Using BGP Communities

Gaurab Raj Upadhaya
Nepal Research and Education Network

APRICOT 2010

BGP Communities

- RFC 1997

  - To facilitate and simplify the control of routing information this document suggests a grouping of destinations so that the routing decision can also be based on the identity of a group.

  - A community is a group of destinations which share some common property.

Other RFCs

- RFC 4360

    - BGP Extended Communities Attribute

- RFC 4384

    - BGP Communities for Data Collection

- Few others describing usage

    - Extension for 4byte ASN soon

Communities

- 32 bit field

- New format is two 16 bit fields separated by a colon

  - <asn>: < string>

  - 45170:64001; 42:1 etc

Nepal Research and Education Network

- Nepal Research and Education Network (NREN) has been established as a facilitator to support for advanced research and education network through the means of Information and Communication Technology.

- Extending national networking and connecting soon to Internet2/GEANT.

- Facilitation of different research activities

Base Principle

- Tag routes

- Use the Tags to filter outbounds

1st Design

- Based on tagging all prefixes centrally

  - Tag routes, setup per peer filters

  - Didn't scale so well

- 45170:1001 - NREN global

- 45170:1110 - NPIX routes

- 45170:1101 - NREN internal only

  - ip community-list 100 permit _45170:1..._

  - ip community-list 101 permit _45170:1..1_

- later added few more like..

  - 45170:1111 - R&E networks only

- It just started to get crazy and unscalable.

2nd Design

- Take a longer term view

  - Separate peer types and accordingly route types

    - Look at possible combinations

  - multiple tags per prefix

- Also looked at what others were doing, mainly at AARNet.

NREN Customers

- Members with 'commodity' Transit

- Members without 'commodity' Transit

- Members with and without NPIX

- Special RFC1918 networks

- Internal Routes/networks

NREN Upstreams/ Peers

- International R&E Network (TEIN / APAN)

- Potential bilateral R&E Networks (ERNET, CERNET)

- Local IXP Connection

- Internet Transit Providers (two)

Simplified Assumptions

- Everyone will receive R&E Routes, and on-net routes.

- For now, everyone will also receive NPIX routes - but may be optional in future

- We assign a /29 to each site without any PI space, and not route any PA space longer then /24 internally

Second Design Matrix

Session and Route Type	Peer Group	Inbound Actions		Outbound Actions		
		Community Tag	Route-map	Community match		Route-map
				List	String	
Internal Networks						
NREN Prefixes / iBGP sessions	iBGP-PEER	45170:10000	NREN-ROUTES	100	45170:…..	None
Subscriber Networks						
Members with Transit	TRANSIT-MEMBERS	45170:10000 45170:20xxx	MEMBER-TRANSIT-ROUTES-IN	101	45170:.0…	MEMBER-TRANSIT-ROUTES-OUT
Members without Transit	NOTRANSIT-MEMBERS	45170:20000	MEMBER-NON-TRANSIT-ROUTES-IN	102	45170:.00.. 45170:20…	MEMBER-NON-TRANSIT-ROUTES-OUT
NWP / SPECIAL-RFC1918	SPECIAL1918-PEERS	45170:11000	RFC1918-SPECIAL-ROUTES-IN	110	45170:..000	RFC1918-SPECIAL-ROUTES-OUT
Upstream Networks						
R&E Networks/ TEIN etc	EDU-PEERS	45170:10001	EDU-ROUTES-IN	104	45170:.000.	EDU-ROUTES-OUT
NPIX	NPIX-PEERS	45170:10010	NPIX-ROUTES-IN	105	45170:.0000	NPIX-ROUTES-OUT
Transit	TRANSIT	45170:10100	TRANSIT-IN	106	45170:10000	NREN-ROUTES
Default	TRANSIT	45170:10111	TRANSIT-IN	107	45170:10111	none

Expanding on the Matrix

- Let's look at the first one - iBGP Peers

  - Inbound (done on origination of iBGP routes)

    - 45170:10000 = iBGP Routes

  - Outbound

    - We used regular expressions matching

    - 45170: . . . . . = matches almost everything, as iBGP routers should see everything

    - The community tag is useful for informative purposes

Expanding the Matrix

- Members with Transit

  - Are easy, they are like your regular ISP customer

  - But tag inbound routes with 45170:10000 and 45170:20zzz, where zzz=location ID

  - Outbound only send routes with 45170:.0 . . .

    - Will see everything but the RFC1918 and special networks

      - As policy, we won't send RFC1918 out.

Expanding the Matrix

- Members without Transit
  - Inbound Tagged with 45170:20000
  - Outbound only see routes tagged with
    - 45170:.00.. or 45170:20zzz
    - will send all R&E, NPIX and NREN members

Design Wise

- The number of '1' in the string indicate how restrictive the incoming routes are

  - Members get tagged with string starting with '2', to differentiate from external routes

    - Community 45170: 10000 (iBGP Routes)

    - Community 45170: 20000 (Members without Transit)

    - Community 45170: 10001 (International R&E Nets)

Design wise

- the number of '0' in the string indicate how restrictive the out going routes are. i.e,

- Community 100 = 45170: . . . . . (iBGP Routes)

- Community 101 = 45170: .0... (Members with Transit)

- Community 102 = 45170: .00... (members without Transit)

- Community 104 = 45170: .000. (International R&E Nets)

Configuration snippets

```
    ip bgp-community new-format
ip community-list 100 permit _45170:....._
ip community-list 101 permit _45170:.0..._
ip community-list 102 permit _45170:.00.._
ip community-list 102 permit _45170:20..._
ip community-list 104 permit _45170:.000._
ip community-list 105 permit _45170:.0000_
ip community-list 106 permit _45170:10000_
ip community-list 107 permit _45170:10111_
ip community-list 110 permit _45170:..000_
```

Configuration Snippets

```
!member-transit-route
route-map MEMBER-TRANSIT-ROUTES-IN permit 10
set community 45170:10000 45170:20999
!
route-map MEMBER-TRANSIT-ROUTES-OUT permit 10
match community 101
!
!member-non-transit-route
!
route-map MEMBER-NON-TRANSIT-ROUTES-IN permit 10
set community 45170:20000
!
route-map MEMBER-NON-TRANSIT-ROUTES-OUT permit 10
match community 102
!
```

Benefits

- Much easier to mange new members

- Standardized configuration on all core/edge and customer end routers

- Routes tagged on the ingress as much as possible

- Can recognize routes based on tags

- Can co-relate v4 and v6 routes easily.

Plan

- Outbound communities for tagged routes accepted by the upstreams

    - DDoS mitigation and routing policy implementation

    - Signal routing policy changes and adapt to outages.

Lessons

- Plan carefully..

- Regular expression match makes things simple but also complex to design

  - Use all the 16bits you can play with.

- Keep good documentation

  - Deploy same setting on all router (use template)

Questions

- gaurab @ nren.net.np