## Chapter Goals

- Understand the purpose of the Border Gateway Protocol.

- Explain BGP attributes and their use in route selection.

- Examine the BGP route selection process.

Border Gateway Protocol

Introduction

The *Border Gateway Protocol (BGP)* is an interautonomous system routing protocol. An autonomous system is a network or group of networks under a common administration and with common routing policies. BGP is used to exchange routing information for the Internet and is the protocol used between Internet service providers (ISP). Customer networks, such as universities and corporations, usually employ an Interior Gateway Protocol (IGP) such as RIP or OSPF for the exchange of routing information within their networks. Customers connect to ISPs, and ISPs use BGP to exchange customer and ISP routes. When BGP is used between autonomous systems (AS), the protocol is referred to as External BGP (EBGP). If a service provider is using BGP to exchange routes within an AS, then the protocol is referred to as Interior BGP (IBGP). Figure 39-1 illustrates this distinction.

Figure 39-1   External and Interior BGP
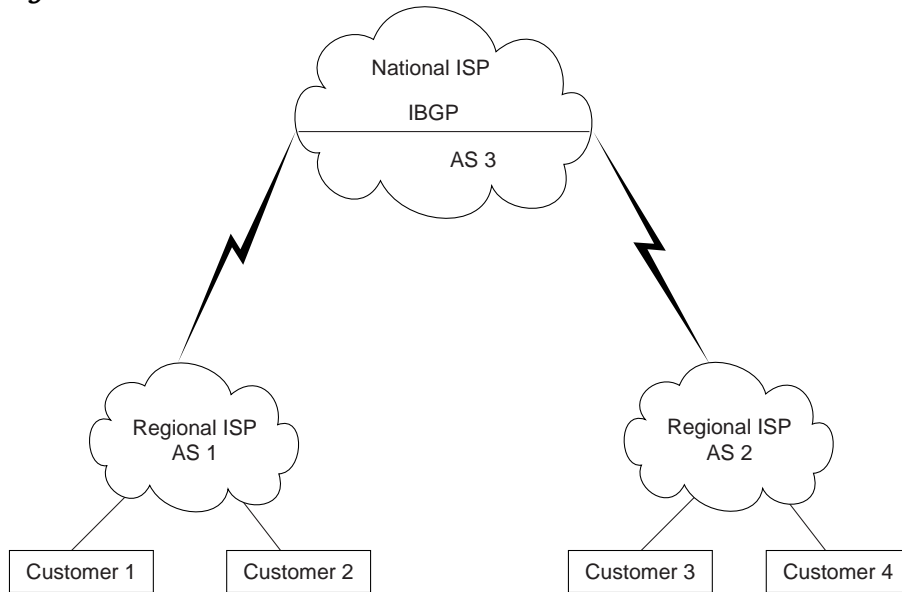
BGP is a very robust and scalable routing protocol, as evidenced by the fact that BGP is the routing protocol employed on the Internet. At the time of this writing, the Internet BGP routing tables number more than 90,000 routes. To achieve scalability at this level, BGP uses many route parameters, called attributes, to define routing policies and maintain a stable routing environment.

In addition to BGP attributes, classless interdomain routing (CIDR) is used by BGP to reduce the size of the Internet routing tables. For example, assume that an ISP owns the IP address block 195.10.x.x from the traditional Class C address space. This block consists of 256 Class C address blocks, 195.10.0.x through 195.10.255.x. Assume that the ISP assigns a Class C block to each of its customers. Without CIDR, the ISP would advertise 256 Class C address blocks to its BGP peers. With CIDR, BGP can supernet the address space and advertise one block, 195.10.x.x. This block is the same size as a traditional Class B address block. The class distinctions are rendered obsolete by CIDR, allowing a significant reduction in the BGP routing tables.

BGP neighbors exchange full routing information when the TCP connection between neighbors is first established. When changes to the routing table are detected, the BGP routers send to their neighbors only those routes that have changed. BGP routers do not send periodic routing updates, and BGP routing updates advertise only the optimal path to a destination network.

BGP Attributes

Routes learned via BGP have associated properties that are used to determine the best route to a destination when multiple paths exist to a particular destination. These properties are referred to as BGP attributes, and an understanding of how BGP attributes influence route selection is required for the design of robust networks. This section describes the attributes that BGP uses in the route selection process:
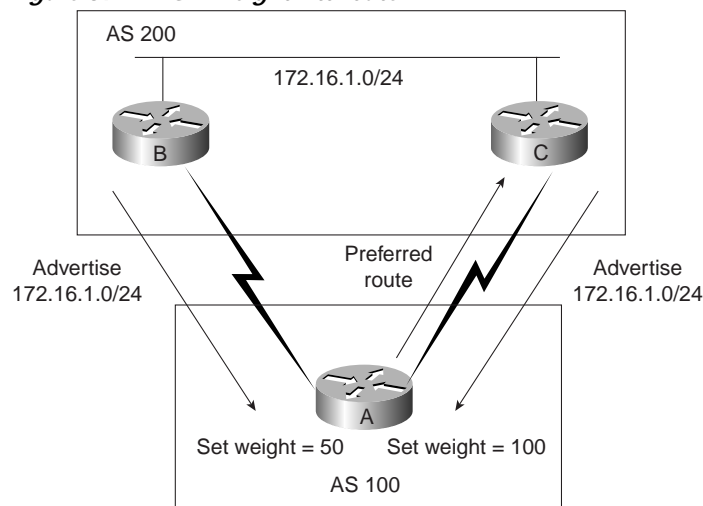
- Weight
- Local preference
- Multi-exit discriminator
- Origin

- AS_path
- Next hop
- Community

Weight Attribute

Weight is a Cisco-defined attribute that is local to a router. The weight attribute is not advertised to neighboring routers. If the router learns about more than one route to the same destination, the route with the highest weight will be preferred. In Figure 39-2, Router A is receiving an advertisement for network 172.16.1.0 from routers B and C. When Router A receives the advertisement from Router B, the associated weight is set to 50. When Router A receives the advertisement from Router C, the associated weight is set to 100. Both paths for network 172.16.1.0 will be in the BGP routing table, with their respective weights. The route with the highest weight will be installed in the IP routing table.

Figure 39-2   BGP Weight Attribute

Local Preference Attribute

The *local preference* attribute is used to prefer an exit point from the local autonomous system (AS). Unlike the weight attribute, the local preference attribute is propagated throughout the local AS. If there are multiple exit points from the AS, the local preference attribute is used to select the exit point for a specific route. In Figure 39-3, AS 100 is receiving two advertisements for network 172.16.1.0 from AS 200. When Router A receives the advertisement for network 172.16.1.0, the corresponding local preference is set to 50. When Router B receives the advertisement for network 172.16.1.0, the corresponding local preference is set to 100. These local preference values will be exchanged between routers A and B. Because Router B has a higher local preference than Router A, Router B will be used as the exit point from AS 100 to reach network 172.16.1.0 in AS 200.

Figure 39-3   BGP Local Preference Attribute
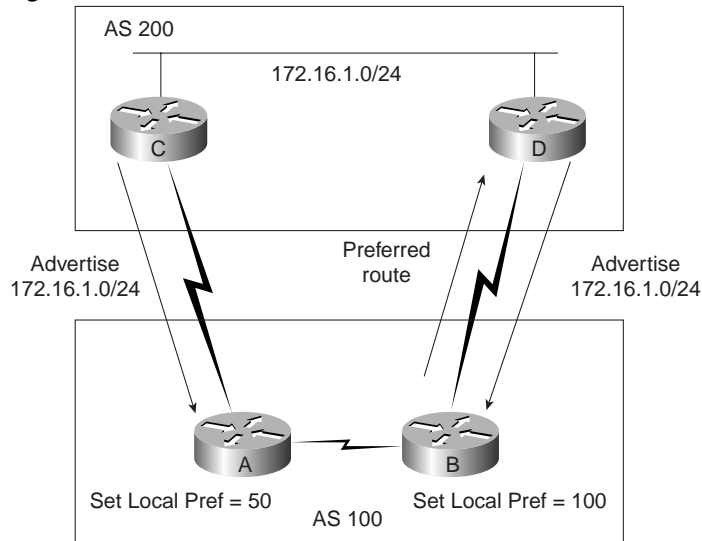
Multi-Exit Discriminator Attribute

The *multi-exit discriminator (MED)* or *metric attribute* is used as a suggestion to an external AS regarding the preferred route into the AS that is advertising the metric.

The term *suggestion* is used because the external AS that is receiving the MEDs may be using other BGP attributes for route selection. We will cover the rules regarding route selection in the next section. In Figure 39-4, Router C is advertising the route 172.16.1.0 with a metric of 10, while Route D is advertising 172.16.1.0 with a metric of 5. The lower value of the metric is preferred, so AS 100 will select the route to router D for network 172.16.1.0 in AS 200. MEDs are advertised throughout the local AS.

Origin Attribute

The *origin attribute* indicates how BGP learned about a particular route. The origin attribute can have one of three possible values:

- **IGP**—The route is interior to the originating AS. This value is set when the network router configuration command is used to inject the route into BGP.

- **EGP**—The route is learned via the Exterior Border Gateway Protocol (EBGP).

- **Incomplete**—The origin of the route is unknown or learned in some other way. An origin of incomplete occurs when a route is redistributed into BGP.

The origin attribute is used for route selection and will be covered in the next section.

Figure 39-4   BGP Multi-Exit Discriminator Attribute
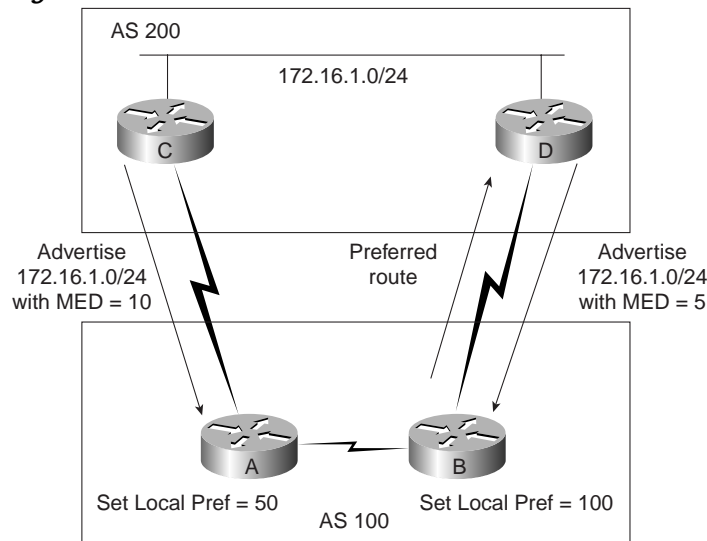
AS_path Attribute

When a route advertisement passes through an autonomous system, the AS number is added to an ordered list of AS numbers that the route advertisement has traversed. Figure 39-5 shows the situation in which a route is passing through three autonomous systems.

AS1 originates the route to 172.16.1.0 and advertises this route to AS 2 and AS 3, with the AS_path attribute equal to {1}. AS 3 will advertise back to AS 1 with AS-path attribute {3,1}, and AS 2 will advertise back to AS 1 with AS-path attribute {2,1}. AS 1 will reject these routes when its own AS number is detected in the route advertisement. This is the mechanism that BGP uses to detect routing loops. AS 2 and AS 3 propagate the route to each other with their AS numbers added to the AS_path attribute. These routes will not be installed in the IP routing table because AS 2 and AS 3 are learning a route to 172.16.1.0 from AS 1 with a shorter AS_path list.

Next-Hop Attribute

The EBGP *next-hop* attribute is the IP address that is used to reach the advertising router. For EBGP peers, the next-hop address is the IP address of the connection between the peers. For IBGP, the EBGP next-hop address is carried into the local AS, as illustrated in Figure 39-6.
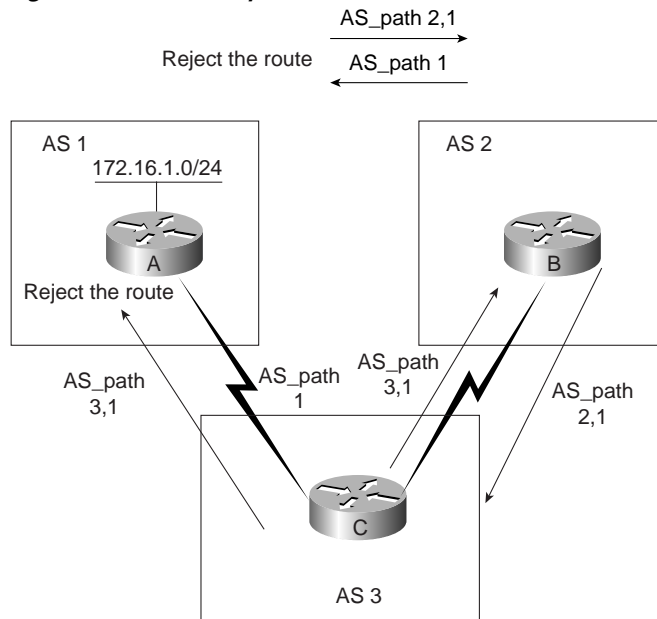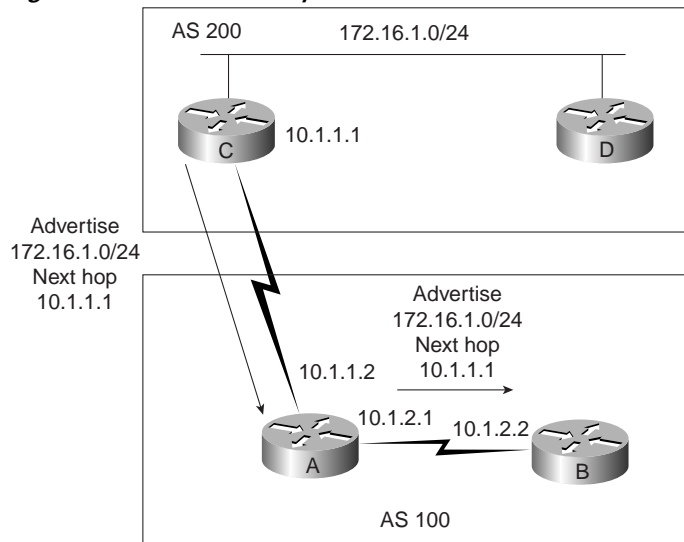
Figure 39-5   BGP AS-path Attribute

Figure 39-6   BGP Next-Hop Attribute

Router C advertises network 172.16.1.0 with a next hop of 10.1.1.1. When Router A propagates this route within its own AS, the EBGP next-hop information is preserved. If Router B does not have routing information regarding the next hop, the route will be discarded. Therefore, it is important to have an IGP running in the AS to propagate next-hop routing information.
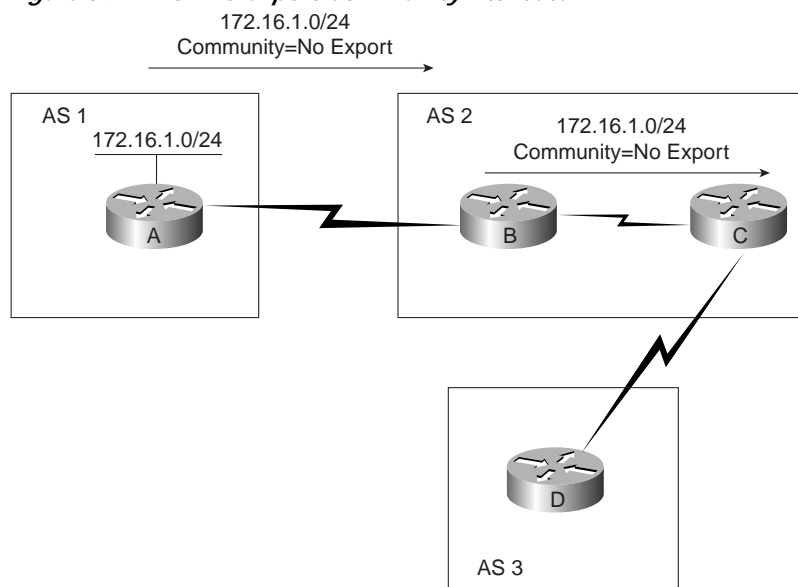
Community Attribute

The community attribute provides a way of grouping destinations, called communities, to which routing decisions (such as acceptance, preference, and redistribution) can be applied. Route maps are used to set the community attribute. Predefined community attributes are listed here:

- **no-export**—Do not advertise this route to EBGP peers.

- **no-advertise**—Do not advertise this route to any peer.

- **internet**—Advertise this route to the Internet community; all routers in the network belong to it.

Figure 39-7 illustrates the no-export community. AS 1 advertises 172.16.1.0 to AS 2 with the community attribute no-export. AS 2 will propagate the route throughout AS 2 but will not send this route to AS 3 or any other external AS.

Figure 39-7   BGP no-export Community Attribute

In Figure 39-8, AS 1 advertises 172.16.1.0 to AS 2 with the community attribute no-advertise. Router B in AS 2 will not advertise this route to any other router.
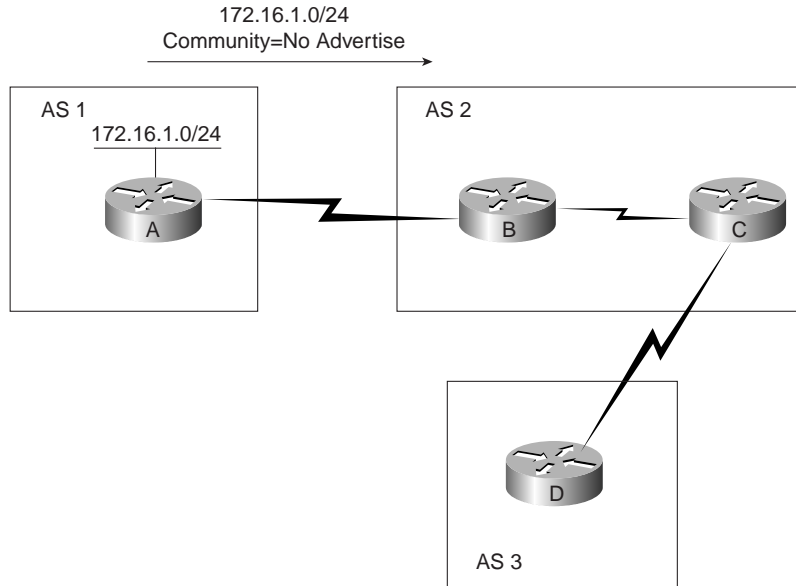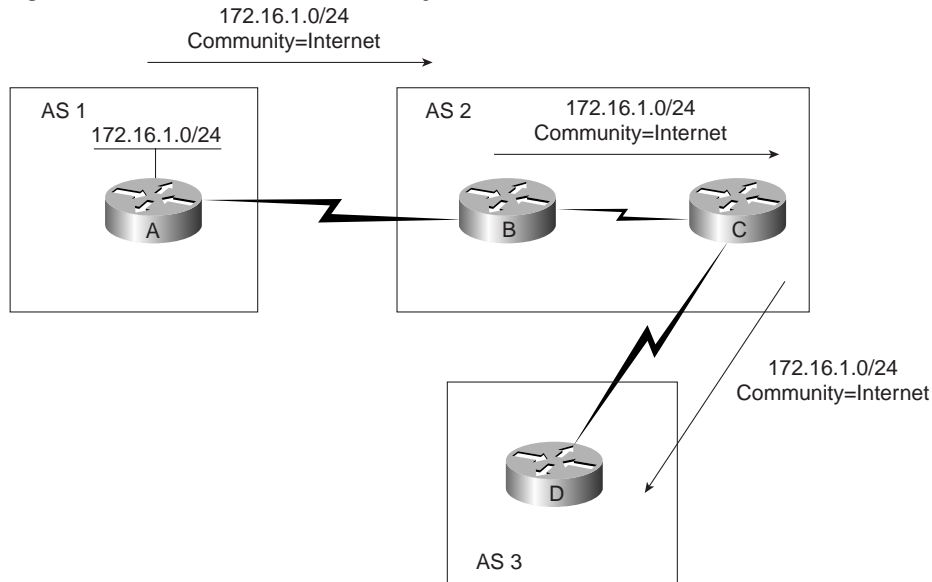
Figure 39-8   BGP no-advertise Community Attribute

172.16.1.0/24
Community=No Advertise

Figure 39-9 demonstrates the internet community attribute. There are no limitations to the scope of the route advertisement from AS 1.

Figure 39-9   BGP internet Community Attribute

172.16.1.0/24
Community=Internet

BGP Path Selection

BGP could possibly receive multiple advertisements for the same route from multiple sources. BGP selects only one path as the best path. When the path is selected, BGP puts the selected path in the IP routing table and propagates the path to its neighbors. BGP uses the following criteria, in the order presented, to select a path for a destination:

•   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, prefer the route that has the shortest AS_path.

- If all paths have the same AS_path length, prefer the path with the lowest origin type (where IGP is lower than EGP, and EGP is lower than incomplete).

- If the origin codes are the same, prefer the path with the lowest MED attribute.

- 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.

Review Questions

Q—*Can IBGP be used in place of an IGP (RIP, IGRP, EIGRP, OSPF, or ISIS)?*

A—Yes and no. Remember that the next-hop information from EBGP is carried into IBGP. If IBGP does not have a route to reach the next hop, then the route will be discarded. Typically an IGP needs to be used to exchange routes to the next hop, but this can be achieved by using static routes on all the routers running IBGP. So, the answer is yes if you want to use and maintain static routes. Otherwise, the answer is no.

Q—*Assume that a BGP router is learning the same route from two different EBGP peers. The AS_path information from peer 1 is {2345,86,51}, and the AS_path information from peer 2 is {2346,51}. What BGP attributes could be adjusted to force the router to prefer the route advertised by peer 1?*

A—Weight and local preference. Both have a higher preference than AS_path length.

Q—*Can BGP be used only by Internet service providers?*

A—No. BGP can be used to scale large enterprise networks. A large network can be divided into segments, with each segment running an IGP. Routing information between segments could then be exchanged using BGP.

Q—*If a directly connected interface is redistributed into BGP, what value will the origin attribute have for this route?*

A—Any redistributed route will have an origin of incomplete.

For More Information

- RFC 1771, "BGP4"

- Halabi, Bassam. *Internet Routing Architectures*. Cisco Press: Indianapolis, 1997.

- Parkhurst, William R., and David R. Jackson. *Practical BGP for Internet Routing*. Cisco Press: Indianapolis, in press.

- BGP4 Case Studies/Tutorial Section 1, http://www.cisco.com/warp/customer/459/13.html

- BGP4 Case Studies/Tutorial Section 2, http://www.cisco.com/warp/customer/459/14.html

- BGP4 Case Studies/Tutorial Section 3, http://www.cisco.com/warp/customer/459/15.html

- BGP4 Case Studies/Tutorial Section 4, http://www.cisco.com/warp/customer/459/16.html
- BGP4 Case Studies/Tutorial Section 5, http://www.cisco.com/warp/customer/459/17.html