



# A Hands-on Tutorial on BGP

University of South Carolina (USC)

Front Range GigaPop (FRGP)

Engagement and Performance Operations Center (EPOC)



Internet2 Technology Exchange

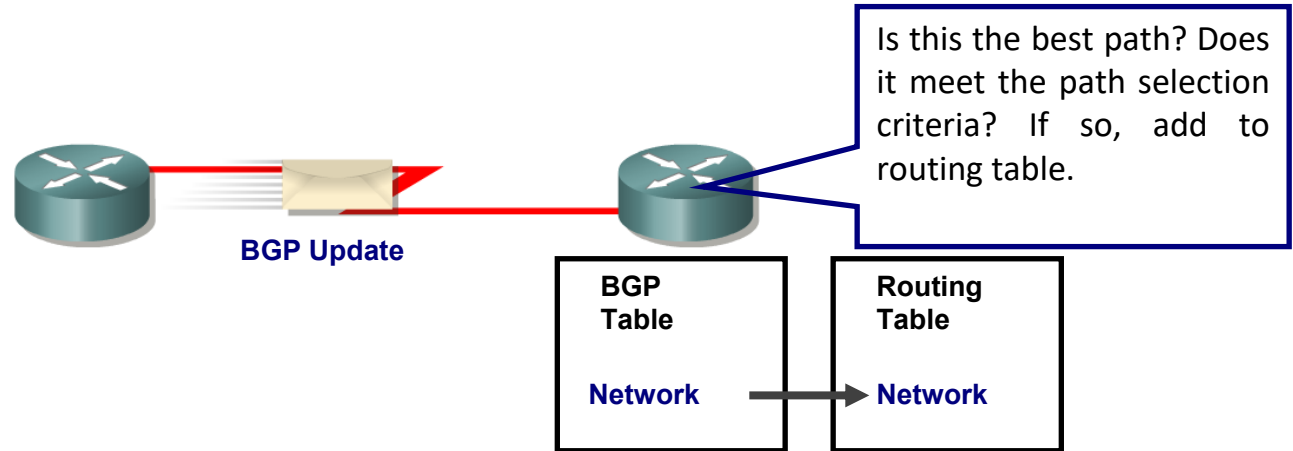
Denver, Colorado – December 5, 2022



# Overview of BGP Attributes

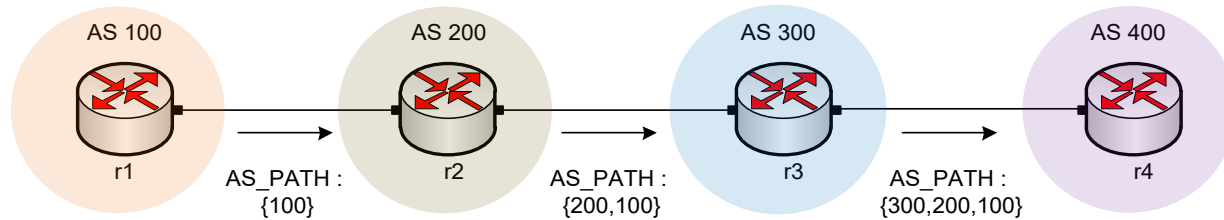
# BGP – Best Path

- The main goal is to provide interdomain routing
- BGP selects one path as the best path
- It places the selected path in its routing table and propagates the path to its neighbors



# BGP Route Advertisements between ASs

- In BGP route advertisements, each border router prepends its own AS number to the route before advertising the route to the next AS



# BGP – Best Path

---

1. Prefer the path with **highest weight** (configured locally; set to 0 (default) for routes not originated by the router)
2. If weights are the same, prefer the path with **highest local preference** (set to 100 by default)
3. If the local preferences are the same, prefer the **path that was originated by BGP running on the router or redistributed from an Interior Gateway Protocol (IGP)**
4. If no route was originated, prefer the path with the **shortest AS\_PATH**
5. If the paths have the same AS\_PATH length, prefer the path with the **lowest origin type** (IGP is lower than Exterior Gateway Protocol (EGP), and EGP is lower than Incomplete)
6. If the origin codes are the same, prefer the path with the **lowest MED attribute** (set to 0 by default)
7. If the paths have the same MED, prefer the **External path (EBGP) over the Internal path (IBGP)**
8. If the paths are still the same, prefer the path through the **closest IGP neighbor (lowest IGP metric)**
9. If both paths are external, prefer the path that was **received first (oldest one)**
10. If the paths are still the same, prefer the path from the **BGP router with the lowest router ID**
11. If the router ID is the same for multiple paths, prefer the path with the **lowest IP address**

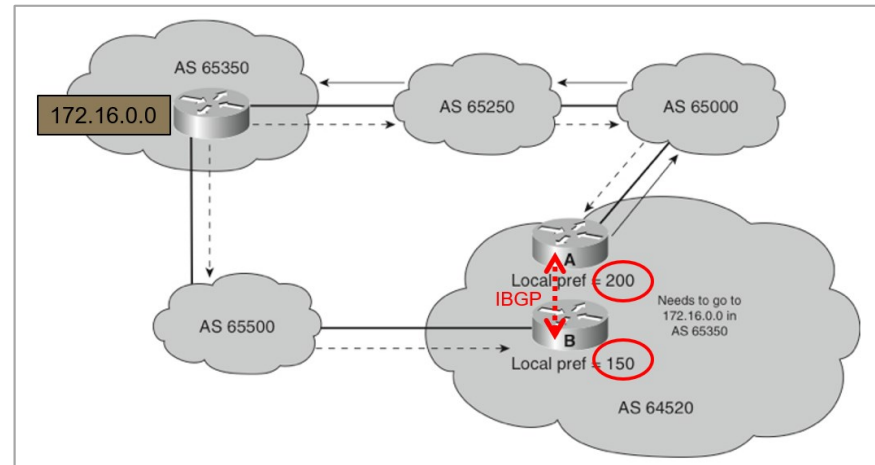
# BGP – Best Path

---

- Without route manipulation, the most common reason for path selection is Step 4
  - Prefer the path with the **shortest AS\_PATH**
- If multiple paths have the same length, the second most common decision point is Step 7
  - If the paths have the same MED, prefer **EBGP over internal IBGP**
  - We sometimes refer to this as “hot potato” routing
- The network administrator can manipulate the local preference to change the outbound path, and the MED attribute for the inbound path

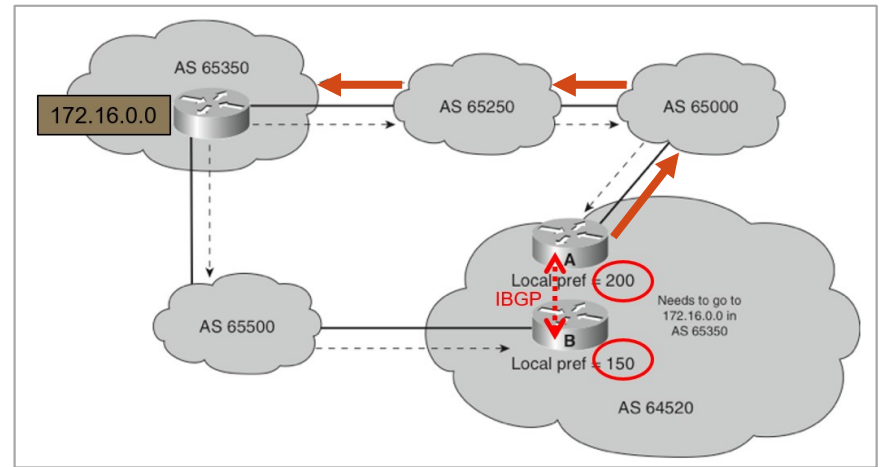
# Local Preference Attribute – For “Outbound Route”

- A well-known discretionary attribute
- Indicates to routers in the AS which path is preferred to exit the AS (higher is better)
- Configured on a router
- Exchanged only among routers within the same AS (passed only via IBGP, not via EBGP)
- Default value on a Cisco/FRR router is 100
- Local Preference takes precedence over AS\_PATH
- This is very frequently used in provider networks



# Local Preference Attribute – For “Outbound Route”

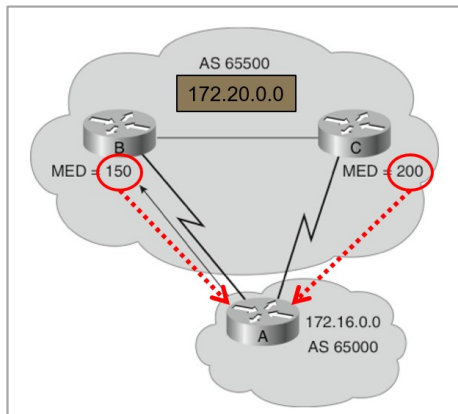
- A well-known discretionary attribute
- Indicates to routers in the AS which path is preferred to exit the AS (higher is better)
- Configured on a router
- Exchanged only among routers within the same AS (passed only via IBGP, not via EBGP)
- Default value on a Cisco/FRR router is 100
- Local Preference takes precedence over AS\_PATH
- This is very frequently used in provider networks
- All traffic in AS 64520 to 172.16.0.0 is sent via router A





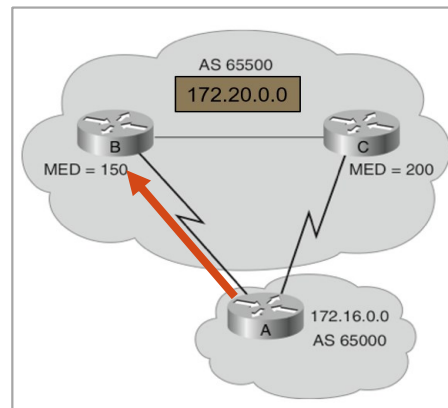
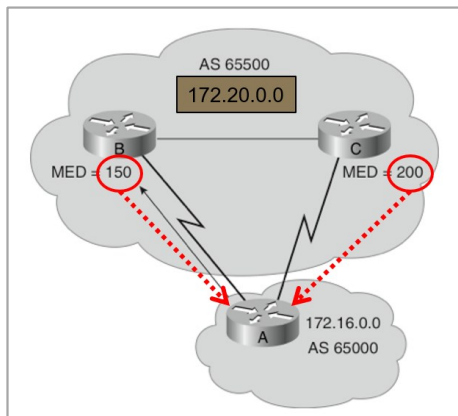
# The Med Attribute – For “Inbound Route”

- Multi-Exit Discriminator (MED), also called metric
- Indicates to external neighbors the preferred path into an AS
- By default, a router compares MED for paths from neighbors in the same AS (lowest is better)
- MED is sent to EBGP peers:
  - Those routers propagate the MED within their AS, but do not pass it on to the next AS
  - This may, or may not, be honored by the neighboring AS



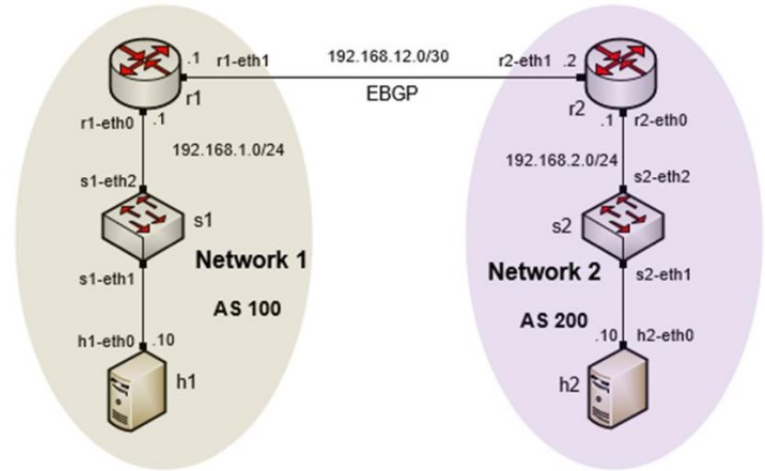
# The Med Attribute – For “Inbound Route”

- Router B has set the MED to 150 for 172.20.0.0
- Router C has set the MED to 200 for 172.20.0.0
- Router A receives EBGP updates from routers B and C
  - Router A selects router B as the best next hop to get to 172.20.0.0 because of the lower MED



# Next-hop Attribute

- A well-known mandatory attribute
- Unlike IGPs, BGP routes AS by AS, not router by router
- The next-hop address for a network from another AS is an IP address of the entry point of the next AS along the path to that destination network
- This default behavior is sometimes overridden through an iBGP export/outbound policy known as “next-hop self”



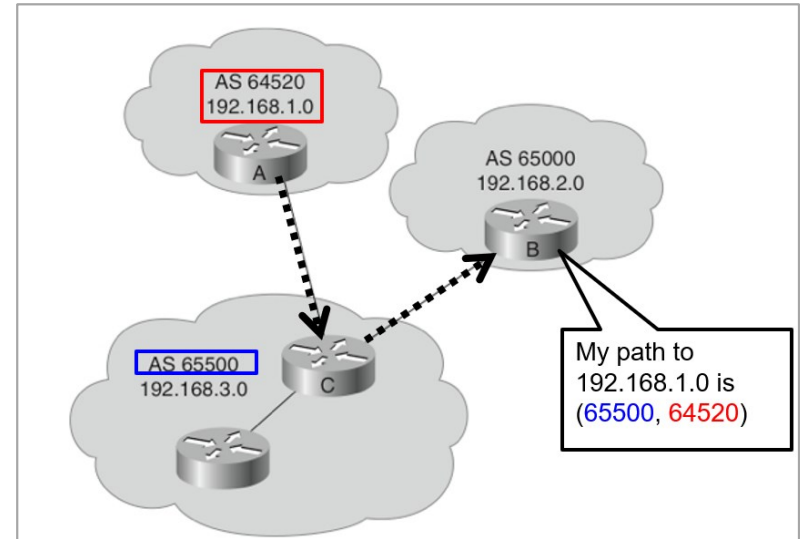
BGP table router r1

```
frr-pc# show ip bgp
BGP table version is 2, local router ID is 192.168.12.1, vrf id 0
Default local pref 100, local AS 100
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes: i - IGP, e - EGP, ? - incomplete

Network        Next Hop          Metric LocPrf Weight Path
*> 192.168.1.0/24 0.0.0.0           0         32768 i
*> 192.168.2.0/24 192.168.12.2     0         0 200 i
```

# The AS-Path Attribute

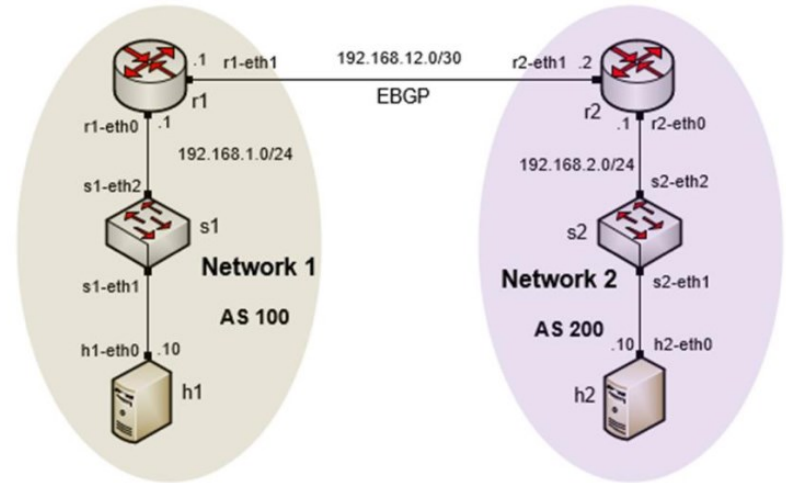
- Well-known mandatory attribute
- Whenever a route update passes through an AS, the AS number is prepended to that update
- Router A: advertises network 192.168.1.0 in AS 64520
- Router C: prepends its own AS number to it and advertises the route to Router B
- Router B: the path to reach 192.168.1.0 is:
  - 65500, 64520



# Additional Slides

# Origin Attribute

- A well-known mandatory attribute
- Defines the origin of the path information
- The origin attribute can be one of three values:
- **IGP (“i”)**
  - The route is interior to the originating AS
  - Normally when the **network command** is used
- **EGP (“e”)**
  - The route is learned via EGP
  - EGP is legacy and no longer supported
- **Incomplete (“?”)**
  - The route’s origin is unknown / some other means
  - It usually occurs when a route is **redistributed into BGP**



BGP table router r1

```
frr-pc# show ip bgp
BGP table version is 2, local router ID is 192.168.12.1, vrf id 0
Default local pref 100, local AS 100
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, r RIB-failure, S Stale, R Removed
NextHop codes: @NNN nextHop's vrf id, < announce-nh-self
Origin codes:  i - IGP, e - EGP, ? - incomplete

Network          Next Hop          Metric LocPrf Weight Path
*> 192.168.1.0/24  0.0.0.0           0       32768  i
*> 192.168.2.0/24 192.168.12.2     0       0 200  i
```

# Administrative Distance

- A router may run multiple routing protocols / static routes
- If BGP and OSPF are configured on a router, both protocols may provide different best paths (analogous to google-maps and mapquest)
- How does the router know which protocol to choose?
  - The route with lower Administrative Distance is installed in the routing table

Route Source	Administrative Distance
Connected	0
Static	1
EIGRP summary route	5
External BGP	20
Internal EIGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
External EIGRP	170
Internal BGP	200

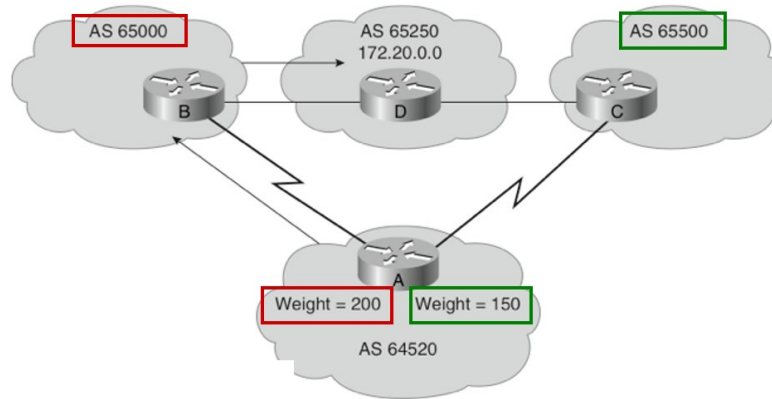
```
frr-pc# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP,
       O - OSPF, I - IS-IS, B - BGP, E - EIGRP, N - NHRP,
       T - Table, v - VNC, V - VNC-Direct, A - Babel, D - SHARP,
       F - PBR, f - OpenFabric,
       > - selected route, * - FIB route, q - queued route, r - rejected route

O> 192.168.1.0/24 [20/0] via 192.168.13.1, r3-eth2, 00:34:48
B  192.168.2.0/24 [200/0] via 192.168.23.1, r3-eth1, 00:34:38
O> * 192.168.2.0/24 [110/20] via 192.168.23.1, r3-eth1, 00:49:22
O  192.168.3.0/24 [110/10] is directly connected, r3-eth0, 00:49:04
C> * 192.168.3.0/24 is directly connected, r3-eth0, 00:52:03
C> * 192.168.13.0/30 is directly connected, r3-eth2, 00:52:03
O  192.168.23.0/30 [110/10] is directly connected, r3-eth1, 00:49:32
C> * 192.168.23.0/30 is directly connected, r3-eth1, 00:52:03
```

# The Weight Attribute - For “Outbound Route”

---

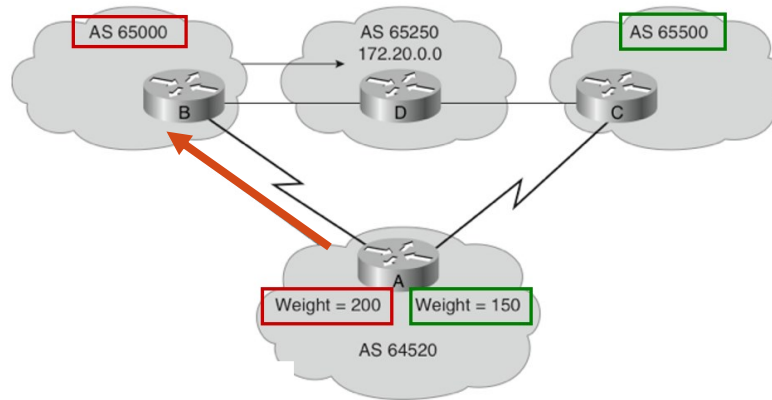
- Configured locally and not propagated to any other routers (higher is better)
- This is a vendor-specific attribute – Cisco supports it
- Juniper has a different mechanism to achieve a similar result
- Weight takes precedence over Local Preference
- Value from 0 to 65535; default is 32768
- Default is 0 for routes not originated by this router





# The Weight Attribute - For “Outbound Route”

- Router A has two ways to reach 172.20.0.0
  - via Router B (AS 65000)
  - via Router C (AS 65500)
- Router A is configured to set the weight of updates coming from:
  - Router B to 200
  - Router C to 150
- Weight for Router B is higher, so Router A uses Router B as a next hop to reach 172.20.0.0



# BGP Table

- Internal version number of the table
- This number is incremented whenever the table changes

```
frr-pec# show ip bgp
BGP table version is 3, local router ID is 192.168.23.2, vrf id 0
Default local pref 100, local AS 200
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes:  i - IGP, e - EGP, ? - incomplete

   Network        Next Hop        Metric LocPrf Weight Path
  *> 192.168.1.0/24 192.168.12.1    0      100     0 100 i
  *>                192.168.13.1    0           0 100 i
  *>i192.168.2.0/24 192.168.23.1    0      100     0 i
  *> 192.168.3.0/24 0.0.0.0         0           32768 i
```

# Status Code

- Displayed at the beginning of each line in the table

```
frr-pc# show ip bgp
BGP table version is 3, local router ID is 192.168.23.2, vrf id 0
Default local pref 100, local AS 200
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes:  i - IGP, e - EGP, ? - incomplete
```

```
Network        Next Hop        Metric LocPrf Weight Path
i> 192.168.1.0/24 192.168.12.1    0      100      0 100 i
* > 192.168.13.1 192.168.13.1    0      100      0 100 i
* > i 192.168.2.0/24 192.168.23.1    0      100      0 i
* > 192.168.3.0/24 0.0.0.0         0      32768    1
```

Code	Meaning
s	Table entry is suppressed
d	Table entry is dampened
h	Table entry history
*	Table entry is valid
>	Table entry is the best entry to use for this network
i	Table entry was learned via an internal BGP session
r	Table entry is a RIB-failure
S	Table entry is stale
=	Table entry has multipath to use for this network
b	Table entry has a backup path to use for this network
x	The table entry has a best external route to use for this network