BGP
Protocol & Configuration

AfNOG

Border Gateway Protocol (BGP4)

- Case Study 1, Exercise 1: Single upstream
- Part 6: BGP Protocol Basics
- Part 7: BGP Protocol - more detail
- Case Study 2, Exercise 2: Interior BGP
- Part 8: Routing Policy and Filtering
- Exercise 3: Filtering on AS-path
- Exercise 4: Filtering on prefix-list
- Part 9: More detail than you want
- Part 10: BGP and Network Design

BGP Part 9

More detail than you want

BGP Attributes

Synchronization

Path Selection

BGP Path Attributes: Why ?

- Encoded as Type, Length & Value (TLV)
- Transitive/Non-Transitive attributes
- Some are mandatory
- Used in path selection
- To apply policy for steering traffic

BGP Attributes

- Used to convey information associated with NLRI
  - AS path
  - Next hop
  - Local preference
  - Multi-Exit Discriminator (MED)
  - Community
  - Origin
  - Aggregator

Local Preference

- Not used by eBGP, mandatory for iBGP
- Default value of 100 on Cisco IOS
- Local to an AS
- Used to prefer one exit over another
- Path with highest local preference wins

Local Preference
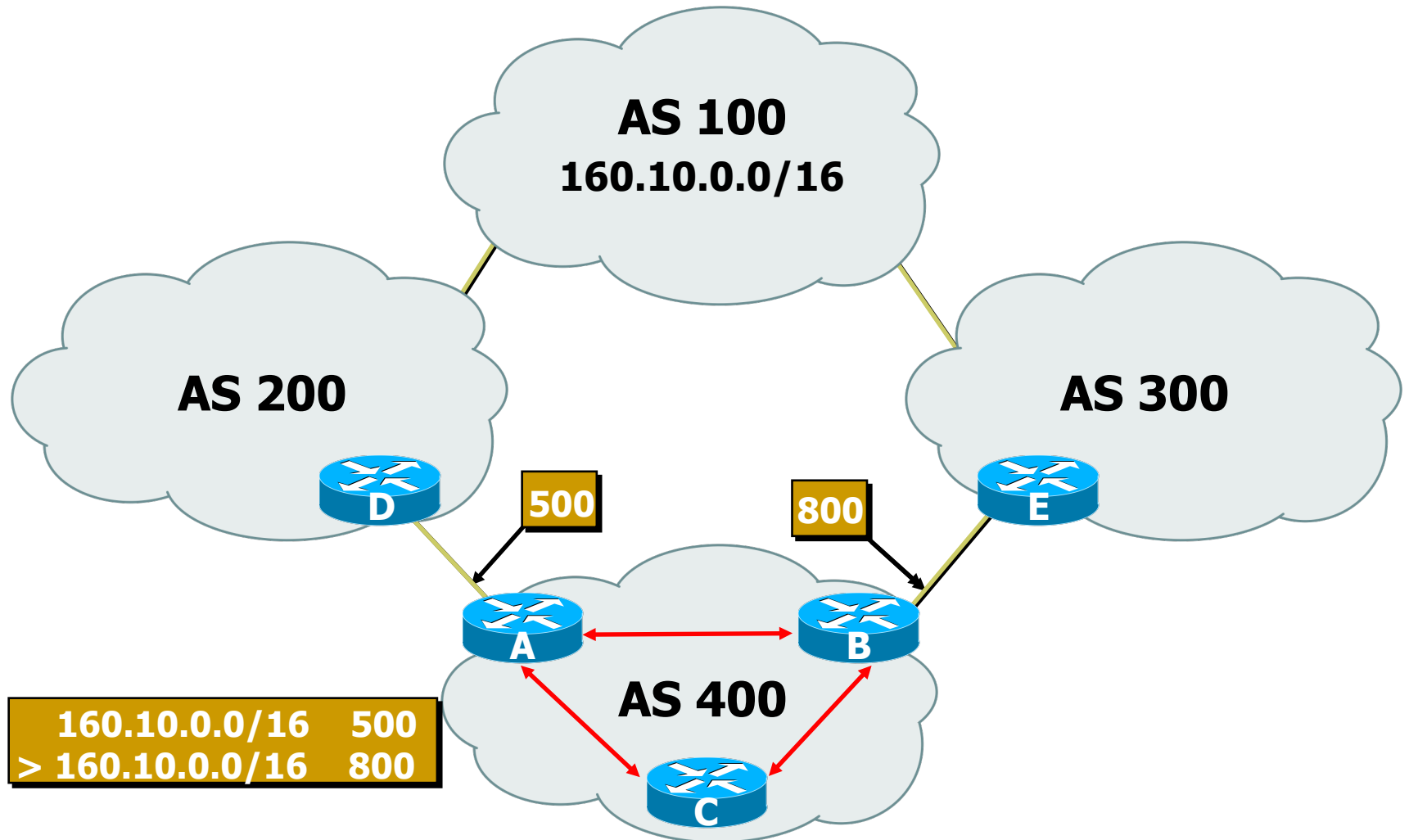
Multi-Exit Discriminator

- Non-transitive and optional
- Represented as a numerical value
  - Range 0x0 – 0xffffffff (if unset 0 is assumed)
- Used to convey relative preference of entry points to an AS
- Comparable if the paths are from the same AS
  - `bgp always-compare-med` allows comparisons of MEDs from different ASes
- Path with the lowest MED wins
- IGP metric can be conveyed as MED

Multi-Exit Discriminator (MED)
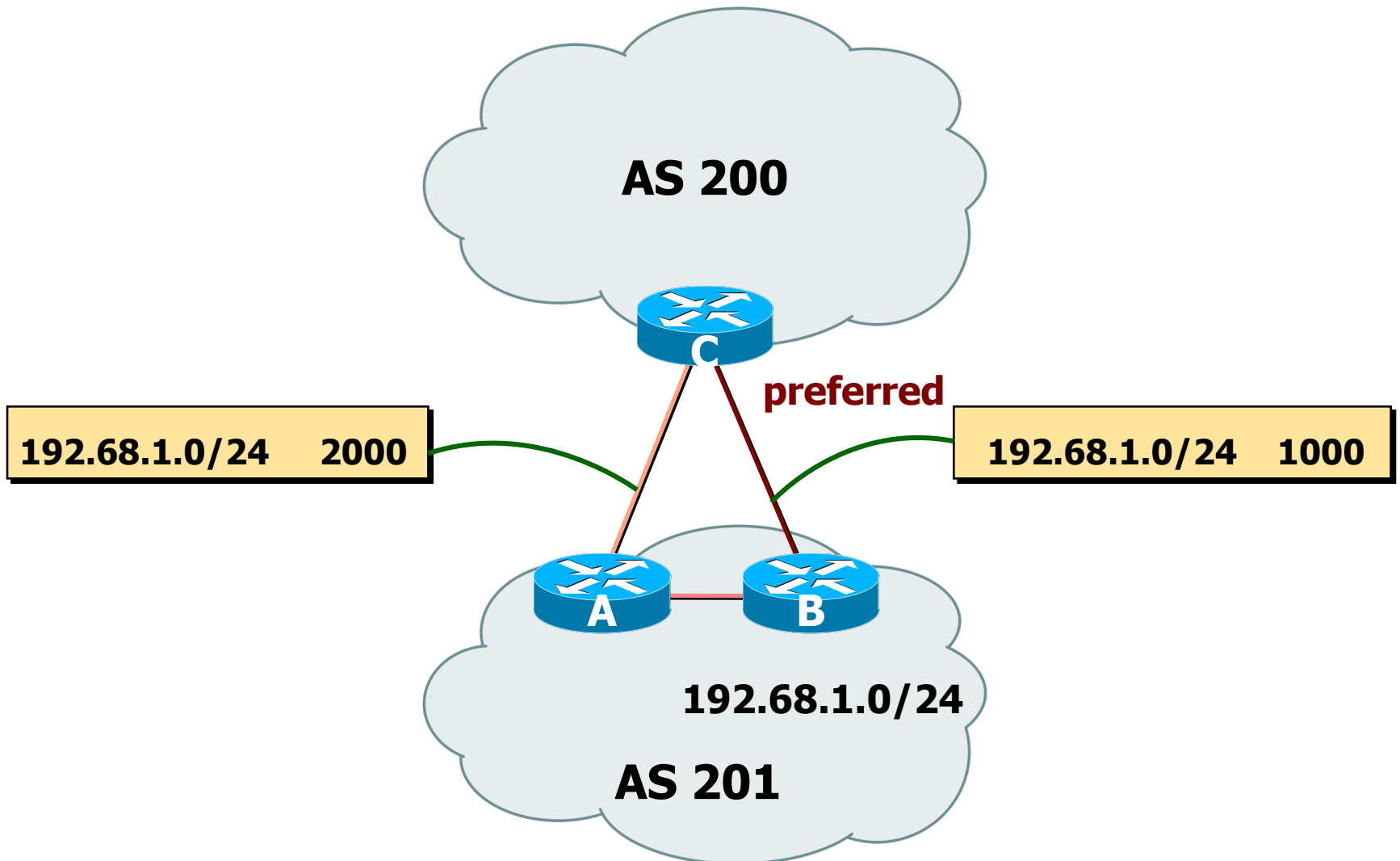
Origin

- Conveys the origin of the prefix
  - Historical attribute
- Three values:
  - IGP – from BGP network statement
    - E.g. – *network 35.0.0.0*
  - EGP – redistributed from EGP (not used today)
  - Incomplete – redistributed from another routing protocol
    - E.g. – *redistribute static*
- IGP < EGP < incomplete
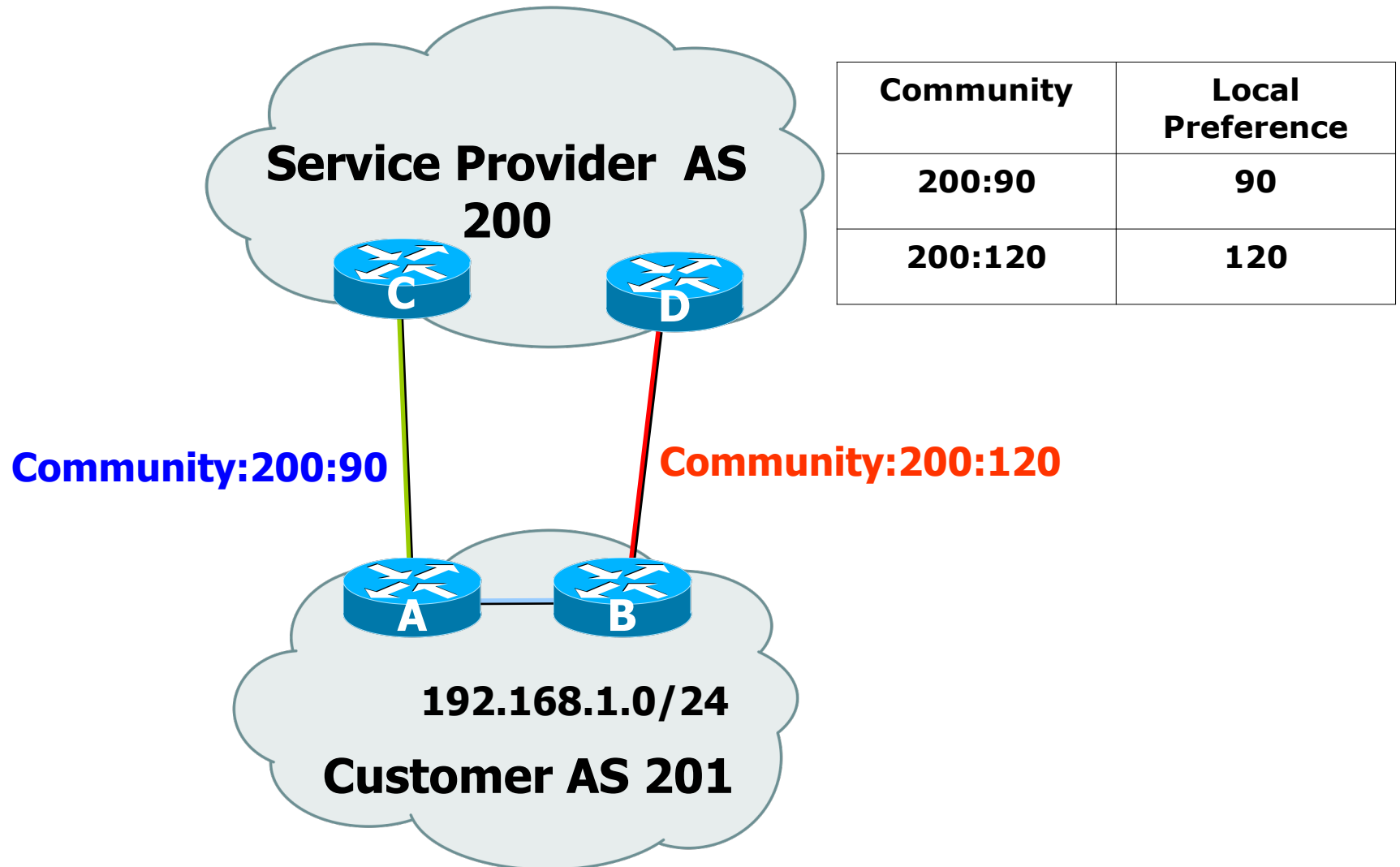 - Lowest origin code wins

Weight

Not really an attribute (Cisco proprietary)

- Used when there is more than one route to same destination
- Local to the router on which it is assigned, and not propagated in routing updates
- Default is 32768 for paths that the router originates and zero for other paths
- Routes with a higher weight are preferred when there are multiple routes to the same destination

Communities

- Transitive, Non-mandatory
- Represented as a numeric value
  - 0x0 – 0xffffffff
  - Internet convention is ASN:<0-65535>
- Used to group destinations
- Each destination could be member of multiple communities
- Flexibility to scope a set of prefixes within or across AS for applying policy

Communities

Community	Local Preference
200:90	90
200:120	120

Service Provider AS 200

Community:200:90

Community:200:120

192.168.1.0/24

Customer AS 201

Well-Known Communities

- Several well known communities
www.iana.org/assignments/bgp-well-known-communities
- no-export   65535:65281
  - do not advertise to any eBGP peers
- no-advertise   65535:65282
  - do not advertise to any BGP peer
- no-export-subconfed   65535:65283
  - do not advertise outside local AS (only used with confederations)
- no-peer   65535:65284
  - do not advertise to bi-lateral peers (RFC3765)

No-Export Community

105.7.0.0/16
105.7.X.XNo-Export

105.7.X.X

AS 100

AS 200

105.7.0.0/16
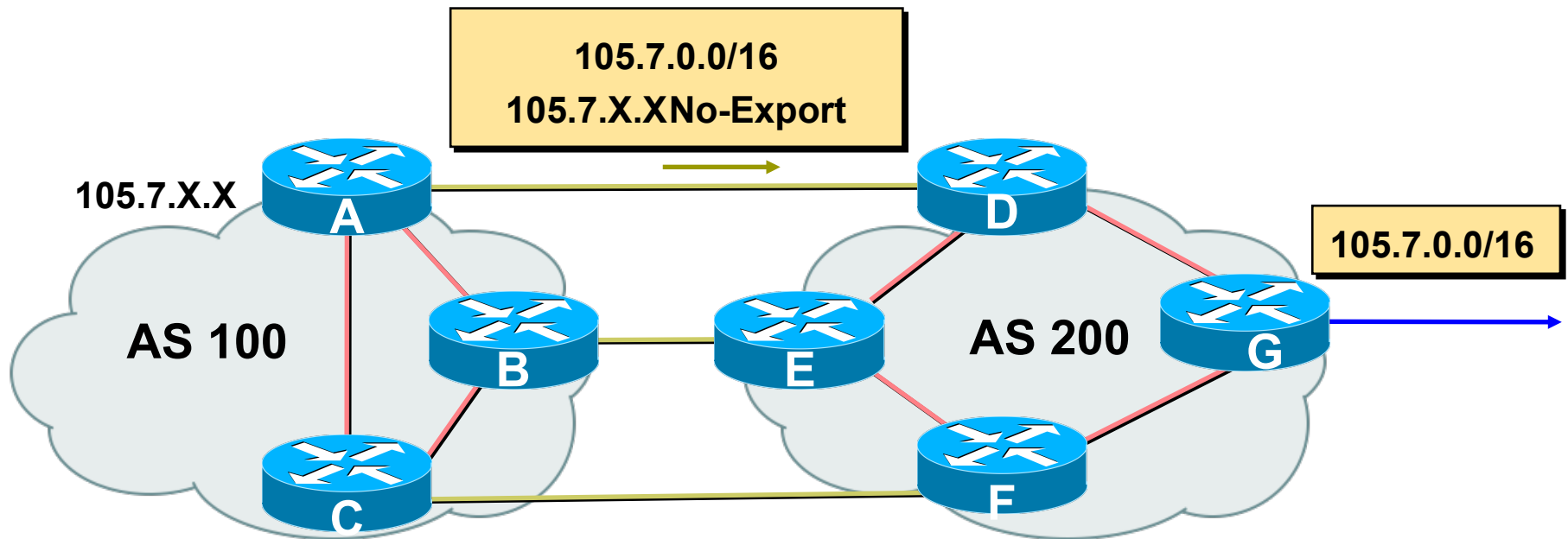
- AS100 announces aggregate and subprefixes
  - Intention is to improve loadsharing by leaking subprefixes
- Subprefixes marked with no-export community
- Router G in AS200 does not announce prefixes with no-export community set

Administrative Distance

- Routes can be learned via more than one protocol
  - Used to discriminate between them
- Route with lowest distance installed in forwarding table
- BGP defaults
  - Local routes originated on router: 200
  - iBGP routes: 200
  - eBGP routes: 20
- Does not influence the BGP path selection algorithm but influences whether BGP learned routes enter the forwarding table

FYI:
Default Administrative Distances

Route Source	Cisco	Juniper	Huawei	Brocade	Nokia/ALU
Connected Interface	**0**	**0**	**0**	**0**	**0**
Static Route	**1**	**5**	**60**	**1**	**1**
EIGRP Summary Route	5	N/A	?	N/A	N/A
External BGP	**20**	**170**	**255**	**20**	**170**
Internal EIGRP Route	90	N/A	?	N/A	N/A
IGRP	100	N/A	?	N/A	N/A
OSPF	**110**	**10**	**10**	**110**	**10**
IS-IS	**115**	**18**	**15**	**115**	**18**
RIP	120	100	100	120	100
EGP	140	N/A	N/A	N/A	N/A
External EIGRP	170	N/A	?	N/A	N/A
Internal BGP	**200**	**170**	**255**	**200**	**130**
Unknown	255	255	?	255	?

17

Cisco BGP Administrative Distance (fix)

- On all your Cisco BGP routers you should set the admin distance for internal, external, and locally originated routes (per address family) to the same distance as follows: (200 is a good choice based on the previous table)
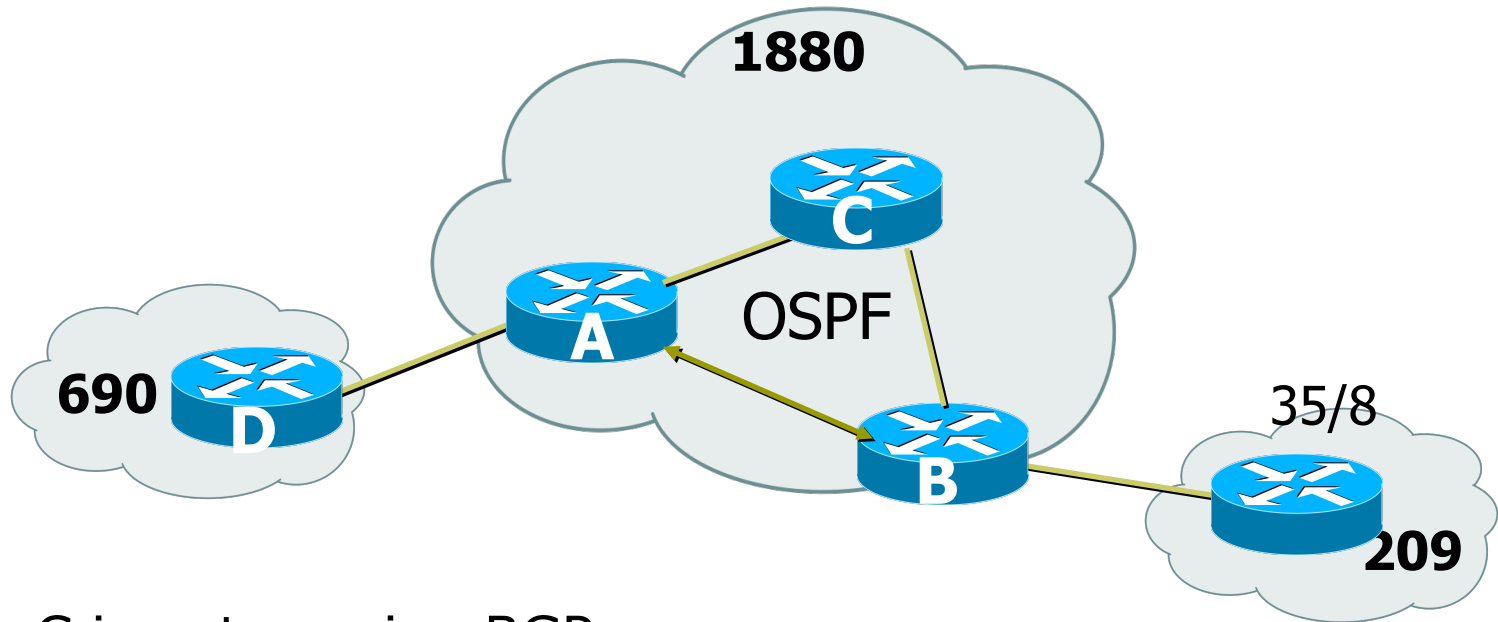
  - Syntax: distance bgp <eBGP> <iBGP> <local>

```
!
router bgp 100
    address-family ipv4
        distance bgp 200 200 200
    address-family ipv6
        distance bgp 200 200 200
!
```

Synchronization

1880

C

A   OSPF

690

D

35/8

B

209

- C is not running BGP
- A won't advertised 35/8 to D until the IGP is in sync
- Turn synchronization off!

```
router bgp 1880
 no synchronization
```

Synchronization

- In Cisco IOS, BGP does not advertise a route before all routers in the AS have learned it via an IGP
  - Default in IOS prior to 12.4; very unhelpful to most ISPs
- Disable synchronization if:
  - AS doesn't pass traffic from one AS to another, or
  - All transit routers in AS run BGP, or
  - iBGP is used across backbone
- You should always use iBGP
  - so, always use "no synchronization"

BGP route selection (bestpath)

- Route has to be synchronized
  - Only if synchronization is enabled
  - Prefix must be in forwarding table
- Next-hop has to be accessible
  - Next-hop must be in forwarding table
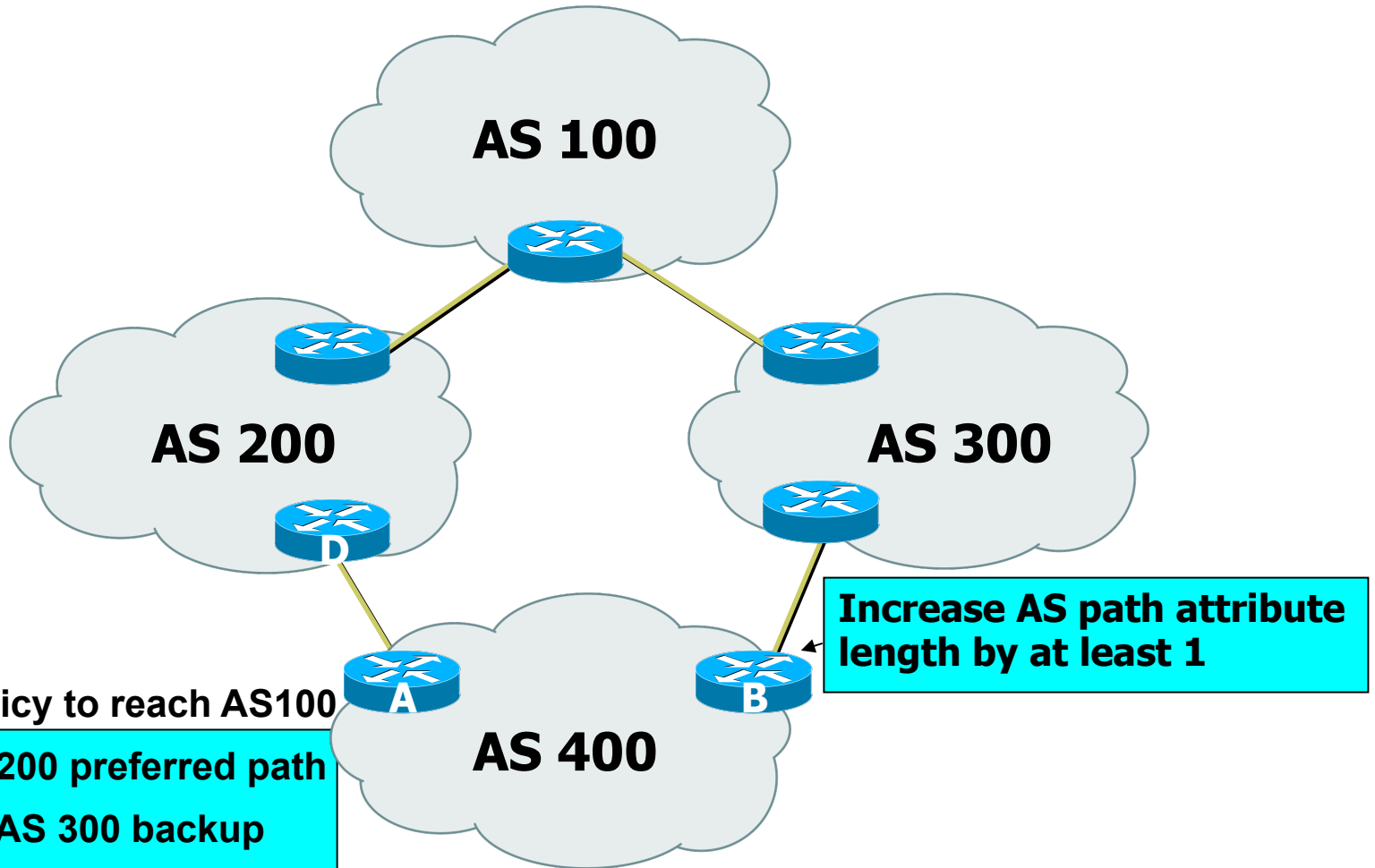- Largest weight
- Largest local preference

BGP route selection (bestpath)

- Locally sourced
  - Via redistribute or network statement
- Shortest AS path length
  - Number of ASes in the AS-PATH attribute
- Lowest origin
  - IGP < EGP < incomplete
- Lowest MED
  - Compared from paths from the same AS

BGP route selection (bestpath)

- External before internal
  - Choose external path before internal
- Closest next-hop
  - Lower IGP metric, nearest exit to router
- Lowest router ID
- Lowest IP address of neighbour

BGP Route Selection...

AS 100

AS 200

AS 300

D

Increase AS path attribute length by at least 1

A

B

AS 400's Policy to reach AS100

AS 400

AS 200 preferred path

AS 300 backup

Summary: Attributes displayed in a BGP RIB

```
Router1>sh ip bgp

BGP table version is 16, local router ID is 10.10.15.241

Status codes: s suppressed, d damped, h history, * valid, > best, i — internal,

              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,

              x best-external, a additional-path, c RIB-compressed,

Origin codes: i – IGP, e – EGP, ? – incomplete

RPKI validation codes: V valid, I invalid, N Not found


      Network          Next Hop          Metric LocPrf Weight Path
 *>   10.10.0.0/26     0.0.0.0                0          32768 i
 * i  10.10.0.0/20     10.10.15.226           0     100      0 i
 * i                   10.10.15.225           0     100      0 i
 *>                    0.0.0.0                0          32768 i
 *>i  10.10.0.64/26    10.10.15.225           0     100      0 i
 *>i  10.10.0.128/26   10.10.15.226           0     100      0 i
 * i  10.20.0.0/26     10.10.15.226           0     100      0 20 i
 *>i                   10.10.15.225           0     100      0 20 i
 * i  10.20.0.0/20     10.10.15.226           0     100      0 20 i
 *>i                   10.10.15.225           0     100      0 20 i
```