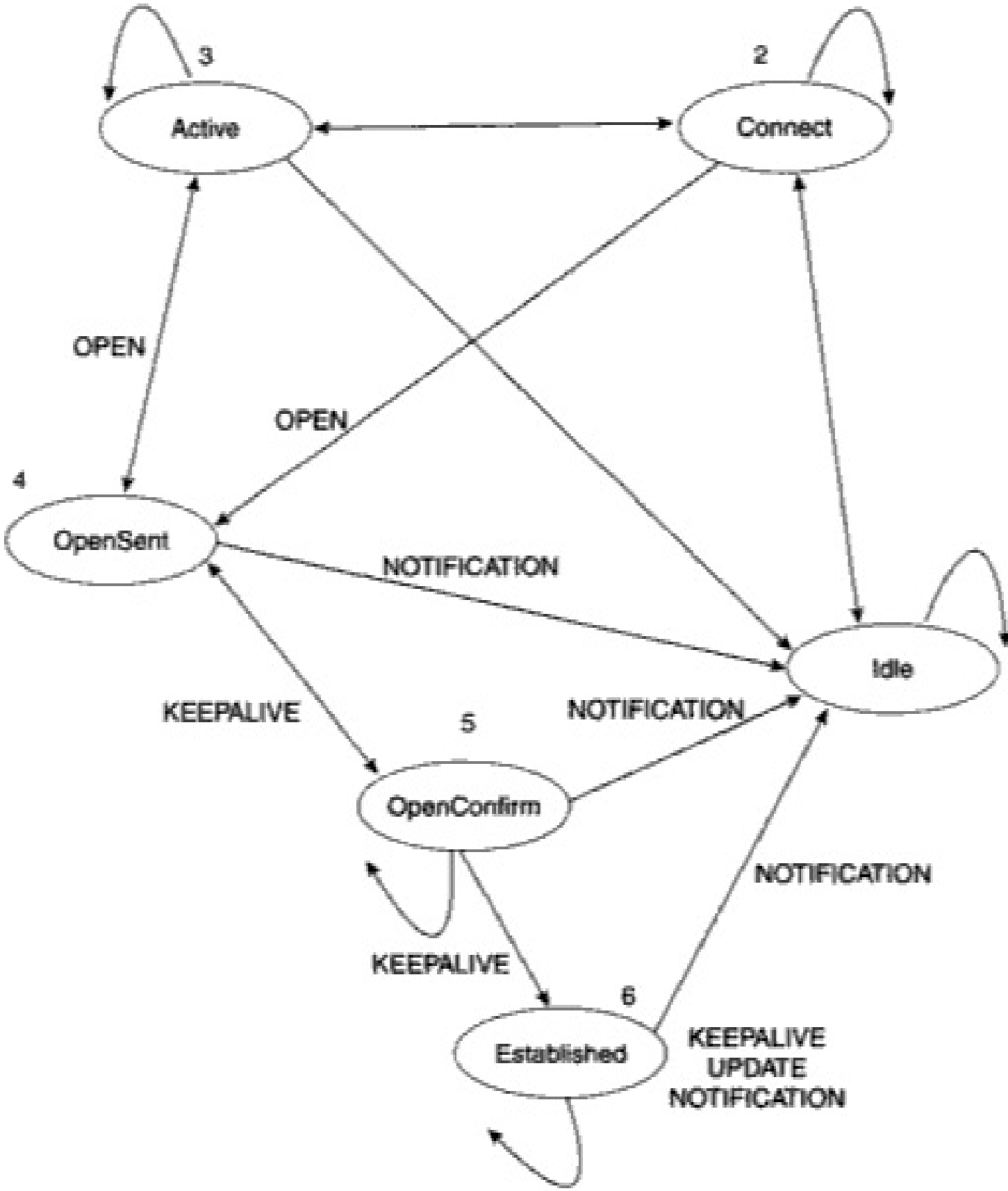
e State Machine Perspective

BGP neighbor negotiation proceeds through different stages before the connectio
established. Figure 5-8 illustrates a simplified finite state machine (FSM) that high
major events in the process with an indication of messages (OPEN, KEEPALIVE
NOTIFICATION) sent to the peer in the transition from one state to the other.

5-8 *BGP Neighbor Negotiation Finite State Machine*



Foundation Summary

The "Foundation Summary" section of each chapter lists the most important facts from the chapter. Although this section does not list every fact from the chapter that will be on your exam, a well-prepared candidate should, at a minimum, know all the details in each "Foundation Summary" before going to take the exam.

The key features of BGP include the following:

- BGP is an enhanced path vector protocol.
- BGP supports VLSM, CIDR, and summarization.
- BGP creates and maintains connections between peers, using the destination TCP port 179 to initiate the connection.
- The connection is maintained by periodic keepalives.
- The failure to see a keepalive, an update, or the receipt of a notification is the means by which destination networks and paths to those destinations are tracked. Any change in the network results in a triggered update.
- The metric used in BGP is intricate and is the source of its complexity and its strength. The metric, referred to as *attributes*, allows great granularity in path selection.
- The use of hierarchical addressing and the capability to manipulate traffic flow results in a network that is designed to grow.
- BGP has its own routing table, although it is capable of both sharing and inquiring about the interior IP routing table.
- It is possible to manipulate the traffic flow by using the complex metric called *attributes*. Despite the complexity offered in path selection using policy-based routing, the traffic is still forwarded using the hop-by-hop paradigm. This means that no router can send traffic on a route that the next-hop router would not choose for itself.

Table 15-5 explains the four categories of attributes.

Table 15-5 *The Four Categories of Attributes*

Category	Description
Well-Known:	
Mandatory (required by all routers)	These attributes are required and are therefore recognized by all BGP implementations.
Discretionary (required by all routers and recognized by all routers)	These attributes are not required to be present in the update messages, but if they are present, all routers running BGP will recognize and act on the information contained.
Optional:	
Transitive	The router might not recognize these attributes, but if this is the case, it marks the update as partial and sends the update, complete with attributes, to the next router. The attributes traverse the router unchanged, if they are not recognized.
Nontransitive	Nontransitive attributes are dropped if they fall onto a router that does not understand or recognize the attribute. These attributes will *not* be propagated to the BGP peers. Unrecognized nontransitive optional attributes must be quietly ignored and not passed along to other BGP peers. New transitive optional attributes might be attached to the path by the originator or by any other autonomous system in the path (see RFC 1771).

The attributes are appropriately carried in the updates that inform the routers of the routes.

The attributes and a description of their characteristics are shown in Table 15-6.

Table 15-6 *The BGP Attributes*

Attribute Name	Category	Code	Preference	Description
Origin	Well-known, mandatory	1	Lowest origin code Where: IGP < EGP < Incomplete	This path attribute identifies the source of the routing update. The possible sources of routing information are as follows: • The path originates from within the autonomous system. It was created with the iBGP **network** command. The route will be marked in the BGP routing table with an "i." • If the source is an exterior routing protocol, it will be identified with an "e" in the routing table. • The route could have been redistributed into BGP; as such, there is incomplete information. The route is marked by a question mark (?).
AS_Path	Well-known, mandatory	2	Shortest path	This is a sequence of the autonomous systems that the prefix has traversed.
Next hop	Well-known, Mandatory	3	Shortest path or IGP metric	The next hop attribute states the next hop on the path for the router to take. In eBGP, this will be the source address of the router that sent the update from the other autonomous system. In iBGP, for routes that originated outside the autonomous system, the address will still be the source address of the router that sent the update. The protocol states that the next hop advertised by eBGP should be carried into the iBGP. Therefore, it is important that the IGP is aware of this network so that any router within the autonomous system can reach the next hop.

Table 15-6 *The BGP Attributes (Continued)*

Attribute Name	Category	Code	Preference	Description
Multiple Exit Discriminator (MED)	Optional, nontransitive	4	Lowest value	This attribute informs routers outside the autonomous system which path to take into the autonomous system. It is known as the external metric of a route. Therefore, it is passed between the autonomous systems, but it will not be propagated into a third autonomous system.
Local preference	Well-known, discretionary	5	Highest value	This attribute is used to tell routers within the autonomous system how to exit the autonomous system in the case of multiple paths. It is the opposite of the MED attribute. This value is passed solely between iBGP peers.
Atomic aggregate	Well-known, discretionary	6	Information not used in path selection	This attribute states that the routes have been aggregated and that some information has been lost.
Aggregator	Optional, transitive	7	Information not used in path selection	This attribute states the BGP router ID and the autonomous system number of the router that was responsible for aggregating the route.
Community	Optional, transitive	8	Information not used in path selection	This is the capability to tag certain routes that have something in common. They are thereby made members of the same "club" or community. This is often used in conjunction with another attribute that will affect route selection for the community. For example, the use of the local preference and community attributes would allow the network administrators and other privileged beings to use the high-speed link to the Internet, while others shared a fractional T1. Communities have no geographical or logical limits. BGP can filter on incoming or outgoing routes for filtering, redistribution, or path selection.

(continues)

Table 15-6 *The BGP Attributes (Continued)*

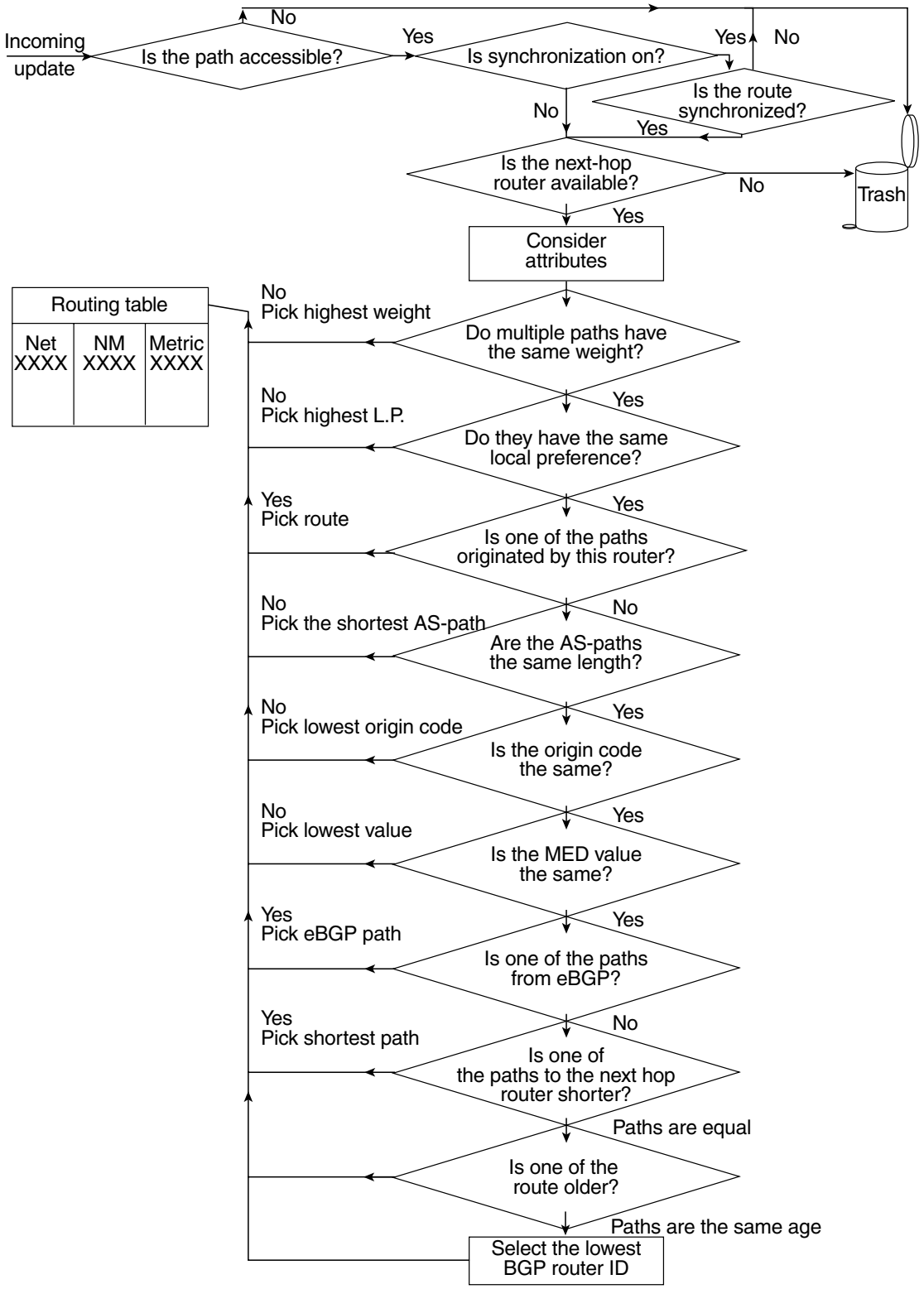
Attribute Name	Category	Code	Preference	Description
Originator ID	Optional, nontransitive	9	Information not used in path selection	The route reflector (described in the following chapter) appends this attribute. It carries the router ID of the originating router in the local autonomous system. It is used to prevent loops.
Cluster ID	Optional, nontransitive	10	Information not used in path selection	The cluster identifies the routers involved in the route reflection. The cluster list shows the reflection path that has been taken. This is used to prevent looping errors.
Weight	Cisco-defined		Highest value	This is proprietary to Cisco and is used in route selection. It is local to the router, and because it is not propagated to other routers, there is no problem with compatibility. When there are multiple paths, it selects a path to a destination with different next hops to the same destination. Note that the weight attribute has no code. Because it is a local attribute and is not propagated to other routers, no code is needed.

Figure 15-10 shows the logic of the path selection used in BGP.

Use Figure 15-10 in association with the following list of the selection process steps:

1. If the router has a valid route to the destination, use that route.
2. If there is more than one valid route to the destination, take the route with the highest weight (Cisco proprietary).
3. If the weights are the same, select the route with the highest local preference.
4. If the routes have the same local preference, prefer the route that originated on the local router.
5. If there are no routes that originated on the router, examine the AS_Path and select the shortest path.
6. If the AS_Path is the same, examine and choose the lowest origin code.
7. If the origin codes are the same, select the path with the lowest MED (the MED values must have been sent from the same neighboring autonomous system).

Figure 15-10 *Path Selection in BGP*



8. If the MED values are the same, choose an external BGP route over an internal BGP route.
9. If there is no external route, choose the path with the lowest IGP metric or cost to the next-hop router for iBGP.
10. If the paths are equal, select the oldest path.
11. If all else fails, choose the router with the lowest BGP router ID. The router ID is either the highest IP address, with preference given to the loopback interface, or manually configured.

Table 15-7 summarizes the commands covered in this chapter.

Table 15-7 *Summary of BGP Commands*

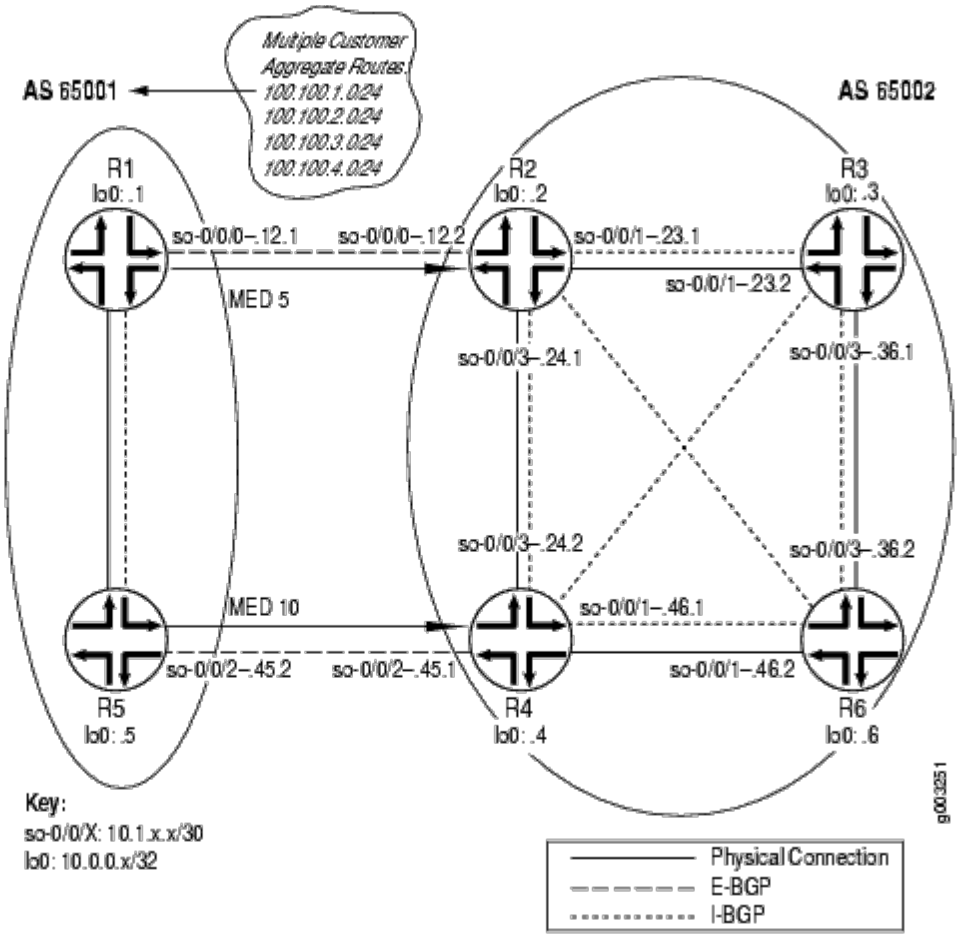
Command	Function
router bgp *autonomous-system-number*	Starts the BGP routing process.
network *network-number* **mask** *network-mask*	Identifies the networks to be advertised by the BGP process.
neighbor {*ip-address* | *peer-group-name*} **remote-as** *autonomous-system-number*	Identifies the neighbor with whom the router is synchronizing its routing table and activates a TCP session with the neighbor. It also configures the **remote-as** option for a peer group.
neighbor {*ip-address* | *peer-group-name*} **next-hop-self**	To avoid the problem of selecting the next-hop router on a NBMA network inappropriately, this command is used to force the router to use its own IP address as the next hop when advertising to neighbors.
no synchronization	Turns off synchronization and the need for the IGP to know of a route before BGP can advertise it. This is used when the iBGP network is fully meshed.
aggregate-address *ip-address mask* [**summary-only**] [**as-set**]	Used to create an aggregate address. The **summary-only** option advertises the summary, and the **as-set** option lists the autonomous system numbers that the more specific routes have traversed.
debug ip bgp [**dampening** | **events** | **keepalives** | **updates**]	Enables you to be very specific about the BGP **debug** parameters.
clear ip bgp {* | *address*} [**soft** [**in** | **out**]]	Resets the session between the neighbors and reestablishes it with the new configuration that has been entered. The soft option does not tear down the sessions, but it resends the updates. The **in** and **out** options allow the configuration of inbound or outbound soft updates. The default is for both.
show ip bgp [**summary** | **neighbors**]	Shows the BGP connections. A network can be specified to retrieve information on the lone network. The **summary** option will give the status of the BGP connections. The **neighbors** option gives both TCP and BGP connections.

Examine BGP Routes and Route Selection

Purpose

You can examine the BGP path selection process to determine the single, active path when BGP receives multiple routes to the same destination prefix.

Figure 16: BGP Network Topology



The network in Figure 16 shows that R1 and R5 announce the same aggregate routes to R2 and R4, which results in R2 and R4 receiving two routes to the same destination prefix. The route selection process on R2 and R4 determines which of the two BGP routes received is active and advertised to the other internal routers (R3 and R6).

Before the router installs a BGP route, it must make sure that the BGP `next-hop` attribute is reachable. If the BGP next hop cannot be resolved, the route is not installed. When a BGP route is installed in the routing table, it must go through a path selection process if multiple routes exist to the same destination prefix. The BGP path selection process proceeds in the following order:

1. Route preference in the routing table is compared. For example, if an OSPF and a BGP route exist for a particular destination, the OSPF route is selected as the active route because the OSPF route has a default preference of 10, while the BGP route has a default preference of 170.
2. Routes are compared for local preference. The route with the highest local preference is preferred. For example, see Examine the Local Preference Selection.
3. The AS path attribute is evaluated. The shorter AS path is preferred.
4. The origin code is evaluated. The lowest origin code is preferred (`I (IGP) < E (EGP) < ? (Incomplete)`).
5. The MED value is evaluated. By default, if any of the routes are advertised from the same neighboring AS, the lowest MED value is preferred. The absence of a MED value is interpreted as a MED of 0. For an example, see Examine the Multiple Exit Discriminator Route Selection.
6. The route is evaluated as to whether it is learned through EBGP or IBGP. EBGP learned routes are preferred to IBGP learned routes. For an example, see Examine the EBGP over IBGP Selection
7. If the route is learned from IBGP, the route with the lowest IGP cost is preferred. For an example, see Examine the IGP Cost Selection. The physical next hop to the IBGP peer is installed according to the following three rules:

a. After BGP examines the `inet.0` and `inet.3` routing tables, the physical next hop of the route with the lowest preference is used.
b. If the preference values in the `inet.0` and the `inet.3` routing tables are a tie, the physical next hop of the route in the `inet.3` routing table is used.
c. When a preference tie exists in the same routing table, the physical next hop of the route with more paths is installed.

8. The route reflection cluster list attribute is evaluated. The shortest length cluster list is preferred. Routes without a cluster list are considered to have a cluster list length of 0.
9. The router ID is evaluated. The route from the peer with the lowest router ID is preferred (usually the loopback address).
10. The peer address value is examined. The peer with the lowest peer IP address is preferred.

Steps To Take

To determine the single, active path when BGP receives multiple routes to the same destination prefix, enter the following JUNOS CLI operational mode command:

```
user@host> show route destination-prefix < detail >
```

The following steps illustrate the inactive reason displayed when BGP receives multiple routes to the same destination prefix and one route is selected as the single, active path:

1. Examine the Local Preference Selection
2. Examine the Multiple Exit Discriminator Route Selection
3. Examine the EBGP over IBGP Selection
4. Examine the IGP Cost Selection

[Contents] [Prev] [Next] [Index] [Report an Error]