TATA COMMUNICATIONS

BGP Routing Support in AS6453

GIMEC International

August 3, 2016

Summary

The optional and transitive Border Gateway Protocol (BGP) community attribute, introduced in RFC1997, enables an efficient and flexible mechanism for implementing BGP routing policies.

Autonomous Systems classify their routes: by neighbor type (customer, peer, provider), by geographical import location, or by some other criteria us of choice, using BGP communities, and apply policies to these route classes.

In addition to such internal use of BGP communities, a provider may use them to enrich its customers' inter-domain routing toolbox.

Tata Communications' AS6453 provides customer facing BGP community support that covers what Tier-1 peers offer their customers:

- the Policy Tuning Communities, sent by a customer along with his routes as a request for other than the default policy in the AS6453 network, and
- the Information Communities, sent by AS6453 as hints to a customer.

The resulting enhanced routing service is provided to transit customers only, after their explicit request, and Tata Communication restricts its exchange of community values to the ones published in this document.

A shorter version of this information is available in the RIPE database 'aut-num AS6453' object.

If operationally needed, Tata Communications may (without notice) override any of these hints and policy tuning requests. Generally, though, a notification is sent to the customer.

1 Basic Customer Interconnection Requirements

The following is required to share a BGP session with Tata Communications' AS6453 network:

- The customer's gateway router must be using BGP version 4.

- The customer must use a Regional Internet Registry (RIR) assigned Autonomous System Number (ASN). In the case of the customer being multi-homed with AS6453 only, Tata Communications may assign the customer a private ASN.
- The customer must have its customer-routes registered with an instance of the Internet Routing Registry (IRR). The customer's 'as-set' must include the ASNs originating routes and for each route exported a relevant 'route' object is required. Details about this at the end of this document.
- The customer must export its customer-routes reasonably aggregated, and AS6453 will not accept any prefix that is longer than an IPv4 /24 or an IPv6 /48. An exception is the "black-holing" for denial-of-service mitigation described below.
- The customer may signal return path preference in AS6453 by means of BGP Multi-Exit Discriminator (MED). The customer may request AS6453 to send IGP sourced BGP MED.
- The customer must operate a responsive Network Operation and Abuse Prevention Center.

More details about the third and fourth items at the end of this document.

2 Policy Tuning Communities

Tata Communications provides its transit customers with a set of BGP community values each one of them corresponding to a request for a non-default policy in AS6453:

- Adjustment of the local preference of a customer's customer-route in AS6453.
- Influence of how AS6453 redistributes the customer's customer-routes to its peers.
- Black-holing of a customer's customer-route, as a means for denial-of-service attack mitigation.

2.1 Globally Significant Communities

Assigned by the Internet Assigned Numbers Authority (IANA), some BGP community values are considered well-known:

Well-known communities	
community	action
NO_EXPORT or LOCAL_AS	keep local to AS6453
NO_ADVERTISE	keep local to this router

The LOCAL_AS community value is known as NO_EXPORT_SUBCONFED in RFC1997, and since AS6453 is not implementing a BGP confederation architecture its operational value is the same as NO_EXPORT.

More recently RFC3765 introduced the well-known NOPEER community value. This value is not used by AS6453, but as you may see below an operationally equivalent value 65009:0 is provided.

2.2 Request for Local Preference Adjustment

In AS6453 the default local preference (LOCAL_PREF) value for customer-routes is 100, and for peer-routes it is 90 .

Along the lines of RFC1998, a Tata Communications customer may request other than the default local preference:

Adjust Local Preference	
community	action
6453:n, $n \in \{70, 80, 90, 110\}$	assign local preference n in AS6453

An example of how this local preference adjustment can be used, in cases where MED is not a sufficiently strong signal for managing the return path, is given in RFC1998.

2.3 Tuning of Redistribution to AS6453 Peers

In some multi-homing scenarios it may be useful to a customer having means to influence the way its customer-routes are re-distributed to AS6453 peers. The following BGP community values are recognized by AS6453 only to its ISP (Internet Service Provider) peers excluding CP (Content Providers), for such use:

Redistribution to Peers	
community	action
6500n:ASN, $n \in \{0, 1, 2, 3\}$	to peer ASN, prepend 6453 n times
65009:ASN	do not redistribute to peer ASN
6500n:0, $n \in \{1, 2, 3\}$	to all peers, prepend 6453 n times
65009:0	do not redistribute to any peer

ASN specific signals are processed before global ones, allowing a customer to build policies of the kind described in the example (4.1) below.

2.4 Mitigation of a Denial-of-service Attack

A customer may want to have AS6453 black-hole a subset of its customer-routes, as a means of denial-of-service attack mitigation:

Denial-of-service Attack Mitigation	
community	action
64999:0	black-hole this route

The route in question may be a host-route or any other subset of the customer's legitimate customer-routes. An illustration is given below (4.2).

See RFC3882 for an extensive discussion about ways of using BGP to block against denial-of-service attacks.

3 Information Communities

Among the BGP community values used internally in AS6453, a subset is "leaked" to transit customers as hints, that they may use as a basis for enforcing their import policies. Often, as indicated in an example below (4.1), such import policies are implemented to harmonize with policy tuning requests signaled to AS6453.

3.1 Geographical Origin of an AS6453 Route

The following table define a set of BGP community values indicating where (geographically) a route was imported into AS6453.

Geographical Origin of a Route

community	continent/sub-continent/site	site code
6453:1000	North America	
6453:1100	North America, East Coast	
6453:1102	Newark, 165 Halsey	njy
6453:1103	Ashburn, Equinix	aeq
⋮		
6453:1106	Montréal, Bonaventure	mtt
	Montréal, CANIX	w2c
	Montréal, IDS, CANIX 2	w3c
	Montréal, CANIX 3	w6c
6453:1107	Laurentides	lau
⋮		
6453:1110	Miami, NAP of Americas	mln
6453:1111	New York, Switch and Data	nto
6453:1112	New York, 32 AOA	nw8
6453:1113	New York, 60 Hudson	n0v
6453:1114	New York, 111 8th Avenue	n75
6453:1115	Manassas,EvoSwitch	vn5
6453:1116	New Jersey, Wall	wv1
6453:1117	New Jersey, Secaucus	eai
6453:1200	North America, Central North	
6453:1202	Chicago, Equinix	ct8
6453:1203	Toronto	ttt
6453:1204	Toronto, Equinix	tnk
6453:1205	Toronto, ORANO	t7g
6453:1206	Toronto, Neutral Data	t6n

Continued on Next Page...

Geographical Origin of a Route, continued...

community	continent/sub-continent/site	site code
6453:1207	Toronto, Cologix	tgs
6453:1300	North America, West Coast	
6453:1301	Palo Alto, Equinix	pdi
6453:1302	Los Angeles, Equinix	eql
6453:1303	Los Angeles, 1 Wilshire	laa
	Los Angeles, 1 Wilshire	lmr
	Los Angeles, Coresite	lvw
6453:1304	Hillsboro,Telx	eaq
6453:1305	San Jose, Equinix	sqn
6453:1306	Seattle	00s
6453:1307	Vancouver, West Hastings	vcw
6453:1308	Lake Cowichan	lcn
6453:1309	Santa Clara	sv1
6453:1310	Phoenix	un0
6453:1311	Las Vegas	w40
6453:1312	San Francisco	sf9
6453:1400	North America, Central South	
6453:1402	Atlanta, Telx	a56
6453:1403	Dallas, Equinix	dtx
	Dallas, Equinix	dt8
6453:1404	Denver, Confluent	ddv
6453:1405	Mcallen	xw7
6453:1406	Laredo	xw8
6453:1407	Dallas, Bryan St	xz3
6453:2000	Europe	
6453:2100	United Kingdom	
6453:2101	Londond, Telehouse North	ldn
6453:2102	London, Harbour Exchange	lhx
6453:2103	London, Redbus Interhouse	lrs
6453:2104	London, Stratford	l78
6453:2105	London, Telecity Redbus	lrt
6453:2106	London, Telehouse East	lvx
	London, Telehouse East	ly9
6453:2107	London, High Wycombe, Cressex	hw1

Continued on Next Page...

Geographical Origin of a Route, continued...

community	continent/sub-continent/site	site code
6453:2108	Somerset, Highbridge Cable Station	sv8
6453:2109	London, Slough Equinix	ld5
6453:2200	France	
6453:2201	Courbevoie, LDCom Co-location	pg1
6453:2202	Paris, Telehouse 1	pv0
	Paris, Telehouse 1	pv4
	Paris, Telehouse 2	pvu
6453:2203	Saint Denis, Equinix	pye
6453:2204	Marseille, Netcenter	wyn
6453:2205	Vitry-sur-Seine, Iliad DC3	vi8
6453:2300	Austria, Germany and Switzerland	
6453:2302	Frankfurt, InterXion	fr0
	Frankfurt, InterXion	fr1
6453:2304	Frankfurt, Ancotel	f2c
6453:2305	Frankfurt, Itenos	fnm
	Frankfurt, Itenos	fv0
6453:2306	Zurich	z3z
6453:2400	Benelux	
6453:2401	Amsterdam, SARA	ad1
	Haarlem, Evoswitch	hnn
6453:2402	Brussels, Interxion	b1d
6453:2403	Amsterdam, Telecity 2	av2
6453:2404	Amsterdam, Equinix-AM2	avu
6453:2500	Portugal and Spain	
6453:2502	Madrid, ESPANIX Co-location	mx2
6453:2503	Barcelona, Telvent	bjz
6453:2504	Derio, Cable Station	dvs
6453:2505	Madrid, InterXion	mdo
6453:2506	Madrid, Carrier House	wv6
⋮		
6453:2511	Lisbon, Prior Velho	pv9
6453:2512	Seixal	sz5
6453:2700	Norway and Sweden	
6453:2701	Oslo, Digiplex Co-location	os1

Continued on Next Page...

Geographical Origin of a Route, continued...		
community	continent/sub-continent/site	site code
6453:2702	Stockholm, TeleCity	stk
6453:2800	Italy	
6453:2801	Milan, INET	mlt
	Milan, Infracom	wi3
6453:2802	Rome, Caspur	rct
6453:2900	"Far East, Europe"	
6453:2901	Warsaw, Energis Polska	w1t
6453:2902	Warsaw, Netia	wzn
6453:2903	Moscow, MSK-IX	1m9
6453:2904	Istanbul	it5
6453:2905	Ankara	it6
6453:3000	Asia Pacific	
6453:3100	Hong-Kong	
6453:3101	Hong-Kong, Kowloon	kth
6453:3102	Hong-Kong, Mega-I	hk2
6453:3103	Hong-Kong, Equinix HK1	h71
	Hong-Kong, Equinix HK1	h81
6453:3104	Hong Kong, Billion Center	7b8
6453:3200	The Philippines and Guam	
6453:3201	Manila, Quezon City	qby
6453:3202	–	–
6453:3203	Guam	pv4
6453:3300	Australia	
6453:3301	Sydney, Mascot, Equinix	m3h
6453:3302	Sydney, Mascot, Equinix	1mh
6453:3303	Sydney,Global switch	0pp
6453:3400	Malaysia, Singapore and Thailand	
6453:3401	Kuala Lumpur, AIMS	kt1
6453:3402	Singapore, Equinix	svq
6453:3403	Singapore, Global Switch	svw
	Singapore, Global Switch	7sr
6453:3404	Bangkok	bk7
6453:3405	Singapore, TCX	ih4
	Singapore, TCX	ih5
6453:3406	Singapore, DRT	w42

Continued on Next Page...

Geographical Origin of a Route, continued...		
community	continent/sub-continent/site	site code
6453:3407	Singapore, Equinix SG2	40b
6453:3500	Indonesia	
6453:3501	Djakarta	–
6453:3600	Japan	
6453:3601	Tokyo, Equinix	tv2
6453:3602	Tokyo, Otemachi	ovc
6453:3603	Osaka, Equinix	e14
6453:3604	Chiba, EMI	kv8
6453:3700	Taiwan	
6453:3701	Taipei	tj5
6453:4000	Middle East and Africa	
6453:4100	Egypt	
6453:4101	Cairo	cyr
6453:4200	Saudi Arabia	
6453:4201	Riyadh	rsd
6453:4202	Jeddah	jsd
6453:4203	Riyadh	rmz
6453:4204	Fujairah	n71
6453:4300	South Africa	
6453:4301	Johannesburg	jso
6453:4302	Cape Town	klt
6453:4400	Kenya	
6453:4401	Nairobi	2n1
6453:4500	Tanzania	
6453:4501	Dar es Salaam	zia
6453:4502	Dar es Salaam	2n1
6453:6000	India	
6453:6100	Maharashtra	
6453:6101	Mumbai, LVSB	mlv
6453:6102	Mumbai, LVSB	w1u
6453:6200	Tamil Nadu	
6453:6201	Chennai, VSB	cfo
6453:6202	Chennai, VSB	cxr
6453:6300	Kerala	
6453:6301	Cochin	cov

3.2 Neighbor Type of an AS6453 Route

AS6453, having no transit provider, is simply classifying its routes into peer-routes and customer-routes.

Neighbor Type of Route	
community	type of route
6453:86	peer-route
6453:50	customer-route

This information is sent to AS6453 transit customers, along with the “geo community values” defined above (3.1).

4 Examples and Hints

BGP being a fairly rich protocol for implementing inter-domain traffic engineering, follows an example of how Tata Communications' enhanced services may help you along your way in a multi-homed scenario. An example of denial-of-service mitigation is also given.

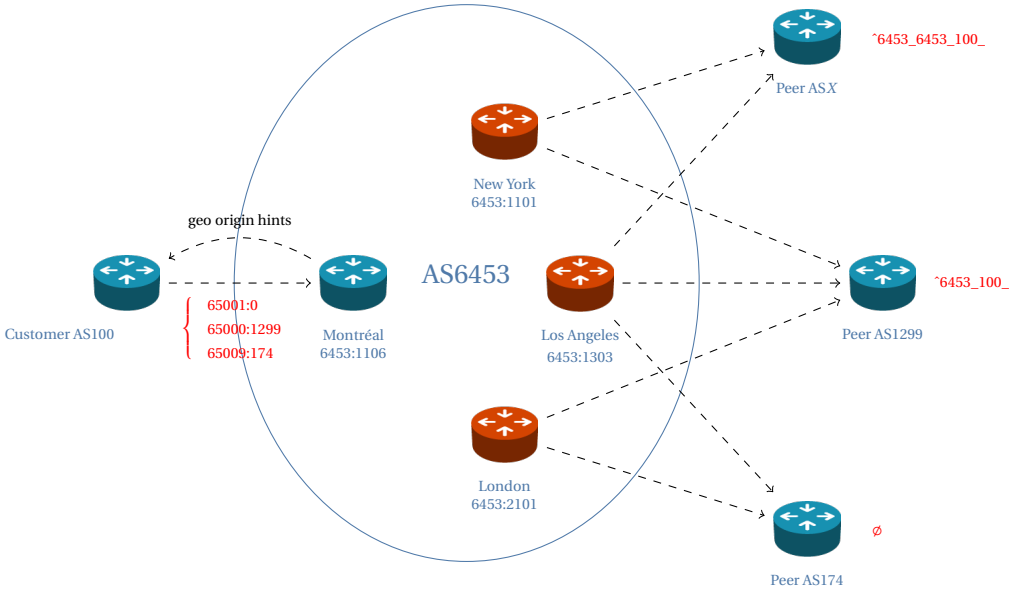
In the examples the signal processing is happening in the red routers, and the comments hopefully obvious by the context.

4.1 Tuning of Redistribution to AS6453 Peers

Suppose, for example, that you – AS100 – after having evaluated your multi-homing situation

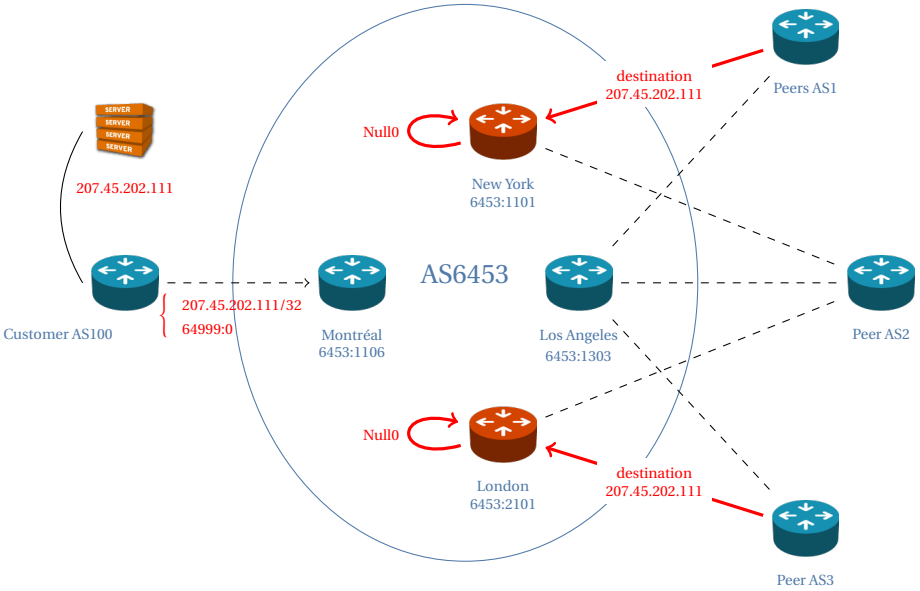
- want to have AS6453 prepend 6453 once to the AS_PATH when redistributing (some subset of) your customer-routes to its peers,
- except to AS1299 to whom you want AS6453 to redistribute (some subset of) your customer-routes as-is, and
- AS174 to whom you don't want AS6453 to redistribute (some subset of) your customer-routes.

Then you would attach the community values 65001:0, 65000:1299 and 65009:174 to (that subset of) your routes when exporting them to AS6453, with the following result:



4.2 Denial-of-Service Mitigation

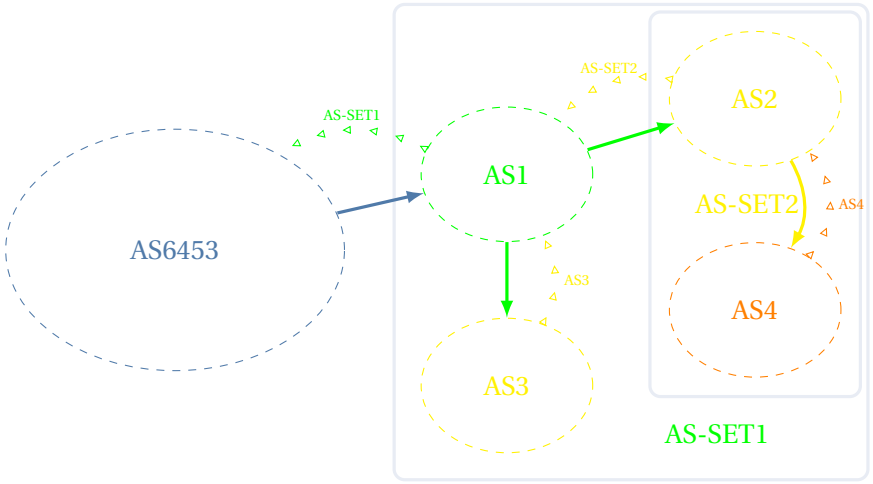
Suppose for this example that you – AS100 – find yourself subject to a denial-of-service attack targeting your server at 207.45.202.111 (being in your assigned and registred IP space). Then you may attach the BGP community value 64999:0 to your specific prefix 207.45.202.111/32, and inject this export along with your regular ones, with the following result:



Your prefixes injected for black-holing are kept local to AS453 (as members of NO_EXPORT). Please note that it is your responsability to remove the black-holing request, once you judge that the denial-of-service attack is over.

5 Customer's Routes and the Internet Routing Registry

Consider the following interconnection topology, where each link points from a provider to a customer and the "arrow edges" show intended route export.



AS6453 is (generally) establishing a BGP session only with customer's having been assigned an ASN by a RIR, visible in the form of a relevant `aut-num` object in the IRR:

```
aut-num:      AS1
as-name:      AS1
descr:        Company 1
<snip/>
source:       RIPE # Filtered
```

AS1 is providing transit service to AS2 and AS3; AS2 is providing transit to AS4. So, AS1 generally wants AS6453 to send them traffic to themselves and to their customers. In order to make their relation official they register it in the IRR:

```
as-set:       AS-SET1
descr:        Company 1 Customers
members:      AS1
members:      AS-SET2
members:      AS3
<snip/>
source:       RIPE # Filteredd
```

where

```
as-set:       AS-SET2
descr:        Company 2 Customers
members:      AS2
members:      AS4
<snip/>
source:       RIPE # Filteredd
```

is registred by AS2. Based on this declaration, a BGP policy import filter

```
ip as-path access-list n deny _65[0-9][0-9][0-9]_
ip as-path access-list n deny _64[6-9][0-9][0-9]_
ip as-path access-list n deny _645[2-9][0-9]_
ip as-path access-list n deny _6451[2-9]_
ip as-path access-list n deny _(209|701|1239|1299|1668|2914|3320|3356|3549|3561|5511|6461|7018)_
ip as-path access-list n permit ^1( .*|)$
ip as-path access-list n permit ^1 ([^ ]+ )*(1|2|3|4)( .*|)$
ip as-path access-list n deny .*
```

covering AS-SET1 = AS1 ∪ AS-SET2 ∪ AS3 = AS1 ∪ AS2 ∪ AS4 ∪ AS3 is applied to the AS6453 BGP session with AS1. In addition a prefix based import filter is applied to the session. Suppose, for example, that AS1 is advertizing, to AS6453, the prefix p with origin AS4. Then AS6453 is accepting p provided that $|p| \leq 24$ and that there is a `route` object in IRR such that

```
route:        r
descr:        Company 4
origin:       AS4
<snip/>
source:       RIPE
```

and $p \subseteq r$. Based on this declaration, a BGP policy import filter

```
ip prefix-list AS1 description AS-SET1
<snip/>
ip prefix-list AS1 permit r le 24
<snip/>
ip prefix-list AS1 deny 0.0.0.0/0 le 32
```

is applied to the AS6453 BGP session with AS1. Summing it up, a prefix p advertized by AS1 is accepted by AS6453 if and only if it is originated by AS1 itself or behind AS1 by one of AS1's registred customers (for example AS4), there is a route object with corresponding origin (AS4) registred for r and such that $p \subseteq r$, and $|p| \leq 24$. The same policy hold for import of IPv6 route prefixes, but with $|p| \leq 48$.

These import filters are updated four times a day. An update process is started at 02:00, 08:00, 14:00 and 20:00 UTC, but it can take several hours before the actual update hits a given router.

6 Further Information

Technical assistance and further information is provided via the Tata Communications' Global Customer Service Center (GCSC).