BGP

****

- ✓ It fall under the categories of Exterior Gateway Protocols (EGP).
- ✓ The Main Purpose of BGP is to exchange routes between 2 Autonomous Systems.
- ✓ It serves as a gateway protocol that facilitates the exchange of routing information between autonomous systems (AS).

- ✓ There is a flavor of BGP that is run within the Autonomous System. This is called iBGP.
- ✓ Imagine you're dealing with a large number of external BGP (eBGP) routes (500,000 routes) across multiple locations.
- ✓ BGP can handle a significantly larger number of routes than IGP protocols like OSPF , IS-IS & EIGRP.
- ✓ iBGP is to allow the propagation of BGP routes from one edge to another maintaining BGP Attributes like the AS-Path.

- ➢ Path Attributes: BGP includes various path attributes that describe the characteristics of the advertised routes.
- ✓ These attributes include things like AS path, next-hop, origin, local preference, and others.
- ✓ Path attributes used for BGP route selection and policy enforcement.

- ✓ AS-Path attribute represents the sequence of AS through which a route advertisement has traversed. It helps prevent routing loops

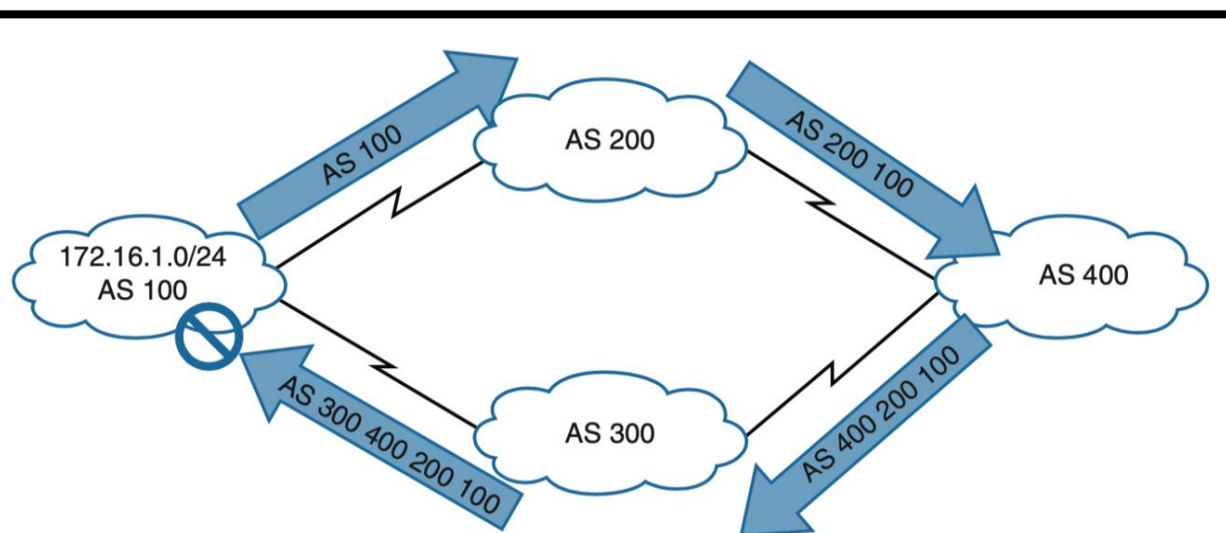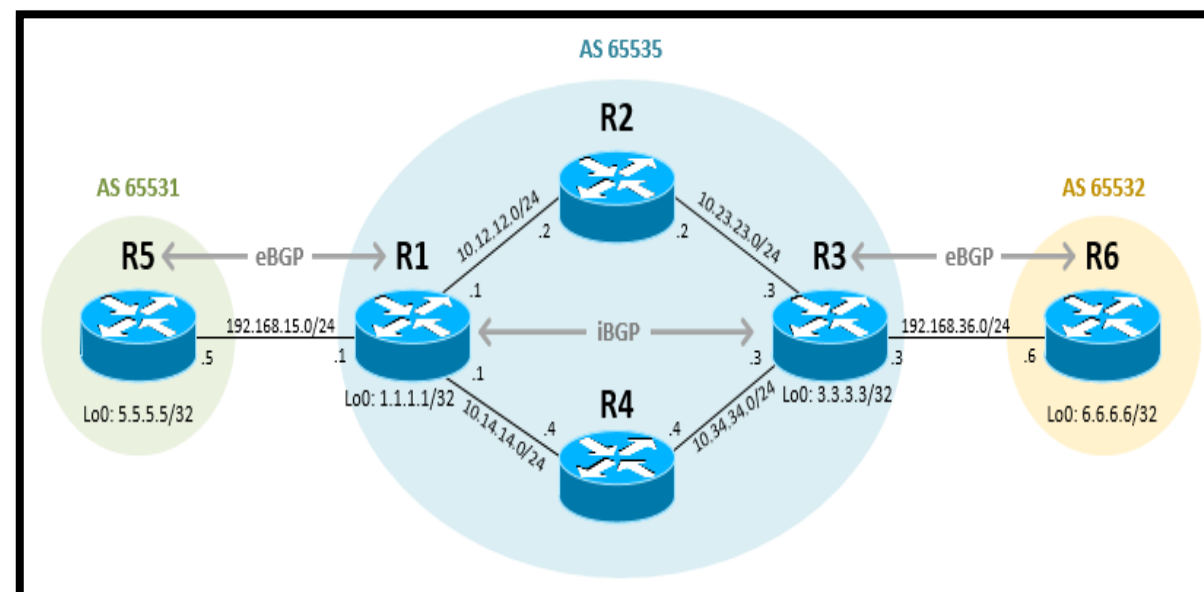Figure 11-1 *Path Vector Loop Prevention*

BGP Neighbour:

✓ BGP Neighbor router is also known **as BGP Peer.**

✓ **As Neighbor** is **manually** defined, unlike other routing protocols, **it doesn't form neighborship automatically.**

✓ TCP Connection is formed with the manually defined neighbor **using TCP port 179.**

✓ TCP is used for **retransmitting lost data**, **avoiding duplicate data**, **and ensuring accurate transmission**

✓ **Administrative Distance:**

- **20 for External BGP (EBGP)**
- **200 for Internal BGP (IBGP).**

BGP Tables:

➤ **Neighbour Table:**

•Information about all the **manually defined neighbors will be** managed in the neighbor **table**

· **# show ip BGP neighbors**

· **# show ip BGP summary**

➤ **BGP Database Table:**

•**The list of all possible networks learned by BGP & their attributes and paths in this table.**

· **# show ip BGP**

➤ **Routing Table:**

- **Only the best path will be available in the IP routing table.**
- **# show ip Route BGP**

Open:

It is used to initiate a BGP session between 2 routers. Includes the below:

    BGP version

    Sender AS number

    Hold Time (Cisco default 180 seconds)

    BGP identifier

    optional parameter **(such as Authentication)**

Default Timers
Keepalive: 60 seconds
Holdtime: 180 Seconds

Keepalive Message:

- This Message is used to maintain the BGP session between the BGP Peers
- It will be sent once in one-third the "hold time". The default timer in cisco device is 60 seconds.
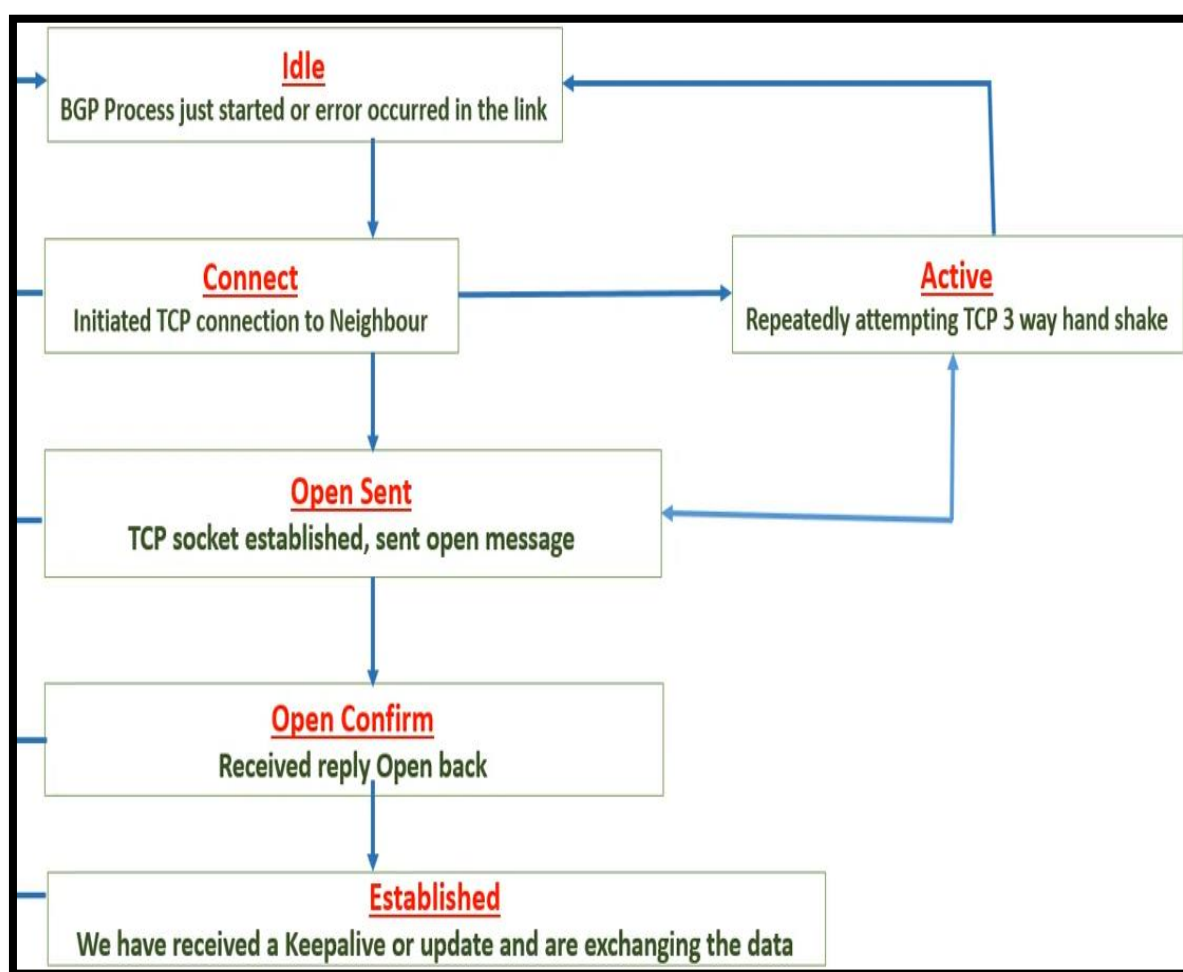
Update:

    If any new route added and existing route deleted. **(NLRI) Network layer reachability information**

    Path Attributes

Notification:

    Used to report errors or terminate the session.

    Can include error message (" BGP version not supported")

BGP State Diagram

Idle
BGP Process just started or error occurred in the link

Connect
Initiated TCP connection to Neighbour

Active
Repeatedly attempting TCP 3 way hand shake

Open Sent
TCP socket established, sent open message

Open Confirm
Received reply Open back

Established
We have received a Keepalive or update and are exchanging the data

---

1-Idle:
- The initial state where BGP resources are initialized.
- In this state, inbound connection attempts are refused.

2-Connect:
- BGP waits for the 3-way handshake to complete.
- If successful, the OPEN message is sent to the peer
- BGP transitions to the Open Sent state.

3-Active:
- If the 3-way handshake fails, BGP moves to this state.
- It retries the connection.

4-OpenSent:
- BGP has initiated the TCP connection and sent the OPEN message.
- It awaits acknowledgment from the peer.

5-Open Confirm:
- The peer acknowledges the OPEN message,
- BGP confirms the session establishment.

6-Established:
- The BGP neighbor adjacency is complete,
- routers exchange UPDATE packets containing routing information.

•**Following reasons**:
- Neighbor is peering with the wrong address.
- Neighbor does not have a neighbor statement for this router.
- AS number is misconfigured.
- Neighbor does not have a route to the source IP address of the BGP open packet generated by this router.
- Any extra command missing (update source).

BGP Attributes and Path Selection

- (BGP), route advertisements consist of the (**Network Layer Reachability Information (NLRI) and the path attributes (PAs).**
- The NLRI consists of the network **prefix and prefix length**; BGP attributes such as AS Path and origin are stored in the PAs.
- A BGP route may contain **multiple paths** to the same destination network.
- Every path's attributes impact the desirability of the route when a router **selects the best path**.

```
˅ Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 56
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 28
  ˅ Path attributes
    > Path Attribute - ORIGIN: IGP
    > Path Attribute - AS_PATH: empty
    > Path Attribute - NEXT_HOP: 10.1.7.7
    > Path Attribute - MULTI_EXIT_DISC: 0
    > Path Attribute - LOCAL_PREF: 100
  ˅ Network Layer Reachability Information (NLRI)
    ˅ 150.1.7.7/32
        NLRI prefix length: 32
        NLRI prefix: 150.1.7.7
˅ Border Gateway Protocol - UPDATE Message
```

- ✓ BGP use attributes to decide the best route.
- ✓ BGP attributes are similar to metrics in OSPF and EIGRP.
- ✓ BGP selects the best path based on a list of attributes.
- ✓ BGP has 0 to 14 attributes for Cisco.
- ✓ BGP does not use metrics but use set of attributes.
- ✓ BGP has four main types of attributes.
- ✓ **Well-Known Mandatory.**
- ✓ **Well-Known Discretionary.**
- ✓ **Optional Transitive.**
- ✓ **Optional Non-Transitive.**

Well-Known Mandatory:

✓ These attributes are required by all BGP peers. (BGP Implementation)
✓ These attribute must be included in all BGP updates.
✓ **If these attribute are missing a Notification, error is generated.**
✓ **If these attribute are missing the session will be closed.**
✓ Well-Known mandatory attributes are **AS Path**, **Next Hop Address**, **& Origin**

Type Code	Attribute Name	Category
1	Origin	Well-Known Mandatory
2	AS Path	Well-Known Mandatory
3	Next Hop	Well-Known Mandatory
4	Multi Exit Disc (MED)	Optional Non-Transitive
5	Local Pref	Well-Known Discretionary
6	Atomic Aggregate	Well-Known Discretionary
7	Aggregator	Optional Transitive
8	Community	Optional Transitive
9	Originator ID	Optional Non-Transitive
10	Cluster List	Optional Non-Transitive

Well-Known Discretionary: (Not need to be in update)

✓ These attributes are recognized & supported by all BGP speakers.
✓ May or may not appear in every BGP Update message **(Not mandatory)**
✓ Well-Known Discretionary attributes are **Local Preference & Atomic Aggregate**.

Optional Transitive:

✓ These attributes are **recognized by some BGP routers** but **not necessarily by all**.
✓ Will be passed on if not recognized by the receiver.
✓ Transitive, these attributes are across AS boundaries **(inter-AS routing)**
✓ Optional Transitive attributes are **Aggregator and Community**.

Optional Non-Transitive:

✓ These attributes are **specific to a particular vendor** or **implementation** and are **not understood by all BGP routers**
✓ The attribute should be ignored and not passed on to other peers.
✓ Non-transitive, these attributes are restricted **to the same AS**.
✓ Optional Non-Transitive attributes are **MED, Originator ID and Cluster List.**

Order of BGP Path Attributes

Order	Preference	Description
0. Synchronized	TRUE	Use only routes that meet the synchronization requirement
1. Weight	Highest	Administrative override
2. Local Preference	Highest	Used internally to pick path out of AS
3. Self originated	TRUE	Used to prefer paths originated on this router
4. AS-Path	Shortest	Minimize AS-hops
5. Origin	i<?	Prefer stability
6. MED	Lowest	Used external to come in
7. External	EBGP<IBGP	External path preferred over internal path
8. IGP cost	Lowest	Look for more information
9. EBGP Peering	Oldest	Prefer stability
10. RID	Lowest	Choose one with lowest BGP router ID

BGP 12 Step Program

Prerequisite: no AS loop, good next hop, synchronized if necessary.

1. Highest weight
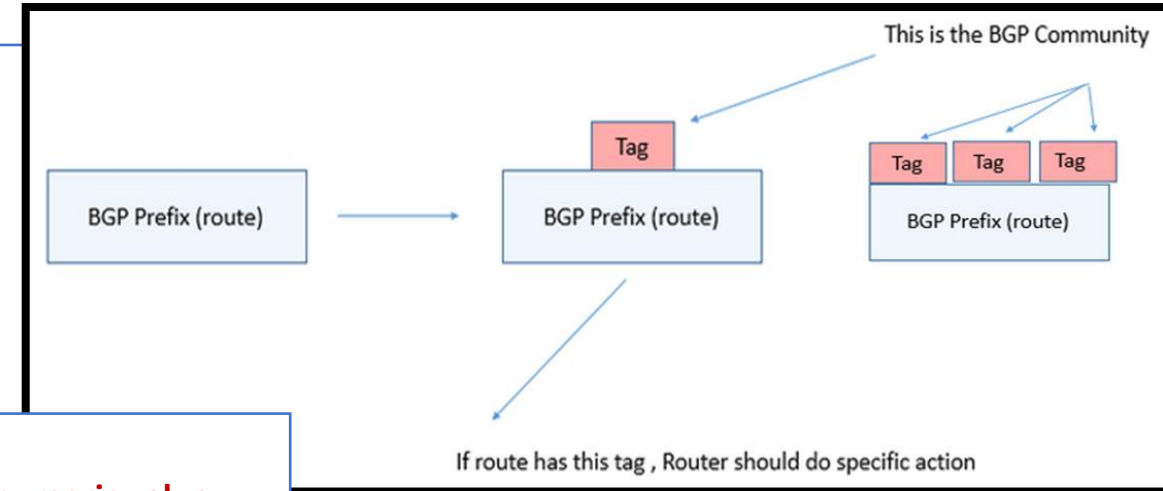2. Highest local preference
3. Locally originated   Next hop 0.0.0.0
4. Shortest AS path length
5. Origin code
6. Lowest MED
7. EBGP over IBGP
8. If routed to neighbor, prefer the path with lowest IGP metric to next hop

*Consider multipath in RIB

9. If external, prefer older one (> 1 min)
10. Lowest router ID or originator ID
11. Minimum cluster list length
12. Lowest neighbor address
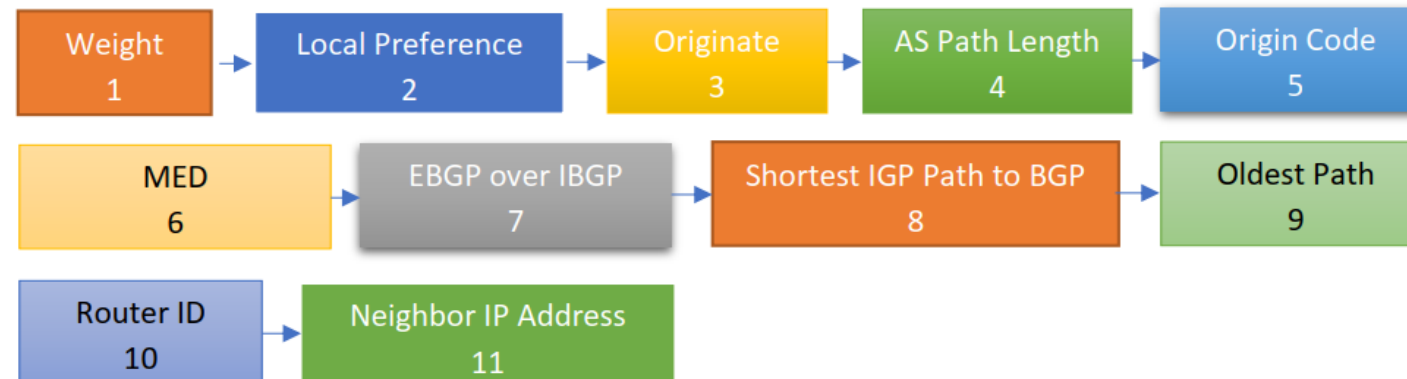
- ✓ **The BGP community** is the tagging mechanism **we use to mark our BGP prefixes**.
- ✓ When the neighbor receives the prefix, it will **examine** the community value and **take specific action**
- ✓ They are often used for **traffic engineering**, **dynamic routing policies**, and **route filtering.**
- ✓ It is an Optional, Transitive BGP attribute, By default, **are not sent with BGP updates**,
- ✓ you need to explicitly use the **send-community command**.
- ✓ Any BGP prefix **can have more than one tag (BGP community);** leads to specific unique action.
- ✓ Can attach up to **32 communities** to a **single** route

There are three types of BGP communities:
- ➢ Standard communities **Well-known & user defined**
- ➢ Extended communities
- ➢ • Large communities

This is the BGP Community

If route has this tag , Router should do specific action

Well-known Standard BGP communities
There are four well-known Standard BGP communities that **can use instead of picking a numeric value**

•**No-Export:** Don't advertise this tagged route to **EBGP peers**, in other meaning
(Only advertise to IBGP neighbors & BGP confederations.)

•**No-Advertise:** Don't advertise this tagged route to any peer, **(eBGP& iBGP)**
(Tell receiving router to keep the tag to yourself and do not pass to any neighbor).
****

•**Local-as:** prevent sending tagged routes outside the local AS within the Confederation.

•**Internet:** advertise these tagged routes to all neighbors **(IBGP & EBGP).**

There are three types of BGP communities:
- ➢ Standard communities **Well-known & user defined**
- ➢ Extended communities
- ➢ • Large communities

This is the BGP Community

Tag

BGP Prefix (route) → BGP Prefix (route)

Tag | Tag | Tag

BGP Prefix (route)

If route has this tag , Router should do specific action
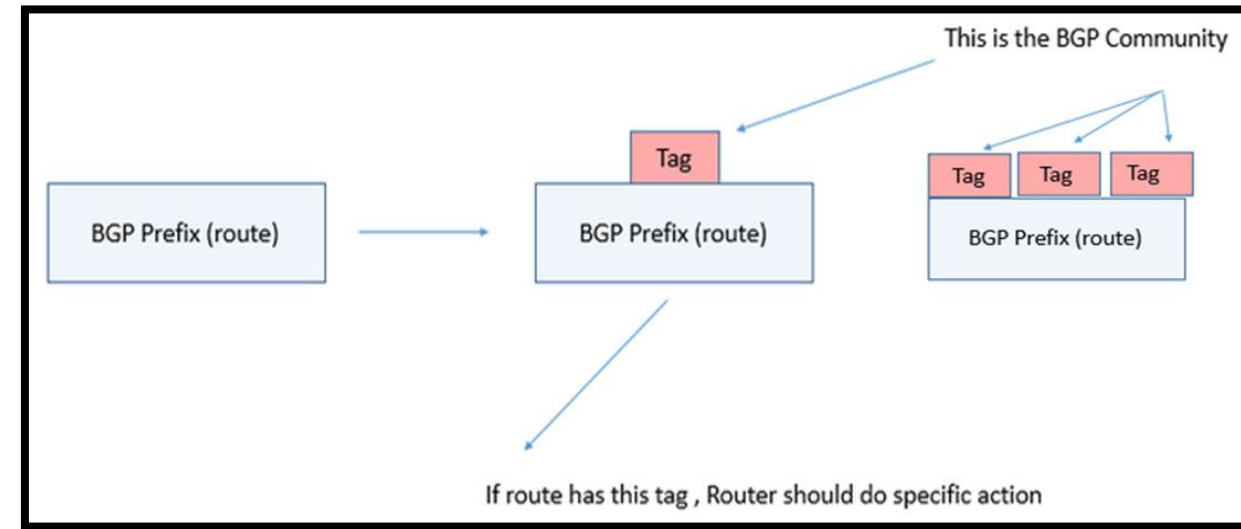
Standard communities user defined :-written as numeric 32-bit tags
in (AS : Action) format
The first 16 bits is the (AS) number   **65000 : 10**
The second 16 bits have the local significance **(Action)**
The primary purpose of standard communities is to the group
and tag routes so that Actions can perform.

Extended communities:- written as numeric 64-bit tags in (AS: Membership)
Route Target (RT) extended community use to **control the exporting and importing of VPN routes.**
(RT) used in MPLS L3 VPN with MP-BGP
Other types of extended communities, such as to **Site of Origin (SOO), Ethernet VPN (EVPN), OSPF Domain Identifier**.

Large communities :- 96-bit tags in (AS: Action : Target AS)

Well-known Standard BGP communities

There are four well-known Standard BGP communities that **can use instead of picking a numeric value**

•**No-Export:** Don't advertise this tagged route to **EBGP peers**, in other meaning

(Only advertise to IBGP neighbors & BGP confederations.)

•**No-Advertise:** Don't advertise this tagged route to any peer, **(eBGP& iBGP)**

(Tell receiving router to keep the tag to yourself and do not pass to any neighbor).

****

•**Local-as:** prevent sending tagged routes outside the local AS within the Confederation.

LOCAL AS (only for IBGP (NOT eBGP & eBGP Confederation)

•**Internet:** advertise these tagged routes to all neighbors **(IBGP & EBGP).**
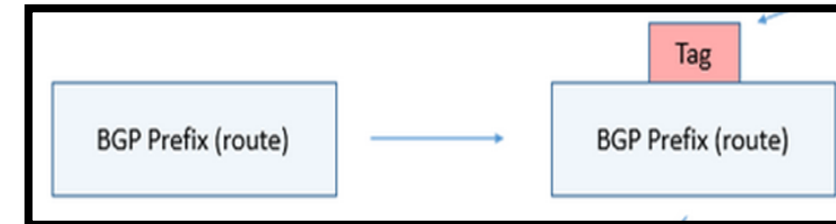
```
R1(config)#route-map map1
R1(config-route-map)#set community ?
  <1-4294967295>  community number
  aa:nn           community number in aa:nn format
  gshut           Graceful Shutdown (well-known community)
  internet        Internet (well-known community)
  local-AS        Do not send outside local AS (well-known community)
  no-advertise    Do not advertise to any peer (well-known community)
  no-export       Do not export to next AS (well-known community)
  none            No community attribute
```

Full Mesh Problem:

➢ In a typical IBGP setup, all routers within an AS must be fully meshed
➢ Establish peering session with every other router.
➢ Large number of TCP sessions
➢ Manual configuration

Solutions:-

1-Route Reflector Clusters
2-Confederations

BGP Confederations:-

✓ **Feature used** to **split** large Autonomous System into Sub-Autonomous Systems
✓ **Internal BGP:** Each sub-AS, **routers use (IBGP) to share** routing information. to **keep routing updates** within the sub-AS efficient
✓ **Confederation Peers: Routers at the borders** of sub-ASes **communicate with each other**
✓ **Using a special type** of **BGP session called a confederation peer session**.
✓ **These sessions exchange routes between sub-ASes**.

BGP Confederations Benefits :-

✓ Management.
✓ Using different BGP Attributes within sub-AS
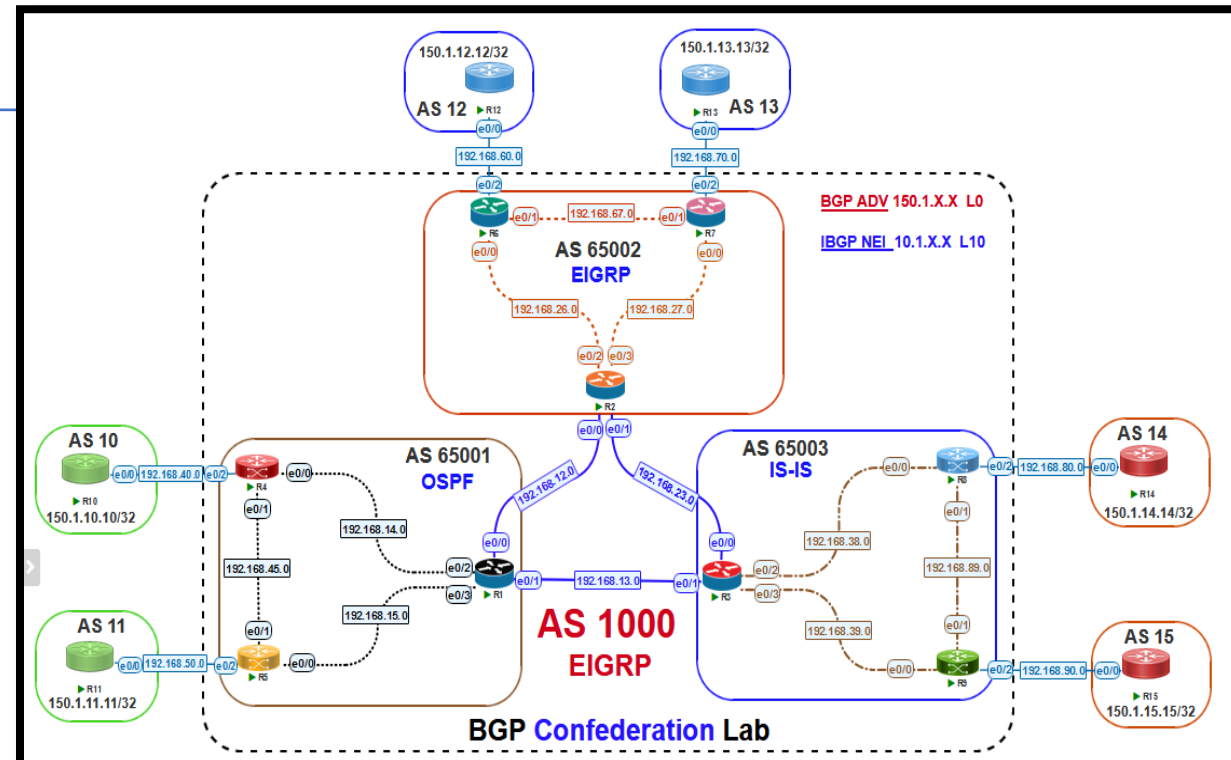
BGP **Confederation** Lab

BGP 12 Step Program

Prerequisite: no AS loop, good next hop, synchronized if necessary.

1. Highest weight
2. Highest local preference
3. Locally originated   Next hop 0.0.0.0
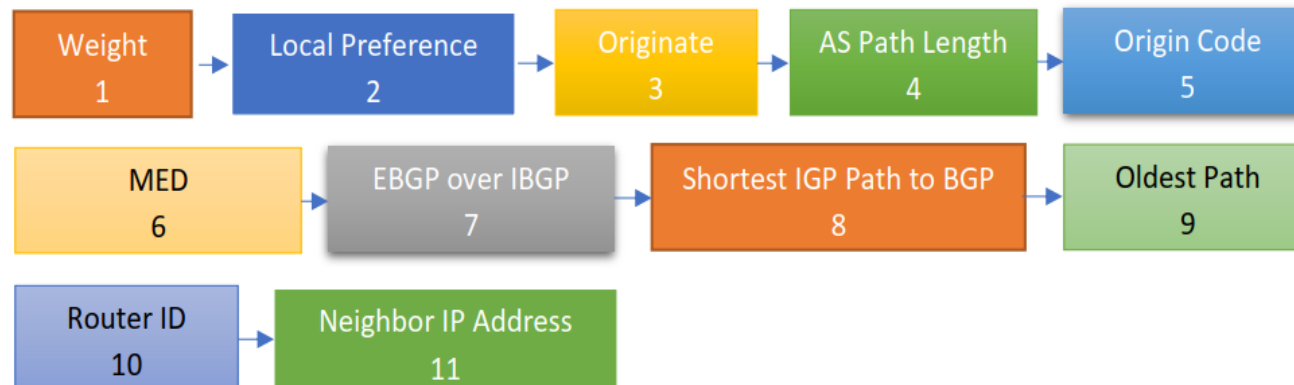4. Shortest AS path length
5. Origin code
6. Lowest MED
7. EBGP over IBGP
8. If routed to neighbor, prefer the path with lowest IGP metric to next hop

*Consider multipath in RIB

9. If external, prefer older one (> 1 min)
10. Lowest router ID or originator ID
11. Minimum cluster list length
12. Lowest neighbor address

BGP Best-path Algorithm

The following list provides the attributes that the BGP best-path algorithm uses for the process of selecting the best route. These attributes are processed in the order listed:

1. Prefer the highest *weight*
2. Prefer the highest *local preference*
3. Prefer the route *originated* by the local router
4. Prefer the path with the shorter Accumulated Interior Gateway Protocol (AIGP) metric attribute
5. Prefer the shortest *AS_Path*
6. Prefer the best *origin* code
7. Prefer the lowest multi-exit discriminator (*MED*)
8. Prefer an *external* path over an *internal* path
9. Prefer the path through the *closest IGP neighbor*
10. Prefer the *oldest route* for eBGP paths
11. Prefer the path with the lowest *neighbor BGP RID*
12. Prefer the path with the lowest *neighbor IP address*

cisco

| Weight 1 | Local Preference 2 | Originate 3 | AS Path Length 4 | Origin Code 5 |

| MED 6 | EBGP over IBGP 7 | Shortest IGP Path to BGP 8 | Oldest Path 9 |

| Router ID 10 | Neighbor IP Address 11 |

- **Large networks sometimes use** multiple, separate **BGP Autonomous Systems (ASes) for administrative reasons,**
- even though they're **under a single management domain.**

- **Traditional BGP relies on attributes like AS path length**, which **doesn't consider the "internal" costs** within each AS.
- This can lead BGP to choose **suboptimal paths across the entire network.**

- **AIGP is the Solution: (Accumulated Interior Gateway Protocol)** **(metric attribute)**
- AIGP is an **optional non-transitive BGP attribute**. This means it's **not passed** on indefinitely **between BGP routers.**
- The AIGP attribute carries the accumulated metric from an Interior Gateway Protocol (IGP) within an AS.
- All ASes must be under the control of single domain
- BGP routers **can advertise this AIGP metric to their neighbors**, providing **visibility into the** "internal costs" of paths within other ASes.

- **AIGP metric:-** **can make decisions and choose the best path** for traffic flow **even if it traverses multiple Ases under the same administrative control.**

- ✓ Prior to **January 2009**, BGP AS numbers were **16-Bit** **2-octet numbers** in the range from **1 to 65535**
- ✓ Due to increased demand, the Internet Assigned Number Authority (IANA) started allocating **4-octet AS numbers**
- ✓ An extension has been created that supports **32-Bit or 4-Octed or 4-Bytes AS numbers** total about **4294967296** .

AS Range	Purpose	Total AS
0	Reserved	1
1 – 64495	Public AS Numbers	64495
64496 – 64511	Reserved to use in Documentation	15
64512 – 65534	Private AS Numbers	1023
65535	Reserved	1
4200000000 - 4294967294	Private AS Numbers	94,967295

- **AS 23456, also known as AS_TRANS**, is a reserved **specifically for use in BGP communication between routers**.
- When a **new router** (supporting 4-byte AS numbers) communicates with an **old router** (supporting only 2-byte AS numbers), it can use AS_TRANS instead of its actual 4-byte AS number.

2   . 18928

150000/65536=2

2*65536=131,072