The Border Gateway Protocol (BGP) is an inter-autonomous system routing protocol. An autonomous system is a group of networks under common administrative control and routing policies. This sequence diagram describes the sequence of messages exchanged when a new BGP router is made operational. The steps involved are:

1. Establish TCP connections
2. Exchange BGP Open messages.
4. Exchange routing information with the BGP Update message.

After the initial handshake, the routers exchange the BGP Update messages. The attributes exchanged during BGP Update coupled with router specific configuration govern the route selection. Important factors are:
- Router level configuration of the weights.
- Local preference settings on the routers.
- Metric suggestions from the advertising router. (Multi-exist discriminator)
- Origin of the route (EGP, IGP or Unknown-Origin)
- AS_Path: Autonomous System (AS) Path of the advertised route (i.e. the list of Autonomous Systems in the route advertisement path.)
- Next Hop: IP Address used to reach the advertising EBGP router.
- Community:
  - No-Export: Routes learnt with this community setting cannot be advertised to other AS.
  - No-Advertise: Routes learnt with this attribute cannot be advertised to IGPs.
  - Internet: Routes can be advertised to any BGP router in the Internet.

**BGP UPDATE**

Border Gateway Protocol, 
Unfeasible routes length: 0 bytes, 
Total path attribute length: 72 bytes, 
ORIGIN: IGP (4 bytes), 
Flags: 0x80 (Optional, Non-transitive, Complete), 
AS_PATH: {500, 500} 65211 (13 bytes), 
NEXT_HOP: 192.168.0.15 (7 bytes), 
LOCAL_PREF: 100 (7 bytes), 
AGGREGATOR: AS: 65210 origin: 192.168.0.10 (9 bytes), 
COMMUNITIES: 65215:1 790:4 340:250 (15 bytes), 
ORIGINATOR_ID: 192.168.0.15 (7 bytes), 
CLUSTER_LIST: 192.168.0.250 (7 bytes), 
NLRI prefix: 192.168.4.0 (192.168.4.0)

BGP Router 1 advertises routers to BGP Router 2.

**BGP UPDATE**

Border Gateway Protocol, 
Unfeasible routes length: 0 bytes, 
Total path attribute length: 39 bytes, 
ORIGIN: EGP (4 bytes), 
Flags: 0xc0 (Optional, Transitive, Complete), 
AS_PATH: empty (3 bytes), 
NEXT_HOP: 192.168.0.33 (7 bytes), 
LOCAL_PREF: 100 (7 bytes), 
COMMUNITIES: 65033:500 65033:600 (11 bytes), 
NLRI prefix: 10.0.0.0 (10.0.0.0)

BGP Router 2 advertises routers to BGP Router 1.

Update IP Routing Table

Rules for updating the IP Routing table are listed below.

- If the path specifies a next hop that is inaccessible, drop the update.
- Prefer the path with the largest weight.
- If the weights are the same, prefer the path with the largest local preference.
- If the local preferences are the same, prefer the path that was originated by BGP running on this router.
- If no route was originated by this router, prefer the route that has the shortest AS_path.
- If all paths have the same AS_path length, prefer the path based on the origin (IGP is preferred over EGP, and EGP is preferred over Origin-Unknown).
- If the origin codes are the same, prefer the path with the lowest MED attribute (Metric suggestion from the advertising router).
- If the paths have the same MED, prefer the external path over the internal path.
- If the paths are still the same, prefer the path through the closest IGP neighbor.
- Prefer the path with the lowest IP address, as specified by the BGP router ID.
We have looked at the message flow when a new BGP router starts up. We will now look at the same flow with full message details.

BGP Router 1 enters the network

Idle

TCP SYN

Connect

TCP SYN+ACK

TCP ACK

BGP OPEN

The neighbor peering state starts as Idle.

BGP Router 1 comes up and establishes a TCP connection with BGP Router 2.

TCP connection is being setup with the neighbor.

The BGP routers exchange the BGP OPEN messages with important parameters and the autonomous system number of the router.

TCP connection has been established with the neighbor and a BGP OPEN message has been sent to the peer.

Open Sent

OPEN Message

Marker: 16 bytes
Length: 29 bytes
Type: OPEN Message (1)
Version: 4
My AS: 65033
Hold time: 180
BGP identifier: 192.168.0.15
Optional parameters length: 0 bytes

TCP ACK

BGP OPEN

The peering state now moves to "Open Confirm" as the BGP OPEN message has been received from the neighbor.

Open Confirm

OPEN Message

Marker: 16 bytes
Length: 29 bytes
Type: OPEN Message (1)
Version: 4
My AS: 65033
Hold time: 180
BGP identifier: 192.168.0.33
Optional parameters length: 0 bytes

TCP ACK

BGP KEEPALIVE

The BGP routers initiate the exchange of periodic health messages.

KEEPALIVE Message

Marker: 16 bytes
Keep alive messages have been exchanged with the neighbor. The BGP peering state now moves to "Established" state. In this state, the routers will now exchange routing information using the BGP UPDATE messages.

After the initial handshake, the routers exchange the BGP Update messages. The attributes exchanged during BGP Update coupled with router specific configuration govern the route selection. Important factors are:

- Router level configuration of the weights.
- Local preference settings on the routers.
- Metric suggestions from the advertising router. (Multi-exist discriminator)
- Origin of the route (EGP, IGP or Unknown-Origin)
- AS Path: Autonomous System (AS) Path of the advertised route (i.e. the list of Autonomous Systems in the route advertisement path.)
- Next Hop: IP Address used to reach the advertising EBGP router.
- Community:
  - No-Export: Routes learnt with this community setting cannot be advertised to other AS.
  - No-Advertise: Routes learnt with this attribute cannot be advertised to IGPs.
  - Internet: Routes can be advertised to any BGP router in the Internet.

BGP UPDATE

Border Gateway Protocol,
Unfeasible routes length: 0 bytes,
Total path attribute length: 72 bytes,
ORIGIN: IGP (4 bytes),
Flags: 0x80 (Optional, Non -transitive, Complete),
AS _PATH: {500, 500} 65211 (13 bytes),
NEXT _HOP: 192 .168 .0 .15 (7 bytes),
LOCAL _PREF: 100 (7 bytes),
AGGREGATOR: AS: 65210 origin: 192 .168 .0 .10 (9 bytes),
COMMUNITIES: 65215:1 790:4 340:250 (15 bytes),
ORIGINATOR _ID: 192 .168 .0 .15 (7 bytes),
CLUSTER _LIST: 192 .168 .0 .250 (7 bytes),
NLRI prefix: 192 .168 .4 .0 (192 .168 .4 .0)

BGP Router 1 advertises routers to BGP Router 2.

UPDATE Message

Marker: 16 bytes
Length: 98 bytes
Type: UPDATE Message (2)
Unfeasible routes length: 0 bytes
Total path attribute length: 72 bytes
Path attributes
ORIGIN: INCOMPLETE (4 bytes)
Flags: 0x40 (Well-known, Transitive, Complete)
Type code: ORIGIN (1)
Length: 1 byte
Origin: INCOMPLETE (2)
AS _PATH: (500, 500) 65211 (13 bytes)
Flags: 0x40 (Well-known, Transitive, Complete)
Type code: AS_PATH (2)
Length: 10 bytes
AS path: (500, 500) 65211
AS path segment: (500, 500)
Path segment type: AS_SET (1)
Path segment length: 2 ASs
Path segment value: 500 500
AS path segment: 65211
Path segment type: AS_SEQUENCE (2)
Path segment length: 1 AS
Path segment value: 65211
NEXT_HOP: 192.168.0.15 (7 bytes)
  Flags: 0x40 (Well-known, Transitive, Complete)
  Type code: NEXT_HOP (3)
  Length: 4 bytes
  Next hop: 192.168.0.15 (192.168.0.15)

LOCAL_PREF: 100 (7 bytes)
  Flags: 0x40 (Well-known, Transitive, Complete)
  Type code: LOCAL_PREF (5)
  Length: 4 bytes
  Local preference: 100

ATOMIC_AGGREGATE: 3 bytes
  Flags: 0x40 (Well-known, Transitive, Complete)
  Type code: ATOMIC_AGGREGATE (6)
  Length: 0 bytes

AGGREGATOR: AS: 65210 origin: 192.168.0.10 (9 bytes)
  Flags: 0x00 (Optional, Transitive, Complete)
  Type code: AGGREGATOR (7)
  Length: 6 bytes
  Aggregator AS: 65210
  Aggregator origin: 192.168.0.10 (192.168.0.10)

COMMUNITIES: 65215:1 790:4 340:250 (15 bytes)
  Flags: 0x00 (Optional, Transitive, Complete)
  Type code: COMMUNITIES (8)
  Length: 12 bytes
  Communities: 65215:1 790:4 340:250
    Community: 65215:1
    Community AS: 65215
    Community value: 1
    Community: 790:4
    Community AS: 790
    Community value: 4
    Community: 340:250
    Community AS: 340
    Community value: 250

ORIGINATOR_ID: 192.168.0.15 (7 bytes)
  Flags: 0x80 (Optional, Non-transitive, Complete)
  Type code: ORIGINATOR_ID (9)
  Length: 4 bytes
  Originator identifier: 192.168.0.15 (192.168.0.15)

CLUSTER_LIST: 192.168.0.250 (7 bytes)
  Flags: 0x80 (Optional, Non-transitive, Complete)
  Type code: CLUSTER_LIST (10)
  Length: 4 bytes
  Cluster list: 192.168.0.250
  Cluster List: C0A800FA

Network layer reachability information: 3 bytes
  172.16.0.0/16
  NLRI prefix length: 16
  NLRI prefix: 172.16.0.0 (172.16.0.0)

- If the path specifies a next hop that is inaccessible, drop the update.
- Prefer the path with the largest weight.
- If the weights are the same, prefer the path with the largest local preference.
- If the local preferences are the same, prefer the path that was originated by BGP running on this router.
- If no route was originated by this router, prefer the route that has the shortest AS_path.
- If all paths have the same AS_path length, prefer the path based on the origin
  (IGP is preferred over EGP, and EGP is preferred over Origin-Unknown).
- If the origin codes are the same, prefer the path with the lowest MED attribute (Metric suggestion from the advertising router).
- If the paths have the same MED, prefer the external path over the internal path.
- If the paths are still the same, prefer the path through the closest IGP neighbor.
- Prefer the path with the lowest IP address, as specified by the BGP router ID.
BGP Router 1 advertises routers to BGP Router 2.

**UPDATE Message**
- Marker: 16 bytes
- Length: 64 bytes
- Type: UPDATE Message (2)
- Unfeasible routes length: 0 bytes
- Total path attribute length: 39 bytes
- Path attributes
  - ORIGIN: EGP (4 bytes)
    - Flags: 0x0 (Well-known, Transitive, Complete)
    - Type code: ORIGIN (1)
    - Length: 1 byte
    - Origin: EGP (1)
  - AS_PATH: empty (3 bytes)
    - Flags: 0x0 (Well-known, Transitive, Complete)
    - Type code: AS_PATH (2)
    - Length: 0 bytes
    - AS path: empty
  - NEXT_HOP: 192.168.0.33 (7 bytes)
    - Flags: 0x0 (Well-known, Transitive, Complete)
    - Type code: NEXT_HOP (3)
    - Length: 4 bytes
    - Next hop: 192.168.0.33 (192.168.0.33)
  - MULTI_EXIT_DISC: 0 (7 bytes)
    - Flags: 0x0 (Optional, Non-transitive, Complete)
    - Type code: MULTI_EXIT_DISC (4)
    - Length: 4 bytes
    - Multiple exit discriminator: 0
  - LOCAL_PREF: 100 (7 bytes)
    - Flags: 0x0 (Well-known, Transitive, Complete)
    - Type code: LOCAL_PREF (5)
    - Length: 4 bytes
    - Local preference: 100
  - COMMUNITIES: 65033:500 65033:600 (11 bytes)
    - Flags: 0x0 (Optional, Transitive, Complete)
    - Type code: COMMUNITIES (8)
    - Length: 8 bytes
    - Communities: 65033:500 65033:600
      - Community: 65033:500
      - Community AS: 65033
      - Community value: 500
      - Community: 65033:600
      - Community AS: 65033
      - Community value: 600
  - Network layer reachability information: 2 bytes
    - 10.0.0.0/8
    - NLRI prefix length: 8
    - NLRI prefix: 10.0.0.0 (10.0.0.0)

**KEEPALIVE Message**
- Marker: 16 bytes
- Length: 19 bytes
- Type: KEEPALIVE Message (4)