Multihomed BGP Configurations
Alvaro Retana
Cisco IOS®—Deployment and Scalability

Agenda

• General Considerations
• Multihomed Networks
• Best Current Practices
The Basics

General Considerations

General Considerations

Agenda

• What is a multihomed network?
  Motivation and Drivers

• BGP Route Selection Process
  It’s effect on multihomed policy
**Stub Network**

- No need for BGP
  - ISP advertises the stub network
  - Policy confined within ISP policy
- Default to the border

**Load-Sharing—Single Path**

Router A:
```
interface loopback 0
ip address 10.60.0.1 255.255.255.255
!
router bgp 100
neighbor 10.200.0.2 remote-as 200
neighbor 10.200.0.2 update-source loopback0
neighbor 10.200.0.2 ebgp-multi-hop 2
```
Load-Sharing—Multiple Paths from Same AS

Router A:
router bgp 100
neighbor 10.200.0.1 remote-as 200
neighbor 10.300.0.1 remote-as 200
maximum-paths 6

What Is Multihoming?

• Connecting to two or more ISPs to increase:

  Reliability -- Ensure external connectivity even if there’s a upstream failure

  Performance -- Better paths to common Internet destinations (policy settings)

  Load Sharing -- Distribute traffic between the different connections
Multihomed Network

- Many situations possible
  - Separate connections to the same ISP
  - Secondary for only backup
  - Load sharing between primary and secondary
  - Selectively use different ISPs

BGP Route Selection Process

- Only **ONE** path is considered the best at any time
- Policy can be set per router to guarantee loop-free forwarding, the policy should be consistent across the AS
Basic Decision Algorithm

Consider only (synchronized) routes with no AS loops and valid next-hop, then prefer:

- Highest WEIGHT
- Highest LOCAL PREFERENCE
- Locally Originated (e.g., network/aggregate)
- Shortest AS-PATH
- Lowest ORIGIN (IGP < EGP < incomplete)
- Lowest MED
- EBGP
- IBGP
- Lowest IGP METRIC to next-hop
- Lowest router-id

AS_PATH

- A list of AS to be traversed to reach a prefix
  - 1880 1883 (AS_SEQUENCE)
- Path includes one or more members of a set
  - {1880, 1881, 1882} (AS_SET)
- Used for loop detection!
**AS_PATH**

- Configuration (rtr B):
  
  ```
  router bgp 690
  neighbor x.x.x.x remote-as 666
  neighbor x.x.x.x route-map prepend out
  
  route-map prepend permit 10
  match as-path 2
  set as-path prepend 690 690
  
  ip as-path access-list 2 permit ^$
  ```

**Multi-Exit Discriminator (MED)**

- Indication (to external peers) of the preferred path into an AS used in multiple entry AS non-transitive

- Compared only for routes received from the same AS
  
  unless `bgp always-compare-med` is enabled

- Lower MED value is more preferable (default = 0)
**MED**

- Configuration (rtr B):
  
  ```
  router bgp 1755
  neighbor x.x.x.x remote-as 1880
  neighbor x.x.x.x route-map set_MED out
  
  route-map set_MED permit 10
  match as-path 2
  set metric 2
  
  ip as-path access-list 2 permit _690$
  ```

  Set MED = 2 for all routes originated in AS 690.

**MED & IGP Metric**

- `set metric-type internal`

  enable BGP to advertise a MED which corresponds to the IGP metric values

  changes are monitored (and re-advertised if needed) every 600s

  `bgp dynamic-med-interval <secs>`
MED - Example

C Chooses the shortest path to the destination.

Deterministic MED

- By default, the prefixes are compared in order of arrival
  it may result in inconsistent decisions
  use bgp deterministic-med
  the bestpath is recalculated as soon as the command is entered
**LOCAL_PREF**

- Indication of preferred path to exit the local AS
- Global to the local AS
- Paths with highest LOCAL-PREF are most desirable (default = 100)

```
bgp default local-preference value
```

**LOCAL_PREF (Cont.)**

- Configuration (rtr A):
  
  ```
  router bgp 109
  neighbor x.x.x.x remote-as 1880
  neighbor x.x.x.x route-map foo in
  route-map foo permit 10
  match as-path 2
  set local-preference 120
  
ip as-path access-list 2 permit ^1880_
  ```

- Set LOCAL_PREF = 120 for all the routes received from AS 1880.
Community Attribute

- Used to group destinations to help scale policy application
- Each prefix can belong to multiple communities
- Not propagated by default

```
neighbor ip-address send-community
```

Problem: Scale Routing Policy
Solution: COMMUNITY

Communities:
1:100—customer routes
1:80—peer routes

CORE

Customer A
Full Routes

Set Community
1:80

Match Community
1:100

Customer B
Customer Routes

Set Community
1:100

Match Community
1:100

Peer A
Double the Fun!

Multihomed Networks

Multihomed Networks Agenda

- Types of Multihoming
  Outbound Traffic
- Address Allocation Issues
  Inbound Traffic
Typical Multihomed E-Commerce Network

ISP1

ISP2

Border Routers (BGP and IGP)

Internal Routers (BGP and/or IGP)

Types of Multihoming

• Three common cases:
  Default from all providers
  Customer+Local routes from all providers (partial routes)
  Full routes from all providers
Default from All Providers

- Low memory/CPU solution
- Provider sends BGP default exit point selected by IGP metrics to reach default
- All the local routes are advertised to the provider
  inbound path decided by Internet influence using AS-path prepend

C Chooses Lowest IGP Metric to Default
Default from All Providers

AS 400

AS 200

AS 100

160.10.0.0/16

AS 300

Closest exit used - “watershed effect”

Partial Routes from All Providers

- Medium memory and CPU
- Best path is usually determined by the shortest AS-path
- Use local-preference to override based on prefix, as-path, or community
- IGP metric to default used for all other destinations
Partial Routes from All Providers

Provider AS 200

Provider AS 300

Customer AS 100
160.10.0.0/16

IP prefix-list AS100 permit
16.10.0.0/16
route-map AS300in permit 10
match ip address prefix-list AS100
set local-preference 800

C Chooses Shortest AS Path

C Chooses Highest Local-Preference
Full Routes from All Providers

- Higher memory/CPU solution
- Reach all destinations by selecting the best path—usually shortest AS path
- May still manually tune using local-pref and as-path/community/prefix matches
Controlling Inbound Traffic?

• Very difficult due to lack of transitive metric.
• Dependency on ownership of the IP address space.
• Can divide outgoing updates across providers, but what happens to redundancy?

Address Allocation Cases

• Provider Derived Address Space
• Provider Independent Address Space (portable block)
• Address Space owned by other Providers
• Address Space owned by both Providers
A multihomed network cannot use a provider’s connection unless the provider advertises the network’s routes to the rest of the Internet.

Praveen Akkiraju,
Multihoming Strategies for Internet Connectivity

Provider Derived Address Space

- Simple addressing scheme
- Major providers own CIDR blocks.
  The assign part of it to their customers.
  The provider may aggregate the customer routes back into its CIDR block.
- Must use with provider’s policy
Aggregation

- Example:
  router bgp 1880
  network 200.200.1.0
  aggregate-address 200.200.0.0 255.255.0.0 as-set

Provider Derived Address Space

The more specific route of taken back.
Portable Addressing

- The address space is owned by the local network
- Policy set locally
- Can’t be aggregated by any provider
- Prefix must be “leaked”
- May be filtered upstream

Portable Addressing

Any path may be chosen by the remote AS.
Using AS-PATH prepend

Customer AS 100

Provider AS 200

EE

BB

CC

AA

DD

210.210.1.0/24 300 400 400
210.210.1.0/24 200 400 (best)

Provider AS 300

Address Space owned by other Providers

- One of the providers may aggregate, but the other can’t
- Policy may be set locally restricted if aggregated
The more specific route of taken back.

- Complex addressing scheme needed
- The advertisements must be leaked by both providers
- Symmetry is not easily achievable
  must leak all routes on all providers
  use as-path prepend
Address Space owned by both Providers

Provider AS 200
- 200.200.1.0/24
- 220.220.1.0/24

Provider AS 300
- 200.200.0.0/16

Provider AS 400
- 200.200.0.0/16
- 200.200.1.0/24

Be a Good Neighbor
Best Current Practices
Best Current Practices

Agenda

- BGP Specific Issues
  - MD5 Authentication
  - Route Dampening
- General Filtering
  - Use the same filters for packet forwarding and received updates

MD5 Authentication

- Used in BGP to authenticate the TCP session
  - the contents of the session are not authenticated
- Configure authentication to avoid establishing sessions with rogue peers

```bash
router bgp 109
eighbor 1.1.1.1 remote-as 1
neighbor 1.1.1.1 password xxxx
```
Route Flap Dampening

- Route flaps ripple through the entire Internet
  up and down of path
  change in attributes
- Constant changes in the BGP table wastes CPU and bandwidth
- Objective: reduce the scope of route flap propagation
Selective Dampening

- Selective dampening based on
  - AS-PATH
  - Community
  - Prefix
- Variable dampening

Dampening Configuration

- `bgp damping <halflife-time> <reuse> <suppress> <maximum-suppress-time>`

- Example:
  - `router bgp 109`
  - `bgp dampening route-map SELECTIVE_DAMPENING`!
  - `access-list 110 permit ip any 255.255.255.0 0.0.0.255`
  - `access-list 111 permit ip any any`!
  - `route-map SELECTIVE_DAMPENING permit 10`
    - `match ip address 110`
    - `set dampening 30 125 2000 120`!
  - `route-map SELECTIVE_DAMPENING permit 20`
    - `match ip address 111`
    - `set dampening 25 750 2000 45`!
RFC 1918 Filtering

router bgp 109
neighbor 1.1.1.1 prefix-list rfc-1918 in

! prefix-list rfc-1918 deny 10.0.0.0/8 le 32
prefix-list rfc-1918 deny 192.168.0.0/16 le 32
prefix-list rfc-1918 deny 172.16.0.0/12 le 32
prefix-list rfc-1918 permit 0.0.0.0/0 le 32

ISP Network

Ingress to Internet

Customer Network

Limit the Updates

router bgp 1
neighbor 1.1.1.2 prefix-list customer in
neighbor 1.1.1.2 maximum-prefix 500

! prefix-list customer permit 200.200.1.0/24
prefix-list customer deny 0.0.0.0/0 le 32

ISP Network

Customer Network: 200.200.1.0/24

router bgp 109
neighbor 1.1.1.1 prefix-list local in
neighbor 1.1.1.1 filter-list 1 out

! prefix-list local deny 200.200.1.0/24 le 32
prefix-list local any 0.0.0.0/0 le 32

Ip as-path access-list 1 permit ^$
References - Registration

- American Registry for Internet Numbers (ARIN)
  http://www.arin.net/
- The Internet Corporation of Assigned Names and Numbers (ICANN)
  http://www.icann.org

References - BGP

- Requests for Comments (RFC)
  1918  Address Allocation for Private Internets
  1998  An Application of the BGP Community Attribute in Multi-home Routing
  2008  Implications of Various Address Allocation Policies for Internet Routing
2260  Scalable Support for Multi-homed Multi-provider Connectivity

2827  Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing

• Enabling Enterprise Multihoming with Cisco IOS Network Address Translation
  

• Internet Routing Architectures. Sam Habali
Empowering the Internet Generation℠