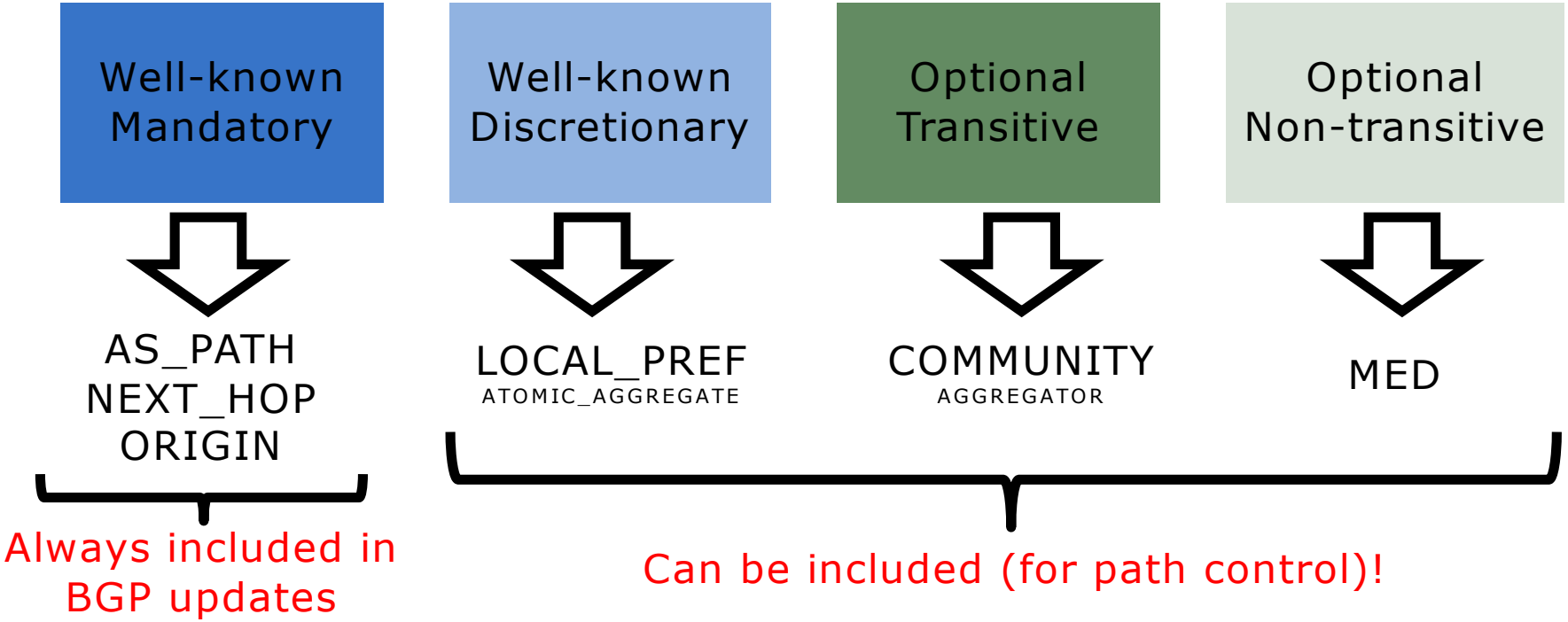
BGP Attributes & Path Selection

APNIC

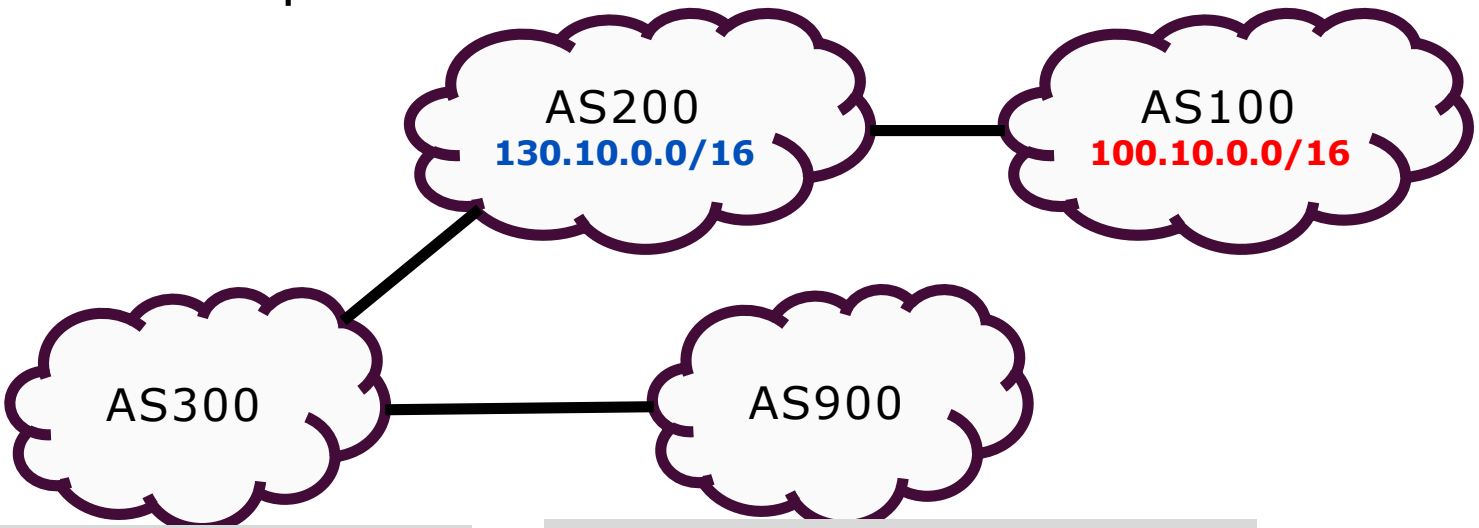
BGP Path Attributes

- Attributes describe the path to a network(s)/NLRI
  - Used to enforce routing policies for path control!



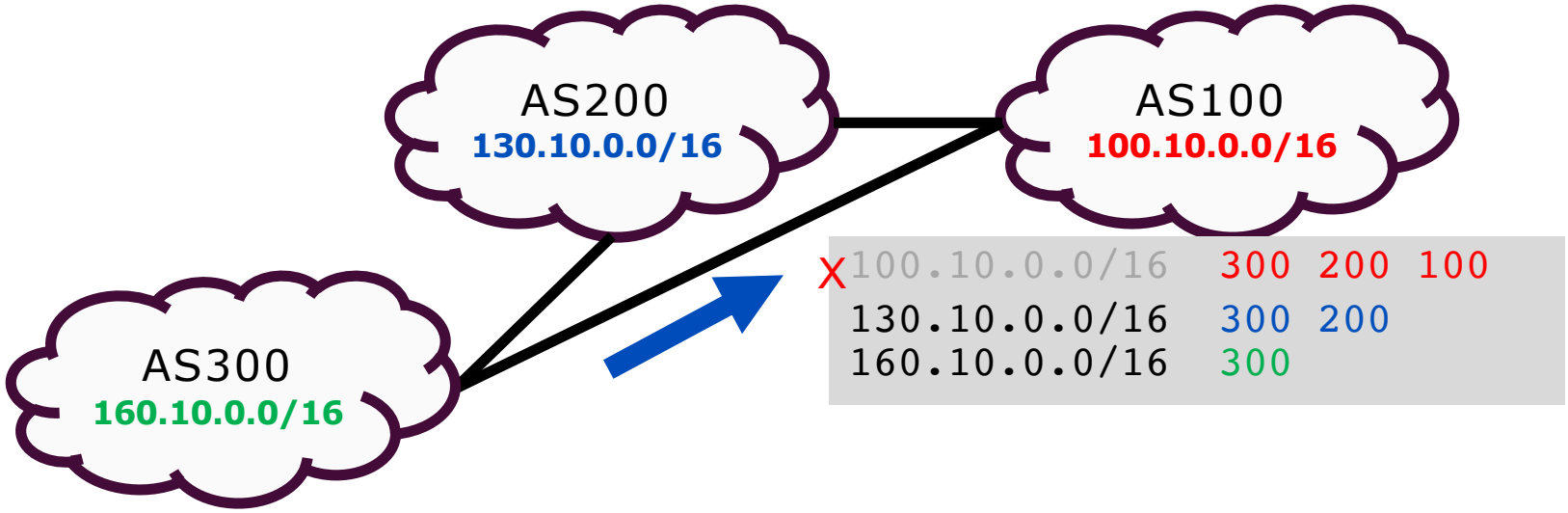
AS_PATH

- Indicates the list of ASes a route has passed through to reach the local AS
  - the list of ASes to reach a destination
  - can influence path selection!



AS_PATH

- Used to ensure a loop-free exchange of routing info between ASes
  - If own AS is seen in an update from an eBGP peer, loop is detected (Update is dropped)!

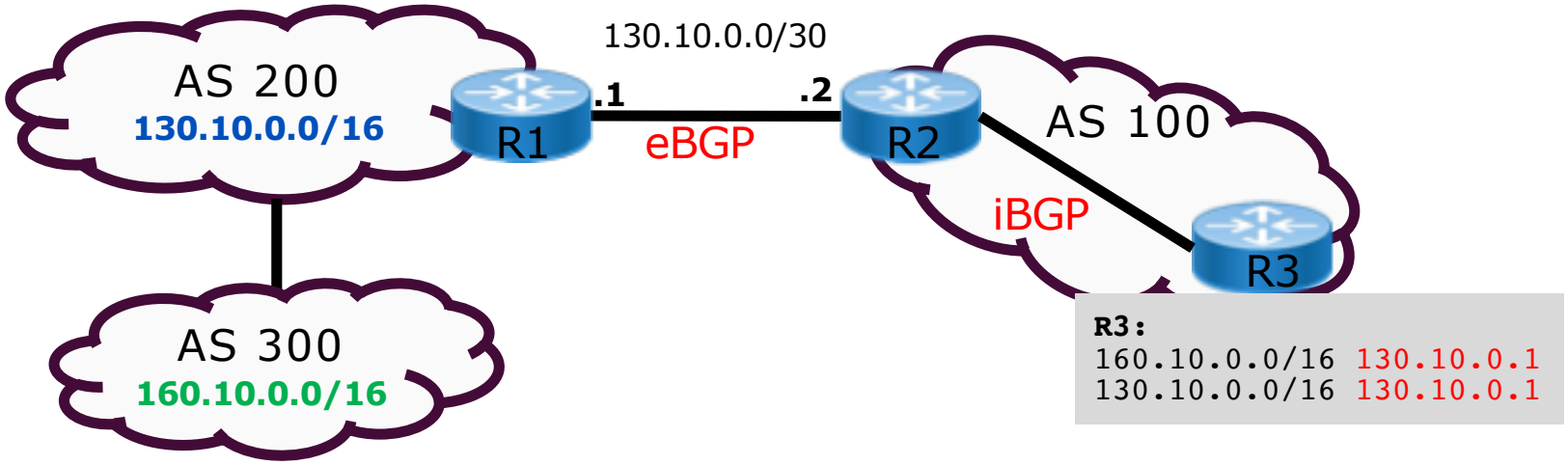


NEXT_HOP

- Indicates the next hop address to reach the destination
  - Source of the update packet!
- For eBGP
  - eBGP neighbor address (to reach the next AS)
- For iBGP
  - Generally the loopback address

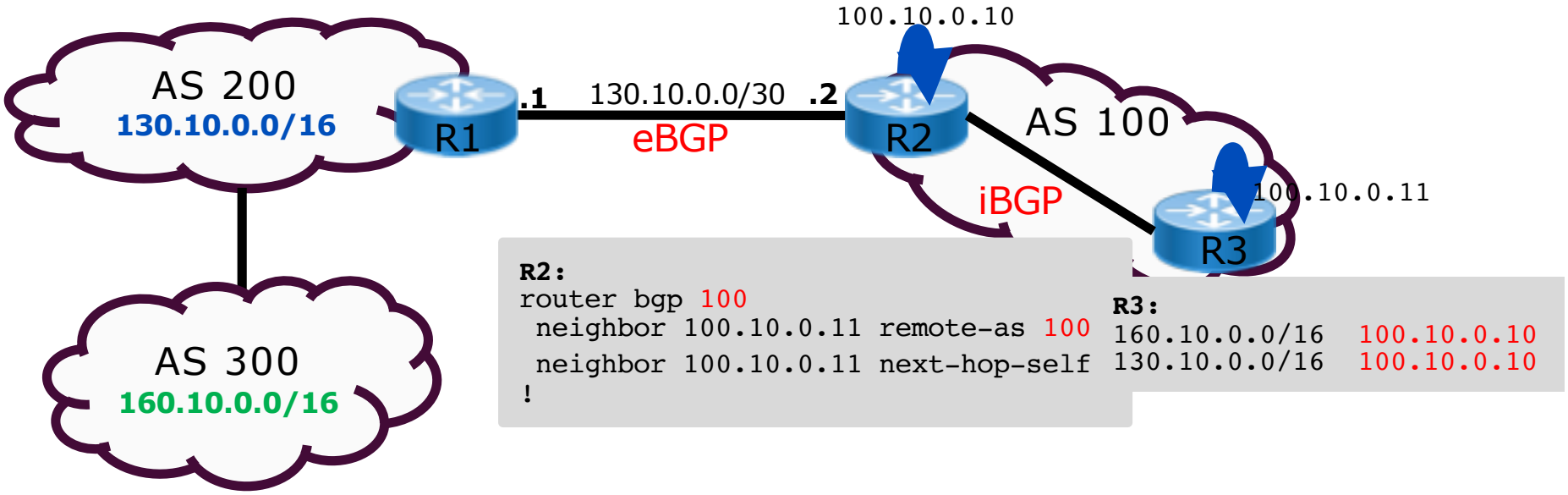
NEXT_HOP

- eBGP learned routes are advertised to iBGP peers without changing the next hop
  - Routers within the AS need to be able to reach the next hop (IGP or static)
  - Else, external routes not installed in the routing table!



NEXT-HOP-SELF

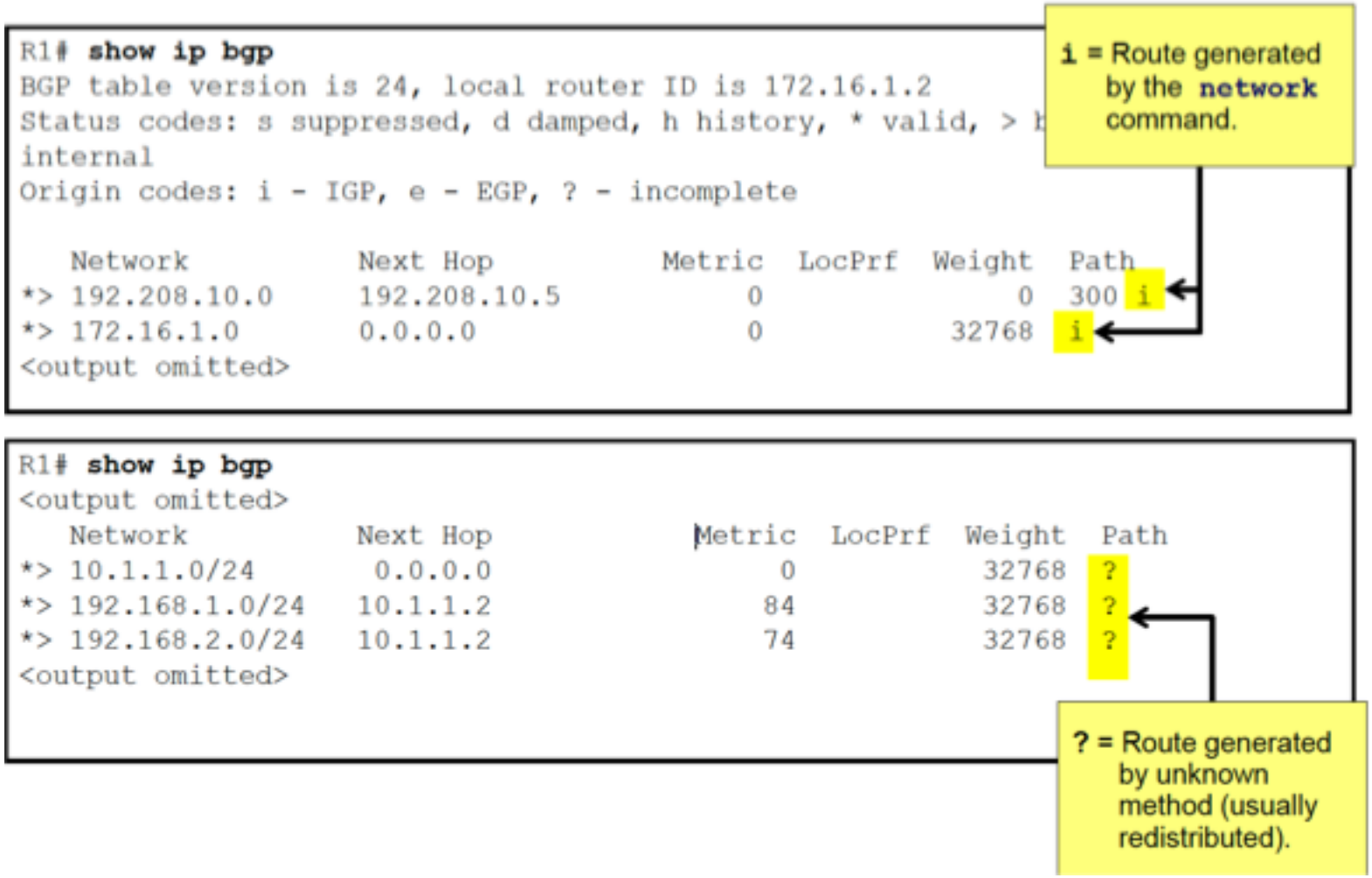
- Override the eBGP next hop default behavior with `next-hop-self` command
  - Advertises itself as the next hop for external routes
    - Reachable through IGP



ORIGIN

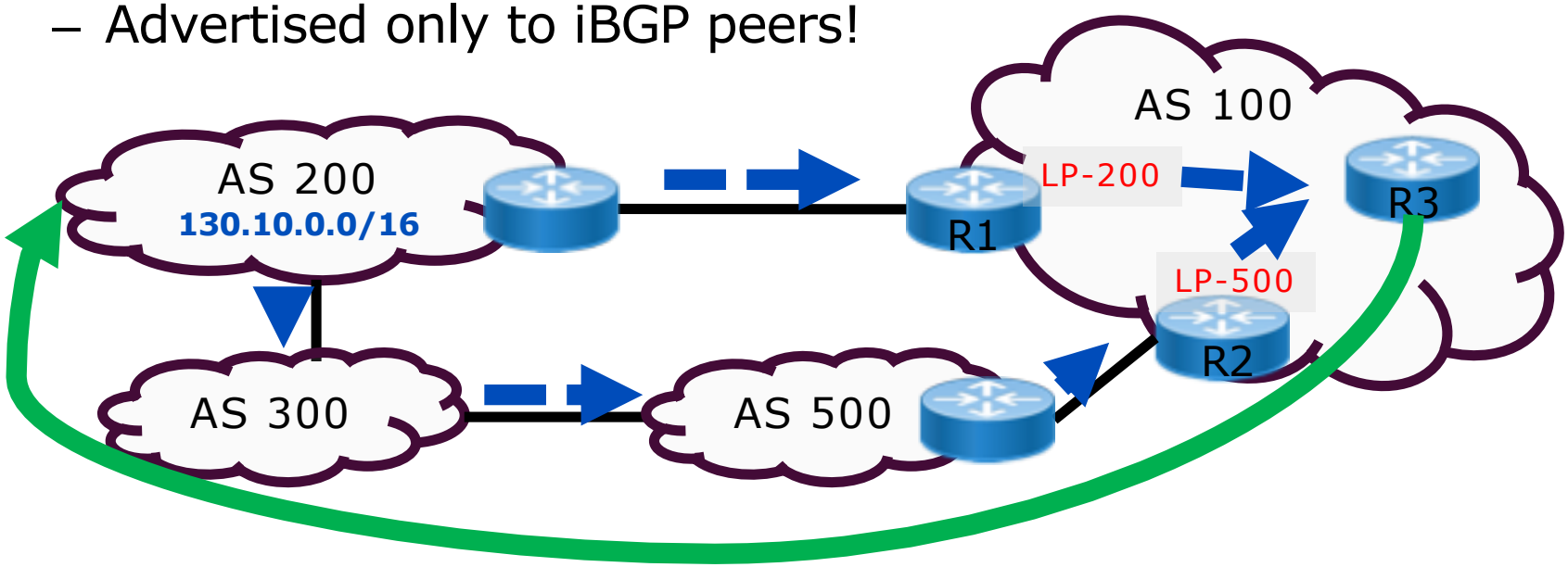
- Indicates the origin of the route
  - IGP (i)
    - Interior to the originating AS (advertised with the **`network`** command)
  - EGP (e)
    - Generated by EGP (obsolete!)
  - Incomplete (?)
    - Route's origin is unknown (usually redistributed)

ORIGIN


R1# show ip bgp
BGP table version is 24, local router ID is 172.16.1.2
Status codes: s suppressed, d damped, h history, * valid, > b
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
Network Next Hop Metric LocPrf Weight Path
*> 192.208.10.0 192.208.10.5 0 0 300 i
*> 172.16.1.0 0.0.0.0 0 32768 i
<output omitted>
i = Route generated by the network command.
R1# show ip bgp
<output omitted>
Network Next Hop Metric LocPrf Weight Path
*> 10.1.1.0/24 0.0.0.0 0 32768 ?
*> 192.168.1.0/24 10.1.1.2 84 32768 ?
*> 192.168.2.0/24 10.1.1.2 74 32768 ?
<output omitted>
? = Route generated by unknown method (usually redistributed).

LOCAL_PREF

- Local preference tells routers within the AS (local) the preferred path to exit the AS
  - Path with highest local_pref wins
    - Outbound traffic!
- Local to the AS
  - Advertised only to iBGP peers!



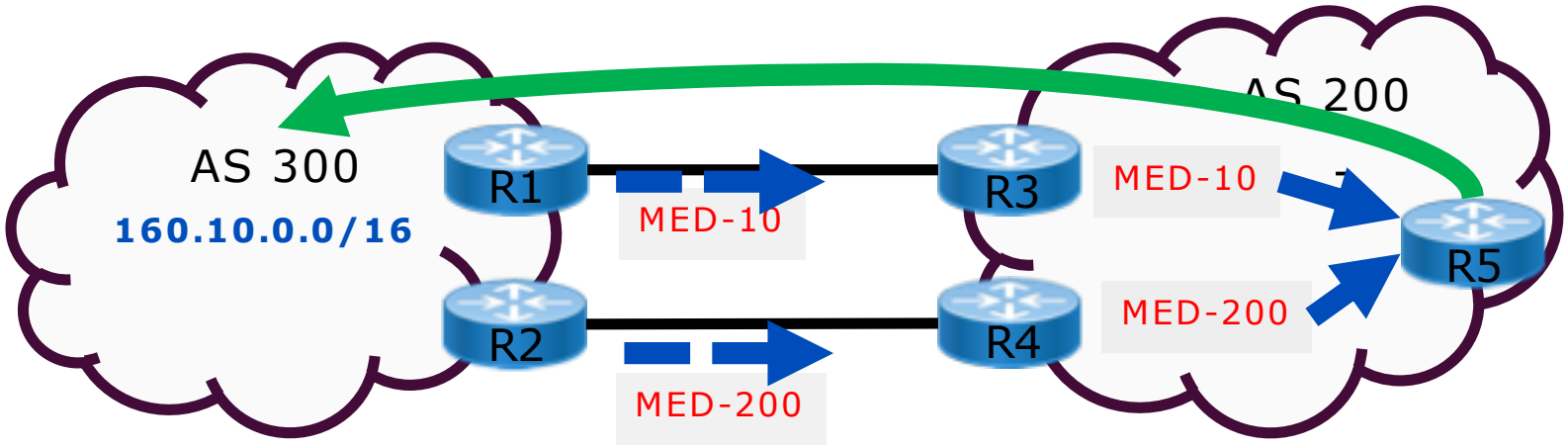
COMMUNITY

- Used to group prefixes (incoming/outgoing) and apply policies to the communities
  - A prefix can belong to more than one community
- Is (was?) a 32-bit integer
  - Represented as two 16-bit integers [`ASN:number`]
    - Works well for 2-byte ASN
- With 4-byte ASNs
  - Common to see [`private-ASN:number`]
  - RFC 8092 (BGP Large Communities): 96-bit integer
    - [32-bit ASN:32-bit:32-bit]

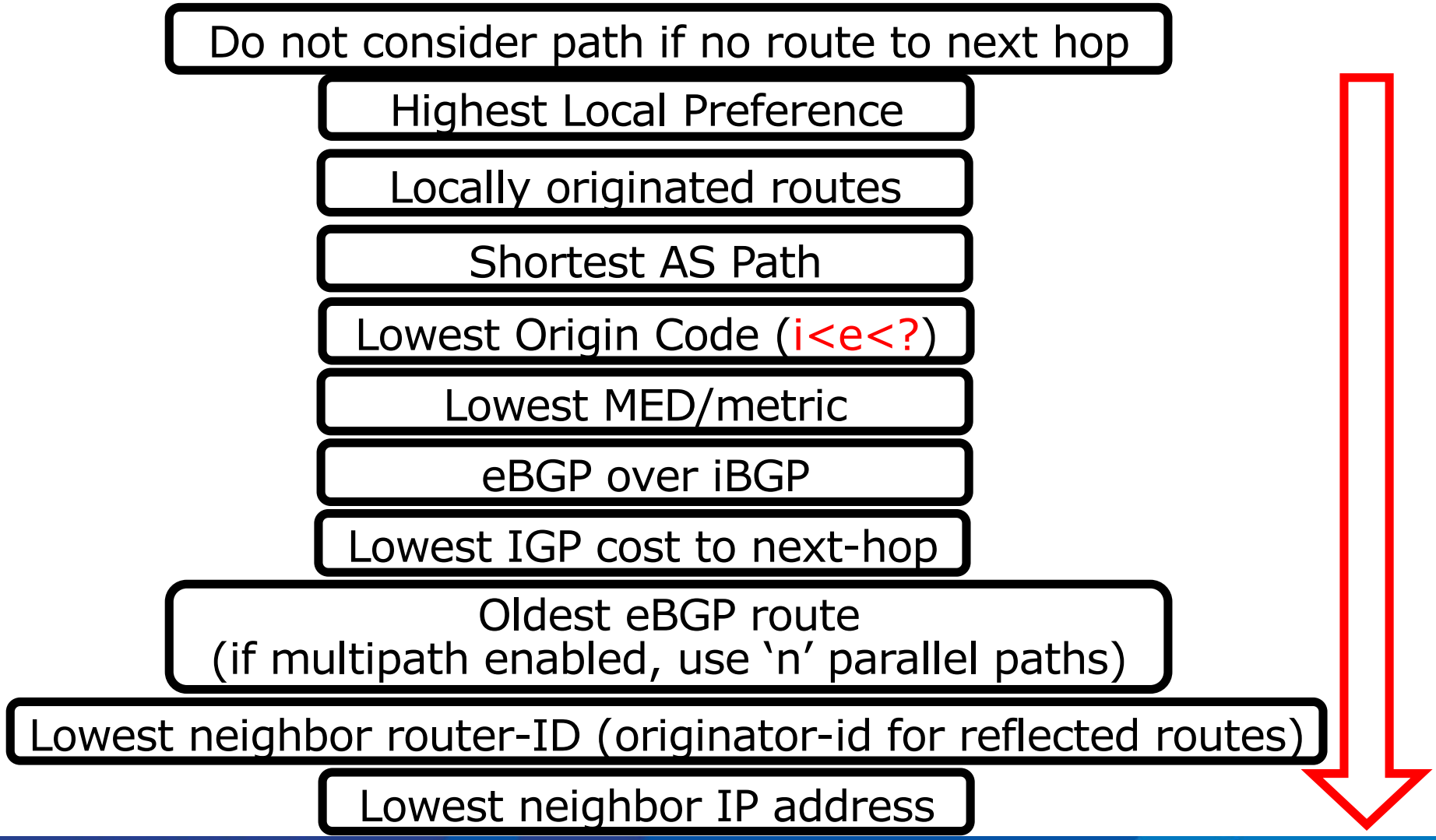
MED

- Multi-exit discriminator is inter-AS non-transitive
  - Indicates to neighbor AS about the preferred entry points into the local AS (incoming traffic)

- The path with lowest MED wins!



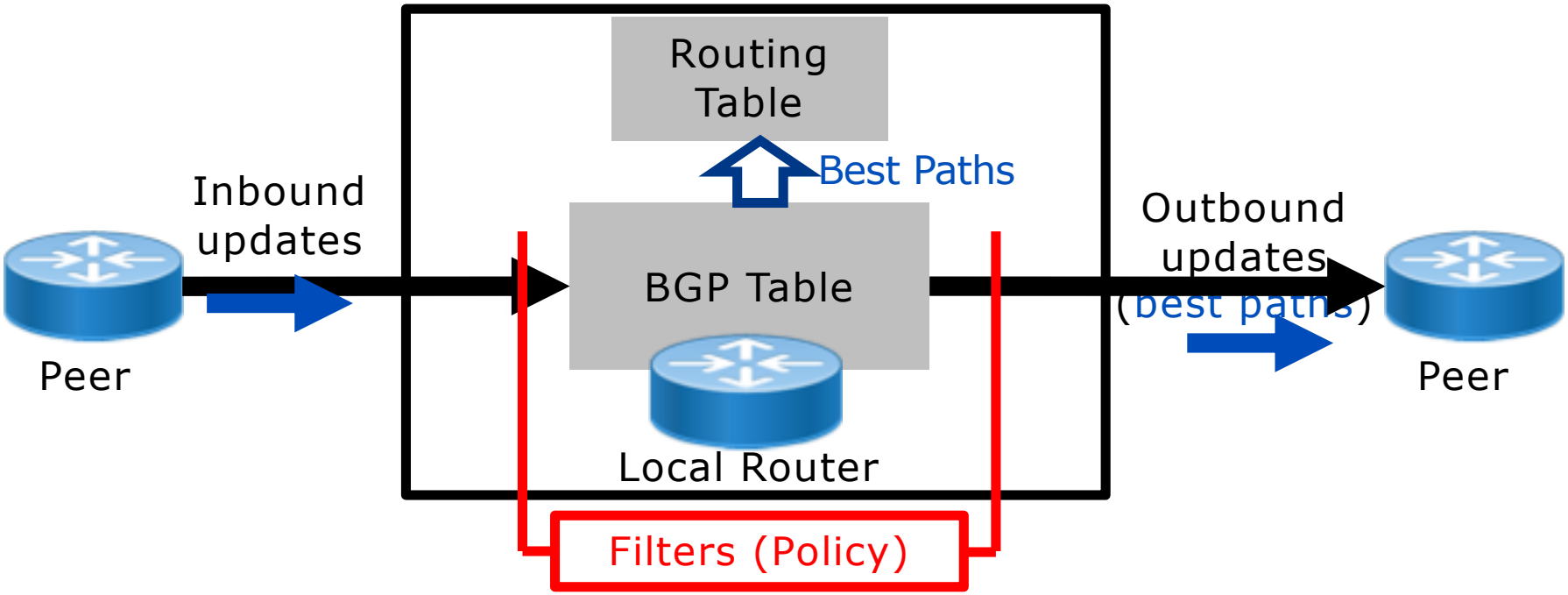
BGP Best Path Selection



BGP Operation

- BGP learns routes from iBGP and eBGP peers
  - Selects best path based on the attributes
  - Installs best path in the routing table
  - Advertises the best paths to its other BGP peers
    - eBGP learned routes to iBGP peers
    - iBGP learned routes to eBGP peers

BGP Operation



BGP Tables

- Neighbor Table
  - List of all BGP neighbors
- BGP Table
  - List of routes learned from all BGP neighbors
  - (And locally originated routes!)
- Routing (Forwarding) Table
  - All best paths
    - selected based on attributes and whose next-hops are reachable!

Acknowledgement:

- Philip Smith
- Cisco Systems


Questions