BGP Communities for Data Collection
draft-ietf-grow-collection-communities-08

Status of this Memo

Copyright Notice

Abstract

BGP communities (RFC 1997) are used by service providers for many purposes, including tagging of customer, peer, and geographically originated routes. Such tagging is typically used to control the scope of redistribution of routes within a provider's network, and to its peers and customers. With the advent of large scale BGP data collection (and associated research), it has become clear that the information carried in such communities is essential for a deeper understanding of the global routing system. This memo defines standard (outbound) communities and their encodings for export to BGP

route collectors.

Table of Contents

1. Introduction

BGP communities [RFC1997] are used by service providers for many purposes, including tagging of customer, peer, and geographically originated routes. Such tagging is typically used to control the scope of redistribution of routes within a providers network, and to its customers and peers. Communities are also used for a wide variety of other applications, such as allowing customers to set attributes such as LOCAL_PREF [RFC1771] by sending appropriate communities to their service provider. Other applications include signaling various types of VPNs (e.g., VPLS [I-D.ietf-ppvpn-vpls-requirements]), and carrying link bandwidth for traffic engineering applications [I-D.ietf-idr-bgp-ext-communities].

With the advent of large scale BGP data collection [RV][RIS] (and associated research), it has become clear that the geographical and topological information, as well as the relationship the provider has to the source of a route (e.g., transit, peer, or customer), carried in such communities is essential for a deeper understanding of the global routing system. This memo defines standard communities for export to BGP route collectors. These communities represent a significant part of information carried by service providers as of this writing, and as such could be useful for internal use by service providers. However, such use is beyond the scope of this memo. Finally, those involved in BGP data analysis are encouraged to verify with their data sources as to which peers implement this scheme (as there is a large amount of existing data as well as many legacy peerings).

The remainder of this memo is organized as follows. Section 2 provides both the definition of terms used as well as the semantics of the communities used for BGP data collection, and section 3 defines the corresponding encodings for RFC 1997 [RFC1997] communities. Finally, section 4 defines the encodings for use with extended communities [I-D.ietf-idr-bgp-ext-communities].

2. Definitions

In this section, we define the terms used and the categories of routes that may be tagged with communities. This tagging is often refered to as coloring, and we refer to a route's "color" as its community value. The categories defined here are loosely modeled on those described in [WANG] and [HUSTON].

2.1. Peers and Peering

Consider two network service providers, A and B. Service providers A

and B are defined to be peers when (i). A and B exchange routes via BGP, and (ii). traffic exchange between A and B is settlement-free. This arrangement is also typically known as "peering". Peers typically exchange only their respective customer routes (see "Customer Routes" below), and hence exchange only their respective customer traffic. See [HUSTON] for a more in-depth discussion of the business models surrounding peers and peering.

2.2. Customer Routes

Customer routes are those routes which are heard from a customer via BGP and are propagated to peers and other customers. Note that a customer can be an enterprise or another network service provider. These routes are sometimes called client routes [HUSTON].

2.3. Peer Routes

Peer routes are those routes heard from peers via BGP, and not propagated to other peers. In particular, these routes are only propagated to the service provider's customers.

2.4. Internal Routes

Internal routes are those routes that a service provider originates and passes to its peers and customers. These routes are frequently taken out of the address space allocated to a provider.

2.5. Internal More Specific Routes

Internal more-specific routes are those routes which are frequently used for circuit load balancing purposes, IGP route reduction, and also may correspond to customer services which are not visible outside the service provider's network. Internal more specific routes are not exported to any external peer.

2.6. Special Purpose Routes

Special purpose routes are those routes which do not fall into any of the other classes described here. In those cases in which such routes need to be distinguished, a service provider may color such routes with a unique value. Examples of special purpose routes include anycast routes, and routes for overlay networks.

2.7. Upstream Routes

Upstream routes are typically learned from upstream service provider as part of a transit service contract executed with the upstream provider.

2.8. National Routes

These are route sets that are sourced from and/or received within a particular country.

2.9. Regional Routes

Several global backbones implement regional policy based on their deployed footprint, and on strategic and business imperatives. Service providers often have settlement-free interconnections with an AS in one region, and that same AS is a customer in another region. This mandates use of regional routing, including community attributes set by the network in question to allow easy discrimination among regional routes. For example, service providers may treat a route set received from another service provider in Europe differently than the same route set received in North America, as it is common practice to sell transit in one region while peering in the other.

3. RFC 1997 Community Encoding and Values

In this section we provide RFC 1997 [RFC1997] community values for the categories described above. RFC 1997 communities are encoded as BGP Type Code 8, and are treated as 32 bit values ranging from 0x0000000 through 0xFFFFFFF. The values 0x0000000 through 0x0000FFFF and 0xFFFF0000 through 0xFFFFFFFF are reserved.

The best current practice among service providers is to use the high order two octets to represent the provider's AS number, and the low order two octets to represent the classification of the route, as depicted below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                <AS>           |           <Value>             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where <AS> is the 16 bit AS number. For example, the encoding 0x2A7C029A would represent the AS 10876 with value 666.

4. Community Values for BGP Data Collection

In this section we define the RFC 1997 community encoding for the route types described above for use in BGP data collection. It is anticipated that a service provider's internal community values will be converted to these standard values for output to a route

collector.

This memo follows the best current practice of using the basic format <AS>:<Value>. The values for the route categories are described in the following table:

Category	Value
Reserved	<AS>:0000000000000000
Customer Routes	<AS>:0000000000000001
Peer Routes	<AS>:0000000000000010
Internal Routes	<AS>:0000000000000011
Internal More Specific Routes	<AS>:0000000000000100
Special Purpose Routes	<AS>:0000000000000101
Upstream Routes	<AS>:0000000000000110
Reserved	<AS>:0000000000000111- <AS>:0000011111111111
National and Regional Routes	<AS>:0000100000000000- <AS>:1111111111111111
Encoded as	<AS>:<R><X><CC>
Reserved National and Regional values	<AS>:0100000000000000- <AS>:1111111111111111

Where

<AS> is the 16-bit AS
<R> is the 5-bit Region Identifier
<X> is the 1-bit satellite link indication
  X = 1 for satellite links, 0 otherwise
<CC> is the 10-bit ISO-3166-2 country code [ISO3166]

and <R> takes the values:

Africa (AF)	00001
Oceania (OC)	00010
Asia (AS)	00011
Antarctica (AQ)	00100
Europe (EU)	00101
Latin America/Caribbean Islands (LAC)	00110
North America (NA)	00111
Reserved	01000-11111

Figure 2: Initially Assigned Community Values

That is:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             <AS>              |   <R>   |X|        <CC>       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

For example, the encoding for a national route over a terrestrial link in AS 10876 from the Fiji Islands would be:

```
 <AS>  = 10876 = 0x2A7C
 <R>   = 00010
 <X>   = 0
 <CC>  = Fiji Islands Country Code = 242 = 0011110010
```

In this case, the low order 16 bits are 0001000011110010 = 0x10F2

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |            0x2A7C             |            0x10F2             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Note that a configuration language might allow the specification of this community as 10876:4338 (0x10F2 == 4338 decimal).

Finally, note that these categories are not intended to be mutually exclusive, and multiple communities can be attached where appropriate.

4.1. Extended Communities

In some cases, the values and their encodings described in Section 4 may clash with a service provider's existing community assignments. Extended communities [I-D.ietf-idr-bgp-ext-communities] provide a convenient mechanism that can be used to avoid such clashes.

The Extended Communities Attribute is a transitive optional BGP attribute with the Type Code 16, and consists of a set of extended communities of the following format:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  Type high    |  Type low(*)  |                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+          Value                |
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

For purposes of BGP data collection, we encode the communities described in Section 4 using the two-octet AS specific extended community type, which has the following format:

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      0x00     |   Sub-Type    |    Global Administrator       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                     Local Administrator                       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The two-octet AS specific extended community attribute encodes the service provider's two octet Autonomous System number (as assigned by a Regional Internet Registry, or RIR) in the Global Administrator field, and the Local Administrator field may encode any information.

This memo assigns Sub-Type 0x0008 for BGP data collection, and specifies that the <Value> field, as defined in section 3.1, is carried in the low order octets of the Local Administrator field. The two high order octets of the Local Administrator field are reserved, and are set to 0x00 when sending and ignored upon receipt.

For example, the extended community encoding for 10876:4338 (representing a terrestrial national route in AS 10876 from the Fiji Islands) would be:

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      0x00     |     0x0008    |           0x2A7C              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      0x00     |      0x00     |           0x10F2              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

4.2. Four-octet AS specific extended communities

The four-octet AS specific extended community is encoded 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      0x02     |    0x0008     |    Global Administrator       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Global Administrator (cont.)  |           0x10F2              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

In this case, the 4 octet Global Administrator sub-field contains a

4-octets Autonomous System number assigned by the IANA.

5. Note on BGP UPDATE Packing

Note that data collection communities have the potential of making the attribute set of a specific route more unique than it would be otherwise (since each route collects data that is specific to it's path inside one or more ASes). This, in turn, can affect whether multiple routes can be grouped in the same BGP update message, and may lead to increased use of bandwidth, router CPU cycles, and memory.

6. Acknowledgments

The community encoding described in this memo germinated from an interesting suggestion from Akira Kato at WIDE. In particular, the idea would be to use the collection community values to select paths that would result in (hopefully) more efficient access to various services. For example, in the case of RFC 3258 [RFC3258] based DNS anycast service, BGP routers may see multiple paths to the same prefix, and others might be coming from the same origin with different paths, but others might be from different region/country (with the same origin AS).

Joe Abley, Randy Bush, Sean Donelan, Xenofontas Dimitropoulos, Vijay Gill, John Heasley, Geoff Huston, Steve Huter, Michael Patton, Olivier Marce, Ryan McDowell, Rob Rockell, Rob Thomas, Pekka Savola, Patrick Verkaik and Alex Zinin all made many insightful comments on early versions of this draft. Henk Uijterwaal suggested the use of the ISO-3166-2 country codes.

7. Security Considerations

While this memo introduces no additional security considerations into the BGP protocol, the information contained in the communities defined in this memo may in some cases reveal network structure that was not previously visible outside the provider's network. As a result, care should be taken when exporting such communities to route collectors. Finally, routes exported to a route collector should also be tagged with the NO_EXPORT community (0xFFFFFF01).

7.1. Total Path Attribute Length

The communities described in this memo are intended for use on egress to a route collector. Hence an operator may choose to overwrite its

internal communities with the values specified in this memo when exporting routes to a route collector. However, operators should in general ensure that the behavior of their BGP implementation is well-defined when the addition of an attribute causes a PDU to exceed 4096 octets. For example, since it is common practice to use community attributes to implement policy (among other functionality such as allowing customers to set attributes such as LOCAL_PREF), the behavior of an implementation when the attribute space overflows is crucial. Among other behaviors, an implementation might usurp the intended attribute data or otherwise cause indeterminate failures. These behaviors can result in unanticipated community attribute sets, and hence result in unintended policy implications.

8. IANA Considerations

This memo assigns a new Sub-Type for the AS specific extended community type in the First Come First Served extended transitive category. In particular, the IANA should assign Sub-type 0x0008 as defined in Section 4.1.

In addition, this memo instructs the IANA to create two registries for BGP Data Collection Communities, one for standard communities and one for extended communities. Both of these registries should initially be populated by the values described in Section 4. IETF Consensus, usually through the Global Routing Operations Working Group (grow) is required for the assignment of new values in these registries (in particular, for <Value> or <R>), as described in Figure 2 [RFC2434].

9. References

9.1. Normative References

[RFC1771] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995.

[RFC1997] Chandrasekeran, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, August 1996.

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[I-D.ietf-idr-bgp-ext-communities]
Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute",

draft-ietf-idr-bgp-ext-communities-07 (work in progress), March 2004.

[ISO3166] "ISO 3166 Maintenance agency (ISO 3166/MA)", Web Page: http://www.iso.org/iso/en/prods-services/iso3166ma/index.html, 2004.

9.2. Informative References

[I-D.ietf-ppvpn-vpls-requirements]
Augustyn, W., "Requirements for Virtual Private LAN Services (VPLS)", draft-ietf-ppvpn-vpls-requirements-00 (work in progress), March 2002.

[RIS] "The RIPE Routing Information Service", Web Page: http://www.ripe.net/ris, 2004.

[RV] Meyer, D., "The Routeviews Project", Web Page: http://www.routeviews.org, 2002.

[WANG] Wang, F. and L. Gao, "Inferring and Characterizing Internet Routing Policies", ACM SIGCOMM Internet Measurement Conference, 2003.

[HUSTON] Huston, G., "Interconnection, Peering, and Settlements", Web Page: http://www.isoc.org/inet99/proceedings/1e/1e_1.htm, 2003.

Author's Address

David Meyer

Email: dmm@1-4-5.net

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.