BGP Communities

BGP for networks who peer: Part 6

Wolfgang Tremmel
wolfgang.tremmel@de-cix.net
BGP (new) Webinars Overview

- 01 - Prefixes and AS numbers
- 02 - BGP Introduction
- 03a - Setting up iBGP
- 03b - Setting up eBGP
- 04 - Becoming multi-homed
- 05 - BGP Best Path Selection
- 06 - BGP Communities
We learned about BGP announcing prefixes
...and about some attributes of prefix announcements
Attributes of BGP prefixes

- **Mandatory** attributes: have to be there
  - Example: AS-Path

- **Optional** attribute: are, well, optional
  - Example: MED

- **Transitive** attributes
  - are kept on the prefix and forwarded via BGP

- **Non-transitive** attributes
  - are added to a prefix and not forwarded by the receiver
About BGP attributes

- **AS path**: mandatory
- **Origin**: mandatory
- **Next Hop**: mandatory
- **MED**: optional, non-transitive
- **Local Preference**: required for iBGP, not sent via eBGP

- IANA keeps track of attribute codes
  - Currently there are more than 40 attributes registered
Introducing:

BGP Communities
**BGP Communities**

- A transitive, optional BGP attribute
- **Transitive**: Once attached, it stays until removed
- **Optional**: it does not have to be there
- "BGP Communities are like a sticker on a suitcase"

Source: https://openclipart.org/detail/273877/suitcase-with-stickers-no-trademarks-remix
"Original" BGP Communities

- Definition:
  "A community is a group of destinations which share some common property"

- Introduced in RFC1997 in year 1996

- A community is expressed by a 32Bit-Number

- High 16 bit are the AS defining the low 16 bits
  - Notation: "6695:1000", "5669:32000"

- You can attach as many communities as you like (within reason)

- BGP max message size is 4096 Bytes
What are they useful for? Information!

198.51.100.0/24  80.81.192.15  from 80.81.192.15

Path: 1301 286 517

Origin IGP, metric 0, localpref 100, valid, external
Informational Communities

198.51.100.0/24  80.81.192.15  from 80.81.192.15

Path: 1301 286 517

Origin IGP, metric 0, localpref 100, valid, external

Received from: Upstiegen
Example: Encode geographical information

65010:1

Example: "1" here means geographical community.

ISO-Country-Codes here …
250 - France
276 - Germany
840 - USA

You may encode the continent here (if you are global) like:
• 1 = Europe
• 2 = North America
• 3 = Asia …

Just an Example!
Example: Encode logical information

65010:2

Example: "2" here means logical source

Upstream? Peering? Customer?
1 = Upstream
2 = Private Peer
3 = Peer at an IXP
4 = Customer

More details here, like:
- Customer ID
- Upstream location
- up to you!
What are they useful for? Action!

198.51.100.0/24
Path: 65010
Origin IGP, metric 0, local pref 100, valid, external
Announce to Partners DE CIX

Encoding up to you!
Action Communities: Encoding

- Again - you only have two 16bit numbers ... (with original BGP Communities)
- Some Ideas ...
  - If you want your customers to send you "actions"
    - You really should have them put your AS number into the first 16bit number
    - You **must scrub** everything they should not send on incoming
  - Possible actions:
    - (not) announce to upstream, peers, customers
    - fine granular announcement control (geographically, by IXP, ...)
    - announce with longer AS path
    - change *local preference*
    - Blackhole
Action Communities: Well-Known

A couple of communities are pre-defined by RFCs

**NO-EXPORT**
- Do not send the prefix to eBGP neighbours (other ASes)

**NO-ADVERTISE**
- Do not send the prefix to anyone (not even internal via iBGP)

**NO-PEER**
- Do not send to any peers

**BLACKHOLE**
- Sink all traffic to prefixes tagged with this community
- Most commonly used with host routes
- Implies NO-EXPORT
32Bit AS? No luck with original communities

65010:12345

- Two 16-bit numbers
- No way to encode a 32Bit AS number and something else ...
  - RFC4360 - Extended Communities
- Extended Communities - Lots of new features
  - In total 2*32Bits
  - Introducing a "type" field
  - Possible to encode 16Bit Type, 32Bit AS, 16Bit Data
**Extended Communities**

- **I** = Type is IANA assigned (= well known) or private
- **T** = 0: Transitive across AS borders
- **T** = 1: Non-Transitive - should be removed before forwarding to another AS
- **Type**: Types are either IANA-assigned or experimental. For a list of assigned types see the RFC
- **Value**: 48 Bits, meaning is dependent on type
- Standardized in 2006
### Extended Communities and 32Bit ASes

<table>
<thead>
<tr>
<th>I</th>
<th>T</th>
<th>32Bit-AS Number (continued)</th>
<th>32Bit-AS Number</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>01</td>
<td>0x02</td>
<td>0x02 or 0x03</td>
<td>32Bit-AS Number</td>
</tr>
</tbody>
</table>

- You can encode a 32Bit AS-Number
- and a 16 Bit value
### Extended Communities and 32Bit ASes

You can encode a 32Bit AS-Number
- and a 16 Bit value
- or a 16Bit AS-Number
  - and a 32 Bit value

32Bit AS and 32Bit Value?
- not possible!
Extended communities use cases

Notation:
- Similar to original communities: \textbf{RT:6500000:1234} or \textbf{RT:1234:6500000}

Disadvantages
- Only 48 bits in total
- Only one 32-bit value is possible (and one 16-bit value)
- RT, RO and other types confusing to many operators

Conclusion
- Another community version was needed
- It took the IETF a while to realize that (11 years)
Introducing: Large Communities

- Very simple - three 32Bit values (finally something useful)
- Global Administrator:
  - An AS number (in 32Bit notation)
  - Has defined meaning of two other fields
  - May have published that meaning
- Local Data
  - Can be seen as "just two 32Bit numbers"
  - Or as "Function" / "Parameter"
Large BGP Communities

Notation:
- Similar to Original Communities: 196610:100:65000010
- Defined in two RFCs:
  - RFC8092: BGP Large Communities Attribute
  - RFC8195: Use of BGP Large Communities
- A dedicated website exists: [http://largebgpcommunities.net](http://largebgpcommunities.net)
  - Keeping track of Implementations, News etc.
Experiment: Working with BGP Communities
Thank you!

academy@de-cix.net

Interested in more webinars? Please subscribe to our mailing list at https://lists.de-cix.net/wws/subscribe/academy
Links and further reading

Interested in more webinars? Please subscribe to our mailing list at https://lists.de-cix.net/wws/subscribe/academy
Links and further reading

• BGP attribute types:
  • Registering new types: RFC2042
  • Published in BGP Parameters database at IANA
• Well-known communities:
  • Standard: see IANA website
  • Extended: see IANA website
• Provider examples:
  • DE-CIX Communities: https://www.de-cix.net/en/resources/informational-bgp-communities
  • DE-CIX Routeserver Guides: https://www.de-cix.net/en/resources/route-server-guides
  • KPN - AS286 - Community page: https://as286.net/AS286-communities.html
  • NTT - AS2914 - Community page: https://www.us.ntt.net/support/policy/routing.cfm
  • NTT Looking Glass: https://www.us.ntt.net/support/looking-glass/
• Lab: Download the DE-CIX Academy lab here: https://bitbucket.org/decix-academy/dockerbgp/src/master/

Interested in more webinars? Please subscribe to our mailing list at https://lists.de-cix.net/wws/subscribe/academy
# BGP Communities

<table>
<thead>
<tr>
<th></th>
<th>Original Communities</th>
<th>Extended Communities</th>
<th>Large Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Defined in</strong></td>
<td>RFC1997</td>
<td>RFC4360</td>
<td>RFC8092</td>
</tr>
<tr>
<td><strong>Published</strong></td>
<td>August 1996</td>
<td>February 2006</td>
<td>February 2017</td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>32bit</td>
<td>64bit</td>
<td>96bit</td>
</tr>
<tr>
<td><strong>Commonly used</strong></td>
<td>16Bit AS : 16Bit Value</td>
<td>Type : 32Bit AS : 16Bit Value -or- Type : 16Bit Value : 32Bit AS</td>
<td>32Bit AS : 32Bit Value : 32Bit Value</td>
</tr>
<tr>
<td><strong>Example</strong></td>
<td>6695:65010</td>
<td>RT:6695:2010223112</td>
<td>6695:65010:2010223112</td>
</tr>
<tr>
<td><strong>Intention</strong></td>
<td>Replace original standard</td>
<td></td>
<td>Supplement original standard</td>
</tr>
</tbody>
</table>

- **BGP Communities**
- **Original Communities**
- **Extended Communities**
- **Large Communities**

- **Defined in**: RFC1997
  - **Published**: August 1996
  - **Additional RFCs**: RFC1998, RFC3765, RFC7999

- **Defined in**: RFC4360
  - **Published**: February 2006
  - **Additional RFCs**: RFC4384, RFC5668, RFC5701, RFC7153, RFC8097

- **Defined in**: RFC8092
  - **Published**: February 2017
  - **Additional RFCs**: RFC8195 also: [http://largebgpcommunities.net](http://largebgpcommunities.net)

- **Size**
  - **Original Communities**: 32bit
  - **Extended Communities**: 64bit
  - **Large Communities**: 96bit

- **Commonly used**
  - **Original Communities**: 16Bit AS : 16Bit Value
  - **Extended Communities**: Type : 32Bit AS : 16Bit Value -or- Type : 16Bit Value : 32Bit AS
  - **Large Communities**: 32Bit AS : 32Bit Value : 32Bit Value

- **Example**
  - **Original Communities**: 6695:65010
  - **Extended Communities**: RT:6695:2010223112
  - **Large Communities**: 6695:65010:2010223112

- **Intention**
  - **Original Communities**: Replace original standard
  - **Extended Communities**: supplement original standard
## BGP Communities for DE-CIX Route Servers

Example is for Frankfurt

For other sites, see DE-CIX website

<table>
<thead>
<tr>
<th>Action</th>
<th>Original Communities</th>
<th>Extended Communities</th>
<th>Large Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Announce to all peers</td>
<td>6695:6695</td>
<td>rt:6695:6695</td>
<td>6695:1:0</td>
</tr>
<tr>
<td>Do not announce to any peer</td>
<td>0:6695</td>
<td>rt:0:6695</td>
<td>6695:0:0</td>
</tr>
<tr>
<td>Redistribute to PEERAS (16Bit only)</td>
<td>6695:PEERAS</td>
<td>rt:6695:PEERAS</td>
<td>6695:1:PEERAS</td>
</tr>
<tr>
<td>Do not redistribute to PEERAS</td>
<td>0:PEERAS</td>
<td>rt:0:PEERAS</td>
<td>6695:0:PEERAS</td>
</tr>
<tr>
<td>Add NO-EXPORT</td>
<td>6695:65281</td>
<td></td>
<td>6695:901:0</td>
</tr>
<tr>
<td>Add NO-ADVERTISE</td>
<td>6695:65282</td>
<td></td>
<td>6695:902:0</td>
</tr>
<tr>
<td>Add NO-EXPORT to PEERAS</td>
<td></td>
<td></td>
<td>6695:901:PEERAS</td>
</tr>
<tr>
<td>Add NO-ADVERTISE to PEERAS</td>
<td></td>
<td></td>
<td>6695:902:PEERAS</td>
</tr>
<tr>
<td>Prepend 1 times to all peers</td>
<td>65001:0</td>
<td></td>
<td>6695:101:0</td>
</tr>
<tr>
<td>Prepend 2 times to all peers</td>
<td>65002:0</td>
<td></td>
<td>6695:102:0</td>
</tr>
<tr>
<td>Prepend 3 times to all peers</td>
<td>65003:0</td>
<td></td>
<td>6695:103:0</td>
</tr>
<tr>
<td>Prepend 1 times to PEERAS</td>
<td>65001:PEERAS</td>
<td>rt:65001:PEERAS</td>
<td>6695:101:PEERAS</td>
</tr>
<tr>
<td>Prepend 2 times to PEERAS</td>
<td>65002:PEERAS</td>
<td>rt:65002:PEERAS</td>
<td>6695:102:PEERAS</td>
</tr>
<tr>
<td>Prepend 3 times to PEERAS</td>
<td>65003:PEERAS</td>
<td>rt:65003:PEERAS</td>
<td>6695:103:PEERAS</td>
</tr>
<tr>
<td>Blackhole (if supported by Peer)</td>
<td>BLACKHOLE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
https://de-cix.net/academy

Interested in more webinars? Please subscribe to our mailing list at https://lists.de-cix.net/wws/subscribe/academy