



How Big Networks Can Peer Regionally

Steve Gibbard

Packet Clearing House
APRICOT, 2007

Agenda

- › Brief lecture
 - › History
 - › Goals
 - › Configuration examples
 - › Results
- › Questions/Clarifications/Discussion



Importance of Sane Peering Policies

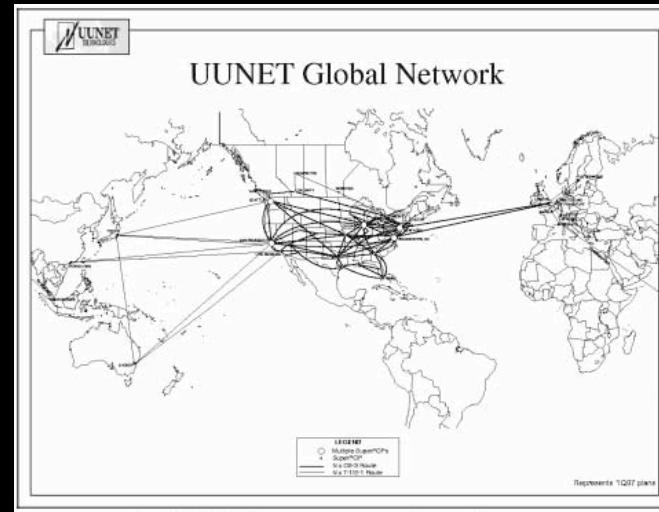
- › Historical peering policies:
 - › Some small networks peered openly.
 - › Other small networks bought transit and didn't bother.
 - › Big networks used peering policies to try to protect monopolies.
 - › Not peering with anybody smaller than them.
 - › Not peering with anybody who didn't cover at least their full coverage area.
 - › Not peering with anybody who didn't have the same sorts of customers.

Competitors Respond to Incentives

- › Some of the small networks really wanted to peer.
- › Small networks worked to meet big network requirements.
 - › Small networks needed to get lots of traffic.
 - › Small networks needed to copy the coverage areas of the bigger networks.
 - › Small networks needed to match traffic profiles of big networks.
 - › If big networks had lots of end users, small networks needed end users.
 - › If big networks had lots of content, small networks needed content.

Not Good for Big Networks

- › Networks were created with identical footprints and identical target markets.
- › Prices declined towards zero.
- › Many of them went bankrupt.
- › Those that haven't aren't generally making money.
- › The battles continue.

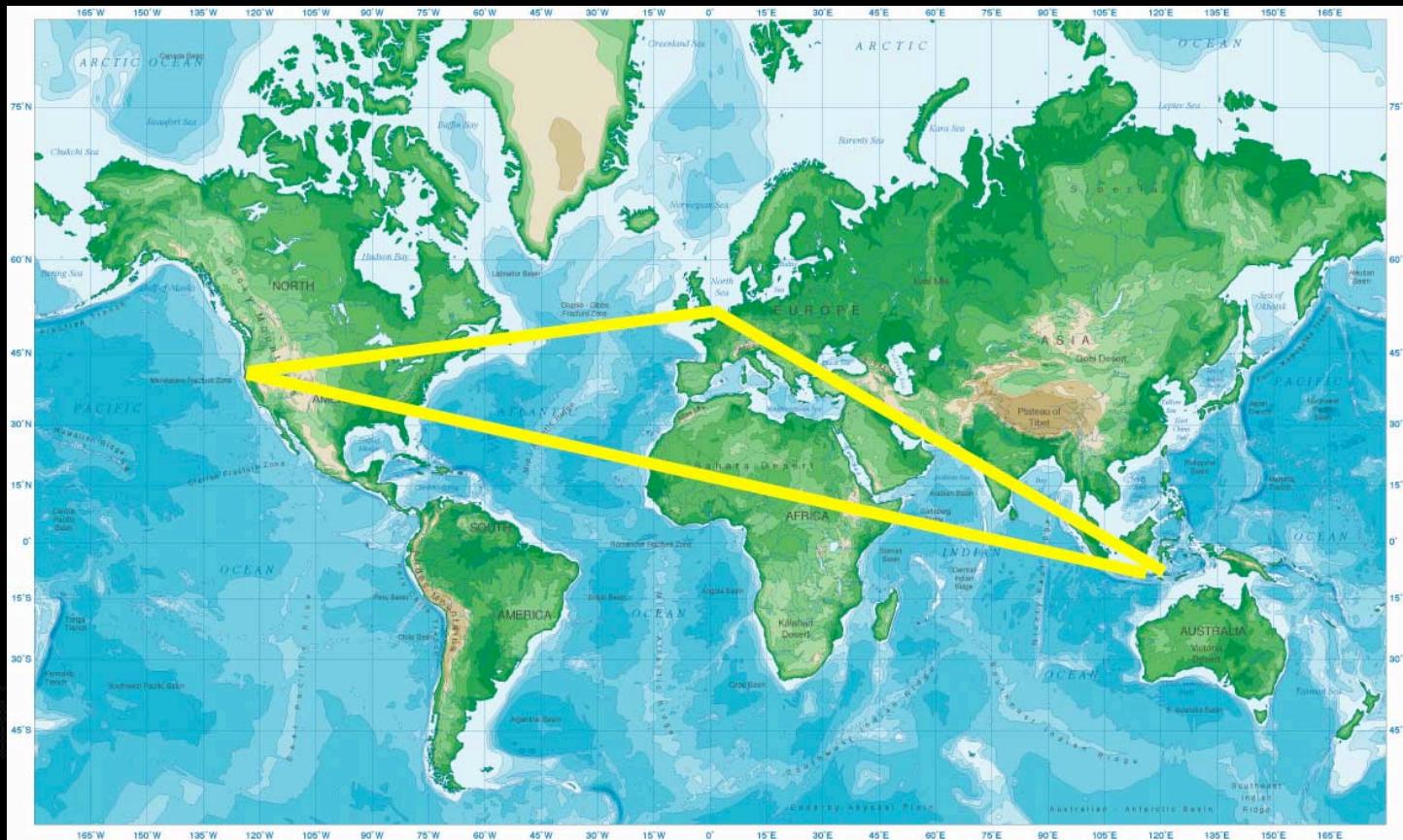




Incumbent Telco Alternative

- › No in-region peering:
 - › Attempt to protect monopoly.
 - › All domestic non-incumbent traffic gets hauled out of country, handed off, and hauled back in.
 - › Raises prices for everybody -- this is the point.
 - › Hurts reliability and performance.

Tromboning Traffic

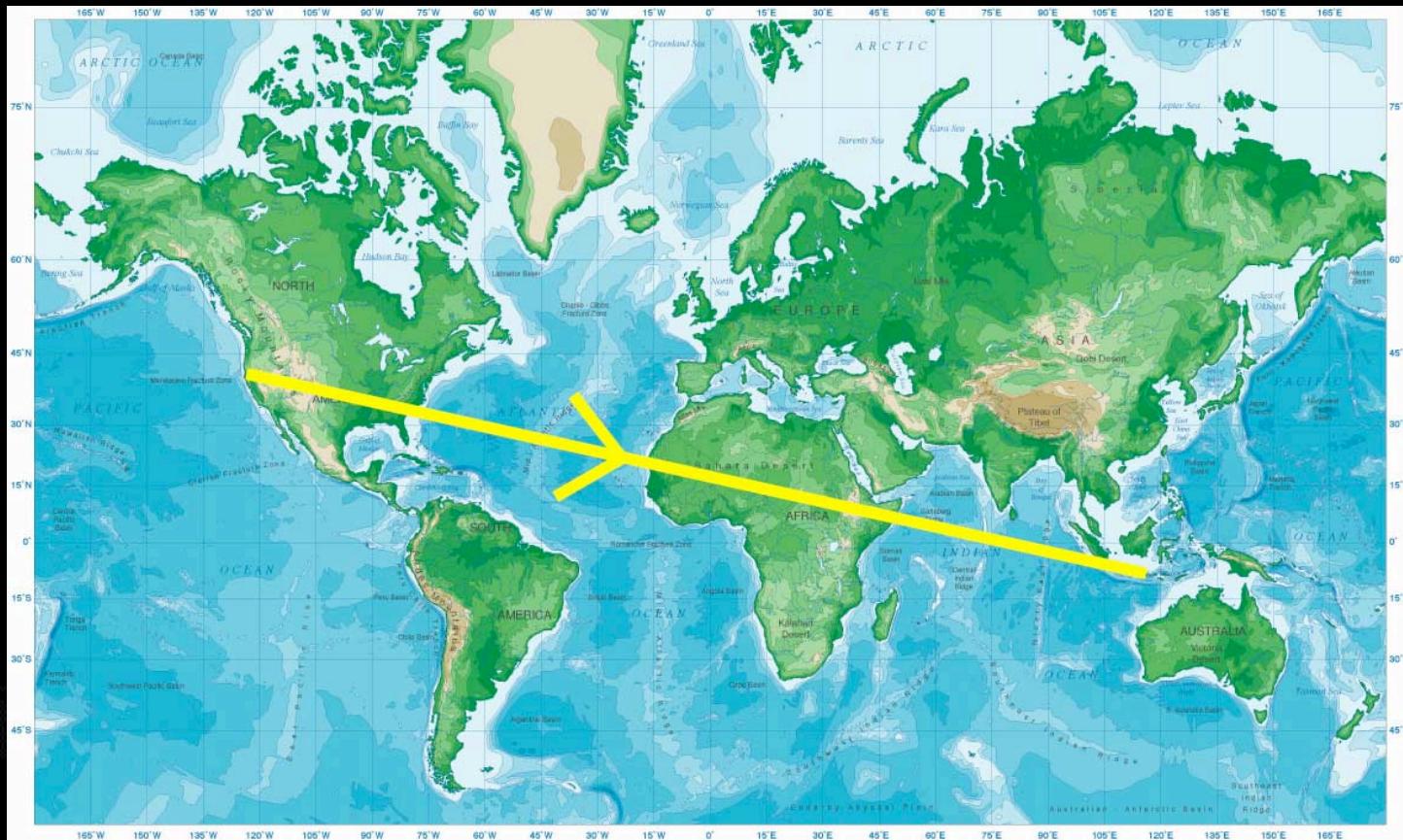




Creates Unhelpful Incentives

- › Incentives for competitors:
 - › Get big network's big customers on their networks, so they don't have to pay to get to them.
 - › Put small provider content in places where only big network has to pay.
 - › Charge customers more to talk to incumbent's customers than to others.
 - › Create alternate international capacity.

Off-Shore Hosting Traffic Flow

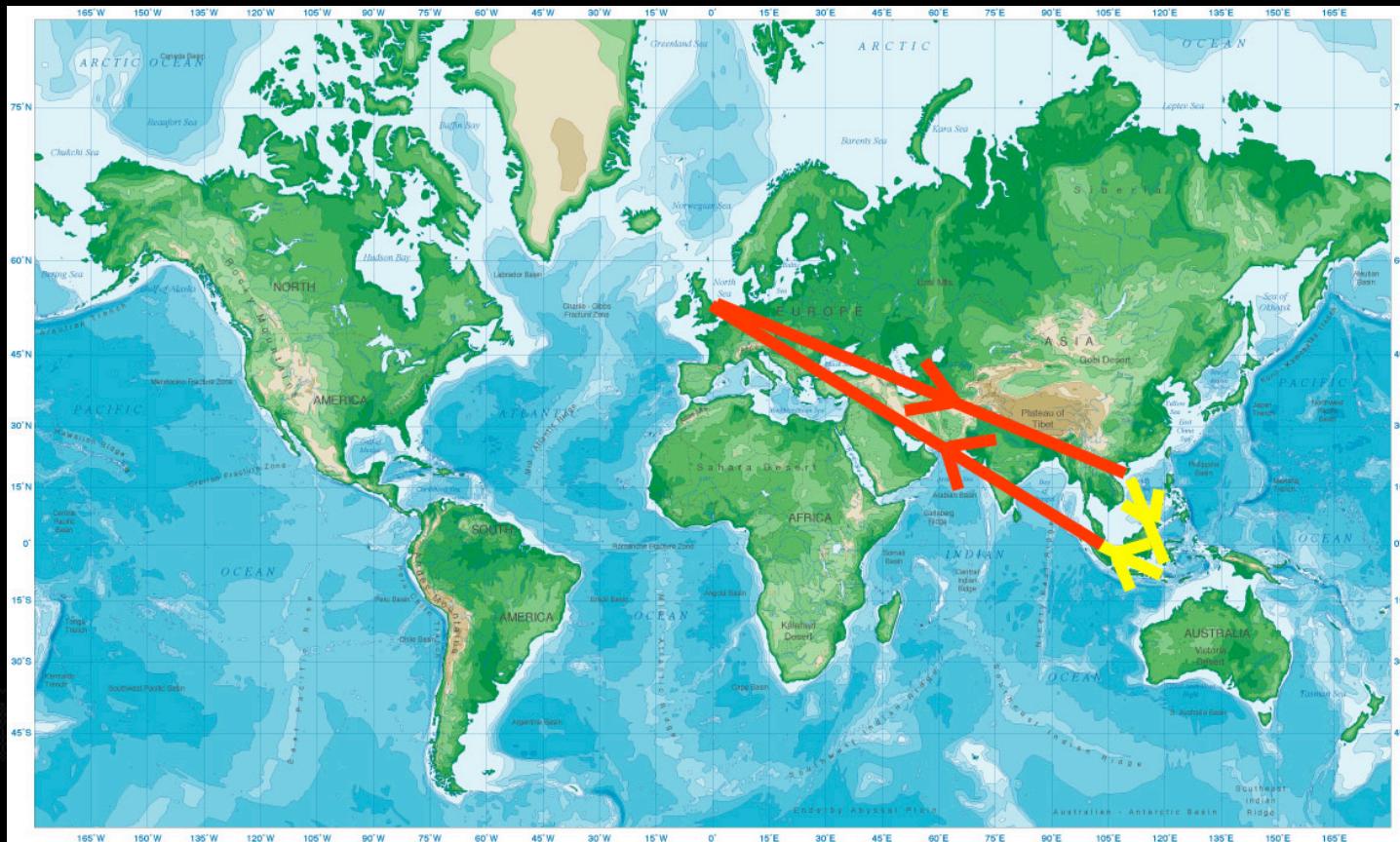




Some Limitations Make Sense

- › Big providers probably shouldn't be hauling traffic around the world for free for smaller providers.
- › Some volumes of traffic are too small to be worth dealing with.

Little/Big Network Traffic Flow





Elements of a Sane Policy

- › Low barriers to entry. Encourage competitors to stay small.
- › Keep local traffic local. Lowers costs and increases performance.
- › Encourage the really small to buy transit.
 - › Cost of loops and work of setting up sessions take care of this.
- › Don't carry long distance traffic for your peers if they won't carry it for you.

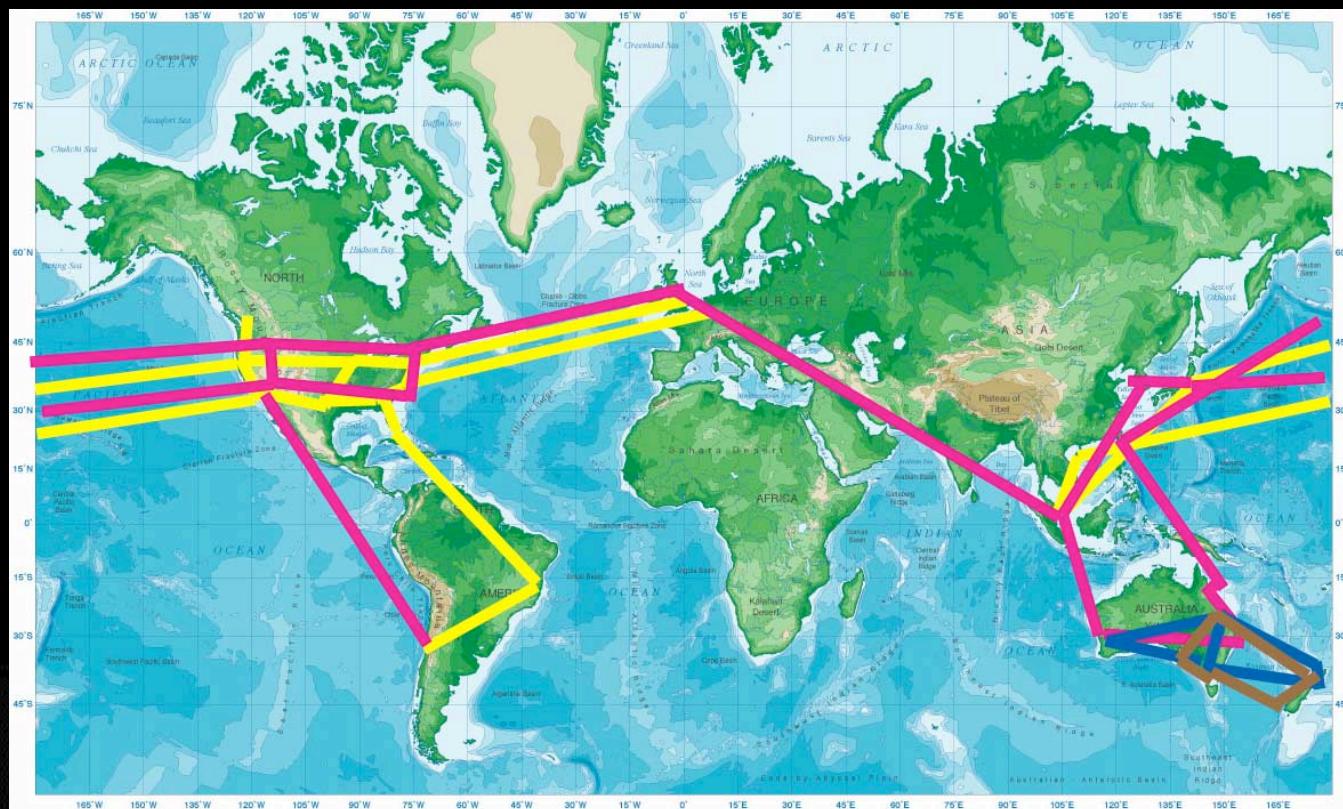


Parts of Status Quo Good

- › Global networks fully peer with global networks.
- › Networks within a region peer with each other.
- › Networks with adjacent territories peer at edges (sometimes).
- › Those who can't reach some network types by peering do so through transit.

Layers of Peering Illustration

Purple and yellow are natural full peers, as are blue and brown.





Peering Across Network Types

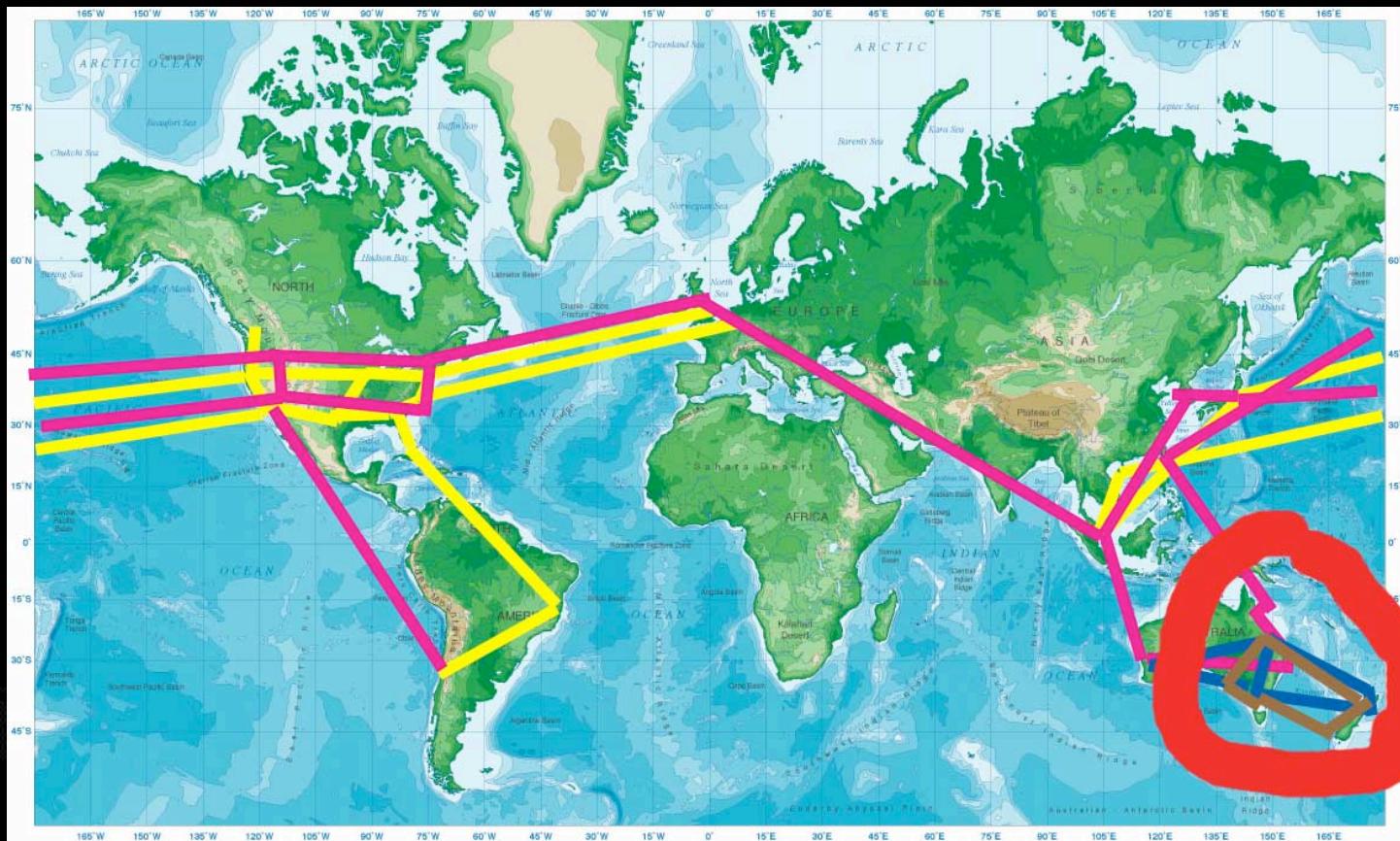
- › Limitations:
 - › Global networks won't exchange full routing tables with regional networks due to long distance traffic flows.
 - › Small networks still need transit from bigger networks to get farther away.



Answer: Regional Peering

- › Divide big networks up into regions.
- › Tag routes with communities.
- › Peer with regional networks.
- › Exchange only routing data from overlapping regions.

Regional Peering Illustration



Single vs. Multiple AS

- Multiple AS:
 - Some networks do this with multiple ASes.
 - UUNet, as an example.
 - Separate networks for separate regions -- really simple.
 - Not much granularity.
- Single AS:
 - Route-maps more complicated.
 - Allows more granularity.
 - Focus here on single AS.

Concepts

- › BGP Communities -- used to label (or “tag”) routes.
- › Route-maps -- filters and manipulates routes.
- › Communities set by route-maps.
- › Communities interpreted by community-lists.
- › Local-preference -- determines route priority.



Example Community Numbers

- › Region 1 peers: 65535:11000
- › Region 2 peers: 65535:12000
- › Region 3 peers: 65535:13000
- › Region 1 internal: 65535:21000
- › Region 2 internal: 65535:22000
- › Region 3 internal: 65535:23000
- › Global peers: 65535:3x000
- › Transit: 65535:4x000
- › Customer: 65535:5x000



Example: Internal Route-Map for Sourcing Routes

```
router bgp 65535
    network 192.168.0.0 mask 255.255.255.0 route-map REGION-1-SOURCE

route-map REGION-1-SOURCE permit 10
    set community 65535:21000
    set local-preference 400
```



Example: Inbound Route-Map, Regional Peers

```
route-map REGION-1-PEER-IN  
  set community 65535:11000  
  set local-preference 300
```



Example: Outbound Route-Map, Regional Peers

```
route-map REGION-1-PEER-OUT  
match community 101
```

```
ip community-list 101 permit ^65535:21...$
```



Example: Inbound Route-Map, Global Peer

```
router bgp 65535
  neighbor GLOBAL-PEER route-map GLOBAL-PEER-IN in

  route-map GLOBAL-PEER-IN permit 10
    set community 65535:31000
    set local-preference 300
```



Example: Outbound Route-Map, Global Peer

```
router bgp 65535
    neighbor GLOBAL-PEER route-map GLOBAL-PEER-OUT out

    route-map GLOBAL-PEER-OUT permit 10
        match community 103

    ip community-list 103 permit ^65535:2....$
```



Example: Internal Inter-Region Route-Map

```
router bgp 65535
    neighbor 192.168.5.2 remote-as 65535
    neighbor 192.168.5.2 route-map REGION-1-IBGP-IN in
    neighbor 192.168.5.2 desc IBGP peer

    route-map REGION-1-IBGP-IN permit 10
        match community 104

        community-list 104 permit ^65535:11...$  
Community-list 104 permit ^65535:3....$
```



BGP Table In-Region

```
router.region1>sh ip bgp 192.168.0.0
BGP routing table entry for 192.168.0.0/19, version 34338463
Paths: (4 available, best #1, table Default-IP-Routing-Table)
Multipath: eBGP
    Advertised to update-groups:
        11
        26228
            198.32.176.78 from 198.32.176.78 (10.0.101.2)
                Origin IGP, localpref 300, valid, external, best
                Community: 65535:21000
        6461 26228
            198.32.176.11 from 198.32.176.11 (64.125.0.137)
                Origin IGP, localpref 300, valid, external
                Community: 65535:31000
        2914 26228
            129.250.10.57 from 129.250.10.57 (129.250.0.45)
                Origin IGP, localpref 300, valid, external
                Community: 65535:31000
        6453 6461 26228
            63.243.133.21 from 63.243.133.21 (64.86.80.253)
                Origin IGP, localpref 50, valid, external
                Community: 65535:41000
```



BGP Table: Out of Region

```
router.region2>sh ip bgp 192.168.0.0
BGP routing table entry for 192.168.0.0/19, version
10600365
Paths: (2 available, best #1, table Default-IP-Routing-
Table)
Multipath: eBGP
Advertised to update-groups:
    11
    6461 26228
        198.32.160.22 from 198.32.160.22 (209.249.254.55)
            Origin IGP, localpref 300, valid, external, best
Community: 65535:41000
    6453 6461 26228
        63.243.133.21 from 63.243.133.21 (64.86.80.253)
            Origin IGP, localpref 50, valid, external
Community: 65535:41000
```



Regional-peer Outbound, R1

```
router.region1>sh ip bgp 10.5.3.0
BGP routing table entry for 10.5.3.0/24, version 41140050
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Multipath: eBGP
Advertised to update-groups:
      1          4          6          11
32978
204.61.210.11 from 204.61.210.11 (204.61.210.11)
  Origin IGP, localpref 350, valid, external, multipath
Community: 65535:51000
```

```
router.region1>sh ip bgp 10.20.40.0/24
BGP routing table entry for 10.20.40.0/24, version 41702869
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Multipath: eBGP
Advertised to update-groups:
      1          4          6          11
38052
204.61.210.12 from 204.61.210.12 (204.61.210.12)
  Origin IGP, localpref 350, valid, external, best
Community: 65535:52000
```

Regional-peer Outbound, R1

```
router.region1>sh ip bgp route-map REGION-1-PEER-OUT
BGP table version is 42032745, local router ID is 204.61.209.254
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
      S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

      Network          Next Hop          Metric LocPrf Weight Path
*  10.5.3.0/24    204.61.210.11        0 32978 i
*>                    204.61.210.10        0 32978 i
```

```
router.region1#sh ip bgp neigh 198.32.176.33 adv
BGP table version is 42032546, local router ID is 204.61.209.254
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
      S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.5.3.0/24	204.61.210.10	350	0	32978	i



Optional: Multiple-Region Peer

```
route-map REGION-1-2-PEER-IN
set community 65535:11000 65535:12000
    set local-preference 300
```

```
route-map REGION-1-PEER-OUT
match community 101
```

```
ip community-list 110 permit ^65535:21000$
Ip community-list 110 permit ^65535:51000$
Ip community-list 110 permit ^65535:22000$
Ip community-list 110 permit ^65535:52000$
```



Questions/Discussion?

Steve Gibbard

Packet Clearing House

<http://www.pch.net>

scg@pch.net