IOS XR's special treatment list, internet == 0:0,

is not formally a Well-Known Community as it is not in [IANA-WKC]; but taken from the Reserved range [0x00000000-0x0000FFFF].

This merely notes an inconsistency. It is not a plea to 'protect' the entire IANA list from "set community."

5. Note for Those Writing RFCs for New Community-Like Attributes

When establishing new [RFC1997]-like attributes (large communities, wide communities, etc.), RFC authors should state explicitly how the new attribute is to be handled.

6. Action Items

Network operators are encouraged to limit their use of the "set" directive (within reason), to improve consistency across platforms.

Unfortunately, it would be operationally disruptive for vendors to change their current implementations.

Vendors MUST clearly document the behavior of "set" directive in their implementations.

Vendors MUST ensure that their implementations' "set" directive treatment of any specific community does not change if/when that community becomes a new Well-Known Community through future standardization. For most implementations, this means that the "set" directive MUST continue to remove the community; for those implementations where the "set" directive removes no communities, that behavior MUST continue.

Given the implementation inconsistencies described in this document, network operators are urged never to rely on any implicit understanding of a neighbor ASN's BGP community handling. I.e., before announcing prefixes with NO_EXPORT or any other community to a neighbor ASN, the operator should confirm with that neighbor how the community will be treated.

7. Security Considerations

Surprising defaults and/or undocumented behaviors are not good for security. This document attempts to remedy that.

8. IANA Considerations

The IANA is requested to list this document as an additional reference for the [IANA-WKC] registry.

9. Acknowledgments

The authors thank Martijn Schmidt, Qin Wu for the Huawei data point, Greg Hankins, Job Snijders, David Farmer, John Heasley, and Jakob Heitz.

10. Normative References

[IANA-WKC]

IANA, "Border Gateway Protocol (BGP) Well-Known Communities", <<https://www.iana.org/assignments/bgp-well-known-communities>>.

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <<http://www.rfc-editor.org/info/rfc1997>>.

[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>>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<http://www.rfc-editor.org/info/rfc8174>>.

Authors' Addresses

Jay Borkenhagen
AT&T
200 Laurel Avenue South
Middletown, NJ 07748
United States of America

Email: jayb@att.com

Randy Bush
IIJ & Arrcus
5147 Crystal Springs
Bainbridge Island, WA 98110
US

Email: randy@psg.com

Ron Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, VA 20171
US

Email: rbonica@juniper.net

Serpil Bayraktar
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
United States of America

Email: serpil@cisco.com