Large BGP Communities

Matsuzaki ‘maz’ Yoshinobu
<maz@iij.ad.jp>
A Brief History of BGP Communities

• BGP Communities Attribute (RFC 1997, August 1996)
  – Designed to simplify Internet routing policies
  – Encodes a 32-bit value displayed as “16-bit ASN:16-bit value”
  – Broad support in BGP implementations, and widely deployed by network operators for Internet routing
  – For example: 2914:420 2914:1206 2914:2203 2914:3200

• BGP Extended Communities Attribute (RFC 4360, February 2006)
  – Adds label, value, longer range
  – Useful for L3VPNs, fewer implementations available
  – Slow adoption rate
  – Cannot see the forest for the trees (RFC 7153)
What Network Operators Use

- RFC 1997 style communities, as they have been used for the past 20 years
- Widely documented in training material, operations procedures, policy documentation
- Required in RFPs and documented in contracts

Sources: https://www.us.ntt.net/support/policy/routing.cfm (AS 2914), https://onestep.net/communities/
2016/11/18
BTN OG3, Thimphu Bhutan
Along Came a Problem

- We knew we’d run out of 16-bit ASNs eventually
- 32-bit ASN work started in January 2001
  - RFC 4893 in May 2007
  - RFC 6793 in December 2012
- RIRs started allocating 32-bit ASNs by request in 2007
- No distinction between 16-bit and 32-bit ASNs now
  - Widely used as edge and transit ASNs
- However, you can’t fit a 32-bit value into a 16-bit field
  - Can’t use native 32-bit ASNs at all
  - 32-bit ASN owners use private ASNs in communities or some other kludge
  - Creates namespace collisions between ASNs

32-bit ASNs in a 16-bit Field
Abstract

This document describes the Large BGP 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 in 4-octet ASNs.
Related Work for 32-Bit ASNs in Communities

• 4-Octet AS Specific BGP Extended Community ([RFC 5668], October 2009)
  – RFC 4360 style extended community for 32-bit ASNs
  – Perceived as a micro optimization
• Flexible BGP Communities ([draft-lange-flexible-bgp-communities])
  – December 2002 – August 2010
  – BGP peer community grouping, 32-bit ASNs, plus other stuff
  – No consensus or implementations
• Wide BGP Communities Attribute ([draft-ietf-idr-wide-bgp-communities])
  – July 2010 – September 2016
  – Complementary and comprehensive solution
  – Generalized BGP peer community grouping, 32-bit ASNs, plus other stuff
  – No consensus or implementations, needs time to develop
• No Internet routing communities solution for almost 10 years
IETF Support for Large BGP Communities

• Overwhelming interest on the IDR mailing list
  – Network operators
  – Implementers
• Hundreds of messages and counting on the Working Group adoption thread

Working Group adoption thread: https://mailarchive.ietf.org/arch/search/?email_list=idor&gbi=1&index=vea3744yi5sj8bub_i54uay-fE
2016/11/18
Like RFC 1997 Communities, but Larger
Design Goals

• Simply “larger”, that’s it...
  – No added complexity or functionality
  – Extend RFC 1997 communities for 32-bit ASNs
  – Signal an action without losing information about either the origin or the target

• Broadly deployable solution that is available quickly
  – Transitive

• Flexibility for network operators to define their own communities
  – Opaque, may be ignored

• A unique namespace for all 16-bit and 32-bit ASNs
  – Parity and fairness as everyone now can use their globally unique ASN
  – No namespace collisions between ASNs

• Easy to implement
• Easy to adopt
• Easy to remember and tell each other on the phone
  – Canonical representation
  – Especially in an international community with many different languages
Things That are “Out of Scope”

• No RFC 1997 to Large BGP Communities mapping
  – Out of scope because routing policies differ widely between network operators

• No TLV or header
  – Just use BGP Path Attributes code 30 (0x1E)
  – Purposely kept simple to meet the specific use requirements

• No well-known communities
  – Not needed, since RFC 1997 well-known communities like “no-advertise”, “no–export”, “blackhole”, etc. can still be used
Encoding and Usage

- Large BGP Communities are encoded as a 96-bit quantity and displayed as “32-bit ASN:32-bit value:32-bit value”
- Canonical representation is $Me:$Action:$You
Large BGP Community Examples

<table>
<thead>
<tr>
<th>RFC 1997 (Current)</th>
<th>Large BGP Communities</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>65400:peer-as</td>
<td>2914:65400:peer-as</td>
<td>Do not Advertise to peer-as in North America (NTT)</td>
</tr>
<tr>
<td>0:peer-as</td>
<td>6667:0:peer-as</td>
<td>Do not Announce to Route Server peer-as (AMS-IX)</td>
</tr>
<tr>
<td>65520:nnn</td>
<td>2914:65520:nnn</td>
<td>Lower Local Preference in Country nnn (NTT)</td>
</tr>
<tr>
<td>2914:410</td>
<td>2914:400:10</td>
<td>Route Received From a Peering Partner (NTT)</td>
</tr>
<tr>
<td>2914:420</td>
<td>2914:400:20</td>
<td>Route Received From a Customer (NTT)</td>
</tr>
</tbody>
</table>

- No namespace collisions or use of reserved ASNs
- Enables us to use 32-bit ASNs in $Me and $You values
## Major Milestones Towards an RFC Standard

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2, 2016</td>
<td>Published draft-heitz-idr-large-community-03</td>
</tr>
<tr>
<td>September 6, 2016</td>
<td>Requested IDR WG Adoption</td>
</tr>
<tr>
<td>September 24, 2016</td>
<td>IDR Working Group Adoption of draft-ietf-idr-large-community-00</td>
</tr>
<tr>
<td>September 29, 2016</td>
<td>Early IANA BGP Path Attributes Code (30) Allocation</td>
</tr>
<tr>
<td>October 1, 2016</td>
<td>Published draft-ietf-idr-large-community-01</td>
</tr>
<tr>
<td>October 8, 2016</td>
<td>Published draft-ietf-idr-large-community-02</td>
</tr>
<tr>
<td>October 11, 2016</td>
<td>Large BGP Communities Beacon Prefixes Announced</td>
</tr>
<tr>
<td>October 16, 2016</td>
<td>Published draft-ietf-idr-large-community-03</td>
</tr>
<tr>
<td>October 17, 2016</td>
<td>Start of IDR Working Group Last Call</td>
</tr>
<tr>
<td>October 26, 2016</td>
<td>Early IANA BGP Path Attributes Code (32) Allocation</td>
</tr>
<tr>
<td>November 2, 2016</td>
<td>Start of IETF Last Call and IESG Review</td>
</tr>
</tbody>
</table>
Timeline Overview

**IETF**
Consensus Building, Progression from I-D to RFC, Publication

**Implementers**
Feature Design, Implementation, Testing, Documentation, Shipping

**Network Operators**
Evangelism, Training, Preparation, Testing, Deployment

Days/Months
Months/Years
Weeks/Months
18 Months
∞
12 Months
## BGP Speaker Implementation Status

<table>
<thead>
<tr>
<th>Implementation</th>
<th>Software</th>
<th>Status</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arista</td>
<td>EOS</td>
<td>Planned</td>
<td>Feature Requested BUG169446</td>
</tr>
<tr>
<td>Cisco</td>
<td>IOS XR</td>
<td>✔️ Done!</td>
<td>Engineering Release</td>
</tr>
<tr>
<td>cz.nic</td>
<td>BIRD</td>
<td>✔️ Done!</td>
<td>BIRD 1.6.3 (<a href="#">commit</a>)</td>
</tr>
<tr>
<td>ExaBGP</td>
<td>ExaBGP</td>
<td>✔️ Done!</td>
<td>PR482</td>
</tr>
<tr>
<td>Juniper</td>
<td>JunOS</td>
<td>Planned</td>
<td>Second Half 2017</td>
</tr>
<tr>
<td>MikroTik</td>
<td>RouterOS</td>
<td>Won’t Implement Until RFC</td>
<td>Feature Requested 2016090522001073</td>
</tr>
<tr>
<td>Nokia</td>
<td>SR OS</td>
<td>Planned</td>
<td></td>
</tr>
<tr>
<td>OpenBSD</td>
<td>OpenBGPD</td>
<td>✔️ Done!</td>
<td>OpenBSD 6.1 (<a href="#">commit</a>)</td>
</tr>
<tr>
<td>OSGR</td>
<td>GoBGP</td>
<td>✔️ Done!</td>
<td>PR1094</td>
</tr>
<tr>
<td>rtbrick</td>
<td>Fullstack</td>
<td>Planned</td>
<td>ETA: <a href="#">December 2016</a></td>
</tr>
<tr>
<td>Quagga</td>
<td>Quagga</td>
<td>✔️ Done!</td>
<td>Patch Provided for 1.1.0 <a href="#">875</a></td>
</tr>
<tr>
<td>VyOS</td>
<td>VyOS</td>
<td>Requested</td>
<td>Feature Requested <a href="#">T143</a></td>
</tr>
</tbody>
</table>

Visit [http://largebgpcommunities.net/implementations/](http://largebgpcommunities.net/implementations/) for the Latest Status
## Tools and Ecosystem Implementation Status

<table>
<thead>
<tr>
<th>Implementation</th>
<th>Software</th>
<th>Status</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE-CIX</td>
<td>pbgpp</td>
<td>✔️ Done!</td>
<td>PR16</td>
</tr>
<tr>
<td>FreeBSD</td>
<td>tcpdump</td>
<td>✔️ Done!</td>
<td>PR213423</td>
</tr>
<tr>
<td>pmacct.net</td>
<td>pmacct</td>
<td>✔️ Done!</td>
<td>PR61</td>
</tr>
<tr>
<td>OpenBSD</td>
<td>tcpdump</td>
<td>✔️ Done!</td>
<td>OpenBSD 6.1 (patch)</td>
</tr>
<tr>
<td>tcpdump.org</td>
<td>tcpdump</td>
<td>✔️ Done!</td>
<td>PR543 (commit)</td>
</tr>
<tr>
<td>Wireshark</td>
<td>Dissector</td>
<td>✔️ Done!</td>
<td>18172 (patch)</td>
</tr>
</tbody>
</table>

Visit [http://largebgpcommunities.net/implementations/](http://largebgpcommunities.net/implementations/) for the Latest Status
Large BGP Communities Beacon Prefixes

- The following prefixes are announced with AS path 2914_15562$
  - 192.147.168.0/24 (looking glass)
  - 2001:67c:208c::/48 (looking glass)
- Large BGP Community: 15562:1:1

Cisco IOS Output (Without Large BGP Communities Support)

```
route-views>sh ip bgp 192.147.168.0
BGP routing table entry for 192.147.168.0/24, version 98399100
Paths: (39 available, best #30, table default)
  Not advertised to any peer
  Refresh Epoch 1
  701 2914 15562
    137.39.3.55 from 137.39.3.55 (137.39.3.55)
      Origin IGP, localpref 100, valid, external
      unknown transitive attribute: flag 0xE0 type 0x20 length 0xC
      value 0000 3CCA 0000 0001 0000 0001
      rx pathid: 0, tx pathid: 0
```

BIRD Output (With Large BGP Communities Support)

```
COLOCLUE1 11:06:17 from 94.142.247.3] (100/-) [AS15562i]
Type: BGP unicast univ
BGP.origin: IGP
BGP.as_path: 8283 2914 15562
BGP.next_hop: 94.142.247.3
BGP.med: 0
BGP.local_pref: 100
BGP.community: (2914,410) (2914,1206) (2914,2203) (8283,1)
BGP.large_community: (15562, 1, 1)
```
BGP Implementer To Do List

• Add support for BGP Path Attributes code 32 (0x20) to BGP
  – Optional CLI command to enable
• Extend your routing policies
  – Set and match
  – Regular expressions
• Extend your show commands
  – Including the debug commands and packet dump output
• Update your documentation
• Update your training material
• Educate your technical staff
Network Operator To Do List

• The entire network ecosystem needs to support Large BGP Communities in order to provision, deploy and troubleshoot
• Ask your routing vendors and implementers for software support
• Update your tools and provisioning software
• Extend your routing policies, and openly publish this information
• Train your technical staff
Questions?

Presentation created by:

Greg Hankins
Nokia
greg.hankins@nokia.com
@greg_hankins

Job Snijders
NTT Communications
job@ntt.net
@JobSnijders

Visit http://LargeBGPCommunities.net/ for the Latest Info

Reuse of this slide deck is permitted and encouraged!