IDR   J. Heitz, Ed.
Internet-Draft   Cisco
Intended status: Standards Track   J. Snijders, Ed.
Expires: May 17, 2017   NTT
                                                              K. Patel
                                                                Arrcus
                                                           I. Bagdonas
                                                               Equinix
                                                           N. Hilliard
                                                                  INEX
                                                     November 13, 2016

BGP Large Communities
draft-ietf-idr-large-community-07

Abstract

   This document describes the BGP Large Communities attribute, an
   extension to BGP-4. This attribute provides a mechanism to signal
   opaque information within separate namespaces to aid in routing
   management. The attribute is suitable for use with four-octet
   Autonomous System Numbers.

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 [RFC2119].

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 May 17, 2017.

Table of Contents

[1](1). **Introduction**

   BGP implementations typically support a routing policy language to
   control the distribution of routing information. Network operators
   attach BGP communities to routes to associate particular properties
   with these routes. These properties may include information such as
   the route origin location, or specification of a routing policy
   action to be taken, or one that has been taken, and is applied to all
   routes contained in a BGP Update Message where the Communities
   Attribute is included. Because BGP communities are optional
   transitive BGP attributes, BGP communities may be acted upon or
   otherwise used by routing policies in other Autonomous Systems (ASes)
   on the Internet.

BGP Communities attributes are a variable length attribute consisting of a set of one or more four-octet values, each of which specify a community [RFC1997]. Common use of the individual values of this attribute type split this single 32-bit value into two 16-bit values. The most significant word is interpreted as an Autonomous System Number (ASN) and the least significant word is a locally defined value whose meaning is assigned by the operator of the Autonomous System in the most significant word.

Since the adoption of four-octet ASNs [RFC6793], the BGP Communities attribute can no longer accommodate the above encoding, as a two-octet word cannot fit a four-octet ASN. The BGP Extended Communities attribute [RFC4360] is also unsuitable. The six-octet length of the Extended Community value precludes the common operational practise of encoding four-octet ASNs in both the Global Administrator and the Local Administrator sub-fields.

To address these shortcomings, this document defines a BGP Large Communities attribute encoded as an unordered set of one or more twelve-octet values, each consisting of a four-octet Global Administrator field and two four-octet operator-defined fields, each of which can be used to denote properties or actions significant to the operator of the Autonomous System assigning the values.

2. BGP Large Communities Attribute

This document defines the BGP Large Communities attribute as an optional transitive path attribute of variable length. All routes with the BGP Large Communities attribute belong to the communities specified in the attribute.

Each BGP Large Community value is encoded as a 12-octet quantity, as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Global Administrator                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Local Data Part 1                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Local Data Part 2                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Global Administrator: A four-octet namespace identifier.

Local Data Part 1: A four-octet operator-defined value.

   Local Data Part 2: A four-octet operator-defined value.

   The Global Administrator field is intended to allow different
   Autonomous Systems to define BGP Large Communities without collision.
   This field SHOULD be either one of the reserved values as defined
   below, or an Autonomous System Number (ASN). If it is a reserved
   value, then the Local Data Parts are as defined by the reserved
   value. If it is an ASN then the Local Data Parts are to be
   interpreted as defined by the owner of the ASN.

   There is no significance to the order in which twelve-octet Large
   Community Attribute values are encoded in a Large Communities
   attribute, A BGP speaker can transmit them in any order.

   Duplicate BGP Large Community values SHOULD NOT be transmitted. A
   receiving speaker SHOULD silently remove duplicate BGP Large
   Community values from a BGP Large Community attribute.

3. Aggregation

   If a range of routes is aggregated, then the resulting aggregate
   should have a BGP Large Communities attribute which contains all of
   the BGP Large Communities attributes from all of the aggregated
   routes.

4. Canonical Representation

   The canonical representation of BGP Large Communities is three
   separate unsigned integers in decimal notation in the following
   order: Global Administrator, Local Data 1, Local Data 2. Numbers
   MUST NOT contain leading zeros; a zero value MUST be represented with
   a single zero. Each number is separated from the next by a single
   colon. For example: 64496:4294967295:2, 64496:0:0.

   BGP Large Communities SHOULD be represented in the canonical
   representation.

5. Reserved BGP Large Community values

   The following Global Administrator values are reserved: 0, 65535, and
   4294967295. Operators SHOULD NOT use these Global Administrator
   values.

   Although this document does not define any Special-Use BGP Large
   Communities, the Global Administrator values specified above could be
   used if there is a future need for them.

6. Error Handling

   The error handling of BGP Large Communities is as follows:

   o A BGP Large Communities attribute SHALL be considered malformed if
      the length of the BGP Large Communities Attribute value, expressed
      in octets, is not a non-zero multiple of 12.

   o A BGP UPDATE message with a malformed BGP Large Communities
      attribute SHALL be handled using the approach of "treat-as-
      withdraw" as described in section 2 [RFC7606].

   The BGP Large Communities Global Administrator field MAY contain any
   value, and a BGP Large Communities attribute MUST NOT be considered
   malformed if the Global Administrator field contains an unallocated,
   unassigned or reserved ASN or is set to one of the reserved BGP Large
   Community values defined in Section 5.

7. Security Considerations

   This extension to BGP has similar security implications as BGP
   Communities [RFC1997].

   This document does not change any underlying security issues
   associated with any other BGP Communities mechanism. Specifically,
   an AS relying on the BGP Large Communities attribute carried in BGP
   must have trust in every other AS in the path, as any intermediate
   Autonomous System in the path may have added, deleted, or altered the
   BGP Large Communities attribute. Specifying the mechanism to provide
   such trust is beyond the scope of this document.

   BGP Large Communities do not protect the integrity of each community
   value. Operators should be aware that it is possible for a BGP
   speaker to alter BGP Large Community Attribute values in a BGP Update
   Message. Protecting the integrity of the transitive handling of BGP
   Large Community attributes in a manner consistent with the intent of
   expressed BGP routing policies falls within the broader scope of
   securing BGP, and is not specifically addressed here.

   Network administrators should note the recommendations in Section 11
   of BGP Operations and Security [RFC7454].

8. Implementation status - RFC EDITOR: REMOVE BEFORE PUBLICATION

   This section records the status of known implementations of the
   protocol defined by this specification at the time of posting of this
   Internet-Draft, and is based on a proposal described in [RFC7942].
   The description of implementations in this section is intended to

assist the IETF in its decision processes in progressing drafts to
RFCs. Please note that the listing of any individual implementation
here does not imply endorsement by the IETF. Furthermore, no effort
has been spent to verify the information presented here that was
supplied by IETF contributors. This is not intended as, and must not
be construed to be, a catalog of available implementations or their
features. Readers are advised to note that other implementations may
exist.

As of today these vendors have produced an implementation of BGP
Large Communities:

o Cisco IOS XR

o ExaBGP

o GoBGP

o BIRD

o OpenBGPD

o pmacct

o Quagga

The latest implementation news is tracked at
http://largebgpcommunities.net/ [1].

9. IANA Considerations

IANA has made an Early Allocation of the value 32 (LARGE_COMMUNITY)
in the "BGP Path Attributes" registry under the "Border Gateway
Protocol (BGP) Parameters" group and is now asked to make that
Permanent.

10. Contributors

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

John Heasley
NTT Communications
Email: heas@shrubbery.net

Adam Simpson
Nokia
Email: adam.1.simpson@nokia.com

11. Acknowledgments

12. References

12.1. Normative References

[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>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet
           Autonomous System (AS) Number Space", RFC 6793,
           DOI 10.17487/RFC6793, December 2012,
           <http://www.rfc-editor.org/info/rfc6793>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K.
           Patel, "Revised Error Handling for BGP UPDATE Messages",
           RFC 7606, DOI 10.17487/RFC7606, August 2015,
           <http://www.rfc-editor.org/info/rfc7606>.

12.2. Informative References

   [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>.

   [RFC7454] Durand, J., Pepelnjak, I., and G. Doering, "BGP Operations
              and Security", BCP 194, RFC 7454, DOI 10.17487/RFC7454,
              February 2015, <http://www.rfc-editor.org/info/rfc7454>.

   [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running
              Code: The Implementation Status Section", BCP 205,
              RFC 7942, DOI 10.17487/RFC7942, July 2016,
              <http://www.rfc-editor.org/info/rfc7942>.

12.3. URIs

   [1] http://largebgpcommunities.net

Authors' Addresses

   Jakob Heitz (editor)
   Cisco
   170 West Tasman Drive
   San Jose, CA 95054
   USA

   Email: jheitz@cisco.com

   Job Snijders (editor)
   NTT Communications
   Theodorus Majofskistraat 100
   Amsterdam 1065 SZ
   The Netherlands

   Email: job@ntt.net

   Keyur Patel
   Arrcus, Inc

   Email: keyur@arrcus.com

Ignas Bagdonas
Equinix
London
UK

Email: ibagdona.ietf@gmail.com

Nick Hilliard
INEX
4027 Kingswood Road
Dublin   24
IE

Email: nick@inex.ie