

BGP Attributes and Path Selection



ISP Training Workshops

BGP Attributes



The “tools” available for the job

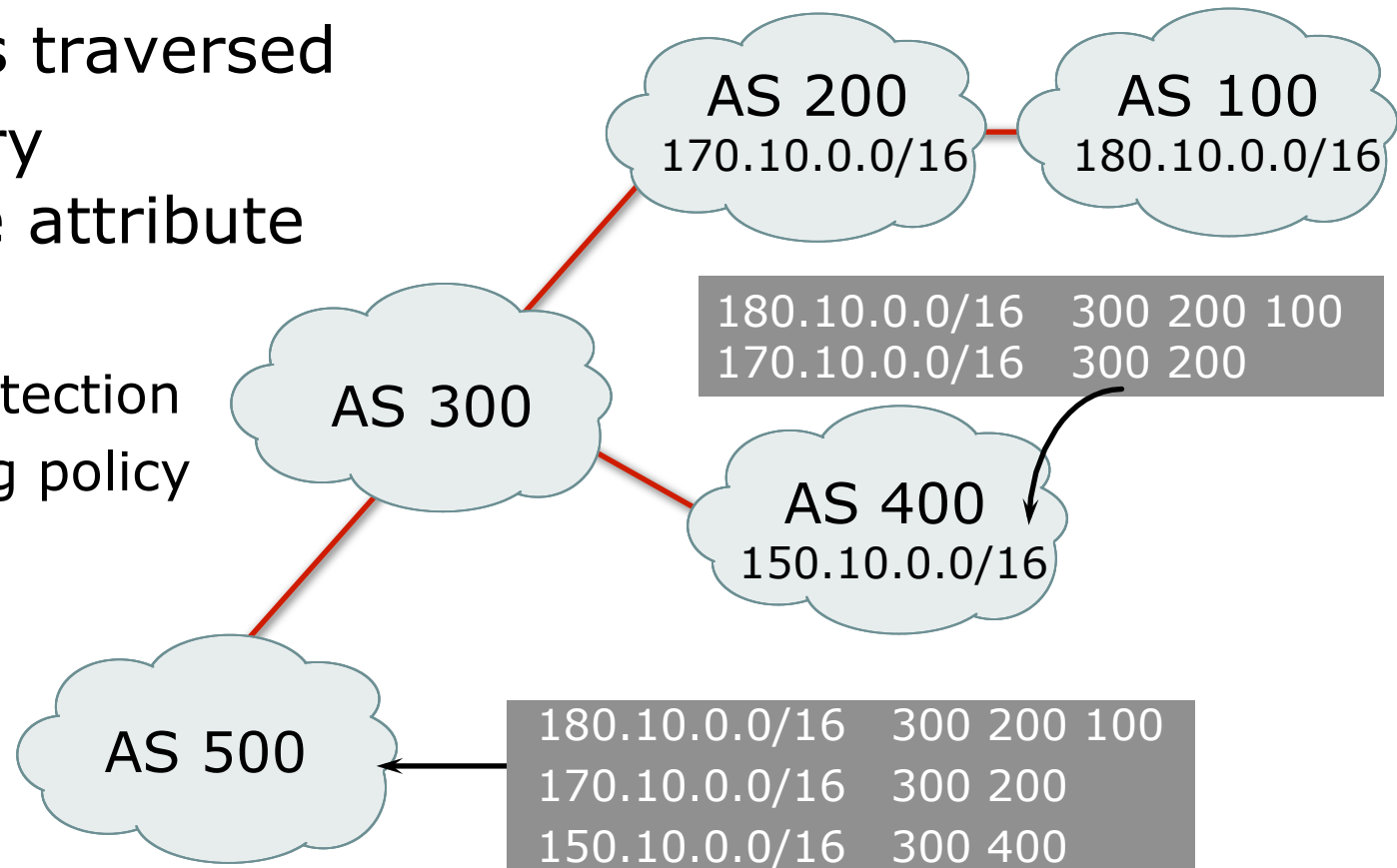
What Is an Attribute?

...	Next Hop	AS Path	MED
-----	----------	---------	-----	-----	-----

- ❑ Part of a BGP Update
- ❑ Describes the characteristics of prefix
- ❑ Can either be transitive or non-transitive
- ❑ Some are mandatory

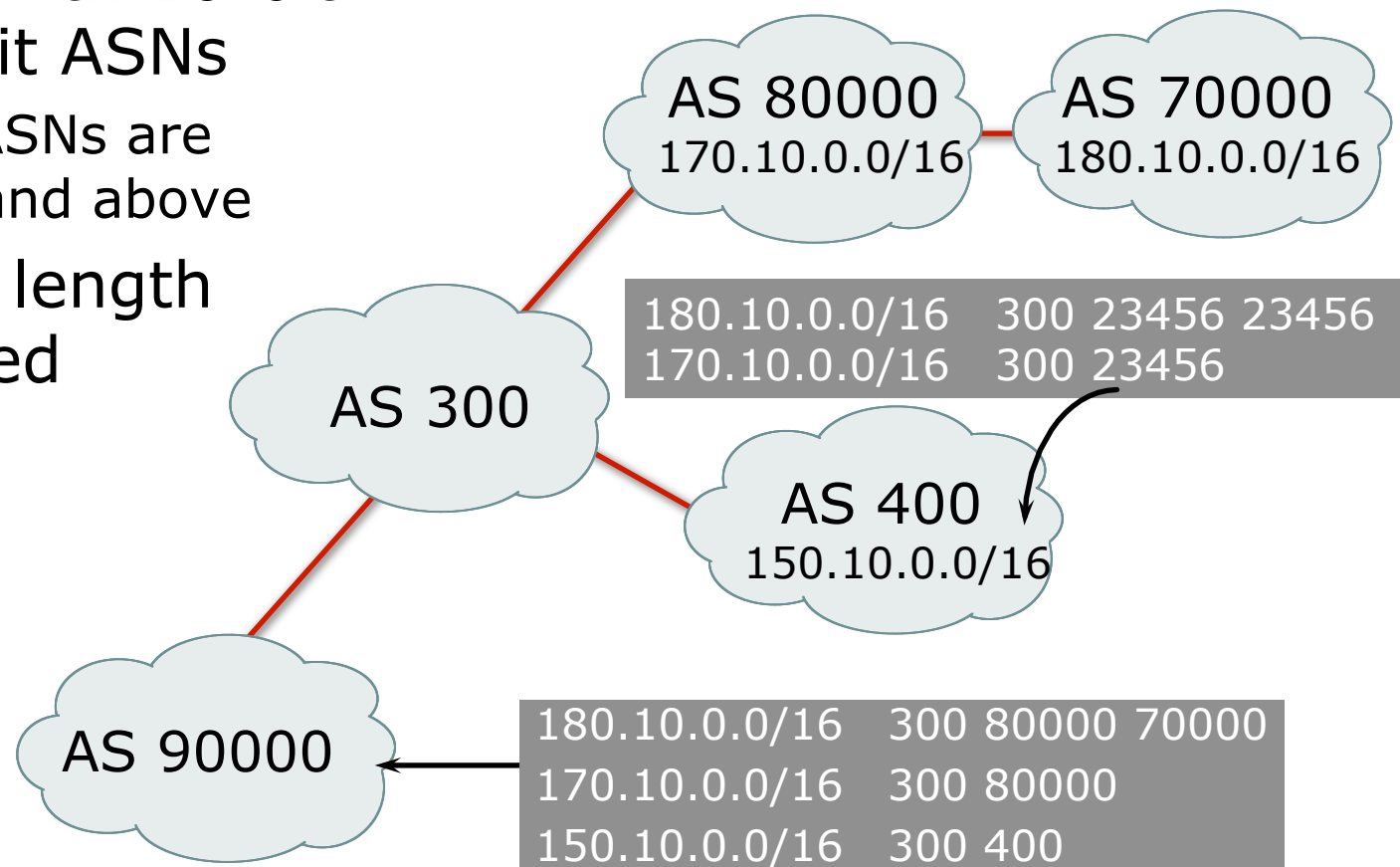
AS-Path

- ❑ Sequence of ASes a route has traversed
- ❑ Mandatory transitive attribute
- ❑ Used for:
 - Loop detection
 - Applying policy

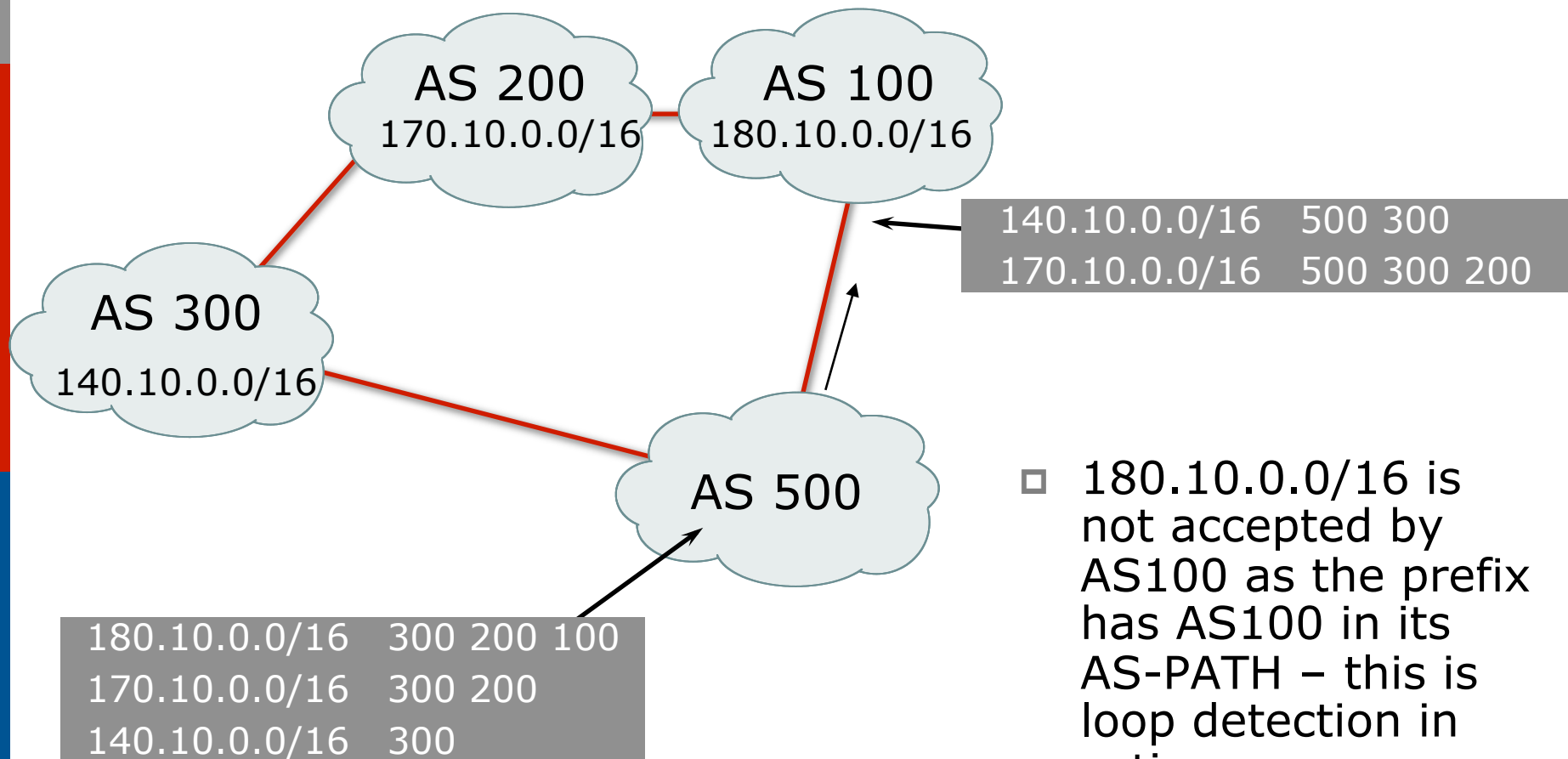


AS-Path (with 16 and 32-bit ASNs)

- ❑ Internet with 16-bit and 32-bit ASNs
 - 32-bit ASNs are 65536 and above
- ❑ AS-PATH length maintained

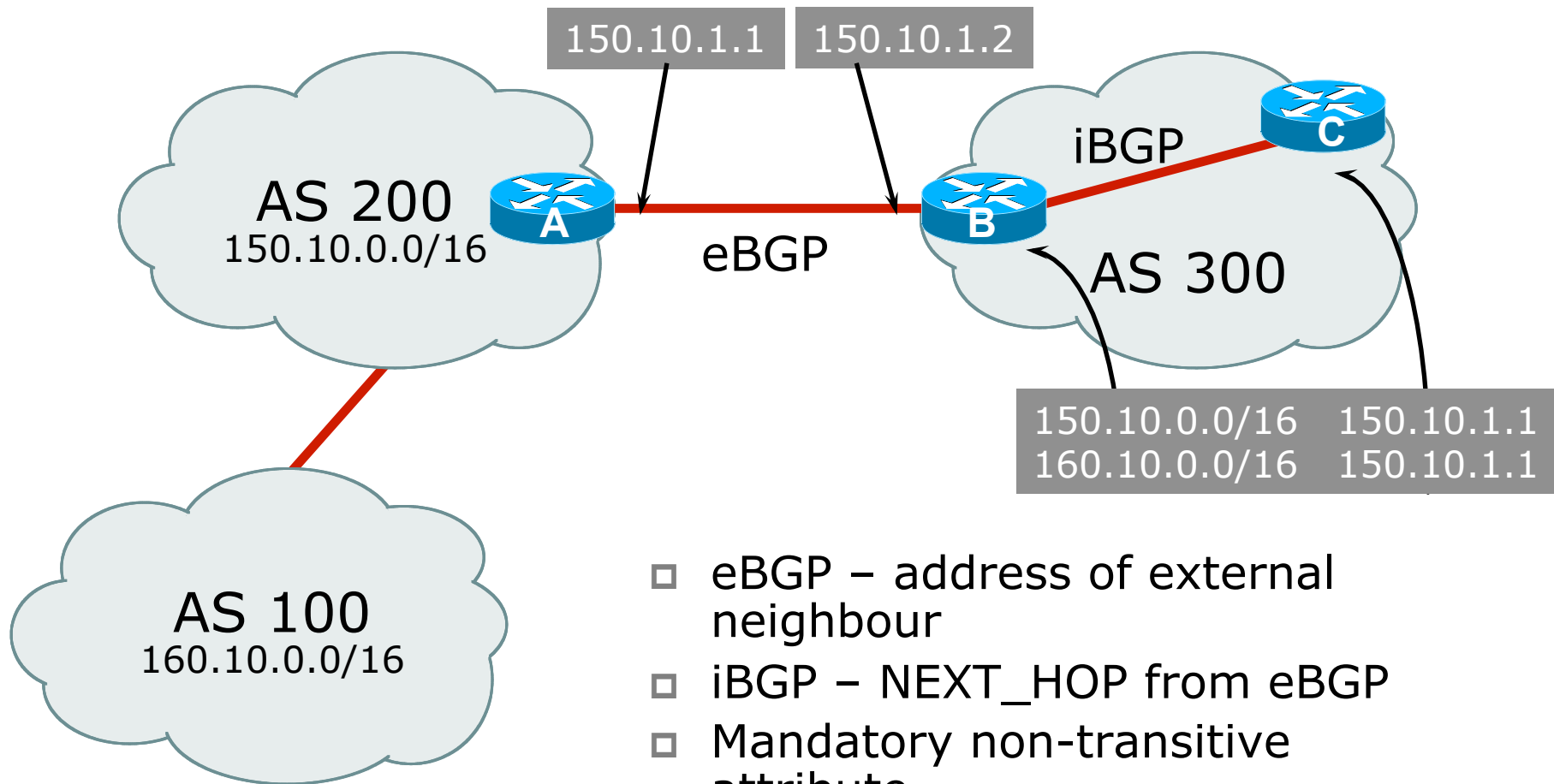


AS-Path loop detection



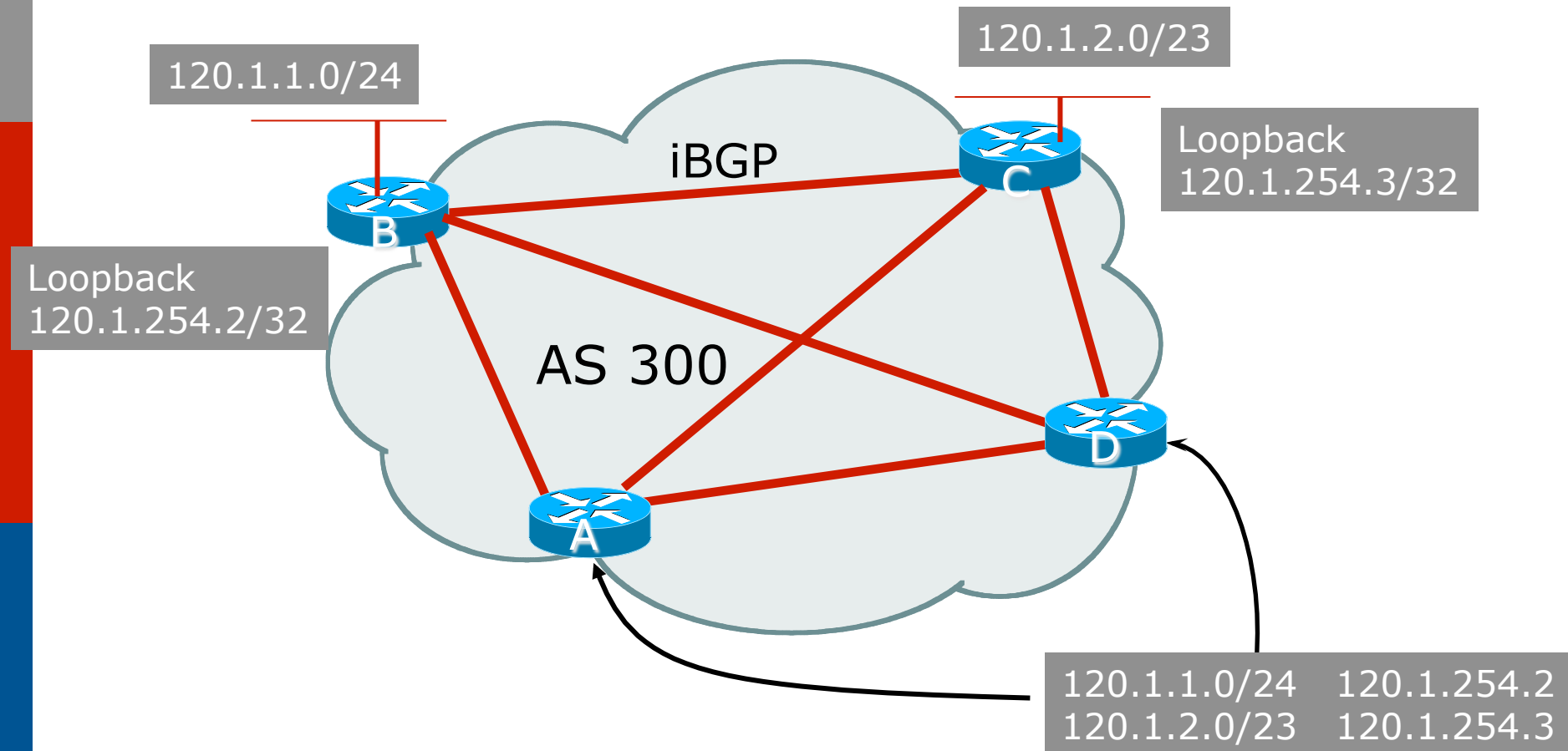
- ❑ 180.10.0.0/16 is not accepted by AS100 as the prefix has AS100 in its AS-PATH – this is loop detection in action

Next Hop



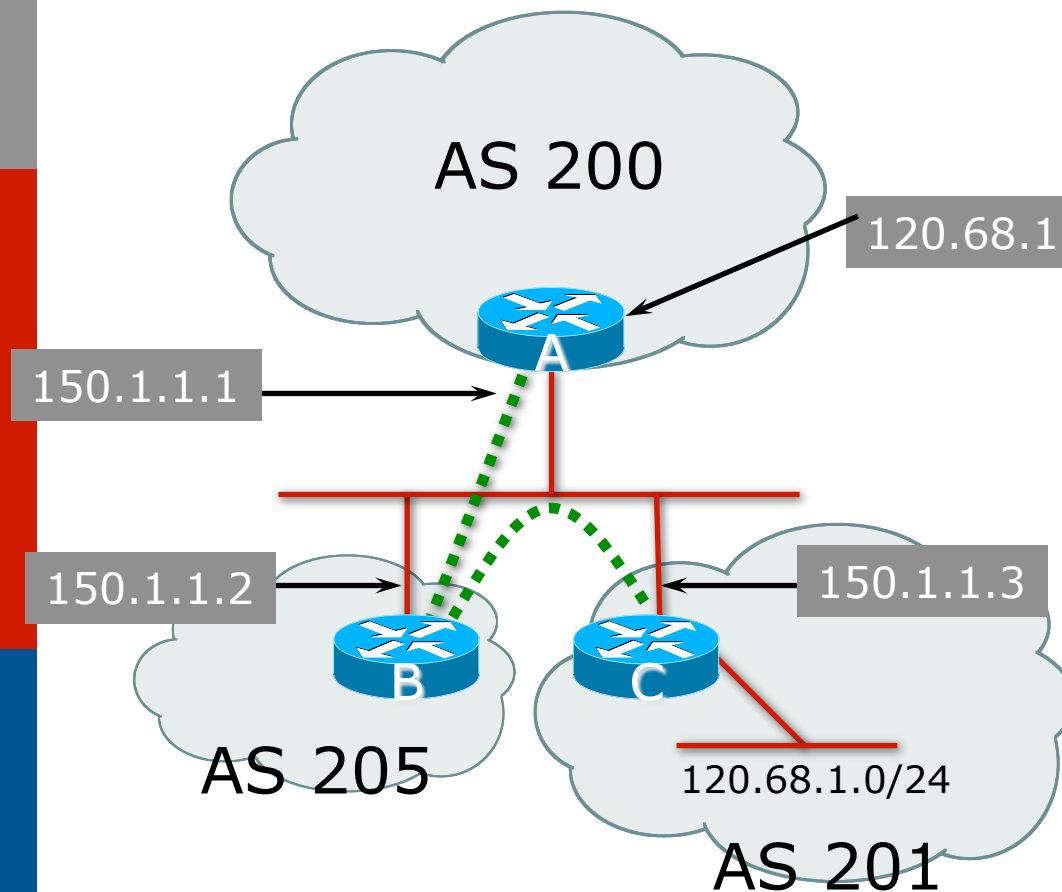
- ❑ eBGP – address of external neighbour
- ❑ iBGP – NEXT_HOP from eBGP
- ❑ Mandatory non-transitive attribute

iBGP Next Hop



- ❑ Next hop is ibgp router loopback address
- ❑ Recursive route look-up

Third Party Next Hop



- ❑ eBGP between Router A and Router B
- ❑ eBGP between Router B and Router C
- ❑ 120.68.1/24 prefix has next hop address of 150.1.1.3 – this is used by Router A instead of 150.1.1.2 as it is on same subnet as Router B

Next Hop Best Practice

- ❑ Cisco IOS default is for external next-hop to be propagated unchanged to iBGP peers
 - This means that IGP has to carry external next-hops
 - Forgetting means external network is invisible
 - With many eBGP peers, it is unnecessary extra load on IGP
- ❑ ISP Best Practice is to change external next-hop to be that of the local router

```
neighbor x.x.x.x next-hop-self
```

Next Hop (Summary)

- ❑ IGP should carry route to next hops
- ❑ Recursive route look-up
- ❑ Unlinks BGP from actual physical topology
- ❑ Use “next-hop-self” for external next hops
- ❑ Allows IGP to make intelligent forwarding decision

Origin

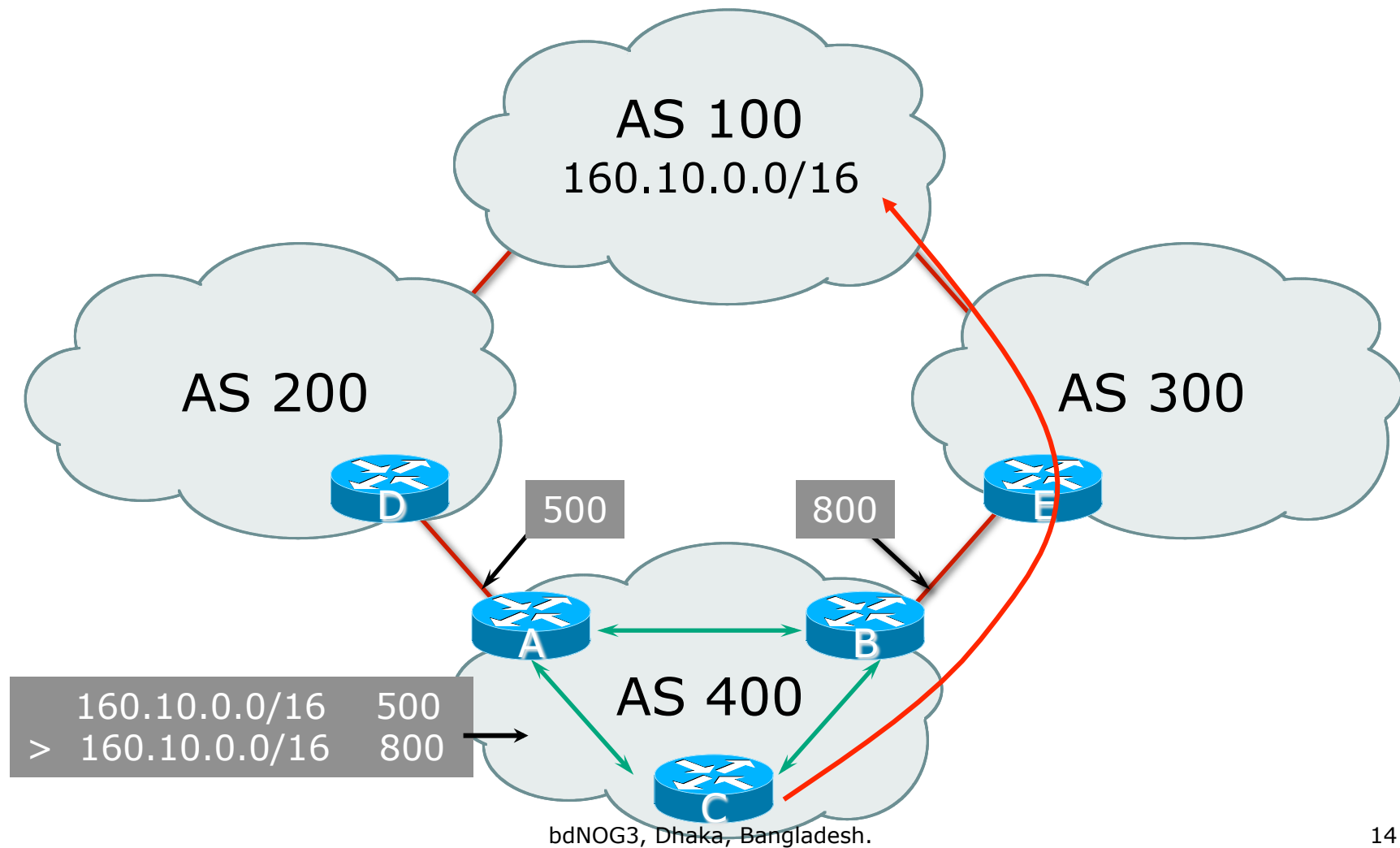
- ❑ Conveys the origin of the prefix
- ❑ **Historical** attribute
 - Used in transition from EGP to BGP
- ❑ Transitive and Mandatory Attribute
- ❑ Influences best path selection
- ❑ Three values: IGP, EGP, incomplete
 - IGP – generated by BGP network statement
 - EGP – generated by EGP
 - incomplete – redistributed from another routing protocol

Aggregator

- ❑ Conveys the IP address of the router or BGP speaker generating the aggregate route
- ❑ Optional & transitive attribute
- ❑ Useful for debugging purposes
- ❑ Does not influence best path selection
- ❑ Creating aggregate using “aggregate-address” sets the aggregator attribute:

```
router bgp 100  
  aggregate-address 100.1.0.0 255.255.0.0
```

Local Preference



Local Preference

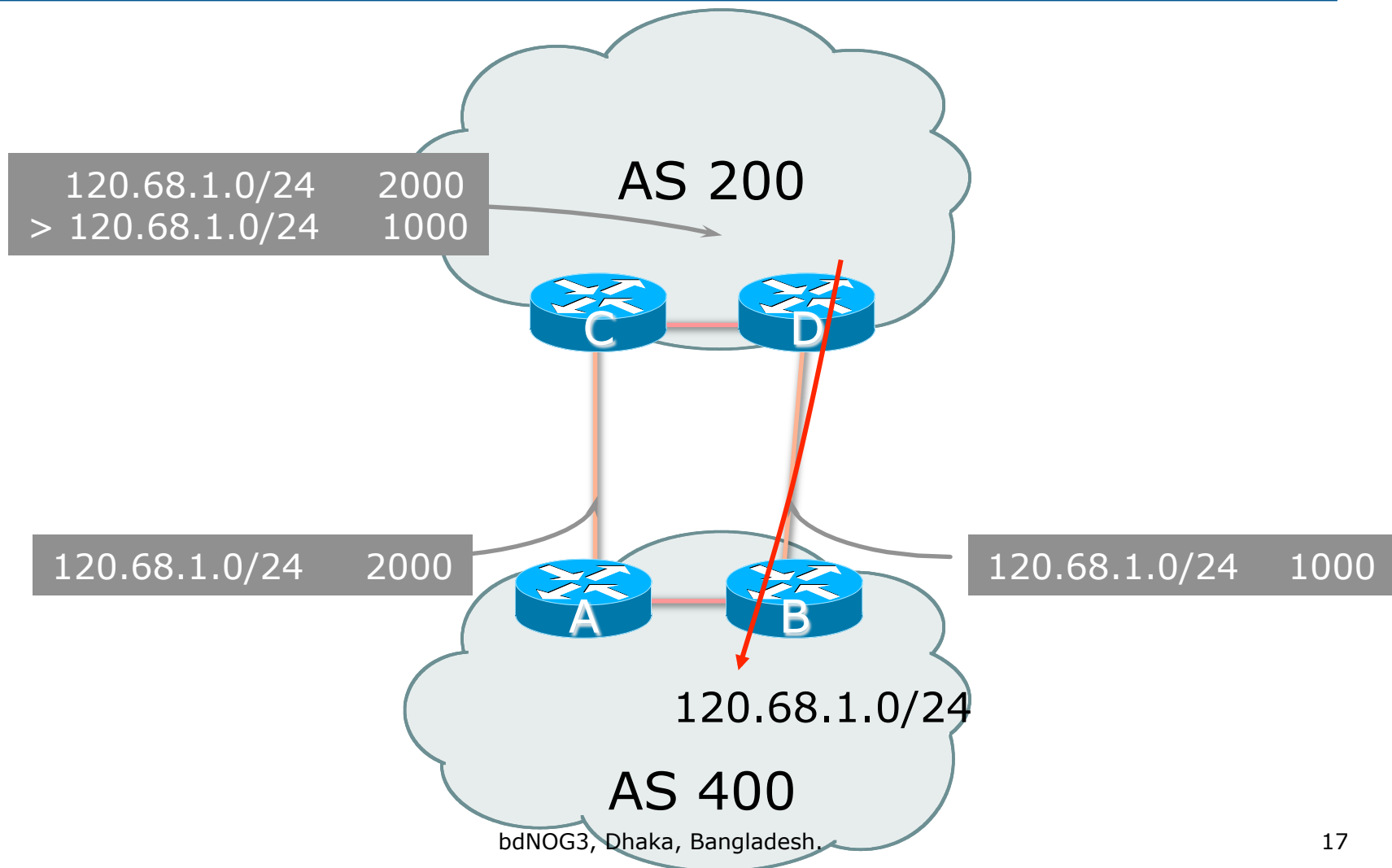
- ❑ Non-transitive and optional attribute
- ❑ Local to an AS only
 - Default local preference is 100 (IOS)
- ❑ Used to influence BGP path selection
 - determines best path for *outbound* traffic
- ❑ Path with highest local preference wins

Local Preference

❑ Configuration of Router B:

```
router bgp 400
  neighbor 120.5.1.1 remote-as 300
  neighbor 120.5.1.1 route-map local-pref in
!
route-map local-pref permit 10
  match ip address prefix-list MATCH
  set local-preference 800
route-map local-pref permit 20
!
ip prefix-list MATCH permit 160.10.0.0/16
```


Multi-Exit Discriminator (MED)



Multi-Exit Discriminator

- ❑ Inter-AS – non-transitive & optional attribute
- ❑ Used to convey the relative preference of entry points
 - determines best path for inbound traffic
- ❑ Path with lowest MED wins
- ❑ Absence of MED attribute implies MED value of **zero** (RFC4271)

Multi-Exit Discriminator

❑ Configuration of Router B:

```
router bgp 400
  neighbor 120.5.1.1 remote-as 200
  neighbor 120.5.1.1 route-map set-med out
!
route-map set-med permit 10
  match ip address prefix-list MATCH
  set metric 1000
route-map set-med permit 20
!
ip prefix-list MATCH permit 120.68.1.0/24
```

Weight

- ❑ Not really an attribute – local to router
- ❑ Highest weight wins
- ❑ Applied to all routes from a neighbour

```
neighbor 120.5.7.1 weight 100
```

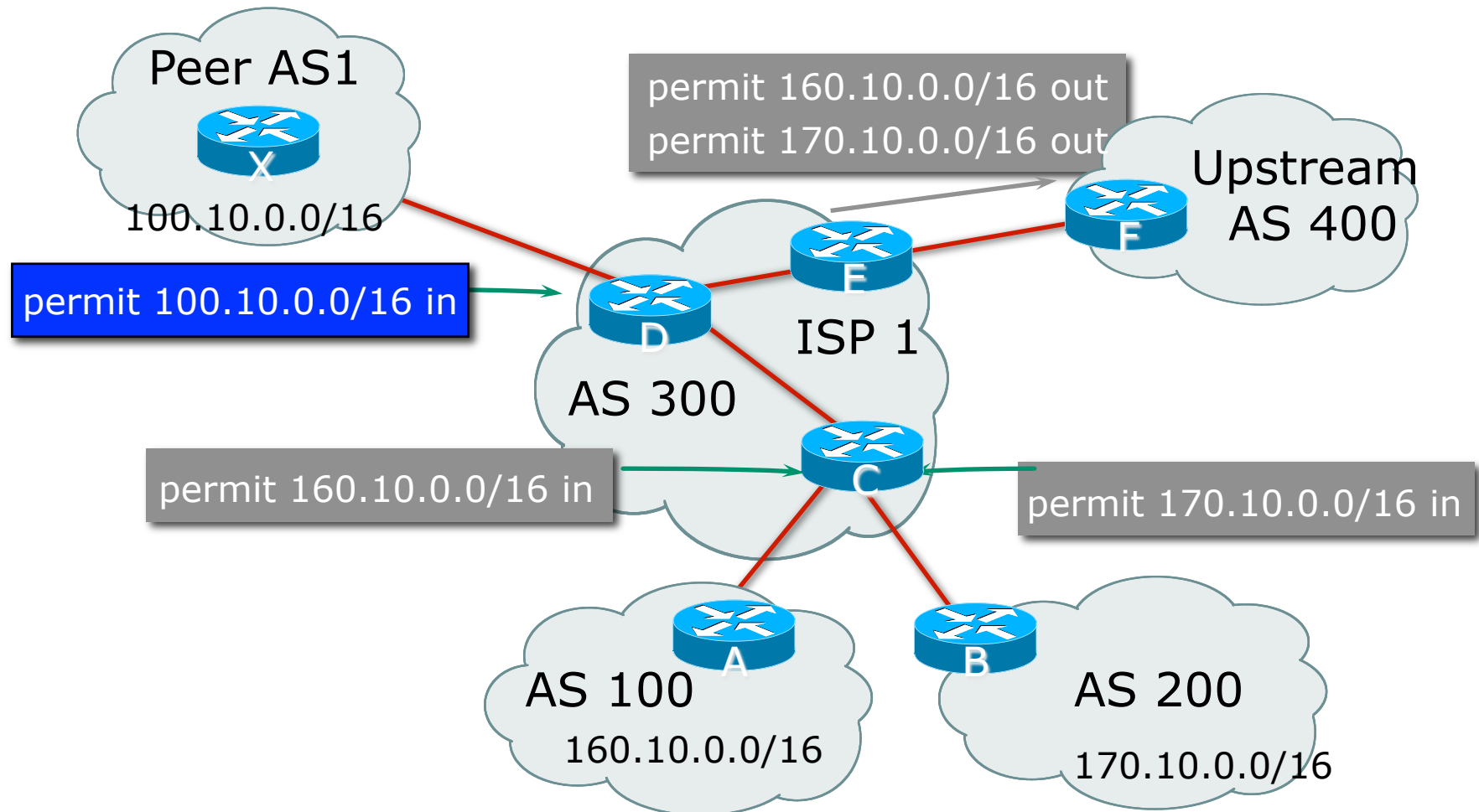
- ❑ Weight assigned to routes based on filter

```
neighbor 120.5.7.3 filter-list 3 weight 50
```

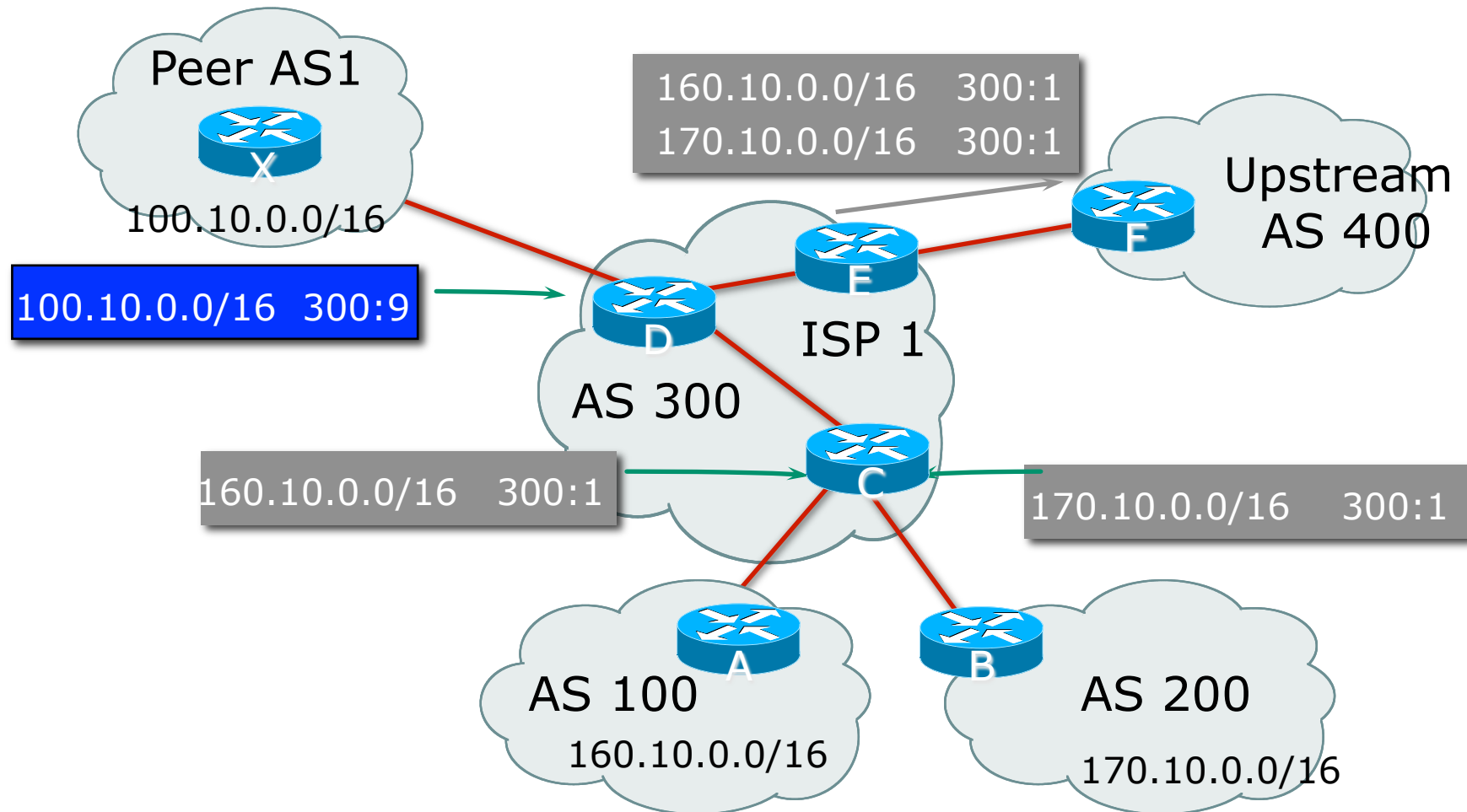
Community

- ❑ Communities are described in RFC1997
 - Transitive and Optional Attribute
- ❑ 32 bit integer
 - Represented as two 16 bit integers (RFC1998)
 - Common format is <local-ASN>:xx
 - 0:0 to 0:65535 and 65535:0 to 65535:65535 are reserved
- ❑ Used to group destinations
 - Each destination could be member of multiple communities
- ❑ Very useful in applying policies within and between ASes

Community Example (before)



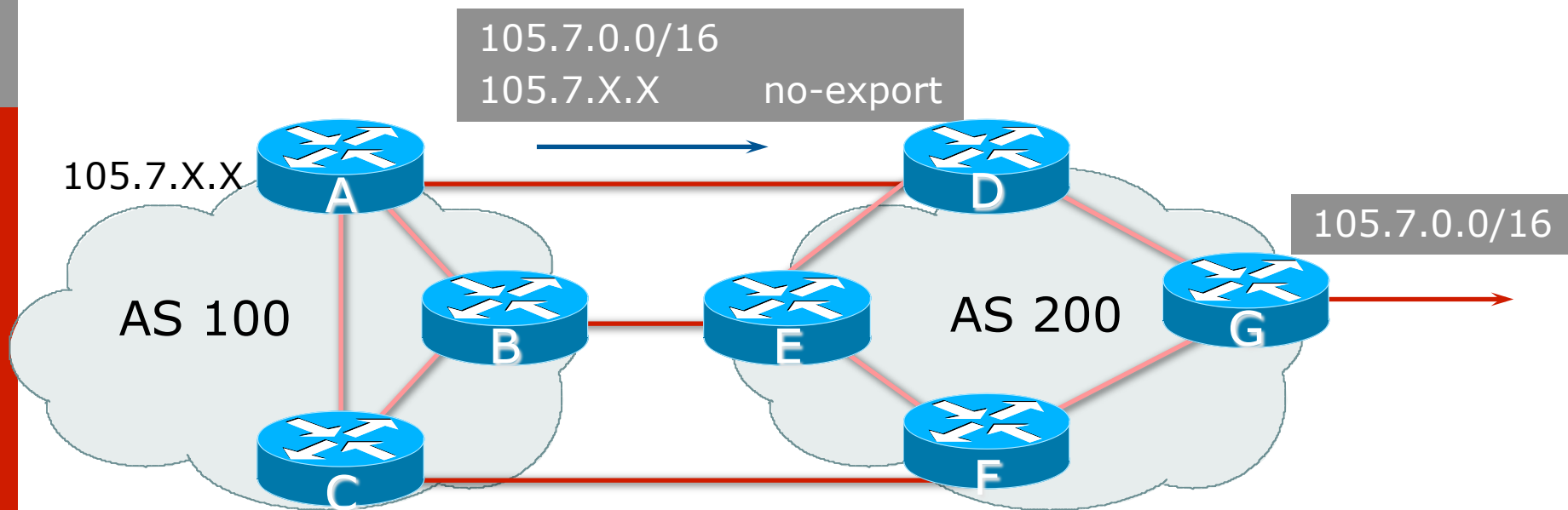
Community Example (after)



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



- ❑ 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

What about 4-byte ASNs?

- ❑ Communities are widely used for encoding ISP routing policy
 - 32 bit attribute
- ❑ RFC1998 format is now “standard” practice
 - ASN:number
- ❑ Fine for 2-byte ASNs, but 4-byte ASNs cannot be encoded
- ❑ Solutions:
 - Use “private ASN” for the first 16 bits
 - Wait for <http://datatracker.ietf.org/doc/draft-ietf-idr-as4octet-extcomm-generic-subtype/> to be implemented

Summary

Attributes in Action

Router6>sh ip bgp

BGP table version is 16, local router ID is 10.0.15.246

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
*>i 10.0.0.0/26	10.0.15.241	0	100	0	i
*>i 10.0.0.64/26	10.0.15.242	0	100	0	i
*>i 10.0.0.128/26	10.0.15.243	0	100	0	i
*>i 10.0.0.192/26	10.0.15.244	0	100	0	i
*>i 10.0.1.0/26	10.0.15.245	0	100	0	i
*> 10.0.1.64/26	0.0.0.0	0		32768	i
*>i 10.0.1.128/26	10.0.15.247	0	100	0	i
*>i 10.0.1.192/26	10.0.15.248	0	100	0	i
*>i 10.0.2.0/26	10.0.15.249	0	100	0	i
*>i 10.0.2.64/26	10.0.15.250	0	100	0	i

...

BGP Path Selection Algorithm



Why is this the best path?

BGP Path Selection Algorithm for Cisco IOS:

- Step 1: Prefer highest weight (local to router)
- Step 2: Prefer highest local preference (global within AS)
- Step 3: Prefer route originated by the local router
- Step 4: Prefer shortest AS path
- Step 5: Prefer lowest origin code (IGP < EGP < incomplete)
- Step 6: Prefer lowest MED (from other AS)
- Step 7: Prefer EBGP path over IBGP path
- Step 8: Prefer the path through the closest IGP neighbor
- Step 9: Prefer oldest route for EBGP paths
- Step 10: Prefer the path with the lowest neighbor BGP router ID

BGP Attributes and Path Selection



ISP Training Workshops