**How to use
the BGP community attribute**

TWNIC OPM 3
Routing SIG
24 November, 2004
Taipei

Introduction

- Presenter
  – Miwa Fujii miwa@apnic.net

Overview

- Review of routing policy
  – Routing preference – current common practices
  – Impact to routing system load
- Allocations vs advertisements
- Review of community attribute
  – "no_export" usage
- Conclusion

Review of routing policy

- Routing preferences can be expressed via routing policy:
  - How can you direct incoming traffic to your site in a way that does not overload any single link?
  - How can you put VoIP traffic in a high quality, low delay link and put other traffic on cheaper links?
  - Minimise your costs by maximising your use of links with the lowest unit cost?
  - Set up primary and backup links with dynamic failover?

Routing preferences – common practices

- Advertise specifics as well as aggregates
- Use selective advertising of specifics to create preferred primary paths
- Use AS prepending to make relative AS path lengths a primary path selector

- These common practices do impact routing system load

Routing system load

- More specifics and AS path prepending:
  - are coarse-grained tools
  - impose a load on the global inter-domain routing system

- How big is this imposed overhead?

What is going on?

- Global routing table size is steadily growing
  – How fast is it growing?
  – What is the major cause of the growth?
  – What can we do to suppress the growth?

  Suppressing growth of the global routing table is in everyone's best interest! ☺

Global routing table growth
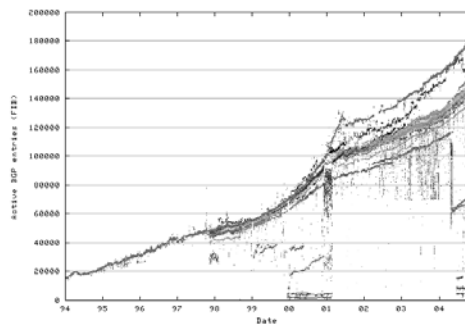
Reference: http://bgp.potaroo.net/ last updated 22/10/04

Allocations vs advertisements

- RIRs made 4506 IPv4 allocations (Jan 2003 – Feb 2004)
  – 3641 allocations announced
  – 865 allocations not yet announced
- 10904 routing advertisements used to span the 3641 allocations
  – 2938 advertisements precisely match the RIR allocation
  – 7966 advertisements are more specific advertisements of 1206 RIR allocations

Reference: "Allocation vs advertisements" presented by Geoff Huston at APNIC 17
http://www.apnic.net/meetings/17/docs/sigs/routing/routing-pres-huston-allocvsannouncement.pdf

Allocations vs advertisements

4506 allocations made by RIRs

Not announced
865

Announced
3641

10904 routing
Announcements

Advertisements
that matched the
RIR allocations:
2938
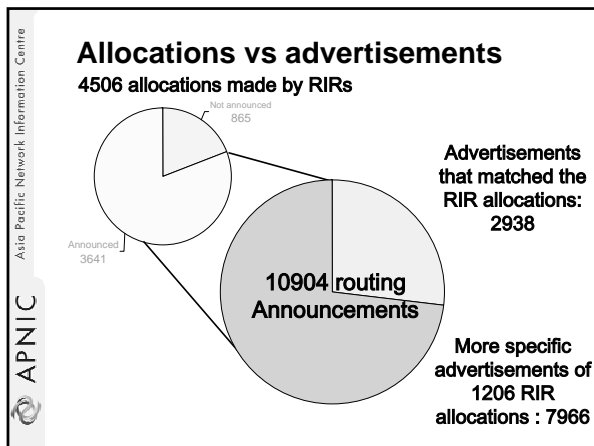
More specific
advertisements of
1206 RIR
allocations : 7966

APNIC

Analysis of statistics

- Advertising more specific /24 address prefixes within an allocated address block
  - This is the predominant form of advertising a split allocation block in fragments
  - Many of these more specific advertisements appear to be local
- One fifth of allocations are fragmented in this fashion
  - On average there are 6.6 additional advertisements of fragments of the address block

APNIC

Limits to routing

- The routing system does not have infinite capacity
- Too many routing entries will cause widespread routing failure
- How many is too many?
  - We don't know precisely
  - We will know when we see widespread routing failure, but by then it will be too late!

APNIC

4

What can we do?

- Look after the routing system
  - Use aggregate routing announcements wherever possible
  - Use explicit signalling to transmit policy preferences – community attributes
  - Limit the propagation of more specific routes to the local domain where they will have their effect – community attributes
- Effective use of the BGP community attribute
  - will reduce unnecessary announcements of fragmented prefixes

Review of community attribute

- A BGP route object (a unit of routing information carried by BGP) is composed of:
  - IP prefix value and prefix size
  - An AS-path attribute
  - Nexthop IP address
  - Community attribute (optional)

Note: A BGP route object is different to an IRR route object

Review of community attribute

- Community attribute:
  - is an optional component of a BGP route object
  - is a way for a route advertiser to signal to a route receiver some additional information about the BGP route object
  - may be bilateral or transitive
  - are intended to:
    - alter the way the receiver makes decisions about forwarding
    - alters the further propagation of the BGP route object
  - improve the capability of BGP speaker to describe the policy intention regarding distribution of routing information

BGP community attribute

- Optional transitive variable length attribute of a BGP route object
- 32 bit value
  - Format = asn:n
    - asn = 1 – 65,535
    - n = 1 - 65,535
  - Standards-defined values
    http://www.iana.org/assignments/bgp-well-known-communities
  - Commonly agreed values
    - Agreed value by a community
    - Amongst ISPs
    For example: http://info.us.bb.verio.net/routing.html

RFC 1997

RFC
1997

- Specifies the BGP community attribute
- Community
  - a means to specify a property of a BGP route object that affects the interpretation or manipulation of the BGP route object
- Communities are attached to a unit of BGP routing information by a BGP speaker
- Communities are received by the BGP listener:
  - Communities may be left attached
  - Stripped off such routing information
  - Translated to another community

Example of using communities

- A transit AS may allow its customers to selectively determine how a route is readvertised by the transit provider:
  - A customer can associate community values with each route object to limit the extent to which the route is readvertised by the transit provider
  - The customer controls the transit service

Community	Action
0	ReAdvertise to all regions
100	ReAdvertise in Asia
101	ReAdvertise in North America
102	ReAdvertise in Western Europe
103	Readvertise in Eastern Europe

Some well-known communities

- NO_EXPORT
  - All routes received carrying a community attribute containing this value MUST NOT be advertised to EBGP peers
- NO_ADVERTISE
  - All routes received carrying a community attribute containing this value MUST NOT be advertised to any BGP peers (internal and external)

How does "no_export" work?

AS1

192.200.1.0/24 Community = no_export

AS2

A

B

C

AS1 advertises 192.200.1.0/24 to AS2 with community attribute **no_export**

AS2 will propagate the route within AS2 but will not send this route to AS3 or any other external AS

D

AS3

Reference: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
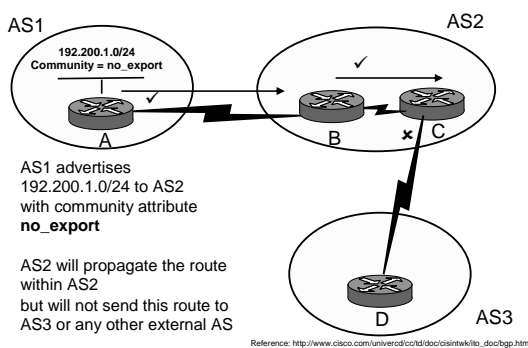
What do you need to do?

- ISP
  - Think what routing policies are relevant to you
  - Publishes multiple values of user-settable communities
  - Filters incoming route announcements to match them
  - Modifies route parameters accordingly
- Customer
  - Marks their announcements with wanted communities

Common use of communities

- Customer control of readvertisement
  - Regional-based transit
  - Peering control
- Customer control of preferences
  - Primary / Backup preference for routes
- Supplier information to customer
  - Where the route object was learned
  - Relationship to supplier (peer, customer, upstream)
  - Desired preference (primary / backup)

Further reading

- Cisco Internet Protocol Journal
  - http://www.cisco.com/warp/public/759/ipj_6-2/ipj_6-2_bgp_communities.html
- Using BGP Community Values to Control Routing Policy in Upstream Provider Network
  - http://www.cisco.com/en/US/tech/tk365/tk80/technologies_configuration_example09186a00801475b2.shtml

---

Conclusion

- Review your routes announcements
  - Are you announcing fragmented, more specific prefixes unnecessarily?
    - If so, consider use of no_export
  - Refer your router vendor's manual to learn how to use the community attribute
- Your efforts will help to slow down the growth of the global routing table
- And you will have better control of your external relationships with your IP suppliers and customers
  - it will benefit your business positively

Questions ?

Extra information

Example 1: community attribute "no_export" usage

Customer of AS1   **ISP**

AS 2   **AS 1**

Customer of AS2

AS 4

2 of /24 assignments out of AS2's address block

community NO_EXPORT

AS2 only announces aggregated
routes to AS1 (no need to announce AS4's two /24s to EBGP peers)

How can AS2 express such policy in RPSL?

Example of community attribute usage in IRR

```
aut-num: AS2
import:   from AS1 accept ANY
import:   from AS4 accept AS4
export:   to AS1 action community {NOT_EXPORT};
          announce ANY
          AND NOT {0.0.0.0/0}
          AND NOT fltr-bogons
export:   to AS4 announce ANY AND
          NOT {0.0.0.0/0} AND NOT fltr-bogons
```

How does "no_export" work?

192.200.1.0/24
Community = no_export

AS1   AS2

192.200.1.0/24   **192.200.1.0/24**
Community = no_export

3.3.3.3

A   B 3.3.3.1   C

AS1 advertises 192.200.1.0/24
to AS2 with community attribute
no_export

AS2 will propagate the route
within AS2 but will not send this
route to AS3 or
any other external AS

2.2.2.2

D   AS3

Reference: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
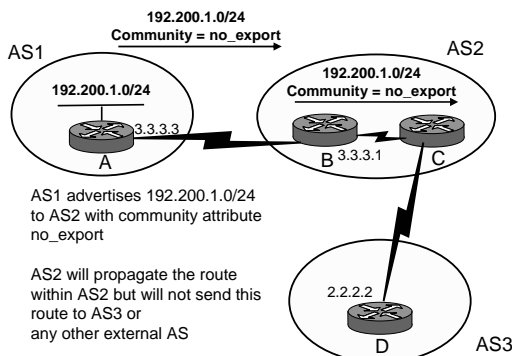
Example of Cisco command to use no_export

```
route-map communitymap
match ip address 1
set community no-export
```

- Even if we set the community attribute, this attribute will not be sent to neighbors by default.
    - In order to send the attribute to our neighbor we have to use the following:

```
neighbor {ip-address|peer-group-name} send-community
```

Example:
Router A
```
    router bgp 1
    neighbor 3.3.3.1 remote-as 2
    neighbor 3.3.3.1 send-community
    neighbor 3.3.3.1 route-map communitymap out
```

Reference: http://www.cisco.com/en/US/tech/tk365/tk80/technologies_tech_note09186a00800c95bb.shtml

Cisco command for no_export

```
RouterA#
router bgp 1
network 192.200.1.0
neighbor 3.3.3.1 remote-as 2
neighbor 3.3.3.1 send-community
neighbor 3.3.3.1 route-map SET_NO_EXPORT out
!
access-list 1 permit 192.200.1.0 0.255.255.255
route-map SET_NO_EXPORT permit 10
match ip address 1
set community no_export
```

APNIC

Reference: http://www.cisco.com/warp/public/459/aggregation.pdf

Juniper command for no_export

```
Bgp {
        local-as 1;
                        group test {
                        type external;
        passive;
                        import bgp-in;
                        peer-as 2;
                        neighbor 3.3.3.1;
                        }
        }
policy-options {

    policy-statement bgp-in {
            term 10 {
                    from {
                                protocol bgp;
                                community example;
                    }
                    then accept;
            }
            term 20 {
                    then reject;
            }
    community example members [ no-export] ;
```

APNIC

11