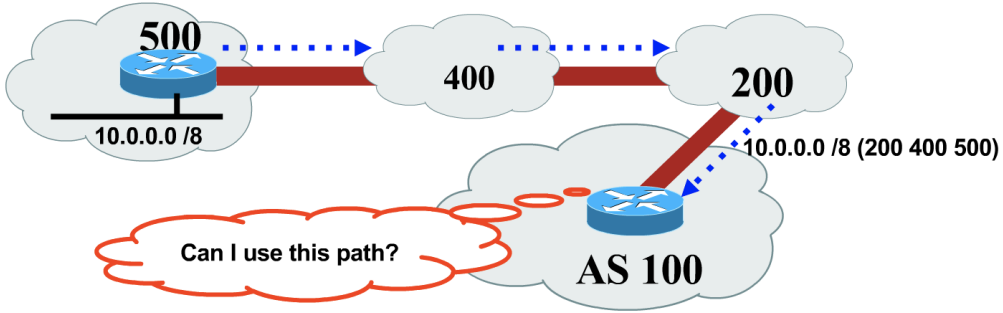
Tópicos Avançados
de Redes

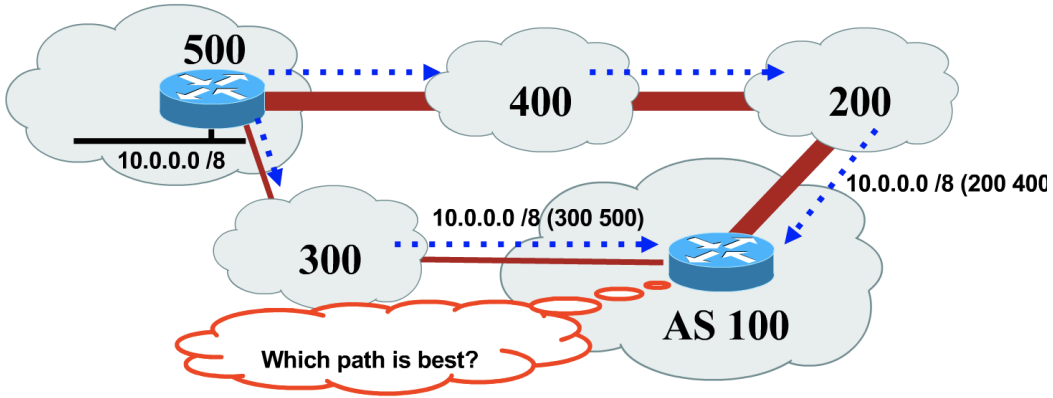
POLITÉCNICO DE LEIRIA

BGP Advanced

Ano Letivo 2019/2020

BGP Updates – Selecting a Path





BGP Path Attributes

- BGP selects the “best path” for any given prefix based upon received, and locally configured, Path Attributes associated with that Prefix.
  - Some Path Attributes MUST be sent when generating a BGP Update
  - Some Path Attributes must NOT be sent (are locally significant only).
  - Some received Path Attributes are at the router’s discretion whether to forward on to peers or not.

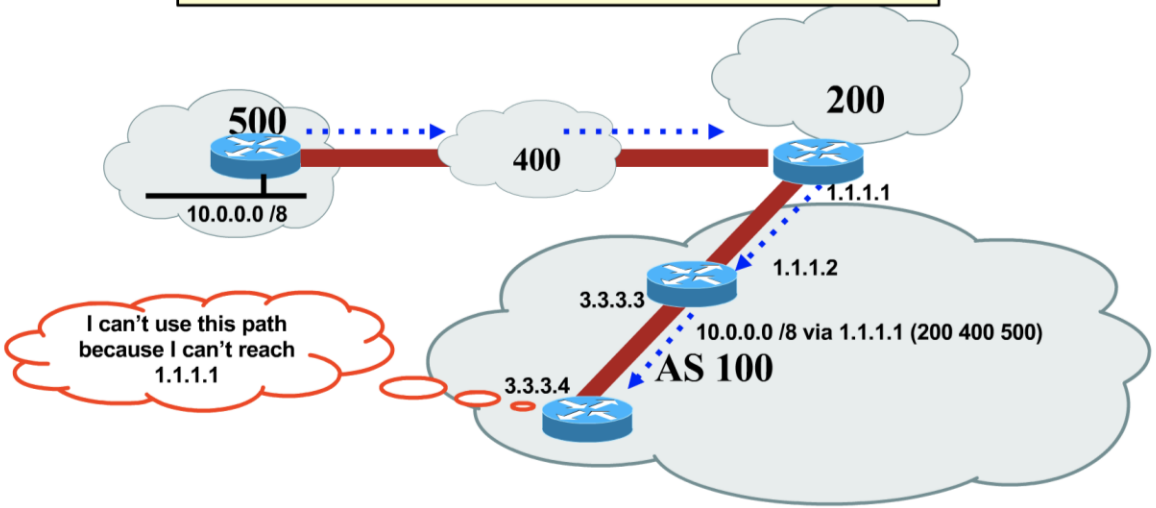
BGP Path Attributes

- *Attributes fall into four categories:*
- Well-known mandatory attributes "I know what this is and I have to give it to you".
  - ✓must be recognized by all BGP speakers
  - ✓must be included in all update messages.
- Well-known discretionary attributes; "I know what this is, but I might not give it to you".
  - ✓must be recognized by all BGP speakers
  - ✓may be carried in updates but are not required in every update.
- Optional transitive attributes; "I'm not sure what this is, but I'll give it to you anyway".
  - ✓may be recognized by some BGP speakers, but not all.
  - ✓should be preserved and advertised to all peers whether or not they are recognized.
- Optional non-transitive attributes; "What the heck is this? I'm discarding it!"
  - ✓may be recognized by some BGP speakers, but not all.
  - ✓If an update containing an optional transitive attribute is received, the update should be advertised to peers without the unrecognized attributes.

BGP Updates – Selecting a Path

- **Path Selection Criteria # 1:**
  Only consider paths with reachable NEXT_HOPs
  "Next_Hop" is a Mandatory Path Attribute

- **Path Selection Criteria # 2:**
  Do not consider iBGP path if not synchronized



BGP Path Selection – Weight

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT

BGP Path Selection – Weight

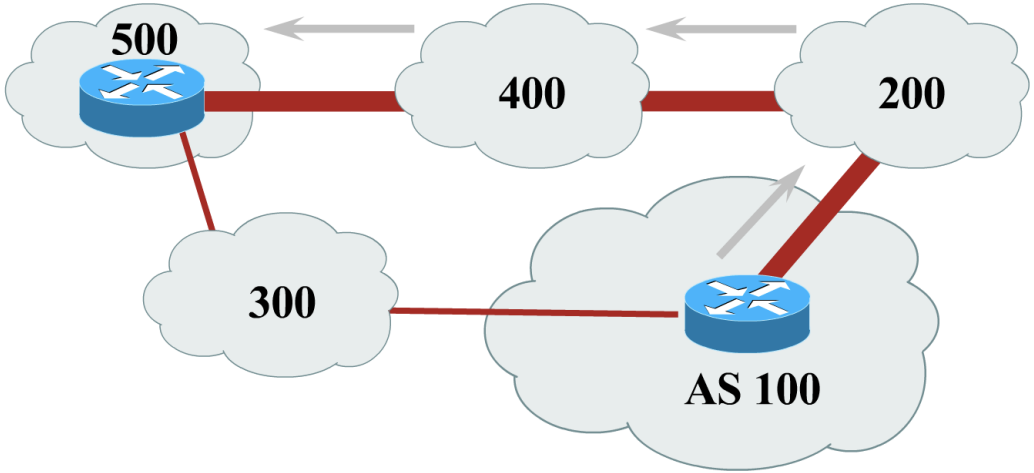
- Cisco specific
- Local to the router

  Not propagated
- value: 0 - 65535
- Default:

  originated locally = 32768

  other = 0

BGP Path Selection – Weight



- Higher Weight is preferred over lower Weight
- Weight is set via a route-map or neighbor statement
- Mainly used when an AS has only one router peering to multiple ASs

BGP Path Selection – Local Preference

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF

BGP Path Selection – Local Preference

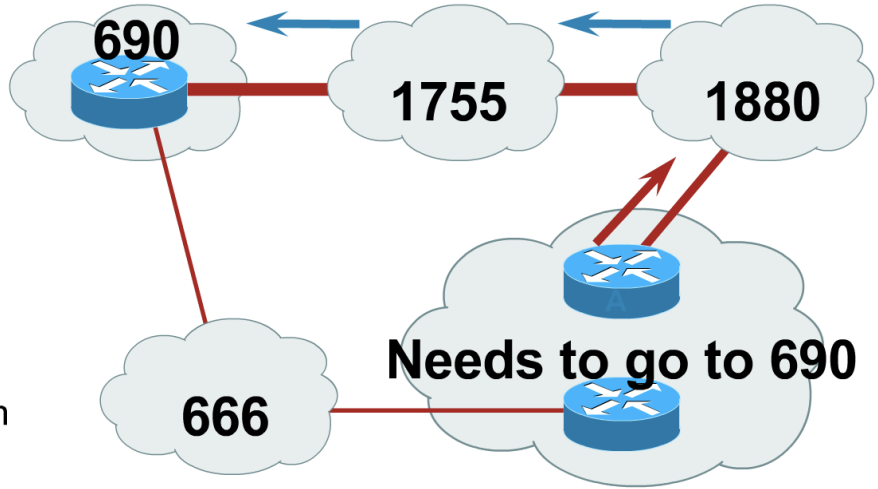
- Indication of preferred path to exit the local AS (passed on to other BGP peers in same AS)
- Global to the local AS
- Paths with highest LOCAL-PREF are most desirable (default = 100)

  bgp default local-preference *value*

BGP Path Selection – Local Preference

- Configuration (rtr A):

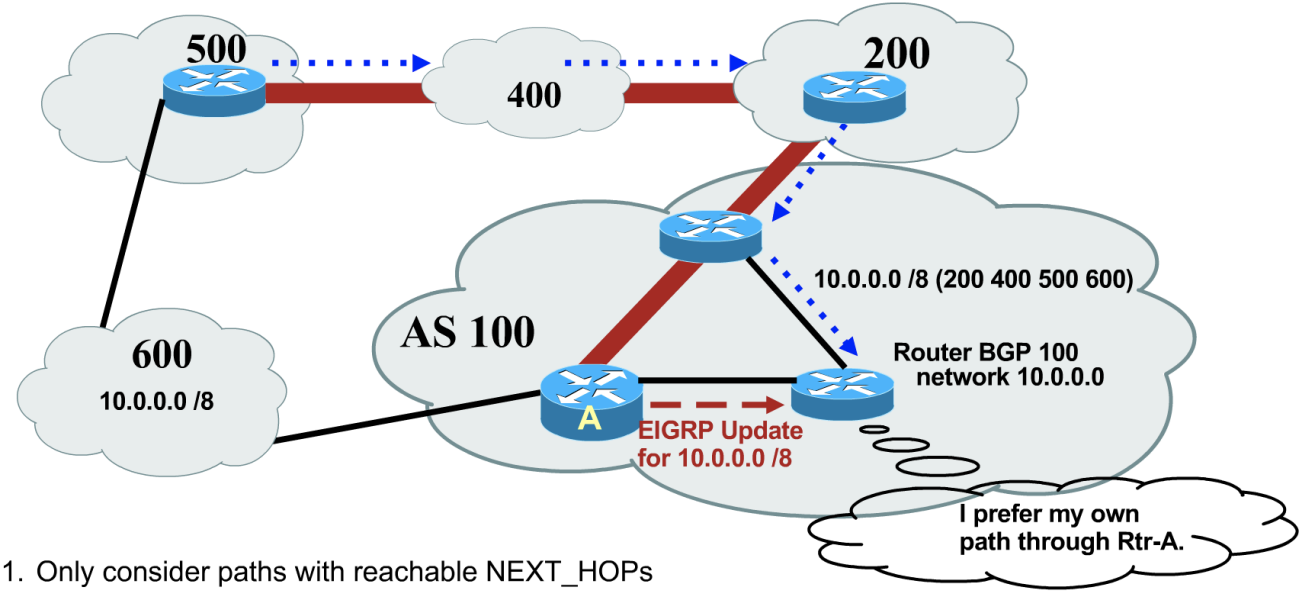
```
router bgp 109
neighbor x.x.x.x remote-as 1880
neighbor x.x.x.x route-map foo in
!
route-map foo permit 10
 match as-path 2
 set local-preference 120
!
ip as-path access-list 2 permit  ^1880_
```



Local Preference Displayed

```
R3#sho ip bgp 22.0.0.0/8
BGP routing table entry for 22.0.0.0/8, version 8
Paths: (2 available, best #1, table Default-IP-Routing-Table)
Flag: 0x820
  Advertised to update-groups:
        1    2
  7
    7.7.7.7 from 7.7.7.7 (7.7.7.7)
      Origin IGP, metric 0, localpref 200, valid, external, best
  2
    2.2.2.2 from 2.2.2.2 (22.22.22.2)
      Origin IGP, metric 0, localpref 100, valid, external
R3#
```

BGP Path Selection – Locally Originated



1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF

5. Prefer locally originated route

BGP Path Selection – Locally Originated

- R3#show run
- <output truncated>
- !
- router bgp 3
- synchronization
- bgp log-neighbor-changes
- network 22.0.0.0
- neighbor 2.2.2.2 remote-as 2
- neighbor 6.6.6.6 remote-as 3
- neighbor 7.7.7.7 remote-as 7

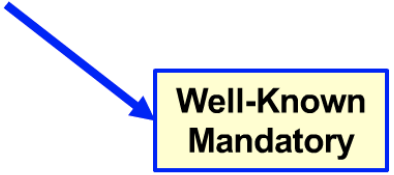
```
R3#sho ip bgp
BGP table version is 11, local router ID is 2.2.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 22.0.0.0         0.0.0.0                  0         32768 i
*                   7.7.7.7                  0    200      0 7 i
*                   2.2.2.2                  0             0 2 i
*> 33.0.0.0         2.2.2.2                  0             0 2 ?
R3#_
```

BGP Path Selection – AS Path

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF

5. Prefer locally originated route

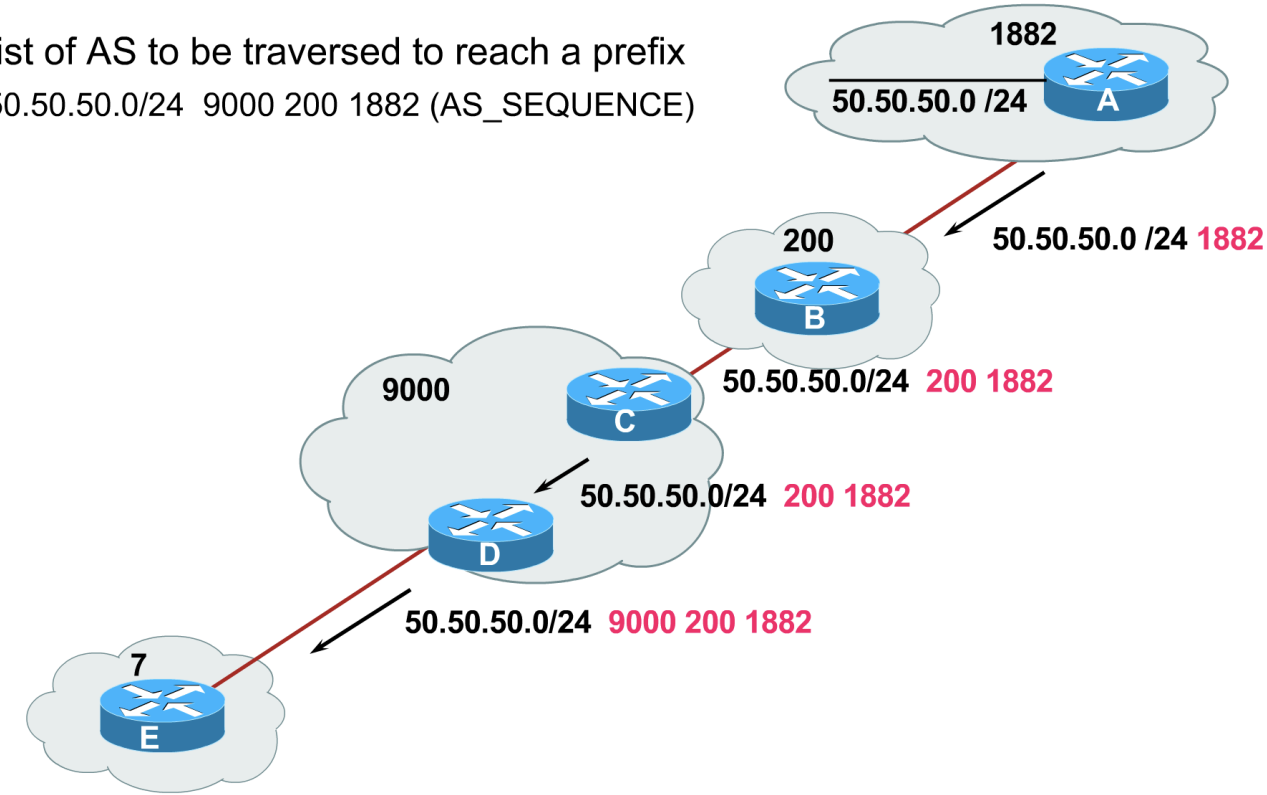
6. Shortest AS_PATH



BGP Path Selection – AS Path

- A list of AS to be traversed to reach a prefix

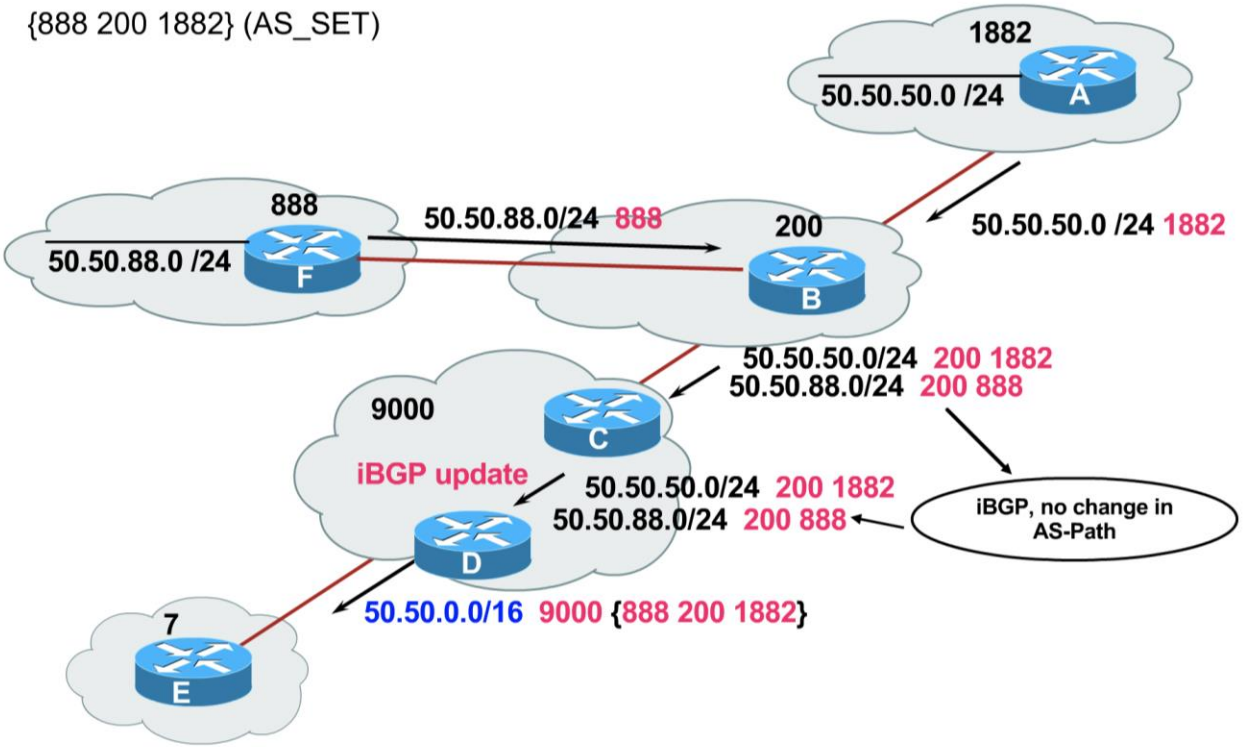
  50.50.50.0/24 9000 200 1882 (AS_SEQUENCE)



AS Path (Aggregation: AS_Set)

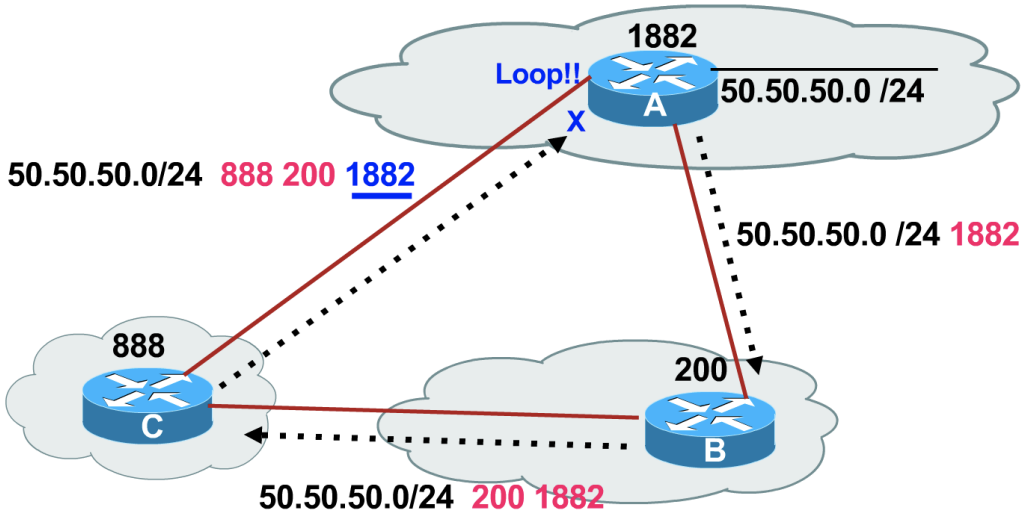
- Path includes one or more members of a set

  {888 200 1882} (AS_SET)

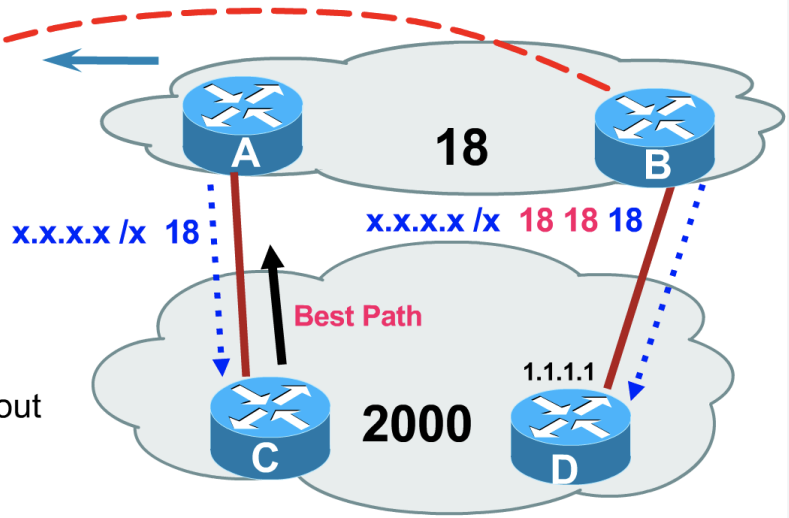


AS Path (Aggregation: AS_Set)

- Used for loop detection!



AS Path Prepending



- Configuration (rtr B):

```
router bgp 18
neighbor 1.1.1.1 remote-as 2000
neighbor 1.1.1.1 route-map prepend out
!
route-map prepend permit 10
 match as-path 2
 set as-path prepend 18 18
!
ip as-path access-list 2 permit  .*
```

All Prefixes

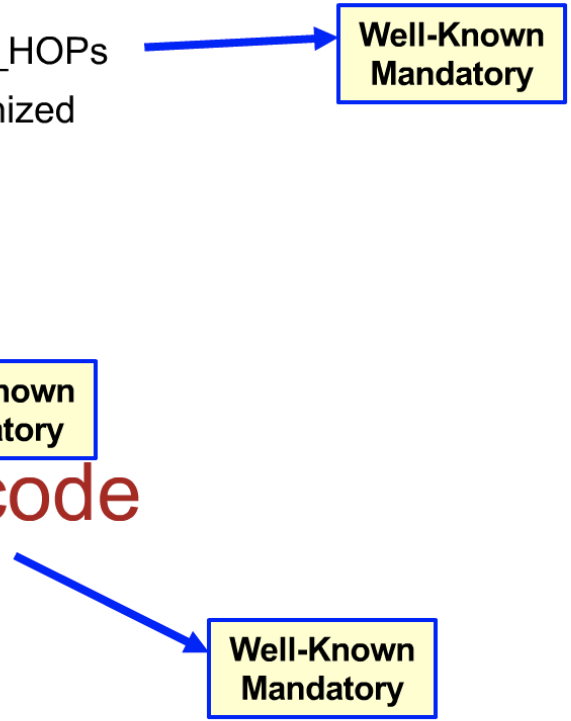
BGP Path Selection – Origin

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF

5. Prefer locally originated route
6. Shortest AS_PATH

7. Lowest ORIGIN code

IGP < EGP < incomplete



BGP Path Selection – Origin

- Origin of the prefix
- Values:

  IGP (i) = via network command

  EGP (e) = learned from EGP

  incomplete (?) = redistribution

```
R3#sho ip bgp
BGP table version is 11, local router ID is 2.2.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 22.0.0.0         0.0.0.0                  0         32768 i
*                   7.7.7.7                  0    200      0 7 i
*                   2.2.2.2                  0             0 2 i
*> 33.0.0.0         2.2.2.2                  0             0 2 ?
R3#_
```

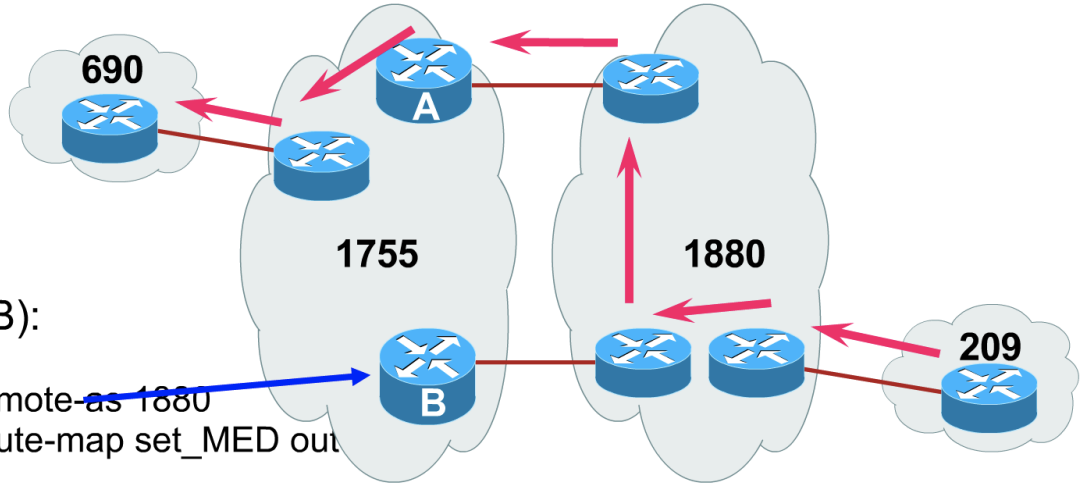
BGP Path Selection – MED

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF
5. Prefer locally originated route
6. Shortest AS_PATH
7. Lowest ORIGIN code
   IGP < EGP < incomplete
8. Lowest Multi-Exit Discriminator (MED)

BGP Path Selection – MED

- Indication (to external peers) of the preferred path into an AS
  - used in multiple entry AS
  - non-transitive
- Compared only for routes from the same AS
- Lower MED value is more preferable (default = 0)

BGP Path Selection – MED



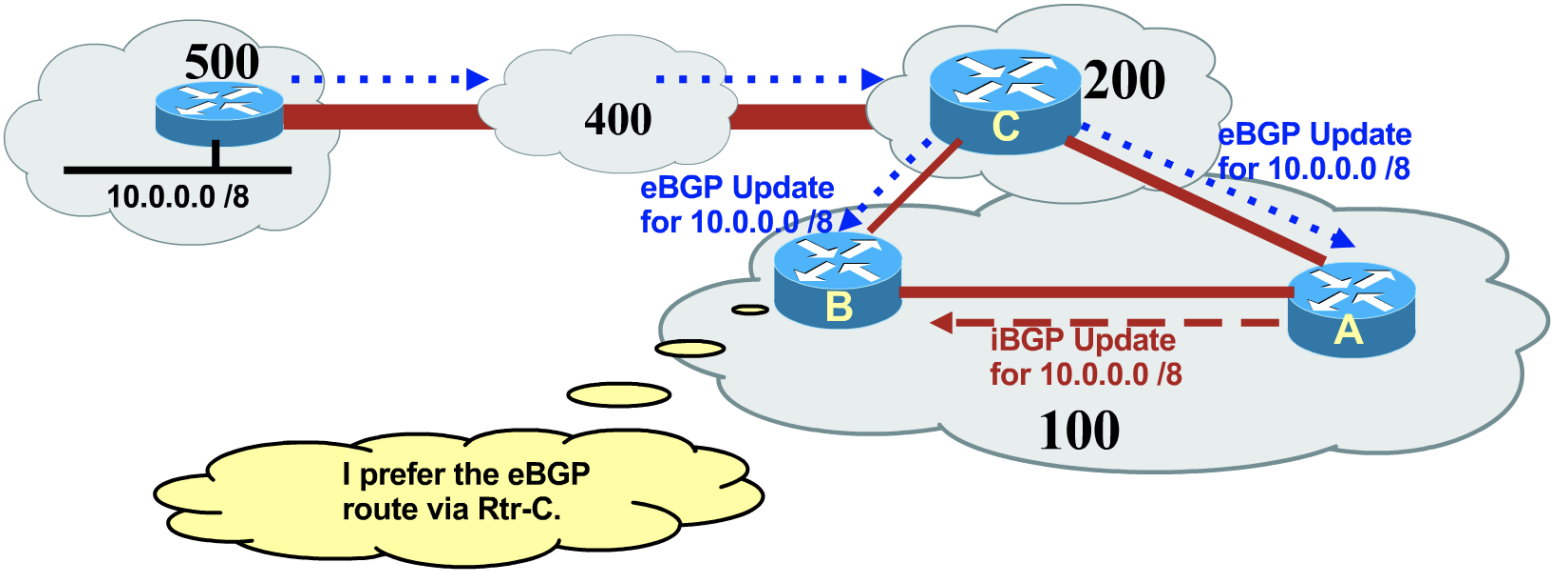
- Configuration (rtr B):

```
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$
```

BGP Path Selection – eBGP vs iBGP

1. Only consider paths with reachable NEXT_HOPs
2. Do not consider iBGP path if not synchronized
3. Highest WEIGHT
4. Highest LOCAL_PREF
5. Prefer locally originated route
6. Shortest AS_PATH
7. Lowest ORIGIN code
   IGP < EGP < incomplete
8. Lowest Multi-Exit Discriminator (MED)
9. Prefer an *External* path over an *Internal* one

BGP Path Selection – eBGP vs iBGP



BGP Path Selection (Cont.)

7 Lowest ORIGIN code
IGP < EGP < incomplete

8 Lowest Multi-Exit Discriminator (MED)…default is 0.
Default: Considered only if paths are from the same neighbor AS

9 Prefer an *External* path over an *Internal* one

10 Lowest IGP metric to the NEXT_HOP

11 For eBGP paths

IF multipath is enabled, the router may install up to N parallel paths in the routing table but best-path must still be determined.

Select the Oldest unless updates arrived at the same time.

IF updates arrived at the same time, see next bullet

BGP Path Selection (Cont.)

12 Lowest Router-ID

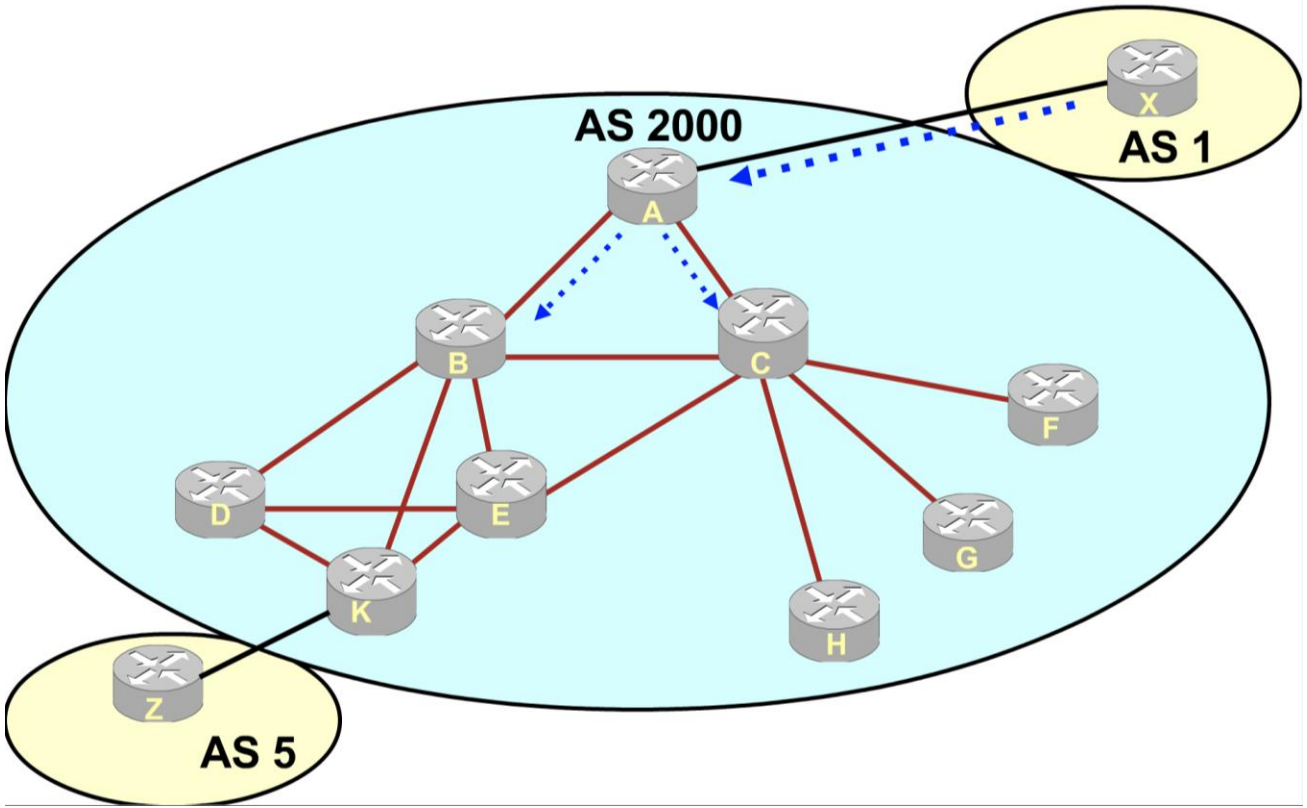
Originator-ID is considered for reflected routes

13 Shortest Cluster-List

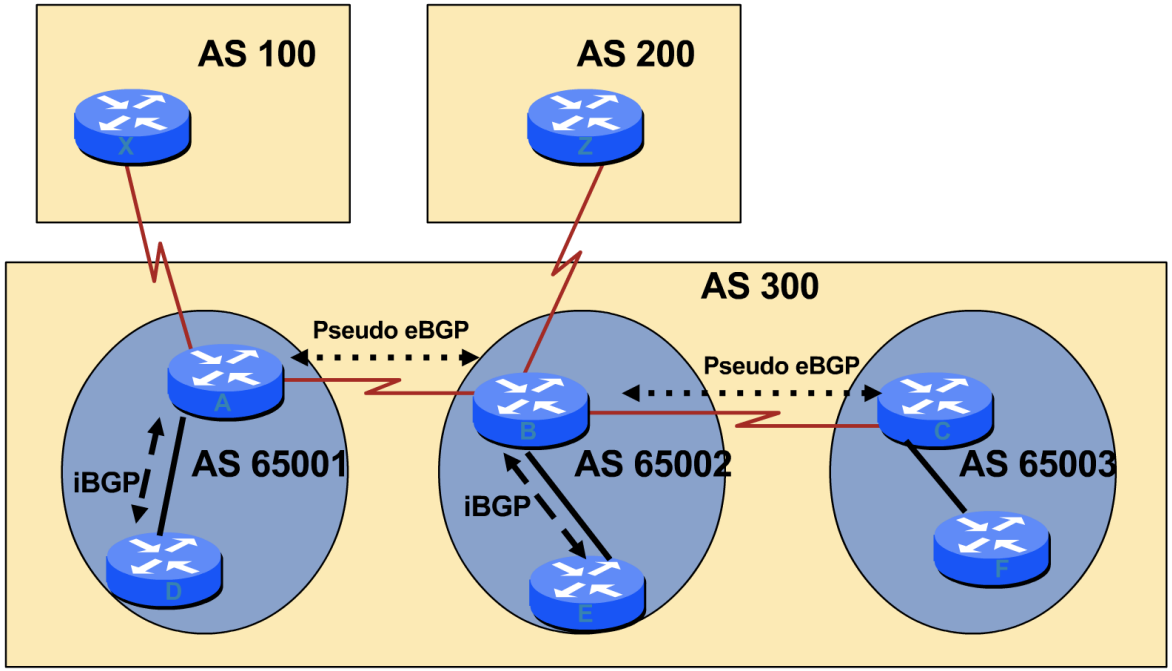
Client must be aware of RR attributes!

14 Lowest neighbor IP address

The iBGP Problem…



Confederations



- Divide the AS into sub-AS's

Confederations

- Solves iBGP mesh problem
- Divide the AS into sub-AS's
  - Use private AS#s for Sub AS's
- Visible to outside world as single AS
- Preserve local preference, MED, and NEXT_HOP
- iBGP speakers within a sub-AS are fully meshed
- Route-reflectors can be used within a Sub AS

BGP with Confederations

- AS_Path attribute changed by adding:

  AS_Confederation_Set attribute (sub-AS listing in an unordered manner…used for aggregated routes).

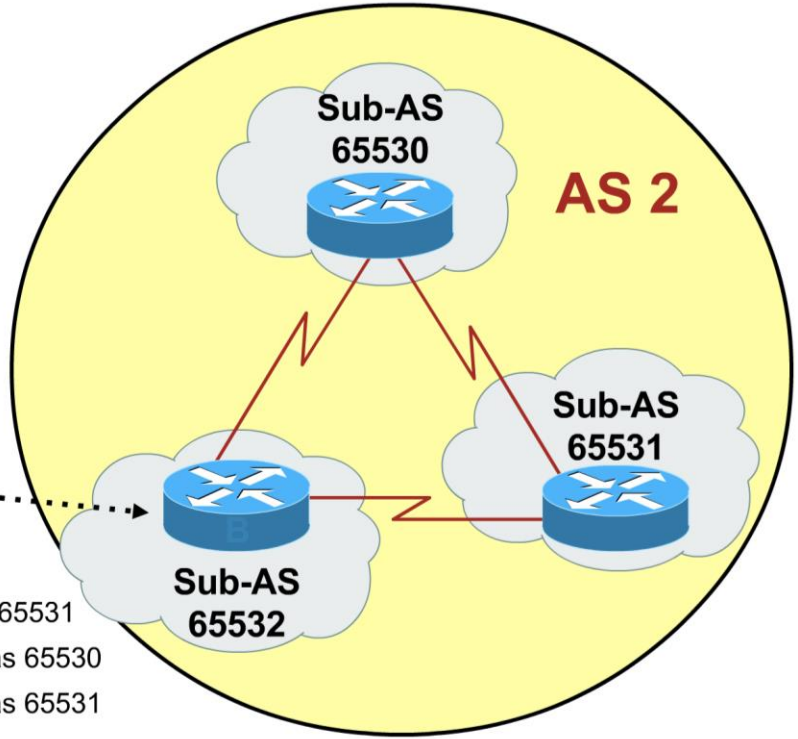
  - 10.0.0.0 /16 (65531 65500) 1 246 5000 ?

  AS_Confederation_Sequence (sub-AS listing in the order of visit that the update traversed…used for regular non-aggregated prefixes)

  - 10.0.0.0 /16 (65531 65500) 1 246 5000 ?

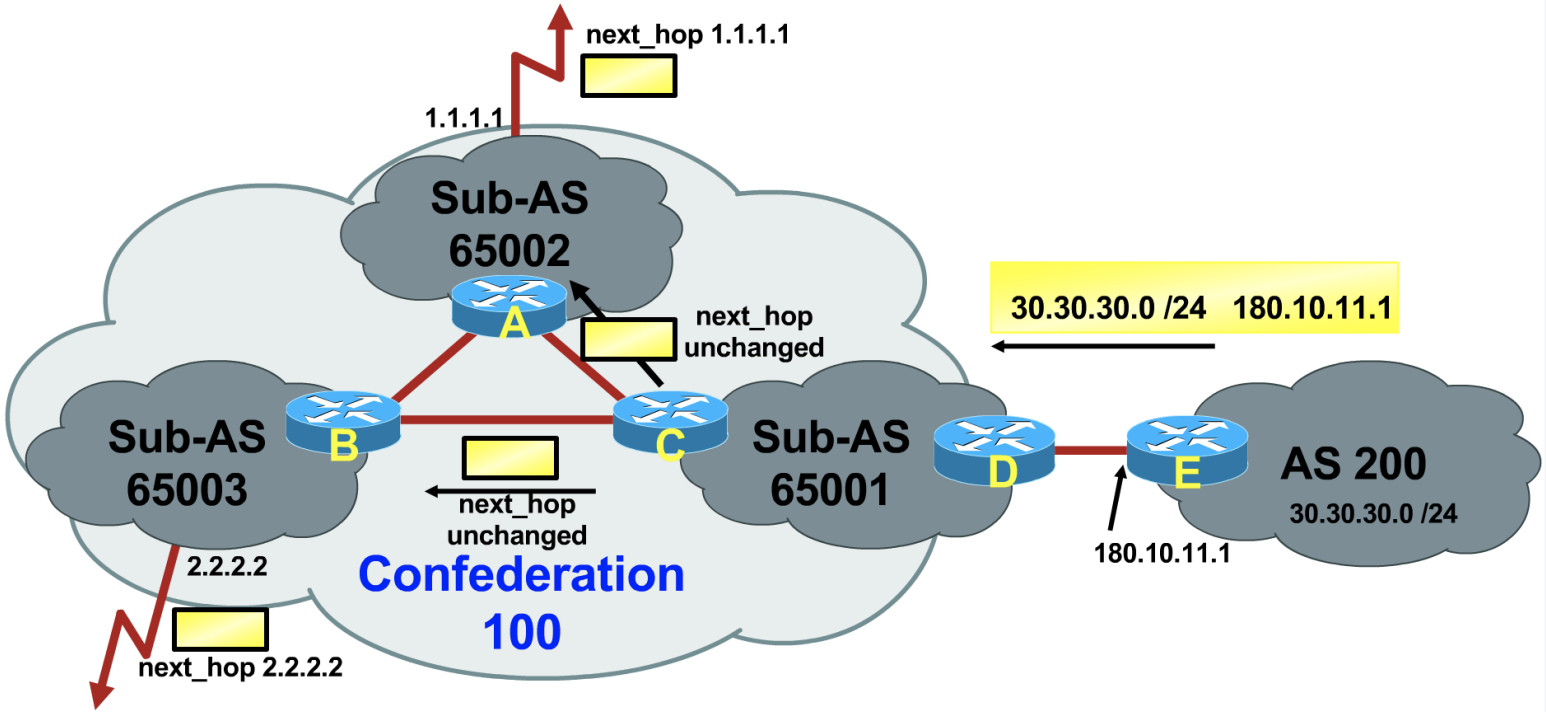
Both of the above will be removed from the AS-Path before sending an update to another external AS.

Confederations



- B#
- !
- router bgp 65532
- bgp confederation identifier 2
- bgp confederation peers 65530 65531
- neighbor 141.153.12.1 remote-as 65530
- neighbor 141.153.17.2 remote-as 65531

Confederations – Next Hop



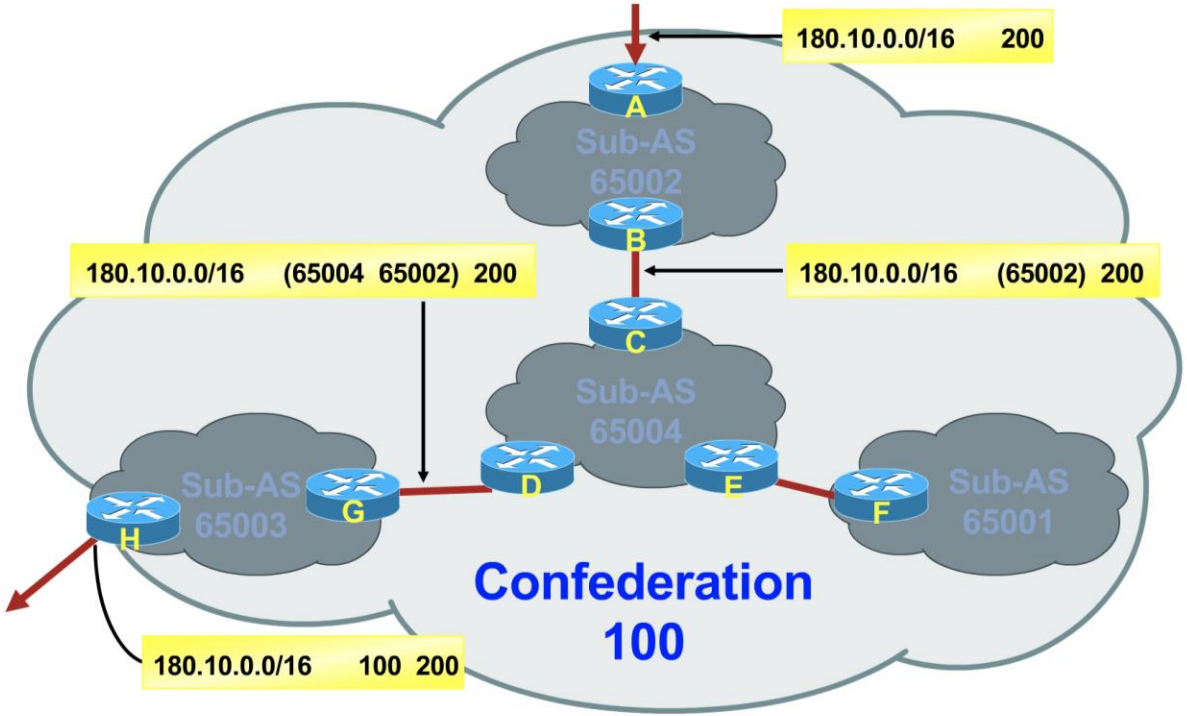
Route Propagation Decisions

- Same as with “normal” BGP:
  - From peer in same sub-AS → only to external peers
  - From external peers → to all neighbors
- “External peers” refers to
  - Peers outside the confederation
  - Peers in a different sub-AS
    - Preserve LOCAL_PREF, MED and NEXT_HOP

Confederations – AS Path

- Sub-AS traversed are carried as part of AS_PATH (AS_CONFED_SEQUENCE or AS_CONFED_SET) for loop avoidance
  - Not counted as regular AS when comparing AS_PATH
  - Paths with only confederation ASNs in the AS_PATH are skipped during MED comparison
    - bgp bestpath med confed

Confederations – AS Path (Cont.)



Confederations – AS Path (Cont.)

```
show ip bgp 199.227.0.22
BGP routing table entry for 199.227.0.0/16, version 3027825
Paths: (3 available, best #2, advertised over IBGP, EBGP)

  2548 1239 3064, (aggregated by 3064 199.227.255.1)
    207.22.128.22 from 207.22.128.22 (209.4.29.1)
      Origin IGP, localpref 100, valid, internal, atomic-aggregate

  (65502) 701 1239 3064, (aggregated by 3064 199.227.255.1)
    209.4.3.1 from 209.4.3.1 (209.4.3.9)
      Origin IGP, localpref 100, valid, confed-external, atomic-aggregate, best

  (65500) 2548 1239 3064, (aggregated by 3064 199.227.255.1)
    207.100.210.22 from 207.100.210.22
      Origin IGP, localpref 50, valid, confed-external, atomic-aggregate
```

BGP