Internet Engineering Task Force
INTERNET-DRAFT
draft-quoitin-bgp-comm-survey-00.txt

Bruno Quoitin and Olivier Bonaventure
University of Namur, Belgium

February, 2002
Expires August, 2002

A survey of the utilization of the BGP community attribute

Abstract

In this document, we describe the two most common utilizations of the BGP community attribute, namely to tag routes and indicate how a route should be redistributed by external peers. We then discuss how often these two types of community attribute are used on the basis of the RIPE whois database and of BGP table dumps.

Status of this Memo

1 Introduction

The BGP Community attribute defined in [TCL96] is a powerful mechanism that can be used to build more scalable BGP configurations. This attribute consists of a set of four octet values, each of which specifies a community. [TCL96] reserves the community values ranging from 0x0000000 through 0x0000ffff and 0xffff0000 through 0xffffffff. Furthermore, three communities are defined with global significance:

- NO_EXPORT (0xffffff01): routes with this community attached should not be advertised outside a BGP confederation;
- NO_ADVERTISE (0xffffff02): routes with this community attached must not be advertised to other peers;
- NO_EXPORT_SUBCONFED (0xffffff03): routes received with this community attached must not be advertised to peers outside the boundary of a subconfederation.

Besides these reserved community values, [TCL96] proposed to divide the community space by using an AS number in the two high-order octets. This proposal can be considered as a delegation of 65536 values of the community space to each AS. Thus, ASx is free to use community values ranging from ASx:0 to ASx:0xffff. However, [TCL96] did not discuss how the community values corresponding to the private AS space [HB96] (i.e. community values 64512:00 - 65534:65535) could be used in the global Internet.

In this document, we describe the most common utilizations of the BGP community attribute in the global Internet. We base our analysis on the information available in the RIPE whois database and the BGP table dumps collected by the RIPE RIS (Réseaux IP Européens - Routing Information Service) [RIS02] and the Route Views projects [Mey02]. This document is organised as follows. First, we discuss in section 2 utilizations of this attribute on the basis of the RIPE whois database. In section 3 we briefly discuss the communities found in the BGP tables in the global Internet and present our conclusions.

2 Common utilizations of the BGP Community attribute

A classical application of the Community attribute is for multi-homing purposes as discussed in [CB96]. However, since the publication of [CB96], the Community attribute has been used for other purposes, including the support of VPNs [RY99]. We do not discuss this application to VPNs in this document.

Two of the most common utilizations of the Community attribute in the global Internet are to tag the routes received from a specific peer or at a specific location and to influence the redistribution of specific routes in order to perform some kind of interdomain traffic engineering.

2.1 Route tagging communities

In this case, the community value is used by an Autonomous System to indicate the location where the route was received from an external peer. These community values are inserted by the BGP router that receives a route at a given location. Many AS rely on such communities in today's Internet. Based on the needs of each AS, different types of locations are used in practice today : geographic, interconnection point, autonomous system (AS). We provide in the following sections some examples based on the information found in the RIPE whois database in January 2002.

2.1.1 Type of peer

In this case, the AS defines a few types of BGP peers (typically customer, (national or international) peering partner and transit provider) and tags each received route with a community indicating the type of peer from which the route was received.

2.1.2 Geographic location

AS often need to know the geographic location where a given route was received. The types of geographic locations used by each AS depend on the AS size. A national AS might want to know the city where each route was learned, while an international AS would instead need to know the country or continent where a given route was learned. Often, an AS that utilizes such community values relies on an unstructured list of values and associates a location to each value. For example, AS13129 (Global Access Telecommunications, Inc.) defines in [RIW02] the values shown in table 1 to tag routes learned from specific cities.

13129:3010	Frankfurt
13129:3020	Munich
13129:3030	Hamburg
13129:3040	Berlin
13129:3050	Dusseldorf
13129:3210	London
13129:3220	Paris
13129:3610	New York

Table 1: Tagging communities published by AS13129

Some ASs have devised structured encodings of those route tagging community values such as the one of AS286 (EUnet) shown in table 2 where the value used to tag a received route is based on the telephone country code. These communities are documented in [RIW02].

286:1000 + *countrycode*	Public peer routes
286:2000 + *countrycode*	Private peer routes
286:3000 + *countrycode*	Customer routes
where *countrycode* is the E.164 international dial prefix.	

Table 2: Tagging communities published by AS286

Another example is the encoding chosen by AS3561 (Cable & Wireless) shown in table 3 based on the ISO 3166 codes for countries. The resulting communities are documented in [CW02].

3561:*SRCCC*	**S**	is the source (peer or customer)
	R	is the regional code
	CCC	is the ISO 3166 country code

Table 3: Tagging communities published by AS3561

2.1.3 Interconnection point

In some cases, AS also need to remember the interconnection point where a given route was received. For instance, AS13129 defines communities used to tag routes learned at specific interconnection points. These communities, published in [RIW02], are shown in table 4. We have not encountered structured encodings for the community values used to tag the interconnection point where routes where learned.

2.1.4 Autonomous system (AS)

A few AS also use communities to remember the AS from which each route was learned. This utilization of the community attribute is redundant with the AS Path attribute, but could be useful in confederations or to simplify the configuration of some routers. For instance, AS8938

13129:2110	DE-CIX
13129:2120	INXS
13129:2130	SFINX
13129:2140	LINX

Table 4: Tagging communities published by AS13129

(Energis (Switzerland) AG) defines communities used to tag routes learned from specific autonomous systems. These communities [RIW02] are shown in table 5.

8938:2100	Genuity US (AS1)
8938:2200	Level3 US (AS3356)
8938:2300	Ebone (AS1755)
8938:2400	Sprint (AS1239)

Table 5: Tagging communities used by AS8938

Another example is AS1899 (KPNQwest France) that has chosen to reuse community values in the private AS space (64512:0 - 65534:65535) to tag routes received from other ASs as shown in table 6. These communities are documented in [RIW02].

64675:*AS*	Routes received from Peer *AS* on PARIS

Table 6: Tagging communities published by AS1899

2.2 Communities affecting the redistribution of routes

Another important utilization of BGP Community attribute is for traffic engineering purposes. In this case, the community is typically inserted by the originator of the route in order to influence its redistribution by downstream routers.

Three types of communities are often used today to influence the redistribution of routes towards specific peers or interconnection points:

1. Do not announce the route to a specified peer(s);
2. Prepend *n* times to the AS-Path (where we have found values for n generaly ranging from 1 to 3) when announcing the route to specified peer(s);
3. Set the LOCAL_PREF value in the AS receiving the route;

We discuss these three types of communities in more details and show how often they are used based on the RIPE whois database in the following sections.

2.2.1 "Do not announce the route" community

In this case, the community is attached to a route to indicate that this route should not be announced to a specified peer or at a specified interconnection point. This is the case in the example shown in figure 1, where AS10 and AS20 have a private peering contract and AS20 does not want that the routes announced to AS10 be redistributed to AS10's upstream peers.

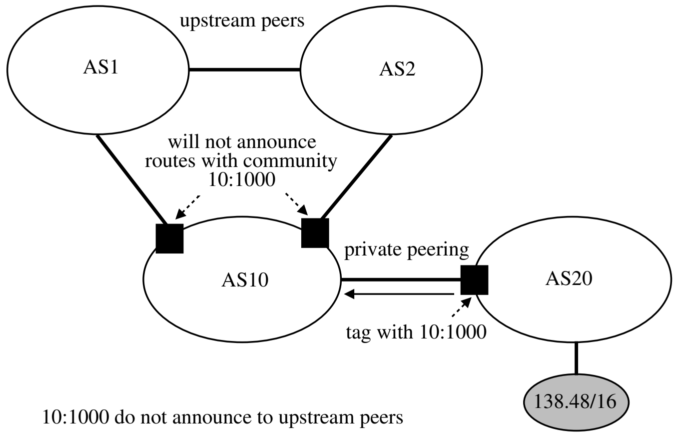

Figure 1: Do not announce to upstream peers

For this, AS20 tags these routes with a community published by AS10 that will prevent the redistribution of such routes.

A large number of AS have documented their support for this kind of community values. Table 7 summarizes the documented utilizations of those communities according to the RIPE whois database in October 2001. This table shows that while many AS utilize community values to indicate that a route should not be announced to a given AS or at a given interconnection point, some also allow the utilizaton of such communities to indicate that a route should not be announced outside a given region or continent.

Most of the AS that support this type of community values rely on an structured list of community values for this purpose. For example, table 8 shows some of the community values used by AS1755 (OpenTransit) and documented in [RIW02].

However, a few AS rely on a more structured enconding of the community values used for this purpose. For example, AS9057 (Level3) has chosen to reuse a range community values of the private AS space as “do not announce” community values as shown in table 9.

2.2.2 Prepend to AS-Path

AS-Path prepending is a manipulation that makes the AS-Path artificially longer when announcing a route to specific peers. The announced route will not be preferred but can still be used as a backup route. Although in theory AS-Path prepending is considered as a rough solution because “*it is virtually impossible to compute the AS-Path length needed to induce the upstream to make the desired choice*” [CAI02], this is a popular solution to control the interdomain traffic received by stub ISPs. The analysis of the BGP table dumps [Hus02] shows that AS-Path prepending is very frequently used in the Internet today.

For instance, in the ridiculous network shown in figure 2, AS10 could provide limited backup transit service to its peer AS20 by announcing routes learned from AS1 and prepending 3 times AS10 to the AS-Path. So, in a normal state, the path from AS1 to AS20 is shorter via AS2. If this path is not available anymore, then the path through AS10 can be used.

Another use of AS-Path prepending is to force some incoming traffic to follow a given path. In the example shown in figure 3, AS1 offers the possibility to its peers to influence the redistribution of their routes by the use of the community attribute. Because AS2 and AS3 carry a lot of traffic towards AS10, AS10 want to achieve some kind of load balancing by forcing the traffic coming from AS2 to follow another path and ask AS1 to prepend two times to the AS-

AS number	Do not announce to
AS1755	US upstreams/peers, European peers, specified AS
AS8437	All upstreams, all peerings, specified AS, specified IX
AS2683	Specified AS, specified IX
AS13299	Specified IX
AS13297	Specified IX
AS3303	any US peers/upstreams, specified AS
AS5571	Specified IX
AS12458	Specified IX
AS8918	Specified AS, Specified IX
AS8235	All peers, specified AS
AS13300	Specified IX
AS2118	Outside AS2118 country
AS16186	Specified transit, specified IX
AS8627	US-Upstream Peers, specified AS, private peers
AS6735	Specified AS, specified IX
AS1557	Specified AS, specified IX
AS15366	Specified IX
AS9032	Specified AS
AS8228	US upstreams/peers, specified AS, peers in country/continent
AS6705	Specified AS, specified IX
AS5400	AS in specified continent, specified AS
AS5511	AS in specified continent, specified AS
AS8472	Specified IX, specified AS
AS1901	AS in continent, specified AS, specified IX
AS12329	Specified AS
AS12306	Specified IX
AS12976	Specified AS
AS517	Specified AS, specified IX
AS3215	Specified IX, specified AS
AS286	AS in specified continent, specified AS
AS8470	Foreign AS, AS in country
AS12541	Specified IX, specified AS
AS13129	Specified AS, specified IX
AS2820	Inside and outside country
AS8246	Specified AS
AS1273	Specified AS, specified IX
AS8938	Any upstream, specified AS
AS8708	Upstreams, peers
AS6728	US, specified AS, specified IX
AS8933	Any commercial peer, specified AS
AS3259	Outside Continent
AS12779	Upstreams, specified AS
AS8210	Upstreams in specified continent, specified AS
AS12832	Downstream AS
AS15444	Specified IX
AS9057	Customers but not peers, specified AS
AS5430	Specified peering
AS12359	Transit providers, customers, specified IX, AS in country
AS702	Only within AS702 and customers, outside continent

Table 7: “Do not announce” communities documented in the RIPE database in October 2001

Value	Meaning
1755:1000	Do not announce to US upstreams/peers
1755:1101	Do not announce to Sprintlink(US)/AS1239
1755:1102	Do not announce to UUNET(US)/AS701
1755:1103	Do not announce to Abovenet(US)/AS6461
...	
1755:2000	No announcement to european peers
...	

Table 8: "Do not announce" communities used by AS1755

Value	Meaning
65000:*XXX*	do not announce on peerings to *AS XXX*
64970:*XXX*	do not announce on Asian/Pacific peerings to *AS XXX*
64980:*XXX*	do not announce on European peerings to *AS XXX*
64990:*XXX*	do not announce on North American peerings to *AS XXX*

Table 9: "Do not announce" communities used by AS9057

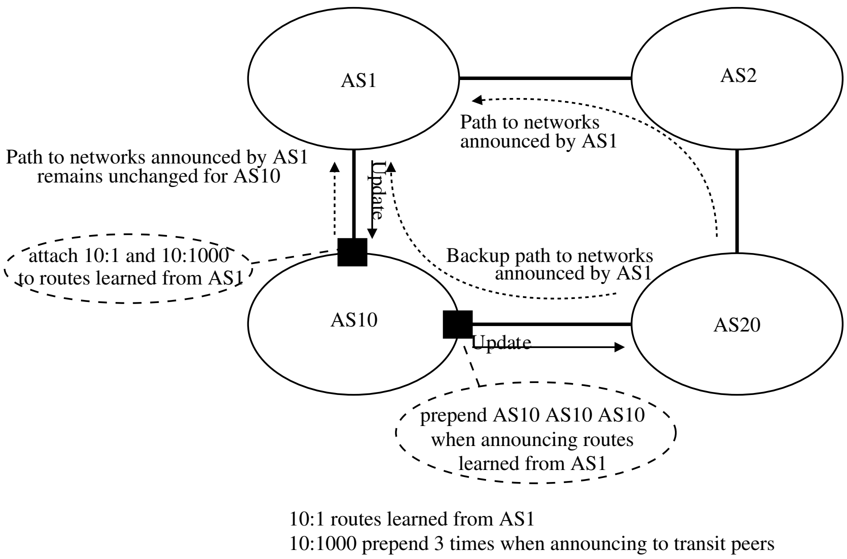

Figure 2: Providing a backup path

Path of routes announced by AS10 when they are forwarded to AS3. Without this change, all the traffic from both AS2 and AS3 would have come through AS1. With the prepending, the path AS20:AS30:AS10 is shorter than AS1:AS1:AS1:AS10 and is then preferred.

Based on the RIPE whois database in October 2001, many ISPs rely on communities to allow their peers (mainly customers) to request the utilization of AS-Path prepending when announcing some routes to specified external peers, at specified interconnection points or in specified regions. A summary of the RIPE whois database may in found in table 10.

AS1755	US upstreams/peers, European peers, specified AS
AS8437	All upstreams, specified AS, specified IX
AS2683	Specified AS, specified IX
AS13299	Specified IX
AS13297	Specified IX
AS3303	All US peers/upstreams, specified AS
AS5571	Specified IX
AS12458	Specified IX
AS8918	Specified AS, Specified IX
AS8235	Specified AS
AS13300	Specified IX
AS8627	All, specified AS, specified IX, private peers
AS6735	Specified AS
AS1557	Specified AS, specified IX
AS15366	Specified IX
AS9032	Specified AS
AS8228	Specified AS, peers inside given country or continent
AS9109	prepend as update crosses continent boundaries
AS12868	All
AS6705	Specified AS, specified IX
AS5400	AS in specified continent, specified AS
AS5511	AS in specified continent, specified AS
AS8472	Specified IX, specified AS
AS1901	AS in continent, specified AS, specified IX
AS12329	Specified AS
AS12306	Specified IX
AS12552	All peers
AS12976	All peers
AS517	Specified AS, specified IX
AS8582	Specified AS
AS3215	Specified IX, specified AS
AS286	AS in specified continent
AS8470	AS in country
AS12541	Specified IX, specified AS
AS3316	All peers
AS13129	Specified AS, specified IX
AS8246	Specified AS
AS1273	Specified AS, specified IX
AS8938	Any upstream, specified AS
AS8708	Upstreams, peers
AS5568	All peers
AS8933	Specified AS
AS3259	Peers
AS12779	Upstreams, specified AS
AS8210	Upstreams in specified continent, specified AS
AS12832	All
AS3292	US transit providers
AS2116	Specified AS
AS8503	Peers
AS9057	Specified AS
AS702	All peers

Table 10: prepend communities documented in the RIPE database in October 2001

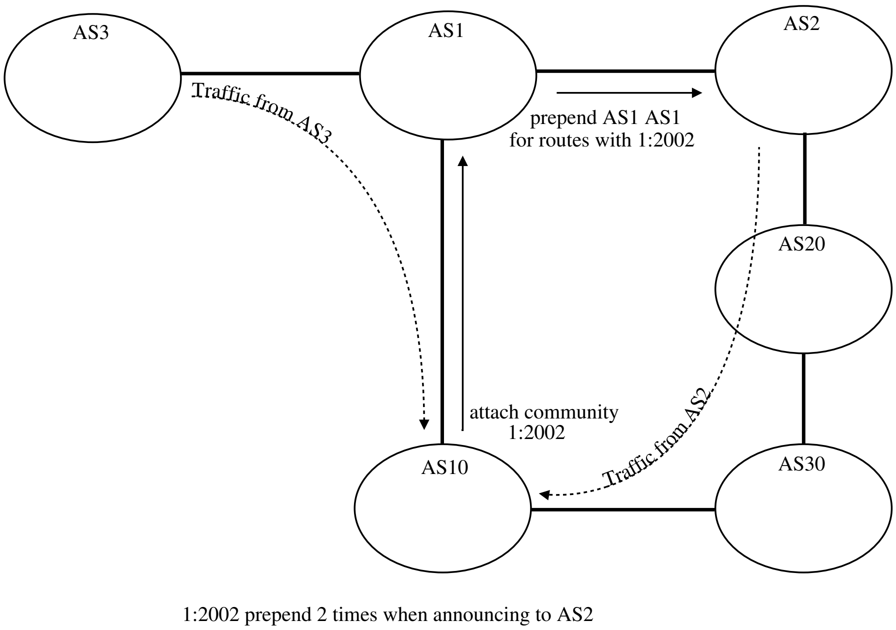

Figure 3: Engineering routes to local prefixes

Usually, an AS that provides such communities relies on an unstructured set of communities. There are however a few exceptions. AS3561 (Cable & Wireless) has devised an interesting set of communities to allow peers to ask not to export or ask to prepend. This set can be found in table 11 (the list of peers to which these community values are supported may be found in [CW02]).

3561:30*PPN*	PP	is the peer code
		= 1, prepend once
		= 2, prepend twice
		= 3, prepend three times

Table 11: AS-Path prepend communities published by AS3561

Some AS have gone one step further by reusing the community values in the private AS space. For example, AS8235 has chosen to use community values $6550N : AAAA$ to allow its customers to request AS8235 to prepend its AS number N times when the associated route is announced to AS $AAAA$.

AS9057 relies even more on the community values in the private AS space. It uses community values from 20 different private AS numbers to allow its customers to indicate whether a route should or should not request path prepending when a route is announced to a specified peer. For example, community value 65001:*XXX* indicates that the associated route should be prepended once when announced to peer *XXX*.

2.2.3 Setting of the local preference

A final utilization of the communities is to set the LOCAL_PREF of the receiving router as documented [CB96]. This utilization of the BGP community attribute is still present in the RIPE whois database and we have found that different levels of preference are provided. For instance: low local preference for customer (backup), normal local preference for customer, high local preference for customer, reduced peering, normal peering, preferred interconnect (private peering), upstream peer and other specific preferences.

In October 2001, 19 AS have documented their utilization of such communities in the RIPE whois database. For example, AS702 (UUNET Europe) defines the 2 communities shown in table 12.

702:80	Set Local Pref 80 within AS702
702:120	Set Local Pref 120 within AS702

Table 12: Communities defined by AS702

3 Analysis of BGP routing tables

Section 2 has described the most common utilizations of the BGP community attribute. From the description above, one could expect that community values should rarely appear in the global Internet routing tables since most communities are used to tag routes inside a given AS or to influence the redistribution of routes by a given AS.

To verify this assumption, we have conducted an analysis of BGP routing tables collected by RIPE RIS project [RIS02] and the Route Views project (University of Oregon, [Mey02]) during the period January 2001 - January 2002. The detailed results of this analysis can be found in [QB02].

A first observation of those BGP table dumps shows that the BGP community attribute is widely used, even in the global Internet. For instance, at RIPE NCC, Amsterdam, the number of communities has increased to more than 1000 distinct values at the beginning of the year 2002 while nearly 50% of the routes advertised to the test router maintained by RIPE had at least one community attached ! We could see the same evolution at other sites except at Otemachi, Japan where no community appears. A short summary can be found in table 13.

While thuse numbers clearly indicate the widespread utilization of the BGP community attribute, they do not distinguish between route tagging and redistribution communities. To understand the types of communities that are used, we have built a database with the communities documented in the RIPE whois database [RIW02] and various web sites of ISPs [TI02, JI02, NE02, CPL00, SPR02, CW02]. However, it should be noted that our database is far from complete since some ASs do publish the description of the communities that their peers can use. Despite of this, we can already find some interesting results.

In table 14, we have classified the communities in three classes. The "Tagging" class corresponds to the communities discussed in section 2.1 while the "TE" class corresponds to the the communities that affect the redistribution of the routes as discussed in section 2.2. The unknown class contains the community values that are not in our database. A graphical evolution of this classification can be found in [QB02] for the period January 2001 to January 2002. Our analysis shows that the "Tagging" and "TE" communities represent a great large of fraction the total number of communities found in the studied BGP routing tables.

The large number of "Unknown" communities in table 14 is due to our incomplete database. However, a closer look at those "Unknown" communities reveals some interesting points. First,

Site	Percentage of routes containing communities	Number of distinct communities
RIPE NCC, Amsterdam	41 %	1233
LINX, London	7 %	668
SFINX, Paris	19 %	38
AMS-IX, Amsterdam	0.4 %	134
CIXP, Geneva	2.3 %	259
VIX, Vienna	84 %	529
JPIX, Otemachi (Japan)	0 %	0
University of Oregon	62.1 %	1774

Table 13: Utilization of communities (Jan 2002).

AS	TE	Tagging	Unknown
RIPE NCC, Amsterdam	60345	331316	758089
LINX, London	14371	16283	13315
SFINX, Paris	31	8	261
AMS-IX, Amsterdam	462	356	1868
CIXP, Geneva	11879	5473	3270
VIX, Vienna	39626	42056	14006
JPIX, Otemachi (Japan)	0	0	0
University of Oregon	314841	388406	2125204

Table 14: Classification of routes on (Jan 2002).

some AS using community values in the space considered as reserved (0x00000000 - 0x0000ffff and 0xffff0000 - 0xffffffff) by [TCL96]. We have seen routes from multiple peers using community values in this range and one peer had announced more than 60k routes with such a community value. Second, we also see some utilization of community values in the private AS space range (i.e. 64512:0 - 65534:65534), but the number of routes with such communities is smaller than those with reserved community values.

4 Conclusion

In this document, we have described two of the main utilizations of the BGP community attribute in the global Internet. The first common utilization of this attribute is to tag the routes received through an eBGP session with an explicit indication of the location (city, country, interconnection point, . . .) where the route was learned. The main reason to utilize route tagging communities is that when it is used on all border routers of a given AS, then all routers of the AS can be configured to make their routing decisions mainly on the basis of those communities. Our analysis of the BGP table dumps and the RIPE whois database shows that this type of BGP communities is often used in today's Internet.

A second common utilization is to affect the redistribution of the associated route by downstream routers. In this case, the community value is associated to a route by the router sending the router to indicate to the remote eBGP peer how the route should be redistributed. We have seen several types of such communities. The two most common cases are used to request that a route should not be announced to a specified (set of) peer(s) and to request the route to be prepended when announced to a specified (set of) peer(s). Our analysis of the RIPE whois database has shown that a large number of AS are using such communities today. Furthermore, some AS have chosen to rely on BGP community values in the private in order to have more structured community values. If this utilization of the BGP community values in the private space would become a widely used solution since there is no coordination between the AS about the utilization of those communities. A much better solution would be to define a set of "well-known" structured community values to support the needs of those AS. A proposal based on the utilization of the extended communities attribute may be found in [BCH+02].

Acknowledgments

This work was partially funded by the European Commission within the IST ATRIUM project.

This work would not have been possible without the very useful information provided by the RIPE RIS project, the RIPE WHOIS database and the RouteViews project.

Author's Addresses

Olivier Bonaventure, Bruno Quoitin
Infonet group (FUNDP)
Rue Grandgagnage 21, B-5000 Namur, Belgium
Email: Olivier.Bonaventure@info.fundp.ac.be, Bruno.Quoitin@info.fundp.ac.be
URL: http://www.infonet.fundp.ac.be

References

[BCH+02] O. Bonaventure, S. De Cnodder, J. Haas, B. Quoitin, and R. White. Controlling the redistribution of bgp routes. Internet draft, draft-bonaventure-bgp-redistribution-02.txt, work in progress (to appear), February 2002.

[CAI02] Cooperative Association for Internet Data Analysis CAIDA. Analysis of BGP data from Oregon Route Views. available from http://www.caida.org/ broido/bgp/bgp.html, January 2002.

[CB96] E. Chen and T. Bates. An Application of the BGP Community Attribute in Multi-home Routing. Internet Engineering Task Force, RFC1998, August 1996.

[CPL00] Connect.com.au Pty Ltd. FAQ on Multi-homing and BGP. available from http://info.connect.com.au/docs/routing/general/multi-faq.shtml, June 2000.

[CW02] Communities defined by Cable & Wireless. available from http://infopage.cary.cw.net/Routing_Registry/communities.htm, January 2002.

[HB96] J. Hawkinson and T. Bates. Guidelines for creation, selection, and registration of an Autonomous System (AS). Internet Engineering Task Force, RFC1930, March 1996.

[Hus02] G. Huston. Telstra Internet Network Performance Reports - BGP Table size. available from http://www.telstra.net/ops/bgp, January 2002.

[JI02] Communities defined by Jippii. available from http://www.jippii.net/communities.html, January 2002.

[Mey02] D. Meyer. Route Views Archive project. University of Oregon, http://archive.routeviews.org, January 2002.

[NE02] Communities defined by Nescalibur. available from http://www.noc.esib.net/bgp_communities, January 2002.

[QB02] B. Quoitin and O. Bonaventure. Utilization of the BGP Communities attribute. Infonet Group, University of Namur, http://alpha.infonet.fundp.ac.be/anabgp, January 2002.

[RIS02] Routing Information Service project. Réseaux IP Européens, http://www.ripe.net/ripencc/pub-services/np/ris-index.html, January 2002.

[RIW02] Whois database. RIPE NCC, http://abcoude.ripe.net/ris/rawdata, January 2002.

[RY99] E. Rosen and Y.Rekhter. Bgp/mpls vpns. Internet Engineering Task Force, RFC2547, March 1999.

[SPR02] Sprintlink BGP FAQ. available from http://www.sprint.net/faq/bgp.html#control, January 2002.

[TCL96] P. Traina, R. Chandrasekeran, and T. Li. BGP Communities Attribute. Internet Engineering Task Force, RFC1997, August 1996.

[TI02] Communities defined by Tiscali Intl Network. available from http://www.as3257.net/html/communities.htm, January 2002.