BGP

Configuring BGP on Cisco Routers

Volume 1

Version 3.2

Student Guide

CLS Production Servies: 12.29.05

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica
Croatia • Cyprus • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece
Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia
Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania
Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland
Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Students, this letter describes important course evaluation access information!

Welcome to Cisco Systems Learning. Through the Cisco Learning Partner Program, Cisco Systems is committed to bringing you the highest-quality training in the industry. Cisco learning products are designed to advance your professional goals and give you the expertise you need to build and maintain strategic networks.

Cisco relies on customer feedback to guide business decisions; therefore, your valuable input will help shape future Cisco course curricula, products, and training offerings. We would appreciate a few minutes of your time to complete a brief Cisco online course evaluation of your instructor and the course materials in this student kit. On the final day of class, your instructor will provide you with a URL directing you to a short post-course evaluation. If there is no Internet access in the classroom, please complete the evaluation within the next 48 hours or as soon as you can access the web.

On behalf of Cisco, thank you for choosing Cisco Learning Partners for your Internet technology training.

Sincerely,

Cisco Systems Learning

Table of Contents

Volume 1

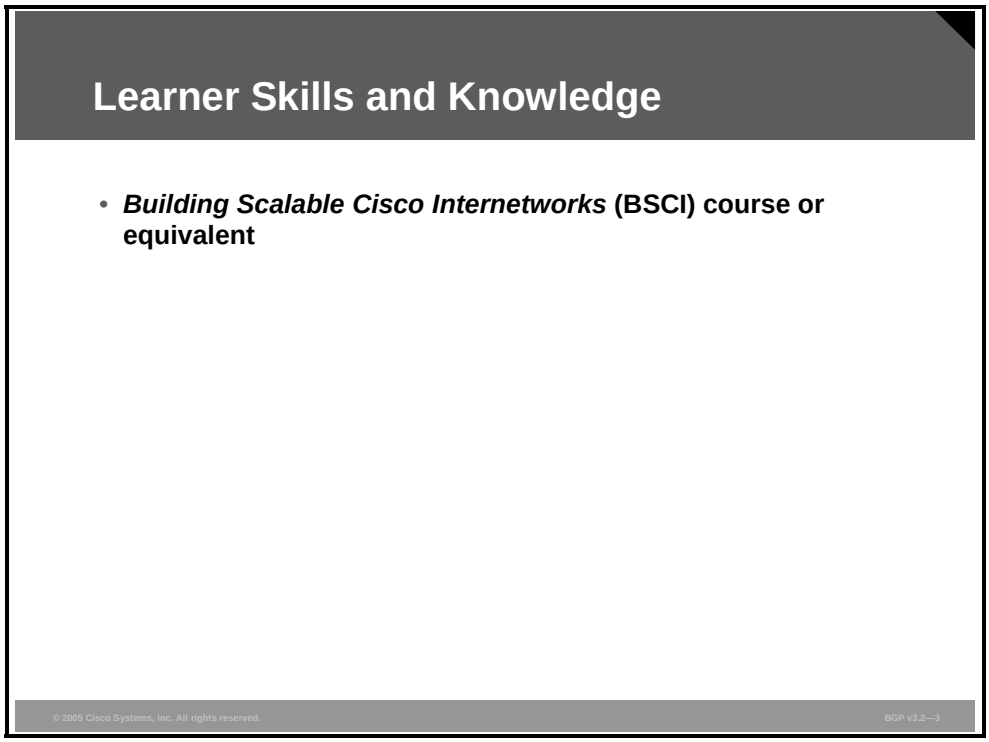
Course Introduction

Overview

Configuring BGP on Cisco Routers (BGP) v3.2 provides students with in-depth knowledge of Border Gateway Protocol (BGP), the routing protocol that is one of the underlying foundations of the Internet and New World technologies such as Multiprotocol Label Switching (MPLS). This curriculum covers the theory of BGP, configuration of BGP on Cisco IOS routers, detailed troubleshooting information, and hands-on exercises that provide learners with the skills that they need to configure and troubleshoot BGP networks in customer environments. Different service solutions in the curriculum cover BGP network design issues and usage rules for various BGP features, preparing learners to design and implement efficient, optimal, and trouble-free BGP networks.

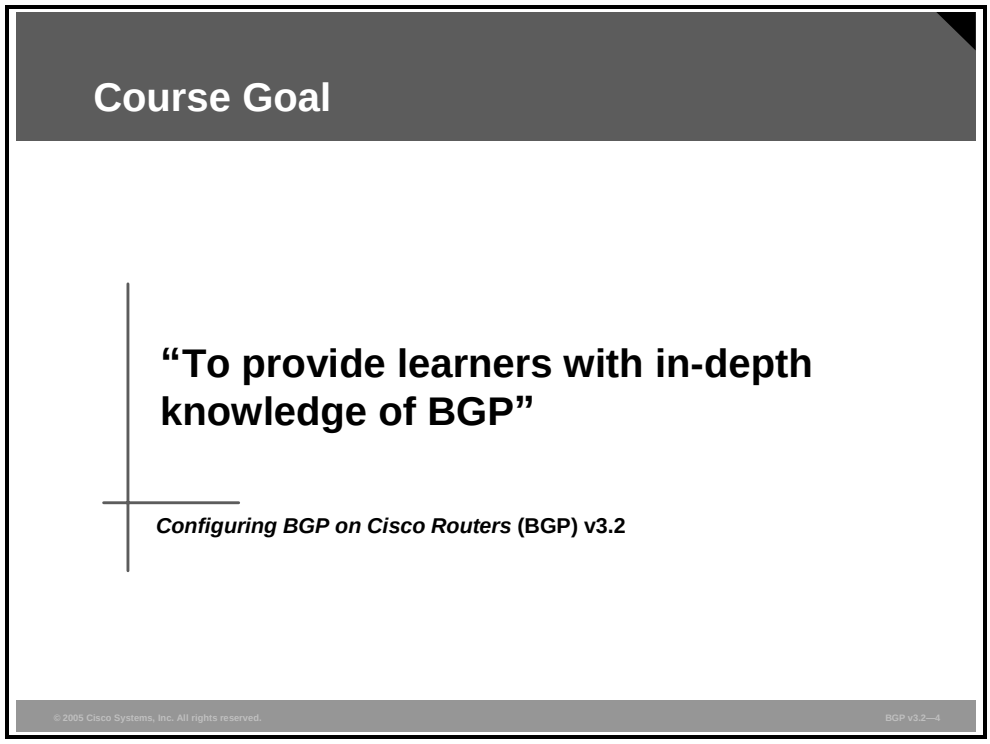
Learner Skills and Knowledge

This subtopic lists the skills and knowledge that learners must possess to benefit fully from the course. The subtopic also includes recommended Cisco learning offerings that learners should complete in order to benefit fully from this course.



Course Goal and Objectives

This topic describes the course goal and objectives.

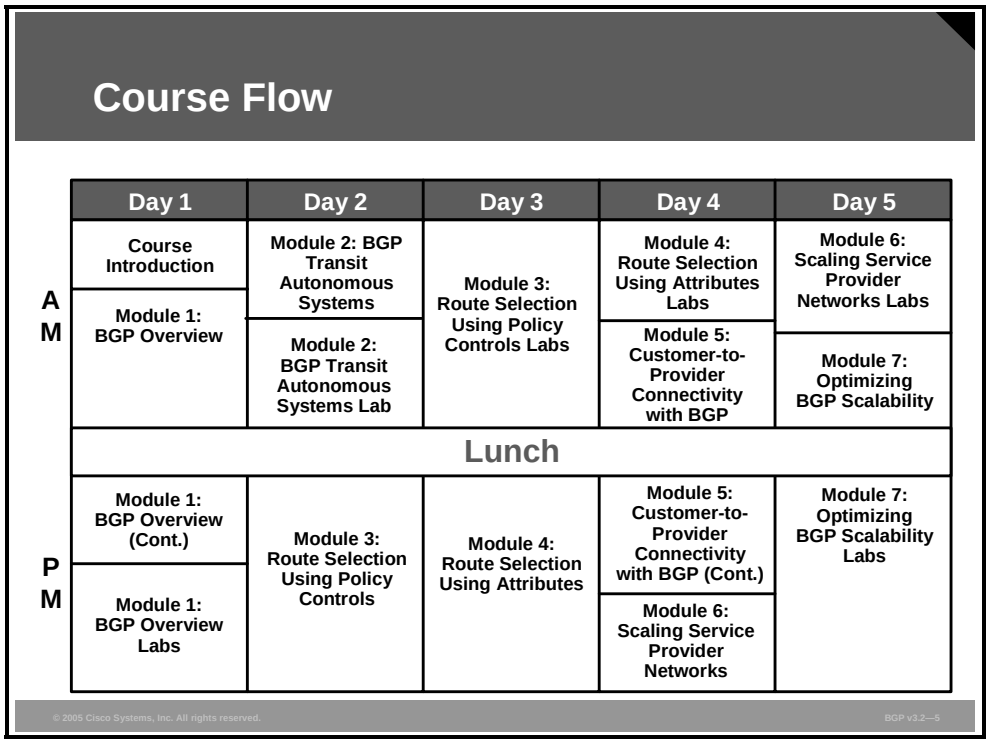


Upon completing this course, you will be able to meet these objectives:

- Configure, monitor, and troubleshoot basic BGP to enable interdomain routing in a network scenario with multiple domains
- Use BGP policy controls to influence the route selection process with minimal impact on BGP route processing in a network scenario where you must support connections to multiple ISPs
- Use BGP attributes to influence the route selection process in a network scenario where you must support multiple connections
- Implement the correct BGP configuration to successfully connect the customer network to the Internet in a network scenario where you must support multiple connections
- Enable the provider network to behave as a transit AS in a typical service provider network with multiple BGP connections to other autonomous systems
- Identify common BGP scaling issues and enable route reflection and confederations as possible solutions to these issues in a typical service provider network with multiple BGP connections to other autonomous systems
- Use available BGP tools and features to optimize the scalability of the BGP routing protocol in a typical BGP network

Course Flow

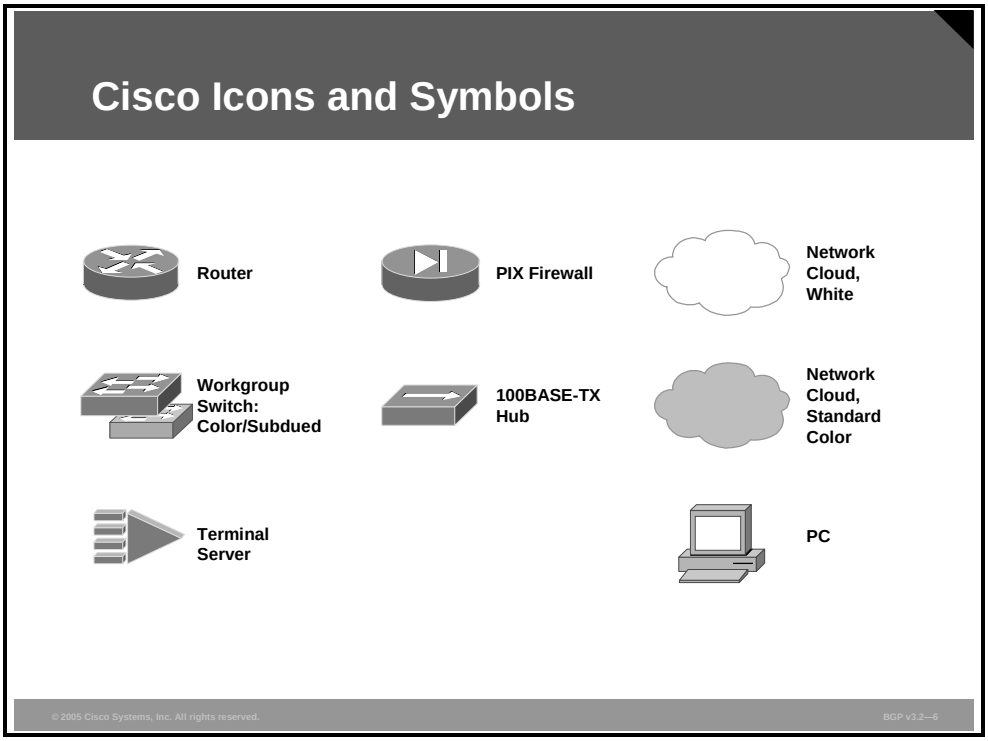
This topic presents the suggested flow of the course materials.



The schedule reflects the recommended structure for this course. This structure allows enough time for the instructor to present the course information and for you to work through the lab activities. The exact timing of the subject materials and labs depends on the pace of your specific class.

Additional References

This topic presents the Cisco icons and symbols used in this course, as well as information on where to find additional technical references.



Cisco Glossary of Terms

For additional information on Cisco terminology, refer to the *Cisco Internetworking Terms and Acronyms* glossary of terms at http://www.cisco.com/univercd/cc/td/doc/cisintwk/ita/index.htm.

Your Training Curriculum

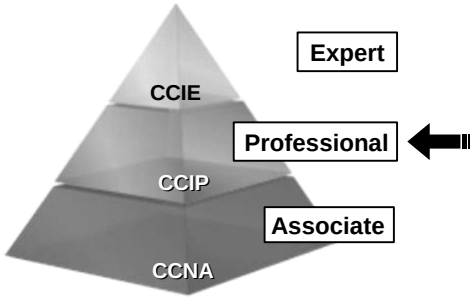
This topic presents the training curriculum for this course.

You are encouraged to join the Cisco Certification Community, a discussion forum open to anyone holding a valid Cisco Career Certification (such as Cisco CCIE®, CCNA®, CCDA®, CCNP®, CCDP®, CCIP™, or CCSP®). It provides a gathering place for Cisco-certified professionals to share questions, suggestions, and information about Cisco Career Certification programs and other certification-related topics. For more information, visit http://www.cisco.com/web/learning/le3/le2/le37/le8/learning_certification_type_home.html

Cisco Career Certifications: Service Provider Internetworking

Expand Your Professional Options and Advance Your Career

Professional-level recognition in service provider internetworking



Service Provider Internetworking

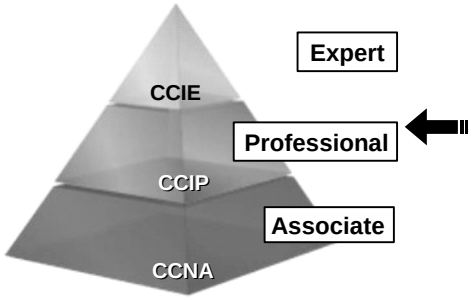
Required Exam	Recommended Training Through Cisco Learning Partners
BSCI	Building Scalable Cisco Internetworks (BSCI)
BGP	Configuring BGP on Cisco Routers (BGP)
QOS	Implementing Cisco Quality of Service (QoS)
MPLS	Implementing Cisco MPLS (MPLS)

http://www.cisco.com/go/certifications

Cisco Career Certifications: Service Provider Internetworking (Cont.)

Expand Your Professional Options and Advance Your Career

Professional-level recognition in service provider internetworking



Service Provider Internetworking

Another option:

Required Exam	Recommended Training Through Cisco Learning Partners
Composite	Building Scalable Cisco Internetworks (BSCI) and Building Cisco Multilayer Switched Networks (BCMSN)
QOS	Implementing Cisco Quality of Service (QoS)
BGP + MPLS	Configuring BGP on Cisco Routers (BGP) and Implementing Cisco MPLS (MPLS)

http://www.cisco.com/go/certifications

Module 1

BGP Overview

Overview

Border Gateway Protocol (BGP) is an interdomain (interautonomous system) routing protocol that is used to exchange routing information for the Internet. BGP, by design, is a very robust and scalable routing protocol. Because BGP is deployed as an interdomain routing protocol, it has many rich features that allow administrators to implement a variety of routing policies. This module covers basic BGP technology, details BGP session establishment and routing information exchange, and describes basic Cisco IOS configuration and troubleshooting tasks.

Module Objectives

Upon completing this module, you will be able to configure, monitor, and troubleshoot basic BGP to enable interdomain routing in a network scenario with multiple domains. This ability includes being able to meet these objectives:

- Identify appropriate BGP usage and limitations
- List BGP path attributes and the functionality of each attribute
- Describe the concept of BGP neighbors and neighbor session establishment procedures
- Describe BGP route processing
- Configure a router for BGP
- Perform the steps to correct basic BGP configuration and session errors

Lesson 1

Introducing BGP

Overview

The Border Gateway Protocol (BGP) is a very robust and scalable routing protocol, which is demonstrated by the fact that it is the routing protocol that is used on the Internet. Service providers and customer networks, such as universities and corporations, usually use an Interior Gateway Protocol (IGP) such as Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP), or Open Shortest Path First (OSPF) for the exchange of routing information within their networks. Any communication between these IGPs and the Internet or between service providers will be accomplished through BGP. This lesson introduces basic BGP characteristics and features.

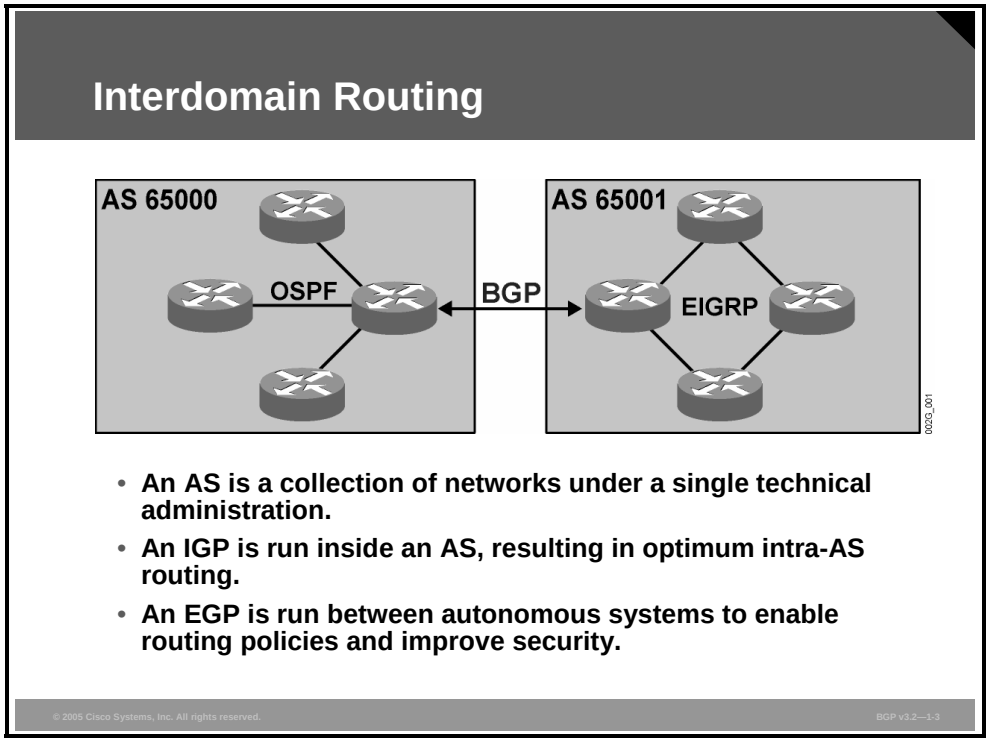
Objectives

Upon completing this lesson, you will be able to identify appropriate BGP use and limitations. This ability includes being able to meet these objectives:

- Describe interdomain routing in relation to the design goals of interdomain routing protocols
- List the basic characteristics of BGP
- Identify when a single-homed customer should use BGP as an interdomain routing protocol
- Describe when BGP is appropriate for the multihomed customer
- Describe the use of BGP in a transit autonomous backbone
- List some of the limitations of BGP

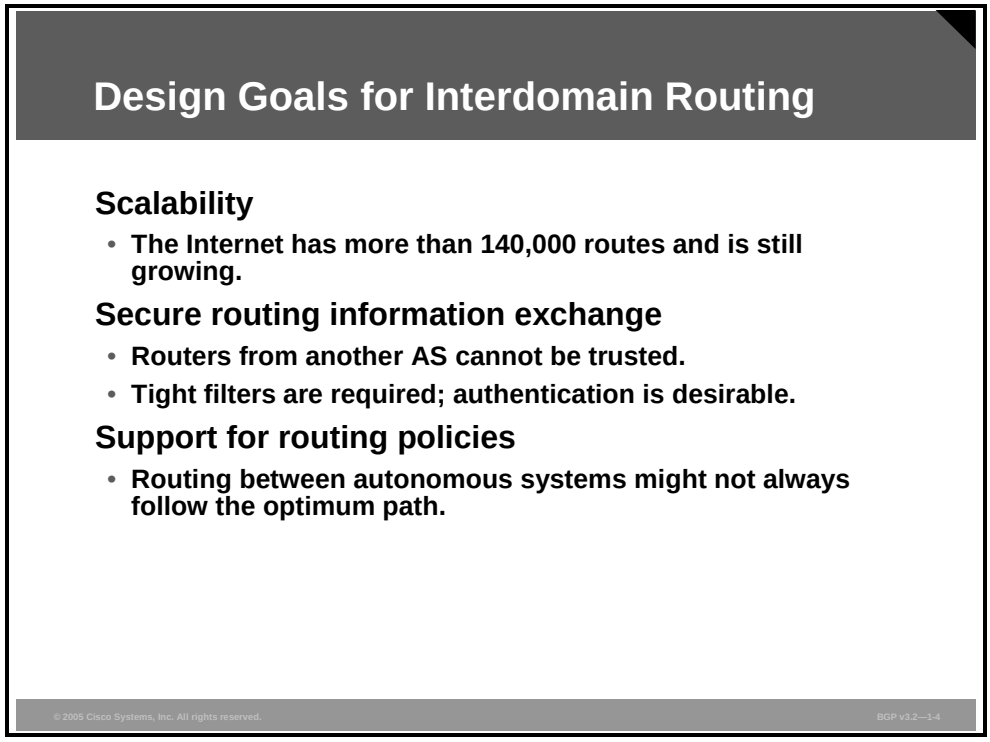
Interdomain Routing

This topic describes interdomain routing in relation to the design goals of interdomain routing protocols.



When talking to people who are involved with Internet routing, network administrators commonly use the terms “autonomous system,” “interdomain routing,” “interior routing protocol,” and “exterior routing protocol.” These terms, which may be confusing for a novice, are defined as follows:

- An autonomous system (AS) is a collection of networks under a single technical administration. Other definitions refer to a collection of routers or IP prefixes, but in the end they are all essentially the same thing. The important principle is the technical administration, which means sharing the same routing protocol and routing policy. Legal and administrative ownership of the routers does not matter with autonomous systems. Autonomous systems are identified by AS numbers. AS numbers are 16-bit, unsigned integers ranging from 1 to 65535. Public AS numbers (1 to 64511) are assigned and managed by an Internet registry. A range of private AS numbers (64512 to 65535) has been reserved for customers that need an AS number to run BGP in their private networks.
- Interdomain routing is routing between autonomous systems. It is usually based on a set of policies, not just the technical characteristics of the underlying infrastructure.
- Exterior routing protocols (BGP being the only exterior routing protocol that is used today) are protocols that have the right set of functions to support various interdomain routing policies. Such protocols are contrary to interior routing protocols (for example, OSPF, RIP, or EIGRP), which focus only on finding the optimum (usually fastest) route between two points, without respect to routing policies.

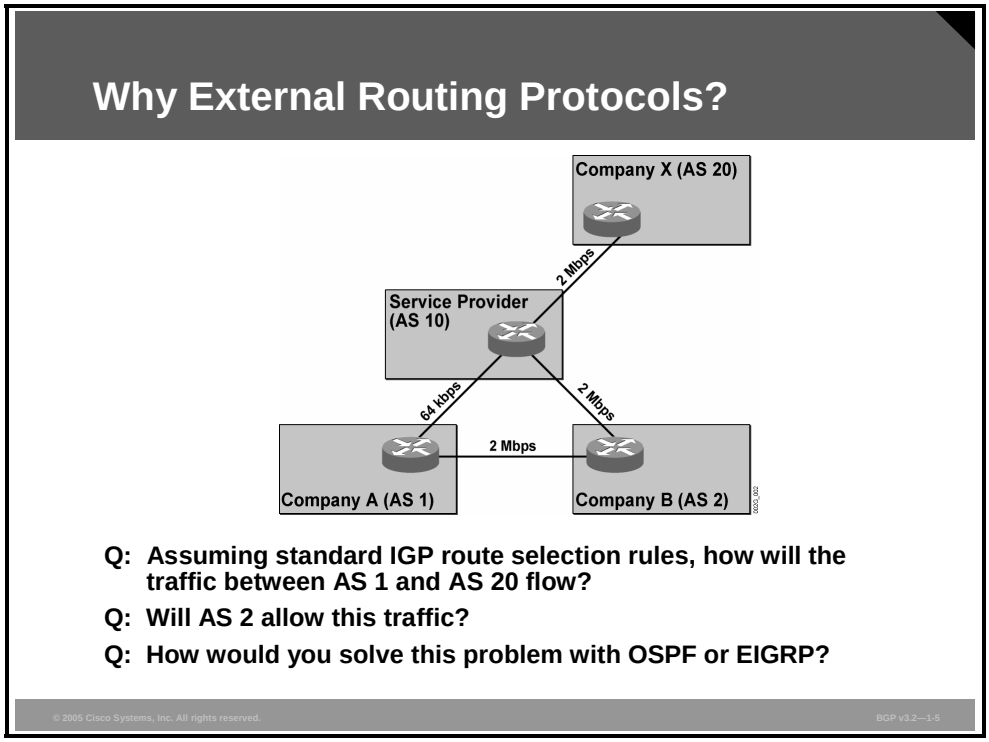


The design goals for any interdomain routing protocol include the following:

- **Scalability:** An interdomain routing protocol has to be able to support Internet routing, which consists of more than 140,000 routes.
- **Secure information exchange:** Because the routers from other autonomous systems cannot be trusted, tight filters on routing updates and router authentication are desirable features.
- **Support for routing policies:** Routing between autonomous systems might not always follow the optimum path, and exterior routing protocols have to support a wide range of customer requirements.

Example: Why External Routing Protocols?

The example illustrates the need for an interdomain routing protocol. It depicts two companies that are connected to the Internet via leased lines of differing speeds.



In routing protocols other than BGP, routing decisions are normally made to take advantage of the highest bandwidth available. Doing so would make traffic between AS 1 and AS 20 flow via AS 2. This situation is not desirable for AS 2, because it would allow the users in Company A to generate traffic on the Internet access line that was purchased and paid for by Company B.

Company B is unlikely to allow traffic from Company A to reach the Internet using the Company B access line. Company B, in fact, could create an access-list blocking all IP packets from AS 1 from being transmitted on the 2-Mbps serial line from Company B to the Internet. That action would create a black hole because Company A would send its packets to Company B and then Company B would drop them.

To avoid this situation, Company B must make sure that the packets from Company A that are destined for the Internet are never sent to Company B. Also, Company B must make sure that packets from the Internet that are destined for Company A are never sent using the Internet access line to Company B. Company B could implement a routing policy that indicates that AS 2 will receive reachability information from AS 1 for its own use but that AS 2 will not forward that particular information to the Internet. Also, AS 2 will receive reachability information about the Internet from its Internet service provider (ISP) but will never forward that information to AS 1. Only networks local to AS 2 will be sent to AS 1.

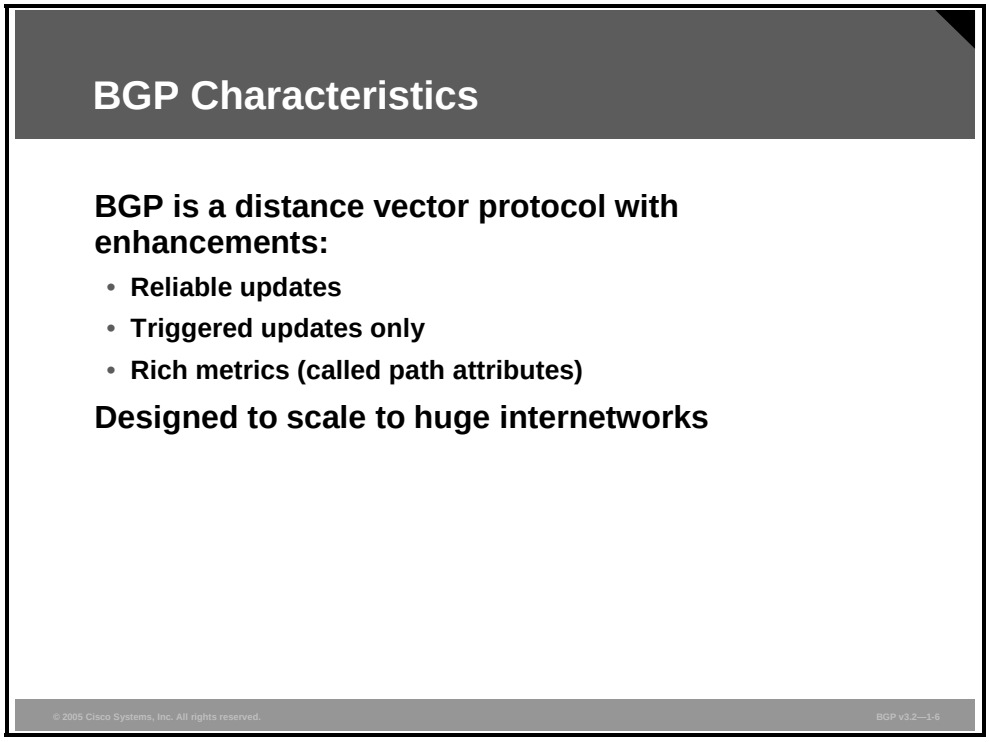
The result of this routing policy would be that AS 1 sees all networks within AS 2 as reachable over the 2-Mbps link that directly connects AS 1 with AS 2. The routers in AS 1 will not see the rest of the Internet as reachable through AS 2. Therefore, AS 1 forwards packets toward the Internet directly over the 64-kbps link.

Also, the IP networks in AS 1 will appear reachable by AS 2 over the 2-Mbps link, which directly connects AS 1 with AS 2. However, the ISP will not receive that reachability information from AS 2; it will receive it only from AS 1. Therefore, traffic from the Internet to Company A will be transmitted over the 64-kbps link.

This routing policy is easy to implement when network administrators are using BGP but impossible to implement with any other routing protocol. EIGRP, for example, can do route filtering only on individual IP subnets, not on all prefixes belonging to an AS. Link-state protocols, such as OSPF, cannot do powerful route filtering at all. BGP can do this routing based on AS numbers, which makes it possible to scale BGP over the Internet.

BGP Characteristics

This topic lists the basic characteristics of BGP.



BGP is a distance vector protocol. This means that BGP will announce to its neighbors those IP networks that it can reach itself. The receivers of that information will say, “If that AS can reach those networks, then I can reach them via the AS.”

If two different paths are available to reach the same IP subnet, then the shortest path is used. This determination requires a mechanism capable of measuring the distance. All distance vector protocols have such mechanisms, called “metrics.” BGP contains a very sophisticated method of computing the shortest path by using attributes that are attached to the reachable IP subnet.

BGP sends routing updates to its neighbors by using a reliable transport. This technique means that the sender of the information always knows that the receiver has actually received the information. As a result, there is no need for periodic updates or routing information refreshes. In BGP, only information that has changed is transmitted.

The reliable information exchange, combined with the batching of routing updates that is also performed by BGP, allows BGP to scale to large, Internet-sized networks.

BGP Characteristics (Cont.)

Reliable updates

- **TCP used as transport protocol**
- **No periodic updates**
- **Periodic keepalives to verify TCP connectivity**
- **Triggered updates batched and rate-limited**
  - **Every 5 seconds for internal peer**
  - **Every 30 seconds for external peer**

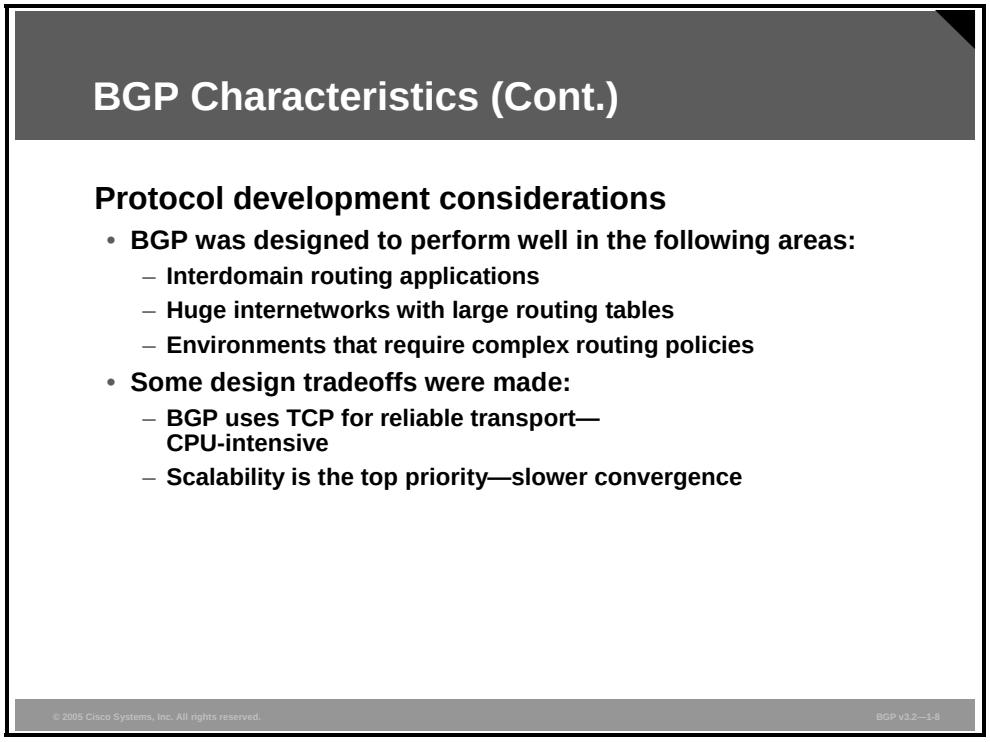
The reliable transport mechanism that is used by BGP is standard TCP. BGP is an application protocol that uses both the TCP and IP protocols for reliable connections.

Because BGP uses a reliable transport, the sender knows that the receiver has actually received the transmitted information. This capability makes periodic updates unnecessary.

A router that has received reachability information from a BGP peer must be sure that the peer router is still there. Otherwise, the router could route traffic toward a next-hop router that is no longer available, causing the IP packets to be lost in a black hole. TCP does not provide the service to signal that the TCP peer has been lost, unless some application data is actually transmitted between the peers. In an idle state, where there is no need for BGP to update its peer, the peer could be unreachable without TCP detecting it. Therefore, BGP takes care of detecting the presence of neighbors by periodically sending small BGP keepalive packets to them. These packets are considered application data by TCP and therefore must be transmitted reliably. According to the BGP specification, the peer router also must reply with a BGP keepalive packet.

When BGP was created, a key design goal was to be able to handle enormous amounts of routing information in very large and complex networks. In this environment, many links could go up and down (flapping), causing topology changes, which must be considered by the routing protocol. But low convergence time and quick responses to topology changes require fast updates and high CPU power to process both incoming and outgoing updates. The larger the network, the more updates per second can be expected if immediate response is required. The presence of too many updates in large networks can jeopardize network scalability.

The designers of BGP decided that scalability was a more important issue than low convergence time, so BGP was designed to batch updates. Any changes that are received within the batch interval time are saved. At the end of the interval, only the remaining result is forwarded in an outgoing update. If a network flaps several times during the batch interval, only the state at the end of the interval is sent in an update. The batching feature avoids an uncontrolled flood of updates all over the Internet because the number of updates is limited by the batching procedure.

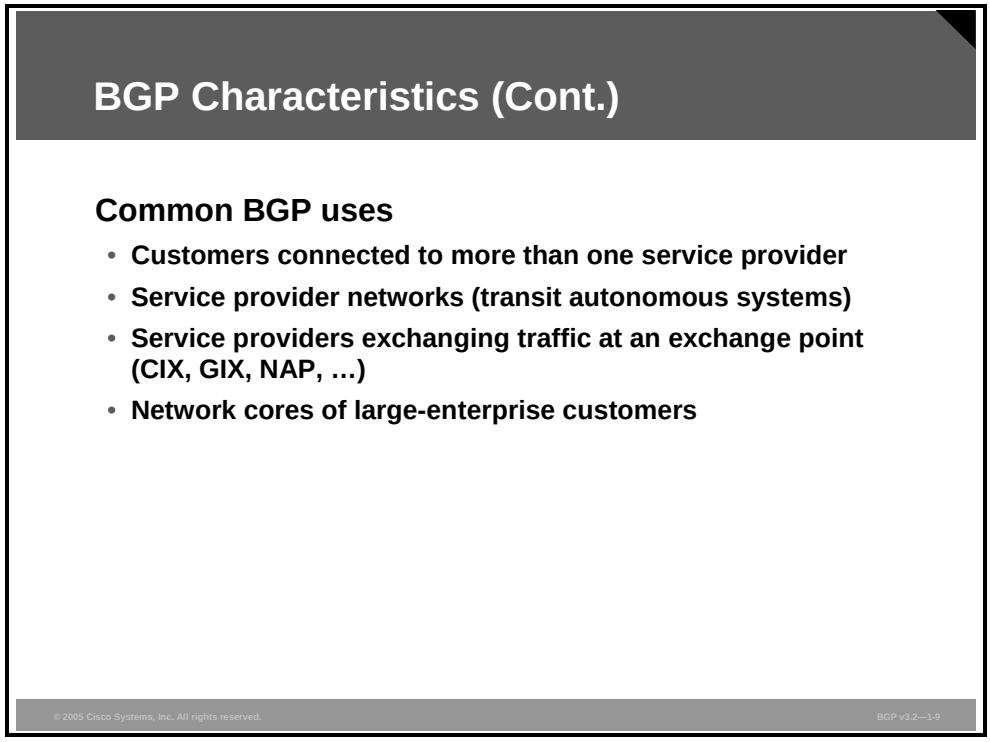


The designers of the BGP protocol have succeeded in creating a highly scalable routing protocol, which can forward reachability information between autonomous systems (also known as routing domains). The designers had to consider an environment with an enormous number of reachable networks and complex routing policies that were driven by commercial rather than technical considerations.

TCP, a well-known and widely proven protocol, was chosen as the transport mechanism. That decision kept BGP simple, but it increased the CPU resource requirements for routers running BGP. The point-to-point nature of TCP also introduces a slight increase in network traffic, because any update that should be sent to many receivers has to be multiplied into several copies, which are then transmitted on individual TCP sessions to the receivers.

Whenever there was a design choice between fast convergence and scalability, scalability was the top priority. The batching of updates and the relatively low frequency of keepalive packets are examples of designers placing convergence time second to scalability.

Note	BGP convergence times can be modified with the configuration of nondefault values for BGP scan and advertisement timers. Refer to the “Optimizing BGP Scalability” module for more information on tuning BGP convergence.

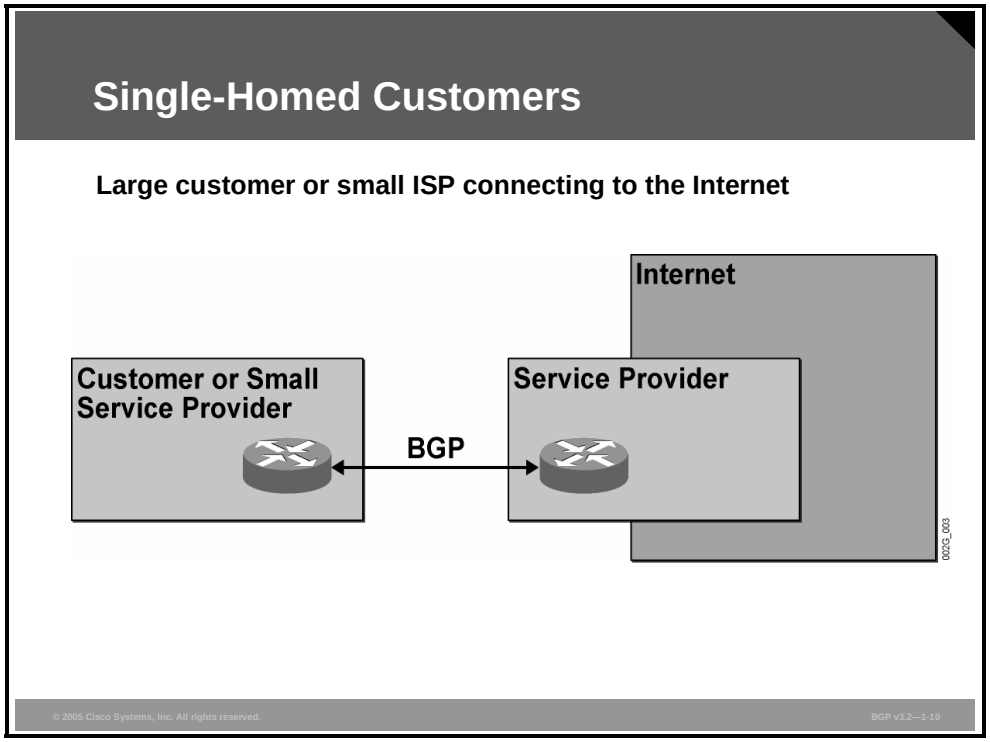


The figure shows typical scenarios in which BGP is usable. These scenarios include the following:

- Customers connected to more than one service provider.
- ISP networks themselves acting as transit systems and forwarding external traffic.
- Exchange points, which can be defined by the network access point (NAP) between region and core. International exchange points can be defined by either Commercial Internet eXchange (CIX) or Global Internet eXchange (GIX) points.
- Very large enterprises using BGP as their core routing protocol.

Single-Homed Customers

This topic identifies when a single-homed customer should use BGP as an interdomain routing protocol.



The figure shows a customer network connected to the Internet using a single ISP, but such a scenario is generally not the case when BGP is used. Normal Internet access to a single ISP does not require BGP; static routes are more commonly used to handle this situation. Small ISPs buying Internet connectivity from other ISPs use this type of connectivity more often, especially if they want to start their business the proper way—by using their own AS number and having their own address space.

Use Guidelines? Single-Homed Customers

Use BGP between the customer and the service provider in these situations:

- **Customers multihomed to the same service provider**
- **Customers that need dynamic routing protocol with the service provider to detect failures**
  - **Hint: Use private AS number for these customers.**
- **Smaller ISPs that need to originate their routes in the Internet**

Use static routes in all other cases:

- **Static routes always simpler than BGP**

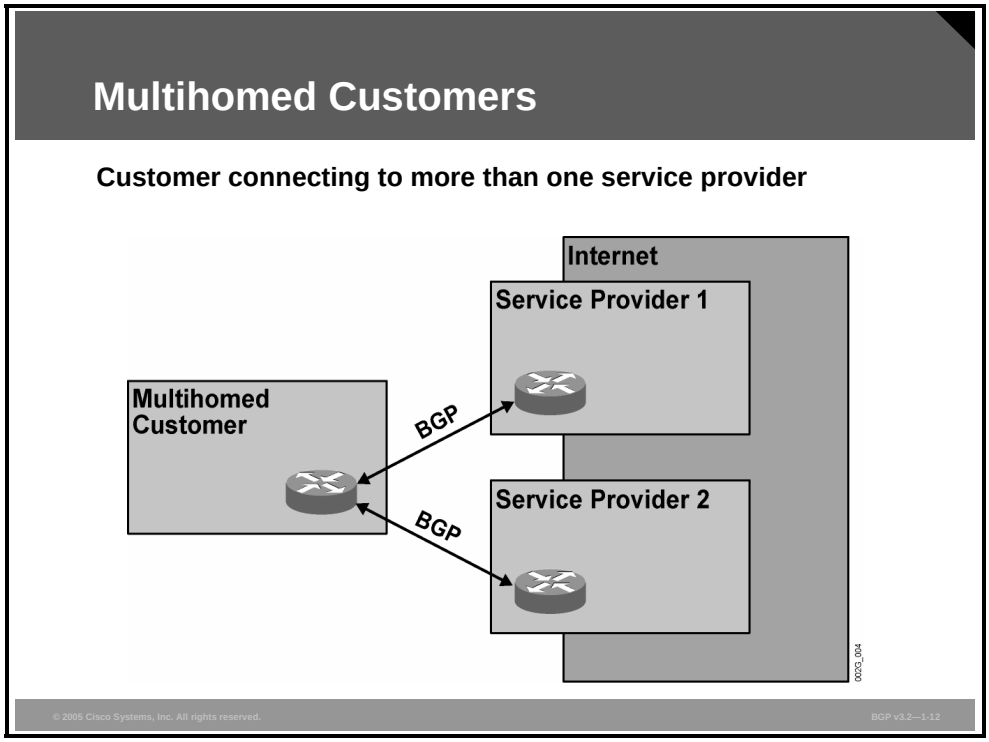
Under certain conditions, BGP must be configured between the customer and the service provider. For example, BGP is needed when customers are multihomed to the same service provider (that is, the customer networks have multiple links connecting them with the service provider network) and require dynamic routing protocol interaction with the service provider to detect link failures. Private AS numbers (AS numbers above 64512) are usually implemented in BGP configurations for these customers.

Customers that plan to connect to more than one ISP, and small ISPs that plan to have multiple Internet connections in the future, usually use BGP with their service provider. They use this option even when they have a single link with the service provider in order to be prepared for future upgrades.

In all other cases, using static routes from the service provider toward the customer and using a default static route from the customer toward the service provider is the preferred method of provider-to-customer routing in the Internet.

Multihomed Customers

This topic describes when BGP is appropriate for the multihomed customer. BGP use guidelines for multihoming are also discussed.

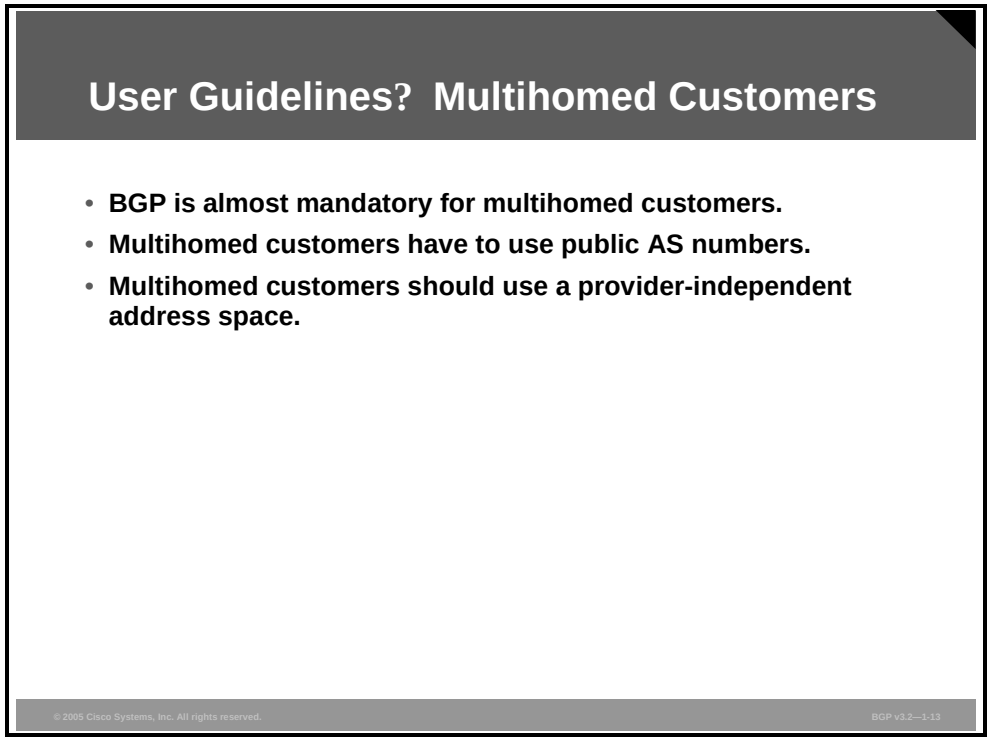


The figure illustrates a customer network that is connected to two different ISPs, requiring the use of BGP for full redundancy.

The customer must have its own officially assigned AS number. The customer is also responsible for announcing its own IP networks to both ISPs. Both ISPs forward all routes received from the Internet to the customer network. The customer should avoid forwarding any routing information received from one ISP to the other. Otherwise, the customer becomes a transit provider between the two ISPs. This is a situation that most customers like to avoid because it creates a resource drain on routers and network links.

Full redundancy is achieved in this setup. If either of the two access links fails, the reachability information that was previously transmitted on the now-failed link is withdrawn. But BGP reachability information is still announced by the customer router over the remaining link. Thus, the ISP still sees all networks within the customer AS as reachable but only over the remaining path. Also, received routes from the Internet are withdrawn when the link fails, but routes received over the remaining link are not affected. Thus, the Internet, including the ISP to which the direct connection has failed, is still reachable over the remaining link.

This design can also handle other problems. A case where both access links are available, but the connection between one of the ISPs and the rest of the Internet is lost, works as follows: The ISP that has a problem reaching the rest of the Internet withdraws all those routes and tells the customer AS that it can no longer reach the Internet. But the networks local to the ISP with the Internet reachability problem are still reachable by the customer, so those routes are not withdrawn. The networks in the customer AS are still reachable by the ISP in trouble, but that ISP can no longer forward the announcement to the rest of the Internet. The rest of the Internet will, however, see the customer networks as reachable over the path to the other ISP, which is fully functional.

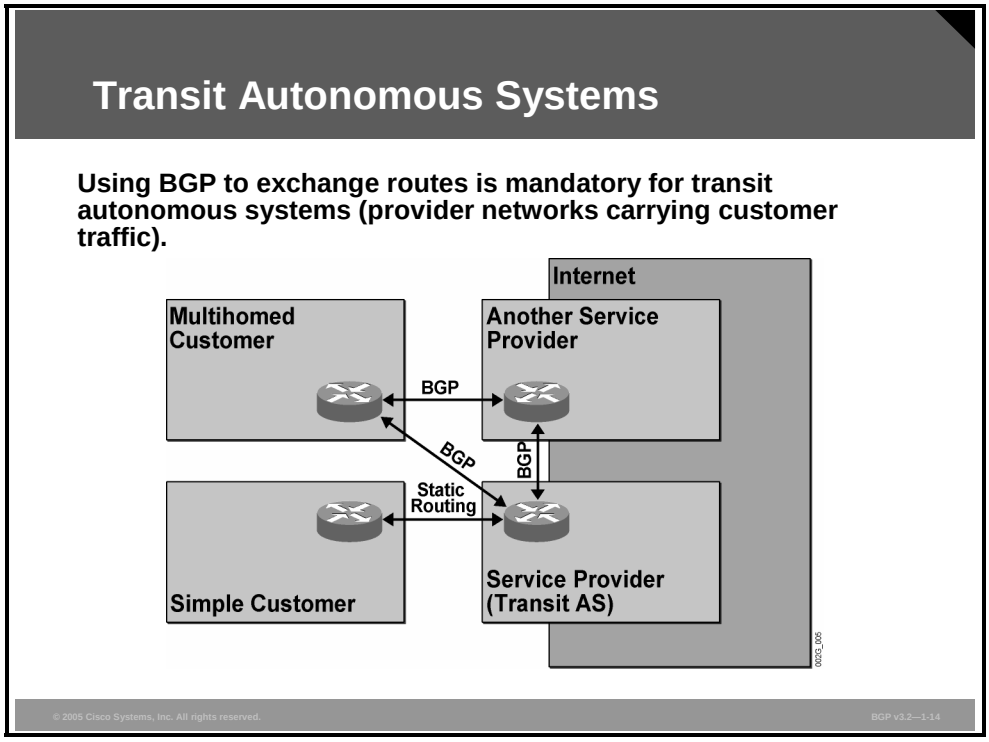


The following use guidelines apply to multihomed customers:

- Although there are designs where BGP could be avoided, most multihomed customers need to use BGP with their service providers.
- Multihomed customers must have their own AS numbers, and it is recommended to use a public AS number.
- Multihomed customers should use a provider-independent address space, which is allocated to them directly by an Internet registry.

Transit Autonomous Systems

This topic describes the use of BGP in a transit autonomous backbone.



BGP is most commonly implemented in service provider networks to ensure connectivity between customers and the rest of the Internet. An ISP might exchange BGP updates with the customers or use static routing toward them. That ISP also connects to other ISPs and is required to forward the routes that are received from customers to the rest of the Internet, as well as in the other direction. As a result, user data traffic starts to flow between the customers and the rest of the Internet. Such a network, providing transit services to traffic that is originated in other networks, is called a “transit autonomous system,” or “transit AS.” A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information received from one AS to another AS.

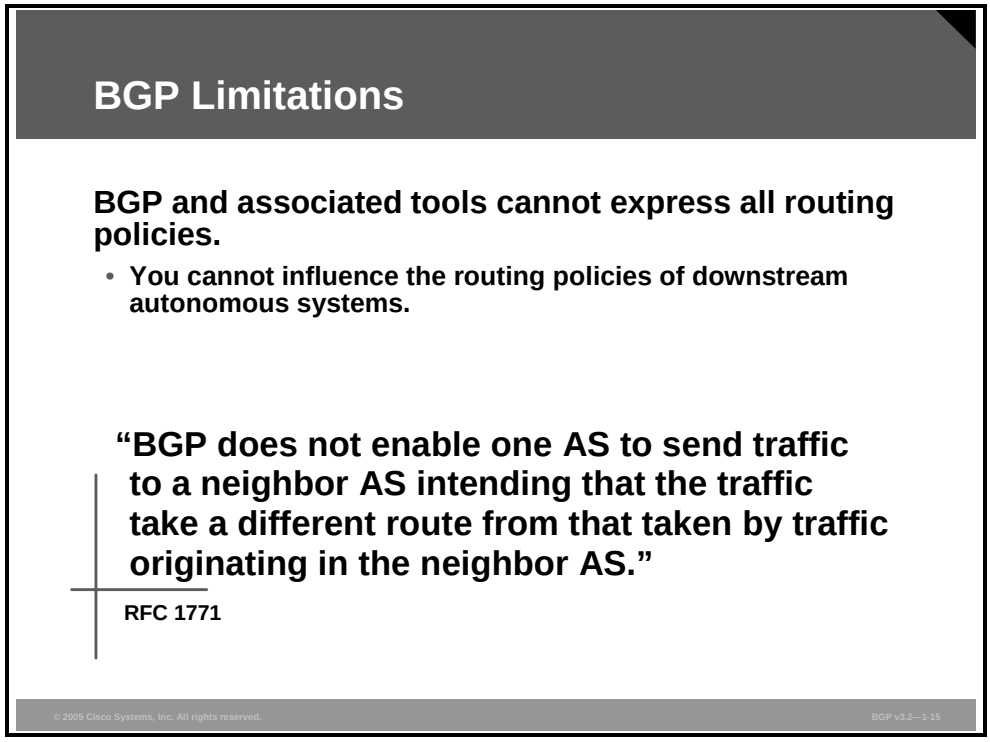
When routing information is forwarded, the receiver will see an available path to a destination and start transmitting user data toward the destination using that path. The transit AS must be prepared to relay the user data, as explained later in this course.

ISP networks can sometimes have dedicated peer-to-peer connections, using, for example, packet over SONET (POS). These connections are sometimes called private peering. ISPs also interconnect at exchange points. Technically, an exchange point is just a multiaccess subnet: a LAN (for example, a Gigabit Ethernet or Fast Ethernet switch), a Dynamic Packet Transport (DPT) ring, or an ATM switch. Many ISPs can connect to an exchange point and establish BGP sessions.

The benefit of an exchange point is that it is highly scalable. There is no need for additional physical interfaces in the ISP border router when a new ISP is launched. If the already established ISPs want to, they can open a BGP session with the new ISP. When this session is opened, they start to exchange routing information and then user data traffic over the exchange point.

BGP Limitations

This topic lists some of the limitations of BGP.



BGP-enabled routers make forwarding decisions based on the destination IP address only; the source IP address does not affect the decision. If an AS acts as a transit AS for other autonomous systems, the IP packets that are created and transmitted from the other autonomous systems are not treated differently from the IP packets that are created and transmitted from the local AS. If the local AS has decided that the best path to reach a certain destination is via a specific next-hop router, then it will route all user data traffic toward the final destination via that specific next-hop router. The local AS makes its decision based on destination address only, regardless of which IP host has sourced the IP packets.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **BGP has the right set of functions to support the various interdomain routing policies. Contrary to BGP, interior routing protocols focus only on finding the optimum (usually fastest) route between two points, without respect to routing policies.**
- **BGP is an enhanced distance vector protocol with reliable transport provided by TCP, a rich set of metrics called BGP path attributes, and scalability features such as batched updates that make it suitable for very large networks.**
- **Customers that plan to connect to more than one ISP, and small ISPs that plan to have multiple Internet connections in the future, usually use BGP with their service provider.**

Summary (Cont.)

- **Although there are designs where BGP could be avoided, most multihomed customers use BGP with their service providers.**
- **A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information received from one AS to another AS.**
- **BGP is bound by IP hop-by-hop, destination-only routing. Routing policies that deviate from this model cannot be implemented with BGP.**

Lesson 2

Understanding BGP Path Attributes

Overview

To aid routers in calculating the best route to select when multiple paths to a particular destination exist, routes that are learned via Border Gateway Protocol (BGP) have properties that are associated with them. These properties are referred to as BGP path attributes. An understanding of how BGP path attributes influence route selection is required to design robust BGP networks.

This lesson introduces BGP path attributes and their purpose. The lesson also discusses classifications that are used to describe attributes and the properties of each classification. The functionality of the autonomous system (AS) path and next-hop attributes are also explained in detail in this lesson.

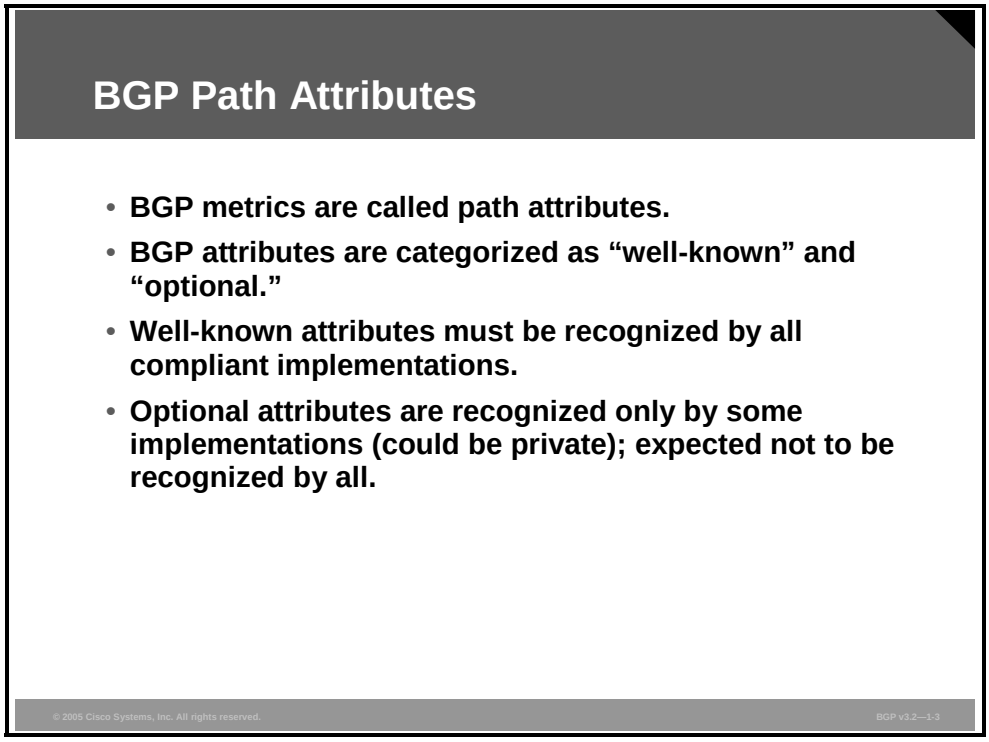
Objectives

Upon completing this lesson, you will be able to list BGP path attributes and the functionality of each attribute. This ability includes being able to meet these objectives:

- Describe the concept of BGP path attributes
- Explain the difference between mandatory and discretionary well-known BGP attributes
- Explain the difference between nontransitive and transitive optional BGP attributes
- Describe the functionality of the AS-path attribute
- Describe the functionality of the next-hop attribute

BGP Path Attributes

This topic describes the concept of BGP path attributes.

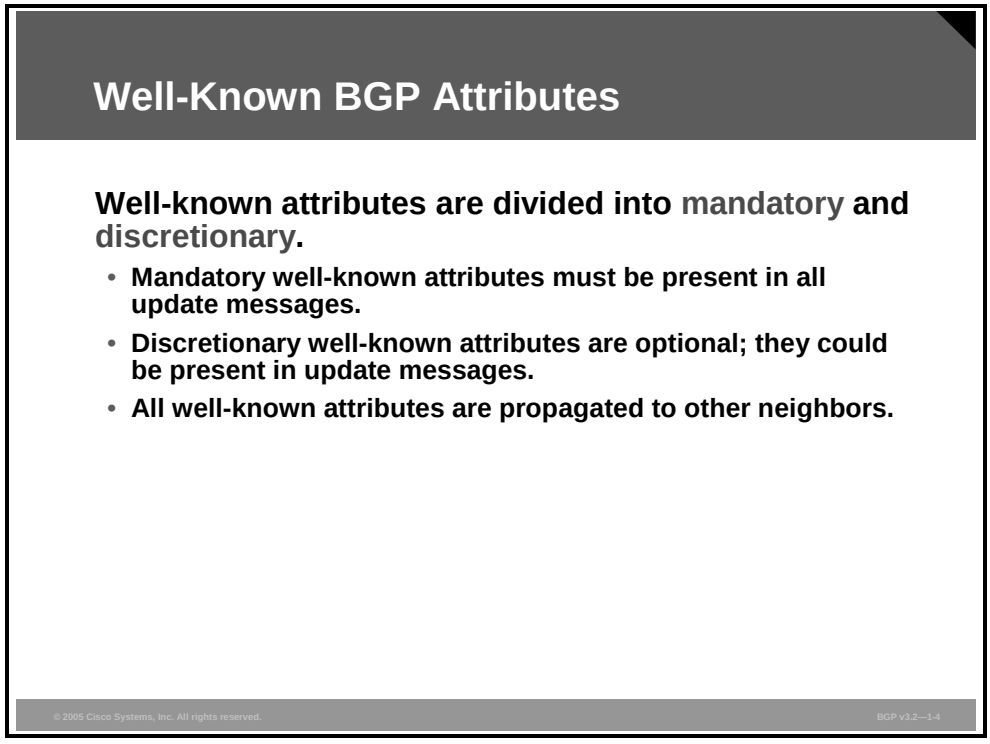


Each BGP update consists of one or more IP subnets and a set of attributes that are attached to them. Some of the attributes are required to be recognized by all BGP implementations. Those attributes are called “well-known BGP attributes.”

Attributes that are not well-known are called “optional.” These could be attributes that are specified in a later extension of BGP or even in private vendor extensions that are not documented in a standard document.

Well-Known BGP Attributes

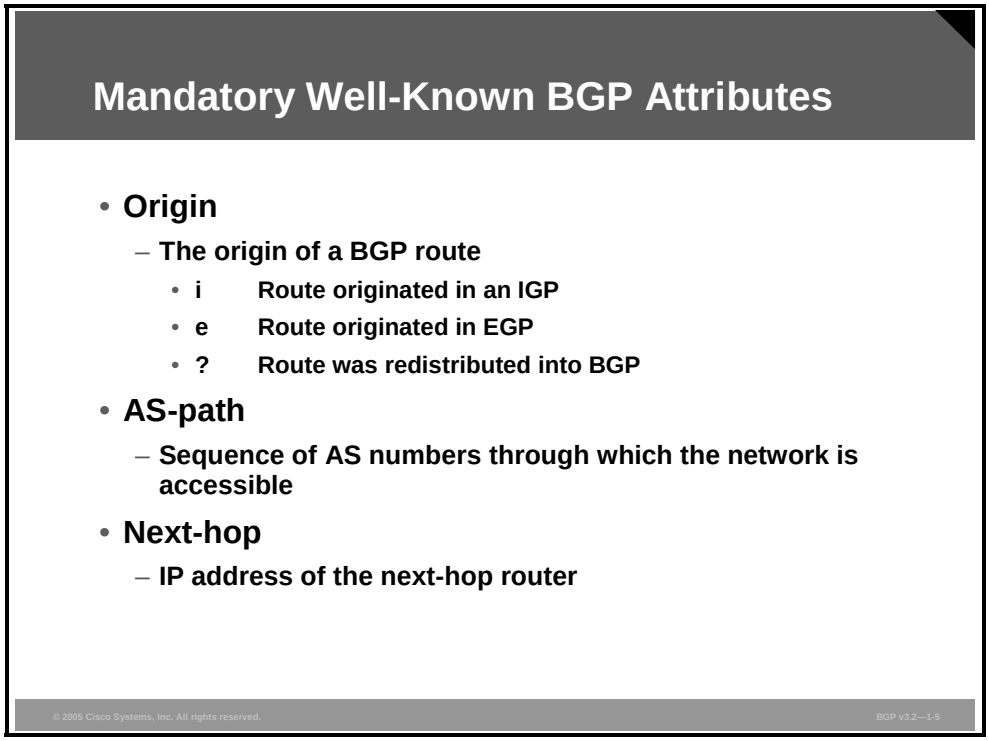
This topic describes the differences between mandatory and discretionary well-known BGP attributes.



There is a small set of three specific well-known attributes that are required to be present on every update. These are the next-hop, AS-path, and origin attributes and are referred to as “mandatory well-known attributes.”

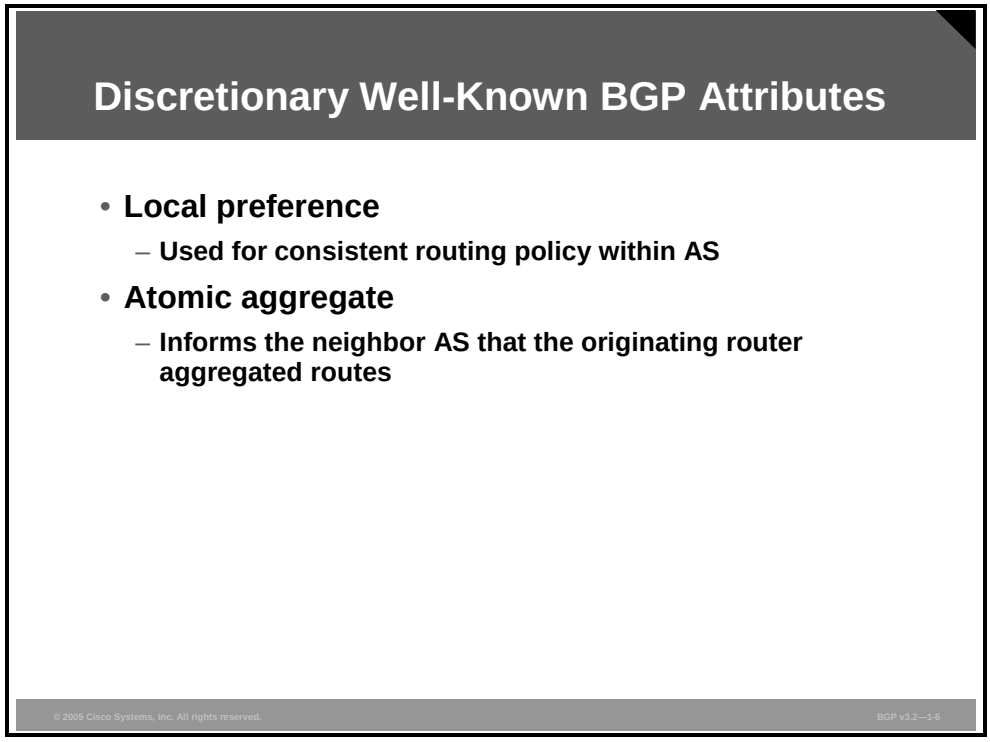
Other well-known attributes may or may not be present, depending on the circumstances under which the updates are sent and the desired routing policy. The well-known attributes that could be present, but are not required, are called “discretionary well-known attributes.”

When a router receives a BGP update, it analyzes the attached attributes and compares them with the attributes that were attached to the same IP subnet when it was received from a different source. The router then makes a decision about which source indicates the best path to the particular IP subnet. The best route is propagated, along with its well-known attributes, to other BGP-speaking neighbors.



The three mandatory well-known attributes are origin, AS-path, and next-hop:

- **Origin:** When a router first originates a route in BGP, it sets the origin attribute. If information about an IP subnet is injected using the **network** command or via aggregation (route summarization within BGP), the origin attribute is set to "i" for Interior Gateway Protocol (IGP). If information about an IP subnet is injected using redistribution, the origin attribute is set to "?" for unknown or incomplete information (these two words have the same meaning). The origin code "e" was used when the Internet was migrating from exterior gateway protocol (EGP) to BGP and is now obsolete.
- **AS-path:** The egress router modifies the AS-path attribute every time information about a particular IP subnet passes over an AS border. When a router first originates a route in BGP, the AS-path attribute is empty. Each time that the route crosses an AS boundary, the transmitting AS prepends its own AS number to appear first in the AS path. You can track the sequence of autonomous systems through which the route has passed by using the AS-path attribute.
- **Next-hop:** The router also modifies the next-hop attribute as the route passes through the network. This attribute indicates the IP address of the next-hop router—the router to which the receiving router should forward the IP packets toward to reach the destination that is advertised in the routing update.

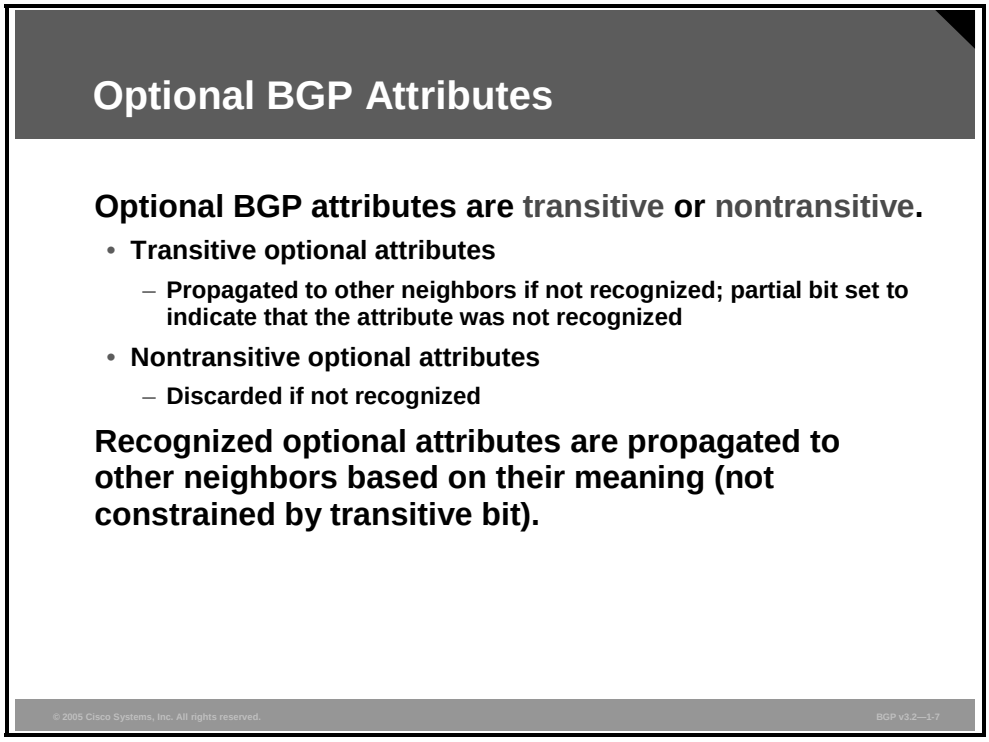


Discretionary well-known attributes must be supported by all BGP implementations but do not have to be present in all BGP updates. Routers use discretionary well-known attributes only when those functions are required. The following are descriptions of these two attributes:

- **Local preference:** Local preference is used in the route selection process. This attribute is carried within an AS only. The router prefers a route with a high local preference value to a route with a low value. By default, routes that are received from a peer AS are tagged with the local preference set to a value of 100 before they are entered into the local AS. If this value is changed through BGP configuration, the BGP selection process is influenced. Because all routers within the AS get the attribute along with the route, a consistent routing decision is made throughout the AS.
- **Atomic aggregate:** The atomic aggregate attribute is attached to a route that is created as a result of route summarization (called “aggregation” in BGP). This attribute signals that information that was present in the original routing updates may have been lost when the updates were summarized into a single entry.

Optional BGP Attributes

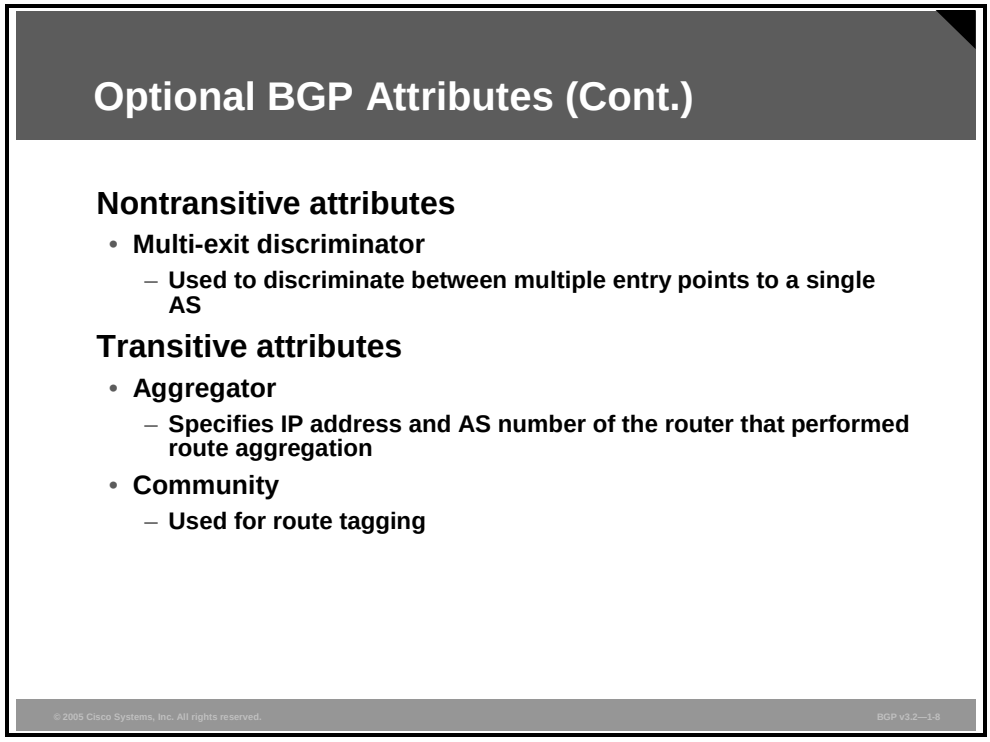
This topic explains the difference between nontransitive and transitive optional BGP attributes.



When a router receives an update that contains an optional attribute, the router checks to see whether its implementation recognizes the particular attribute. If it does, then the router should know how to handle it and whether to propagate it.

If the router does not recognize the attribute, the BGP implementation should look for the transitive bit in the attribute code. Some attributes, although not recognized by the router, might still be helpful to upstream routers and should be propagated. These attributes (called “transitive optional attributes”) are propagated even when they are not recognized. If a router propagates an unknown transitive optional attribute, it sets an additional bit in the attribute header, called the “partial bit,” to indicate that at least one of the routers in the path did not recognize the meaning of a transitive optional attribute.

Other attributes, called “nontransitive optional attributes,” might be of no value to upstream routers if a router earlier in the path does not recognize them. Routers that do not recognize these attributes drop them.



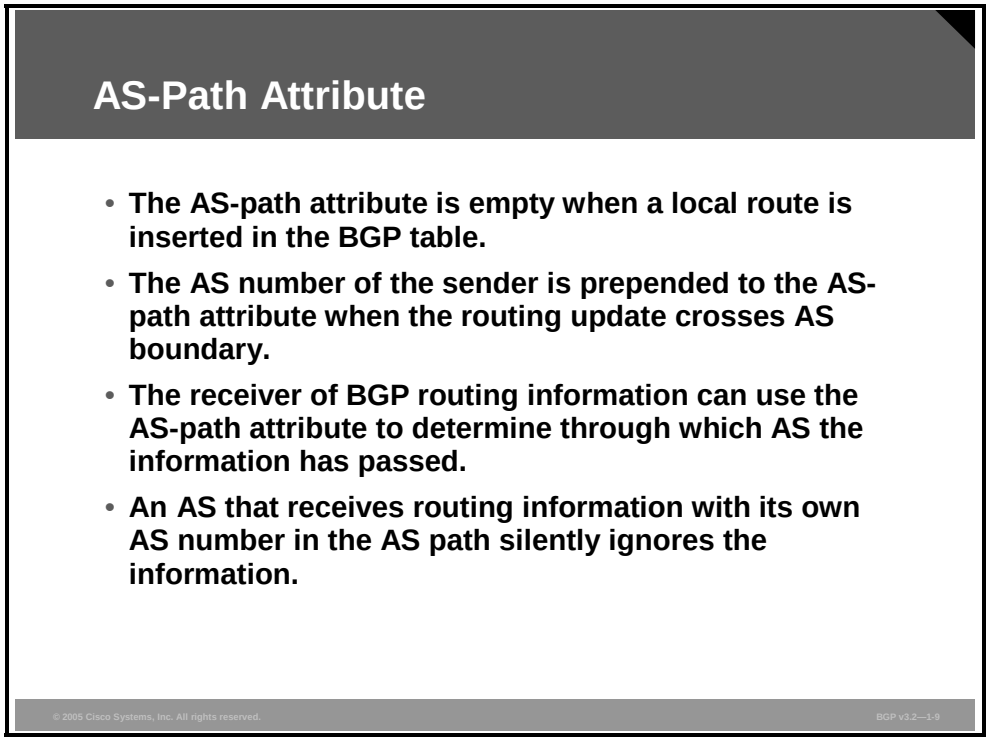
One of the nontransitive optional attributes is the multi-exit discriminator (MED) attribute, which also influences the BGP route selection process. Whenever there are several links between two adjacent autonomous systems, one AS can use the MED attribute to tell another AS to prefer one of the links for specific destinations.

Transitive optional attributes include the following:

- **Aggregator:** Identifies the AS and the router within that AS that created a route summarization, or aggregate.
- **Community:** A numerical value that can be attached to certain routes as they pass a specific point in the network. The community value can later be examined by other routers at different points in the network for filtering or route selection purposes. BGP configuration may cause routes with a specific community value to be treated differently than others.

AS-Path Attribute

This topic describes the functionality of the BGP AS-path attribute.



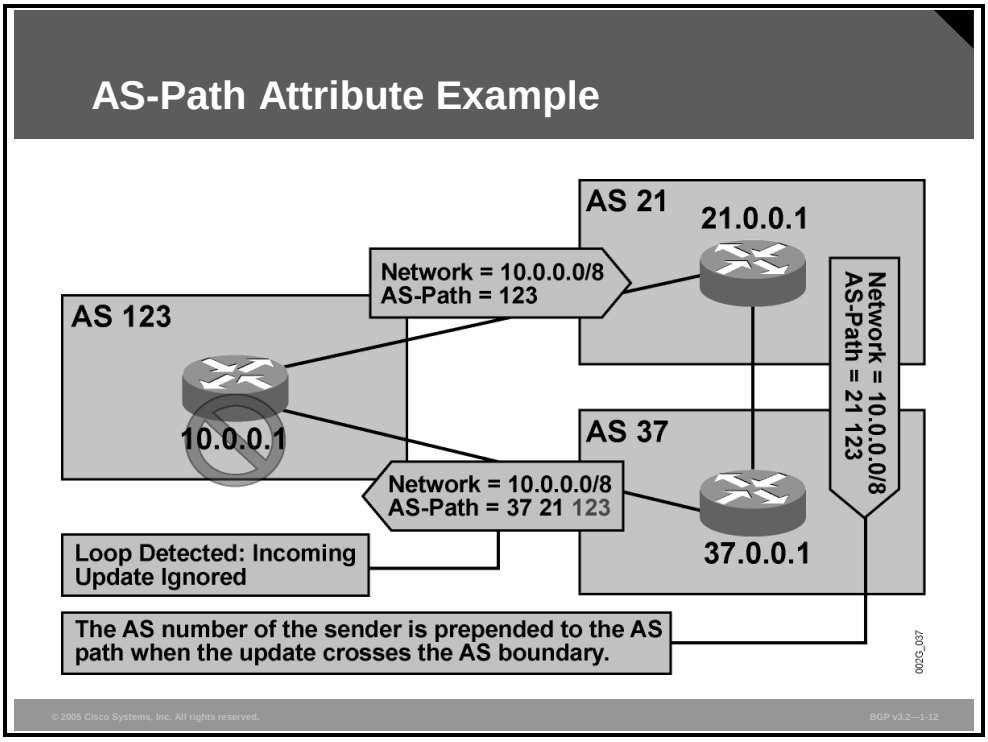
The AS-path attribute is modified by an edge router every time information about a particular IP subnet passes over an AS border. When a router first originates a route in BGP, the AS-path attribute is empty. The local AS number is prepended to the AS path each time that the route crosses an AS boundary. There are several consequences of this behavior:

- When you examine BGP routes, the AS path can be interpreted as the sequence of autonomous systems that must be passed through to reach the indicated network. The AS that originally injected the route into BGP is always found at the rightmost end of the AS path.
- It is easy to distinguish local routes from routes that have been received from other autonomous systems—BGP routes with an empty AS path were injected into BGP from within the local AS.

The AS-path attribute is also used to avoid routing loops. When a router receives a BGP update, it checks the AS-path attribute and looks for its own AS number. If that number is found in the AS path, then the route has already crossed the local AS and the router is now faced with a routing information loop. To avoid this situation, the route is silently ignored.

Example: AS-Path Attribute

The figure shows how BGP loop prevention works.



The network 10.0.0.0/8 is local to AS 123. The router in AS 123 injects the route 10.0.0.0/8 into BGP with an empty AS-path attribute.

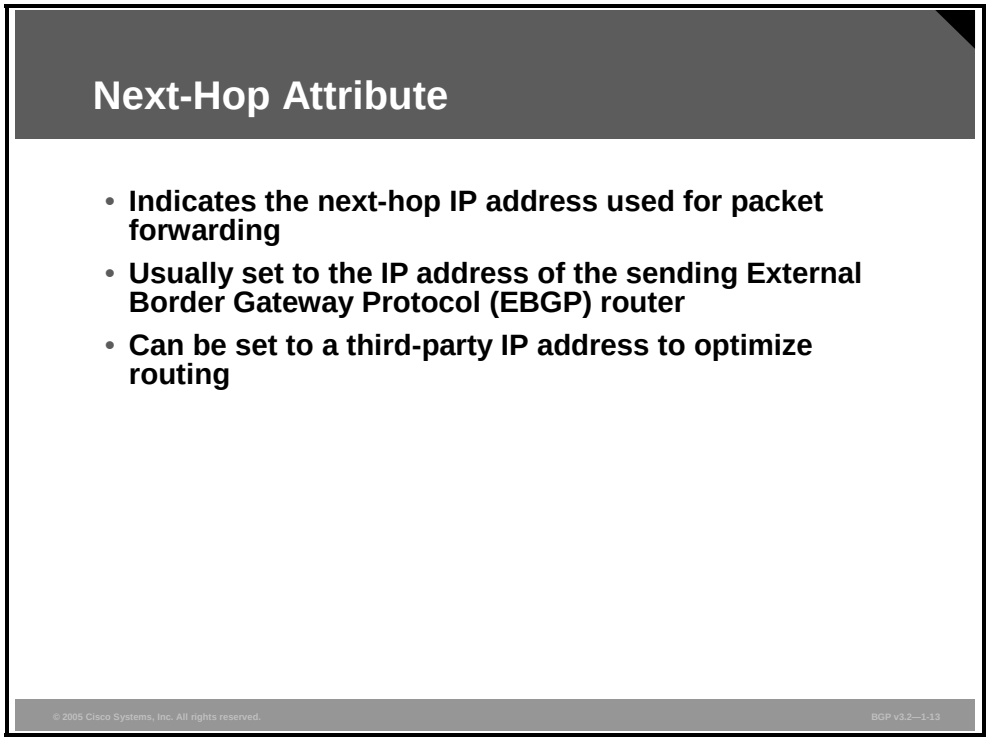
When the routing update about network 10.0.0.0/8 is sent by the edge router in AS 123 to AS 21, the AS number 123 is prepended to the empty AS path, resulting in an AS path consisting of only 123. The sending router does the prepending as part of the outgoing BGP update processing. While the route is still within AS 123, the AS-path entry for AS 123 does not appear in the AS path.

The router in AS 21 propagates the information about the network 10.0.0.0/8 to AS 37. Because it is sending the BGP update to AS 37, it prepends its own AS number to the AS path, resulting in an AS path consisting of the sequence of 21 123.

AS 37 also propagates the received route to AS 123. To avoid a routing loop, where AS 123 might try to reach its own network (10.0.0.0/8) via AS 37, BGP has a built-in mechanism that causes the router in AS 123 to drop the incoming update as soon as it finds its own AS (123), in the AS path. No error will be signaled, because nothing is really wrong. It is merely the procedure that is used by BGP to avoid a routing information loop.

Next-Hop Attribute

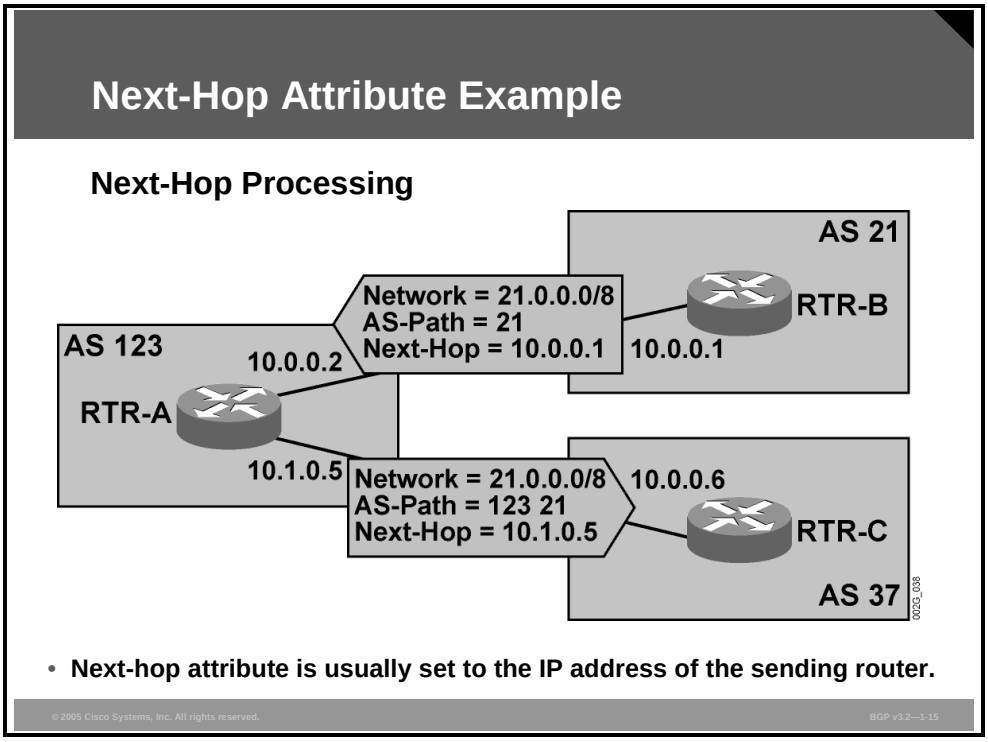
This topic describes the functionality of the next-hop attribute in BGP.



The BGP next-hop attribute identifies the IP address that a router should use to forward packets toward the destination that is announced in a BGP routing update. In most cases, the sending router sets the next-hop attribute to its own IP address. There are cases, however, where the next-hop IP address points to a third router.

Example: Next-Hop Attribute

The figure shows the usual next-hop processing.

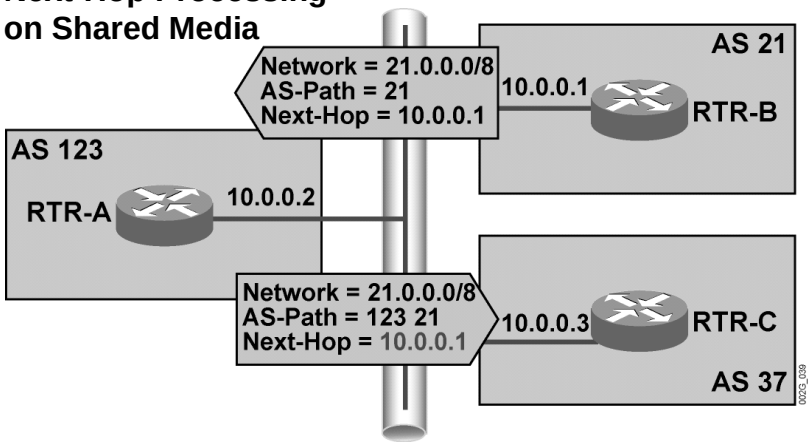


The processing is as follows:

1. RTR-B announces network 21.0.0.0/8 to RTR-A. The outgoing IP address of RTR-B (the address that is used to establish the BGP TCP session) is used as the BGP next hop.
2. RTR-A receives the routing update and installs it in its BGP table and routing table. Should RTR-A need to forward packets toward network 21.0.0.0/8, it would send those packets toward the IP address 10.0.0.1 (RTR-B).
3. When RTR-A propagates the information about 21.0.0.0/8 to RTR-C, it sets the BGP next-hop attribute to its own IP address.

Next-Hop Attribute Example

Next-Hop Processing on Shared Media



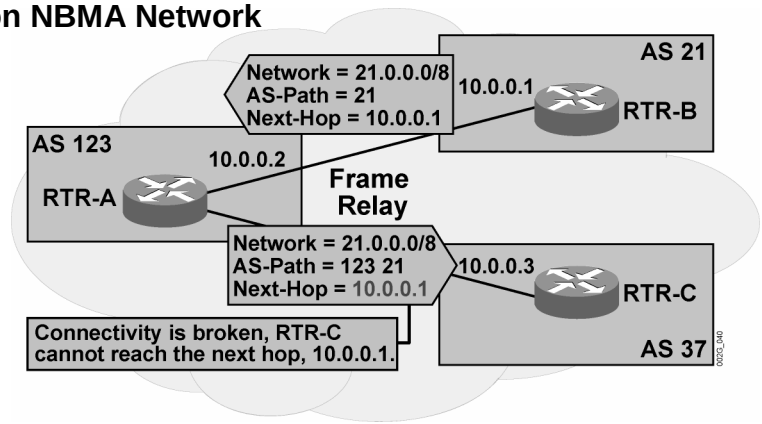
- **If the receiving BGP router is in the same subnet as the current next-hop address, the next-hop address remains unchanged to optimize packet forwarding.**

The next-hop processing changes if the BGP routers connect to a shared subnet. In the figure here, if RTR-A announces the network 21.0.0.0/8 to RTR-C with the BGP next-hop address set to RTR-A, the packets from AS 37 toward network 21.0.0.0/8 will have to cross the shared LAN twice. RTR-A thus sends the routing update toward RTR-C with the BGP next-hop address unchanged (still pointing toward RTR-B), allowing optimal data transfer across the shared LAN.

Note	More formally, the BGP next-hop rule states that if the current BGP next hop is in the same IP subnet as the receiving router, the next-hop address is not changed; otherwise, the next-hop attribute is changed to the IP address of the sending router.

Next-Hop Attribute Example

Next-Hop Processing on NBMA Network



- **BGP next-hop processing can break connectivity with improper network designs over partially meshed WAN networks.**

BGP next-hop processing results in optimum data transfer over shared media (for example, a LAN subnet). In partially meshed networks (such as Frame Relay), BGP next-hop processing can break IP connectivity. Consider, for example, the network diagram in the figure: RTR-A sends a routing update about network 21.0.0.0/8 to RTR-C with RTR-B set to the next-hop address (as they are all in the same subnet). Because there is no direct connection (virtual circuit) between RTR-C and RTR-B but RTR-C still tries to send packets directly toward RTR-B, the connectivity between AS 37 and AS 21 is broken.

There are two ways to solve the connectivity loss that is introduced by this design:

- Use the subinterfaces on RTR-A to make sure that RTR-B and RTR-C are in different subnets (and BGP next-hop processing would ensure that RTR-A is the BGP next hop in the outgoing BGP updates).
- Disable the BGP next-hop processing on RTR-A in an existing multipoint Frame Relay design that shares a common subnet. (This option is strongly discouraged in normal BGP design because routing problems should be solved with a proper network design of point-to-point subinterfaces.)

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **BGP metrics attached to a BGP route are called “path attributes.”**
- **Some path attributes are well-known and should be recognized by every BGP implementation. Some of the well-known attributes are mandatory and have to be present in every BGP update. These are the AS-path, next-hop, and origin attributes. Other well-known attributes are discretionary.**
- **Attributes that are not required to be recognized by every BGP implementation are called “optional.” These attributes could be transitive (propagated if not recognized) or nontransitive (dropped).**

Summary (Cont.)

- **The AS-path attribute lists the autonomous systems that the routing update has already crossed. This attribute is used for BGP loop detection and BGP route selection.**
- **The next-hop attribute specifies the IP address that is to be used for packet forwarding. The next hop is usually set to the IP address of the BGP router sending the update.**

Lesson 3

Establishing BGP Sessions

Overview

Understanding the Border Gateway Protocol (BGP) neighbor session establishment process is a key component to understanding the fundamental operation of the BGP protocol. It also forms a knowledge base for later lessons, including configuring basic BGP.

BGP is an exterior gateway protocol (EGP) that has been designed for scalability and policy control. As a result, BGP requires neighboring routers to be explicitly configured before BGP routing updates can be sent between them. This situation differs from Interior Gateway Protocols (IGPs) such as Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF), that discover neighbors through the use of a broadcast packet or a hello protocol. In this lesson, BGP neighbor session establishment procedures are discussed.

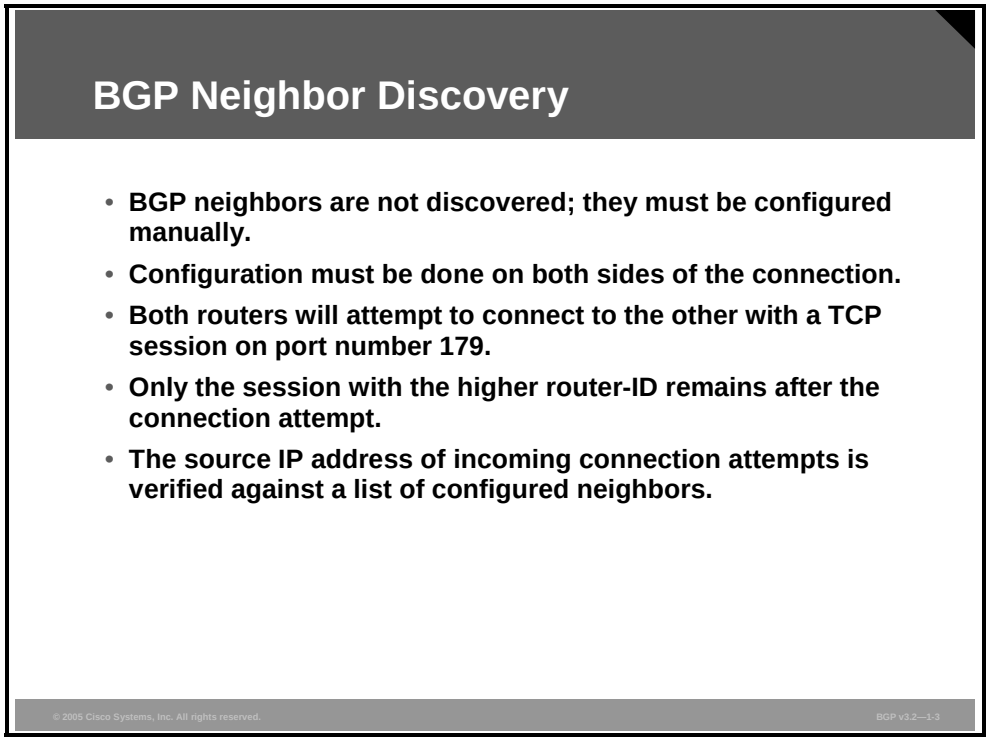
Objectives

Upon completing this lesson, you will be able to describe the concept of BGP neighbors and neighbor session establishment procedures. This ability includes being able to meet these objectives:

- Explain how BGP discovers neighbors
- Describe the BGP session establishment process
- Describe the role of the BGP keepalive in session establishment and maintenance
- Explain how optional MD5 authentication can protect sessions between BGP peers

BGP Neighbor Discovery

This topic explains how the BGP routing protocol discovers neighbors.



Unlike other routing protocols, BGP has no means of automatically detecting neighbors. The BGP protocol is carried in a TCP session, which must be opened from one router to the other. To do so, the router attempting to open the session must be manually configured with neighbor information indicating to which IP address to direct its connection attempts.

The router that receives the incoming connection attempts does not answer them if the attempts are not from one of the configured neighbors. The IP source address of the connection attempt packet (TCP SYN packet) is verified against the list of IP addresses that the router itself would direct its connection attempts to.

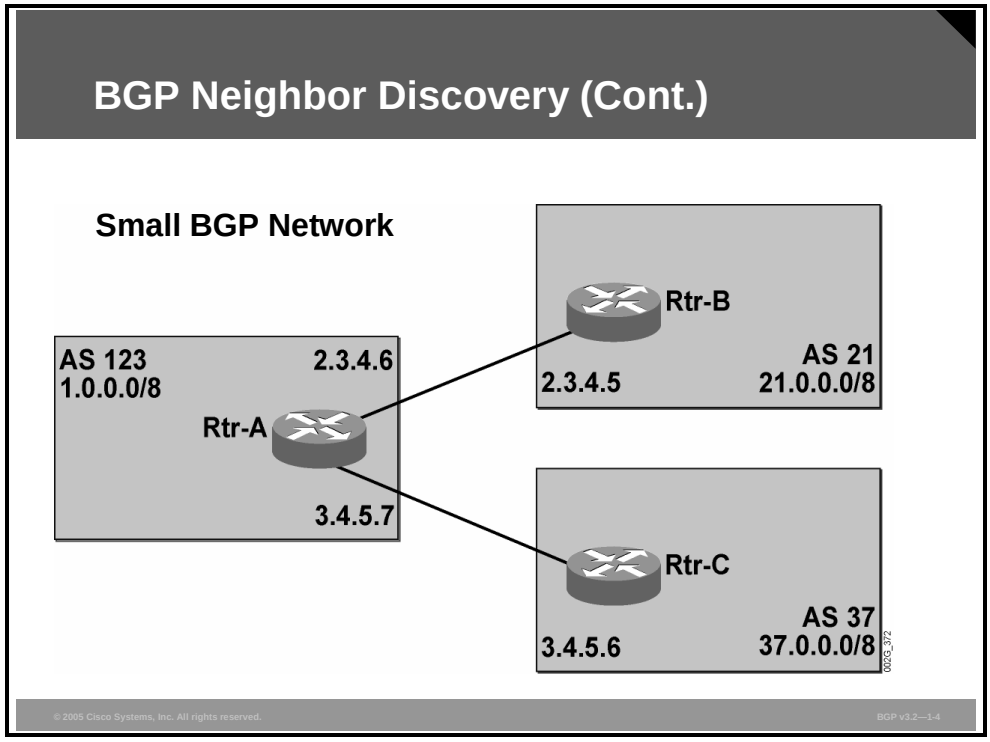
To succeed in the connection attempts, both routers must be configured to reach each other. A side effect of this situation is that they will both attempt to connect. This side effect adds robustness to the session establishment process, but it also introduces the risk that two BGP sessions will be established between a pair of BGP routers.

Two routers should have only a single BGP session between them. The router-ID values that are exchanged when the BGP session is established allow the BGP routers to detect when two parallel sessions exist. Only the session that was initiated by the router with the numerically higher router-ID will be retained. The other session is dropped.

A router may not open a BGP session to itself. If the configured neighbor IP address is, in fact, an IP address of the local router, the router recognizes the problem and tears down the session. The router-ID is also used for this verification.

Example: BGP Neighbor Discovery

This example illustrates a small BGP network.



The network displayed in the figure serves as the sample network to generate printouts in the following examples.

BGP Neighbor Discovery (Cont.)

Initially, all BGP sessions to the neighbors are idle.

```
Rtr-A#show ip bgp summary
BGP table version is 1, main routing table version 1

Neighbor        V    AS MsgRcvd MsgSent  TblVer  InQ OutQ Up/Down  State/PfxRcvd
2.3.4.5         4    21       0       0       0    0    0 never    Idle
3.4.5.6         4    37       0       0       0    0    0 never    Idle
```

0003_373

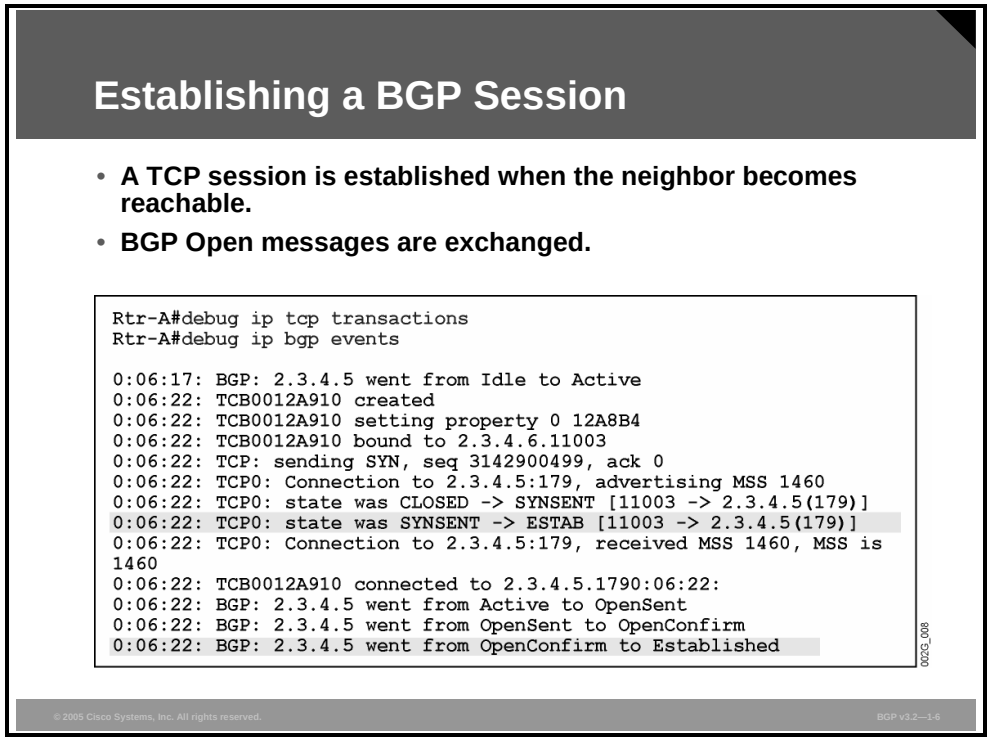
The **show ip bgp summary** command gives an overview of the BGP status. Each configured neighbor is listed in the output of the command. The IP address to which the connection attempts are directed is also displayed, along with the BGP version number, the remote AS number, some counter values, the status of the session, and how long ago the session changed state.

The “Idle” state indicates that the router is currently not attempting any connection establishments.

The various states for a BGP connection are Idle, Active, OpenSent, OpenConfirm, and Established.

Establishing a BGP Session

This topic describes the BGP session establishment process.



Before any connection attempt is made, the BGP peer relation must have left the Idle state and entered the Active state. For a BGP session between two routers in different autonomous systems, this status results when the IP address of the remote router becomes reachable on a directly connected interface.

The debug output shows how the router creates a socket data structure and binds it to its local IP address 2.3.4.6 and a high port number 11003. Then the router sends a TCP SYN packet to the configured peer router IP address of 2.3.4.5 and the well-known destination port 179. The connection attempt succeeds, and the TCP session is then ready to transfer the BGP information.

The first BGP information sent is the BGP Open message. The BGP session then goes from the Active state to the OpenSent state while waiting for the other router to respond. If the peer router accepts the parameters in the Open message, it responds with its own Open message. When the local router receives this message, the state goes from OpenSent to OpenConfirm. The local router then verifies the peer router parameters in its Open message. If they are accepted, a keepalive packet is sent to signal this acceptance. The state is then “Established.”

Establishing a BGP Session (Cont.)

The BGP Open message contains the following:

- **BGP version number**
- **AS number of the local router**
- **Holdtime**
- **BGP router identifier**
- **Optional parameters**

The parameters in the BGP Open message are as follows:

- **Version number:** The suggested version number. The highest common version that both routers support is used. Most BGP implementations today use BGP version 4 (BGP4).
- **AS number:** The AS number of the local router. The peer router verifies this information. If it is not the AS number that is expected, the BGP session is torn down.
- **Holdtime:** The number of seconds that may elapse between reception of successive BGP messages. If the time is exceeded, the peer is considered dead. The two routers agree to use the lowest suggested value. When the session is established, both routers use keepalive messages to make sure that the hold timer does not expire. A suggested hold-timer value of 0 indicates that the timer never expires and no keepalives should be sent.
- **BGP identifier:** A number uniquely identifying the router. The Cisco router uses one of its IP addresses for this number, the router-ID. The router-ID is selected as the numerically highest IP address of any loopback interface. If there is no loopback interface, the router uses the highest IP address of any interface that is up at the time of the start of the BGP process.
- **Optional parameters:** Type, length, value (TLV) encoded. An example of optional parameters is session authentication.

Establishing a BGP Session (Cont.)

BGP neighbors? steady state

- **All neighbors shall be up (no state information).**

```
Rtr-A#show ip bgp summary
BGP table version is 10, main routing table version 10
3 network entries (3/6 paths) using 516 bytes of memory
3 BGP path attribute entries using 284 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory

Neighbor  V  AS MsgRcvd MsgSent  TblVer  InQ OutQ Up/Down  State/PfxRcvd
2.3.4.5   4  21      17      22      10    0    0 0:01:47          27
3.4.5.6   4  37      11      17      10    0    0 0:07:07          35
```


When the BGP sessions are in the Established state, routing information exchange can take place. The **show ip bgp summary** command output here indicates that a session is established by not displaying any information at all in the “State” column.

The counter values show how many messages have been received and sent in the session. “InQ” shows how many messages have been received but not yet processed. A high InQ number indicates lack of CPU resources to process the input. “OutQ” shows how many outgoing messages are queued. A high OutQ number indicates lack of bandwidth to transmit the outgoing messages or CPU overload of the other router.

“TblVer” (table version) is used by the BGP router to track the changes that need to be sent to the neighbors. There is a major table version number for the local BGP table. The table version number is displayed on the first line of output from this **show** command. There is also one table version number maintained for each of the neighbors of the BGP router; this number is displayed on the information line of the neighbors.

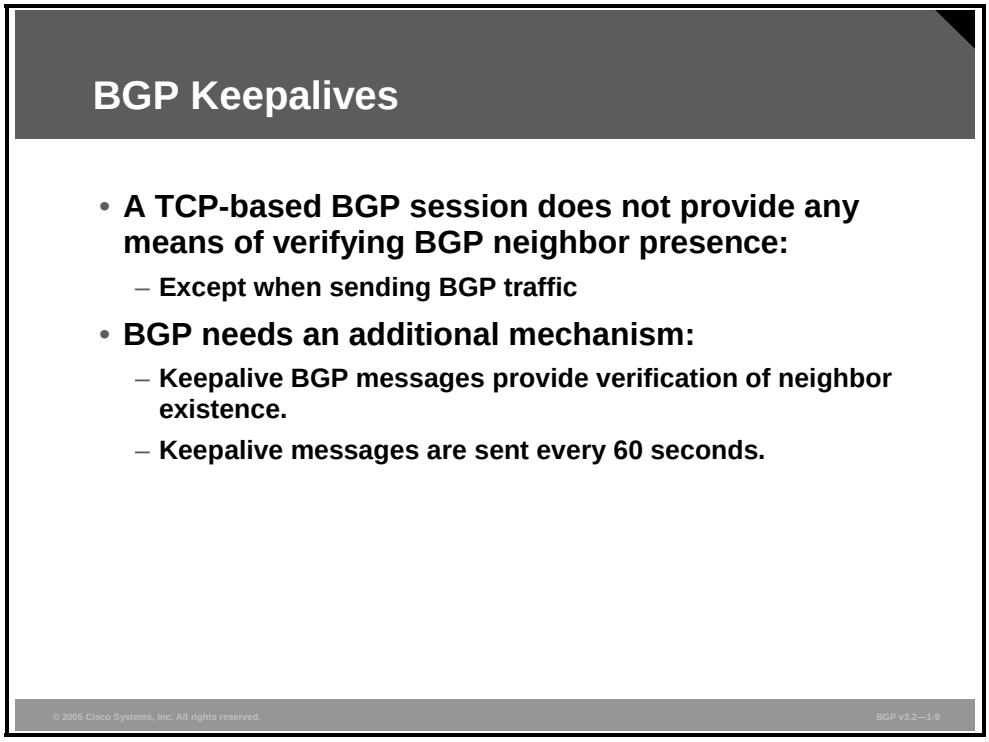
Whenever a BGP router enters a change into its BGP table, the major table version number is incremented and the changed route is tagged with this number. When the time comes to update a specific neighbor, the router scans the BGP table, and all changes that it finds where the version number is between the neighbor version and the current table version are sent to the BGP neighbor in a single BGP routing update. After the entire table is scanned and all changes have been sent to the neighbor, the table version number of the neighbor is set to the highest value of the routes being sent.

A table version of a neighbor that is lower than the major table version indicates that the neighbor is not yet fully updated. The update interval for a neighbor in another AS is normally 30 seconds (the default value of the BGP advertisement timer).

In addition to the information about all sessions to all neighbors, the output also shows the amount of memory that is being used for the BGP data structures.

BGP Keepalives

This topic describes the role of the BGP keepalive in session establishment and maintenance.

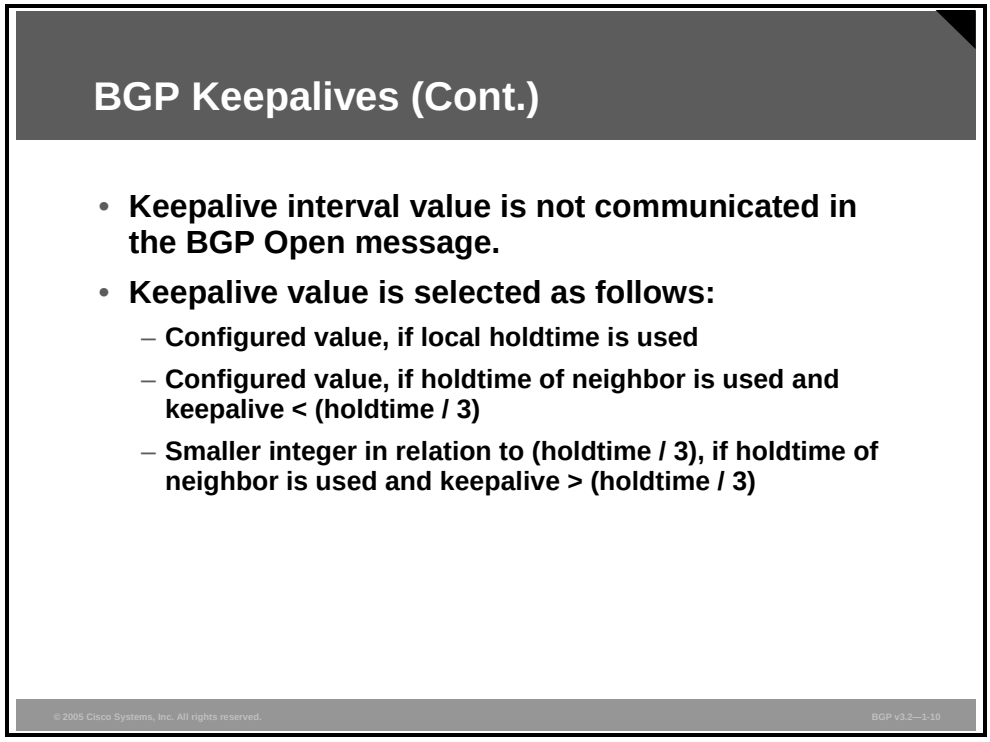


TCP-based BGP sessions do not provide any means of verifying the presence of a BGP neighbor. After BGP has established the TCP session, the only method of verifying neighbor presence is to actually send BGP traffic. BGP traffic is sent over the TCP session with acknowledgments (ACKs), and is therefore reliable. Successfully sending BGP traffic confirms the existence of a BGP neighbor.

However, there are often long periods of time when no BGP traffic is sent between neighbors. During those periods, TCP implements no mechanism to check for the existence of the configured neighbor. BGP neighbors could therefore easily be disconnected during times of session inactivity. This situation would lead to incorrect routing information on the other side of the BGP session.

To avoid routing packets to a router that is no longer there, BGP needs an additional mechanism to make sure that a neighbor exists. BGP sends special keepalive messages during every keepalive interval to inform its peer of its presence. By default, this interval is every 60 seconds. If no BGP traffic is received within the selected holdtime interval, the BGP router sends a BGP notification message to the inactive peer and tears down the BGP session between them. The default BGP holdtime value is 180 seconds.

When changing the default values of keepalive and holdtime intervals, you must take care not to configure too big a keepalive interval in comparison to the holdtime. Too big a difference could result in resetting of the BGP session after only one keepalive message has been missed, making a network unstable. The suggested ratio of keepalive-to-holdtime interval is 1:3.



As opposed to the holdtime interval, BGP peers do not communicate the keepalive interval in the Open message. The selection of a keepalive interval is therefore based on the selected holdtime value. The selected holdtime value that is used by both peers is the smaller of both configured values.

The BGP process selects the keepalive interval value using the following steps:

Step 1 If the locally configured value of holdtime is selected (being the lower of two), the peers use the locally used keepalive interval.

Step 2 If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is less than a third of the holdtime interval, the peers use the locally configured keepalive.

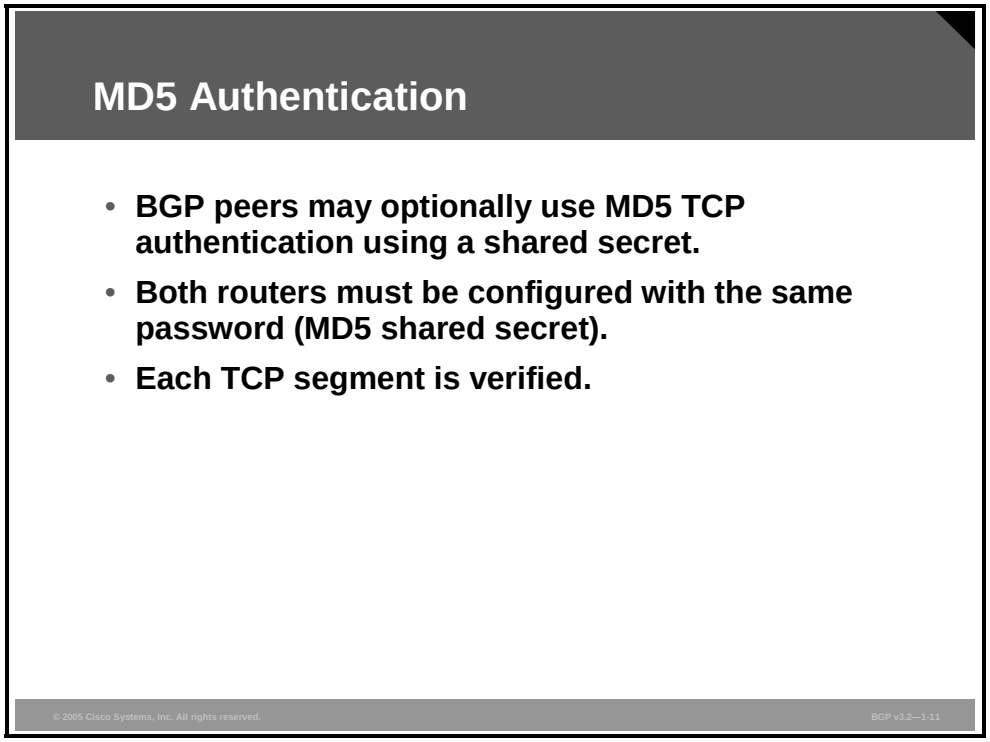
Step 3 If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is more than a third of the holdtime interval, the peers use the smaller integer value in relation to (holdtime / 3).

Example: Keepalive Value

If the selected holdtime equals 17 seconds and the configured keepalive equals 10 seconds, the (holdtime / 3) rule will be used to select the keepalive value. Therefore, (17 / 3) = 5.67, and the keepalive value used by BGP will equal 5 seconds.

MD5 Authentication

This topic explains how Message Digest 5 (MD5) authentication protects a BGP neighbor session.

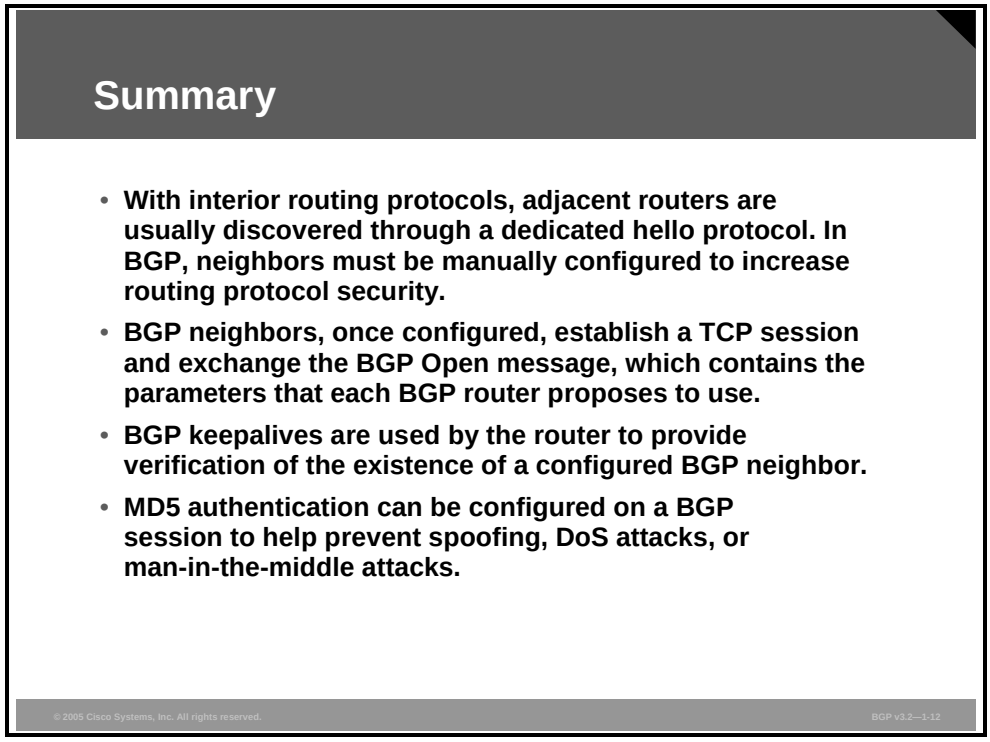


Authentication between BGP neighbors can be negotiated between BGP-speaking routers using optional parameters in the Open message. If you are using MD5 authentication, every TCP segment on the BGP session will be transmitted to the configured neighbor along with a checksum. The checksum is calculated together with a secret known by the two routers using the MD5 algorithm. The common secret is never transmitted on the network. If the receiver, which is using the same common secret, calculates the same checksum from the TCP segment, then the receiver can be pretty sure that the information is transmitted from the correct source and the information has not been altered.

Authentication of BGP sessions is a vital tool to avoid denial-of-service (DoS) attacks.

Summary

This topic summarizes the key points discussed in this lesson.



Lesson 4

Processing BGP Routes

Overview

Route processing is fundamental to the operation of Border Gateway Protocol (BGP). Knowledge of the BGP route selection process, route propagation, and how the BGP and IP routing tables are built is key to properly configuring BGP and troubleshooting BGP routing issues.

This lesson explains the details of processing IP routing information in BGP. It describes how a router builds the BGP routing table, how BGP selects the best route, and how BGP routes are propagated to other BGP neighbors. The lesson also discusses how a router builds the IP routing table when it is using BGP.

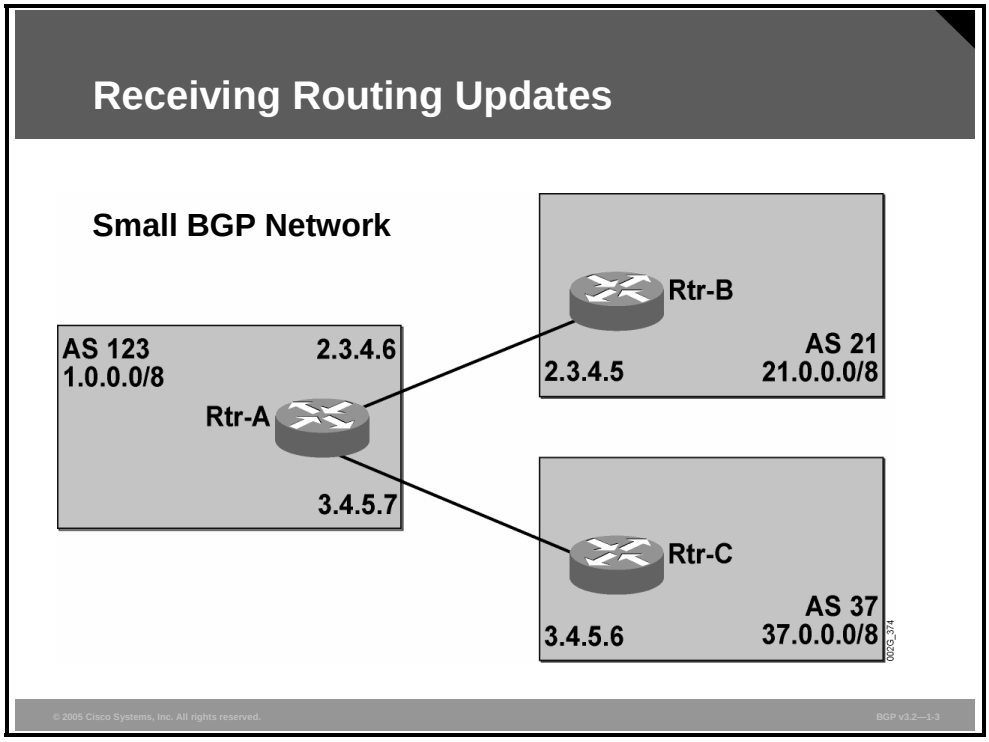
Objectives

Upon completing this lesson, you will be able to describe BGP route processing. This ability includes being able to meet these objectives:

- Describe BGP routing updates
- Explain how a router builds BGP tables
- Describe the route selection process in BGP
- Explain how a router propagates BGP routes to other BGP neighbors
- Explain how a router builds an IP routing table when it is using BGP
- Explain how BGP advertises local networks
- Describe the role of automatic summarization in BGP route processing

Receiving Routing Updates

This topic describes BGP routing updates.



The network displayed in the figure serves as the sample network for generating printouts in the following examples.

Receiving Routing Updates (Cont.)

Information from the BGP tables is exchanged after adjacency establishment.

```
Rtr-A#debug ip bgp update
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 37.0.0.0 255.0.0.0, next hop
2.3.4.5, path 21 37 metric 0
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 1.0.0.0 255.0.0.0 -- denied
1:24:11: BGP: 2.3.4.5 rcv UPDATE about 21.0.0.0 255.0.0.0, next hop
2.3.4.5, path 21 metric 0
1:24:11: BGP: nettable_walker 21.0.0.0/255.0.0.0 calling revise_route
1:24:11: BGP: revise route installing 21.0.0.0/255.0.0.0 -> 2.3.4.5
```

0023_010

After a BGP session is established, routing updates start to arrive. Each BGP routing update consists of one or more entries (routes). Each route is described according to the IP address and subnet mask, along with any number of attributes. The next-hop, autonomous system (AS)-path, and origin attributes must always be present. Other BGP attributes are optionally present.

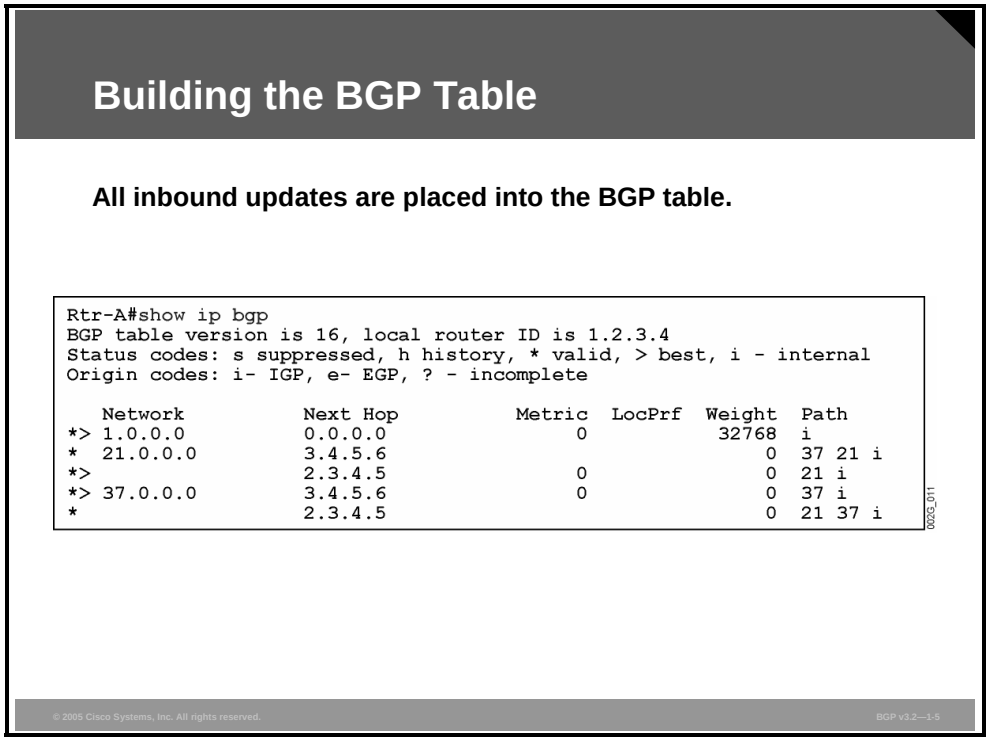
The debug output shows how information about network 37.0.0.0/8 is received from neighbor 2.3.4.5. The neighbor indicates that IP packets to destination IP addresses in network 37.0.0.0/8 can be forwarded to the next-hop address 2.3.4.5. The AS path 21 37 indicates that the final destination is in AS 37 but the packets have to pass through AS 21 to get there. The metric is the multi-exit discriminator (MED) value.

Network 21.0.0.0/8 also has the next-hop address of 2.3.4.5, but the AS path of 21 indicates that that network is inside the directly connected AS 21.

Network 1.0.0.0/8 is denied. The reason is not obvious from the debug output, but the network topology information indicates that network 1.0.0.0 is local to (inside) AS 123. AS 123 has advertised the network to AS 37, which propagates to AS 21 and returns to AS 123. This information loop is detected by the content in the AS-path attribute. The receiving router detects its own AS number in the AS path and silently discards (denies) the route.

Building the BGP Table

This topic explains how a router that is enabled for BGP builds a BGP routing table.



All routes that are received from a neighbor are saved in the router memory. Therefore, there is no need to retransmit or refresh any unchanged information.

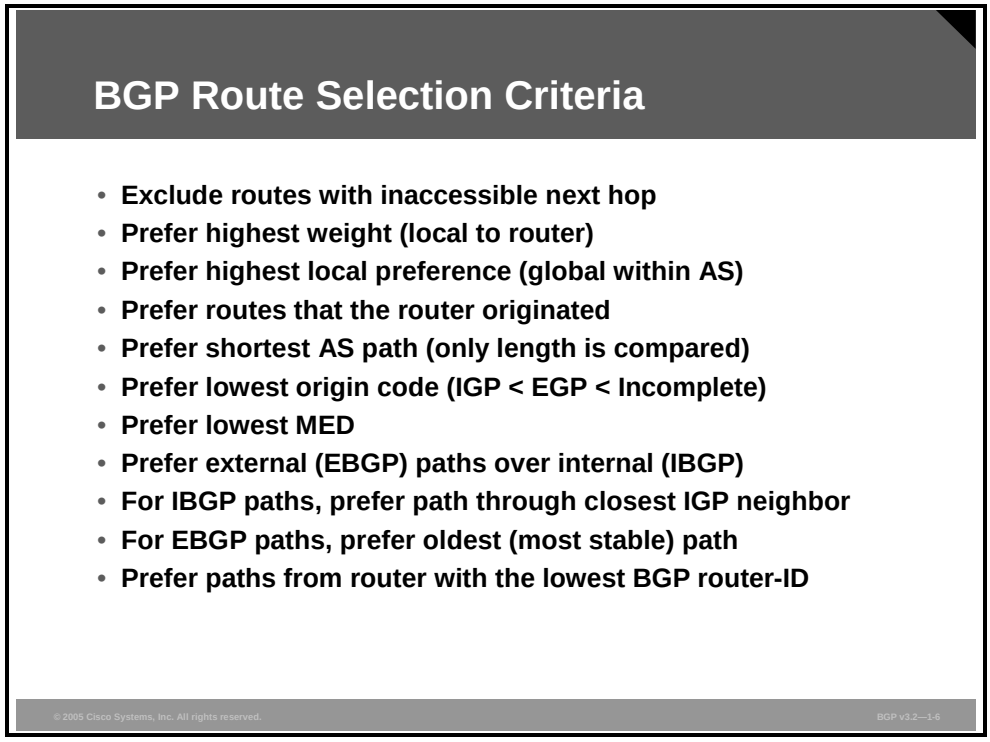
When there is more than one way to reach a particular network, the local router selects one of those as the best. If that alternative is later lost because the neighboring router withdraws the route (or the neighboring router is no longer reachable), the remaining alternatives are still stored in memory and a new alternative is selected as the best without involving other BGP routers.

The **show ip bgp** router command gives an overview of all routing information received from all neighbors. The command displays basic information about each route on a single line. The output is sorted—alternatives to reach the same network are displayed on consecutive lines. The network number is displayed only on the first lines indicating the same network. The network column is left blank on the consecutive lines indicating alternatives to reach the same network.

The router selects only one of the alternatives as the best path toward the destination. This alternative is indicated with the “>” sign.

BGP Route Selection Criteria

This topic describes the route selection process in BGP.



When a router has more than one alternative route to reach the same IP subnet (network and mask), the router has to select one of the routes as best in its default mode of operation. To make this selection, the router uses the BGP attributes that are attached to the various updates. The selection criteria are checked in the order that is indicated in the following steps. The first of the checks that indicates a difference is used, and no further testing is done.

Step 1 The router checks whether the next-hop attribute indicates an IP address that is reachable according to the current routing table. It is not necessary to have a direct connection to the next hop. It can very well be several router hops away and the route to it learned by the Interior Gateway Protocol (IGP). If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best.

Step 2 The router prefers the route with the higher weight. The weight is not carried with the updates; it is a value that is assigned to the route by the local router and considered only within the router itself.

Step 3 If the local preference attributes are different, the route with the highest value is selected as best.

Step 4 If one of the routes is injected into the BGP table by the local router, the local router prefers it to any routes that it receives from other BGP routers.

Step 5 At this point, the lengths of the AS paths are compared (the content is not checked; only the number of autonomous systems in each AS path is counted). The route with the shortest length is selected.

Step 6 If the AS-path lengths are the same, the origin code is checked. BGP will prefer the path with the lowest origin type: IGP is lower than exterior gateway protocol (EGP), and EGP is lower than Incomplete.

Step 7 The router next compares MED values but only if it receives the updates from the same neighboring AS. Routes with a lower MED are preferred.

Step 8 At this point it is clear that the destination network is outside the local AS and that there is not much difference among the alternatives. Because the IP packets to the destination network must leave the AS, it is better that they do so as quickly as possible. If any of the alternatives are received from a BGP peer in another AS, that alternative is preferred.

Step 9 If the router receives all alternatives from peer routers in the local AS, each of them will indicate an exit point, and the closest exit is used. Distance to the exit point is calculated by comparing the IGP costs against the BGP next hops, as indicated in the routing table.

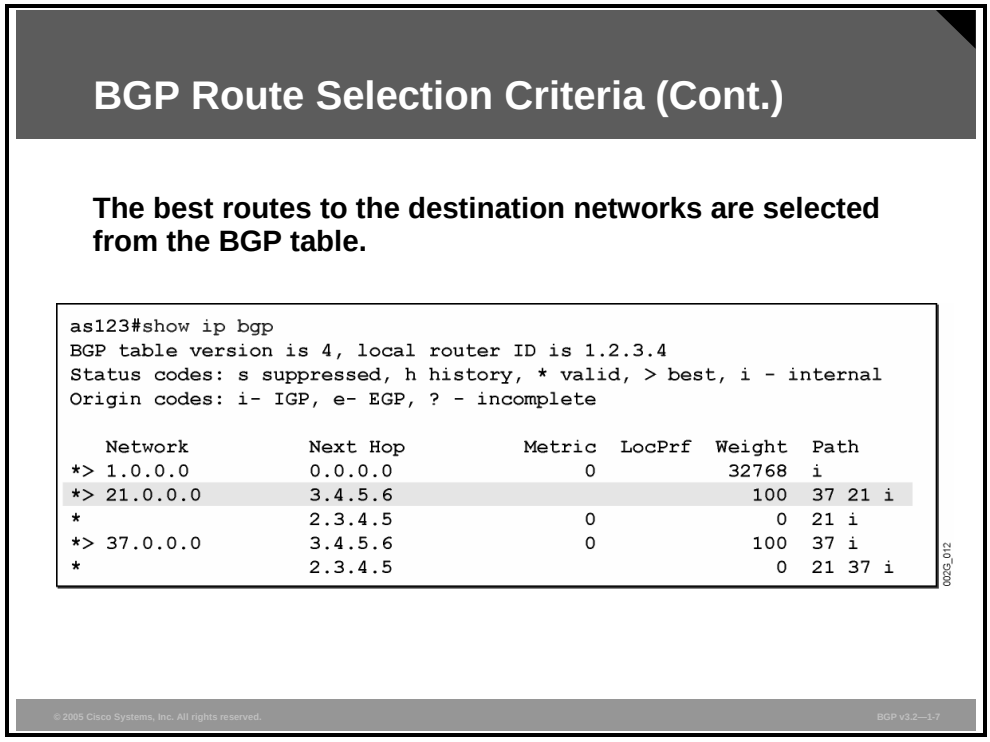
Step 10 If the router receives all alternatives from External Border Gateway Protocol (EBGP) neighbors, the most stable path (the oldest path) is preferred.

Step 11 If the router still cannot differentiate among the routes, it nevertheless has to make a decision and select the best route. It checks the BGP sessions on which it received the updates and chooses the route that was received on the session for which the peer router has the lowest BGP router-ID.

The router makes the final test only after it has made all other checks and determined that all alternative routes are equally good.

Example: BGP Route Selection Criteria

In this example, the router in AS 123 can reach network 21.0.0.0/8 via two paths.

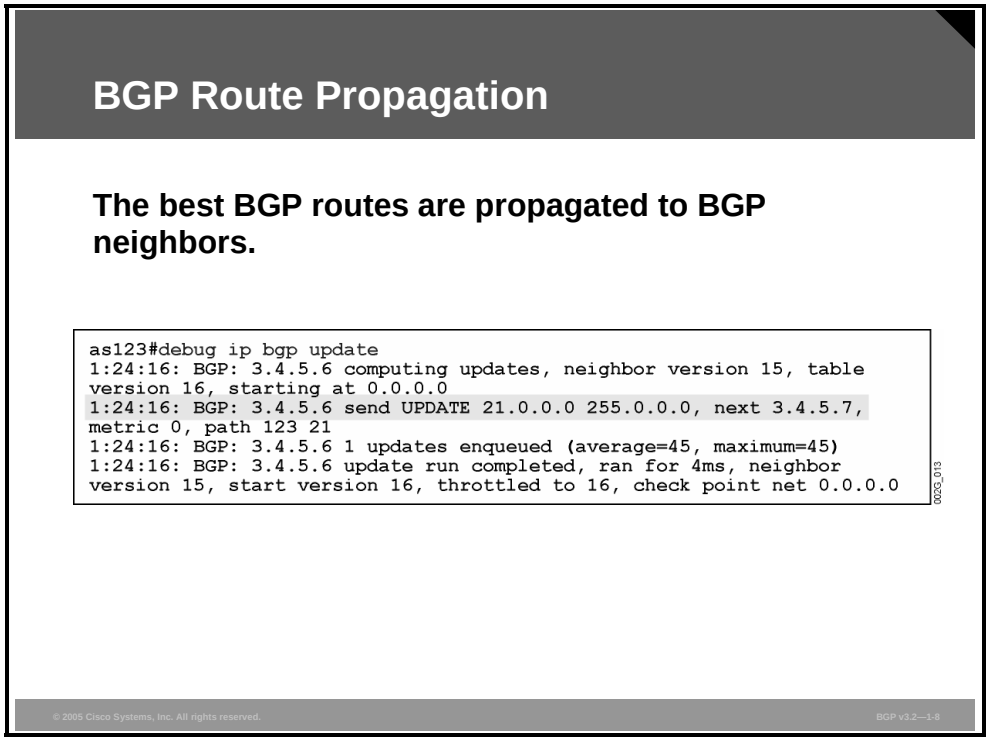


The first path is via neighbor 3.4.5.6 in AS 37 and then to AS 21, and the second path is straight to AS 21 through neighbor 2.3.4.5. In this example, the weight is set to 100 for the path via AS 37, and the other alternative path does not have a weight set. Thus, when checked against BGP path selection rules, the route via AS 37 is selected as the best because it has a higher weight attribute.

Likewise, network 37.0.0.0/8 is reached via AS 37 because the weight indicates that it is the best route.

BGP Route Propagation

This topic explains how a router propagates BGP routes to other BGP neighbors.



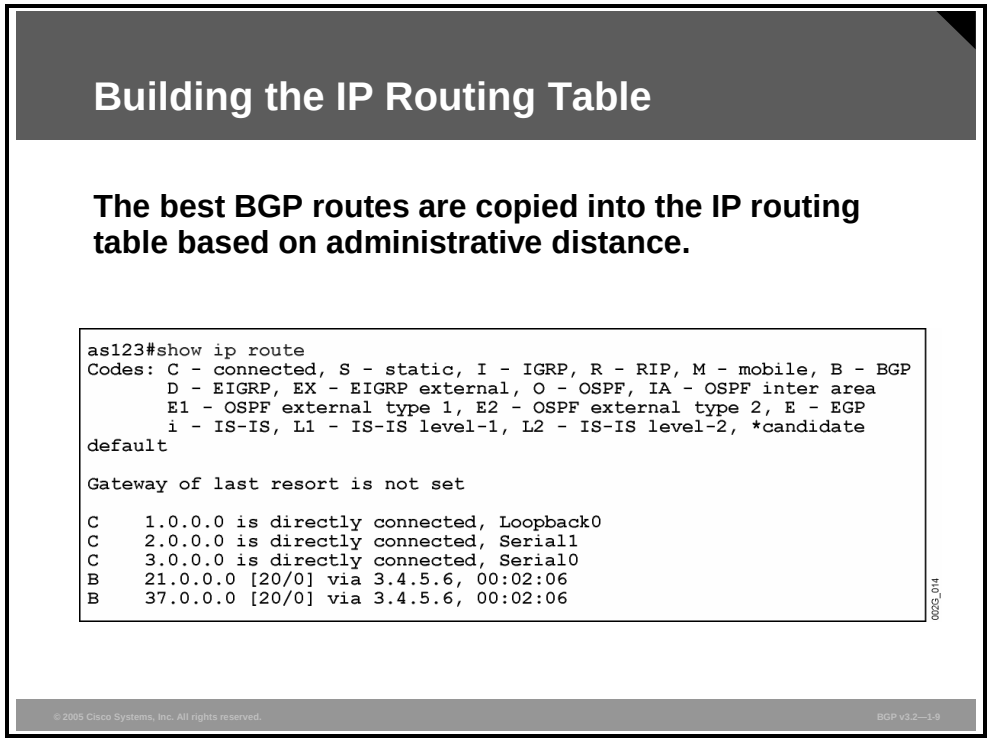
A local router propagates only the route that it selected as best to the neighbors. However, the router never sends a route back on the same BGP session upon which it was received. On the contrary, when it selects a neighbor as the best next hop, the router makes sure that the neighbor is not pointing back to the local router; it accomplishes this task by “poisoning” the route (marking the route unreachable) and sending a withdraw message to that neighbor.

The router conducts route poisoning to avoid a potential routing loop problem in which a neighbor router selected as the best next hop might rely on the local router as the best next hop.

The process of preventing routing information from being sent back to the source of information is called “split horizon.”

Building the IP Routing Table

This topic describes the process of building an IP routing table from the BGP table and from other sources of routing information, such as IGPs.



The route in the BGP table that BGP selects as the best is a candidate for installation in the IP forwarding, or routing, table.

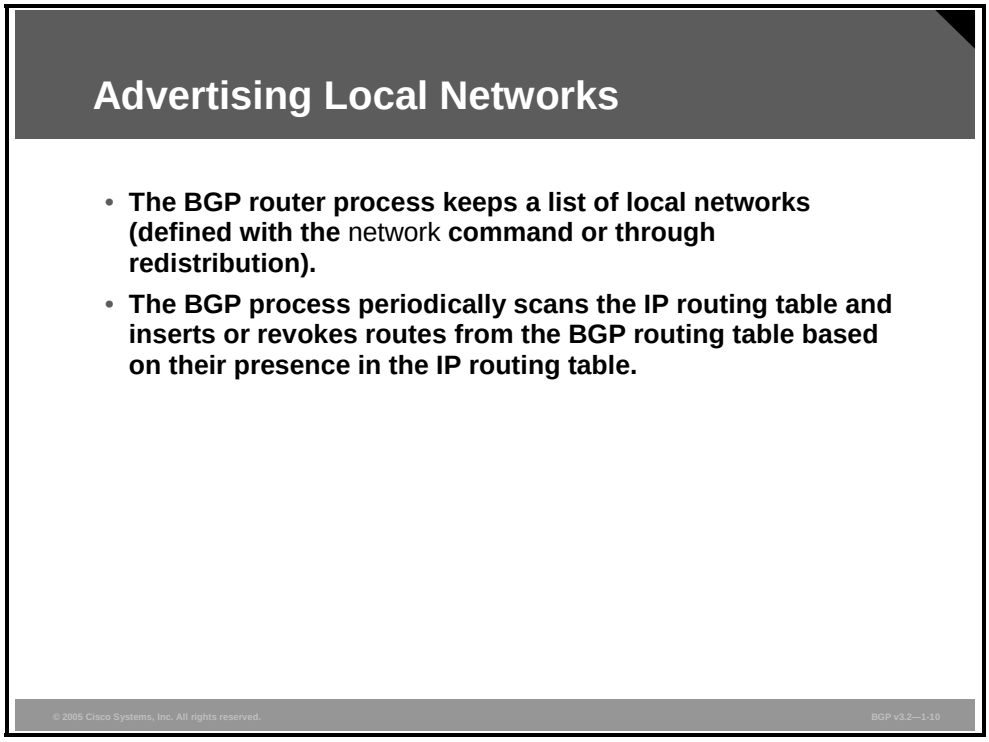
Before a route can be installed, the router has to check whether there is any other routing protocol that has information about the same subnet (network and mask). If the subnet is known via different sources, the router uses the administrative distance (AD) to determine which source to use. AD is a rating of the trustworthiness of a routing information source. AD is often expressed as a numerical value between 0 and 255. The higher the value, the lower the trustworthiness rating. In this case, the router will install the route with the lowest AD.

The output from the **show ip route** command indicates which routes in the routing table were installed using the BGP information. Those routes are denoted with the letter “B.” The AD is shown in the command output as the first number within the brackets.

In this example, networks 21.0.0.0/8 and 37.0.0.0/8 are both reachable via 3.4.5.6. After the router has installed the routes in the routing table, user data traffic starts to be forwarded.

Advertising Local Networks

This topic explains how BGP advertises local networks.



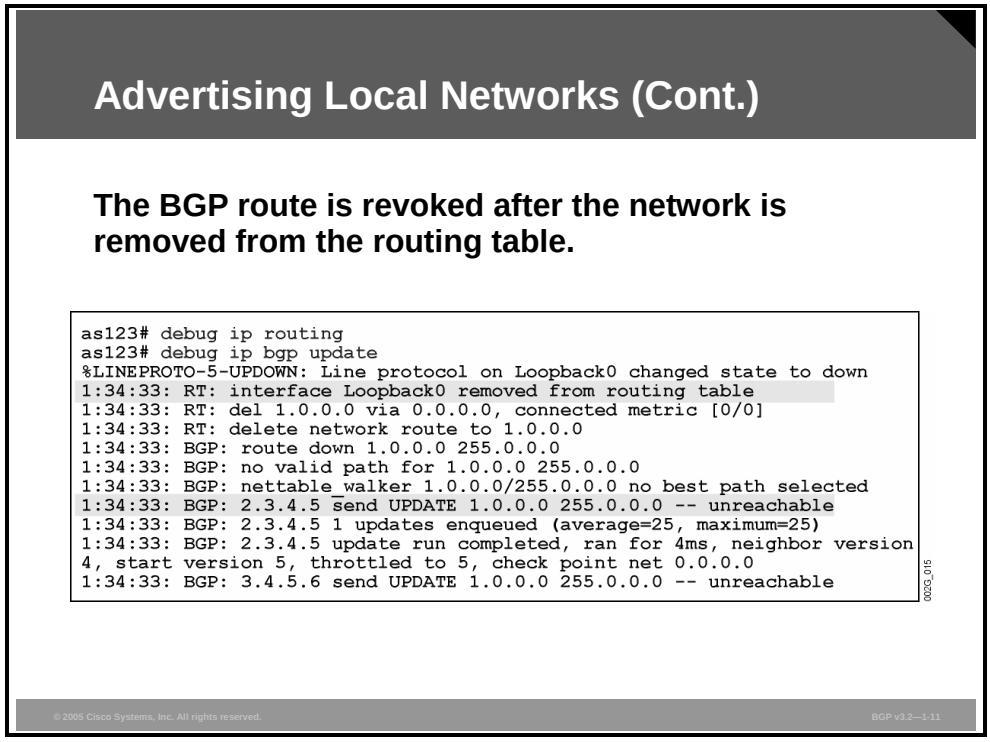
The BGP routing process can inject new routes into the BGP table. A router will propagate newly injected routes to neighboring BGP peers if it selects them as best, giving neighboring autonomous systems information about networks that are reachable in the local AS. This process is called advertising, originating, or announcing local routes.

The BGP process can inject local routes in two different ways:

- A list of networks is configured on the router under the BGP router process using the **network** configuration command. The networks listed are candidates for being injected. Networks are injected only if they appear in the routing table. In the case where the IGP that is used within the AS finds a valid path to them, the routes will be in the routing table.
- Routes that are learned by another routing protocol are redistributed. The IGP that is used with the AS can also act as a source of routing information about local networks.

Example: Advertising Local Networks

In this example, network 1.0.0.0/8 is directly connected to interface Loopback0.



The route to 1.0.0.0/8 was previously installed in the BGP table because it was listed with a network statement and it was in the routing table as directly connected. When the Loopback0 interface goes down, the router removes the directly connected route from its routing table. Because the route no longer exists in the routing table, it must also be removed from the BGP table.

Because there has been a change in the BGP table, the BGP neighbors must be informed. The router sends a BGP update message to both neighbors indicating that network 1.0.0.0/8 is now unreachable.

Advertising Local Networks (Cont.)

The BGP route is advertised after the network appears in the routing table.

```
1:36:42: RT: add 1.0.0.0 255.0.0.0 via 0.0.0.0, connected metric [0/0]
1:36:42: RT: interface Loopback0 added to routing table
1:36:42: BGP: route up 1.0.0.0 255.0.0.0
1:36:42: BGP: nettable_walker 1.0.0.0/255.0.0.0 route sourced locally
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback0, changed state
to up
1:36:43: BGP: 2.3.4.5 computing updates, neighbor version 5, table
version 6, starting at 0.0.0.0
1:36:43: BGP: 2.3.4.5 send UPDATE 1.0.0.0 255.0.0.0, next 2.3.4.6,
metric 0, path 123
1:36:44: BGP: 2.3.4.5 1 updates enqueued (average=50, maximum=50)
1:36:44: BGP: 2.3.4.5 update run completed, ran for 4ms, neighbor
version 5, start version 6, throttled to 6, check point net 0.0.0.0
```

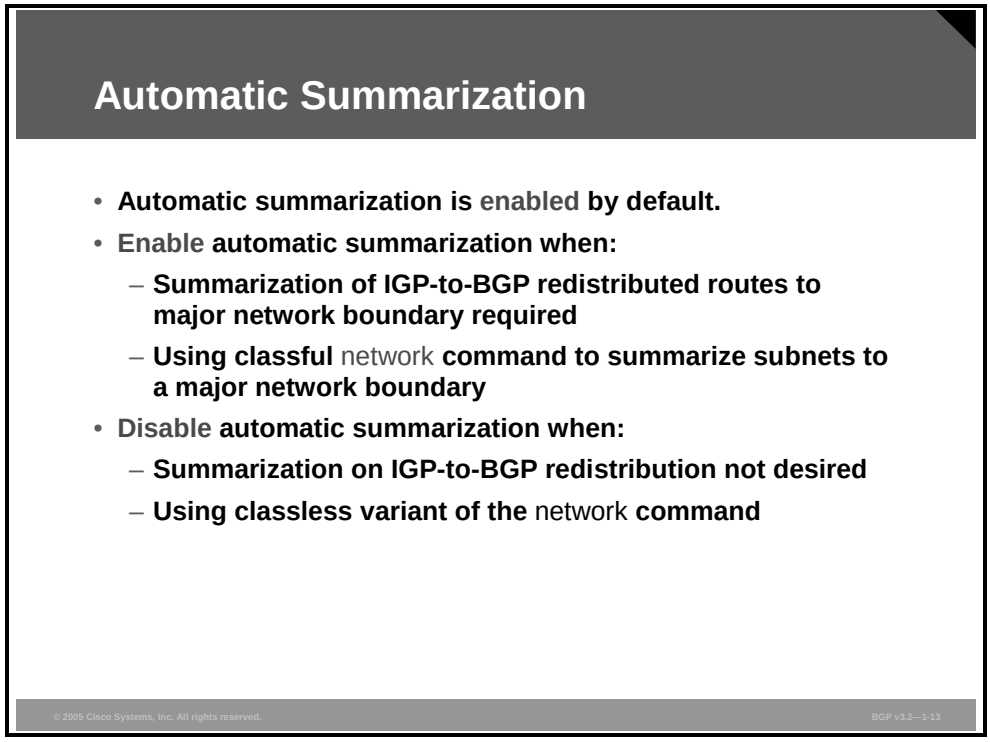
0025_016

In this example, network 1.0.0.0/8 is listed with a network statement in the BGP process. However, the network was not in the routing table of the router, so the network was not injected into its BGP table.

Later, the Loopback0 interface comes back up again. This reappearance means that the network 1.0.0.0/8 is now in the routing table as a directly connected route. As a result, the router once again injects the 1.0.0.0/8 network into its BGP table and subsequently updates its configured neighbors.

Automatic Summarization

This topic describes the role of automatic summarization in BGP route processing.



When a BGP router is configured to locally announce routes into BGP, the behavior of the **network** command varies depending on whether automatic summarization is enabled or disabled. When automatic summarization is enabled, BGP summarizes the locally originated BGP networks (network *x.x.x.x*) to their classful boundaries. Automatic summarization is enabled by default in BGP.

When a subnet exists in the routing table and the following three conditions are satisfied, then any subnet (component route) of that classful network in the local routing table will prompt BGP to install the classful network into the BGP table:

- A classful network statement for that network exists in the routing table.
- A classful mask has been configured on that network statement.
- Automatic summarization is enabled.

When automatic summarization is disabled, the routes that are introduced locally into the BGP table are not summarized to their classful boundaries.

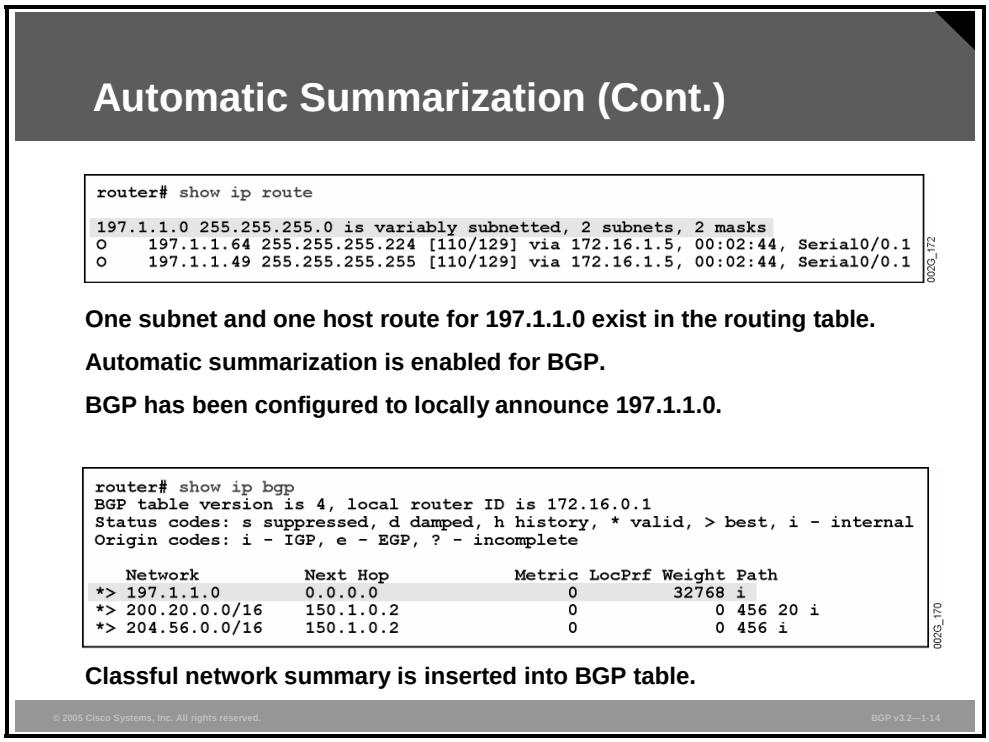
The behavior of the redistribution procedure in BGP is also influenced by the configuration of automatic summarization on the router. When enabled, all redistributed subnets will be summarized to their classful boundaries in the BGP table. When disabled, all redistributed subnets will be present in their original form in the BGP table.

Enable automatic summarization in BGP when the summarization of subnets to their classful boundaries will not introduce flawed information into the BGP table. In other words, leave automatic summarization enabled only when you are using a fully assigned classful network matching the network that was summarized in BGP.

Whenever possible, use the classless variant of the **network** command, specifying the subnet mask length of the network. When you are redistributing networks into BGP, the preferred method is to disable automatic summarization. Disabling automatic summarization ensures that correct information is inserted into the BGP table of the router.

Example: Automatic Summarization

In this example, one subnet and one host route of the major class C network 197.1.1.0/24 (197.1.1.64/27 and 197.1.1.49/32) exist in the routing table.



When you are inserting networks into the BGP table with the classful **network** command and automatic summarization is disabled, no insertion into the BGP table will occur unless an exact match exists in the IP routing table (meaning that a classful network has to be present in the IP routing table).

When automatic summarization is enabled, the major **network** command will summarize all subnets in the IP routing table to their major network boundary.

There is a classful **network** command, and automatic summarization is enabled for BGP. This setup results in the insertion of a classful network summary into the BGP table, instead of separate subnets.

Subnet 197.1.1.64/27 and host route 197.1.1.49/32 were summarized during insertion into the BGP table to the classful network 197.1.1.0/24. This action occurred because a classful **network** command and automatic summarization were configured on the router. If automatic summarization were disabled, no insertion into the BGP table would occur at all.

The locally sourced summary has all the attributes of a locally sourced BGP route (next hop = 0.0.0.0, weight = 32768, empty AS-path list), and is marked as having an IGP origin (being sourced with the **network** command).

Automatic Summarization (Cont.)

```
router# show ip route

172.16.0.0 255.255.0.0 is variably subnetted, 5 subnets, 2 masks
O    172.16.1.0 255.255.255.252 [110/128] via 172.16.1.5, 00:36:35, Serial0/0.1
O    172.16.0.2 255.255.255.255 [110/65] via 172.16.1.5, 00:36:35, Serial0/0.1
O    172.16.0.3 255.255.255.255 [110/129] via 172.16.1.5, 00:36:35, Serial0/0.1
```

0023_173

One subnet and two host routes for 172.16.0.0 exist in the routing table.

Automatic summarization is enabled for BGP.

BGP has been configured to redistribute Open Shortest Path First (OSPF) into BGP.

```
router# show ip bgp
BGP table version is 8, local router ID is 172.16.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.0.0       0.0.0.0                  0         32768 ?
*> 200.20.0.0/16    150.1.0.2                0             0 456 20 i
*> 204.56.0.0/16    150.1.0.2                0             0 456 i
```

0023_171

Classful network summary is inserted into BGP table.

In this example, automatic summarization is enabled, resulting in the summarization of redistributed subnets to their classful boundaries. Subnet 172.16.1.0/30 and the two host routes 172.16.0.2/32 and 172.16.0.3/32 will be summarized into the single class B network 172.16.0.0/16. The network 172.16.0.0/16 is a locally sourced summary with all the attributes of a locally sourced BGP route (next hop = 0.0.0.0, weight = 32768, empty AS-path list). The origin of the route is marked as incomplete because the route is sourced through redistribution.

If automatic summarization were disabled, more specific routes would be present in the BGP table instead of the summary prefix 172.16.0.0/16.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **After BGP sessions are established between BGP routers, they can start exchanging routing updates.**
- **All updates that are received from BGP neighbors are stored in the BGP table, regardless of whether they are used.**
- **The route selection process takes into account various BGP attributes that are attached to the route, as well as local decisions (indicated with weights).**
- **Only the best BGP routes are propagated to other BGP routers.**

Summary (Cont.)

- **Only the best BGP routes are installed in the local IP routing table.**
- **Every BGP router can also originate the routes in BGP. The routes to be originated are entered manually in the BGP routing process or redistributed into BGP from an IGP.**
- **Automatic summarization is enabled by default in BGP.**

Lesson 5

Configuring Basic BGP

Overview

Basic Border Gateway Protocol (BGP) configuration is critical to any successful BGP implementation. Network administrators use the Cisco IOS commands that are included in this lesson in all BGP implementations. Thorough knowledge of the commands in this lesson is therefore crucial to ensuring a successful implementation using BGP.

This lesson introduces the Cisco IOS commands that are required to configure a router for basic BGP operation. Included are the commands that are used to enable the BGP routing protocol process, establish neighbors, and advertise local routes. This lesson concludes with basic commands that network administrators can use to monitor the BGP configuration.

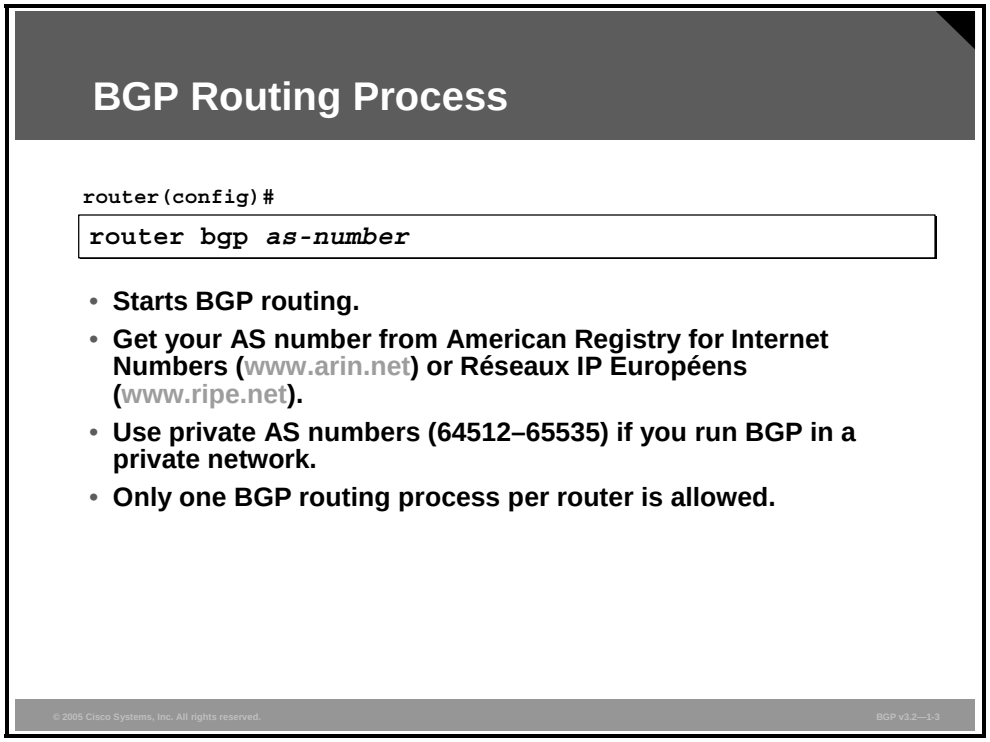
Objectives

Upon completing this lesson, you will be able to configure a router for BGP. This ability includes being able to meet these objectives:

- Identify the Cisco IOS command that is required to configure the BGP routing process
- Identify the Cisco IOS commands that are required to configure external BGP neighbors
- Identify the Cisco IOS commands that are required to configure the basic timers that are used in BGP
- Identify the Cisco IOS command that is required to configure MD5 authentication for BGP
- Identify the commands that are required to announce local networks in BGP
- Describe BGP route redistribution, including the commands that are required to configure BGP route redistribution
- Describe the classless behavior of BGP and identify the Cisco IOS command that is required to configure BGP for classless operation
- Describe BGP route aggregation, including the Cisco IOS commands that are required to configure basic BGP route aggregation
- Describe the BGP Conditional Route Injection feature
- Describe the BGP Support for TTL Security Check feature
- Determine when BGP route aggregation is not appropriate in multihomed topologies

BGP Routing Process

This topic describes the command that is required to initially configure the BGP routing process on a Cisco IOS router.



router bgp

To configure the BGP routing process, use the **router bgp** global configuration command.

- **router bgp** *as-number*

To remove a routing process, use the **no** form of this command.

- **no router bgp** *as-number*

Syntax Description

Parameter	Description
as-number	Number of an autonomous system (AS) that identifies the router to other BGP routers and tags the routing information that is passed along

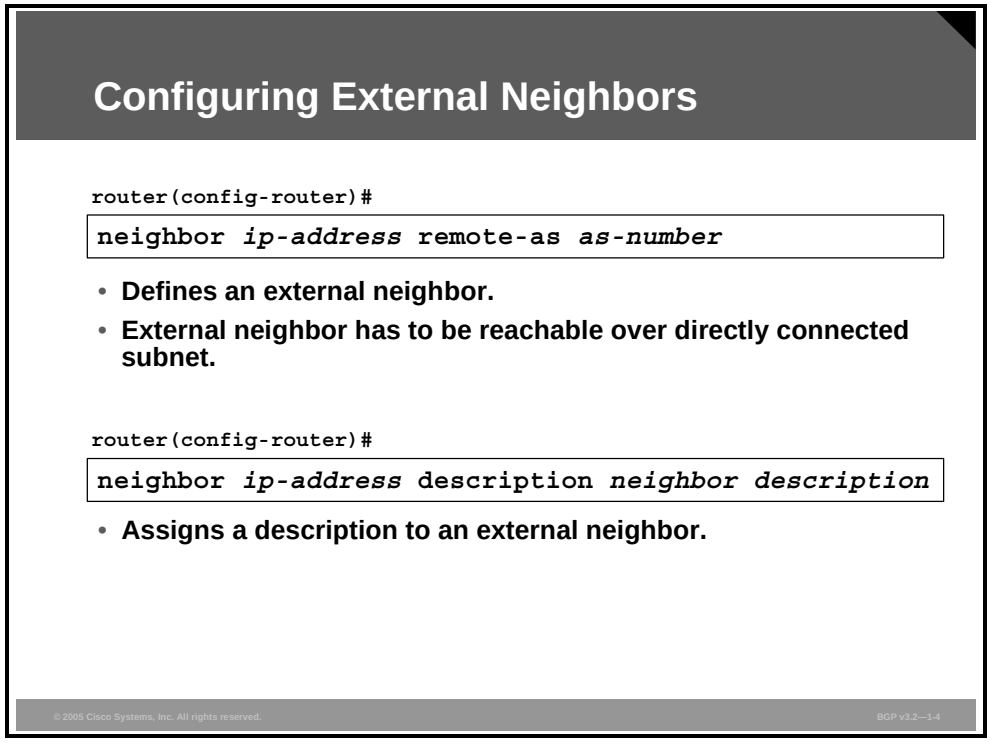
This command starts the BGP routing process in the router. There can be, at most, one BGP process in a router. It must be assigned the local AS number.

The AS number is a 16-bit unsigned integer. It must uniquely identify the AS among all routers that are exchanging BGP routing information, either directly or indirectly. This requirement means that the AS numbers must be unique when BGP information is exchanged with the Internet.

The AS number can be a public AS number (ranging from 1 to 64511) that is assigned by an Internet registry (American Registry for Internet Numbers [ARIN]: www.arin.net or Réseaux IP Européens [RIPE]: www.ripe.net), or a private AS number (ranging from 64512 to 65535). Private AS numbers are never propagated onto the public Internet.

Configuring External Neighbors

This topic describes the commands that are required to configure external BGP neighbors on a Cisco router.



BGP does not automatically discover neighbors. They have to be explicitly configured. The local router will try to connect to the indicated IP address and also accept incoming connection attempts from the indicated IP address.

The first attribute that you must configure with a new neighbor is the remote AS number in which the neighbor is taking part. When the TCP session is established between BGP routers, the configured remote AS number is verified by each router with the exchange of BGP Open messages.

You may optionally configure other attributes with the neighbor. Do this on successive configuration lines, referring to the same neighbor IP address but indicating different attributes. With the **neighbor description** command, a description (text string) can be entered that describes the neighbor.

neighbor remote-as

To add an entry to the BGP neighbor table, use the **neighbor remote-as** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **remote-as** *number*

To remove an entry from the table, use the **no** form of this command.

- **no neighbor** {*ip-address* | *peer-group-name*} **remote-as** *number*

Syntax Description

Parameter	Description
ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
number	AS to which the neighbor belongs

neighbor description

To associate a description with a neighbor, use the **neighbor description** router configuration command.

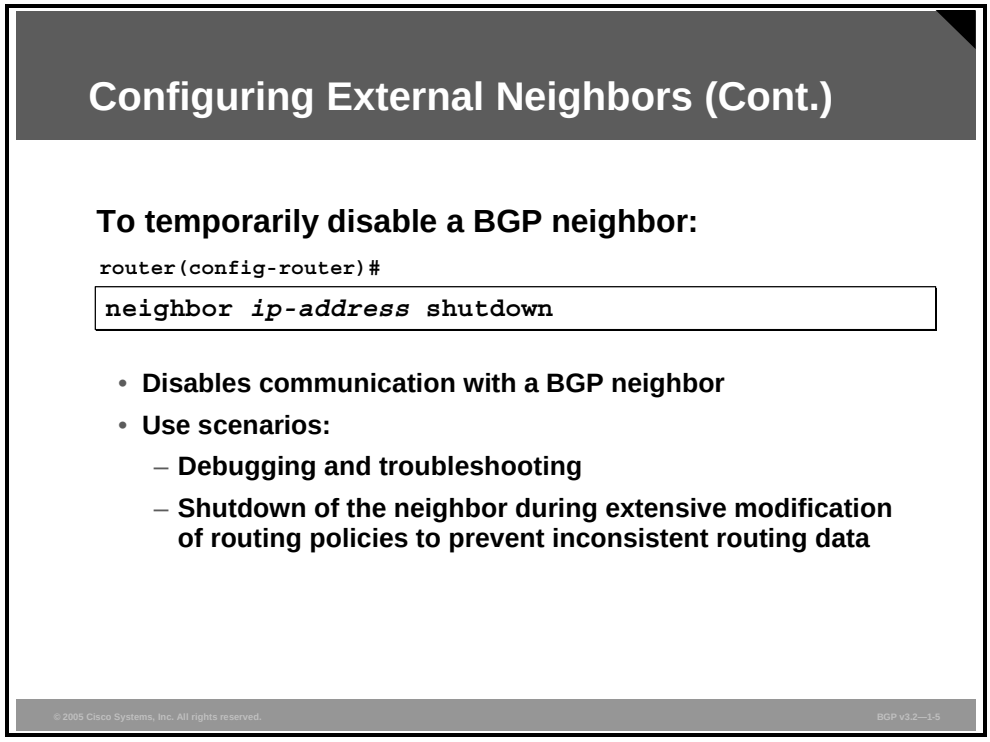
- **neighbor** {*ip-address* | *peer-group-name*} **description** *text*

To remove the description, use the **no** form of this command.

- **no neighbor** {*ip-address* | *peer-group-name*} **description** *text*

Syntax Description

Parameter	Description
ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
text	Text (up to 80 characters) that describes the neighbor



neighbor shutdown

To disable a neighbor, use the **neighbor shutdown** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **shutdown**

To re-enable the neighbor or peer group, use the **no** form of this command.

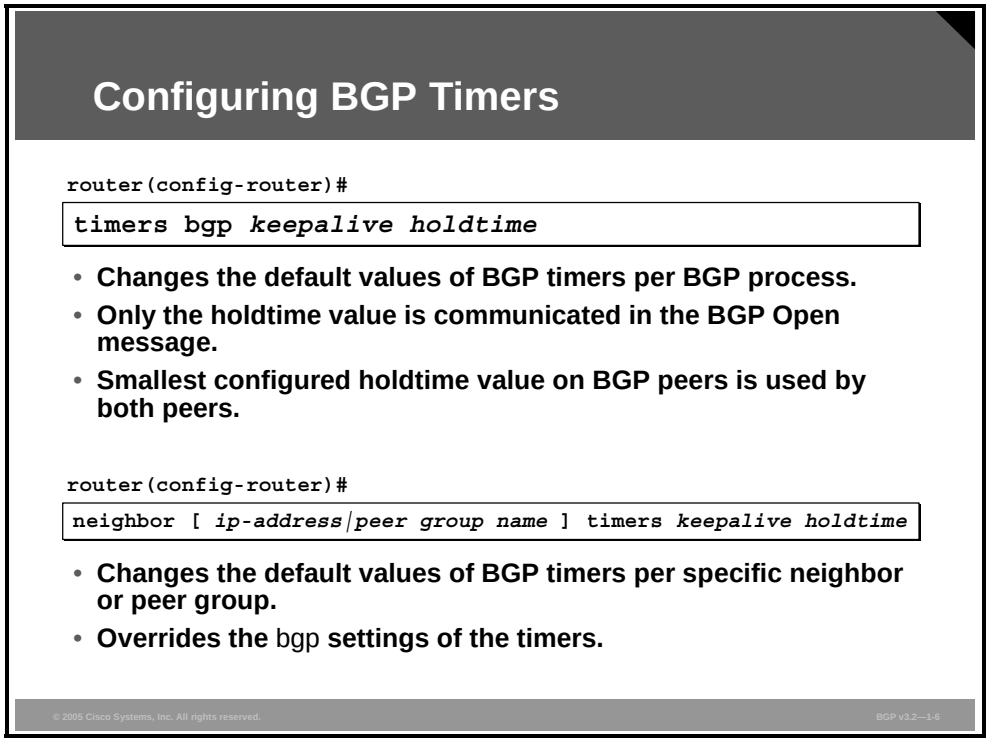
- **no neighbor** {*ip-address* | *peer-group-name*} **shutdown**

Syntax Description

Parameter	Description
ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group

Configuring BGP Timers

This topic describes the commands that are required to modify the default keepalive and holdtime timers in BGP for the BGP process or for the TCP session between BGP neighbors.



Changing the BGP default holdtime and keepalive timers is usually not recommended. The defaults (keepalive: 60 seconds; holdtime: 180 seconds) should work fine in most situations. If for any reason a faster BGP response to a peer down event is needed (for example, in scenarios where multiple paths toward destinations are available), the neighbor timers on the router can be reduced. This reduction will result in a faster detection of a lost peer and faster switching to the alternate path in the BGP table, thus improving convergence.

A BGP router with an expired holdtime (no BGP traffic was received within the holdtime interval) sends a notification to its BGP peer, notifying it as to the reason for closing the session. The BGP router on which the holdtime has expired moves the inactive peer into the Idle state. After a certain time interval, determined by auto-enable and connection timers, a BGP router again tries to reconnect to the previously disconnected BGP peer and will also accept connection attempts from that peer.

timers bgp

To adjust BGP network timers, use the **timers bgp** router configuration command.

- **timers bgp** *keepalive holdtime*

To reset the BGP timing defaults, use the **no** form of this command.

- **no timers bgp** *keepalive holdtime*

Syntax Description

Parameter	Description
keepalive	Frequency (in seconds) with which the Cisco IOS software sends keepalive messages to its peer. The default is 60 seconds.
holdtime	Interval (in seconds) after not receiving a keepalive message that the software declares a peer dead. The default is 180 seconds.

neighbor timers

To set the timers for a specific BGP peer or peer group, use the **neighbor timers** router configuration command. This command overrides the values that have been set by the **timers bgp** command.

- **neighbor** [*ip-address* | *peer-group-name*] **timers** *keepalive holdtime*

To clear the timers for a specific BGP peer or peer group, use the **no** form of this command.

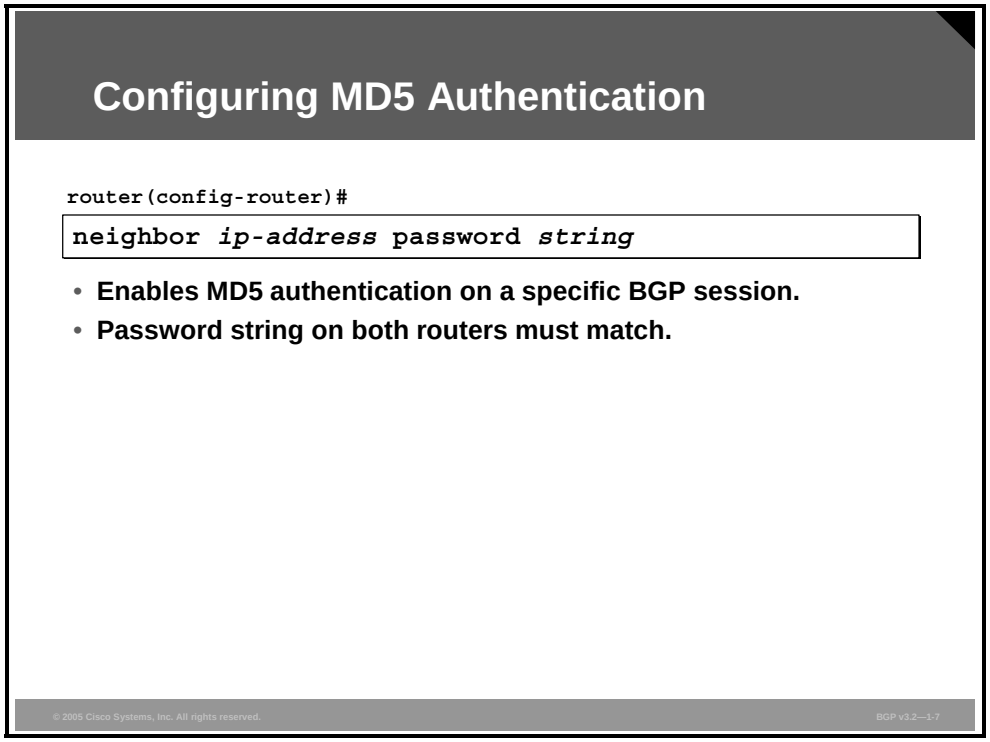
- **no neighbor** [*ip-address* | *peer-group-name*] **timers** *keepalive holdtime*

Syntax Description

Parameter	Description
ip-address	IP address of neighbor.
peer-group-name	Name of a BGP peer group.
keepalive	Frequency (in seconds) with which the Cisco IOS software sends keepalive messages to its peer. The default is 60 seconds.
holdtime	Interval (in seconds) after not receiving a keepalive message that the software declares a peer dead. The default is 180 seconds.

Configuring MD5 Authentication

This topic describes the command that is required to configure Message Digest 5 (MD5) authentication on a session between BGP neighbors.



neighbor password

To enable MD5 authentication on a TCP connection between two BGP peers, use the **neighbor password** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **password** *string*

To disable this function, use the **no** form of this command.

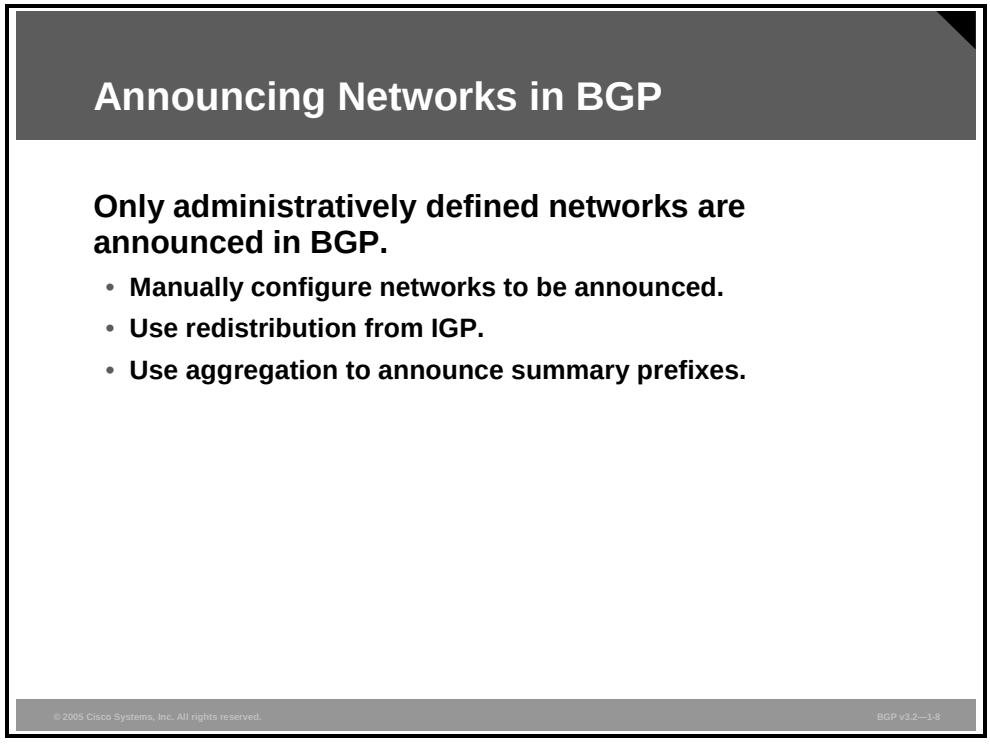
- **no neighbor** {*ip-address* | *peer-group-name*} **password**

Syntax Description

Parameter	Description
ip-address	IP address of neighbor.
peer-group-name	Name of a BGP peer group.
string	Case-sensitive password of up to 80 characters. The first character cannot be a number. The string can contain any alphanumeric characters, including spaces. You cannot specify a password in the format “number-space-anything.” The space after the number causes problems.

Announcing Networks in BGP

This topic describes the Cisco IOS commands that are required to announce local networks to other BGP neighbors.



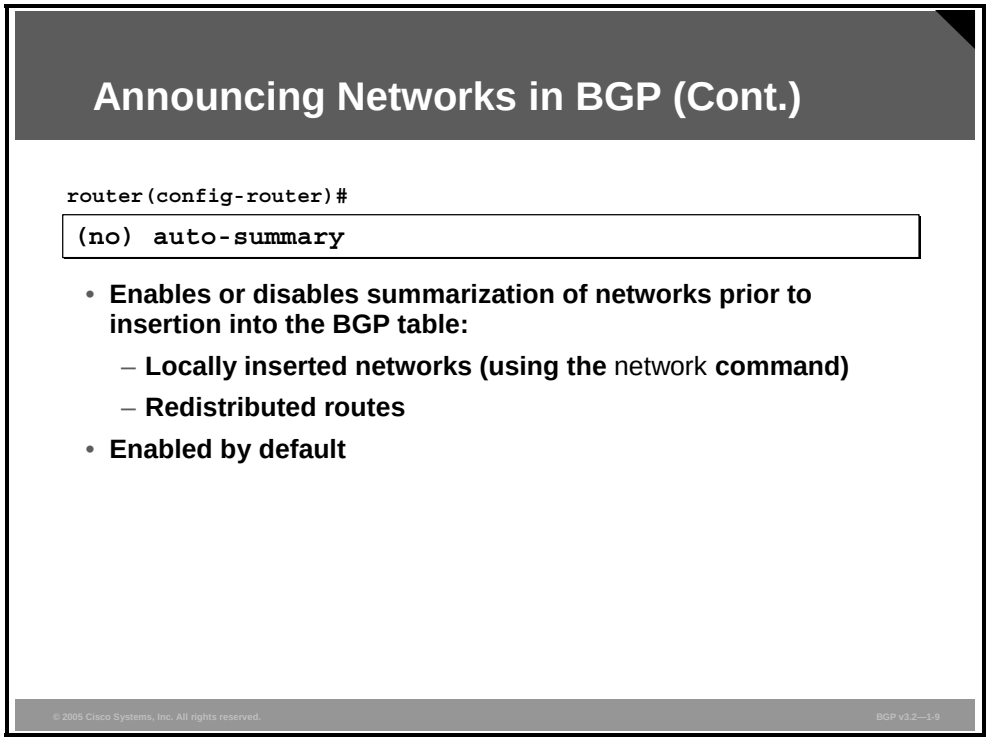
Before any local routing information can be injected by a router into its BGP table for advertising to other BGP routers, some basic configuration is required.

There are two ways to do this configuration:

- List the network numbers that are candidates to be advertised using the **network** configuration command. If any of the listed networks are reachable by the local router, according to its routing table, then the network is injected as a route into the BGP table.
- Redistribute routing information that has been learned by other routing protocols into the BGP table. You can use the Interior Gateway Protocol (IGP) that is used within the AS. Any route that is known by the local IGP can be injected into the BGP table using route redistribution between the IGP and BGP on the local router.

A router can also introduce new routing information into the BGP table by summarizing routes already there. This activity is called route aggregation and also requires configuration.

Any route that is introduced by the router into the BGP table will appear as a new route. The AS-path attribute for such a route will be empty, indicating a local route. The AS path changes later as the route passes AS boundaries.



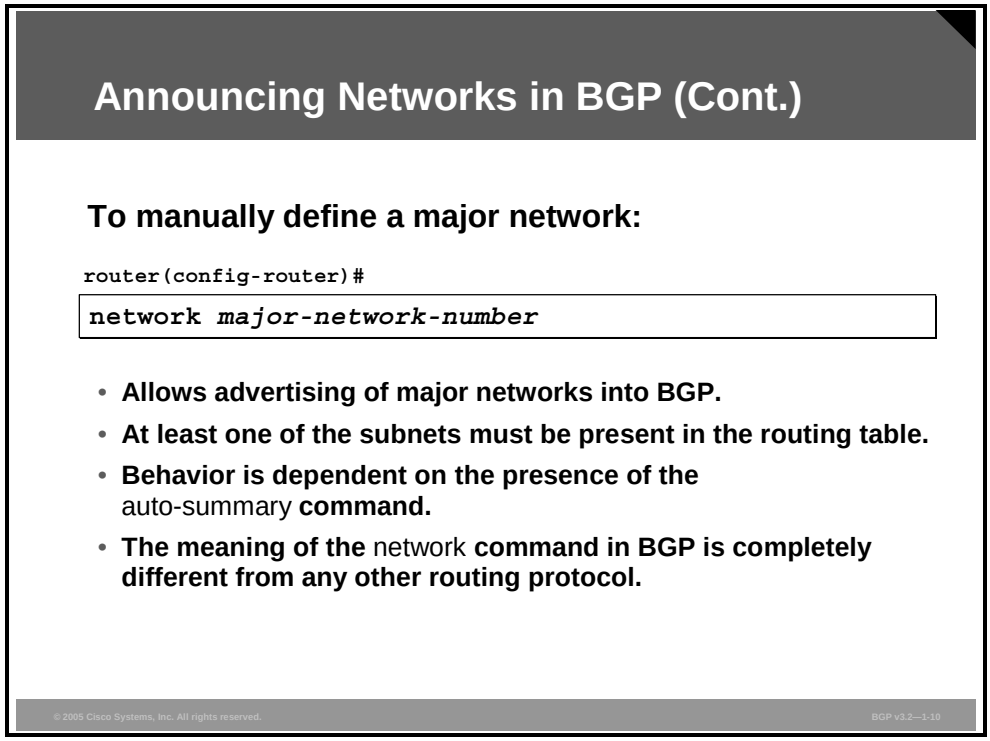
When the router is configured to locally announce routes into BGP, the behavior of the **network** command varies depending on whether automatic summarization is enabled or disabled. When automatic summarization is enabled, the command summarizes locally originated BGP networks to their classful boundaries. By default, automatic summarization is enabled for BGP.

When a subnet exists in the routing table and the following three conditions are satisfied, then any subnet (component route) of that classful network in the local routing table will prompt BGP to install the classful network into the BGP table:

- A classful network statement for that network exists in the routing table.
- A classful mask has been configured on that network statement.
- Automatic summarization is enabled.

When automatic summarization is disabled, the routes that are introduced locally into the BGP table are not summarized to their classful boundaries.

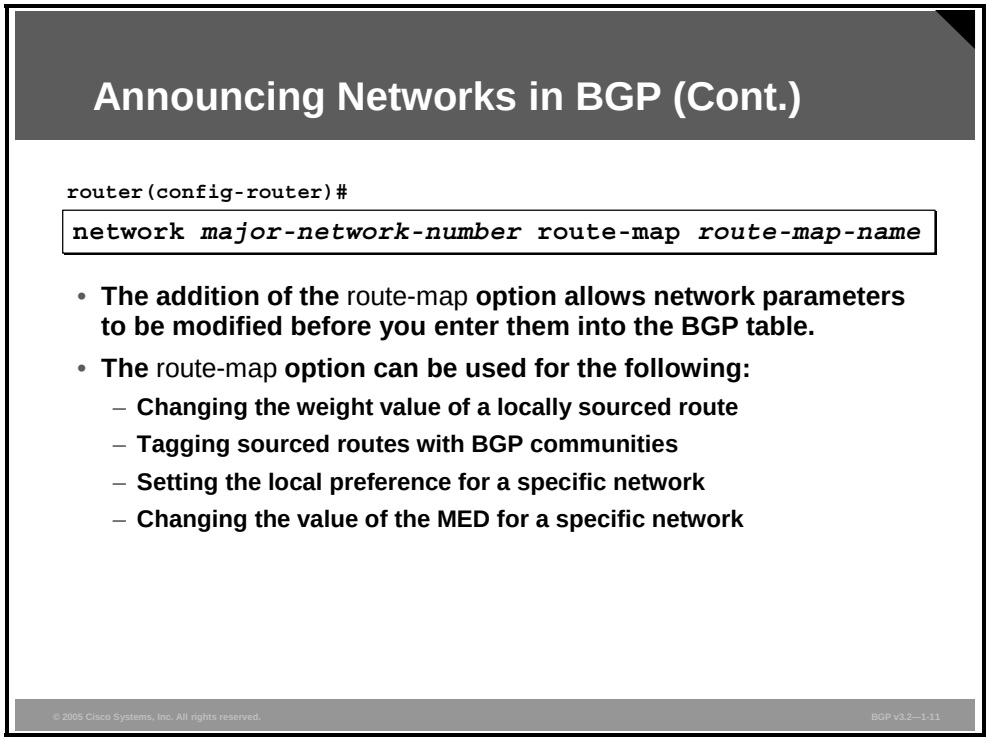
The BGP **auto-summary** command is also responsible for the behavior of the redistribution procedure in BGP. When the command is enabled, all redistributed subnets will be summarized to their classful boundaries in the BGP table. When it is disabled, all redistributed subnets will be present in their original form in the BGP table.



To specify the networks to be advertised by the BGP routing process, use the **network** router configuration command. To remove an entry, use the **no** form of this command.

Note	The meaning of the **network** command in BGP is radically different from the meaning of the command in other routing protocols. In all other routing protocols, the **network** command indicates interfaces over which the routing protocol will be run. In BGP, it indicates only which routes should be injected into the BGP table on the local router. Also, BGP never runs over individual interfaces; it is run over TCP sessions with manually configured neighbors.

The **network** command with no **mask** option uses the classful approach to insert a major network into the BGP table. At least one subnet of the specified major network needs to be present in the IP routing table to allow BGP to start announcing the major network as a BGP route. If automatic summarization is disabled, an exact match is required.



When the router is configured to insert routes into the BGP table, the default attributes of locally sourced routes can be modified with the inclusion of the **route-map** option in the basic **network** command.

The attached route-map can change the following attributes of locally sourced networks with the **network** command:

- **Weight (default value = 32768):** The weight attribute is a special Cisco attribute that is used in the path selection process when there is more than one route to the same destination. Because weight is considered before local preference in BGP route selection, locally sourced routes are always preferred, unless the weight value is modified.
- **Community (default value = nonexistent):** Used for tagging routes at their source.
- **Local preference (default value = 100):** Used for AS-wide BGP best-path selection.
- **Multi-exit discriminator (MED) (default value = 0):** Used for return-path selection in topologies where multiple exit points to the same neighbor AS exist.

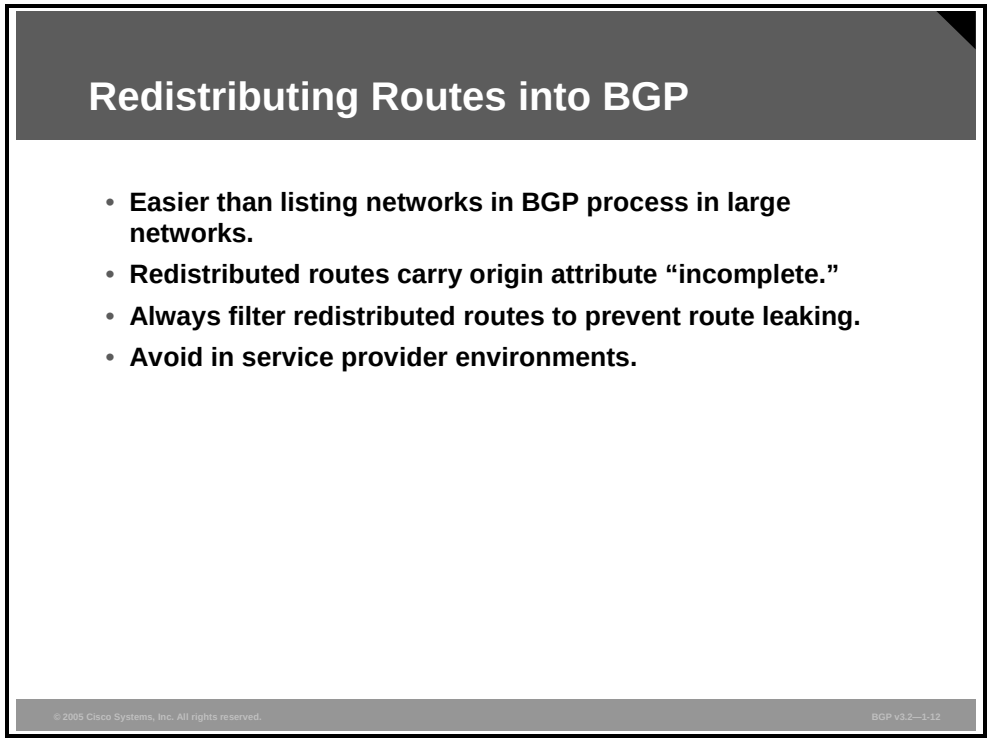
Example: Announcing Networks in BGP

If a subnet existing in the routing table is 75.75.75.0/24, and network 75.0.0.0 is configured under the **router bgp** command (assuming that automatic summarization is enabled), BGP will introduce the classful network 75.0.0.0/8 in the BGP table. If the following three conditions are not all met, then BGP will not install any entry in the BGP table unless there is an exact match in the IP routing table:

- A classful network statement for the network exists in the routing table.
- A classful mask has been configured on that network statement.
- Automatic summarization is enabled.

Redistributing Routes into BGP

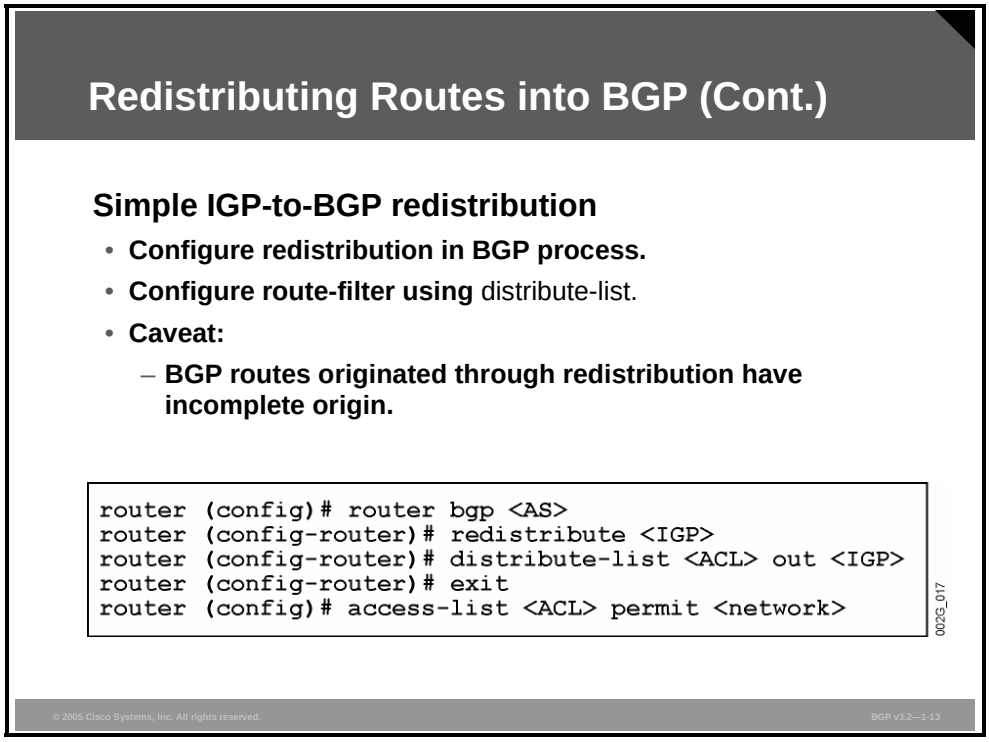
This topic describes route redistribution in BGP and identifies the Cisco IOS commands that are required to configure BGP route redistribution.



There are two alternatives for injecting local routes into the BGP table: list them using the **network** command or redistribute them. Listing the routes gives you total control over networks that could possibly be advertised by BGP. This option is very desirable for multihomed customers or Internet service providers (ISPs). On the other hand, this approach requires a lot of configuration commands that could be hard to maintain.

If there are a lot of networks to be advertised, and BGP is used primarily to achieve scalability, not routing security (for example, in enterprise networks), it could be easier to let the local IGP find the routes and then redistribute them into BGP. However, this approach introduces the risk that the IGP may find some networks that are not supposed to be advertised. Private network numbers, such as network 10.0.0.0/8, are often used within an AS for various reasons but must never be advertised out to the Internet. Careful filtering must be done to prevent unintentional advertising.

When the router injects a route that is listed with a **network** command into its BGP table, the origin code is set to “IGP.” If the route is injected into the BGP table through redistribution, the origin code is set to “unknown/incomplete.”



Routes redistributed into BGP will carry the origin attribute “incomplete.” In most cases this situation does not jeopardize BGP functionality. It could pose a problem if the route selection process has to decide on the best route toward a particular destination based on the MED attribute. In the case of receiving two routes, one with the “IGP” origin (inserted with the **network** command), and another one with the “incomplete” origin, the first route would always be selected, no matter what value the MED attribute is set to (according to the BGP route selection rules).

redistribute (IP)

To redistribute routes from one routing process into another routing process, use the **redistribute** router configuration command.

- **redistribute** *protocol* [*process-id*] {**level-1** | **level-1-2** | **level-2**} [**metric** *metric-value*] [**match** {**internal** | **external 1** | **external 2**}] [tag *tag-value*] [**route-map** *map-tag*] [**weight** *weight*] [**subnets**]

To disable redistribution, use the **no** form of this command.

- **no redistribute** *protocol* [*process-id*] {**level-1** | **level-1-2** | **level-2**} [**metric** *metric-value*] [**match** {**internal** | **external 1** | **external 2**}] [tag *tag-value*] [**route-map** *map-tag*] [**weight** *weight*] [**subnets**]

Syntax Description

Parameter	Description
protocol	Source protocol from which routes are being redistributed. It can be one of the following keywords: **bgp, isis, ospf, eigrp, rip.**
process-id	(Optional) For **bgp**, **egp**, or e**igrp**, this is an AS number, which is a 16-bit decimal number. For **isis**, this is an optional tag that defines a meaningful name for a routing process. You can specify only one Intermediate System-to-Intermediate System (IS-IS) process per router. Creating a name for a routing process means that you use names when configuring routing. For **ospf**, this is an appropriate Open Shortest Path First (OSPF) process ID from which routes are to be redistributed. This ID identifies the routing process. This value takes the form of a nonzero decimal number. For **rip**, no process ID value is needed.
level-1	For IS.
level-1-2	For IS.
level-2	For IS.
metric *metric-value*	(Optional) Metric that is used for the redistributed route. If a value is not specified for this option, and no value is specified, the **default metric is used**.
match {**internal** | **external 1** | **external 2**}	(Optional) For OSPF, the criterion by which OSPF routes are redistributed into other routing processes. It can be one of the following: ■ **internal:** Routes that are internal to a specific AS. ■ **external 1:** Routes that are external to the AS but are imported into OSPF as a type 1 external route. ■ **external 2:** Routes that are external to the AS but are imported into OSPF as a type 2 external route.
route-map	(Optional) The route-map should be interrogated to filter the importation of routes from this source routing protocol to the current routing protocol. If not specified, all routes are redistributed. If this keyword is specified, but no route-map tags are listed, no routes will be imported.
map-tag	(Optional) Identifier of a configured route-map.
weight *weight*	(Optional) Network weight when you are redistributing into BGP. An integer from 0 to 65535.
subnets	Indicates that not only networks with a natural mask should be redistributed but also subnets.

distribute-list out (IP)

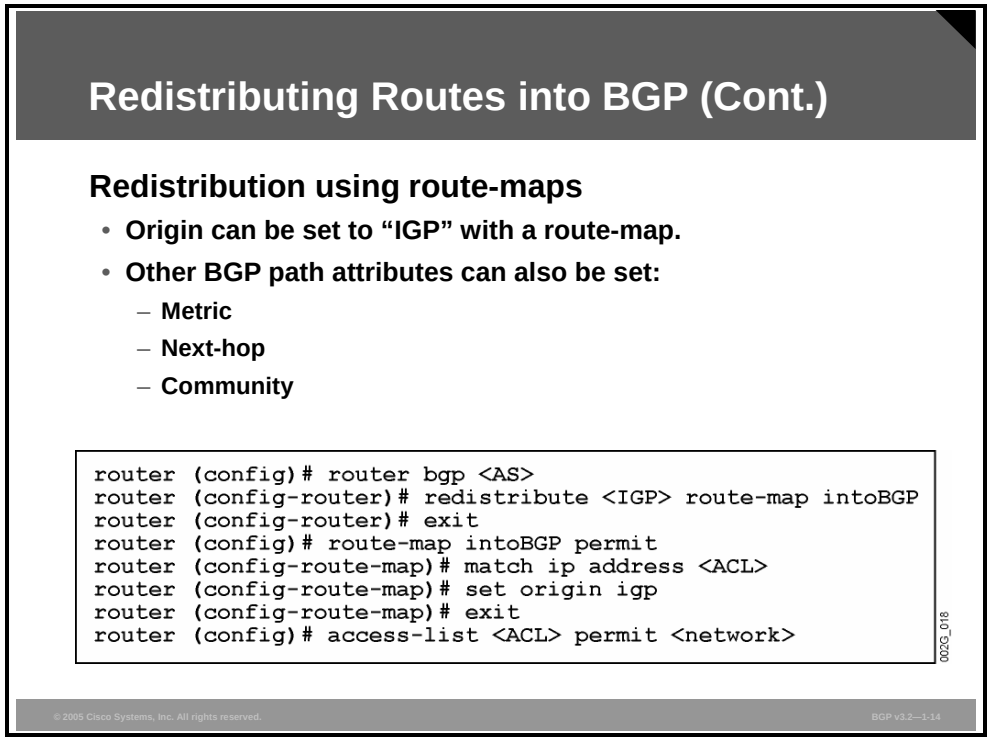
To suppress networks from being advertised in updates, use the **distribute-list out** router configuration command with *routing-process* specified.

- **distribute-list** {*access-list-number* | *access-list-name*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

To cancel this function, use the **no** form of this command.

- **no distribute-list** {*access-list-number* | *access-list-name*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

The access-list referred to by the **distribute-list** command permits the routes that should be redistributed.



Route-maps can be configured on the router to filter updates and modify various attributes. A configured route-map can be applied to routes being redistributed from the IGP.

Only the routes permitted by the route-map will be redistributed. Using the **set** command in the route-map, you can modify specific path attributes that are attached to the redistributed routes. Thus, only selected routes will be advertised, and they will have the desired attribute values.

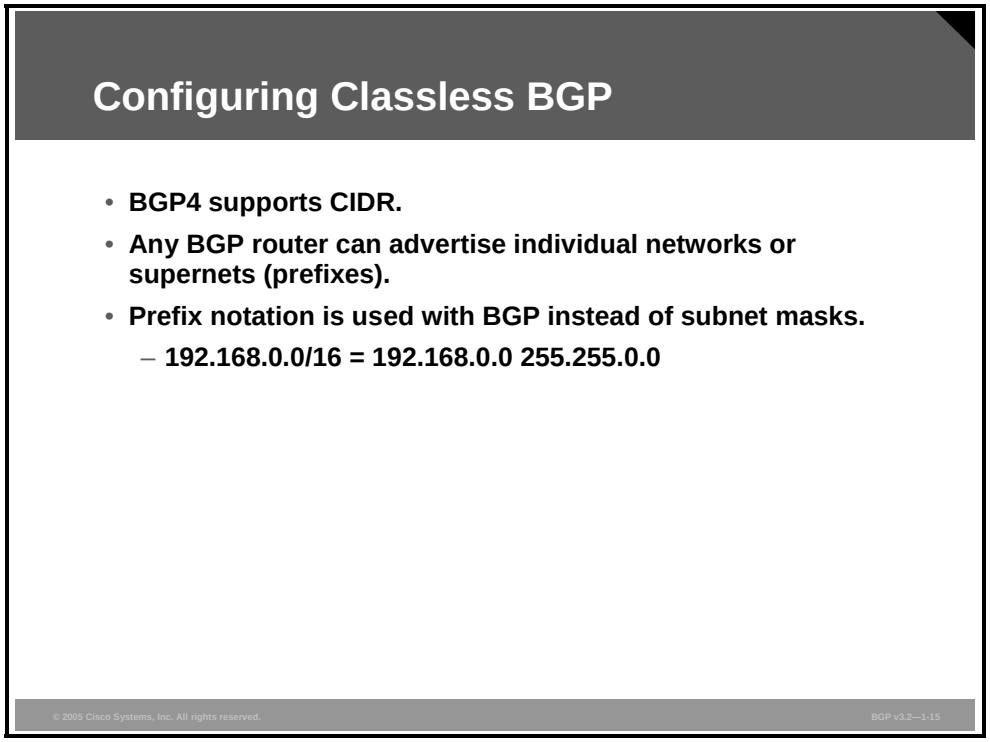
The route-map must be given a name. This name is a case-sensitive string, which is used when you are referring to the route-map. Any string could be used, but a meaningful name is suggested.

Use the **route-map** global configuration command and the **match** and **set** route-map configuration commands to define the conditions for redistributing routes. Each repetition of the **route-map** command has a list of **match** and **set** commands that are associated with it. The **match** commands specify the match criteria—the conditions under which redistribution is allowed for the current **route-map** command. The **set** commands specify the set actions—the particular redistribution actions to perform if the criteria enforced by the **match** commands are met.

When you are passing routes through a route-map, it can have several parts. Any route that does not match at least one match clause relating to a **route-map** command will be ignored; that is, the route will not be advertised. If you want to modify only some data, you must configure a second route-map section with an explicit match specified.

Configuring Classless BGP

This topic describes the classless behavior of BGP and the command that is required to advertise a classless BGP supernet prefix.

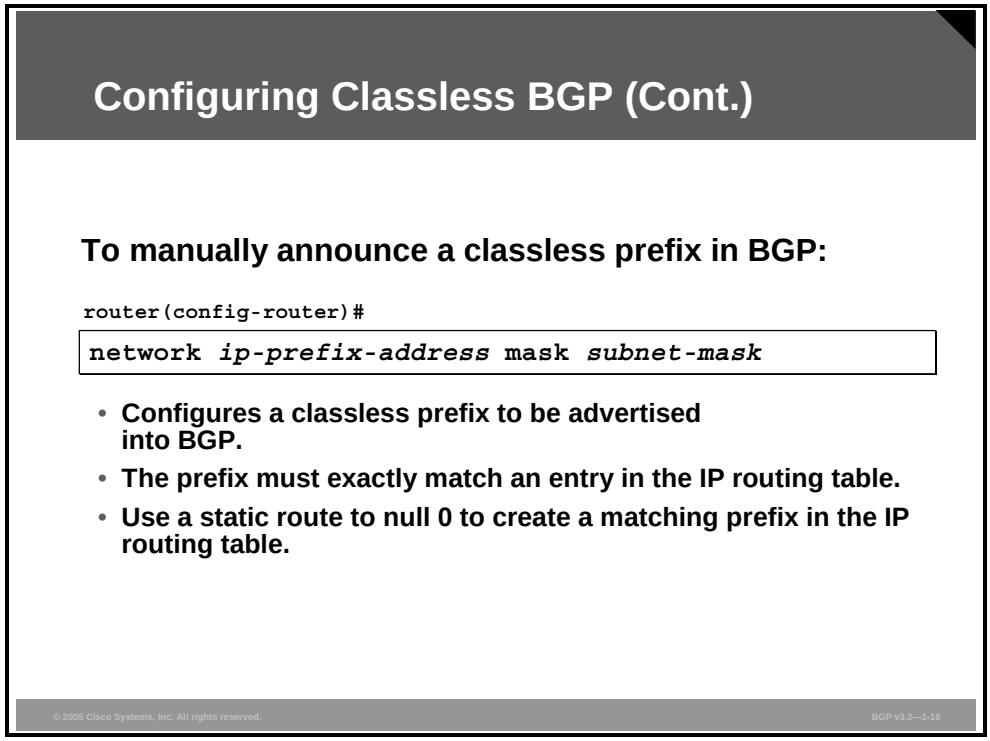


BGP version 4 (BGP4) is a classless protocol, meaning that its routing updates include the IP address and the subnet mask. The combination of the IP address and the subnet mask is called an IP prefix. An IP prefix can be a subnet, a major network, or a supernet.

BGP uses prefix notation (address/number of bits) to display IP prefixes. The number following the slash (/) in the 192.168.0.0/16 notation in the figure refers to the number of bits in the subnet mask being set to 1. The subnet mask 255.255.0.0 starts with 16 consecutive bits set to 1, and the rest of the bits set to 0.

As another example, the subnet 172.16.1.0 with mask 255.255.255.0 can be written using the prefix notation as 172.16.1.0/24.

When classless prefix notation is used, an old class A network, for example, 10.0.0.0, with the natural mask, is written as 10.0.0.0/8. A class B network, 172.17.0.0 with natural mask, is written as 172.17.0.0/16, and a class C network, 192.168.1.0 with natural mask, is written as 192.168.1.0/24.



To advertise classless networks into BGP (a subnet or a supernet), you can use the **network** command with the **mask** keyword and the subnet mask specified. When an exact match is not found in the IP routing table (for example, when you are creating a summary or when you are advertising only a part of your address space), a matching prefix has to be manually configured on the router in the form of a static route pointing to the null 0 interface; otherwise, the advertisement will not succeed.

network (BGP)

To specify the networks to be advertised by the BGP routing process, use the **network** router configuration command.

- **network** *network-number* [**mask** *network-mask*]

To remove an entry, use the **no** form of this command.

- **no network** *network-number* [**mask** *network-mask*]

Syntax Description

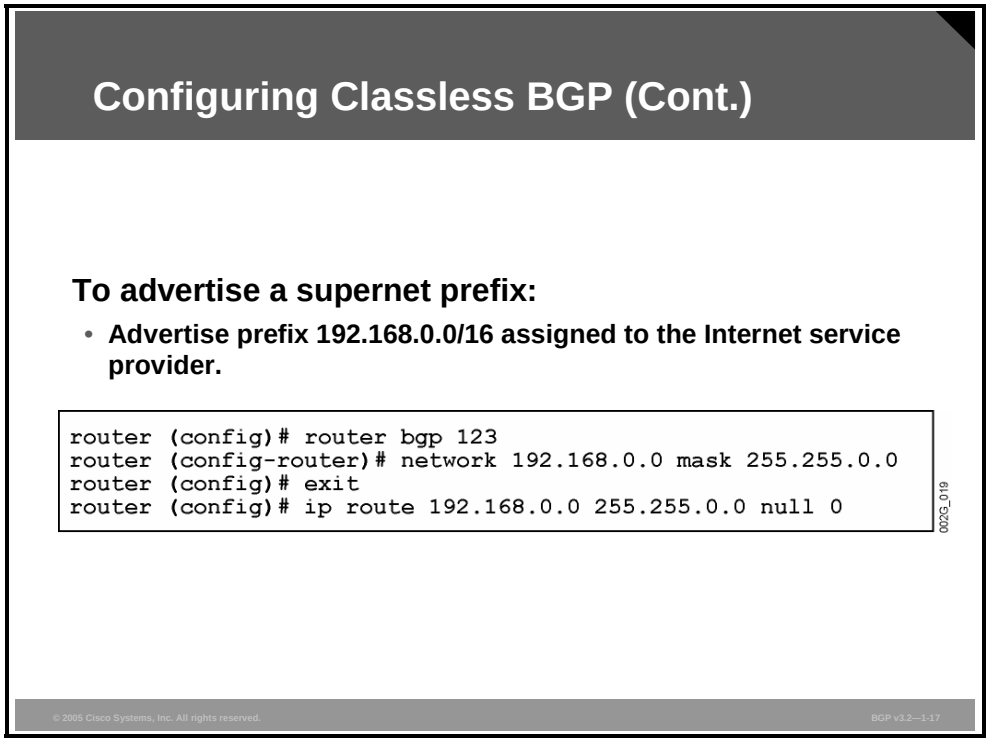
Parameter	Description
network-number	Network that BGP will advertise
mask	(Optional)
network-mask	(Optional) Network mask address

If the keyword **mask** and the subnet mask are omitted, the network is assumed to have its natural mask according to the network class. In this case, the route will still be injected into the BGP table on the router if there is any subnet of the major network that is reachable according to the routing table.

If the network mask is specified, the behavior changes slightly, and it is required that an exact match of network number and subnet mask appear in the routing table before the route is injected into the BGP table.

Example: Configuring Classless BGP

In this example, the IP address space 192.168.0.0/16 is assigned to a service provider, and the service provider would like that address space to be constantly advertised by BGP.



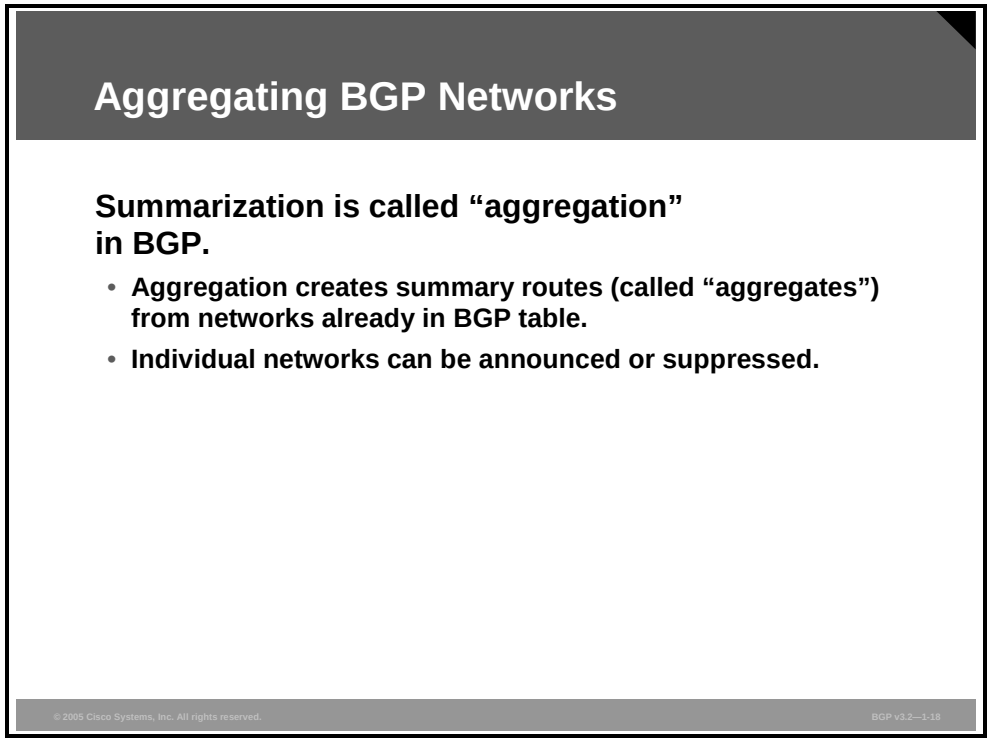
The **network** command with the **mask** option tells BGP that 192.168.0.0/16 is a candidate for being advertised. The **mask** keyword and the mask 255.255.0.0 are required because the mask is not the natural one. However, before the candidate route is actually advertised, the router checks the routing table for an exact match (both network number and mask). It will always be found because there is a static route for it. This static route points to the null interface, which is always available.

The conclusion is that 192.168.0.0/16 will always be advertised by this router. All other BGP routers will use this information and forward any IP packets with the destination IP address in the interval 192.168.0.0 to 192.168.255.255 (inclusive) in the direction of this router. When those packets arrive, the router, in this example, must have more explicit routes to the different parts of the 192.168.0.0/16 address range. This need could be answered by the IGP, which is not shown in the configuration example.

If, however, an IP packet arrives with a destination address to which this router does not have a more explicit route, the static route will route the packet to the null interface, where it is dropped. This routing is a safety precaution that will prevent a routing loop, which might occur when route summaries are used in combination with default routing. If, for example, a packet arrives from the Internet to a subnet of 192.168.0.0/16, which is currently not reachable, the packet might otherwise follow the default route toward the Internet because there is no more explicit route. Of course, the packet would immediately be routed back again, and a routing loop would occur.

Aggregating BGP Networks

This topic describes route summarization in BGP. It also lists the configuration commands that are required to configure summary routes in BGP.



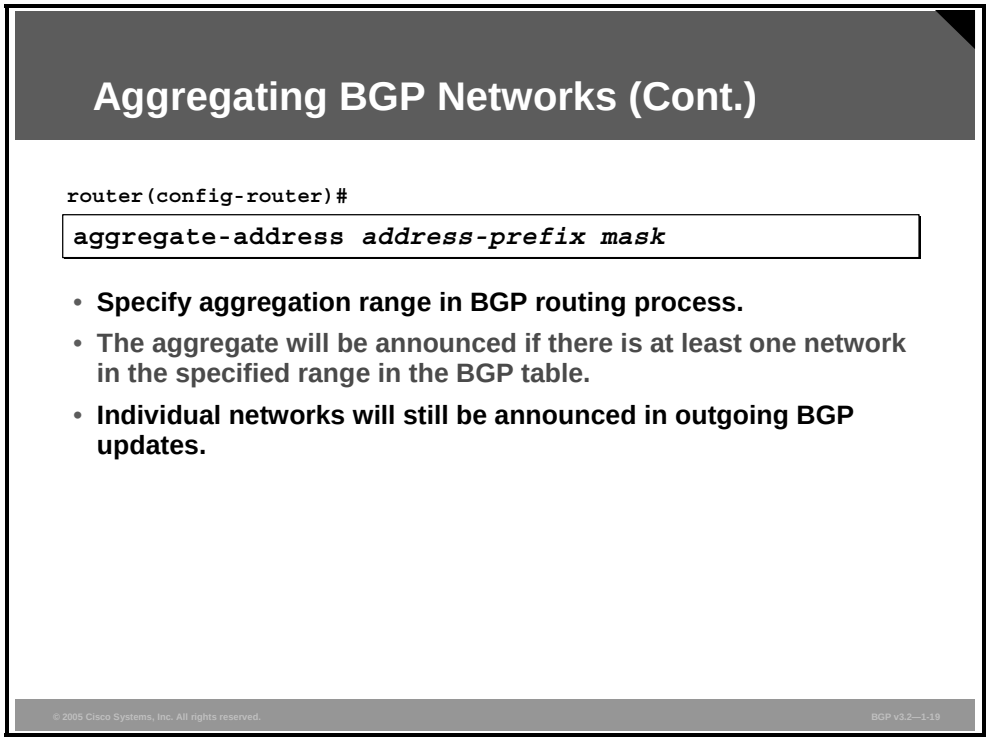
When the BGP table is already populated with routes that should be summarized, you must configure a router to do so. The summarization of BGP routes is called “aggregation.”

Use aggregation when a group of more specific routes has been injected into the BGP table at one stage but can be summarized at a later stage. The routes to be summarized could be IGP routes that have been redistributed into BGP. Before BGP advertises these routes to the rest of the network, an aggregation of the subnets into a larger announcement would be appropriate.

In some networks, more specific routes are injected into the BGP table by some routers, and aggregation is done in another router or even in another AS. This is called “proxy aggregation.”

When a router is configured to do aggregation, you must configure the route summary. If any route that is already in the BGP table is within the range that is indicated by the summary, then the summary route is also injected into the BGP table on the route and advertised to other routers. This action creates more information in the BGP table. To get any benefits from the aggregation, you must suppress the more specific routes that are covered by the route summary. This suppression is an option to the **aggregate** configuration command.

When you suppress the more specific routes through configuration, they are still present in the BGP table of the router doing the aggregation. However, because the routes are marked as suppressed, they are never advertised to any other router.



In this configuration command syntax, where the keyword **summary-only** is not used, both the route summary and the more specific routes will be advertised. This approach is generally not desirable. Therefore, suppression of individual routes, described next, is used in most cases.

aggregate-address

To create an aggregate entry in a BGP routing table, use the **aggregate-address** router configuration command.

- **aggregate-address** *address mask* [**as-set**] [**summary-only**] [**suppress-map** *map-name*][**advertise-map** *map-name*] [**attribute-map** *map-name*]

To disable this function, use the **no** form of this command.

- **no aggregate-address** *address mask* [**as-set**] [**summary-only**] [**suppress-map** *map-name*][**advertise-map** *map-name*] [**attribute-map** *map-name*]

Syntax Description

Parameter	Description
address	Aggregate address
mask	Aggregate mask
summary-only	(Optional) Suppresses more specific routes

Aggregating BGP Networks (Cont.)

An alternative method to configure aggregation:

```
router(config-router)#
aggregate-address address-prefix mask summary-only
```

- **Configure aggregation of BGP routes.**
- **Advertise only the aggregate and not the individual networks.**

Benefits:

- **Smaller BGP routing tables**
- **More stable internetworks (less route flapping)**

Drawback:

- **Problems with multihomed customers**

When the **summary-only** option is used, only the route summary will be advertised, not the more specific routes.

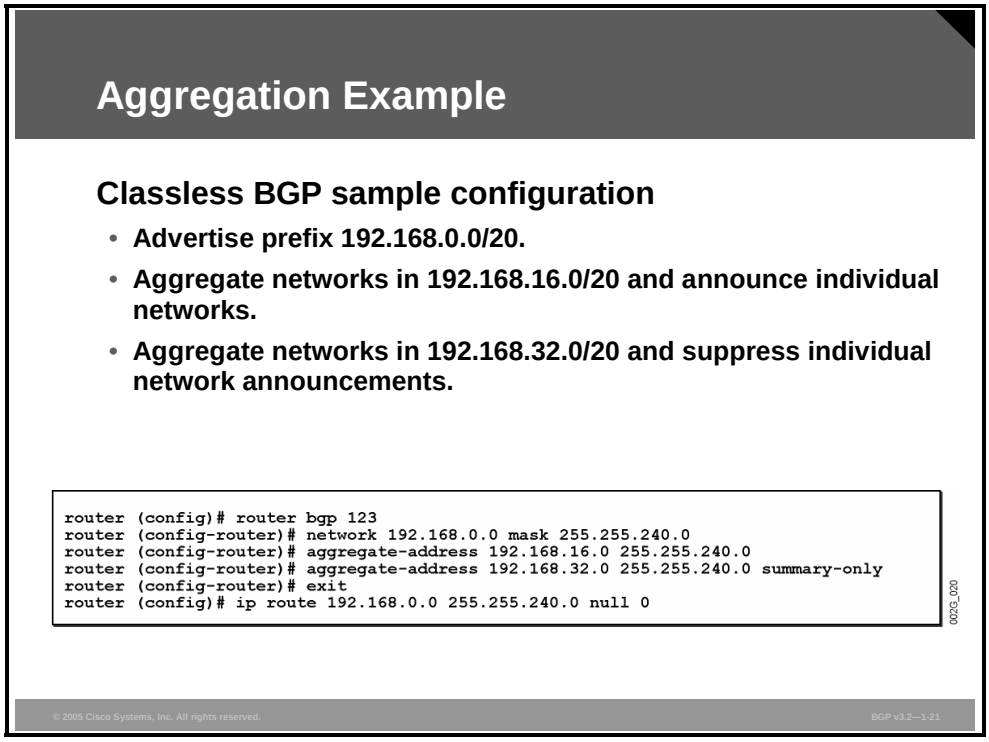
One of the benefits of this approach is that the rest of the routers will receive only one route instead of many more specific routes. It eases the burden on the other routers by reducing the amount of memory that is required to hold the BGP table.

Another benefit is that route flapping is reduced. The router doing the aggregation continues advertising the aggregate as long as there is at least one specific route within the range still available. If one of the more specific routes is lost but at least one remains, the aggregate itself is not lost. The flap of the more specific route is not visible to the rest of the network. This approach reduces the number of updates necessary and the CPU power that is required to process them.

However, all route summarization in any routing protocol causes a loss of granularity (that is, lack of more detailed network or subnet visibility). Suboptimal routing could be introduced when redundant paths are available to reach a group of networks that are advertised by a single route summary. Some of the networks could be more reachable via one of the paths, while others may be more reachable another way. From outside the immediate network, multiple paths may not be visible because only summary routes are advertised. Therefore, there is a risk that the least optimal path will be chosen.

Example: Aggregation

This example illustrates a classless BGP sample configuration.



The configuration example in the figure shows three different ways of advertising a route summary:

- The prefix 192.168.0.0/20 is always advertised. It is injected into the BGP table as a summary. The network statement makes it a candidate for being advertised. Because the mask is specified, an exact match in the routing table is a required condition before the route is injected into the BGP table. The matching route is inserted in the IP routing table by the static IP route statement to the null 0 interface.
- The prefix 192.168.16.0/20 is conditionally advertised. It is injected into the BGP table whenever there is a more specific route within the route summary range that is already in the BGP table. However, the more specific route is still advertised.
- The prefix 192.168.32.0/20 is also conditionally advertised. It is injected into the BGP table whenever there is a more specific route within the route summary range that is already in the BGP table. However, any more specific routes are suppressed and not advertised to any neighbors.

Aggregation Example (Cont.)

Viewing the BGP table

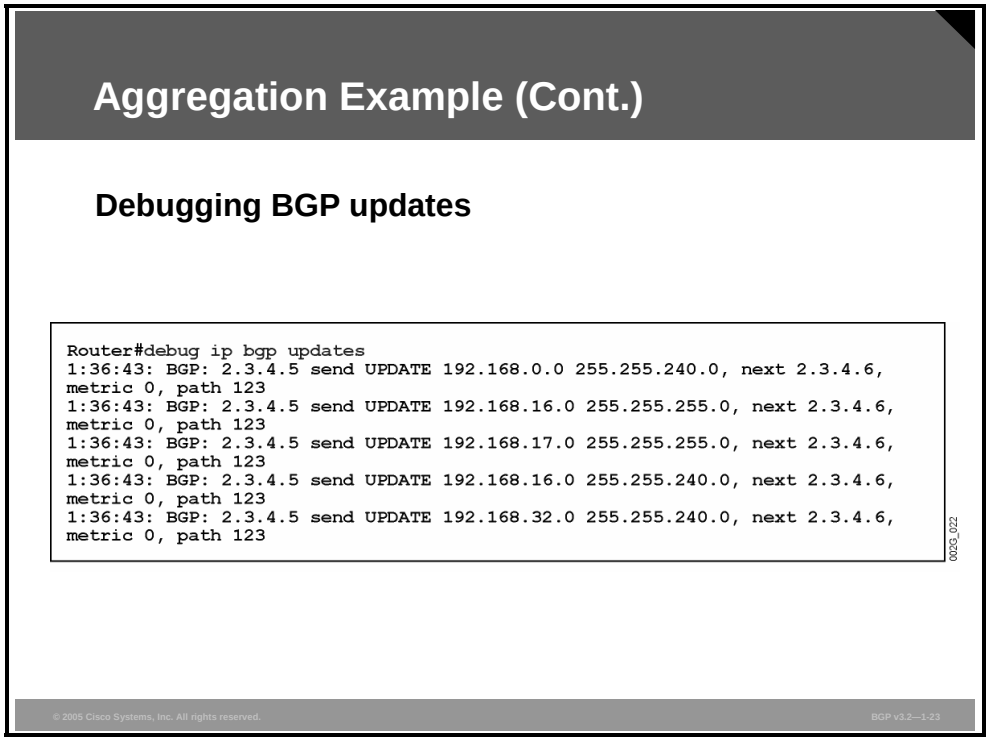
```
as123#show ip bgp
BGP table version is 16, local router ID is 1.2.3.4
Status codes: s suppressed, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop       Metric  LocPrf  Weight  Path
*> 1.0.0.0          0.0.0.0             0           32768  i
*  21.0.0.0         3.4.5.6                             0  37 21 i
*>                  2.3.4.5             0               0  21 i
*> 37.0.0.0         3.4.5.6             0               0  37 i
*                   2.3.4.5                             0  21 37 i
*> 192.168.0.0/20   0.0.0.0             0           32768  i
*> 192.168.16.0/20  0.0.0.0             0           32768  i
*> 192.168.16.0     0.0.0.0             0           32768  ?
*> 192.168.17.0     0.0.0.0             0           32768  ?
*> 192.168.32.0/20  0.0.0.0             0           32768  i
s> 192.168.32.0     0.0.0.0             0           32768  ?
s> 192.168.33.0     0.0.0.0             0           32768  ?
```

002G_021

The **show ip bgp** command prints the BGP table. As shown in the figure, all three prefixes are injected:

- The prefix 192.168.0.0/20 is always injected.
- The prefix 192.168.16.0/20 is injected because there is at least one more specific route within the summary range. In this case, both 192.168.16.0/24 and 192.168.17.0/24 are within the range. Nothing is changed with the more specific routes, so they are still advertised.
- The prefix 192.168.32.0/20 is injected because there is at least one more specific route within the summary range. In this case, both 192.168.32.0/24 and 192.168.33.0/24 are within the range. The more specific routes are marked as suppressed using the lowercase letter “s.” The “s” means that they are still present and available in the BGP table of the router, but they are not advertised on any BGP session.

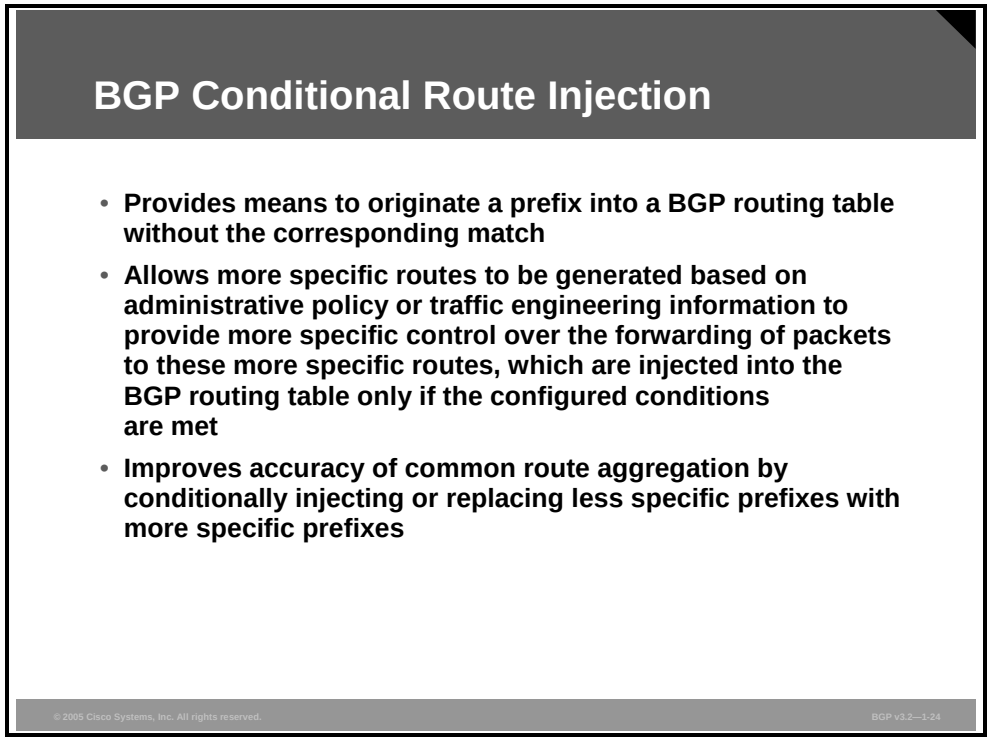
Note Because the prefixes 192.168.16.0/24, 192.168.17.0/24, 192.168.32.0/24, and 192.168.33.0/24 all have natural masks as applied to class C networks, the prefix length is not displayed in the **show ip bgp** printout. The network mask is, however, stored in the BGP table and sent on any BGP update.



The debug output shows the BGP updates that have been sent to a neighbor. All three route summary prefixes, 192.168.0.0/20, 192.168.16.0/20, and 192.168.32.0/20, are included in the updates. Also, the nonsuppressed more explicit routes, 192.168.16.0/24 and 192.168.17.0/24, are included in the update. However, the suppressed more explicit routes, 192.168.32.0/24 and 192.168.33.0/24, are never sent as updates on the BGP session.

BGP Conditional Route Injection

This topic describes the BGP Conditional Route Injection feature.



Routes that are advertised through the BGP are commonly aggregated to minimize the number of routes that are used and reduce the size of global routing tables. However, common route aggregation can obscure more specific routing information that is more accurate but not necessary to forward packets to their destinations. Routing accuracy is obscured by common route aggregation because a prefix that represents multiple addresses or hosts over a large topological area cannot be accurately reflected in a single route. Cisco IOS software provides several methods by which you can originate a prefix into BGP. The methods include redistribution and using the **network** or **aggregate-address** command. These methods assume the existence of more specific routing information (matching the route to be originated) in either the routing table or the BGP table.

The BGP Conditional Route Injection feature allows you to originate a prefix into a BGP routing table without the corresponding match. This feature allows more specific routes to be generated based on administrative policy or traffic engineering information to provide more specific control over the forwarding of packets to these more specific routes, which are injected into the BGP routing table only if the configured conditions are met. Enabling this feature allows you to improve the accuracy of common route aggregation by conditionally injecting or replacing less specific prefixes with more specific prefixes. Only prefixes that are equal to or more specific than the original prefix may be injected.

The BGP Conditional Route Injection feature is enabled with the **bgp inject-map exist-map** command. This command uses two route maps (inject-map and exist-map) to install one (or more than one) more specific prefix into a BGP routing table. The exist-map specifies the prefixes that the BGP speaker will track. The inject-map defines the prefixes that will be created and installed into the local BGP table.

bgp inject-map exist-map

To inject a more specific route into a BGP routing table, use the **bgp inject-map exist-map** command in address family or router configuration mode.

- **bgp inject-map** {*inject-map-name*} **exist-map** {*exist-map-name*}[**copy-attributes**]

To disable this function, use the **no** form of this command.

- **No bgp inject-map** {*inject-map-name*} **exist-map** {*exist-map-name*}[**copy-attributes**]

Syntax Description

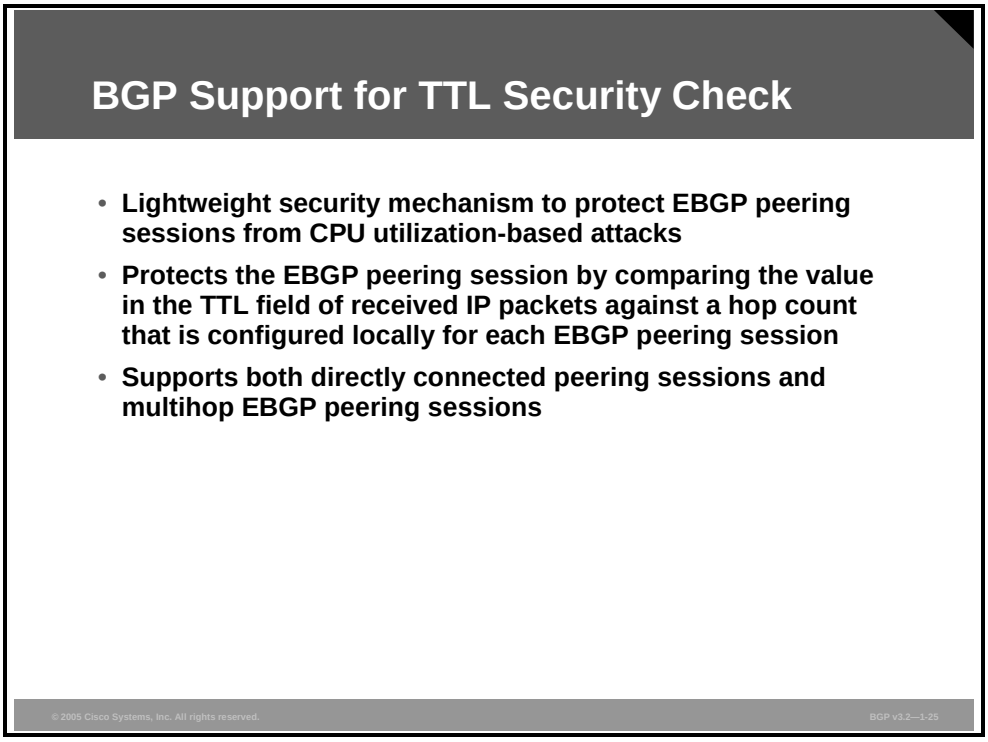
Parameter	Description
inject-map-name	Defines the prefixes that will be created and installed into the local BGP table
exist-map-name	Specifies the prefix that the BGP speaker will track
copy-attributes	(Optional) Configures the injected route to inherit the attributes of the aggregate route

This configuration example configures conditional route injection for the inject-map named ORIGINATE and the exist-map named LEARNED_PATH:

```
router bgp 109
 bgp inject-map ORIGINATE exist-map LEARNED_PATH
 !
 route-map LEARNED_PATH permit 10
  match ip address prefix-list ROUTE
  match ip route-source prefix-list ROUTE_SOURCE
 !
 route-map ORIGINATE permit 10
 set ip address prefix-list ORIGINATED_ROUTES
 set community 14616:555 additive
 !
 ip prefix-list ROUTE permit 10.1.1.0/24
 !
 ip prefix-list ORIGINATED_ROUTES permit 10.1.1.0/25
 ip prefix-list ORIGINATED_ROUTES permit 10.1.1.128/25
 !
 ip prefix-list ROUTE_SOURCE permit 10.2.1.1/32
```

BGP Support for TTL Security Check

This topic describes the BGP Support for TTL Security Check feature.



The BGP Support for TTL Security Check feature introduces a lightweight security mechanism to protect External Border Gateway Protocol (EBGP) peering sessions from CPU utilization-based attacks. These types of attacks are typically brute-force denial of service (DoS) attacks that attempt to disable the network by flooding the network with IP packets that contain forged source and destination IP addresses. This feature protects the EBGP peering session by comparing the value in the Time to Live (TTL) field of received IP packets against a hop count that is configured locally for each EBGP peering session. If the value in the TTL field of the incoming IP packet is greater than or equal to the locally configured value, the IP packet is accepted and processed normally. If the TTL value in the IP packet is less than the locally configured value, the packet is silently discarded and no Internet Control Message Protocol (ICMP) message is generated. This is designed behavior; a response to a forged packet is unnecessary.

Accurately forging the TTL count in an IP packet is generally considered to be impossible. It is possible to forge the TTL field in an IP packet header. However, accurately forging a packet to match the TTL count from a trusted peer is not possible unless the network to which the trusted peer belongs has been compromised.

This feature supports both directly connected peering sessions and multihop EBGP peering sessions. The BGP peering session is not affected by incoming packets that contain invalid TTL values. The BGP peering session remains open, and the router silently discards the invalid packet. The BGP session, however, can still expire if keepalive packets are not received before the session timer expires.

The BGP Support for TTL Security Check feature should be configured on each participating router. It provides an effective and easy-to-deploy solution to protect EBGP peering sessions

from CPU utilization-based attacks. When this feature is enabled, a host cannot attack a BGP session if the host is not a member of the local or remote BGP network or if the host is not directly connected to a network segment between the local and remote BGP networks. This solution greatly reduces the effectiveness of DoS attacks against a BGP AS.

neighbor ttl-security

To secure a BGP peering session and to configure the maximum number of hops that separate two EBGP peers, use the **neighbor ttl-security** command in address-family or router configuration mode.

- **neighbor** *neighbor-address* **ttl-security hops** *hop-count*

To disable this function, use the **no** form of this command.

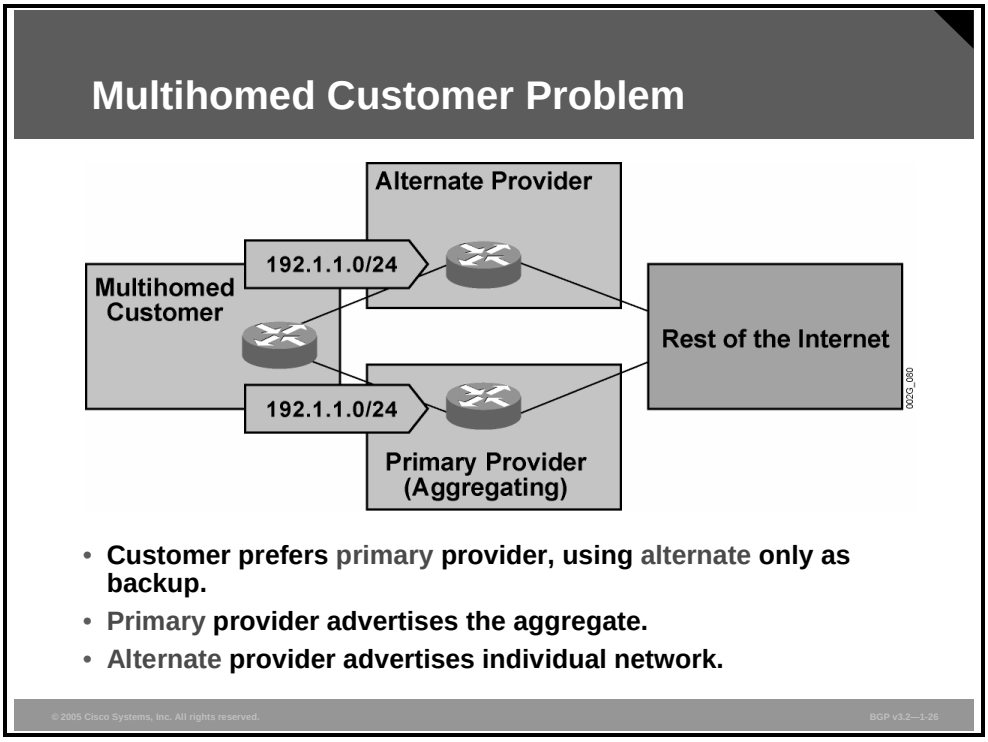
- **no neighbor** *neighbor-address* **ttl-security hops** *hop-count*

Syntax Description

Parameter	Description
neighbor address	IP address of the neighbor.
hops *hop-count*	Maximum number of hops that can separate the EBGP peer from the local router. The value for the *hop-count* argument is a number from 1 to 254.

Multihomed Customer Problem

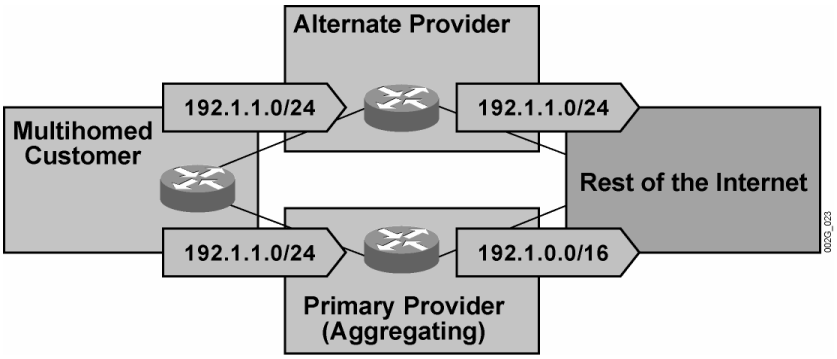
This topic describes a situation in which route aggregation in BGP is not appropriate.



In this example, the primary provider is doing aggregation of 192.1.0.0/16 before sending it to the rest of the network. This situation means that the primary provider is also doing proxy aggregation for the route 192.1.1.0/24 that is advertised by the multihomed customer. The rest of the Internet will not see the route 192.1.1.0/24 via the primary provider.

The multihomed customer also advertises 192.1.1.0/24 to the alternate provider. In this case, the provider does not do any aggregation of any routes starting with 192.1 (and should not do so). This situation means that the alternate provider will propagate 192.1.1.0/24 to the rest of the Internet.

Multihomed Customer Problem (Cont.)



- **Customer prefers primary provider, using alternate only as backup.**
- **Primary provider advertises the aggregate.**
- **Alternate provider advertises individual network.**

The rest of the Internet now sees overlapping routes. It sees 192.1.1.0/24 as reachable via the alternate provider and 192.1.0.0/16 as reachable via the primary provider. These two routes are treated as different routes. They are not compared with each other in a route selection process because they indicate different destinations. Because the router views them as different destinations, both routes will be injected into the routing table.

If a packet arrives with a destination address in the 192.1.1.0/24 network, the rest of the Internet will follow the “longest matching prefix” rule and forward the packet to the alternate provider.

To avoid this issue, the primary provider must turn off aggregation. If the primary provider does so, the rest of the Internet will see 192.1.1.0/24 both ways. And, because exactly the same route (network and mask) is reachable in two ways, route selection processing starts. Depending on the attribute values, the rest of the Internet could be advised to use the primary provider instead of the alternate one.

However, turning off aggregation will also cause the primary provider to advertise all routes within the aggregate, and all benefits of aggregation will be lost.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **The BGP process in a Cisco router is started with the** router bgp **command.**
- **The** neighbor remote-as **router configuration command adds an entry to the BGP neighbor table, the** neighbor description **router configuration command associates a description with a neighbor, and the** neighbor shutdown **router configuration command disables a neighbor.**
- **The BGP keepalive and holdtime timers can be changed for the BGP process (using the** timers bgp **router configuration command) or on a per-neighbor basis (using the** neighbor timers **router configuration command).**
- **MD5 authentication can be used to secure a connection between two BGP neighbors. The** neighbor password **router configuration command enables MD5 authentication on a TCP connection between two BGP peers.**

Summary (Cont.)

- **Local networks are announced in BGP by listing them with the** network **command or by redistributing them with the** redistribute **command. The** network **command can be used to announce any IP prefix. If you use the classless version of the** network **command, a matching route has to reside in the IP routing table.**
- **If there are a lot of networks to be advertised, and BGP is used primarily to achieve scalability, it may be easier to let the local IGP find the routes and then redistribute them into BGP. To redistribute routes from one routing process into another routing process, use the** redistribute **router configuration command.**
- **BGP4 supports CIDR, and any BGP router can advertise individual networks or supernets (prefixes). To specify the networks to be advertised by the BGP routing process, use the** network **router configuration command.**

Summary (Cont.)

- **There are cases where routes that are already in the BGP table have to be summarized. This process is called "aggregation" in BGP and is configured with the** aggregate-address **command.**
- **The BGP conditional route injection feature provides a means to originate a prefix into a BGP routing table without the corresponding match, allowing more specific routes to be generated based on administrative policy or traffic engineering information to provide more specific control over the forwarding of packets to these more specific routes.**
- **The BGP Support for TTL Security Check feature introduces a lightweight security mechanism to protect EBGP peering sessions from CPU utilization-based attacks; a host cannot attack a BGP session if the host is not a member of the local or remote BGP network or if the host is not directly connected to a network segment between the local and remote BGP networks.**
- **BGP route aggregation is not appropriate in multihomed topologies.**

Lesson 6

Monitoring and Troubleshooting BGP

Overview

Border Gateway Protocol (BGP) monitoring commands are important to ensuring that basic BGP configurations are operating correctly. If basic BGP configurations are not functioning as expected, BGP troubleshooting skills are critical to successful problem resolution.

This lesson introduces the Cisco IOS commands that are available for monitoring and troubleshooting basic BGP configurations. The commands that are required to monitor the status of BGP, neighbor connections, and the BGP table are discussed. The lesson also discusses techniques for troubleshooting the most common BGP session startup issues.

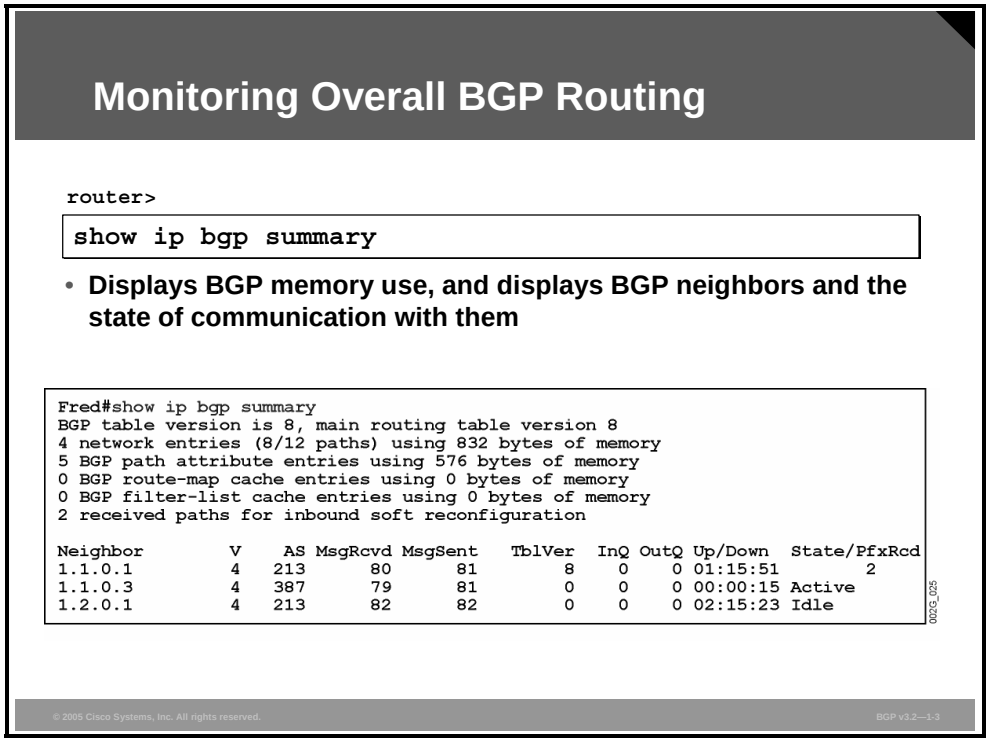
Objectives

Upon completing this lesson, you will be able to perform the steps to correct basic BGP configuration and session errors. This ability includes being able to meet these objectives:

- Identify the Cisco IOS command that is required to monitor the overall status of the BGP routing process
- Identify the Cisco IOS command that is required to monitor BGP neighbors
- Identify the Cisco IOS commands that are required to monitor the BGP table
- Identify the Cisco IOS commands that are required to perform basic BGP debugging
- List common BGP session startup problems
- Troubleshoot basic BGP session startup problems when the neighbor is not reachable
- Troubleshoot basic BGP session startup problems when the neighbor is not configured
- Troubleshoot basic BGP session startup problems when an AS number mismatch exists

Monitoring Overall BGP Routing

This topic describes the command that is used to monitor the overall status of the BGP routing protocol process.



This command is very useful when you are troubleshooting BGP. The output in the figure provides a short summary of the status of the BGP process in the router.

The first section of the output describes the BGP table and its content:

- The BGP table version is the version number of the local BGP table. This number is increased every time that the table is changed.
- The main routing table version shows the last version of the BGP database that was injected into the main routing table.
- The subsequent lines of text indicate the amount of memory that has been allocated to hold the table. These lines of text display how many networks are known and how many different paths and attribute values are associated with them.

The second section of the output is a table in which the current neighbor statuses are shown. There is one line of text for each neighbor that has been configured. The columns are as follows:

- IP address of the neighbor as configured in the local router
- BGP version number that is used by the router when communicating with the neighbor
- Autonomous system (AS) number of the remote neighbor
- Number of messages and updates that have been received from the neighbor since the session was established
- Number of messages and updates that have been sent to the neighbor since the session was established
- Version number of the local BGP table that has been included in the most recent update to the neighbor
- Number of messages that are waiting to be processed in the incoming queue from this neighbor
- Number of messages that are waiting in the outgoing queue for transmission to the neighbor
- How long the neighbor has been in the current state and the name of the current state (the state “Established” is not printed out, so no state name indicates “Established”)

You can use this information to verify that BGP sessions are up and established. If they are not, you will have to further investigate the BGP configuration to locate the problem. You can also verify the IP address and AS number of the configured BGP neighbor with the **show ip bgp summary** command.

If the session state is “Established,” the number of messages that have been sent and received, as displayed in the output of the **show ip bgp summary** command, can indicate BGP stability. Use the command a few times, with a time interval between the printouts, and calculate how many messages have been exchanged during that period.

A large number of messages in the incoming queue indicates a lack of CPU resources in the local router. A large number of messages in the outgoing queue indicates a lack of bandwidth to the remote router or a lack of CPU resources in the remote router.

show ip bgp summary

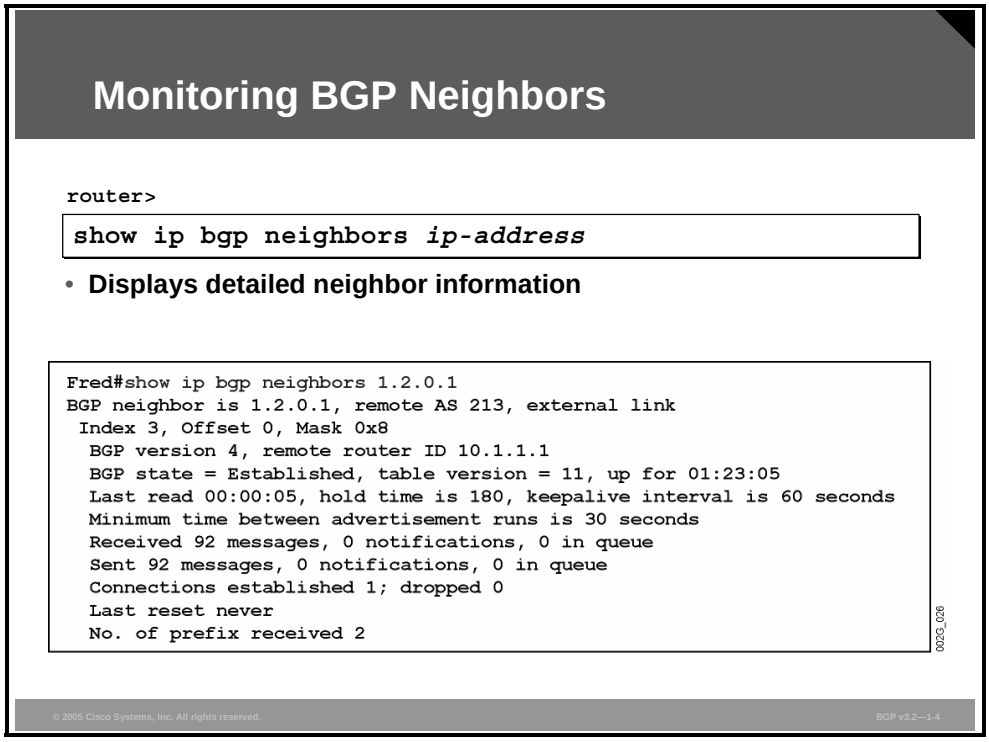
To display the status of all BGP connections, use the **show ip bgp summary** EXEC command.

- **show ip bgp summary**

This command has no arguments or keywords.

Monitoring BGP Neighbors

This topic describes the Cisco IOS command that is used to monitor BGP neighbors.



You can use this command for two different purposes. The general purpose, as shown in the figure, is to get information about the TCP session and the BGP parameters of the session. All BGP session parameters are displayed. In addition, TCP timers and counters are also displayed.

The other use is not shown in this example. If any of the optional qualifiers referring to routes or paths are given, the BGP routing information that was sent or received on this session is displayed. This feature is useful when you are troubleshooting path selection.

show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

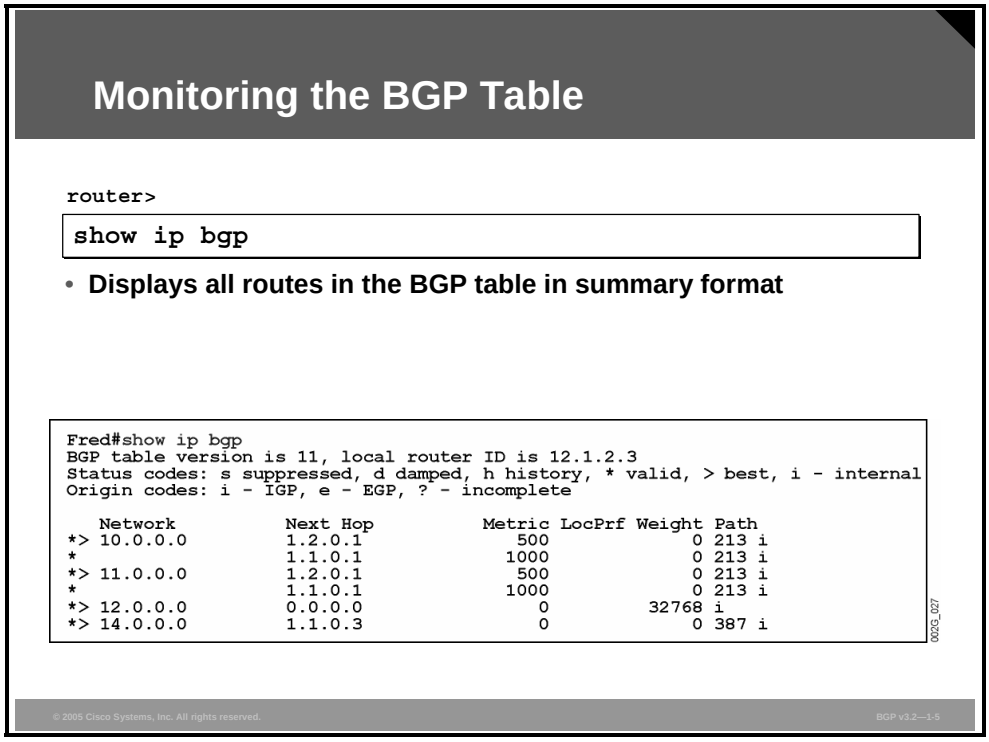
- **show ip bgp neighbors** [*address*] [**received-routes** | **routes** | **advertised-routes** | {**paths** *regular-expression*} | **dampened-routes**]

Syntax Description

Parameter	Description
address	(Optional) Address of the neighbor whose routes you have learned from. If you omit this argument, all neighbors are displayed.
received-routes	(Optional) Displays all received routes (both accepted and rejected) from the specified neighbor.
routes	(Optional) Displays all routes that are received and accepted. This is a subset of the output from the **received-routes** keyword.
advertised-routes	(Optional) Displays all the routes that the router has advertised to the neighbor.
paths *regular-expression*	(Optional) Regular expression that is used to match the paths received.
dampened-routes	(Optional) Displays the dampened routes to the neighbor at the IP address specified.

Monitoring the BGP Table

This topic describes the Cisco IOS commands that are used to monitor the BGP routing table.



In most cases, when the **show ip bgp** command is given without optional qualifiers, the entire BGP table is displayed. An abbreviated list of information about each route is displayed, one line per prefix. The output is sorted in network number order. Therefore, if the BGP table contains more than one route to the same network, the routes are displayed on successive lines. The network number is printed on the first of these lines only. The following lines, which refer to the same network, have the network number field left blank.

Some, but not all, of the BGP attributes that are associated with the route are displayed on the line. Next-hop, multi-exit discriminator (MED) (displayed as “Metric”), local preference, and weight each have their own columns. The AS-path attribute is displayed as the sequence of AS numbers in the “Path” column. Immediately following the AS path, but not part of the AS-path attribute, the origin attribute is displayed. The lowercase letter “i” means Interior Gateway Protocol (IGP), “e” means exterior gateway protocol (EGP), and “?” means incomplete or unknown.

The BGP path selection process selects one of the available routes to each of the networks as the best. This route are pointed out by the character “>” in the left column.

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

- **show ip bgp** [*network*] [*network-mask*] [**longer-prefixes**]

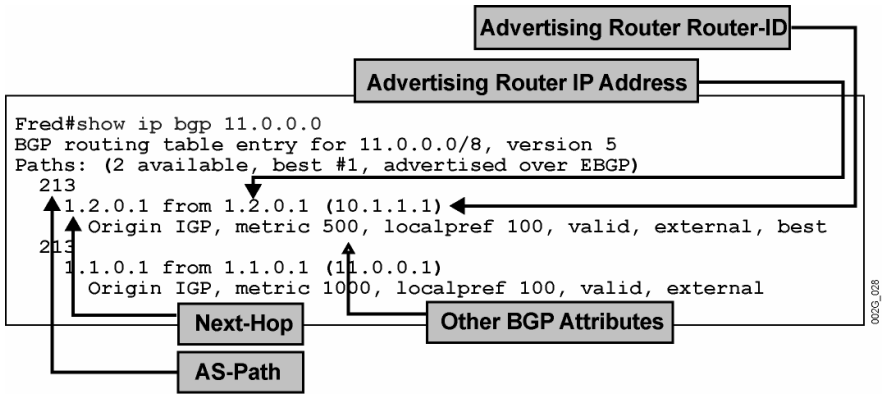
Syntax Description

Parameter	Description
network	(Optional) Network number, which is entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes matching the address and mask pair
longer-prefixes	(Optional) Displays the network and its more specific networks or prefixes.

Monitoring the BGP Table (Cont.)

```
router>
show ip bgp ip-prefix [mask subnet-mask]
```

- **Displays detailed information about all paths for a single prefix**



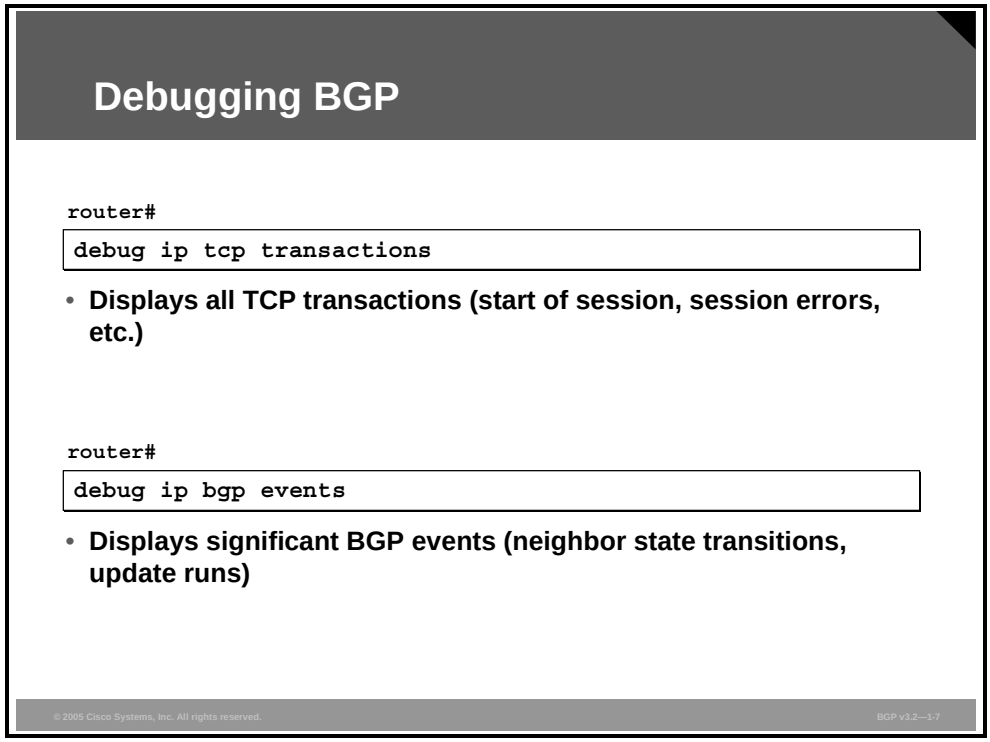
If more information and the complete set of BGP attributes are required, the **show ip bgp** command should be entered with the network number on the command line. This command displays all relevant BGP information about that specific network.

In this example, the information about network 11.0.0.0 is displayed. There are two routes to 11.0.0.0. One is received from neighbor 1.2.0.1 and the other from 1.1.0.1.

The BGP route selection process has selected the route via 1.2.0.1 as the best. This is thus the route that BGP will try to install in the routing table. Installation of routes in the routing table is made based on the administrative distance (AD).

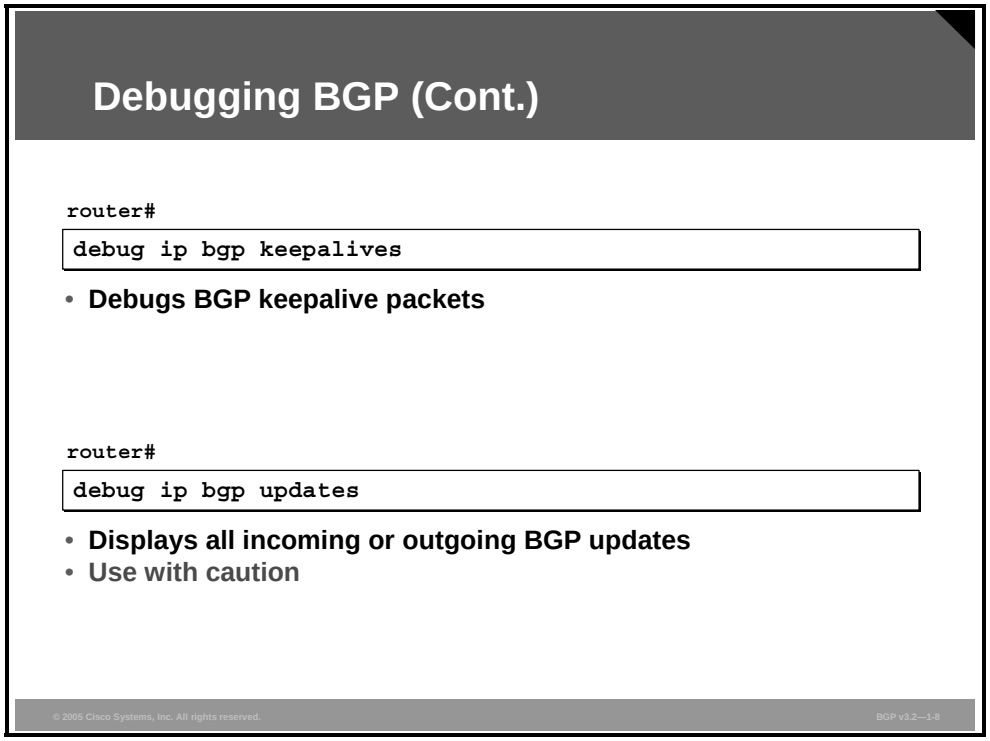
Debugging BGP

This topic describes the Cisco IOS commands that are used to perform debugging of basic BGP configurations.



If a BGP session stays in the Active state, where it is actively sending connection attempts to the neighbor, **debug ip tcp transactions** can provide valuable information about failed connection attempts. All TCP transactions in the router are displayed on the console as they happen. The network administrator can then determine whether the TCP session is being established, and, if not, what the probable cause of the problem might be.

If the TCP session succeeds but is torn down within a short period of time, you might find the reason if you use **debug ip bgp events**. All BGP events will be displayed on the console as they happen if this debug command is enabled.



In a stable state with no network topology changes, no BGP updates are sent between neighboring routers. When a BGP session has been idle for some time, the BGP protocol will exchange keepalive packets between BGP neighbors. The keepalive timer has a default value of 60 seconds.

Use the **debug ip bgp keepalives** command to get a printout on the console for every keepalive packet that is sent or received. Successful keepalive exchanges indicate that the session is working and is in a stable state.

If no keepalives have been sent or received, the session might still be working. The reason for not seeing any keepalives would be that the session is never idle long enough.

Use the **debug ip bgp updates** command to get a printout on the console for every update message that is sent or received. The successful exchange of updates indicates that the session is working and is not in the Idle state.

In a large network, updates are sent and received in large volumes. Starting the **debug ip bgp updates** command might cause extensive output on the console. In some cases, the CPU resources that are used to generate those outputs are so great that few CPU resources remain to actually forward traffic. In a case with very busy BGP sessions, it is actually possible to set the router in a condition where all CPU resources are consumed with the debugging printouts.

Debugging BGP (Cont.)

```
router#
debug ip bgp updates acl
```

- **Displays all incoming or outgoing BGP updates for routes matching an IP access-list**

```
router#
debug ip bgp ip-address updates [acl]
```

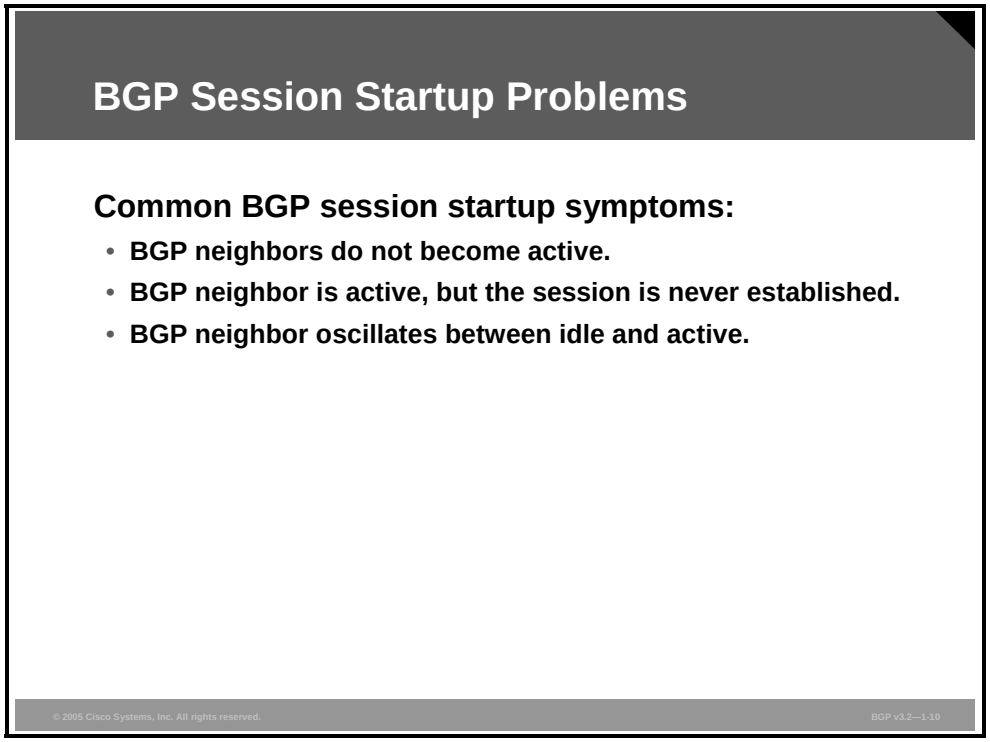
- **Displays all BGP updates received from or sent to a BGP neighbor (optionally matching an IP access-list)**

To avoid debug printouts for every update that is sent or received, you can create and associate an access-list with the **debug** command. When you use this command, the console displays only the updates that refer to a network number that is permitted by the access-list. The command is extremely useful in a live network with busy BGP sessions where the troubleshooter is interested only in updates for specific networks.

Indicating a specific neighbor can even further restrict the debugging. The console displays only the updates on the session with the indicated neighbor. Optionally, you can combine this debug command with an access-list.

BGP Session Startup Problems

This topic describes the most common session startup issues that you can experience when configuring basic BGP.

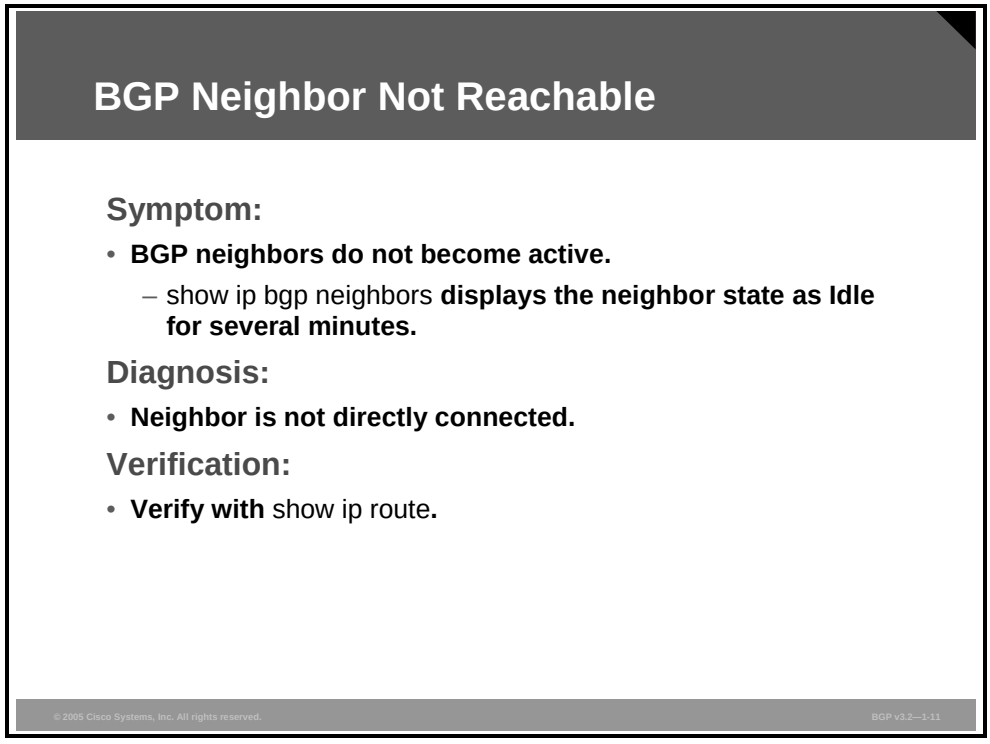


There are a number of common BGP session startup symptoms:

- A BGP neighbor never becomes active.
- A BGP neighbor is active, but the BGP session is not established.
- The BGP neighbor state oscillates between Idle and Active.

BGP Neighbor Not Reachable

This topic describes basic BGP troubleshooting for BGP session startup problems where the neighbor is not reachable.



BGP sessions to a router in another AS should normally run across directly connected interfaces (routers that share a common IP subnet). You must configure neighboring routers to reach each other using the IP address belonging to this shared subnet so that no other routing protocol is required to set up the BGP session.

If a router is configured with a BGP neighbor that is in another AS but not directly connected, the session will stay in the Idle state. The router will not even attempt to set up the session.

The normal way to fix this problem is to change the neighbor reference so that it is referred by an IP address that is directly connected. However, in some odd cases, the neighbor is intentionally reachable using an interface that is not directly connected. In that rare case, the local router must have routing information on how to reach that address. Also, you must configure the BGP session with the **ebgp-multihop** option.

If the session goes into the Active state, the router will attempt to establish the session. If session establishment is unsuccessful, you will have to troubleshoot the problem. The **debug ip tcp transactions** command will display the connect attempts.

BGP Neighbor Not Reachable (Cont.)

Symptom:

- **BGP neighbor is active; session is not established.**
  - debug ip tcp transactions **display shows that the TCP SYN packet is not answered with a SYN-ACK packet.**

Diagnosis:

- **Neighbor is not reachable.**

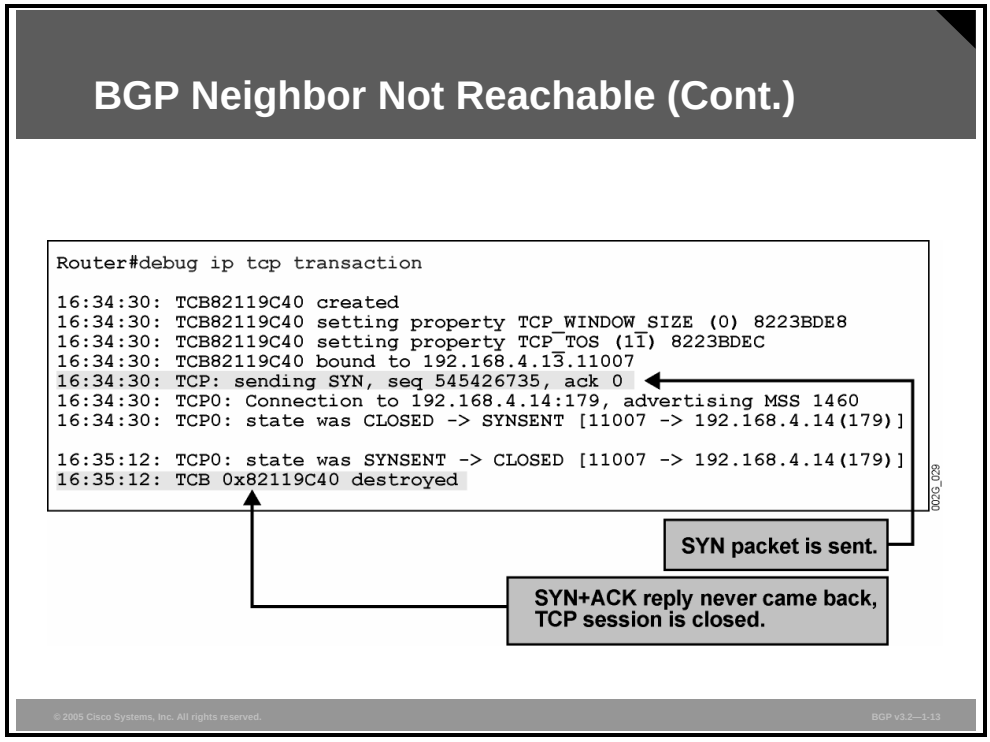
Verification:

- **Verify connectivity with** ping.
- **Check for the presence of an access-list.**

TCP session establishment starts with the router sending a TCP SYN packet. If the TCP SYN packet is never answered, the remote router might be dead or not reachable. Try to use the **ping** command and verify the existence of the remote router and the IP packet exchange between the local and remote router.

Example: BGP Neighbor Not Reachable

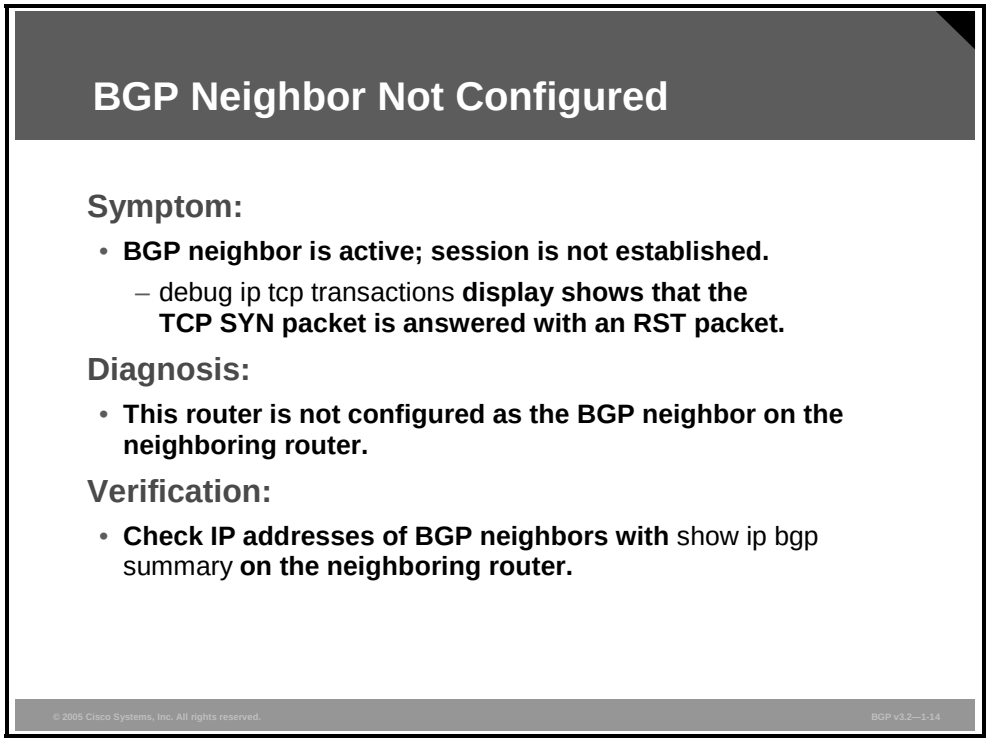
In this example, the remote BGP router is not available.



The sending router never receives the reply to the SYN packet and aborts the TCP session in approximately 45 seconds (changing the state from synsent to closed).

BGP Neighbor Not Configured

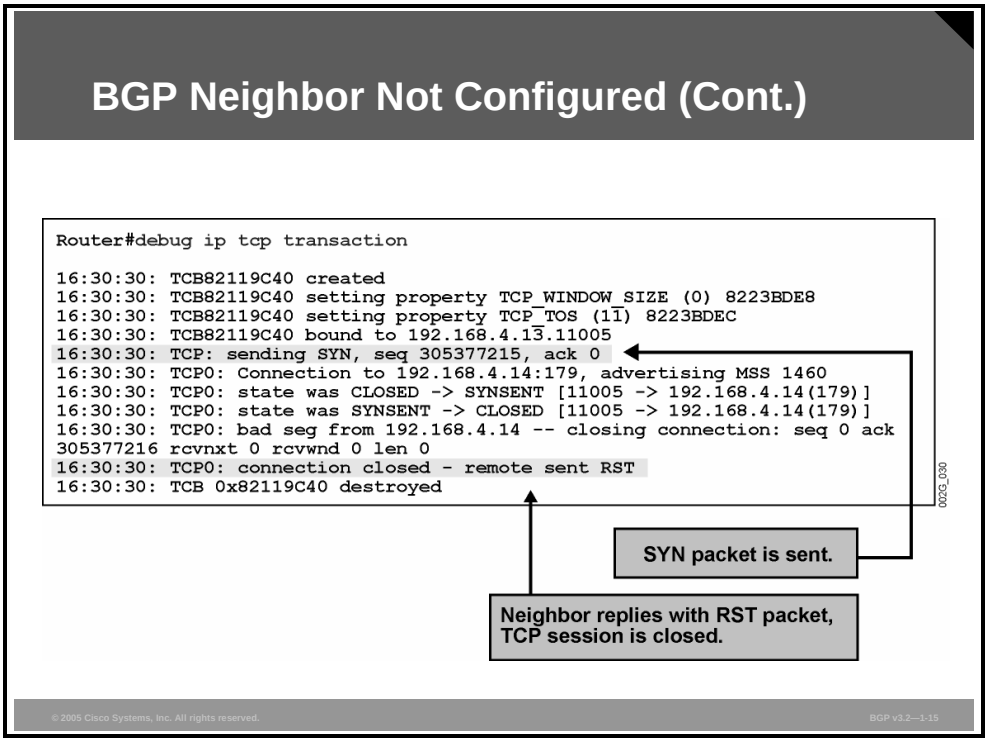
This topic describes basic BGP troubleshooting for BGP session startup problems where the neighbor is not configured.



If the TCP SYN packet is answered with a TCP RST packet, the remote router is alive and reachable but is not willing to grant the connection attempt. The reason for this refusal may be that BGP has not been fully configured on the remote router or that the source IP address that is used by the local router in the connection attempt is not in the list of valid neighbors for the remote router.

Example: BGP Neighbor Not Configured

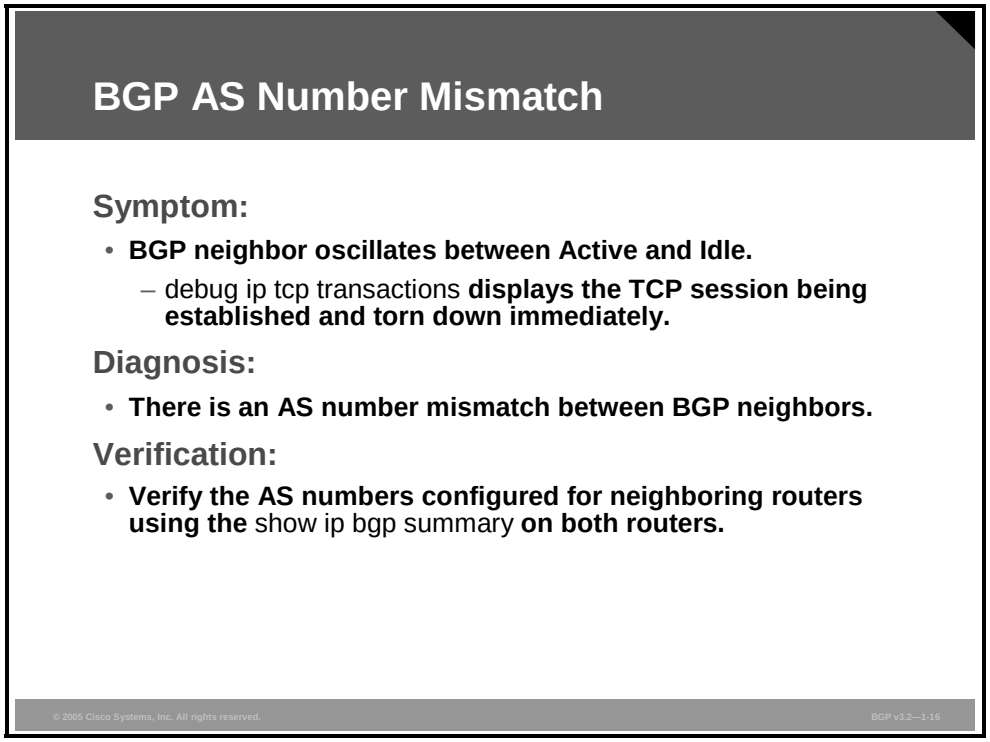
In this example, the remote router is not configured for BGP or there was a mismatch in the neighbor IP addresses.



The remote router responds with an RST packet as soon as it receives the initial SYN packet, terminating the BGP session.

BGP AS Number Mismatch

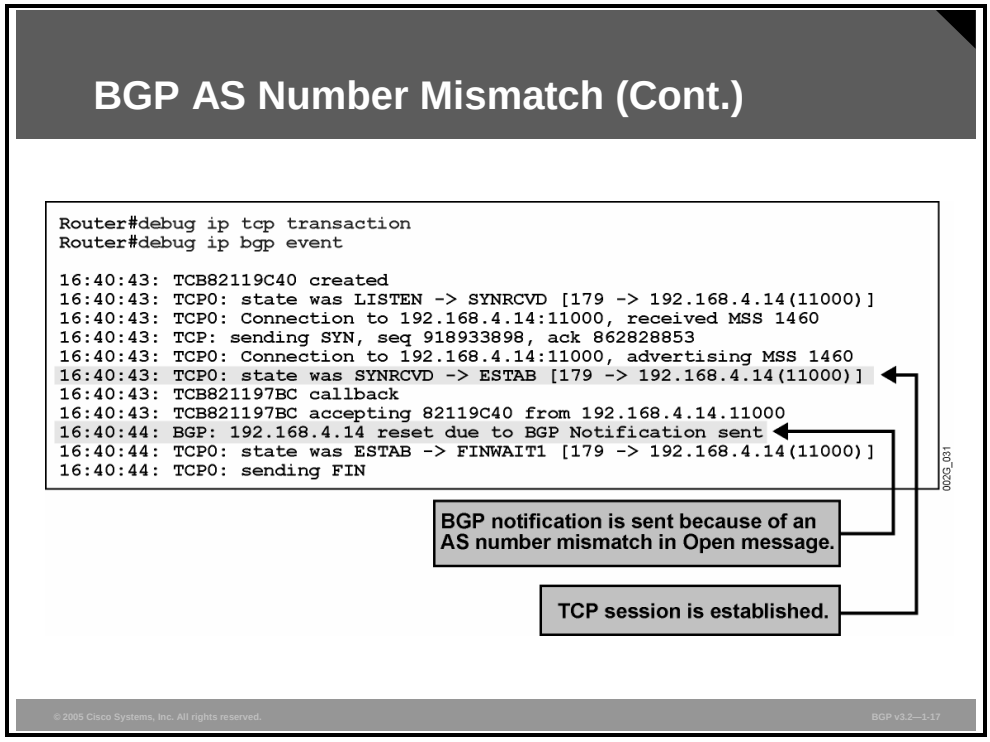
This topic describes basic BGP troubleshooting for BGP session startup problems where the AS numbers are not properly configured.



If the TCP session is established using the specified three-way handshake of SYN, SYN-ACK, and ACK, but the router drops the session after a short packet exchange, the BGP parameters are mismatched. Make sure that the remote AS that is configured on the router matches the local AS that is configured on the neighbor. If the AS numbers do not match, the router drops the session after exchanging BGP Open messages.

Example: BGP AS Number Mismatch

This example illustrates a mismatch in an AS number.



Whenever there is a mismatch in AS numbers (or any other BGP parameters that are necessary for proper BGP operation), the BGP session is terminated with a BGP notification, and the TCP session is terminated as well.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **The** show ip bgp summary **command displays the overall status of BGP and shows configured neighbors and their state.**
- **You can use the** show ip bgp neighbors **command to get more in-depth information about a specific BGP neighbor.**
- **All entries in the BGP table can be displayed with the** show ip bgp **command. You can also use** show ip bgp **to display an extended printout about a specific route in the BGP table.**
- **You can use the** debug ip tcp transactions **command to troubleshoot BGP session establishment problems. The command** debug ip bgp events **displays significant BGP events, while** debug ip bgp updates **displays the routing information being exchanged between BGP neighbors.**

Summary (Cont.)

- **Three common BGP session startup symptoms are that BGP neighbors never become active, that the BGP neighbor is active but the BGP session is not established, and that the BGP neighbor state oscillates between idle and active.**
- **If a router is configured with a BGP neighbor that is in another AS but not directly connected, the session stays in the Idle state.**
- **If a BGP neighbor is unreachable, no reply is sent for the TCP SYN packet, causing the session to time out.**
- **If the TCP session is established using the three-way handshake (SYN, SYN-ACK, ACK), but the session is dropped after a short packet exchange, BGP parameters are mismatched.**

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **BGP has reliable transport provided by TCP, a rich set of metrics called BGP path attributes, and scalability features such as batched updates that make it suitable for very large networks.**
- **Configured BGP neighbors establish a TCP session and exchange the BGP Open message, which contains the parameters that each BGP router proposes to use.**
- **Some path attributes are well-known and should be recognized by every BGP implementation. Some of the well-known attributes, such as AS-path, next-hop, and origin, are mandatory and have to be present in every BGP update.**

Module Summary (Cont.)

- **The route selection process takes into account various BGP attributes that are attached to the route, as well as local decisions.**
- **When you are configuring BGP neighbors, you will enable the BGP routing protocol process, establish neighbors, and advertise local routes.**
- **To ensure that basic BGP configurations are operating correctly, there are a number of Cisco IOS commands to monitor the status of BGP, neighbor connections, and the BGP table, as well as to troubleshoot the most common BGP session startup issues.**

This module presented an overview of BGP, a very robust and scalable routing protocol. The first lesson covered the uses of BGP and its limitations. The second lesson moved to describing the concept of BGP neighbors and the procedures for establishing neighbor sessions. Then, the third lesson listed BGP path attributes and their functionality. The fourth lesson described interdomain route processing. When you had the foundation of a general understanding of BGP, the fifth lesson explained how to configure it. The module ended with a lesson that identified the steps to monitor the operation of BGP and correct basic configuration errors.

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Border Gateway Protocol*. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm.
- Cisco Systems, Inc. *Configuring BGP*. http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/ip_c/ipcprt2/1cdbgp.htm.
- Cisco Systems, Inc. *Using the Border Gateway Protocol for Interdomain Routing*. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Which three items are BGP enhancements to traditional distance vector routing protocols? (Choose three.) (Source: Introducing BGP)

A) reliable updates
B) use of triggered updates only
C) enhanced security
D) rich metrics
E) route summarization
F) snapshot updates

Q2) Which protocol facilitates reliable update capabilities in BGP? (Source: Introducing BGP)

A) TCP
B) UDP
C) HSRP
D) ICMP

Q3) What are three characteristics of an AS? (Choose three.) (Source: Introducing BGP)

A) uses IGPs for intradomain routing
B) uses EGPs for interdomain routing
C) is a collection of networks under a common administrative authority
D) consists of a group of network domains
E) automatically summarizes addresses
F) regulated by the IETF

Q4) Which three scenarios are common scenarios where BGP is used? (Choose three.) (Source: Introducing BGP)

A) a customer with a connection to multiple service providers
B) service provider networks acting as transit systems and forwarding external traffic through their network
C) a single-site customer intranet with complex administrative policies between departments
D) as the core routing protocol in very large enterprise networks
E) as the routing protocol in an IS-IS backbone area
F) as the core routing protocol in an SNA network

Q5) What are three recommended BGP use guidelines for multihomed customer networks? (Choose three.) (Source: Introducing BGP)

A) Most multihomed customers should use BGP with their service providers.
B) Most multihomed customers should forward routing information that is received from one provider to the other provider.
C) The multihomed customer must have its own public AS number.
D) Multihomed customers should use a provider-independent, public address space.
E) The multihomed customer may use and advertise RFC 1918 addresses.
F) Multihomed customers should use the AS number of their primary ISP.

Q6) What is a limitation of the BGP routing protocol? (Source: Introducing BGP)

A) You cannot use BGP to implement hop-by-hop routing policy controls.
B) You cannot use BGP to influence the routing policy in a downstream AS.
C) BGP cannot control forwarding of packets based on their destination address.
D) BGP cannot scale to very large networks with more than 110,000 routes.

Q7) Which three statements are true of BGP mandatory well-known attributes? (Choose three.) (Source: Understanding BGP Path Attributes)

A) They must be present in all BGP updates.
B) All BGP-compliant implementations must recognize them.
C) All BGP-compliant routers must adhere to policies specified in mandatory attributes.
D) All well-known attributes are propagated to other neighbors.
E) They must be present in some BGP updates.

Q8) Which three attributes are BGP mandatory well-known attributes? (Choose three.) (Source: Understanding BGP Path Attributes)

A) next-hop
B) weight
C) AS-path
D) origin
E) MED
F) local preference

Q9) Which three possible values are assigned to the BGP origin attribute? (Choose three.) (Source: Understanding BGP Path Attributes)

A) IGP
B) EGP
C) unknown
D) internal
E) external
F) MED

Q10) Which nontransitive optional BGP attribute is useful in assisting with the route selection process when multiple links to another AS exist? (Source: Understanding BGP Path Attributes)

A) next-hop
B) local preference
C) MED
D) AS-path

Q11) In which two ways can the BGP next-hop attribute be modified? (Choose two.) (Source: Understanding BGP Path Attributes)

A) If the next-hop attribute is in the same IP subnet as the receiving router, the attribute is unchanged; otherwise, it is set to the IP address of the sending router.
B) The next-hop attribute is always set to the IP address of the sending router.
C) The next-hop attribute is modified only when BGP packets exit an AS.
D) The BGP next-hop attribute is modified only when BGP packets traverse point-to-point links.

Q12) Which three statements regarding the BGP AS-path attribute are true? (Choose three.) (Source: Understanding BGP Path Attributes)

A) The local AS number is prepended to the AS path each time that the route crosses an AS boundary.
B) The AS that originally injected the route into BGP is always found at the rightmost end of the AS path.
C) The AS-path attribute can be used to avoid routing loops.
D) BGP routes with an empty AS path were injected into BGP from outside the local AS.
E) The local AS number is appended to the end of the AS path each time that the route crosses an AS boundary.
F) The AS that originally injected the route into BGP is always found at the leftmost end of the AS path.

Q13) What is indicated by a state of "Idle" in the output of the **show ip bgp summary** command? (Source: Establishing BGP Sessions)

A) The router is currently not attempting to establish a connection with a neighbor.
B) The connection to the configured neighbor has timed out.
C) The connection to a BGP neighbor has been established, and no errors have been received on the connection.
D) The connection to a BGP neighbor has been established, and no packets have been sent.

Q14) What happens if two TCP connection attempts between configured BGP neighbors succeed? (Source: Establishing BGP Sessions)

A) Both connections will be terminated, and the neighbors will re-establish a neighbor relationship.
B) One connection will be maintained as primary and the other as backup.
C) One of the two connections will be torn down.
D) The router with the lower router-ID will determine if the second connection is torn down or used as a backup TCP connection.

Q15) Given the following BGP session states:

1. OpenConfirm
2. Established
3. Idle
4. OpenSent
5. Active

What is their order of progression during the creation of a successful neighbor session? (Source: Establishing BGP Sessions)

A) 5, 1, 4, 2, 3
B) 3, 4, 1, 5, 2
C) 5, 4, 1, 3, 2
D) 3, 5, 4, 1, 2

Q16) What does the field “TblVer” indicate in the output of the **show ip bgp summary** command? (Source: Establishing BGP Sessions)

A) the current version of BGP in use by the router
B) the number of route prefixes that are contained in the BGP update of the router
C) BGP messages that have been received from that neighbor
D) the last version of the BGP database that was sent to that neighbor

Q17) What occurs when you use MD5 between two BGP neighbors? (Source: Establishing BGP Sessions)

A) Every packet is encrypted with MD5.
B) The IP header is encrypted using MD5.
C) An MD5 checksum is calculated and sent with each packet so that its source can be verified.
D) A username and password are embedded in an IP datagram that is matched to a username and password on the remote neighbor.

Q18) What does a router that is running BGP do with a BGP update that contains its own AS path? (Source: Processing BGP Routes)

A) The router checks to see whether the information that is contained in the update is better than its current information. If it is, it will update its BGP table.
B) The router accepts the route update.
C) The router silently discards (denies) the route.
D) The router returns an error to the router that sent the update.

Q19) How many alternate paths to a single destination will a BGP router maintain in the BGP table? (Source: Processing BGP Routes)

A) The router will maintain only the best path to the destination.
B) The router will maintain two paths, the best path and a backup route.
C) The BGP table will hold up to four routes by default and a maximum of six configurable routes.
D) The BGP table will store all valid, advertised routes to the destination in the BGP table.

Q20) When a router has more than one alternative route to reach the same IP subnet (network and mask), the router has to select one of them as best in its default mode of operation Match the following steps to the correct step in the process. (Source: Processing BGP Routes)

A) The router compares MED values, but only if it receives the updates from the same neighboring AS. Routes with a lower MED are preferred.

B) The router checks whether the next-hop attribute indicates an IP address that is reachable according to the current routing table. If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best. (Source: Processing BGP Routes)

C) The router prefers the route with the higher weight.

D) If the local preference attributes are different, the route with the highest value is selected as best.

E) The lengths of the AS paths are compared (the content is not checked; only the number of autonomous systems in each AS path is counted). The route with the shortest length is selected.

F) If one of the routes is injected into the BGP table by the local router, the local router prefers it to any routes that it receives from other BGP routers.

G) If the AS-path lengths are the same, the origin code is checked. BGP prefers the path with the lowest origin type: IGP is lower than EGP, and EGP is lower than Incomplete.

_____ 1. Step 1

_____ 2. Step 2

_____ 3. Step 3

_____ 4. Step 4

_____ 5. Step 5

_____ 6. Step 6

_____ 7. Step 7

Q21) What are two ways in which local networks are advertised into the BGP routing protocol process? (Choose two.) (Source: Processing BGP Routes)

A) automatically, after a BGP neighbor session is established

B) manually, with the **network** command

C) through redistribution into the BGP process

D) by advertising them to the BGP table on the router after Cisco Discovery Protocol discovers connected networks

Q22) What are two situations when is it appropriate to disable automatic summarization in BGP? (Choose two.) (Source: Processing BGP Routes)

A) when BGP neighbors are not configured to advertise aggregate routes to upstream providers

B) when the classless variant of the **network** command is used

C) when you are using a classless IGP in the AS

D) when the effects of automatic summarization of IGP-to-BGP redistribution are not desired

Q23) What is the AD of BGP routes in the IP routing table that were learned from BGP neighbors in a different AS? (Source: Processing BGP Routes)

A) 1
B) 20
C) 90
D) 120

Q24) Which three BGP attributes are displayed for each route in the BGP table when you are using the **show ip bgp** command? (Choose three.) (Source: Processing BGP Routes)

A) weight
B) communities
C) origin
D) AS-path

Q25) What is the valid AS number range for a BGP process on a Cisco router? (Source: Configuring Basic BGP)

A) 1 to 256
B) 1 to 32768
C) 1 to 65535
D) 1 to 131072

Q26) Which two parameters must you configure with the **neighbor** command to establish a BGP session with an external neighbor? (Choose two.) (Source: Configuring Basic BGP)

A) neighbor IP address
B) subnet mask of the IP network
C) remote AS number
D) local AS number
E) description of the neighbor

Q27) What is the best method to temporarily disable a BGP neighbor session? (Source: Configuring Basic BGP)

A) remove the **neighbor** command from the BGP router process
B) remove the BGP router process from the configuration
C) terminate the neighbor connection with the **neighbor shutdown** command
D) disconnect the neighbor by initiating a router reload

Q28) Which two of the following statements about configuring BGP timers are accurate? (Choose two.) (Source: Configuring Basic BGP)

A) Changing the BGP default holdtime and keepalive timers is usually not recommended.
B) The **neighbor timers** command sets the timers for a specific BGP peer or peer group.
C) The **timers bgp** command sets the timers for a specific BGP peer or peer group.
D) Holdtime indicates the frequency (in seconds) with which the Cisco IOS software sends messages to its peer.

Q29) Which two of the following are characteristics of the *string* component of the **neighbor** *{ip-address | peer-group-name}* **password** *string* command? (Choose two.) (Source: Configuring Basic BGP)

A) can contain any alphanumeric characters, including spaces
B) case-sensitive password of up to 100 characters
C) first character can be a number
D) cannot specify a password in the format “number-space-anything”

Q30) Which three steps must you complete to advertise a classless prefix into BGP? (Choose three.) (Source: Configuring Basic BGP)

A) configure the prefix with the **network** command
B) specify the **mask** keyword with the locally advertised route
C) configure the **redistribute connected** command under the BGP router process
D) use a static route pointing to null 0 that matches the prefix

Q31) Which origin code is carried with routes that are redistributed into BGP? (Source: Configuring Basic BGP)

A) internal
B) external
C) unknown
D) incomplete

Q32) Which two of the following statements about the classless behavior of BGP are correct? (Choose two.) (Source: Configuring Basic BGP)

A) When an exact match is not found in the IP routing table a matching prefix is automatically configured on the router.
B) In the **network *ip-prefix-address* mask *subnet-mask*** command, the prefix does not have to match an entry in the IP routing table
C) To advertise classless networks into BGP (a subnet or a supernet), you can use the **network** command with the **mask** keyword and the subnet mask specified.
D) If the keyword **mask** and the subnet mask are omitted, the network is assumed to have its natural mask according to the network class.

Q33) What are two benefits of using route aggregation in BGP? (Choose two.) (Source: Configuring Basic BGP)

A) It ensures that even if aggregate networks are down, the aggregate is advertised, which eliminates black holes.
B) It reduces the amount of memory that is used in the router to store the BGP table.
C) It reduces route flapping and its effects on router CPU resources.
D) BGP attribute granularity is maintained, which ensures optimal path selection.

Q34) Which two of the following are characteristics of the BGP Conditional Route Injection feature? (Choose two.) (Source: Configuring Basic BGP)

A) allows you to originate a prefix into a BGP routing table without the corresponding match
B) enabled with the **bgp inject-map exist-map** command
C) allows conditional injecting or replacing more specific prefixes with less specific prefixes
D) allows origination of a prefix into a BGP routing table only with the corresponding match

Q35) Which two of the following are characteristics of the BGP Support for TTL Security Check feature? (Choose two.) (Source: Configuring Basic BGP)

A) should be configured on only one participating router
B) prevents BGP sessions from expiring even if keepalive packets are not received before the session timer expires
C) protects the EBGP peering session by comparing the value in the TTL field of received IP packets against a hop count that is configured locally for each EBGP peering session.
D) supports both directly connected peering sessions and multihop EBGP peering sessions

Q36) Which two of the following are functions of the **show ip bgp summary** command? (Choose two.) (Source: Monitoring and Troubleshooting BGP)

A) displays BGP memory use
B) displays BGP neighbors and status of communication with them
C) locates problems in BGP sessions that are up and established
D) displays the BGP routing table

Q37) Which command do you use to display detailed BGP neighbor information? (Source: Monitoring and Troubleshooting BGP)

A) **show ip bgp summary**
B) **show ip bgp**
C) **show ip bgp neighbors address**
D) **show ip bgp detail**

Q38) Which two of the following statements about the **show ip bgp** command that is used to monitor the BGP routing table are accurate? (Choose two.) (Source: Monitoring and Troubleshooting BGP)

A) The **show ip bgp** command shows an abbreviated list of information about each route, displaying one line per prefix.
B) The **show ip bgp** command shows a full list of information about each route, displaying one line per prefix.
C) All of the BGP attributes that are associated with the route are displayed on the line.
D) If the BGP table contains more than one route to the same network, the routes are displayed on successive lines of the command output.

Q39) Which debug command should you enable to troubleshoot BGP session startup issues where the TCP connection never succeeds? (Source: Monitoring and Troubleshooting BGP)

A) **ip bgp updates**
B) **ip packets**
C) **ip bgp keepalives**
D) **ip tcp transactions**

Q40) What are the three most common session startup issues that you can experience when configuring basic BGP? (Choose three.) (Source: Monitoring and Troubleshooting BGP)

A) BGP neighbors do not become active.
B) BGP routing loops cause black holes.
C) A BGP neighbor is active, but the BGP session is not established.
D) The BGP neighbor state oscillates between Idle and Active.
E) The BGP session is active, but the neighbor cannot be reached.
F) BGP keepalives experience intermittent failures.

Q41) What is the most common reason for a BGP session not leaving the Idle state? (Source: Monitoring and Troubleshooting BGP)

A) The TCP port for the connection is not configured.
B) The external neighbor is not directly connected.
C) The TCP SYN packet is answered with an RST packet.
D) The neighbors have been configured with the same AS number.

Q42) What will result from attempting to open a BGP connection with a neighbor that has not been properly configured for BGP? (Source: Monitoring and Troubleshooting BGP)

A) The BGP session will remain in the Idle state.
B) The neighbor session will be established, and the session startup parameters will be negotiated over the TCP session.
C) The BGP session will be immediately terminated with a TCP RST packet.
D) The BGP session will become “stuck in Active state.”

Q43) When a BGP neighbor oscillates between Active and Idle, what is the likely diagnosis? (Source: Monitoring and Troubleshooting BGP)

A) There are mismatched keepalive intervals.
B) There is an AS number mismatch between BGP neighbors.
C) One router is not configured as the BGP neighbor on the neighboring router.
D) The BGP neighbor is not reachable.

Module Self-Check Answer Key

Q1) A, B, D
Q2) A
Q3) A, B, C
Q4) A, B, D
Q5) A, C, D
Q6) B
Q7) A, B, D
Q8) A, C, D
Q9) A, B, C
Q10) C
Q11) A, B
Q12) A, B, C
Q13) A
Q14) C
Q15) D
Q16) D
Q17) C
Q18) C
Q19) D
Q20) A-7, B-1, C-2, D-3, E-5, F-4, G-6
Q21) B, C
Q22) B, D
Q23) B
Q24) A, C, D
Q25) C
Q26) A, C
Q27) C
Q28) A, B
Q29) A, D
Q30) A, B, D
Q31) D
Q32) C, D
Q33) B, C
Q34) A, B
Q35) C, D
Q36) A, B
Q37) C
Q38) B, D
Q39) D
Q40) A, C, D
Q41) B
Q42) C
Q43) B

Module 2

BGP Transit Autonomous Systems

Overview

This module is one of the focal points of the Border Gateway Protocol (BGP) curriculum: a discussion of BGP issues in a transit autonomous system (AS). The module covers basic BGP transit AS issues, ranging from synchronization between an Interior Gateway Protocol (IGP) and BGP to Internal Border Gateway Protocol (IBGP) full-mesh and next-hop requirements.

Module Objectives

Upon completing this module, you will be able to use BGP policy controls to influence the BGP route selection process in a network scenario in which you must support connections to multiple ISPs. This ability includes being able to meet these objectives:

- Describe the function of a transit AS and the need for IBGP
- Describe the interaction in a transit AS between EBGP and IBGP in relation to relevant attributes
- Describe the function of an IGP in forwarding packets through an AS
- Configure an AS to act as a transit backbone in a BGP network
- Verify proper operation of a configured BGP transit network by performing the steps necessary to correct basic IBGP configuration errors

Lesson 1

Working with a Transit AS

Overview

All transit autonomous systems are required to carry traffic originating from or destined for locations outside of that autonomous system (AS). For the transit AS to meet this requirement, a degree of interaction and coordination between Border Gateway Protocol (BGP) and the Interior Gateway Protocol (IGP) that is used by that particular AS is necessary. Such a configuration requires special care to ensure consistency of routing information throughout the AS.

The topology of the Internet can be viewed as a series of connections between stub networks, multihomed networks, and transit autonomous systems. A multihomed AS containing more than one connection to the outside world and allowing traffic not originating in that AS to travel through it is a transit AS. This lesson introduces the concept of the multihomed transit AS and how BGP exchanges routing information inside the AS and between neighboring autonomous systems. It also explains the requirement for Internal Border Gateway Protocol (IBGP) within the multihomed transit AS.

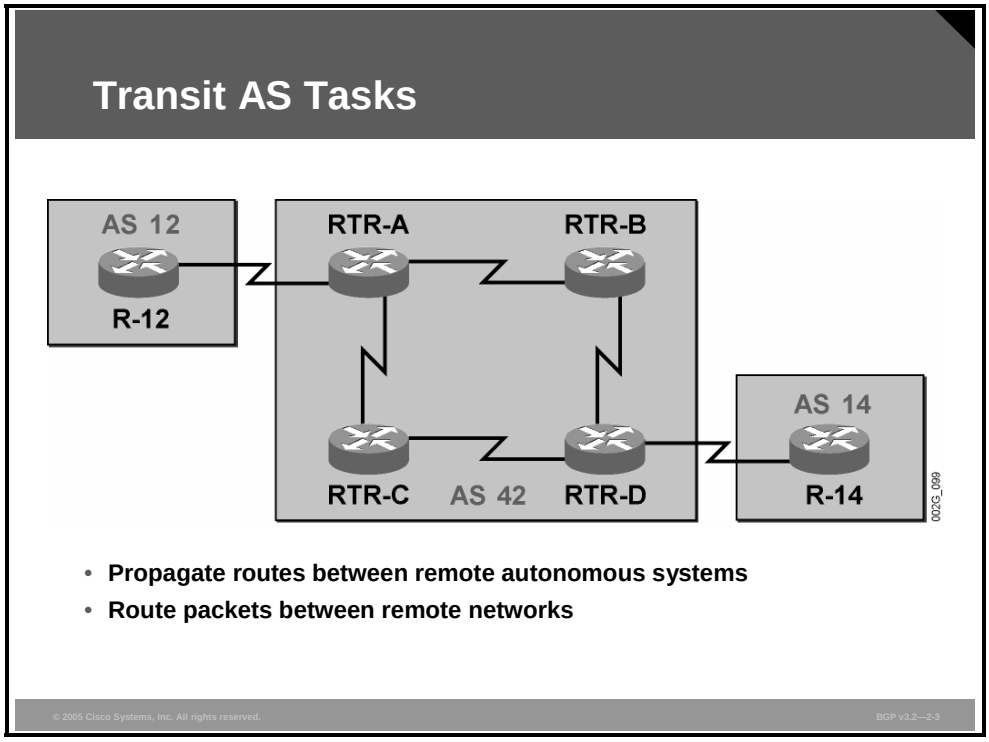
Objectives

Upon completing this lesson, you will be able to describe the function of a transit AS and the need for IBGP. This ability includes being able to meet these objectives:

- List the functions of a transit AS
- Describe external route propagation between autonomous systems in a BGP network
- Describe internal route propagation within a BGP AS
- Explain how transiting packets are forwarded inside a transit AS
- Explain the need for deploying IBGP on all core routers

Transit AS Tasks

This topic describes the functions of a transit AS.

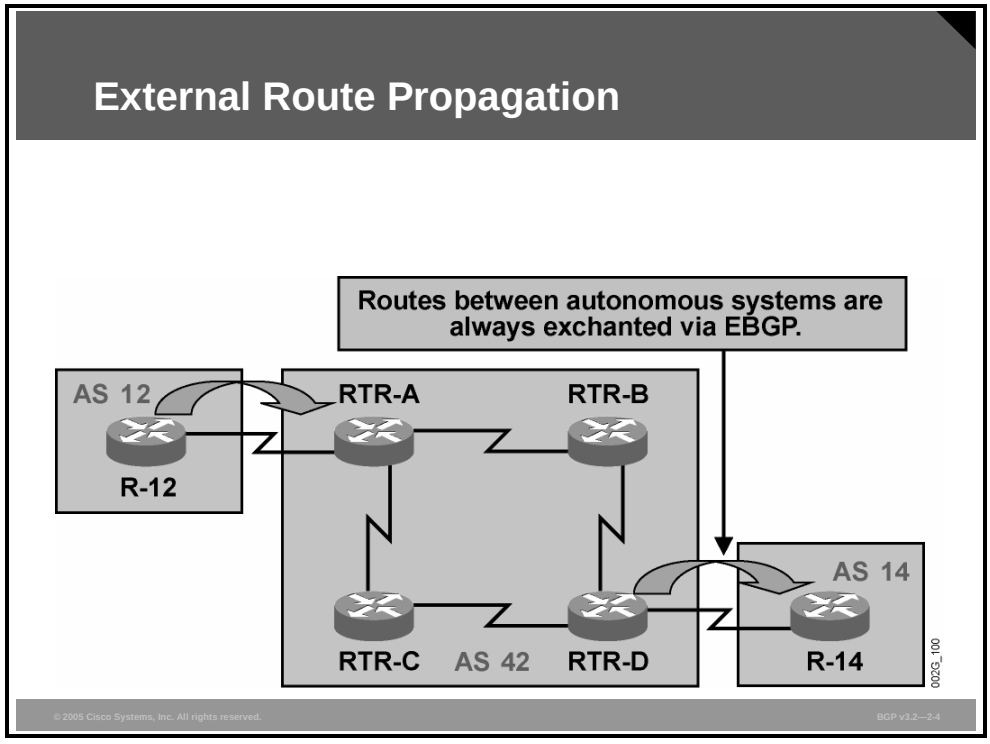


Routers in a transit AS have to perform two tasks:

- Receive routing information updates about reachable networks from neighboring autonomous systems, propagate the information through their own AS, and send it to other neighboring autonomous systems.
- Forward IP packets that they have received from a neighboring AS through their own AS to a downstream neighboring AS. The routers in the transit AS perform this task using the routing information that they have received as part of the first task.

External Route Propagation

This topic describes external route propagation between autonomous systems in a BGP network.



Two autonomous systems usually exchange routing information about reachable networks using BGP. There is currently no alternative routing protocol that has the scalability and security characteristics of BGP.

In the figure, the BGP session between R-12 and RTR-A is called an External Border Gateway Protocol (EBGP) session because R-12 and RTR-A are in different autonomous systems.

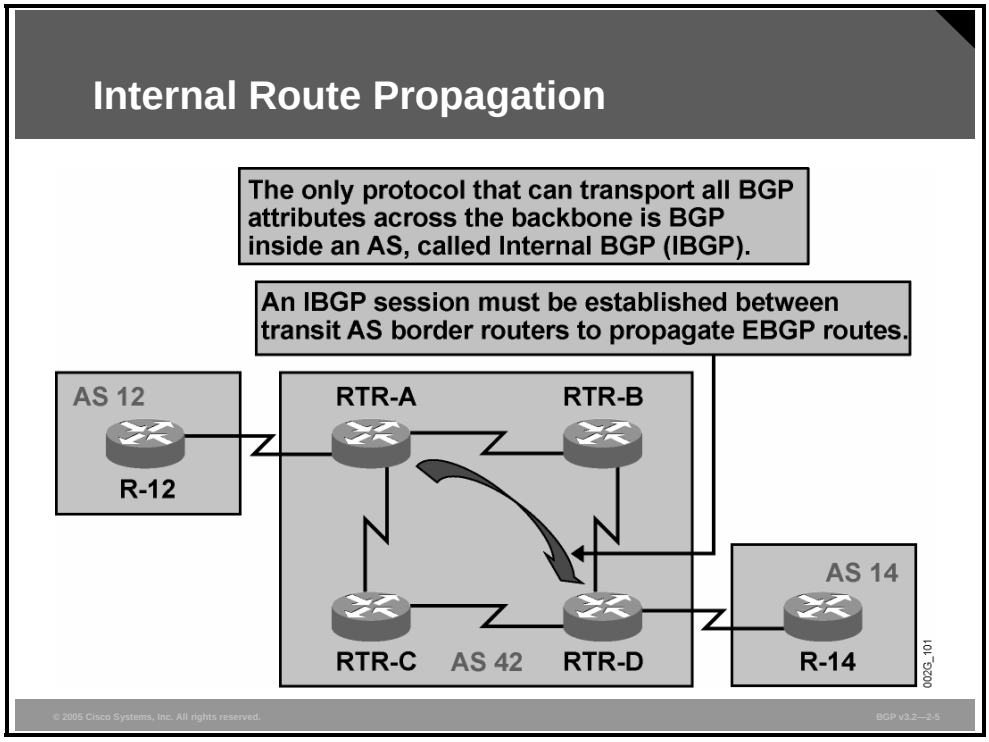
BGP routing information updates consist of the network address, subnet mask, and any number of BGP attributes. No other routing protocol provides the same richness of route attributes as BGP. Translating BGP route attribute information into any other protocol would likely cause a loss of information. Therefore, the EBGP information that RTR-A receives is not translated; it is just forwarded to other BGP-speaking routers (RTR-D in the figure) within the AS.

Likewise, RTR-D has BGP information and can propagate it to R-14 in AS 14 over the EBGP session.

EBGP sessions are, in general, established between directly connected neighbors. BGP-speaking routers, therefore, need no additional routing information to establish a session.

Internal Route Propagation

This topic describes internal route propagation within a BGP AS.

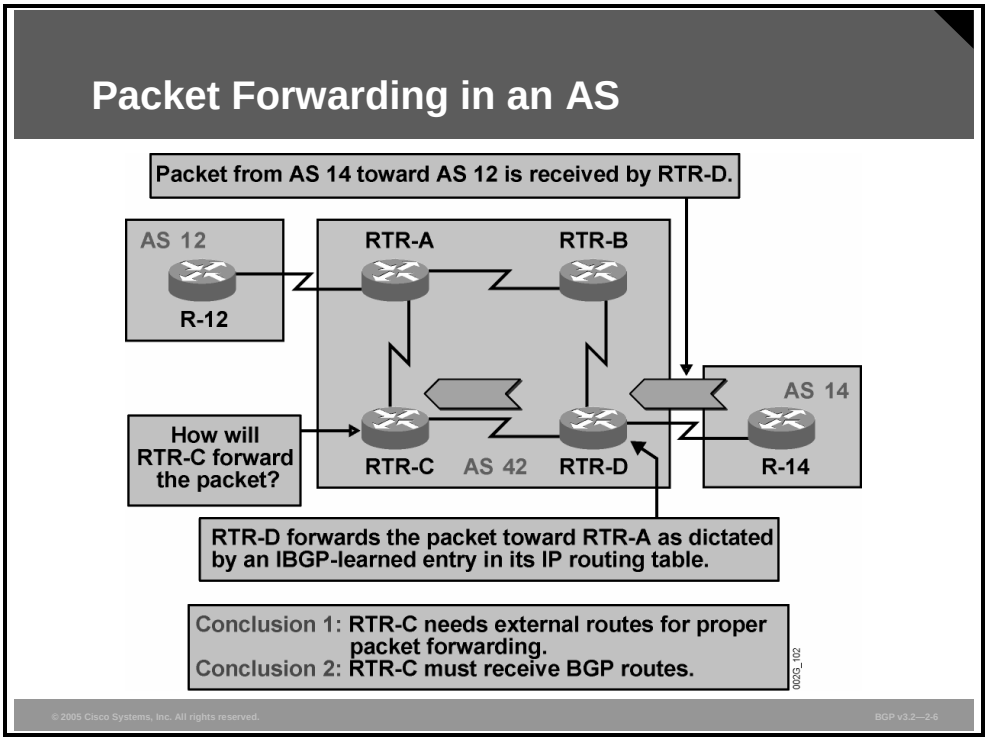


In this example, the BGP session between RTR-A and RTR-D, which are both in the same AS, is an IBGP session.

IBGP sessions are, in general, established between distant routers in the same AS. These routers need additional routing information to establish the session, because there is no requirement that IBGP neighbors be directly connected. This information typically comes from the IGP, which is running within the AS independently of BGP.

Packet Forwarding in an AS

This topic describes how transiting packets are forwarded inside a transit AS.



In this example, after AS 14 has received the routing information about reachable networks inside AS 12, IP packets can start to flow (in the figure, from AS 14 toward AS 12). R-14, the egress router in AS 14, forwards IP packets with destinations in AS 12 toward RTR-D, according to information received through EBGP.

RTR-D now uses the IBGP information that it received from RTR-A and forwards the packets in the direction of RTR-A, which in this case means via RTR-C.

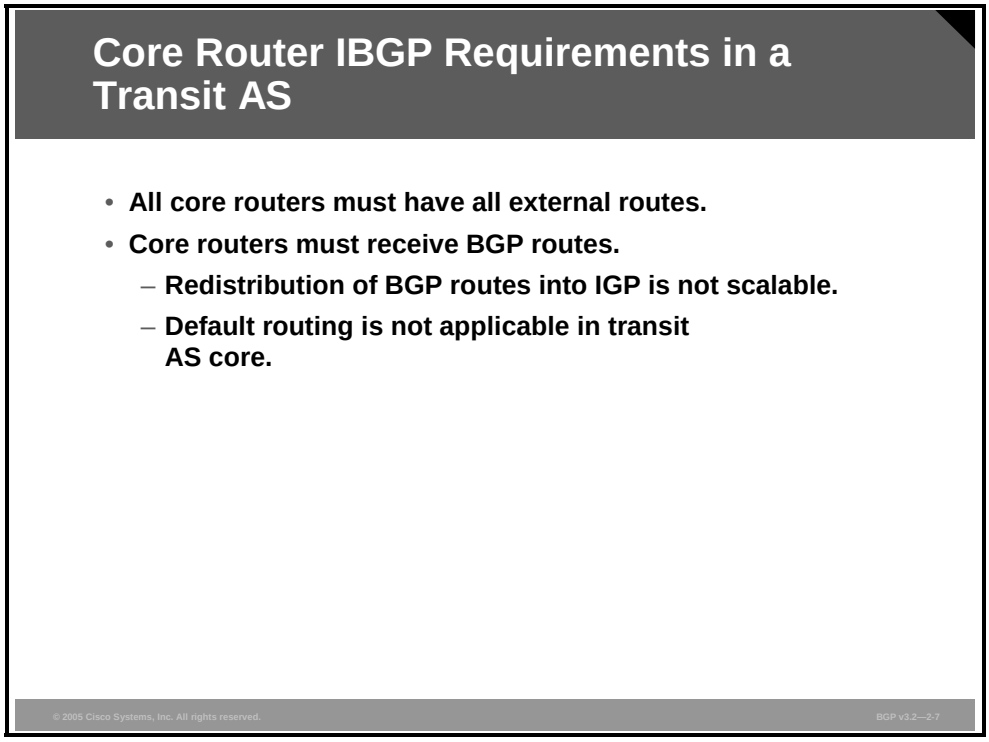
When the IP packets reach RTR-C, the router checks its routing table for a matching entry, but it fails to find one. The packet is dropped because the destination network is unreachable from the perspective of RTR-C.

This situation is, of course, unacceptable. To prevent dropped packets resulting from unreachable networks, RTR-C must also have routing information about the networks reachable inside AS 12. The same information that RTR-D received from RTR-A over the IBGP session must be propagated to RTR-C.

Note	RTR-B has the same network reachability requirements as RTR-C, because RTR-D could forward the packets via RTR-B as well as via RTR-C.

Core Router IBGP Requirements in a Transit AS

This topic describes the need for deploying IBGP on all core routers.



Within a transit AS, all routers that are in a theoretical transit path between external destinations should have information about all external routes that are received from any neighboring AS. If a single router on a transit path does not have this information, there is always a possibility that an IP packet that is received from a neighboring AS will not be able to be forwarded by that router through the transit AS. The router lacking routing information about the final destination of the IP packet drops it into what effectively becomes a black hole.

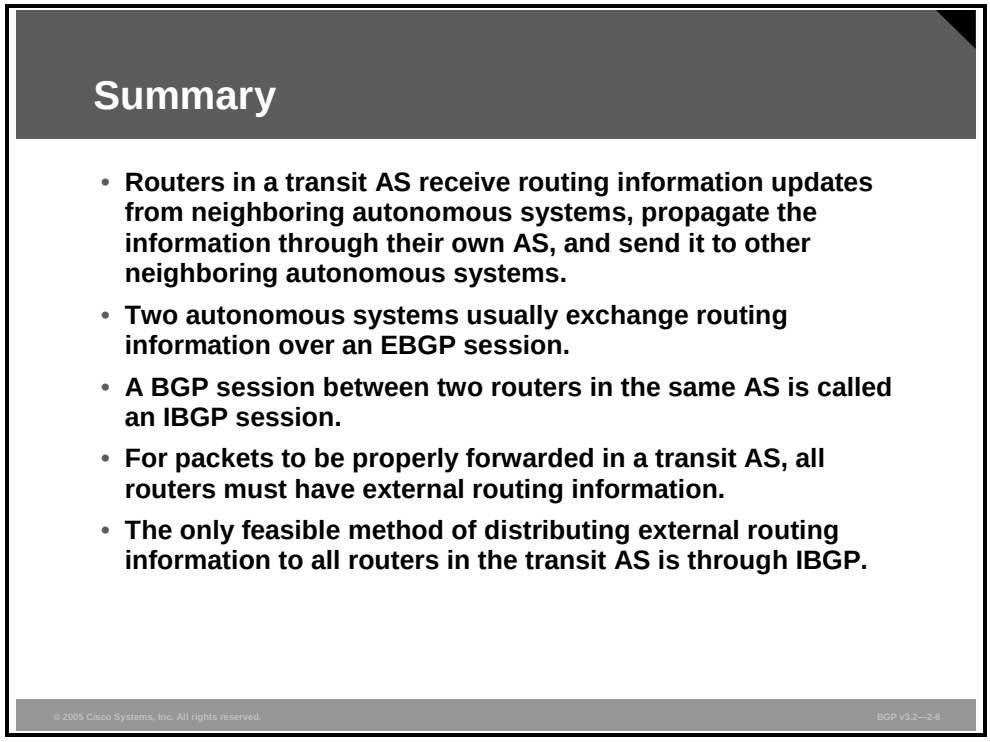
The only feasible way for the router to distribute all external routing information is by using IBGP. Redistribution of the EBGP routes into an IGP is not viable because no IGP can carry the volume of information that BGP currently carries in the Internet.

Note The risk of losing information during redistribution of EBGP routes into an IGP is not the reason why BGP is used to update intermediate routers in the transit path instead of an IGP. Redistribution into an IGP is not used because of the scalability issues that would arise from doing so.

Default routing or a gateway of last resort cannot be used by routers within the transit path when transit services are provided to other autonomous systems. If some routes were to be filtered out and the default route used instead, full routing flexibility would be lost. The transit AS would not be able to forward packets to all destinations at all times. In fact, routing loops and black holes might be easily introduced.

Summary

This topic summarizes the key points discussed in this lesson.



Lesson 2

Interacting with IBGP and EBGP in a Transit AS

Overview

Configuring a Border Gateway Protocol (BGP) network in a transit services configuration requires special care to ensure consistency of routing information throughout the autonomous system (AS). Understanding the interaction between External Border Gateway Protocol (EBGP) and Internal Border Gateway Protocol (IBGP) is crucial to successfully configuring and troubleshooting the transit autonomous network.

This lesson introduces the requirements of IBGP and describes how routers residing in the transit AS process the next-hop attribute. Changes to the normal processing of the next-hop attribute are also described in this lesson. The lesson concludes with a comparison between EBGP and IBGP.

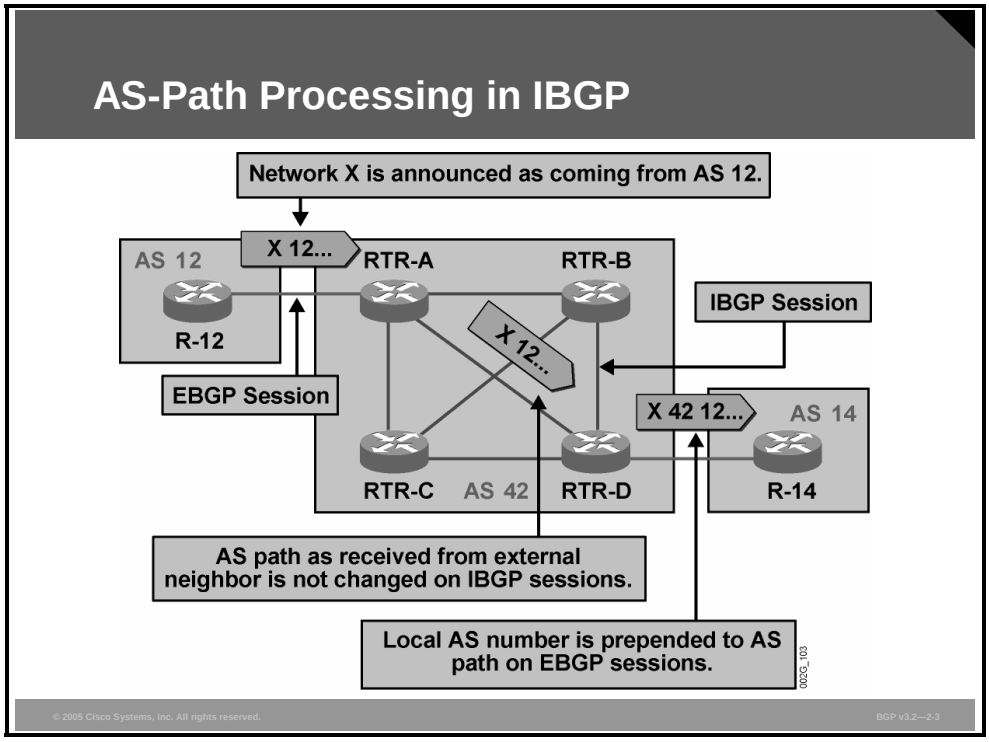
Objectives

Upon completing this lesson, you will be able to describe the interaction in a transit AS between EBGP and IBGP in relation to relevant attributes. This ability includes being able to meet these objectives:

- Describe AS-path processing in IBGP
- Describe BGP multipath load sharing
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers and the implications of that need
- List the benefits of establishing IBGP neighbor sessions using loopback interfaces
- Describe next-hop processing in IBGP
- Explain why all EBGP peers must be reachable by all BGP-speaking routers within the AS
- Describe how to configure edge routers to announce themselves as the next hop in IBGP updates
- Describe the differences between EBGP and IBGP sessions

AS-Path Processing in IBGP

This topic describes AS-path processing in IBGP.



All BGP routing updates carry the mandatory well-known attribute AS-path, which lists the autonomous systems that the routing update has already crossed.

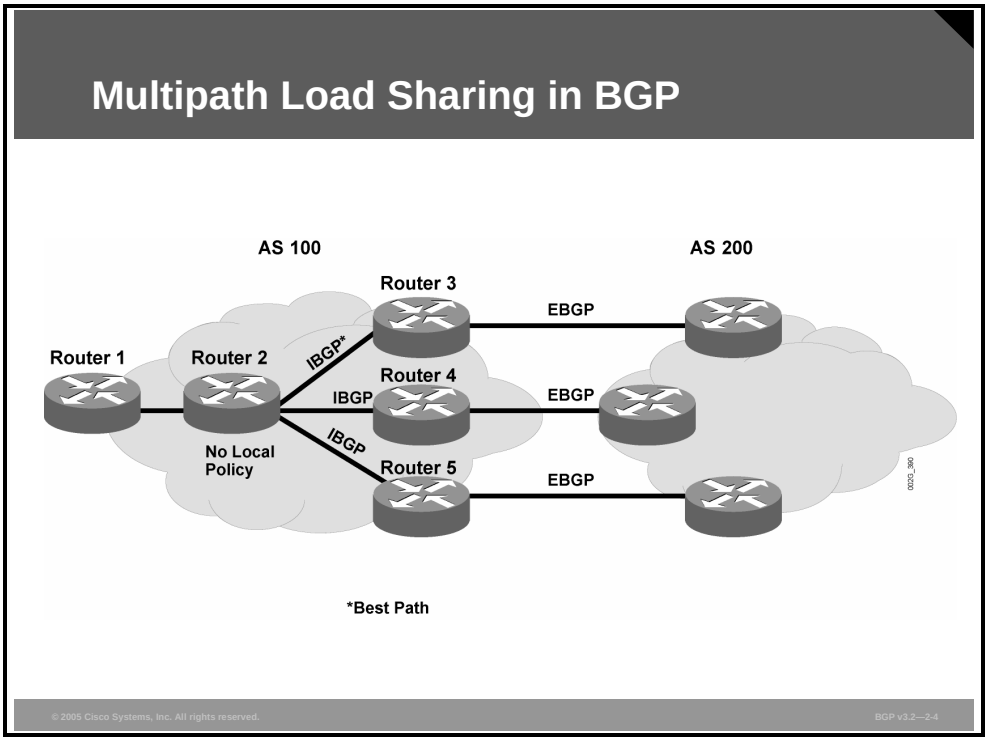
When a router originates a BGP prefix (network X in this example), the AS path is empty. Whenever a BGP prefix is announced over an EBGP session, the AS number of the router that is sending the information is prepended to the AS path. In the example, R-12 inserts “12” in the AS path before forwarding the routing update to RTR-A.

The AS path is not changed when the BGP prefix is propagated across IBGP sessions because the routing update has not crossed an AS boundary. In the figure, RTR-A forwards the information over an IBGP session to RTR-D with the AS path unchanged. The AS-path information about network X will be the same in all routers within AS 42, because all the routers are updated using IBGP sessions from RTR-A.

When RTR-D forwards the information about network X to R-14, RTR-D prepends its own AS number (42) to the AS path. Thus, R-14 receives the routing information about network X with an AS-path attribute of “42 12.”

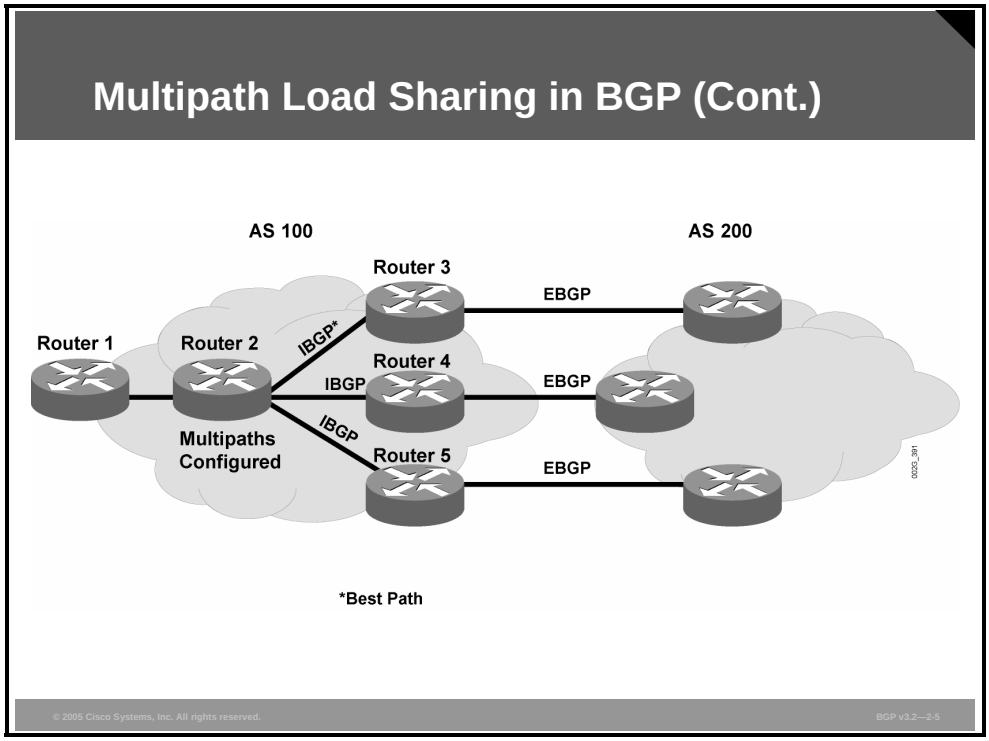
Multipath Load Sharing in BGP

This topic describes multipath load sharing in BGP.



When a BGP-speaking router with no local policy configured receives Network Layer Reachability Information (NLRI) from multiple IBGP sources for the same destination, the router chooses one IBGP path as the best path. The best path is then installed in the IP routing table of the router. For example, the figure illustrates that with three paths to AS 200, router 2 determines that one of the paths to AS 200 is the best path and uses only this path to reach AS 200.

The IBGP multipath load-sharing feature enables the BGP-speaking router to select multiple IBGP paths as the best paths to a destination. The best paths, or multipaths, are then installed in the IP routing table of the router.



For example, on router 2 in the figure, the paths to routers 3, 4, and 5 are configured as multipaths and can be used to reach AS 200, equally sharing the load to AS 200.

For multiple paths to the same destination to be considered as multipaths, the following criteria must be met:

- All attributes must be the same. The attributes include weight, local preference, AS path (entire attribute and not just length), origin code, multi-exit discriminator (MED), and IGP distance.
- The next hop router for each multipath must be different.

Even if the criteria are met and multiple paths are considered multipaths, the BGP-speaking router still designates one of the multipaths as the best path and advertises this best path to its neighbors.

Configuring multiple IBGP best paths enables a router to evenly share the traffic destined for a particular site.

maximum-paths ibgp

To control the maximum number of parallel internal BGP routes that can be installed in a routing table, use the **maximum-paths ibgp** command in router configuration mode.

- **maximum-paths ibgp** *maximum-number*

To disable this feature, use the **no** form of this command.

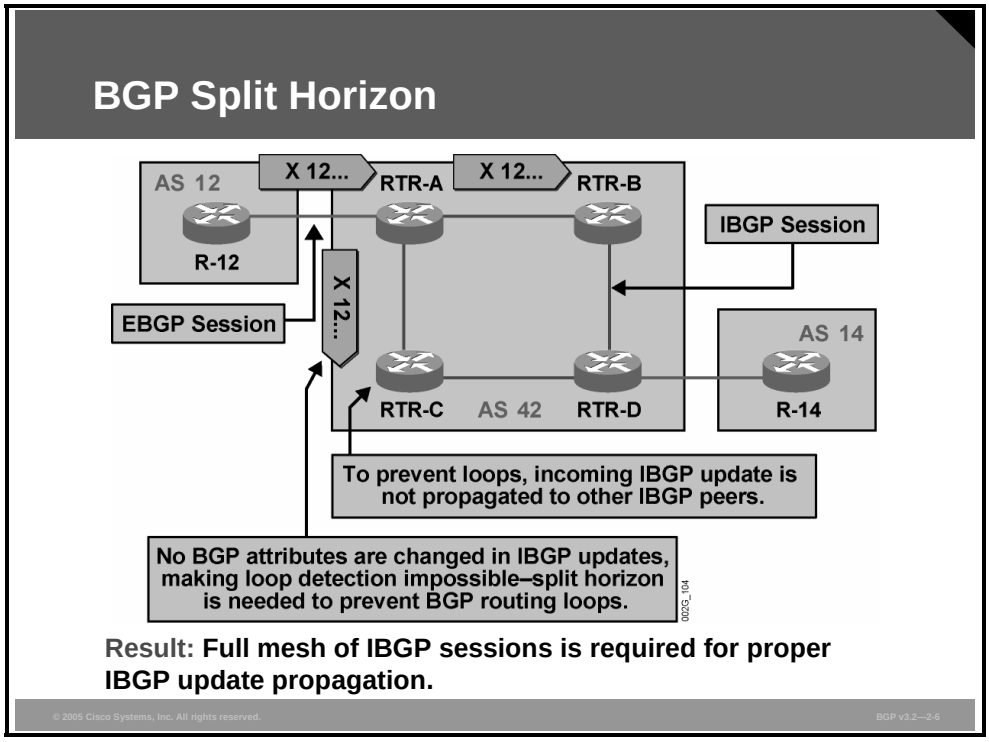
- **no maximum-paths ibgp**

Syntax Description

Parameter	Description
maximum-number	A number from 1 to 6. The maximum number of parallel routes that an IP routing protocol installs in a routing table.

BGP Split Horizon

This topic explains the need for BGP split horizon as a mechanism to prevent routing loops.



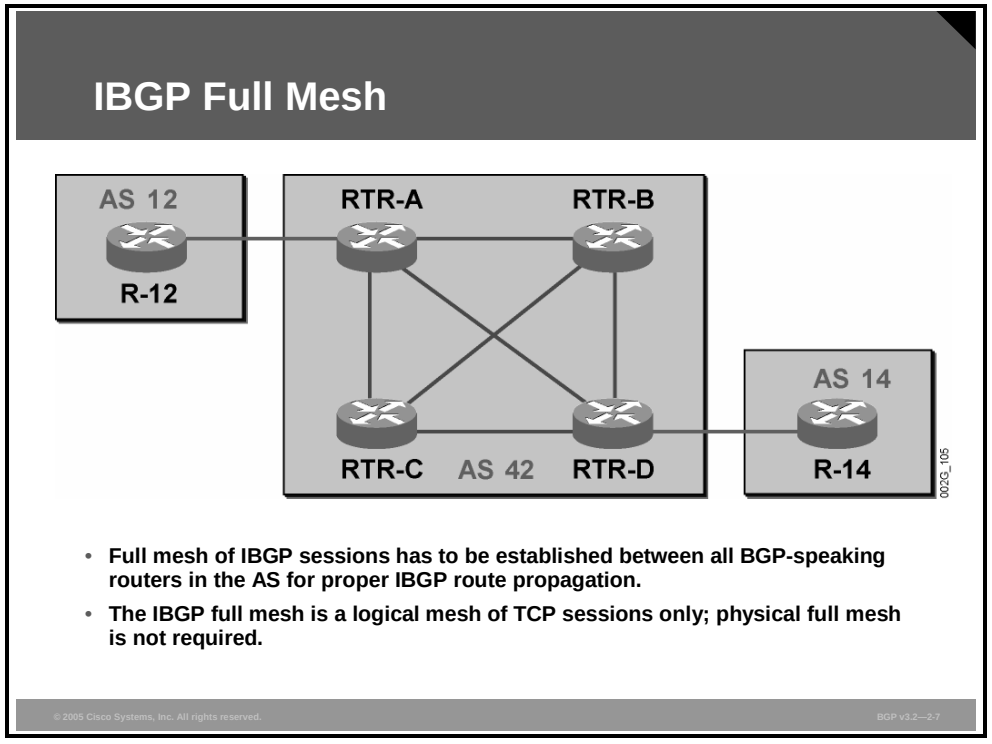
All routers within an AS must make routing decisions in a consistent way. They must have access to the same routing information with the same attributes in order to come to the same conclusion about which exit point of the AS to use. In other words, the BGP attributes should not be changed within the AS.

The AS-path attribute is not changed over an IBGP session, because the BGP update has not crossed the AS boundary. However, the AS-path attribute is the primary means of detecting routing information loops—a BGP router that encounters its own AS in the AS path of an incoming BGP update silently ignores the information. Because the AS path is modified by BGP-speaking routers only on EBGP sessions, this loop-preventing mechanism is useful between autonomous systems only, not within them.

Routing information loops within the AS are prevented by IBGP split horizon—routing information that is received through an IBGP session is never forwarded to another IBGP neighbor, only toward EBGP neighbors. Because of BGP split horizon, no router can relay IBGP information within the AS—all routers must be directly updated from the border router that received the EBGP update.

IBGP Full Mesh

This topic describes the need for a full-mesh topology between IBGP routers and the implications of this need.



Because every router on the transit path within the AS must have routing information about all external networks that are received by any of the border routers, RTR-B and RTR-C must lave IBGP sessions to all border routers. This level of communication is not enough, though, because any of the internal routers could also create new BGP routing information (for example, originate a customer network). These updates must also reach all the routers within the AS. The conclusion is that all BGP routers within an AS must have IBGP sessions with every other BGP router in the AS, resulting in a full mesh of BGP sessions between BGP-speaking routers in an AS.

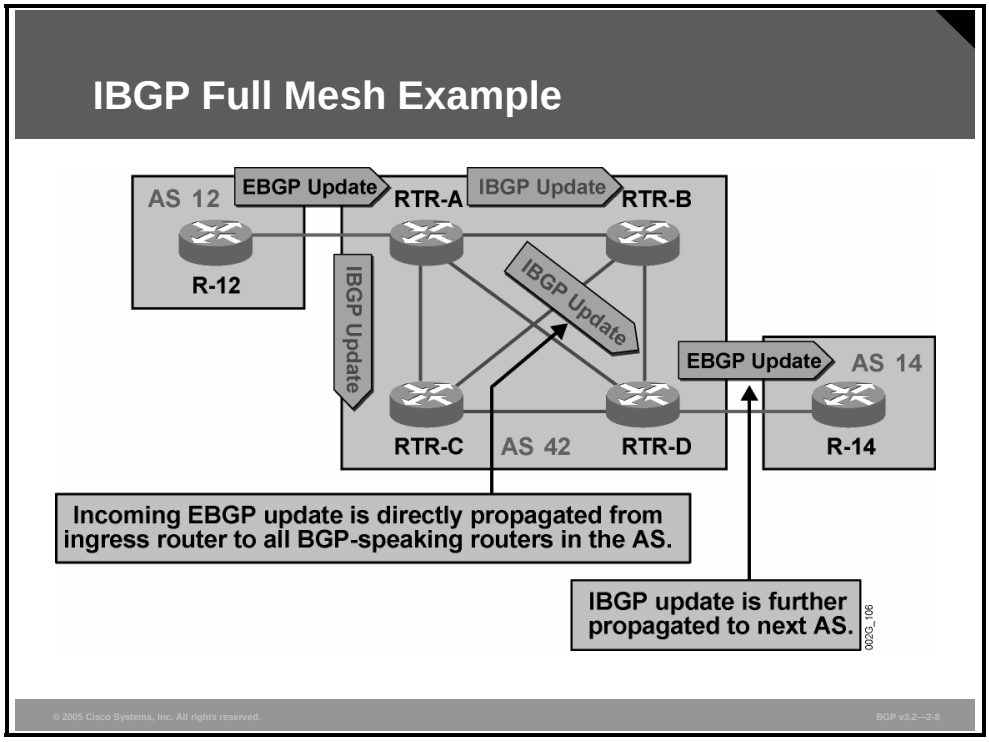
In the network shown in the figure, RTR-A must have IBGP sessions with RTR-B, RTR-C, and RTR-D to propagate routes that are received from AS 12 to all routers within AS 42. Similarly, RTR-D must have IBGP sessions with RTR-A, RTR-B, and RTR-C to be able to propagate routes that are received from AS 14 to all routers within AS 42.

Note The IBGP session between RTR-B and RTR-C is not strictly necessary for proper forwarding of IP packets between external destinations. It does become mandatory if RTR-B or RTR-C starts to originate BGP networks. To prevent potential future connectivity issues, it is a good practice to establish a full mesh of IBGP sessions regardless of whether they are needed at the time of network deployment or not.

The IGP that runs within AS 42 provides enough information to any BGP router within AS 42 to send IP packets to any other router in the AS. Having enough router reachability information makes it possible to establish IBGP sessions between routers even though they are not physically connected. The IBGP full mesh is a logical full mesh of TCP sessions and will run on an arbitrary physical topology.

Example: IBGP Full Mesh

The figure illustrates IBGP split-horizon and IBGP full-mesh principles in a sample network.



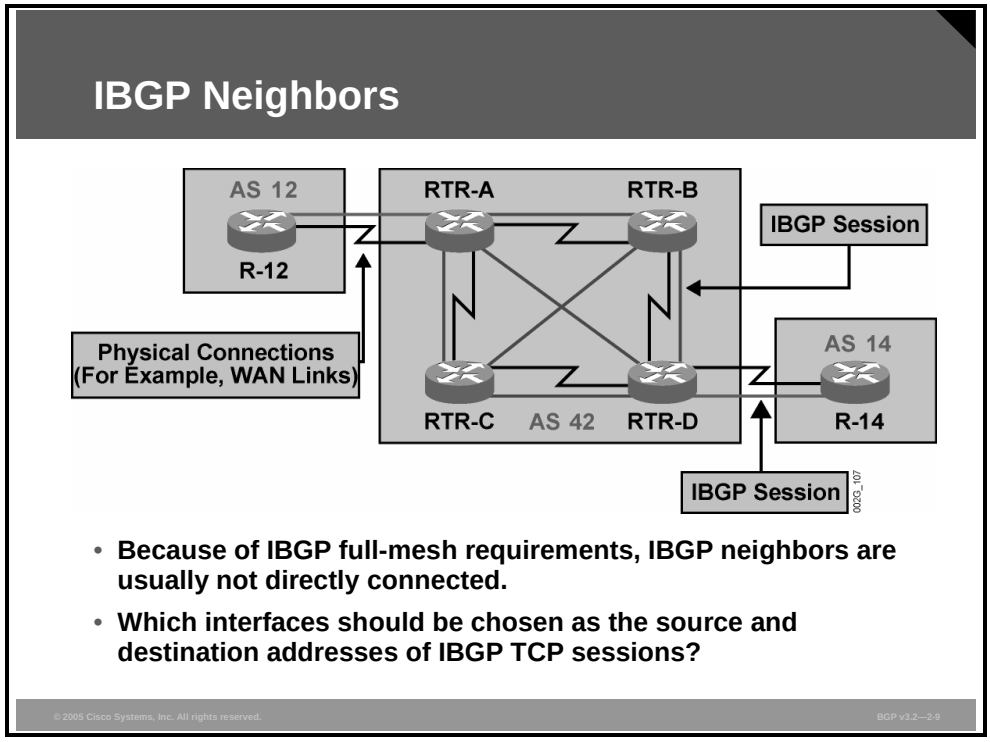
R-12 is sending an update to RTR-A over an EBGP session. Updates that are received on an EBGP session should be forwarded on all other IBGP sessions, so RTR-A updates RTR-B, RTR-C, and RTR-D. All routers within AS 42 are updated directly by RTR-A.

RTR-B and RTR-C are prevented from forwarding the update that they received from RTR-A because of BGP split horizon.

RTR-D, which received the information on an IBGP session, is prevented from updating RTR-B and RTR-C because of the same split-horizon rule. But RTR-D will update R-14 over an EBGP session.

IBGP Neighbors

This topic lists the benefits of establishing IBGP neighbor sessions using loopback interfaces.



In the figure, the transit AS 42 has a redundant physical topology. The IGP provides reachability information for all routers and networks within AS 42, allowing all routers in the AS to establish IBGP sessions to all other routers, even if the routers are not directly connected.

If the IBGP session between RTR-A and RTR-D was established using IP addresses that belong to the physical WAN interfaces, the IBGP session would go down if either of the WAN interfaces went down. As a result, the router would tear down the TCP session that is used for BGP between the routers because the IP address of an interface that is in the down state is invalid. Subsequently, all IP packets that are received with a destination address pointing to that interface will also be dropped.

Network designers must be careful during the network design and implementation phase that those IBGP sessions remain established for as long as the two BGP routers have any usable path between them.

IBGP Neighbors (Cont.)

Always run IBGP sessions between loopback interfaces.

- **IBGP sessions can always be established, even if some physical interfaces are down.**
- **IBGP sessions are stable—physical interface failure will not tear down IBGP sessions.**
- **There is no BGP recovery after a failure inside the transit AS.**
  - **The configured IGP will re-establish the path between loopback interfaces.**
  - **IBGP sessions are not affected.**

The best choice when you are configuring IBGP sessions is to establish each session between loopback interfaces on each BGP router.

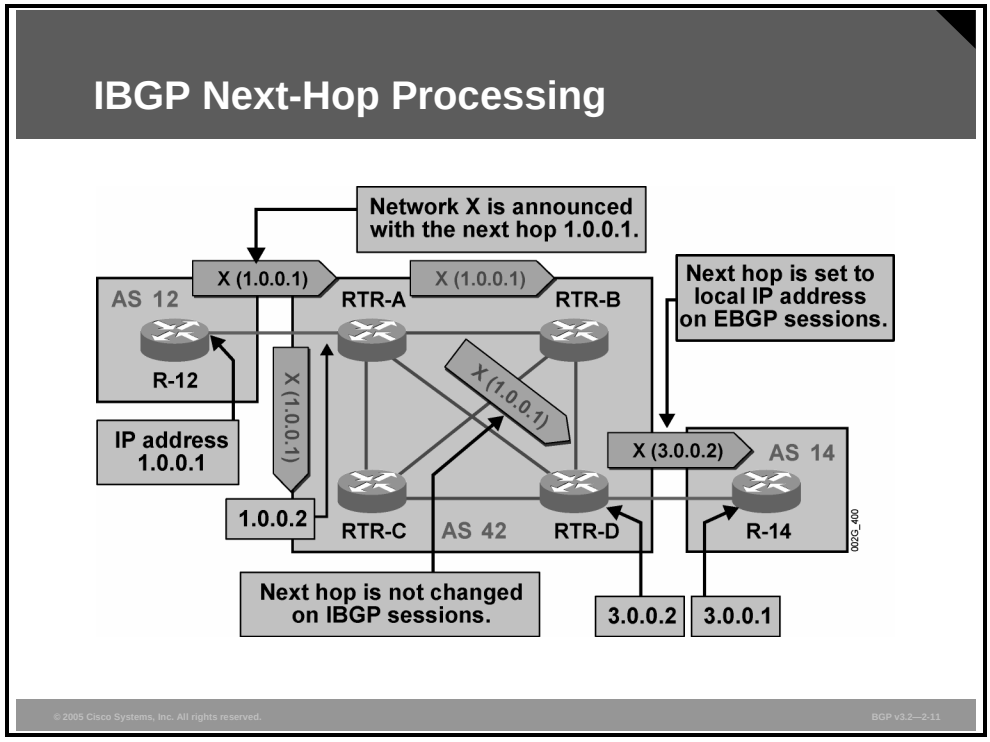
To establish BGP connectivity between the loopback interfaces, the IP addresses of these interfaces have to be reachable by both routers. It is important that the IGP carry information about the subnets that are assigned to each loopback interface so that the interfaces are reachable by all BGP routers in the AS.

The IBGP sessions that are established between loopback interfaces have increased stability. These sessions will not go down if a single physical interface goes down. As long as the IGP can find any path between the two routers, the IBGP session will remain up. BGP will not notice that the IGP has changed the traffic path between the two routers.

Note Because BGP sessions run over TCP, they can survive even a short loss of connectivity between BGP routers with no impact to the BGP routing protocol. The only requirement placed on the IGP is that the network must converge before the BGP keepalive timer expires.

IBGP Next-Hop Processing

This topic describes how the next-hop attribute is processed in IBGP.



Every BGP update carries the mandatory well-known attribute next-hop, which specifies the IP address that should be used by the router as the forwarding next hop for packets sent toward the announced destination address. In most cases, the next hop is set to the IP address that the sending router is using as its source IP address for EBGP sessions. The receiving BGP router will use the information and route IP packets toward the announced destination via the indicated next hop, which is normally directly connected.

The next-hop attribute is not changed on IBGP updates, meaning that when the border router forwards the BGP update on IBGP sessions, the next-hop address is still set to the IP address of the far end of the EBGP session. Therefore, the receiver of IBGP updates will see the next-hop information indicating a destination that is not directly connected. To resolve this problem, the router will check its routing table and see if and how it can reach the next-hop address. The router can then route IP packets with destination addresses matching the network in the BGP update in the same direction as it would have routed an IP packet with a destination address equal to the IP address stated in the next-hop attribute. This process is known as recursive routing.

In the figure, R-12 sends a BGP update about network X. Because it is sending this update over an EBGP session to RTR-A, the next-hop attribute is set to the IP address that is used at the R-12 side of the EBGP session, 1.0.0.1.

RTR-A can use this information and route packets to network X by forwarding them to R-12.

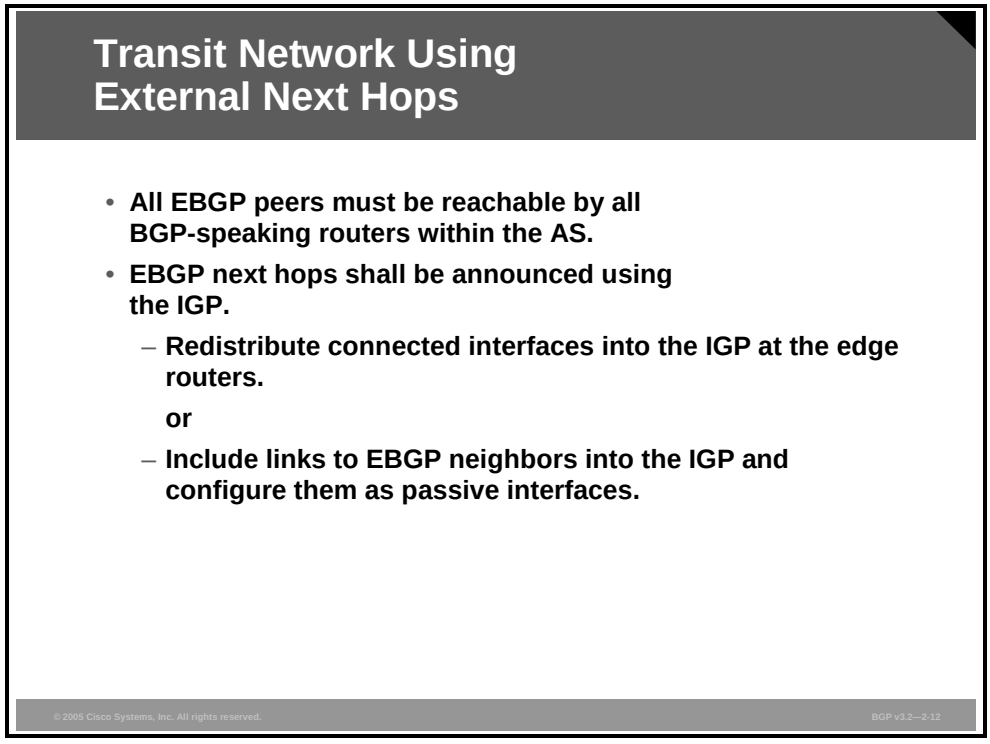
RTR-A also forwards the BGP update over all its IBGP sessions. It does not change the next-hop attribute, so RTR-B, RTR-C, and RTR-D get information that they can reach network X by forwarding packets to 1.0.0.1. But that IP address is not directly connected, so the routers must

look in their routing tables to see if and how they can reach 1.0.0.1. If the recursive route lookup is successful, each router can then route packets to network X in the same direction as it would route packets to 1.0.0.1.

RTR-D also forwards the BGP update about network X to R-14. The connection between these routers is an EBGP session, meaning that RTR-D sets the next-hop attribute to its own IP address, 3.0.0.2, which is used by RTR-D on the EBGP session toward R-14.

Transit Network Using External Next Hops

This topic describes why all EBGP peers must be reachable by all BGP-speaking routers within the transit AS.



All BGP-speaking routers within the AS get information about external networks with the next-hop attribute, which is set to the far end of the EBGP sessions reaching the border routers of the AS.

Routers use a recursive routing mechanism when they determine how to forward IP packets toward external destinations. When BGP routes are used in the routing table, the router checks how it would have reached the next-hop address, and it installs the BGP route with the same forwarding indication as for the route that is used to reach the next-hop IP address.

To get the recursive routing to work, the router must resolve all possible next-hop references that use information in the routing table, which is already there. The IGP that is used within the AS must carry this information.

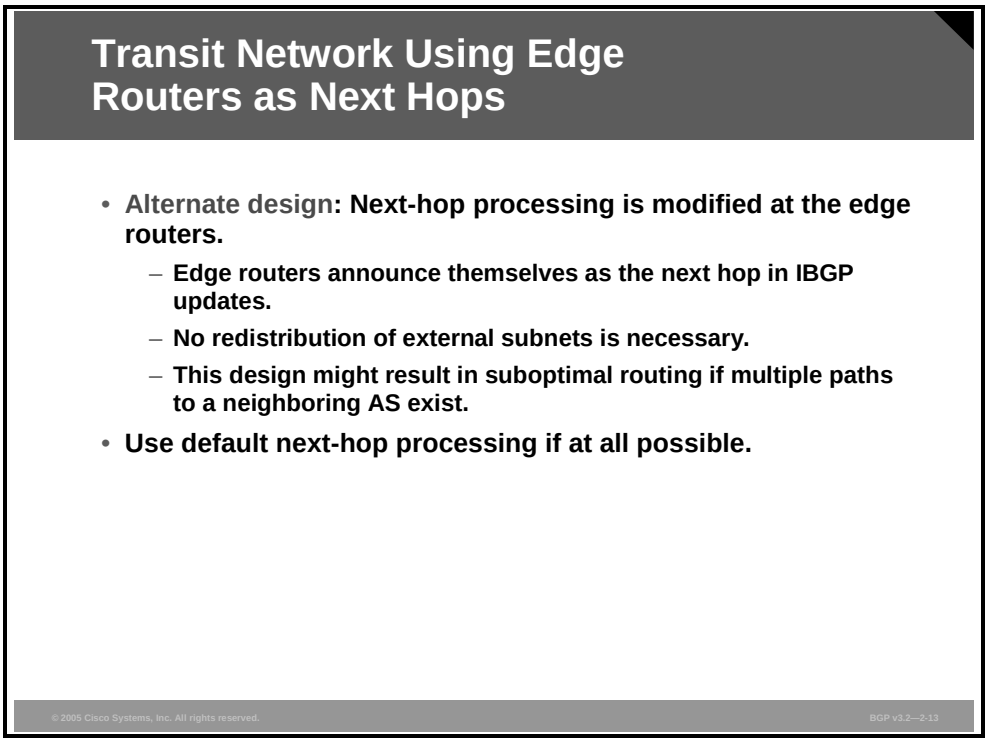
One way of making the IGP carry the information that is necessary to resolve the BGP next-hop addresses is to make sure that all the border routers, which contain the EBGP sessions, redistribute connected subnets into the IGP using the **redistribute connected** routing protocol configuration command. Because EBGP sessions are established between routers using a directly connected interface, the far end of the EBGP sessions is an IP address within the directly connected subnet. By redistributing the connected interfaces into the IGP, the border routers allow next-hop references to be resolvable by all routers within the AS.

External subnets that are redistributed into the IGP might appear as external IGP routes, depending on what IGP is configured within the AS. There are several scalability issues that are associated with external routes in some routing protocols. For example, Open Shortest Path First (OSPF) carries each external subnet in a separate link-state advertisement (LSA) object. If

route redistribution is not desirable for any reason, an alternative method is to include the subnet on which the EBGP session is running in the IGP configuration using the **network** command. To prevent the border router from exchanging IGP routing with the border router of the other AS, you must configure the interface as a passive interface. Failure to do so could cause the two autonomous systems to exchange routes using the IGP. In that case, all benefits of having separate autonomous systems would be lost.

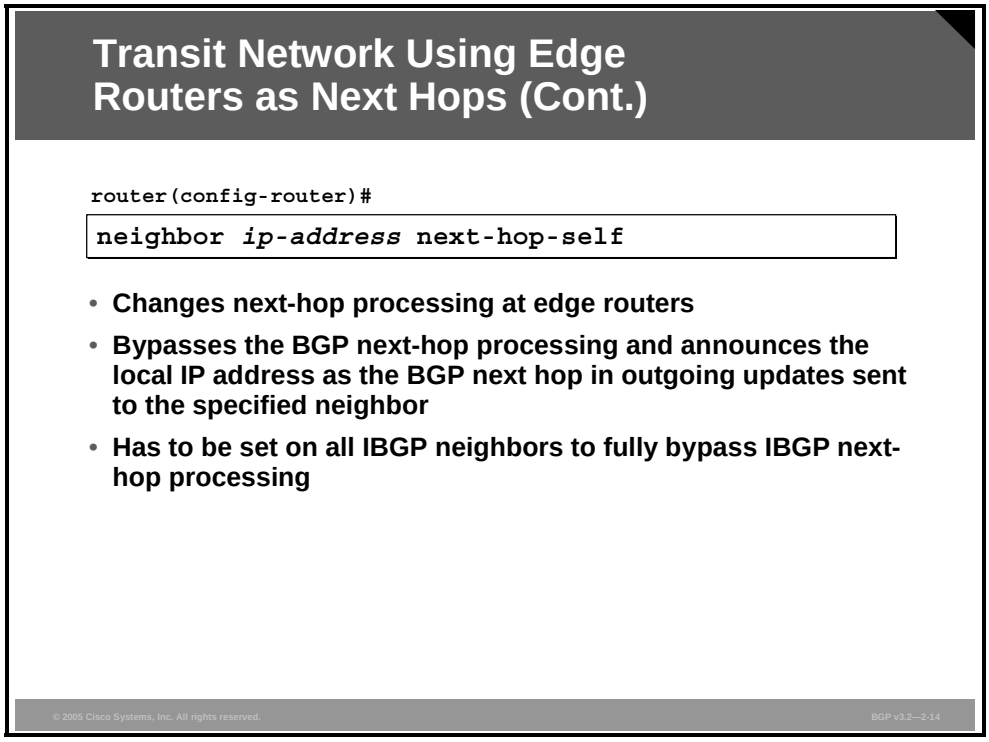
Transit Network Using Edge Routers as Next Hops

This topic describes how to configure edge routers to announce themselves as the next hop in IBGP updates.



The next-hop attribute is usually not modified by an IBGP peer when the BGP update is propagated across IBGP sessions. However, you could configure the BGP router to have a different behavior and set its IP address as the next-hop address even when the BGP updates are sent across IBGP sessions (emulating behavior on EBGP sessions). If you do configure an IBGP router to emulate the behavior of EBGP sessions on the IBGP sessions of the border routers, the BGP updates that are received on the EBGP sessions will be forwarded on the IBGP sessions and the next-hop attribute will be set to the IP address that is used on the local side of the IBGP session. The original next hop, set by the far end of the EBGP session, will be lost.

The receiver of the IBGP information will do recursive routing in the normal way. But the next-hop address that is used will be the IP address of the far end of the IBGP session, because the border router has changed it. The IP address of the far-end IBGP peer is always known in the routing table; otherwise, the IBGP session would not have been established. There is no need for the receiver of the IBGP information to have knowledge of how to reach the far end of the EBGP session, because that IP address is no longer set as the next hop.



neighbor next-hop-self

To configure the router as the next hop for a BGP-speaking neighbor or peer group, use the **neighbor next-hop-self** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **next-hop-self**

To disable this feature, use the **no** form of this command.

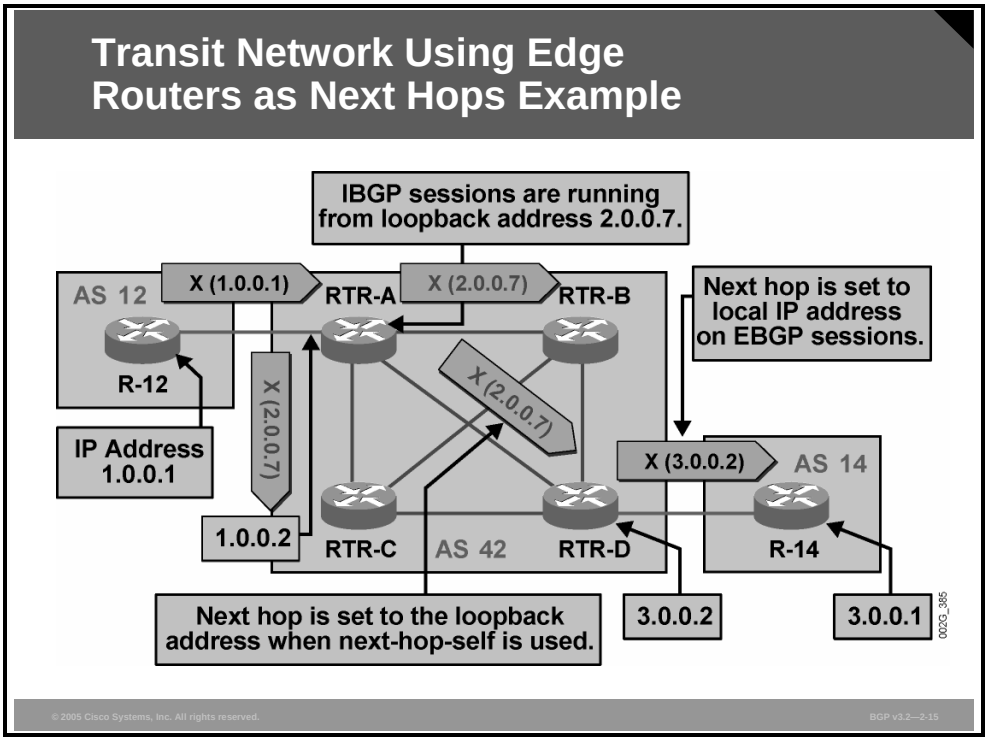
- **no neighbor** {*ip-address* | *peer-group-name*} **next-hop-self**

Syntax Description

Parameter	Description
ip-address	IP address of the BGP-speaking neighbor
peer-group-name	Name of a BGP peer group

Example: Transit Network Using Edge Routers as Next Hops

In the figure, the **next-hop-self** configuration has been used on all IBGP sessions.



The next-hop attribute is normally not changed on IBGP updates. When the border router forwards the incoming EBGP update over an outgoing IBGP session, the border router changes the next-hop address to the IP address that is used as the source address of the IBGP session.

The receiver of IBGP updates will see next-hop information that indicates a destination, which might not be directly connected. To resolve this problem, it will check its routing table and see if and how the next-hop address can be reached. Then it will route IP packets with destination addresses that match the network in the BGP update in the same direction as it would have routed an IP packet with the destination address equal to the IP address in the next-hop attribute. In this case, it is obvious that the next-hop address can be reached, because the IBGP session would not have been established otherwise.

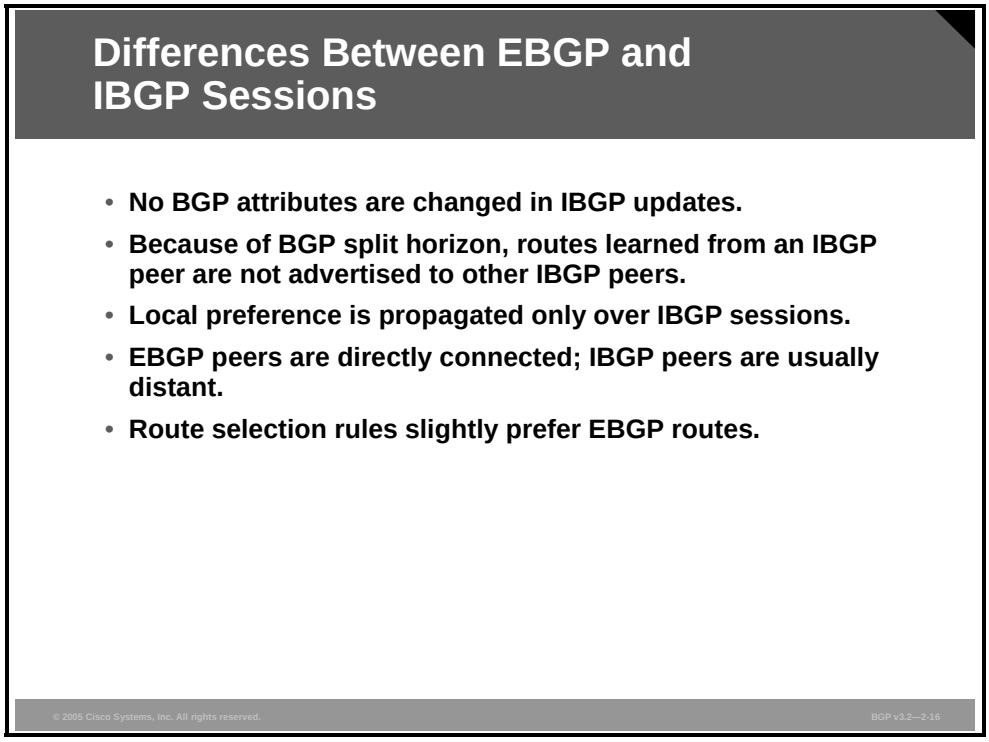
In the figure, R-12 sends a BGP update about network X. Because it is sending a BGP update over an EBGP session to RTR-A, the next-hop attribute is set to the IP address that is used at the R-12 side of the EBGP session, 1.0.0.1. RTR-A can use this information and route packets to network X by forwarding them to R-12.

RTR-A also forwards the BGP update on all its IBGP sessions. It changes the next-hop attribute to the IP address of its own loopback interface, so RTR-B, RTR-C, and RTR-D will get information that they can reach network X by forwarding packets to 2.0.0.7. But that address is not directly connected. The routers will inspect the routing table to see if and how they can reach 2.0.0.7. They can then route packets to network X in the same direction that they would use to route packets to 2.0.0.7.

RTR-D also forwards the BGP update about network X to R-14. This is an EBGP session, which means that RTR-D will set the next-hop attribute to its own IP address that is used on that EBGP session, 3.0.0.2.

Differences Between EBGP and IBGP Sessions

This topic describes the differences between EBGP and IBGP sessions.

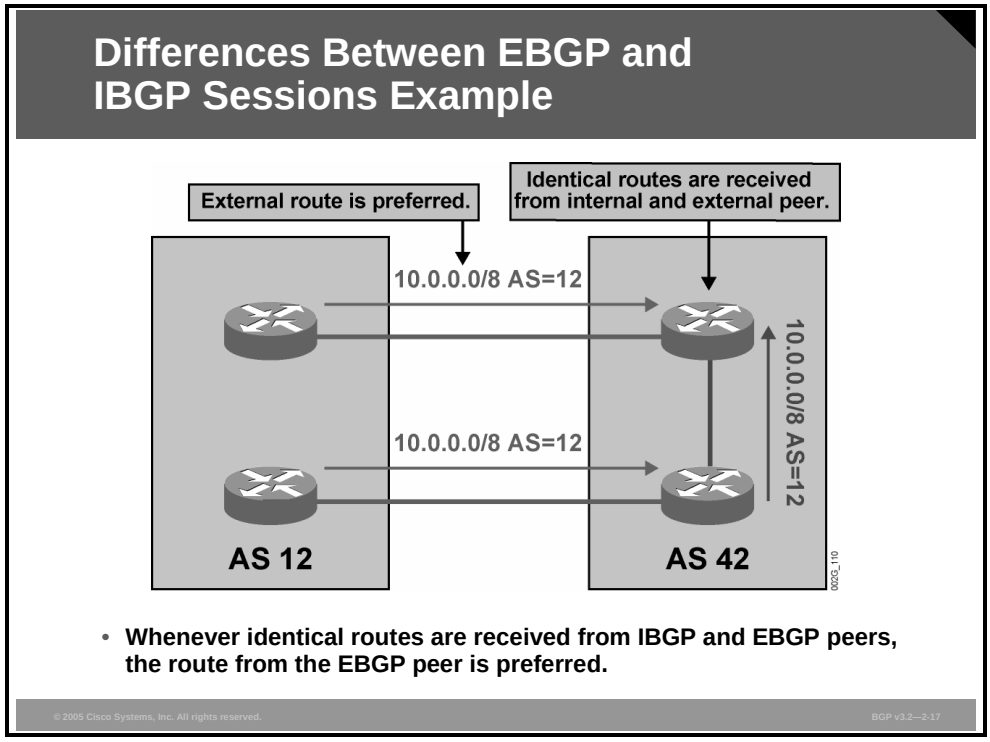


Both EBGP and IBGP sessions forward BGP updates; however, they do it in slightly different ways:

- The router does not change BGP attributes when an update is sent across an IBGP session, unless **next-hop-self** is configured. When a BGP-speaking router sends an update across an EBGP session, the next-hop attribute is always set and the AS number of the router is prepended to the AS-path attribute.
- IBGP uses split horizon to prevent routing information loops. EBGP does not use split horizon and instead uses the AS path to detect loops. In both cases, a router forwards only the best route and never sends a route back on the session from which it was received. But IBGP split-horizon rules also prohibit a router from forwarding any information that is received on an IBGP session to another IBGP session.
- IBGP border routers remove the local preference attribute from a BGP route before the BGP update is sent over an EBGP session. This difference means that the local preference attribute is distributed on IBGP sessions only.
- Two routers with an EBGP session between them normally establish the session using the IP addresses from a common, shared subnet. Using the shared subnet to establish the session guarantees that the two routers can exchange IP packets without any IGP running between them. Also, recursive routing will always succeed because the next-hop address is reachable using a directly connected route.
- IBGP sessions are normally established between all routers in the AS in a full mesh. But all routers in an AS might not have physical connections to every other router within the AS. Because IBGP sessions are established between routers using IP addresses of different subnets, an IGP must be running within the AS in order to establish IBGP sessions.
- BGP route selection rules slightly favor EBGP routes over equivalent IBGP routes.

Example: Differences Between EBGP and IBGP Sessions

This example illustrates the preference of the EBGP route.



One of the default goals of transit packet forwarding is to propagate the transit packet toward the downstream AS as soon as possible. A border router that receives otherwise equivalent routes to the same destination over both an EBGP session and an IBGP session will prefer the information that is received through the EBGP session.

Note	Equivalent routes are routes that have equal BGP path attributes used in the BGP route selection rules (weight, local preference, AS-path length, origin, MED).

In the figure, the upper router in AS 42 receives BGP updates about network 10.0.0.0/8 over two different paths. One update is received over the EBGP session to AS 12. The other update is received over the IBGP session to the lower router in AS 42. All essential attributes are the same, so route selection cannot be made easily.

The upper router in AS 42 realizes that IP packets with destination addresses within network 10.0.0.0/8 should sooner rather than later leave AS 42. It is better to make them leave the AS right away. So the update that was received on the EBGP session is preferred over the update that was received on the IBGP session.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- All BGP routing updates carry the mandatory well-known attribute AS-path, which lists the autonomous systems that the routing update has already crossed. The AS-path attribute is not changed when the BGP prefix is propagated across IBGP sessions.
- The IBGP multipath load sharing feature enables the BGP speaking router to select multiple IBGP paths as the best paths to a destination. The best paths, or multipaths, are then installed in the IP routing table of the router.
- Routing information loops within the AS are prevented by IBGP split horizon—routing information that is received through an IBGP session is never forwarded to another IBGP neighbor, only toward EBGP neighbors.
- All BGP routers within an AS must have IBGP sessions with every other BGP router in the AS, resulting in a full mesh of BGP sessions.

Summary (Cont.)

- For recursive routing to work, a router must resolve all possible next-hop references that use information in the routing table. The IGP that is used in the AS must carry this information.
- For stability, the best choice when you are configuring IBGP sessions is to establish the session between loopback interfaces of BGP routers.
- The next-hop attribute is typically set to the IP address that the sending router is using as its source IP address for an EBGP session. Recursive routing is done to resolve the next hop inside an AS because the next-hop attribute is not changed on IBGP updates.

Summary (Cont.)

- **You can configure an edge router to set its IP address as the next-hop address even when the BGP updates are sent across IBGP sessions. As a result, there is no need for the receiver of the IBGP information to know how to reach the far end of the EBGP session, because that IP address is no longer set as the next hop.**
- **Both EBGP and IBGP sessions forward BGP updates but in slightly different ways. BGP attributes are not changed when an update is sent across an IBGP session unless** next-hop-self **is configured.**

Lesson 3

Forwarding Packets in a Transit AS

Overview

A transit autonomous system (AS) requires interaction between External Border Gateway Protocol (EBGP) and Internal Border Gateway Protocol (IBGP) and between IBGP and an Interior Gateway Protocol (IGP) in the transit AS. This lesson describes packet forwarding through a transit AS and discusses the requirements for successful packet forwarding, such as recursive route lookup and an IGP in the transit AS. This lesson concludes with a discussion of the interaction between IBGP and an IGP running within the transit AS.

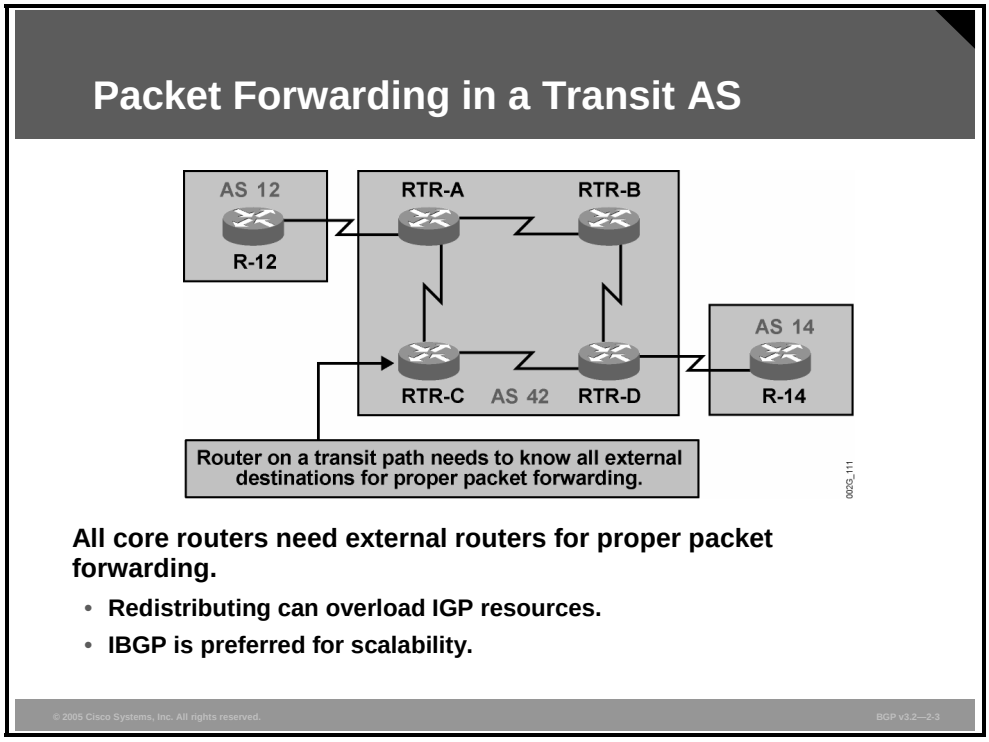
Objectives

Upon completing this lesson, you will be able to describe the function of an IGP in forwarding packets through an AS. This ability includes being able to meet these objectives:

- Describe packet forwarding in a transit AS
- Explain how recursive lookup functions in Cisco IOS software
- Explain the need for an IGP in a transit backbone that is running BGP on all routers
- Describe interactions between BGP and IGP in a transit AS
- Explain the potential problems that might arise from BGP and IGP interaction

Packet Forwarding in a Transit AS

This topic describes packet forwarding in a transit AS.



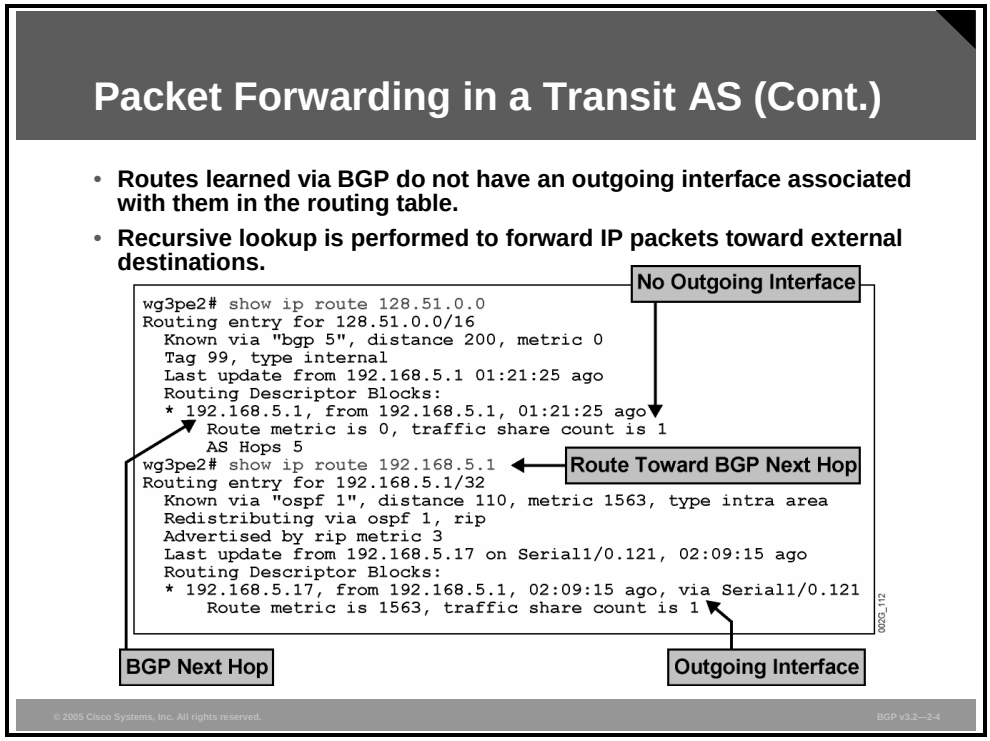
When Border Gateway Protocol (BGP) updates have propagated through the transit AS to all neighboring autonomous systems, the IP traffic can start to flow.

In the figure, Router R-14 forwards to RTR-D IP packets with the destination address matching a network in AS 12. RTR-D checks its routing table and finds that there is a BGP route for that destination. The BGP route has a next-hop reference, which points to the far end of the EBGP session between R-12 and RTR-A. So RTR-D once again checks the routing table and finds that it should forward the packet to RTR-C in this case.

Thus, RTR-C receives the IP packet with a destination address indicating a host within AS 12. To be able to forward this packet, RTR-C must have a matching route in its routing table. A default route or gateway of last resort is not appropriate because in the next instant RTR-C could receive another packet, coming from the other direction and destined for AS 14.

The conclusion is that both RTR-C and RTR-B, to handle all possible cases, must have routing information to all the external networks that RTR-A and RTR-D have. The only scalable way of providing routers with this information is to update RTR-C and RTR-B with IBGP from both RTR-A and RTR-D.

In theory, the external information that is received by RTR-A and RTR-D could be redistributed by these ingress routers into the IGP in use within the transit AS. However, no IGP can handle the volume of information that BGP can. So there would always be a risk that the IGP would break because of information overload, causing a total network meltdown in the AS. The volume of routing information that is carried by BGP in the contemporary Internet long ago passed the limits of what it is possible to carry in any IGP.

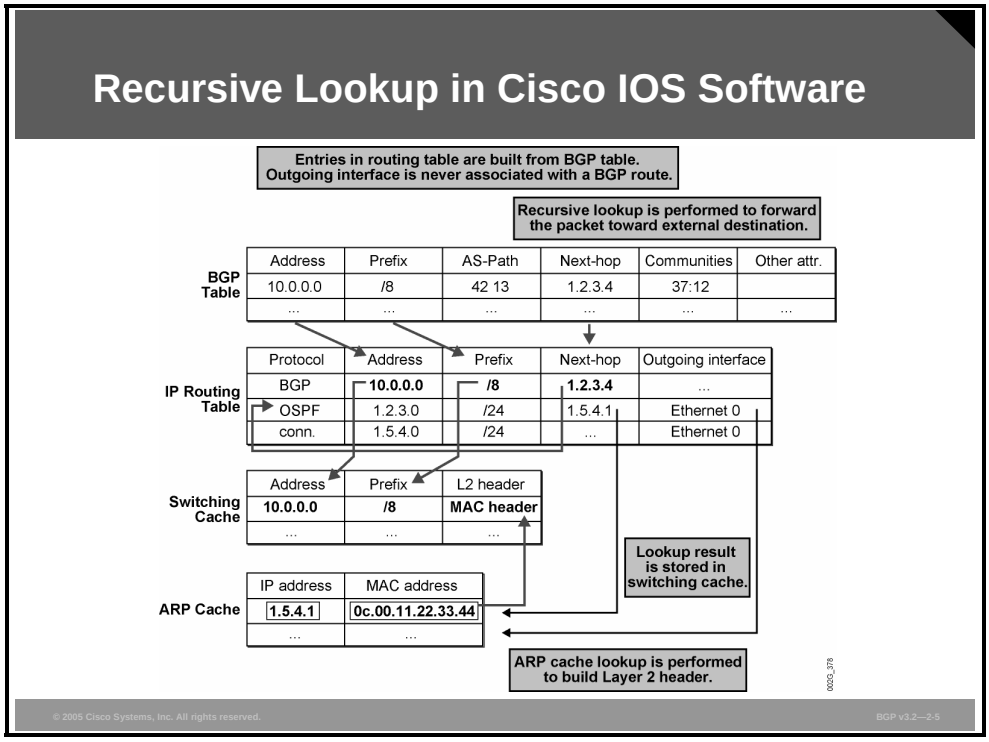


A BGP route is installed in the IP routing table of a router only if the IP address in the next-hop attribute is reachable according to the information already in the routing table. The installed BGP route contains a reference to that next-hop address. So, the network will be reachable via an IP address, which may or may not be directly connected. Because there is no clear reference to a physical interface, the BGP route is installed in the IP routing table without any information about outgoing interface.

The router must evaluate the recursive reference to the BGP next hop sooner or later in order to allow packet forwarding toward external destinations. The point in time when the recursive reference is resolved depends on the IP switching mechanism that is used by the router. At the latest, the router performs the recursive route lookup when an IP packet with a destination address that matches the BGP route should be forwarded. The router determines which outgoing interface should be used and which Layer 2 address to assign (if applicable). The router creates a cache entry so that successive IP packets to the same destination can be routed using the same outgoing interface and Layer 2 address.

Recursive Lookup in Cisco IOS Software

This topic describes how recursive lookup functions in Cisco IOS software.



The figure presents the steps in the recursive lookup process in Cisco IOS software. The router has received a BGP update about network 10.0.0.0/8. It was associated with an AS-path attribute set to 42 13, a next-hop attribute set to the IP address 1.2.3.4, and a community value 37:12. Some other attributes were also carried with the update.

Because the next-hop address 1.2.3.4 is reachable according to the routing table, the BGP route is also installed in the routing table. Network number, subnet mask, and next-hop attributes are inherited from the BGP table. No outgoing interface is assigned.

When an IP packet with a destination in network 10.0.0.0 is received, the router searches the routing table and finds the installed BGP route. The router takes the indicated next-hop address 1.2.3.4 and searches the routing table again. It now finds a match with the Open Shortest Path First (OSPF) route to subnet 1.2.3.0/24. The 1.2.3.0/24 route has an outgoing interface set to interface Ethernet 0 and a next hop set to 1.5.4.1, meaning that packets that are destined for network 10.0.0.0 should be forwarded via 1.5.4.1, which is directly reachable over Ethernet 0. The Address Resolution Protocol (ARP) table is used to find the MAC address for IP address 1.5.4.1. The MAC address is used to forward the IP packet to network 10.0.0.0 out the Ethernet 0 interface. The MAC header is stored in the cache for successive packets to network 10.0.0.0.

Note	The example illustrates the recursive lookup performed when the router uses cache-based IP switching mechanisms (for example, fast switching or optimum switching).

Recursive Lookup in Cisco IOS Software (Cont.)

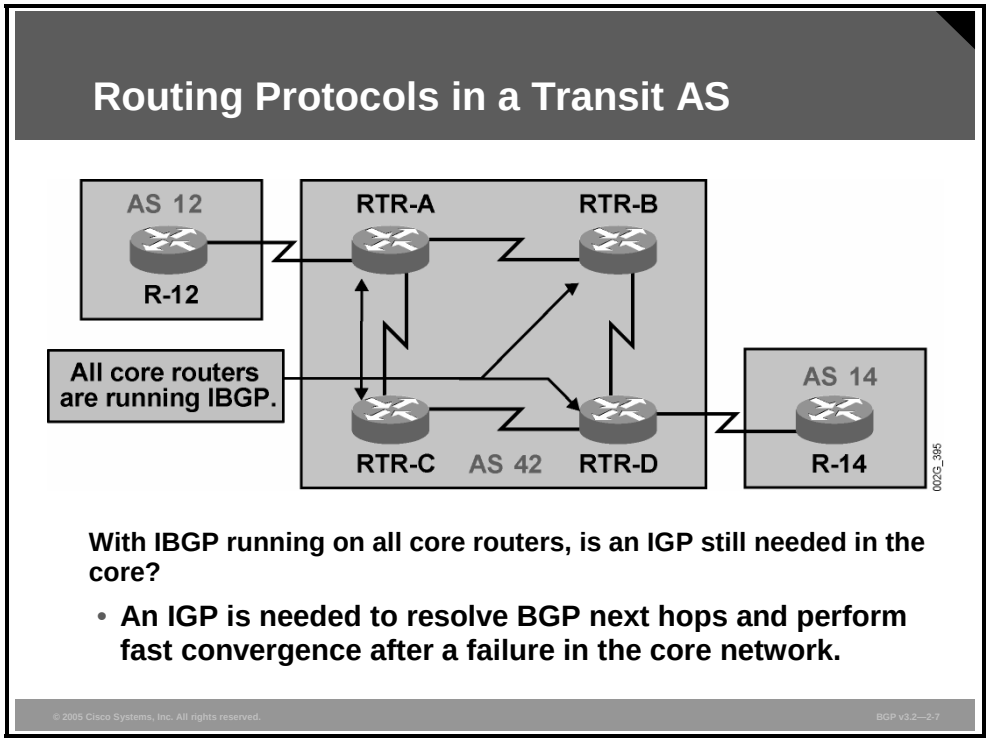
- **Traditional Cisco IOS software switching mechanisms perform recursive lookup when forwarding the first packet.**
  - **Fast switching, optimum switching.**
- **CEF precomputes the routing table.**
  - **All recursive lookups are performed while the routing table is built.**

Traditional Cisco IOS switching mechanisms used the traffic-driven, cache-based switching approach. Both fast switching and optimum switching populate the IP switching cache on demand, meaning that before any IP packets are forwarded, the cache is empty. After the first packet to a specific destination arrives, all routing table lookups are done, including recursive lookup in the case of a BGP route. The result of the lookup is cached for later use when successive packets for the same destination arrive. The process is repeated for every specific destination.

Cisco Express Forwarding (CEF) prebuilds a complete IP forwarding table, called the Forwarding Information Base (FIB), that is based on the IP routing table. After the router installs a routing entry into its routing table, incoming routing information updates trigger the recursive lookup, and the outgoing interface and the actual physical next hop of the route are determined. MAC address resolution and MAC header generation are still traffic-driven and stored in the cache.

Routing Protocols in a Transit AS

This topic describes the need for an IGP in a transit backbone that is running BGP on all routers.



Some network designers base their network design on the wrong assumption that an internal routing protocol is not needed in a transit AS where all routers run BGP. However, the internal routing protocol is still needed inside an AS for two reasons:

- To provide routing information that is needed to establish the IBGP sessions
- To resolve next-hop references (recursive routing)

For example, when RTR-D in the figure receives an IP packet with the destination in AS 12, it does a recursive lookup to find the outgoing interface to be used for packet forwarding. It performs the recursive lookup based on IGP information. If there is suddenly an internal problem within AS 42, and the next-hop address is reachable a different way, the IGP determines this fact. The IGP route to the next-hop network is changed by the router because of newly received IGP route information, and all cache entries that rely on the old information are invalidated. The next recursive lookup that RTR-D performs will indicate a different outgoing interface than before the problem occurred.

During the IGP convergence process, the BGP routing is not affected. The only routing updates that are exchanged during the transit AS convergence are IGP updates describing how to reach internal destinations (including the far ends of the EBGP sessions).

The packet forwarding to external destinations thus benefits from the high-speed convergence that is offered by the IGP. The faster the IGP determines that it should use an alternate path within the AS to reach the next-hop address, the faster it will re-establish IP connectivity toward external destinations.

The conclusion is that an IGP is still needed inside a transit AS, and the network will work better if it is an IGP with fast convergence.

Routing Protocols in a Transit AS (Cont.)

- **Core routers need to run BGP and an IGP.**
- **BGP carries all external routes.**
- **The IGP propagates BGP next hops and other core subnets only.**
- **All customer routes are also carried in BGP.**
  - **Reduces IGP topology database**
  - **Removes customer-caused route flaps from IGP; IGP becomes more stable**

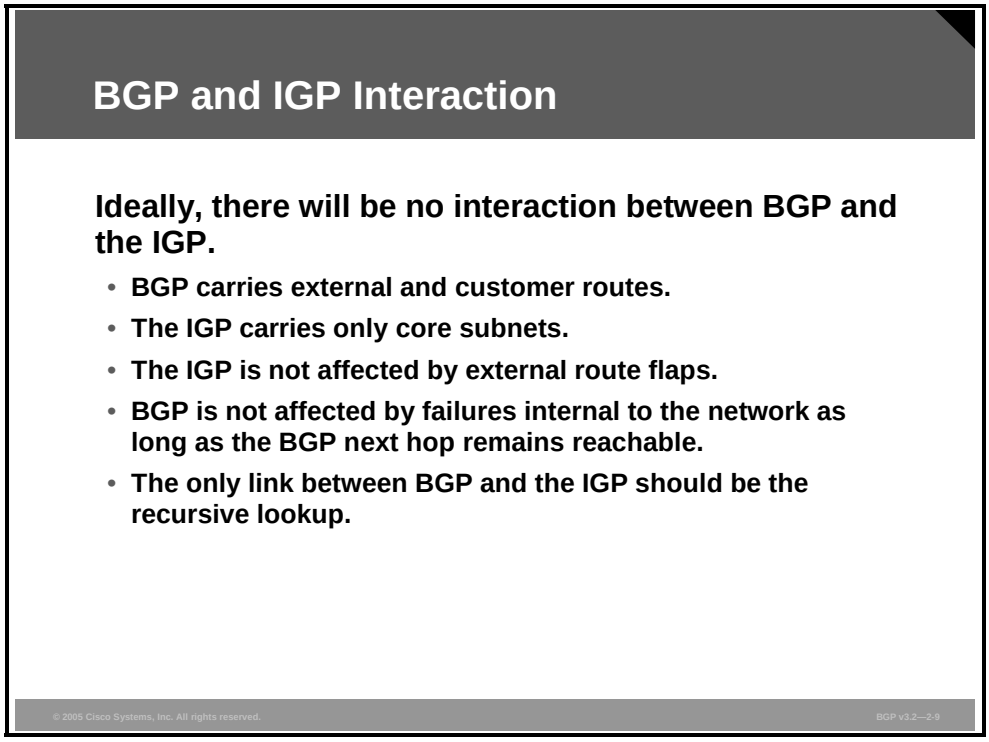
Both BGP and the configured IGP should be configured on all core routers inside the transit AS. The IGP should carry as little information as possible—ideally only the links within the core network, the loopback interfaces, and the external subnets that are used in EBGP sessions with neighboring autonomous systems. This information is enough to establish IBGP sessions and resolve next-hop addresses. The IGP will also work better if it carries less routing information.

No routes external to the transit AS should ever be redistributed by any router from BGP into the IGP. All external routes should be in BGP only.

In autonomous systems that provide customer connectivity (not only transit service), it is also highly recommended that the customer networks be carried in BGP to reduce the amount of information in the IGP and increase IGP stability.

BGP and IGP Interaction

This topic describes the interaction between BGP and IGP in a transit AS.



Ideally, BGP and the IGP carry two different sets of routing information. BGP carries the routes that are received from other autonomous systems and the routes that belong to the local AS and should be announced to other autonomous systems. The IGP carries only enough information to establish IBGP sessions and resolve the EBGP next-hop addresses via the IGP routing tables.

The IGP will provide reachability toward the BGP next-hop addresses only if it is not disturbed by external updates from other autonomous systems.

BGP should take care of the external information. As long as the IGP finds a usable way to the BGP next hops, the BGP does not need to do any recalculation because of internal problems within the AS.

BGP and IGP Interaction (Cont.)

Sometimes, BGP and the IGP will propagate the same route.

- **Usually stems from bad network design.**
- **In this case, routes are determined in EBGP/IGP/IBGP order based on administrative distances of the routes.**

Routing Protocol	Default Administrative Distance
EBGP	20
IGP	90 – 170
IBGP	200

Sometimes the interaction between BGP and the IGP is not ideal, for a number of reasons, including bad network design. In the worst case, the same networks might be carried in both the IGP and BGP. For example, the subnets connecting the AS with neighboring autonomous systems have to be announced via the IGP to enable next-hop resolution but may also be announced via BGP by the remote AS or the local AS. In any case, information about the same IP prefix will appear in both the IGP and the BGP data structures.

When the router installs routing information into the routing table, it checks to see whether there are several sources of information for a particular IP prefix. If so, the router installs the information that it determines is most reliable. The administrative distance (AD) determines which source to use.

BGP considers both EBGP and IBGP routes in the BGP selection process. BGP will therefore never try to install both an EBGP route and an IBGP route for the same destination. Comparison between ADs will thus occur only when two different protocols carry the same destination network.

If BGP selects an EBGP route as the best route for a given destination network, it will try to install that route with a very low AD, meaning that routes that are learned via EBGP have a high likelihood of being installed in the routing table.

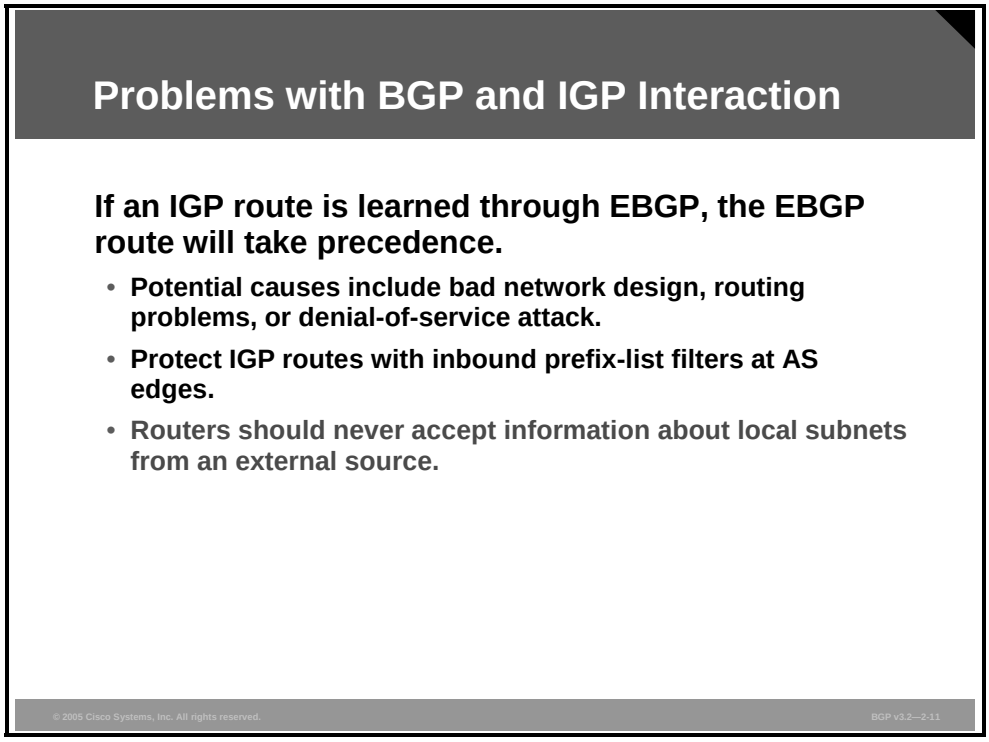
If BGP selects an IBGP route as the best, it will try to install it with a high AD, meaning that routes that are learned via IBGP have a low likelihood of being installed in the routing table.

All IGPs, such as Enhanced Interior Gateway Routing Protocol (EIGRP), OSPF, Intermediate System-to-Intermediate System (IS-IS), and so on, have a medium likelihood of being installed. The ADs for IGPs fall between the ADs of EBGP and IBGP.

Note The reason for giving EBGP a low default AD is because EBGP indicates routes external to the local AS. IP packets with destination addresses to those networks should leave the AS sooner rather than later. It is, in most cases, better that they leave the AS right away.

Problems with BGP and IGP Interaction

This topic describes the potential problems that might arise from BGP and IGP interaction.

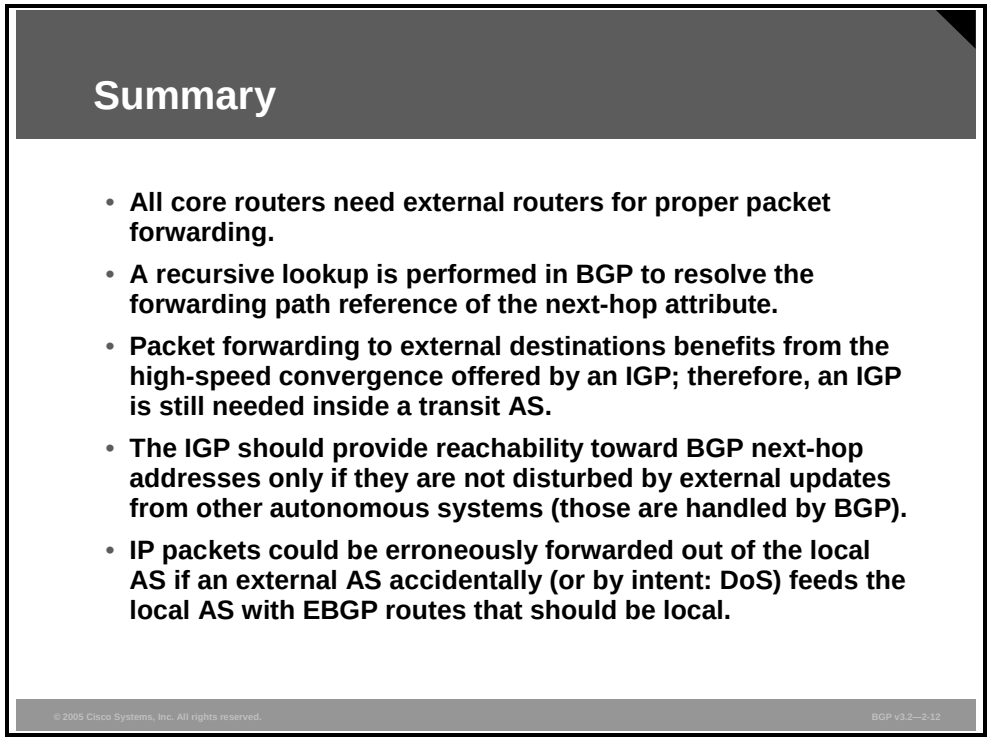


If routing information about the same IP prefix is learned via both EBGP and an IGP, the router will use the EBGP information. If an external AS is feeding the local AS with EBGP routes that actually should be local, routers within the AS will erroneously forward IP packets that are destined to those local networks out of the local AS.

There are several potential reasons for this behavior; the most common is that the remote AS is improperly configured or there is a denial-of-service (DoS) attack. To protect a local AS from this undesired behavior, network administrators should install inbound filters on all EBGP sessions to filter incoming routes and reject routing information about networks that are actually local to the AS.

Summary

This topic summarizes the key points discussed in this lesson.



Lesson 4

Configuring a Transit AS

Overview

Specifying an autonomous system (AS) as a transit backbone introduces specific requirements in the design, scaling, and configuration of Border Gateway Protocol (BGP). This lesson introduces the configuration requirements of Internal Border Gateway Protocol (IBGP) to implement a transit AS. Configuration details of IBGP are discussed in this lesson, including IBGP neighbor configuration, using loopback interfaces for IBGP neighbors, disabling BGP synchronization, and modifying the default administrative distances (ADs) of BGP. This lesson concludes with a discussion of the scalability concerns of BGP in the transit backbone.

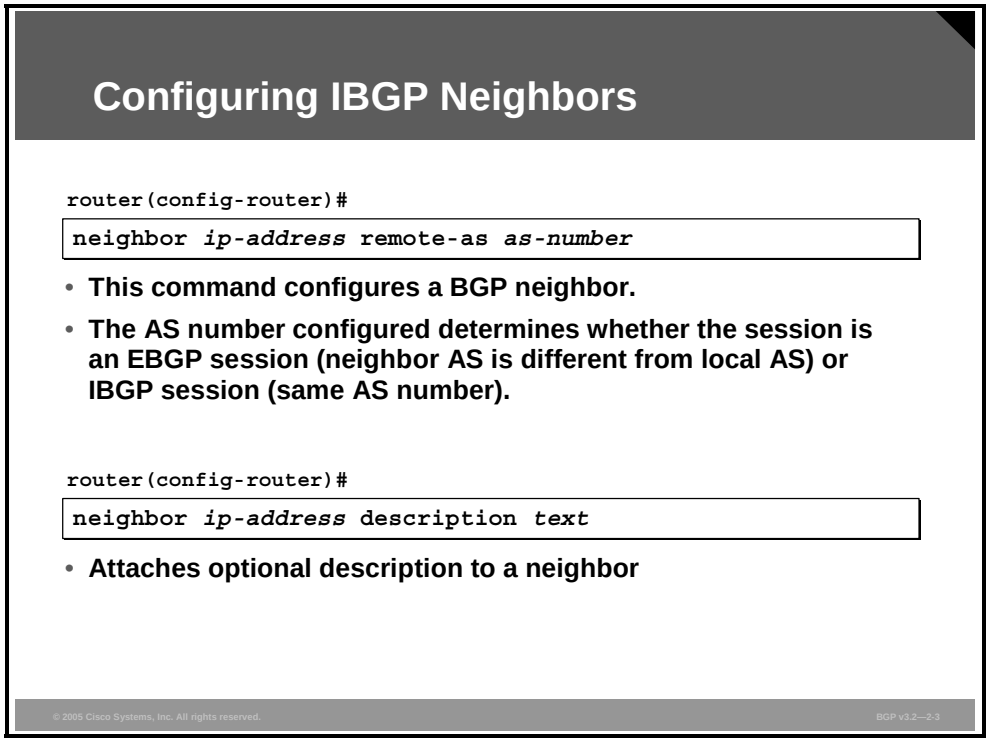
Objectives

Upon completing this lesson, you will be able to configure an AS to act as a transit backbone in a BGP network. This ability includes being able to meet these objectives:

- Identify the Cisco IOS commands that are required to configure IBGP neighbors in an AS
- Identify the Cisco IOS command that is required to configure IBGP sessions between loopback interfaces in a common AS
- Identify the Cisco IOS command that is required to configure BGP synchronization to ensure successful IBGP operation of the transit AS
- Identify the Cisco IOS command that is required to change the AD of BGP routes
- List the scalability limitations of IBGP-based backbones

Configuring IBGP Neighbors

This topic describes the Cisco IOS commands that are required to configure IBGP neighbors in an AS.



Configuring an IBGP neighbor involves two simple steps:

- Configure a BGP neighbor, specifying the same AS number.
- Attach a description to the neighbor to help in documentation and troubleshooting efforts (optional).

neighbor remote-as

To add an entry to the BGP neighbor table, use the **neighbor remote-as** router configuration command.

- **neighbor** [*ip-address* | *peer-group-name*] **remote-as** *as-number*

To remove an entry from the table, use the **no** form of this command.

- **no neighbor** [*ip-address* | *peer-group-name*] **remote-as** *as-number*

Syntax Description

Parameter	Description
ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
as-number	AS to which the neighbor belongs

neighbor description

To associate a description with a neighbor, use the **neighbor description** router configuration command.

- **neighbor** [*ip-address* | *peer-group-name*]**description** *text*

To remove the description, use the **no** form of this command.

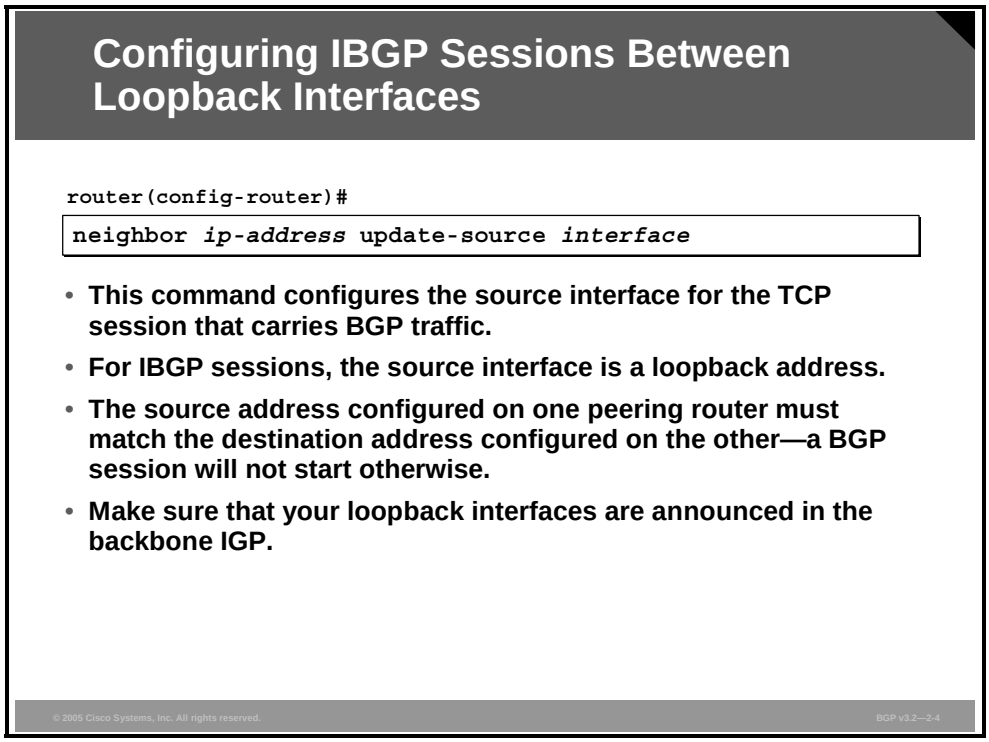
- **no neighbor** [*ip-address* | *peer-group-name*] **description** *text*

Syntax Description

Parameter	Description
ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
text	Text (up to 80 characters) that describes the neighbor

Configuring IBGP Sessions Between Loopback Interfaces

This topic describes the Cisco IOS command that is required to configure IBGP sessions between loopback interfaces on routers in a common AS.



When a BGP session is established between two routers, both routers attempt to set up the TCP connection by sending TCP SYN packets to each other. If both succeed, one of the sessions is brought down so that only one remains. The TCP packets have a destination IP address that is configured with the **neighbor** command. But they must also have a source IP address assigned. If no update source is configured, the router sets the source IP address of the outgoing TCP session to the IP address of the outgoing physical interface.

When a TCP SYN packet with the BGP well-known port number arrives at the peer router, the receiver checks to determine if the connection attempt is coming from one of the configured peers. If the source IP address is not in the list of configured neighbors, the receiver denies the connection attempt.

As a general rule, IBGP sessions should be established between loopback interfaces of BGP-speaking routers. The destination IP address that is configured in the neighbor statement should therefore be the IP address of the loopback interface of the peer router. But the local router must also make sure that the source address of the outgoing TCP connection attempt is the IP address that the peer router has listed. Configuring BGP neighbors using **neighbor update-source** ensures that the source address of the outgoing TCP connection is correct by referring to the interface that has the correct IP address. Normally, this interface is the loopback interface of the local router.

neighbor update-source

To instruct Cisco IOS software to allow IBGP sessions to use any operational interface for TCP connections, use the **neighbor update-source** router configuration command.

- **neighbor** [*ip-address* | *peer-group-name*] **update-source** *interface*

To restore the interface assignment to the closest interface, which is called the “best local address,” use the **no** form of this command.

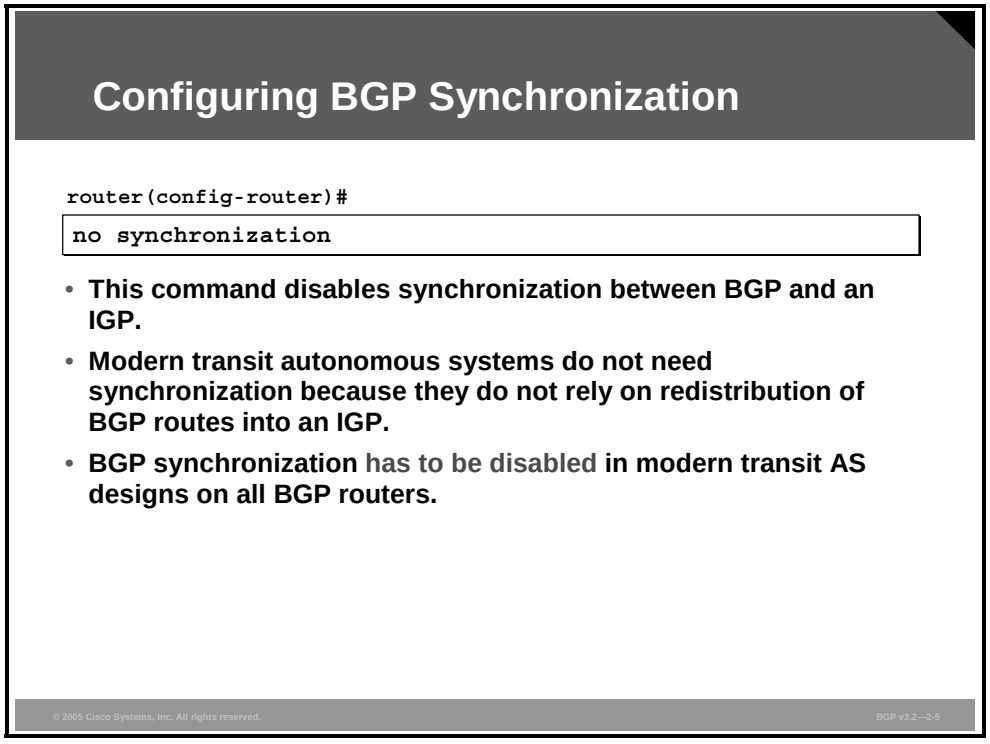
- **no neighbor** [*ip-address* | *peer-group-name*] **update-source** *interface*

Syntax Description

Parameter	Description
ip-address	Neighbor IP address
peer-group-name	Name of a BGP peer group
interface	Loopback interface

Configuring BGP Synchronization

This topic describes the Cisco IOS command that is required to configure BGP synchronization to ensure successful IBGP operation of a transit AS.



The BGP synchronization rule states that if an AS provides transit service to another AS, BGP should not advertise a route until all of the routers within the AS have learned about the route via an Interior Gateway Protocol (IGP). Network designers used synchronization in older transit AS designs that relied on BGP route redistribution into the IGP. Modern AS designs do not rely on this feature anymore because the number of routes carried in the Internet exceeds the scalability range of any known IGP. Redistribution into the IGP is thus no longer applicable, and you must disable the synchronization feature for your transit AS to work.

synchronization

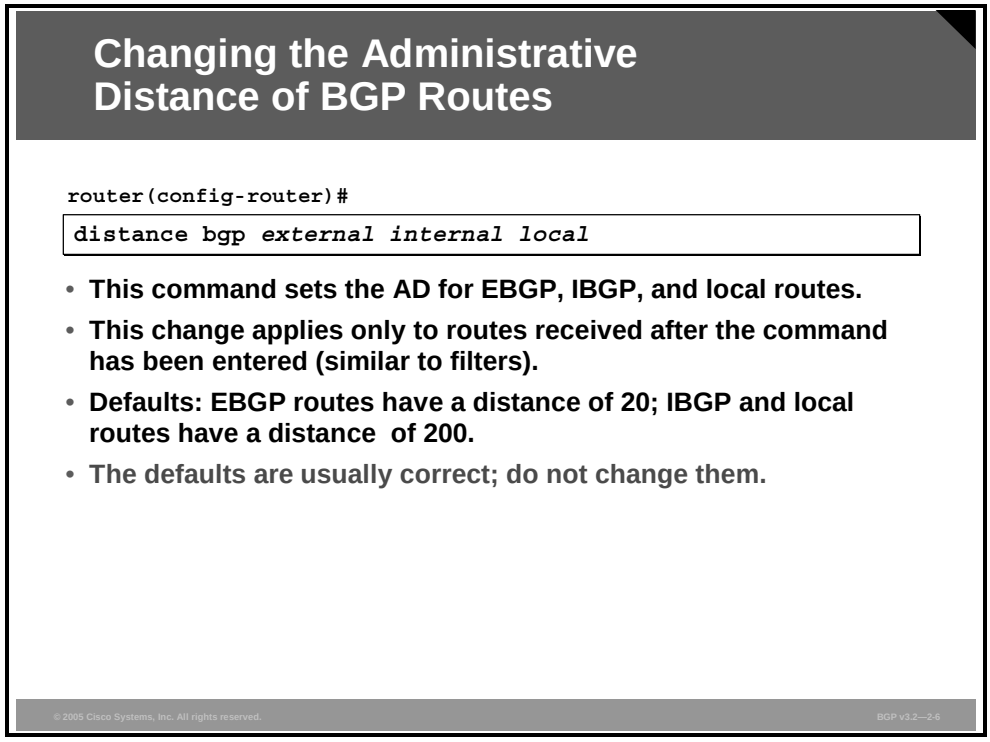
To enable the Cisco IOS software to advertise a network route without waiting for the IGP (that is, to disable synchronization between BGP and your IGP), use the **no** form of the **synchronization** command. Note that in Cisco IOS Software Release 12.2(8)T and later, the default changed to disable synchronization.

- **no synchronization**

This command has no arguments or keywords.

Changing the Administrative Distance of BGP Routes

This topic describes the Cisco IOS command that is required to change the AD of BGP routes.



distance bgp

To allow the use of external, internal, and local ADs that could be a better route to a node, use the **distance bgp** router configuration command.

- **distance bgp** *external-distance internal-distance local-distance*

To return to the default values, use the **no** form of this command.

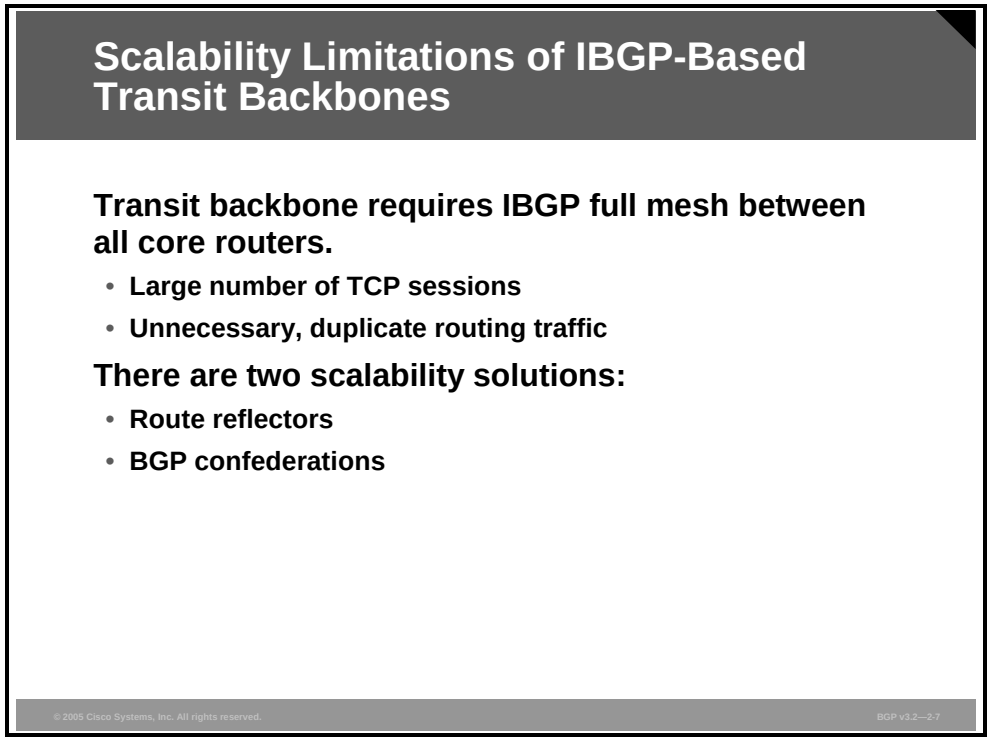
- **no distance bgp** *external-distance internal-distance local-distance*

Syntax Description

Parameter	Description
external-distance	AD for BGP external routes. External routes are routes for which the best path is learned from a neighbor external to the AS. Acceptable values are from 1 to 255. The default is 20. Routes with a distance of 255 are not installed in the routing table.
internal-distance	AD for BGP internal routes. Internal routes are routes that are learned from another BGP entity within the same AS. Acceptable values are from 1 to 255. The default is 200. Routes with a distance of 255 are not installed in the routing table.
local-distance	AD for BGP local routes. Local routes are the networks that are listed with a **network** router configuration command, often as back doors (BGP back door makes the IGP route the preferred route) for that router or for networks that are being redistributed from another protocol. Acceptable values are from 1 to 255. The default is 200. Routes with a distance of 255 are not installed in the routing table.

Scalability Limitations of IBGP-Based Transit Backbones

This topic describes the scalability limitations of IBGP-based backbones.



IBGP split-horizon rules, as documented in previous lessons, mandate an IBGP connection between every border router and every other BGP router in an AS.

The general design rule in IBGP design is to have a full mesh of IBGP sessions. But, a full mesh of IBGP sessions among *n* number of routers would require ($n * (n - 1)$) / 2 IBGP sessions. For example, a full mesh between 10 routers would require (10 * 9) / 2 = 45 IBGP sessions.

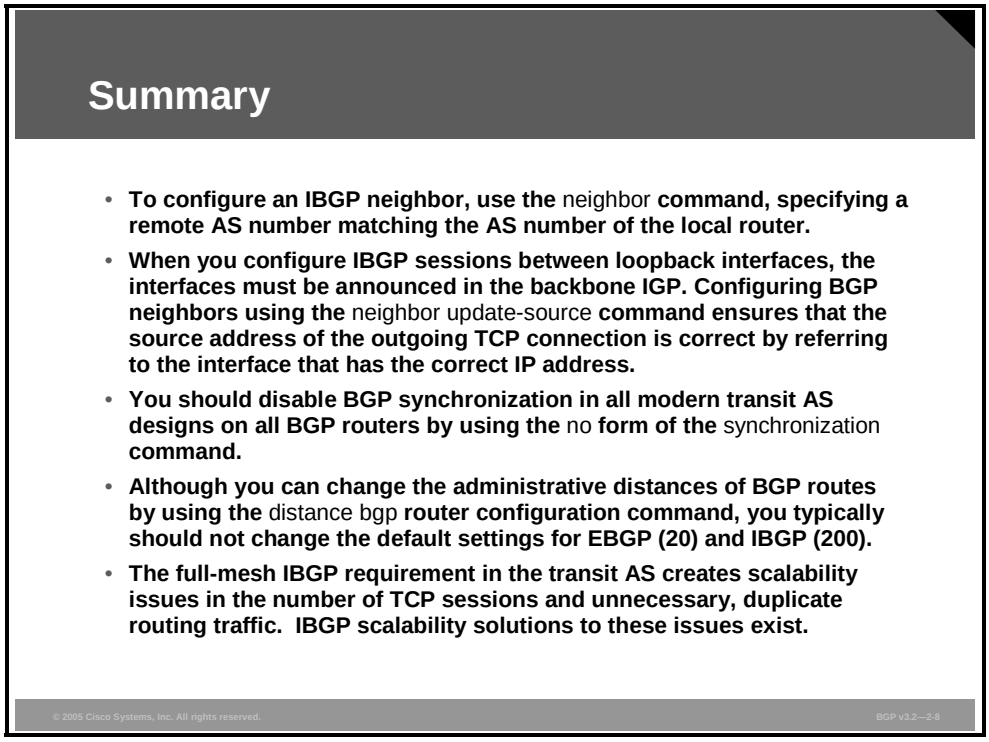
Because every IBGP session on a router uses a separate TCP session, an update that must be sent by the router to all IBGP peers must be sent on each of the TCP sessions. If a router is attached to the rest of the network over just a single link, this single link has to carry all TCP/IP packets for all IBGP sessions. This situation results in duplication of the update over the single link.

Two solutions are available:

- The route reflector solution modifies the IBGP split-horizon rules and allows a particular router to forward (under certain conditions) incoming IBGP updates to a select group of IBGP neighbors. The router performing this function is the “route reflector.”
- The BGP confederations solution introduces the concept of a number of smaller autonomous systems within the original AS. These smaller autonomous systems exchange BGP updates among themselves using intraconfederation EBGP sessions.

Summary

This topic summarizes the key points discussed in this lesson.



Lesson 5

Monitoring and Troubleshooting IBGP in a Transit AS

Overview

Introduction of a transit backbone into a Border Gateway Protocol (BGP) network can create unique troubleshooting challenges. This lesson introduces Internal Border Gateway Protocol (IBGP) monitoring commands and troubleshooting techniques for solving the most common IBGP problems that you might encounter in a transit backbone. Common problems with IBGP, as discussed in this lesson, occur when IBGP sessions do not reach the Established state, when routing information that is received via IBGP is never selected, and when the best BGP route is never installed in the routing table.

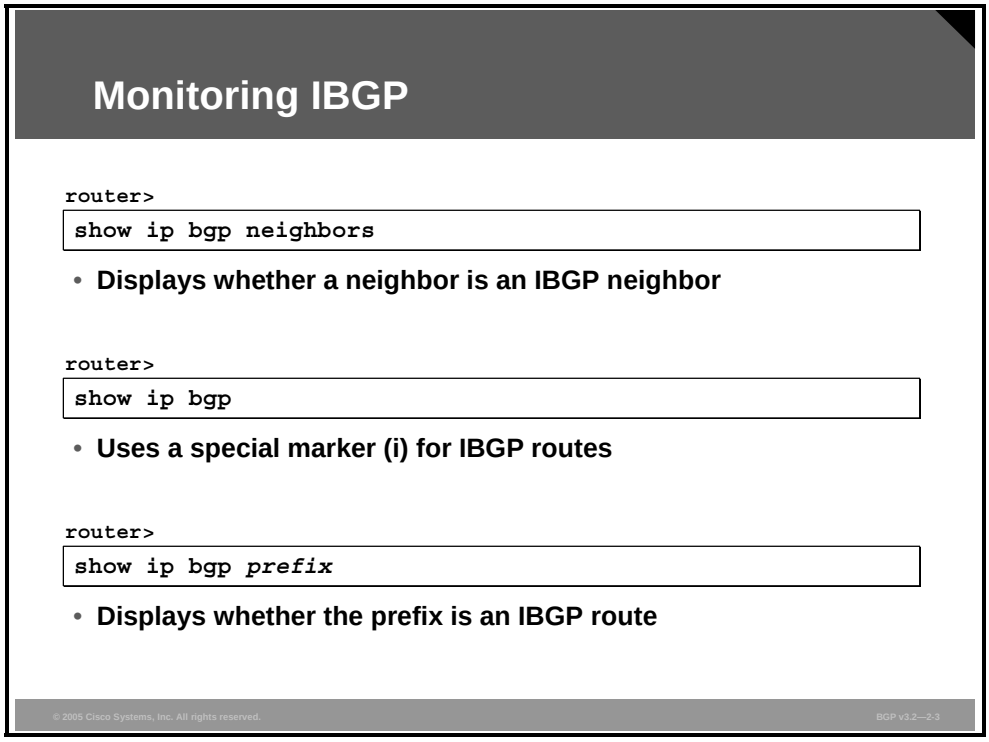
Objectives

Upon completing this lesson, you will be able to verify proper operation of a configured BGP transit network by performing the steps necessary to correct basic IBGP configuration errors. This ability includes being able to meet these objectives:

- Identify the Cisco IOS commands that are required to monitor IBGP operation
- Describe common IBGP configuration problems
- Explain how to troubleshoot IBGP session startup issues
- Explain how to troubleshoot IBGP route selection issues
- Explain how to troubleshoot IBGP synchronization issues

Monitoring IBGP

This topic describes the Cisco IOS commands that are required to monitor IBGP operation.



show ip bgp neighbors

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors** EXEC command.

- **show ip bgp neighbors** [*ip-address*] [**received-routes** | **routes** | **advertised-routes** | {*paths regular-expression*} | **dampened-routes**]

Syntax Description

Parameter	Description
ip-address	(Optional) Address of the neighbor to display neighbor information about. If you omit this argument, all neighbors are displayed.
received-routes	(Optional) Displays all received routes (both accepted and rejected) from the specified neighbor.
routes	(Optional) Displays all routes that are received and accepted. The display output when you are using this keyword is a subset of the output from the **received-routes** keyword.
advertised-routes	(Optional) Displays all the routes that the router has advertised to the neighbor.
paths regular-expression	(Optional) Regular expression that the router uses to match the paths that are received.
dampened-routes	(Optional) Displays the dampened routes to the neighbor at the IP address that is specified.

show ip bgp

To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

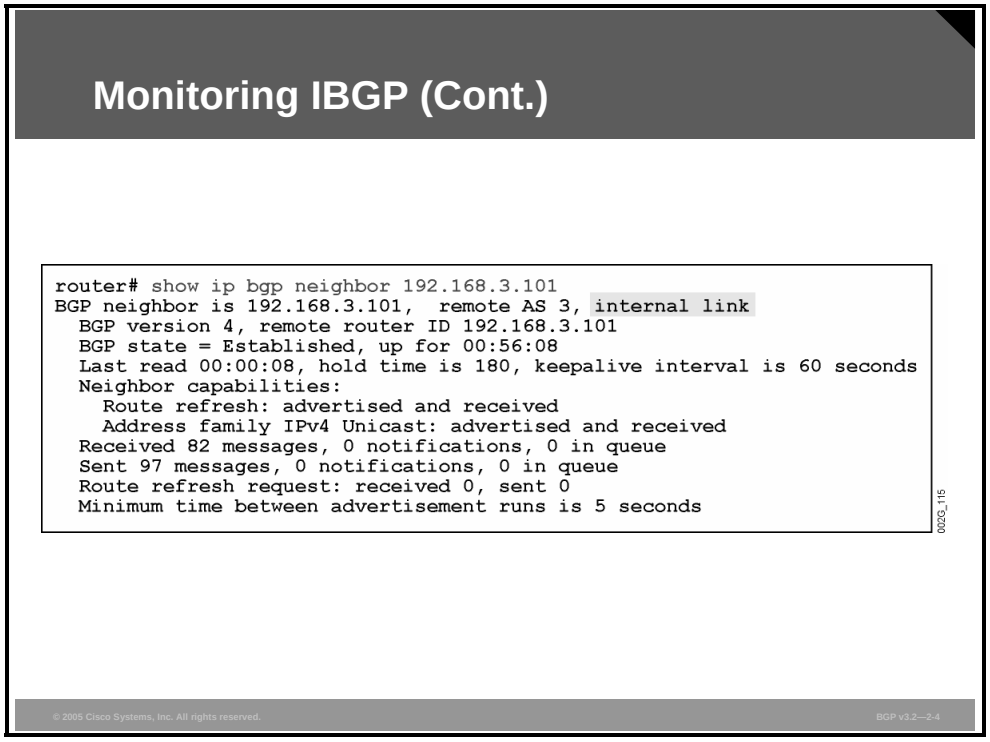
- **show ip bgp** [*network*] [*network-mask*] [**longer-prefixes**]

Syntax Description

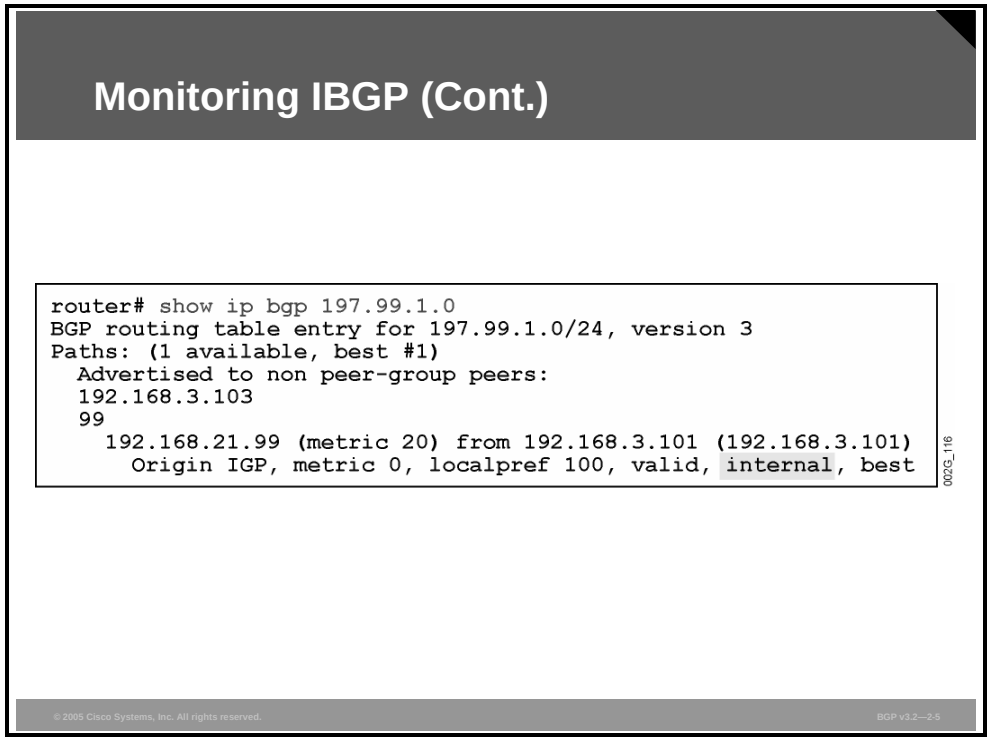
Parameter	Description
network	(Optional) Network number that is entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes that match the address-mask pair
longer-prefixes	(Optional) Displays the route and more specific routes

Example: Monitoring IBGP

This example shows the commands to monitor IBGP.



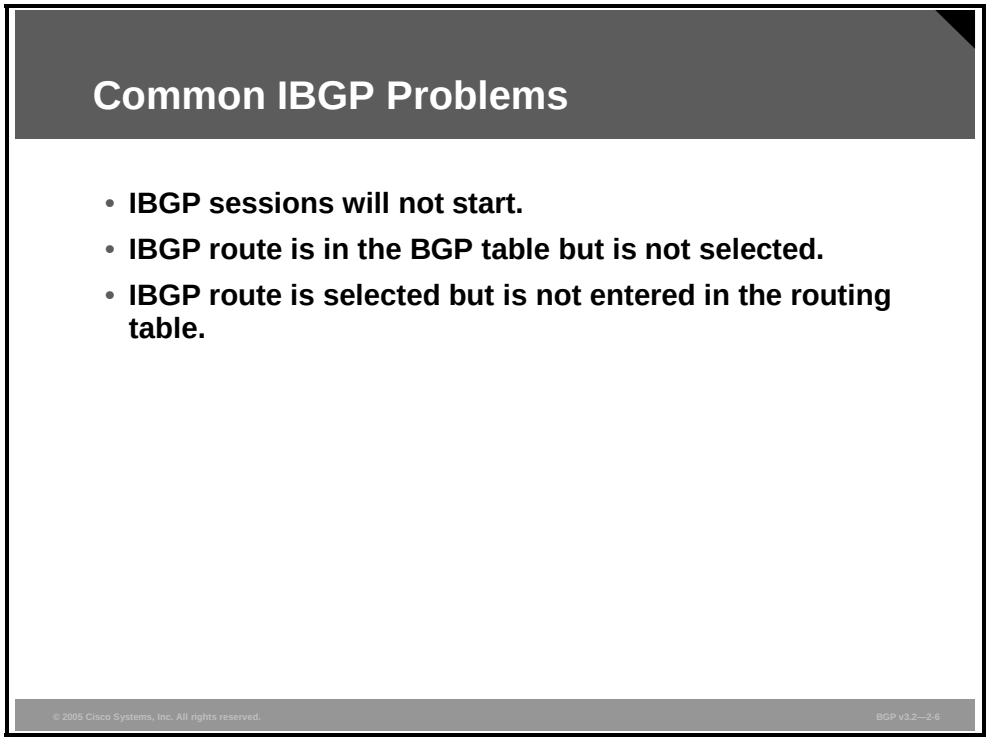
The **show ip bgp neighbors** command displays whether a router is running an IBGP (internal) or External Border Gateway Protocol (EBGP) (external) session with a BGP neighbor. The indication is given by the “internal link” phrase (highlighted in the second line of the figure).



The **show ip bgp** *prefix* command displays whether a BGP route was received from an IBGP (internal) or EBGP (external) neighbor. The indication is given by the word “internal” that is displayed in the last line of the printout (highlighted in the last line of the figure).

Common IBGP Problems

This topic describes configuration problems that are common to IBGP implementations.

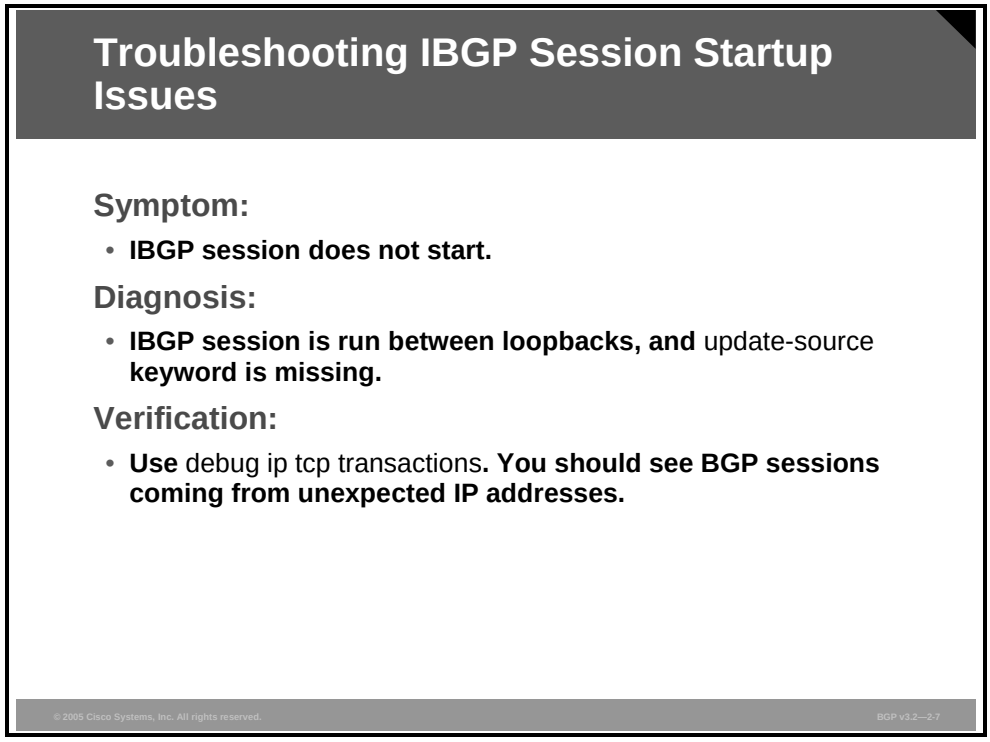


Troubleshooting the BGP configuration of a transit AS can be cumbersome, because there are a number of common pitfalls that you might encounter. Three of the most common problems are the following:

- IBGP sessions do not reach the Established state.
- Routing information that is received via IBGP is never selected.
- The best BGP route is never installed in the routing table.

Troubleshooting IBGP Session Startup Issues

This topic describes how to troubleshoot IBGP session startup issues.



A common mistake when you are configuring IBGP sessions is to forget the **neighbor update-source loopback 0** configuration command.

When you are configuring IBGP neighbors on the router, it is easy to remember to make a correct reference to the loopback interface of the remote router. But it is equally important to make sure that the correct source IP address of the outgoing TCP session is set. The peer router will not accept the session if the incoming source address does not match the peer router list of IBGP neighbors.

To verify that this issue is causing the problem, use the **debug ip tcp transactions** command. The output of the **debug ip tcp transactions** command should display TCP SYN packets coming from unexpected IP addresses on the receiving router and TCP sessions being reset with TCP RST packets on the sending (misconfigured) router.

Troubleshooting IBGP Session Startup Issues (Cont.)

Symptom:

- **IBGP session does not start.**

Diagnosis:

- **Loopback interfaces are not reachable.**

Verification:

- **Do extended** ping **between loopback addresses to verify reachability.**

An IBGP session between two routers can be established from the loopback interface of one router to the loopback interface of the other router only if the two routers can exchange IP packets using those addresses as source and destination. This exchange is possible only if the Interior Gateway Protocol (IGP) carries the subnets that are assigned to each of the loopback interfaces.

When you are verifying reachability with the **ping** command, make sure that the ping packets are sourced from the loopback interface. Use an extended **ping** and explicitly refer to the IP address of the loopback interface to ensure that packets are sourced from the loopback interface.

Troubleshooting IBGP Session Startup Issues (Cont.)

Symptom:

- **IBGP session does not start.**

Diagnosis:

- **Packet filters prevent establishment of BGP sessions.**

Verification:

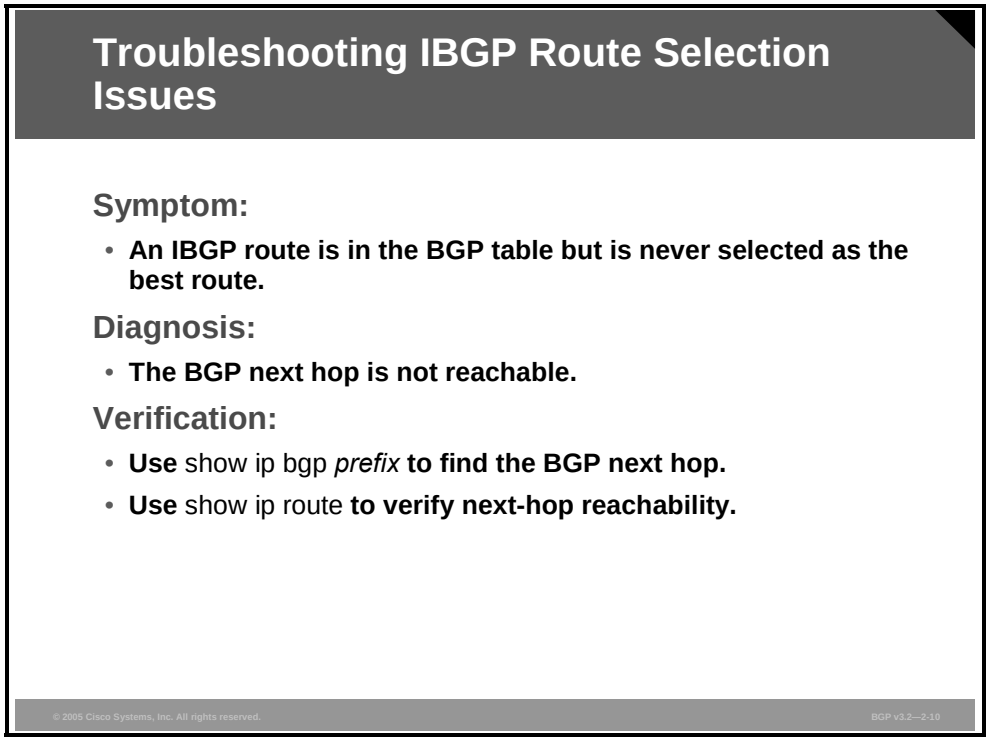
- **Use** debug ip tcp transactions **and** debug ip icmp **to see whether the initial TCP SYN packets are rejected.**

Packet filters can stop the BGP sessions. The path between the two BGP peer routers must be free from filters blocking the BGP traffic.

BGP runs on the well-known TCP port 179. Both routers will make connection attempts to that destination port. They will use a high-numbered TCP port as source. It is enough that one of the connection attempts succeeds. But for better performance during recovery from network failure, both attempts should have the possibility of succeeding. If both attempts do succeed, one of the connections will be brought down.

Troubleshooting IBGP Route Selection Issues

This topic describes how to troubleshoot IBGP route selection issues.



A BGP update can be used by the router to reach network destinations only if the next-hop address specified in the BGP update is reachable. A BGP update that refers to a next hop that is currently not reachable according to the routing table will be saved in the BGP table, but it cannot be installed by the router into its routing table. If the next-hop address later becomes reachable, the BGP route will become a candidate route that could be used by that router for packet forwarding to that destination.

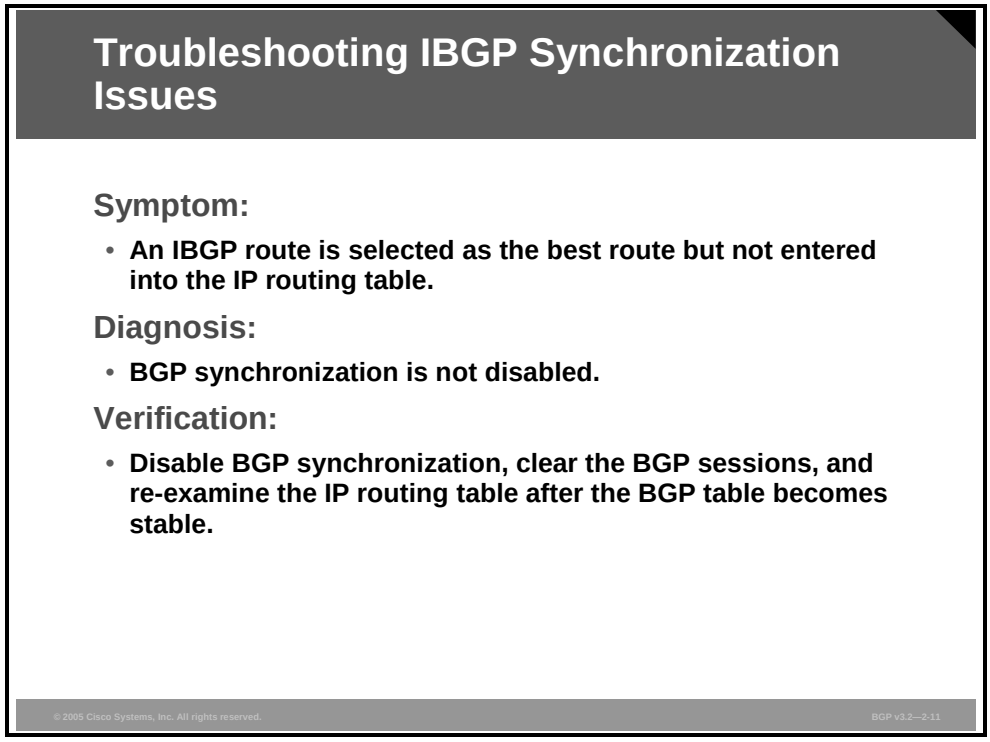
To verify the next-hop reachability, check the BGP route in the BGP table by using the **show ip bgp** *prefix* command. The next hop is referred to as “inaccessible” if it is not currently reachable according to the routing table.

A common mistake is to forget to let the IGP announce the reachability of subnets that physically connect the local autonomous system (AS) with a neighboring AS. These subnets are used by the router to establish the EBGP session, and the next hop that is received in an incoming BGP update will be the far end of the EBGP session. If all routers in the local AS do not have a path to that subnet, the next-hop address will be inaccessible.

You can prevent this problem by including the subnet that links the transit AS to neighboring autonomous systems in the IGP by using either the **redistribute connected** command or the **network** and **passive-interface** configuration commands.

Troubleshooting IBGP Synchronization Issues

This topic describes how to troubleshoot IBGP synchronization issues.



In old BGP designs, redistribution between BGP and an IGP was common practice, and these protocols had to be synchronized to ensure proper packet forwarding. In modern designs, redistribution is no longer used and synchronization has to be turned off. However, the default value is to have synchronization enabled.

Routers with BGP synchronization enabled will not be able to install IBGP routes in the routing table or propagate them to other EBGP neighbors.

You can fix this problem by configuring **no synchronization** in the router BGP configuration.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

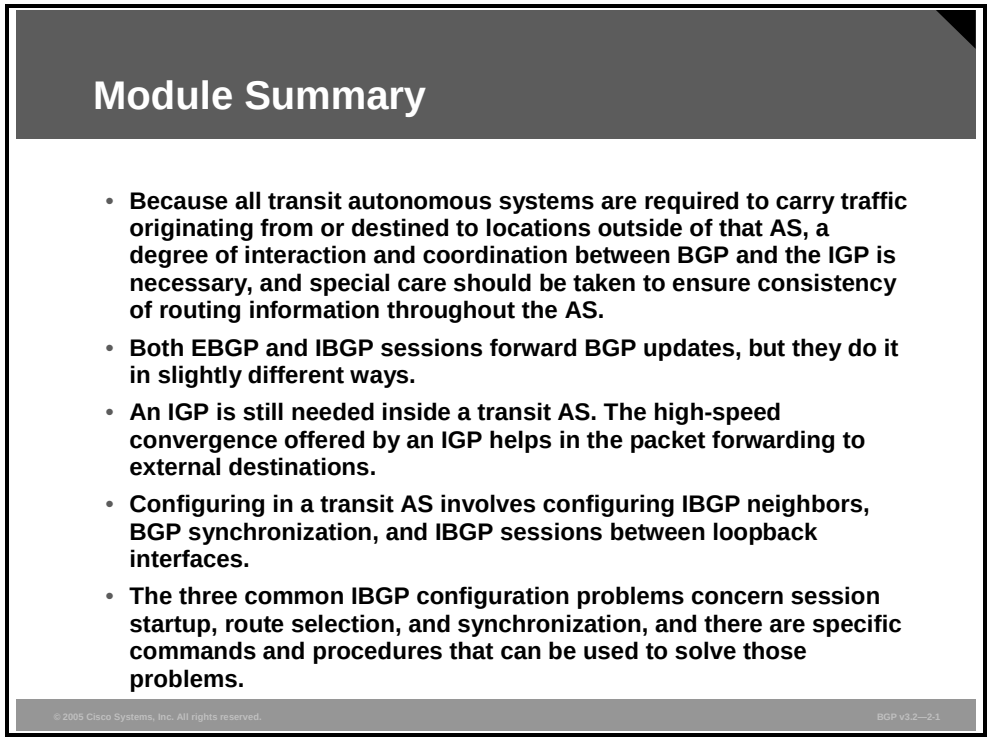
- **You can use the** show ip bgp neighbors **and** show ip bgp *prefix* **commands to monitor IBGP operation.**
- **Common IBGP configuration problems include IBGP sessions that do not reach the Established state, routing information that is received via IBGP that is never selected, and the best BGP route never being installed in the routing table.**
- **There are a number of problems that can occur during IGBP session startup. You can use** debug ip tcp transactions **to see BGP sessions coming from unexpected IP addresses, use an extended** ping **and explicitly refer to the IP address of the loopback interface to ensure that packets are sourced from the loopback interface, or use** debug ip tcp transactions **and** debug ip icmp **to see whether the initial TCP SYN packets are rejected.**

Summary (Cont.)

- **It is important to include the subnet linking the transit AS to an external AS in the IGP to prevent the BGP next hop from being unreachable. To verify next-hop reachability, check the BGP route in the BGP table by using the** show ip bgp *prefix* **command.**
- **Routers with BGP synchronization enabled will not be able to install IBGP routes in the routing table or propagate them to other EBGP neighbors. Configure** no synchronization **in the router BGP configuration to solve this problem.**

Module Summary

This topic summarizes the key points discussed in this module.



This module discussed BGP issues in a transit AS. The first lesson introduced the requirements of IBGP and described processing the next-hop attribute with the routers that reside in the transit AS. The next lesson explained the interaction between EBGP and IBGP. The following lesson described the function of an IGP in forwarding packets through an AS. Configuring the AS as a transit backbone was discussed in the next lesson. The final lesson presented IBGP monitoring commands and troubleshooting techniques to solve the most common IGBP problems in a transit backbone.

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *BGP Case Studies*. http://www.cisco.com/warp/public/459/bgp-toc.html.
- Cisco Systems, Inc. *Using the Border Gateway Protocol for Interdomain Routing*. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm.
- Cisco Systems, Inc. *Troubleshooting TCP/IP*. "Troubleshooting IP Connectivity and Routing Problems." http://www.cisco.com/univercd/cc/td/doc/cisintwk/itg_v1/tr1907.htm#xtocid27.
- Cisco Systems, Inc. *Configuring BGP*. http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm.
- Cisco Systems, Inc. *Cisco IOS IP Routing Configuration Guide, Release 12.2*

  http://www.cisco.com/en/US/partner/products/sw/iosswrel/ps1835/products_configuration_guide_book09186a0080087fa9.html.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Why is IBGP a mandatory component of a transit AS? (Source: Working with a Transit AS)

A) It is the only feasible way to ensure that all routers in the AS have consistent external routing information.
B) It eliminates the scalability issues of running an IGP within the transit AS.
C) Running IBGP on all routers is the only way to satisfy the filtering requirements of the transit AS.
D) An IGP is not capable of handling the potential routing loops in the transit AS.

Q2) How is EBGP used in a transit AS? (Source: Working with a Transit AS)

A) as a means of transporting customer routes across the transit backbone
B) to exchange routes between different autonomous systems and the transit AS
C) to enhance scalability by transporting IGP routes for the transit AS
D) as a means of injecting local routes into the transit backbone

Q3) Why is redistributing BGP routes into an IGP for use in a transit backbone NOT recommended? (Source: Working with a Transit AS)

A) Redistribution removes all BGP attributes that are needed to ensure optimal routing within the transit AS.
B) An IGP cannot enforce complex administrative policies and route selection rules.
C) IGPs cannot scale to the demands that are presented by the number of routes on the Internet.
D) IGPs are not stable when faced with a flapping network.

Q4) What are the two key functions of a transit AS? (Choose two.) (Source: Working with a Transit AS)

A) to filter out routes that do not belong to customers of the service provider
B) to provide Internet connectivity to customers of the service provider
C) to propagate routes between remote autonomous systems
D) to route packets between remote networks

Q5) How are BGP routes sent across the transit backbone? (Source: Working with a Transit AS)

A) by redistributing BGP into an IGP and then back into BGP
B) through the use of IBGP
C) by establishing EBGP sessions between all routers in the transit backbone
D) by redistributing connected routes at the edge of the transit backbone

Q6) Which two statements are true regarding the AS-path attribute as it relates to IBGP? (Choose two.) (Source: Interacting with IBGP and EBGP in a Transit AS)

A) Each router in the AS appends its AS number to the AS path on outgoing BGP updates.
B) The AS path inside an AS will be empty for routes originating inside a neighboring AS.
C) The AS-path attribute is not used to detect routing loops inside an AS.
D) The AS-path attribute is not modified within the AS.

Q7) What is the primary function of the IBGP multipath load-sharing feature? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) to choose one IBGP path as the best path to a destination
B) to choose multiple IBGP paths as the best paths to a destination
C) to designate one path as the best path and advertise this best path to its neighbors.
D) to enable a router to handle all the traffic destined for a particular site

Q8) Which of the following statements about a BGP split horizon is accurate? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) Because of the BGP split horizon, any router can relay IBGP information within the AS.
B) BGP split horizon is useful only within autonomous systems.
C) Routing information loops within the AS are prevented by an IBGP split horizon.
D) With a BGP split horizon, routing information that is received through an IBGP session is forwarded to another IBGP neighbor.

Q9) Why is it recommended that loopback interfaces be used to form IBGP neighbor sessions? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) reduces router memory resource requirements
B) reduces router CPU resource requirements
C) ensures IBGP session stability
D) is more secure than using the physical interface

Q10) How is the BGP next-hop attribute processed over an IBGP connection? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) The next-hop address is set to the address of the receiving router.
B) The next-hop address is not modified over the IBGP session.
C) The next-hop address is set to the IP address of the nearest EBGP peer.
D) The next-hop attribute is set to the IP address of the nearest EBGP peer; if no external AS connection has been configured, the next hop is set to the default gateway that is configured on the router.

Q11) Which two statements are true of the full-mesh requirement in IBGP? (Choose two.) (Source: Interacting with IBGP and EBGP in a Transit AS)

A) The IBGP mesh must be a logical full mesh.
B) A physical full mesh must be maintained within the IBGP AS.
C) Because of BGP split horizon, no router can relay IBGP information within the AS.
D) All routers within the AS must be directly connected to ensure correct delivery of BGP routing information.

Q12) Which three statements regarding the next-hop-self configuration in BGP are true? (Choose three.) (Source: Interacting with IBGP and EBGP in a Transit AS)

A) Changing the next-hop attribute might cause suboptimal routing.
B) The configuration changes how the next-hop attribute is processed at edge routers.
C) The configuration announces the local IP address as the BGP next hop in outgoing updates that are sent to the specified neighbor.
D) The configuration removes the requirement for the IGP to carry reachability information for intra-AS destinations.

Q13) Why must all EBGP peers be reachable by all BGP-speaking routers within the transit AS? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) BGP-speaking routers in a transit AS use the next-hop-self attribute to find their EBGP neighbors.
B) EBGP peers in a transit AS use the length of the AS path to decide which BGP route to install in the routing table.
C) When BGP routes are used in the routing table, the router checks how it would have reached the next-hop address, and it installs the BGP route with the same forwarding indication as for the route that is used to reach the next-hop IP address.
D) All BGP peers do not need to speak to each other within a transit AS.

Q14) Which command is used to configure the router as the next hop for a BGP-speaking neighbor or peer group? (Source: Interacting with IBGP and EBGP in a Transit AS)

A) neighbor description
B) neighbor remote as
C) maximum-paths ibgp
D) neighbor next-hop-sf

Q15) What are three differences between IBGP and EBGP sessions? (Choose three.) (Source: Interacting with IBGP and EBGP in a Transit AS)

A) Route selection rules slightly prefer IBGP routes.
B) Routes that are learned from IBGP peers are not advertised to other IBGP peers.
C) EBGP peers are directly connected, and IBGP peers are usually distant.
D) By default, no BGP attributes are changed in IBGP updates.

Q16) What are two reasons why you must run IBGP on all routers within a transit backbone? (Choose two.) (Source: Forwarding Packets in a Transit AS)

A) so routers can properly forward packets toward all external destinations
B) to ensure that a full mesh exists among all routers in the AS
C) to allow routers to properly process the BGP next-hop attribute
D) because IGPs cannot scale large enough to handle redistribution of BGP routes

Q17) If a transit backbone has IBGP running on all routers, what are two reasons why it is still necessary to use an IGP? (Choose two.) (Source: Forwarding Packets in a Transit AS)

A) to provide routing information that is needed to establish the IBGP sessions
B) to resolve next-hop references that are used in recursive routing
C) so that BGP routes can be properly transported through the AS
D) to provide user workstations with a network default gateway

Q18) What is the AD of the following protocols? (Fill in the blanks.) (Source: Forwarding Packets in a Transit AS)

A) IBGP _____
B) EBGP _____
C) OSPF _____
D) IS-IS _____
E) RIP _____

Q19) What are two reasons why the AD is an important consideration for BGP network design? (Choose two.) (Source: Forwarding Packets in a Transit AS)

A) The AD affects how routes are selected for use in the IP routing table.
B) The AD controls how routing information is entered into the BGP table.
C) If a route is advertised by both an IGP and through EBGP, the router will prefer the external route.
D) The AD is not a large concern to BGP design, because the router will always choose the route that is advertised by the protocol that is best suited to reach the destination.

Q20) With regard to recursive route lookups, what are two ways in which CEF is different from traditional Cisco IOS switching mechanisms such as route caching? (Choose two.) (Source: Forwarding Packets in a Transit AS)

A) Traditional Cisco IOS switching mechanisms wait for the first packet to arrive before recursive lookup can take place.
B) New entries in the IP routing table will trigger a recursive lookup in traditional Cisco IOS switching mechanisms.
C) CEF prebuilds a complete IP forwarding table based on the IP routing table.
D) CEF will build a FIB directly from the entries in the BGP table before any BGP packets arrive at the router.

Q21) When you are configuring the BGP neighbor session, what differentiates an EBGP neighbor from an IBGP neighbor? (Source: Configuring a Transit AS)

A) The keyword **internal** is at the end of the **neighbor** command.
B) IBGP neighbors will have the same AS number specified.
C) A description for the internal or external neighbor must be attached with the **neighbor description** command.
D) Directly connected neighbors will automatically form an EBGP session.

Q22) Which two steps are required when you use a loopback interface for IBGP peering sessions? (Choose two.) (Source: Configuring a Transit AS)

A) Ensure that the loopback interfaces are reachable through an IGP.
B) Ensure that the two neighbors are directly attached.
C) Verify that each router has multiple physically redundant paths.
D) Configure a neighbor statement with the **update-source** command.

Q23) Why is it important to disable BGP synchronization in a transit backbone? (Source: Configuring a Transit AS)

A) IGPs can support the routing requirements of full Internet routing, and hence synchronization is no longer necessary.
B) Because BGP redistribution into an IGP is no longer practical, enabling the synchronization feature is no longer applicable.
C) Synchronization requires all BGP transit routes to be explicitly mapped to an exit point, creating too much administrative overhead.
D) Synchronization requires BGP attributes to be properly mapped to IGP metrics for BGP routing across the transit backbone to function properly, creating too much overhead.

Q24) What are two negative ramifications of the full-mesh requirement that is imposed by IBGP? (Choose two.) (Source: Configuring a Transit AS)

A) administrative difficulty of applying an AS-wide routing policy
B) requirement to use next-hop-self for proper routing to external destinations
C) large number of TCP sessions
D) unnecessary duplication of routing traffic

Q25) What are two scalability tools that you can use to overcome the full-mesh requirement for IBGP sessions? (Choose two.) (Source: Configuring a Transit AS)

A) confederations
B) floating static routes
C) route reflectors
D) disabling BGP synchronization

Q26) Which Cisco IOS **show** command indicates that a BGP route is an IBGP route? (Source: Monitoring and Troubleshooting IBGP in a Transit AS)

A) **show ip route**
B) **show ip route bgp**
C) **show ip bgp**
D) **show ip bgp internal**

Q27) Which three of the following are the most common BGP implementation problems? (Choose three.) (Source: Monitoring and Troubleshooting IBGP in a Transit AS)

A) IBGP sessions do not reach the Established state.
B) TCP window size is set incorrectly.
C) Routing information that is received via IBGP is never selected.
D) The best BGP route is never installed in the routing table.

Q28) What are three common situations that prevent IBGP sessions from starting? (Choose three.) (Source: Monitoring and Troubleshooting IBGP in a Transit AS)

A) The IBGP session has been configured to peer to a loopback interface, but **update-source** has not been configured on the neighbor.
B) An access-list filter is blocking access to TCP port 179.
C) The IBGP session has been configured to peer to a loopback interface, but the loopback interface has not been administratively enabled with the **no shutdown** command.
D) The IBGP session has been configured to peer to a loopback interface, but the interfaces are not reachable via the IGP.

Q29) Which common issue could prevent IBGP best routes from being inserted into the IP routing table? (Source: Monitoring and Troubleshooting IBGP in a Transit AS)

A) failure to disable BGP synchronization
B) failure to disable BGP split horizon
C) lack of a route to the BGP next hop for the IGP
D) failure to inject a default route into the IGP

Q30) Which two of the following statements about solving IBGP synchronization problems are accurate? (Choose two.) (Source: Monitoring and Troubleshooting IBGP in a Transit AS)

A) Routers with BGP synchronization enabled will be able to install IBGP routes in the routing table and propagate them to other EBGP neighbors.
B) Routers with BGP synchronization enabled will not be able to install IBGP routes in the routing table or propagate them to other EBGP neighbors.
C) The default value is to have synchronization enabled.
D) The default value is to have synchronization disabled.

Module Self-Check Answer Key

Q1) A

Q2) B

Q3) C

Q4) C, D

Q5) B

Q6) C, D

Q7) B

Q8) C

Q9) C

Q10) B

Q11) A, C

Q12) A, B, C

Q13) C

Q14) D

Q15) B, C, D

Q16) A, D

Q17) A, B

Q18) A-200
B -20
C- 110
D -115
E -120

Q19) A, C

Q20) A, C

Q21) B

Q22) A, D

Q23) B

Q24) C, D

Q25) A, C

Q26) C

Q27) A, C, D

Q28) A, B, D

Q29) A

Q30) B, C

Module 3

Route Selection Using Policy Controls

Overview

Border Gateway Protocol (BGP) enables traffic in Internet backbones to determine an optimal path to its destination across networks comprising more than one autonomous system (AS). Routes that are learned via BGP have properties that are associated with them that aid BGP in determining the best route to a particular destination. There are many instances in which the default BGP route selection does not match administrative or business policies. Likewise, redundant network designs often require enterprises to run BGP when they are connected to more than one Internet service provider (ISP). In these situations, full BGP routing tables and default BGP route selection are not desirable.

This module provides information on how to connect Internet customers to multiple service providers. It introduces the need for filtering BGP updates and changing BGP route selection policies. In addition, this module describes different Cisco IOS mechanisms (AS-path filters, prefix-lists, route-maps) available for BGP route filtering.

Module Objectives

Upon completing this module, you will be able to use BGP attributes to influence the route selection process in a network scenario where you must support multiple connections. This ability includes being able to meet these objectives:

- Describe the need for influencing BGP route selection in a customer scenario where connections to multiple ISPs must be supported
- Successfully configure BGP to influence route selection using AS-path filters in a customer scenario where connections to multiple ISPs must be supported
- Successfully configure BGP to influence route selection using prefix-list filters in a customer scenario where connections to multiple ISPs must be supported
- Use outbound route filtering to minimize the impact of BGP routing updates on router resources in an operational BGP network
- Correctly configure BGP to influence route selection using route-maps in a typical BGP network
- Configure the soft reconfiguration feature to minimize the impact of expediting BGP policy updates in a typical BGP network

Lesson 1

Using Multihomed BGP Networks

Overview

In some circumstances, it is important to have multiple paths to an Internet service provider (ISP). There are business and technical reasons to configure a Border Gateway Protocol (BGP) network in a multihomed configuration. Mission-critical applications often call for redundant network designs. When access to applications is provided over the Internet, enterprises typically use multihomed BGP networks to achieve their goals of high availability.

Full BGP routing tables and default BGP route selection might ordinarily be considered as desirable characteristics for a network. However, the overhead of full BGP routing tables is not warranted in these situations. Furthermore, the default route selection in BGP often does not match the business and technical requirements for multihomed enterprise networks that use BGP. This lesson discusses these business and technical issues and the requirement to use filters to influence route selection and to apply a routing policy.

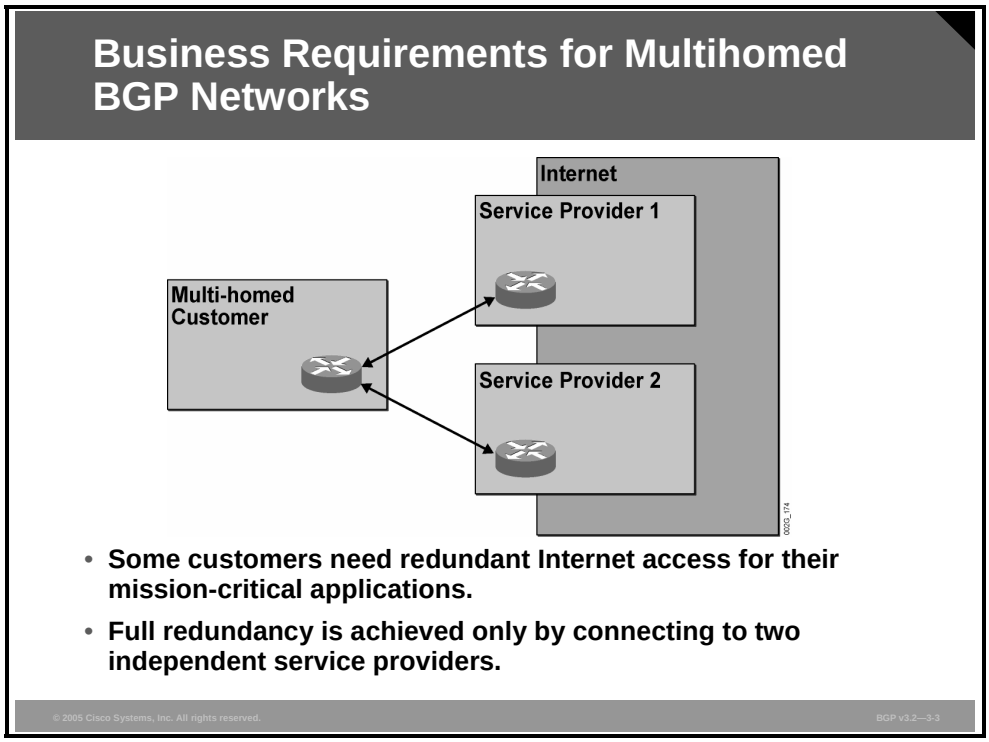
Objectives

Upon completing this lesson, you will be able to describe the need for influencing BGP route selection in a customer scenario where connections to multiple ISPs must be supported. This ability includes being able to meet these objectives:

- List the business requirements for multihomed BGP networks in service provider environments
- Describe the technical requirements for multihomed BGP networks in service provider environments
- Explain the need for BGP policies that influence route selection in a multihomed BGP network
- Describe typical routing policies for multihomed BGP customers
- Explain the need to influence BGP route selection in a service provider environment
- Explain the need for BGP filters in a service provider environment

Business Requirements for Multihomed BGP Networks

This topic describes the business requirements for multihomed BGP networks in service provider environments.



Companies with web servers (or similar servers) offering mission-critical business services over the Internet often like to have their networks redundantly connected to the Internet. When the companies calculate the expected loss of business because of an unexpected disconnection from the Internet, they may conclude that having two connections to the Internet is profitable.

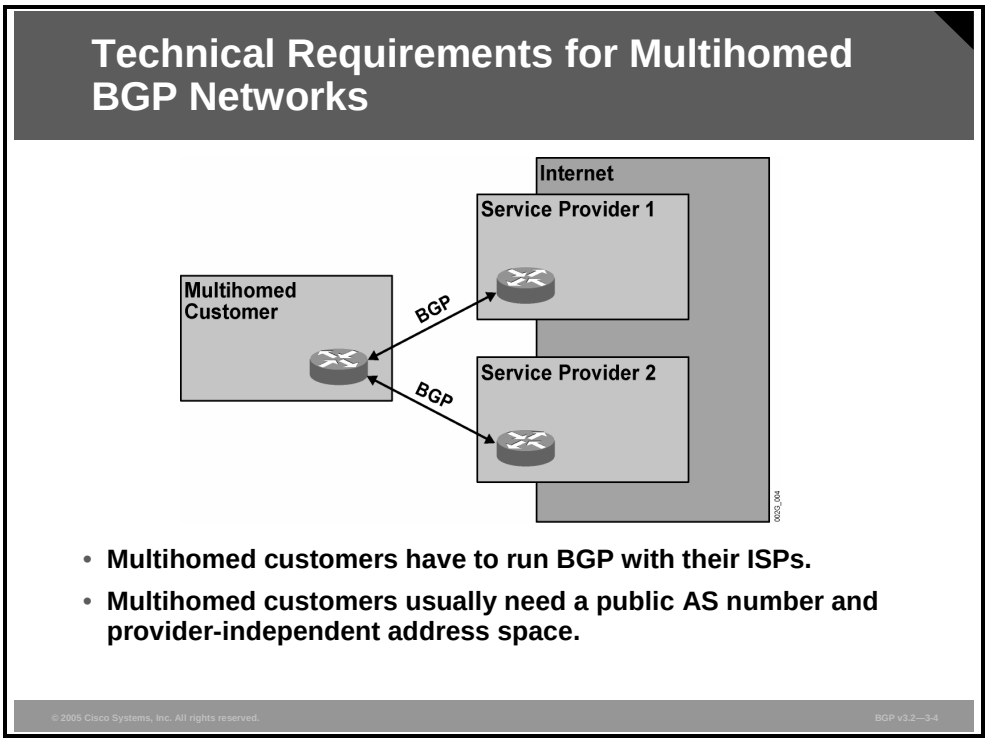
In such cases the company may consider being a customer to two different ISPs or having two separate connections to one ISP.

With two connections to one single ISP, BGP is usually not required. This solution provides backup for link failure and router failure. However, this solution does not provide backup for problems in the network of the ISP or the connection of the ISP to the rest of the Internet.

Full redundancy is achieved only by connecting to two independent ISPs. If one of the ISP networks loses its connection to the rest of the Internet, the customer will still reach the rest of the Internet via the other service provider. At the same time, the customer will still reach those users directly connected to the failing ISP via its direct connection.

Technical Requirements for Multihomed BGP Networks

This topic describes the technical requirements for multihomed BGP networks in service provider environments.



The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.

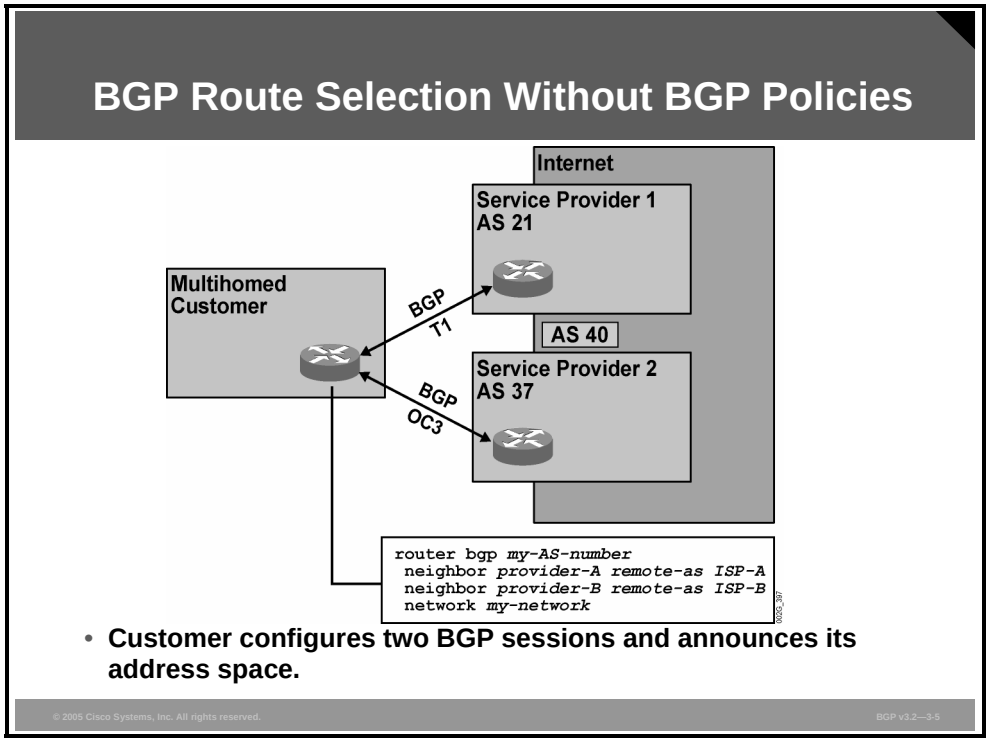
The customer must, in most cases, have its own public autonomous system (AS) number and announce its own IP networks to both ISPs. The ISPs will propagate customer announcements to the rest of the Internet, and the customer will be seen as reachable via both ISP networks. The customer network also receives full Internet routing from both ISPs. This capability gives the customer network the opportunity to choose the best connection at that time to reach any destination on the Internet.

Most customers are not multihomed. They do not exchange BGP with their ISP. Instead, they use default routing to the ISP, and the ISP does static routing to the customer. ISPs use this fact to optimize the number of prefixes that they announce into the Internet. IP network numbers are usually assigned to customers from a range of IP networks that are delegated to the ISP. This situation means, in the ideal case, that all customers that are connected to one single ISP can have their IP networks summarized in a few BGP updates.

In the multihomed scenario, however, the ISP cannot benefit from IP network number assignment from the delegated range. The customer is connected to two different ISPs, and it is not obvious from which provider-assigned address space it should get the IP addresses. The best solution is to do the assignment from a range completely independent of the providers, a provider-independent address space.

BGP Route Selection Without BGP Policies

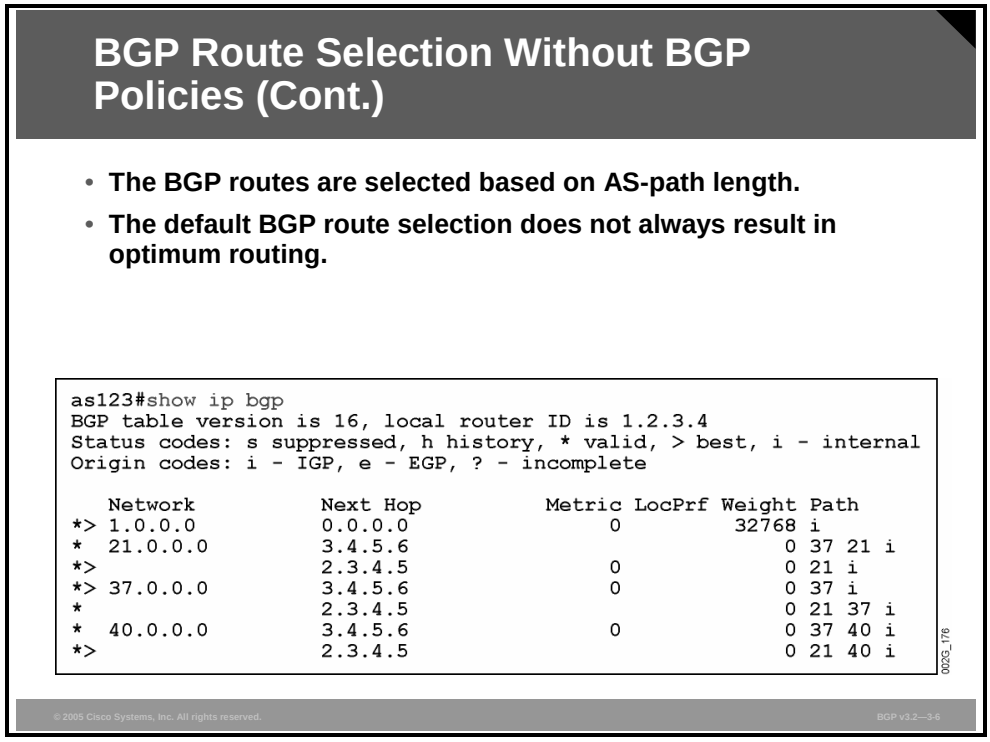
This topic describes the need for BGP policies that influence route selection in a multihomed BGP network.



The simple approach illustrated in the figure may be the source of many problems. By simply starting BGP sessions with both ISPs, and announcing the customer's networks to both ISPs, that customer could experience difficulties as a result of the default behaviors of BGP. The following example illustrates problems that may occur in this environment.

Example: BGP Route Selection Without BGP Policies

This is an example of a multihomed customer with AS 123.



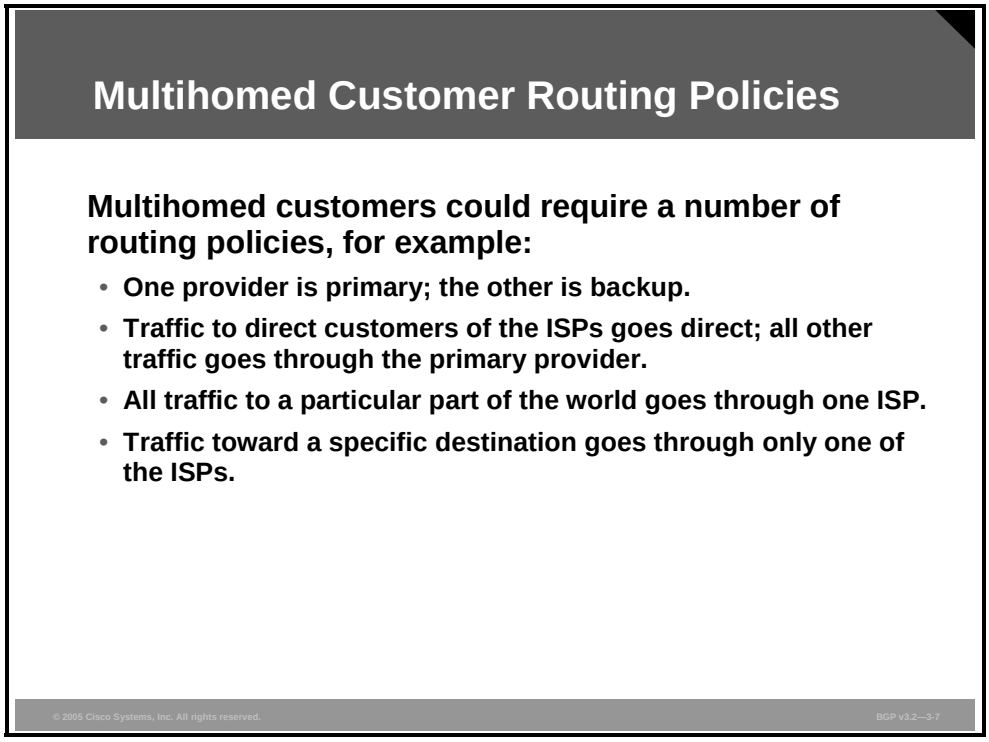
The multihomed customer is connected to two ISPs: AS 37 and AS 21. The two ISPs are interconnected, and both are also connected to AS 40.

The customer receives all routes from both service providers, giving redundancy. The default BGP route selection prefers the shortest AS path. If the AS-path lengths are equal, BGP prefers the most stable route, or the route that is received from the peer with the lower router-ID.

In many cases, however, this route is not the most optimal way to reach all destinations. For example, the bandwidth that is available to reach the ISPs has not been taken into consideration. To change the route selection behavior, some BGP parameters must be configured to support more complex routing policies.

Multihomed Customer Routing Policies

This topic lists typical routing policies for multihomed BGP customers.



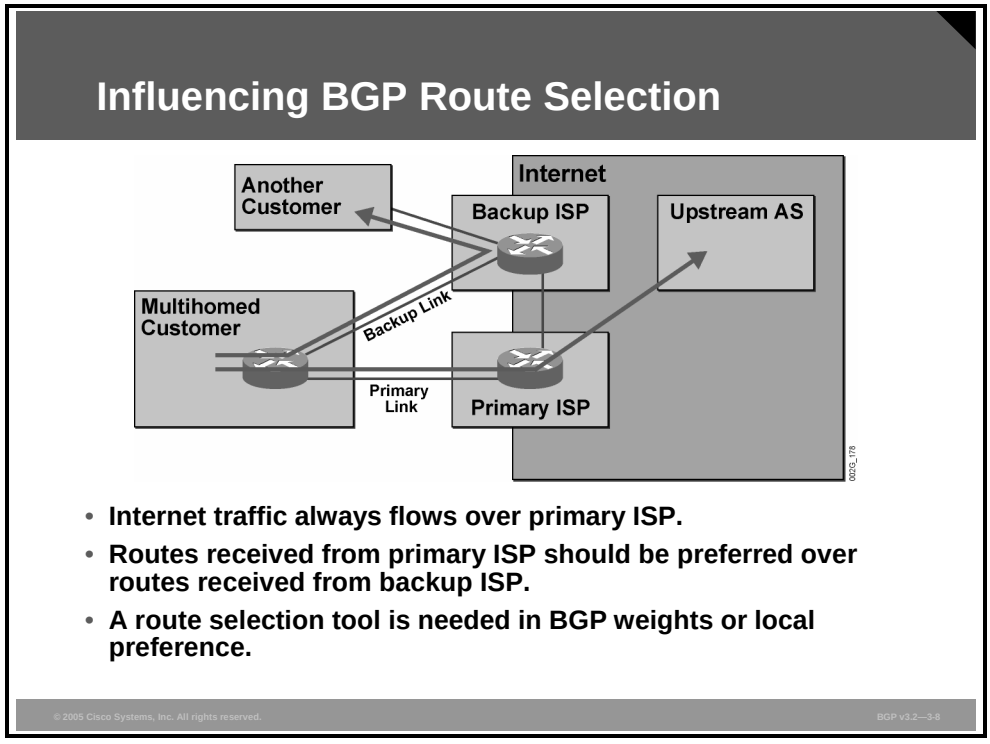
Depending on the circumstances, here are the different polices that a multihomed customer might require:

- One of the two ISPs can be considered the primary connection. This distinction can be the result of available bandwidth or commercial agreements. However, although one of the ISPs is considered the primary connection, some users may have direct connections to the secondary ISP. Therefore, going via the primary ISP to reach users that are connected to the secondary ISP may be suboptimal.
- Destinations in one part of the world may be reached more optimally via one of the ISPs, rather than via the other, because the two ISPs may have different infrastructures and peering agreements with other ISPs.

It is virtually impossible to establish a routing policy that gives optimal routing to every destination on the Internet. Optimization can be done only with the most common destinations in mind. This situation can result in specific rules on how to reach specific destination networks or the AS.

Influencing BGP Route Selection

This topic describes the need to influence BGP route selection in a service provider environment.

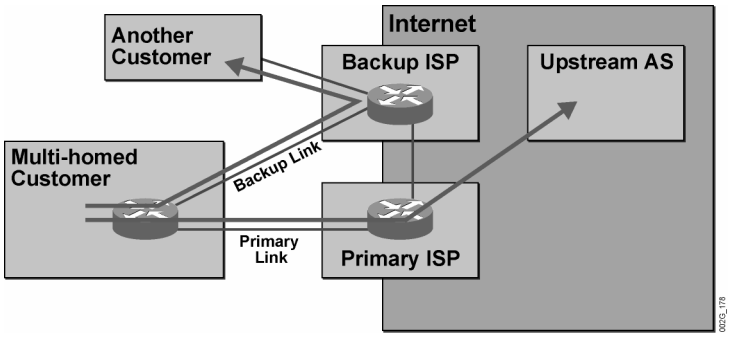


When one of the two ISPs is designated as a primary ISP and the other as a backup, BGP attributes must be configured as a means of influencing BGP route selection rules. If both ISP connections terminate in one single customer router, all routes that are received from the primary ISP can be assigned a BGP weight. A higher weight indicates a more preferred path.

However, the weight value is local to one router. The weight value is not shared between routers. If one ISP connection terminates in one of the customer routers and the other ISP connection terminates in another, the two customer routers must agree on which link to use. Using local preference instead of weight can do this. All routes that are received from the primary ISP over the primary link are assigned a local preference value, which is higher than the default value of 100. The customer router that receives the routes from the primary ISP completes the assignment and communicates the information to the other routers within the AS of the customer.

The result of using either weight or local preference is that the AS of the customer reaches all its destinations on the Internet via the primary link as long as it is available and reaches those destinations within the AS of the secondary ISP. In the case of link failure or failures within the network of the primary ISP, some of the routes, or all of the routes, will no longer be received over the primary link. In that case, the AS of the customer no longer sees those destinations as reachable over the primary link. The only remaining choice is the backup link. Therefore, the backup link is used by the customer network only to reach destinations that are not reachable over the primary link.

Influencing BGP Route Selection (Cont.)



- **Internet traffic flows over primary ISP; traffic to customers of backup ISP goes direct.**
- **Route selection has to be performed based on AS numbers in the AS path.**

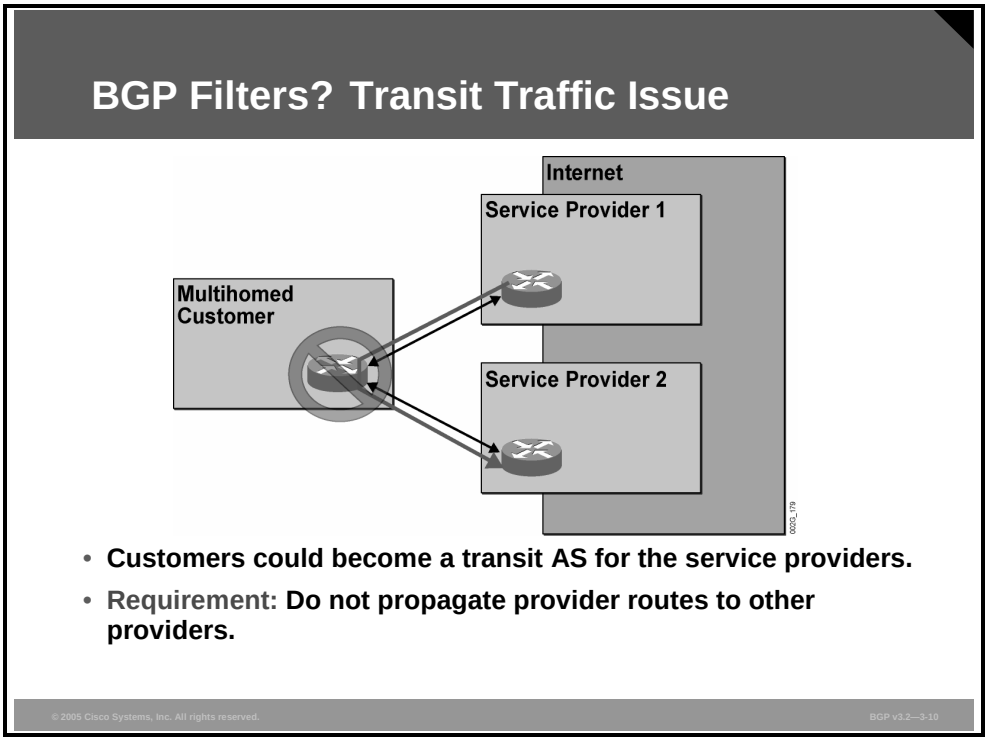
In most cases, it is optimal to reach other customers connected to the backup ISP via the backup link, compared with reaching them via the primary link.

The routing policy described previously, where routes are blindly preferred if they are received on the primary link, can easily be modified to use the backup link for destinations in the AS of the backup ISP. On the router, filtering tools can be configured to select routing information that is based on the content in the AS-path attribute. Those routes, with an AS-path attribute matching specific selection criteria, can be assigned an even higher weight or local preference.

This approach results in a routing policy that gives precedence to reaching destinations within the AS of the primary ISP and within all autonomous systems upstream of the primary ISP over the primary link. Destinations within the AS of the backup ISP receive precedence over the backup link.

BGP Filters

This topic describes the need for BGP filters in a service provider environment.

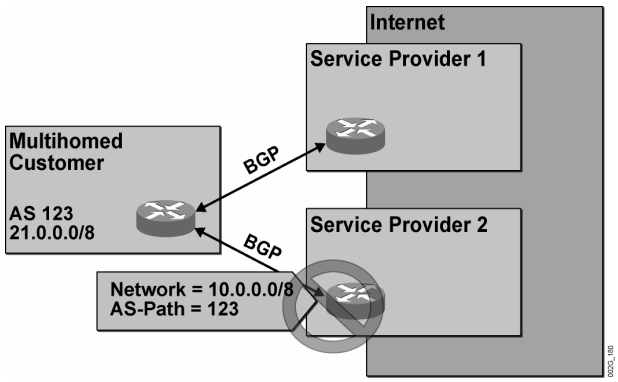


When BGP has selected the best path, the information is advertised by the router to all neighboring autonomous systems, except for the session that it was received on (called “BGP split-horizon functionality,” which prevents near-range routing loops). This situation causes the customer AS to become a transit AS between the two ISPs, and should be avoided.

Most customers do not intend to create transit traffic between ISP networks. The access lines to the ISPs are not suited to carry this volume of traffic, and the customer certainly does not want to have its bandwidth consumed by transit traffic.

The solution to this problem is to filter outgoing information to both ISPs. Filtering of routing information is performed based on the content of the AS-path attribute that is assigned to every BGP route. Only routes having an AS-path attribute that indicates that they are sourced by the AS of the customer are allowed to be sent to either of the two ISPs.

BGP Filters? Routing Update Reliability Issue



- **Customers running BGP could announce any route to the service providers.**
- **Requirement: Service providers have to filter IP prefixes in incoming updates.**

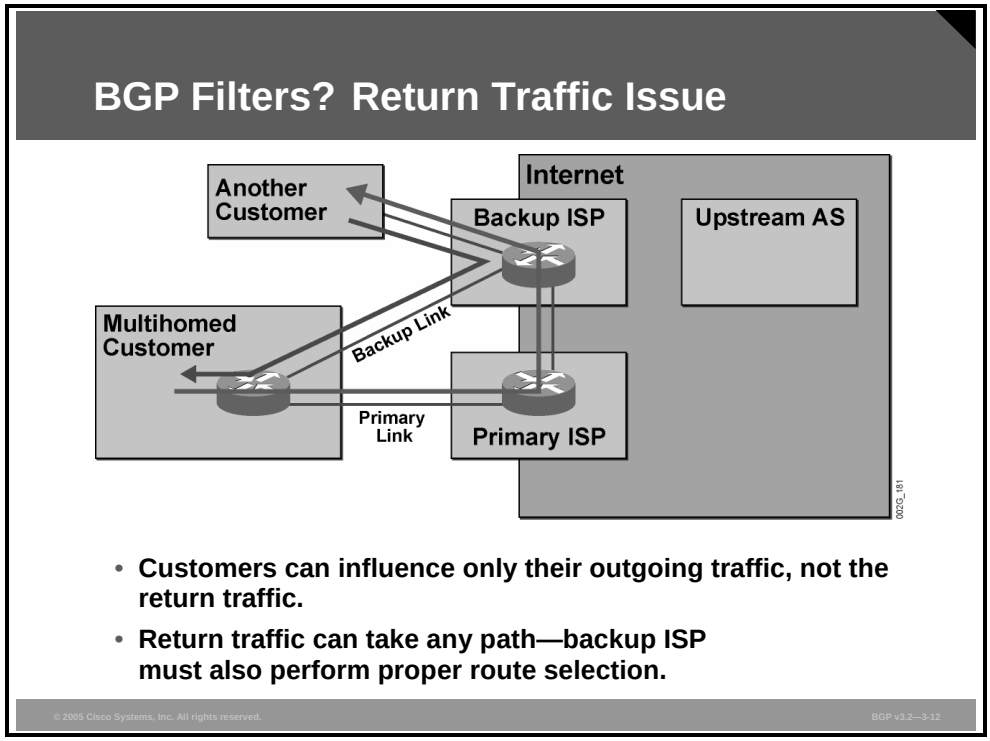
Without some sort of filtering, BGP routing information that is created by the AS of the customer can potentially be propagated all over the Internet. In this way, the customer can inject erroneous information into the Internet routing tables.

Customers are much less experienced in avoiding these kinds of problems than are service providers. There is much more risk of errors being introduced when a customer is assigned its own AS and uses BGP with the ISP, as compared with the single-homed scenario in which the ISP has sole responsibility to announce BGP routes to the rest of the Internet.

Almost all of the Internet problems that a customer can cause by improperly configuring its BGP can be stopped by the ISP. The ISP should filter all incoming information from the customer and accept only what is supposed to arrive. The ISP should discard anything outside strict limits. In this way, the ISP prevents the propagation of erroneous information to the rest of the Internet.

The ISP can maintain a list of the IP network numbers that the customer is announcing and filter out any other route. If this approach is not possible because of the volume of those lists, the ISP should at least be able to filter out the most obvious erroneous announcements.

Note Private addresses, according to RFC 1918, should never be announced to the Internet.



The customer can easily define a policy about how to send outgoing IP packets on the correct link. It is much harder to influence the neighboring AS about how to direct the IP packets coming into the customer network.

A customer that creates a routing policy in which one of the two ISPs is always preferred may see that the return traffic is arriving on what the customer thought was the backup link. This situation means that the customer has configured the weight or local preference to make sure that all outgoing traffic is leaving the customer AS over the primary link, but the backup ISP does not have any such configuration. Therefore, return traffic enters the customer AS by using the shortest AS path as its selection criterion.

The best way to solve this problem is for the customer to ask the backup ISP to change its routing policy. The change should cause the backup ISP to prefer reaching the customer AS via the AS of the primary ISP. The backup ISP must implement this change in its own AS.

Note	Sometimes the backup ISP administrator might be reluctant to change the configuration for a single customer. In this case, the customer should use another BGP feature, the AS-path prepending tool, to influence the selection of the primary or backup link by lengthening the AS path of routes that are sent to the backup provider.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Some customers need redundant Internet access for their mission-critical applications and address this need by having two separate connections to one ISP or implementing a multihomed configuration (connecting to two different Internet service providers).**
- **The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.**
- **An approach to multihoming that is too simple can be a source of problems. Starting BGP sessions and announcing customer networks to multiple ISPs by using the default behavior of BGP may not result in optimal routing.**

Summary (Cont.)

- **Depending on the circumstances, a multihomed customer may require different polices, such as one of the two ISPs being considered the primary connection or reaching destinations in one part of the world more optimally via one of the ISPs. Optimization should be done with the most common destinations in mind, resulting in specific rules on how to reach specific destination networks or the AS.**
- **In BGP route selection, a routing policy may be created that gives precedence to reaching destinations within the AS of the primary ISP and all upstream autonomous systems over the primary link and reaching destinations within the AS of the backup ISP over the backup link.**
- **When BGP has selected the best path and the information has been propagated to all neighboring autonomous systems, the customer AS may become a transit AS between the two ISPs. The customer must avoid this situation by using BGP filters.**

Lesson 2

Employing AS-Path Filters

Overview

In network implementations that require connections to multiple Internet service providers (ISPs), network operators typically use autonomous system (AS)-path filters to influence Border Gateway Protocol (BGP) route selection. It is important for a network administrator to understand the syntax of an AS-path regular expression and how string-matching operators function when they are using AS-path regular expressions to match BGP routes.

BGP allows connectivity between multiple ISPs for redundancy and scalability. Service providers employ AS-path filters to remedy the problems that are associated with the various connectivity methods that are used within BGP. This lesson explains the methods that are used to implement BGP AS-path filters.

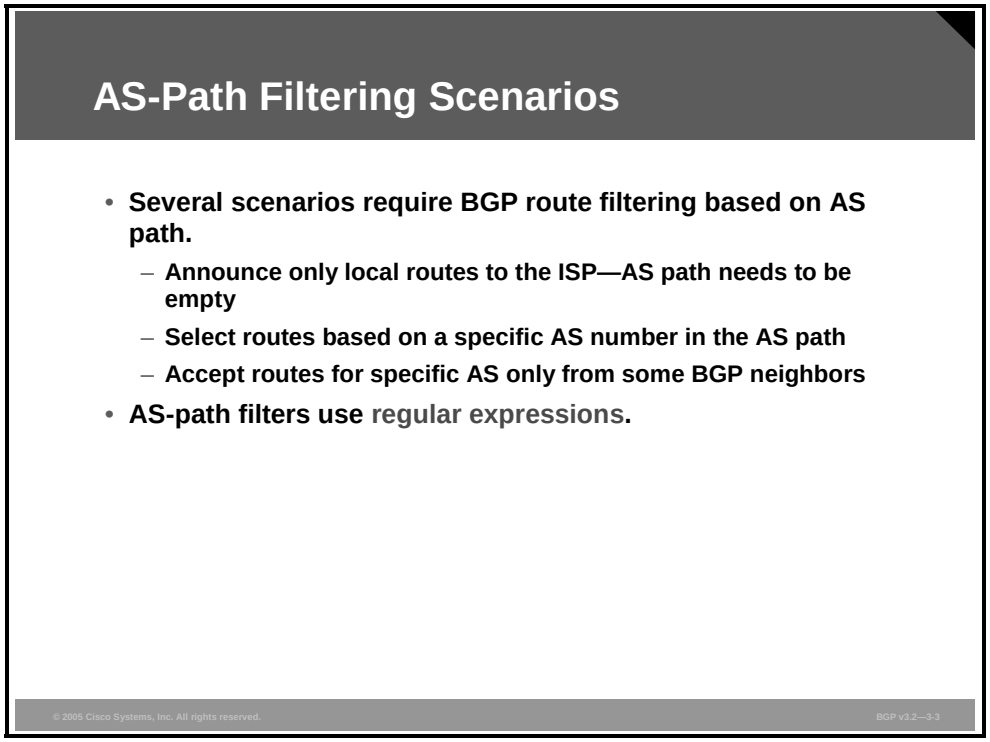
Objectives

Upon completing this lesson, you will be able to successfully configure BGP to influence route selection using AS-path filters in a customer scenario in which connections to multiple ISPs must be supported. This ability includes being able to meet these objectives:

- Identify network scenarios in which you must support connections to multiple ISPs and in which AS-path filters can be used to influence route selection
- Describe the function of an AS-path regular expression
- Explain how string-matching operators function when you are using AS-path regular expressions to match BGP routes
- Identify where you can apply an AS-path filter when configuring a router to influence route selection
- Identify the Cisco IOS commands that are required to configure AS-path filters to influence route selection
- Identify the Cisco IOS commands that are required to monitor the operation of configured AS-path filters

AS-Path Filtering Scenarios

This topic identifies network scenarios that require connections to multiple ISPs where route selection must be influenced with AS-path filters.



Several scenarios require filtering and selection of routing information, based on the content of the AS-path attribute. Each BGP route must have an AS-path attribute. It is a well-known mandatory attribute and must therefore be present in each BGP update.

Using selection criteria that are based on the AS-path attribute, a router can identify a set of specific routes from the total set of routes that it receives. Those routes where the AS-path contents match the criteria are selected. Routes that do not match the criteria are not selected.

The AS path is a sequence of numbers. Each number indicates an AS. When a route is sourced by means of a **network** command in a BGP process or redistribution into a BGP process, the AS-path attribute is created and left empty. Each time the route is advertised by an egress router to another AS, the AS-path attribute is modified by the egress router, which prepends its AS number to the AS-path attribute.

While a newly sourced route is still within the AS in which it was created, the AS path is empty. When the AS has a requirement to filter out all but the routes that are local to itself before sending them to a neighboring AS, the AS will permit sending of the routes with the empty AS path and will deny all others.

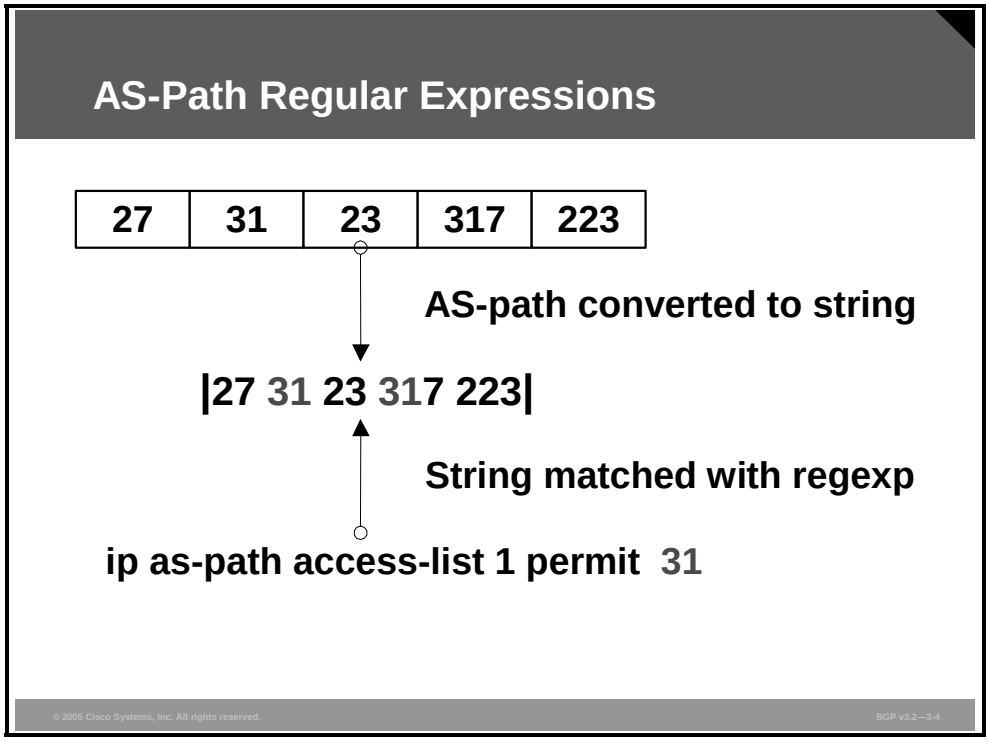
Routers can also filter incoming routes based on their AS-path attributes. Some destination autonomous systems should not be received from a certain neighbor. Therefore, the routes matching that AS in the AS path can be filtered on the receiving router in case they are accidentally sent.

Selection based on the AS path is also a tool that you can use when changing the weight or local preference attributes for some destination autonomous systems but not for others.

When routers filter BGP updates based on the content of the AS-path attribute, they use regular expressions. Regular expressions are commonly found in the UNIX environment and also in some Microsoft Windows-based applications. Regular expressions are a string-matching tool. A regular expression consists of a string of characters. Some of these characters have special meanings, such as functioning as wildcards and operators, and some of these characters simply mean themselves, for example, A to Z, a to z, or 0 to 9. A regular expression is said to match a string if the ordinary characters and the applied meaning of the special operator characters can be translated into the matched string. When a regular expression matches, the selection test is said to be true. If it does not match, the test is false.

AS-Path Regular Expressions

This topic describes the function of an AS-path regular expression.



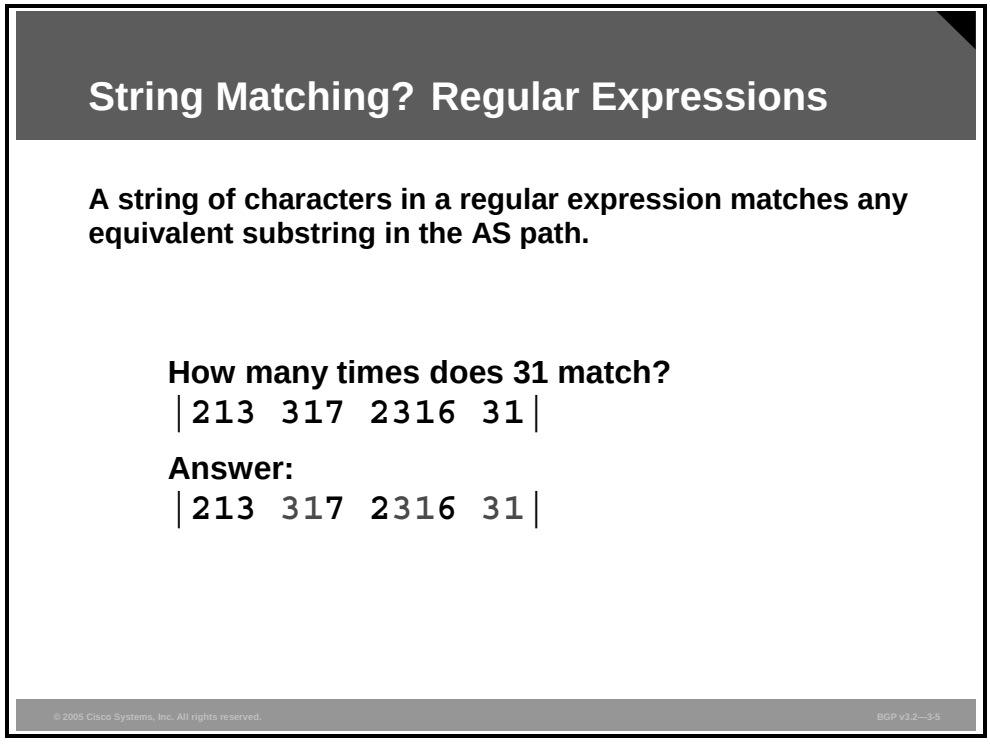
The AS-path attribute carried with all BGP routes in a BGP update is a very compact binary encoding of a sequence of integer numbers; it is not a sequence that can be tested by using a regular expression.

Cisco IOS software internally translates the binary encoding into a character string. Each AS number in the sequence is converted into a string using decimal representation. The space character separates each AS number in the AS-path attribute. The router applies the regular expression test to this internally created character string.

Characters in a regular expression that are not assigned a specific operation match themselves. The regular expression “31” matches all occurrences of the character “3” followed by the character “1” in the AS path. In this example, “31” matches at two occurrences. One occurrence is sufficient to make the test true. No occurrence means that the test failed.

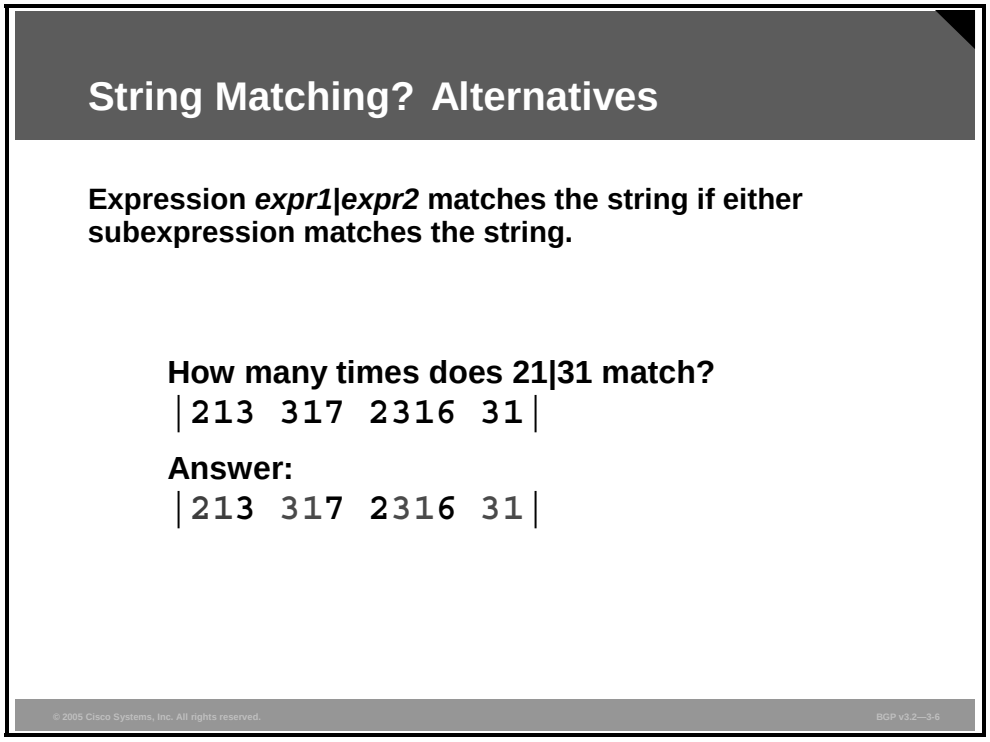
String Matching

This topic describes how string matching functions when you are using AS-path regular expressions to match BGP routes.



The regular expression “31” will match any occurrence of “3” followed by “1” regardless of the characters immediately preceding the “3” and immediately following the “1”. Therefore, “31” will match an occurrence of “3” and “1” in the middle of an AS number.

The regular expression “31” matches the AS-path string “213 *31*7 2*31*6 *31*” three times, because “31” matches a part of “317”, “2316”, and “31”.



The character "|" (vertical bar) has a special meaning. It is an operator that means "or." The regular expression "21|31" matches the sequence of "2" followed by "1" or the sequence of "3" followed by "1". Therefore, this sample regular expression will match a two-character sequence: the "21" or the "31".

The regular expression "21|31" matches the AS-path string "*21*3 *31*7 2*31*6 *31*" four times, because "21" matches a part of "213" and "31" matches a part of both "317" and "2316" and also "31".

String Matching? Ranges and Wildcard Characters

A range of characters matches any single character in the range.

Examples: [1234] or [1-4]

Dot (.) matches any single character

How many times does [1-3].[34] match?

```
|213 317 2316 31|
```

Answer:

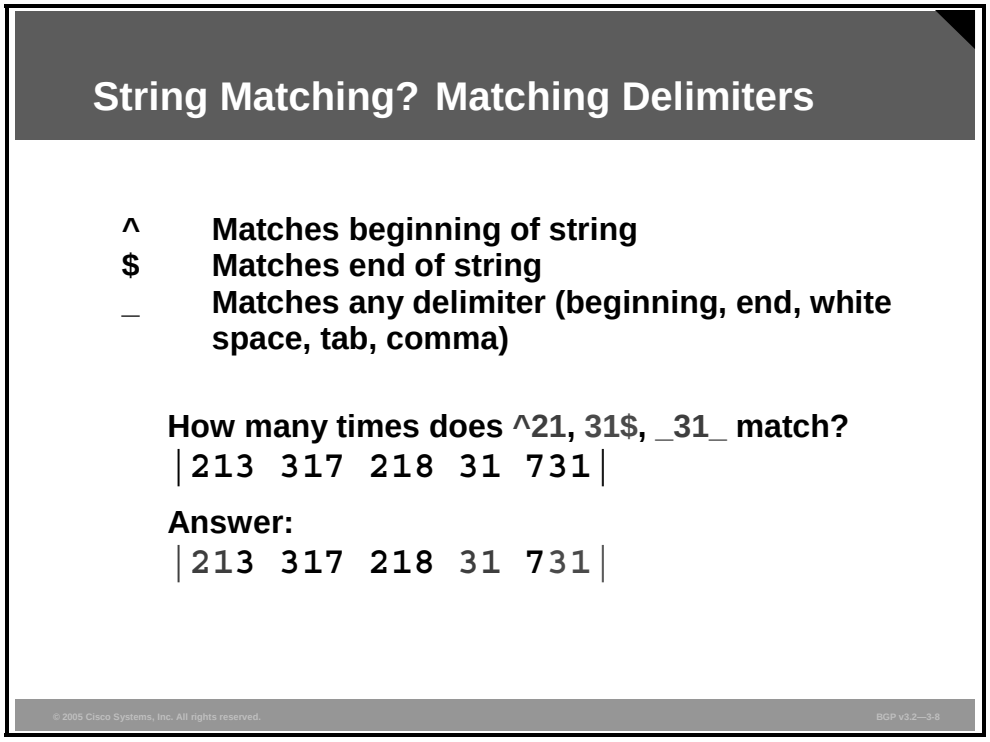
```
|213 317 2316 31|
|213 317 2316 31|
```


The pair of brackets “[” and “]” has a special meaning. Brackets surround a set of characters of which any one matches. The set of characters is either expressed as the list of the matches (for example, “[1234]”) or the sequence with the starting character, a hyphen, and the ending character (for example, “[1-4]”). Both examples match one single character, which must be any one in the set of the four characters “1”, “2”, “3”, and “4”.

The character “.” (a dot) matches any single character. Small regular expressions can be combined into a larger expression. Such a combination is matching if all of the parts match one after the other. The sample regular expression “[1-3].[34]” matches a sequence of three characters, of which the first must be either “1”, “2”, or “3”, the second character can be any character, and the third must be either “3” or “4”.

Note The space character delimiting two AS numbers is just a character. The dot (“.”) for example, matches this character.

The regular expression “[1-3].[34]” matches the AS-path string “213 317 2316 31” twice. Initially, it matches “213”. The leading “[1-3]” matches the leading “2”. The dot, which matches any character, matches the “1”, and “[34]” matches the trailing “3”. Secondly, the regular expression also matches in “213 317 2316 31”. This is a little harder to see, because the dot (“.”) matches the space character between “213” and “317”.



A character string must have a start and an end. The character with the special meaning "^" matches the beginning of a string. Because all strings have a beginning, the "^" character matches all strings. However, the "^" character is used to position the part of the regular expression that follows. The character following the "^" character must be the first character of the string; otherwise, that character would not match the beginning of the string.

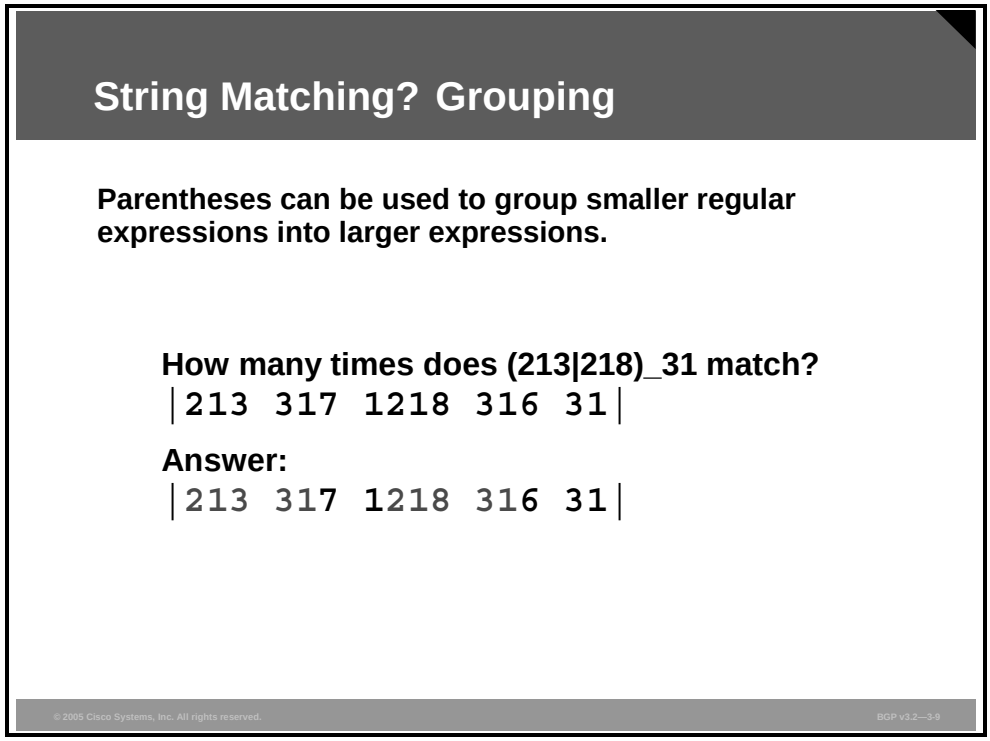
The special character "$" is used analogously, but it means the end of the string. The character preceding the "$" must be the last character in the string; otherwise, the "$" does not match the end of the string.

The underscore ("_") matches any delimiter. The space character between two AS numbers is an example of a delimiter. The beginning of the string and the end of the string are also considered delimiters. Other delimiters are the tab and the comma (","). The underscore ("_") is used to ensure that the desired AS number is found in an AS-path string but not as part of some other AS number. For example, the regular expression "31" will match the AS number string "317", but the regular expression "_31_" will not. Both "31" and "_31_" will match the AS number string "31".

The regular expression "^21" can match the AS-path string "213 317 218 31 731" only one time because there is only one beginning of the string. The regular expression "^21" matches only if the string starts with the sequence "21", which it does.

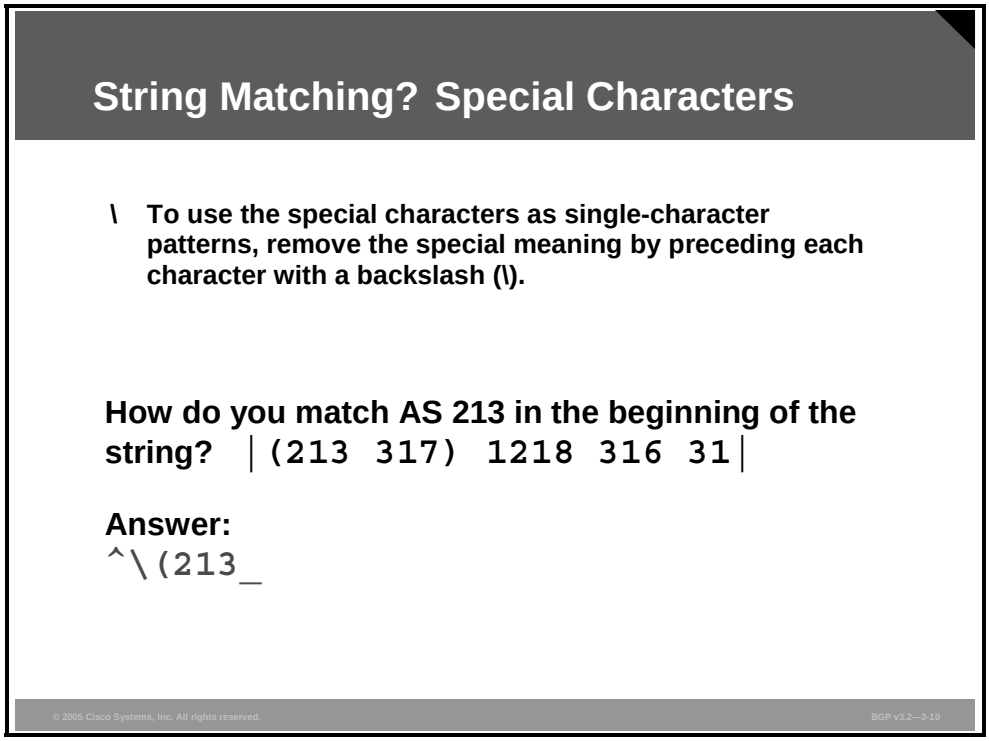
The regular expression "31$" can match the AS-path string "213 317 218 31 731" only one time because there is only one end of the string. The regular expression "31$" matches only if the string ends with the sequence "31", which it does.

The regular expression "_31_" can, in theory, match an AS-path string several times. However, in this case, when matched against the string "213 317 218 31 731", the regular expression "_31_" matches only the AS number "31" in the AS path.



Complicated expressions must sometimes be grouped with parentheses, "(“and")". This feature can be useful when you are searching for a sequence of two or more AS numbers of which the first can match any of the specific autonomous systems (in this example, "213" or "218"), but the last must be a specific AS ("31" in this example). If the parentheses were not used here, the expression would match either the single AS "213" or the sequence of the two (218 31).

The regular expression "(213|218)_31" matches the AS-path string "213 317 1218 316 31" twice. The first match is "*213 31*7 1218 316 31"; the second match is "213 317 1*218 31*6 31".



Sometimes the target string that you are trying to match with a regular expression contains some of the characters that also have special meanings in the regular expression. To match these characters in the target string, use the backslash ("\") together with the character in the regular expression.

Note	This type of regular expression syntax is used for matching AS-path strings inside a BGP confederation. A confederation is used to eliminate the scaling problem of full-mesh Internal Border Gateway Protocol (IBGP) by splitting the AS into smaller regional autonomous systems. The example shows that 213 and 317 were part of a confederation by its use of "(" and ")". Confederations are explained further in the "Scaling Service Provider Networks" module.

String Matching? Repeating Operators

*	Matches zero or more atoms
?	Matches zero or one atom
+	Matches one or more atoms

An atom is a single character or a grouping.

How do you match AS sequences “23 45” and “23 78 45” in a single regular expression?

Answer:

```
_23(_78)?_45_
```


The special characters, star (asterisk), “*”, question mark, “?”, and plus, “+”, all apply repetition of the expression that immediately precedes them.

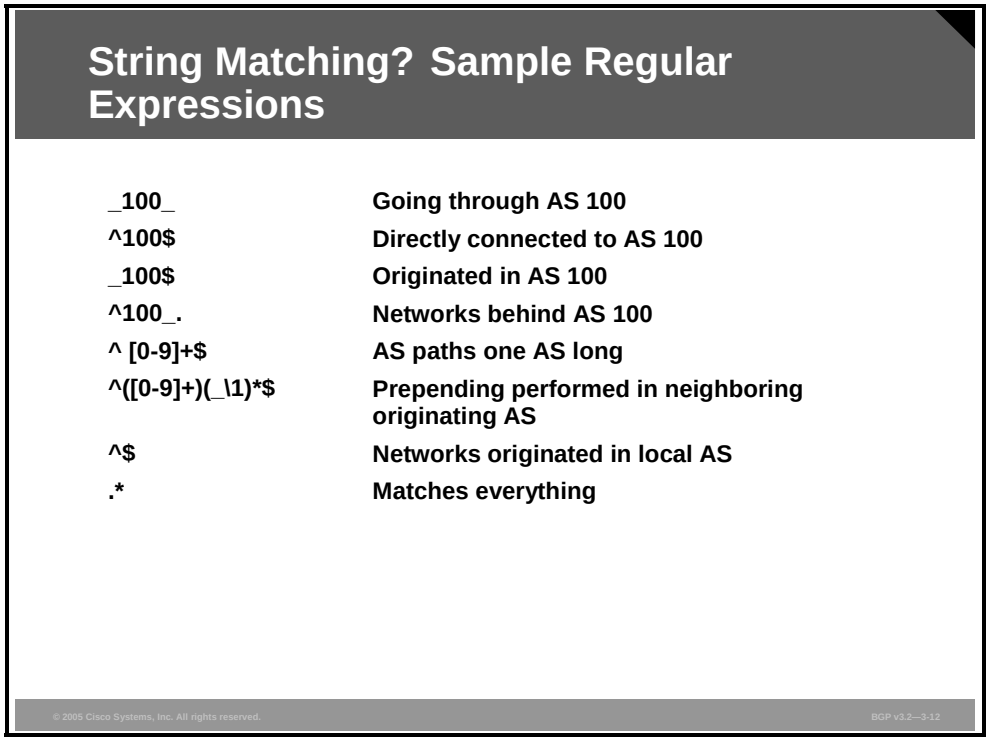
The star (asterisk), “*”, means that the expression that immediately precedes it is repeated zero or more times. This means that the expression may not be there, but it may also be there any number of times. The expression “1*” will match a sequence of no characters or a sequence of any number of the character “1”.

A question mark (“?”) means that the expression that immediately precedes it is repeated zero or one time. This means that the expression may not be there, but it may also be there once. The expression “1?” will match a sequence of no characters or the single character “1”.

The plus sign (“+”) means that the expression that immediately precedes it is repeated one or more times. This means that the expression must be there at least once. The expression “1+” will match a sequence of one or more of the character “1”.

Example: String Matching

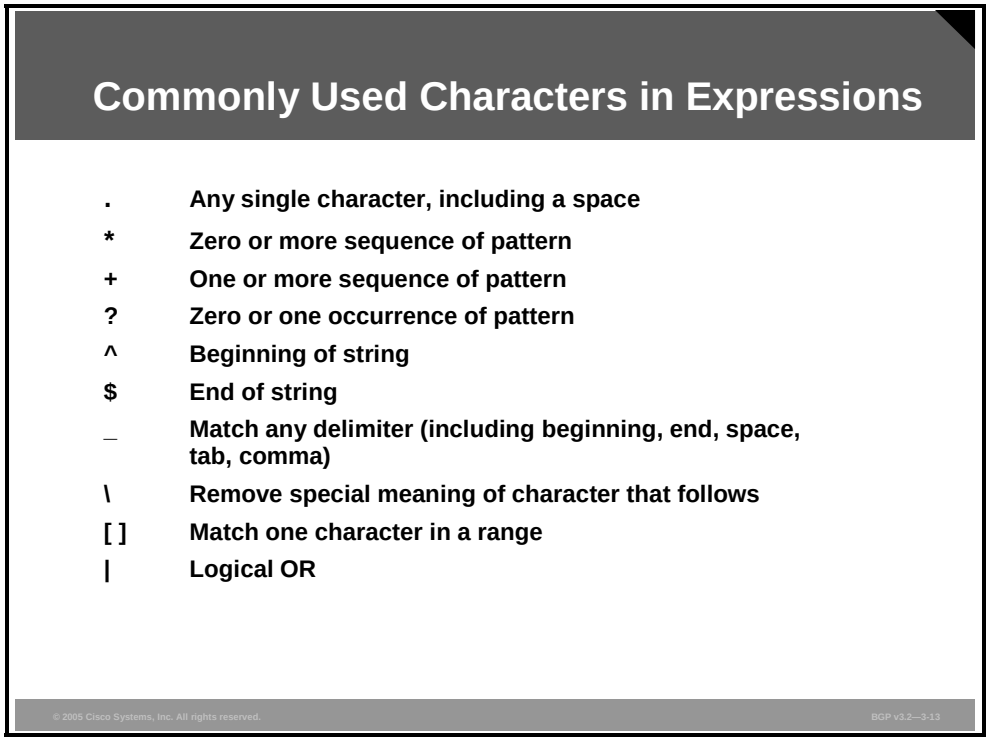
This example shows a sample of regular expressions.



Regular expressions can be arbitrarily complex. However, most searching is accomplished using regular expressions similar to those in these examples:

- If you are searching for all routes that have AS 100 in their AS paths, the regular expression to use is “_100_”.
- If you are searching for all routes that are sourced in your directly connected neighboring AS 100, the regular expression to use is “^100$”. Use an expression with both a carat (“^”) and dollar sign (“$”) present when you are searching for an exact match.
- If you are searching for all routes that are sourced in AS 100, but that AS is not necessarily a directly connected neighboring AS, the regular expression to use is “_100$”. The dollar sign (“$”) indicates that the AS path must end with AS 100. This is an indication that the route was sourced in AS 100. The underscore (“_”) is used to make sure that it is AS 100 at the end of the string and not, for example, AS 2100.
- If you are searching for all the routes that are reachable behind AS 100, the regular expression to use is “^100_.”. The carat (“^”) indicates that the AS path must start with AS 100. The underscore (“_”) is used to make sure that this number is not matching with, for example, AS 1001. The dot (“.”) is used to indicate that the AS path does not end with AS 100, and that there must be something following AS 100.
- If you are searching for all routes that are sourced in any AS directly neighboring your AS, the regular expression to use is “^[0-9]+$”. The “[0-9]” part means any digit. The plus sign (“+”) repetition character means one or more times. Therefore, the combination “[0-9]+” means a sequence of one or more of digits. The carat (“^”) and dollar sign (“$”) mean the beginning and the end of the string. Therefore, the string may consist only of a sequence of one or more digits.

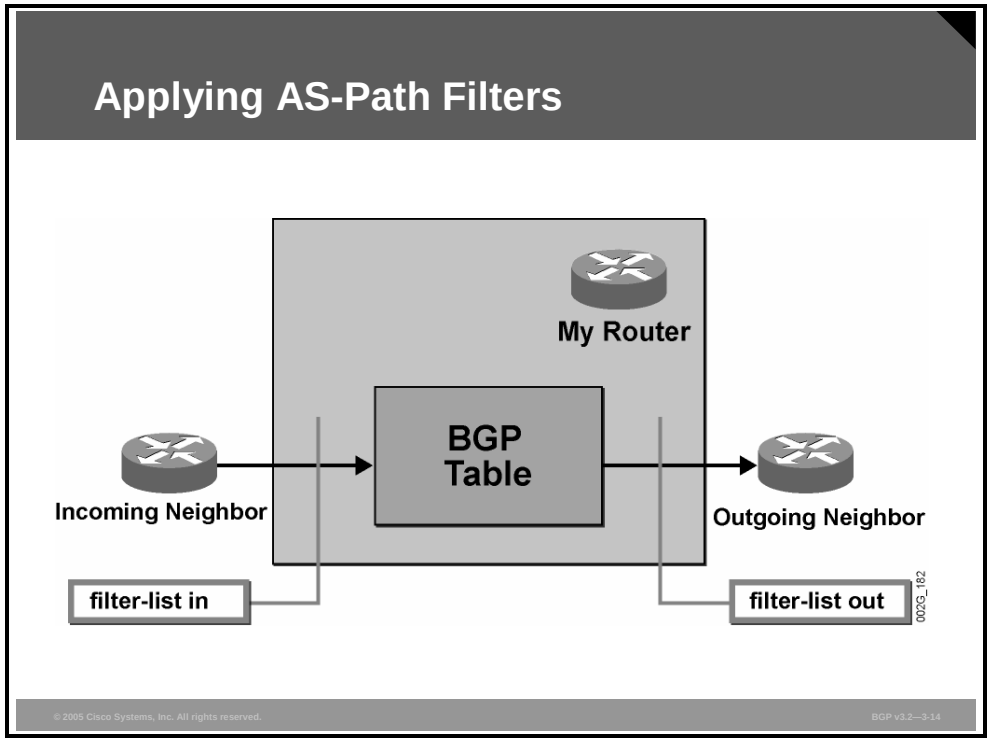
- If you are searching for all routes that are sourced in any AS directly neighboring your AS, and possibly performing AS-path prepending (multiplication of a directly connected AS number), the regular expression to use is “^([0-9]+)(_\1)*$”. The expression in the first set of parentheses matches any AS number. The parentheses store the value of the matched AS, and this value is then recalled by the second part of the regular expression, including a variable. The variable “\1” is put into parentheses for the purpose of the multiplier operator “*”, meaning that this part can match any number of successive occurrences of the same AS number that was matched by the “[0-9]+” expression. For example, this regular expression matches AS paths “99 99 99”, “200”, “101 101”, or “5 5 5 5 5”, but it does not match the AS path “101 99”.
- The combination “^$” means an empty string and is used when you are searching for all routes that are sourced in the local AS.
- Sometimes a search is made to select a few specific routes and do something special with them, while the rest of the routes will be handled in a different way. To search for all routes, regardless of the content of their AS-path attribute, use the regular expression “.*”. The dot (“.”) matches any single character. The repetition character, star (asterisk), “*”, means that the match should be repeated zero or more times. Thus, the combination (“.*”) matches any string.



The figure lists the most commonly used characters in expressions.

Applying AS-Path Filters

This topic identifies where you can apply an AS-path filter when configuring a router to influence route selection.

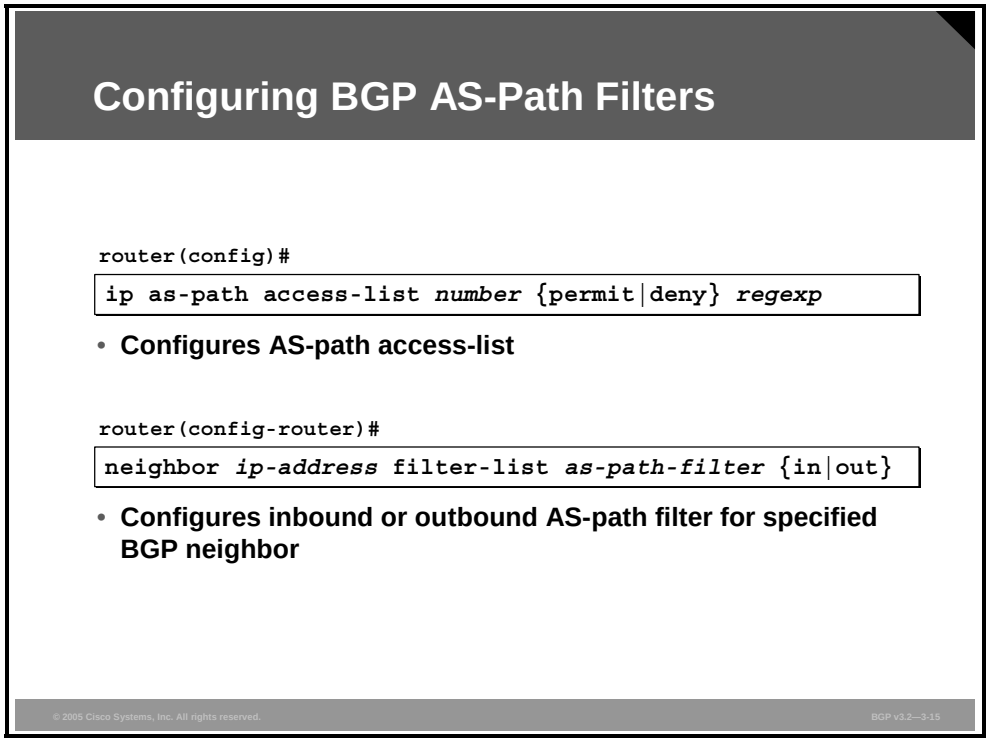


AS-path filters that are configured on a router select those routes that are allowed. Routes that are selected behave as described here:

- The selected routes enter the local BGP table when the selection is applied on the incoming routes from a neighbor; routes that are not selected are silently dropped.
- The selected routes are transmitted to the neighbor when the selection is applied on the outgoing routes to the neighbor; routes that are not selected are used locally but are never sent to the neighbor.

Configuring BGP AS-Path Filters

This topic identifies the commands that you can use to configure AS-path filters to influence route selection.



An AS-path filter is created by an AS-path access-list. This access-list is applied to a set of routes from which to select a subset. Routes that are permitted by the access-list are included in the subset, and those that are denied are not included. As in all access-lists, the candidate to be permitted or denied membership in the subset is tested against all the lines in the access-list, in the order in which the list is configured. The first match indicates “permit” or “deny,” as specified. If the end of the access-list is reached without any explicit match, the candidate is implicitly denied.

The test by the AS-path access-list is performed by using regular expressions that are applied on the AS-path attribute of the route.

The access-list can, for example, be applied on the routes received from, or those sent to, a specific BGP neighbor.

ip as-path access-list

To define a BGP AS-path access-list, use the **ip as-path access-list** global configuration command.

- **ip as-path access-list** *access-list-number* {**permit** | **deny**} *as-regular-expression*

To disable use of the access-list, use the **no** form of this command.

- **no ip as-path access-list** *access-list-number* {**permit** | **deny**} *as-regular-expression*

Syntax Description

Parameter	Description
access-list-number	Integer from 1 to 199 that indicates the regular expression access-list number.
permit	Permits access for matching conditions.
deny	Denies access to matching conditions.
as-regular-expression	AS in the access-list using a regular expression (See the "Regular Expressions" appendix in the *Cisco IOS Dial Services Command Reference* for information about forming regular expressions.)

neighbor filter-list

To set up a BGP filter, use the **neighbor filter-list** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **filter-list** *access-list-number* {**in** | **out**}

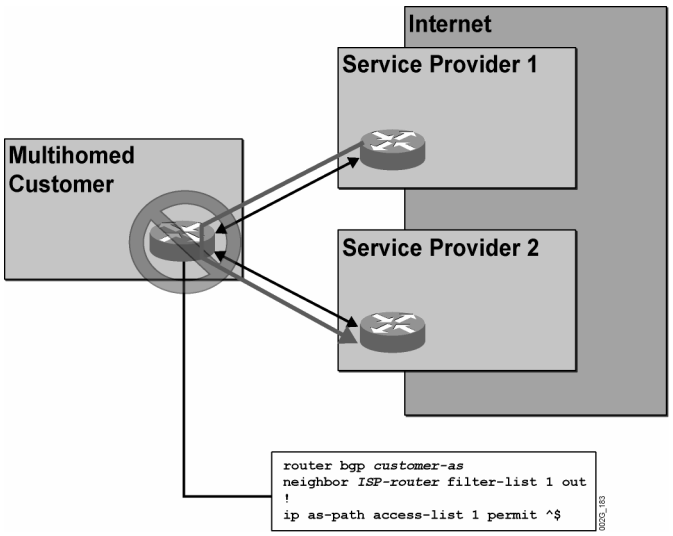
To disable this function, use the **no** form of this command.

- **no neighbor** {*ip-address* | *peer-group-name*} **filter-list** *access-list-number* {**in** | **out**}

Syntax Description

Parameter	Description
ip-address	IP address of the neighbor.
peer-group-name	Name of a BGP peer group.
access-list-number	Number of an AS path access-list. Define this access-list with the **ip as-path access-list** command.
in	Access-list to incoming routes.
out	Access-list to outgoing routes.

Configuring BGP AS-Path Filters (Cont.)

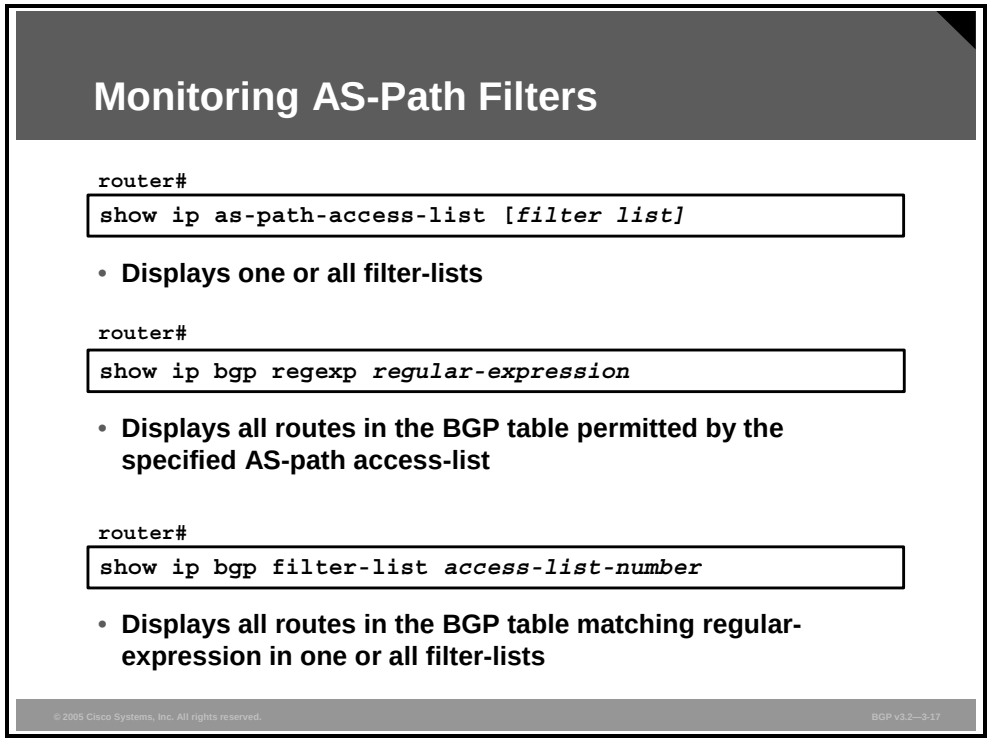


Multihomed customers do not want to act as a transit AS between their service providers. The customer avoids this situation by not transmitting all its routes to its service providers. The service providers send IP packets to the customer only if the IP packets have destination addresses that match one of the routes that the customer has sent by BGP to the service provider. By making sure that only locally sourced routes are sent, the customer avoids receiving IP packets for destinations outside its own AS.

Within the customer AS, the locally sourced routes have empty AS paths. The empty string is matched by the regular expression "^$". The command **ip as-path access-list** *1* permits only the routes that are locally sourced and implicitly denies the rest. By applying this filter-list on outgoing information to all neighbors, the customer will announce local routes only.

Monitoring AS-Path Filters

This topic identifies the Cisco IOS commands that are required to monitor the operation of configured AS-path filters.



The Cisco IOS commands that are most frequently used to monitor the operation of configured AS-path filters include **show ip as-path-access-list**, **show ip bgp regexp**, and **show ip bgp filter-list**.

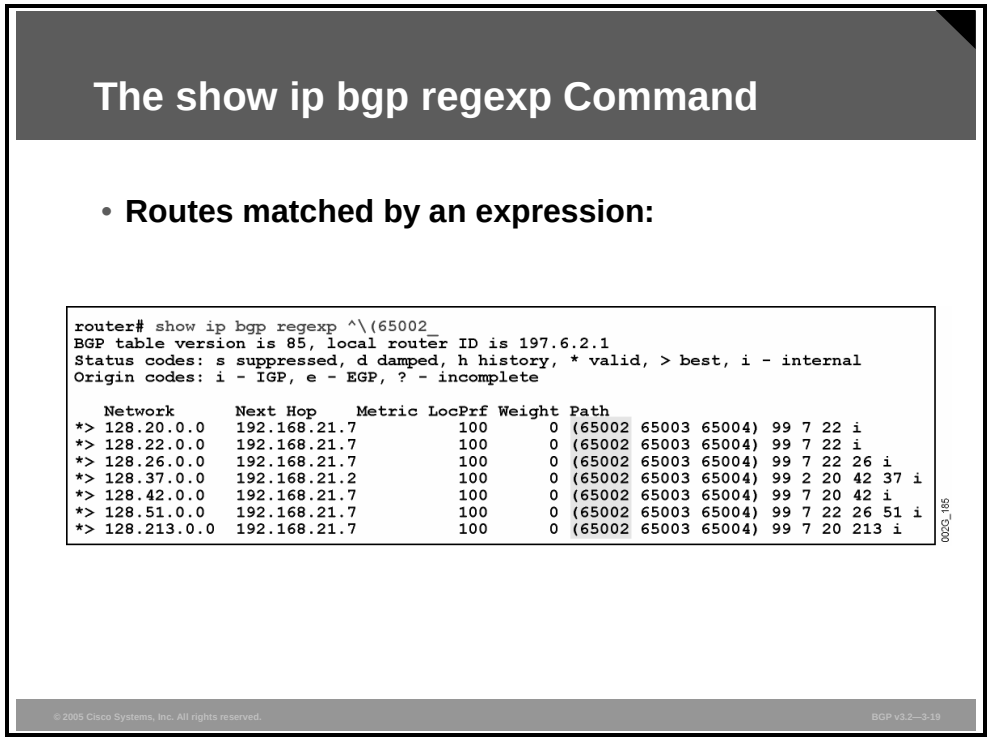
The show ip as-path-access-list Command

- **Displaying configured filters:**

```
router# show ip as-path-access-list
AS path access list 6
    permit ^$
AS path access list 7
    deny _213_
    permit .*
AS path access list 8
    permit _214_
AS path access list 25
    permit _42_
AS path access list 27
    deny _22_|_51$
    permit .*
```

002G_184

The **show ip as-path-access-list** command displays a specific access-list or all AS-path access-lists in the router.



Because regular expressions sometimes get complex, thorough testing of them is required. The **show ip bgp regexp** command displays all routes currently in the BGP table that have an AS-path attribute that matches the typed-in regular expression. Use the **show ip bgp regexp** command to test a regular expression that is typed in on the command line. The result is a printout on the screen of all those routes currently in the BGP table that had an AS-path attribute matching the typed-in regular expression.

In the example figure (which shows BGP confederations, a scalability feature), you wish to find all BGP confederation routes from AS number 65002. To search for this character in the beginning of the string, use the character "\", the backslash. The regular expression "^\(65002" matches all the routes that are received from the intra-confederation AS number 65002.

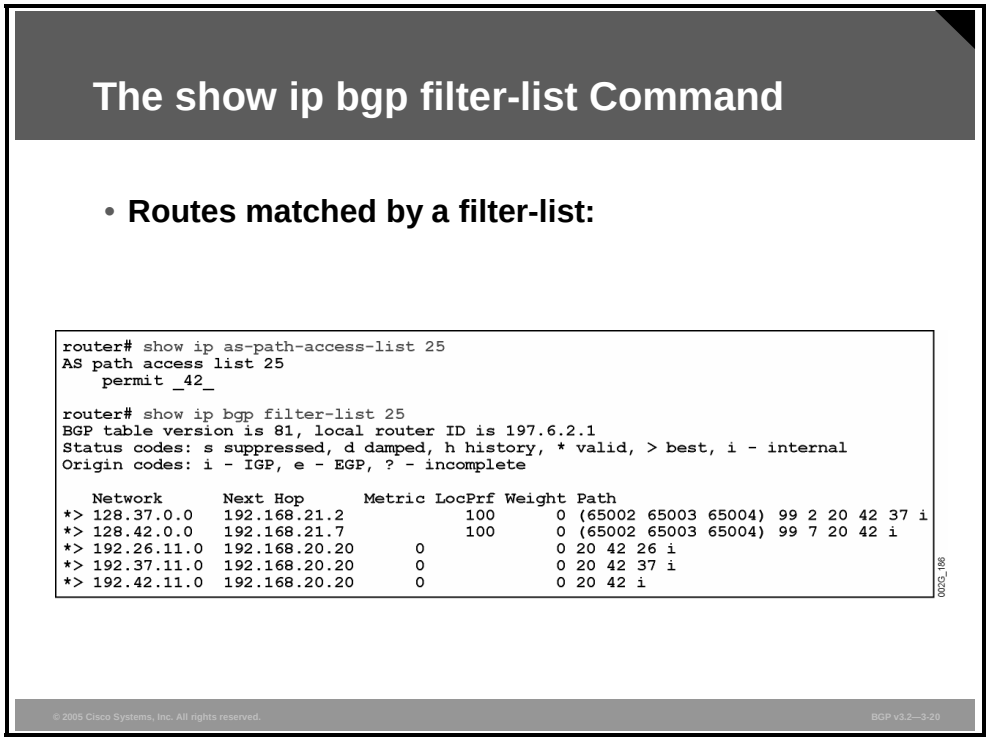
show ip bgp regexp

To display routes that match an AS-path regular expression, use the **show ip bgp regexp** privileged EXEC command.

- **show ip bgp regexp** *regular-expression*

Syntax Description

Parameter	Description
regular-expression	Regular expression to match the BGP AS paths



An AS-path access-list is even more complex because it is a combination of several regular expressions. There is one expression on each access-list line. Use the **show ip bgp filter-list** command to test the entire AS-path access-list. The result is a printout on the screen of all the routes currently in the BGP table that had an AS-path attribute permitted by the access-list.

The AS-path access-list number 25 in the example figure consists of one single line. It permits the routes that have an AS-path attribute containing the AS number 42 somewhere in their AS paths. All other routes are implicitly denied. The **show ip bgp filter-list** command displays all the routes currently in the BGP table that are permitted by AS-path access-list 25. As a result of configuring BGP confederations, the AS path contains some AS numbers enclosed in parentheses. BGP confederations and their usage are explained in the “Scaling Service Provider Networks” module.

show ip bgp filter-list

To display routes that conform to a specified filter-list, use the **show ip bgp filter-list** privileged EXEC command.

- **show ip bgp filter-list** *access-list-number*

Syntax Description

Parameter	Description
access-list-number	Integer from 1 to 199 that indicates the regular expression access-list number

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Several scenarios require BGP route filtering based on AS path, including announcing only local routes to the ISP (AS path needs to be empty), selecting routes based on a specific AS number in the AS path, and accepting routes for a specific AS only from some BGP neighbors. By applying specific selection criteria to the contents of the AS-path attribute, routers can select a subset of routes from the total set of routes that are received.**
- **Cisco IOS software internally translates the AS-path encoding, which is carried with all BGP routes into a character string. This string is then tested against the regular expression.**
- **String matching operates when you are using AS-path regular expressions to match BGP routes.**

Summary (Cont.)

- **You can use AS-path filters to select those routes that will be allowed.**
- **An AS-path filter is created by an AS-path access-list, which is applied to a set of routes from which to select a subset. The** ip as-path access-list **global configuration command defines a BGP AS-path access-list, and the** neighbor filter-list **router configuration command sets up a BGP filter.**
- **There are a number of Cisco IOS commands that are required to monitor the operation of configured AS-path filters, including** show ip as-path-access-list**,** show ip bgp regexp**, and** show ip bgp filter-list**.**

Lesson 3

Filtering with Prefix-Lists

Overview

Where multiple paths between a customer and an Internet service provider (ISP) exist, there is a requirement to filter certain information during Border Gateway Protocol (BGP) updates to influence route selection or to enforce an administrative policy. To meet this requirement, you must use filters. Using prefix-lists is typically easier than using standard IP access-lists and provides performance benefits. It is important to understand the commands to apply filtering of inbound or outbound updates with prefix-lists and where they should be applied.

This lesson discusses the requirement for using prefix-based filters in customer implementations where connections to multiple ISPs must be supported and describes the advantages of prefix-lists over IP access-lists. The commands to apply filtering of inbound or outbound updates with prefix-lists and to configure prefix-list filters are discussed, and also where network administrators should apply them.

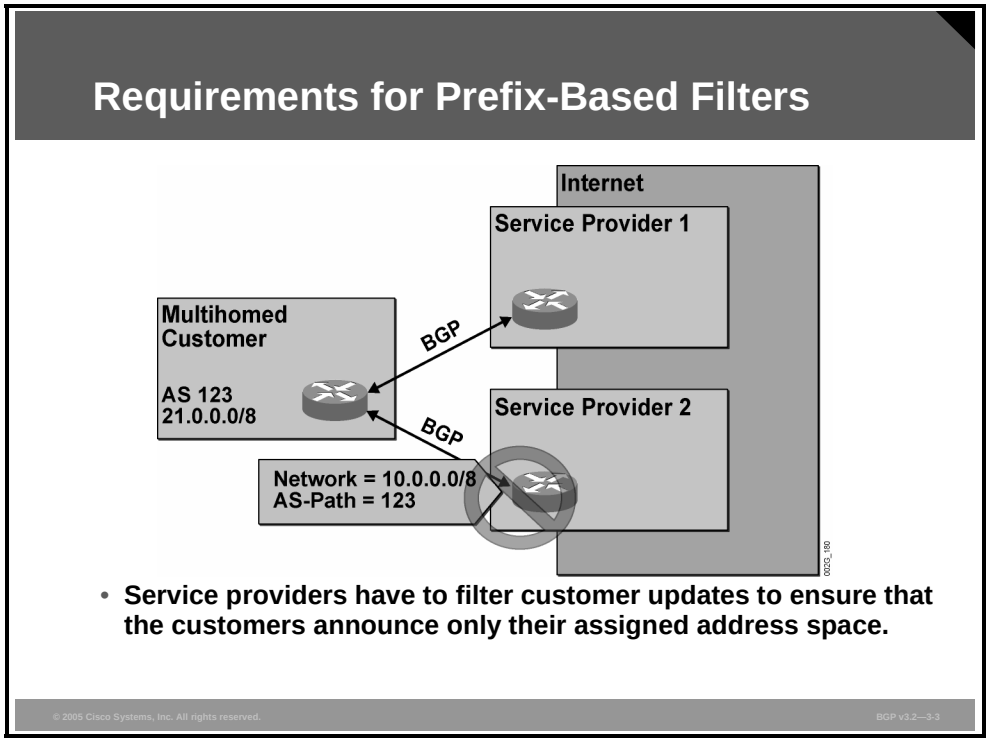
Objectives

Upon completing this lesson, you will be able to successfully configure BGP to influence route selection using prefix-list filters in a customer scenario in which connections to multiple ISPs must be supported. This ability includes being able to meet these objectives:

- Identify the requirement for prefix-based filters in network implementations where multiple connections between a customer and ISPs exist
- List the advantages of prefix-lists versus IP access-lists
- Identify the Cisco IOS command that is required to configure prefix-list filters
- Describe where you can implement prefix-lists in a BGP network
- Identify the Cisco IOS commands that are required to apply filtering of inbound or outbound updates with prefix-lists
- Identify the Cisco IOS commands that are required to modify configured prefix-list filters
- Identify the Cisco IOS commands that are required to monitor the operation of configured prefix-list filters

Requirements for Prefix-Based Filters

This topic identifies the requirement for prefix-based filters in network implementations where multiple connections between a customer and ISPs exist.



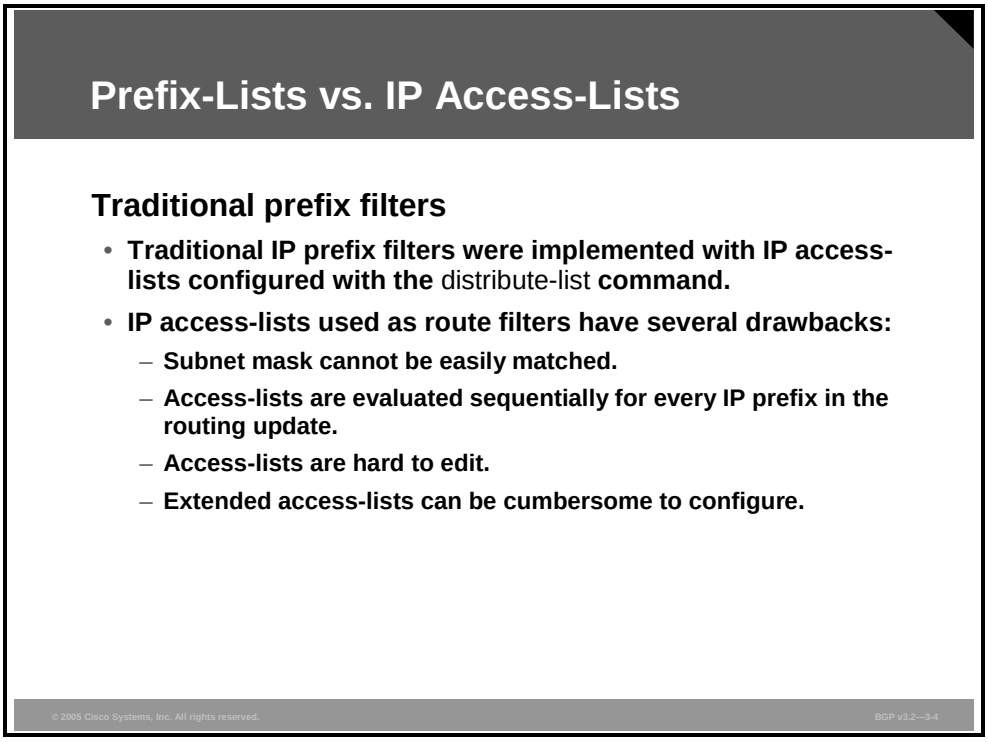
Customers with multihomed networks are responsible for announcing their own networks using BGP. Typically, customers are not as experienced with BGP as service providers, and therefore problems are more likely to occur. A service provider with a multihomed customer must take precautions not to accept, use, or forward any erroneous routing information that is received from the customer.

The customer is assigned a set of IP network numbers that it should announce. If the customer announces any additional networks, something is wrong. The customer may have forgotten not to act as a transit autonomous system (AS) and may have started propagating routes that it has received from the other service provider. Or, the customer may have accidentally started to announce its private address space, which the customer may use for address links, loopback interfaces, or other devices that should never access the Internet.

To avoid problems, the service provider can apply an IP prefix filter on the incoming information from the customer. The service provider will accept only network numbers permitted by an access-list or prefix-list.

Prefix-Lists vs. IP Access-Lists

This topic lists the advantages of prefix-lists versus IP access-lists.

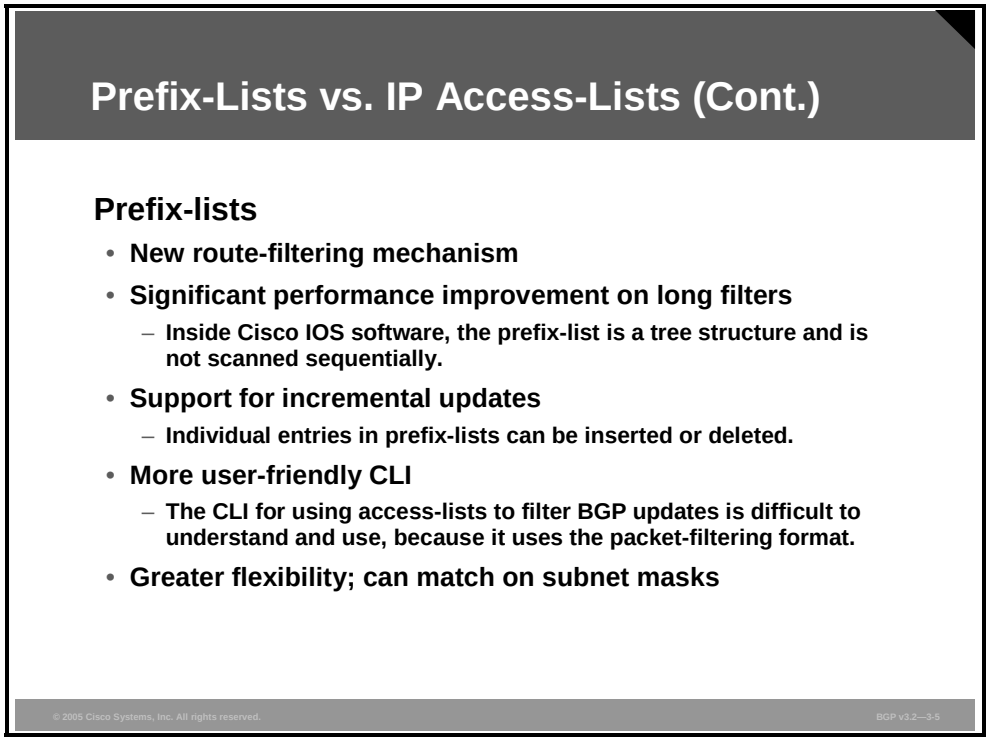


Traditionally, the filtering of IP network numbers has been accomplished using an access-list. The access-list is then bound to either the incoming or outgoing information of a neighbor by the **neighbor distribute-list** command. A BGP update about a network number that is permitted by the access-list will be accepted, and those denied will be dropped.

However, standard access-lists do not support the testing of the subnet masks. If the access-list permits 10.0.0.0/16, it would also permit 10.0.0.0/8.

Extended access-lists can do testing on both an IP network number and a subnet mask, but the syntax is cumbersome.

Finally, access-lists are difficult to edit. The router automatically adds new access-list entries to the end of the list. Because the router evaluates the list sequentially, and the first match results in a “permit” or “deny” statement, the order of the lines in the access-list is of utmost importance. The inability to add a line in the middle of a list has been an administrative burden.



The **ip prefix-list** configuration command has several benefits compared to using the **access-list** command. The intended use of prefix-lists is limited to route filtering, whereas access-lists were initially intended for packet filtering, which was then extended to filter routes.

The prefix-list is internally transformed into a tree structure, with each branching of the tree serving as a test. Cisco IOS software determines a verdict of either “permit” or “deny” much faster this way, compared to sequentially interpreting an access-list.

The configuration command-line interface (CLI) that you use when configuring the **ip prefix-list** command provides the ability to assign a line number to each line of the prefix-list. The router uses this number to sort the entries in the prefix-list. If the lines are initially assigned line numbers, with some spacing in between them, administrators can insert additional lines at a later time. Individual lines can also be removed without removing the entire list.

Routers match network numbers in a routing update against the prefix-list, using as many bits as indicated. For example, a prefix-list can be specified to be 10.0.0.0/16, which will match 10.0.0.0 routes but not 10.1.0.0 routes.

Optionally, the prefix-list can also specify the size of the subnet mask. In addition, the prefix-list can indicate that the subnet mask must be in a specified range.

Prefix-Lists vs. IP Access-Lists (Cont.)

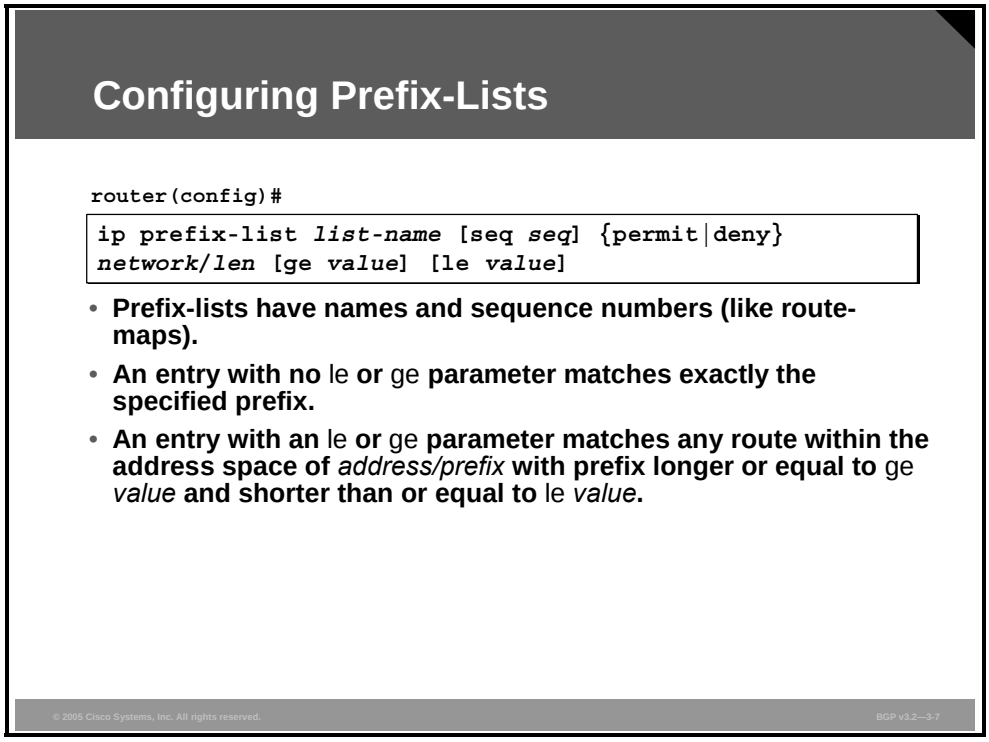
- **Key access-list features are preserved.**
  - **Filtering using “permit” or “deny”**
  - **Order dependency (first match wins)**
  - **Security-focused: no match means “deny”**
- **The matching mechanism has changed.**
  - **Matches routes in a part of address space with subnet mask longer or shorter than a set number**

The prefix-list shares several similarities with the access-list. It can consist of any number of lines, each of which indicates a test and a result. The router can interpret the lines in the specified order, although this process is optimized in the Cisco IOS software. When a router evaluates a route against the prefix-list, the first line that matches results in either a “permit” or “deny.” If none of the lines in the list match, the result is “implicitly deny.”

Testing is done using prefixes. The indicated number of bits in the prefix is compared with the same number of bits in the network number in the update. If the bits match, testing continues with an examination of the number of bits set in the subnet mask. The prefix-list line can indicate a range within which the number must fall to pass the test. If no range is indicated, the subnet mask must match the prefix size.

Configuring Prefix-Lists

This topic identifies the Cisco IOS command that is required to configure prefix-list filters.



ip prefix-list

To create an entry in a prefix-list, use the **ip prefix-list** global configuration command.

- **ip prefix-list** *list-name* [**seq** *seq-value*] {**permit** | **deny**} *network/len* [**ge** *ge-value*] [**le** *le-value*]

To delete the entry, use the **no** form of this command.

- **no ip prefix-list** *list-name* [**seq** *seq-value*] {**permit** | **deny**} *network/len* [**ge** *ge-value*] [**le** *le-value*]

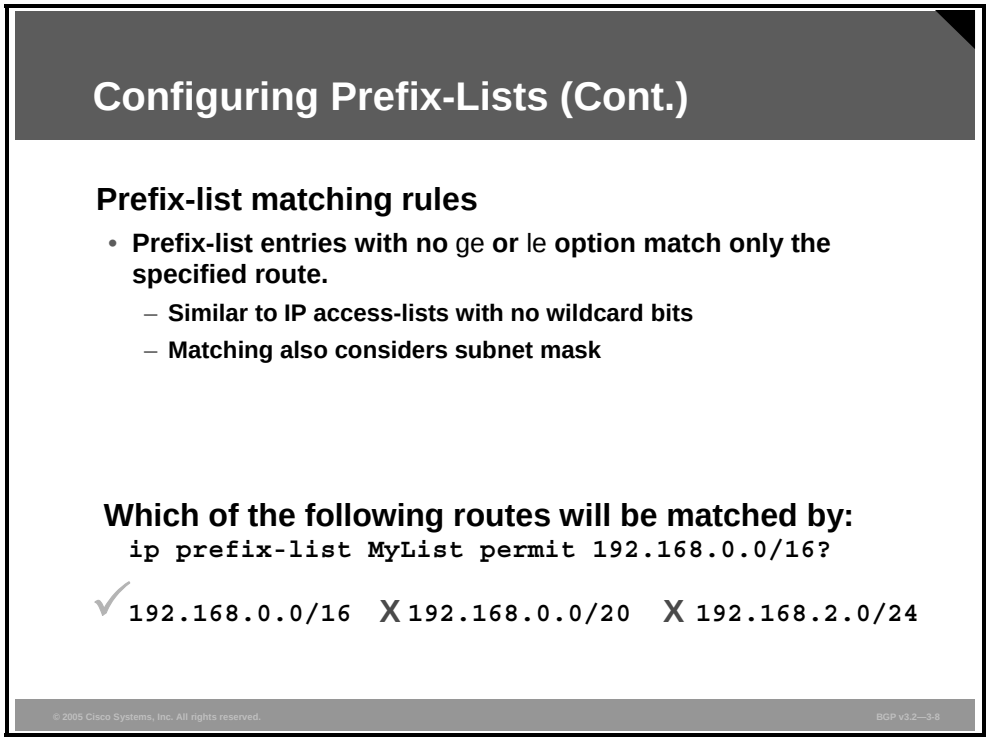
Syntax Description

Parameter	Description
list-name	Name of a prefix
seq	(Optional) Applies the sequence number to the prefix
seq-value	(Optional) Specifies the sequence number for the prefix
permit	Permits access for matching conditions
deny	Denies access to matching conditions
network/len	(Mandatory) The network number and length (in bits) of the subnet mask
ge	(Optional) Applies the *ge-value* to the range specified
ge-value	(Optional) Specifies the lesser value of a range (the "from" portion of the range description)
le	(Optional) Applies the *le-value* to the range specified
le-value	(Optional) Specifies the greater value of a range (the "to" portion of the range description)

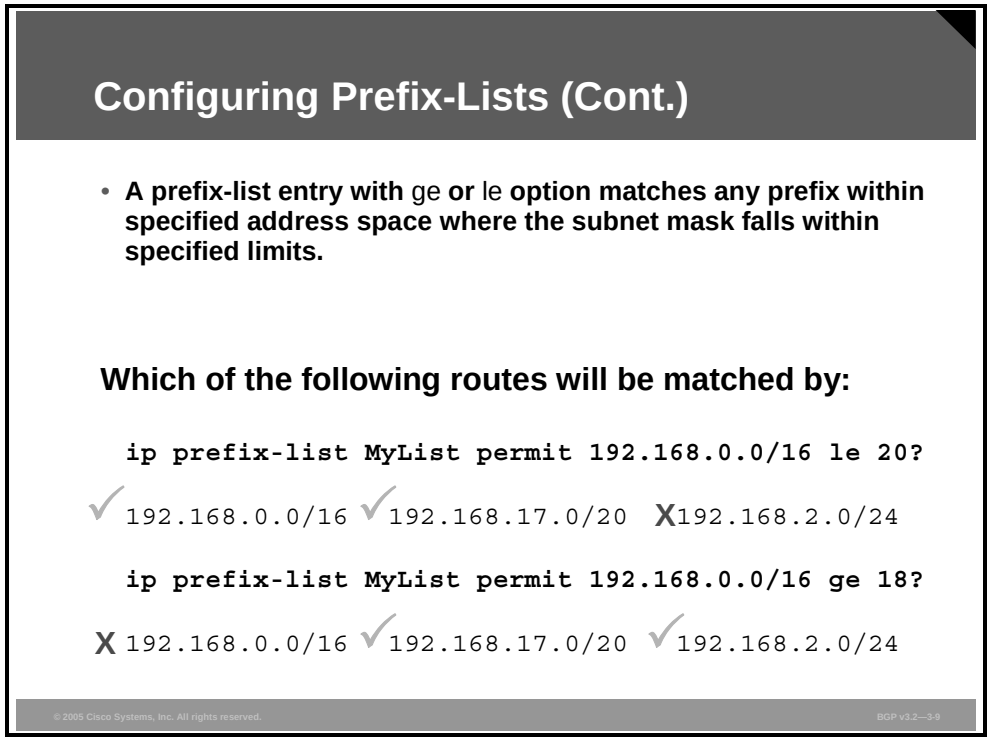
When multiple entries of a prefix-list match a given prefix, the sequence number of a prefix-list entry identifies the entry with the lowest sequence number. In this case, the entry with the smallest sequence number is considered to be the “real” match.

Note You can specify sequence values for prefix-list entries in any increments that you want (the automatically generated numbers are increased in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert additional entries into the prefix-list. If you choose very large increments, you could run out of sequence values.

You can use the parameters **ge** and **le** to specify the range of the prefix length to be matched for prefixes that are more specific than *network/len*. The exact match is assumed when neither **ge** nor **le** is specified. The range is assumed to be from *ge-value* to 32 only if the **ge** attribute is specified. The range is assumed to be from *len* to *le-value* only if the **le** attribute is specified.



Prefix-list entries without the **ge** or **le** option match only the route with the specified IP address and subnet mask. In the example here, the prefix-list entry **permit 192.168.0.0/16** will not match the route 192.168.2.0/24 because of the mismatch in the IP address. It will also not match the route 192.168.0.0/20 because of the mismatch in the subnet mask.



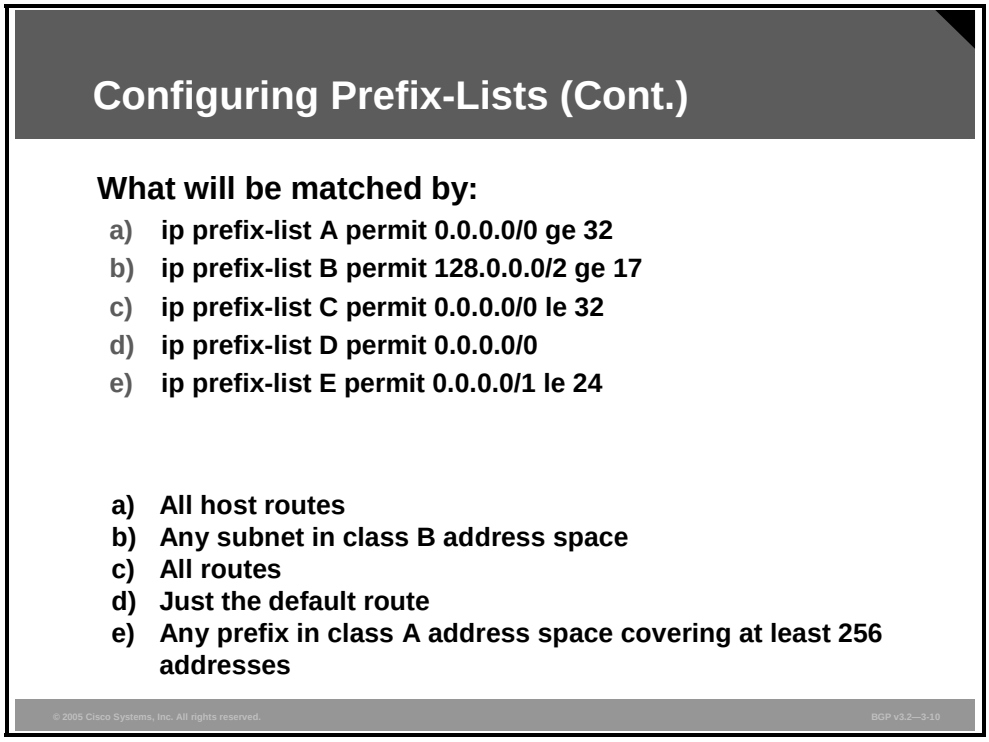
Prefix-list entries with the **ge** or **le** option specified match any prefix within the address space that is specified by the *network/len* parameter, as long as the subnet mask length of the route falls within the range that is specified by the **le** and **ge** parameters.

In the first example in the figure, the route 192.168.2.0/24 is not matched by prefix-list entry **permit 192.168.0.0/16** even though the IP address falls within the specified address range, because the subnet mask is too long.

In the second example, the route 192.168.0.0/16 is not matched by prefix-list entry **permit 192.168.0.0/18** because the subnet mask is too short.

Example: Configuring Prefix-Lists

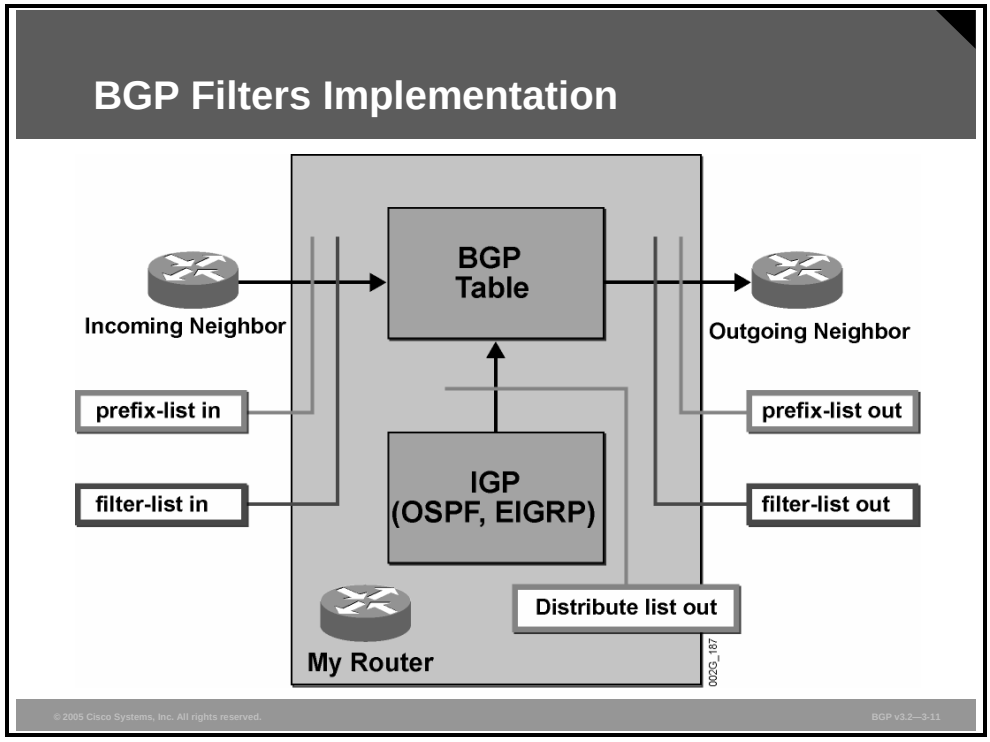
The figure contains some commonly used prefix-list examples.



In the figure, all host routes will be matched by prefix-list A **permit 0.0.0./0 ge 32**, while any subnet in a class B address space will be matched by prefix-list B **permit 128.0.0.0/2 ge 17**. Prefix-list C **permit 0.0.0.0/0 le 32** will match all routes, but only the default route will be matched by prefix-list **D permit 0.0.0.0/0**. Finally, any prefix in a class A address space that covers at least 256 addresses will be matched to prefix-list E **permit 0.0.0.0/1 le 24**.

BGP Filters Implementation

This topic describes where you can implement prefix-lists in a BGP network.

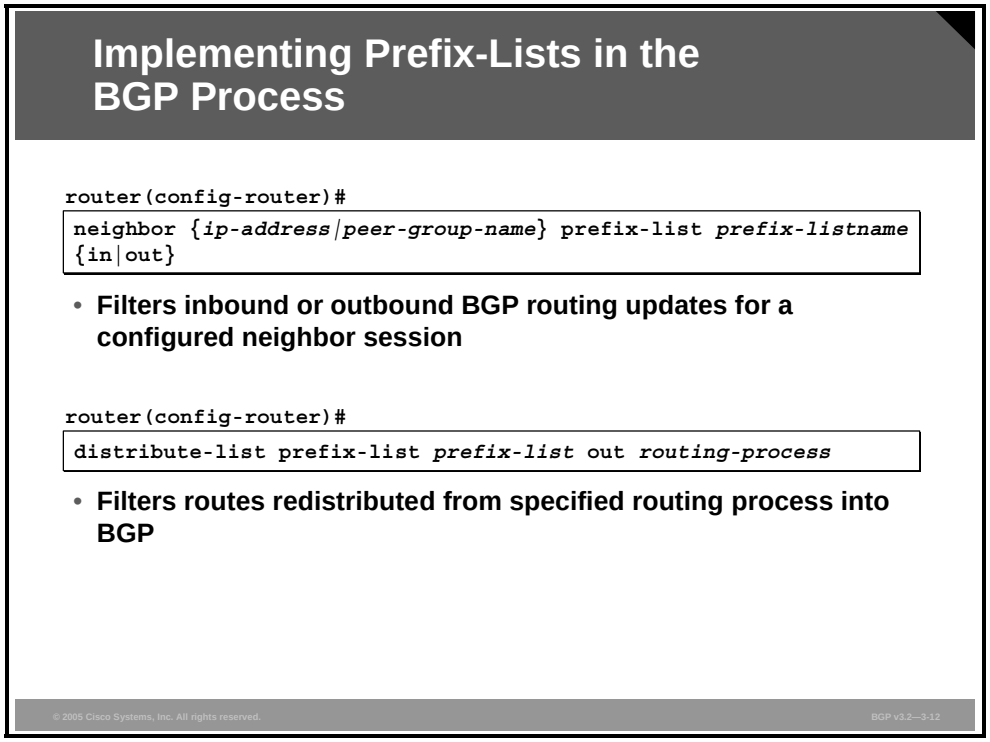


You can optionally apply filter-lists and prefix-lists on either incoming or outgoing neighbors in any combination. Both the incoming prefix-list and the incoming filter-list must permit the routes that are received from a neighbor before they are accepted into the BGP table. Outgoing routes must pass both the outgoing filter-list and the outgoing prefix-list before being transmitted to the neighbor.

When a router is configured to redistribute routing information from an Interior Gateway Protocol (IGP) into BGP, the routes must successfully pass any prefix-list or access-list that is applied to the redistribution before a route is injected into the BGP table.

Implementing Prefix-Lists in the BGP Process

This topic identifies the Cisco IOS commands that are required to apply prefix-lists for filtering inbound or outbound updates.



You can use prefix-lists to filter incoming or outgoing BGP updates to neighbors. You can also use prefix-lists to filter routes that are being redistributed into the BGP process from other routing protocols.

neighbor prefix-list

To distribute BGP neighbor information as specified in a prefix-list, use the **neighbor prefix-list** router configuration command.

- **neighbor** {*ip-address* | *peer-group-name*} **prefix-list** *prefix-listname* {**in** | **out**}

To remove an entry, use the **no** form of this command.

- **no neighbor** {*ip-address* | *peer-group-name*} **prefix-list** *prefix-listname* {**in** | **out**}

Syntax Description

Parameter	Description
ip-address	IP address of neighbor
peer-group-name	Name of a BGP peer group
prefix-listname	Name of a prefix-list
in	Access-list is applied to incoming advertisements to that neighbor
out	Access-list is applied to outgoing advertisements from that neighbor

Note A BGP peer group is a group of BGP neighbors with the same update policies. Route-maps, distribute-lists, filter-lists, and so on usually set update policies. Instead of defining the same policies for each separate neighbor, a peer group name is configured on the router, and these policies are assigned to the peer group. BGP peer groups are discussed in a later module.

distribute-list out

To suppress networks from being advertised in updates, use the **distribute-list out** router configuration command.

- **distribute-list** {*access-list-number* | *name* | **prefix-list** *prefix-listname*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

To disable this function, use the **no** form of this command.

- **no distribute-list** {*access-list-number* | *name* | **prefix-list** *prefix-listname*} **out** [*interface-name* | *routing-process* | *autonomous-system-number*]

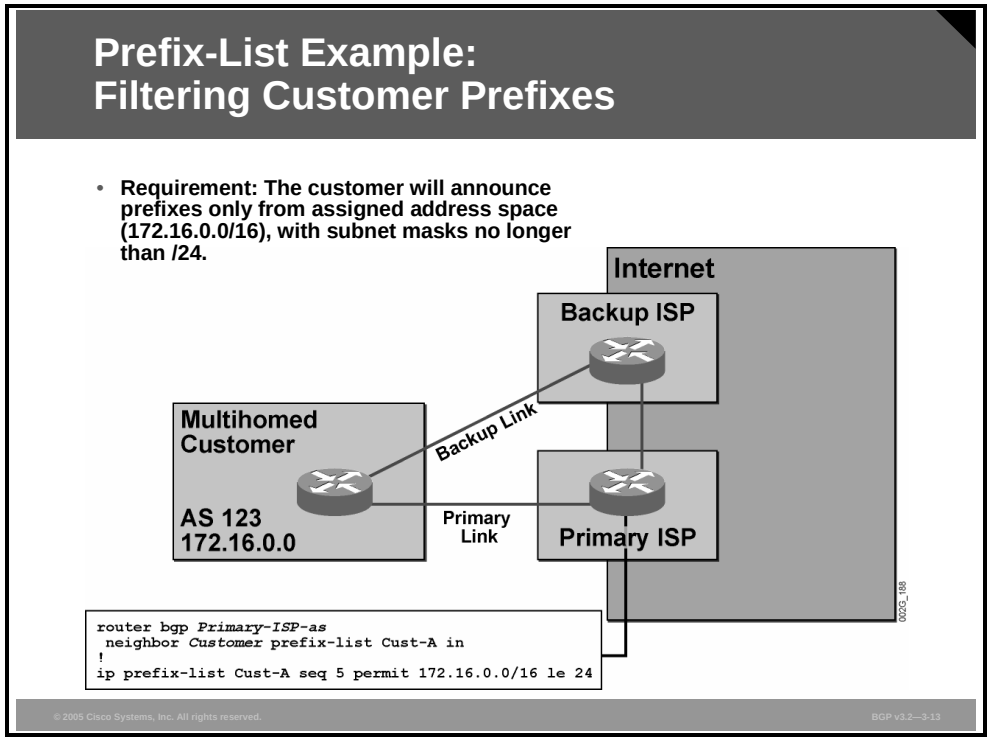
Syntax Description

Parameter	Description
access-list-number | *name*	Standard IP access-list number or name. The list defines which networks are to be received and which are to be suppressed in routing updates.
prefix-listname	Name of a prefix-list. The list defines which networks are to be received and which are to be suppressed in routing updates, based upon matching of the network prefix to the prefixes in the list.
out	Applies the access-list to outgoing routing updates.
interface-name	(Optional) Name of a particular interface.
routing-process	(Optional) Name of a particular routing process, or the keyword **static** or **connected.**
autonomous-system-number	(Optional) AS number.

Note Although you can use the **neighbor prefix-list** router configuration command as an alternative to the **neighbor distribute-list** command, do not use both the **neighbor prefix-list** and **neighbor distribute-list** command filtering for the same neighbor in any given direction. These two commands are mutually exclusive, and only one command (**neighbor prefix-list** or **neighbor distribute-list**) can be applied for each inbound or outbound direction.

Example: Filtering Customer Prefixes

The figure illustrates filtering customer prefixes.

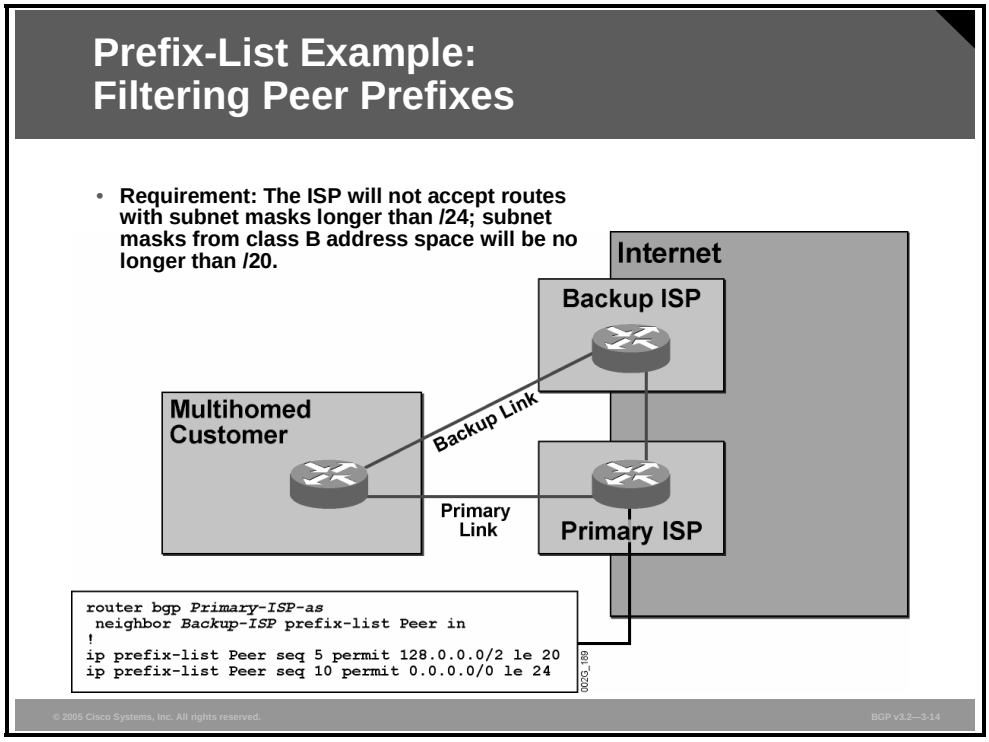


In this example, a multihomed customer has been assigned the address space 172.16.0.0/16. The customer may subnet this address space but may not announce subnets smaller than a subnet mask of 255.255.255.0. Larger subnets are accepted. If the customer has subnetted the network into smaller subnets, it must summarize the routing information about those subnets into at least /24 prefixes before announcing them.

The primary ISP implements a prefix-list named Cust-A to perform the filtering of incoming information from the multihomed customer. The prefix-list permits all routes that are received from the customer that have 172.16 in the first 16 bits and have a subnet mask of 24 bits or less. Any other routes from the customer are denied and silently ignored.

Example: Filtering Peer Prefixes

The figure illustrates filtering peer prefixes.



In this example, the primary ISP will not accept any route from the customer that indicates a subnet smaller than a 255.255.255.0 subnet mask. The class B network, however, must not be subnetted into subnets smaller than a 255.255.240.0 subnet mask.

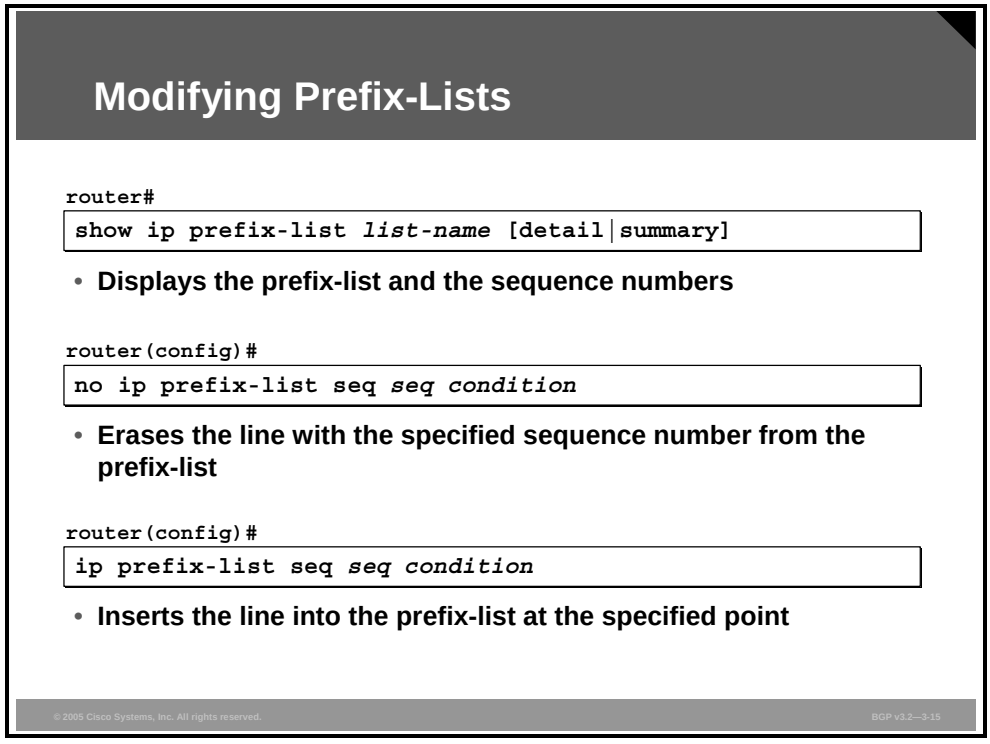
The primary ISP implements this route by using a prefix-list named Peer. The first line in the prefix-list checks whether it is a class B network. Remember that a class B address always has the binary sequence 10 as the first 2 bits in the first byte. The second line matches any prefix.

When the primary ISP receives a route from the customer, it compares the route with both lines. If the route is a class B network, both lines match. Testing continues with checking the subnet mask. An upper bound is explicitly indicated, giving a maximum prefix length of 20 bits.

If the received route is not a class B network, only the second line matches. In this case, the subnet mask length must be greater than or equal to 0 and less than or equal to 24, providing a route less explicit than a /24 prefix.

Modifying Prefix-Lists

This topic identifies the Cisco IOS commands that are required to modify configured prefix-list filters.



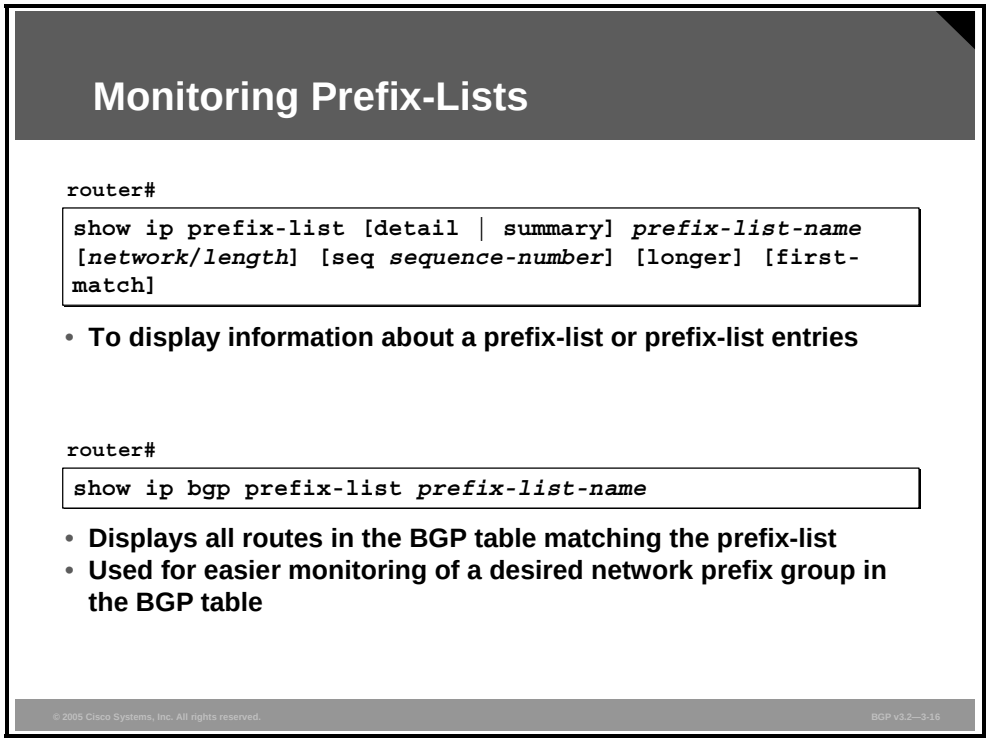
Lines in a prefix-list are assigned sequence numbers. These assignments significantly improve the manageability of the list. These sequence number assignments provide the opportunity to remove a specific line, and, if spacing between the sequence numbers allows, they provide the ability to insert a line between two existing lines.

To display a currently configured prefix-list and its sequence numbers, use the **show ip prefix-list** command with the **detail** keyword.

You can specify sequence values for prefix-list entries in any increments that you want (the automatically generated numbers are increased in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert additional entries into the prefix-list. If you choose very large increments, you could run out of sequence values.

Monitoring Prefix-Lists

This topic lists the Cisco IOS commands that are required to monitor the operation of configured prefix-list filters.



show ip prefix-list

To display information about a prefix-list or prefix-list entries, use the **show ip prefix-list** EXEC command.

- **show ip prefix-list** [**detail** | **summary**] *name* [*network/len*] [**seq** *seq-num*] [**longer**] [**first-match**]

Syntax Description

Parameter	Description
detail | **summary**	(Optional) Displays detailed or summarized information about all prefix-lists
name	(Optional) The name of a specific prefix-list
network/len	(Optional) The network number and length (in bits) of the network mask
seq	(Optional) Applies the sequence number to the prefix-list entry
seq-num	The sequence number of the prefix
longer	Displays all entries of a prefix that are more specific than the given *network/len*
first-match	Displays the entry of a prefix that matches the given *network/len*

The **show ip bgp prefix-list** command displays selected routes from a BGP routing table based on the contents of a prefix-list. Use this command for selective filtering of BGP table output on Cisco IOS devices on the basis of network prefix groups.

To perform prefix-list-based BGP table filtering, follow these steps:

Step 1 Configure a prefix-list that permits ranges of networks meant to be displayed in the BGP table output.

Step 2 Include a reference to a configured prefix-list in the **show ip bgp prefix-list** command.

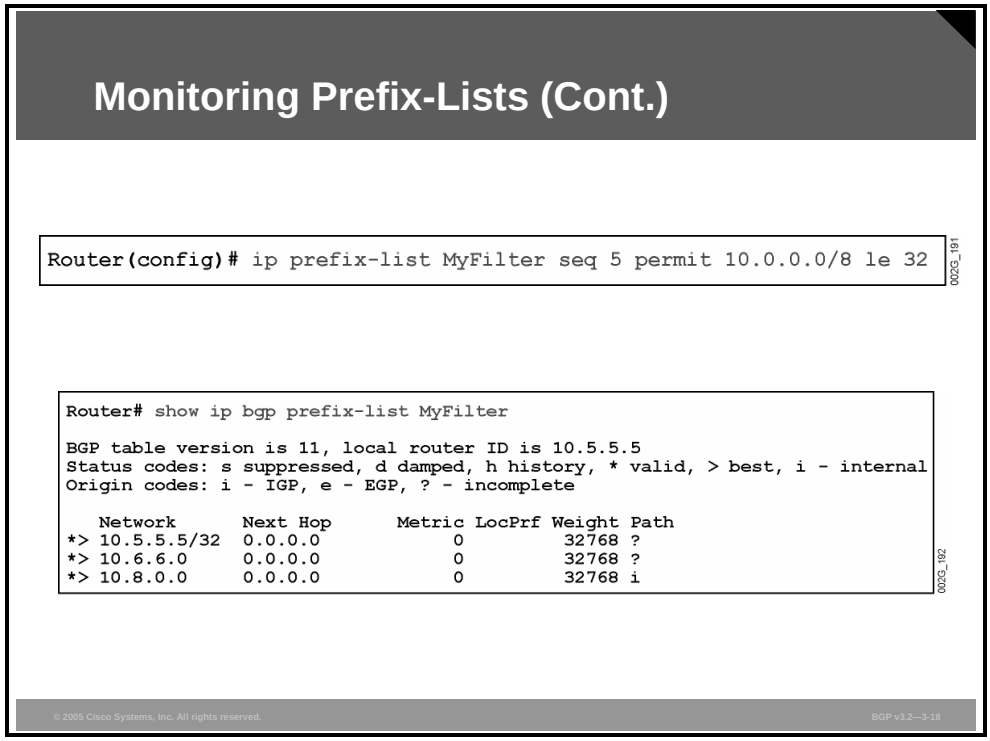
Note The support for prefix-list BGP table filtering was added in Cisco IOS Software Release 12.2(11)T and 12.0(14)ST.

Monitoring Prefix-Lists (Cont.)

```
router# show ip prefix-list detail
Prefix-list with the last deletion/insertion: InFilter
ip prefix-list InFilter:
   count: 4, range entries: 3, sequences: 5 - 20, refcount: 2
   seq 5 deny 128.0.0.0/2 le 15 (hit count: 0, refcount: 2)
   seq 10 deny 192.0.0.0/3 ge 25 (hit count: 0, refcount: 1)
   seq 15 deny 193.0.0.0/8 ge 21 (hit count: 0, refcount: 1)
   seq 20 permit 0.0.0.0/0 (hit count: 0, refcount: 1)
```

002G_190

In this example, the **show ip prefix-list** command has been issued with the **detail** keyword. The output of the command displays detailed information about configured prefix-lists, including sequence numbers, the prefix-list entries, and the number of times that each entry has been matched by a corresponding prefix.



This example shows a simple prefix-list-based filtering of the BGP table. The prefix-list filter permits all networks with the first octet equal to 10 and any length of a subnet mask (**le 32**). In the **show ip bgp prefix-list** command output, only the networks permitted by the prefix-list filter are displayed.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Customers with multihomed networks are responsible for announcing their own networks using BGP, and service providers with multihomed customers must take precautions not to accept, use, or forward any erroneous routing information that is received from their customers.**
- **Prefix-lists have a number of advantages over access-lists, including faster "permit" or "deny" determinations and easier CLI editing.**
- **Prefix-lists are configured using the** ip prefix-list **global configuration command.**
- **Filter-lists and prefix-lists can be optionally applied on either incoming or outgoing neighbors in any combination.**

Summary (Cont.)

- **Prefix-lists can filter incoming or outgoing BGP updates to neighbors and filter routes that are being redistributed into the BGP process from other routing protocols. Use the** neighbor prefix-list router configuration command **to distribute BGP neighbor information as specified in a prefix-list.**
- **Certain Cisco IOS commands (such as the** show ip prefix-list **command) are used to modify configured prefix-list filters.**
- **To display or monitor statistics about a prefix-list or prefix-list entries, you can use the** show ip prefix-list **EXEC command.**

Lesson 4

Using Outbound Route Filtering

Overview

An outbound route filter (ORF) is an additional mechanism that is used to minimize the number of updates that are requested from a neighbor, which reduces link bandwidth consumption and CPU use when a router requests a route refresh. An ORF also allows filtering of information that external networks should not receive (such as RFC 1918 information). Understanding how to monitor outbound route filtering capabilities is also important because a Border Gateway Protocol (BGP) neighbor that supports specific ORF capabilities will report those capabilities to a monitoring neighbor and can then send a filter of the supported type to the neighbor.

This lesson discusses the function of outbound route filtering in a BGP network. The format and function of ORF messages are discussed, as well as the commands that enable ORF negotiations and the activation of an ORF prefix-list. The commands that are used to trigger a route refresh are also detailed. Finally, there is a discussion on how to monitor the operations of a configured ORF in a BGP network.

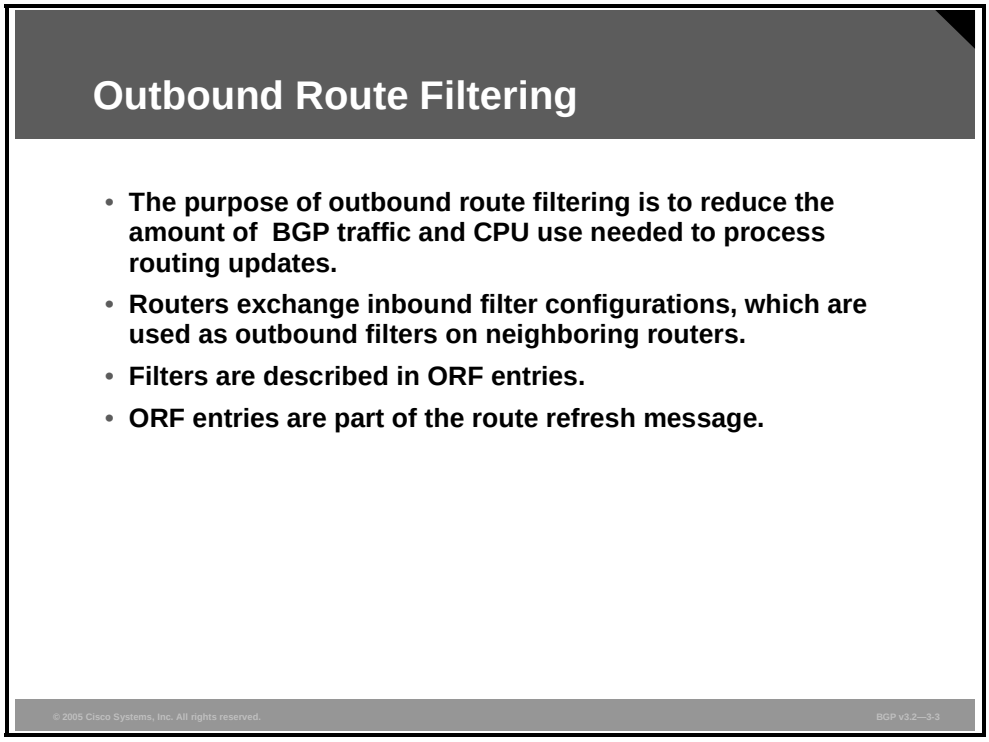
Objectives

Upon completing this lesson, you will be able to use outbound route filtering to minimize the impact of BGP routing updates on router resources in an operational BGP network. This ability includes being able to meet these objectives:

- Describe the function of outbound route filtering in a BGP network
- Describe the function of prefix-based outbound route filtering
- Describe the format and function of an ORF message
- Identify the Cisco IOS command that is required to enable ORF negotiations and activate an ORF prefix-list
- Identify the Cisco IOS command that is used to trigger a route refresh
- Identify the Cisco IOS command that is required to monitor the operation of a configured ORF

Outbound Route Filtering

This topic describes the function of outbound route filtering in a BGP network.



Outbound route filtering is a prefix-based BGP feature that is enabled through the advertisement of ORF capabilities to peer routers integrated in Cisco IOS Software Release 12.2(4)T. The advertisement of the ORF capability indicates that a BGP-speaking router will accept a prefix-list from a neighbor and apply the prefix-list to locally configured ORFs (if any exist). When this capability is enabled, the BGP speaker can install an inbound prefix-list filter to the remote peer as an outbound filter, which reduces unwanted routing updates.

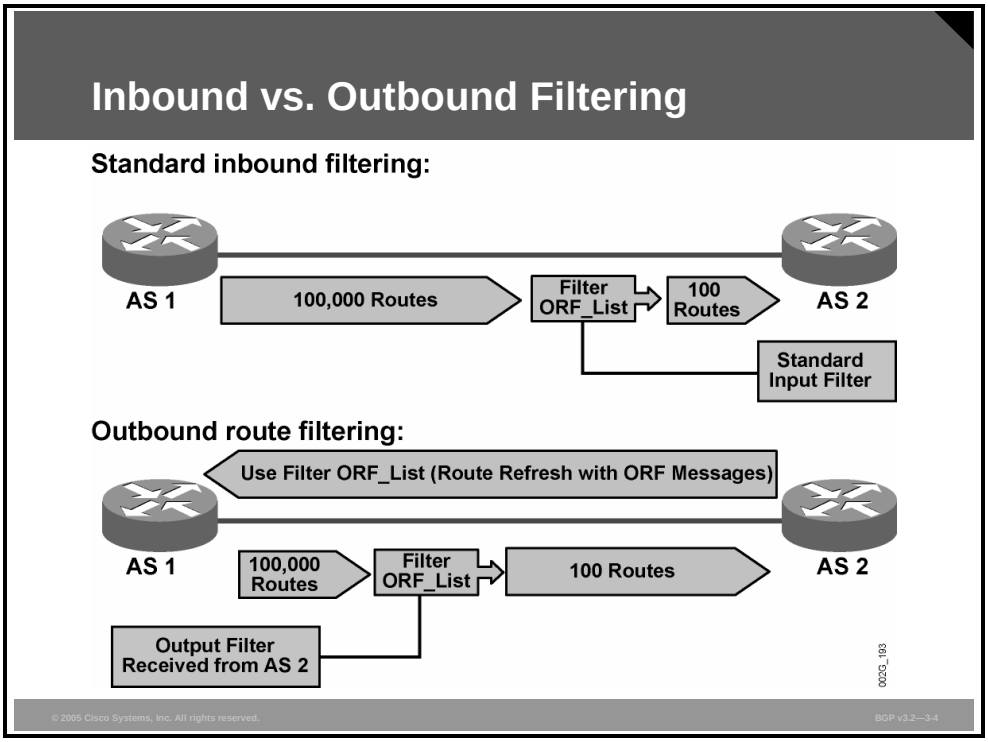
The standard route refresh message contains the Address Family Information (AFI) for which the refresh is needed. Outbound route filtering is an additional mechanism that is used to minimize the number of updates that are requested from a neighbor.

This mechanism reduces link bandwidth consumption and CPU use when a router requests a route refresh. Filters that should be used by routers with the route refresh are described in ORF entries that are part of the route refresh message.

You can configure the ORF feature with send, receive, or send and receive capabilities. The local peer advertises the ORF capability in send mode. The remote peer receives the ORF capability in receive mode and applies the filter as outbound policy. The local and remote peers exchange updates to maintain the ORF for each router. Peer routers exchange updates depending on the ORF prefix-list capability that is advertised. The remote peer starts sending updates to the local peer after it receives a route refresh request or an ORF prefix-list with an immediate status.

Example: Inbound vs. Outbound Filtering

The figure illustrates the comparison between standard inbound filtering and outbound route filtering.



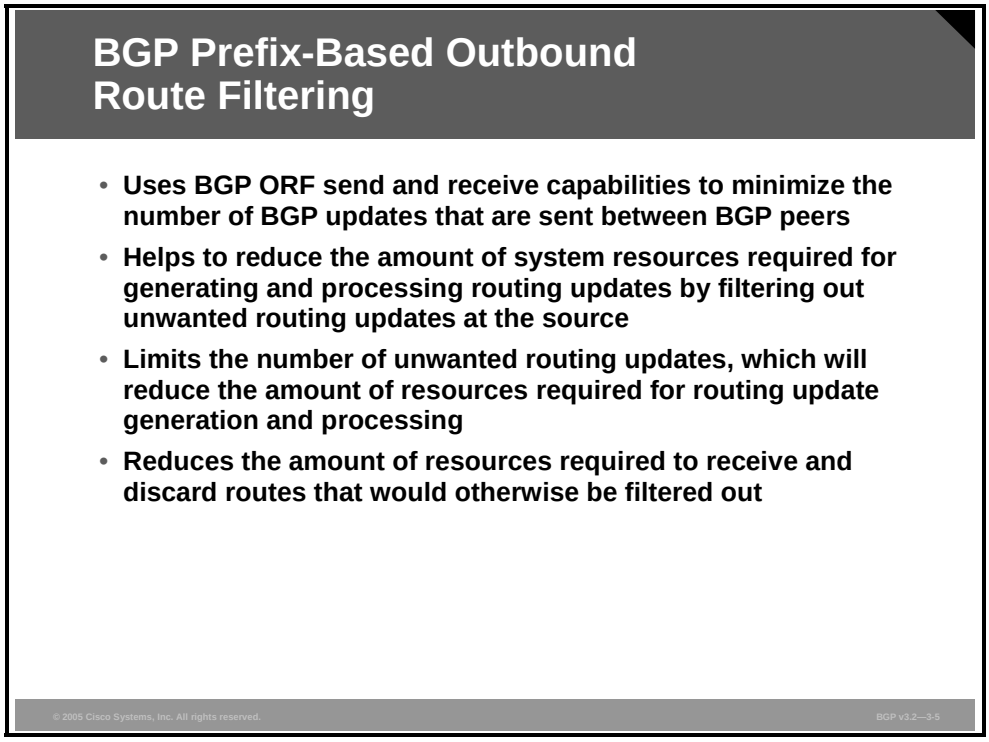
Outbound route filtering can limit the number of unwanted routing updates, which will reduce the amount of resources that are required for routing update generation and processing. This feature also reduces the amount of resources that are required to receive and discard routes that would otherwise be filtered out by the receiving router if the ORF feature were not available.

The example shows two scenarios:

- The first example shows that 100,000 routes are sent to a neighbor, and the input filter permits only 100 of these routes.
- The second example shows how a route refresh with a filter is sent to the neighbor. The neighbor then uses the filter before sending the updates. This way, only 100 updates are sent.

BGP Prefix-Based Outbound Route Filtering

This topic describes the function of BGP prefix-based outbound route filtering.



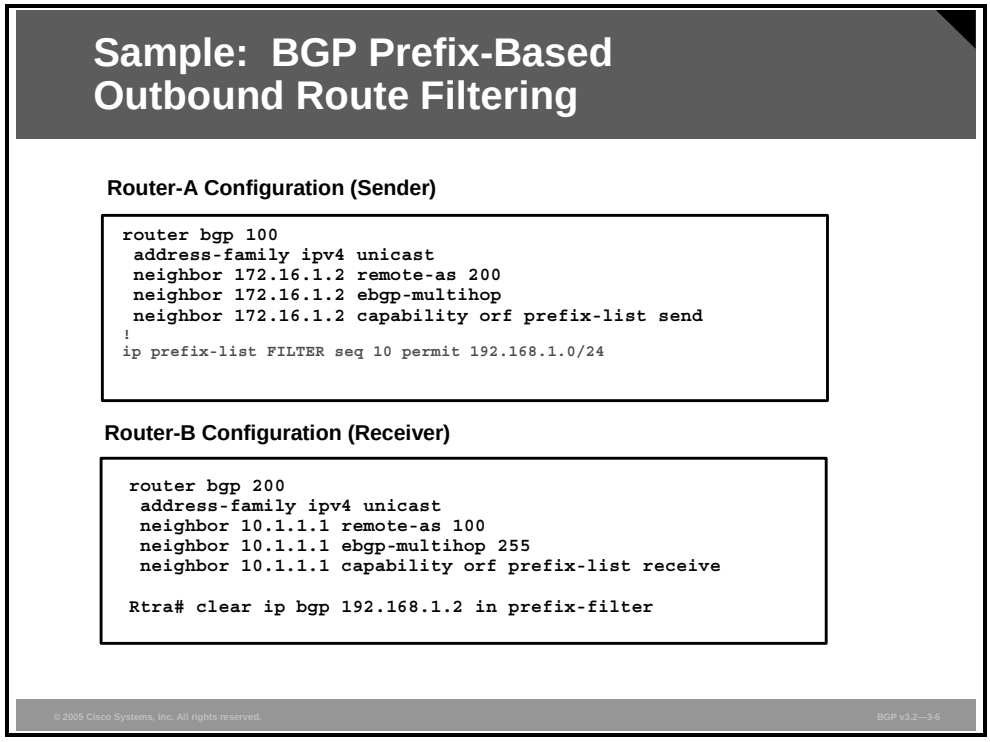
The BGP prefix-based outbound route filtering feature uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers. Configuring this feature can help reduce the amount of system resources required for generating and processing routing updates by filtering out unwanted routing updates at the source.

The BGP prefix-based outbound route filtering feature is enabled through the advertisement of ORF capabilities to peer routers. The advertisement of the ORF capability indicates that a BGP speaker will accept a prefix-list from a neighbor and apply the prefix-list to locally configured ORFs (if any exist). When this capability is enabled, the BGP speaker can install the inbound prefix-list filter to the remote peer as an outbound filter, which reduces unwanted routing updates.

The BGP prefix-based outbound route filtering feature can be configured with send, receive, or send and receive ORF capabilities. The local peer advertises the ORF capability in send mode. The remote peer receives the ORF capability in receive mode and applies the filter as an outbound policy. The local and remote peers exchange updates to maintain the ORF on each router. Updates are exchanged between peer routers by address family depending on the ORF prefix-list capability that is advertised. The remote peer starts sending updates to the local peer after a route refresh has been configured with the **clear ip bgp** command or after an ORF prefix-list with immediate status is processed. The BGP speaker will continue to apply the inbound prefix-list to received updates after the speaker pushes the inbound prefix-list to the remote peer.

Example: BGP Prefix-Based Outbound Route Filtering

The figure provides an example of the function of BGP prefix-based outbound route filtering.

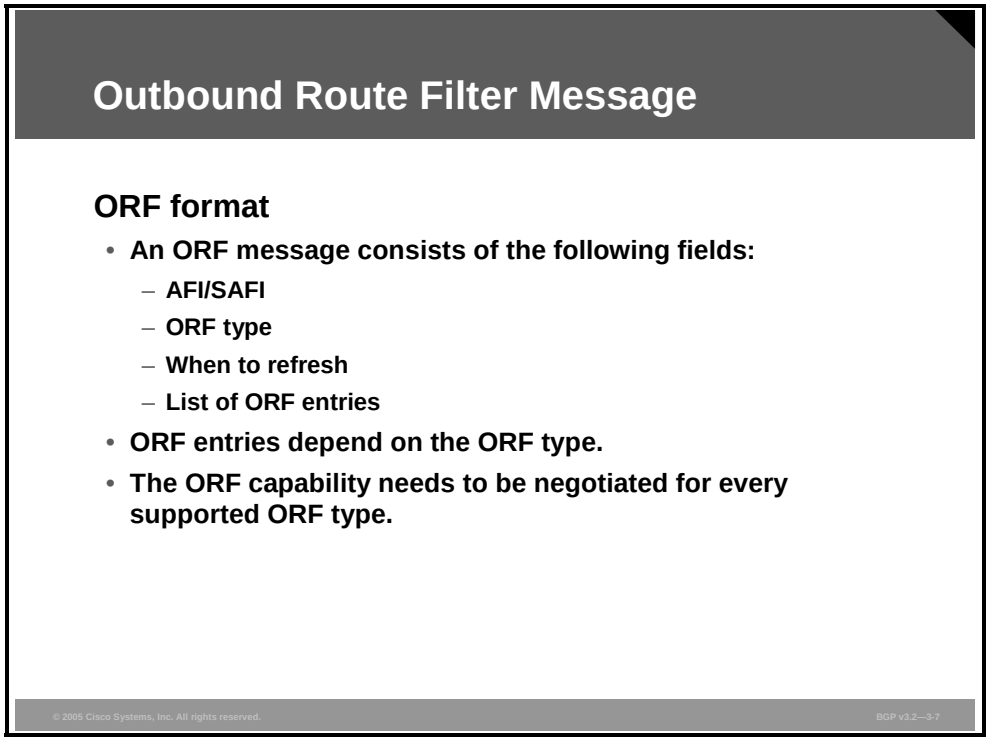


In the example, an ORF has been configured on Router-A (10.1.1.1) to advertise the filter to Router-B (172.16.1.2). An IP prefix-list named FILTER is created to specify the 192.168.1.0/24 subnet for outbound route filtering. The ORF send capability is configured on Router-A so that Router-A can advertise the ORF to Router-B.

Router-B is configured to advertise the ORF receive capability to Router-A. Router-B will install the ORF, defined in the FILTER prefix-list, after the ORF capabilities have been exchanged. An inbound soft reset is initiated on Router-B at the end of this configuration to activate the ORF.

Outbound Route Filter Message

This topic describes the format and function of the ORF message.



An ORF message contains the following information:

- AFI and Subsequent Address Family Information (SAFI) for which the filter should be used
- ORF type, which identifies the type of filter
- When to refresh (immediate or deferred refresh)
- List of ORF entries where the actual filter is defined

You can use the AFI/SAFI component of the ORF message to provide a coarse level of granular control by limiting the ORF to only the routes whose Network Layer Reachability Information (NLRI) matches the configured AFI/SAFI component.

The ORF capability has to be negotiated by the router for each ORF type that is supported in the ORF message.

Outbound Route Filter Message (Cont.)

ORF types:

- **NLRI (ORF type = 1)**
  - **Filters based on the prefix**
- **Communities (ORF type = 2)**
  - **Filters based on standard BGP community attributes**
- **Extended communities (ORF type = 3)**
  - **Filters based on extended BGP community attributes**
- **Prefix-list (ORF type = 128)**
  - **Filters based on Cisco implementation of prefix filtering**

The value contained in the ORF type determines the content that is contained in the ORF message.

Currently, ORF type 0 is reserved, ORF types 1 to 127 are assigned by the Internet Assigned Numbers Authority (IANA), and ORF types 128 to 255 are vendor-specific (and not assigned by the IANA). Commonly used ORF types are as follows:

- ORF type 1 is used to filter based on the NLRI.
- ORF type 2 is used to filter based on standard BGP community attributes.
- ORF type 3 is used to filter based on extended BGP community attributes.
- ORF type 128 is used to filter based on the Cisco proprietary implementation of prefix filtering (prefix-lists).

Outbound Route Filter Message (Cont.)

AFI/SAFI is IPv4 unicast.

ORF type is NLRI:

- **Action:** **ADD, DELETE, or DELETE ALL**
- **Match:** **PERMIT or DENY**
- **Scope:** **EXACT or REFINE**
- **NLRI:** **Prefix**
- **When:** **IMMEDIATE or DEFER**

The content of the ORF value is determined by the ORF-type setting. An ORF type of NLRI-based filtering (type 1) uses the following actions:

- **ADD:** Adds a line to a prefix-list filter on the remote peer
- **DELETE:** Removes a line from a filter that was previously installed on a remote peer
- **DELETE ALL:** Removes all previously installed filters on the remote peer

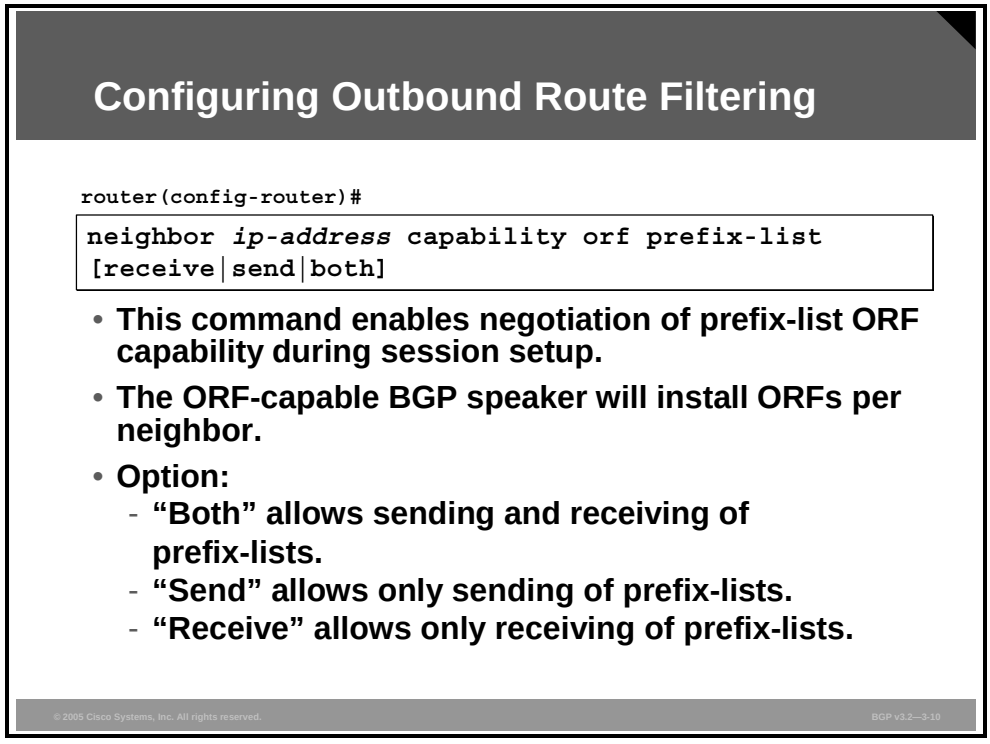
For each filter entry, there is a match component that specifies either PERMIT or DENY. A PERMIT asks the peer to send updates with routes that match the set of entries as specified in the ORF. DENY specifies that the remote peer should not send updates for the entries matching those specified in the ORF.

For prefixes specified with a match component of PERMIT, the remote peer is asked to pass a prefix with a scope of EXACT (an exact match) or REFINE (its subnets).

Also contained within the ORF message is the when-to-refresh field. A router can set this field to IMMEDIATE (asking the remote peer to refresh as soon as it has finished processing the ORF message) or DEFER (asking the remote peer to wait until it receives a subsequent route refresh message with the same AFI/SAFI).

Configuring Outbound Route Filtering

This topic identifies the Cisco IOS command that is required to enable ORF negotiations and activate an ORF prefix-list.



Cisco routers support the uploading of their prefix-lists to a neighbor. You need to use the **neighbor** *ip-address* **capability orf prefix-list receive** command to advertise this capability, and you need to use **neighbor** *ip-address* **capability orf prefix-list send** command to upload the inbound prefix filter to the neighbor. The uploaded filter is then used on the neighboring router after a statically configured outbound prefix-list (if it exists) is applied.

The **neighbor** *ip-address* **capability orf prefix-list** command enables the negotiation of the prefix-list ORF capability during BGP session setup. The prefix-list-based ORF (ORF type = 128) is the only ORF type that Cisco IOS software supports.

neighbor orf prefix-list

To advertise ORF capabilities to a peer router, use the **neighbor orf prefix-list** command in address family or router configuration mode.

- **neighbor** {*ip-address*} [**capability**] **orf prefix-list** [**receive** | **send** | **both**]

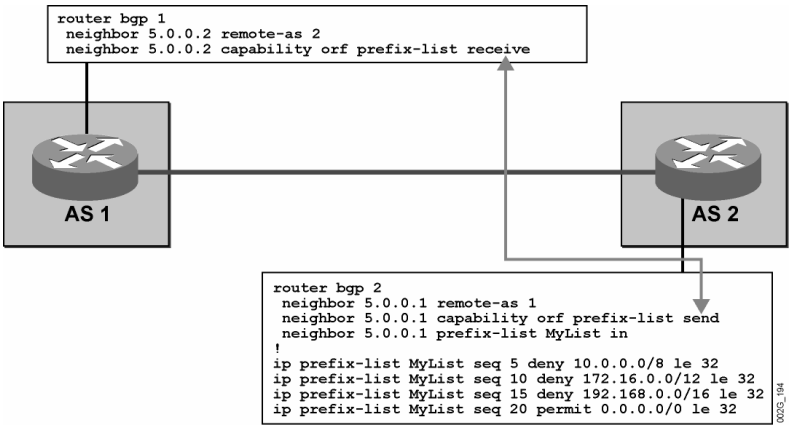
To disable ORF capabilities, use the **no** form of this command.

- **no neighbor** {*ip-address*} [**capability**] **orf prefix-list** [**receive** | **send** | **both**]

Syntax Description

Parameter	Description
ip-address	The IP address of the neighbor router
capability	(Optional) Informs the specified neighbor that this router has ORF capabilities
receive	(Optional) Enables the ORF prefix-list capability in receive mode
send	(Optional) Enables the ORF prefix-list capability in send mode
both	(Optional) Enables the ORF prefix-list capability in both receive and send modes

Configuring Outbound Route Filtering (Cont.)



- **The command** capability orf prefix-list send **on one router requires** capability orf prefix-list receive **on a neighboring router.**

The example shows the configuration of two routers where one router has uploaded an input prefix-list to the neighbor to be used as an output filter.

The following configuration steps are necessary to enable outbound route filtering:

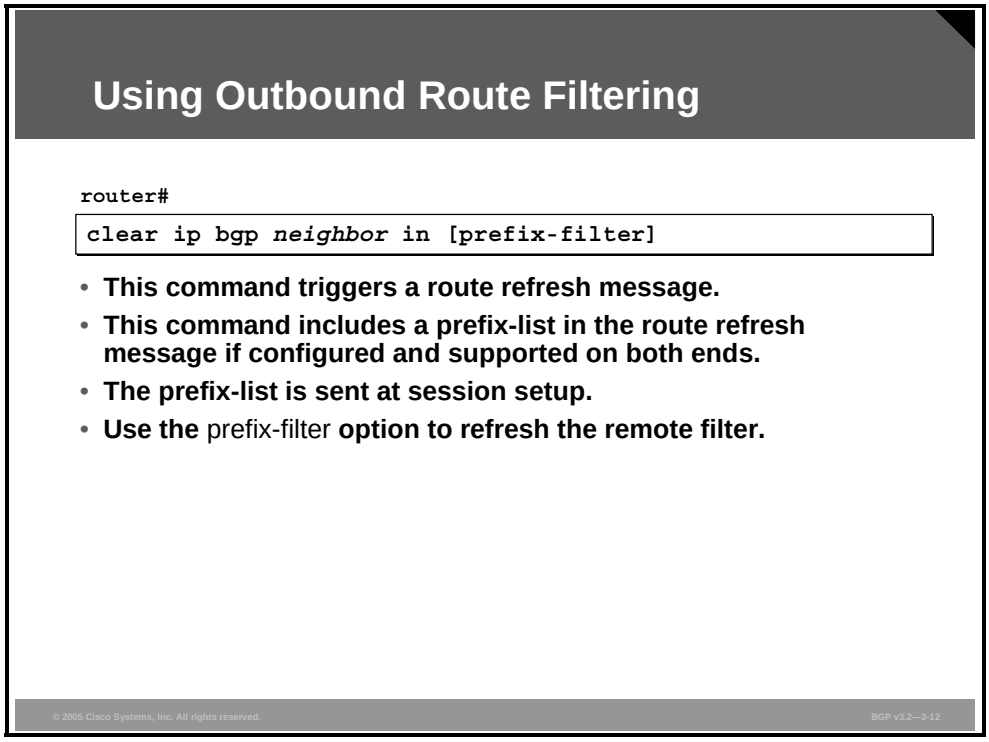
Step 1 Enable negotiation of outbound filtering based on prefix-lists.

Step 2 Attach an input prefix-list to a neighbor.

Step 3 Enable sending of input prefix-list to the neighbor.

Using Outbound Route Filtering

This topic identifies the Cisco IOS command that is used to trigger a route refresh message.



Use the **clear ip bgp** *neighbor* command with the **prefix-filter** keyword to push out the existing ORF prefix-list so that a new route refresh will be received from a neighbor. The neighbor will use the ORF prefix-list that was previously negotiated.

You need to use the **clear ip bgp** *neighbor* command only when the filter has been modified because the neighbor will store the filter for subsequent route refresh requests. The neighbor will then use the filter on all updates toward the router that originated the filter.

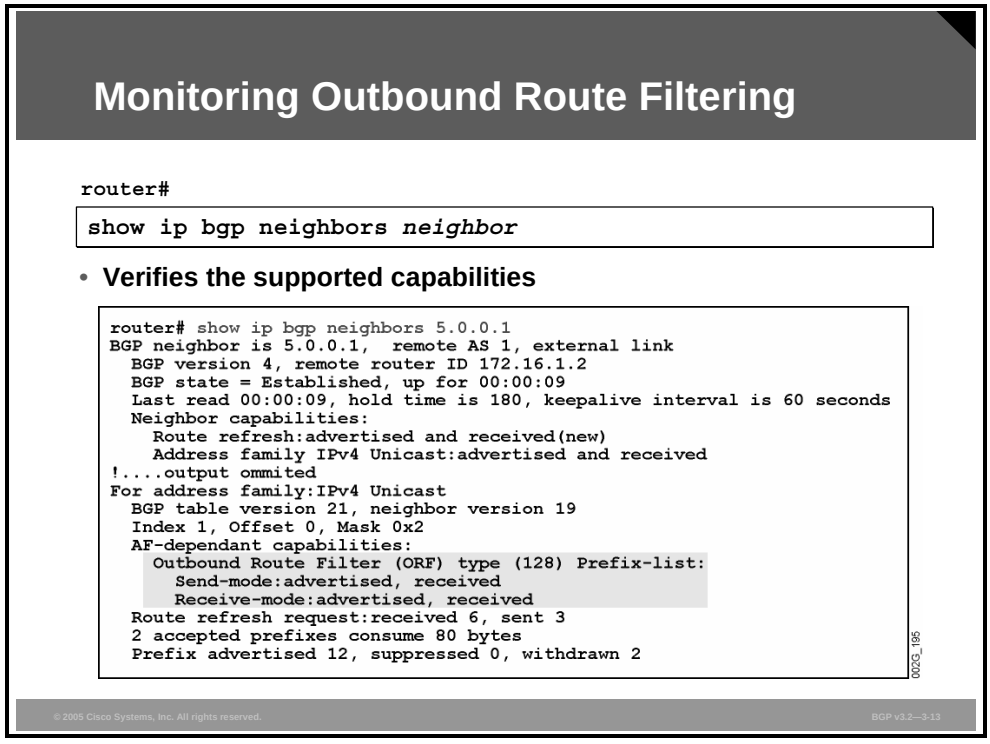
Note	You should enter the **in** keyword when you are using the **clear ip bgp** *neighbor* command because inbound route refresh is desired; only the inbound prefix-list filter is pushed to the neighbor and used by the neighbor in the outbound direction.

The router will ignore the **prefix-filter** keyword if ORF capability has not been received or the send capability has not been enabled.

When the **clear ip bgp** *neighbor* command is used without the **prefix-filter** keyword, a normal route refresh is performed. You should always use the **prefix-filter** keyword when ORF inbound routing policy changes occur.

Monitoring Outbound Route Filtering

This topic identifies the Cisco IOS command that is required to monitor the operation of an ORF that you have configured and activated.



Use the **show ip bgp neighbors** *neighbor* command to display the supported capabilities.

If the neighbor supports a certain ORF capability, it is shown as “advertised, received” and a filter of the supported type can be sent by that router to its neighbor.

The example output from the **show ip bgp neighbors** command shows that neighbor 5.0.0.1 is configured with the prefix-based ORF feature in both send and receive modes. ORF capabilities negotiation has been completed and is displayed per address family. The ORF type that has been negotiated by this router with its peer is 128 (Cisco proprietary, prefix-list-based).

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Outbound route filtering is a mechanism that is used to minimize the number of updates that are requested from a neighbor.**
- **The BGP prefix-based outbound route filtering feature uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers. Configuring this feature can help reduce the amount of system resources required for generating and processing routing updates by filtering out unwanted routing updates at the source.**
- **The ORF message contains the information that is used to determine which updates will be passed. The format of an ORF message includes AFI and SAFI for which the filter should be used; ORF type, which identifies the type of filter; when to refresh (immediate or deferred refresh); and a list of ORF entries where the actual filter is defined.**

Summary (Cont.)

- **The** neighbor *ip-address* capability orf prefix-list **command with the** send **and** receive **keywords enables ORF negotiations and activates an ORF prefix-list.**
- **Use the** clear ip bgp *neighbor* **command to trigger a BGP route refresh.**
- **With the** show ip bgp neighbors **command, neighbor-supported ORF capabilities are displayed as "advertised, received," and a filter of the supported type can be sent to the neighbor.**

Lesson 5

Applying Route-Maps as BGP Filters

Overview

Border Gateway Protocol (BGP) is a powerful routing protocol that supports a wide variety of administrative policy controls and route selection features. Many complex filtering goals and administrative policies cannot be achieved by using only single-purpose filtering methods or by compounding multiple filtering methods together. Route-maps provide a method to perform a variety of compound, complex filtering operations within a single tool. Understanding the operation and use of route-maps is a critical component in the successful implementation of any large-scale BGP deployment.

This lesson describes route-maps and how you can use them for BGP filtering. Included in this lesson are the commands that are required to use route-maps with prefix-lists and discussion of how to use them as BGP filters and how to monitor previously configured route-maps.

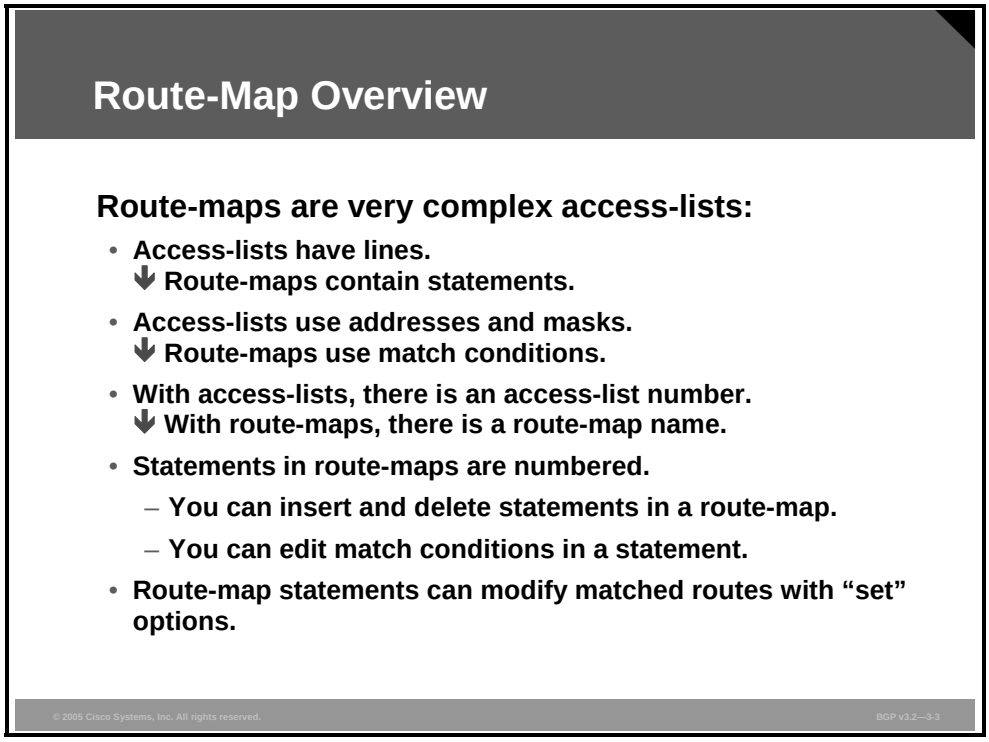
Objectives

Upon completing this lesson, you will be able to use outbound route filtering to minimize the impact of BGP routing updates on router resources in an operational BGP network. This ability includes being able to meet these objectives:

- Describe the high-level function of a route-map
- Describe the function of the BGP Route-Map Policy List Support feature
- Describe the function of the BGP Route-Map Continue feature
- Identify the Cisco IOS commands that are required to configure a route-map to match against a prefix-list
- Identify where you can apply route-maps as route filters in a BGP network
- Identify the Cisco IOS command that is required to enable a route-map as a BGP route filter
- Identify the Cisco IOS commands that are required to monitor the operation of a configured route-map that is used as a BGP filter

Route-Map Overview

This topic describes the high-level function of a route-map.



A route-map is a filter, and what is denied by the route-map is dropped. Additionally, you can use the route-map to modify attributes of the permitted routes.

Route-maps are similar to access-lists. Both have a set of tests to be performed, and several tests can be done in sequence. The first match produces the result of either “permit” or “deny.”

An access-list has a number of lines, each indicating a testing condition. The route-map is more complex than the access-list. The route-map consists of several groups of configuration lines; each group is called a statement. The statement has a sequence number that provides the opportunity to remove or modify an explicit statement without removing the entire route-map. There is also an opportunity to add a new statement between two existing statements.

Each route-map statement starts with a configuration line indicating the name of the route-map, the sequence number, and whether the result should be permitted or denied if the testing matches. The statement then continues, following configuration lines with the match clauses. Matching can be done in several ways: testing on the prefix, the autonomous system (AS) path, or some other attribute. The statement concludes with optional “set” statements, where attributes may be modified or set.

Route-Map Overview (Cont.)

```
route-map name [permit|deny sequence]
match condition
match condition
set parameter
```

002G_196

- **The default statement action is “permit.”**
- **A route not matched by any statement is dropped.**
- **“Permit all” is achieved by specifying “permit” without a “match” clause.**
- **Match conditions in one statement are ANDed together.**
- **The first matching statement permits or denies the route.**

A route-map consists of several statements. Each statement starts with the route-map configuration line, on which the name of the route-map must be indicated. A good practice is to always indicate the **permit** or **deny** keyword followed by a sequence number.

The matching clauses for the statements are listed on the match lines following the route-map line. There may be several match lines, each referring to a different test to be performed. All tests must be passed for the statement to be matched. If any of the match line tests fails, the next route-map statement is tried. Statements are tried in sequence number order. If there are no more statements in the route-map, the result is, implicitly, “deny.”

If all of the match clauses succeed, there is a match for the statement and the indicated result is used. If the result is to deny, the route is then silently ignored. If the result is to permit, the route is accepted and the set clauses are applied. The set clauses allow one or more attributes to be changed or set to specific values before the route is accepted.

Route-Map Overview (Cont.)

- **Route-map conditions are specified in the match statement.**
- **Route-maps can match on:**
  - **Network number and subnet mask matched with an IP prefix-list**
  - **Route originator**
  - **BGP next-hop address**
  - **BGP origin**
  - **Tag attached to IGP route**
  - **AS-path**
  - **BGP community attached to BGP route**
  - **IGP route type (internal/external …)**

Each route-map statement can have several match clauses, and each match clause is given its own configuration line. The match clause refers to the tests to be made on the candidate route. Tests of the candidate route can be based on the following criteria:

- IP network numbers and subnet masks, by referring to a prefix-list or access-list that will be applied on the route
- Route originator, by referring to a prefix-list or access-list that will be applied on the value of the originator BGP attribute
- Next hop, by referring to a prefix-list or access-list that will be applied on the value of the next-hop BGP attribute
- Origin code, by testing the value of the origin BGP attribute
- Tag value that is attached to an Interior Gateway Protocol (IGP) route—used only when redistribution from an IGP into BGP occurs
- AS path, by referring to an AS-path access-list that will be applied on the value of the AS-path BGP attribute
- Community, by referring to a community-list that will be applied on the value of the Community BGP attribute
- IGP route type, by testing if the IGP route is internal or external—used only when redistribution from an IGP to BGP occurs

Route-Map Overview (Cont.)

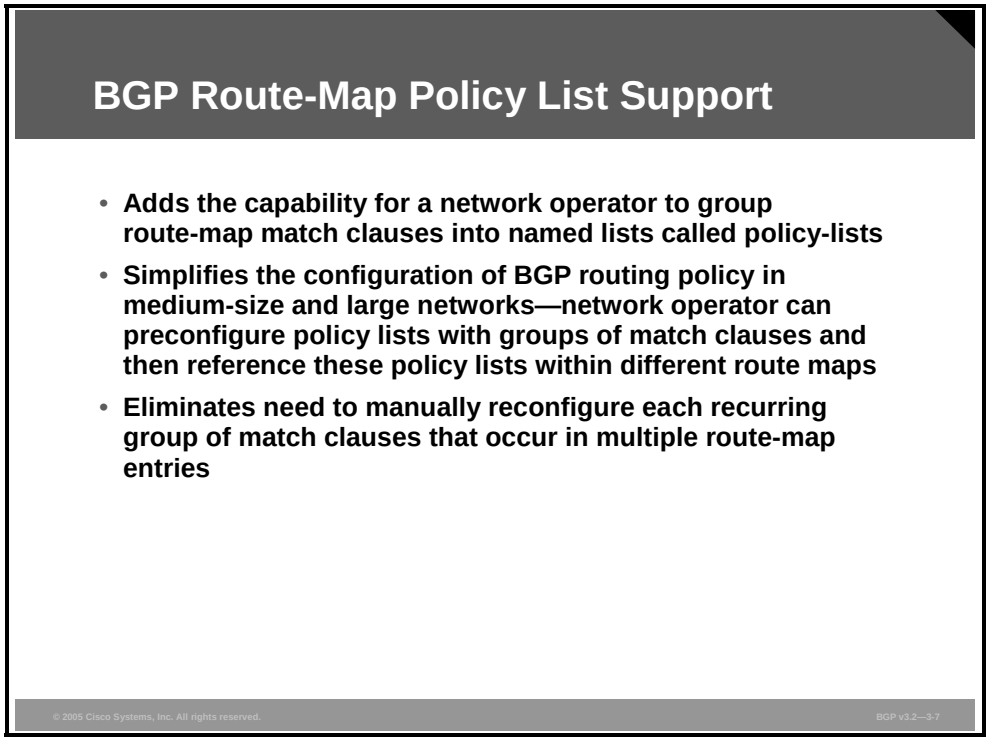
- **Route-maps can also change the attributes of BGP routes.**
- **Route-maps can set:**
  - **Origin**
  - **BGP next-hop**
  - **Weight**
  - **BGP community**
  - **Local preference**
  - **MED**

Each route-map statement may have several set clauses. Each set clause is applied to the route when the route-map statement permits the route. With a route-map, the following can be set:

- Origin BGP attribute
- Next-hop BGP attribute
- Weight
- Community BGP attribute
- Local preference BGP attribute
- Multi-exit discriminator (MED) BGP attribute, by setting the metric

BGP Route-Map Policy List Support

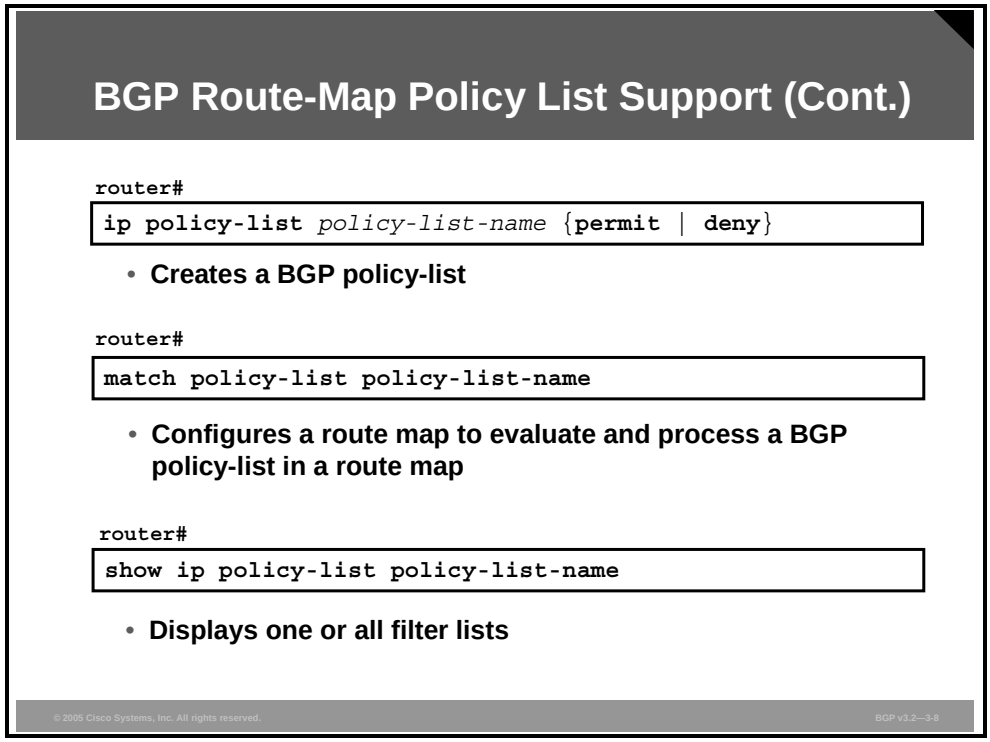
This topic describes the function of the BGP Route-Map Policy List Support feature.



The BGP Route-Map Policy List Support feature introduces new functionality to BGP route maps, allowing a network operator to group route-map match clauses into named lists called policy-lists. A policy-list functions like a macro.

When a policy-list is referenced in a route-map, all of the match clauses are evaluated and processed as if they had been configured directly in the route-map. The BGP Route-Map Policy List Support feature simplifies the configuration of BGP routing policy in medium-size and large networks because a network operator can preconfigure policy-lists with groups of match clauses and then reference these policy-lists within different route-maps. The network operator no longer needs to manually reconfigure each recurring group of match clauses in multiple route-map entries.

A policy-list is like a route-map that contains only match clauses. The policy-list is created and then referenced within a route-map. There are no changes to match clause semantics and route-map functions. Match clauses are configured in policy-lists with permit and deny statements. The route-map evaluates and processes each match clause and permits or denies routes based on the configuration. AND and OR semantics in the route-map function the same way for policy-lists that they do for match and set clauses. There are some commands that are related to the BGP Route-Map Policy List Support feature: the **ip policy-list** command, the **match policy-list** command, and the **show ip policy-list** command.



ip policy-list

To create a BGP policy-list, use the **ip policy-list** command in policy-map configuration mode.

- **ip policy-list** *policy-list-name* **{permit | deny}**

To remove a policy-list, use the **no** form of this command

- **no ip policy-list** *policy-list-name*

Syntax Description

Parameter	Description
policy-list-name	Name of the configured policy-list
permit	Permits access for matching conditions
deny	Denies access to matching conditions

match policy-list

To configure a route-map to evaluate and process a BGP policy-list in a route-map, use the **match policy-list** command in route-map configuration mode.

- **match policy-list** *policy-list-name*

To remove a path list entry, use the **no** form of this command.

- **no match policy-list** *policy-list-name*

Syntax Description

Parameter	Description
policy-list-name	Name of the configured policy-list

show ip policy-list

To display information about a configured policy-list and policy-list entries, use the **show ip policy-list** command in user EXEC mode.

- **show ip policy-list** *policy-list-name*

Syntax Description

Parameter	Description
policy-list-name	Name of the configured policy-list

Configuring Policy-List Examples

The following configuration example creates a BGP policy-list that permits matches on the AS path and multi-exit discriminator (MED) of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-1 permit
Router(config-policy-list)# match as-path 1
Router(config-policy-list)# match metric 10
Router(config-policy-list)# end
```

The following configuration example creates a BGP policy-list that permits matches on the specified BGP community and the next hop of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-2 permit
Router(config-policy-list)# match community 20
Router(config-policy-list)# match metric 10
Router(config-policy-list)# ip community-list 20 permit 20:1
Router(config-policy-list)# end
```

The following configuration example creates a BGP policy-list that denies matches on the specified BGP community and the next hop of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-3 deny
Router(config-policy-list)# match community 20
Router(config-policy-list)# match metric 10
Router(config-policy-list)# end
```


Configuring Route-Maps to Reference Policy-List Examples

The configuration examples in this section create BGP route-maps that reference BGP policy-lists with the **route-map** route-map configuration command.

The following configuration example creates a route-map that references policy-lists and separate match and set clauses in the same configuration. This example uses AND semantics between POLICY-LIST-NAME-1 and POLICY-LIST-NAME-2.

```
Router(config)# route-map MAP-NAME-1 10
Router(config-route-map)# match ip-address 1
Router(config-route-map)# match policy-list POLICY-LIST-NAME-1
Router(config-route-map)# match policy-list POLICY-LIST-NAME-2
Router(config-route-map)# set community 10:1
Router(config-route-map)# set local-preference 140
Router(config-route-map)# end
```

The following configuration example creates a route-map that references policy-lists and separate match and set clauses in the same configuration. This example uses OR semantics between POLICY-LIST-NAME-3 and POLICY-LIST-NAME-4.

```
Router(config)# route-map MAP-NAME-2 10
Router(config-route-map)# match policy-list POLICY-LIST-NAME-3 POLICY-
LIST-NAME-4
Router(config-route-map)# set community 10:1
Router(config-route-map)# set local-preference 140
Router(config-route-map)# end
```

Verifying BGP Route-Map Policy List Support

To verify that a policy-list has been created, use the **show ip policy-list** command. The output of this command displays the policy-list name and configured match clauses. The following sample output is similar to the output that will be displayed:

```
Router# show ip policy-list
policy-list POLICY-LIST-NAME-1 permit
  Match clauses:
    metric 20
policy-list POLICY-LIST-NAME-2 permit
  Match clauses:
    as-path (as-path filter): 1
```

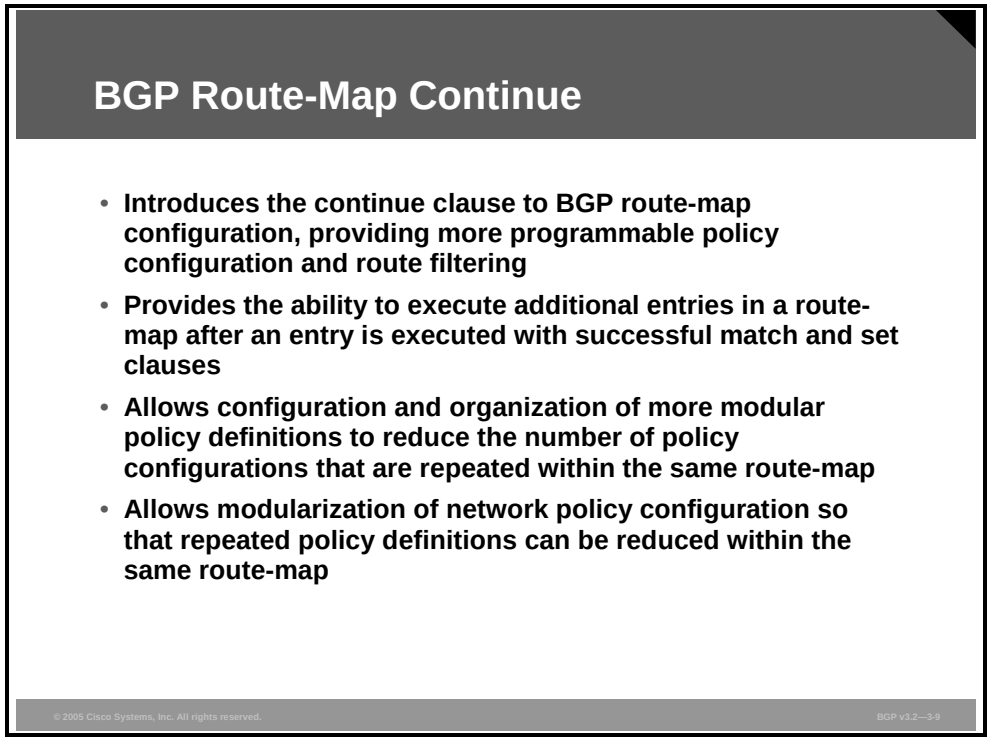
Note A policy-list name can be specified when the **show ip policy-list** command is entered. This option can be useful for filtering the output of this command and verifying a single policy-list.

To verify that a route-map has been created and a policy-list is referenced, use the **show route-map** command. The output of this command displays the route-map name and policy-lists that are referenced by the configured route-maps. The following sample output is similar to the output that will be displayed:

```
Router# show route-map
route-map ROUTE-MAP-NAME-1, deny, sequence 10
  Match clauses:
  Set clauses:
  Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME-1, permit, sequence 10
  Match clauses:
    IP Policy lists:
        POLICY-LIST-NAME-1
  Set clauses:
  Policy routing matches: 0 packets, 0 bytes
```


BGP Route-Map Continue

This topic describes the continue clause used in BGP route-map configuration.



The BGP Route-Map Continue feature introduces the continue clause to BGP route-map configuration. The continue clause provides more programmable policy configuration and route filtering. It introduces the ability to execute additional entries in a route-map after an entry is executed with successful match and set clauses. Continue clauses allow you to configure and organize more modular policy definitions to reduce the number of policy configurations that are repeated within the same route-map.

Continue clauses provide a programmable method to organize and control the flow of a route-map. Route-map configuration was linear before this feature was introduced. Continue clauses also allow you to modularize network policy configuration so that repeated policy definitions can be reduced within the same route-map.

Route-Map Operation Without Continue Clauses

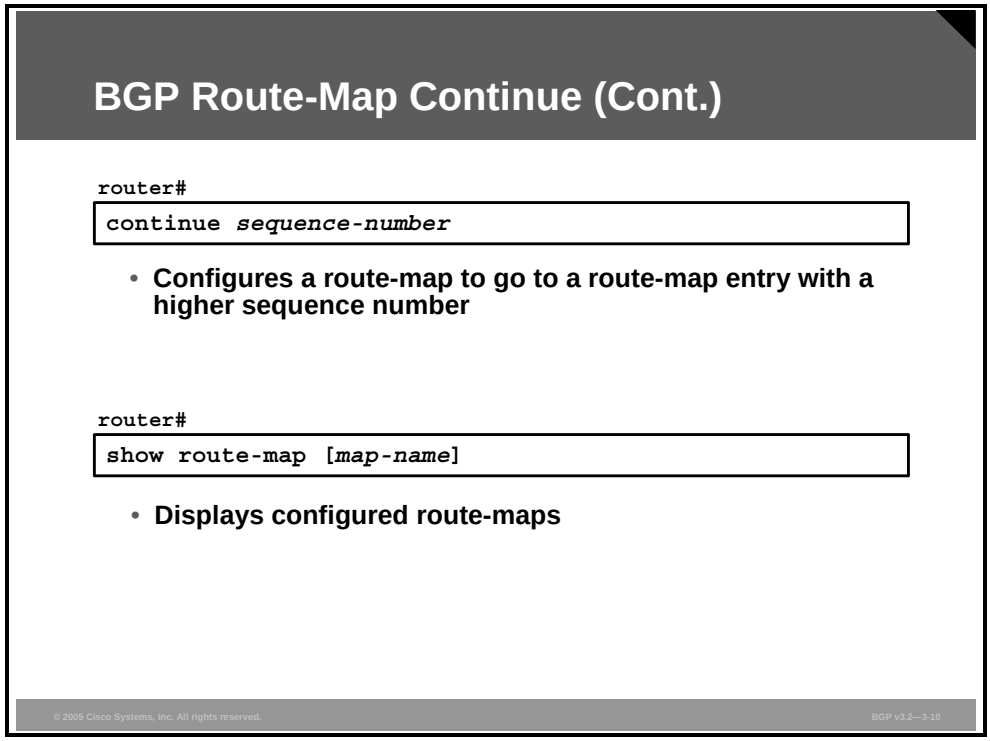
A route-map evaluates match clauses until a successful match occurs. After the match occurs, the route-map stops evaluating match clauses and starts executing set clauses, in the order in which they were configured. If a successful match does not occur, the route-map “falls through” and evaluates the next sequence number of the route-map until all configured route-map entries have been evaluated or a successful match occurs. Each route-map sequence is tagged with a sequence number to identify the entry. Route-map entries are evaluated in order, starting with the lowest sequence number and ending with the highest sequence number. If the route-map contains only set clauses, the set clauses are executed automatically, and the route-map does not evaluate any other route-map entries.

Route-Map Operation with Continue Clauses

When a continue clause is configured, the route-map continues to evaluate and execute match clauses in the specified route-map entry after a successful match occurs. The continue clause can be configured to go to (or jump to) a specific route-map entry by specifying the sequence number, or if a sequence number is not specified, to go to the next sequence number. This behavior is called an “implied continue.” If a match clause exists, the continue clause is executed only if a match occurs. If no successful matches occur, the continue clause is ignored.

If a match clause does not exist in the route-map entry but a continue clause does, the continue clause is automatically executed and goes to the specified route-map entry. If a match clause exists in a route-map entry, the continue clause is executed only when a successful match occurs. When a successful match occurs and a continue clause exists, the route-map executes the set clauses and then goes to the specified route-map entry. If the next route-map contains a continue clause, the route-map executes the continue clause if a successful match occurs. If a continue clause does not exist in the next route-map, the route-map is evaluated normally. If a continue clause exists in the next route-map but a match does not occur, the route-map does not continue and falls through to the next sequence number, if one exists.

Note A continue clause can be executed, without a successful match, if a route-map entry does not contain a match clause.



You will use two commands with the BGP Route-Map Continue feature, the **continue** command and the **show route-map** command.

continue

To configure a route-map to go to a route-map entry with a higher sequence number, use the **continue** command in route-map configuration mode.

- **continue** *sequence-number*

To remove a continue clause from a route-map, use the **no** form of this command.

- **no continue**

Syntax Description

Parameter	Description
sequence-number	(Optional) Route-map sequence number. If a route-map sequence number is not specified when configuring a continue clause, the continue clause continues to the route-map entry with the next sequence number. This behavior is referred to as an "implied continue."

show route-map

To display configured route-maps, use the **show route-map** command in EXEC mode.

- **show route-map** [*map-name*]

Syntax Description

Parameter	Description
map-name	(Optional) Name of a specific route-map

BGP Route-Map Continue Clause Example Configuration

The following example shows continue clause configuration in a route-map sequence.

The first continue clause in route-map entry 10 indicates that the route map will go to route-map entry 30 if a successful match occurs. If a match does not occur, the route map will fall through to route-map entry 20. If a successful match occurs in route-map entry 20, the set action will be executed and the route map will not evaluate any additional route-map entries.

If a successful match does not occur in route-map entry 20, the route map will fall through to route-map entry 30. This sequence does not contain a match clause, so the set clause will be automatically executed and the continue clause will go to the next route-map entry because a sequence number is not specified.

If there are no successful matches, the route-map will fall through to route-map entry 30 and execute the set clause, and route-map entry 40 will not be evaluated.

```
route-map ROUTE-MAP-NAME permit 10
 match ip address 1
 match metric 10
 set as-path prepend 10
 continue 30
 !
route-map ROUTE-MAP-NAME permit 20
 match ip address 2
 match metric 20
 set as-path prepend 10 10
 !
route-map ROUTE-MAP-NAME permit 30
 set as-path prepend 10 10 10
 continue
 !
route-map ROUTE-MAP-NAME permit 40
 match community 10:1
 set local-preference 104
```

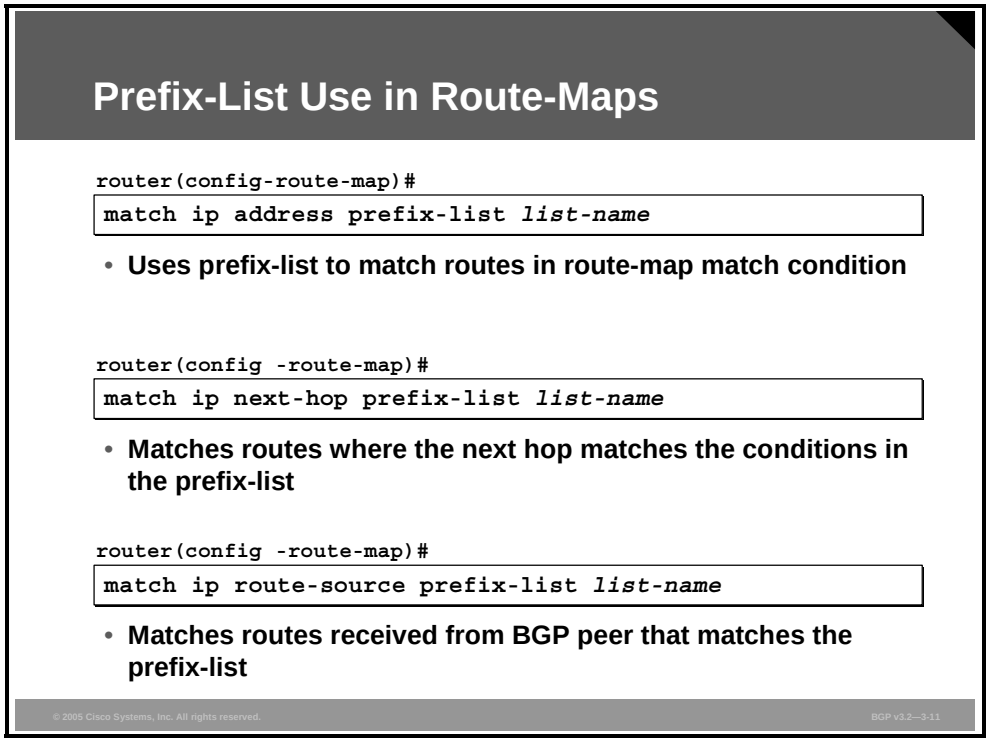
BGP Route-Map Continue Clause Verification Example

To verify the configuration of continue clauses, use the **show route-map** command. The output of this command displays configured route-maps, match, set, and continue clauses. The following sample output is similar to the output that will be displayed:

```
Router# show route-map
route-map ROUTE-MAP-NAME, permit, sequence 10
  Match clauses:
    ip address (access-lists): 1
    metric 10
  Continue: sequence 40
  Set clauses:
    as-path prepend 10
  Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 20
  Match clauses:
    ip address (access-lists): 2
    metric 20
  Set clauses:
    as-path prepend 10 10
  Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 30
  Match clauses:
  Continue: to next entry 40
  Set clauses:
    as-path prepend 10 10 10
  Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 40
  Match clauses:
    community (community-list filter): 10:1
  Set clauses:
    local-preference 104
  Policy routing matches: 0 packets, 0 bytes
route-map LOCAL-POLICY-MAP, permit, sequence 10
  Match clauses:
  Set clauses:
    community 655370
  Policy routing matches: 0 packets, 0 bytes
```

Prefix-List Use in Route-Maps

This topic identifies the Cisco IOS commands that are required to configure a route-map to match against a prefix-list.



match ip address

To distribute any routes that have a destination network number address that is permitted by a standard access-list, an extended access-list, or a prefix-list, or to perform policy routing on packets, use the **match ip address** command in route-map configuration mode.

- **match ip address** *{access-list-number [access-list-number... | access-list-name...] | access-list-name [access-list-number...| access-list-name] | prefix-list prefix-list-name [prefix-list-name...]}*

To remove the **match ip address** entry, use the **no** form of this command.

- **no match ip address** *{access-list-number [access-list-number... | access-list-name...] | access-list-name [access-list-number...| access-list-name] | prefix-list prefix-list-name [prefix-list-name...]}*

Syntax Description

Parameter	Description
access-list-number...	Number of a standard or extended access-list. It can be an integer from 1 to 199. The ellipsis indicates that multiple values can be entered.
access-list-name...	Name of a standard or extended access-list. It can be an integer from 1 to 199. The ellipsis indicates that multiple values can be entered.
prefix-list	Distributes routes based on a prefix-list.
prefix-list-name...	Name of a specific prefix-list. The ellipsis indicates that multiple values can be entered.

match ip next-hop

To redistribute any routes that have a next-hop router address passed by one of the access-lists specified, use the **match ip next-hop** command in route-map configuration mode.

- **match ip next-hop** *{access-list-number | access-list-name}[...access-list-number | ...access-list-name]*

To remove the next hop entry, use the **no** form of this command.

- **no match ip next-hop** *{access-list-number | access-list-name}[...access-list-number | ...access-list-name]*

Syntax Description

Parameter	Description
access-list-number	Number of a standard or extended access-list. It can be an integer from 1 to 199.
access-list-name	Name of a standard or extended access-list. It can be an integer from 1 to 199.

match ip route-source

To redistribute routes that have been advertised by routers and access servers at the address specified by the access-lists, use the **match ip route-source** command in route-map configuration mode.

- **match ip route-source** *{access-list-number | access-list-name}[...access-list-number | ...access-list-name]*

To remove the route-source entry, use the **no** form of this command.

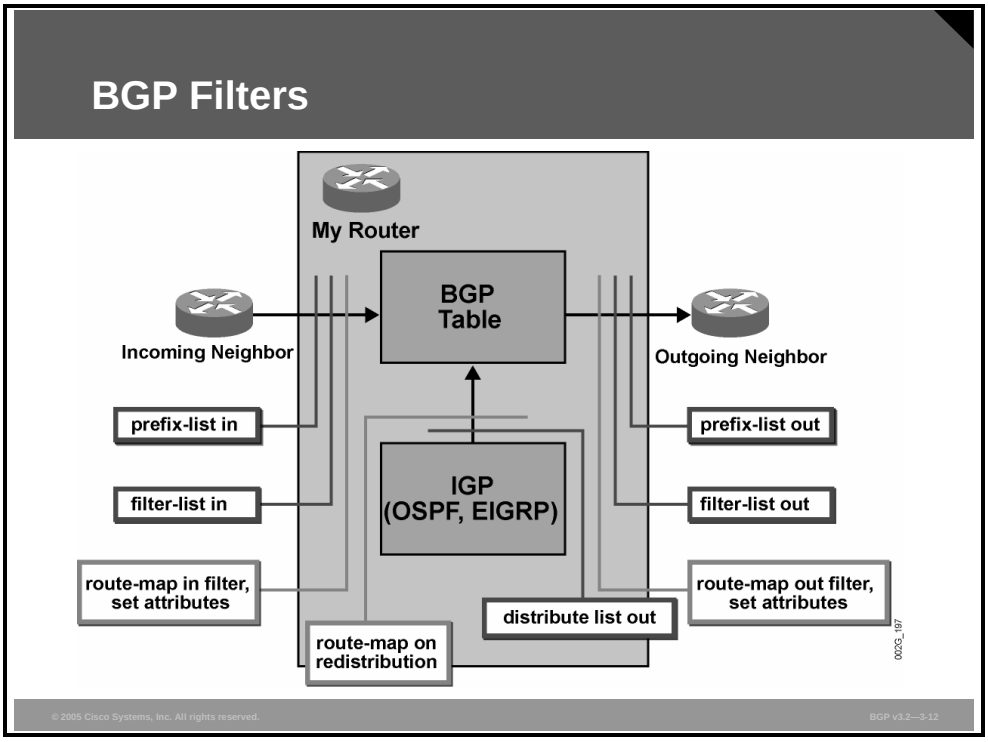
- **no match ip route-source** *{access-list-number | access-list-name}[...access-list-number | ...access-list-name]*

Syntax Description

Parameter	Description
access-list-number	Number of a standard or extended access-list. It can be an integer from 1 to 199.
access-list-name	Name of a standard or extended access-list. It can be an integer from 1 to 199.

BGP Filters

This topic identifies where you can apply route-maps as route filters in a BGP network.

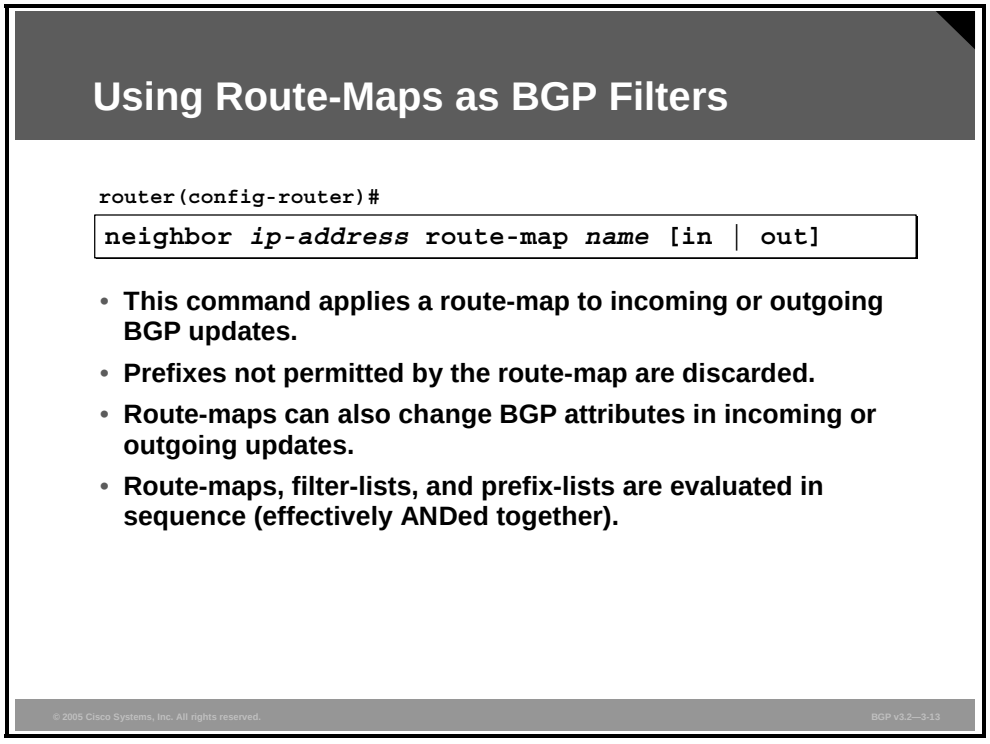


You can optionally apply filter-lists, prefix-lists, and route-maps to either incoming or outgoing information or any combination of the two. The incoming prefix-list, the incoming filter-list, and the incoming route-map must all permit the routes that are received from a neighbor before being accepted into the BGP table. Outgoing routes must pass the outgoing filter-list, the outgoing prefix-list, and the outgoing route-map before being transmitted to the neighbor.

When a router is configured to redistribute routing information from an IGP into BGP, the routes must successfully pass any prefix-list or route-map that is applied to the redistribution before a route is injected into the BGP table.

Using Route-Maps as BGP Filters

This topic identifies the Cisco IOS command that is required to enable a route-map as a BGP route filter.

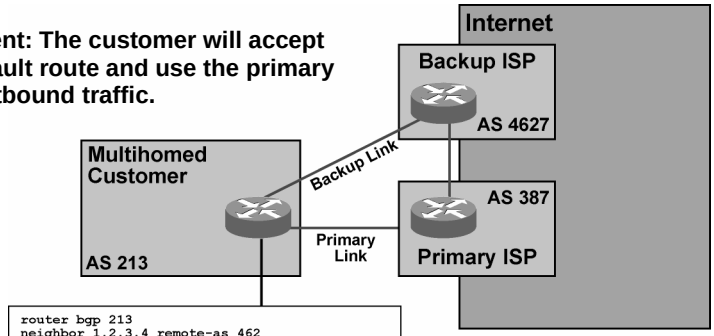


You can apply a route-map on incoming or outgoing routing information for a neighbor. The routing information must be permitted by the route-map in order to be accepted. If there is no statement in the route-map explicitly permitting a route, then the route will be implicitly denied and dropped.

The permitted routes may have their attributes set or changed by the set clauses in the route-map. Setting attributes on routes is useful when you are influencing route selection. Some routes can be permitted by one of the statements in the route-map and have their attributes changed. Another statement in the route-map could permit other routes and not have their attributes altered. When route selection is performed, the attribute values indicate that one route is preferred over the other.

Using Route-Maps as BGP Filters (Cont.)

- **Requirement: The customer will accept only a default route and use the primary link for outbound traffic.**



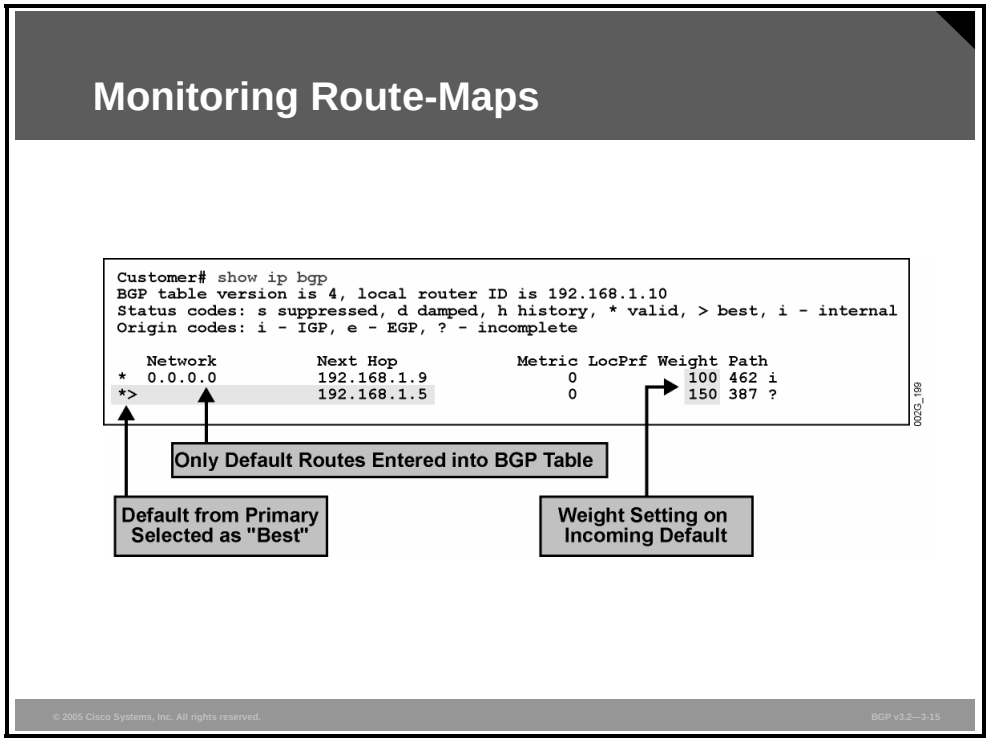

```
router bgp 213
neighbor 1.2.3.4 remote-as 462
neighbor 1.2.3.4 route-map filter in
neighbor 3.4.5.6 remote-as 387
neighbor 3.4.5.6 route-map filter in
!
route-map filter permit 10
 match ip address prefix-list defonly
 match as-path 10
 set weight 150
!
route-map filter permit 20
 match ip address prefix-list defonly
 set weight 100
!
ip as-path access-list 10 permit _387$
ip prefix-list defonly seq 10 permit 0.0.0.0/0
```

0020_196

In this example, the customer will accept only a default route and use the primary link that is connected to AS 387 for outbound traffic.

Monitoring Route-Maps

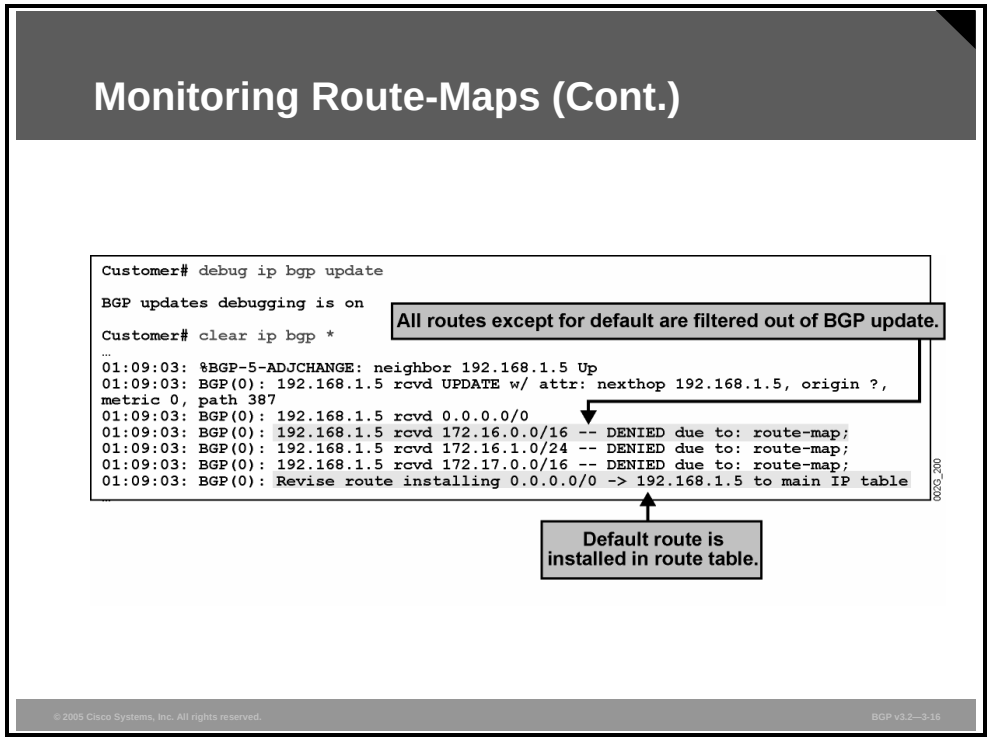
This topic identifies the Cisco IOS commands that are required to monitor the operation of a configured route-map that is used as a BGP filter.



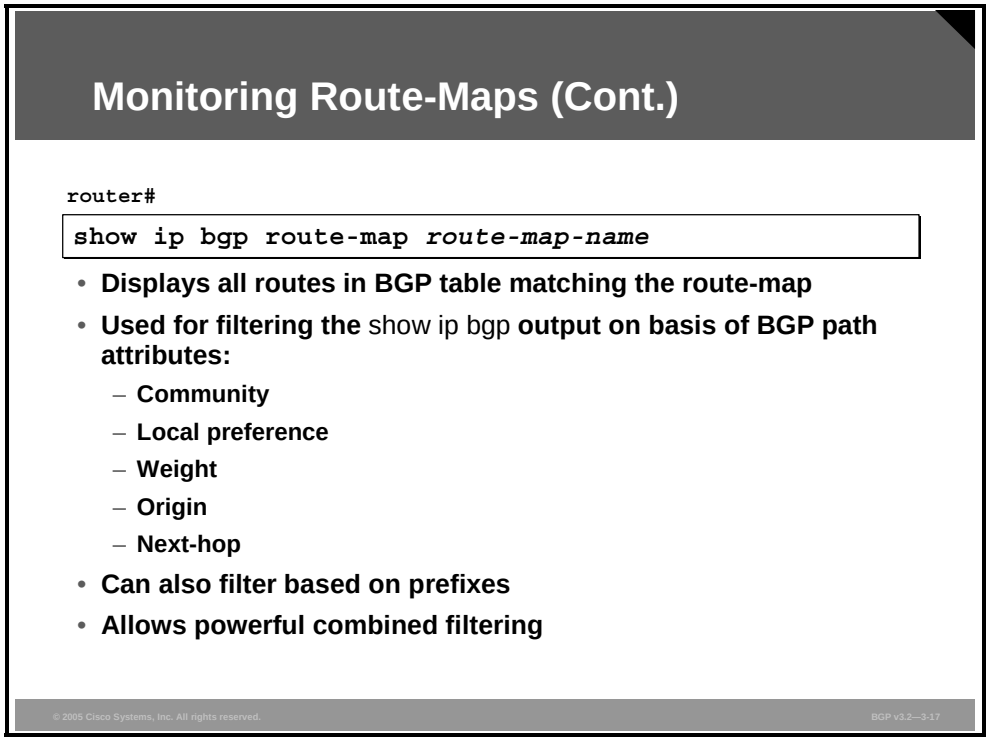
Use the **show ip bgp** command to display the configured route-map characteristics.

The example shows that only default routes are entered into the BGP table. The default route from the primary link has been selected by BGP as the “best” route. The BGP route selection rules have been modified based on the configuration of the BGP weight attribute in the route-map. As part of that configuration, the weight of the primary link has been set to 150 and the weight of the backup link has been set to 100.

Because BGP path selection prefers the highest weight, the router uses the primary link as the outgoing path.



Here you see that all routes except for the default route are being filtered out of the BGP update (denied). The default route is installed in the route table.



You can also use route-maps for selective and powerful filtering of the BGP table. The **show ip bgp route-map** command displays selected routes from a BGP routing table based on the contents of a route-map.

A route-map can match the routes on the basis of BGP path attributes (local preference, community, weight, origin, next-hop) or prefix-lists and access-lists (matching IP prefixes). The power of route-map filtering lies in the possibility of combining different filters (for example, filtering on community, prefix, and next-hop values).

Note	Support for route-map filtering was added in Cisco IOS Software Release 12.2(11)T and 12.0(14)ST.

Monitoring Route-Maps (Cont.)

```
Customer# show ip bgp route-map filter
BGP table version is 9, local router ID is 192.168.1.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  0.0.0.0          192.168.1.9              0             100 462 i
*>                  192.168.1.5              0             150 387 ?
```

Default routes are the only routes matching the route-map filter.

- **Networks matched by the route-map are displayed.**

In this example, a customer is using a simple route-map to filter the BGP table. By using the **show ip bgp route-map** command, the customer can display the filtered BGP table. The customer router configuration from which this output is collected is shown here for reference:

```
router bgp 213
neighbor 1.2.3.4 remote-as 462
neighbor 1.2.3.4 route-map filter in
neighbor 3.4.5.6 remote-as 387
neighbor 3.4.5.6 route-map filter in
!
route-map filter permit 10
 match ip address prefix-list defonly
 match as-path 10
 set weight 150
!
route-map filter permit 20
 match ip address prefix-list defonly
 set weight 100
!
ip as-path access-list 10 permit _387$
ip prefix-list defonly seq 10 permit 0.0.0.0/0
```

The route-map “filter” matches incoming networks from two service providers. For all routes that are sent by the primary provider (AS 387), the local router accepts the default route only, and it is marked as the preferred route with a weight of 150. Only a default route is accepted from the backup provider, and its weight metric has been set to 100.

The customer then applies the route-map to the output of the **show ip bgp route-map** command, and only the networks that conform to the AS-path and prefix-list filters are displayed (network 0.0.0.0/0 in the example).

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **A route-map is a filter that has the ability to drop denied routes as well as modify attributes of the permitted routes.**
- **The BGP Route-Map Policy List Support feature introduces new functionality to BGP route-maps, adding the ability for a network operator to group route-map match clauses into named lists called policy-lists.**
- **The BGP Route-Map Continue feature introduces the continue clause to BGP route-map configuration. Continue clauses provide a programmable method to organize and control the flow of a route-map.**

Summary (Cont.)

- **You can configure a route-map to match against a prefix-list by using the** match ip address**,** match ip next-hop**, and** match ip route-source **commands.**
- **Filter-lists, prefix-lists, and route-maps can optionally all be applied on either incoming or outgoing information in any combination.**
- **A route-map can be applied on incoming or outgoing routing information to or from a neighbor, but the routing information must be permitted by the route-map in order to be accepted.**
- **Monitoring route-maps is possible using the** show ip bgp **and** debug ip bgp update **commands.**

Lesson 6

Implementing Changes in BGP Policy

Overview

Because of the huge volumes of routing information that Border Gateway Protocol (BGP) is capable of handling, traditional routing update methods are not feasible. Routing policies for a BGP neighbor may include filtering mechanisms such as route-maps, distribute-lists, prefix-lists, and autonomous system (AS)-path filter-lists. Each of these filters may impact inbound or outbound routing table updates.

Whenever there is an administrative change in routing policy, the BGP session must be reset before the new policy can take effect. To accomplish this task, there are two types of reset: hard reset and soft reset. Clearing a BGP session using a hard reset invalidates the cache and results in a negative impact on the operation of networks, because the information in the cache becomes unavailable. A soft reset is recommended because it allows routing tables to be reconfigured and activated without clearing the BGP session.

This lesson discusses routing updates in a BGP environment and the traditional methods of forcing BGP route updates after changes in a filter policy. The function and benefits of soft reconfiguration and route refresh are also discussed. The lesson also presents the commands that are required to perform a soft reconfiguration and route refresh and explains how to monitor and troubleshoot these features.

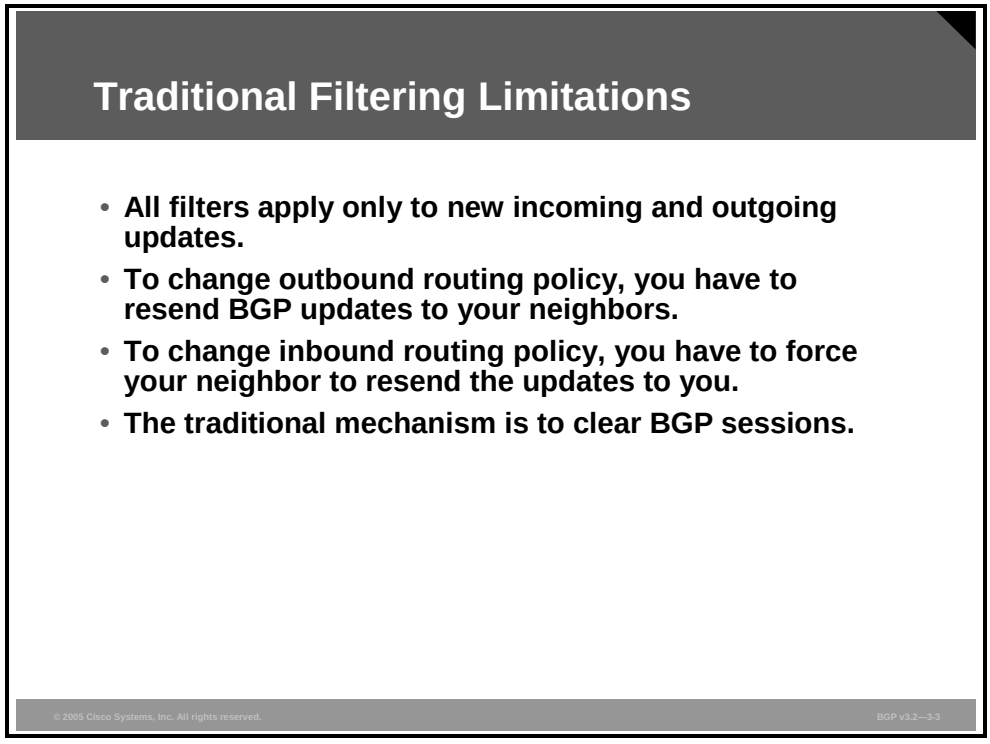
Objectives

Upon completing this lesson, you will be able to configure the soft reconfiguration feature to minimize the impact of expediting BGP policy updates in a typical BGP network. This ability includes being able to meet these objectives:

- Identify the limitations of the traditional methods of forcing BGP route updates after changing a filter policy
- Describe the function of the soft reconfiguration feature
- Identify the Cisco IOS commands that are required to configure and perform a soft reconfiguration
- Identify the Cisco IOS tools that are available to monitor the operation of a soft reconfiguration
- Describe the function of the BGP Soft Reset Enhancement feature
- Describe the function and benefits of the route refresh function
- Identify the Cisco IOS command that is required to trigger a route refresh
- Identify the Cisco IOS commands that are required to monitor route refresh operation
- Explain the benefit of using route-maps as BGP filters

Traditional Filtering Limitations

This topic identifies the limitations of traditional methods when you are forcing BGP route updates after changing filter policies.



BGP can potentially handle huge volumes of routing information. But when network administrators change configuration lines in filters or route-maps, the router cannot go through the huge table of BGP information and calculate which entry is no longer valid in the local table. Nor can the router determine which route or routes, already advertised, should be withdrawn from a neighbor. There is an obvious risk that the first configuration change will be immediately followed by a second, which would cause the whole process to start all over again.

To avoid such a problem, Cisco IOS software applies changes only on the updates that are received or transmitted after the configuration change has been performed. This approach means that the new routing policy, enforced by the new filters, is applied only on routes that are received or sent after the change. If network administrators would like to apply the policy change on all routes, they have to force the router to let all routes pass through the new filter.

If the filter is applied to outgoing information, the router has to resend the entire BGP table through the new filter. If the filter is applied to incoming information, the router needs its neighbor to resend its entire BGP table so that it passes through the new filters.

Traditionally, to accomplish these goals, network administrators have torn down the affected BGP sessions after completing a configuration change. After the sessions are down, all information that is received on those sessions is invalidated and removed from the BGP table. Also, the remote neighbor will detect a session down state, and it likewise will invalidate the routes that are received on the session. After a period of 30 to 60 seconds, the sessions are re-established automatically and the entire BGP table is exchanged again, but through the new filters. This process, however, disrupts packet forwarding.

Traditional Limitations of Clearing the BGP Session

```
router#
clear ip bgp {* | ip-address | peer-group-name}
```

- **This command tears down the BGP session with all neighbors, a specific neighbor, or all neighbors in a peer group.**
- **All BGP routes are lost after the session is torn down; connectivity through the BGP neighbor is lost.**
- **A new session is re-established within 30 to 60 seconds.**
- **A full routing update is exchanged once the session is re-established, resulting in enforcement of new routing policy.**
- **Processing the full Internet routing table can take a long time.**
- **Clearing the BGP session is a very disruptive way to implement routing policies.**

The EXEC command **clear ip bgp** tears down one or several BGP sessions. The BGP sessions are terminated, and the TCP connections closed. The neighbors go into the Idle state and stay there for approximately 30 seconds. Next, the neighbor session goes into the Active state, and the sessions are re-established.

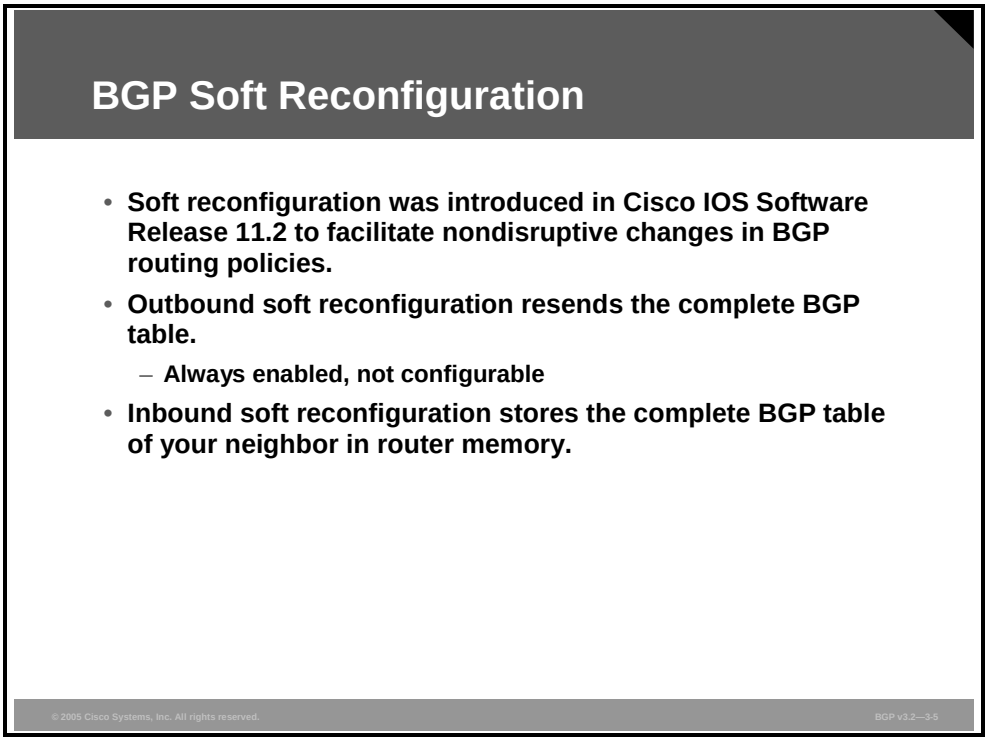
You can implement the **clear ip bgp** command with the ***** (asterisk) argument, which applies to all sessions, or you can make a reference to a specific session or group of sessions to tear down.

When the session is down, all routes that are received over the session by both routers are invalidated. When the session is once again in the Established state, all BGP routes have to be resent by both peers and pass through the new filters, which enforces the new policy.

Exchanging the complete Internet routing table takes time, bandwidth, and CPU resources. IP packet forwarding to and from the neighbor is down for several minutes. Also, revoking and reannouncing the routes will be registered by the rest of the Internet as a flap for each route.

BGP Soft Reconfiguration

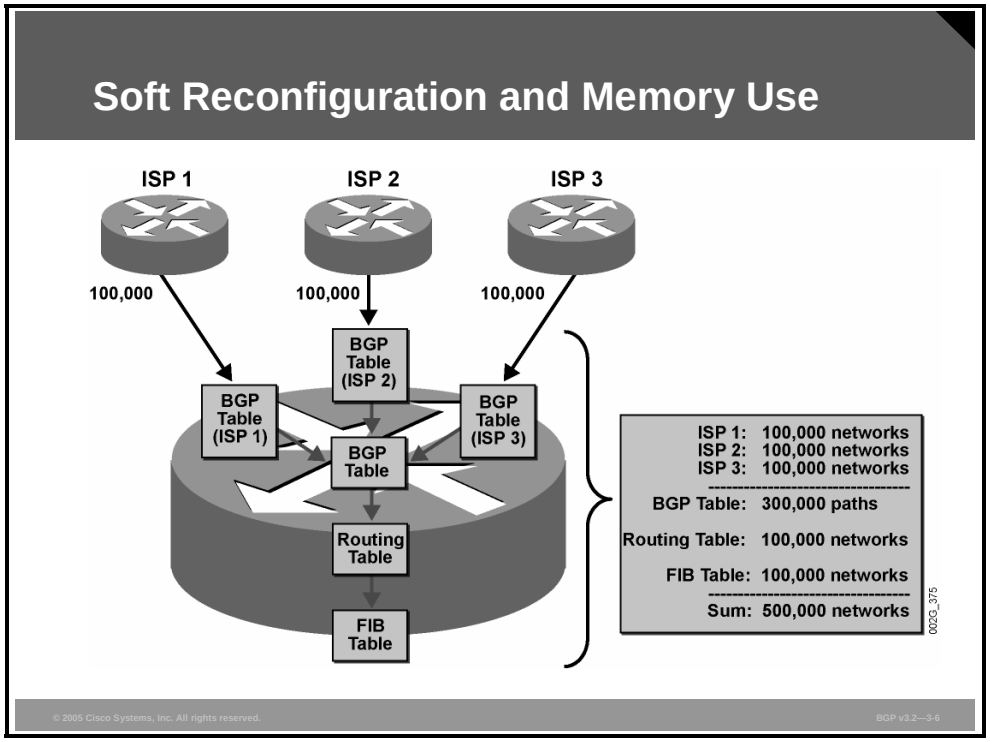
This topic describes the function of the soft reconfiguration feature.



With Cisco IOS Software Release 11.2 came the introduction of the soft reconfiguration feature. Soft reconfiguration provides the ability to run all routes through the filters without tearing down the sessions. Outbound soft reconfiguration was easy to implement because it is a simple resending of all routes in the local BGP table. Inbound soft reconfiguration was more complicated because a copy of all the routes that are received from a neighbor is required. The copy of the routes that are received from the neighbor is saved independently of the BGP table, before any filters are applied. Whenever the incoming filters are changed, a replay of everything that has been received from the neighbor takes place without involving the neighbor. The major drawback of this approach is the amount of memory that is required to hold the copy.

Example: Soft Reconfiguration and Memory Use

This example shows the impact of soft reconfiguration on an Internet service provider (ISP) router with three upstream neighbors sending full Internet routing information.

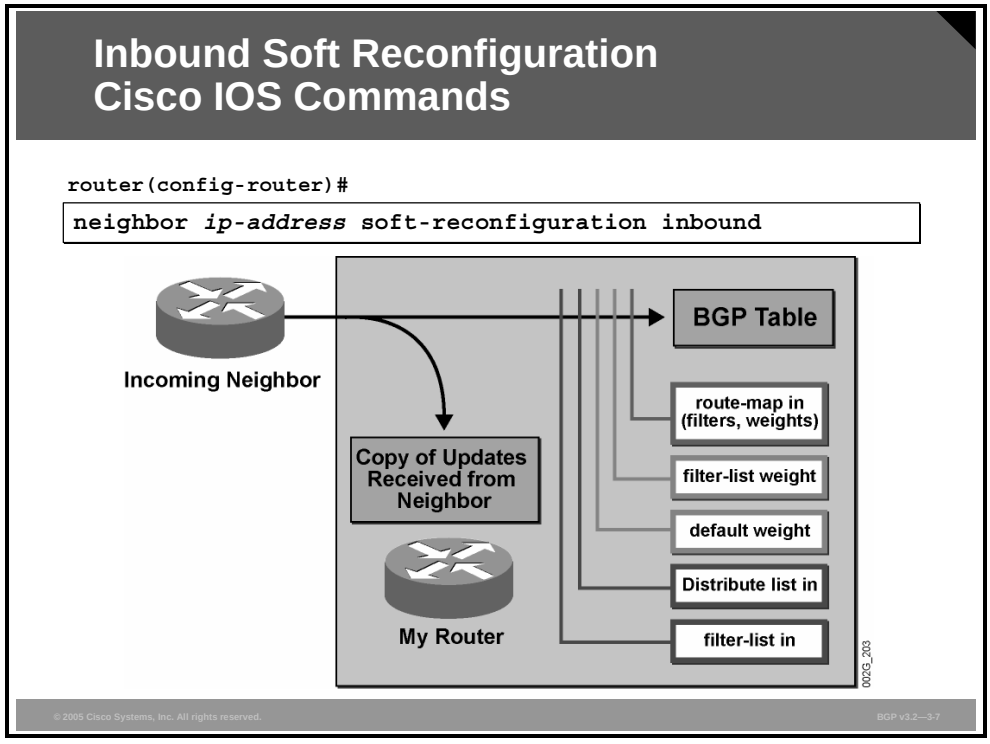


Each neighbor is sending 100,000 prefixes. The router stores each set in a dedicated per-neighbor BGP table. All 300,000 paths will then appear in the main BGP table if there is no filtering. The router will then choose the best path for each prefix and put it into the routing table. If Cisco Express Forwarding (CEF) switching is enabled, the router will store another copy in the Forwarding Information Base (FIB) table.

This solution obviously does not scale in terms of the number of neighbors and prefixes.

Cisco IOS Commands for Soft Reconfiguration

This topic identifies the Cisco IOS commands that are required to configure and perform a soft reconfiguration.



When you configure the **soft-reconfiguration inbound** command for a neighbor, the router stores all routes that are received from that neighbor as an extra copy in memory. This copy is taken before any filtering is applied by the router to routes that it receives.

This process is not enabled by default because it may consume large volumes of memory.

neighbor soft-reconfiguration

To configure Cisco IOS software to start storing updates, use the **neighbor soft-reconfiguration** router configuration command.

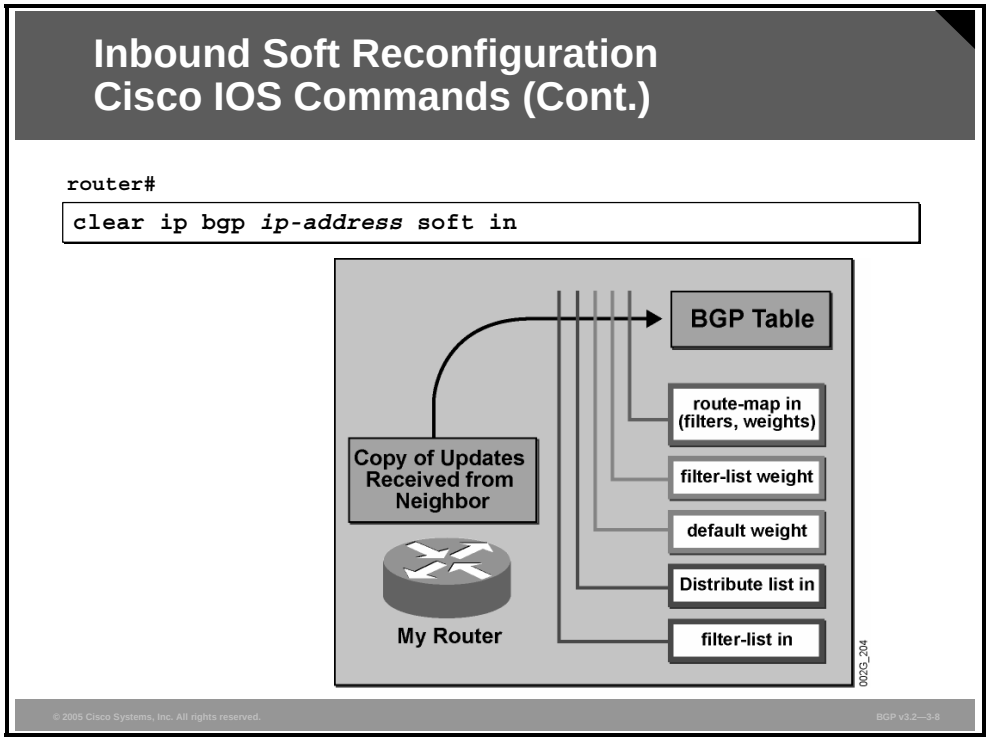
- **neighbor** {*ip-address* | *peer-group-name*} **soft-reconfiguration inbound**

To not store received updates, use the **no** form of this command.

- **no neighbor** {*ip-address* | *peer-group-name*} **soft-reconfiguration inbound**

Syntax Description

Parameter	Description
ip-address	IP address of the BGP-speaking neighbor
peer-group-name	Name of a BGP peer group
inbound	(Optional) Keyword that indicates that the update to be stored is an incoming update



When the network administrator has completed the changes to filters and route-maps that are applied on incoming information (changes that will implement a new routing policy), the **clear ip bgp** *ip-address* **soft in** command is executed on the router in privileged EXEC mode. After the command has been entered, the router will not tear the session down. Instead, the router resends the saved copy of the received routing information through the new filters, and the result is stored in the local BGP table.

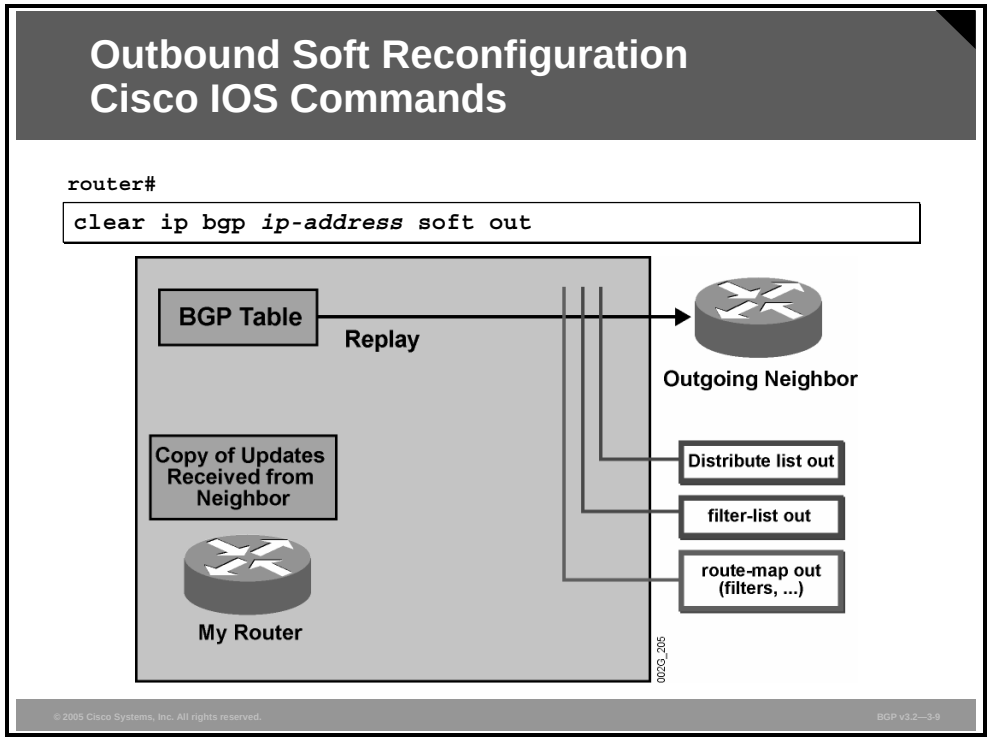
clear ip bgp

To reset a BGP connection using BGP soft reconfiguration, use the **clear ip bgp** EXEC command at the system prompt.

- **clear ip bgp** {***** | *address* | *peer-group-name*} [**soft** [**in** | **out**]]

Syntax Description

Parameter	Description
*****	Resets all current BGP sessions.
address	Resets only the identified BGP neighbor.
peer-group-name	Resets the specified BGP peer group.
soft	(Optional) Soft reset. Does not reset the session.
in | **out**	(Optional) Triggers inbound or outbound soft reconfiguration. If the **in** or **out** option is not specified, both inbound and outbound soft reset are triggered.

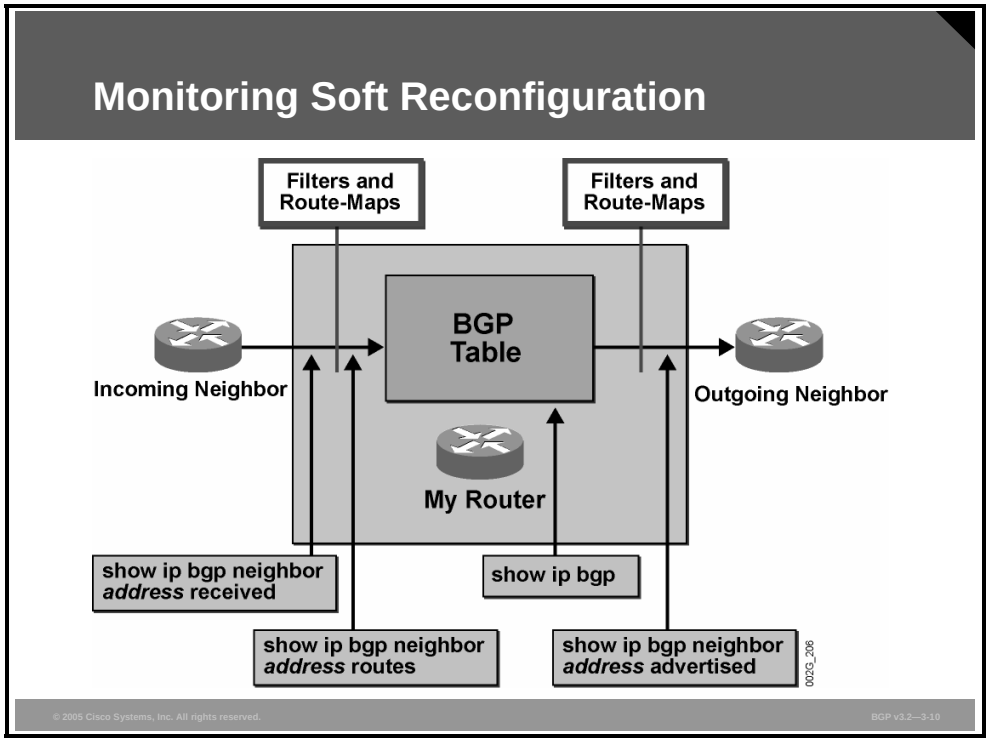


When the network administrator has completed the changes to filters and route-maps that are applied on the outgoing information (changes that will implement a new routing policy), the **clear ip bgp** *ip-address* **soft out** command is executed on the router in privileged EXEC mode. After the command has been entered, the router will not tear the session down. Instead, the table version number of the neighbor is reset to 0. When the next update interval for the neighbor arrives, the local router will go through the entire BGP table and find that all the routes need to be sent to the neighbor because they all have a table version number higher than 0.

This process causes all the BGP routes to be resent through the new filters.

Monitoring Soft Reconfiguration

This topic identifies the Cisco IOS tools that are available to monitor the operation of a soft reconfiguration.



The **show ip bgp** command is used to display the local BGP table. You can check the entries that have been propagated to a specific neighbor with the **show ip bgp neighbor** *ip-address* **advertised** command. It displays the subset of the local BGP table that has passed the split-horizon check and all outgoing filters for the neighbor.

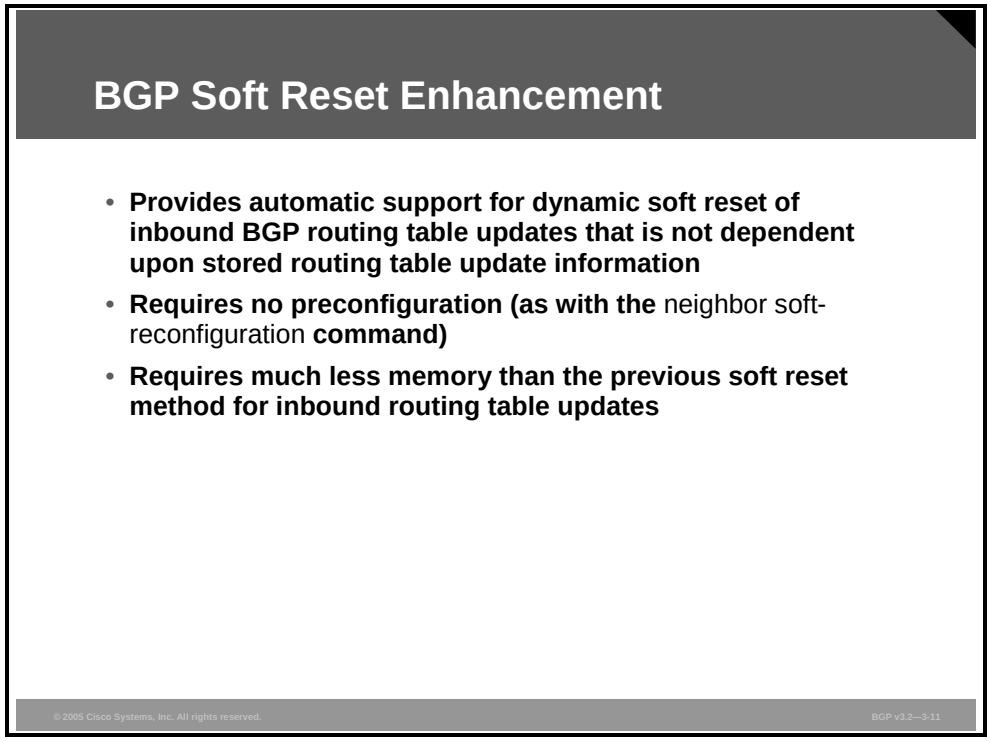
You can check incoming information that is received from a neighbor with the **show ip bgp neighbor** *ip-address* **routes** command. It displays which of the routes in the local BGP table were received (and accepted) from the indicated neighbor. Only routes that are passed by the incoming filter for the neighbor are displayed.

If the **soft-reconfiguration inbound** feature is enabled for a neighbor, the information that is saved in the extra copy outside the filters is displayed using the **show ip bgp neighbor** *ip-address* **received** command.

These commands are useful when you are troubleshooting the routing policy. You can compare routes outside the incoming filters with what was actually accepted into the BGP table from a neighbor. In addition, routes that are transmitted and advertised to a neighbor can be compared to what is inside the outgoing filters in the local BGP table.

BGP Soft Reset Enhancement

This topic describes the function of the BGP Soft Reset Enhancement feature.



Whenever there is a change in the routing policy, the BGP session must be cleared, or reset, for the new policy to take effect. There are two types of reset, hard reset and soft reset. Clearing a BGP session using a hard reset invalidates the cache and results in a negative impact on the operation of networks as the information in the cache becomes unavailable. Soft reset is recommended because it allows routing tables to be reconfigured and activated without clearing the BGP session. Soft reset is done on a per-neighbor basis. There are two types of soft reset:

- **Dynamic inbound soft reset:** When soft reset is used to generate inbound updates from a neighbor
- **Outbound soft reset:** When soft reset is used to send a new set of updates to a neighbor

Previously, to perform a soft reset for inbound routing table updates, the **neighbor soft-reconfiguration** command directed the Cisco IOS software in the local BGP router to store all received (inbound) routing policy updates without modification. This method is memory-intensive and not recommended unless absolutely necessary. (Outbound updates have never required the extra memory and are not affected by this feature.)

The BGP Soft Reset Enhancement feature, however, provides automatic support for dynamic soft reset of inbound BGP routing table updates that is not dependent on stored routing table update information. The new method requires no preconfiguration (as with the **neighbor soft-reconfiguration** command) and requires much less memory than the previous soft reset method for inbound routing table updates.

There are a number of benefits to the BGP Soft Reset Enhancement feature:

- **Allows dynamic route refresh requests:** This feature provides a way to initiate nondisruptive routing policy changes by allowing the dynamic exchange of route refresh requests between BGP routers, and the subsequent readvertisement of the respective outbound routing tables.
- **Requires no preconfiguration:** Because support for the soft reset using the route refresh capability is included in this release of the Cisco IOS software, no further router configuration is required. You can initiate a soft inbound reset using only the **clear ip bgp in** command.
- **Requires no additional memory resources:** Unlike a soft reset using the stored inbound routing table updates provided by the **neighbor soft-reconfiguration** command, when both BGP peers support the route refresh capability, inbound routing table updates are not stored in the local BGP router. The soft reset requests are exchanged dynamically, and no additional memory is required.
- **Provides flexibility:** There are now two available methods for inbound soft reset; the older method, using stored inbound routing table updates, and the method provided by this feature, using dynamic exchange of update information.

When the routing policy of a BGP neighbor changes, the session must be reset (cleared) for the changes to take effect. Because resetting a BGP session can be disruptive to networks, a soft reset method is recommended for reconfiguring the routing table. Previously, in order to reconfigure the inbound routing table, both the local BGP router and the BGP peer first needed to be configured to store incoming routing policy updates using the **neighbor soft-reconfiguration** command. Additional resources, particularly memory, were required to store the inbound routing table updates. The **clear ip bgp** command could then initiate the soft reset, which generated a new set of inbound routing table updates using the stored information.

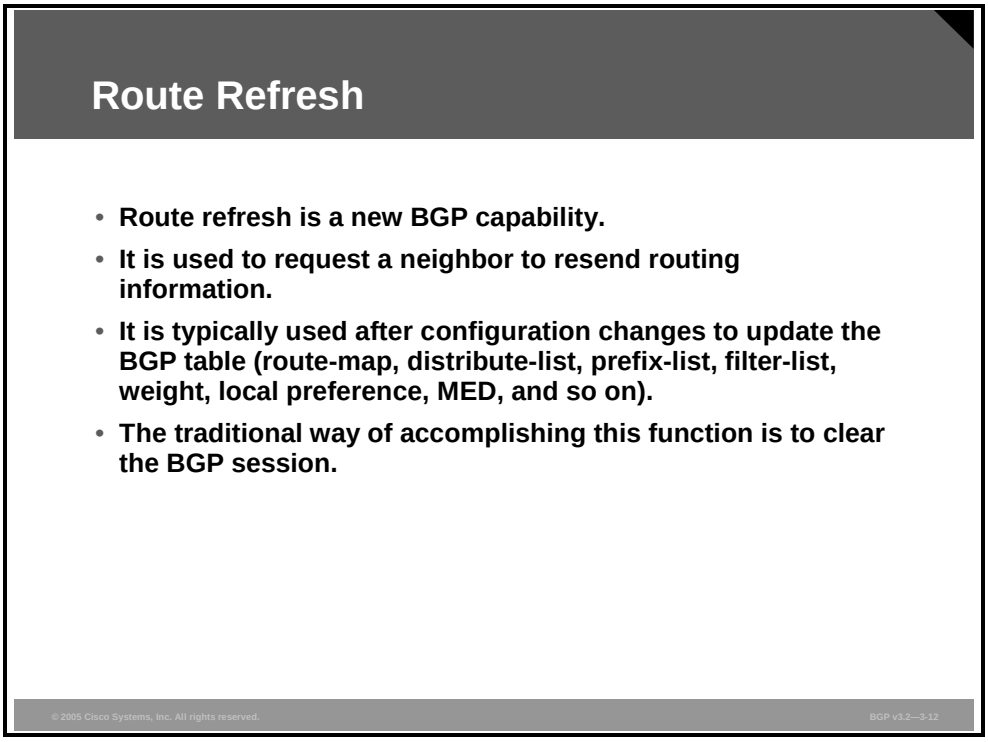
The BGP Soft Reset Enhancement feature provides an additional method for soft reset that allows the dynamic exchange of route refresh requests and routing information between BGP routers and the subsequent readvertisement of the respective outbound routing table. Soft reset using the route refresh capability does not require preconfiguration and consumes no additional memory resources.

To use this new method, both BGP peers must support the soft route refresh capability, which is advertised in the Open message sent when a peer sends its routing table update. Any router running BGP with this software release automatically supports the route refresh capability. Routers running earlier Cisco IOS software releases do not support the route refresh capability and must use the older soft reset method. If the soft reset fails, you can still clear the BGP session, but it will have a negative impact on network operations and should be used only as a last resort.

Note	Outbound resets have never required preconfiguration or storing of routing table updates and remain unchanged by the BGP Soft Reset Enhancement feature.

Route Refresh

This topic describes the function and benefits of the route refresh function.



Route refresh is one of the new capabilities of BGP. Routers use the route refresh feature to request a neighbor to resend all the routing information when it is needed.

There are several ways of refreshing the routing information from a neighbor:

- Clearing the neighbor relationship
- Soft-clearing the neighbor relationship (if soft reconfiguration is enabled for this specific neighbor)
- Using route refresh (if the neighbor supports this capability)

Note	To use soft reset without preconfiguration, both BGP peers must support the soft route refresh capability, which is advertised in the Open message that is sent when the peers establish a TCP session. Routers that run Cisco IOS software releases earlier than Release 12.1 do not support the route refresh capability and must clear the BGP session using the **neighbor soft-reconfiguration** command.

Route Refresh (Cont.)

- **Inbound soft reconfiguration consumes memory on the receiving router.**
  - **It is needed only because there is no mechanism in standard BGP to request retransmission of BGP routes.**
- **BGP route refresh is an optional BGP capability that allows a BGP router to request retransmission of BGP routes from a neighbor.**

The **soft-reconfiguration inbound** feature consumes large volumes of memory in the Internet environment. The number of routes that can be received from a peer router on the Internet is so large that it is not feasible to store an extra copy.

The only reason for making the extra copy is to be able to replay the data through the new routing policy without tearing down the session and re-establishing it.

What is needed is a mechanism to ask the neighbor router to do a “clear soft outbound.” If this were possible, the extra copy would not be needed. The neighboring router, of course, has its own copy in its BGP table, which it could resend to the local router whenever it is signaled to do so.

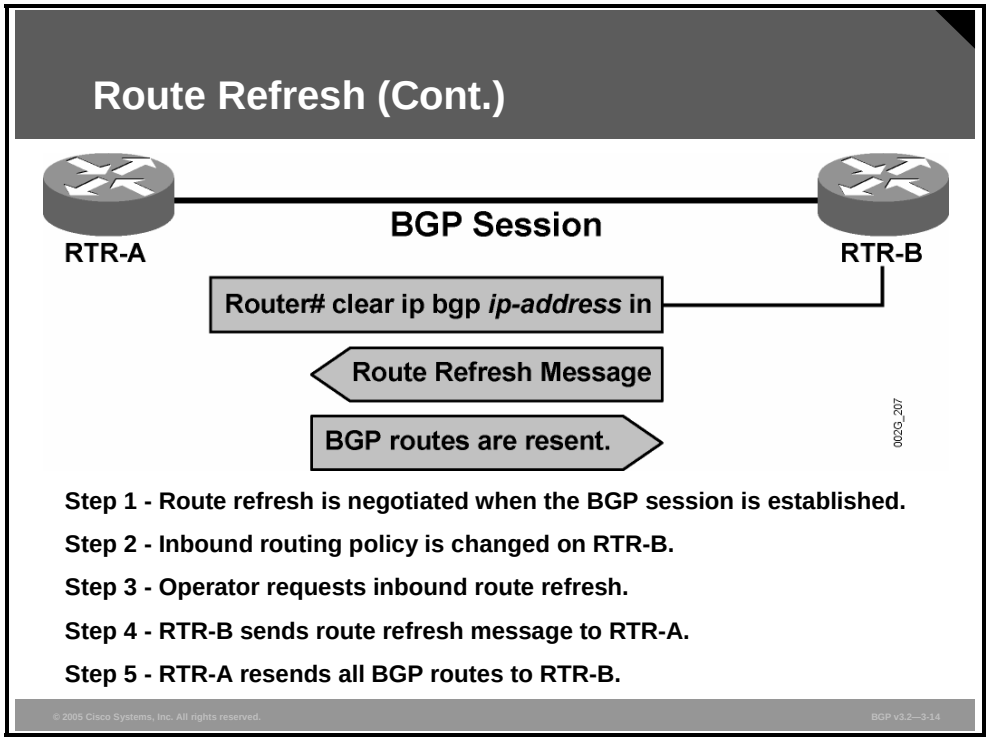
There is no such mechanism in standard BGP, but there is an optional BGP capability that allows one router to request a refresh from its neighbor: route refresh.

The table compares the various methods of BGP session reset, stating the advantages and disadvantages of each.

Type of Reset	Advantages	Disadvantages
Hard reset	No memory overhead.	The prefixes in the BGP, IP, and FIB tables that are provided by the neighbor are lost. Not recommended.
Outbound soft reset	No configuration, no storing of routing table updates.	Does not reset inbound routing table updates.
Dynamic inbound soft reset	Does not clear the BGP session or cache. Does not require storing of routing table updates, and has no memory overhead.	Both BGP routers must support the route refresh capability (Cisco IOS Software Release 12.1 and later releases).
Configured inbound soft reset (uses the **neighbor soft-reconfiguration** command)	Can be used when both BGP routers do not support the automatic route refresh capability.	Requires preconfiguration. Stores all received (inbound) routing policy updates without modification, and is thus memory-intensive. Recommended only when absolutely necessary.

Example: Route Refresh

The example shows the steps in a route refresh.

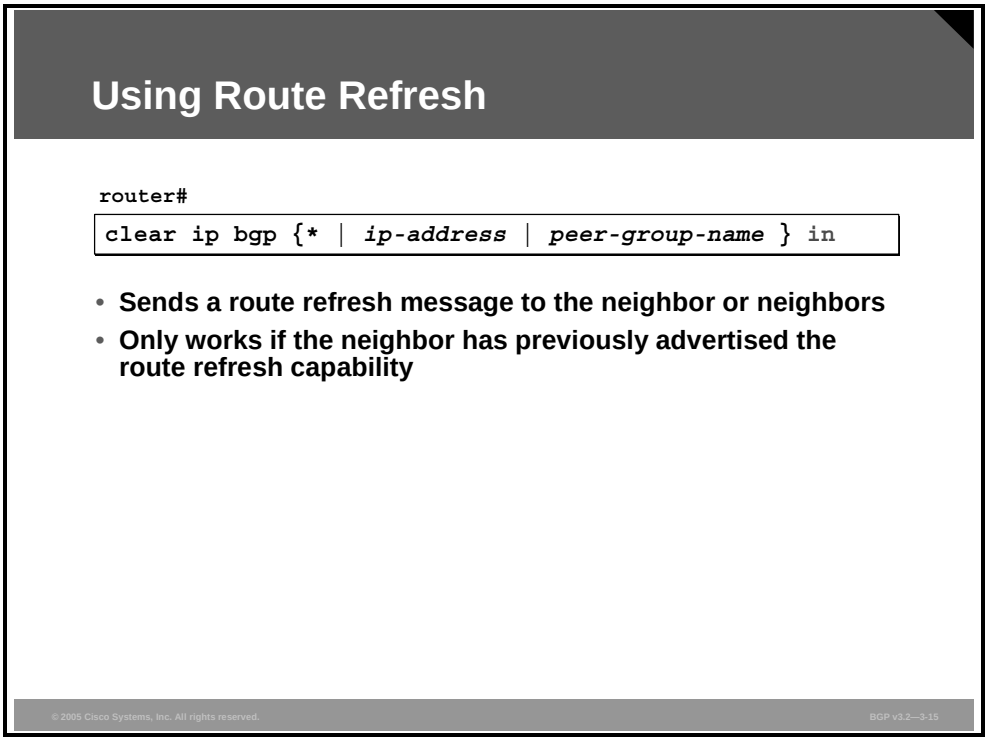


The ability to use the route refresh feature must be negotiated by the router when the BGP session is first established. The local router keeps a record that the capability is available with the neighbor. There is no need to keep a copy of the routing information that is received from the neighbor if it has the ability to refresh.

After reconfiguring the filters and route-maps that will implement a new routing policy, a network administrator can issue the **clear ip bgp** *ip-address* **soft in** command in the local router. The router checks whether the route refresh capability is available, and if it is, requests a resend of the BGP table of the neighbor instead of replaying its own copy.

Using Route Refresh

This topic identifies the Cisco IOS command that is required to perform a route refresh.



Use the **clear ip bgp * in** command to send a route refresh message to all neighbors or **clear ip bgp** *ip-address* **in** to send a route refresh message to a specific neighbor.

You need not use the **soft** keyword, because soft reset is automatically assumed when the route refresh capability is supported.

clear ip bgp

To reset a BGP connection with BGP soft reconfiguration, use the **clear ip bgp** privileged EXEC command at the system prompt.

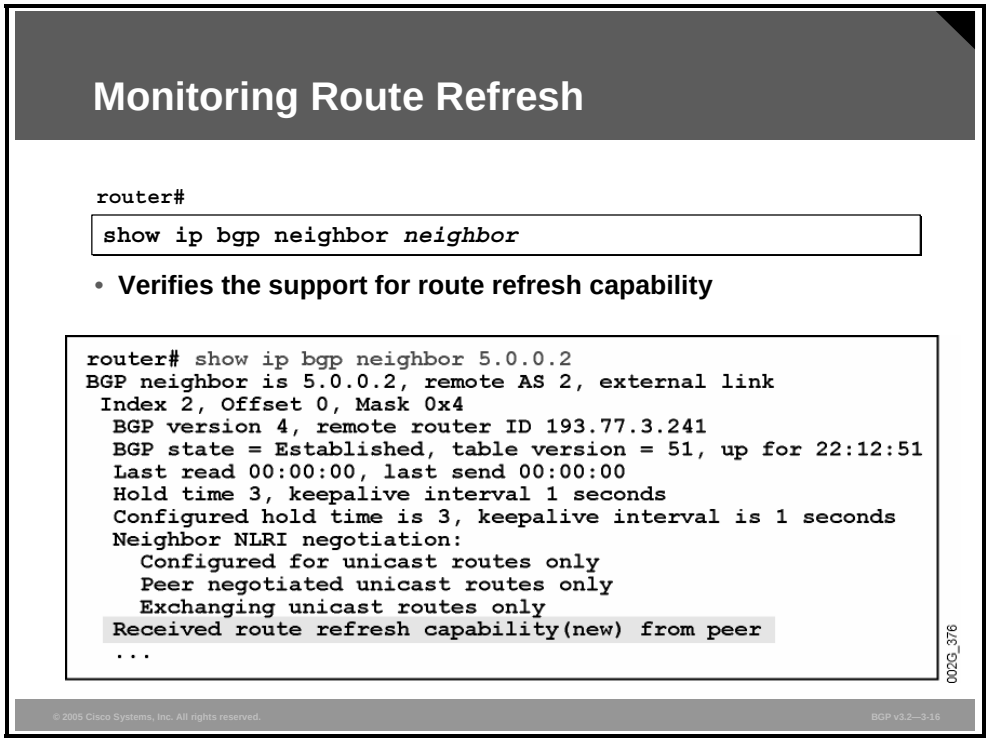
- **clear ip bgp** {*****| *ip-address* | *peer-group-name*} [**soft** [**in** | **out**]]

Syntax Description

Parameter	Description
*****	Resets all current BGP sessions.
ip-address	Resets only the identified BGP neighbor.
peer-group-name	Resets the specified BGP peer group.
soft	(Optional) Soft reset. Does not reset the session.
in | **out**	(Optional) Triggers inbound or outbound soft reconfiguration. If the **in** or **out** option is not specified, both inbound and outbound soft reset are triggered.

Monitoring Route Refresh

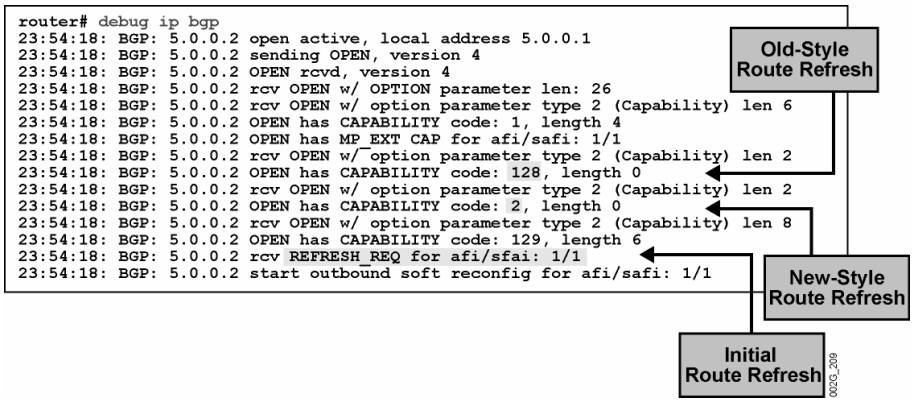
This topic identifies the Cisco IOS commands that are required to monitor route refresh operation.



Use the **show ip bgp neighbor** command to see whether the neighbor supports the route refresh message.

Note	The printout of the **show ip bgp neighbor** command varies among Cisco IOS releases. The printout in the figure here was generated by Cisco IOS Software Release 12.0(1)S and represents a manual configuration of soft reset.

Monitoring Route Refresh (Cont.)



Debug output after BGP session reset

Use **debug ip bgp** to display the negotiation of capabilities. Debugging displays received capabilities.

The example shows that a neighbor is advertising both old-style and standard (new-style) route refresh. After the session has been established, an initial standard route refresh message is sent by the router for the address family 1/1 (IP version 4 [IPv4] unicast).

Monitoring Route Refresh (Cont.)

```
router# debug ip bgp
router# debug ip bgp updates
router# clear ip bgp 5.0.0.2 in
1d00h: BGP: 5.0.0.2 sending REFRESH_REQ(5) for afi/safi: 1/1
1d00h: BGP: 5.0.0.2 rcv UPDATE w/ attr: nexthop 5.0.0.2, origin i, metric 0, path 2
1d00h: BGP: 5.0.0.2 rcv UPDATE about 10.0.0.0/8
1d00h: BGP: bumping version for 10.0.0.0/8 from 0 to 52
1d00h: BGP: nettable_walker 10.0.0.0/8 calling revise_route
1d00h: BGP: revise route installing 10.0.0.0/8 -> 5.0.0.2
1d00h: BGP: 5.0.0.2 computing updates, neighbor version 51, table version 52, starti
ng at 0.0.0.0
1d00h: BGP: 5.0.0.2 update run completed, ran for 0ms, neighbor version 51, start ve
rsion 52, throttled to 52, check point net 0.0.0.0
1d00h: BGP: 3.0.0.2 computing updates, neighbor version 51, table version 52, starti
ng at 0.0.0.0
1d00h: BGP: 3.0.0.2 send UPDATE 10.0.0.0/8, next 3.0.0.1
1d00h: BGP: , metric 0, path 1 2
1d00h: BGP: 3.0.0.2 1 updates enqueued (average=45, maximum=45)
1d00h: BGP: 3.0.0.2 update run completed, ran for 0ms, neighbor version 51, start ve
rsion 52, throttled to 52, check point net 0.0.0.0
```

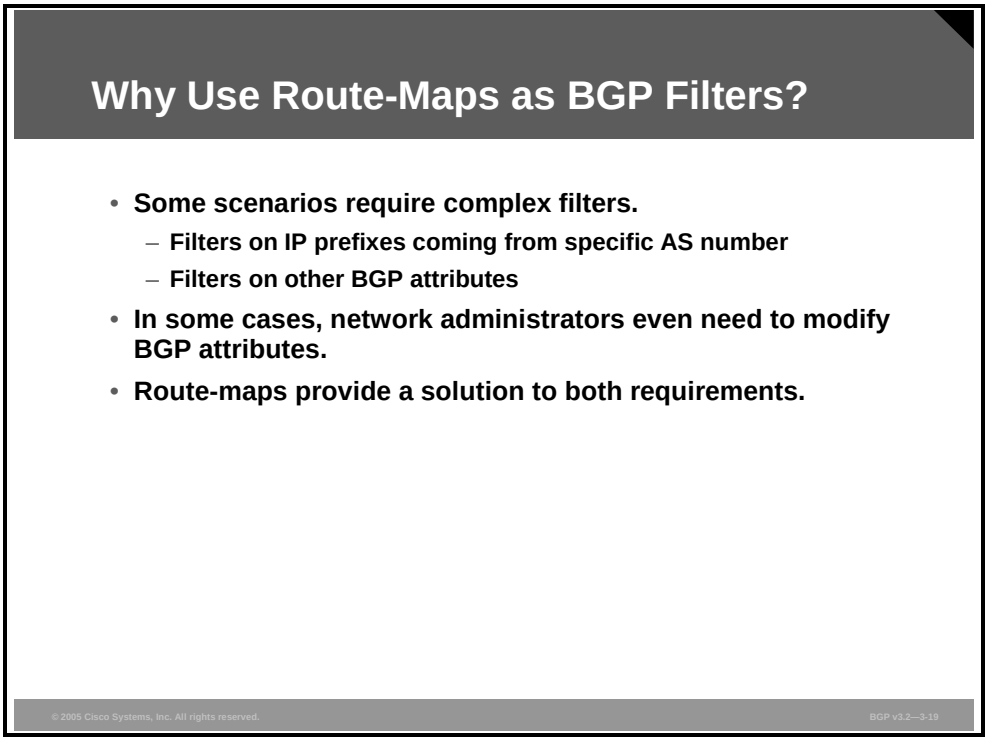
002G_210

Debug output after route refresh

Debugging also shows a route refresh message being sent to a neighbor after the network administrator issues the **clear ip bgp** *ip-address* **in** command from privileged EXEC mode.

Why Use Route-Maps as BGP Filters?

This topic identifies the need to use route-maps to influence route selection in a BGP network.



Network administrators cannot achieve certain complex filtering goals by using a prefix-list only or by using an AS-path filter-list only. Using both of these filters simultaneously means that a route must be permitted by both to be accepted. Sometimes the goal is to permit a specific prefix if it is received with a specific AS-path and to deny it otherwise.

Combinations of tests can be implemented using route-maps. A route-map is a powerful filtering tool that can also modify routing information. Different attributes can have their values set or changed by the route-map.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Because of the huge volumes of routing information that BGP is capable of handling and the effects of a mass routing update, BGP cannot use traditional routing update methods.**
- **Soft reconfiguration provides the possibility to run all routes through filters without tearing down the sessions.**
- **The Cisco IOS commands that are required to configure and perform a soft reconfiguration include the** neighbor soft-reconfiguration **router configuration command, which configures Cisco IOS software to start storing updates and the** clear ip bgp **EXEC command, which resets a BGP connection using BGP soft reconfiguration.**

Summary (Cont.)

- **The Cisco IOS tools that are available to monitor the operation of a soft reconfiguration include the** show ip bgp **command, which displays the local BGP table, the *show ip bgp neighbor ip-address routes* command, which checks incoming information that is received from a neighbor, and the** show ip bgp neighbor *ip-address* received **command, which displays the information that is saved in the extra copy outside the filters.**
- **The BGP Soft Reset Enhancement feature provides automatic support for dynamic soft reset of inbound BGP routing table updates that is not dependent upon stored routing table update information. This method requires no preconfiguration and needs much less memory than the previous soft reset method for inbound routing table updates.**

Summary (Cont.)

- **Route refresh is a new BGP capability that is used to request a neighbor to resend routing information after configuration changes.**
- **The** clear ip bgp *ip-address* soft in **command sends a route refresh message to the neighboring router and executes if the neighbor has previously advertised the route refresh capability.**
- **To verify that a neighbor supports route refresh, you can use the** show ip bgp neighbor **command. To display the negotiation process, you can use the** debug ip bgp **command.**
- **Network administrators cannot achieve certain complex filtering goals by using a prefix-list only or by using an AS-path filter list only. A route-map is a powerful filtering tool that can also modify routing information.**

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.**
- **An AS-path filter is created by an AS-path access-list. The access-list is applied to a set of routes from which to subset can be selected.**
- **Use prefix-lists to filter incoming or outgoing BGP updates to neighbors and to filter routes that are being redistributed into the BGP process from other routing protocols.**

Module Summary (Cont.)

- **Outbound route filtering is a mechanism that is used to minimize the number of updates that are requested from a neighbor.**
- **Route-maps provide a method to perform a variety of compound, complex filtering operations (such as dropping denied routes and modifying attributes of the permitted routes) within a single tool.**
- **Soft reconfiguration provides the ability to run all routes through filters without tearing down the sessions.**

This module discussed BGP route filtering and BGP route selection policies. The module described multihomed BGP networks and identified the need for BGP route selection. This module also addressed configuring BGP to influence route selection by using AS-path filters, prefix-list filters, and route-maps. Outbound route filtering was also explained. In addition, details about soft reconfiguration and route refresh were provided.

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Sample Configuration for BGP with Two Different Service Providers (Multihoming)*. http://www.cisco.com/warp/public/459/27.html.
- Cisco Systems, Inc. *Using Regular Expressions in BGP*. http://www.cisco.com/warp/public/459/26.html.
- Cisco Systems, Inc. *BGP Case Studies* (see “BGP Case Studies 3” section). http://www.cisco.com/warp/public/459/bgp-toc.html.
- Cisco Systems, Inc. *Configuring BGP*. http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cbgp.htm#xtocid15.
- Cisco Systems, Inc. *BGP Prefix-Based Outbound Route Filtering*. http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t11/ft11borf.htm.
- Internet Engineering Task Force (IETF) Network Working Group. *Cooperative Route Filtering Capability for BGP-4*. http://www.ietf.org/internet-drafts/draft-ietf-idr-route-filter-12.txt
- Cisco Systems, Inc. *BGP Case Studies* (see “BGP Case Studies 1” section). http://www.cisco.com/warp/public/459/bgp-toc.html#routemaps.
- Cisco Systems, Inc. *Compatible Systems Setup Guides: BGP Configuration Guide*. http://www.cisco.com/warp/public/707/cscsupport/setup_guides/bgp.html#bgpRouteMap.
- Cisco Systems, Inc. *BGP Soft Reset Enhancement*. http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t7/sftrst.htm.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) What are two reasons why a customer would want to connect to two ISPs? (Choose two.) (Source: Using Multihomed BGP Networks)

A) to expand capacity for Internet traffic
B) to better protect confidential information as it travels through the Internet
C) to provide redundancy to mission-critical services that are offered over the Internet
D) to efficiently route Internet traffic to two different divisions within the company

Q2) What are the two technical requirements for multihomed customers? (Choose two.) (Source: Using Multihomed BGP Networks)

A) The ISPs must assign a range of IP network numbers to the customer.
B) The customer network must exchange BGP information with each ISP network.
C) In most cases, the customer must have its own public AS number.
D) The customer network must not use AS-path filters.

Q3) Which of the following statements best illustrates the importance of BGP policies that influence route selection in a multihomed BGP network? (Source: Using Multihomed BGP Networks)

A) The default BGP route selection does not always result in optimum routing.
B) The default BGP route selection always results in optimum routing.
C) After the route selection behavior has been set, it cannot be changed.
D) The customer receives all routes from both service providers, giving redundancy; therefore, BGP policies are not necessary.

Q4) Which three of the following are potential customer routing policies? (Choose three.) (Source: Using Multihomed BGP Networks)

A) One service provider is designated as primary, and the other is a backup.
B) Traffic is load-balanced across both ISP networks.
C) Traffic toward a specific destination goes through only one of the ISPs.
D) Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary ISP.
E) The two ISPs may have similar peering agreements with other ISPs.
F) The ISPs use default routing to the customer, and the customer uses static routing to the ISP (or ISPs).

Q5) Which statement about the need to influence BGP route selection in a service provider environment is accurate? (Source: Using Multihomed BGP Networks)

A) If both ISP connections terminate in one single customer router, only some routes that are received from the primary ISP can be assigned a BGP weight.
B) In most cases, it is more optimal to reach other customers connected to the backup ISP via the backup link, compared with reaching them via the primary link.
C) All routes that are received from the primary ISP over the primary link are assigned a local preference value, which is lower than the default value of 100.
D) When one ISP connection terminates in one single customer router, all routes that are received from that ISP are assigned a BGP weight.

Q6) Which two potential multihomed network issues can be prevented with IP prefix filters? (Choose two.) (Source: Using Multihomed BGP Networks)

A) the propagation of private AS numbers
B) the propagation of private addresses that are used in the network
C) the propagation of unreachable next-hop addresses
D) the propagation of more specific prefixes from an address range

Q7) Which three goals represent appropriate reasons to apply AS-path filters? (Choose three.) (Source: Employing AS-Path Filters)

A) to ensure that only locally originated routes are announced
B) to limit routes that are advertised from IBGP neighbors
C) to select a subset of all routes based on their originating AS
D) to limit neighbor route updates to specific AS-originated routes
E) to ensure that all destination autonomous systems should be received from a specified neighbor
F) to change the weight or local preference attributes for all destination autonomous systems

Q8) Which AS path is matched by the regular expression “72$”? (Source: Employing AS-Path Filters)

A) 213 72 218 31 727
B) 27 317 271 50 72
C) 315 27 723 19 91
D) 72 591 368 20 87

Q9) What is the difference between the regular expressions “_100_” and “_100$”? (Source: Employing AS-Path Filters)

A) The first expression refers to routes that have the substring “100” in their AS paths; the second expression refers only to routes that are directly connected to AS 100.
B) The first expression refers to routes that have the substring “100” in their AS paths; the second expression refers only to routes that originated in AS 100.
C) The first expression refers to routes that go through AS 100; the second expression refers to routes that originated in AS 100.
D) The first expression refers to routes that are directly connected to AS 100; the second expression refers to routes that originated in AS 100.

Q10) How do you match AS paths that contain exactly two single-digit AS numbers? (Source: Employing AS-Path Filters)

A) use the expression "**"
B) use the expression ".."
C) use the expression "[0-9]_[0-9]"
D) use the expression "^[0-9]_[0-9]$"

Q11) Which three steps are required to apply a new inbound routing policy to a neighbor? (Choose three.) (Source: Employing AS-Path Filters)

A) Define an AS-path access-list.
B) Attach the AS-path filter to inbound or outbound updates for a specific BGP neighbor.
C) Send incoming and outgoing AS-path filters to the BGP neighbor.
D) Force the updates to go through the new filter.

Q12) How can you test your regular expression? (Source: Employing AS-Path Filters)

A) **show ip bgp access-list** command
B) **show ip bgp filter** command
C) **show ip bgp regexp** command
D) **show ip bgp summary** command

Q13) What are two reasons that a multihomed customer needs prefix-lists? (Choose two.) (Source: Filtering with Prefix-Lists)

A) to ensure that only valid IP prefixes are announced to the ISPs
B) to set a limit on the number of prefixes that can be accepted from the ISPs
C) to prevent the customer from receiving its own IP prefixes from the ISP
D) to verify that the customer has received full Internet route tables

Q14) Which three of the following choices are advantages of prefix-lists over access-lists? (Choose three.) (Source: Filtering with Prefix-Lists)

A) significant performance improvement on long filters
B) support for incremental updates
C) ability to consist of any number of lines, each of which indicates a test and a result
D) flexibility
E) more complex command-line interface
F) sequential scanning of prefix-lists within Cisco IOS software

Q15) When you define prefix-lists, what are two reasons to use sequence numbers? (Choose two.) (Source: Filtering with Prefix-Lists)

A) to reference the associated access-list for the prefix-list entry
B) to provide a means of linking an AS-path filter-list to the prefix-list
C) to provide an execution order for prefix-list entries
D) to provide a means of inserting or deleting list entries

Q16) Which of the following statements is accurate about the **ge** and **le** parameters in the **ip prefix-list** global configuration command required to configure prefix-list filters? (Source: Filtering with Prefix-Lists)

A) No match is assumed when neither **ge** nor **le** is specified.
B) The range is assumed to be from *ge-value* to 142 only if the **ge** attribute is specified.
C) The range is assumed to be from *len* to 32 only if the **le** attribute is specified.
D) An exact match is assumed when neither **ge** nor **le** is specified.

Q17) Which of the following statements about implementing prefix-lists in a BGP network is accurate? (Source: Filtering with Prefix-Lists)

A) You can optionally apply filter-lists and prefix-lists on either incoming or outgoing neighbors in any combination.
B) You can optionally apply filter-lists and prefix-lists only on outgoing neighbors in any combination.
C) Either the incoming prefix-list or the incoming filter-list must permit the routes that are received from a neighbor before they are accepted into the BGP table.
D) Outgoing routes must pass the outgoing prefix-list before being transmitted to the neighbor.

Q18) How can you apply the same prefix-list to multiple BGP neighbors on a router? (Source: Filtering with Prefix-Lists)

A) by configuring a neighbor prefix-list statement for each BGP peer
B) by configuring a neighbor distribute-list statement for each neighbor
C) by using the BGP *peer-group* option with the neighbor statement
D) by configuring the prefix-list as a global filter under the BGP routing process

Q19) How can you use the **show ip prefix-list** command to display the prefix-list entry that matches a specific prefix and length? (Source: Filtering with Prefix-Lists)

A) not a feature of the **show ip prefix-list** command
B) by specifying the **detail** keyword
C) with the longer keyword to display all matches except those with more specific entries
D) by specifying the **first-match** keyword

Q20) Which of the following best describes the capabilities of the proprietary ORF type that is supported on Cisco routers? (Source: Using Outbound Route Filtering)

A) standard BGP communities filtering
B) extended BGP communities filtering
C) AS-path filtering
D) prefix-list filtering

Q21) What are two key benefits to using outbound route filtering? (Choose two.) (Source: Using Outbound Route Filtering)

A) conserves CPU cycles
B) improves security
C) reduces bandwidth that is used by unnecessary routing updates
D) increases neighbor availability

Q22) Which two of the following statements about BGP prefix-based outbound route filtering are accurate? (Choose two.) (Source: Using Outbound Route Filtering)

A) uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers
B) can limit the number of unwanted routing updates
C) increases the amount of resources required to receive and discard routes that would otherwise be filtered out
D) can be used to increase the amount of processing on a router that is not accepting full routes from a service provider network

Q23) How should you configure the **neighbor capability orf prefix-list** command on a router that is applying a prefix-list filter as an outbound route policy? (Source: Using Outbound Route Filtering)

A) **send**
B) **receive**
C) **both**
D) **prefix-filter**

Q24) What are two methods of determining that a router has ORF capabilities exchange configured? (Choose two.) (Source: Using Outbound Route Filtering)

A) **show running-config | begin bgp** command
B) **show ip bgp negotiate** command
C) **show ip bgp neighbors** command
D) **show ip prefix-list** command

Q25) What are two prerequisites before you can configure ORF prefix-list functionality? (Choose two.) (Source: Using Outbound Route Filtering)

A) A route refresh must be sent using the **clear ip bgp** command.
B) A BGP peering session between the ORF routers must be up and running.
C) ORF capabilities must be enabled on both routers.
D) You must configure a prefix-list filter on the receiving router.

Q26) Which of the following statements about the function of a route-map is accurate? (Source: Applying Route-Maps as BGP Filters)

A) A route-map cannot be used to modify attributes of the permitted routes.
B) A route-map is a filter that uses a series of match conditions, and that which is denied by the route-map is dropped.
C) A route-map is less complex than the access-list.
D) Each route-map statement starts with a series of match clauses.

Q27) Which three of the following statements are accurate about the BGP Route-Map Policy List Support feature? (Choose three.) (Source: Applying Route-Maps as BGP Filters)

A) The BGP Route-Map Policy List Support feature allows a network operator to group route-map match clauses into named lists called policy-lists.
B) The network operator manually reconfigures each recurring group of match clauses that occur in multiple route-map entries.
C) The AND and OR semantics in the route-map function differently for policy-lists than for match and set clauses.
D) To create a BGP policy-list, use the **ip policy-list** command in policy-map configuration mode.
E) To configure a route-map to evaluate and process a BGP policy-list in a route-map, use the **match policy-list** command in route-map configuration mode.
F) To display information about a configured policy-list and policy-list entries, use the **show ip policy-list** command in route-map configuration mode.

Q28) Which two of the following are functions of the BGP Route-Map Continue feature? (Choose two.) (Source: Applying Route-Maps as BGP Filters)

A) provides the ability to pause if a sequence number is not specified
B) provides the capability to execute additional entries in a route-map after an entry is executed with successful match and set clauses
C) allows modularization of network policy configuration so that repeated policy definitions can be expanded within the same route-map
D) allows configuration and organization of more modular policy definitions to reduce the number of policy configurations that are repeated within the same route-map

Q29) Which of the following commands is used to distribute any routes that have a destination network number address that is permitted by a standard access-list, an extended access-list, or a prefix-list, or to perform policy routing on packets? (Source: Applying Route-Maps as BGP Filters)

A) **match ip next-hop**
B) **match ip route-source**
C) **match ip address**
D) **show ip bgp route-map**

Q30) How do you implement a “permit all” statement when you are using route-maps? (Source: Applying Route-Maps as BGP Filters)

A) By default, a route-map has an “implicit permit any” statement if no match is found.
B) You must configure a route-map with a “permit” parameter and no match clause.
C) You must configure a route-map with a “deny” parameter and a “deny none” clause.
D) You must configure a route-map with a “permit any” match clause.

Q31) What happens to incoming BGP updates that do not match any route-map match clauses? (Source: Applying Route-Maps as BGP Filters)

A) They are entered into the BGP table.
B) They are entered into the BGP table and marked with a weight of 32768.
C) They are not accepted by the router or entered into the BGP table.
D) They are entered into the BGP table if a matching route exists in the IP routing table.

Q32) Which three BGP attributes can you set using route-maps? (Choose three.) (Source: Applying Route-Maps as BGP Filters)

A) MED
B) path origin
C) administrative distance
D) weight metric
E) next-hop
F) atomic aggregate

Q33) What are two reasons for using route-map sequence numbers? (Choose two.) (Source: Applying Route-Maps as BGP Filters)

A) to allow insertion or deletion of route-map entries
B) to order the execution sequence of route-map match clauses
C) to provide an ordered execution sequence for the route-map
D) to map between prefix-list statements and route-map match clauses

Q34) Why is clearing a BGP session a disruptive change in routing policy? (Source: Implementing Changes in BGP Policy)

A) Clearing a BGP session takes a long time and can disrupt packet forwarding.
B) You cannot recover information that is sent while the BGP session is being cleared.
C) You cannot automatically re-establish sessions that are torn down during the clearing operation.
D) You cannot selectively tear down BGP sessions; you must clear sessions with all neighbors.

Q35) What is the impact of inbound soft reconfiguration? (Source: Implementing Changes in BGP Policy)

A) It clears the session after you reconfigure the new routing policy.
B) It creates a copy of all routes that are received from a neighbor after the filters are applied.
C) It requires extra memory to hold a copy of all routes that are received from the neighbor.
D) It resets the table version number of the neighbor to 0.

Q36) Which two steps must you complete to use inbound soft configuration functionality? (Choose two.) (Source: Implementing Changes in BGP Policy)

A) Clear the BGP session inbound on the local router.
B) Clear the BGP session outbound on the remote router.
C) Configure the local neighbor with the **soft-reconfiguration in** command.
D) Configure the remote neighbor with the **soft-reconfiguration out** command.

Q37) Match the functions to the tools used to monitor the operation of a soft reconfiguration. (Source: Implementing Changes in BGP Policy)

A) display which of the routes in the local BGP table were received (and accepted) from the indicated neighbor
B) check the entries that have been propagated to a specific neighbor
C) display the information that is saved in the extra copy outside the filters

_____ 1. **show ip bgp neighbor** *ip-address* **received** command

_____ 2. **show ip bgp neighbor** *ip-address* **advertised** command

_____ 3. **show ip bgp neighbor** *ip-address* **routes** command

Q38) Which two of the following are functions of the BGP Soft Reset Enhancement feature? (Choose two.) (Source: Implementing Changes in BGP Policy)

A) allows dynamic route refresh requests
B) requires no preconfiguration
C) provides newer method for inbound soft reset that uses stored inbound routing table updates
D) uses expanded memory

Q39) Which of the following statements about the command that is required to perform a route refresh is accurate? (Source: Implementing Changes in BGP Policy)

A) You will use the **clear ip bgp * in** command to send a route refresh message to all neighbors.
B) You will use the **clear ip bgp** *ip-address* **in command** to send a route refresh message to all neighbors.
C) You must use the **soft** keyword with the **clear ip bgp** command because soft reset is not automatically assumed when the route refresh capability is supported.
D) The **clear ip bgp** command works even if the neighbor has not previously advertised the route refresh capability.

Q40) What are two situations in which you would prefer inbound soft reconfiguration to route refresh? (Choose two.) (Source: Implementing Changes in BGP Policy)

A) when there is insufficient memory to hold a copy of the BGP table of the neighbor
B) when a route refresh fails
C) when you wish to troubleshoot filters and use the **show ip bgp neighbor** command with the **received-routes** option
D) when the neighboring router does not support the route refresh capability

Q41) How do you determine whether a BGP neighbor supports route refresh? (Source: Implementing Changes in BGP Policy)

A) A flag in the BGP table indicates the presence of route refresh capability.
B) The **show ip bgp neighbor** command indicates whether the option is supported.
C) Initiate the **debug ip bgp negotiation** command to see whether the router has completed a route refresh capabilities exchange.
D) Execute the **clear ip bgp** * command. Command-line BGP status messages will indicate route refresh support capabilities.

Q42) Which of the following statements about using route-maps as filters is accurate? (Source: Implementing Changes in BGP Policy)

A) Network administrators can achieve certain complex filtering goals by using a prefix-list only or by using an AS-path filter list only; therefore, route-maps are not necessarily a good solution.
B) Combinations of tests cannot be implemented using route-maps; therefore, in some cases, an AS-path filter list is preferred.
C) Route-maps have less capability for filtering than access-lists.
D) Route-maps provide both complex filters and a way to modify BGP attributes.

Module Self-Check Answer Key

Q1) A, C
Q2) B, C
Q3) A
Q4) A, C, D
Q5) B
Q6) B, D
Q7) A, C, D
Q8) B
Q9) C
Q10) D
Q11) A, B, D
Q12) C
Q13) A, C
Q14) A, B, D
Q15) C, D
Q16) D
Q17) A
Q18) C
Q19) D
Q20) D
Q21) A, C
Q22) A, B
Q23) B
Q24) A, C
Q25) B, C
Q26) B
Q27) A, D, E
Q28) B, D
Q29) C
Q30) B
Q31) C
Q32) A, D, E
Q33) A, C
Q34) A
Q35) C
Q36) A, C
Q37) 1-C
2-B
3-A
Q38) A, B
Q39) A
Q40) C, D
Q41) B
Q42) D

