Troubleshooting BGP in Large IP Networks

Session 2210
BGP in Large Scale Networks

- Scalable
- Stable
- Simple

Avoid the Problem in the First Place

- Use **simple** configurations
  maintain a consistent policy throughout the AS

- Promote **stable** networks
  nail-down your routes
  use loopback interfaces

- Grow into your network
  use peer-groups and RRs for **scalability**
Agenda

- Basic Tools
- Peer Establishment
- UPDATE Exchange
- Selection Algorithm
- Route Reflectors

Tool Time

Basic Tools
BGP Troubleshooting Tools

- *show* commands
- *debug* output
- Log messages

**show Commands**

```
router#show ip bgp ?
A.B.C.D            IP prefix <network>/<length>, e.g., 35.0.0.0/8
A.B.C.D            Network in the BGP routing table to display
clid-only          Display only routes with non-natural netmasks
community          Display routes matching the communities
community-list     Display routes matching the community-list
dampened-paths     Display paths suppressed due to dampening
filter-list        Display routes conforming to the filter-list
flap-statistics    Display flap statistics of routes
inconsistent-as    Display only routes with inconsistent origin ASs
neighbors          Detailed information on TCP and BGP neighbor connections
paths              Path information
peer-group         Display information on peer-groups
quote-regexp       Display routes matching the AS path "regular expression"
regexp             Display routes matching the AS path regular expression
summary            Summary of BGP neighbor status
|                   Output modifiers
<cr>
```
**show Commands (Cont.)**

```
router#show ip bgp neighbors x.x.x.x ?
adverised-routes          Display the routes advertised to a BGP neighbor
 dampened-routes           Display the dampened routes received from neighbor
 flap-statistics           Display flap statistics of the routes learned from neighbor
 paths                     Display AS paths learned from neighbor
 received                  Display information received from a BGP neighbor
 received-routes          Display the received routes from neighbor
 routes                    Display routes learned from neighbor
 |<cr>
```

**The BGP Table**

```
router#show ip bgp
BGP table version is 9, local router ID is 7.72.6.1
Status codes: s suppressed, d dampened, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

 Network     Next Hop          Metric     LocPrf  Weight  Path
<= 3.0.0.0    0.0.0.0          0          32768 i
<= 5.0.0.0    0.0.0.0          0          32768 i
<= 6.0.0.0    6.72.6.2        4294967294             0 2 i
* i           6.72.6.2        4294967294    100      0 2 i
* i           6.72.6.2        4294967294             0 2 i
* i           6.72.6.2        4294967294             0 2 i
* i           6.72.6.2        4294967294             0 2 i
* i           6.72.6.2        4294967294             0 2 i
```
The BGP Table (Cont.)

```
router#show ip bgp 6.0.0.0
BGP routing table entry for 6.0.0.0/8, version 2
Paths: 2 available, best #1
  Advertised to non peer-group peers:
    7.25.14.4 7.72.6.3 7.75.7.1
  2
    6.72.6.2 from 6.72.6.2 (7.72.6.2)
      Origin IGP, metric 4294967294, localpref 100, valid, external, best
  2
    6.72.6.2 from 7.75.7.1 (7.75.7.1)
      Origin IGP, metric 4294967294, localpref 100, valid, internal
```

show ip bgp Summary

```
router#show ip bgp summary
BGP router identifier 7.72.6.1, local AS number 1
BGP table version is 9, main routing table version 9
8 network entries and 12 paths using 1176 bytes of memory
3 BGP path attribute entries using 144 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
BGP activity 8/0 prefixes, 12/0 paths

Neighbor    V AS M sgRcvd M sgSent Tb M Ver InQ OutQ Up/Down State/P St Rd
  6.72.6.2  4  2  6885  6882  0  0  0  4d18h  4
  7.25.14.4  4  3  6882  6883  0  0  0  4d18h  0
  7.72.6.3  4  1  6880  6886  9  0  0  4d18h  0
  7.75.7.1  4  1  6884  6885  9  0  0  4d18h  4
```
show ip bgp neighbors

router#show ip bgp neighbors 6.72.6.2
BGP neighbor is 6.72.6.2, remote AS 2, external link
Index 1, 0 flbit:0, M ask 0x2
BGP version 4, remote router ID 7.72.6.2
BGP state = Established, table version = 9, up for 4d21h
Last read 00:00:56, last send 00:00:48
Hold time 180, keepalive interval 60 seconds
Neighbor NLRI negotiation:
  Configured for unicast routes only
  Peer negotiated unicast and multicast routes
  Exchanging unicast routes only
  Received route refresh capability from peer
Maximum time between advertisement runs is 30 seconds
Received 7044 messages, 0 notifications, 0 in queue
Sent 7041 messages, 0 notifications, 0 in queue
Prefix advertised 4, suppressed 0, withdrawn 0
Route refresh request: received 0, sent 0
Inbound path policy configured
Route map for incoming advertisements is k
Connections established 1; dropped 0
Last reset never
Number of unicast/multicast prefixes received 4, 0
Exterernal BGP neighbor may be up to 255 hops away.
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 3.72.6.1, Local port: 179
Foreign host: 6.72.6.2, Foreign port: 11014

debug ip bgp

router#debug ip bgp ?
A.B.C.D BGP neighbor address
dampening BGP dampening
events BGP events
keepalives BGP keepalives
updates BGP updates
<cr>
  • Remember—can be dangerous!
    Use only in the lab or if advised by the TAC
  • To make a little safer:
    logging buffered <size>
    no logging console
Session Establishment
(debug ip bgp)

16:06:30: BGP: 7.72.6.1 sending OPEN, version 4
16:06:31: BGP: 7.72.6.1 OPEN rcvd, version 4
16:06:31: BGP: 7.72.6.1 rcv OPEN w/ OPTION parameter len: 12
16:06:31: BGP: 7.72.6.1 rcv OPEN w/ option parameter type 2 (Capability) len 6
16:06:31: BGP: 7.72.6.1 OPEN has CAPABILITY code: 1, length 4
16:06:31: BGP: 7.72.6.1 OPEN has MP_EXT CAP for afi/safi: 1/1
16:06:31: BGP: 7.72.6.1 rcv OPEN w/ option parameter type 2 (Capability) len 2
16:06:31: BGP: 7.72.6.1 OPEN has CAPABILITY code: 128, length 0

16:06:31: BGP: 7.75.7.1 passive open
16:06:31: BGP: 7.75.7.1 OPEN rcvd, version 4
16:06:31: BGP: 7.75.7.1 sending OPEN, version 4
16:06:31: BGP: 7.75.7.1 rcv OPEN w/ OPTION parameter len: 12
16:06:31: BGP: 7.75.7.1 rcv OPEN w/ option parameter type 2 (Capability) len 6
16:06:31: BGP: 7.75.7.1 OPEN has CAPABILITY code: 1, length 4
16:06:31: BGP: 7.75.7.1 OPEN has MP_EXT CAP for afi/safi: 1/1
16:06:31: BGP: 7.75.7.1 rcv OPEN w/ option parameter type 2 (Capability) len 2
16:06:31: BGP: 7.75.7.1 OPEN has CAPABILITY code: 128, length 0

Session Establishment
(debug ip bgp events)

17:31:39: BGP: 7.72.6.1 went from Idle to Active
17:32:00: BGP: 7.72.6.1 went from Active to OpenSent
17:32:00: BGP: 7.72.6.1 went from OpenSent to OpenConfirm
17:32:00: BGP: 7.72.6.1 went from OpenConfirm to Established

17:31:59: BGP: 7.75.7.1 went from Idle to Active
17:32:00: BGP: 7.75.7.1 went from Active to Idle
17:32:00: BGP: 7.75.7.1 went from Idle to Connect
17:32:00: BGP: 7.75.7.1 went from Connect to OpenSent
17:32:00: BGP: 7.75.7.1 went from OpenSent to OpenConfirm
17:32:00: BGP: 7.75.7.1 went from OpenConfirm to Established
Looking at the Updates

router#debug ip bgp updates?
<1-199>      Access list
<1300-2699>  Access list (expanded range)
<cr>

router#debug ip bgp x.x.x.x updates?
<1-199>      Access list
<1300-2699>  Access list (expanded range)
<cr>

Use an access-list to limit the output!

debug ip bgp Updates

<table>
<thead>
<tr>
<th>Peer Address</th>
<th>Prefix Being Advertised</th>
<th>NEXT_HOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP: 6.72.6.2 computing updates, neighbor version 0, table version 13, starting at 0.0.0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: 6.72.6.2 send UPDATE 3.0.0.0/8, next 3.72.6.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: , metric 0, path 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: 6.72.6.2 send UPDATE 5.0.0.0/8 (negflags: 0x0), next 3.72.6.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: 6.72.6.2 send UPDATE 7.0.0.0/8 (negflags: 0x0), next 3.72.6.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: 6.72.6.2 1 updates enqueued (average=56, max um=56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGP: 6.72.6.2 update run completed, ran for 0m s, neighbor version 0, start version 13, throttled to 13, check point next 0.0.0.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### debug ip bgp Updates (Cont.)

BGP: 6.72.6.2 rcv UPDATE w/attr: nexthop 6.72.6.2, origin i, metric 494, path 2
BGP: 6.72.6.2 rcv UPDATE about 6.0.0.0/8
BGP: 6.72.6.2 rcv UPDATE about 17.0.0.0/8
BGP: 6.72.6.2 rcv UPDATE about 23.0.0.0/8
BGP: 6.72.6.2 rcv UPDATE about 35.0.0.0/8

**Peer Address** | **Prefixes in the Same UPDATE** | **Attributes Apply to All Prefixes**

BGP: 6.72.6.2 rcv UPDATE w/attr: nexthop 6.72.6.2, origin i, metric 294, path 2 1
BGP: 6.72.6.2 rcv UPDATE about 3.0.0.0/8 —DENIED due to: as-path contains our own AS;
BGP: 6.72.6.2 rcv UPDATE about 7.0.0.0/8 —DENIED due to: as-path contains our own AS;

---

### Logging Neighbor Changes

- Generate a log message whenever a BGP neighbor changes state, also indicate reason for reset

- Syntax (router subcommand):

  ```
  [no] bgp log-neighbor-changes
  ```

  Typical log messages:

  `%BGP-5-ADJCHANGE: neighbor x.x.x.x Up`
  `%BGP-5-ADJCHANGE: neighbor x.x.x.x Down-Remote AS changed`
show ip bgp neighbors x.x.x.x

router#show ip bgp neighbors 7.75.7.1
BGP neighbor is 7.75.7.1, remote AS 2, external link
... Received 194 messages, 1 notifications, 0 in queue
Sent 194 messages, 0 notifications, 0 in queue
Prefix advertised 0, suppressed 0, withdrawn 0
Route refresh request: received 0, sent 0
Connections established 7; dropped 7
Last reset 00:04:11, due to BGP Notification received, hold time expired
Number of unicast/multicast prefixes received 0/0
External BGP neighbor may be up to 255 hops away.
No active TCP connection

Come Meet the Neighbors!

Peer Establishment
Peer Establishment

- Routers establish a TCP session
  Port 179—permit in ACLs
  IP connectivity (route from IGP)
- OPEN messages are exchanged
  Peering addresses must match the TCP session
  Local AS configuration parameters
  Capabilities negotiation

Common Problems

- Sessions are not established
  No IP reachability
  Incorrect configuration
  Peering addresses
  OPEN parameters
Can’t Establish Session - Symptoms

```
routerA#show ip bgp summary
BGP router identifier 7.72.6.1, local AS number 1
BGP table version is 4, main routing table version 4
6 network entries and 6 paths using 774 bytes of memory
2 BGP path attribute entries using 96 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
BGP activity 6,0 prefixes, 6,0 paths
```

```
Neighbor     V  AS M sgRcvd M sgSent Tbl Ver InQ OutQ Up/Down State/PrfxRcd
6.72.6.2     4  2  0  0  0  0  0 0 never  Idle
7.25.14.4    4  3  4  5  4  0 00:01:43 0
7.72.6.3     4  1  0  0  0  0  0 never  Active
7.75.7.1     4  1  7  5  4  0 00:01:55 3
```

• The peering session is not **established**!
  
State may change between **active**, **idle** and **connect**

Can’t Establish Session—Troubleshooting I

• **Is the remote-as** assigned correctly?

```
router bgp 1
neighbor 6.72.6.2 remote-as 2
neighbor 7.72.6.3 remote-as 1
```
Can’t Establish Session—Troubleshooting I (Cont.)

- Verify IP connectivity
  - check the routing table
  - use ping/trace to verify two way reachability
  - inspect for ACLs in the path to the neighbor

```
routera#show ip route 7.72.6.3
Routing entry for 7.72.6.3/32:
Known via "ospf123", distance 110, metric 87, type intra area
Last update from 27.27.27.254 on Pos5/0, 00:09:33 ago
Routing Descriptor Blocks:
  27.27.27.254, from 7.72.6.3, 00:09:33 ago, via Pos5/0
  Route metric is 87, traffic share count is 1
```

```
routera#ping 7.72.6.3
Sending 5, 100-byte ICMP Echos to 7.72.6.3, timeout is 2 seconds:
!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/30/32 ms
```

```
routera#debug ip bgp
BGP debugging is on
10:51:02: BGP: 7.72.6.3 open active, delay 6864 ms
10:51:09: BGP: 7.72.6.3 open active, local address 27.27.27.253
10:51:09: BGP: 7.72.6.3 open failed: Connection refused by remote host
```

- Is the remote router configured for BGP?
  What IP address is the remote router configured to receive?
The TCP session is always sourced from the closest IP address to the destination!

- Configuration:
  
  Router A
  
  `router bgp 1`
  `neighbor 27.27.27.254 remote-as 1`

  Router C
  
  `router bgp 1`
  `neighbor 27.27.27.253 remote-as 1`

  If redundant paths exist, use loopback interfaces to establish the session.

- Solution:

  make sure both routers source the information from the appropriate interface
Can’t Establish Session—Symptoms

RouterA #show ip bgp summary
BGP router identifier 7.72.6.1, local AS number 1
BGP table version is 4, main routing table version 4
6 network entries and 6 paths using 774 bytes of memory
2 BGP path attribute entries using 96 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
BGP activity 6/0 prefixes, 6/0 paths

Neighbor  V  AS  M sgRcvd M sgSent TblVer InQ OutQ Up/Down State/PfxRcd
6.72.6.2  4 2 0 0 0 0 0 never  Idle
7.72.6.2  4 3 385 385 4 0 06:22:17 0
7.72.6.3  4 1 42 49 4 0 00:00:15 0
7.75.7.1  4 1 388 385 4 0 06:22:30 3

- The eBGP session is still having trouble!

Can’t Establish Session - Troubleshooting II

- Verify IP connectivity
  check the routing table
  use ping/trace to verify two way reachability

RouterA #show ip route 6.72.6.2
% Network not in table

RouterA(config)#ip route 6.72.6.2 255.255.255.255 1.1.1.5

RouterA #ping 6.72.6.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 6.72.6.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 m s
Can’t Establish Session—Troubleshooting II (Cont.)

- Peering with a loopback interface

  **Advantages**
  - Interface is always up
  - Multiple physical paths may exist to reach it

  **Disadvantages**
  - Physical link failure may take longer to detect

```
routerA debug ip bgp
routerrA debug ip tcp transactions
13:25:30: TCP: sending RST, seq 0, ack 2030100669
```

```
router bgp 1
neighbor 6.72.6.2 remote-as 2
neighbor 6.72.6.2 update-source Loopback1
```

- The debug output indicates the neighbor’s configured peering address
Hint: by default, eBGP peers should be directly connected
in this case, the peering address doesn’t match a connected interface in the local router

routerA #show ip bgp neighbors 6.72.6.2
BGP neighbor is 6.72.6.2, remote AS 2, external link
Index 1, offset 0, mask 0x2
BGP version 4, remote router ID 0.0.0.0
BGP state = Idle, table version = 0
Last read 00:00:06, last send never
Hold timer 180, keepalive interval 60 seconds
Neighbor NLRI negotiation:
  configured for unicast routes only
  minimum time between advertisement runs is 30 seconds
Received 0 messages, 0 notifications, 0 in queue
Sent 0 messages, 0 notifications, 0 in queue
Prefix advertised 0, suppressed 0, withdrawn 0
Route refresh request: received 0, sent 0
Connections established 0; dropped 0
Last reset never
Number of unicast/multicast prefixes received 0/0
External BGP neighbor not directly connected.
No active TCP connection
Can’t Establish Session—Troubleshooting II (Cont.)

- At this point, the session should come up

Can’t Establish Session—Symptoms

- Still having trouble!
  Connectivity issues have already been checked and corrected.
If an error is detected, a **notification** is sent and the session is closed.

In this case the remote router had a bad configuration.
Capabilities Negotiation

- Allows for the advertisement of capabilities (type 2)

- Backwards compatible
  
  New error subcode introduced to indicate which capabilities are not supported—the session must be reset

draft-ietf-idr-bgp4-cap-neg, Mar. 2000

Where’s the Beef?

UPDATE Exchange
UPDATE Exchange

- Once the session has been established, UPDATEs are exchanged all the locally known routes only the bestpath is advertised
- Incremental UPDATE messages are exchanged afterwards

Propagation Decisions

- bestpath received from eBGP peer advertise to all peers
- bestpath received from iBGP peer advertise only to eBGP peers a full iBGP mesh must exist
Common Problems

- Missing routes
  - No iBGP full mesh
  - Filters: routes are not received/sent
- Slow convergence

UPDATE Filters

- Type of filters
  - Prefix filters
  - AS_PATH filters
  - Community filters
  - Any attribute may be used in a route-map
- Applied incoming and/or outgoing
Missing Routes—Troubleshooting Steps

- Determine which filters are applied to the BGP session
  
  `show ip bgp neighbors x.x.x.x`

  Look at the configuration

- Examine the route and pick out the relevant attributes
  
  `show ip bgp x.x.x.x`

Missing Routes—Troubleshooting Steps (Cont.)

- Compare the route against the filters

- If no match is found
  
  Use route-refresh or soft-reconfiguration

  Filter the updates through an ACL to determine where the problem is
**Missing Routes—Symptoms**

- Missing 4.0.0.0/8 in 7.75.7.1 (routerA) not received from 7.72.6.3 (routerB)

```
routerB#sh ip bgp nei 7.75.7.1 advertised-routes | include 4.0.0.0
*> 4.0.0.0 0.0.0.0 0 32768 i
```

routerB shows that the route was advertised to routerA!

**Missing Routes—Troubleshooting**

```
routerA#show access-lists 10
Standard IP access list 10
  permit 4.0.0.0

routerA#debug ip bgp 7.72.6.3 updates 10
BGP updates debugging is on for access list 10 for neighbor 7.72.6.3

routerA#clear ip bgp 7.72.6.3 in
01:22:41: BGP: 7.72.6.3 rcv UPDATE w/ attr: nexthop 7.72.6.3, origin i, metric 0, path 2
01:22:41: BGP: 7.72.6.3 rcv UPDATE about 4.0.0.0/8 -- DENIED due to: distribute/prefix-list;
```
Missing Routes—Troubleshooting (Cont.)

router bgp 1
   no synchronization
bgp log-neighbor-changes
neighbor 7.72.6.3 remote-as 2
neighbor 7.72.6.3 ebgp-multihop 255
neighbor 7.72.6.3 update-source Loopback0
neighbor 7.72.6.3 prefix-list filter in
! ip prefix-list filter seq 5 deny 4.0.0.0/8
ip prefix-list filter seq 10 permit 0.0.0.0/0 les 32

Slow Convergence—Symptoms

• The eBGP peering is established, but convergence is not complete even after several hours

routerA #show ip bgp summary
...
Neighbor     V AS M sgR cvd M sgSent TbM v er InQ  OutQ Up/Down State/Perf Rcd
150.10.10.1  1 41 3550 3570 847 0 206 05:53:51 100

• Possible causes
  Remote router is not healthy (OutQ)
  Lower layer problems (IP)
**Slow Convergence—Troubleshooting**

```
router bgp 1
neighbor 150.10.10.1 remote-as 2
neighbor 150.10.10.1 ebgp-multihop 2
neighbor 150.10.10.1 update-source Loopback0
```

```
routerA# show ip route 150.10.10.1
Routing entry for 150.10.10.1/32
Routing Descriptor Blocks:
10.105.1.71, from 150.20.20.1, 00:06:14 ago, via POS2/1/0
* 156.1.1.1, from 150.20.20.1, 00:06:14 ago, via POS2/1/1
```

```
routerA# ping 150.10.10.1
Sending 5, 100-byte ICMP Echos to 150.10.10.1: !!!!!
Success is 100 percent, round-trip min/avg/max = 4/64/296 ms
```

```
eBGP Peering
```

```
router bgp 1
neighbor 150.10.10.1 remote-as 2
neighbor 150.10.10.1 ebgp-multihop 2
neighbor 150.10.10.1 update-source Loopback0
```

```
Longest path has more than 2 hops to the destination. Use higher TTL!
```
A common policy should be maintained across the AS to guarantee loop-free operation. Not all routers may select the same path. Filters may be used to modify or add attributes, affecting the selection algorithm.
Common Problems

- Inconsistent decision/policy
  - MED
  - External paths
  - Communities
    - By default, communities are not propagated
    
    ```
    neighbor x.x.x.x send-community
    ```

Inconsistent Decision—Symptom I

- The bestpath changes every time the peering is reset.

```
routerA#sh ip bgp 160.100.0.0
BGP routing table entry for 160.100.0.0/16, version 40
Paths: (3 available, best #3, advertised over IBGP, EBG P)
  1
    204.146.33.10 from 204.146.33.10 (204.146.33.1)
      Origin IGP, metric 0, localpref100, valid, internal
  3
    204.146.33.66 from 204.146.33.66 (204.146.33.2)
      Origin IGP, metric 20, localpref100, valid, internal
  3
    204.146.33.6 from 204.146.33.6 (10.4.1.1)
      Origin IGP, metric 30, valid, external, best
```
Inconsistent Decision—Symptom I (Cont.)

• Same paths, but different result!

routerA #sh ip bgp 160.100.0.0
BGP routing table entry for 160.100.0.0/16, version 2
Paths: (3 available, best #3, advertised over EBP)
  1
   204.146.33.10 from 204.146.33.10 (204.146.33.1)
     Origin IG P, metric 0, localpref100, valid, internal
  3
   204.146.33.6 from 204.146.33.6 (10.4.1.1)
     Origin IG P, metric 30, valid, external
  3
   204.146.33.66 from 204.146.33.66 (204.146.33.2)
     Origin IG P, metric 20, localpref100, valid, internal, best

Inconsistent Decision—Symptom I (Cont.)

• Different result...again!!

routerA #sh ip bgp 160.100.0.0
BGP routing table entry for 160.100.0.0/16, version 12
Paths: (3 available, best #3, advertised over EBP)
  3
   204.146.33.6 from 204.146.33.6 (10.4.1.1)
     Origin IG P, metric 30, valid, external
  3
   204.146.33.66 from 204.146.33.66 (204.146.33.2)
     Origin IG P, metric 20, localpref100, valid, internal
  1
   204.146.33.10 from 204.146.33.10 (204.146.33.1)
     Origin IG P, metric 0, localpref100, valid, internal, best
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
  enable in all the routers in the AS

Deterministic MED—Operation

- The paths are ordered by peer-AS
- The bestpath for each group is selected
- The overall bestpath results from comparing the winners in each group
Deterministic MED—Result

The bestpath will always be the same!

```
routerA#show ip bgp 160.100.0.0
BGP routing table entry for 160.100.0.0/16, version 15
Paths: (3 available, best #1, advertised over EBG P)
  1
  204.146.33.10 from 204.146.33.10 (204.146.33.1)
    Origin: IG P, metric 0, localpref 100, valid, internal, best
  3
  204.146.33.66 from 204.146.33.66 (204.146.33.2)
    Origin: IG P, metric 20, localpref 100, valid, internal
  3
  204.146.33.6 from 204.146.33.6 (10.4.1.1)
    Origin: IG P, metric 30, valid, external
```

Inconsistent Decision—Symptom II

The bestpath changes every time the peering is reset

```
routerA#show ip bgp 7.0.0.0
BGP routing table entry for 7.0.0.0/8, version 15
Paths: (2 available, best #2)
Not advertised to any peer
  2
  1.1.1.5 from 1.1.1.5 (1.1.1.1)
    Origin: IG P, metric 0, localpref 100, valid, external
  2
    Origin: IG P, metric 0, localpref 100, valid, external, best
```
Inconsistent Decision—Symptom II (Cont.)

The “oldest” external is the bestpath. All other attributes are the same Stability enhancement!

routeA #show ip bgp 7.0.0.0
BGP routing table entry for 7.0.0.0/8, version 17
Paths: (2 available, best #2)
Not advertised to any peer
2
Origin IGP, metric 0, localpref 100, valid, external
2
1.1.1.5 from 1.1.1.5 (1.1.1)
Origin IGP, metric 0, localpref 100, valid, external, best

Route Reflectors

Playing with Mirrors
Route Reflectors

- Provide additional control to allow router to advertise (reflect) iBGP learned routes to other iBGP peers
  - Method to reduce the size of the iBGP mesh
- Normal BGP speakers can coexist
  - Only the RR has to support this feature

Route Reflectors—Terminology

Lines Represent Both Physical Links and BGP Logical Connections
Reflection Decisions

Once the best path is selected:

From non-client reflect to all clients

From client → reflect to all non-clients AND other clients

From eBGP peer → reflect to all clients and non-clients

Common Problems

- Missing routes
- Routing loops and “close calls”
Missing Routes—Symptoms

- At least one route is missing from at least one router in the network.

```
routerA # show ip bgp 4.0.0.0
% Network not in table
routerA # show ip bgp summary
BGP router identifier 7.25.14.4, local AS number 1
BGP table version is 1, main routing table version 1
...
Neighbor VAS msgRcvd Msgrcvd InQ OutQ Up/DOWN State/PfxRcd
7.72.6.2 4 1 7 7 1 0 00:04:18 0
```

Missing Routes—Troubleshooting

- Check routers for filters

```
routerB #
router bgp 1
no synchronization
neighbor 7.72.6.2 remote-as 1

routerB # show ip bgp 4.0.0.0
% Network not in table
```

```
routerC
```
Missing Routes—Troubleshooting I

• Follow the path where the routes should have been learned

```plaintext
routerC #
rout er bgp 1
no synchronization
bgp cluster-id 0.0.0.5
neighbor 7.72.6.2 remote-as 1
neighbor 7.75.7.1 remote-as 1
neighbor 7.75.7.1 route-reflector-client

RouterC # show ip bgp | include 4.0.0.0
*>4.0.0.0 7.72.6.3 0 100 0 2 i
```

Same Cluster-ID!

Missing Routes—Troubleshooting II

• Alternative way to find duplicate cluster-id

use route-refresh + `debug ip bgp updates ACL`

```plaintext
routerB # clear ip bgp 7.72.6.1 in
21:45:40: BGP: 7.72.6.1 rcv UPDATE w/attr: nexthop 7.72.6.3,
origin i, localpref 100, metric 0, path 2
21:45:40: BGP: 7.72.6.1 rcv UPDATE about 4.0.0.0/8 — DENIED due
 to: reflected from the same cluster;
```
Clusters with multiple RRs

If the RRs have the same cluster-id, all the clients must peer with all the reflectors.

If not needed for administration, don’t assign a cluster-id for added flexibility!

Routing Loop—Symptom

```
rtrBrtrB
rtrCrtrC

routerD#traceroute 7.1.1.1
1 1.1.1.2 24 msec 24 msec 40 msec
2 156.1.1.1 28 msec 48 msec 24 msec
3 156.1.1.2 24 msec 24 msec 24 msec
4 156.1.1.1 28 msec 24 msec 24 msec
5 156.1.1.2 28 msec 28 msec 28 msec
6 156.1.1.1 28 msec 28 msec 32 msec

Loop!
```
Routing Loop—Troubleshooting

Verify Routing Information

routerC# show ip bgp 7.0.0.0
BGP routing table entry for 7.0.0.0/8
1
150.10.10.1 (metric 115) from 150.10.10.1 (150.20.20.1)
Origin IGP, valid, external, best
routerC# show ip route 150.10.10.1
Routing entry for 150.10.10.1/32
Routing Descriptor Blocks:
  * 156.1.1.1, from 150.20.20.1, via Ethernet2/1/1

routerB# show ip bgp 7.0.0.0
BGP routing table entry for 7.0.0.0/8
1
156.1.1.2 from 156.1.1.2 (212.212.212.1)
Origin IGP, valid, internal, best
routerB# show ip route 156.1.1.2
Routing entry for 156.1.1.0/24
Routing Descriptor Blocks:
  * directly connected, via Ethernet1

Addresses on Same Subnet!

Routing Loop—Troubleshooting (Cont.)

• Check configuration

  routerC#
  router bgp 134
  neighbor 150.10.10.1 remote-as 1
  neighbor 150.10.10.1 ebgp-multihop 255
  neighbor 150.10.10.1 update-source Loopback0
  neighbor 156.1.1.1 remote-as 134
  neighbor 156.1.1.1 route-reflector-client
  neighbor 156.1.1.1 next-hop-self
  !
  ip route 150.10.10.1 255.255.255.255 s0 250

Lines Represent Physical Connections
Routing Loop—Troubleshooting (Cont.)

- **Solution**
  
  Establish the eBGP peering permanently through the “backup” link
  
  Use LOCAL_PREF or MED to break any tie!

Close Call—Symptom

- The bestpath is not being followed to an external destination

```
routera#show ip route 4.4.4.4
Routing entry for 4.0.0.0/8
  Known via "bgp 1", distance 200, metric 0
  Tag 2, type internal
  Last update from 6.72.63.0 00:25:45 ago
  Routing Descriptor Blocks:
    * 6.72.63.0, from 7.75.71.0 00:25:45 ago
    Route metric is 0, traffic share count is 1
    AS Hops 1, BGP network version 0
```

Expected to go out through the NEXT_HOP in the update.
Close Call—Symptom (Cont.)

All eBGP peers are configured with their interface address

The NEXT_HOP is expected in the trace

Close Call—Troubleshooting (Cont.)

Verify configuration

Check for alternate routes
• **routerC is a RR with two clients**

```
router bgp 1
  no synchronization
  bgp log-neighbor-changes
  neighbor 7.72.6.1 remote-as 1
  neighbor 7.72.6.1 update-source Loopback0
  neighbor 7.72.6.1 route-reflector-client
  neighbor 7.72.6.2 remote-as 1
  neighbor 7.72.6.2 update-source Loopback0
  neighbor 7.72.6.2 route-reflector-client
  neighbor 6.72.6.3 remote-as 2

routerC #sh show ip bgp 4.0.0.0
BGP routing table entry for 4.0.0.0/8, version 2
  (Received from a RR-client)
  8.25.14.3 (metric 3) from 7.72.6.1 (7.72.6.1)
    Origin IGP, metric 0, localpref100, valid, internal
  2
  6.72.6.3 (metric 2) from 6.72.6.3 (6.72.6.3)
    Origin IGP, metric 0, localpref100, valid, external, best

routerC #traceroute 4 A A A
  1 6.72.6.3  0 m sec 4 m sec 0 m sec
  2 161.44.0.56 0 m sec 4 m sec 0 m sec
  3 161.44.0.18 0 m sec 4 m sec 0 m sec
  4 A A A  0 m sec 4 m sec 0 m sec

Expected Path!
```
**Output from the other client follows the same exit as routerA**

```
routerB#traceroute 4.4.4.4
1 8.25.14.3 8 m sec 16 m sec 8 m sec
2 172.18.176.1 16 m sec 12 m sec 16 m sec
3 161.44.0.48 12 m sec 16 m sec 12 m sec
4 161.44.0.15 16 m sec 12 m sec 16 m sec
5 4.4.4.4 8 m sec 8 m sec 8 m sec
```

**routerB is following the correct path!**

```
routerB#show ip bgp 4.0.0.0
BGP routing table entry for 4.0.0.0/8, version 13
  2
    8.25.14.3 (metric 2) from 8.25.14.3 (8.25.14.3)
      Origin IGP, metric 0, localpref 100, valid, external, best
  2
    6.72.6.3 (metric 50) from 7.75.7.1 (7.75.7.1)
      Origin IGP, metric 0, localpref 100, valid, internal
```
The logical connection between routerC (RR) and routerA provides the route.

The physical path is followed.

Logical Connection

4.0.0.0/8

Summary/Tips

- Isolate the problem!
- Use ACLs when enabling debug commands
- Enable `bgp log-neighbor-changes`
- IP reachability must exist for sessions to be established
  - Learned from IGP
- Make sure the source and destination addresses match the configuration
Summary/Tips

- Use loopback interfaces for stability and where multiple paths exist
- Use common filters
  - Keep them simple!
- Maintain a consistent policy throughout the AS
- Use deterministic-med

Summary/Tips

- Select the appropriate knob/attribute for the job
  - Learn the decision algorithm
- Route reflectors
  - Follow the physical topology
  - Define a cluster-id only if administratively needed
Troubleshooting BGP in Large IP Networks

Session 2210

Please Complete Your Evaluation Form

Session 2210