

IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 19, 2011

R. Raszuk
Cisco Systems
J. Haas
Juniper Networks
R. Steenbergen
nLayer Communications, Inc.
B. Decraene
France Telecom
P. Jakma
DCS, Uni. of Glasgow
October 16, 2010

Wide BGP Communities Attribute
draft-raszuk-wide-bgp-communities-01

Abstract

Communicating various routing policies via route tagging plays an important role in external BGP peering relations. It is also a very common best practice among operators to propagate various additional information about routes intra domain. The most common tool used today to attach various information about routes is realized with the use of BGP communities.

Such information is important for the BGP speakers to perform some mutually agreed actions without the need to maintain a separate offline database for each pair of prefix and an associated with it requested set of action entries.

This document defines a new encoding which will enhance and simplify what can be accomplished today with the use of BGP communities. The most important addition this specification brings over currently defined BGP communities is the ability to specify, carry as well as use for execution operator's defined set of parameters. Specification also provides an extensible platform for any new community encoding needs in the future.

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 19, 2011.

Copyright Notice

Copyright (c) 2010 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.

Table of Contents

1. Introduction	4
2. Wide BGP Community Attribute	4
3. Wide BGP Community Attribute Containers	5
3.1. Fixed size container template	6
3.2. Variable size container template	6
4. Container Type 1: Wide Community	7
4.1. Container Type 1 - TTL	7
4.2. Container Type 1 - Length	8
4.3. Container Type 1 - Community Value	8
4.4. Container Type 1 - Source AS number	8
4.5. Container Type 1 - Community Parameters	8
5. Well Known Standard BGP Communities	9
6. Operational considerations	9
7. Example	10
8. Security considerations	11
9. IANA Considerations	11
10. Contributors	12
11. Acknowledgments	13
12. References	13
12.1. Normative References	13
12.2. Informative References	13
Authors' Addresses	14

1. Introduction

RFC 1997 [RFC1997] defines a BGP Community Attribute to be used as a tool to contain in BGP update message various additional information about routes which may help to automate peering administration. As defined in RFC 1997 [RFC1997] BGP Communities Attribute consists of one or more sets of four octet values, where each one of them specifies a different community. Except two reserved ranges the encoding of community values mandates that first two octets are to contain the Autonomous System number followed by next two octets containing locally defined value.

With the introduction of 4-octet Autonomous System numbers by RFC 4893 [RFC4893] it became obvious that BGP Communities as specified in RFC 1997 will not be able to accommodate new AS encoding. In fact RFC 4893 explicitly recommends use of four octets AS specific extended communities as a way to encode new 4 octet AS numbers.

While encoding of 4 octet AS numbers are being addressed by [draft-ietf-idr-as4octet-extcomm-generic-subtype] neither the base BGP communities (both standard or extended) nor as4octet-extcomm-generic document define sufficient level of encoding freedom which could be of practical use. Authors believe that defining a new BGP Path Attribute which will provide ability to contain locally defined parameters will enhance current level of network policies as well as simplify BGP policy management. Proposed simple encoding will also enable to deliver a set of new network services without a need to define a new BGP extension each time.

While defining a new type of any tool there is always a unique opportunity to specify a subset of well recognized behaviors. List of the most commonly used today BGP communities as well as provision for a new registry for future definitions will be contained in a separate document.

2. Wide BGP Community Attribute

For the purposes of encoding for Wide BGP Communities a new BGP Path Attribute has been defined. The attribute type code is of the value (TBC by IANA).

Wide BGP Community Attribute is an optional, transitive BGP attribute, and may be present only once in the update message.

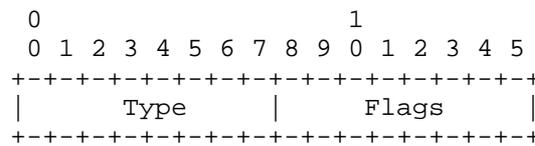
The attribute contains a number of typed containers, which are either fixed or variable in size. Any given container type may appear multiple times, unless that container type's definition says

otherwise.

3. Wide BGP Community Attribute Containers

Two container templates are defined for carrying BGP community information, to hold fixed or variably sized data. All container definitions MUST conform with one of these two templates.

Containers always start with the following header:



Container header

Flags are defined globally, to apply to all community container types.

- Bit 0: 0 => local community value
 - 1 => registered community value
- 1: 0 => do not decrement TTL field across confederation boundaries
 - 1 => decrement TTL across confederation boundaries
- 2...7: => ignored, preserve or set to zero.

Bit 0 set (value 1) indicates that the given container carries a Wide BGP Community which is registered with IANA. When not set (value 0) it indicates that community value which follows is locally assigned with a local meaning. Ignored bits SHOULD be preserved in any received containers, or set to 0 otherwise. Bit 1 is used to manage propagation scope of given community across confederation boundaries. When not set (value of 0) TTL field is not consider at the sub-AS boundaries. When set (value of 1) sub-AS border router follows the same procedure reg handling TTL field as applicable to ASBR at the domain boundary.

3.1. Fixed size container template

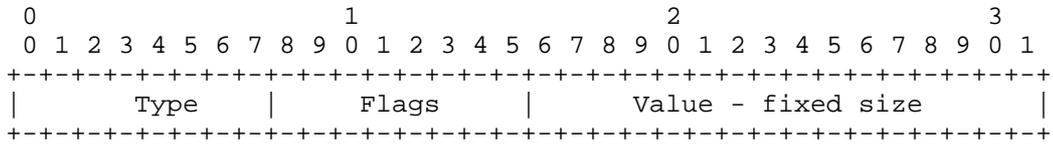


Figure 3: Fixed size type container

3.2. Variable size container template

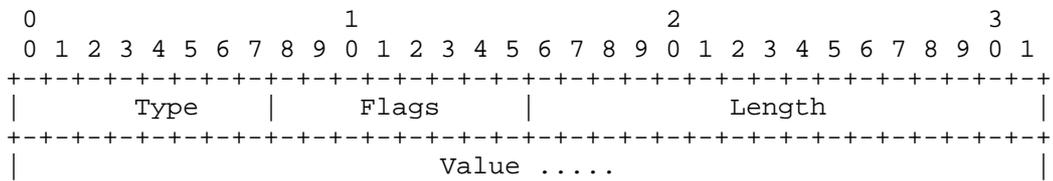
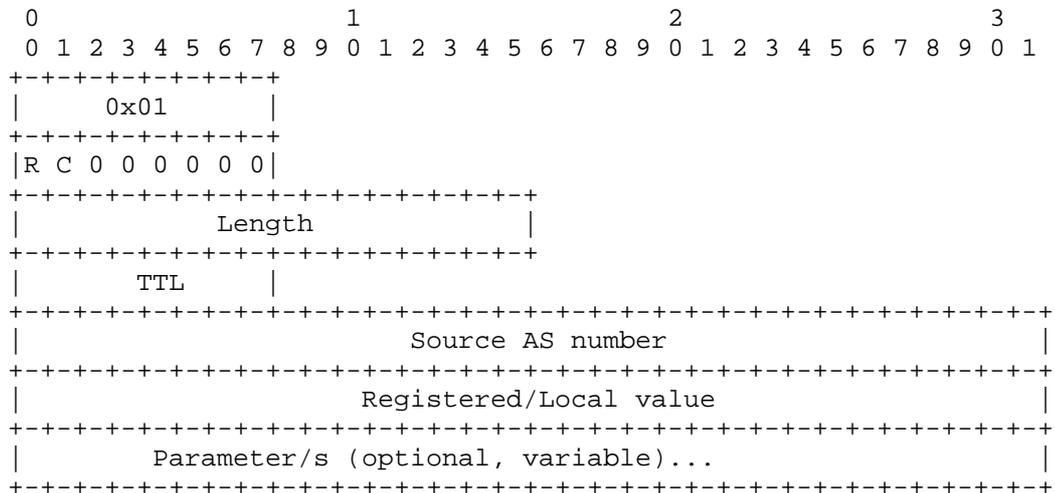


Figure 4: Variable size type container (TLV Format)

4. Container Type 1: Wide Community

Wide BGP Community Type 1 container is of variable size and is encoded as follows:



R is the value of the registered/local bit. C is the value indicating how to treat TTL field across confederation boundaries.

Figure 4: Wide BGP Community Type 1

4.1. Container Type 1 - TTL

TTL: 1 octet

This field represents the forwarding radius in the unit of AS hops for given Wide BGP Community. At each AS boundary when propagating given community over an EBGp session the TTL field must be decremented by value of 1 by the sending EBGp speaker. TTL with value of zero received to the ASBR over IBGP session indicates that this community must not cross an AS boundary.

The special value of 0xFF indicates that the enclosed community may be always propagated over EBGp boundary. Value of 0xFF must not be decremented during propagation.

The exact same procedures as described above applies also to sub-confederation boundaries when the global C flag is set to 1.

4.2. Container Type 1 - Length

The length represents the total lengths of a given container in octets. The minimum length when no optional parameters are attached is 13 octets.

4.3. Container Type 1 - Community Value

Community Value: 2 octets

The Wide BGP Community value encoded in this field indicates private/local or registered Wide BGP Community type which defines what set of actions a router is requested or recommended to take upon reception of routes with such BGP communities.

4.4. Container Type 1 - Source AS number

Source Autonomous System number: 4 octets

The Autonomous System number which indicates the originator of given Wide BGP Community.

When Autonomous System is a two octet number the first two octets of this 4 octet value are to be filled with zeros.

4.5. Container Type 1 - Community Parameters

Parameters: variable size

Community parameter are defined to contain additional data for execution of given BGP community.

Community parameter field could consist of an autonomous system number(s) which should be conditionally compared when executing given community, AS PATH prepend count to be added, local preference value to be inserted under some conditions, markers indicating number of BGP speakers traversed, cumulative IGP metrics to be used for transparent redistribution, etc...

For consistent Autonomous System treatment all encoded AS numbers SHOULD be encoded as 4 octet values. When such AS is a two octet number the first two octets of this 4 octet value are to be filled with zeros.

Two special values are reserved in the Parameter Autonomous System number field: 0x00000000 - to indicate "None of Autonomous Systems" and value of 0xFFFFFFFF - to indicate "All of Autonomous Systems".

The detailed interpretation of each set of parameters will be provided when describing given community type in a separate document or when locally defined by an operator.

5. Well Known Standard BGP Communities

According to RFC 1997 as well as to IANA's Well-Known BGP Communities registry today the following BGP communities are defined to have global significance:

0xFFFF0000	planned-shut	[draft-francois-bgp-gshut]
0xFFFFFFFF01	NO_EXPORT	[RFC1997]
0xFFFFFFFF02	NO_ADVERTISE	[RFC1997]
0xFFFFFFFF03	NO_EXPORT_SUBCONFED	[RFC1997]
0xFFFFFFFF04	NOPEER	[RFC3765]

This document recommends for simplicity as well as for avoidance of backward compatibility issues the continued use of BGP Standard Community Attribute type 8 as defined in RFC 1997 to distribute non Autonomous System specific Well-Known BGP Communities.

For the same reason the described registry does not intended to obsolete BGP Extended Community Attribute and any already defined and already deployed extended communities.

6. Operational considerations

Having two different ways to propagate locally assigned BGP communities, one via use of Standard BGP Community attribute and the other one via use of Wide BGP Community may seem to potentially cause problems when considering propagation of conflicting actions.

However it needs to be noticed and pointed out that today even within Standard BGP Communities operator or operators may append similar conflicting information to already existing community propagation tool set.

It is therefor recommended that any implementation when supporting both standard and wide BGP communities will allow for their easy inbound and outbound policing. For the actual execution all communities should be treated as union and if supported by an implementation their execution permission are to be a local configuration matter.

When advertising as well as during insertion of Wide BGP Communities

which are predefined as range of values - only use of one value of selected range is allowed.

7. Example

An operator wishes to tag incoming routes with a policy specifying that during their advertisement to two peering ASes 2424 and 8888 or during their advertisement to peers marked as RED (0xFF0000) the routes carrying such community will be advertised with AS_PREPEND equal to 4.

That can be easily accomplished by locally defining by an operator a new wide community value using type 1 proposed encoding as below:

```
PREPEND 4 TIMES TO AS 2424 or 8888 or to peers marked as RED
```

```
TTL - 0x00
LENGTH - 26 octets
VALUE - 01 / 0x12
PARAMETERS - 2 x 4 octets AS number
              1 x class of peers
              1 octet prepend's number
```

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
|           0x1           |
+-----+-----+-----+-----+
| 0 0 0 0 0 0 0 0 0 |
+-----+-----+-----+-----+
| Length:      26      |
+-----+-----+-----+-----+
|           TTL: 0           |
+-----+-----+-----+-----+
|                               Own ASN                               |
+-----+-----+-----+-----+
|           Community: LOCAL PREPEND ACTION CATEGORY I           |
+-----+-----+-----+-----+
|           Target ASN# 2424 (0x00000978)           |
+-----+-----+-----+-----+
|           Target ASN# 8888 (0x000022B8)           |
+-----+-----+-----+-----+
|           Peer color RED 0x00FF0000           |
+-----+-----+-----+-----+
|           Prepend #: 4 |
+-----+-----+-----+-----+

```

8. Security considerations

All the security considerations for BGP Communities as well as for BGP RFCs apply here.

9. IANA Considerations

This document defines a new BGP Path Attribute called Wide BGP Communities Attribute. For this new type IANA is to allocate new type value in the corresponding registry:

Registry Name: BGP Path Attributes

This document makes the following assignments for the optional, transitive Wide BGP Communities Attribute:

Name	Type Value
----	-----
Wide BGP Community Attribute	27

This document requests IANA to define and maintain a new registry named: "Wide BGP Communities Attribute Container Types".

The pool of: 0x00-0xFF has been defined for its allocations. The allocation policy is on a first come first served basis.

This document makes the following assignments for the Wide BGP Communities Attribute Types values:

Name	Type Value
----	-----
Reserved	0x00
Type 1	0x01
Types 2-254 to be allocated on FCFS basis	
Reserved	0xFF

10. Contributors

The following people contributed significantly to the content of the document:

Shintaro Kojima
 OTEMACHI 1st. SQUARE EAST TOWER, 3F
 1-5-1, Otemachi,
 Chiyoda-ku, Tokyo 100-0004
 Japan
 Email: koji@mfeed.ad.jp

Juan Alcaide
 Cisco Systems
 Research Triangle Park, NC
 United States
 Email: jalcaide@cisco.com

Burjiz Pithawala
Cisco Systems
170 West Tasman Dr
San Jose, CA
United States
Email: bpithaw@cisco.com

Saku Ytti
TDC Oy
Mechelininkatu 1a
00094 TDC
Finland
Email: ytti@tdc.net

11. Acknowledgments

Authors would like to thank Enke Chen, Pedro Marques and Alton Lo for their valuable input.

12. References

12.1. Normative References

- [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.
- [RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, February 2006.

12.2. Informative References

- [RFC1997] Chandrasekeran, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, August 1996.
- [RFC1998] Chen, E. and T. Bates, "An Application of the BGP Community Attribute in Multi-home Routing", RFC 1998, August 1996.
- [RFC4384] Meyer, D., "BGP Communities for Data Collection", BCP 114, RFC 4384, February 2006.
- [RFC4893] Vohra, Q. and E. Chen, "BGP Support for Four-octet AS

Number Space", RFC 4893, May 2007.

[RFC5668] Rekhter, Y., Sangli, S., and D. Tappan, "4-Octet AS Specific BGP Extended Community", RFC 5668, October 2009.

Authors' Addresses

Robert Raszuk
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
US

Email: raszuk@cisco.com

Jeffrey Haas
Juniper Networks
1194 N.Mathilda Ave
Sunnyvale, CA 94089
US

Email: jhaas@pfrc.org

Richard A Steenbergen
nLayer Communications, Inc.
209 W Jackson Blvd
Chicago, IL 60606
US

Email: ras@nlayer.net

Bruno Decraene
France Telecom
38-40 rue du General Leclerc
Issi Moulineaux cedex 9 92794
France

Email: bruno.decraene@orange-ftgroup.com

Paul Jakma
School of Computing Science, Uni. of Glasgow
Sir Alwyn Williams Building
University of Glasgow
Glasgow G1 5AE
UK

Email: paulj@dcs.gla.ac.uk

