BGP Fundamentals

APNIC

Border Gateway Protocol - BGP

- Runs over TCP (port 179)
  - TCP connection required before BGP session
  - Need to be reachable!
- Path vector routing protocol
  - Best path selection based on path attributes
  - Route: destination and the attributes of the path to reach the destination
- Incremental BGP updates

Internal & External BGP

- eBGP used to:
  - Exchange networks/routes between ASes
    - Aggregates and sub-aggregates
  - Implement routing policies
    - To manipulate inbound and outbound traffic
- iBGP is used to:
  - Carry customer networks/prefixes
  - Internet routes (some or all) across the AS backbone

BGP Message Types

- Open:
  - After a TCP connection has been established between two BGP routers, an Open message is sent
    - Once the open message is confirmed (keepalive), the BGP session is established – become BGP peers/neighbors!
  - Contains:
    - Sender's ASN
    - BGP version
    - BGP router ID
    - Hold-time (3 x keepalive interval)

BGP Message Types

- Keepalive:
  - Exchanged initially to acknowledge Open messages
  - Exchanged periodically (60 secs) to maintain BGP session
    - Dataless packet

- Update:
  - BGP peers exchange network information through Update messages
    - One update for each path!
  - Contains:
    - Withdrawn routes – no more reachable
    - Path attributes – attributes for this path to reach the destinations specified by the NLRI
    - NLRI – list of networks reachable through this path <prefix, length>

BGP Message Types

- Notification:
  - Sent when an error condition is detected
  - The BGP session is torn down immediately!
  - Contains:
    - Error code
    - Error sub-code
    - Data related to error

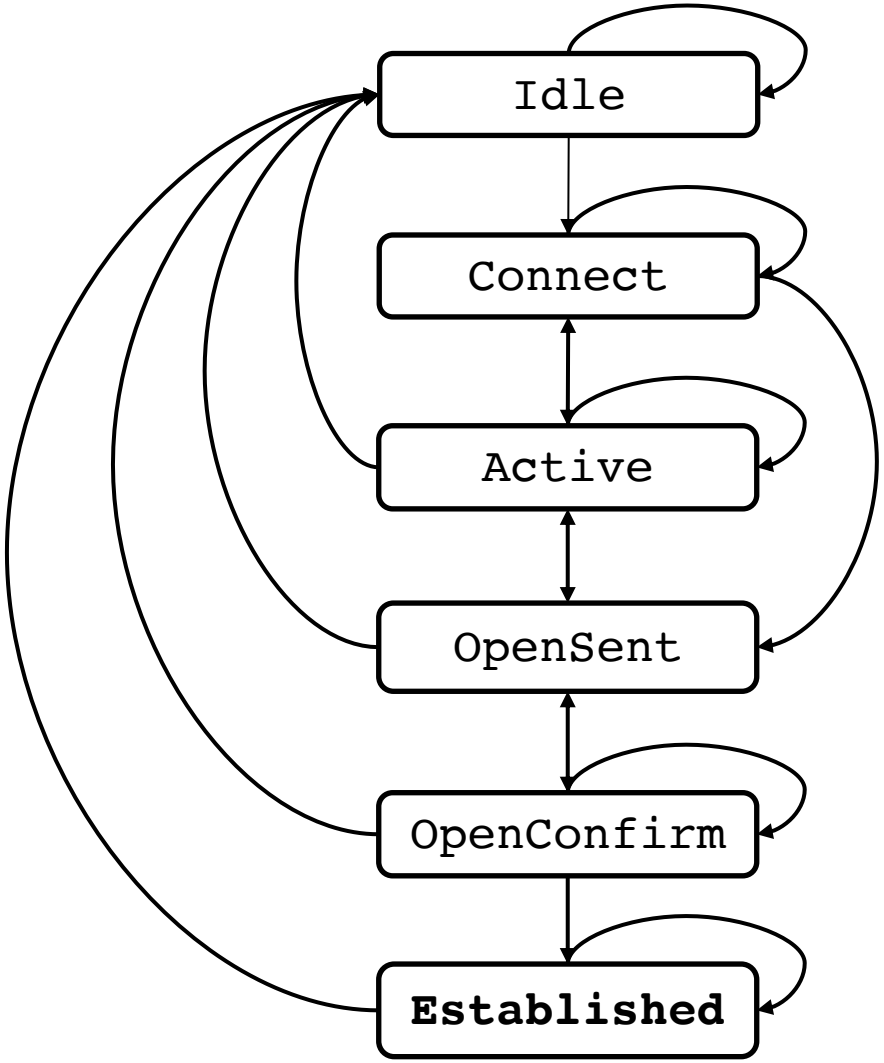
BGP Neighbor States

- A BGP router goes through six different states
  - Idle
    - The router is looking for a route to its neighbor
  - Connect
    - BGP router moves from Idle to Connect state if it has found a route to its neighbor, and has started the TCP handshake
    - If the TCP session successful, sends an Open message (and transitions to OpenSent)
    - Else, move to Active state
  - Active
    - A router transitions to Active state if the initial TCP connection was not successful (in Connect state)
    - Restarts the TCP connection
    - If successful, sends an Open message
    - Else, falls back to Idle state

BGP Neighbor States

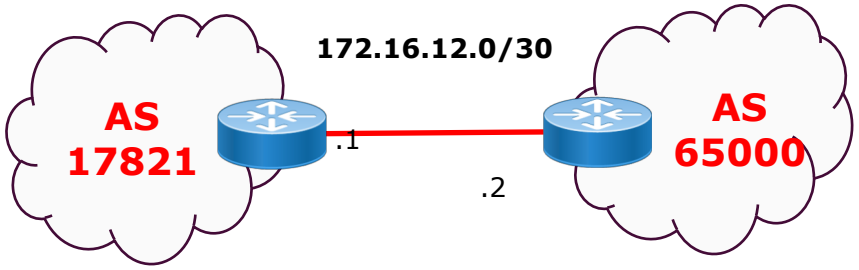
- A BGP router goes through six different states
  - OpenSent
    - An Open message has been sent to the neighbor
    - Waiting for Open message from neighbor
    - If it receives an Open message and there are no mismatches (*version, source addr same as TCP addr, ASN, router-ID, TTL, md5*), sends KeepAlive, moves to OpenConfirm
    - Else (if mismatches/errors), sent Notification and falls back to Idle
  - OpenConfirm
    - waiting for the initial KeepAlive
    - If received, transitions to Established
    - If holdtimer expires or Notification received, moves to Idle
  - Established
    - The BGP neighbor relationship (session) is established!
    - Routing information can now be exchanged
    - If holdtimer expires/error, moves back to Idle

BGP State Machine



BGP Neighbor Relationship

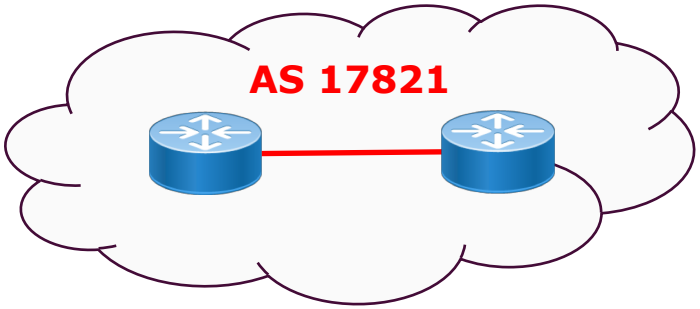
- eBGP neighbors/peers
  - BGP session established between routers in different ASes
  - Generally directly connected!
    - Session established using directly connected intf IP
    - Peering address must match the TCP session!
  - Else, we need a static route to reach the neighbor and change the eBGP TTL value (default 1)




```
router bgp 17821
 neighbor 172.16.12.2 remote-as 65000
!
 address-family ipv4
  neighbor 172.16.12.2 activate
 !
```

BGP Neighbor Relationship

- iBGP neighbors/peers
  - BGP session established between routers within the same AS
  - Does not need to be directly connected
    - IGP ensure reachability (TCP connection)
  - Generally using loopback addresses

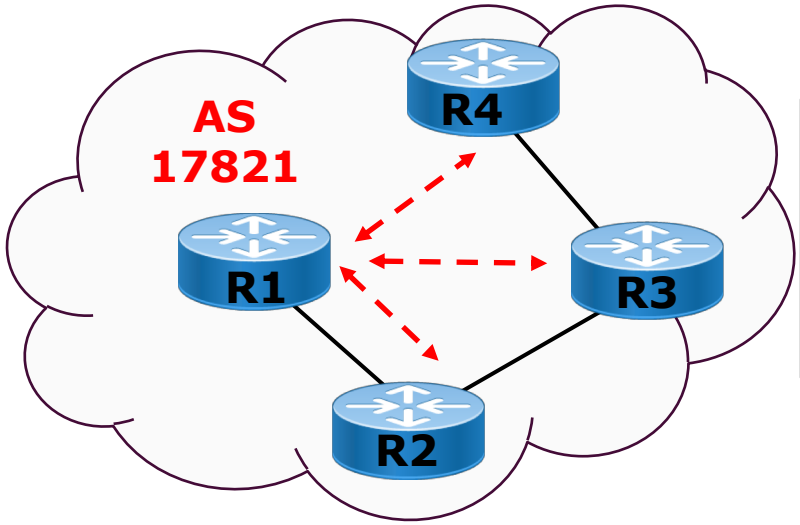



```
router bgp 17821
 neighbor 10.10.10.2 remote-as 17821
!
```

iBGP Operation

- iBGP routers must:
  - Originate directly connected routes
  - Carry routes learned from outside the AS to all routers within the AS
    - Fully-meshed instead of redistributing!
    - Advertise routes learned from eBGP peers to all iBGP peers!
  - To prevent routing loops (in a fully-meshed network)
    - iBGP routers are not allowed to advertise iBGP learned routes to other iBGP peers!

iBGP full-mesh




```
router bgp 17821
 neighbor 10.10.10.2 remote-as 17821
 neighbor 10.10.10.3 remote-as 17821
 neighbor 10.10.10.4 remote-as 17821
 !
```

Sourcing iBGP from Loopback

- By default, routers use the exit-interface address as the source address for locally originated packets (updates)
  - If the BGP TCP session was established using any other interface (loopbacks) addresses, the source address for BGP updates must match!

- The `update-source loopback` command achieves this

```
router bgp 17821
 neighbor 10.10.10.1 remote-as 17821
 neighbor 10.10.10.1 update-source loopback 0
!
```

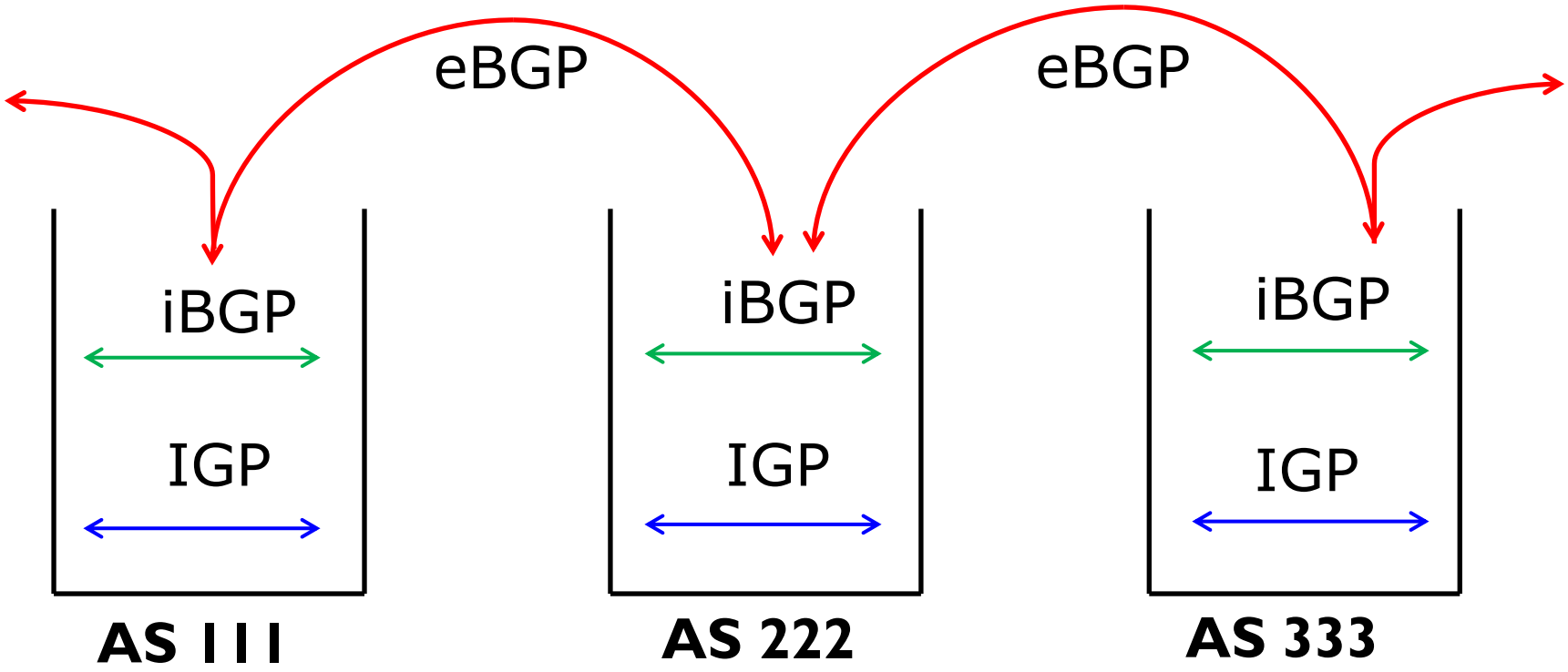
Advertising Networks in BGP

- The `network` statement
  - allows BGP to inject routes into BGP table and advertise to neighbors only if it already exists in the routing table!

```
router bgp 17821
 address-family ipv4 unicast
  network <prefix> mask <subnet-mask>
 address-family ipv6 unicast
  network <prefix/length>
```

- BGP "Synchronization Rule":
  - iBGP learned routes should not be installed in the routing table nor advertised to eBGP peers unless the route was learned through an IGP first!
    - Prevents black-hole routes!

How it all works?



Barry Greene & Philip Smith “Cisco ISP Essentials”

Acknowledgement:

- Philip Smith
- Cisco Systems


Questions