BGP

Configuring BGP on Cisco Routers

Volume 2

Version 3.0

Student Guide

Text Part Number: 97-1428-01

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

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica Croatia • 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

Table of Contents

Volume 1

Volume 2

Route Selection Using Attributes

Overview

Routes learned via the Border Gateway Protocol (BGP) have properties associated with them that aid a router in determining the best route to a destination when multiple paths to that particular destination exist. These properties are referred to as BGP attributes. This module introduces the role of BGP attributes, and how their presence influences route selection in BGP. Understanding how BGP attributes influence route selection is required for the design of robust networks.

This module provides advanced information on how to connect Internet customers to multiple service providers. It includes an in-depth description of BGP attributes used in route selection, including weight, local preference, AS-path prepending, multi-exit discriminator (MED), and BGP communities.

Upon completing this module, you will be able to:

- List BGP path attributes and the functionality of each attribute.
- Successfully configure BGP to influence route selection using the weight attribute, given a customer scenario where you must support multiple connections.
- Use the local preference attribute to influence route selection, given a customer scenario where you must support multiple connections.
- Use AS-path prepending to influence the return path selected by the neighboring autonomous systems, given a customer scenario where you must support multiple connections.

- Use the MED attribute to influence route selection, given a customer scenario where you must support multiple connections.

- Use BGP community attributes to influence route selection, given a customer scenario where you must support multiple connections.

Outline

The module contains these lessons:

- BGP Path Attributes
- Influencing BGP Route Selection with Weights
- BGP Local Preference
- AS-Path Prepending
- BGP Multi-Exit Discriminator
- BGP Communities

BGP Path Attributes

Overview

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

Importance

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

Objectives

Upon completing this lesson, you will be able to:

- Describe the purpose 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

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

This lesson includes these topics:

- Overview
- BGP Path Attributes
- Well-Known BGP Attributes
- Optional BGP Attributes
- AS-Path Attribute
- Next-Hop Attribute
- Summary
- Assessment (Quiz): BGP Path Attributes

BGP Path Attributes

This topic describes the concept of BGP path attributes.

BGP Path Attributes

Cisco.com

- BGP metrics are called path attributes
- BGP attributes are categorized as well-known and optional
- Well-known attributes must be recognized by all compliant implementations
- Optional attributes are only recognized by some implementations (could be private), expected not to be recognized by everyone

 BGP v3.0—3-4

Each BGP update consists of one or more IP subnets and a set of attributes 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 specified in a later extension of BGP or even private vendor extensions not documented in a standard document.

Practice

Q1) What do network administrators use in BGP to define the metrics used for best route selection?

A) link states

B) path attributes

C) distance vectors

D) cost

Well-Known BGP Attributes

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

Well-Known BGP Attributes

Cisco.com

Well-known attributes are divided into mandatory and discretionary

- **Mandatory well-known attributes must be present in all update messages**
- **Discretionary well-known attributes are optional, they could be present in update messages**
- **All well-known attributes are propagated to other neighbors**

 BGP v3.0—3-6

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 will analyze the attached attributes and compare them with the attributes attached to the same IP subnet when 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.

Well-Known BGP Attributes (Cont.)

Cisco.com

Mandatory Well-Known Attributes

- Origin
  - Specifies the origin of a BGP route
    - IGP   Route originated in an IGP
    - EGP   Route originated in EGP
    - Unknown Route was redistributed into BGP
- AS-path
  - Sequence of AS numbers through which the network is accessible
- Next-hop
  - IP address of the next-hop router

 BGP v3.0—3-6

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

- 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 "IGP." If information about an IP subnet is injected using redistribution, the origin attribute is set to "unknown" or "incomplete" (these two words have the same meaning). The origin code, "EGP," was used when the Internet was migrating from exterior gateway protocol (EGP) to BGP and is now obsolete.

- The egress router modifies the AS-path attribute every time information about a particular IP subnet passes over an autonomous system (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.

- The router also modifies the next-hop attribute as the route passes through the network. It indicates the IP address of the next-hop router—the router to which the receiving router should forward the IP packets toward the destination advertised in the routing update.

Well-Known BGP Attributes (Cont.)

Cisco.com

Discretionary Well-Known Attributes

- **Local preference**
  - **Used for consistent routing policy within AS**
- **Atomic aggregate**
  - **Informs the neighbor AS that the originating router aggregated routes**

BGP v3.0—3-7

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 their functions are required.

- 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 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.

- The atomic aggregate attribute is attached to a route that is created as a result of route summarization (called aggregation in BGP). It signals that information that was present in the original routing updates may have been lost when the updates were summarized into a single entry.

Practice

Q1) Which three BGP path attributes must be carried with each update? (Choose three.)

A) origin

B) AS-path

C) local preference

D) next-hop

Q2) Which two attributes are discretionary well-known BGP attributes? (Choose two.)

A) multi-exit discriminator

B) local preference

C) origin

D) atomic aggregate

Optional BGP Attributes

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

Optional BGP Attributes

Cisco.com

Optional BGP attributes are transitive or nontransitive

Transitive optional attributes

- **Propagated to other neighbors if not recognized; partial bit set to indicate that the attribute was not recognized**

Nontransitive optional attributes

- **Discarded if not recognized**

Recognized optional attributes are propagated to other neighbors based on their meaning (not constrained by transitive bit)

BGP v3.0—3-8

When a router receives an update that contains an optional attribute, the router checks if its implementation recognizes the particular attribute. If it does, then the router should know how to handle it and whether to propagate it or not.

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 will set 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 some router in the path does not recognize them. Routers that do not recognize these attributes will drop them.

Optional BGP Attributes (Cont.)

Cisco.com

Nontransitive attributes

- **Multi-Exit Discriminator**
  - **Used to discriminate between multiple entry points to a single autonomous system**

Transitive attributes

- **Aggregator**
  - **Specifies IP address and AS number of the router that performed route aggregation**
- **Community**
  - **Used for route tagging**

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:

- **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.

Practice

Q1) How are recognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q2) How are nonrecognized transitive optional attributes propagated between BGP neighbors?

A) Optional attributes are converted to transitive well-known attributes.

B) With the partial bit set.

C) Based on their meaning.

D) Optional attributes are not propagated to neighbors.

Q3) Which is a nontransitive optional BGP path attribute?

A) local preference

B) weight

C) MED

D) community

AS-Path Attribute

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

AS-Path Attribute

Cisco.com

- **The AS-path attribute is empty when a local route is inserted in the BGP table**
- **The AS number of the sender is prepended to the AS-path attribute when the routing update crosses AS boundary**
- **The receiver of BGP routing information can use the AS-path to determine through which AS the information has passed**
- **An AS that receives routing information with its own AS number in the AS-path silently ignores the information**

 BGP v3.0—3-10

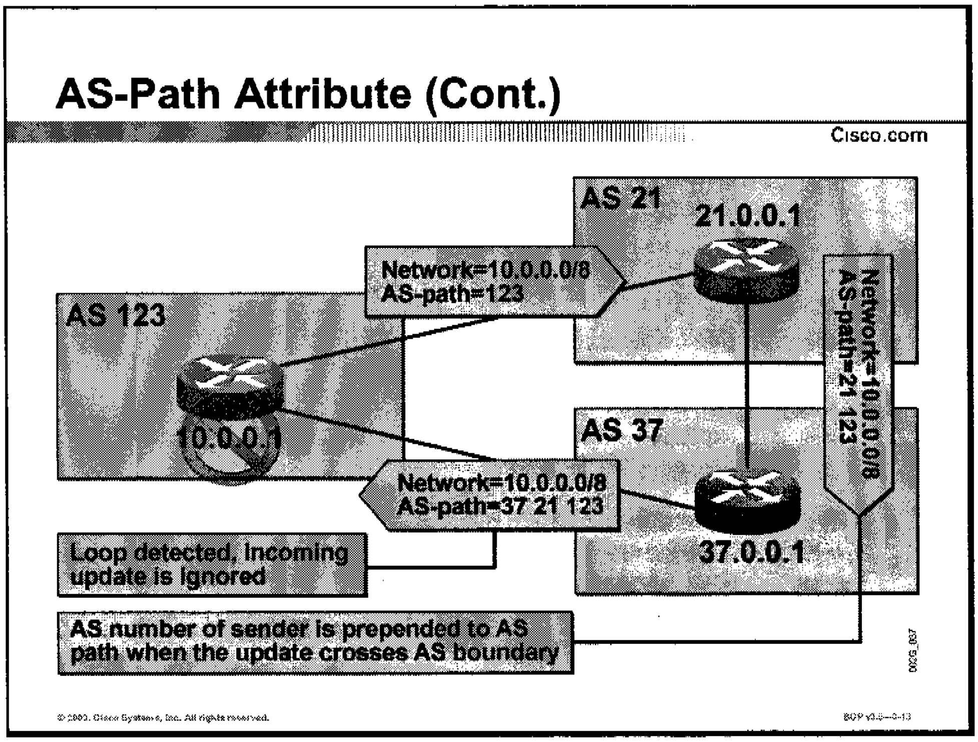
The AS-path attribute is modified by an edge router every time information about a particular IP subnet passes over an AS border. When a router first originates a route in BGP, the AS-path attribute is empty. The local AS number is prepended to the AS path each time the route crosses an AS boundary. 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 in order to reach the indicated network. The AS that originally injected the route into BGP is always found in the rightmost end of the AS path.

- It is easy to distinguish local routes from routes 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 will check the AS-path attribute and look for its own AS number. If it 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



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, 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 will 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. As 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 where the router in AS 123 drops 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 used by BGP to avoid a routing information loop.

Practice

Q1) How do BGP routers detect routing loops?

A) BGP routers check for the longest AS path in all routing updates.

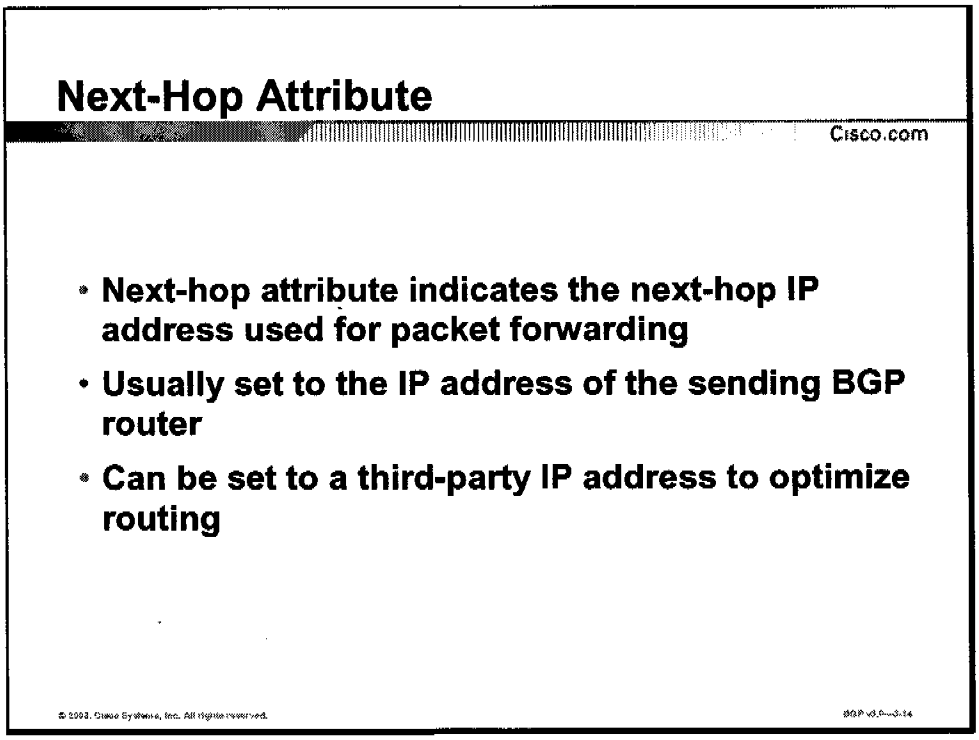
B) BGP routers ignore incoming routes with an AS path containing their own AS number.

C) Multiple paths with the same origin AS are considered routing loops.

D) Multiple paths with the same destination AS are considered routing loops.

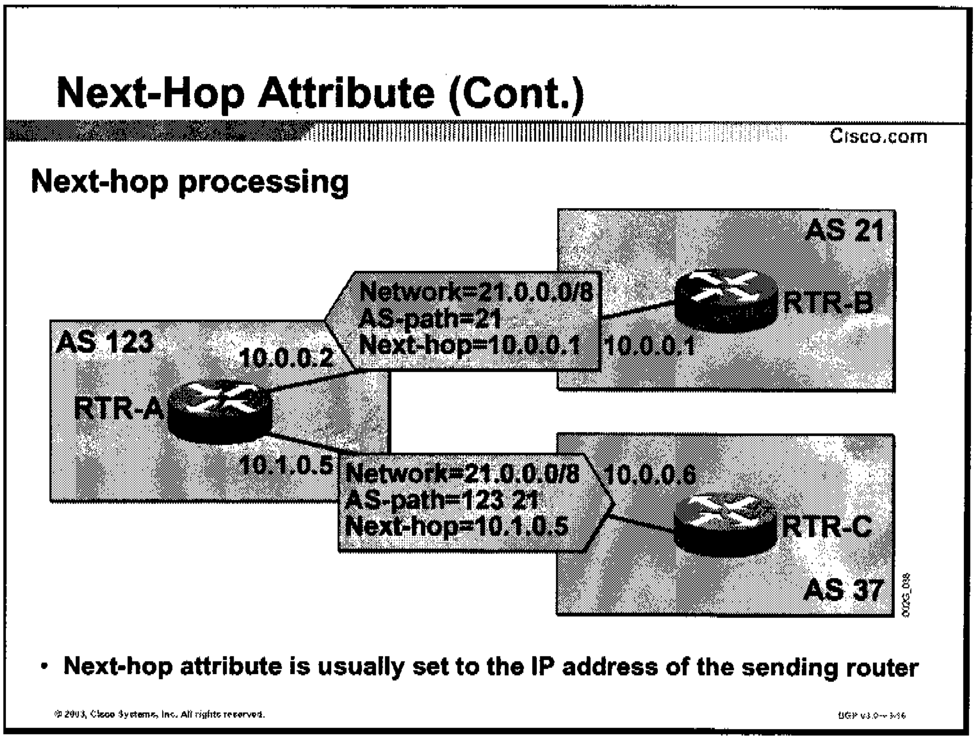
Next-Hop Attribute

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 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



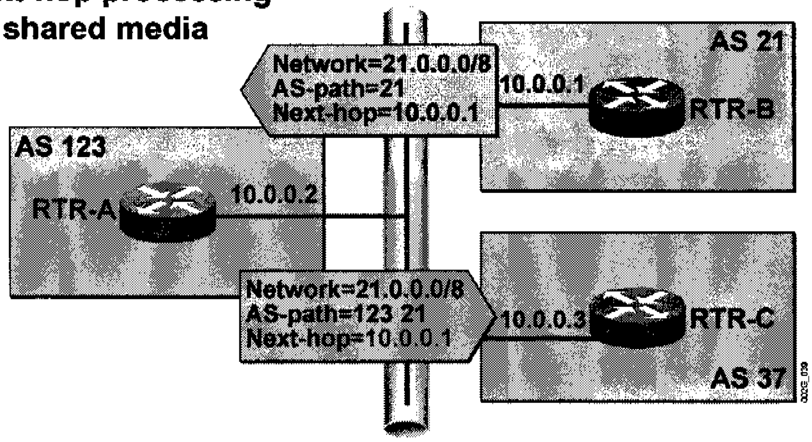
The figure shows the usual next-hop processing:

- RTR-B announces network 21.0.0.0/8 to RTR-A. The outgoing IP address of RTR-B (the address used to establish BGP TCP session) is used as the BGP next hop.
- 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).
- 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 (Cont.)

Cisco.com

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 is not changed to optimize packet forwarding**

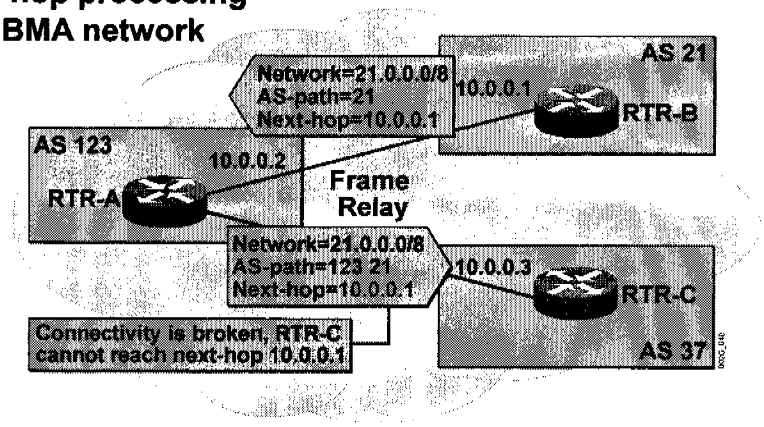
The next-hop processing changes if the BGP routers connect to a shared subnet. In the figure here, if RTR-A announces the network 21.0.0.0/8 to RTR-C with the BGP next-hop address set to RTR-A, the packets from AS 37 toward network 21.0.0.0/8 will have to cross the shared LAN twice. RTR-A thus sends 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 (Cont.)

Cisco.com

Next-hop processing on NBMA network



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

 BGP v3.0—3-20

BGP next-hop processing results in optimum data transfer over shared media (for example, a LAN subnet). In partially meshed networks (like Frame Relay), BGP next-hop processing can break IP connectivity. Consider, for example, the network diagram above: RTR-A will send 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 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. (This option is strongly discouraged in normal BGP designs because routing problems should be solved with a proper network design.)

Practice

Q1) When is the next-hop attribute different from the IP address of the sending router?

A) In situations where static routing is used to reach nondirectly connected IBGP peers.

B) In cases where the BGP next hop is in a different subnet than the receiving router.

C) The next-hop attribute is always different, because it points to the exit gateway.

D) If the current BGP next hop is in the same IP subnet as the receiving router.

Q2) How can improperly designed, partially meshed networks break BGP connectivity?

A) A direct connection may not be available to the same subnet next hop.

B) A design using separate IP networks and subinterfaces can cause BGP to set the incorrect next hop.

C) The BGP next hop is always set to the sending router, causing it to act as the hub of the meshed network.

D) Partially meshed networks automatically disable BGP next-hop processing and should be avoided.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

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

 BGP v3.0—3-21

Next Steps

After completing this lesson, go to:

- Influencing BGP Route Selection with Weights lesson

References

For additional information, refer to these resources:

- For more information on BGP attributes, refer to "Border Gateway Protocol" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

Quiz: BGP Path Attributes

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the purpose 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

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Which three statements are true of BGP mandatory well-known attributes? (Choose three.)

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.

Q2) Which three attributes are BGP mandatory well-known attributes? (Choose three.)

A) next-hop

B) weight

C) AS-path

D) origin

Q3) What three possible values are assigned to the BGP origin attribute? (Choose three.)

A) IGP

B) EGP

C) unknown

D) internal

Q4) What nontransitive optional BGP attribute is useful in assisting with the route selection process when multiple links to another AS exist?

A) next-hop

B) local preference

C) MED

D) AS-path

Q5) How is the BGP next-hop attribute modified?

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.

Q6) Which three statements are true regarding the BGP AS-path attribute? (Choose three.)

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 in the rightmost end of the AS path.

C) The AS-path attribute is also used to avoid routing loops.

D) BGP routes with an empty AS path were injected into BGP from outside the local AS.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Influencing BGP Route Selection with Weights

Overview

This lesson discuses how to influence Border Gateway Protocol (BGP) route selection by setting the weight attribute of incoming BGP routes. Three methods used to set the weight attribute are discussed in this lesson as follows: default weight, filter-list weight, and setting the weight attribute with route-maps. This lesson also explains how to monitor the BGP table to verify correct weight configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The "optimum" or "best" route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection. One of these tools is the weight attribute.

Objectives

Upon completing this lesson, you will be able to:

- List BGP route selection criteria
- Describe the use of BGP weights to influence the BGP route selection process
- Influence the BGP route selection process by configuring per-neighbor weights
- Influence the BGP route selection process by configuring BGP weights with AS-path filters
- Influence the BGP route selection process by configuring BGP weights with route-maps

- Identify the Cisco IOS® commands required to monitor BGP route selection and weights
- Summarize BGP route selection and filtering tools

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

This lesson includes these topics:

- Overview
- BGP Route Selection Criteria
- Influencing BGP Route Selection
- Configuring Per-Neighbor Weights
- Changing Weights with AS-Path Filters
- Changing Weights with Route-Maps
- Monitoring BGP Route Selection and Weights
- BGP Route Selection and Filtering Tools Summary
- Summary
- Assessment (Lab): Influencing BGP Route Selection with Weights

BGP Route Selection Criteria

This topic lists the different criteria used by BGP for best-path route selection.

BGP Route Selection Criteria

Cisco.com

- **Prefer highest weight (local to router)**
- **Prefer highest local preference (global within AS)**
- **Prefer routes that the router originated**
- **Prefer shorter AS paths (only length is compared)**
- **Prefer lowest origin code (IGP < EGP < Incomplete)**
- **Prefer lowest MED**
- **Prefer external (EBGP) paths over internal (IBGP)**
- **For IBGP paths, prefer path through closest IGP neighbor**
- **For EBGP paths, prefer oldest (most stable) path**
- **Prefer paths from router with the lower BGP router-ID**

BGP v3.0—3-4

BGP route selection criteria take the weight parameter into consideration first. If a router has two alternative paths to the same destination, and their weight values are different, BGP selects the route with the highest weight value as the best. Only when the two alternatives have equal weight is the next criterion, local preference, checked.

A high local preference value is preferred before a low value. Only when the two alternatives have an equal local preference is the next criterion checked.

Practice

Q1) What is the difference between local preference and weight?

A) Local preference has a higher priority in BGP path selection.

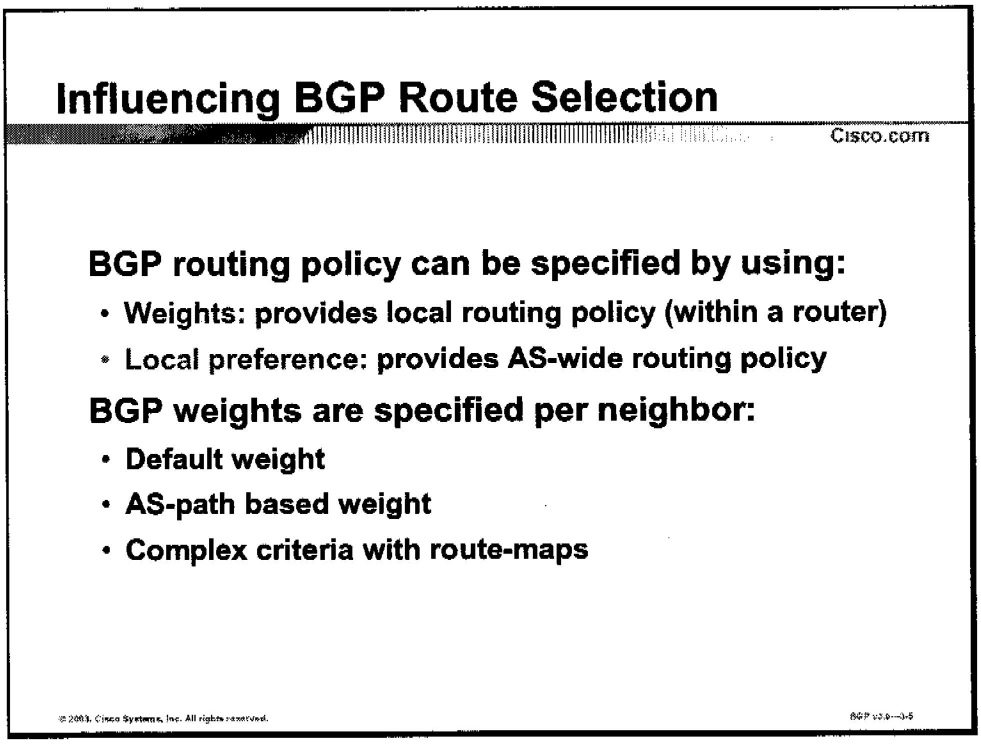
B) Local preference is used AS-wide while weight is local to a single router.

C) Local preference is local only to a specific BGP-speaking router.

D) Local preference is used to influence incoming path selection.

Influencing BGP Route Selection

This topic describes how network administrators can use BGP weights to influence the BGP route selection process.



The **Weight** attribute is local to a single router only. The weight value is never propagated by the BGP protocol. It constitutes a routing policy local to the router.

Local preference is assigned to a route as an attribute. It is carried with the route on all internal BGP sessions. This situation means that all other BGP-speaking routers within the autonomous system (AS) receive the same information. Normally, a router assigns a local preference to a route when it is received on an external BGP session, before it is accepted and entered in the BGP table of the border router. Routers propagate the local preference attribute on internal BGP sessions only. This policy constitutes a routing policy for the entire AS.

The router can assign the weight attribute to a route three ways:

- All routes received from a specific neighbor can be assigned a weight value. This weight value indicates that the neighbor is more preferred than the other neighbors.
- Received routes from a neighbor can be matched against an AS-path filter. Those matched by the filter are assigned a weight value. Those not matched are accepted, but their weight value is not set.
- A route-map applied on incoming routes from a neighbor can be used to select some routes and assign them weight values. Remember that a route-map also acts as a filter and will silently drop those routes not permitted by any statement in the route-map.

If configured, the default weight assignment on routes received from a neighbor is applied first. All routes received from the neighbor are assigned a weight value as defined by the default weight.

Secondly, a filter-list, which does not filter routes out, is applied if configured on the router to select those routes received from the neighbor that should be assigned a specific weight value. Routes permitted by the filter-list have their weight values changed to the value indicated.

Lastly, a route-map is applied, if configured on the router. The route-map can be arbitrarily complex and select routes based on various selection criteria, such as a network number or AS path. The selected routes can have some attributes altered. The route-map can set the weight values of permitted routes. Selection can be done in several route-map statements, giving the opportunity to assign a certain weight value to some routes and another weight value to others. A route-map can also completely filter out routes.

Practice

Q1) What three methods can you use to set the BGP weight attribute? (Choose three.)

A) route-map

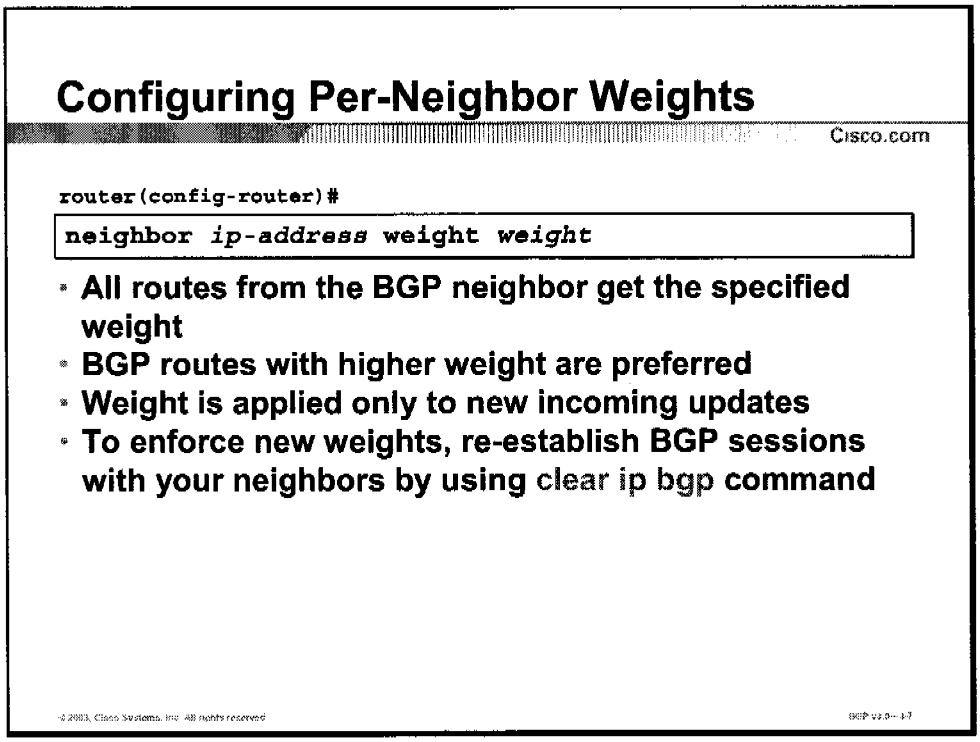
B) AS-path filter-list

C) access control list

D) default weight assigned to a specific neighbor

Configuring Per-Neighbor Weights

This topic describes how to influence the BGP route selection process by configuring per-neighbor weights.



neighbor weight

To assign a weight to a neighbor connection, use the **neighbor weight** router configuration command.

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

To remove a weight assignment, use the **no** form of this command.

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

Syntax Description

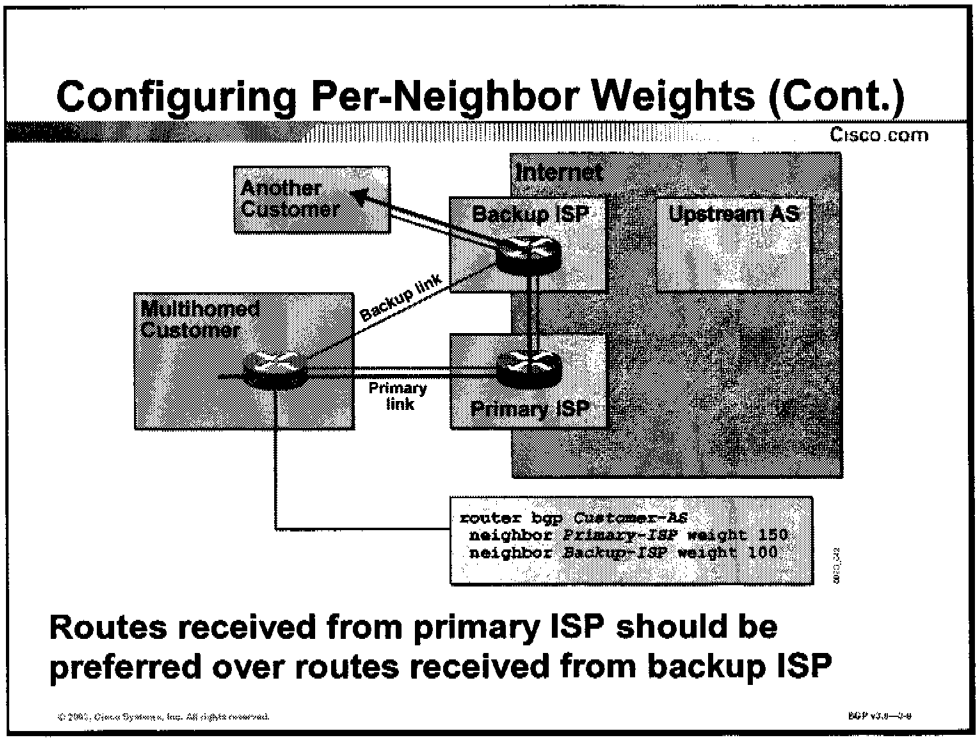
ip-address	IP address of neighbor.
peer-group-name	Name of a BGP peer group.
weight	Weight to assign. Acceptable values are 0 to 65535.

All routes received from the neighbor after the configuration line is in place are assigned the weight value. To make sure that all routes from the neighbor receive the new weight value, you can restart the BGP session, thus forcing the neighbor to resend all routes.

If no weight value is specified, the default value of 0 is applied.

Restarting of BGP sessions might be necessary after making a configuration change in the routing policy. The configuration change itself will not alter the already-received routes. The **clear ip bgp** EXEC command tears down the BGP session, and the session automatically restarts.

Example



In this example, the multihomed customer would like to use the primary link to the primary Internet service provider (ISP) for all destinations. The weight is configured by the customer on both BGP sessions, giving a higher weight to those routes received from the primary ISP compared to those received from the backup ISP.

Any time the multihomed customer receives routing information about the same IP network number from both the ISPs, the customer compares the weights assigned to the routes. Those received from the primary ISP will always win this comparison. The multihomed customer sends the outgoing IP packets to the destination network via the primary ISP regardless of the other BGP attributes assigned to both alternatives.

Consequently, the other customer directly connected to the backup ISP will also be reached via the primary ISP.

Configuring Per-Neighbor Weights (Cont.)

Cisco.com

```
Customer# 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                              150 37 21 i
*                   2.3.4.5                  0           100 21 i
*> 37.0.0.0         3.4.5.6                  0           150 37 i
*                   2.3.4.5                              100 21 37 i
*> 40.0.0.0         3.4.5.6                  0           150 37 40 i
*                   2.3.4.5                              100 21 40 i
```

 BGP v3.0—3-19

In this example, the multihomed customer has received routes to three different class A networks outside of its own AS (network 21.0.0.0/8, network 37.0.0.0/8, and network 40.0.0.0/8). The customer received all three routes from both the primary ISP and the backup ISP.

When the routes were received from the primary ISP, the weight value 150 was assigned to each of the routes. When the routes were received from the backup ISP, the weight value 100 was assigned to each of the routes.

The customer router now makes the route selection. It has two alternative paths for each destination network. For each of them, the router selects the path via the primary ISP as the best. It makes this selection regardless of other BGP attributes, such as AS-path length.

The network 21.0.0.0/8 is reached via the primary ISP although it is actually a network in the AS of the backup ISP (AS 21).

The class A network 1.0.0.0/8 in this example is injected into the BGP table by this router. By default, locally sourced routes are assigned a weight of 32768.

Practice

Q1) What is the default weight for routes received from a BGP neighbor?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

Q2) What default weight is applied to locally sourced BGP routes?

A) 0

B) 100

C) 32768

D) depends on the Cisco IOS release

Q3) When are the weights configured on a neighbor enforced?

A) Before the new weights can take effect, the BGP process on the router must be removed and reconfigured.

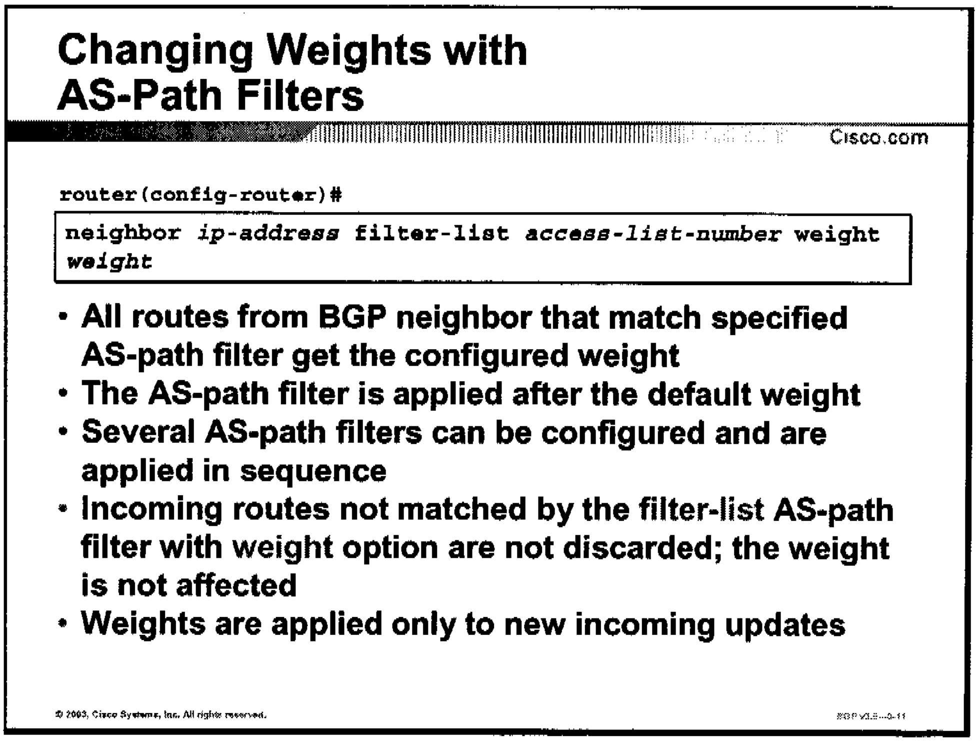
B) The router must first be rebooted for the new weights to take effect.

C) The new weights will be applied after the BGP update interval of 30 minutes expires.

D) The new weight configuration is applied to all routes received following the configuration change.

Changing Weights with AS-Path Filters

This topic describes how to influence the BGP route selection process by configuring BGP weights with AS-path filters.



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** | **weight** *weight*}

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** | **weight** *weight*}

Syntax 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. You define this access list with the **ip as-path access-list** command.
in	Access list to incoming routes.
out	Access list to outgoing routes.

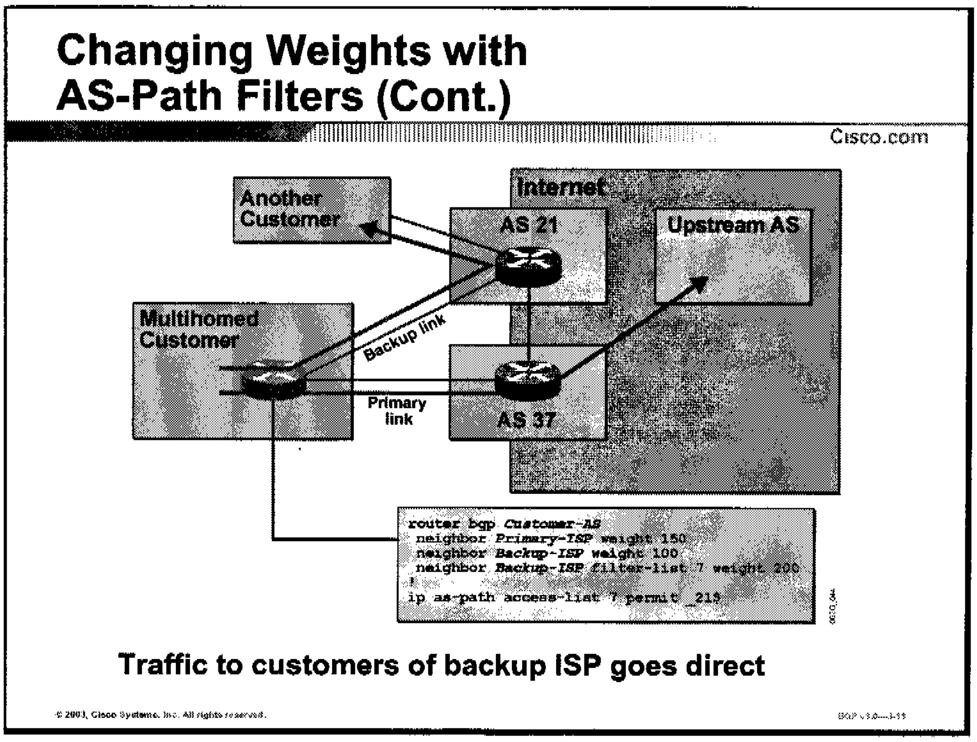
weight *weight*	Assigns a relative importance to incoming routes matching AS paths. Acceptable values are 0 to 65535.

When you use the keyword **weight**, the router does not use the filter-list to discard any routes. Instead, the filter-list is applied on the incoming routes and selects those that should have their weight value altered. Those that the filter-list does not select are accepted without changing the weight.

Restarting of BGP sessions might be necessary after configuring a change in the routing policy. BGP applies the new configuration only on routes coming in after the configuration change.

Note Specifying weights with filter-lists is no longer supported in Cisco IOS Release 12.1, and the command has already been removed from Cisco IOS release 12.1T. These releases use an incoming route-map, where you match an AS path with the **match as-path** command and set weight with the **set weight** command. When using a route-map as a replacement for the filter-list with **weight** option, make sure that specifying a "permit" entry in the route-map without an associated match condition does not filter all other routes. Using route-maps as a weight-setting mechanism is explained later in this lesson.

Example



Sometimes a blind preference to use a specific provider in a multihomed network can cause strange results. The case where a customer reaches networks in the AS of the backup ISP via the primary ISP even though it has a direct connection to the backup ISP is an example of a network configuration leading to strange results.

To change BGP path selection into a more predictable behavior, the blind preference has to be replaced with a more selective preference. In the example, the backup ISP has the AS number 21. Within that AS there are many single-homed customer networks that do not have their own AS number but are part of the AS of the backup ISP.

All routes received from the primary ISP are assigned the weight 150 by the customer edge router. But routes received from the backup ISP are assigned weight values in a more selective way. First, all routes are assigned the value 100. Then those routes selected by the AS-path access-list 7 have their values changed. The filter-list 7 selects those routes having an AS path containing only 21, indicating that they were created in AS 21.

The result is that most routes are preferred to reach the network via the primary ISP. Only those routes created in the AS of the backup ISP are preferred to reach the network via the direct link to the backup ISP.

Changing Weights with AS-Path Filters (Cont.)

Cisco.com

```
Customer# 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                              150 37 21 i
*>                  2.3.4.5                  0           200 21 i
*> 37.0.0.0         3.4.5.6                  0           150 37 i
*                   2.3.4.5                              100 21 37 i
*> 40.0.0.0         3.4.5.6                  0           150 37 40 i
*                   2.3.4.5                              100 21 40 i
```


BGP v3.0—3-14

The **show ip bgp** command output shows the BGP table of a multihomed customer in AS 123. It has direct links to two different ISPs. The primary ISP is AS 37, and the backup ISP is AS 21.

In this example, the customer would like to reach the destination networks within AS 21 directly via the backup ISP. To accomplish this, the customer must configure a more selective assignment of weights. First, the customer assigns the default weight 150 to all routes received from the primary ISP and the weight 100 to all routes received from the backup ISP. Additionally, the routes received from the backup ISP must be checked by a filter-list. Those permitted by the list have their weight values changed to 200. The filter-list selects those routes having an AS path indicating that they were created in AS 21.

As a result, the customer reaches all networks outside its own AS via the primary ISP, except for the class A network 21.0.0.0/8, which it reaches via the backup ISP.

Practice

Q1) How could you implement a primary/backup ISP routing policy using weights?

A) Assign higher weights to all routes received from the backup ISP.

B) Assign lower weights to all routes received from the backup ISP. - default = 0

C) Assign higher weights to all routes received from the primary ISP.

D) Assign lower weights to all routes received from the primary ISP.

Q2) What is the difference between a **filter-list in** configuration command and a **filter-list weight** configuration command?

A) The filter-list without the **weight** keyword can manipulate AS paths or the weight attribute, but the **weight** keyword is limited to the weight attribute.

B) **Filter-list weight** will not filter out any route but will assign the weight value to those routes permitted by the filter-list.

C) Using **filter-list weight** forces the router to use process switching.

D) **Filter-list weight** can match an AS path with a single AS number entry only.

Changing Weights with Route-Maps

This topic describes how to influence the BGP route selection process by configuring BGP weights with route-maps.

Changing Weights with Route-Maps

Cisco.com

- **Weights can be set with route-maps in complex scenarios**
- **Routes can be matched on any combination of prefix-lists, AS-path filters, or other BGP attributes**
- **Routes not matched by the route-map are discarded**

```
router bgp as-number
 neighbor ip-address route-map route-map-name in
!
route-map route-map-name permit sequence
 match condition
 set weight weight
```

 BGP v3.0—3-15

The route-map is a powerful tool to select and alter routing information. When a route-map is applied to incoming information from a BGP neighbor, each received update is examined as it passes through the route-map. Statements in the route-map are executed in the order specified by their sequence numbers

The first statement in the route-map that has all the match clauses indicating a match is the one used. If the route-map says "permit," the set clauses are applied to the route, the route is accepted, and the weight is changed.

Match clauses can be arbitrarily complex. One of them can refer to an AS-path access-list that does matching on AS paths. Another can refer to a prefix-list that does matching on the announced network number. Only when all configured match clauses permit is the route-map statement used and its result, "permit" or "deny," applied.

If a received route is not matched by any of the route-map statements, and the end of the route-map is reached, the route-map logic has an "implicit deny" rule. This rule means that if no statement selects a route, the route is discarded.

If the "implicit deny" rule is not desired, an "explicit permit all" at the end of the route-map can overrule it. To ensure that such a route-map statement is the last statement, you should assign it a very high sequence number. It should not have any match clause at all. The lack of a match clause means, "match all." By not configuring any set clause, you can ensure that no attributes are altered by the statement.

Example

Changing Weights with Route-Maps (Cont.)

Cisco.com

Set weight 200 to networks coming from 2.3.4.5 originated in AS 21

```
router bgp 123
neighbor 2.3.4.5 route-map w200 in
!
route-map w200 permit 10
match as-path 47
set weight 200
!
route-map w200 permit 20
set weight 100
!
ip as-path access-list 47 permit _21$
```


BGP v3.0—3-10

This is an example of a route-map that sets the weight value to each route received from a neighbor.

All received routes have their AS paths checked against the AS-path access-list 47. Those routes having an AS path indicating that originated in AS 21 are permitted by the AS-path access-list 47 as referenced by route-map statement number 10. Routes permitted and selected by route-map statement number 10 in the w200 route-map will have their weight set to 200 as indicated by the set clause in the route-map.

The routes that are not originated in AS 21 (routes not permitted by AS-path access-list 47) are then tested by route-map statement number 20. This statement does not include a match clause, indicating that all routes are matched. Therefore, all routes not matched by route-map statement 10 are matched by route-map statement 20. The route-map has been configured with an "explicit permit all" statement at the end of the route-map.

Routes matched by route-map statement 20 have their weight set to 100. The result is that the routes originated in AS 21 are accepted by the router and assigned the weight 200. All others are accepted and assigned the weight value 100. No route is discarded by this route-map.

Practice

Q1) When you are using route-maps to modify weights, what happens by default to a route that does not match any of the route-map statements?

A) The route is accepted with the weight attribute unmodified.

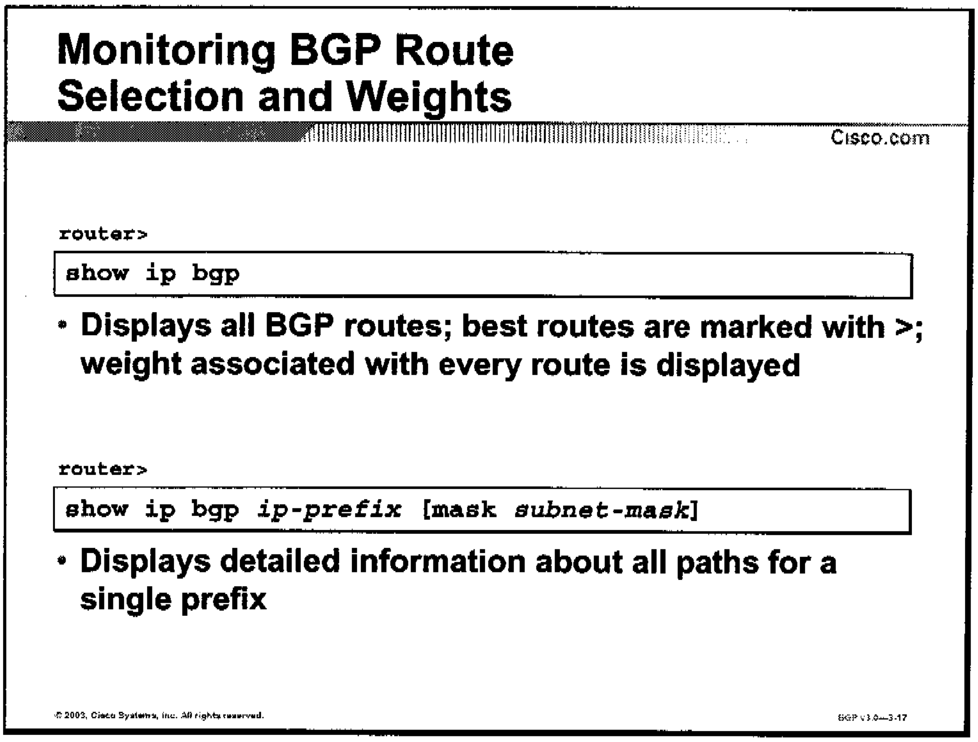
B) The route is discarded.

C) The route will be inserted into the BGP table but not the IP routing table.

D) An error will be displayed on the router console and in router debugs.

Monitoring BGP Route Selection and Weights

This topic lists the Cisco IOS commands required to monitor BGP route selection and weights.



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

network	(Optional) Network number, entered to display a particular network in the BGP routing table
network-mask	(Optional) Displays all BGP routes matching the address/mask pair
longer-prefixes	(Optional) Displays route and more specific routes

Without any argument, the **show ip bgp** command displays the entire BGP table. The routes selected as the best are indicated by the ">" character.

To get more detailed information about routes to a specific destination network, you can use the network number, and optionally the subnet mask, as an argument on the command line. These additions will display more detailed information about that specific network.

Example

Monitoring BGP Route Selection and Weights (Cont.)

Cisco.com

```
router> show ip bgp
BGP table version is 11, local router ID is 12.1.2.3
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
*> 10.0.0.0         1.2.0.1                500           100 37 213 i
*                   1.1.0.1               1000             0 213 i
*> 11.0.0.0         1.2.0.1                500           100 37 48 i
*                   1.1.0.1               1000             0 213 48 i
*> 12.0.0.0         0.0.0.0                  0         32768 i
*> 14.0.0.0         1.1.0.3                  0             0 387 i
```

 BGP v3.0—3-18

The **show ip bgp** command gives a printout of all routes in the BGP table. Each route is displayed on one line. This one-line limitation means that more detailed information about the route cannot be displayed, due to lack of space.

The network number is displayed, and if the subnet mask differs from the natural mask, the prefix length is indicated. The BGP next-hop attribute, multi-exit discriminator (MED; metric), local preference, weight, AS-path, and origin code are displayed on the line. Local preference is displayed only if it is not the default value.

The printout is sorted in network number order. If there is more than one route to the same network, the network number is printed on the first line only. The other routes to the same network have their network field left blank on the output.

Routes selected as the best to reach a certain destination network are indicated by the ">" character.

In this example, weight has been used to prefer routes received from the neighbor in AS 37. Therefore, although the AS path is shorter via AS 213, the class A network 10.0.0.0/8 is reached via AS 37 (because the weight is higher).

Information about network 14.0.0.0/8 is received only from the neighbor in AS 387. As there is no alternative, the route is selected as best.

Monitoring BGP Route Selection and Weights (Cont.)

Cisco.com

```
router> show ip bgp 11.0.0.0
BGP routing table entry for 11.0.0.0/8, version 5
Paths: (2 available, best #1, advertised over EBGP)
  213
    1.2.0.1 from 1.2.0.1 (10.1.1.1)
      Origin IGP, metric 500, localpref 100, valid, external, best
  213
    1.1.0.1 from 1.1.0.1 (11.0.0.1)
      Origin IGP, metric 1000, localpref 100, valid, external
```


BGP v3.0—3-19

The **show ip bgp** command with network number as argument displays more detailed information about that network only. First, a short summary indicating the network number and prefix length is displayed along with the table version number for this route. The next line says how many alternative routes have been received, and which one of them has been selected by the router as the best.

Next, there are a couple of lines for each of the received routes to reach the network. For each of the routes, all attributes are displayed. The one selected as the best also has the word "best" displayed.

In this example, there are two alternatives to reach network 11.0.0.0/8. Each of them is received from different neighbors in AS 213. The network 11.0.0.0 is created in AS 213.

The route selection mechanism has selected the first route listed as the best. It was chosen because the MED (metric) value is lower.

Practice

Q1) What three pieces of information can you obtain from the output of the **show ip bgp** command? (Choose three.)

A) the best route to a destination

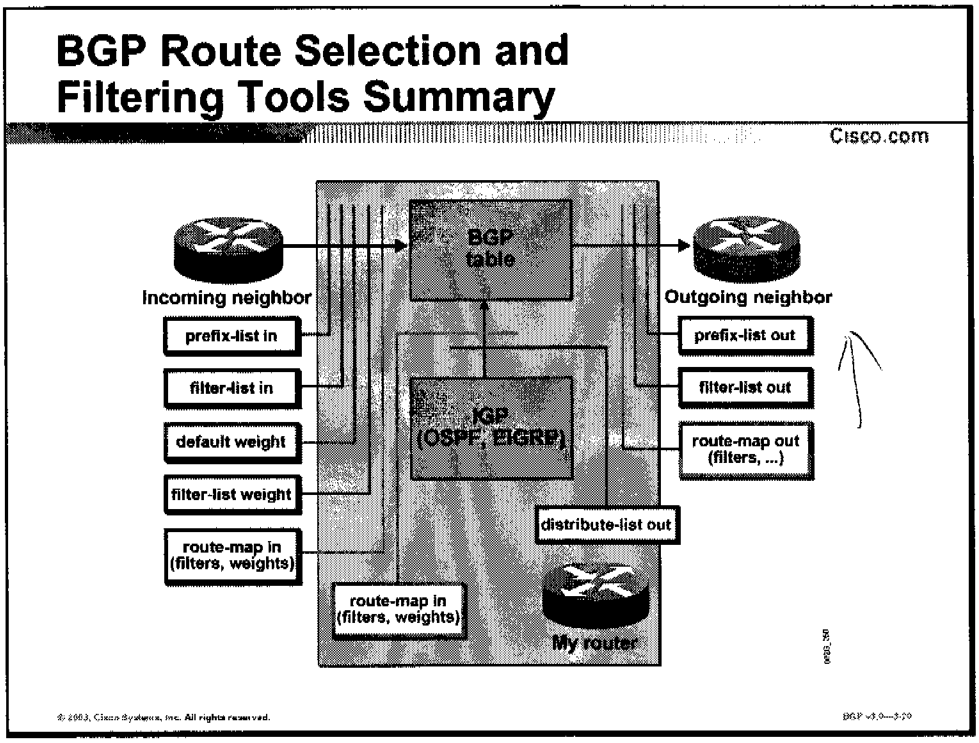
B) the weight attribute

C) the AS path of the route

D) the administrative distance of the route

BGP Route Selection and Filtering Tools Summary

This topic presents a summary of all BGP filtering tools in the order they are applied.



The figure shows all the possible applications of prefix-lists, filter-lists, weights, and route-maps. They are applied in the order indicated.

Prefix-lists and filter-lists, both in and out, filter out routes and discard those not permitted. Weight setting is applicable only on incoming routes because a router never propagates the weight attribute to its neighbors. Route-maps can be filters that discard routes but can also be used to modify and set various attributes on both incoming and outgoing routes.

Practice

Q1) Which method of influencing route selection with weights is the last to be applied on an incoming interface?

A) prefix-list

B) route-map

C) filter-list weight

D) default weight

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- One of the most commonly used ways of influencing BGP path selection is with weights. Weights are the first criteria in BGP route selection.
- Changing weights is an operation that applies only to the local router and is lost when the BGP update is propagated to other BGP neighbors.
- You can use the neighbor weight command to assign a weight value to all routes received from a neighbor.
- Configuring a neighbor filter-list with the weight keyword will apply a weight to routes matching the AS-path list.

BGP v3.0—3-21

Summary (Cont.)

Cisco.com

- Route-maps can be applied to neighbors to set the weight attribute of received routes.
- You can use the show ip bgp command to display all bgp routes, the routes selected by BGP as "best," and the weight attribute setting for each route.
- The weight attribute setting is applicable only on incoming routes because weight is never propagated to other neighbors.

BGP v3.0—3-22

Next Steps

After completing this lesson, go to:

- BGP Local Preference lesson

References

For additional information, refer to these resources:

- For more information on the weight attribute, refer to "Border Gateway Protocol" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For further information on the weight attribute, refer to "Configuring BGP" at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm#xtocid15

Laboratory Exercise: Influencing BGP Route Selection with Weights

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the weight attribute in a situation where you must support connections to multiple ISPs.

After completing this exercise, you will be able to:

- Influence the BGP route selection process by configuring per-neighbor weights
- Influence the BGP route selection process by configuring BGP weights with route-maps
- Monitor BGP route selection and weights

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode
neighbor *ip-address* **weight** *weight*	Assign weight to all received updates from specified neighbor
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	To apply a route-map to incoming or outgoing routes
route-map *name* {**permit** | **deny**} **seq**	Define or modify an existing entry in a route-map
match as-path *list-number*	Specify route-map matching criteria to match an as-path access list
set weight *weight*	Set weight in a route-map
show ip bgp summary	Verify if both BGP sessions are up
show ip bgp	Inspect the contents of the BGP table
clear ip bgp	Clear the BGP session with your neighbor

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Currently, you are using the service provider "Cheap" as your primary provider for Internet connectivity and the service provider "Good" as your backup provider. As the result of this policy, the link toward the "Good" service provider is underused, while the link toward the provider "Cheap" is overloaded.
- An analysis of the Internet structure beyond your service providers indicates that you could improve the link use if you send traffic toward AS 213 and AS 37 directly to the "Good" service provider.
- In this exercise, you will improve the simple routing policy created in the Multihomed BGP Networks lab exercise, by specifying BGP weights with route-maps.
- You must implement the following routing policy:
  - Prefer routes announced from router "Cheap" over those announced from router "Good"
  - Prefer routes going through or originating in AS 213 or AS 37 from router "Good"
- Figure 1 shows the connectivity that is established between your AS and the two service providers "Good" and "Cheap."

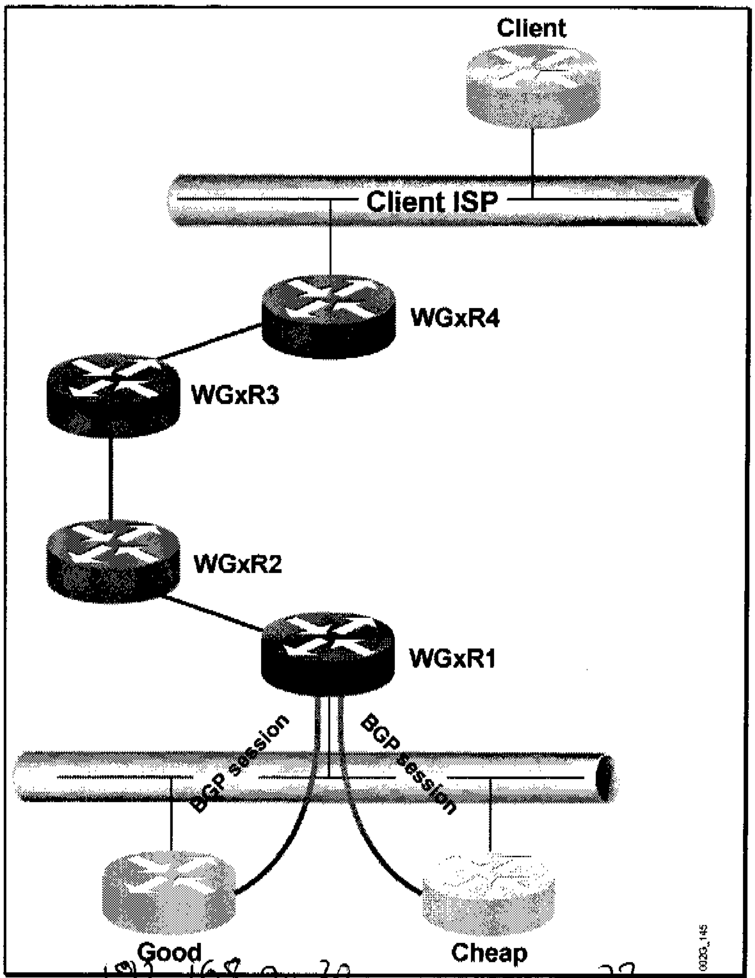


Figure 1: Connections to two different service providers

Exercise Procedure

Acess list ②

Complete these steps:

Configuring AS-path access-lists:

Step 1 Create an AS-path access-list that permits AS numbers 213 and 37 in the AS path.

Setting per-neighbor default weights:

Step 2 Set default weights for both neighboring routers.

Step 3 Inspect your BGP table to verify that you prefer routes coming from "Cheap" to routes coming from "Good."

Using AS-path filters within route-maps to set weights:

Step 4 Create a new route-map. Use the previously configured AS-path access-list as the **match** condition in one of the route-map statements, and set the weight of matched routes as needed.

Step 5 Apply the route-map to incoming updates from router "Good."

Exercise Verification

You have completed this exercise when you attain these results:

- Verify your BGP table to see if all prefixes with two paths prefer the one through router "Cheap."

```
wg1r1#sh ip bgp
BGP table version is 43, local router ID is 197.1.8.1
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
*  128.20.0.0       192.168.20.20            0          100 20 i
*>                  192.168.20.22            0          200 22 i
*> 128.20.12.0/24   192.168.20.20            0          100 20 i
*  128.22.0.0       192.168.20.20            0          100 20 i
*>                  192.168.20.22            0          200 22 i
*> 128.22.12.0/24   192.168.20.22                       100 20 22 i
*  128.26.0.0       192.168.20.20            0          100 20 42 26
i
*>                  192.168.20.22            0          200 22 26 i
*> 128.37.0.0       192.168.20.20            0          300 20 42 37
i
*                   192.168.20.22            0          200 22 26 42
37 i
*  128.42.0.0       192.168.20.20            0          100 20 42 i
*>                  192.168.20.22            0          200 22 26 42
i
*  128.51.0.0       192.168.20.20            0          100 20 42 26
51 i
*>                  192.168.20.22            0          200 22 26 51
i
*> 128.213.0.0      192.168.20.20            0          300 20 213 i
*> 128.214.0.0      192.168.20.22            0          200 22 214 i
*> 192.20.11.0      192.168.20.22            0          200 22 i
*> 192.20.12.0/30   192.168.20.20                       200 22 20 i
*> 192.22.11.0      192.168.20.22            0          200 22 i
*> 192.22.12.0/30   192.168.20.22            0          200 22 i
*> 192.26.11.0      192.168.20.22            0          200 22 26 i
*> 192.37.11.0      192.168.20.22            0          200 22 26 42
37 i
*> 192.42.11.0      192.168.20.22            0          200 22 26 42
i
*> 192.51.11.0      192.168.20.22            0          200 22 26 51
i
*> 192.168.1.0      0.0.0.0                  0        32768 i
*> 192.214.11.0     192.168.20.22            0          200 22 214 i
*> 197.1.0.0/21     0.0.0.0                           32768 i
*> 197.1.0.0/16     0.0.0.0                           32768 i
s> 197.1.1.0        0.0.0.0                  0        32768 i
s> 197.1.2.0        0.0.0.0                  0        32768 i
s> 197.1.3.0        0.0.0.0                  0        32768 i
s> 197.1.4.0        0.0.0.0                  0        32768 i
s> 197.1.5.0        0.0.0.0                  0        32768 i
s> 197.1.6.0        0.0.0.0                  0        32768 i
s> 197.1.7.0        0.0.0.0                  0        32768 i
s> 197.1.8.0        0.0.0.0                  0        32768 i
*> 197.1.8.0/22     0.0.0.0                           32768 i
*  200.20.0.0/16    192.168.20.20            0          100 20 i
*>                  192.168.20.20                       200 22 20 i
*  200.22.0.0/16    192.168.20.22                       100 20 22 i
*>                  192.168.20.22            0          200 22 i
```


Answer these questions:

Q1) Did all paths automatically get a weight of 100 or 200? Why not? What did you have to do?

Q2) Name some parameters and attributes used for best-path selection.

BGP Local Preference

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP local preference attribute of incoming BGP routes. Local preference is similar to the weight attribute but differs from the BGP weight attribute in that weight is local to a specific router on which it is configured. Two methods used to set the local preference attribute are discussed in this lesson as follows: default local preference and setting the local preference attribute with route-maps. This lesson also explains how to monitor the BGP table to verify correct local preference configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The "optimum" or "best" route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection. One of these tools is the local preference attribute.

Objectives

Upon completing this lesson, you will be able to:

- Explain why using BGP weights may not provide consistent BGP route selection in an AS
- Describe how the BGP local preference attribute influences BGP route selection
- Identify the Cisco IOS® commands required to configure default BGP local preference on a router

- Identify the Cisco IOS commands required to configure BGP local preference using route-maps

- Identify the Cisco IOS commands required to monitor BGP local preference

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

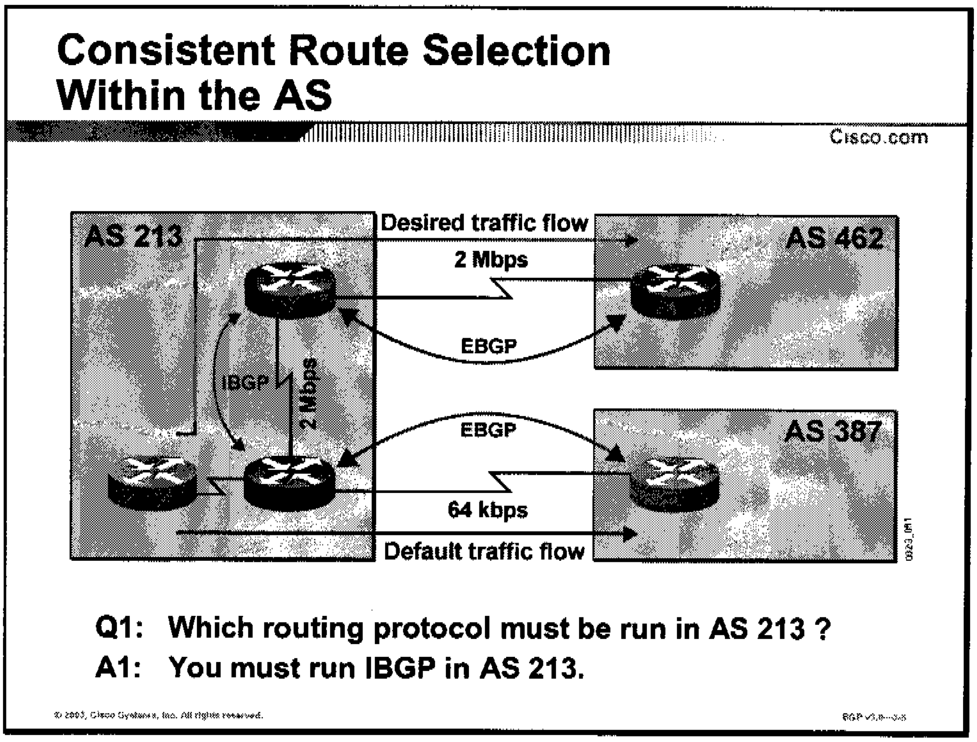
Outline

This lesson includes these topics:

- Overview
- Consistent Route Selection Within the AS
- BGP Local Preference
- Configuring Default Local Preference
- Configuring Local Preference with Route-Maps
- Monitoring Local Preference
- Summary
- Assessment (Lab): BGP Local Preference

Consistent Route Selection Within the AS

This topic explains why using BGP weights may not provide consistent BGP route selection inside of an autonomous system (AS).

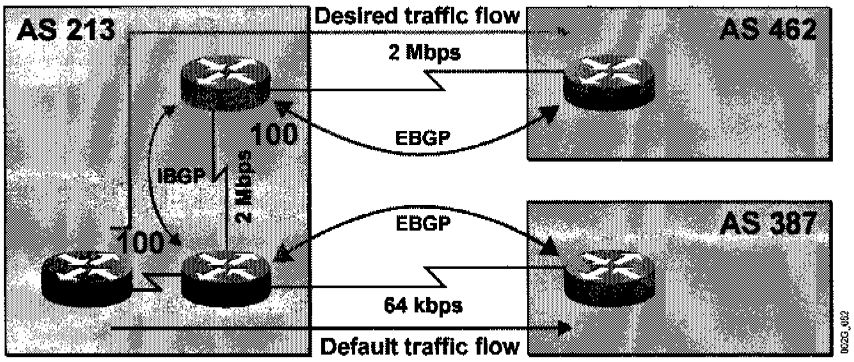


Using BGP in autonomous systems with a single neighbor relationship usually does not require any advanced features. In situations like the one shown in the figure, however, it is important to ensure that customer routers choose the correct link. Obviously, the router should choose the 2-Mbps link and use the 64-kbps link only for backup purposes.

To make sure that the router selects the upper link (2-Mbps link) as its primary link, and has the ability to switch over to the backup if a failure occurs, you must configure an internal BGP session between the two border routers in AS 213.

Consistent Route Selection Within the AS (Cont.)

Cisco.com



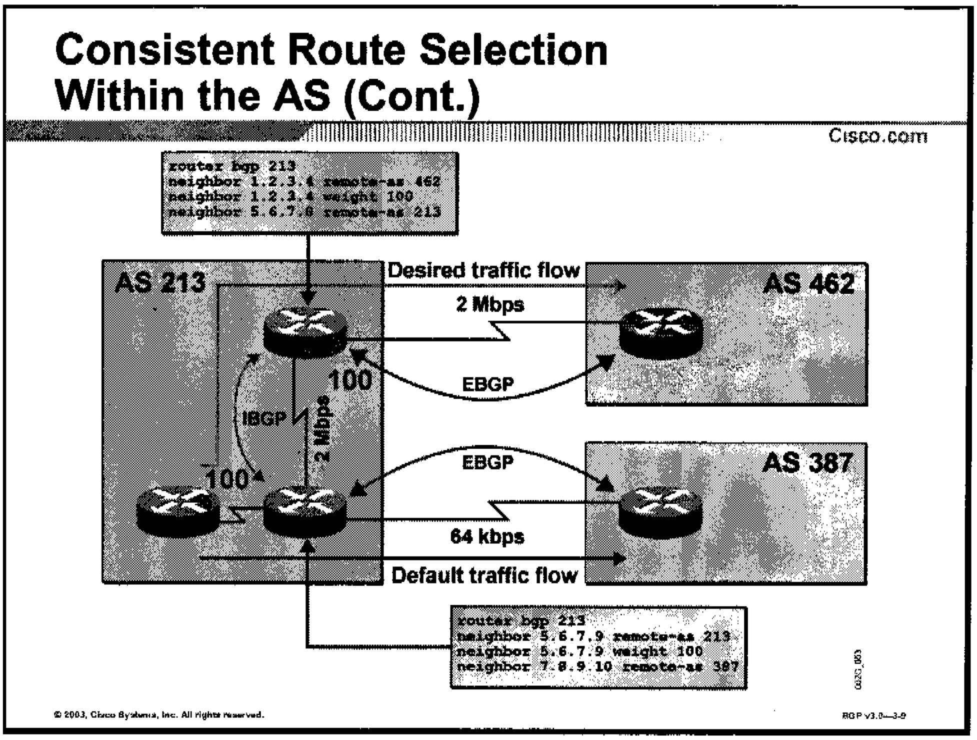
Q2: How will you influence the route selection on routers in AS 213 so that they select the fastest route?

A2: By using weights on EBGP and IBGP sessions.

 BGP v3.0—3-7

One way of changing the default route selection is to use weights. Weight is an attribute that is locally significant to a router. It is a property or parameter and is, therefore, not seen on any neighboring routers. When designing BGP networks using weights, network administrators should set weights on every router. If there is more than one path for the same network, a router will choose the one with the highest weight. The default value for weight is 0.

In this example, the upper router in AS 213 sets a weight of 100 to routes received over the 2-Mbps link from AS 462 (primary link) and prefers them to possible internal updates from the bottom router, where the default weight is zero. The bottom router sets a weight of 100 to internal routes received from the upper router and prefers them to routes received from AS 387. As a result, all packets will leave the AS through the primary 2-Mbps link.



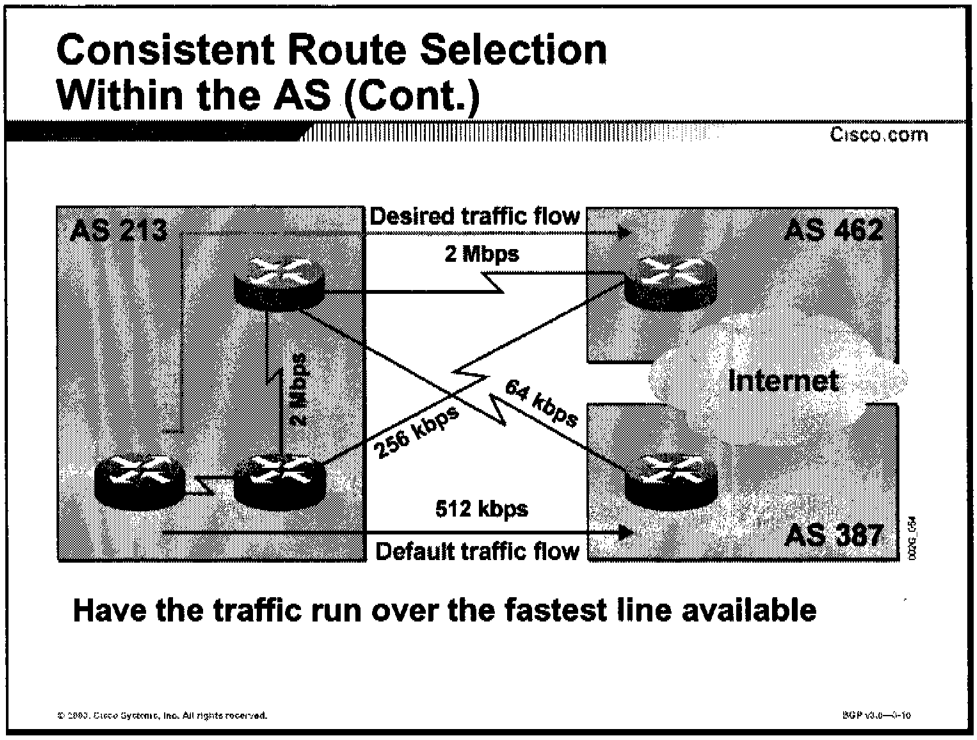
The configurations shown in the figure demonstrate how to change the default weight on a per-neighbor basis. If you use the **neighbor weight** command, all newly arrived updates will have a weight of 100. Updates coming from the other neighbor will still have the default weight of 0.

After you have applied the **neighbor weight** command, a refresh is needed from the neighbor. There are three ways of doing this, depending on the Cisco IOS version:

- Use **clear ip bgp neighbor address** to clear the neighbor relationship and re-establish it to refresh the BGP entries and apply the weight.
- Configure soft reconfiguration for the neighbor and use the **clear** command as shown in the figure. You can perform all subsequent clearing by using **clear ip bgp neighbor address soft in**, which does not reset the neighbor relationship. The soft reconfiguration feature is supported by Cisco starting with IOS version 11.2.
- Use **clear ip bgp neighbor address in** if both neighboring routers support the route refresh. The route refresh feature is available from Cisco starting with IOS version 12.1.

See the Implementing Changes in BGP Policy lesson for a detailed description of the commands here.

Example



This example is more complex. When you are trying to implement this example with weights, it requires two route-maps on each router within AS 213. Luckily, BGP has a similar mechanism that you can use for consistent AS-wide route selection: local preference.

Practice

Q1) What is a key difference between the local preference and weight attributes?

A) Local preference is local to the route on which it is configured.

B) Local preference is local to the AS within which it has been configured.

C) Local preference is local to the BGP administrative domain.

D) Local preference is global to a BGP domain.

Q2) What is an appropriate BGP implementation for the weight attribute?

A) all large-scale BGP implementations requiring AS-wide path selection policies

B) simple routing policies in smaller networks

C) BGP implementations connecting a single-homed customer to provider

D) domains requiring global path selection policies

BGP Local Preference

This topic describes how the BGP local preference attribute influences BGP route selection.

BGP Local Preference

Cisco.com

- **You can use local preference to ensure AS-wide route selection policy**
- **Any BGP router can set local preference when processing incoming route updates, when doing redistribution, or when sending outgoing route updates**
- **Local preference is used to select routes with equal weight**
- **Local preference is stripped in outgoing EBGP updates except in EBGP updates with confederation peers**

BGP v3.0—3-11

Local preference is similar to weight; because it is as an attribute, you can set it once and then view it on neighboring routers without having to reset it. This attribute has a default value of 100, which the router will apply to locally originated routes and updates coming from external neighbors. Updates coming from internal neighbors already have the local preference attribute.

Local preference is the second strongest criteria in the route selection process. If there are two or more paths available for the same network, a router will first compare weight, and if the weights are equal for all paths, the router will then compare the local preference attribute. The path with the highest local preference value is preferred.

The local preference attribute is automatically stripped out of outgoing updates to External Border Gateway Protocol (EBGP) sessions. This fact means that you can use this attribute within a single AS only to influence the route selection process.

BGP Local Preference (Cont.)

Cisco.com

Local preference is the second strongest BGP route selection parameter

Remember the BGP route selection rules:

- **Prefer highest weight (local to router)**
- **Prefer highest local preference (global within AS)**
- **Other BGP route-selection rules**

Weights configured on a router override local preference settings

To ensure consistent AS-wide route selection:

- **Do not change local preference within the AS**
- **Do not use BGP weights**

BGP v3.0—3-12

Local preference is the second strongest BGP route selection parameter. Remember the route selection rules:

1. Prefer highest weight (local to router)

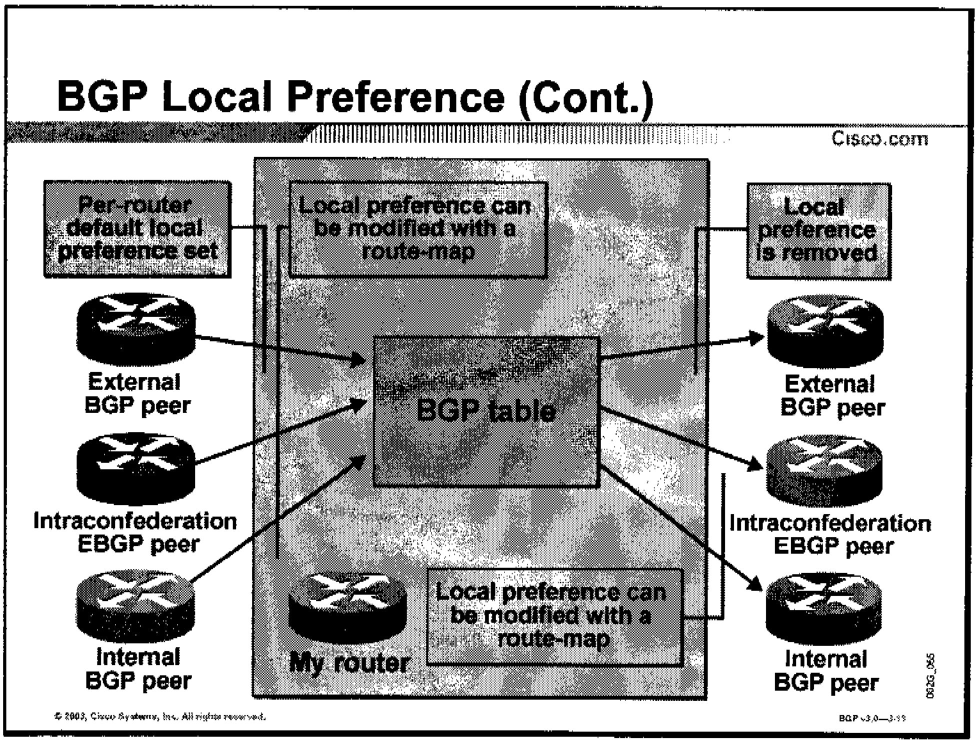
2. Prefer highest local preference (global within AS)

3. Process all remaining BGP route selection rules

Because network administrators can use both weight and local preference to manipulate the route selection process, they must decide which one to use. If local preference is used, the weight should be the same for all paths.

Network administrators can use weight on an individual router to override local preference settings used in the rest of the AS.

In most cases, it is enough to change the default local preference on updates coming from external neighbors. Network administrators should avoid changing the local preference attribute on internal sessions to prevent unnecessary complexity and unpredictable behavior.



Network administrators can apply local preference in the following ways:

- Using a route-map with the **set local-preference** command. You can use the route-map on incoming updates from all neighbors or on outgoing updates to internal neighbors (not recommended).

- Using the **bgp default local-preference** command to change the default local preference value applied to all updates coming from external neighbors or originating locally.

Practice

Q1) If you configure both local preference and weight, which has the highest priority?

A) Local preference always has a higher priority.

B) Weight always has a higher priority.

C) If routes are learned from different neighbors, the local preference will determine path selection.

D) If two routes have the same local preference, weight will determine the selected path.

Configuring Default Local Preference

This topic lists the Cisco IOS commands required to configure default BGP local preference on a Cisco router.

Configuring Default Local Preference

Cisco.com

```
router(config-router)#
bgp default local-preference preference
```

- **Changes the default local preference value**
- **The specified value is applied to all routes that do not have local preference set (EBGP routes)**
- **The default value of this parameter is 100, allowing you to specify more desirable or less desirable routers**

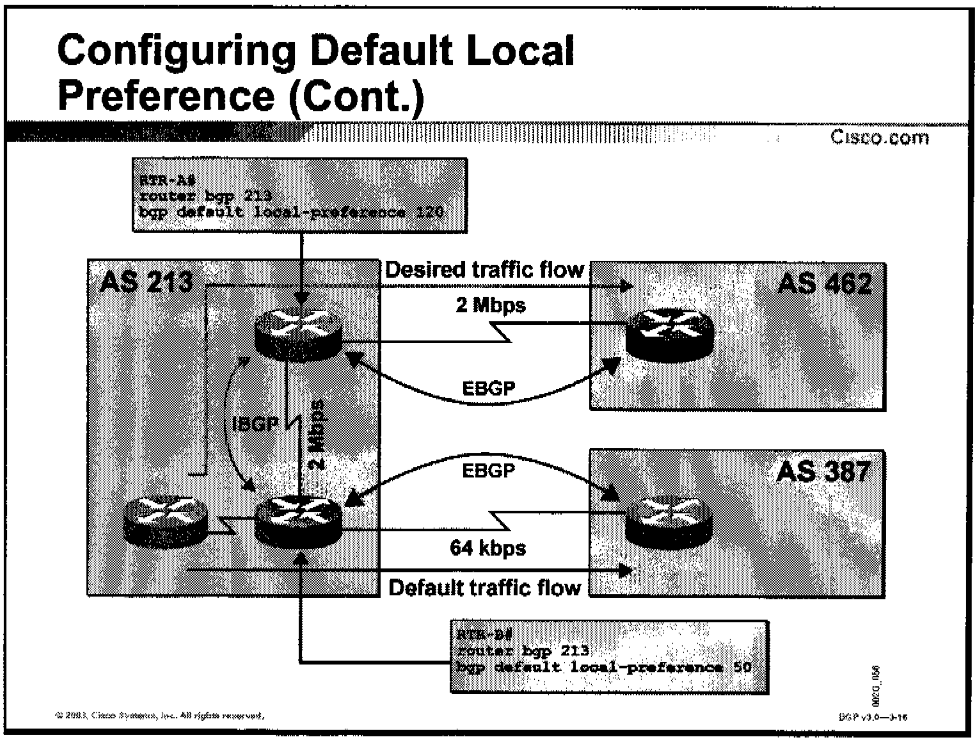
BGP v3.0—3-14

You can use the **bgp default local-preference** command in BGP configuration mode to change the default value of local preference. The new default value applies only to locally originated routes and those received from external neighbors.

Setting a value lower than the default of 100 will result in the router preferring internal paths to external (normally a router would prefer external routes).

Setting a value higher than 100 will result in external paths being preferred to all internal paths (also those with a shorter AS path).

Example



In this example, the local preference attribute is used instead of weights. The two indicated routers in AS 213 have different default local preference values that are applied to external updates. The bottom router receives updates from the external neighbor and applies local preference to them. The same router then receives updates from the upper router, which set a local preference of 120 to all external updates. The bottom router then compares all paths and, where two paths exist, chooses the one with the higher local preference (120).

Practice

Q1) What is the default value of local preference?

A) 0

B) 100

C) 255

D) 32768

Configuring Local Preference with Route-Maps

This topic lists the Cisco IOS commands required to configure BGP local preference using route-map statements.

Configuring Local Preference with Route-Maps

Cisco.com

router(config)#

```
route-map name permit sequence
  match condition
  set local-preference value
```

- **Changes BGP local preference only for routes matched by the route-map entry**

router(config-router)#

```
neighbor address route-map name in | out
```

- **Applies route-map to incoming updates from specified neighbor or outgoing updates to specified neighbor**
- **Per-neighbor local preference is configured by using a route-map with no match condition**

 BGP v3.0—3-17

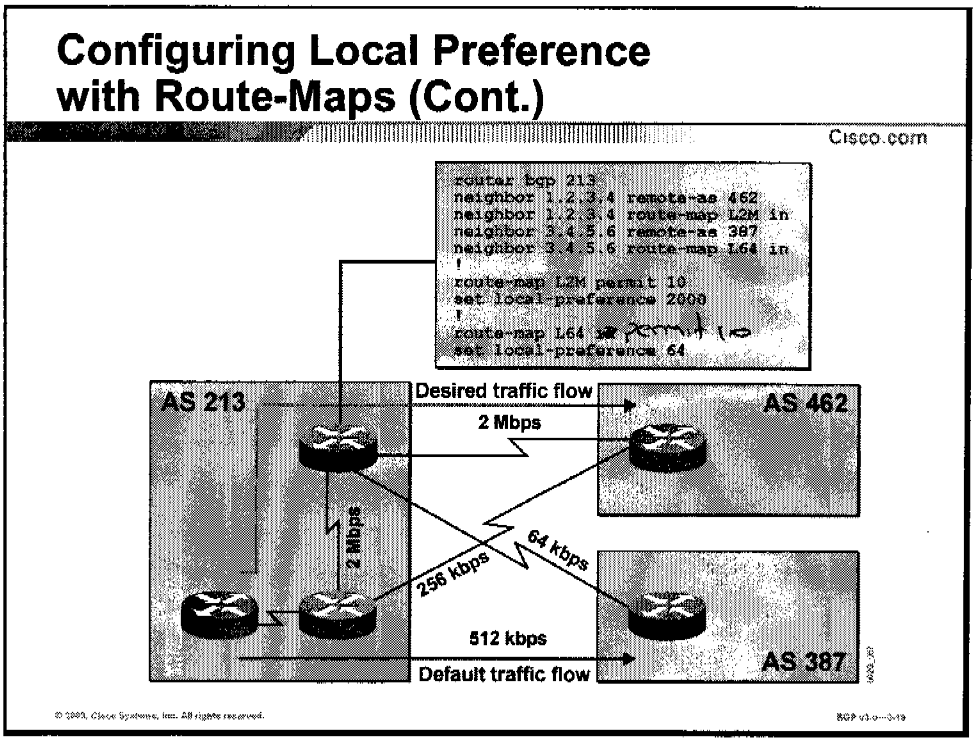
To have more control over setting local preference, you may be forced to use a route-map. A route-map can have more statements, each with a different **set local-preference** command and a different match condition. If there is no **match** command, the route-map statement will apply local preference to all routes. The route-map can then be applied to BGP route updates in either the incoming or outgoing direction.

Note Applying a route-map to outgoing updates on external sessions will have no effect on local preference in the neighboring AS.

When routers use a route-map to set local–preference, the route-map is typically applied to incoming BGP routes advertised by an EBGP neighbor. The local router uses the local preference attribute in BGP route selection. In addition, the router also propagates the attribute to all Internal Border Gateway Protocol (IBGP) sessions in the local AS. Normally, no modifications of local preference are made on IBGP sessions. This restriction ensures that all routers in the local AS will use the same local-preference value and make the same decision in the route selection process.

Note If a network is not matched in any of the route-map statements, the network will be filtered. To permit unmatched networks without setting the local preference attribute, another route-map statement without **match** and **set** commands should be added to the end of the route-map. This statement should simply permit the remaining networks.

Example



In this example, both routers have two external sessions. Using the **bgp default local-preference** command is no longer possible because the second fastest link is on another router.

The configuration here sets local preference according to the bandwidth of the link. A similar configuration exists on the bottom router. If the primary (2-Mbps) link fails, the paths learned through the bottom router in AS 213 (routes with a local preference of 512) will be used.

Practice

Q1) What effect does a route-map have on the local preference setting of outgoing EBGP updates?

A) The local preference can be set for outgoing EBGP updates to notify the neighboring AS about desired path selection.

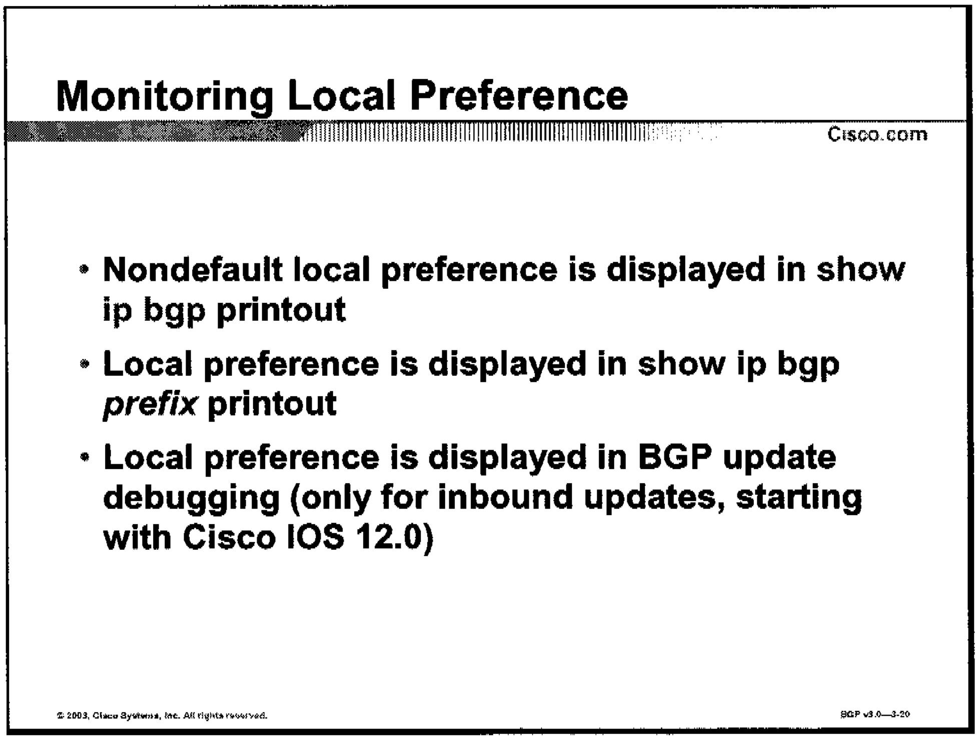
B) Route-maps cannot be used to apply changes to local preference in the outbound direction.

C) You can configure local preference only on IBGP neighbors.

D) Applying a route-map to outgoing updates on EBGP sessions does not affect local preference in the neighboring AS.

Monitoring Local Preference

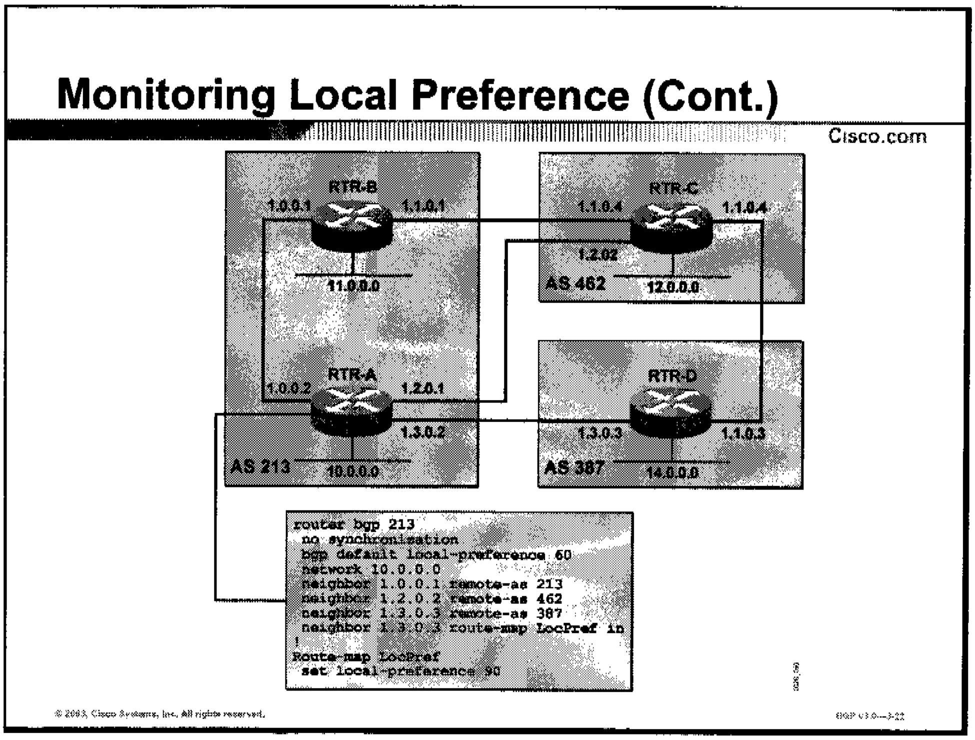
This topic lists the Cisco IOS commands necessary to monitor BGP local preference.



Although local preference is not a mandatory attribute, it is applied to every route. When you are using the **show ip bgp** command, a locally applied default value is not shown. All other values are displayed. You should use the command **show ip bgp** *prefix* to also display the locally applied value.

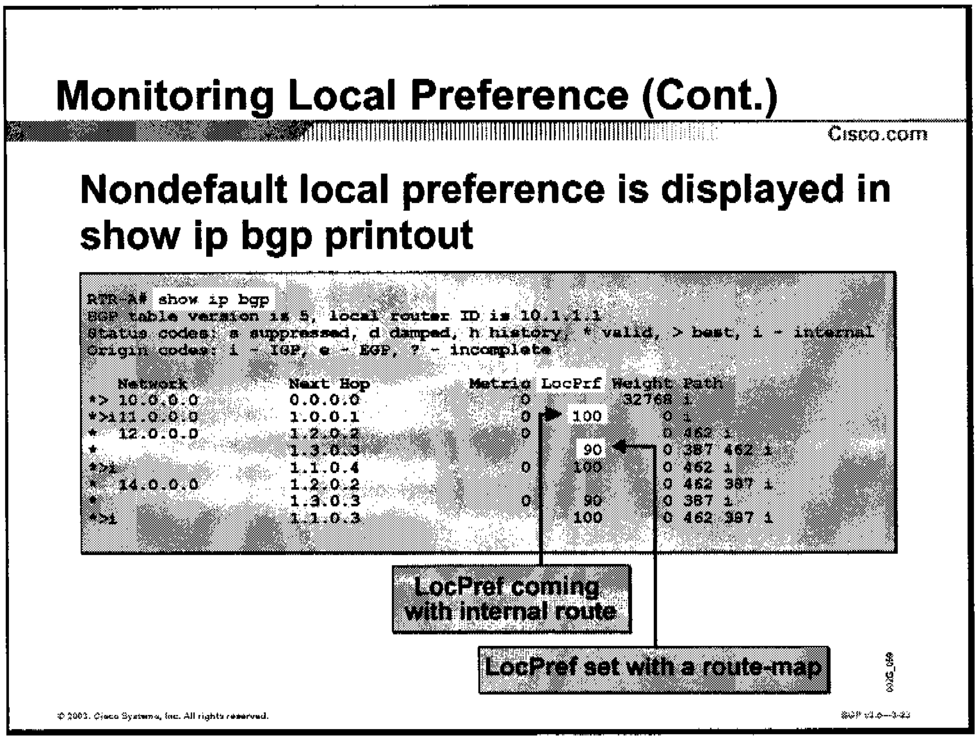
The output displayed from **show** and **debug** commands will vary depending on the Cisco IOS version. Newer versions typically display more information. In Cisco IOS version 12.0 and in later versions, enabling debugging of incoming routing updates will also display the local preference attribute.

Example



The network displayed in the figure was used to collect output from the **show** and **debug** commands in the next few examples. Every physical connection also includes a BGP session. All monitoring and troubleshooting commands were used on router RTR-A.

RTR-A has one internal and two external neighbors. RTR-B is setting local preference 100 to all updates, and RTR-A is setting a default local preference (value 60) for all external updates except for those coming from router RTR-D, where a route-map is used to set a local preference of 90. The following pages show the output of **show** and **debug** commands on router RTR-A.



The output shown in the figure contains routes with three different local preference values:

- Network 10.0.0.0/8 is locally originating on RTR-A, and the applied default local preference 60 is not displayed.
- The second path for network 12.0.0.0/8 was received from RTR-D and received a local preference value of 90 by the route-map.
- All routes received from router RTR-B are marked as internal and have a local preference value of 100 set on RTR-B.

Note	The output of the **show ip bgp** command will not display the local preference value if the value is the same as the **bgp default local-preference** value in the local router. In the example, RTR-B was using its default local preference value (100). But when these routes propagated to RTR-A, then RTR-A displays the local preference value of 100 because it is different from the default local preference value configured on RTR-A.

Monitoring Local Preference (Cont.)

Cisco.com

All values for local preference are displayed in show ip bgp *prefix* printout

Default local preference is displayed

```
RTR-A# show ip bgp 12.0.0.0
BGP routing table entry for 12.0.0.0/8, version 4
Paths: (3 available, best #3)
  462
    1.2.0.2 from 1.2.0.2 (12.1.2.3)
     Origin IGP, metric 0, localpref 60, valid, external, ref 2
  387 462
    1.3.0.3 from 1.3.0.3 (14.1.2.3)
     Origin IGP, localpref 90, valid, external, ref 2
  462
    1.1.0.4 (metric 41024000) from 1.0.0.1 (11.0.0.1)
     Origin IGP, metric 0, localpref 100, valid, internal, best, ref 2
```

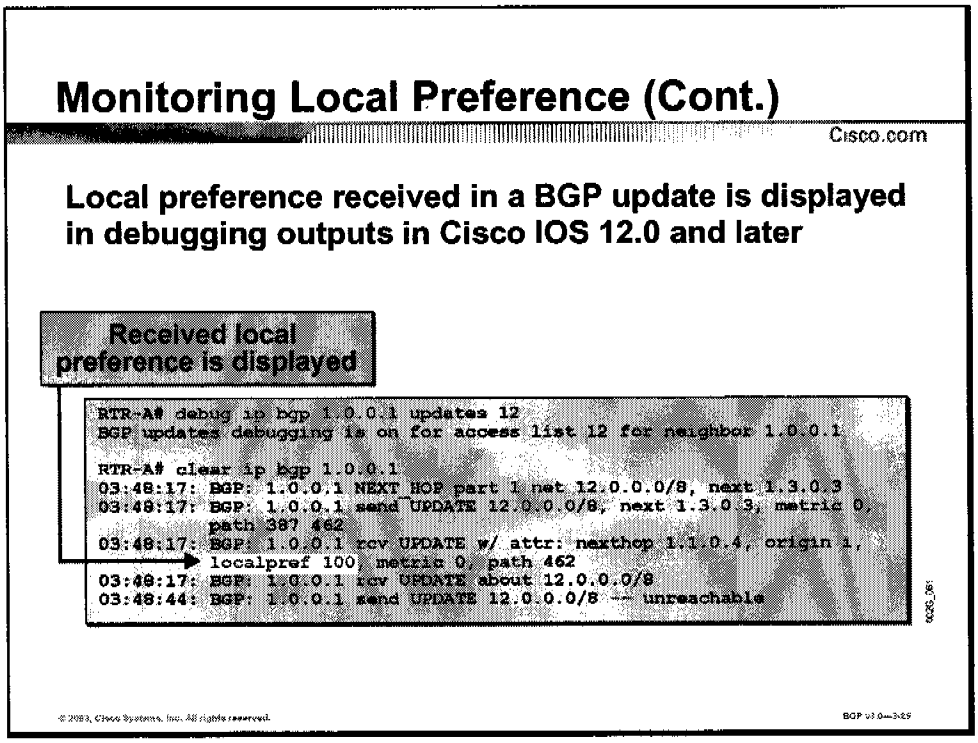

BGP v3.0—3-24

Use the **show ip bgp** *prefix* command to see more detailed information about a specific network, including the locally applied default local preference.

In this example, there are three different paths to reach the same network:

- The first path is external and was received from router RTR-C. The new default local preference value 60 was applied to the update.
- The second path is external and is received from router RTR-D. The route-map was used to set a local preference of 90.
- The third path is internal and was received from RTR-B. The update already contained a local preference attribute with a value of 100.

Router RTR-A chose the last path as best because it has the highest local preference.



This slide shows debugging output of incoming BGP updates. Since a router propagates the local preference attribute to other routers in the same AS only, local preference will be associated with routes sent from internal neighbors.

Practice

Q1) What command should you use to display locally applied local preference settings?

A) **show bgp preference detail**

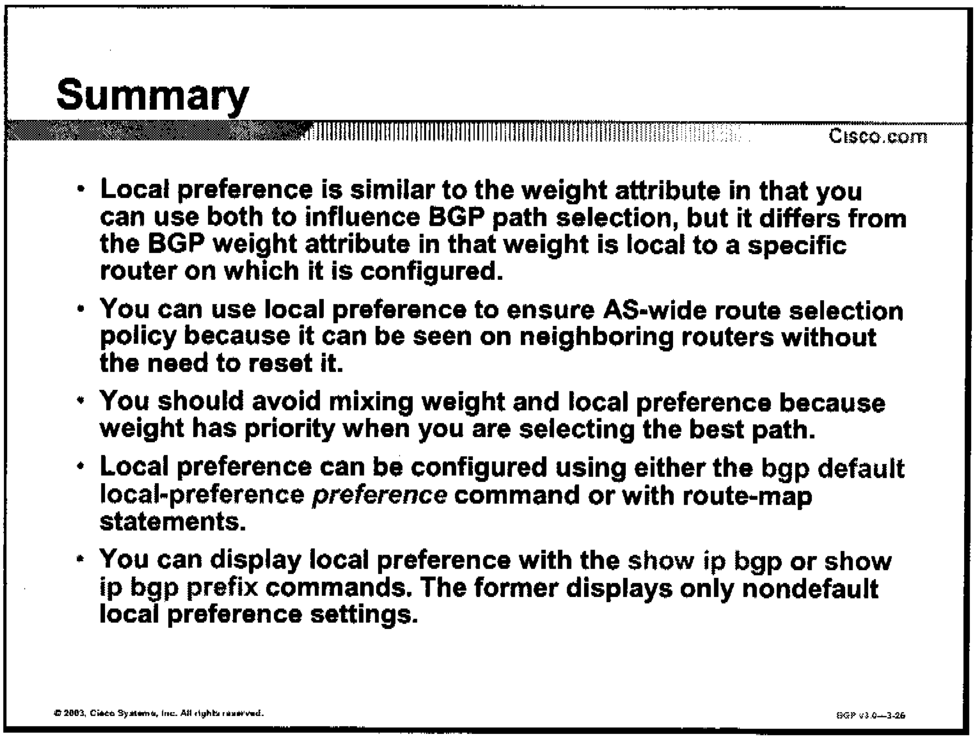
B) **show ip bgp**

C) **show ip bgp detail**

D) **show ip bgp** *prefix*

Summary

This topic summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- AS-Path Prepending lesson

References

For additional information, refer to these resources:

- For more information on BGP local preference, refer to "Border Gateway Protocol" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For more information on BGP local preference, refer to "Using the Border Gateway Protocol for Interdomain Routing" at the following URL: http://www.cisco.com/warp/public/459/14.html

Laboratory Exercise: BGP Local Preference

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the local preference attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure BGP local preference using route-maps
- Monitor BGP local preference

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
no synchronization	Disable BGP synchronization.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
set local-preference *num*	Use this command within a route-map to set the local preference attribute.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
clear ip bgp	Restart the BGP session with your BGP neighbor.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You want to establish two links with the "Good" service provider to increase the reliability of your Internet service. With several links connecting you to the same service provider, you must use local preference in your AS to ensure consistent AS-wide routing policy.

- In this exercise, you will establish the second link toward the "Good" service provider and use the local preference attribute to select the newly established link as the preferred exit point from your network.

- The additional link that you will establish connects WGxR2 and the "Good" router through the Frame Relay network. You will configure a BGP session (private peering) between WGxR2 and "Good" over this link. All traffic from your AS toward "Good" should flow over this link.

- On WGxR2, use data link connection identifier (DLCI) 20x and IP address 192.168.3x.1/30 for the Frame Relay link connection. The other side of this permanent virtual circuit (PVC) is connected to router "Good," which is already configured.

- The BGP routing design contains the following items:

  - AS x should prefer AS 20 as the upstream service provider. Router WGxR2 should be used as the exit point under normal circumstances.

  - Peering to AS 20 through WGxR1 should be used only if the primary link fails.

- Figure 1 shows the new physical connectivity, BGP sessions, and the expected traffic flow in the network.

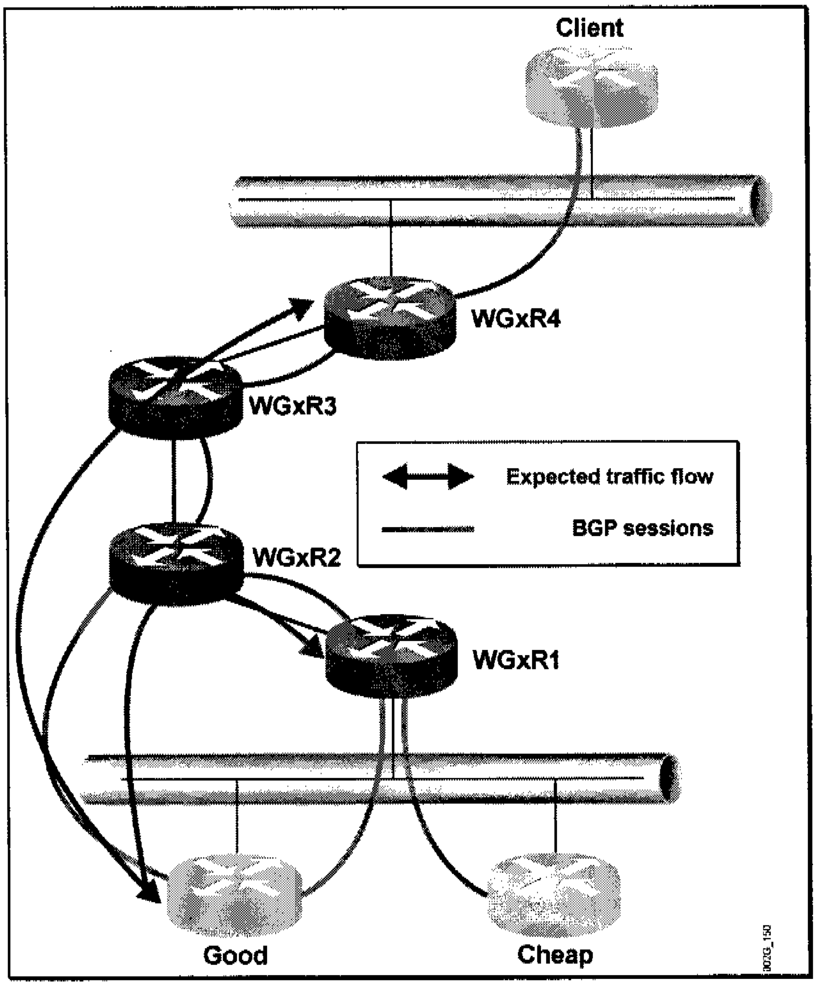


Figure 1: BGP local preference in a service provider network

Exercise Procedure

Complete these steps:

Configure additional IP connectivity:

Step 1 Create another point-to-point subinterface on WGxR2 using the parameters from the following table:

Router	IP address	DLCI
WGxR2	192.168.3x.1/30	20x

192.168.1.1

Establish EBGP peering between WGxR2 and router "Good:"

Step 2 Configure the router "Good" as the BGP neighbor using the parameters from the following table:

Parameter	Value
Service provider IP address	192.168.3x.2
Service provider AS number	20

Establish IBGP peering between WGxR1 and router WGxR2:

Step 3 Establish an IBGP session between routers WGxR1 and WGxR2.

Step 4 Disable BGP synchronization on routers WGxR1 and WGxR2.

Use local preference to ensure AS-wide routing policy:

Step 5 On WGxR1, remove **neighbor weight** statements and any **route-map** statements that modify the weight attribute on WGxR1 for the routers "Good" and "Cheap."

Step 6 Create a new route-map on router WGxR2. Within the route map, set the local preference higher than the default value of 100.

Step 7 Apply the route-map to incoming updates from router "Good."

Step 8 Perform soft clearing of the BGP session by using the **clear ip bgp * [soft] in** command.

Exercise Verification

You have completed this exercise when you attain these results:

- On WGxR2, verify that the local preference has been changed and also note the IBGP routes.

```
wg1r2#sh ip bgp
BGP table version is 29, local router ID is 197.1.3.1
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
*> 10.0.0.0         192.168.31.2             0    200      0 20 i
*> 128.20.0.0       192.168.31.2             0    200      0 20 i
*> 128.20.12.0/24   192.168.31.2             0    200      0 20 i
*> 128.22.0.0       192.168.31.2             0    200      0 20 i
*> 128.22.12.0/24   192.168.31.2                  200      0 20 22 i
*> 128.26.0.0       192.168.31.2             0    200      0 20 42 26
i
*> 128.37.0.0       192.168.31.2             0    200      0 20 42 37
i
*> 128.42.0.0       192.168.31.2             0    200      0 20 42 i
*> 128.51.0.0       192.168.31.2             0    200      0 20 42 26
51 i
*> 128.213.0.0      192.168.31.2             0    200      0 20 213 i
*> 128.214.0.0      192.168.31.2                  200      0 20 22
214 i
*> 192.20.11.0      192.168.31.2             0    200      0 20 i
*> 192.20.12.0/30   192.168.31.2             0    200      0 20 i
*> 192.22.11.0      192.168.31.2             0    200      0 20 i
*> 192.22.12.0/30   192.168.31.2                  200      0 20 22 i
```

```
*> 192.26.11.0      192.168.31.2        0    200       0 20 42 26
i
*> 192.37.11.0      192.168.31.2        0    200       0 20 42 37
i
*> 192.42.11.0      192.168.31.2        0    200       0 20 42 i
*> 192.51.11.0      192.168.31.2        0    200       0 20 42 26
51 i
* i192.168.1.0      192.168.1.1         0    100       0 i
*>                  0.0.0.0             0          32768 i
*>i192.168.31.0     192.168.1.1         0    100       0 i
*> 192.213.11.0     192.168.31.2        0    200       0 20 213 i
*> 192.214.11.0     192.168.31.2             200       0 20 22
214 i
*>i197.1.0.0/21     192.168.1.1              100       0 i
*>i197.1.0.0/16     192.168.1.1              100       0 i
*>i197.1.8.0/22     192.168.1.1              100       0 i
*> 200.20.0.0/16    192.168.31.2        0    200       0 20 i
*> 200.22.0.0/16    192.168.31.2             200       0 20 22 I
```

- Use traceroute from router WGxR4 and WGxR1 to 192.20.11.1.

```
WG1R4#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.9 32 msec 24 msec 20 msec
  2 192.168.1.5 32 msec 45 msec 40 msec
  3 192.168.31.2 56 msec *  52 msec

wg1r1#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.2 40 msec 28 msec 24 msec
  2 192.168.31.2 12 msec *  16 msec
```

- Use traceroute from router "Good" to interface loopback0 on routers WGxR4 and WGxR1.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 wg1 (192.168.20.1) 4 msec 0 msec 4 msec
  2 192.168.1.2 [AS 1] 20 msec 28 msec 24 msec
  3 192.168.1.6 [AS 1] 40 msec 45 msec 36 msec
  4 192.168.1.10 [AS 1] 52 msec *  48 msec

Good#traceroute 197.1.8.1
Type escape sequence to abort.
Tracing the route to 197.1.8.1

  1 wg1 (192.168.20.1) 4 msec *  0 msec
```

- Compare the two outputs of traceroute to determine if the routing is symmetrical. Both traceroute commands should show router WGxR2 in the path.

Answer these questions:

Q1) Is routing between router WGxR1 and "Good" symmetrical?

Q2) Which routers receive the local preference attribute?

AS-Path Prepending

Overview

In networks where connections to multiple providers are required, it is easy to set and control administrative policy for routes leaving an autonomous system (AS) under the same administrative control. A problem arises when administrative policies mandate a specific return path be used for traffic returning to the AS. In this case, administrative policy is difficult to control, because it requires service provider administrators to configure their routers to meet the customer administrative policy. One Border Gateway Protocol (BGP) mechanism that can potentially resolve the administrative policy issue is AS-path prepending. AS-path prepending potentially allows the customer to influence the route selection of its service providers.

This lesson introduces AS-path prepending by describing the problem that AS-path prepending solves. Also included in this lesson are the Cisco IOS® commands required to properly configure and monitor AS-path configurations. In addition, the lesson discusses special filtering requirements for upstream service providers connected to customers wishing to influence route selection using AS-path prepending.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. The "optimum" or "best" route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides many tools for administrators to influence route selection of both incoming and outgoing paths. One of these tools is AS-path prepending.

Objectives

Upon completing this lesson, you will be able to:

- Describe the need to influence BGP return path selection in a service provider environment

- Describe the function of AS-path prepending and how you can use it to facilitate proper return path selection

- Identify design considerations when you are implementing AS-path prepending to influence return path selection
- Identify the Cisco IOS commands required to configure AS-path prepending in a multihomed network
- Identify the Cisco IOS commands required to monitor the operation of AS-path prepending
- Describe concerns with using AS-path filters when neighboring autonomous systems require AS-path prepending

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module

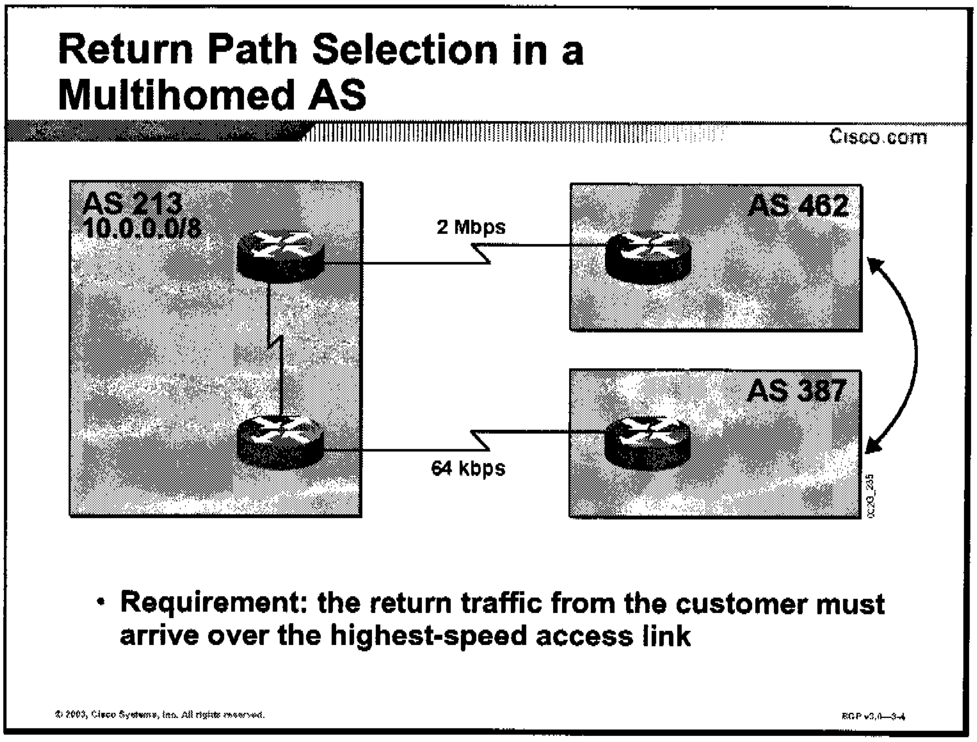
Outline

This lesson includes these topics:

- Overview
- Return Path Selection in a Multihomed AS
- AS-Path Prepending
- AS-Path Prepending Design Considerations
- Configuring AS-Path Prepending
- Monitoring AS-Path Prepending
- AS-Path Filtering Concerns with AS-Path Prepending
- Summary
- Assessment (Quiz): AS-Path Prepending

Return Path Selection in a Multihomed AS

This topic describes the need to influence BGP return path selection in a service provider environment.



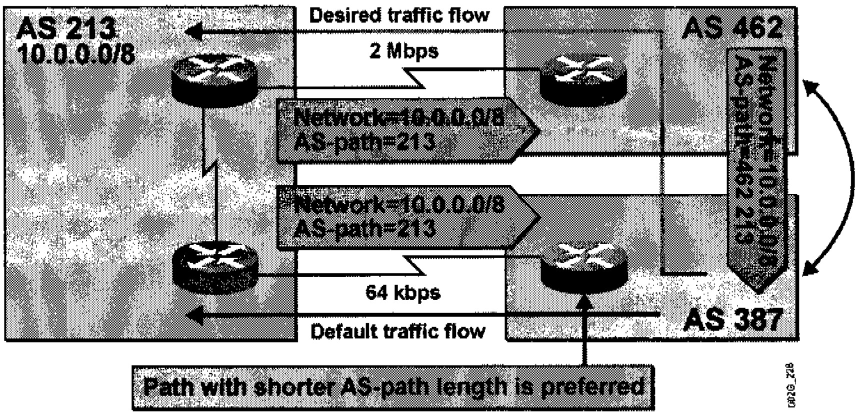
It is fairly easy for an AS to select the appropriate path for outgoing traffic. It is much more complicated to influence other autonomous systems to select the appropriate path for traffic returning to a specific autonomous system.

To configure the preferred path only for outgoing traffic and not for incoming (return) traffic is likely to result in asymmetrical traffic flow as well as suboptimal performance of the return traffic. In the figure, outgoing traffic is directed to the high-speed line (2 Mbps) as a result of configuring local preference or weight. However, the return traffic from AS 387 would take the default path over the low-speed line (64 kbps). The low-speed line would be a limiting factor in the overall performance that the network could achieve.

In this example, AS 213 requests AS 387 to send packets toward network 10.0.0.0/8 via AS 462. The reason for this request is to improve network performance and minimize delay (assuming, of course, that the connectivity between AS 387 and AS 462 is better than the direct 64-kbps link between AS 387 and AS 213).

Default Return Path

Cisco.com



- **Result: the return traffic flows over the path with the shortest AS-path length**

 BGP v3.0—3-9

If no BGP path selection tools are configured on the route to influence the traffic flow, AS 387 will use the shortest AS-path. This action will result in unwanted behavior because the return traffic to AS 213 will be sent over the low-speed WAN link.

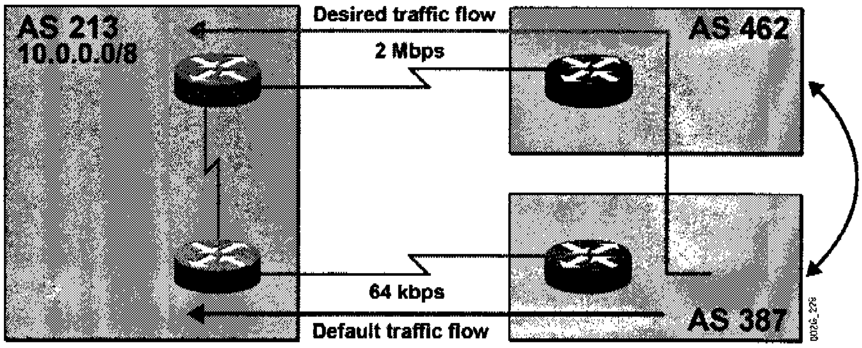
AS 213 announces network 10.0.0.0/8 over External Border Gateway Protocol (EBGP) sessions to both AS 462 and AS 387. When AS 213 sends EBGP updates, it changes the AS-path attribute according to BGP specifications. Both AS 462 and AS 387 receive a BGP update for network 10.0.0.0/8 with the AS path set to 213.

Because AS 462 selects the route for network 10.0.0.0/8 received from AS 213 as its best route, AS 462 will use that route and forward it on to AS 387. According to BGP specifications, AS 462 also changes the AS-path attribute. AS 387 receives the route to network 10.0.0.0/8 from AS 462 with an AS path set to 462 213.

AS 387 has now received two alternative routes to network 10.0.0.0/8 (the direct route from AS 213 and the route through AS 462). Because nothing is configured in AS 387 to influence the flow of traffic, the router will use the BGP route selection rule of shortest AS path to select the best return path to network 10.0.0.0/8.

Proper Return Path Selection

Cisco.com



Q: How do you select the proper return path from AS 387?

A: Use local preference in AS 387

Q: Will the administrator of AS 387 configure it?

A: Unlikely

Remember that the incoming traffic flow (from the perspective of AS 213) will be a result of the route selection for outgoing traffic in AS 387. The traffic that is going out from AS 387 will end up as incoming traffic in AS 213.

If AS 387 configures some changes that cause the route selection process for outgoing traffic to prefer to reach network 10.0.0.0/8 via AS 462, it would result in behavior matching the desired administrative policy for AS 213, which specifies incoming traffic to the AS should be received over the high-speed link.

One way to accomplish the desired administrative policy in AS 213 is to configure the router in AS 387, which is receiving EBGP updates directly from AS 213, to assign a local preference value less than the default value (100) to all routes received from AS 213. The router in AS 387 is also configured specifically not to set local preference on EBGP routes received from AS 462. This results in assignment of the default value of 100 to all routes received from AS 462. When the route selection process in AS 387 selects the best route to reach network 10.0.0.0/8, the difference in local preference values causes AS 387 routers to select the path via AS 462 as the best.

However, all the configuration work to complete this process must be performed in AS 387. The network administrators of AS 387 would be required to modify the router configurations in AS 387 to satisfy the administrative policy requirements of AS 213. All changes must be documented and maintained according to the rules and procedures adopted by AS 387.

If AS 387 is a major Internet service provider (ISP), it is most likely that the network administrators are busy doing things other than tailoring router configurations based on the demand of a single leaf (nontransit) AS that lacks bandwidth on a redundant connection.

BGP Route Selection Rules

Cisco.com

BGP route selection uses the following criteria

- **Prefer largest weight**
- **Prefer largest local preference**
- **Prefer routes that the router originated**
- **Prefer shorter AS paths**
- **Other route selection rules**

Manipulating the outgoing AS-path length could result in proper return path selection

BGP v3.0—3-11

Recall that BGP route selection uses the following criteria:

1. Prefer largest weight
2. Prefer largest local preference
3. Prefer routes that the router originated
4. Prefer shorter AS paths
5. Then prefer all other route selection criteria

It is unlikely that the operator of an AS can request changes in router configurations in another AS. This situation makes it virtually impossible to influence another AS to select the desired path based on the weight and local preference attributes, because both would require configuration changes in the neighboring AS.

But if both the weight and the local preference parameters are left to their default settings, they will not indicate a difference. This situation causes the route selection process to continue down the list of selection criteria. The third criterion for selection will not influence route selection in this scenario, because none of the routes originated at the router performing the route selection. The fourth criterion will apply, however, because the AS paths have different lengths.

If the AS path is not manually manipulated by some administrative means, the path going over the fewest number of autonomous systems is selected by the router regardless of available bandwidth. However, if the AS that is attempting to influence the incoming traffic flow is sending out EBGP updates with a manipulated AS-path attribute over the undesired path, the receiver of this update is less likely to select it as the best because the AS-path now appears to be longer.

The benefit of manipulating AS paths to influence the route selection is that the configuration needed is done in the AS that is requesting a desired return path.

Practice

Q1) What are two important reasons for influencing return path selection in a multihomed AS? (Choose two.)

A) to prevent asymmetrical traffic flow of traffic returning to the AS

B) to prevent network replay and denial-of-service attacks

C) to prevent suboptimal performance of return traffic that prefers to use lower-bandwidth links

D) to eliminate the need to influence outgoing path selection using BGP attributes

Q2) In a multihomed scenario, why does BGP not always use the highest-bandwidth link available?

A) because most administrators fail to configure interface bandwidths using the **bandwidth** command

B) because, ff BGP is not configured to properly redistribute into the IGP, bandwidth metrics are lost from their attached routes

C) because many administrators fail to properly set the metric on each route as a reflection of the link bandwidth

D) because, if not influenced with attributes, the shortest AS path decides which path is used by BGP traffic

AS-Path Prepending

This topic describes the function of AS-path prepending and how you can use it to facilitate proper return path selection.

AS-Path Prepending

Cisco.com

Manual manipulation of AS-path length is called AS-path prepending

AS path should be extended with multiple copies of the AS number of the sender

AS-path prepending is used to

- **Ensure proper return path selection**
- **Distribute the return traffic load for multihomed customers**

 BGP v3.0—3-12

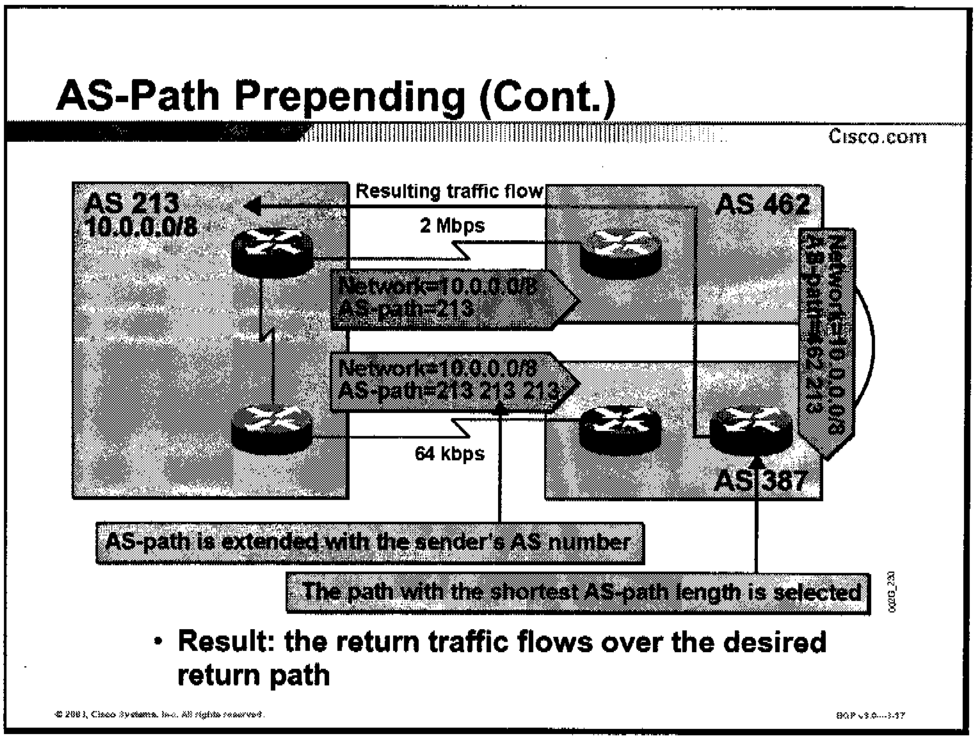
You can manipulate AS paths by prepending AS numbers to already existing AS paths. Normally, you perform AS-path prepending on outgoing EBGP updates over the nondesired return path. Because the AS paths sent out over the nondesired link become longer than the AS path sent out over the preferred path, the nondesired link is now less likely to be used as the return path.

The length of the AS path is extended because additional copies of the AS number of the sender are prepended to (added to the beginning of) the AS-path attribute. To avoid clashes with BGP loop prevention mechanisms, no other AS number, except that of the sending AS, should be prepended to the AS-path attribute.

If another AS number is prepended in the AS path, the routers in the AS that has been prepended will reject the update due to BGP loop prevention mechanisms.

Prepending can be configured on a router for all routing updates sent to a neighbor or only on a subset of them.

Example



In this example, administrative policy in AS 213 prefers that the low-speed link be used for backup purposes only. As long as the high-speed link between AS 213 and AS 462 is available, all traffic should flow toward AS 213 using the high-speed link.

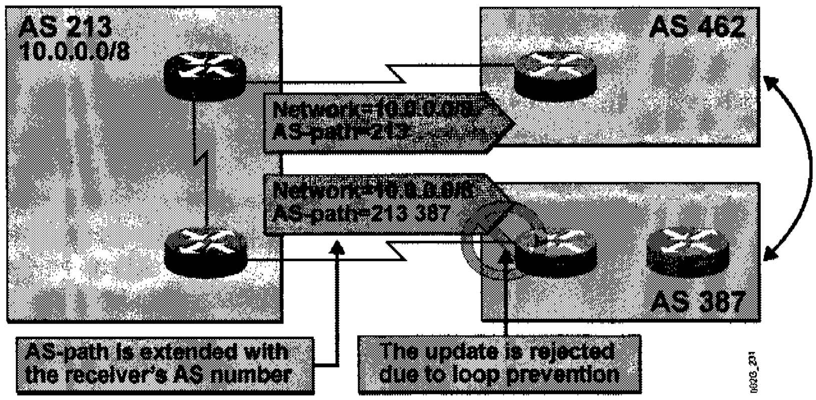
To accomplish this goal, you can configure the router in AS 213 that sends EBGP updates to AS 387 by prepending the AS path with two copies of the AS number 213. AS 387 receives two alternative routes to reach network 10.0.0.0/8: the update received directly from AS 213 (that has a manipulated AS path with a length of three) and the update received via AS 462 (that was not manually manipulated and therefore contains an AS-path length of two).

When AS 387 starts the route selection process to determine which route to use to reach network 10.0.0.0/8, it checks the AS-path length after the weight and local preference parameters. In this case, neither weight nor local preference has been configured, so the length of the AS path will be the deciding factor in the route selection process. Subsequently, AS 387 will prefer the shortest AS path and thus forwards packets toward network 10.0.0.0/8 via AS 462. The desired administrative policy will be met, and AS 213 will receive incoming traffic over the high-speed link.

If the forwarding path from AS 387 via AS 462 to AS 213 and network 10.0.0.0/8 is later broken, the BGP update to reach network 10.0.0.0/8 is revoked. In case of such a network failure, AS 387 will have only one remaining path to reach network 10.0.0.0/8. The route selection process has only one choice, the route directly to AS 213 over the low-speed WAN link. The low-speed link will therefore serve as backup to the high-speed WAN link.

AS-Path Prepending (Cont.)

Cisco.com



- **Prepend the AS path with the AS number of the sender, not the AS number of the receiver.**

 BGP v3.0—3-20

When manually manipulating AS-paths, the only valid AS number that you can prepend is the AS number of the sender. Prepending any other AS number will cause problems.

In the example, AS 213 is prepending AS number 387. The egress router performs AS-path prepending when the route is on its way to be transmitted to AS 387. After the manual manipulation is made, BGP automatically changes the AS path according to the BGP specifications. The local AS number should always be added first when updates are sent over an EBGP session. Therefore, when AS 387 receives the BGP update, the AS path contains the value 213 387. The AS number 387 was set there by the manual manipulation, and the AS number 213 was prepended automatically by BGP because the update was sent over an EBGP session.

When the edge router in AS 387 receives the BGP update, it checks the AS path to verify that the BGP updates were not propagated accidentally by a routing loop. Because the edge router finds its own AS number in the AS path, it assumes that the BGP update has already been in AS 387. According to the BGP specification, the update will be silently ignored.

Now assume that AS 213 had, for the manual manipulation, used yet a different AS number, not its own and not AS number 387. Would AS 387 now have accepted the update? The answer is yes. However, in this scenario, a problem would have appeared at a later stage when the route finally reached the actual AS belonging to the manually prepended AS number. This AS would have rejected the route because it would have found its own AS number somewhere in the AS path.

Practice

Q1) What are two functions of AS-path prepending? (Choose two.)

A) to mask the actual origination point of a route

B) to influence the path that a route takes out of its originating AS

C) to influence the path that a route takes into its originating AS

D) to distribute return traffic load between multihomed routers

Q2) What will happen if you prepend an AS number other than that of the originating AS to the AS path?

A) The upstream AS will automatically reject the route.

B) The route will traverse the Internet and be discarded by the downstream provider.

C) If the route is sent to the AS matching the prepended string, loop prevention mechanisms will silently ignore the update.

D) The AS-path prepending function will allow only the originating AS to be prepended to the AS path.

AS-Path Prepending Design Considerations

This topic identifies design considerations when you are implementing AS-path prepending to influence return path selection.

AS-Path Prepending Design Considerations

Cisco.com

- **There is no exact mechanism to calculate the required prepended AS-path length**
- **If a primary / backup scenario is desired:**
  - **Use a long prepended AS path over the backup link to ensure that the primary AS path will always be shorter**
  - **A long backup AS path consumes memory on every Internet router**
  - **Experiment with various AS-path lengths until the backup link is idle**
  - **Add a few more AS-numbers for additional security (unexpected changes in the Internet)**
- **If traffic load distribution is desired:**
  - **Start with short prepended AS path, monitor link use, and extend the prepended path length as needed**
  - **Continuously monitor the link use and change the prepended AS path length if required**

 BGP v3.0—3-21

How many copies of the AS number of the sender should you prepend to the AS path? The answer depends on the goals of the administrative policy. In the general case, it is not easy to determine the exact number of required AS numbers to prepend. The sending AS does not know what alternative paths are available to other autonomous systems.

Two typical cases where AS-path prepending is used for return path selection:

- Establishing a primary / backup link:

  As an announced backup (prepended) route propagates through the Internet, all the routers along the way that receive the route need to store it together with its AS-path attribute. If this information is long, it will consume extra memory in these routers. However, because routers forward only routes that are selected as "best," an AS that receives multiple alternatives to a destination will select the route with the shortest AS path and forward only that route.

  In the case where both the primary and the secondary link are up, the neighboring AS will receive two routes to the same destination that differ only in the AS path length. The route with the shorter AS path will be subsequently advertised through the Internet.

  In the case where the primary link fails, the route with the longer AS path is the only remaining route. As a result, the primary route is withdrawn, and the prepended route is advertised through the Internet. In this case, extra memory will be consumed in each Internet router due to the storage of the prepended (longer) AS path.

The longer the AS path announced to the EBGP neighbor on the other side of the backup link, the less likely it is that incoming traffic will be received from that neighbor. The network administrator can make a clever guess about how many copies of the AS number to prepend. After the prepending is implemented, the network administrator has to examine the result. If the expected result is not achieved, the configuration can be changed and a few more copies of the AS number can be prepended.

After AS-path prepending has successfully generated the desired results, the network administrator may take a step of precaution by prepending a few more copies of the AS number to the AS path. This action protects the customer from packets being routed over the backup link at a possible later stage when the topology between remote autonomous systems has unexpectedly changed, yielding a longer AS path to reach the primary link.

- Distributing the load of return traffic:

  In a multihomed scenario, there is no way to exactly predetermine the volume of traffic that will be received over a particular link. The traffic load on different links will change depending on where the senders are located (which autonomous systems they belong to). The network topology and the way that different remote autonomous systems are interconnected may also change with time, changing the load distribution. Only constant monitoring and fine-tuning will ensure that the desired results are achieved.

  In a first attempt at load distribution, the network administrator can configure a router connected to an overused link to prepend only a few extra copies of the local AS number. After the network has been given time to converge, the network administrator must check the change in load distribution. Monitoring of the load must be done for a period long enough to be statistically significant (several days or more). If enough volume of traffic has not moved from the overused link to the underused link, the administrator must prepend more copies of the local AS number, and the process of resending local routes and monitoring the results starts all over again.

Practice

Q1) What are two design principles for deciding how many AS number strings to prepend to an AS path? (Choose two.)

A) The AS-path length should be equal to one plus the number of AS-path hops to the destination, if load balancing is desired.

B) The number of AS numbers to append should equal the number of multihomed connections plus one.

C) If load distribution is desired, append as few AS numbers as possible to achieve the desired balancing goal.

D) If a backup link is desired, create a large AS path to ensure that the primary link is always used.

Configuring AS-Path Prepending

This topic identifies the Cisco IOS commands required to configure AS-path prepending in a multihomed network.

Configuring AS-Path Prepending

Cisco.com

router(config)#

```
route-map name permit sequence
  match condition
  set as-path prepend as-number [ as-number … ]
```

- **Prepends the specified AS number sequence to the routes matched by the route-map entry**
- **AS numbers are prepended to the AS path from the BGP table; the AS number of the sender is always prepended to the end result**

router(config-router)#

```
neighbor address route-map name out
```

- **Applies the route-map to outgoing updates sent to the specified BGP neighbor**

 BGP v3.0—3-22

You can configure manual manipulation of the AS-path attribute (prepending) using a route-map with the **set as-path prepend** command. The route-map is used to prepend the specified AS numbers to outgoing EBGP route updates matched with the match condition, as specified in the route-map. AS-path prepending is completed first, and then the route is subject to the normal AS-path modification procedures when it is sent over an EBGP session.

You can also use the route-map to select only a subset of routes that should have their AS path manually manipulated. Use the **set as-path prepend** command with the appropriate **route-map permit** statement.

Note Changing an outgoing route-map affects only the BGP updates sent after the change. In order to propagate the new and longer AS path with all announced routes, the routes must all be resent by the router. To do this, use the privileged EXEC command **clear ip bgp** with the **soft out** qualifier.

set as-path

To modify an AS path for BGP routes, use the **set as-path** route-map configuration command.

set as-path {**tag** | **prepend** *as-path-string*}

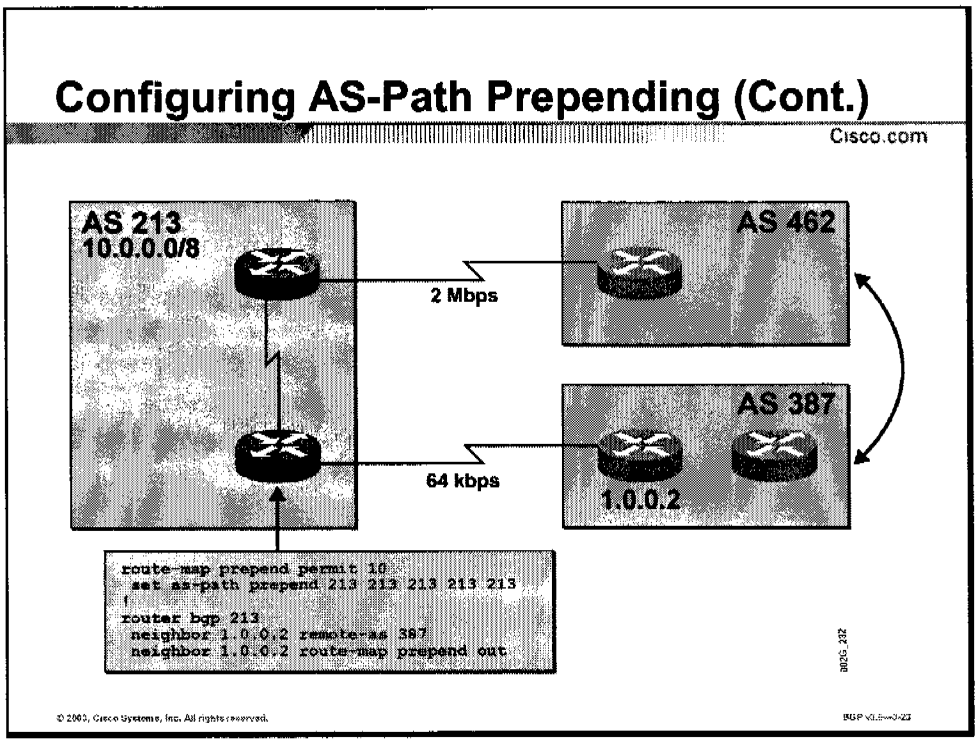
To not modify the AS path, use the **no** form of this command.

no set as-path {**tag** | **prepend** *as-path-string*}

Syntax Description

tag	Converts the tag of a route into an AS path. Applies only when redistributing routes into BGP.
prepend *as-path-string*	Appends the string following the keyword **prepend** to the AS path of the route that is matched by the route-map. Applies to inbound and outbound BGP route-maps.

Example



In this example, the lower router in AS 213 is configured to prepend its own AS number five times for all updates sent to the EBGP neighbor 1.0.0.2 in AS 387. This configuration will result in AS 387 receiving a route to network 10.0.0.0/8 with an AS path containing the AS number 213 six times (213 213 213 213 213 213). Cisco IOS software automatically adds the sixth copy of the AS number as the route leaves AS 213 in accordance with BGP specifications.

The configuration of the AS 213 router is completed in two steps:

Step 1 First, a route-map named "prepend" is created. The route-map selects all BGP routes and prepends five copies of 213 to the existing AS-path attribute already attached to each route. The lack of match conditions in the route-map indicates that all routes are matched.

Step 2 The route-map is applied to all outgoing updates to the EBGP neighbor 1.0.0.2.

Practice

Q1) When you are configuring AS-path prepending with route-maps, when do the changes to the AS path take effect for already announced routes?

A) The changes are applied immediately.

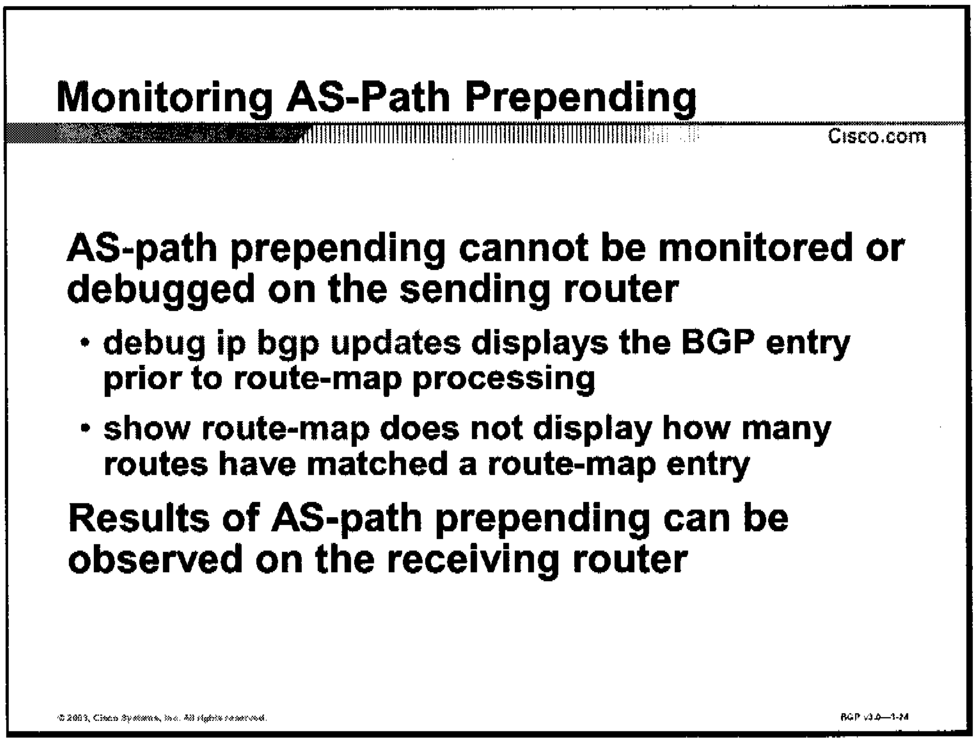
B) The changes will take effect upon clearing the BGP neighbor session of the configured neighbor.

C) The changes will take effect only after reloading the router.

D) Changes are applied only after the upstream AS accepts the AS path.

Monitoring AS-Path Prepending

This topic identifies the Cisco IOS commands required to monitor the operation of AS-path prepending.



When you are monitoring AS-path prepending, the router doing the prepending is not the proper point to observe the results of the AS-path prepend operation. For instance, output from the **debug ip bgp updates** command does not display the prepended paths, because the route-map doing the prepending is applied afterward.

The **show route-map** command displays the configuration details of a route-map. The matching criteria and AS-path manipulation are displayed as output of the command. However, there is no indication of how many routes have been matched by a route-map statement and thus had their AS paths manipulated.

A better place for observing AS-path prepending is on the router receiving the BGP update containing the prepended AS path. At that point, you can use the pattern of AS number sequences in the received AS-path attribute of received routes to find the routes that have a prepended AS path.

Monitoring AS-Path Prepending (Cont.)

Cisco.com

router>

```
show ip bgp regexp regular-expression
```

- **Displays all BGP routes with AS paths matching a regular expression**

```
AS387# show ip bgp regexp ^213 213
BGP table version is 2, local router ID is 1.0.0.2
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
*> 10.0.0.0   1.0.0.1       0              0 213 213 213 i
```


BGP v3.0—3-25

In the figure, the **show ip bgp regexp** command is used to find all the routes in the BGP table of the receiving router in AS 387 that are directly received from AS 213 and have at least one extra copy of AS number 213 in their AS path. Network 10.0.0.0/8 is displayed as the only route meeting this selection criterion. The AS path has been prepended with two extra copies of AS 213. The egress router in AS 213 automatically added the third copy of AS 213 because the route was sent across an EBGP session.

Practice

Q1) Where is the best place to monitor the AS-path prepending operation?

A) on the sending router using the **debug ip bgp updates** command

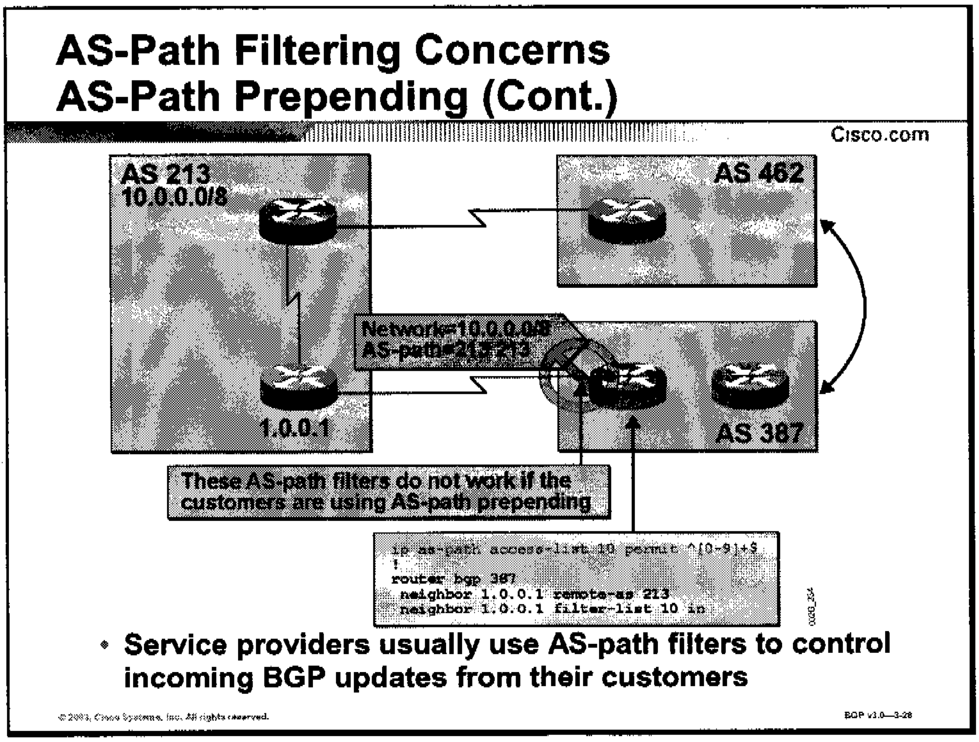
B) on the sending router using the **show route-map** command

C) on the receiving router using the **show ip bgp regexp** command

D) on either the sending or receiving router

AS-Path Filtering Concerns with AS-Path Prepending

This topic describes the concerns of using AS-path filters when neighboring autonomous systems require AS-path prepending.



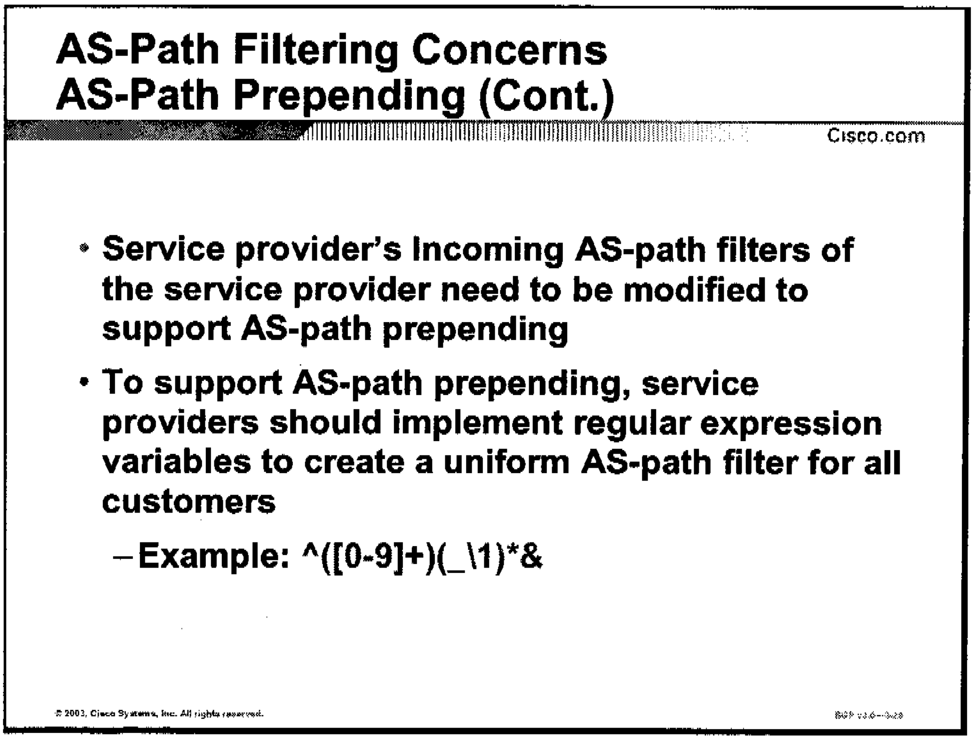
Service providers normally expect their customers to send routes originating only in the AS of the customer. However, because customers might not do so, proactive thinking and care for the rest of the Internet cause the service provider to implement AS-path filters on incoming updates received from their customers.

The network administrator of the service provider in the figure could configure individual filters for each neighbor. However, a single AS-path access-list permitting only AS paths with a length of exactly one AS number would be a better solution because the service provider can uniformly apply it to all incoming routes from all customers (possibly using a peer group).

In the figure, the service provider (AS 387) has configured a filter-list, which allows only AS paths that have a length of one AS number. When the customer changes its router configuration and starts to announce network 10.0.0.0/8 with a prepended AS path, the filter-list for incoming routes to AS 387 in the service provider router will filter those routes out. This filtering results in a situation where the network 10.0.0.0/8 is not reachable over the link between AS 213 and AS 387. Therefore, the backup function is not available.

Network 10.0.0.0/8 is, however, still reachable via the path going through AS 462. This situation means that AS 387 can send packets to network 10.0.0.0/8 but not over the direct link to AS 213. This failure may be hard to detect because, during normal conditions, all autonomous systems in the figure can exchange traffic.

After AS 387 loses the route to network 10.0.0.0/8 via AS 462, possibly because the primary link between AS 213 and AS 462 is gone, the problem will be obvious. AS 387 can now no longer reach network 10.0.0.0/8 at all, although the physical link between AS 213 and AS 387 is available.



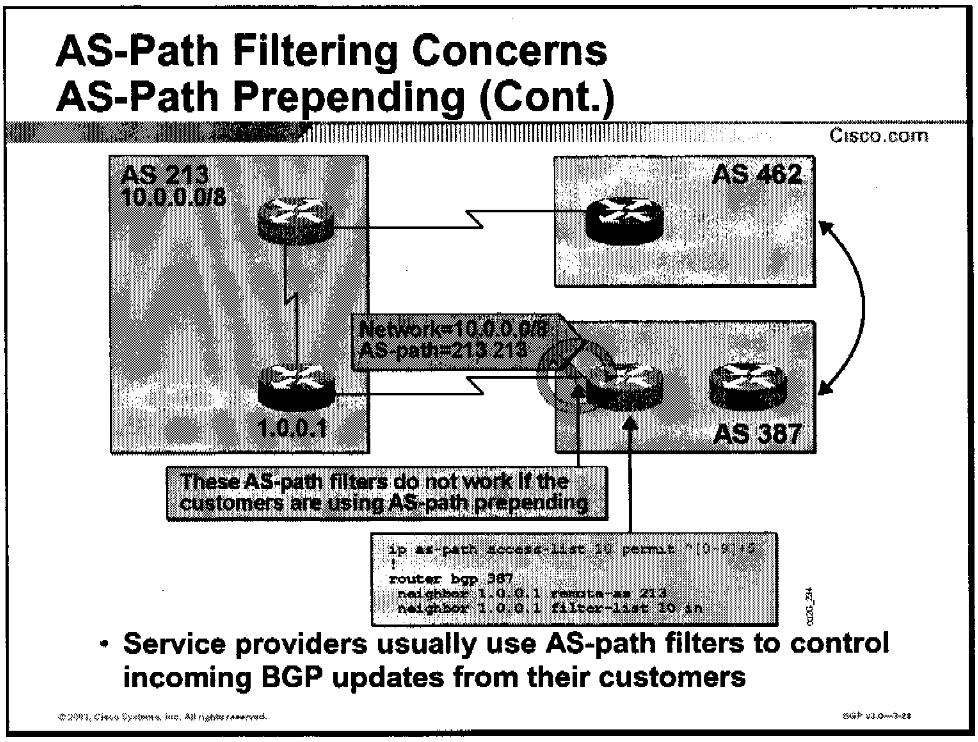
Because the AS of the service provider will receive customer routes with prepended AS-paths that have a length greater than one AS number, the provider must modify its incoming filters.

The service provider needs to create a new inbound regular expression filter, using regular expression variables and parentheses for recall.

What is needed is a filter that will allow any AS path containing one or multiple copies of the same AS number. An example of such a filter is:

- ^([0-9]+)(_\1)*$

This filter matches any AS path beginning with any AS number and continues with none or multiple repetitions of that same AS number (the variable "\1" repeats the value in the brackets). The regular expression would therefore match AS-paths 99 99 99, 2 2 2, or 100, but it would not match AS path 100 99.



In the figure, the service provider (AS 387) has configured an individual filter for all routes received directly from AS 213. The AS path is required to start with 213. Then multiple copies of 213 may follow it. The asterisk allows for zero occurrences, permitting the AS path with a single copy of 213 as well.

If the same service provider router has more customers attached to it, they will all require an individual filter-list because the AS number of the customer is explicitly indicated in the regular expression.

An alternative would be to implement the AS-path filter using regular expression variables.

Practice

Q1) How may the AS of a service provider have to change its configurations when customers manipulate their outgoing AS paths?

A) Upstream providers have to disable all AS-path filtering.

B) The upstream provider must modify any filters based on the AS-path attribute to allow the prepended path.

C) The provider has to create an AS-path filter specific to each customer.

D) Customer-manipulated AS paths typically require no changes in the provider network.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- If the preferred path for incoming (return) traffic is not configured, the likely result is an asymmetrical traffic flow as well as suboptimal performance of the return traffic.
- AS-path prepending is performed on outgoing EBGP updates over the nondesired return path or the path where the traffic load should be reduced.
- You should use a long prepended AS path over the backup link to ensure that the primary AS path will always be shorter. But care should be taken as a long backup AS-path consumes memory.
- Manual manipulation of the AS-path attribute (prepending) is configured using a route-map with the set as-path prepend command.

Summary (Cont.)

Cisco.com

- Monitoring AS-path prepending is best accomplished on the router receiving the prepended routes because the prepended path will not be visible on the prepending router.
- You can use the show ip bgp regexp command can be used to find all the routes on the receiving router with prepended AS paths.
- Service providers with customers using AS-path prepending must create new AS-path filters using specific AS-path entries or with regular expression variables to accommodate AS-path lengths greater than one AS number.

Next Steps

After completing this lesson, go to:

- BGP Multi-Exit Discriminator lesson

References

For additional information, refer to these resources:

- For more information on AS-path prepending, refer to "BGP Case Studies Section 3" at the following URL: http://www.cisco.com/warp/public/459/15.html

Quiz: AS-Path Prepending

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe the need to influence BGP return path selection in a service provider environment
- Describe the function of AS-path prepending and how you can use it to facilitate proper return path selection
- Identify design considerations when you are implementing AS-path prepending to influence return path selection
- Identify the Cisco IOS commands required to configure AS-path prepending in a multihomed network
- Identify the Cisco IOS commands required to monitor the operation of AS-path prepending
- Describe the concerns with using AS-path filters when neighboring autonomous systems require AS-path prepending

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What is AS-path prepending?

A) when a router, sending a BGP update, adds the AS number of the router from which it received the route, to the AS-path attribute

B) when a router, sending a BGP update, adds the AS number of the router to which it is sending the route, to the AS-path attribute

C) when a router, sending a BGP update, adds its AS number to the AS-path attribute multiple times.

D) when a router uses the AS-path attribute in route selection

Q2) AS path will be the route selection criterion used when which of the following is true?

A) It is the first criterion used in BGP route selection.

B) When there is no difference in weight, local preference, or route origination.

C) When the multi-exit discriminator is identical on the candidate routes.

D) The weight, local preference, MED, and origin attributes must be identical before the AS-path attribute is used for route selection.

Q3) What command do you use to manipulate the AS-path attribute?

A) the global configuration command, **set as-path prepend** *as-number*

B) the router configuration command, **set as-path prepend** *as-number*

C) **set as-path prepend** *as-number* in a route-map

D) the interface global command, **set as-path prepend** *as-number*

Q4) Given the following configuration from a router in AS 347, advertising network 11.0.0.0/8 to an EBGP neighbor 2.0.0.2 in AS 529:

```
route-map addAS permit 10
       set as-path prepend 347 347 347

router bgp 347
       neighbor 2.0.0.2 remote-as 529
       neighbor 2.0.0.2 route-map addAS out
```

What are the contents of the AS-path attribute for route 11.0.0.0/8 on a router residing in AS 529?

A) 347 347 347

B) 347 347 347 347

C) 529 347 347 347

D) 529 347 347 347 347

Q5) Why do network administrators need to use AS-path prepending?

A) AS-path prepending allows a customer to potentially influence return path route selection.

B) AS-path prepending is used on a customer router to control outgoing route updates.

C) Service providers use AS-path prepending to control incoming updates from a customer AS.

D) AS-path prepending is used between service providers who are both connected to a customer AS, to determine who will be the primary link to the customer.

Q6) How does AS-path prepending affect a router?

A) AS-path prepending is simply a term used to describe when a router uses the AS-path attribute in route selection and hence does not affect router resources.

B) The longer the AS-path attribute attached to BGP updates, the more router memory requirements increase.

C) AS-path prepending does not impact the router because Cisco IOS software recognizes that AS-path prepending is in use and stores a single AS number with a pointer to the number of AS-path prepends.

D) AS-path prepending causes the router to operate in process-switching mode because the BGP update must be stored, manipulated, and then rewritten to accommodate for the new AS-path attribute.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

BGP Multi-Exit Discriminator

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP multi-exit discriminator (MED) attribute of outgoing BGP routes. The MED attribute is a hint to external neighbors about the preferred path into an autonomous system (AS) when there are multiple entry points into the AS. Two methods used to set the MED attribute are discussed in this lesson as follows: default MED and setting the MED attribute with route-maps. In addition to basic MED attribute configuration, advanced commands to manipulate MED properties are discussed. This lesson also explains how to monitor and troubleshoot the BGP table to verify correct MED configuration and properly influenced path selection.

Importance

When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. It is equally important that the return path selected be the optimum return path into the AS. The "optimum" or "best" route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides a tool for administrators to influence route selection, the MED attribute.

Objectives

Upon completing this lesson, you will be able to:

- Describe how MED can be used to facilitate proper return path selection
- Describe MED propagation inside and between autonomous systems
- Identify the Cisco IOS® commands required to configure default BGP MED on a router
- Identify the Cisco IOS commands required to configure BGP MED using route-maps

- Identify the Cisco IOS commands required to configure advanced MED properties
- Identify the Cisco IOS commands required to monitor BGP MED
- Identify the Cisco IOS commands required to troubleshoot BGP MED

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

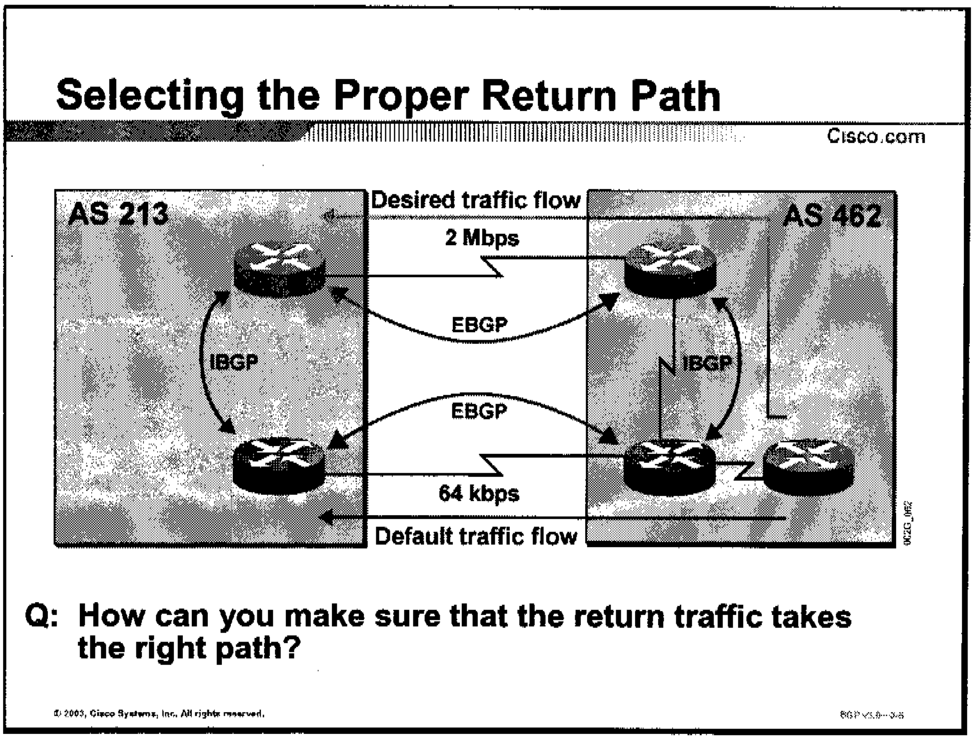
Outline

This lesson includes these topics:

- Overview
- Selecting the Proper Return Path
- MED Propagation in a BGP Network
- Changing Default MED
- Changing MED with Route-Maps
- Advanced MED Configuration
- Monitoring MED
- Troubleshooting MED
- Summary
- Assessment (Lab): BGP Multi-Exit Discriminator

Selecting the Proper Return Path

This topic describes how you can use the MED attribute to facilitate proper return path selection.



When multiple connections between providers are required, BGP attributes such as weight and local preference solve only half the problem: how to choose the right path out of the AS. This topic focuses on the second, more complex half of the same problem: how to influence neighboring autonomous systems to choose the correct return path back into the AS.

Multi-Exit Discriminator

Cisco.com

- You can use MED to influence path selection in neighbor autonomous systems
- An autonomous system can specify its preferred entry point using MED in outgoing EBGP updates
- MED is not propagated outside of a receiving autonomous system
- The default value of the MED attribute is 0
- MED is called "metric" in Cisco IOS Software
- MED is a weak metric
- A lower MED value means more preferred

BGP v3.0—3-6

You can apply the MED attribute on outgoing updates to a neighboring AS to influence the route selection process in that AS. The MED attribute is only useful when there are multiple entry points into an AS.

The MED attribute, sent to an external neighbor, will only be seen within that AS. An AS that receives a route containing the MED attribute will not advertise that MED beyond its local AS.

The default value of the MED attribute is 0. In addition, a lower value of MED is more preferred.

The MED attribute is considered a "weak" metric. In contrast with weight and local preference, a router will prefer a path with the smallest MED value but only if weight, local preference, AS path, and origin code are equal. Using MED may not yield the expected result if the neighboring AS modifies any of the stronger BGP route selection mechanisms.

Note	The term used in Cisco IOS software for MED is "metric." The term "metric" also applies to the **set** command used in route-maps as well as all **show** and **debug** commands.

Practice

Q1) What is the typical application of the MED attribute?

A) to influence path selection out of an originating AS

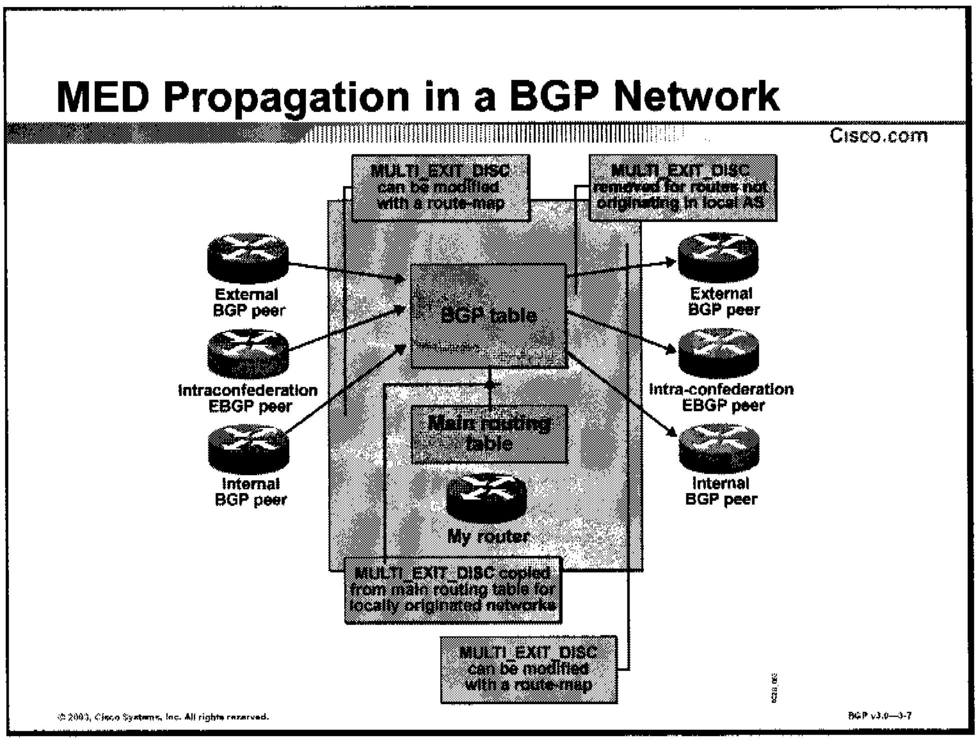
B) to provide a strong metric to select the best path when multiple routes exist

C) to have a BGP attribute traversing many autonomous systems while influencing path selection

D) to influence the return path of traffic back into an AS

MED Propagation in a BGP Network

This topic describes how the value of the MED attribute changes inside a BGP AS and between different BGP autonomous systems.



The figure shows how the value of the MED attribute is assigned depending upon the routing information source. A route-map must be configured on a router to manually assign a value to the MED attribute. For those networks that are also present in the BGP table, the router assigns a default value from the metric in the routing table and copies it into the MED attribute. The MED attribute is automatically removed on external sessions if the attribute did not originate in the local AS.

Practice

Q1) What happens to the MED attribute when it is sent to EBGP peers?

A) The value of MED is ignored and reset to 0.

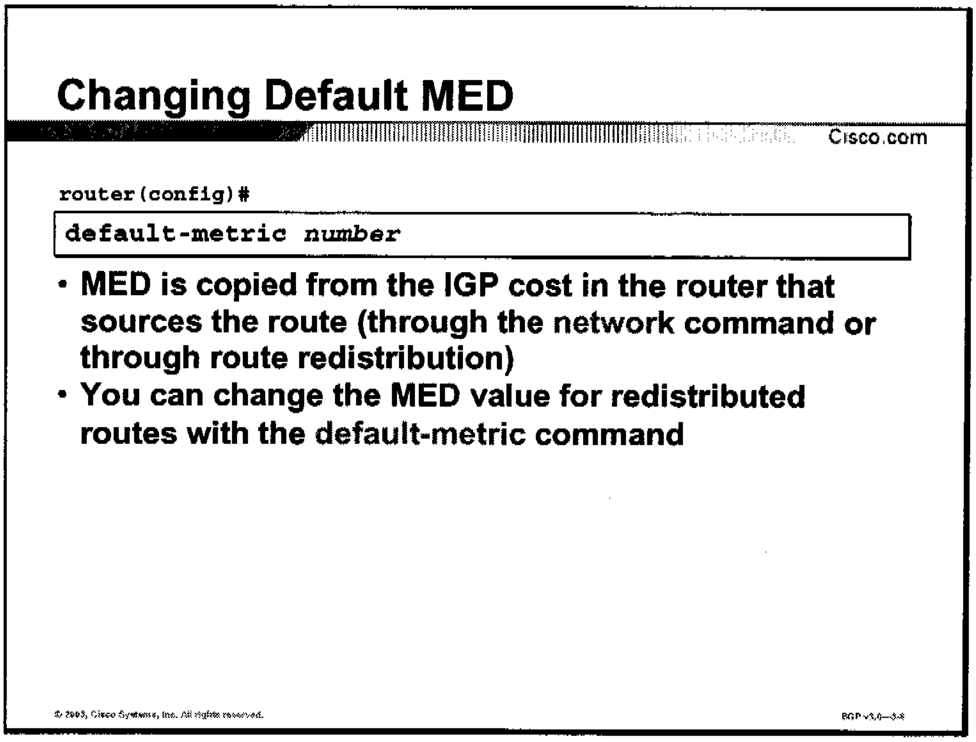
B) All MED values are carried into an AS but are removed by the neighbor AS.

C) The MED is carried into and sent to all EBGP neighbors.

D) MED values originating in the AS are carried into the neighboring AS but do not leave it.

Changing Default MED

This topic lists the Cisco IOS commands required to configure changes to the default BGP MED on a Cisco IOS router.



MED is not a mandatory attribute, and there is no MED attribute attached to a route by default. The only exception is if the router is originating networks that have an exact match in the routing table (through the **network** command or through redistribution). In that case, the router uses the metric in the routing table as the MED attribute value.

Using the **default-metric** command in BGP configuration mode will cause all redistributed networks to have the specified MED value.

default-metric (BGP)

To set the default metric value (MED) for BGP routes, use the **default-metric** command in router configuration mode.

default-metric *number*

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

no default-metric *number*

Syntax Description

number	Default metric value appropriate for the specified routing protocol

Practice

Q1) What is the default value of MED?

A) 0

B) 100

C) 32768

D) MED is unused unless explicitly defined.

Q2) What is the default value of MED when routes are redistributed from another routing protocol?

A) MED is set to 0 for all redistributed routes.

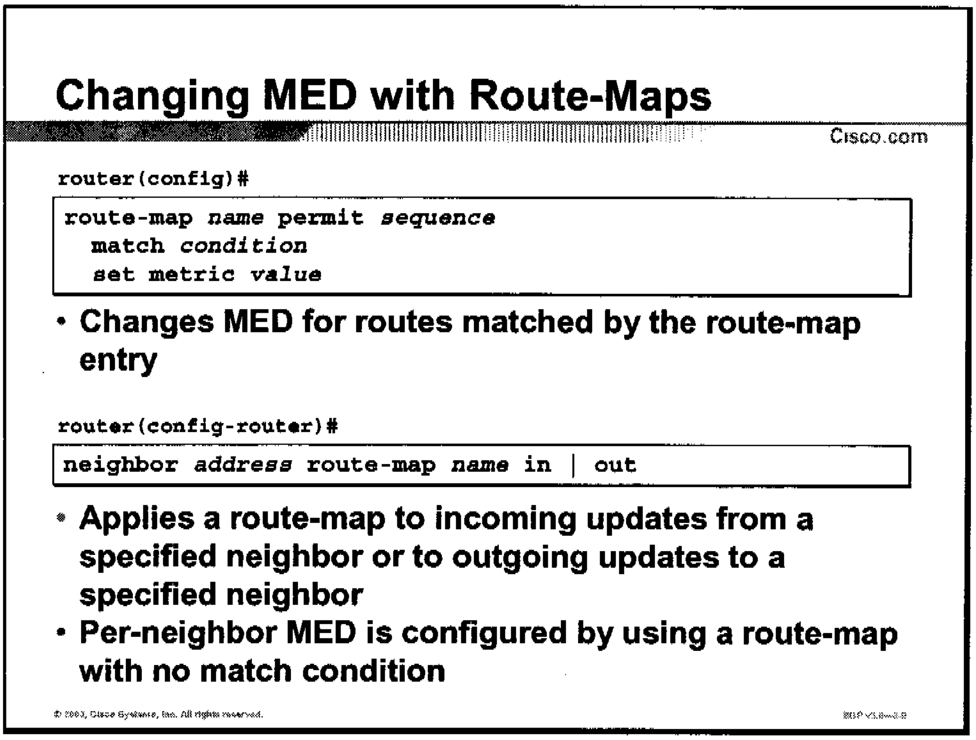
B) MED is set to 100 for all redistributed routes.

C) There is no MED attribute attached to a redistributed route.

D) MED is not attached to redistributed routes except for those matching the IP routing table.

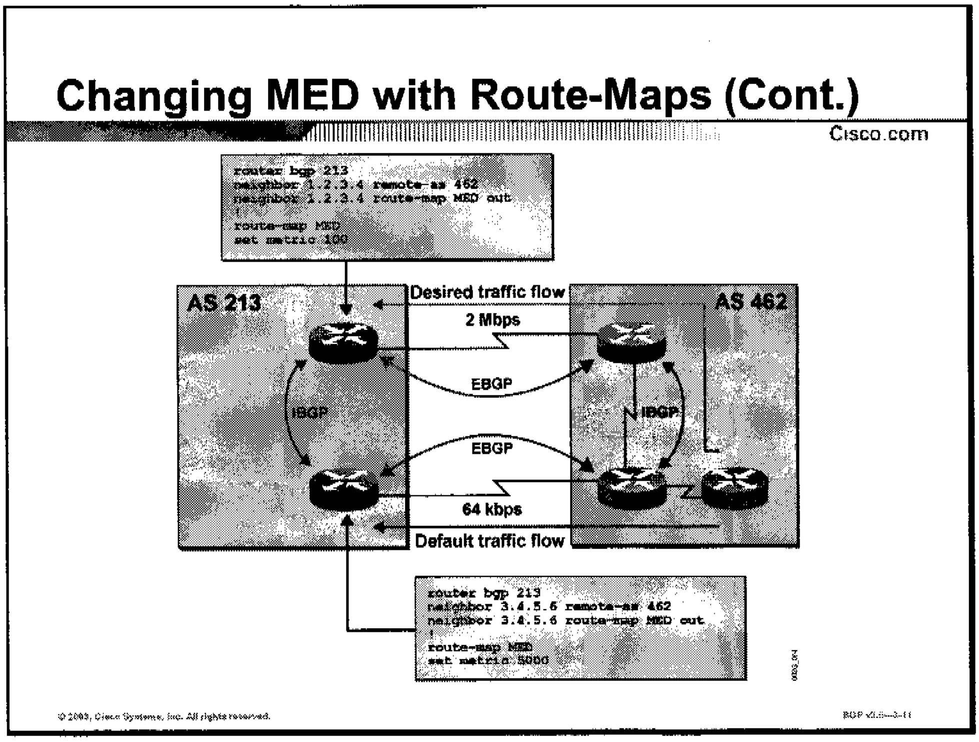
Changing MED with Route-Maps

This topic lists the Cisco IOS commands required to configure changes to the BGP MED attribute using route-map statements.



You can use a route-map to set MED on incoming or outgoing updates. Use the **set metric** command within the route-map configuration mode to set the MED attribute.

Example



The example above shows how to set per-neighbor MED on an outgoing update. The result of this action is that the neighboring AS will prefer the upper link to AS 213. The solution, of course, relies on whether the neighboring AS is not changing the weight, local preference, AS-path, or origin code attributes in updates received from AS 213.

Practice

Q1) What are three BGP attributes that are compared before MED? (Choose three.)

A) largest weight

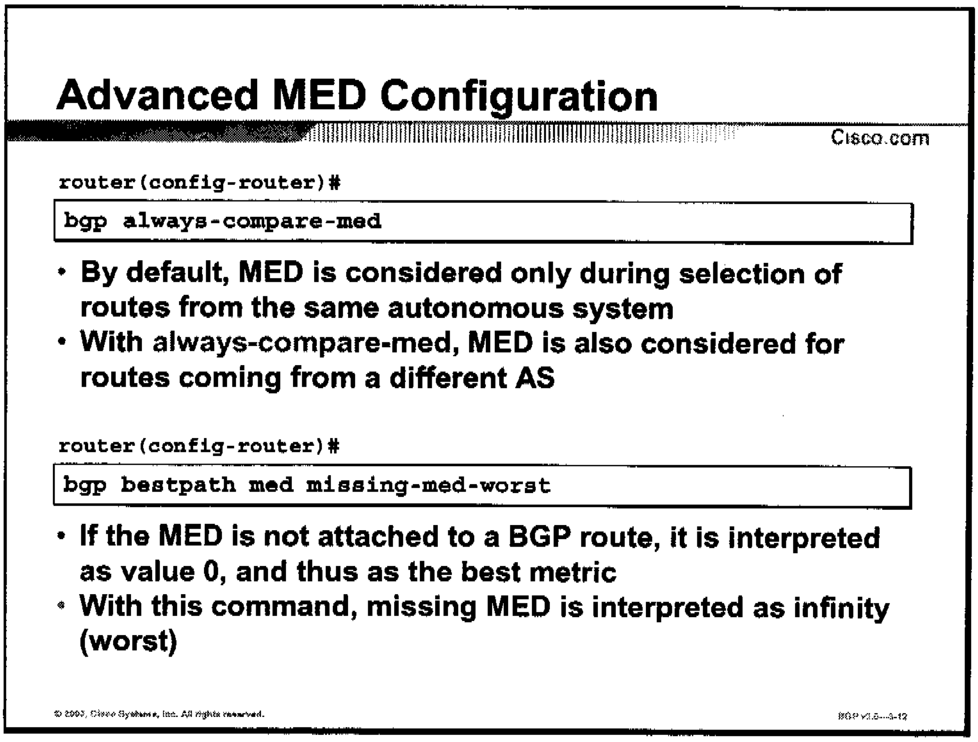
B) originated routes

C) AS-path length

D) lowest IP address

Advanced MED Configuration

This topic lists the Cisco IOS commands required to configure advanced MED features on Cisco routers.



Several rules exist on when and how you should use the MED attribute:

- You should use MED in the route selection process only if both (all) paths come from the same AS. Use the **bgp always-compare-med** command to force the router to compare MED even if paths come from different autonomous systems. You need to enable this option in the entire AS; otherwise routing loops can occur.

- According to a BGP standard describing MED, you should regard a missing MED attribute as an infinite value. Cisco IOS software, on the other hand, regards a missing MED attribute as having value of 0. Use the **bgp bestpath med missing-med-worst** command when combining different vendor equipment. An even better solution is to make sure that every update carries a MED attribute.

Advanced MED Configuration (Cont.)

Cisco.com

router(config-router)#

```
bgp bestpath med confed
```

- **By default, MED is considered only during selection of routes from the same autonomous system, which does not include intraconfederation autonomous systems**
- **Use this command to allow routers to compare paths learned from confederation peers**

router(config)#

```
bgp deterministic-med
```

- **This command changes the BGP route selection procedure to a deterministic but slower one**

 BGP v3.0—3-13

You must use the command **bgp bestpath med confed** when you use MED within a confederation to influence the route selection process. A router will compare MED values for those routes that originate in the confederation.

When enabling deterministic MED comparison, a router will compare MED values before it considers BGP route type (external or internal) and Interior Gateway Protocol (IGP) metric to the next-hop address. The router will compare MED values immediately after the AS-path length.

Note Cisco recommends enabling the **bgp deterministic-med** command in all new network rollouts. For existing networks, you must deploy the command either on all routers at the same time or incrementally, with care to avoid possible Internal Border Gateway Protocol (IBGP) routing loops.

Example

The following example demonstrates how the **bgp deterministic-med** and **bgp always-compare-med** commands can influence MED-based path selection. Consider the following BGP routes for network 172.16.0.0/16 in the order they are received:

entry 1: AS(PATH) 65500, med 150, external, rid 192.168.13.1

entry 2: AS(PATH) 65100, med 200, external, rid 1.1.1.1

entry 3: AS(PATH) 65500, med 100, internal, rid 192.168.8.4

Note BGP compares multiple routes to a single destination in pairs starting with the newest entry and moving toward the oldest entry (starting at top of the list and moving down). For example, entry 1 and entry 2 are compared. The better of these two is then compared to entry 3, and so on.

In the case where both commands are disabled, BGP compares entry 1 and entry 2. Entry 2 is chosen as the best of these two because it has a lower router-ID. The MED is not checked because the paths are from a different neighbor AS. Next, entry 2 is compared to entry 3. BGP chooses Entry 2 as the best path because it is external.

In the case where **bgp deterministic-med** is disabled and **bgp always-compare-med** has been enabled, BGP compares entry 1 to entry 2. These entries are from different autonomous systems, but because the **bgp always-compare-med** command is enabled, MED is used in the comparison. Entry 1 is the better of these two entries because it has a lower MED value. Next, BGP compares entry 1 to entry 3. The MED is checked again because the entries are now from the same AS. BGP chooses entry 3 as the best path.

In the case where **bgp deterministic-med** has been enabled and **bgp always-compare-med** has been disabled, BGP groups routes from the same AS together and compares the best entries of each group. The BGP table looks like the following:

entry 1: AS(PATH) 65100, med 200, external, rid 1.1.1.1

entry 2: AS(PATH) 65500, med 100, internal, rid 192.168.8.4

entry 3: AS(PATH) 65500, med 150, external, rid 192.168.13.1

There is a group for AS 65100 and a group for AS 65500. BGP compares the best entries for each group. Entry 1 is the best of its group because it is the only route from AS 100. BGP compares entry 1 to the best of group AS 65500, entry 2 (because it has the lowest MED). Next, BGP compares entry 1 to entry 2. Because the two entries are not from the same neighbor AS, the MED is not considered in the comparison. The External Border Gateway Protocol (EBGP) route wins over the IBGP route, making entry 1 the best route.

If **bgp always-compare-med** were also enabled, BGP would have taken the MED into account for the last comparison and have selected entry 2 as the best path.

Practice

Q1) What effect does the **bgp deterministic-med** command have on BGP path selection?

A) It forces a MED attribute comparison between MED attributes from different neighbor autonomous systems.

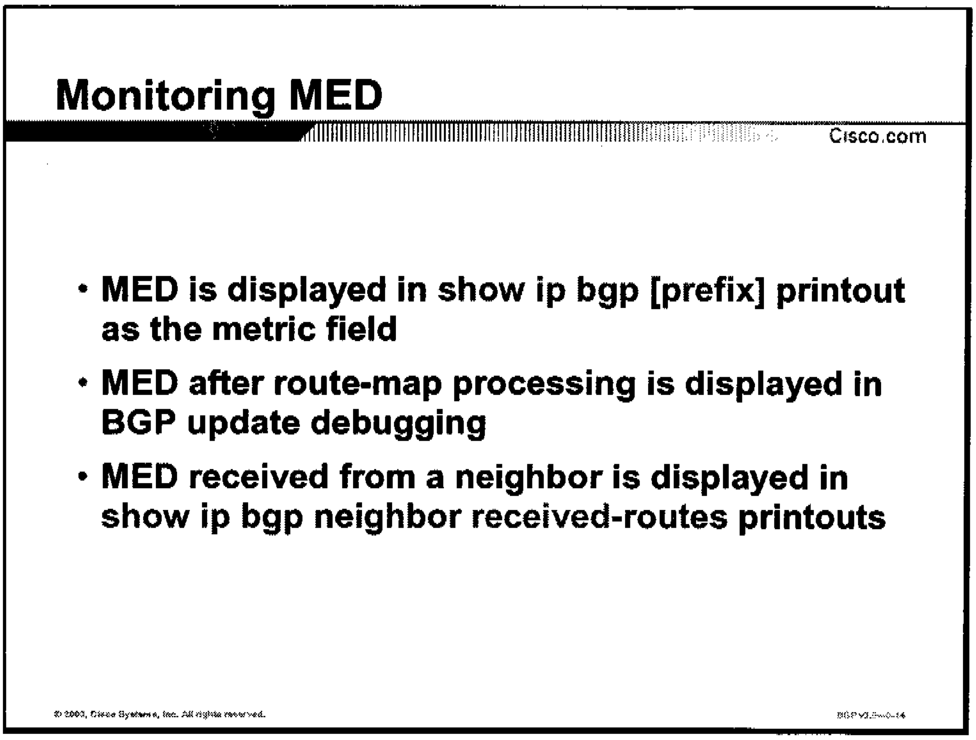
B) It prohibits MED attribute comparison if routes are from a different AS.

C) It ensures that an accurate MED comparison is made across all routes received from the same AS.

D) It allows the MED to determine BGP path selection above other BGP attributes.

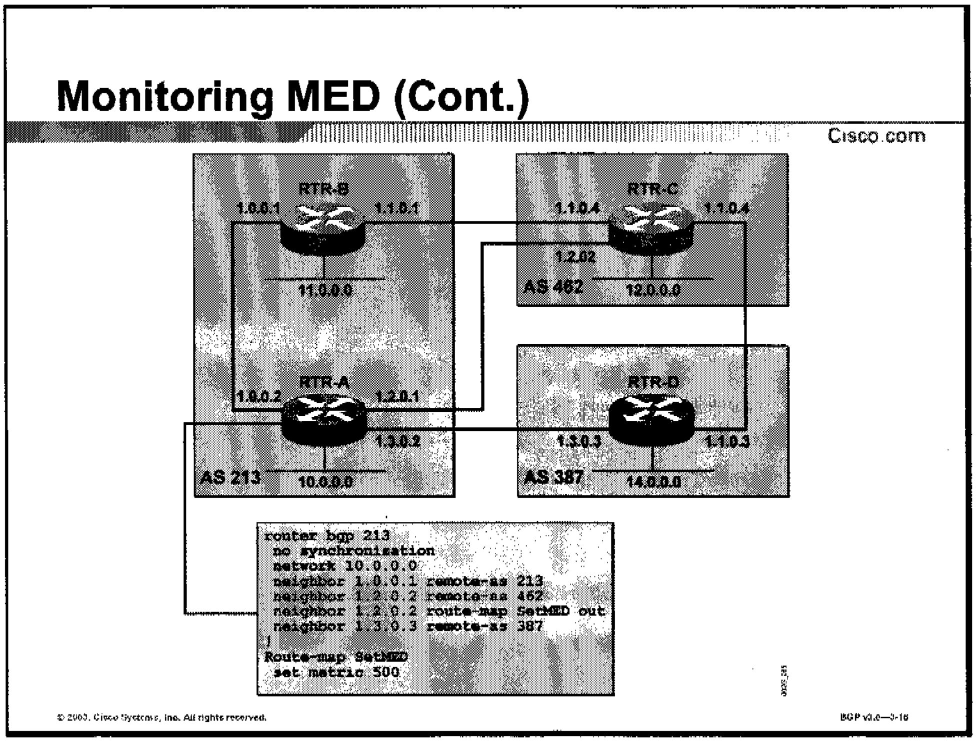
Monitoring MED

This topic lists the Cisco IOS commands required to monitor the BGP MED attribute on a Cisco router.



All BGP-related **show** and **debug** commands display the value of the MED attribute. If the inbound soft reconfiguration feature is enabled on the router, the original MED attribute received by the router is also displayed. The following examples demonstrate command output for Cisco **show ip bgp** commands.

Example



The same network as in the BGP Local Preference lesson is used in this topic to produce sample output shown on the next pages. All commands were executed on router RTR-C.

Some routing updates sent from router RTR-B are sent to router RTR-C with a MED of 500. Some updates sent from RTR-B to RTR-C have the MED set to 0, and some are without a MED attribute. Inbound soft reconfiguration is used on router RTR-C.

Monitoring MED (Cont.)

Cisco.com

- MED is displayed in show ip bgp printout

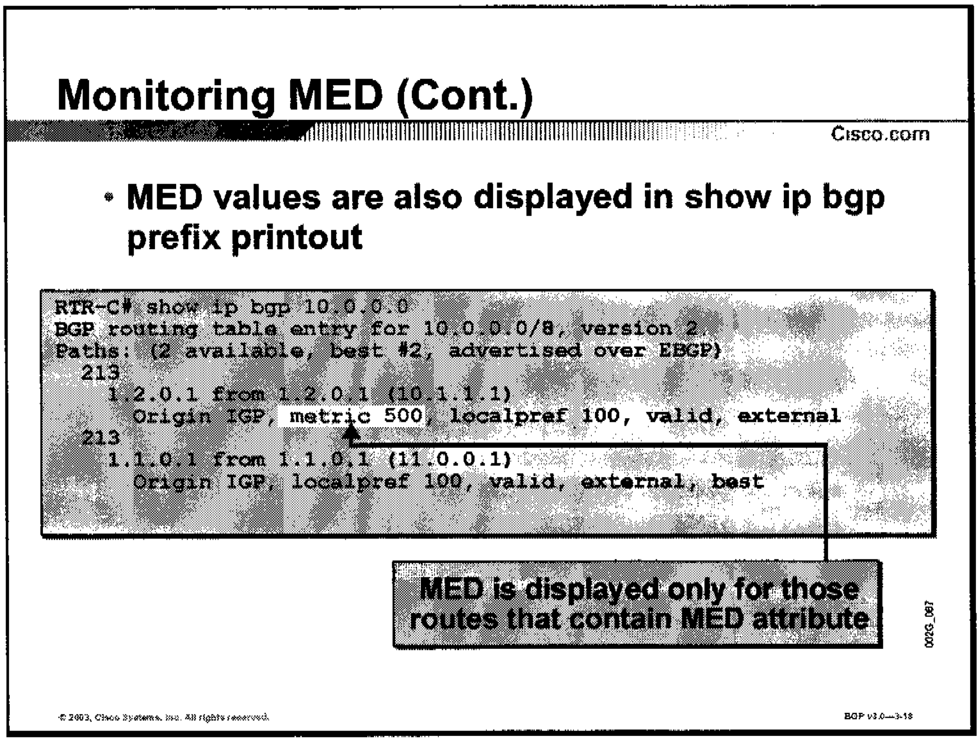
```
RTR-C# show ip bgp
BGP table version is 4, local router ID is 12.1.2.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop          Metric LocPrf Weight Path
*  10.0.0.0         1.2.0.1              500             0 213 i
*>                  1.1.0.1                              0 213 i
*  11.0.0.0         1.2.0.1              500             0 213 i
*>                  1.1.0.1                0             0 213 i
```

MED coming from a neighbor

No MED in this external route

BGP v3.0—3-17

Both networks received from router RTR-B have a MED of 500. Network 10.0.0.0/8 received from RTR-A has no MED attribute while network 11.0.0.0/8 has a MED value of 0.



When looking at detailed information for a specific network, you will see (**show ip bgp** *prefix*) MED only if the attribute exists.

Practice

Q1) Which two statements are true regarding MED in **show ip bgp** commands? (Choose two.)

A) MED is always displayed.

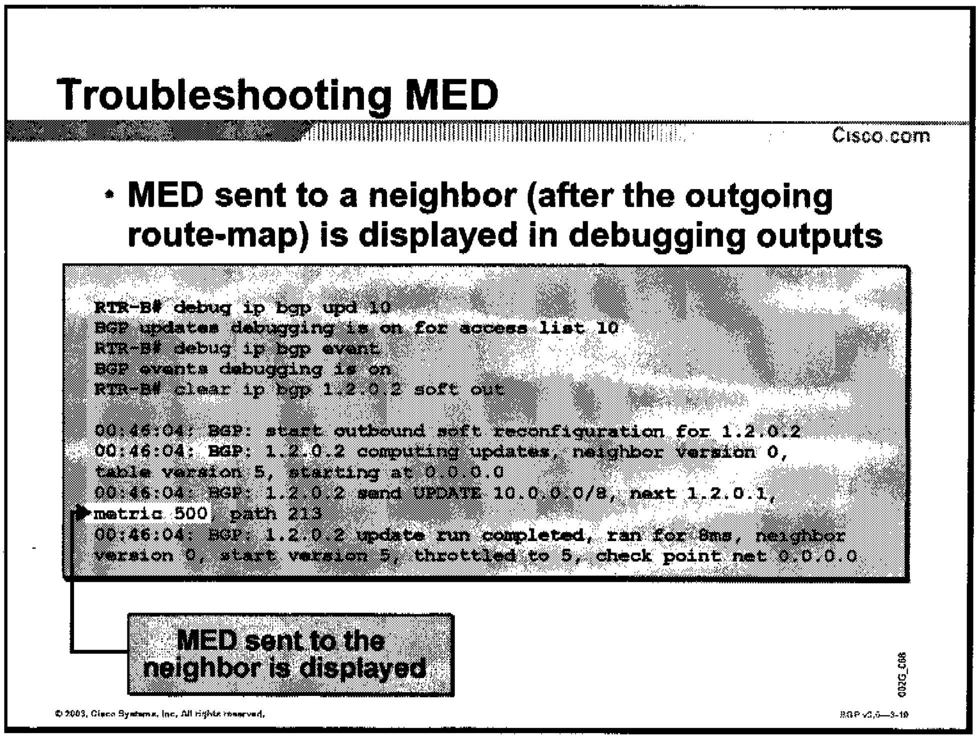
B) MED is displayed only for those routes having a MED attribute.

C) MED is listed as a metric value.

D) MED is listed as a MED value.

Troubleshooting MED

This topic lists the Cisco IOS commands required to troubleshoot BGP MED configurations on a Cisco router.



In case debugging is necessary to troubleshoot a problem, MED, among other attributes, is displayed. This example shows the MED attribute set with an outgoing route-map.

Troubleshooting MED (Cont.)

Cisco.com

- **MED stored in the BGP table (after the incoming route-map processing) is displayed in debugging outputs**

```
RTR-C# debug ip bgp update 10
BGP updates debugging is on for access list 10
RTR-C# clear ip bgp 1.2.0.1

01:03:45: BGP: 1.2.0.1 send UPDATE 10.0.0.0/8, next 1.2.0.2,
metric 0, path 462 213
01:03:45: BGP: 1.2.0.1 rcv UPDATE about 10.0.0.0/8, next hop
1.2.0.1, path 213 metric 500
```

MED stored in the BGP table is displayed

 BGP v3.0—3-20

This debugging example shows the MED attribute value after the update has been processed by an incoming route-map.

Troubleshooting MED (Cont.)

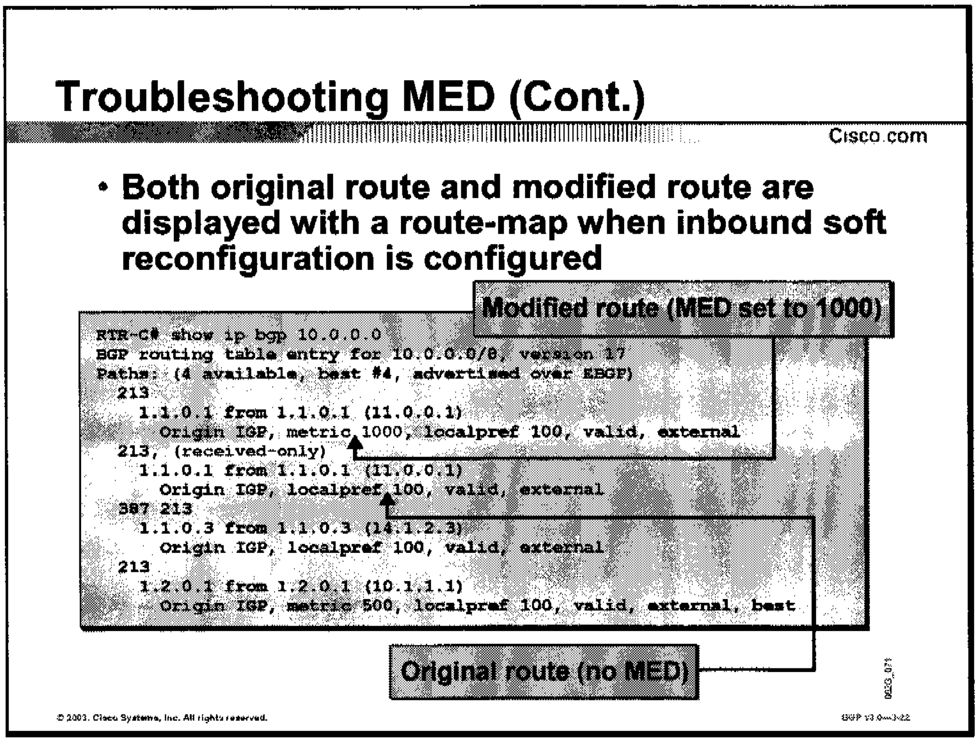
Cisco.com

- **Original MED received from a neighbor (before the incoming route-map processing) is displayed in show ip bgp neighbor received**

```
RTR-C# show ip bgp neighbors 1.1.0.1 received-routes
BGP table version is 19, local router ID is 12.1.2.3
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
*  10.0.0.0         1.1.0.1                               0 213 i
*  11.0.0.0         1.1.0.1                  0            0 213 i

Total number of prefixes 2
```

MED originally received from the neighbor

To see the original MED, you need to enable soft reconfiguration on the router. The command **show ip bgp neighbor *address* received-routes** displays the original updates before any filters or route-maps have filtered or changed them.



If soft reconfiguration is enabled, the original updates to the MED attribute are available by using the **show ip bgp** *prefix* command. The original versions are marked with the **received-only** keyword and follow the version that is in the global BGP table. In the figure, the received update had no MED attribute but was later applied a value of 1000 through a route-map.

Practice

Q1) If you configure inbound soft reconfiguration with a route-map and issue the **show ip bgp** *prefix* command, which value of the MED attribute is displayed?

A) Only the original route (no MED) is displayed.

B) Both the original route and the modified route are displayed.

C) Only the modified route is displayed.

D) The MED attribute is not displayed with the **show ip bgp** *prefix* command.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- In most cases MED is used to achieve symmetrical routing.
- MED is a weak parameter in the route selection process – it is used only if weight, local preference, AS path, and origin code are equal. By default, MED is compared only for paths that were received from the same autonomous system.
- MED is not a mandatory attribute and is normally not present in BGP updates. An exception is when a router originates a network that has an exact match in the routing table – MED is applied a value copied from the metric in the routing table.
- You can use a route-map to set an arbitrary MED value to sent or received routes.

BGP v3.0—3-23

Summary (Cont.)

Cisco.com

- You can configure advanced MED parameters to modify the default MED behaviors. For example, the bgp always-compare-med command forces the router to compare MED even if paths came from different autonomous systems.
- MED is displayed in show commands as the metric field.
- The MED stored in the BGP table after processing the incoming route-map, is displayed in the output of the debug ip bgp update command.

BGP v3.0—3-24

Next Steps

After completing this lesson, go to:

- BGP Communities lesson

References

For additional information, refer to these resources:

- For more information on BGP MED, refer to "Border Gateway Protocol" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

- For more information on BGP communities, refer to "Using the Border Gateway Protocol for Interdomain Routing" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2003.htm#xtocid69

Laboratory Exercise: BGP Multi-Exit Discriminator

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the MED attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure BGP MED using route-maps
- Monitor BGP MED

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
route-map *name* {**permit** | **deny**} **seq**	Define or modify an existing entry in a route-map.
set metric *metric*	Set MED in a route-map.
clear ip bgp *	Reset BGP peer.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You have noticed that the traffic from router "Good" toward your AS sometimes passes through the WGxR1, resulting in asymmetrical routing. You will use MED to indicate to the "Good" router which exit point it should use.

Note This exercise is a continuation of the BGP Local Preference lab exercise.

- The implementation should ensure symmetrical routing. You should use backup peering only when the primary link is down. Use MED to influence the neighboring AS 20 to choose the preferred return path.

- Figure 1 shows the physical connectivity, BGP sessions, and traffic flow in the network.

192.168.71.2

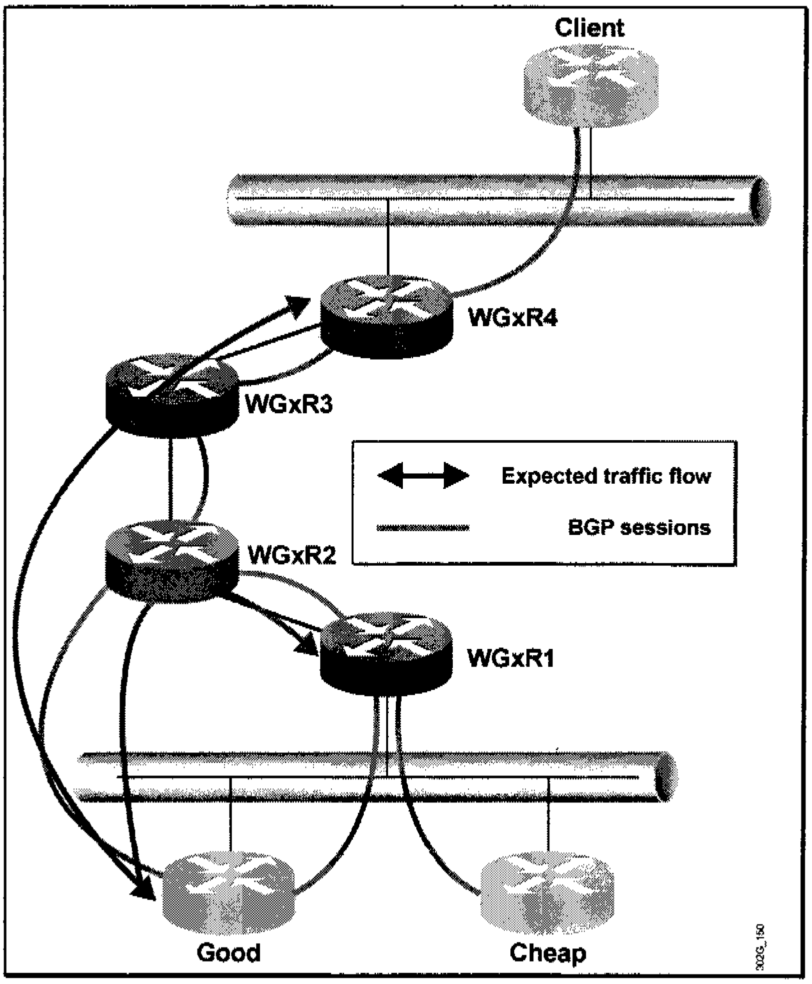


Figure 1: BGP MED in a service provider network

Exercise Procedure

Complete these steps:

Step 1 Before starting the exercise, check the BGP table on router "Good" to see if MED is present in BGP updates received from WGxR1 and WGxR2.

Configuring WGxR1 to set MED:

Step 2 Create a new route-map, which sets the BGP MED to a high value (for example, 200). Apply the route-map to outgoing updates toward the router "Good."

Configuring WGxR2 to set MED:

Step 3 Create a new route-map, which sets the BGP MED to a low value (for example, 100). Apply the route-map to outgoing updates toward the router "Good."

Exercise Verification

You have completed this exercise when you attain these results:

- Check the BGP table on router "Good" to see if MED is present in BGP updates received from WGxR1 and WGxR2.

```
Good#sh ip bgp
BGP table version is 46, local router ID is 199.199.199.199
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
*> 10.0.0.0         192.168.20.233           0         32768 i
*  128.20.0.0       192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 128.20.12.0/24   0.0.0.0                  0         32768 i
*  128.22.0.0       192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 128.22.12.0/24   192.168.20.22            0             0 22 i
*  128.26.0.0       192.168.20.22            0             0 22 26 i
*>                  0.0.0.0                  0         32768 i
*  128.37.0.0       192.168.20.22            0             0 22 26 42
37 i
*>                  0.0.0.0                  0         32768 i
*  128.42.0.0       192.168.20.22            0             0 22 26 42
i
*>                  0.0.0.0                  0         32768 i
*  128.51.0.0       192.168.20.22            0             0 22 26 51
i
*>                  0.0.0.0                  0         32768 i
*> 128.213.0.0      0.0.0.0                  0         32768 i
*> 128.214.0.0      192.168.20.22            0             0 22 214 i
*  192.20.11.0      192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 192.20.12.0/30   0.0.0.0                  0         32768 i
*  192.22.11.0      192.168.20.22            0             0 22 i
*>                  0.0.0.0                  0         32768 i
*> 192.22.12.0/30   192.168.20.22            0             0 22 i
*  192.26.11.0      192.168.20.22            0             0 22 26 i
*>                  0.0.0.0                  0         32768 i
*  192.37.11.0      192.168.20.22            0             0 22 26 42
37 i
*>                  0.0.0.0                  0         32768 i
*  192.42.11.0      192.168.20.22            0             0 22 26 42
i
*>                  0.0.0.0                  0         32768 i
*  192.51.11.0      192.168.20.22            0             0 22 26 51
i
*>                  0.0.0.0                  0         32768 i
*  192.168.1.0      192.168.20.1                           0 22 1 i
*                   192.168.20.1           150             0 1 i
*>                  192.168.31.1            50             0 1 i
*  192.168.31.0     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
*> 192.213.11.0     0.0.0.0                  0         32768 i
*> 192.214.11.0     192.168.20.22            0             0 22 214 i
*  197.1.0.0/21     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
*  197.1.0.0/16     192.168.20.1                           0 22 1 i
*>                  192.168.31.1            50             0 1 i
*                   192.168.20.1           150             0 1 i
```



```
*  197.1.8.0/22      192.168.20.1                         0 22 1 i
*>                   192.168.31.1           50            0 1 i
*                    192.168.20.1          150            0 1 i
*> 200.20.0.0/16     0.0.0.0                 0        32768 i
*> 200.22.0.0/16     192.168.20.22           0            0 22 I
```

- Use traceroute from routers WGxR4 and WGxR1 to 192.20.11.1.

```
wg1r4#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.9 32 msec 24 msec 20 msec
  2 192.168.1.5 28 msec 44 msec 41 msec
  3 192.168.31.2 60 msec 64 msec 56 msec
```

```
wg1r1#traceroute 192.20.11.1
Type escape sequence to abort.
Tracing the route to 192.20.11.1

  1 192.168.1.2 40 msec 28 msec 20 msec
  2 192.168.31.2 32 msec *  36 msec
```

- Use traceroute from router "Good" to interface loopback0 on routers WGxR4 and WGxR1.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 192.168.31.1 [AS 1] 36 msec 28 msec 24 msec
  2 192.168.1.6 [AS 1] 32 msec 36 msec 36 msec
  3 192.168.1.10 [AS 1] 52 msec *  48 msec
```

```
Good#traceroute 197.1.8.1
Type escape sequence to abort.
Tracing the route to 197.1.8.1

  1 192.168.31.1 [AS 1] 36 msec 28 msec 24 msec
  2 192.168.1.1 [AS 1] 40 msec *  36 msec
```

- Compare the two outputs of traceroute to determine if routing is symmetrical. Both trace commands should show router WGxR2 in the path.

Answer these questions:

Q1) Which parameters and attributes have to be equal before MED is compared to select the best path?

Q2) What is the default value of MED?

BGP Communities

Overview

This lesson discusses how to influence Border Gateway Protocol (BGP) route selection by setting the BGP community attribute on outgoing BGP routes. The community attribute provides a method to facilitate and simplify the control of routing information by grouping destinations into a specific community. Routing decisions can then be influenced based on the identity of a group, simplifying the configuration requirements of BGP when applying administrative policies.

In this lesson, BGP communities and their use to facilitate proper return path selection is discussed. The configuration details of BGP communities, and the use of community-lists and route-maps to influence route selection are also discussed. This lesson concludes by explaining how to monitor BGP community attributes.

Importance

The community attribute is a transitive optional BGP attribute, designed to group destinations and allow the easy application of administrative policies. BGP communities provide a mechanism to reduce BGP configuration complexity on a router controlling the distribution of routing information. When connections to multiple providers are required, it is important that BGP select the optimum route for traffic to use. It is equally important that the return path selected be the optimum return path into the autonomous system (AS). The "optimum" or "best" route may not be what the network designer intended based on design criteria, administrative policies, or corporate mandate. Fortunately, BGP provides tools for administrators to influence route selection. The BGP community attribute is one such tool.

Objectives

Upon completing this lesson, you will be able to:

- Describe the issues of return path selection for multihomed customers
- Describe the basic qualities of BGP communities
- Describe how you can use BGP communities to facilitate proper return path selection
- List the steps required to successfully deploy communities in a BGP-based network
- Identify the Cisco IOS® commands required to configure route tagging using BGP communities
- Identify the Cisco IOS commands required to configure BGP community propagation
- Identify the Cisco IOS commands required to match routes based on attached BGP communities using community-lists
- Identify the Cisco IOS commands required to match routes based on attached BGP communities using route-maps
- Identify the Cisco IOS commands required to monitor BGP communities

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

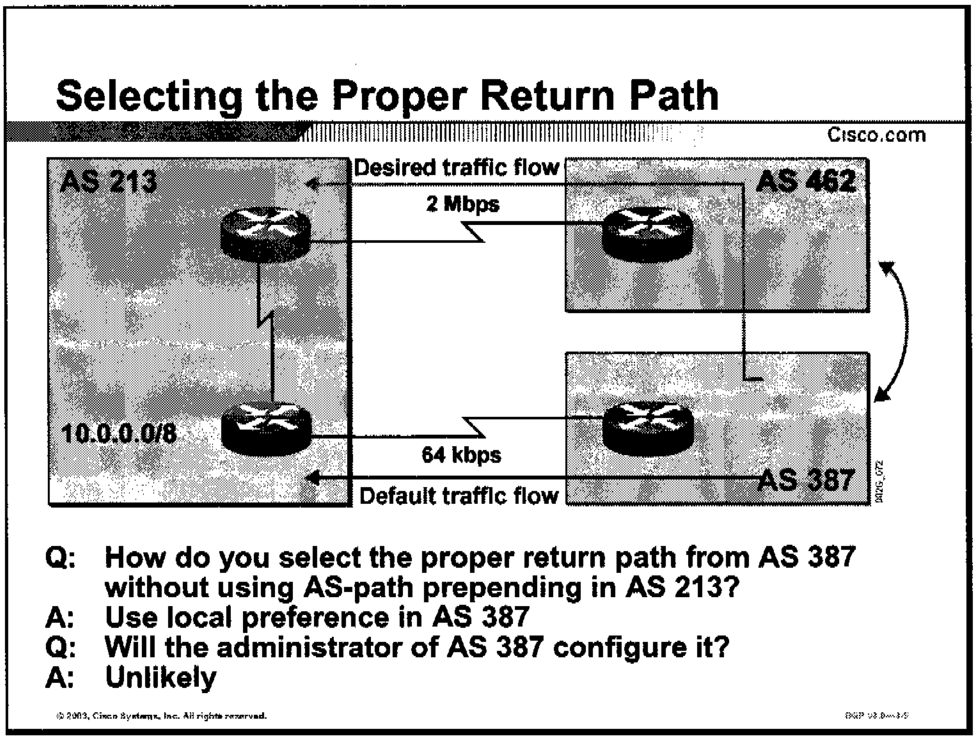
Outline

This lesson includes these topics:

- Overview
- Selecting the Proper Return Path
- BGP Communities Overview
- Using Communities
- Configuring BGP Communities
- Configuring Route Tagging with BGP Communities
- Configuring Community Propagation
- Defining BGP Community-Lists
- Matching BGP Communities in Route-Maps
- Monitoring Communities
- Summary
- Assessment (Lab): BGP Communities

Selecting the Proper Return Path

This topic describes the issues of return path selection for multihomed customers and why you cannot use the BGP attributes of weight, local preference, and MED to solve these issues.



In this example, the customer and the backup service provider would like to avoid AS-path prepending and rely on other BGP tools to properly route the return traffic over the highest-speed WAN link.

Using MED to influence the preferred return path is not possible because MED cannot be propagated across several autonomous systems. AS 387 would, therefore, receive networks from AS 213 directly with the MED attribute, but without a MED attribute from AS 462. In any case, BGP route selection would be based on the length of the AS path, and even if MED was present and used the shortest path, it would still be through the slow 64-kbps link.

The only option for solving this issue is to use local preference in AS 387. The problem with this solution is that service providers normally do not rush to implement every wish that their customers might have.

This lesson describes a solution to this case study using the transitive optional attribute called BGP community in conjunction with local preference.

Practice

Q1) What are two reasons why it is not feasible to use MED to influence return path selection when multiple autonomous systems are involved? (Choose two.)

A) because the MED attribute is designed to influence outbound path selection only

B) because the AS-path attribute would be used for path selection regardless of any configured MED value

C) because the weight attribute will always be used, given that it is first in the BGP route selection process

D) because MED cannot be propagated across several autonomous systems

BGP Communities Overview

This topic describes the basic properties and fundamental qualities of BGP communities.

BGP Communities Overview

Cisco.com

- BGP communities are a means of tagging routes to ensure consistent filtering or route selection policy
- Any BGP router can tag routes in incoming and outgoing routing updates or when doing redistribution
- Any BGP router can filter routes in incoming or outgoing updates or select preferred routes based on communities
- By default, communities are stripped in outgoing BGP updates

BGP v3.0—3-6

A community is an attribute used to tag BGP routes. A router can apply it to any BGP route by using a route-map. Other routers can then perform any action based on the tag (community) that is attached to the route.

There can be more than one BGP community attached to a single route, but the routers, by default, remove communities in outgoing BGP updates.

BGP Communities Overview (Cont.)

Cisco.com

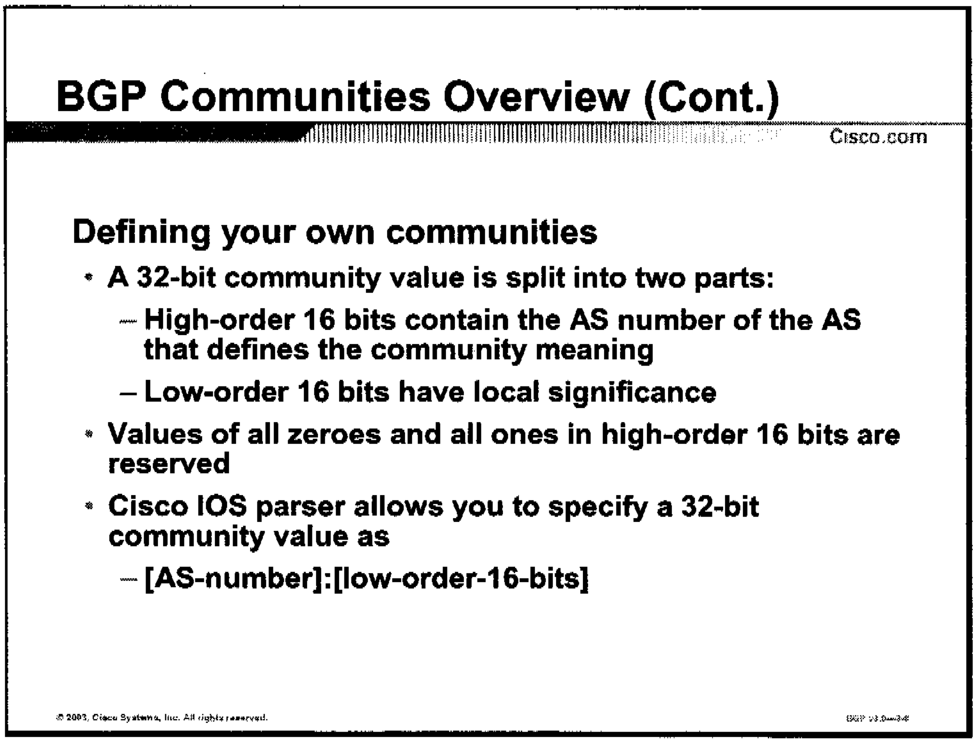
- **The community attribute is a transitive optional attribute. Its value is a 32-bit number (range 0 to 4,294,967,200)**
- **Each network in a BGP routing table can be tagged with a set of communities**
- **The standards define several filtering-oriented communities**
  - **no-export: do not advertise routes to real EBGP peers**
  - **no-advertise: do not advertise routes to any peer**
  - **local-as: do not advertise routes to any EBGP peers**
  - **internet: advertise this route to the Internet community**
- **Routers that do not support communities pass them along unchanged**

 BGP v3.0—3-7

The community attribute is a 32-bit transitive optional attribute that provides a way to group destinations and apply routing decisions (accept, prefer, redistribute, and so on) according to communities. A set of community values has been pre-defined. When a router receives a route marked with a pre-defined community, it will perform a specific, pre-defined action based on that community setting as follows:

- **no-export:** If a router receives an update carrying this community, it will not propagate it to any external neighbors except intraconfederation external neighbors. This is the most widely used predefined community attribute.
- **no-advertise:** If a router receives an update carrying this community, it will not forward it to any neighbor.
- **local-as:** This community has a similar meaning to **no-export**, but it keeps a route within the local AS (or member-AS within the confederation). The route is not sent to external BGP neighbors or to intraconfederation external neighbors.
- **internet:** Advertise this route to the Internet community. All routers belong to it.

Routers that do not support the community attribute will pass the attribute to other neighbors because it is a transitive attribute.



Community attributes are usually used between neighboring autonomous systems. To make sure that the BGP communities are globally unique, a public AS number should be part of the community value. For this reason, you can enter the community value as two 16-bit numbers separated by a colon. The first number (high-order 16 bits) should be the AS number of the AS defining the community value, and the second number is a value that is assigned a certain meaning (for example, translation of a community value into local preference in the neighboring AS).

Communities can also be used internally within an AS (to ensure AS-wide routing policy) in which case the first 16 bits should contain the AS number of the local AS.

Practice

Q2) What is the purpose of the community attribute?

A) to filter incoming BGP route updates

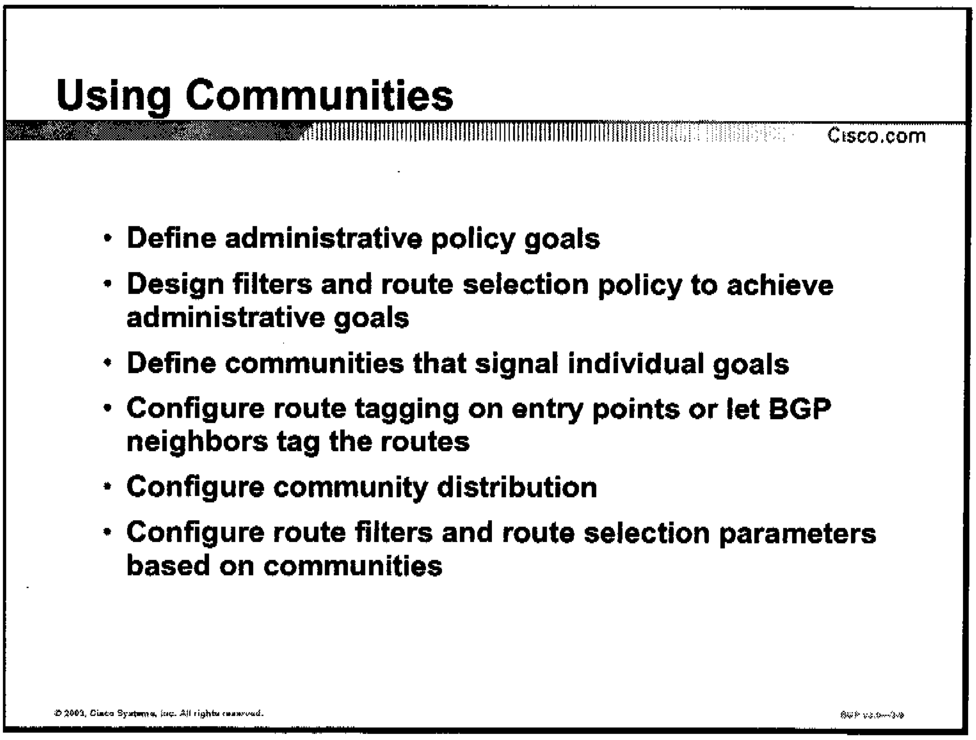
B) to update the BGP table with incoming BGP routes

C) to facilitate selection of the optimum AS exit path

D) to group destinations by tagging BGP updates

Using Communities

This topic describes how BGP communities can facilitate proper return path selection.



Designing a BGP solution around BGP communities usually requires the following steps:

Step 1 Define administrative policy goals that you need to implement.

Step 2 Define the filters and route selection policy that will achieve the required goal(s).

Step 3 Assign a community value to each goal.

Step 4 Apply communities on incoming updates from neighboring autonomous systems or tell the neighbors to set the communities themselves.

Step 5 Enable community distribution throughout your AS to allow community propagation.

Step 6 Match communities with route-maps and route filters, change BGP attributes, or influence route selection process based on the communities attached to the BGP routes.

Example

Using Communities (Cont.)

Cisco.com

Define administrative policy goals

- **Solve asymmetrical customer routing problems**

Design filters and path selection policy to achieve administrative goals

- **Set local preference of customer routes to 50 for customers using the backup ISP**

Define communities that signal individual goals

- **Community 387:17 is used to indicate that the local preference of the route should be lowered to 50**

The following table is an example of how you can define goals and assign communities to them:

Goal	Community value
Set local preference of 50	387:17
Set local preference of 150	387:18
Prepend AS path once when sending the network to external neighbors	387:21
Prepend AS path twice when sending the network to external neighbors	387:22
Prepend AS path three times when sending the network to external neighbors	387:23

All customers of the service provider should know this list so that they can use the BGP communities without having to discuss their use with the service provider.

Practice

Q1) Does the community attribute have any influence on BGP path selection?

A) No, communities are simply tags that are applied to BGP routes.

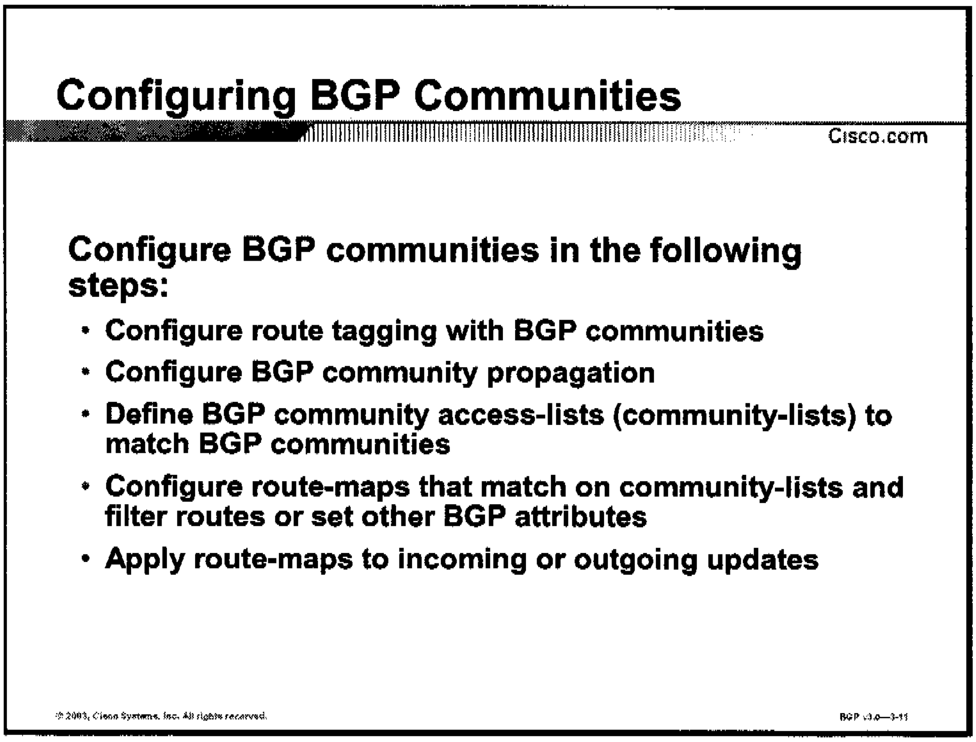
B) No, communities are nontransitive attributes.

C) Yes, BGP paths are selected based on the value in the community tag.

D) Yes, the community attribute is part of the BGP route selection process.

Configuring BGP Communities

This topic lists the steps required to successfully deploy BGP communities in a BGP-based network.



Configuration steps when you are using communities include:

- Setting communities, which requires a route-map.
- Enabling community propagation per neighbor for all internal neighbors. If communities are sent to external neighbors, you must enable community propagation for external neighbors.
- Creating community-lists to be used within route-maps to match on community values.
- Creating route-maps where community-lists are used to match on community values. You can then use route-maps to filter based on community values or to set other parameters or attributes (for example, local preference, MED, or AS-path prepending).
- Applying route-maps to incoming or outgoing updates.

Practice

Q1) What Cisco IOS feature must you use to set the BGP community attribute?

A) distribute-list

B) filter-list

C) route-map

D) access control list

Configuring Route Tagging with BGP Communities

This topic lists the Cisco IOS commands required to configure route tagging using BGP communities.

Configuring Route Tagging with BGP Communities

Cisco.com

```
router(config)#
route-map name
  match condition
  set community value [ value … ] [additive]
```

- **Route tagging with communities is always done with a route-map**
- **You can specify any number of communities**
- **Communities specified in the set keyword overwrites existing communities unless you specify the additive option**

 BGP v3.0—3-12

In a route-map configuration mode, you should use the **set community** command to attach a community attribute (or a set of communities) to a route. You can attach up to 32 communities to a single route with one route-map **set** statement. If the keyword **additive** is used, the original communities are preserved and the router simply appends the new communities to the route. Omitting the **additive** keyword results in overwriting any original community attributes.

Configuring Route Tagging with BGP Communities (Cont.)

Cisco.com

router(config-router)#

```
neighbor ip-address route-map map in | out
```

- **Applies a route-map to inbound or outbound BGP updates**
- **The route-map can set BGP communities or other BGP attributes**

router(config-router)#

```
redistribute protocol route-map map
```

- **Applies a route-map to redistributed routes**

You can apply a route-map to incoming or outgoing updates. You can also use it with redistribution from another routing protocol.

Note A route-map is a filtering mechanism that has an "implicit deny" for all networks not matched in any route-map statement. If a route-map is not intended to filter routes, then you should add another route-map statement at the end to permit all remaining networks without changing it (no **match** and no **set** commands are used within that route-map statement).

Originally, Cisco IOS software accepted and displayed BGP community values as a single 32-bit value in a digital format. Newer Cisco IOS versions support the new format, where you can set or view a community as two colon-separated 16-bit numbers.

The global command **ip bgp-community new-format** is recommended on all routers whenever communities contain the AS number.

After being converted, configuration files with communities in the new as:nn format are not compatible with older versions of Cisco IOS software. For example:

router# show ip bgp 6.0.0.0

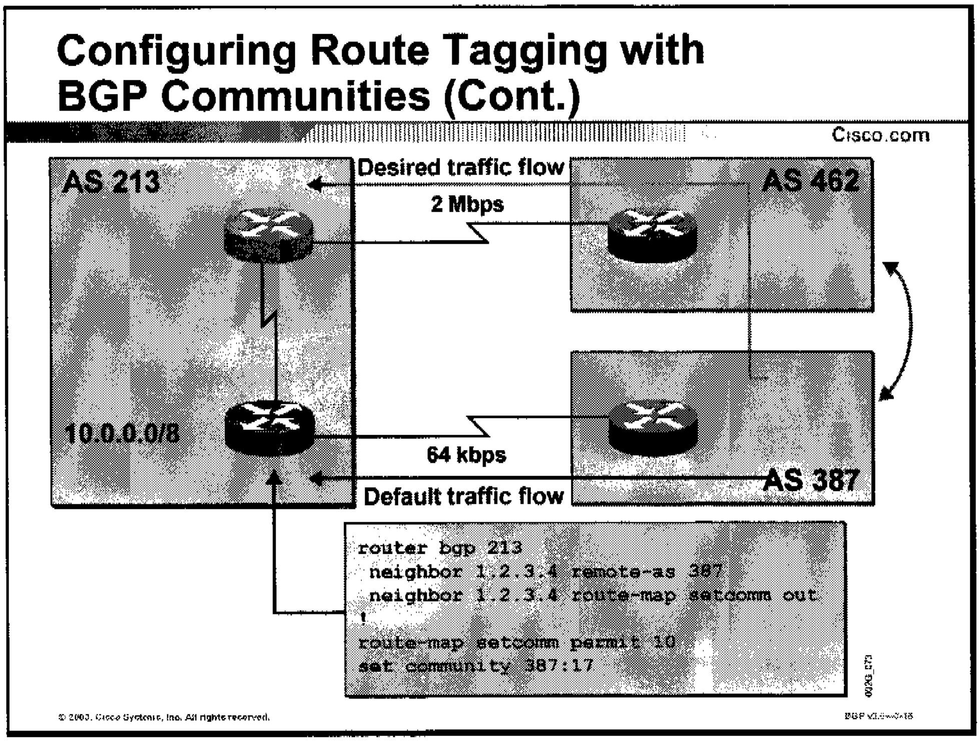
Community: 6553620

ip bgp-community new-format

router# show ip bgp 6.0.0.0

Community: 100:20

Example



In this example, a border router in AS 213 applies a community value 387:17 to all networks sent to neighboring AS 387. In this example, another route-map entry is not needed because the first statement permits all networks (no **match** command means **match all**).

If it is more desirable to set communities on specific routes, you can use a standard access control list (ACL) to match against, with the **match ip address** command in the route-map.

In a later example, networks with community 387:17 will have the local preference changed to a value of 50 within AS 387 to force AS 387 to prefer the other path that carries the default local preference of 100.

Practice

Q1) How many community tags can be attached to a single BGP route?

A) 1

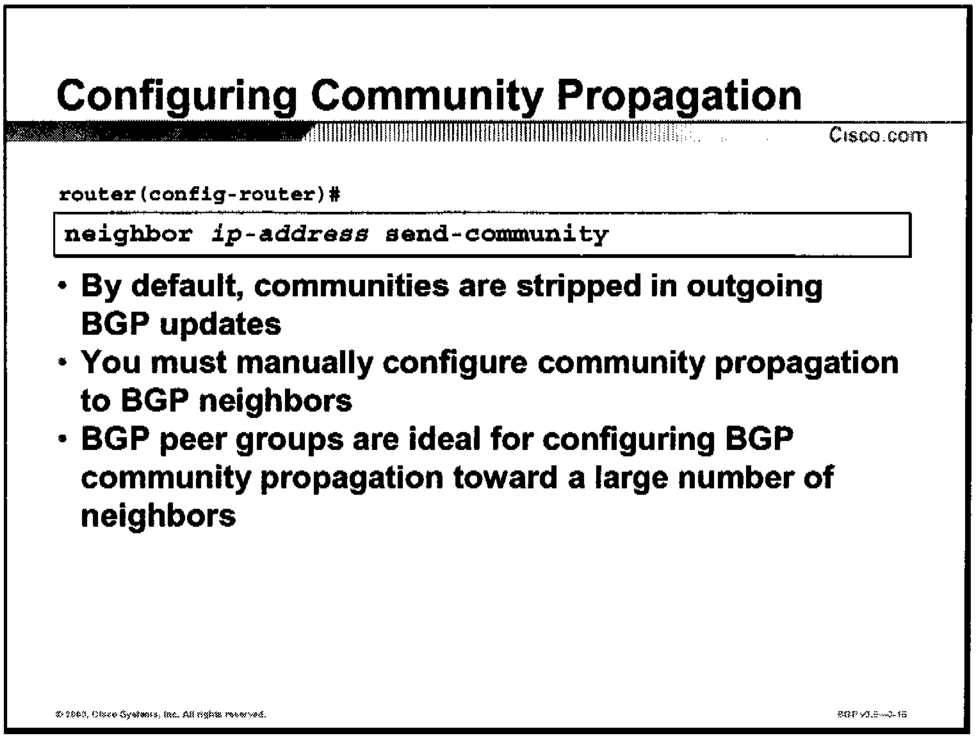
B) 32

C) 255

D) depends on the number configured with the **ip bgp community** command

Configuring Community Propagation

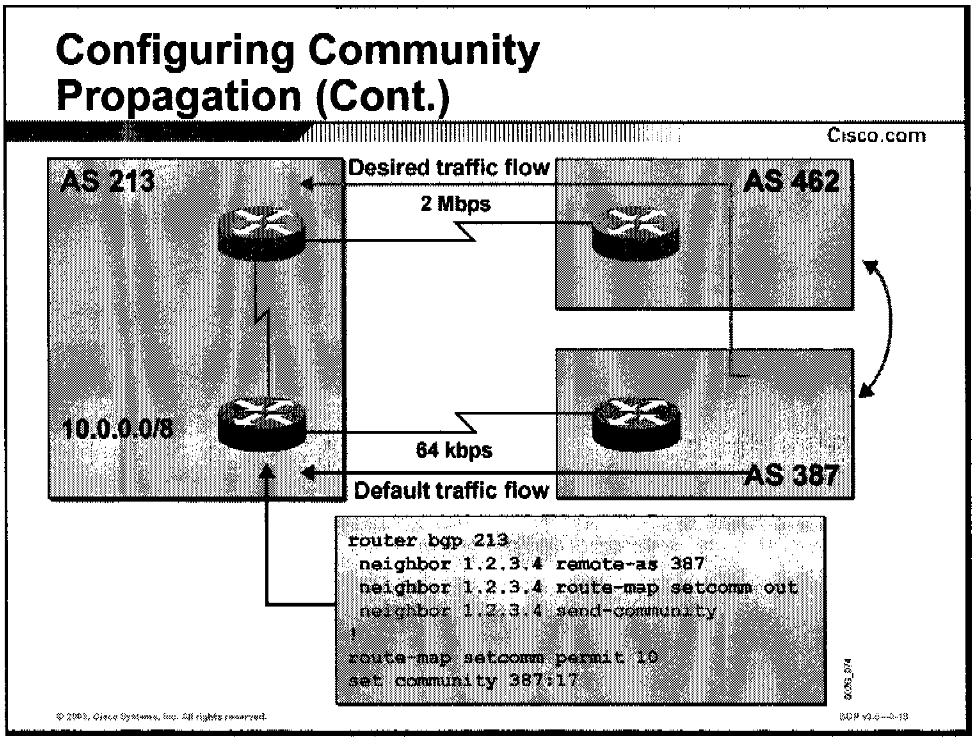
This topic lists the Cisco IOS commands required to enable BGP community propagation to BGP neighbors.



A command commonly forgotten by network administrators when configuring BGP communities is **neighbor** *ip-address* **send-community**. This command is needed to propagate community attributes to BGP neighbors. Even if you use an outgoing route-map to set communities, by default, the router will strip out any community values attached to outgoing BGP updates if you have not configured this command for the specific BGP neighbor.

You can also apply this command to a peer group.

Example



The configuration example discussed earlier in this lesson must include the **send-community** command to enable community propagation from AS 213 to AS 387.

Practice

Q1) What must you configure on a Cisco router to begin the propagation of the community attribute?

A) a BGP peer group

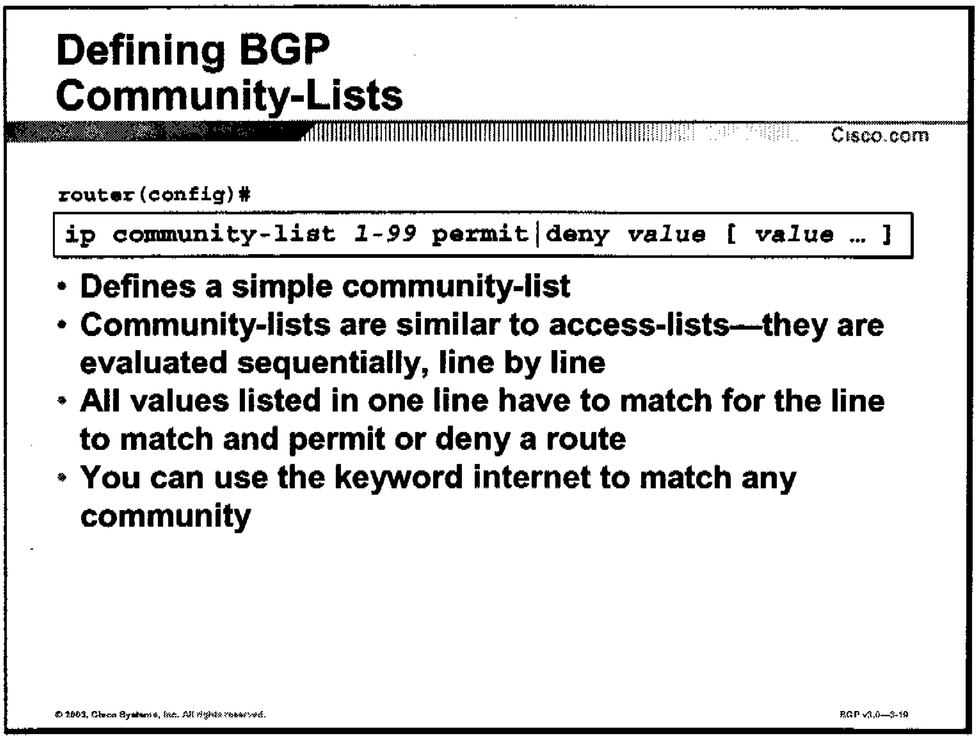
B) **send-community** on the **neighbor** statement

C) a distribute-list under the router BGP process

D) an outgoing route-map on the **neighbor** statement

Defining BGP Community-Lists

This topic lists the Cisco IOS commands required to match routes based on attached BGP communities using community-lists.



You can use a standard community access-list to find community attributes in routing updates. A standard community-list is defined by its assigned list number, which ranges from 1 to 99. Community-lists are similar to standard IP access–lists in these ways:

- The router evaluates the lines in the community-list sequentially.
- If no line matches communities attached to a BGP route, the route will be implicitly denied.

Standard community-lists are different from standard IP access-lists in these ways:

- The keyword **internet** should be used to permit any community value.
- If more values are listed in a single line, they all have to be in an update to have a match.

Defining BGP Community-Lists (Cont.)

Cisco.com

```
router(config)#
```

```
ip community-list 100-199 permit|deny regexp
```

- **Defines an extended community-list**
- **Extended community-lists are like simple community-lists, but they match based on regular expressions**
- **Communities attached to a route are ordered, converted to string, and matched with regexp**
- **Use .* to match any community value**

 BGP v3.0—3-20

An extended community-list is defined by its assigned list number, which ranges from 100 to 199. Regular expressions are used to match community attributes. When a router processes a list of communities, attached to a network update, they are converted into an ordered string of characters. The example below shows how this is done:

1. The original list of communities in an update:

   "10.0.0.0/8, NH=1.1.1.1, origin=I, AS-path="20 30 40", community=10:101, community=10:201, community=10:105, community=10:205"

2. A string of characters containing an ordered list of community values:

   "_10:101_10:105_10:201_10_205_" ("_" represents a space)

3. A regular expression:

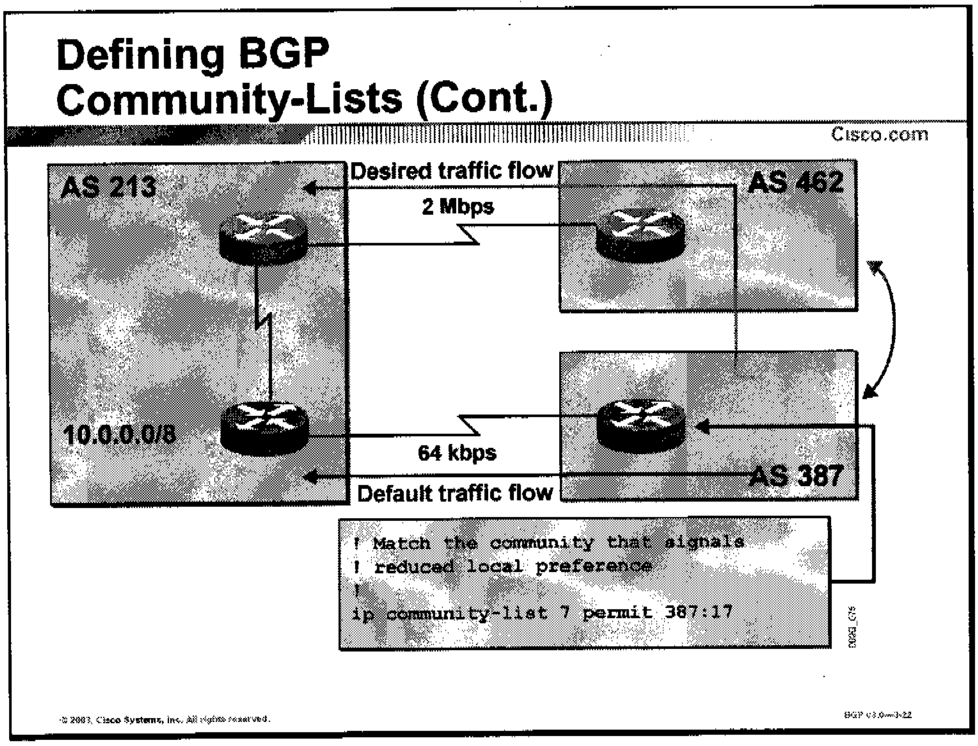
   "permit _10:.0[1-5]_" ("_" represents an underscore that matches spaces)

4. The result:

   This regular expression permits the route because it permits all routes carrying communities with the first 16 bits carrying the AS number 10 and the second 16 bits having 0 as the second digit and 1 to 5 as the third digit; the first digit can be anything (as indicated by the ".").

   Use regular expression ".*" to permit any community.

Example



This example shows a portion of the configuration of the router in AS 387. The access-list has been configured to match communities previously set by the router in AS 213.

Practice

Q1) What match criteria are specified in a standard BGP community-list?

A) destination IP addresses

B) regular expressions

C) community attribute values

D) AS numbers

Q2) What regular expression should you use with an extended BGP community-list to match any community value?

A) internet

B) any

C) .*

D) permit all

Matching BGP Communities in Route-Maps

This topic lists the commands required to match routes based on attached BGP communities using route-maps.

Matching BGP Communities in Route-Maps

Cisco.com

```
router(config)#
route-map name permit | deny
  match community clist-number [exact]
  set attributes
```

- Community-lists are used in match conditions in route-maps to match on communities attached to BGP routes
- A route-map with community-list matches a route if at least some communities attached to the route match the community-list
- With the exact option, all communities attached to the route have to match the community-list
- You can use route-maps to filter routes or set other BGP attributes based on communities attached to routes

BGP v3.0—3-23

Network administrators use route-maps to match networks that carry a subset of communities that are permitted by the community-list. Other parameters or attributes can then be set based on community values. If you use the keyword **exact**, all communities attached to a BGP route have to be matched by the community-list.

Matching BGP Communities in Route-Maps (Cont.)

Cisco.com

Route Selection

- You can use route-maps to set weights, local preference, or metric based on BGP communities attached to the BGP route
- Normal route selection rules apply afterward
- Routes not accepted by route-map are dropped

Default Filters

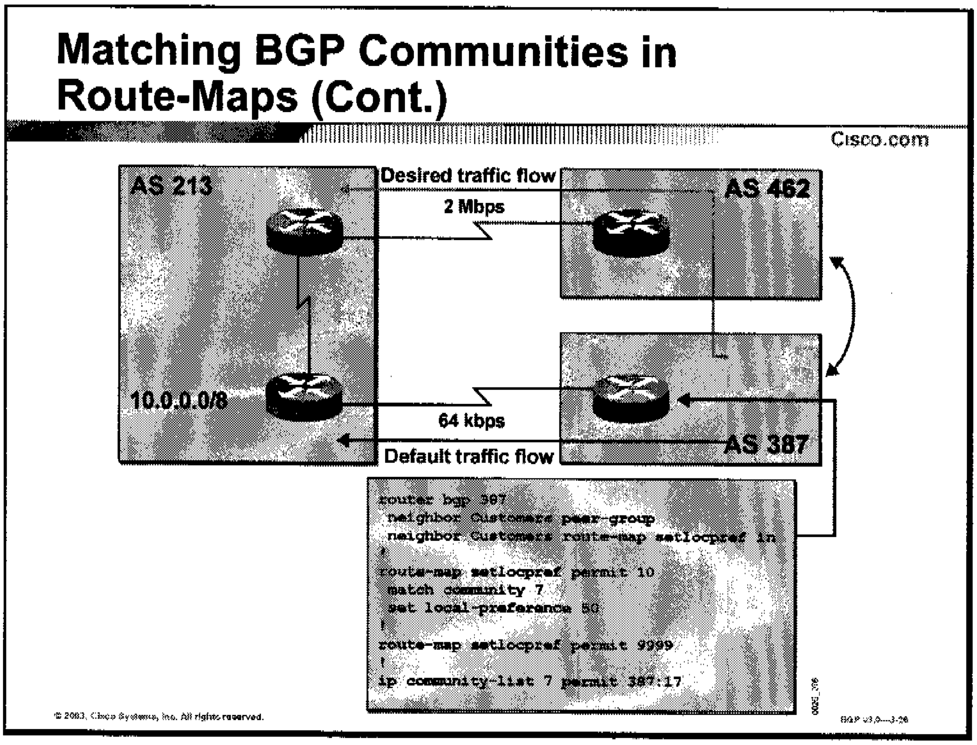
- Routes tagged with community no-export are sent to IBGP peers and intraconfederation EBGP peers
- Routes tagged with local-as are sent to IBGP peers
- Routes tagged with no-advertise are not sent in any outgoing BGP updates

You can use a route-map to filter or modify BGP routing updates. Any BGP-related **set** commands can be used to set BGP parameters and attributes (that is, weight, local preference, multi-exit discriminator [MED]).

As mentioned before, there are some predefined community values that cause routers to automatically filter routing updates:

- **no-export:** If a router receives an update carrying this community, it will not propagate it to any external neighbors except to intraconfederation external neighbors.
- **no-advertise:** If a router receives an update carrying this community, it will not forward it to any neighbor.
- **local-as:** This community has a similar meaning to **no-export**, but it keeps a route within the local subautonomous system. It is not sent to intraconfederation external neighbors or to any other external neighbors.
- **internet:** Advertise this route to the Internet community. All routers belong to it.

Example



This example shows a configuration that translates community 387:17 into local preference 50. All updates received from neighboring AS 213 are processed by the route-map, which uses a community-list to find community 387:17. If the community-list matches one of the community attributes, the **set** command is executed and the route is permitted. If the route does not contain the right community, it is simply permitted by route-map statement 9999 without changing anything in the update.

The result is that AS 387 prefers other paths to AS 213 because they have a default local preference of 100.

Practice

Q1) What is the result of tagging a route with the no-export community?

A) The route will not be advertised within the local AS.

B) The upstream AS will not be allowed to export the route.

C) The route cannot be exported to another routing protocol.

D) The router will not propagate the route to any external neighbors except to intraconfederation external neighbors.

Monitoring Communities

This topic lists the commands required to monitor BGP communities.

Monitoring Communities

Cisco.com

- **Communities are displayed in show ip bgp prefix printout**
- **Communities are not displayed in debugging outputs**
- **Routes in BGP table tagged with a set of communities or routes matching a community-list can be displayed**

Because a community is an attribute that can appear more than once in a single update, the **show ip bgp** command does not show it. You can view communities only if you use the **show ip bgp** *prefix* command.

If you use the **show ip bgp community-list** command, all networks are listed that are permitted by the community-list.

Monitoring Communities (Cont.)

Cisco.com

- **Communities are displayed only in show ip bgp *prefix* printout**

```
Betty#show ip bgp 10.0.0.0
BGP routing table entry for 10.0.0.0/8, version 17
Paths: (3 available, best #3, advertised over EBGP)
  213
    1.3.0.2 from 1.3.0.2 (10.1.1.1)
      Origin IGP, metric 0, localpref 50, valid, external
      Community: 387:17
  213, (received-only)
    1.3.0.2 from 1.3.0.2 (10.1.1.1)
      Origin IGP, metric 0, localpref 100, valid, external
      Community: 387:17
  462 213
    1.1.0.4 from 1.1.0.4 (12.1.2.3)
      Origin IGP, localpref 100, valid, external, best
```

Communities attached to the route in BGP table

Communities received from the neighbor

BGP v3.0—3-28

This example shows the output of the **show ip bgp** *prefix* command where inbound soft reconfiguration was enabled on one of the neighbors. The original update contained one single community attribute (387:17), which can be seen from the second path marked with received-only. This update was then processed by an inbound route map, which matched the community 387:17 and changed the local preference of the received route to 50.

Monitoring Communities (Cont.)

Cisco.com

router>

```
show ip bgp community
```

- Displays all routes in a BGP table that have at least one community attached

router>

```
show ip bgp community as:nn [as:nn ...]
```

- Displays all routes in a BGP table that have all the specified communities attached

Another use of the **show** command is to filter the output of **show ip bgp**.

If the keyword **community** is included, all networks that have at least one community attribute are displayed.

If the keyword **community** is followed by one or more community values, only those networks that carry all those communities are displayed.

Monitoring Communities (Cont.)

Cisco.com

router>

```
show ip bgp community as:nn [as:nn …] exact
```

- **Displays all routes in BGP table that have exactly the specified communities attached**

router>

```
show ip bgp community-list clist
```

- **Displays all routes in BGP table that match community-list *clist***

BGP v3.0—3-30

If the keyword **community** is followed by one or more community values, only those networks that carry all those communities are displayed. If the keyword **exact** is added at the end, only those that match exactly are displayed.

You can also use a community-list to filter the output of the **show ip bgp** command.

Practice

Q1) What command should you use to display the community attribute?

A) **show ip bgp summary**

B) **show community**

C) **show ip bgp** *prefix*

D) **show ip route**

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- You can use the BGP community attribute to create an AS-wide routing policy or to provide services to neighboring autonomous systems.
- Community attributes are usually used between neighboring autonomous systems. Routers that do not support the community attribute will pass the attribute to other neighbors because it is a transitive attribute.
- A community is an attribute used to tag BGP routes that you can use to manipulate path selection and enforce administrative policies.
- To set the community attribute, you must use a route-map
- In route-map configuration mode, you should use the set set community command

 BGP v3.0—3-31

Summary (Cont.)

Cisco.com

- You must configure propagation of BGP communities on the routers on a per-neighbor basis; otherwise, the BGP communities are removed from the outgoing BGP updates.
- You can use community-lists to match against the community attribute as a method of route selection.
- A route-map is used to match networks that carry a subset of communities that are permitted by the community-list.
- You can view communities only if you use the show ip bgp *prefix* command.

 BGP v3.0—3-32

Next Steps

After completing this lesson, go to:

- Customer-to-Provider Connectivity with BGP module

References

For additional information, refer to these resources:

- For more information on BGP communities, refer to "BGP Case Studies" at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#communityattribute

- For further information on BGP communities, refer to "Configuring BGP" at the following URL:
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/np1_c/1cprt1/1cbgp.htm#xtocid34

Laboratory Exercise: BGP Communities

Complete the laboratory exercise to practice what you have learned in this lesson.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will configure BGP to influence route selection using the BGP community attribute in a situation where you must support multiple connections to an Internet service provider (ISP).

After completing this exercise, you will be able to:

- Configure route tagging using BGP communities
- Configure BGP community propagation
- Monitor BGP communities

Command List

The commands used in this exercise are described in the table here.

Table 1: Lab Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	Use this command to apply a route-map to incoming or outgoing routing updates.
route-map *name* {**permit** | **deny**} **seq**	Define or modify an existing entry in a route-map.
match community *c-acl*	Use this command in a route-map to match communities by using a community-list.
set community *community* [**additive**]	Use this command in a route-map to set community attributes. Use the keyword **additive** to append communities instead of replacing them.
set local-preference *num*	Use this command to set local preference attribute.
ip bgp-community new-format	Configure BGP communities using AA:NN format.
ip community-list *num* {**permit|deny**} **community**	Use this command to create a community access-list.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of **show ip bgp** command.
show ip bgp community [community [community …]] [exact-match]	Use this command to view BGP routes that have at least one community attribute or those specified in the command.
show ip bgp community-list *c-list* **[exact-match]**	Use this command to view BGP routes that are permitted by the specified community-list.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- Your second service provider "Cheap" has upgraded its link to your network to match that of the provider "Good." You decide to change your route selection policy to use the newly updated link from "Cheap." As a result of the new service available from "Cheap," you decide to remove the backup link to provider "Good" as you can now achieve both link and provider redundancy.

- You wish for all outgoing traffic to still use the link to "Good," but all return traffic should use the newly updated link from "Cheap." Knowing that you have a connection to the provider "Good," you still wish to use the primary connection to "Good" for return traffic should the link to provider "Cheap" fail.

- You cannot influence return path selection across multiple providers and solicit assistance from provider "Good." The provider "Good" agrees to set local preference for any routes that you mark as follows:
  - Local preference of 50 for routes marked with a community attribute of x:20
- The provider has preconfigured router "Good" to set local preference on IP routes set with the correct community attribute to assist you in influencing return path selection.
- Figure 1 shows the physical connectivity, BGP sessions, and traffic flow in the network.

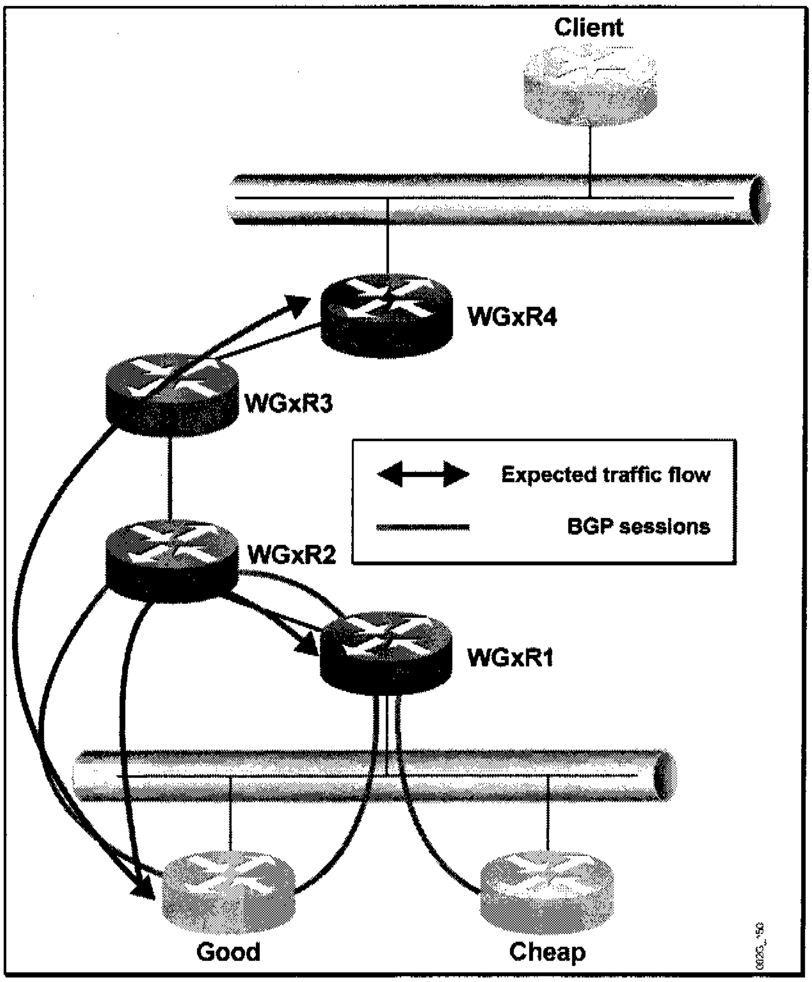


Figure 1: BGP communities physical connectivity

Exercise Procedure

Complete these steps:

Step 1 Before starting this lab, use a trace from router "Good" to 197.x.7.1. The trace should match that below (where x is your WG number).

```
Good#traceroute 197.x.7.1

Type escape sequence to abort.
Tracing the route to 197.x.7.1

1 192.168.3x.1 [AS 1] 16 msec 16 msec 16 msec
2 192.168.x.6 [AS 1] 28 msec 32 msec 28 msec
3 192.168.x.10 [AS 1] 44 msec *  40 msec
```

Step 2 Remove the BGP neighbor statement on WGxR1 for router "Good."

Configure BGP community propagation:

Step 3 Configure BGP community propagation on router WGxR2.

Step 4 Enable your WGxR2 router to configure communities using the AA:NN format.

Configure outbound filters on WGxR2:

Step 5 Create a route-map to set the community attribute on all BGP updates sent from WGxR2 to the provider "Good." Set the community attribute to x:20, where x is your workgroup number.

Step 6 Apply the route-map to the BGP neighbor "Good" in the outbound direction.

Exercise Verification

You have completed this exercise when you attain these results:

- Log on to the "Good" router and verify that the BGP community attribute has been correctly set.

```
Good>show ip bgp community 1:20

Good#sh ip bgp community 1:20
BGP table version is 61, local router ID is 199.199.199.199
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
*  192.168.1.0      192.168.31.1             0     50      0 1 i
*  192.168.31.0     192.168.31.1                   50      0 1 i
*  197.1.0.0/21     192.168.31.1                   50      0 1 i
*  197.1.0.0/16     192.168.31.1                   50      0 1 i
*  197.1.8.0/22     192.168.31.1                   50      0 1 i
```


192.168.31.2

- Log in to router "Good" and perform the same traceroute that was done before starting the configuration in this lab.

```
Good#traceroute 197.1.7.1
Type escape sequence to abort.
Tracing the route to 197.1.7.1

  1 wgl (192.168.20.1) 0 msec 0 msec 0 msec
  2 192.168.1.2 [AS 1] 16 msec 16 msec 16 msec
  3 192.168.1.6 [AS 1] 28 msec 28 msec 36 msec
  4 192.168.1.10 [AS 1] 44 msec *  44 msec
```

Note Normally, you would expect to see the first line of the traceroute showing "Cheap" as the first hop. However, because the routers "Good," "Cheap," and WGxR1 share a common Ethernet segment, the BGP next-hop rule states: If the current BGP next hop is in the same IP subnet as the receiving router, the next hop is not changed; otherwise, it is changed to the IP address of the sending router. The next-hop rule applies even when there is no BGP connection between "Good" and WGxR1.

- On router "Good," execute the command: **show ip bgp 197.1.7.1**.

```
Good#sh ip bgp 197.1.7.1
BGP routing table entry for 197.1.0.0/21, version 22
Paths: (2 available, best #2, table Default-IP-Routing-Table)
  Advertised to peer-groups:
     students2
  1, (aggregated by 1 197.1.8.1)
    192.168.31.1 from 192.168.31.1 (197.1.3.1)
      Origin IGP, localpref 50, valid, external, atomic-aggregate
      Community: 1:20

  22 1, (aggregated by 1 197.1.8.1)
    192.168.20.1 from 192.168.20.22 (192.20.11.1)
      Origin IGP, localpref 100, valid, external, atomic-aggregate,
best
```

Note that the best path, because of local preference, is the bottom route, and the next hop is 192.168.20.22 – "Cheap" – but because of the shared media, the next hop has not been changed. This result is true even though the AS-path length is longer, because local preference is a stronger route selection tool.

Answer these questions:

Q1) What do you have to do to enable community propagation?

Q2) What mechanisms can you use to match or set communities?

20-4520

4

Customer-to-Provider Connectivity with BGP

Overview

Today, many companies use the Internet for a variety of reasons, including increasing employee productivity, increasing sales, increasing customer satisfaction, and reducing cycle time. A key component in connecting companies to the Internet is the service provider. Depending upon business goals, application requirements, and administrative policies, a company will use different methods to connect to a service provider. Company business requirements and policies may even dictate that the company connects to multiple service providers. This module discusses the different requirements for connectivity between customers and service providers. Included in this module is a discussion of physical connection methods, redundancy, load balancing, and technical requirements such as addressing and autonomous system (AS) numbering. This module discusses the configuration requirements to connect a customer to a single service provider using static routes and using the Border Gateway Protocol (BGP). Also discussed in this module are the configuration requirements to connect a customer to multiple service providers using BGP.

Upon completing this module, you will be able to:

- Describe the requirements to connect customer networks to the Internet in a service provider environment
- Implement customer connectivity using static routing, given a service provider network
- Implement customer connectivity using BGP, given a customer scenario where you must support multiple connections to a single ISP
- Implement customer connectivity using BGP, given a customer scenario where you must support connections to multiple ISPs

Outline

The module contains these lessons:

- Customer-to-Provider Connectivity Requirements
- Implementing Customer Connectivity Using Static Routing
- Connecting a Multihomed Customer to a Single Service Provider
- Connecting a Multihomed Customer to Multiple Service Providers

Customer-to-Provider Connectivity Requirements

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses different solutions for connecting customer networks to service providers. Included in this lesson is a discussion of customer network redundancy requirements, routing requirements, IP addressing requirements, and autonomous system (AS) numbering requirements.

Importance

When planning network connectivity to a service provider, network designers must give careful consideration to the different aspects of the connectivity, including physical connection types, redundancy provided by the connection method chosen, IP addressing requirements, and AS numbering considerations, if the network design is going to meet both the business and technical requirements of the applications planned for the network.

Objectives

Upon completing this lesson, you will be able to:

- Identify different physical connections used by customers to connect to a service provider
- Describe the levels of redundancy provided by each physical connection type used by customers to connect to a service provider

- Identify different routing schemes used by customers to connect to a service provider
- Describe routing schemes that are appropriate for each physical connection type used by customers to connect to a service provider
- Describe the addressing schemes used by customers to connect to a service provider
- Describe AS numbering schemes used by customers to connect to a service provider

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

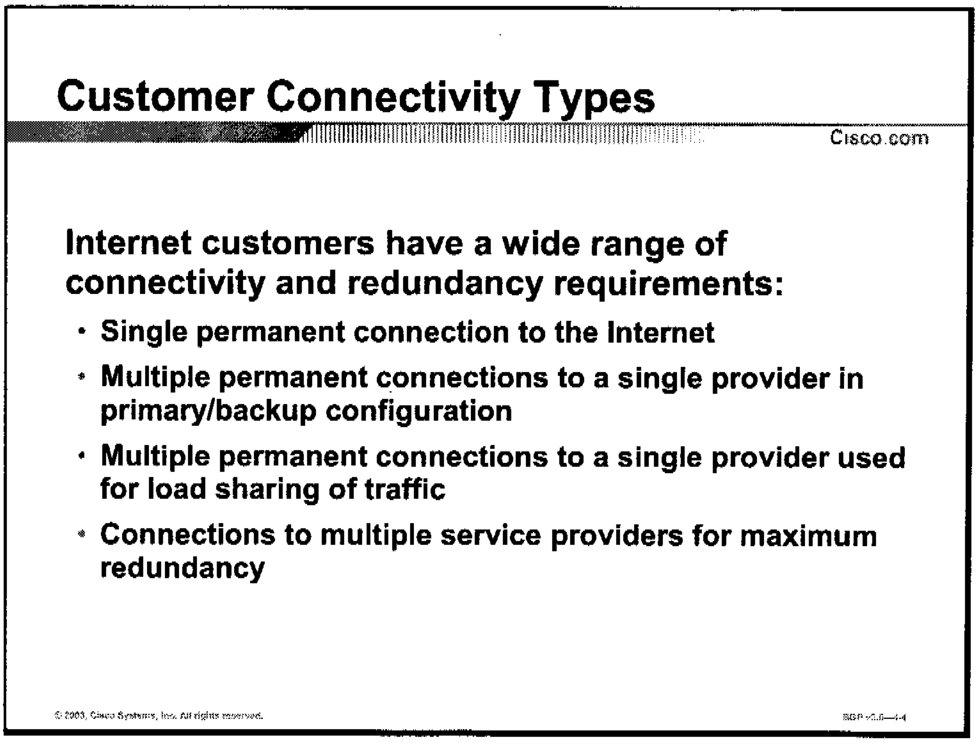
Outline

This lesson includes these topics:

- Overview
- Customer Connectivity Types
- Redundancy in Customer Connections
- Customer-to-Provider Routing Schemes
- Customer Routing
- Addressing Requirements
- AS Number Allocation
- Summary
- Assessment (Quiz): Customer-to-Provider Connectivity Requirements

Customer Connectivity Types

This topic identifies different physical connections used by customers to connect to a service provider.



Service provider customers have different requirements for their Internet connectivity. These different requirements result in different solutions:

- A single permanent connection to one Internet service provider (ISP). This solution meets the requirements for the vast majority of customers.
- Multiple permanent connections where one of the lines is primary and the other line is used for backup only. This setup also provides redundancy on the links. Compared to a dial-up backup, a permanent backup link is preferred for various reasons such as the severe bandwidth limitations on dial-up lines and the time required to establish a dial-up connection.
- Multiple permanent connections to one ISP used for load sharing of traffic. This solution gives redundancy on the links but also provides additional bandwidth.
- Permanent connections to more than one ISP. This solution provides the highest level of redundancy, because it not only can cope with link-level failures but also failures within the network of a service provider.

Practice

Q1) Which method of Internet connectivity gives the highest level of redundancy?

A) a single permanent connection to one ISP

B) multiple permanent connections to one ISP where one of the lines is primary and the other line is used for backup only

C) multiple permanent connections to one ISP used for load sharing of traffic

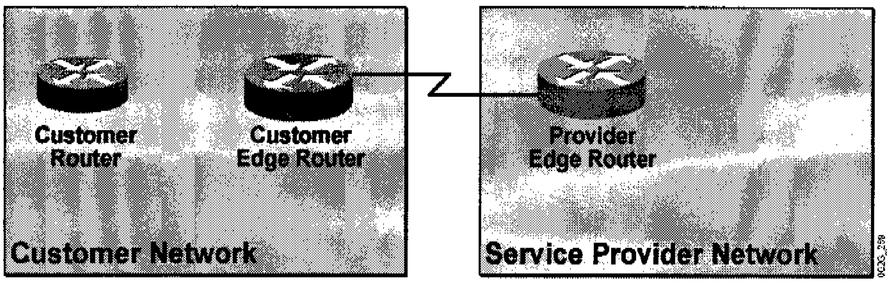
D) multiple permanent connections to more than one ISP

Redundancy in Customer Connections

This topic describes the different levels of redundancy provided by each physical connection type used by customers to connect to a service provider.

Single Permanent Connection to the Internet

Cisco.com



- **The simplest setup—a single link between the customer network and the Internet**
- **No redundancy on link or equipment failure**

BGP v3.0—4-5

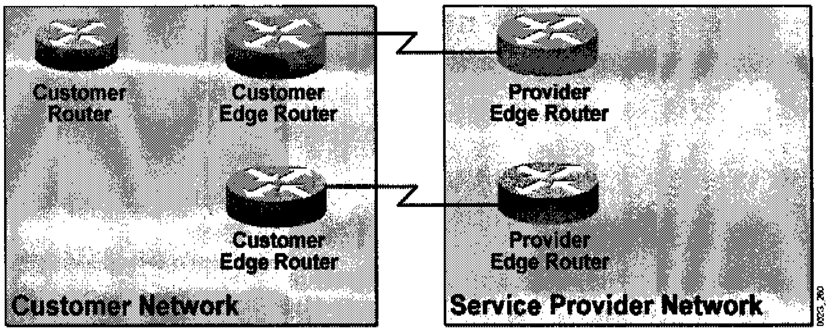
A single permanent connection to one ISP is the most common setup. This setup is also the simplest to implement.

The customer network has an edge router. This router is connected to one of the edge routers of the ISP. The connection is permanent and could be a leased line, a Frame Relay or ATM permanent virtual connection (PVC), a LAN segment, or something equivalent.

There is no redundancy in this solution. Any failure on the permanent link or any of the two edge routers causes a complete outage of the service. Serious failures within the ISP network, which affect all customers of this ISP, also affect the customer in this example.

Multiple Permanent Connections Providing Redundancy

Cisco.com



- **Customers wanting increased redundancy install several physical links to the Internet**
- **Redundant link is used in primary/backup setup or for load sharing**
- **Redundancy on link or equipment failure**
- **No redundancy on service provider failure**

 BGP v3.0—4-6

In this setup, one customer edge router connects to one ISP edge router. A different customer edge router is used to connect to another ISP edge router. If one of these routers fails, only one of the connections breaks down. The other connection is still available.

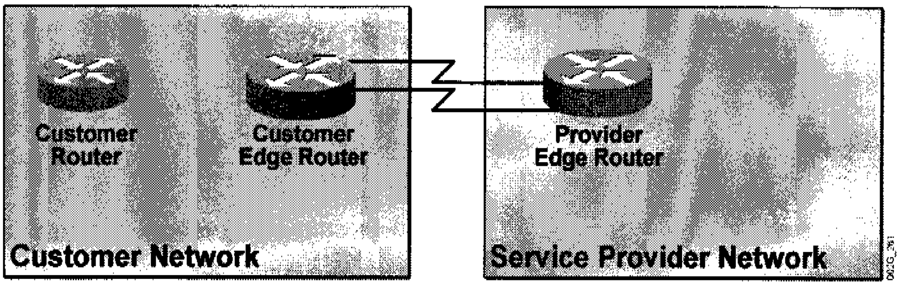
In some cases, the two links may be implemented between the customer and the provider for load sharing and in other cases strictly for backup purposes. For example, backup PVCs in Frame Relay or ATM networks can sometimes be very cost-efficient, provided that these PVCs carry only a very small volume of traffic and that the primary path is available.

In a case where load sharing between both links is a desired network characteristic, the distribution of the load over the links is more complicated compared to a case where both links terminate in the same router.

Again, because the customer is connected to a single ISP, serious ISP network failures, which affect all customers of this ISP, will also affect the customer in this scenario, regardless of the backup link.

Multiple Permanent Connections Providing Load Sharing

Cisco.com



- **Customers wanting to increase their access speed can install several physical links between a pair of routers**
- **Redundancy on link failure**
- **No redundancy on equipment failure**
- **Load sharing in this setup is optimal**

 BGP v3.0—4-7

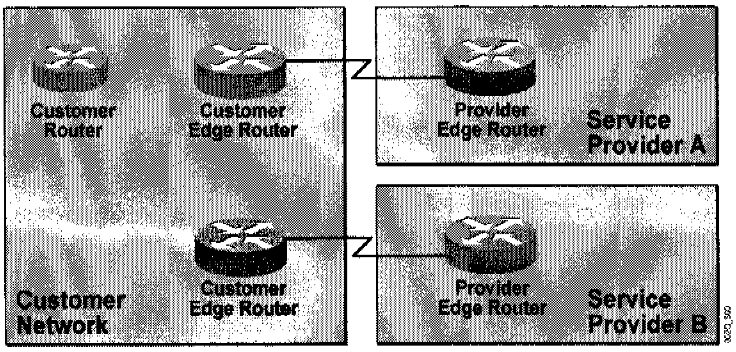
In the example shown, a single router in the customer network is connected to a single router in the ISP network. The redundancy is limited to the link level because router failures are not covered. Using two parallel links between two routers allows for an optimal distribution of load over the links.

Depending on the switching path used in the customer and the ISP routers, load sharing can be performed based on the destination address only (fast switching), based on source-destination address pairs (default behavior for Cisco Express Forwarding [CEF]), or on a packet-by-packet basis (process switching or CEF).

As in the previous examples, serious ISP network failures that affect all customers of this ISP will also affect this customer, regardless of the link backup.

Connections to Multiple Service Providers

Cisco.com



- **Customers with maximum redundancy requirements install physical links to multiple Internet service providers**
- **Redundancy on link, equipment or service provider failure**
- **Primary/backup setup is complex without service provider assistance**
- **Good load sharing is impossible to achieve**

In the example shown, two different edge routers in the customer network have one permanent connection each to different ISPs. Link failures and router failures are covered by the redundancy in exactly the same way as in the example where the two customer routers are connected to two different routers in one ISP network. However, because the two connections in this example go to two different ISPs, the redundancy also covers problems within one ISP network.

The two links may in some cases be implemented by the customer for load sharing and in other cases be used strictly for backup purposes. Controlling load distribution over the links is more complicated in this example. Avoiding any load on the backup link may require assistance from the ISP to which the backup link is connected.

Load sharing between the links in this setup can never be optimal. Equal distribution of the return traffic load from the Internet over the two separate links cannot be done. Distribution of the load of outgoing traffic is done based on destination addresses. Slowly adjusting the appropriate router configuration parameters and observing the link traffic load changes that result, can be used to reach an acceptable distribution of router traffic between the two links.

Practice

Q1) Multiple permanent connections to a single ISP do not provide what level of redundancy?

A) redundancy on link failure

B) redundancy on ISP failure

C) redundancy on equipment failure

D) redundancy on routing protocol failure

Q2) What is a drawback when a customer is installing multiple permanent connections to multiple ISPs?

A) There is no redundancy on ISP failure.

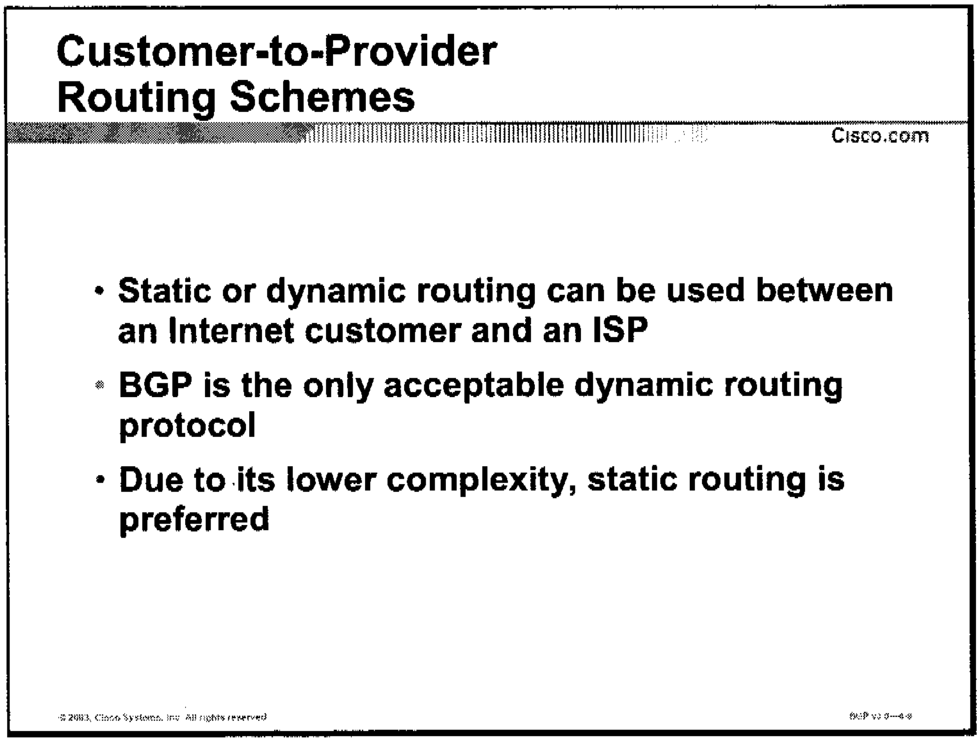
B) Good load sharing is impossible to achieve.

C) The customer can use only Frame Relay PVCs.

D) Equipment failure may cause a complete network outage.

Customer-to-Provider Routing Schemes

This topic identifies different routing schemes that customers use to connect to a service provider.



Different solutions for connecting a customer network to the network of an ISP require different methods of routing information exchange.

- Static routing is preferred due to its lower complexity. In a normal case, the customer network must have a default route to the ISP network and the ISP network must have a route to those IP prefixes that the customer has in its network. As always, static routing provides very low, if any, redundancy.
- Dynamic routing provides redundancy. The customer and the ISP networks must be configured to exchange a common routing protocol. BGP is the only choice due to the large volumes of routing information, the inherent security mechanisms of BGP, and the ability of BGP to handle routing policies.

Practice

Q1) Which form of routing provides the best redundancy?

A) static routing

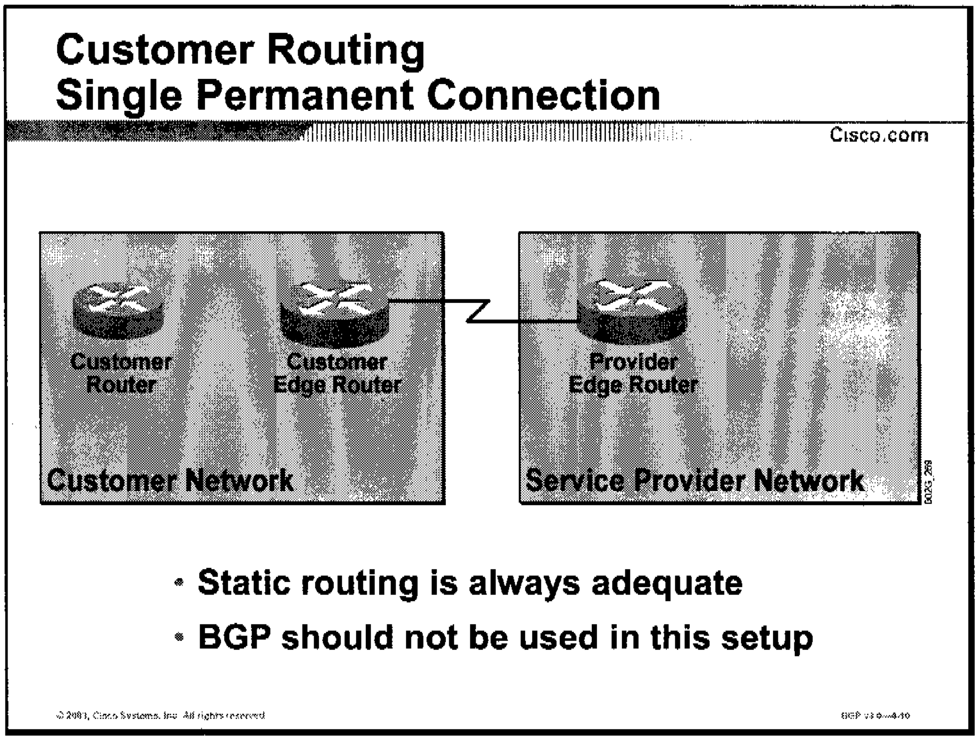
B) content routing

C) dynamic routing

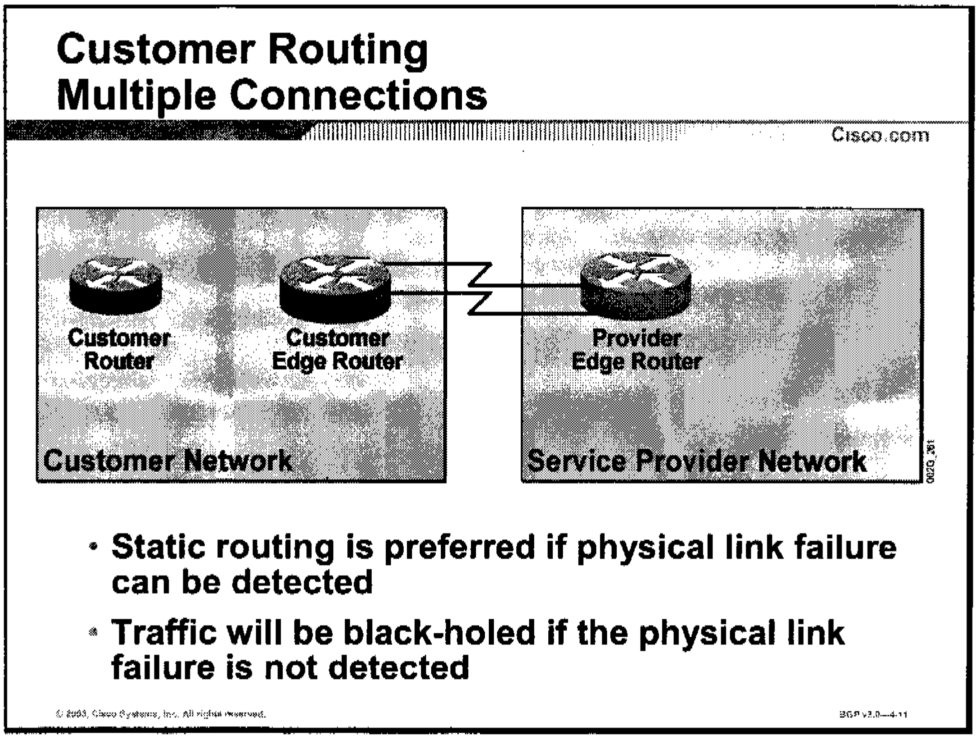
D) embryonic routing

Customer Routing

This topic describes routing schemes that are appropriate for each physical connection type used by customers to connect to a service provider.



In a case where the customer has a single permanent connection to the Internet, static routing is always adequate. The physical topology does not provide any redundancy, and it is therefore unnecessary to add the complexity of dynamic routing. Keep the network simple by avoiding the use of BGP in this case.



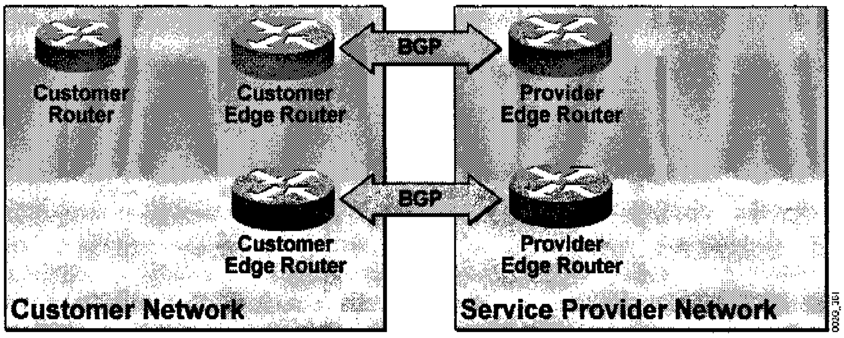
Multiple permanent connections between a single router on the customer network and a single router on the service provider network should be configured with static routing, provided that link failure can be detected by link-level procedures.

With this type of connection, two static routes are configured on each network, pointing to both links between the customer and the ISP. If either of the links fails, the link-level procedures should detect this failure and place the interface in a down state. In this case, the static route is invalid and not used for forwarding packets. The router will subsequently forward all packets over the remaining link.

If the link-level procedures cannot detect a link failure, the static route pointing out over the failed link is still valid. The router continues using this static route to send some of the traffic out on the failed interface. This effectively creates a black hole for some of the traffic.

Customer Routing Multiple Connections (Cont.)

Cisco.com

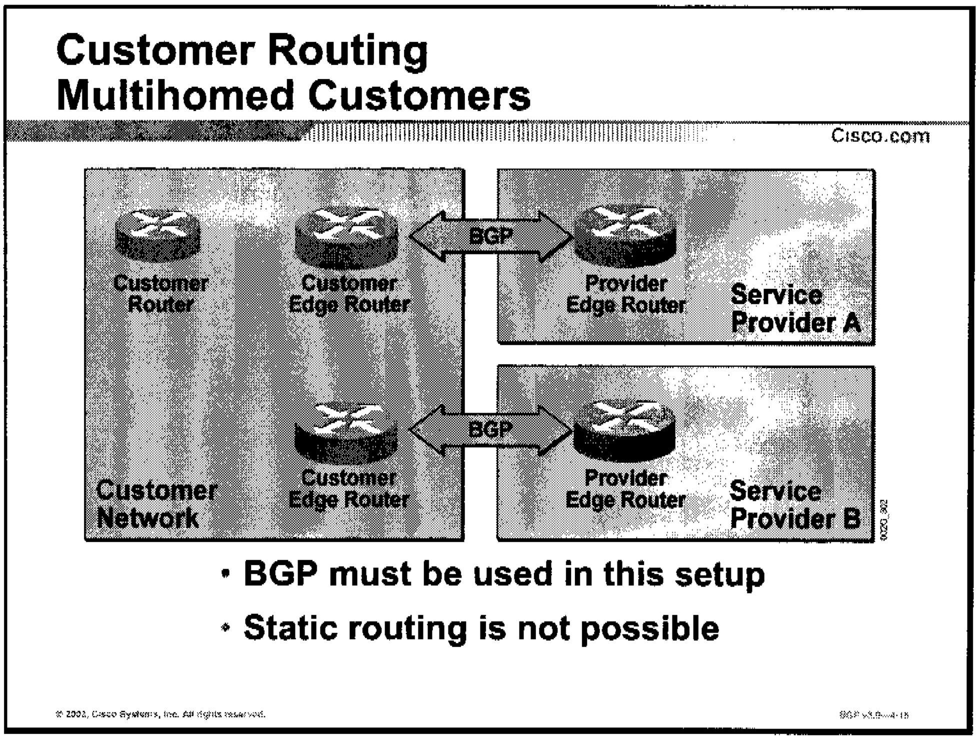


- **You can still use static routing if link and remote equipment failure can be detected reliably**
- **BGP between the customer and the service provider is usually used in this setup**

 BGP v3.0—4-13

You can also use static routing for multiple permanent connections between two different routers on the customer network to two different routers on the service provider network if the failures can be detected by the link-level procedures. When one of the connections is lost, the link-level detects this loss and places the interface in a down state. Because the interface is in the down state, the static route, which points out of the down interface, becomes invalid. As a result, the router stops the redistribution of the static route into BGP.

However, customers requiring the use of multiple connections and multiple routers very often do not rely on the link-level procedures. These customers require a routing protocol such as BGP to detect the failures. Because BGP uses handshaking and reliable transfer, it always detects a failed link or failed remote router.



Multiple permanent connections to more than one ISP always require the use of dynamic routing with BGP. The customers requiring this type of connection do not just want to protect the network connectivity from link failures or remote router failures, they also want to protect their network connectivity from serious problems in the network of an ISP.

Monitoring the link status cannot detect a problem inside one of the ISP networks. If the link is still up and the ISP edge router is still up, the link-level procedures do not indicate any problems. However, the ISP network may suffer from severe problems. An ISP network can be partitioned or disconnected from the rest of the Internet without having any problems with the edge router and the access line to the customer network.

The only way to detect this situation is to use BGP with both ISPs and receive full Internet routing from both of them. When one of the ISPs has problems, the edge router, being the BGP neighbor of the customer, withdraws those routes that it can no longer reach. This action means that the customer routers know which Internet routes each ISP can reach at the moment.

Practice

Q1) What will happen if a link failure is not detected where multiple permanent connections between a single router on the customer side and a single router on the ISP side are configured with static routes?

A) Fast failover will occur.

B) The traffic will encounter a black hole.

C) Customers will get a message that the Internet is down.

D) Nothing happens. All routing continues as usual.

Q2) Multiple permanent connections to more than one ISP always require the use of what?

A) dynamic routing

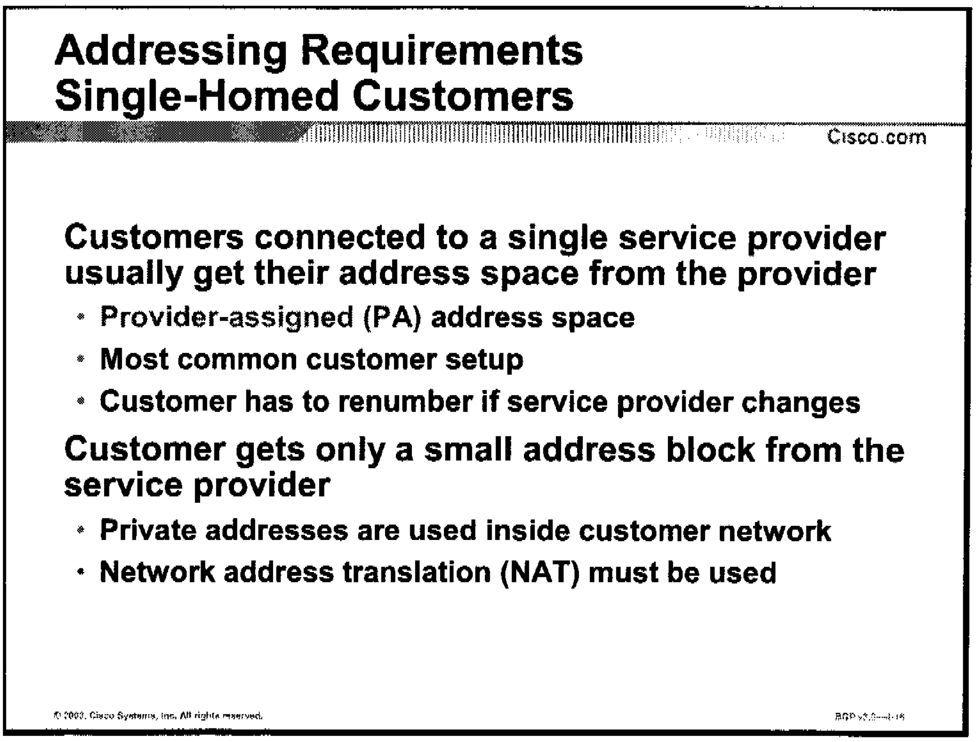
B) mobile routing

C) static routing

D) secure routing

Addressing Requirements

This topic describes the different addressing schemes that customers use to connect to a service provider.



Customers connected to a single ISP usually get their address space assigned by the ISP. An ISP is usually assigned a large address space to delegate to its customers. Because all customers of one ISP get their addresses from one or a few address spaces, it is very likely that the ISP is able to aggregate the customer addresses before sending the routes to the rest of the Internet.

Most customers are connected to a single ISP, which means that they are using provider-assigned (PA) addresses. If the customer should decide to change its service provider, the customer must return its PA addresses to the old ISP and receive a new assignment of PA addresses from the new ISP. Otherwise, the ISPs are no longer able to perform efficient address aggregation.

The consequence for the customer is that the customer has to renumber its network when it changes its service provider.

Some customers decide to use private addresses within their network and do Network Address Translation (NAT) at the connection point to the ISP. This setup means that customers require only a very small portion of public addresses assigned by the ISP. In addition to conserving address space for the benefit of the Internet as a whole, this setup also means that when the customer decides to change its service provider, addresses are renumbered only at the NAT point. The rest of the customer network does not need to be renumbered.

Addressing Requirements Multihomed Customers

Cisco.com

Customers connected to multiple service providers should get their own address space:

- Provider-independent (PI) address space
- No renumbering required on service provider change
- Some service providers might not guarantee routing for small block (for example /24) of PI space

Multihomed customers can sometimes use PA address space:

- Must have a separate public AS number
- The provider must agree to having another ISP advertise its address space

BGP v3.0—4-17

Customers connected to more than one ISP should, if possible, assign their own address space and not have addresses delegated from any of their ISPs. Such assigned addresses are called provider-independent (PI).

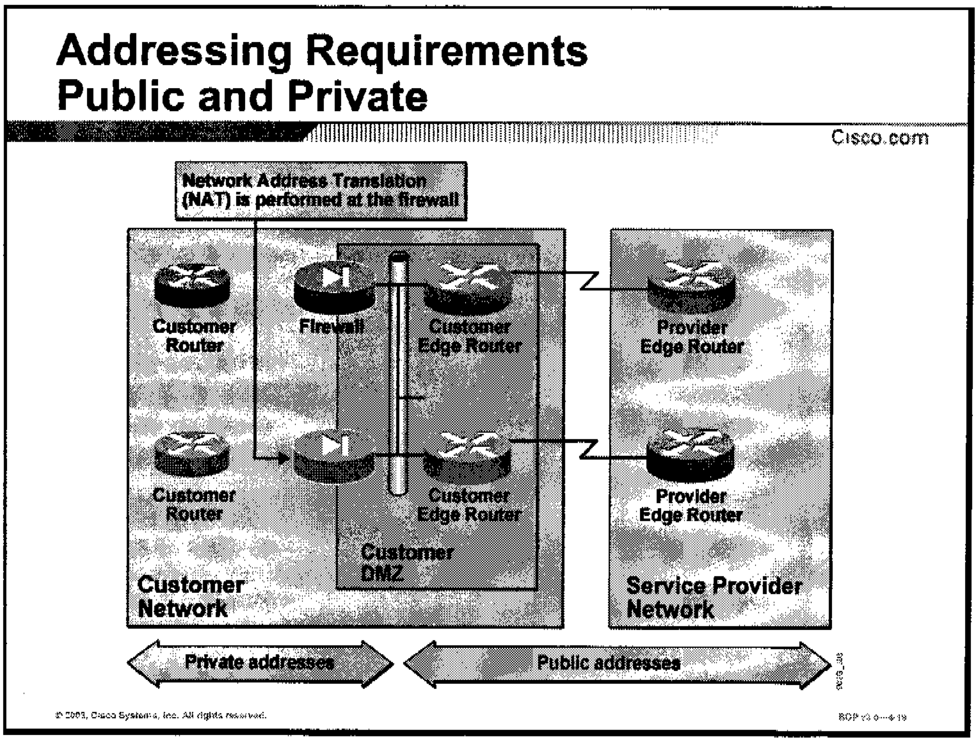
A customer using PI addresses can change its service provider without renumbering its network. The address space is not in any way bound to a particular provider. This arrangement means that no ISP can aggregate the customer routes before sending them to the rest of the Internet. The routes propagate through the Internet with the prefix lengths given.

Some large ISPs filter out routes with long prefixes. ISPs do not want to populate their routing tables with a large number of explicit routes that should have been aggregated into a route summary before they were sent to them. As a result, the customer announcing small blocks of PI addresses, which cannot be aggregated, may not be reachable from all parts of the Internet. A larger block of PI addresses solves the problem.

A multihomed customer can in some cases use PA addresses. The address space must be assigned from one of the ISPs. When the customer announces the block of PA addresses to both ISPs, both should propagate the addresses to the rest of the Internet. The provider that assigned the address space should also announce the larger block of addresses, of which the customer is announcing a subset.

Other ISPs now receive two alternate explicit routes and an overlapping route summary. Filtering out explicit routes is more likely at this time because the other ISPs recognize these as routes that can be aggregated. If the other ISPs filter out the more explicit routes, the customer is still reachable as long as both providers are announcing the overlapping route summary.

Example



In this example, the customer uses private addresses inside its own network. Only a very small network segment, called the customer demilitarized zone (DMZ), has been assigned public addresses.

The customer network is connected to the customer DMZ using two alternate firewalls with both firewalls doing NAT. All packets leaving the customer network have their addresses translated to a public address belonging to the DMZ subnet. The reverse translation is made in the reverse traffic direction.

In this case, the customer requires only a very small block of public addresses. These addresses can be PA addresses. If the customer decides to change its service provider, renumbering is not a problem because only a few devices need to be reconfigured by the customer.

Care must be taken so that traffic flows symmetrically through the firewalls. Otherwise NAT does not work. The easiest way to achieve this symmetry is to let only one firewall be active at a time.

Practice

Q1) When a customer decides to use private addresses within its network, what must it do to connect to its ISP?

A) No action is required.

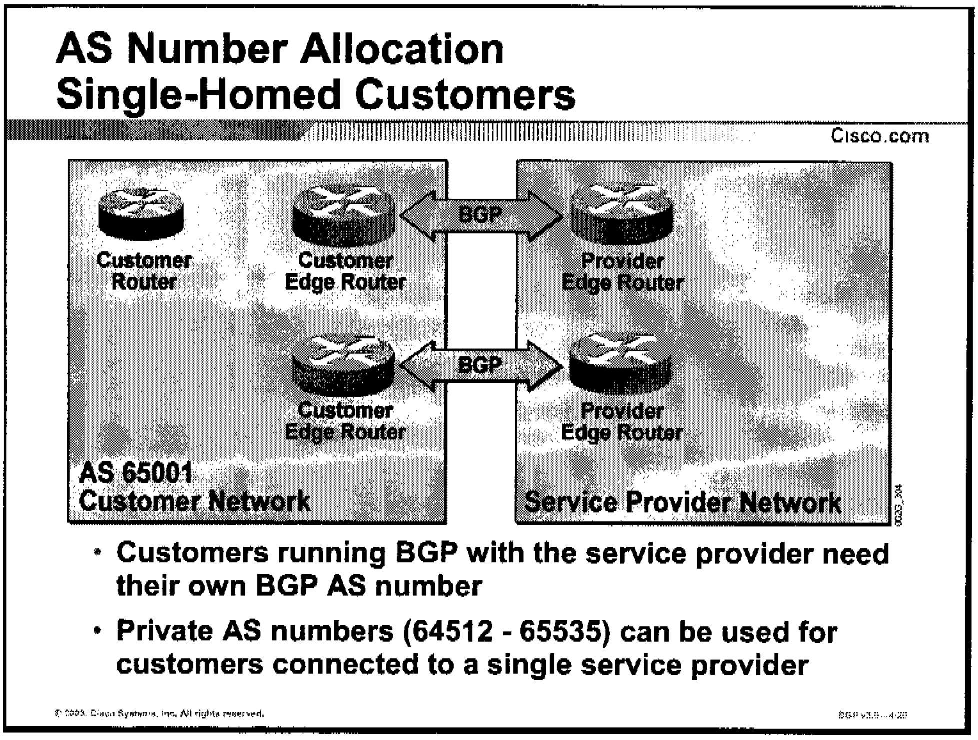
B) The customer must summarize its address space.

C) The customer must configure static addressing.

D) The customer must use NAT.

AS Number Allocation

This topic describes different AS numbering schemes that customers use to connect to a service provider.



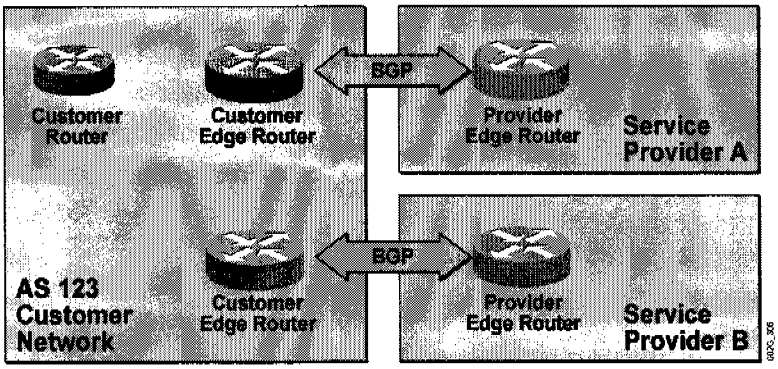
BGP requires the use of AS numbers. When BGP is configured, the AS number is mandatory information. However, public AS numbers are a scarce resource. Customers should use public AS numbers only when required. A customer who uses BGP to exchange routing information with only one ISP does not require a public AS number. This customer can use a private AS number.

An ISP network, which is running BGP with some of its customers, must determine whether a public or a private AS number is required for each customer. When the customer can use a private AS number, the ISP must allocate one from the range of private AS numbers (64512 – 65535). The ISP must make sure not to assign any of the private AS numbers to more than one customer.

When the ISP receives BGP routes from the customer, the ISP routers see the private AS number in the AS path and treat the private number as any other AS number. However, before the ISP propagates any of these routes to the rest of the Internet, it must remove the private AS numbers from the AS path, because the same AS number may be in use by someone else. After the private AS number is removed, the route appears as belonging to the public AS of the ISP.

AS Number Allocation Multihomed Customers

Cisco.com



- **Multihomed customers must run BGP with their service providers**
- **Multihomed customers must use public AS numbers for their autonomous system**

 BGP v3.0—4-21

A multihomed customer requires a public AS number and must run BGP with both of its ISPs. The customer should not use a private AS number because both ISPs must propagate the customer routes to the rest of the Internet. If the customer does use a private AS number, and both ISPs remove the number before sending it to the rest of the Internet, then the customer routes will appear to be local in the public AS of both ISPs. To make BGP work correctly, multihomed customers need to avoid this situation.

Note	With the help of the AS number translation feature, private AS numbers can also be used for multihomed customers, but this type of configuration is not encouraged.

Multihomed customers are correctly connected to the Internet by assigning a public AS number to the customer network. This public AS appears in the AS path and should be propagated by the service provider to the rest of the Internet. The customer network is now reachable by the rest of the Internet through both providers. The route with the shortest AS path is used by Internet endpoints as the best route to the customer network.

Practice

Q1) What number ranges correctly indicate private AS numbers?

A) 1024 – 2048

B) 32768 – 64511

C) 64512 – 65535

D) 65536 – 131072

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- Different customers have different requirements for their Internet connections. These connectivity options include a single connection to a single ISP, multiple connections to the same ISP, and multiple connections to different ISPs.
- The least redundant, and most common, connection is a single permanent connection to a single service provider. Maximum redundancy can be achieved by connecting the customer network to two different ISPs. This setup safeguards against an ISP that has a serious problem within its own network.
- Depending upon the networking requirements of the customer, static and dynamic routing may be used when the customer connects to a service provider.

BGP v3.0—4-22

Summary (Cont.)

Cisco.com

- Customers connected to more than one ISP will need to assign their own provider independent (PI) address space and not have addresses delegated from any of their ISPs.
- Because public addresses are limited, private addresses are commonly used on private networks, and NAT is used to translate the private addresses to public for external connectivity.
- Whenever BGP is in use, an AS number is required. The customer does not need a public AS number if it is connected to a single ISP. The ISP can assign a private AS number in the range 64512-65535 to the customer and remove the information about that AS before the ISP propagates the customer routes to the rest of the Internet. A multihomed customer, however, requires a public AS number.

BGP v3.0—4-23

Next Steps

After completing this lesson, go to:

- Implementing Customer Connectivity Using Static Routing lesson

References

For additional information, refer to these resources:

- For more information on connecting to a service provider, refer to "The Easy Guide to Selecting an Internet Service Provider" at the following URL: http://www.cisco.com/warp/public/cc/so/cuso/smso/crn/ezgd_pl.htm

- For more information on how NAT works, refer to "How NAT Works" at the following URL: http://www.cisco.com/en/US/tech/tk648/tk361/technologies_tech_note09186a0080094831.shtml

- For more information on configuring connections to service providers, refer to "Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments" at the following URL: http://www.cisco.com/warp/public/459/40.html

Quiz: Customer-to-Provider Connectivity Requirements

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify different physical connections that customers use to connect to a service provider
- Describe the levels of redundancy provided by each physical connection type that customers use to connect to a service provider
- Identify different routing schemes that customers use to connect to a service provider
- Describe routing schemes that are appropriate for each physical connection type that customers use to connect to a service provider
- Describe the addressing schemes that customers use to connect to a service provider
- Describe AS numbering schemes that customers use to connect to a service provider

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) If a customer required additional bandwidth as well as redundancy, what method would be preferred?

A) a single permanent connection to one ISP

B) permanent connections to more than one ISP

C) dial-up connections to more than one ISP

D) multiple permanent connections to one ISP

Q2) What type of redundancy do multiple permanent connections providing load-sharing configuration display?

A) link

B) equipment

C) service provider

D) routing protocol

Q3) In a customer-to-provider routing scheme, what method of routing is preferred due to its lower complexity?

A) policy-based routing

B) dynamic routing

C) content routing

D) static routing

Q4) Why is it that with multiple permanent connections to more than one ISP, the use of dynamic routing with BGP is required?

A) When one of the connections is lost, the link level detects this loss and places the interface in a down state.

B) Monitoring the link status cannot detect a problem inside one of the ISP networks.

C) Static routes detect problems inside one of the ISP networks.

D) It is not required, and static routing may be used.

Q5) What can be done when a customer is assigned only a very small subnet of public addresses?

A) Purchase more addresses as required.

B) Use NAT.

C) Add a service provider.

D) Add links to the same service provider.

Q6) What are two different addressing schemes that customers use to connect to a service provider? (Choose two.)

A) provider-independent

B) customer-independent

C) provider-assigned

D) customer-assigned

Q7) Which two of the following criteria are required for a customer to be multihomed to multiple ISPs? (Choose two.)

A) A public AS number.

B) A private AS number.

C) The customer must run BGP with both of its ISPs.

D) The customer must run BGP with one ISP and may use static routing with the other.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Implementing Customer Connectivity Using Static Routing

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses static routing as a solution for connecting customer networks to service providers. Included in this lesson is a discussion of when static routing should be used between a customer and a provider, how to configure static routing in nonredundant, backup, and load-sharing configurations.

Importance

When a single connection to a service provider or multiple connections to the same service provider are the options that a customer has of connecting to the Internet, static routing is the best routing approach to implement between customer and provider. When implementing customer-to-provider connectivity with static routes, it is important that network administrators understand existing guidelines, as discussed in this lesson. Knowledge of static routing implementation guidelines will aid in successfully deploying static routing network configurations.

Objectives

Upon completing this lesson, you will be able to:

- Identify when to use static routing between a customer and a service provider in a BGP environment
- Describe the characteristics of static routing between a customer and a service provider in a BGP environment
- Identify design considerations for propagating static routes in a service provider network
- Configure static route propagation in a BGP environment, given a scenario with different service levels
- Configure a typical backup setup using static routing between a customer and a service provider in a BGP environment
- Describe the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations
- Describe the characteristics of load sharing when you are configuring static routing between a customer and a service provider

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

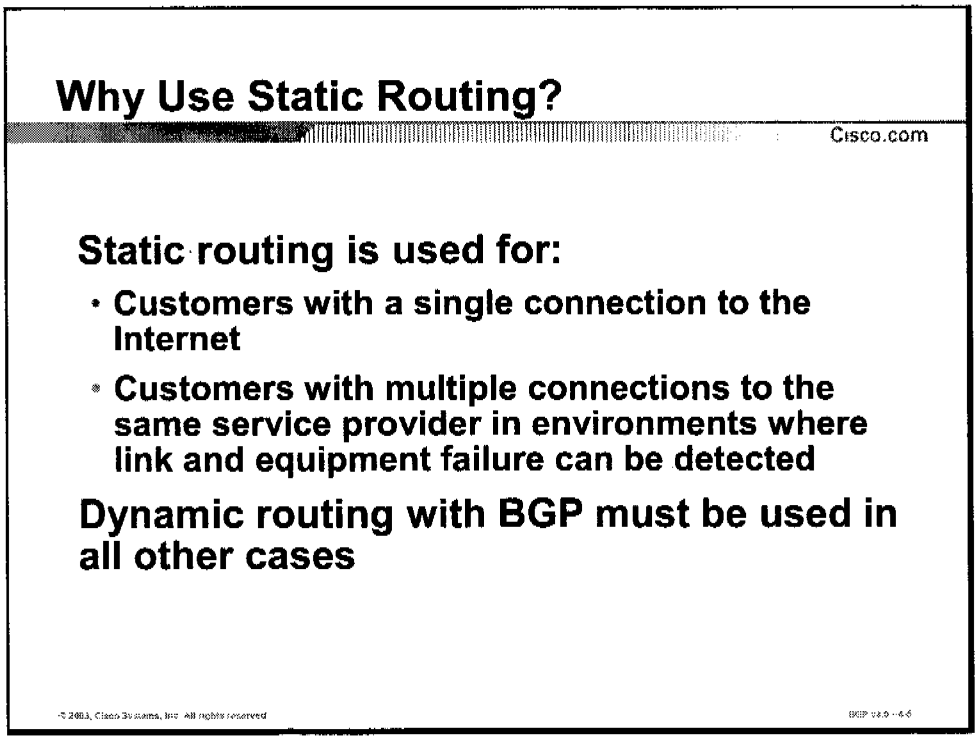
Outline

This lesson includes these topics:

- Overview
- Why Use Static Routing?
- Characteristics of Static Routing
- Designing Static Route Propagation in a Service Provider Network
- Static Route Propagation Case Study Parameters
- BGP Backup with Static Routes
- Floating Static Routes with BGP
- Load Sharing with Static Routes
- Summary
- Assessment (Quiz): Implementing Customer Connectivity Using Static Routing

Why Use Static Routing?

This topic identifies when to use static routing between a customer and a service provider in a Border Gateway Protocol (BGP) environment.



Static routing is the best solution to implement when there is no redundancy in the network topology. A single connection between the customer network and the service provider network does not provide any redundancy. If the link goes down, the connection is lost regardless of what routing protocol is configured in the customer or provider network. When there are redundant connections between the customer network and the network of a single service provider, static routing can be used under specific circumstances.

A static, default route must be conditionally announced by the customer edge routers using an Interior Gateway Protocol (IGP). If the link to one of the customer edge routers goes down, then the router must be able to detect the failure and invalidate the static default route. Announcement of this router as a default gateway using an IGP must now cease. Likewise, on the service provider edge routers, the static routes pointing to the customer networks must be invalidated if the link between them goes down, and redistribution to BGP is therefore stopped.

If link-level procedures cannot detect a link failure, the interface remains in the up state. The static routes are not invalidated, and packets are forwarded into a black hole. In this case, since the router cannot detect a failure at the link-level, BGP must be used between the customer and the provider.

BGP must also be used between the customer and the service provider networks when the customer is multihomed. This is the case regardless of what link failure detection mechanisms are in use.

Practice

Q1) Which of these two situations identify when to use static routing between a customer and a service provider in a BGP environment? (Choose two.)

A) customers with a single connection to the Internet

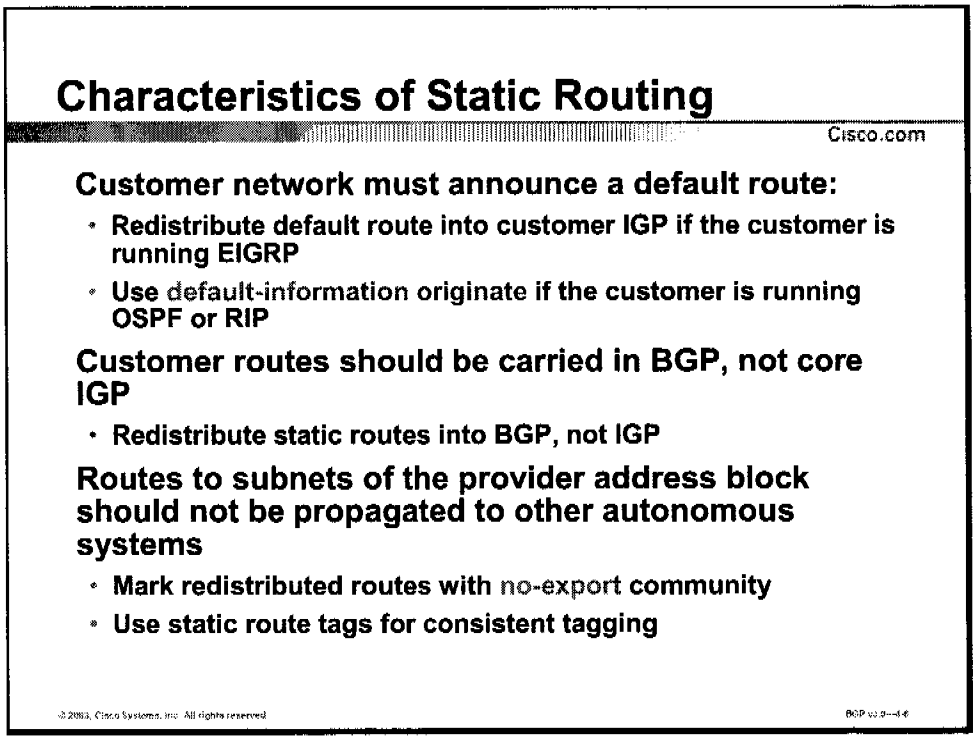
B) customers with multiple connections to multiple service providers

C) customers with multiple connections to the same service provider

D) customers with a single connection to multiple service providers

Characteristics of Static Routing

This topic describes the characteristics of static routing between a customer and a service provider in a BGP environment.



When static routing is implemented between the customer network and the Internet service provider (ISP) network, the edge router of the customer must announce itself as a default gateway or a gateway of last resort. This procedure must be done using the IGP within the customer network because different routers within the customer network must be able to select the best route to the exit point of the network.

Different IGPs use different methods of announcing a router as a gateway of last resort. Enhanced Interior Gateway Routing Protocol (EIGRP) uses the concept of default network while Open Shortest Path First (OSPF) and Routing Information Protocol (RIP) send reachability information about network 0.0.0.0/0. In either case, the network operators of the customer network are responsible for configuring their network to use the customer edge router as a gateway of last resort.

When static routing is used between the customer and the provider, the edge router of the provider must propagate a static route that points to the customer network, to all other routers within the ISP network, and also to the rest of the Internet. The network operators in the ISP network propagate the route using a configuration command to start redistributing the routes into BGP.

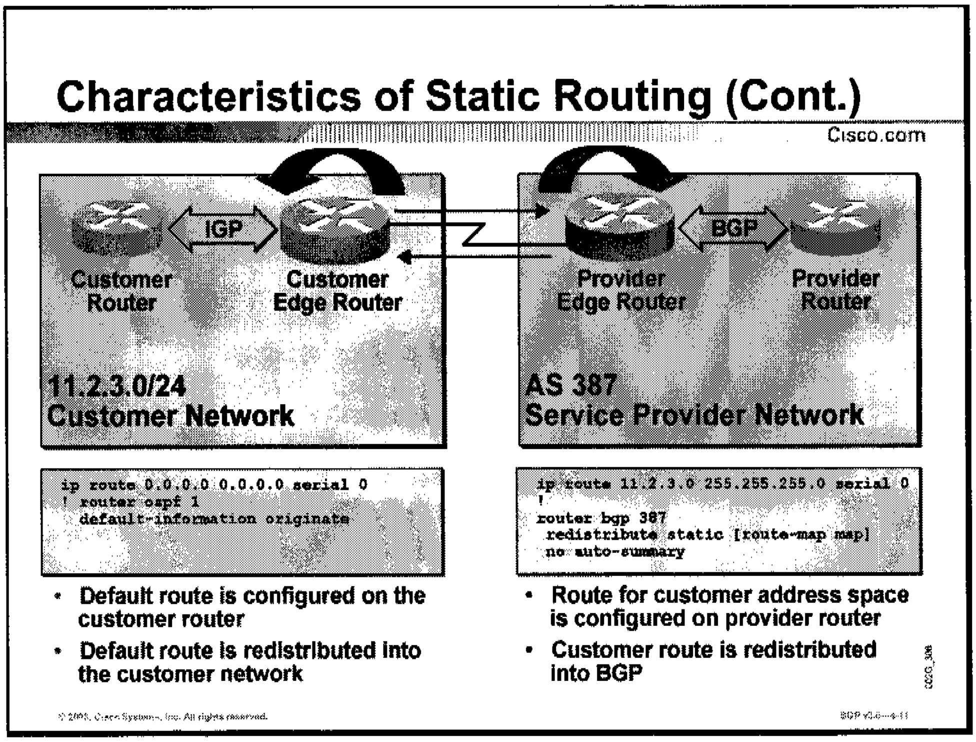
Customer routes should not be redistributed into the IGP of the ISP network. Care should be taken that the IGP of the ISP network does not carry too many routes. Redistributing customer routes into the IGP could potentially cause poor performance and may eventually cause a complete shutdown of IGP routing at the service provider.

If a customer uses provider-assigned (PA) addresses and the ISP announces a large block of addresses where the network of this customer is only a small portion of the block, then the routes of this customer should not be propagated by the service provider to the rest of the Internet. Instead, the rest of the Internet should receive only an announcement containing the larger block of addresses.

An easy way of achieving this setup is to use *communities* within the ISP network. Any customer route that should not be announced to the rest of the Internet is marked using the **no-export** community. To ensure that the BGP communities get propagated, at least over all Internal Border Gateway Protocol (IBGP) sessions, the network operators of the ISP network must configure a **send-community** option for all IBGP neighbors. The edge routers of the ISP network then see the **no-export** community and filter those routes out before sending the update to External Border Gateway Protocol (EBGP) neighbors.

Communities are set using route-maps. A route-map can select routes based on various attributes. One of these attributes is the route tag. Through configuration, a route tag can be assigned by the router to specific static routes. This option means that the network operators of the ISP network can invent a scheme of tagging where all static routes that should not be propagated to other autonomous systems are assigned a specific tag. Then a route-map can select all routes with that tag and assign them the **no-export** community.

Example



In the figure, the customer network is connected to the Internet using a single permanent connection to a single service provider. In this case, a routing protocol does not add any redundancy and would only add complexity.

The customer edge router has a static default route pointing to the interface serial 0. If the serial interface goes down, the route is invalid. The **default-information originate** command is configured in the OSPF process on the customer router; therefore, the router announces a default route into OSPF only as long as it has a valid default route itself.

The service provider edge router also has a static route, declaring the customer IP network number as reachable over the serial 0 interface. It is also invalid if the interface goes into the down state. The ISP edge router must forward this information to all other ISP routers and to the rest of the Internet. This action is accomplished by redistributing the static route into BGP. As long as the static route is valid, BGP announces it. To the rest of the Internet, the customer network appears as reachable within the autonomous system (AS) of the ISP. As far as the rest of the Internet is concerned, the customer is a part of the service provider AS.

Practice

Q1) What could potentially cause poor performance of the service provider IGP routing?

A) when static routing is used

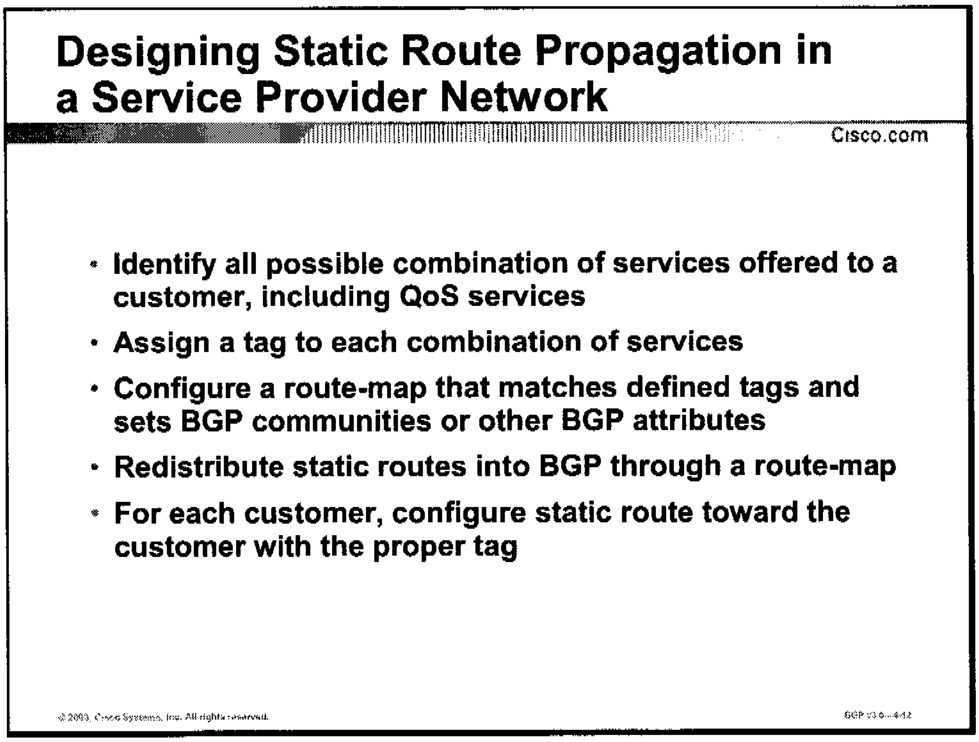
B) redistribution of customer routes into the IGP

C) if a customer uses PA addresses.

D) if the edge router of the customer announces itself as a default gateway

Designing Static Route Propagation in a Service Provider Network

This topic identifies the design considerations for propagating static routes in a service provider network.



You can easily extend the principle of using tags when configuring static routes, and assigning different communities based on those tags, to implement a more complex routing policy:

- Identify all different service levels offered to customers and then all the different combinations of these service levels.
- You must assign each combination its own tag value and its own community.
- You must configure a route-map, which selects routes with each and every one of the assigned tags and sets the corresponding community value. Because the processing of a route-map stops when the match clause of a statement is met, each route should only be assigned a single combination of communities. Therefore, you must take great care to assign a tag and a community combination to each combination of services provided.
- When the provider edge routers redistribute static routes into BGP, these routes must pass through the route-map. BGP assigns the correct community depending on the tag values given on the configuration line for each of the static routes.

- Finally, you need to configure static routes. Before you configure a static route for a specific customer, you must identify the combination of the services provided to this customer. Then you must look up the corresponding tag value. After you have configured the route, you must assign the tag.

With this routing policy, every static route to a customer network is assigned a tag and the redistributed BGP route is assigned a corresponding community. The BGP communities attached to the routes signal to other routers in the ISP network which particular service combination you should use.

Practice

Q1) What must you identify before you configure a static route for a specific customer?

A) the static routes

B) the corresponding tag value

C) the corresponding community value

D) the combination of the services provided to this customer

Static Route Propagation Case Study Parameters

This topic provides a scenario with varied service levels in which static route propagation is configured in a BGP environment.

Static Route Propagation Case Study

Cisco.com

Sample service offering

Addressing

- Provider-assigned address blocks are not propagated to upstream ISPs
- Provider-independent address blocks are propagated to upstream ISP

Quality of Service

- Normal customers
- Gold customers

Define Static Route Tags

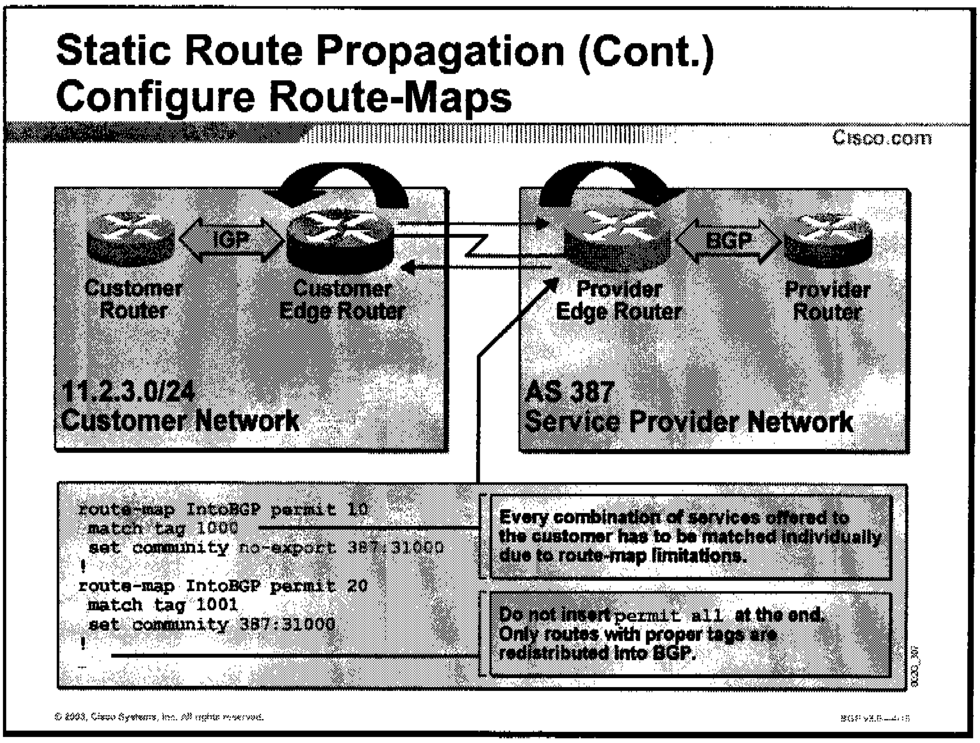
Advertise Customer Route	QoS Type	Route Tag	Community Values
	Normal	1000	no-export 387:31000
Yes	Normal	1001	387:31000
	Gold	2000	no-export 387:32000
Yes	Gold	2001	387:32000

In this case study, the service provider offers two different service levels to its customers: Normal and Gold. Customers are also assigned IP address blocks. Some customers use PA addresses, which the ISP does not announce as explicit routes. The large route-summary block announced by the ISP covers these customers. Other customers use PI addresses that must be explicitly announced to the Internet by the service provider.

Because there are two different quality of service (QoS) services, Normal and Gold, and both PA and PI addresses, the total number of combinations to cover the network policy is four:

- Normal QoS routes, which are assigned by the ISP and should not be explicitly announced
- Normal QoS routes, which are PI routes and should be explicitly announce
- Gold QoS routes, which are assigned by the ISP and should not be explicitly announced
- Gold QoS routes, which are PI and should be explicitly announced

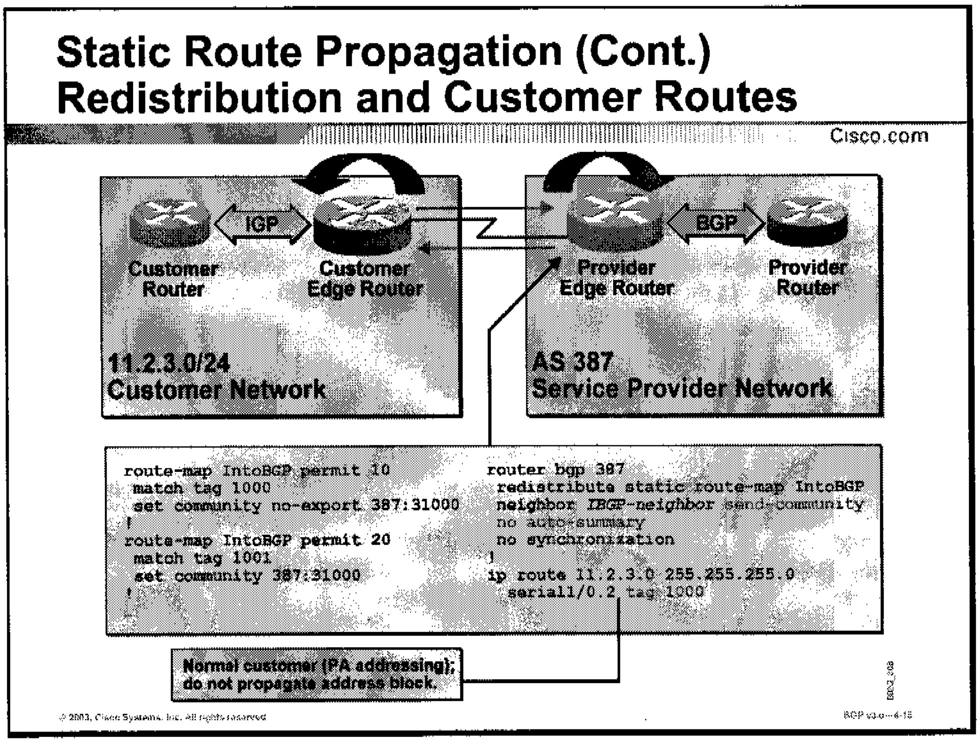
Each of these four combinations receives its own tag value and its own community combination.



Network operators configure a route-map in the ISP edge router that has the static routes to the customer network. Redistribution of the configured customer static routes into BGP is also performed at the ISP edge router.

Because a route-map can match an individual route in one single **route-map** statement only, a single tag value, representing each combination of services, must be assigned to the static routes by the router. When a route is matched, the interpretation of the route-map for that individual route stops. The route-map has one statement for each combination, and each statement matches a tag value and assigns the corresponding community combination for that tag.

The route-map is applied during the redistribution of customer static routes into BGP at the provider edge router. Because the route-map has no **permit any** statement at the end, those static routes not assigned any of the tags being used are not redistributed. The route-map filters these routes out, forcing the network operators to make a tag assignment to all customer routes. Furthermore, the route-map filtering may help catch administrator configuration entry errors, thus giving each and every customer the service combination that they are entitled to.



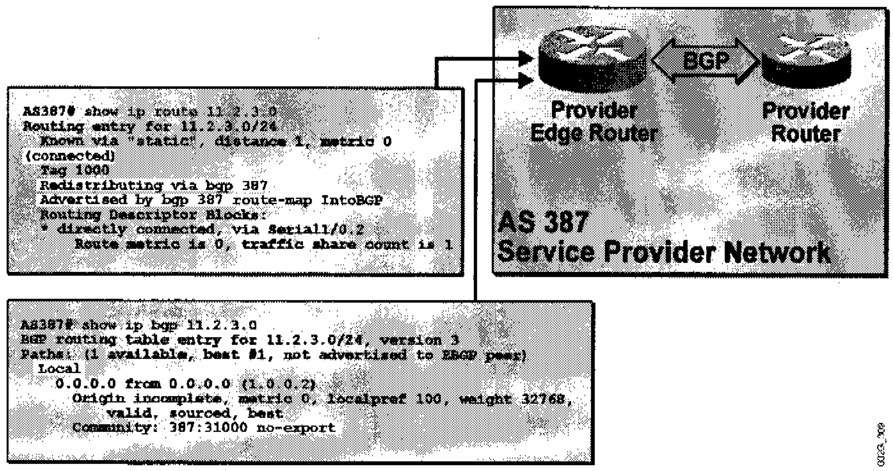
This figure shows how the service provider edge router uses the route-map named *IntoBGP* when redistributing the static routes into BGP. Because the route-map assigns community values that will be used by other routers within the ISP network, network operators must configure all IBGP neighbors with the **send-community** qualifier.

Use the **no auto-summary** BGP configuration command to avoid having the subnet 11.2.3.0/24 automatically summarized into 11.0.0.0/8.

Now, when connecting customers, the network operators identify which service combination to use for this particular customer. The three services associated with this particular customer are: apply normal QoS, use a provider assigned network number, and the provider should not explicitly announce the customer routes. A static route to the customer is configured and assigned the appropriate tag value of 1000, which represents the specified services assigned to the customer.

Static Route Propagation (Cont.)
Static Routes on the Provider Edge Router

Cisco.com



The **show ip route** command displays information from the forwarding table about subnet 11.2.3.0/24. The route is learned by static configuration and is redistributed via BGP. The router, through the use of a statically assigned tag, has assigned a tag value of 1000 to the customer route, and the route must pass through route-map *IntoBGP* before being inserted into the BGP table.

The **show ip bgp** command displays information from the BGP table about subnet 11.2.3.0/24. The route is local within this AS and is sourced by this router. The BGP communities 387:31000 and **no-export** have been assigned by the router to the redistributed customer route using the provider defined route-map prior to inserting the customer route into the BGP table.

Practice

Q1) How many tag value(s) must you assign to each combination of services?

A) 1

B) 2

C) 16

D) 32

Q2) What do you use at the edge router of a service provider to redistribute static routes into BGP?

A) **no-export**

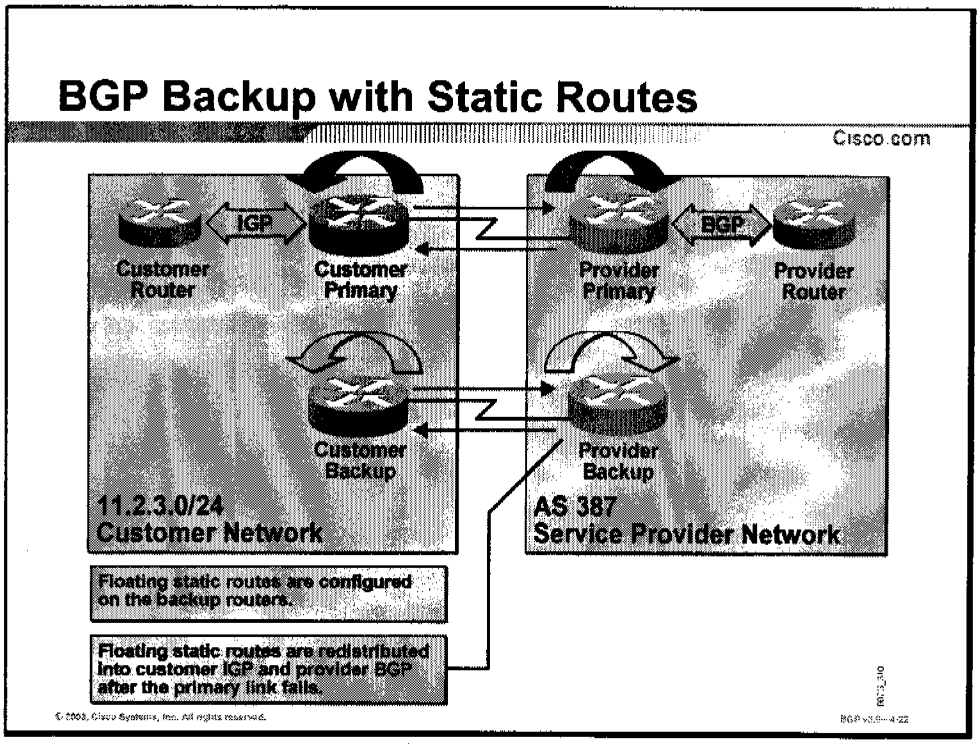
B) **send-community**

C) **route-map**

D) **route-community**

BGP Backup with Static Routes

This topic explains how to configure a typical backup setup using static routing between a customer and a service provider in a BGP environment.



This example illustrates a case where the customer network has two connections to a single service provider. One connection between the customer network and the ISP is the primary connection, and the other connection is used for backup purposes only. If link-level procedures can detect link failures and a failure in the remote router, then static routing can be used instead of a dynamic routing protocol between the customer and provider networks.

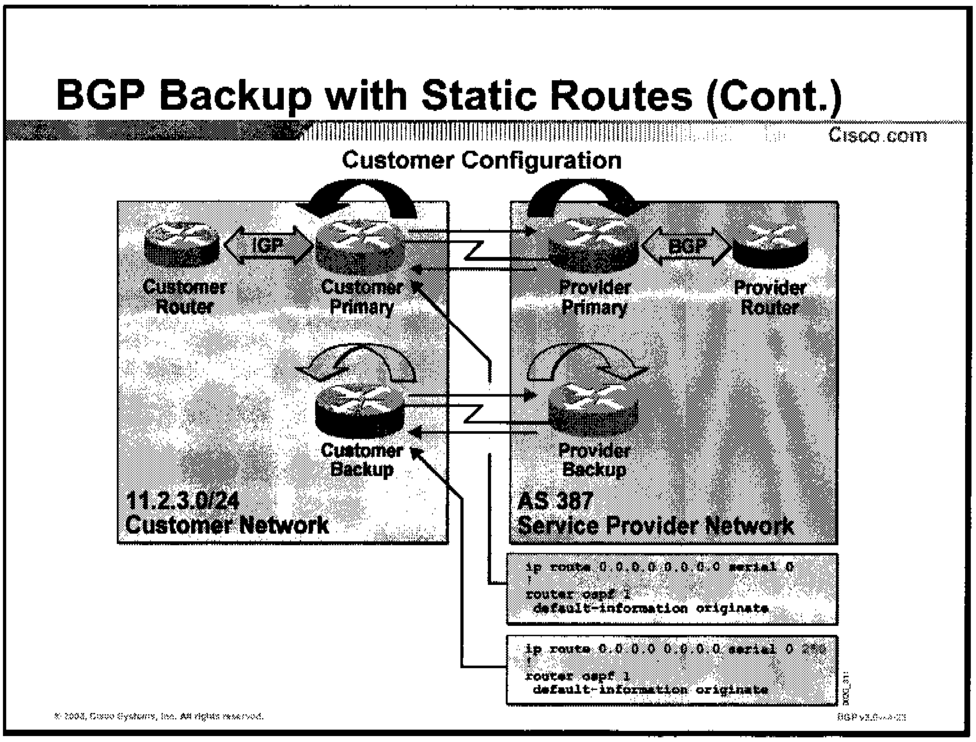
As in the previous example, where no backup link was available, the primary edge router of the customer has a static default route toward the ISP and the primary edge router of the ISP has static routes toward the customer. The customer router redistributes the static default route into its IGP. The ISP router redistributes the static routes into BGP.

If the primary link goes down, the link-level procedures set the interface to the down state, causing the static routes pointing out through the interface to be invalid, and removing the routes from the forwarding table. When the interface changes back to the up state, the static route will reappear in the forwarding table.

Redistribution of routes into any routing protocol is conditioned by the appearance of the route in the forwarding table. Thus, if the interface goes down, the router removes the static route from its forwarding table, and the route is withdrawn from the routing protocol. When the static route reappears, the redistribution process inserts it into the routing protocol again.

The backup edge router of the customer also uses a default static route toward the ISP, via the backup link. The backup edge router is also redistributing the default route into the IGP. However, the static route used is a floating static route, which is assigned a high administrative distance (AD), higher than the AD of the customer IGP. As long as the primary link works, the IGP provides the customer backup edge router with the primary default route. Due to a higher AD, the backup static default route is not installed into the backup router forwarding table. Because the static route is not in the forwarding table, it is not redistributed. If the primary link fails, the IGP no longer feeds the backup edge router with a default route. The backup static default route is the only remaining default route. Therefore, the router will install the floating default route into its forwarding table, and subsequently redistribute it into the IGP.

The backup edge router of the ISP can also use floating static routes, which are redistributed into the ISP BGP process.

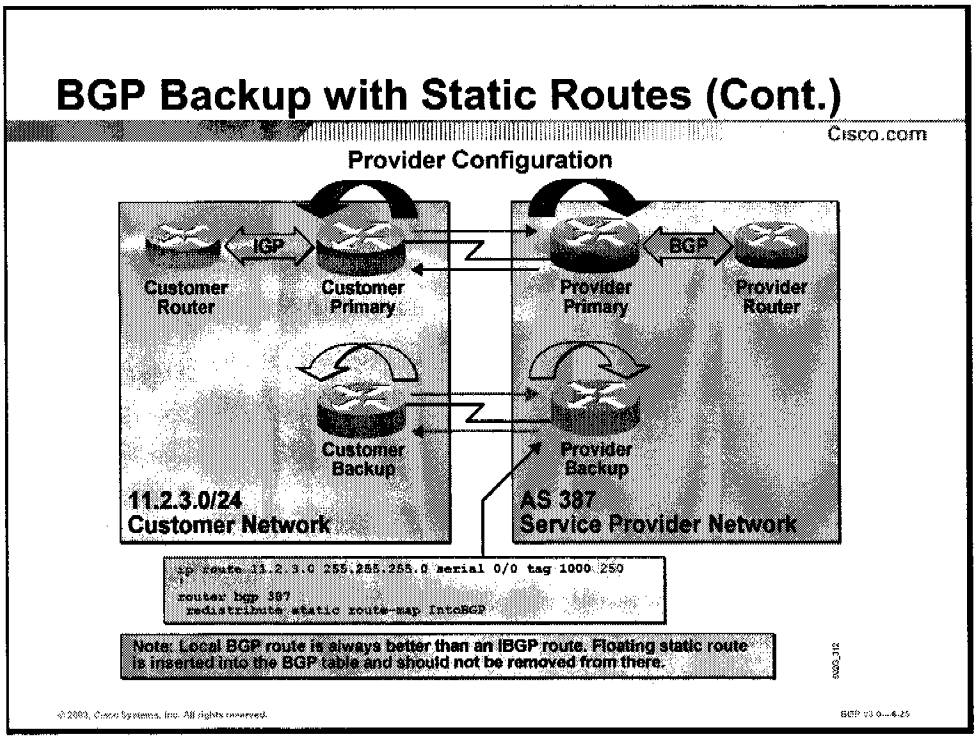


In the figure, the customer network and the ISP network are connected using leased lines with High-Level Data Link Control (HDLC) encapsulation. Both the primary and the backup edge routers in the customer network have a static default route toward the serial interface leading to the ISP. Both routers also do redistribution of the default route into the Open Shortest Path First (OSPF) protocol, which is being used as an IGP within the customer network.

However, the static default route in the backup edge router is configured with an AD value set to 250. This AD value is higher than the AD values of any routing protocol. This configuration means that as long as the backup router receives the default route by OSPF, the static default route is not used.

When the primary link goes down, the static default route in the primary router is not valid. The OSPF protocol stops announcing the default route, because the **default-information originate** command is like a redistribution of the default route from the forwarding table of the router.

The backup router now installs its static default route in the forwarding table. The conditions for announcing the default route by OSPF are met and the rest of the customer routers see the backup router as the gateway of last resort.



When floating static routes are configured on the provider edge routers, they are also redistributed into BGP. This configuration makes things a little bit more complicated.

The network operator configures a floating static route to the customer subnet 11.2.3.0/24. In the provider edge router, the floating static route is assigned the same tag value as the tag value being used in the primary router. The route-map *IntoBGP* is the same as in the primary router and provides the routes to the customer network with the same communities (the same QoS level and indication whether to explicitly announce it to the rest of the Internet).

The floating static route is configured with an AD value set to 250. This value is higher than any routing protocol. When the backup edge router of the ISP no longer receives any routing protocol information about the customer networks, the router will automatically install the floating static route and subsequently redistribute it into BGP.

Based on BGP route selection rules, the redistributed floating static route will always remain the preferred path if additional BGP configuration is not performed on the provider edge router. This preference means that regardless of whether the primary link comes back, the backup router selects the locally sourced route as the best route. Therefore, the backup router continues to announce a path toward the customer network. The backup link does not go back to the Idle state.

BGP Backup with Static Routes (Cont.)

Cisco.com

- **The BGP table on service provider backup router contains the floating static route**

```
AS387_Backup# sh ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 7
Paths: (2 available, best #1, not advertised to EBGP peer)
  Advertised to non peer-group peers:
  10.3.0.5
  Local
    0.0.0.0 from 0.0.0.0 (10.3.0.6)
      Origin incomplete, metric 0, localpref 100, weight 32768, valid,
      sourced, best
      Community: 387:31000 no-export
  Local
    10.3.0.2 (metric 128) from 10.3.0.5 (1.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Originator: 1.0.0.2, Cluster list: 10.3.0.5
      Community: 387:31000 no-export
```

802G_314

In this example, the **show ip bgp** command is used in the backup edge router of the provider to display the information about the customer network 11.2.3.0/24. The primary link has come back, so the backup router now sees two alternate routes. The first route is the route that the router itself has redistributed into BGP using the floating static route. This route is locally sourced by this AS and has been assigned a weight value of 32768. The second route is the one received by IBGP from the primary edge router. This AS also sources this route but no weight value is assigned.

The BGP route selection algorithm selects the route with weight value 32768 as the best. As a result, the route received from the primary edge router is not a candidate to be installed in the forwarding table and never competes with the floating static route. The floating static route stays in the forwarding table, and redistribution of the route continues until the backup link goes down and the route is invalid.

Practice

Q1) What is configured on backup routers and redistributed into the customer IGP and provider BGP after a primary link fails?

A) weighted routes

B) floating static routes

C) floating dynamic routes

D) dynamic forwarding table

Q2) When you are configuring a backup router with BGP static routes, how can you ensure that as long as the backup router is receiving the default route from the IGP that the BGP static route is not used?

A) Assign the static route a low AD value.

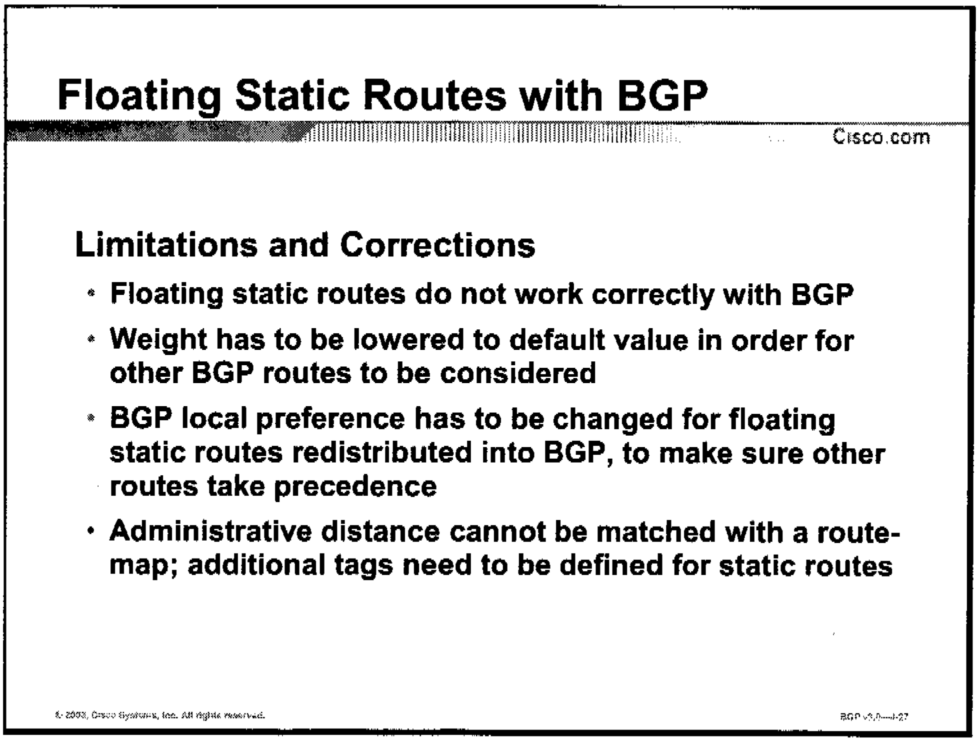
B) Assign the static route a high AD value.

C) Assign the static route an AD equal to that of the IGP.

D) Nothing, the route selection process is automatic.

Floating Static Routes with BGP

This topic describes the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations.



Unfortunately, floating static routes do not work correctly with BGP. After they are inserted, the floating static route is never removed from the forwarding table even if the primary link comes back.

Whenever you use floating static routes in combination with redistribution into BGP, you will need to take additional configuration steps ensure that the BGP route selection algorithm selects the primary route as the best BGP route when it reappears:

- When a router redistributes a floating static route into BGP, the weight value assigned to the floating static route must be reduced. Otherwise, the floating static route will always be selected as the best BGP route after the first failure of the primary link occurs.

- Local preference values must be also be assigned by the router to the floating static route, giving it a lower local preference than the primary route. This assignment makes sure that the primary route is selected as the best BGP route after it comes back.

These two requirements must be specified on the provider edge router in the route-map *IntoBGP* used for the redistribution. The route-map must select the floating static routes and set weight and local preference. However, a route-map cannot do matching based on the AD value assigned to a static route. Some other means are required to make it possible for the route-map to distinguish between normal static routes that should have normal weight and local preference and the floating static ones that should have their values modified.

The solution is to create additional tag values for this set of static routes. The tag value must not only reflect the QoS level and whether to announce the route, but also the tag value must indicate if it is a primary route or a backup route.

Floating Static Routes with BGP (Cont.)

Cisco.com

Sample Static Route Tags with Backup

Advertise Customer Route	Backup	QoS Type	Tag	Community Values	Local Preference
		Normal	1000	no-export 387:31000	100
	Yes	Normal	1010	no-export 387:31000	50
Yes		Normal	1001	387:31000	100
Yes	Yes	Normal	1011	387:31000	50
		Gold	2000	no-export 387:32000	100
	Yes	Gold	2010	no-export 387:32000	50
Yes		Gold	2001	387:32000	100
Yes	Yes	Gold	2011	387:32000	50

There are now eight different tag values identified. Each of tag values indicates a specific combination of explicit route propagation (backup or primary) and QoS level.

When network operators configure static routes in the provider edge router, they must consider which of the combinations that they should use for the route. The route-map that they use when redistributing the static routes into BGP must be configured to recognize all eight combinations and to set the appropriate weight, and community and local preference values.

Floating Static Routes with BGP (Cont.)

Cisco.com

- **The redistribution route-map needs to be updated on all provider edge routers**

```
route-map IntoBGP permit 10
 match tag 1000
 set community no-export 387:31000
 set local-preference 100 !
route-map IntoBGP permit 20
 match tag 1001
 set community 387:31000
 set local-preference 100
```

```
route-map IntoBGP permit 30
 match tag 1010
 set community no-export 387:31000
 set local-preference 50
 set weight 0
!
route-map IntoBGP permit 40
 match tag 1011
 set community 387:31000
 set local-preference 50
 set weight 0
```

Only the first half of the route-map is displayed

BGP v3.0—4-29

The configuration output in this figure displays the first half of the route-map *IntoBGP*. The output shows how four of the eight different tags are identified by match clauses. For each of the tag values, the route-map sets the community, the local preference, and, in some cases, the weight.

Because the displayed half of the route-map deals only with those four tags that indicate QoS Normal, all statements in the configuration display have set the BGP community attribute to 387:31000. The part of the route-map, not shown, deals with the four tags that indicate QoS Gold, which would be configured to set the BGP community attribute to 387:32000.

Tag values of 1000, 1010, 2000, and 2010 indicate that the route should not be explicitly propagated. Those routes that should not be explicitly advertised by the provider to the rest of the Internet are assigned the **no-export** community by the route-map.

Tag values 1010, 1011, 2010, and 2011 all indicate that the route is a backup route. Those tags have their weight value set to 0 and their local preference value set to 50. These settings ensure that upon the return of a failed primary route, the provider edge router will select the primary route as its best path and remove the backup, floating static route, from its route table.

Practice

Q1) Which two of the following statements indicate limitations of floating static routes within BGP? (Choose two.)

A) Weight values cannot be modified.

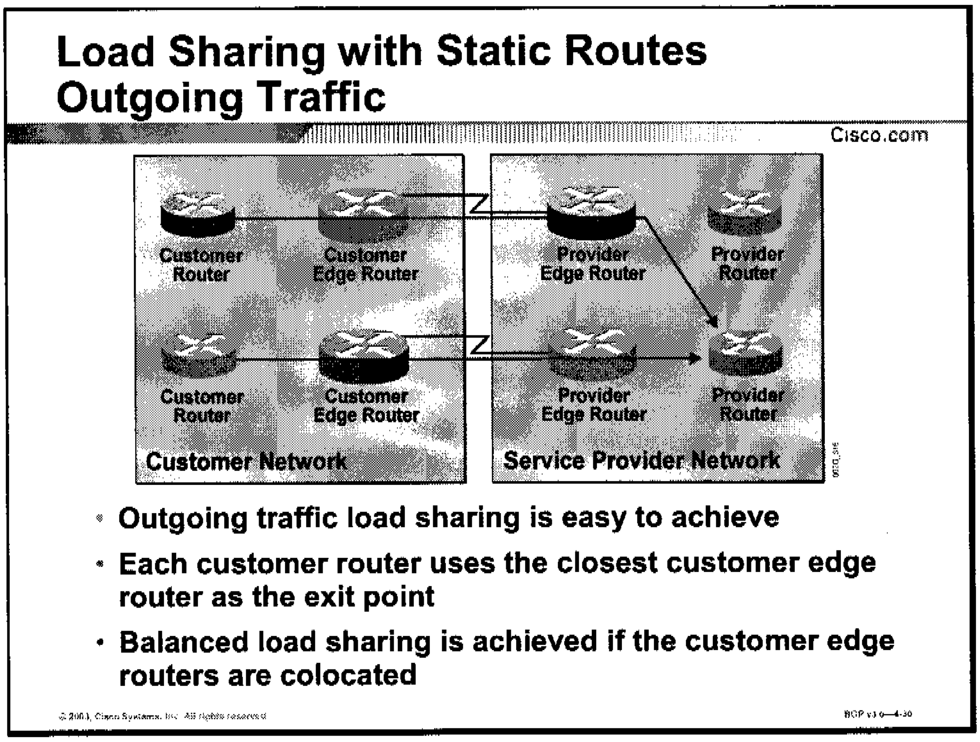
B) Floating routes do not work correctly.

C) AD cannot be matched with a route-map.

D) There is no way to define additional tags for static routes.

Load Sharing with Static Routes

This topic describes the characteristics of load sharing when you are configuring static routing between a customer and a service provider.



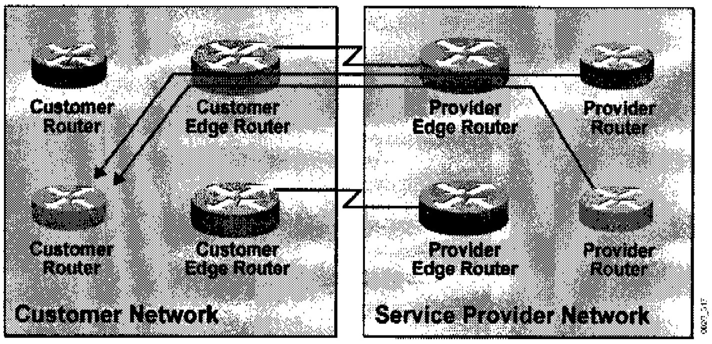
Load sharing of outgoing customer traffic is accomplished by configuring a standard default static route in both customer edge routers. Each static route is valid as long as the serial link in each router is up. When both static routes are valid, both customer edge routers announce the default route into the customer network.

The remaining routers in the customer network see two candidate gateways of last resort. These remaining routers choose the closest one, with respect to the IGP metric. The part of the network that is closer to the uppermost exit point uses that exit point for all outgoing traffic. The other part of the network uses the other (lower) exit point.

If both exit points are colocated, they are equally distant from each of the other routers in the customer network. Each router within the customer network therefore uses load sharing of traffic sent out both exit points.

Load Sharing with Static Routes Return Traffic

Cisco.com



Load sharing of return traffic is impossible to achieve with multiple edge routers:

- **All provider routers select the same BGP route to the destination**
- **All return traffic arrives at the same provider edge router**

The provider routers receive routes toward the customer network via BGP. BGP in its default behavior selects a single route as the best route, allowing no load sharing. The provider routers that receive the same BGP route from two edge routers will always select the closer edge router (if all other BGP attributes are equal, the IBGP route with the closer next hop is selected). The part of the ISP network that is closer to the uppermost connection uses that connection. The other part of the ISP network uses the other (lower) connection.

If both connection points are colocated, all provider routers select the same IBGP route based on router-ID (because the IGP metrics are always equal) and all the return traffic is sent over a single link toward the customer network, resulting in no load sharing.

Note	Since Cisco IOS® version 12.2, 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.

Load Sharing with Static Routes Optimizing Return Traffic

Cisco.com

You can optimize return traffic load sharing

- **Each provider edge router advertises only part of customer address space into the provider backbone**
- **Every provider edge router also advertises the whole customer address space for backup purposes**

Load sharing is not optimal—every link will carry return traffic for part of customer address space

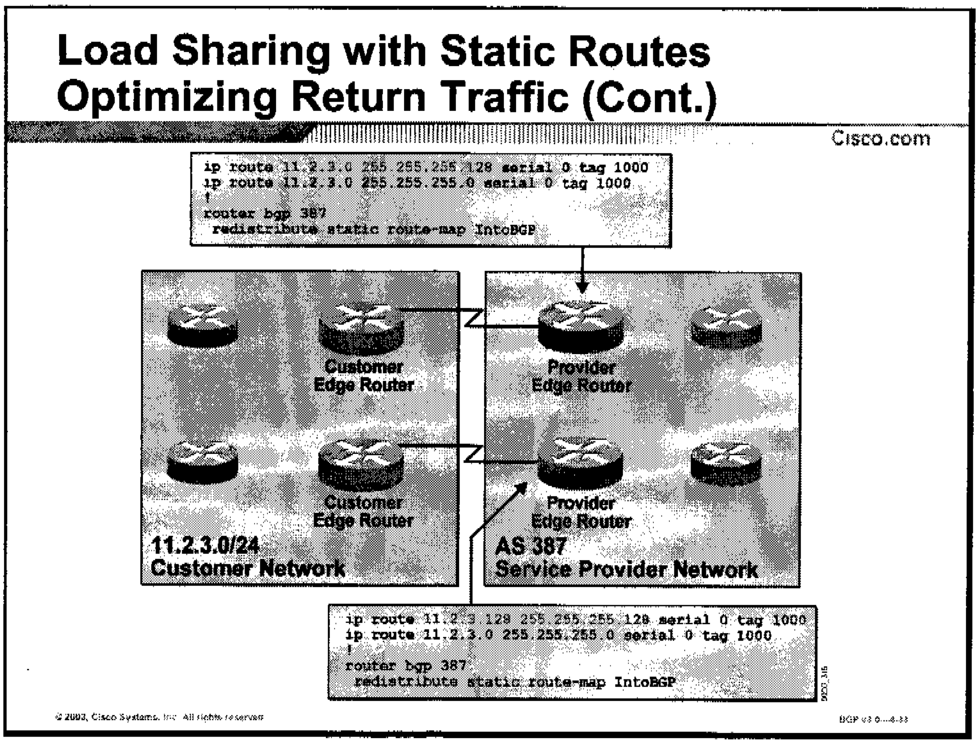
BGP v3.0—4-32

To obtain better control of the return traffic load, the customer address space must be advertised to the provider edge routers using multiple, more explicit routes. The upper edge router could advertise half the address space, and the lower edge router could advertise the other half. For backup reasons, they also should both advertise the entire address space as a larger route summary.

As long as both paths are available, the traffic from the ISP to the customer uses the most explicit route. In this case, two explicit routes are used to send traffic representing one half of the address space over one link and traffic representing the other half of the address space over the other link.

Load sharing in this way does not result in an equal load on the links but rather a statistically based distribution of the traffic load over the links.

Example



In the example here, the customer address space 11.2.3.0/24 is partitioned into two smaller blocks: 11.2.3.0/25 and 11.2.3.128/25.

The upper provider edge router advertises the route to 11.2.3.0/25, and the lower router advertises the route to 11.2.3.128/25. Both edge routers also advertise the entire address space 11.2.3.0/24.

The routers in the ISP network direct traffic with destination addresses in the 11.2.3.0/25 range to the upper connection point. Traffic to destinations in the 11.2.3.128/25 range is directed to the lower connection point.

Practice

Q1) When you are performing load sharing of outgoing traffic with static routes, what is the effect of colocating the edge routers?

A) greater throughput

B) faster convergence

C) higher availability

D) balanced load sharing

Q2) When you are using static routes, with what two routing tricks can you optimize return traffic load sharing? (Choose two.)

A) Each provider edge router advertises only part of the customer address space into the provider backbone.

B) Each provider edge router advertises the entire address space of the customer into the provider backbone.

C) Each provider edge router advertises the entire address space of the customer for backup purposes.

D) Each provider edge router advertises only part of the customer address space for backup purposes.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- You can use static routing in most cases when the customer network is connected to a single ISP. If there is a single connection, you should always use static routing, because there is no redundancy.
- If there are multiple links between the customer network and a single ISP, you can use static routing provided that the static route is invalidated if the link goes down. Detection of down links must be done by link-level procedures.
- When you use static routing, the customer should have a static default route to the ISP and the ISP should have static routes to the entire customer address space.
- The static default route should be redistributed into the customer IGP. The static routes to the customer should be redistributed into BGP. Redistribution must be conditioned by the availability of the connection link if there is redundancy.

 BGP v3.0—4-34

Summary (Cont.)

Cisco.com

- Depending upon the origin of the customer address space, the provider may elect not to advertise the customer space, choosing to advertise a larger aggregate route instead.
- When you are using static routes in a backup scenario, floating static routes are used on the backup routers. After the backup floating static route becomes active, its administrative distance is ignored by BGP because the locally originated route will have a higher weight and be preferred, requiring the use of BGP attributes to ensure proper floating static operation.
- Load balancing can be achieved for outgoing traffic. Return traffic causes problems when multiple connections exist to more than one provider router. The best that can be done is to split the address space for return traffic balancing purposes.

 BGP v3.0—4-35

Next Steps

After completing this lesson, go to:

- Connecting a Multihomed Customer to a Single Service Provider lesson

References

For additional information, refer to these resources:

- For more information on multihoming, refer to "Sample Configuration for BGP with Two Different Service Providers (Multihoming)" at the following URL: http://www.cisco.com/warp/public/459/27.html

- For more information on static routing, refer to "Routing TCP/IP Volume 1," Jeff Doyle, Cisco Press, 1998. ISBN 1-57870-041-8

Quiz: Implementing Customer Connectivity Using Static Routing

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify when to use static routing between a customer and a service provider in a BGP environment
- Describe the characteristics of static routing between a customer and a service provider in a BGP environment
- Identify design considerations for propagating static routes in a service provider network
- Configure static route propagation in a BGP environment, given a scenario with different service levels
- Configure a typical backup setup using static routing between a customer and a service provider in a BGP environment
- Describe the limitations of floating static routes when used in typical backup static routing scenarios and the corrective actions to overcome these limitations
- Describe the characteristics of load sharing when you are configuring static routing between a customer and a service provider

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are two circumstances where you can use static routing as part of installing redundant connections between the customer network and a single service provider network? (Choose two.)

A) The router must be able to detect a link failure.

B) The default route must be announced using the customer IGP.

C) If one link goes down, the interface must remain in an up state.

D) The customer IGP must continue to advertise the static default route.

Q2) A customer route that should not be announced to the rest of the Internet is marked using what?

A) a route tag

B) the **export** community

C) the **no-export** community

D) the public address filter

Q3) When you are designing static route propagation in a service provider network, what three steps must you take? (Choose three.)

A) Assign a tag to each combination of services.

B) Configure a community that matches defined tags.

C) Redistribute static routes into BGP through a route-map.

D) Identify all possible combinations of services offered to a customer.

Q4) What does a route-map assign that will be used by other routers within a network?

A) a tag

B) community values

C) public addressing

D) QoS

Q5) What three key pieces of information can you derive from the following router command output? (Choose three.)

```
AS387_Backup# sh ip bgp 11.2.3.0
BGP routing table entry for 11.2.3.0/24, version 7
Paths: (2 available, best #1, not advertised to EBGP peer)
  Advertised to non peer-group peers:
  10.3.0.5
  Local
    0.0.0.0 from 0.0.0.0 (10.3.0.6)
      Origin incomplete, metric 0, localpref 100, weight 32768,
valid,
      sourced, best
      Community: 387:31000 no-export
  Local
    10.3.0.2 (metric 128) from 10.3.0.5 (1.0.0.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Originator: 1.0.0.2, Cluster list: 10.3.0.5
      Community: 387:31000 no-export
```

A) The primary link has come back, so the backup router now sees two alternate routes.

B) The primary link has not come back up, but the backup router still sees two alternate routes.

C) The first route is the route that the router itself has redistributed into BGP using the floating static route. This route is locally sourced by this AS and has been assigned a weight value of 32768.

D) The second route is the one received by IBGP from the primary edge router. This AS also sources this route, but no weight value is assigned.

Q6) What two things can you do to overcome the problems that occur when a floating static route is redistributed into BGP? (Choose two.)

A) You must raise the weight value.

B) You must lower the weight value.

C) You must set the AD at a higher value than all other routes.

D) You must assign local preference values, giving the floating static route a lower local preference value than the primary route.

Q7) What are three characteristics of using static routes during load sharing of outgoing traffic? (Choose three.)

A) Outgoing traffic load sharing is easy to achieve.

B) Each customer router uses the closest customer edge router as the exit point.

C) Balanced load sharing is achieved if the customer edge routers are colocated.

D) Local preference values must be assigned, giving the floating static route a lower local preference value than the primary route.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Connecting a Multihomed Customer to a Single Service Provider

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses the use of multiple connections between a customer and a single service provider for backup and load-sharing purposes. Included in this lesson is a discussion of how to configure a customer network and a provider network to accommodate multiple connections between them. Also discussed in this lesson are topics specific to networks with multiple connections between a customer and a single provider such as private autonomous system (AS) number removal and configuration of a network to support either backup links or load sharing (balancing).

Importance

When multiple connections to the same service provider are the only means that a customer has of connecting to the Internet, it is important that the connections are correctly configured to ensure proper interaction between the customer and service provider network. It is also important to understand how to configure routing protocols so that customer backup or load-balancing requirements are met.

Objectives

Upon completing this lesson, you will be able to:

- Configure BGP on a customer network to establish routing between a multihomed customer and a single service provider
- Configure conditional advertising of a customer address space when BGP is used to establish routing between a multihomed customer and a single service provider
- Configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider
- Disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path
- Configure a typical backup setup between a multihomed customer and a single service provider in a BGP environment
- Describe how load sharing can be implemented between a multihomed customer and a single service provider in a BGP environment
- Identify the Cisco IOS® command required to configure load sharing between a multihomed customer and a single service provider using BGP multipath
- Configure load sharing between a multihomed customer and a single service provider using EBGP multihop

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

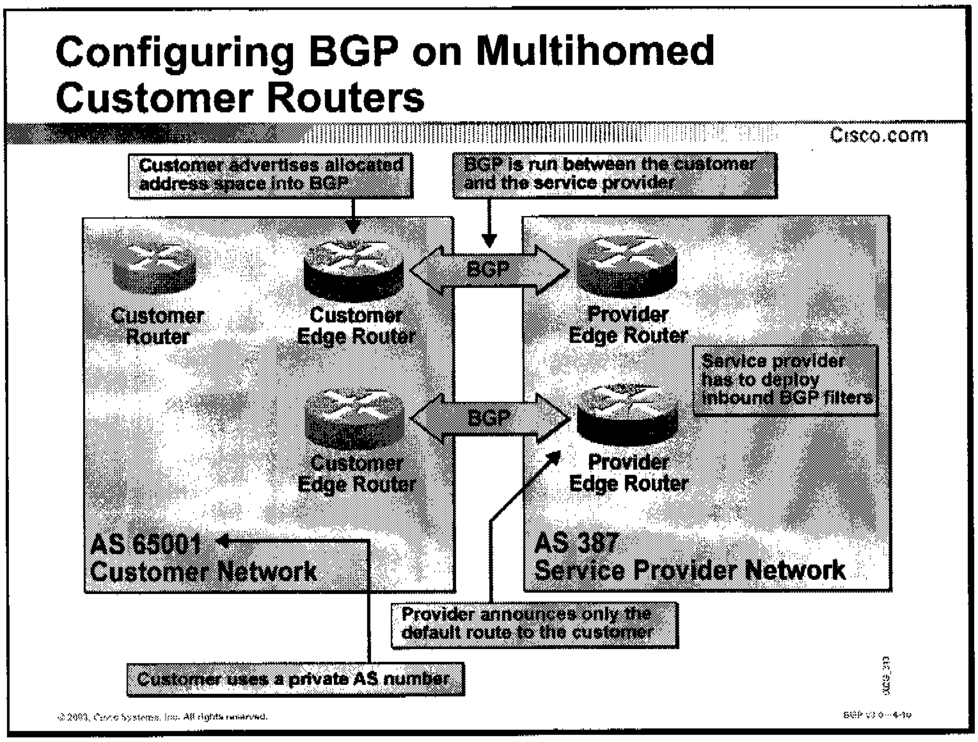
Outline

This lesson includes these topics:

- Overview
- Configuring BGP on Multihomed Customer Routers
- Conditional Advertising in Multihomed Customer Networks
- Configuring BGP on Service Provider Routers
- Removing Private AS Numbers
- Backup Solutions with BGP
- Load Sharing with the Multihomed Customer
- Load Sharing with BGP Multipath
- Load Sharing with EBGP Multihop
- Summary
- Assessment (Quiz): Connecting a Multihomed Customer to a Single Service Provider

Configuring BGP on Multihomed Customer Routers

This topic describes how to configure Border Gateway Protocol (BGP) on a customer network to establish routing between a multihomed customer and a single service provider.



In the case study here, the customer network is connected to a service provider network using multiple permanent links. BGP is used to exchange routing information between the customer and the provider.

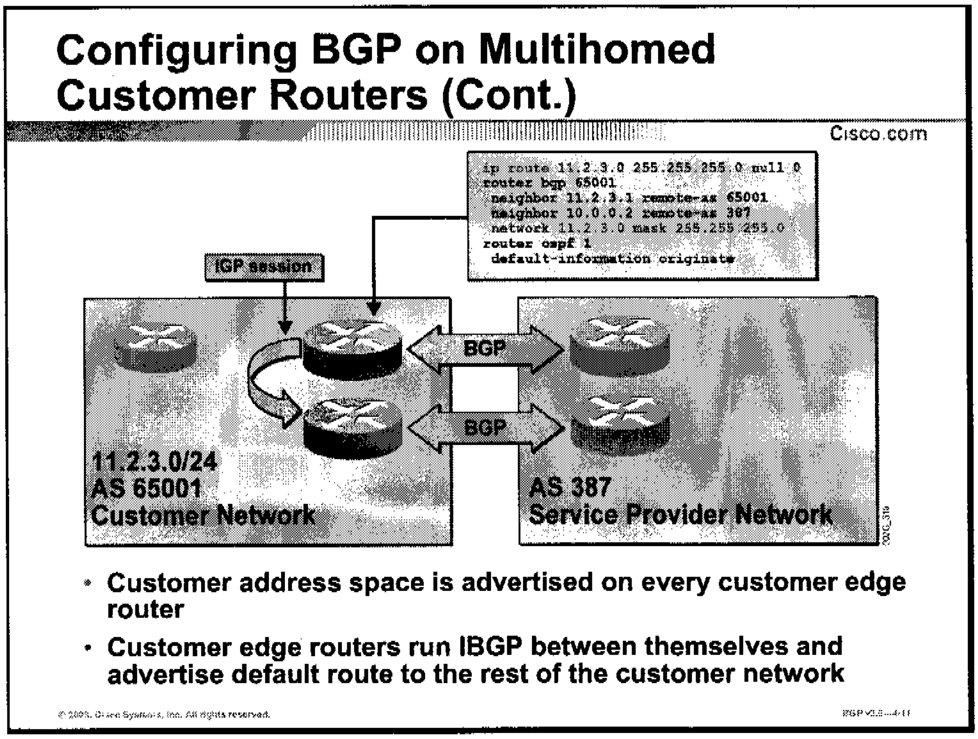
Selecting BGP as the routing protocol between the customer and provider network ensures that a link failure or the failure of a remote router is detected. In this scenario, the customer does not require the use of a public AS number or full Internet routing. Instead, a private AS number is assigned to the customer network, and the Internet Service Provider (ISP) sends a default route to the customer through BGP.

The big difference in this case as compared to a network scenario where static routes and redistribution are used, is that routers within the private AS of the customer now advertise the customer routes via BGP. Thus, the customer is responsible for announcing its own address space. The ISP receives routes from the customer and conditionally propagates them (similar to static routing). If the customer uses provider–assigned (PA) address space, and the ISP can summarize the address space, it will not propagate the explicit routes from the customer to the Internet. The private AS number in the AS-path attribute must first be removed before the ISP can propagate any of the customer routes.

Because the customer is now creating BGP routes that are received by the ISP, any error made by the customer can influence routing operation within the ISP network and, if propagated, within the Internet as whole. Announcing a route to a network, which the customer has not been assigned, may induce routing problems. There is always a risk that such routing problems can occur in a service provider network. However, the risk is much greater when the customer, whose network administrators usually have less experience with BGP, enters the configuration.

To reduce the risk of erroneous route advertising, the ISP should always filter any BGP information received from the customer network. The ISP should reject routes to networks that are not expected to be in the customer AS. Routes that contain an AS path with unexpected AS numbers should also be rejected.

Example



In this figure, the customer has been assigned the private AS number 65001. Both customer edge routers are configured to run BGP and should advertise all of the customer networks with the **network** command. If only one router advertises the network, a single point of failure has been introduced. The two customer edge routers must also run Internal Border Gateway Protocol (IBGP) between them to make common decisions regarding BGP routing information.

Each customer edge router has an External Border Gateway Protocol (EBGP) session with the ISP router on the other side of the link. Over that EBGP session, the ISP announces only a default route to the customer AS. When EBGP receives the default route, it installs it in the forwarding table and redistributes it into the Interior Gateway Protocol (IGP) (in this case Open Shortest Path First [OSPF]) of the customer.

Practice

Q1) Why can you not use static routing in all cases of redundant links between a customer and a single ISP?

A) You cannot use static routing when multihoming to a single provider.

B) You cannot use static routing when load balancing is a design requirement.

C) You cannot use static routing when the ISP conditionally advertises customer routes.

D) You cannot use static routing in scenarios where the failures cannot be detected by Layer 2 protocols.

Conditional Advertising in Multihomed Customer Networks

This topic describes how to configure conditional advertising of a customer address space when you are using BGP.

Conditional Advertising in Multihomed Customer Networks

Cisco.com

- **Customer edge routers should announce the whole customer address space into BGP**
- **The static route covering the whole customer address should point to the core of the customer network, not to null 0**
- **Customer edge router revokes the BGP announcement of customer address space if the edge router loses connectivity with the customer core**

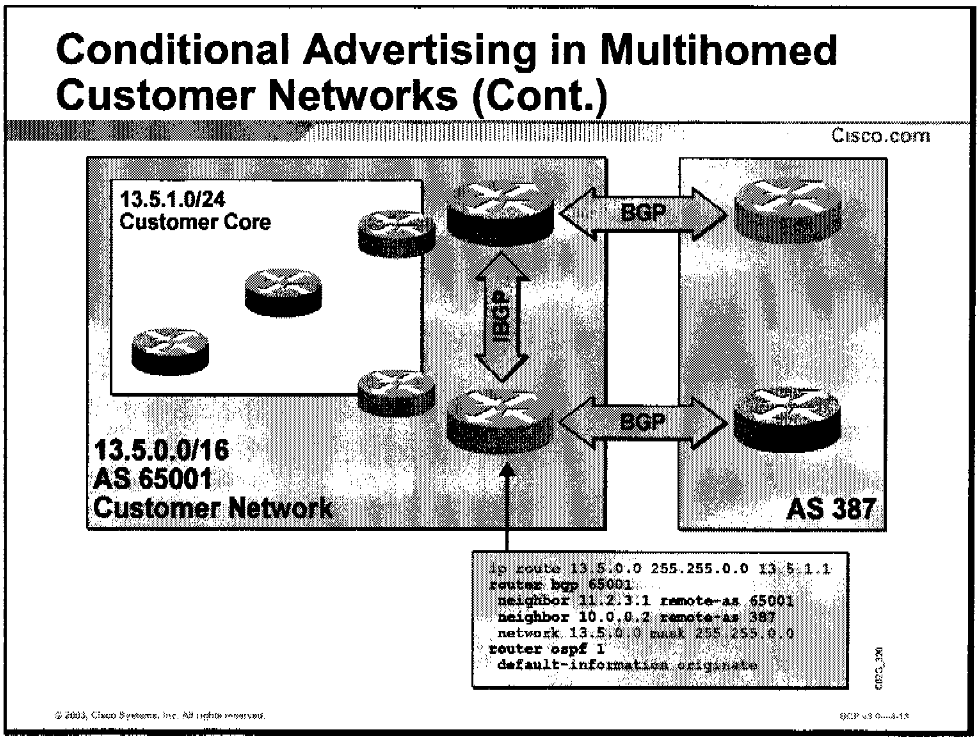
 BGP v3.0—4-12

As a rule of thumb, the customer should announce addresses as large as possible (the larger the address space that can be aggregated, the better). The BGP advertisement is configured on the customer edge routers using the **network** command. Route advertisement is conditioned by the appearance of a corresponding network or subnet in the forwarding table of the edge router. If the network or subnet is manually entered into the forwarding table by configuring a static route to "null 0," the condition is always true because the static route is always there, and the BGP advertisement is always performed.

If the customer edge router loses connectivity to the rest of the customer network but is still connected to the ISP network, BGP advertisement must cease. In this case, BGP advertisement can be stopped if BGP advertisements are bound to the reachability status of a specific subnet in the core of the customer network, according to the customer IGP.

The problem with using a static route to null0 is that it conditions the network statement in the BGP configuration so that BGP always advertises the route. If the customer edge router loses connectivity with the rest of the customer network, the router continues to advertise the entire customer address space. The ISP network receives a valid route from the customer edge router. Traffic is sent to this router, but because it has lost connectivity with the rest of the network, the traffic is dropped (routed to the null0 interface using the static route).

Example



In this example, the customer network uses the address space 13.5.0.0/16. The address space is further subnetted at the customer site. One of the subnets (subnet 13.5.1.0/24) is identified as being a central part of the customer core network.

The customer edge routers participate in the IGP routing of the customer. This participation means that these routers have information about which of the subnets within the address space 13.5.0.0/16 are currently reachable. If these subnets are available, there is an explicit route to each of them. If any of the subnets go down, or the path toward them goes down, the route to that subnet is removed from the forwarding table.

The BGP advertisement in each of the customer edge routers is configured to advertise the full address space used by the customer. When this route is advertised by the customer edge routers, the ISP network, and thus the rest of the Internet, see the complete address space of the customer as one single route, 13.5.0.0/16.

Advertisement of the customer address space by BGP is conditioned by the appearance of the static route, **ip route 13.5.0.0 255.255.0.0 13.5.1.1**. If the static route is valid, then the BGP route 13.5.0.0/16 is advertised. The static route is a recursive route, which means that the router takes another look in the routing table for the address 13.5.1.1 before determining what to do with the static route. The idea is that 13.5.1.1 is reachable via the IGP. The subnet 13.5.1.0/24 is announced by the IGP. If this subnet is reachable by the edge router, then the static route to 13.5.0.0/16 is valid. If there is no route to 13.5.1.1, then the static route is invalid.

Note	The condition, whether or not to advertise the entire customer address space 13.5.0.0/16, is controlled by the IGP reachability of a single subnet, 13.5.1.0/24.

The IGP configuration also includes origination of the default route by both edge routers.

Practice

Q1) Why should you not configure conditional route advertising using a route to null 0?

A) Because the route to null 0 will cause all traffic to de discarded.

B) A route to null 0 will condition BGP to advertise routes only when the interface with a matching assigned network number is in the up state.

C) A route to null 0 will condition BGP such that the network will always be advertised regardless of its state.

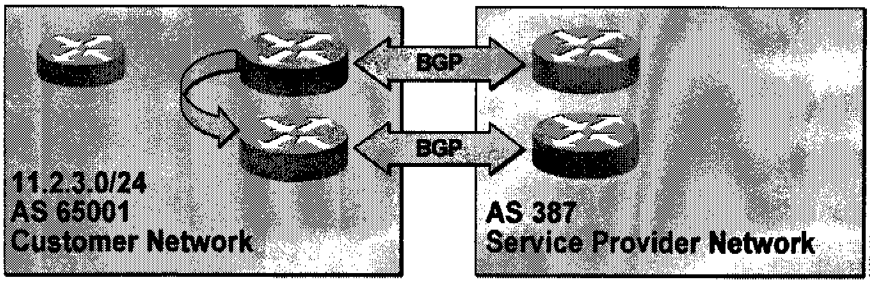
D) Using null 0 as a conditional advertisement will advertise the network only if a matching route exists in the IP routing table.

Configuring BGP on Service Provider Routers

This topic describes how to configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider.

Configuring BGP on Service Provider Routers

Cisco.com



The service provider must:

- **Advertise a default route to the customer through BGP**
- **Filter incoming BGP updates with a prefix-list to verify that the customer announces only the assigned address space**
- **Filter incoming BGP updates with an AS-path filter-list to verify that the customer uses only its own AS number**
- **Optionally, no-export community should be set on customer routes**

 BGP v3.0—1-14

In the ISP network, the two edge routers must have BGP sessions configured for the customer. There is no point in feeding the full Internet routing table to the customer, because the table contains the same set of routes for both links and the customer always uses the ISP for all traffic toward the Internet. A default route injected in the customer network would accomplish the same task.

The customer is responsible for its own advertisements. Assuming that the customers are much less experienced in BGP configuration than the ISP, they are more likely to make errors. Therefore, the ISP must protect itself and the rest of the Internet from those errors.

The service provider should use a prefix-list that allows only customer-assigned routes and denies any other route to ensure that private address space or any other illegal networks erroneously announced by the customer never reach the ISP BGP table. Filtering based on the AS path also provides some protection from customer configuration errors. Only routes originated within the customer AS are allowed. A filter-list performs this check.

If the customer address space is PA address space, and it represents only a small part of a larger block announced by the ISP, the explicit BGP routes received from the customer need not be advertised to the rest of the Internet. The ISP can announce the big block, attracting any traffic toward any subnet within the block. After the traffic enters the ISP network, the more explicit routes to the customer network are available and used. In this case, the provider edge router can tag the BGP routes received from the customer with the **no-export** well-known community, restricting them from being sent by the ISP to any other AS.

Configuring BGP on Service Provider Routers (Cont.)

Cisco.com

```
router(config-router)#
neighbor ip-address default-information
```

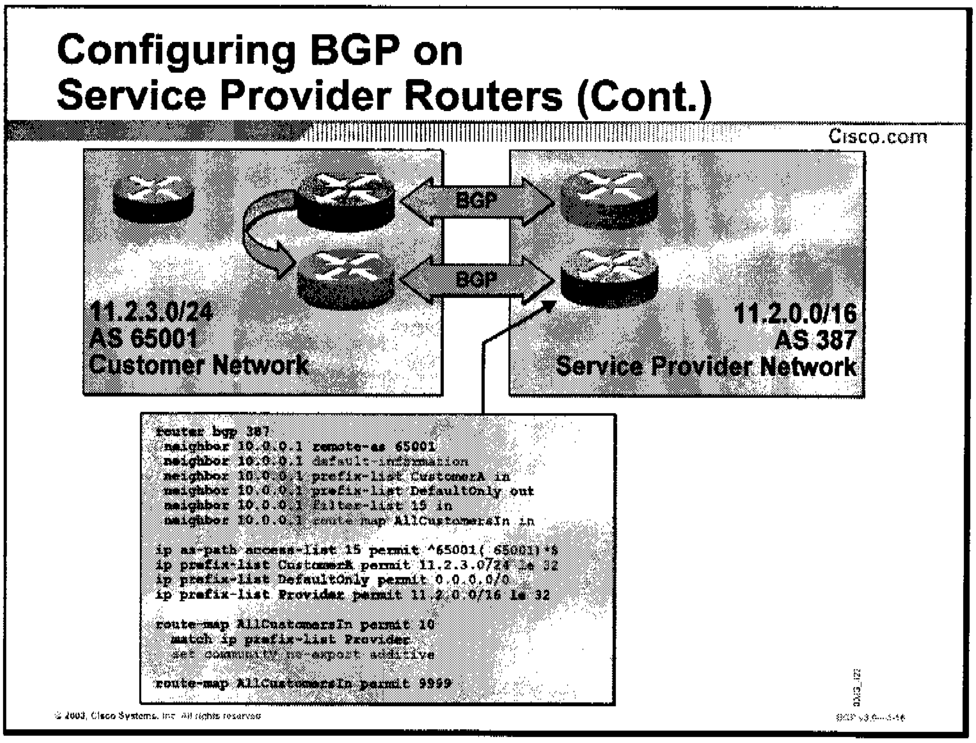
- By default, the default route (0.0.0.0/0) is not advertised in outgoing BGP updates
- The neighbor default-information command advertises default route to a BGP neighbor even if the default route is not present in the BGP table
- Note: the default route is not sent through the outbound BGP filters (prefix-list, filter-list, or route-map)

 BGP v3.0—4-15

The default route, 0.0.0.0/0, is not advertised in outgoing BGP updates unless it is explicitly configured. The **neighbor default-information** router configuration command is used to initiate the advertisement of the default route to a neighbor.

No checking is done by BGP before the default route is advertised. The default route does not need to be present in the BGP table before it is advertised using this command. The default route is also sent without being filtered by any outgoing prefix-lists, filter-lists, or route-maps.

Example



This example shows the configuration of an ISP edge router.

The customer is assigned the private AS number 65001. The BGP session is opened with the customer IP address, 10.0.0.1.

The ISP sends the default route only to the customer. This route is configured first using first the **default-information** command and then the prefix-list *Default-only*.

Received routes from the customer must first pass the prefix-list *CustomerA*. There is one dedicated prefix-list for each individual customer, permitting only those routes that the customer is allowed to announce. If the routes are allowed by the prefix-list, they must also pass the filter-list *15 in*. In this case, the filter allows the private AS of the customer in any number of repetitions, as long as it is the only AS number in the path. This filter-list allows for AS-path prepending configurations on the customer side. If the received route is allowed by both the prefix-list and the filter-list, then the route-map *AllCustomersIn* is applied.

The route-map is a general route-map that is used for all customers. This map checks every route received, via the prefix-list *Provider*, and if the route is within the big block of the PA address space, which the ISP announces to the rest of the Internet, the customer route is marked with the **no-export** community. This mark means that the route is used within the ISP AS only, and is not sent to the rest of the Internet.

Routes that are received from the customer, and are allowed by the prefix-list and filter-list, but do not fall within the PA address space, are allowed by the route-map and are not changed in any way. The ISP propagates these routes to the rest of the Internet.

Practice

Q1) Functionally, what three requirements must you take into consideration when configuring in the service provider network to support a multihomed customer? (Choose three.)

A) The provider should filter incoming BGP updates with a prefix-list to verify customer-announced assigned space.

B) The provider should filter incoming BGP updates with an AS-path filter-list to verify that the customer uses its own AS number only.

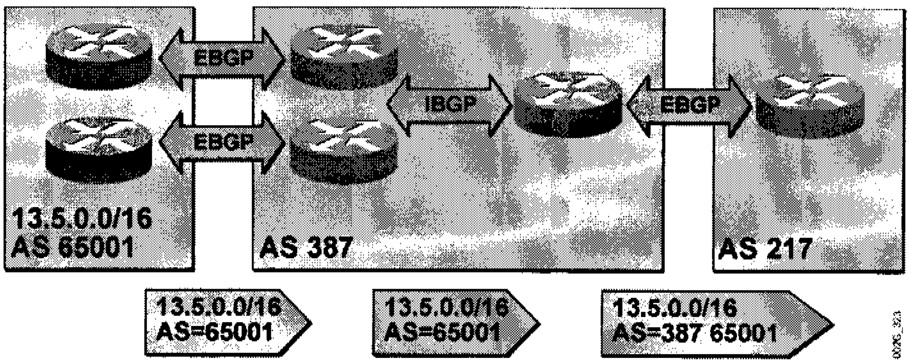
C) The provider should advertise a default route to the customer through BGP.

D) The provider should announce nonsummarized prefixes for the customer networks to the Internet.

Removing Private AS Numbers

This topic describes how to disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path.

Private AS Number Propagation



- Private AS numbers should not be advertised into the Internet
- The private AS numbers must be removed from the AS path before the customer BGP routes are advertised to other service providers

Routes received by the ISP from the customer are propagated to the rest of the Internet only if they are part of the provider-independent (PI) address space.

When the ISP receives BGP routes from the customer, the AS-path attribute of the received routes contains only the AS number of the customer. If the customer uses AS-path prepending, there may be several repetitions of the customer AS number in the AS-path. If customer routes are propagated by the service provider to the Internet, the AS number of the customer will be present in the AS path unless explicitly removed.

Note If the customer has been assigned a private AS number, this AS number must never be advertised by any router to the rest of the Internet.

Removal of a private AS number from the AS path is accomplished by using **remove-private-as** on the ISP EBGP sessions with the rest of the Internet. In the figure, removal of the private AS number takes place on the EBGP session between AS 387 and AS 217.

Removing Private AS Numbers

Cisco.com

```
router(config-router)#
neighbor ip-address remove-private-as
```

- **The command modifies AS-path processing on outgoing updates sent to specified neighbor**
- **Private AS numbers are removed from the tail of the AS path before the update is sent**
- **Private AS numbers followed by a public AS number are not removed**
- **AS number of the sender is prepended to the AS path after this operation**

 BGP v3.0—4-13

neighbor remove-private-as

To remove private AS numbers from the AS path (a list of AS numbers that a route passes through to reach a BGP peer) in outbound routing updates, use the **neighbor remove-private-as** router configuration command.

neighbor {*ip-address* | *peer-group-name*} **remove-private-as**

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

no neighbor {*ip-address* | *peer-group-name*} **remove-private-as**

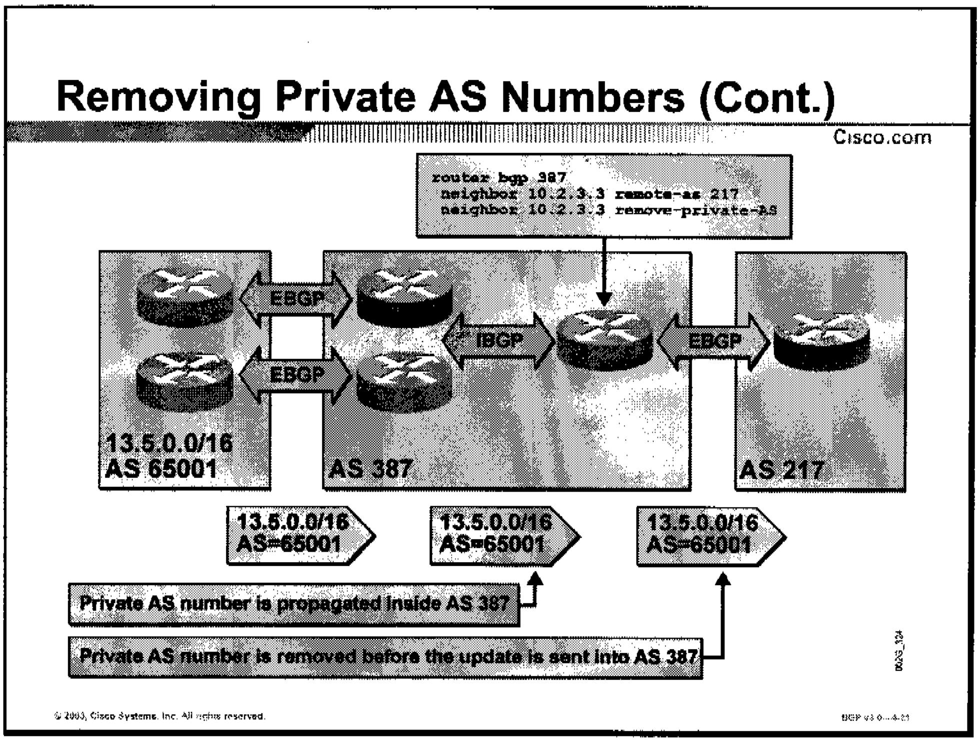
Syntax Description

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

Use this command on the service provider egress routers. Before any of the customer routes of the ISP are advertised by the service provider to the rest of the Internet, the AS numbers in the range 64512-65535 must be removed. The command removes those AS numbers if they are in the tail end of the AS path. Private AS numbers followed by public AS numbers are not removed.

The AS number of the ISP is automatically prepended to the AS-path attribute after the **remove-private-as** operation has completed. This situation means that the AS number of the ISP has not already been prepended to the AS-path attribute when the tail of the AS path is checked for private AS numbers.

Example



In this example, the service provider AS (387) receives routes from the customer. The customer is assigned the private AS number 65001 by the ISP. Therefore, routes received by the provider have an AS path containing only AS 65001. This information should be kept and used within the ISP network and should never be propagated to the rest of the Internet (AS 217 in this example).

The edge router in AS 387 has been configured to remove private AS numbers on EBGP routes toward AS 217. If private AS numbers appear in the tail end of the AS path (before AS 387 is added), they are removed.

This configuration must be applied to all egress router in AS 387 that contain EBGP neighbors leading to other ISPs. No private AS number may be present in an AS path of a route propagated to a network using a public AS number.

Practice

Q1) AS 64525 is connected to AS 229, which in turn is connected to AS 1126. How will configuring private AS removal on AS 229 affect routers in AS 1126?

A) There will be no effect on the routers in AS 1126.

B) The routers in AS 1126 will see all Internet routes as originating in AS 64525.

C) The routers will see networks originating in AS 64525 as originating in AS 229.

D) AS 1126 will receive all of its routes with an AS-path length of two.

Backup Solutions with BGP

This topic describes how to configure a typical backup setup between a multihomed customer and a service provider in a BGP environment.

Backup Solutions with BGP

Cisco.com

The route selection is controlled entirely by the customer routers:

- **Local preference is used to differentiate primary and backup links for the outgoing traffic**
- **Multi-exit discriminator (MED) is used to differentiate primary and backup links for the return traffic**
- **No service provider configuration is required**

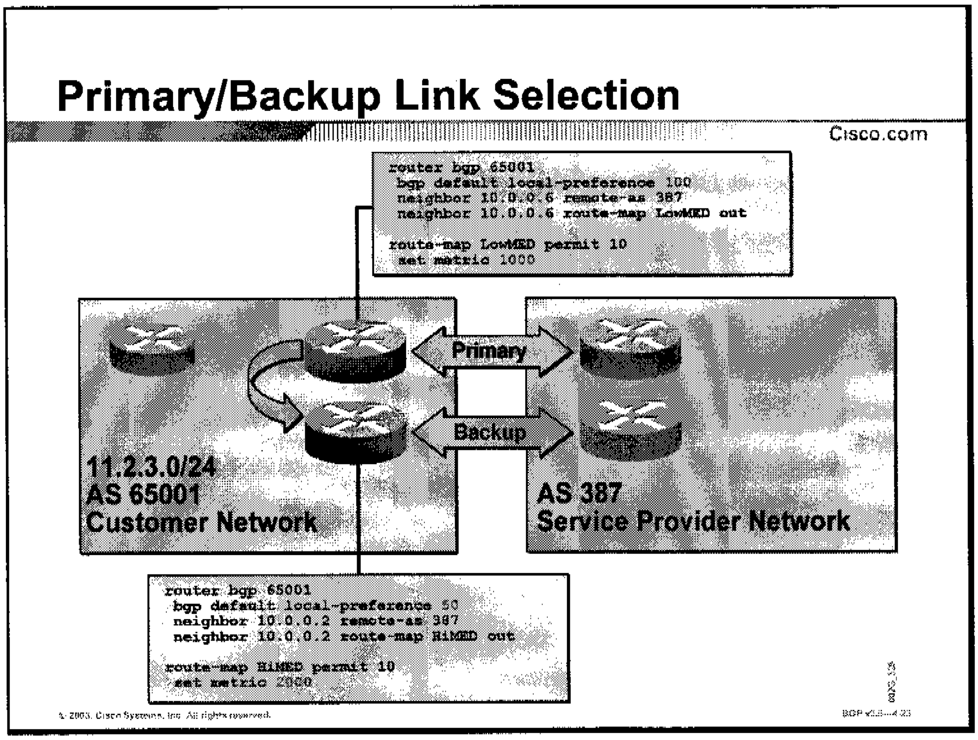
When a customer uses BGP on multiple links between its network and the ISP network, the customer is solely responsible for controlling how it uses the links. The customer can choose to use its links in a primary/backup scenario or in a load-sharing scenario.

If one link is primary, then the other should be used for backup only. The customer can use the local preference configuration to direct all outgoing traffic over the primary link.

Incoming traffic to the customer is controlled using either AS-path prepending or multi-exit discriminator (MED). Because the customer has multiple connections to the same AS, MED is the ideal attribute to use. When the customer announces its routes to the ISP, a bad (high) MED value on the backup link and a good (low) value on the primary link are set.

MED (and AS-path length) is checked by the receiving EBGP peer only if the weight and local preference attributes have not been configured. In this case, the ISP should not use any of these configuration options. The ISP should rely solely on the attributes received from the customer.

Example



In the figure, the customer is connected to the ISP over two permanent connections. The customer uses the upper connection as the primary connection and the lower connection as the backup.

The BGP configuration on the ISP side is transparent. This transparency means that no particular preference is configured to use the upper or lower connection. The ISP relies on the attribute values received from the customer.

The primary edge router on the customer side is configured to set local preference to the value 100 on all EBGP routes received. The backup edge router sets the local preference attribute to a value of 50. This configuration means that the outgoing traffic toward any destination announced by the ISP is primarily sent over the upper link.

Incoming traffic to the customer is directed to the primary link using the MED. In the primary edge router of the customer, all routes that are sent to the ISP have their MED attribute set to the value 100 by the route-map *LowMED out.* In the backup edge router of the customer, all routes that are sent to the ISP have their MED attribute set to the value 2000 by the route-map *HiMED out.* Because the ISP receives the routes with all other attributes set to the same values, the MED values direct traffic, destined to the customer, to the primary link.

Practice

Q1) Which attribute can you use to select the primary/backup link for outgoing traffic?

A) weight

B) local preference

C) AS-path

D) MED

Q2) Which attribute can you use to select the primary/backup link for incoming traffic?

A) weight

B) local preference

C) AS-path

D) MED

Load Sharing with the Multihomed Customer

This topic describes how you can implement load sharing between a multihomed customer and a service provider in a BGP environment.

Load Sharing with the Multihomed Customer

Cisco.com

Load sharing of outgoing customer traffic is identical to the static routing scenario

You can implement load sharing of return traffic in a number of ways:

- **Announce portions of the customer address space to each upstream router**
- **Configure BGP multipath support in the service provider network**
- **Use EBGP multihop in environments where parallel links run between a pair of routers**

 BGP v3.0—4-24

Load sharing of outgoing traffic from the customer network is identical to the static routing scenario. The customer IGP is configured to send information about a gateway of last resort. There is no difference whether the edge router gets its default by static routing or by incoming EBGP updates.

Load sharing of the return traffic coming back to the customer network from the ISP can be implemented in a number of ways:

- The customer could divide its address space into several announcements. The customer edge router can send each announcement over one of its EBGP sessions with the ISP. For backup purposes, the customer should advertise the entire address space over all of its EBGP sessions. The ISP now uses the most explicit route rule, and as long as both links are up, traffic with destinations within one part of the customer address space is routed over one of the links and traffic to the other part is routed over the other link.

- If the customer announces equivalent routes over both links, the ISP routers use the closest connection with respect to the IGP of the ISP. If an ISP router has an equivalent distance to both connection points, the use of the maximum-paths (BGP multipath) option causes load sharing.

- If the multiple links between the customer and the ISP network terminate in one single router on the customer side and one single router on the ISP side, the two routers must establish their EBGP session from loopback interface to loopback interface. Static or dynamic routing is required for one router to get information on how to reach the loopback interface of the other router. The use of the **ebgp-multihop** option is also required because the address of the neighbor is not directly connected.

Practice

Q1) What three options can you use to enable load sharing on parallel links connected to one router? (Choose three.)

A) Split the customer address space into two parts and advertise a portion on each link.

B) Use AS-path prepending on the outgoing routes of the backup path.

C) Use the **ebgp-multihop** option between loopback interfaces of the multihomed routers.

D) Enable BGP multipath support on the multihomed routers.

Load Sharing with BGP Multipath

This topic presents the Cisco IOS command required to configure load sharing between a multihomed customer and a service provider through the use of BGP multipath.

Configuring BGP Multipath Support

Cisco.com

```
router(config-router)#
maximum-paths number
```

- **By default, BGP selects a single path as the best path and installs it in the IP routing table**
- **With maximum-paths configured, a BGP router can select several identical EBGP routes as the best routes and install them in the IP routing table for load-sharing purposes**
- **The BGP router can install up to six BGP routes in the IP routing table**

By default, BGP route selection rules select one, and only one, route as the best. If there are two identical routes, the tiebreaker is the most stable route or the router-ID of the peer router advertising the route. However, by using the **maximum-paths** router configuration command, the BGP route selection process will select more than one route as best if they are exactly identical. The routes are all installed in the forwarding table, and load sharing will take place.

maximum-paths

To control the maximum number of parallel routes that an IP routing protocol can support, use the **maximum-paths** command in address family or router configuration mode.

maximum-paths *number*

To restore the default value, use the **no** form of this command

no maximum-paths

Syntax Description

number	Maximum number of parallel routes that an IP routing protocol installs in a routing table, in the range of 1 to 6

Note	Load sharing between alternative BGP routes is achieved only if the EBGP routes are identical according to all BGP route selection rules and **maximum-paths** is configured with a value larger than 1.

A BGP router can install up to six BGP routes in the IP forwarding table. The actual type (per-session or per-packet) of load sharing done between the routes depends on the switching mode used.

Practice

Q1) By default, BGP can perform load balancing over how many parallel links?

A) one

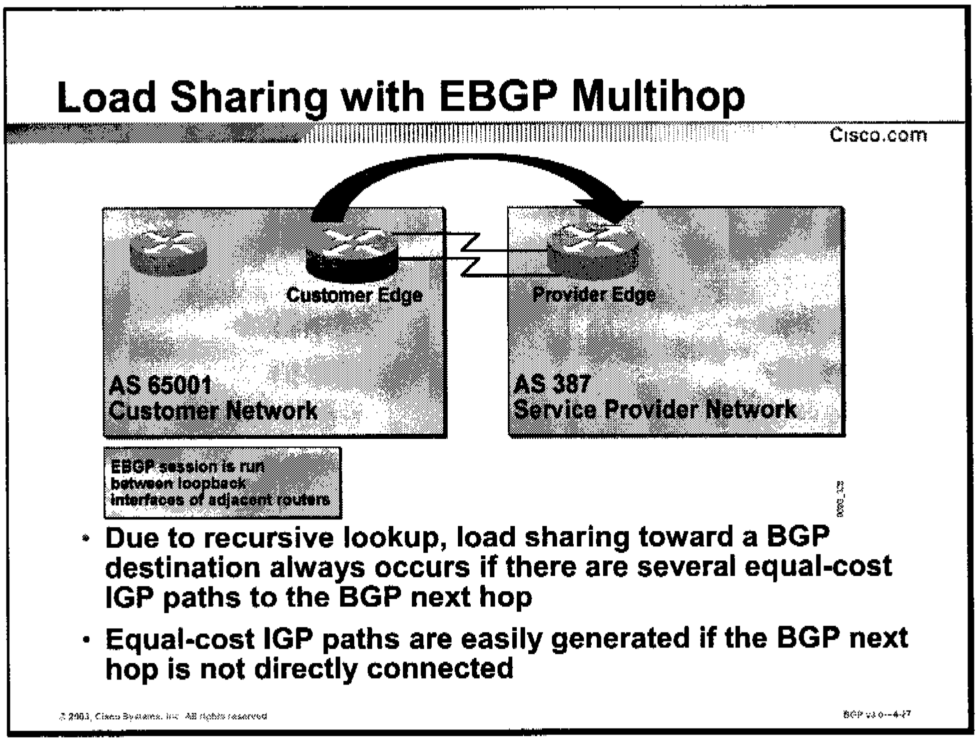
B) four

C) six

D) eight

Load Sharing with EBGP Multihop

This topic describes how to configure load sharing between a multihomed customer and a service provider through the use of EBGP multihop.



When two adjacent routers have multiple links between them, you can configure the EBGP session from loopback interface to loopback interface. In this case, you must use **ebgp-multihop** to make the BGP session go into the active state. There must be static or dynamic routing in use to provide both routers with information on how to reach the loopback interfaces of each other. Otherwise, their EBGP session does not complete establishment.

Routing to the loopback interface of the neighboring router is required to establish the EBGP session and is also used in the recursive lookup when the routes are installed by the router in its forwarding table. The two routes to the loopback interface of the neighboring router should be equivalent for load sharing to occur.

After configuration, one single EBGP session is established between the two routers. This session is used to exchange the routing information. There is only one BGP route to each destination, and it has a next hop referring to the loopback interface of the other router.

Before installing a route to a specific destination in its forwarding table, a router will perform a recursive lookup to resolve the next hop. In this case, the recursive lookup will result in finding two alternative routes. The router will install the BGP route to the final destination twice in the forwarding table (Forwarding Information Base [FIB]). The first time, the route is installed with one of the resolved next-hop addresses, and the second time with the other resolved next-hop address. Since multiple equal-cost paths exist, the router can load share over the two paths depending on the switching mode.

Configuring Multihop EBGP Sessions

Cisco.com

```
router(config-router)#
neighbor ip-address ebgp-multihop [ TTL ]
```

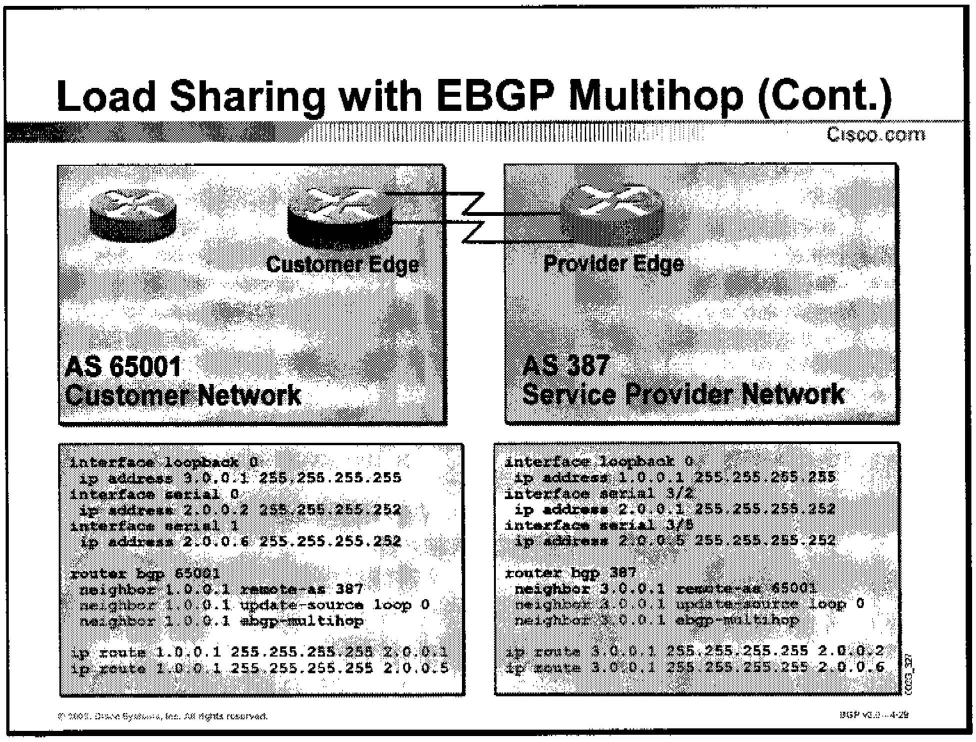
- **By default, EBGP neighbors must be directly connected**
- **The ebgp-multihop command declares an EBGP neighbor to be distant (several hops away)**
- **Number of hops can be specified in the *TTL* parameter**
- **Usually used to run EBGP between loopback interfaces for dial backup or load-sharing purposes**
- Use with extreme caution; routing loops can occur very easily

By default, EBGP neighbors must be directly connected. Cisco IOS software verifies that an EBGP neighbor is reachable as directly connected over one of the router interfaces before the session goes into the active state. For an EBGP session, IP packets carrying the TCP segments with BGP information are also sent using a Time to Live (TTL) value set to the value 1. This value means that they cannot be routed.

The **ebgp-multihop** neighbor configuration command changes this behavior. Although the neighbor is several hops away, the session goes into the active state, and packets start to be exchanged. The TTL value of the IP packets is set to a value larger than 1.If no value is specified on the command line, 255 is used.

Use the **ebgp-multihop** command when establishing EBGP sessions between loopback interfaces for load-sharing purposes. You must take great care when using **ebgp-multihop**, because proper packet forwarding relies on all the intermediate routers along the path to the EBGP peer to make the correct forwarding decision. If the intermediate routers have a correct path to the EBGP peer, but a wrong path to the final destination, the packet may get into a routing loop.

Example



In the figure, the customer network and the ISP network are connected using two parallel links between a single router on the customer side and a single router on the ISP side.

In this case, only one EBGP session is configured between the customer and provider routers. The session should be established from the loopback interface in one router to the loopback interface in the other.

Each of the two edge routers has two static host routes pointing to the loopback interface on the other router. The EBGP session is established from loopback to loopback using **ebgp-multihop**.

The customer receives an EBGP route from the ISP with the next hop set to 1.0.0.1. The customer edge router performs a recursive lookup and finds that it can reach 1.0.0.1 via 2.0.0.1 and via 2.0.0.5. These two routes are equivalent. Therefore, the route to the final destination is installed in the forwarding table of the customer router using both paths.

Depending on the switching mode in use, load sharing is done per packet, per destination, or per source/destination pair.

In this example, link-level procedures ensure that if one of the links goes down, the corresponding static link goes down. All BGP routes in the forwarding table that rely on the static route to the link that went down are invalidated. However, the BGP routes in the forwarding table relying on the remaining link are still valid and used.

Practice

Q1) What three criteria must be met before two routers with parallel links between them can perform load balancing using **ebgp-multihop**? (Choose three.)

A) The **neighbor ebgp-multihop** command must be configured on each router.

B) BGP maximum paths must be set to the number of links between the routers.

C) The routers must have a static route or an IGP containing reachability information for the configured loopback addresses.

D) The **neighbor update-source** option must be configured on both routers.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- When a customer has multiple connections to a single ISP and the link-level procedures cannot detect a link failure, a routing protocol is required. For security reasons, this routing protocol must be BGP.
- The AS number used by the customer does not have to be a public AS number; it can be a private AS number in the range 64512 – 65535.
- When conditionally advertising customer networks to the service provider, you should use a static route covering the whole customer address space pointing to the core of the customer network instead of null 0.
- The service provider should advertise a default route to the customer through BGP. Incoming filters should also be used by the provider to ensure that only the correct address space and AS number are advertised by the customer.

BGP v3.0—4-30

Summary (Cont.)

Cisco.com

- Private AS numbers must never be propagated to the rest of the Internet. The ISP must therefore remove the private AS numbers from the AS path before sending them to another public AS.
- You can use parallel links between the cutomer network and the network of a single ISP for backup or load-sharing purposes. The customer controls the outgoing load using local preference. The customer can control the incoming load using the MED (metric) attribute. With the MED, the links go to a single remote AS.
- By announcing portions of its address space, a customer can use maximum paths and EBGP multihop to provide load sharing over multiple links.
- EBGP multihop can be used for load balancing only if redundant links terminate on the same provider router.

BGP v3.0—4-31

Next Steps

After completing this lesson, go to:

- Connecting a Multihomed Customer to Multiple Service Providers lesson

References

For additional information, refer to these resources:

- For more information on removing private AS numbers, refer to "Removing Private Autonomous System Numbers in BGP" at the following URL: http://www.cisco.com/warp/public/459/32.html

- For more information on using MED for path selection, refer to "How BGP Routers Use the Multi-Exit Discriminator for Best Path Selection" at the following URL: http://www.cisco.com/warp/public/459/37.html

- For more information on load sharing, refer to "Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments" at the following URL: http://www.cisco.com/warp/public/459/40.html

Quiz: Connecting a Multihomed Customer to a Single Service Provider

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Configure BGP on a customer network to establish routing between a multihomed customer and a single service provider
- Configure conditional advertising of a customer address space when BGP is used to establish routing between a multihomed customer and a single service provider
- Configure BGP on a service provider network to establish routing between a multihomed customer and a single service provider
- Disable the propagation of private AS numbers to EBGP peers in a service provider network where a multihomed customer is advertising private numbers in the AS path
- Configure a typical backup setup between a multihomed customer and a single service provider in a BGP environment
- Describe how you can implement load sharing between a multihomed customer and a single service provider in a BGP environment
- Identify the Cisco IOS command required to configure load sharing between a multihomed customer and a single service provider using BGP multipath
- Configure load sharing between a multihomed customer and a single service provider using EBGP multihop

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are three responsibilities of the customer when the customer is multihomed to a single service provider? (Choose three.)

A) Customer edge routers must run IBGP between them.

B) The customer must advertise a default route.

C) The customer must conditionally advertise its assigned address space into BGP.

D) The customer edge routers must run EBGP with the provider.

Q2) Given the following router command output, what method has been used to influence return traffic in a primary/backup link implementation for this multihomed customer?

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
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
*  10.10.20.0/24    192.168.63.3          1000             0 100 100
i
*>                  192.168.64.4          2000             0 400 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 100 I
*> 40.40.40.0/24    192.168.64.4             0             0 400 I
```

A) MED

B) local preference

C) weight

D) AS-path prepending

Q3) What are three responsibilities of the provider router when supporting a multihomed customer? (Choose three.)

A) The provider must advertise a default route to the customer through BGP.

B) The provider must filter customer routes to verify that proper addressing is used.

C) The provider must remove the private AS number, if in use by the customer.

D) The provider must configure new AS-path filters to allow AS-path prepending; otherwise, a primary/backup link cannot be established.

Q4) What will occur if private AS numbers are advertised to the Internet?

A) The Internet will not be able to route packets.

B) Internet routers could drop routes based on BGP loop prevention mechanisms.

C) Customer load balancing will not function.

D) Customer configurations for the primary/backup link using AS-path prepending will not function.

Q5) What BGP configuration is required to properly implement a backup solution for a multihomed customer connected to a single provider? (Choose two.)

A) The customer should set local preference to influence outgoing route selection.

B) The customer should set the weight attribute to influence outgoing path selection.

C) The customer should set MED on each route to influence return path selection.

D) The customer should configure AS-path prepending to ensure proper outgoing path selection.

Q6) A customer router has been configured with maximum paths set to a value of 4. Given the following router command output, over how many links will the router need to perform load balancing?

```
router# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
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
*  10.10.20.0/24    192.168.63.3                           0 300 100
100 i
*>                  192.168.64.4                           0 400 100
i
*                   192.168.65.5                           0 500 100
i
*> 30.30.30.0/24    192.168.63.3             0             0 300 i
*> 40.40.40.0/24    192.168.64.4             0             0 400 i
```


A) The router will use only the path marked as "best" by BGP.

B) The router will perform load balancing over two paths to reach network 10.10.20.0/24.

C) The router will perform load balancing over three paths to reach network 10.10.20.0/24.

D) There is not enough information to determine the correct answer.

Q7) What three methods can you use to provide load sharing over network links between a multihomed customer and a single provider? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring ebgp-multihop between the customer and the provider

C) use of the BGP **maximum-paths** command to perform load balancing over parallel links

D) configuring multiple static routes pointing to the provider

Q8) Why is it not required to configure maximum paths under the BGP routing process when you are performing load balancing with ebgp-multihop?

A) By default, BGP will perform load balancing over up to four paths, configurable up to six.

B) The static route or IGP process is responsible for load balancing in this configuration.

C) Configuring multihop enables maximum paths equal to the TTL setting of the **neighbor ebgp-multihop** command.

D) Configuring ebgp-multipath is a required component of ebgp-multihop load balancing.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Connecting a Multihomed Customer to Multiple Service Providers

Overview

Customers connect to the Internet using service providers to enable different applications such as intranet connectivity with Virtual Private Networks (VPNs), extranet connectivity with suppliers, and other Internet applications. Different customers have different connectivity requirements depending upon their business model, redundancy requirements, and even network budget.

This lesson discusses using multiple connections between a customer and multiple service providers for backup and load-sharing purposes. Included in this lesson is a discussion of the Border Gateway Protocol (BGP) characteristics used to configure customer and provider networks to accommodate the multiple connections between them. Also discussed in this lesson are topics specific to networks with multiple connections between a customer and multiple providers such as address selection, private autonomous system (AS) number translation, and configuration of the network to support either backup links or load sharing.

Importance

When a customer requires the maximum redundancy in its network design, it should implement a multihomed strategy that uses multiple service providers. This configuration requires specific considerations to be implemented properly. Addressing and AS number selection are important considerations that affect the implementation of the network. It is also important to understand how to configure routing protocols so that customer backup or load-sharing requirements are met.

Objectives

Upon completing this lesson, you will be able to:

- Describe BGP configuration characteristics used to establish routing between a multihomed customer and multiple service providers
- Describe addressing strategies available to a multihomed customer connected to multiple service providers
- Describe AS numbering strategies available to a multihomed customer connected to multiple service providers
- Describe the operation of AS number translation
- Describe how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment
- Describe the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers
- Describe how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module
- Route Selection Using Policy Controls module
- Route Selection Using Attributes module

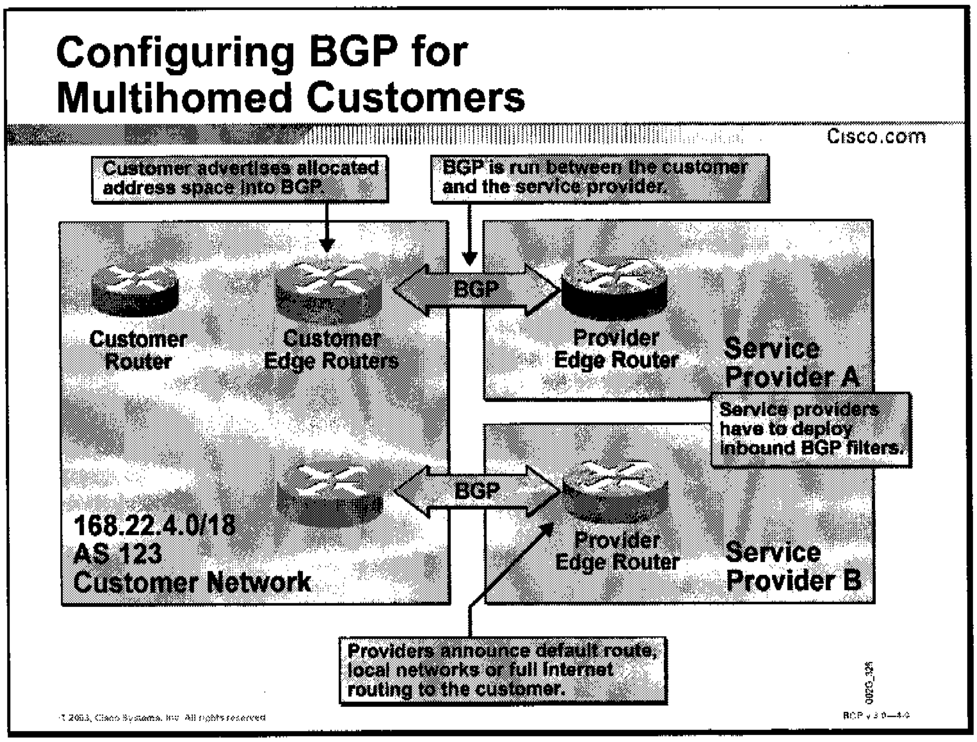
Outline

This lesson includes these topics:

- Overview
- Configuring BGP for Multihomed Customers
- Multihomed Customer Address Space Selection
- Multihomed Customer AS Number Selection
- AS Number Translation
- Primary/Backup Link Selection
- BGP Incoming Link Selection
- Load Sharing with Multiple Providers
- Summary
- Assessment (Quiz): Connecting a Multihomed Customer to Multiple Service Providers

Configuring BGP for Multihomed Customers

This topic describes the different characteristics of a BGP configuration used to establish routing between a multihomed customer and multiple service providers.



The highest level of resilience to network failures is achieved in network designs that connect the customer network to two different service provider networks. Customers use this option when the requirement for resilient Internet connectivity is very high. This requirement also involves duplication of equipment to make the customer network fully redundant.

BGP must be used between the customer and both service providers, because static routing will not work in this type of network. It is not enough to detect link failures or a failure in the remote router by link-level procedures. Failures that occur beyond the directly connected router must also be detected, and the only means of detecting these failures is by using a routing protocol. The only routing protocol suited for the Internet environment is BGP. Correctly configured, BGP takes care of rerouting in the following situations:

- Link failure between the customer network and the network of one of the ISPs.
- Edge router failure on either the customer or the ISP side.
- Link failure or router failure within the customer network causing the customer edge router to lose connectivity with the customer network core. This situation requires correct configuration of route advertisement as described in an earlier lesson.
- Link failure or router failure within the ISP network causing the ISP edge router to lose connectivity with the rest of the Internet.

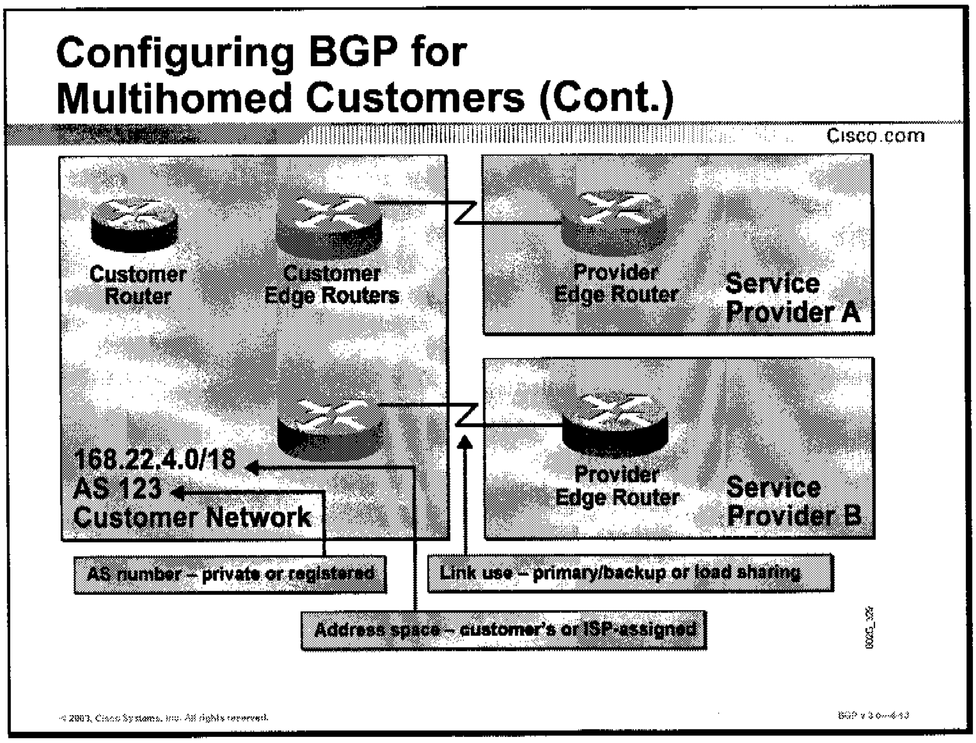
Multihomed customers have multiple permanent links to different ISPs. The links should terminate in different edge routers in the customer network. Otherwise, one of the major advantages, resilience to router failure, is lost.

Multihomed customers should use BGP with both ISPs. The customer should advertise its address space to both providers. Route advertisement should be configured in both customer edge routers. The advertisement should be conditioned with the appropriate route policies by the edge router connectivity to the core of the customer network. This setup is analogous to what is configured when you are connecting a multihomed customer network to a single provider.

The customer should take care not to move any routing information between the two ISPs. It must use outgoing filters to prevent any route received from one of the ISPs to be propagated to the other. Otherwise, the customer network appears as a transit network between the two ISPs.

Both ISP must apply filters on the incoming BGP information from the customer to protect themselves and the rest of the Internet from errors in the BGP configuration of the customer. Each of the service providers must accept routes from the customer that indicate networks within the customer address space only. AS-path filter-lists should be implemented on the provider edge routers to allow incoming routes only if they have the correct AS-path attribute value. If the incoming filters on the ISP edge router accept customer routes, then the service provider should propagate those routes to the rest of the Internet.

Both ISPs must provide the customer with at least some BGP routes. Depending on customer requirements, the volume of BGP routes provided by the ISP could range anywhere from the default route only, to the full Internet routing table.



Before configuring the multihomed network, you need to consider the following questions:

- Should any of the links be used as primary and the others as backup?
- Should both links share the load?
- What address space is the customer using? (Is the customer address space provider-assigned [PA] or provider-independent [PI]?)
- What AS number is the customer using? (Is the customer using a public or a private AS number?)

Practice

Q1) Why should you apply outbound filters to a multihomed customer network?

A) to provide for maximum security at the customer site

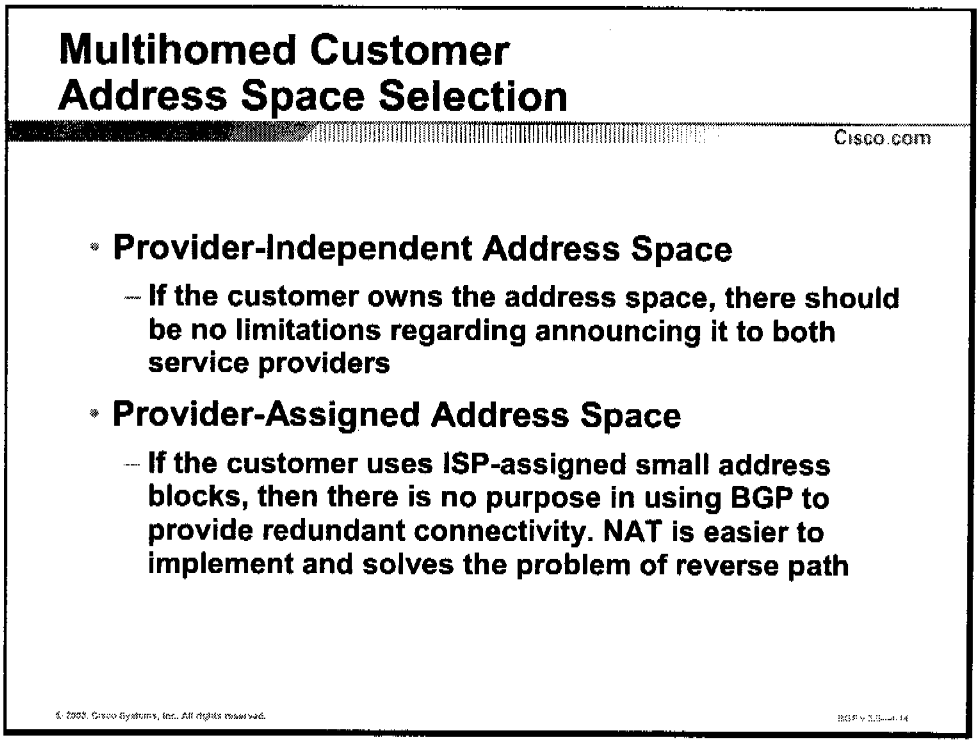
B) to guarantee that return traffic has proper load balancing

C) to prevent the customer network from becoming a transit AS

D) to ensure that outbound traffic has proper load balancing from the customer site

Multihomed Customer Address Space Selection

This topic describes the different addressing strategies available to a multihomed customer connected to multiple service providers.



If the customer has its own address space, it should announce it to both service providers. Both providers are responsible for propagating the customer routes to the rest of the Internet without doing any summarization.

However, if the customer uses a small block of addresses assigned by one of the ISPs, an alternative design, not involving BGP, is to use two different PA address spaces and do Network Address Translation (NAT). With NAT, the router translates traffic going out over one of its connections to one of the PA addresses. If traffic goes out the other way, the addresses are translated to an address from the address space of the other provider.

Practice

Q1) Why is PI address space recommended for customers connected to multiple providers?

A) PI addressing facilitates easier implementation of backup link policies.

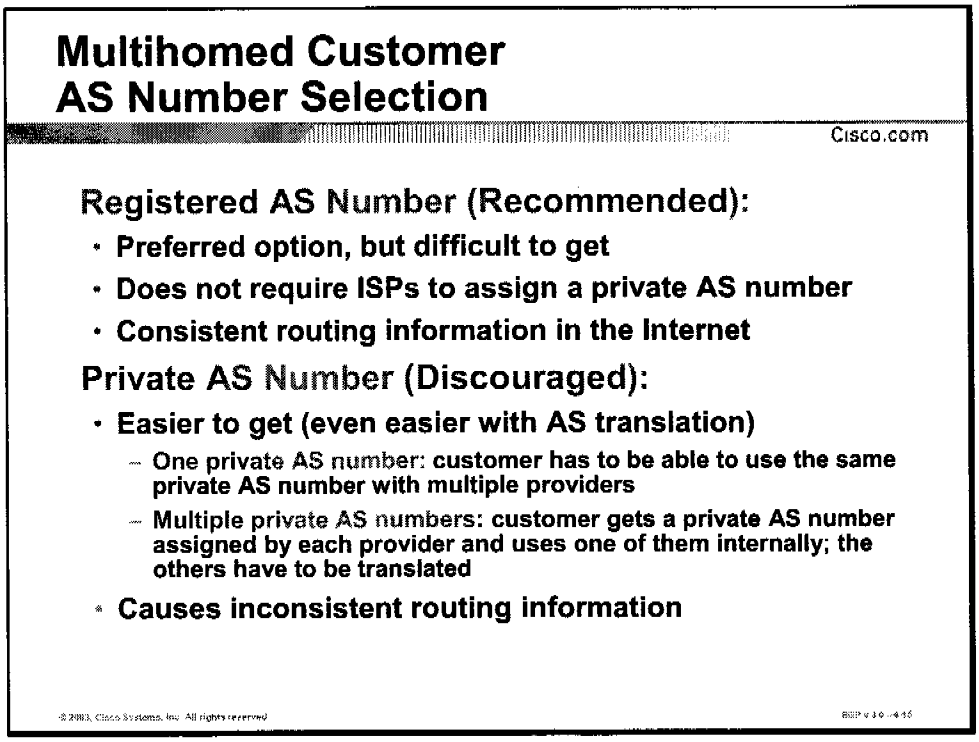
B) PI addressing facilitates easier implementation of load-sharing configurations.

C) PI addressing removes any limitations around advertising the address space.

D) PI addressing enables the use of automatic summarization at upstream ISPs.

Multihomed Customer AS Number Selection

This topic describes the different AS numbering strategies available to a multihomed customer connected to multiple service providers.



The use of BGP requires an AS number. The preferred option is to use a registered, public AS number. However, registered AS numbers are assigned only to those who really need it because public AS numbers are a scarce resource. An ISP with BGP sessions to multiple ISPs must use a registered, public AS number. A customer connected to only one ISP does not require a public AS number. In that case, a private AS number in the range 64512-65535 is sufficient.

Whenever the customer has a public AS number assigned, there are no conflicts in the BGP setup, because the number is guaranteed to be unique within the Internet. Route announcements are made by both the customer and service provider without tampering with the AS path. As a result, consistent AS path information is propagated by the service provider to the rest of the Internet.

In those cases where the customer does not have a public AS number, it must use a private AS number. Because private AS numbers are not propagated to the Internet, several network administrators can, independently of each other, make this assignment. In this case, AS numbers are reused, conserving AS number space. A service provider normally assigns private AS numbers to its customers. This arrangement makes sure that unique private AS numbers are used among the customers of a single ISP.

In the case where a customer is going to be multihomed and the private AS number already assigned from one of the ISPs comes in conflict with AS numbers assigned by the other ISP, renumbering of the customer AS should be considered by the customer. If the two service providers can reach a common agreement on which private AS number the multihomed customer should use, renumbering is a solution. If no common agreement can be made or if renumbering, for some reason, is not an option, AS translation must be configured on the customer network.

No router should ever propagate private AS numbers to the rest of the Internet. An ISP can keep track of which private AS numbers it has assigned to its customers and avoid reuse or conflicts within that scope. However, as soon as the scope is widened to include other ISPs, conflicts will happen. Each ISP, therefore, removes private AS numbers from the AS path before sending routes outside its own AS.

When the routes, with the private AS numbers removed, are propagated to the rest of the Internet, the AS path looks like the routes were originated within the public AS of the ISP. All information about the private AS lying behind the public AS is lost. In the case of a multihomed customer, the customer routes are, in the first step, propagated into each of the autonomous systems of its ISPs. In the next step, the routes have the private AS number removed as the routes are propagated to the rest of the Internet. Now the customer routes appear to be originating in the autonomous systems of both ISPs. To an outside observer, there is now an AS-path inconsistency because it looks like the same route belongs to different autonomous systems.

Practice

Q1) What are three AS number implementation options available to customers connecting to multiple providers? (Choose three.)

A) The customer can obtain a registered, public AS number that is advertised to all upstream providers.

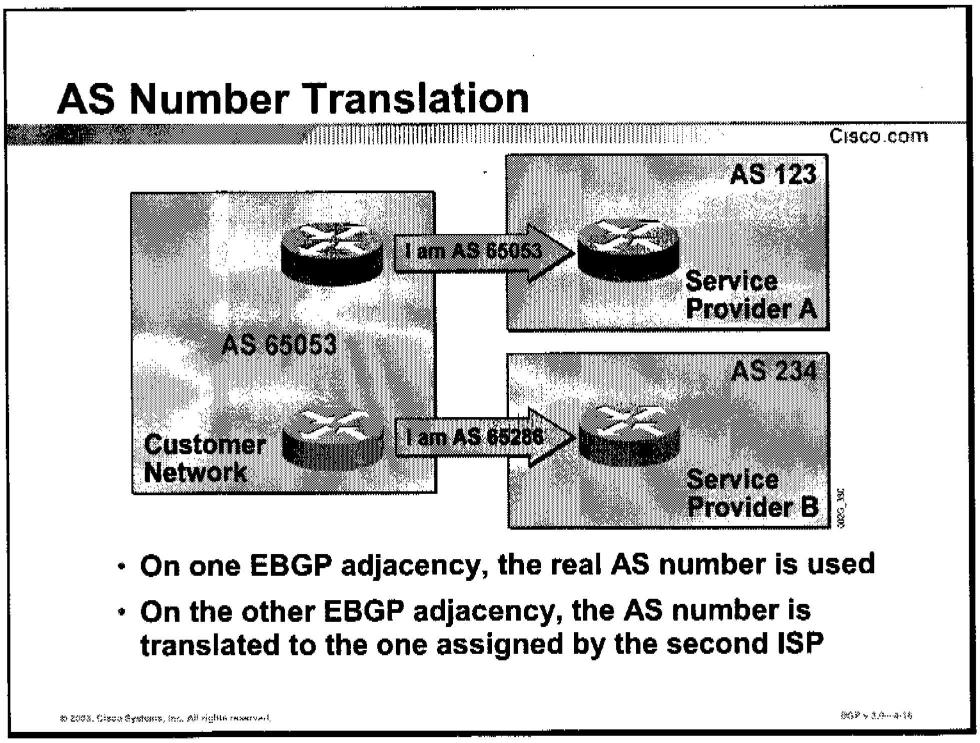
B) The customer can use a single private AS number as long as all upstream providers agree to support the same AS number.

C) The customer can use two different private AS numbers by translating one of the private addresses at the customer edge.

D) The customer can use two different PA AS numbers by configuring EBGP internally at the customer site.

AS Number Translation

This topic describes the operation of AS number translation.



The figure shows a case where a customer is multihomed but forced to use two private AS numbers (for example, because of the scarcity of public AS numbers).

In the figure, service provider A has assigned the private AS number 65053 to the customer. Service provider B did not agree to use this private AS number when connecting to the customer. Instead, service provider B has assigned the private AS number 65286.

The customer now has two different private AS numbers: 65053 and 65286. The customer decides to use 65053 internally. All router BGP configuration lines have 65053 as the AS number. The customer uses AS number 65286 only when establishing the External Border Gateway Protocol (EBGP) session to AS 234.

In the example, service provider A (AS 123) has an EBGP session to the customer where the AS number 65053 is used at the customer end. Service provider B (AS 234) has an EBGP session to the customer where the AS number 65286 is used at the customer end. Translation between these two private AS numbers takes place in the customer edge router as part of the EBGP session to AS 234.

AS Number Translation (Cont.)

Cisco.com

```
router(config-router)#
neighbor ip-address local-as private-as
```

- **Optionally, the customer can get two different private AS numbers assigned by the service providers**
- **Internally, the customer can use an ISP-assigned AS number or even any other private AS number**
- **Externally, the customer is seen as one private AS number to ISP 1, and as a different AS to ISP 2**
- **Note: When you are using this option, the AS path of the customer network contains two AS numbers. ISP has to adapt the incoming AS-path filters.**

The **neighbor** *ip-address* **local-as** *private-as* router configuration command is used to indicate the AS number that the local router uses as its local AS number in the BGP Open message. The remote router is assumed to have an EBGP session to the indicated local AS.

Internally, the customer network uses another private AS number. When routes are sent to the neighbor, the internal AS number is automatically prepended in the AS path first, and then the specified local AS number is prepended as well. As a consequence, the ISP receives the routes with an AS path with both AS numbers in it. The ISP has to adapt its incoming filter-lists as a result of this situation.

Note Some service providers might be unwilling to change their AS-path input filters, leaving the customer no other option than using a public AS number or connecting to a single ISP with a private AS number.

Practice

Q1) What are two issues that arise when you are using AS number translation? (Choose two.)

A) The upstream provider must not filter routes based on a single AS.

B) AS-path prepending is not supported when you are using AS number translation.

C) AS number translation causes two AS numbers to be prepended to the AS path.

D) You can use only private AS numbers with AS number translation.

Primary/Backup Link Selection

This topic describes how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment.

Primary/Backup Link Selection

Cisco.com

Outgoing link selection:

- **You can use the same solution as with multihomed customers connected to one service provider**

Incoming link selection:

- **You cannot use MED because it can be sent only to the neighboring AS and no farther**
- **You must use other means such as BGP communities or AS-path prepending to achieve incoming link selection**

BGP v3.0—4-16

When using BGP on multiple links between a customer and several service provider networks, the customer is solely responsible for controlling the use of the links between them for outgoing traffic. The customer chooses whether to use these links in a primary/backup or a load-sharing configuration.

If one link is primary and the other used for backup purposes only, the customer can use the local preference attribute in the configuration to direct all outgoing traffic over the primary link. This configuration is no different than the configuration used for customers with multiple connections running BGP to a single service provider.

Controlling the load distribution of incoming traffic over multiple links is more difficult in the multihomed scenario when links to multiple service providers are used. You cannot use MED when the customer connects to multiple providers because the updates are sent to two different autonomous systems. Recall that MED is used only when comparing routes received from a single directly connected AS over two parallel links. Therefore, route selection decisions will most likely use the AS-path attribute and prefer the route with the shortest AS-path length.

Practice

Q1) In what two ways is the primary/backup design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) You cannot accomplish incoming route selection using the MED attribute.

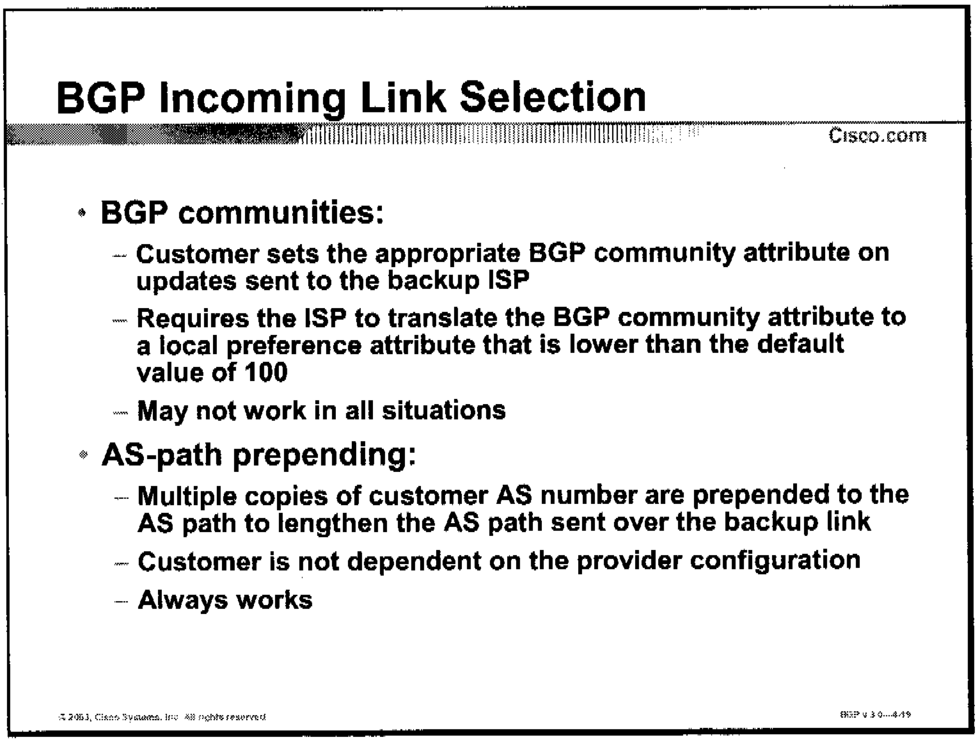
B) You can accomplish outgoing route selection only by using weights.

C) The customer should use local preference to direct traffic to the correct outgoing link.

D) BGP communities and AS-path prepending are used to influence incoming route selection.

BGP Incoming Link Selection

This topic describes the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers.



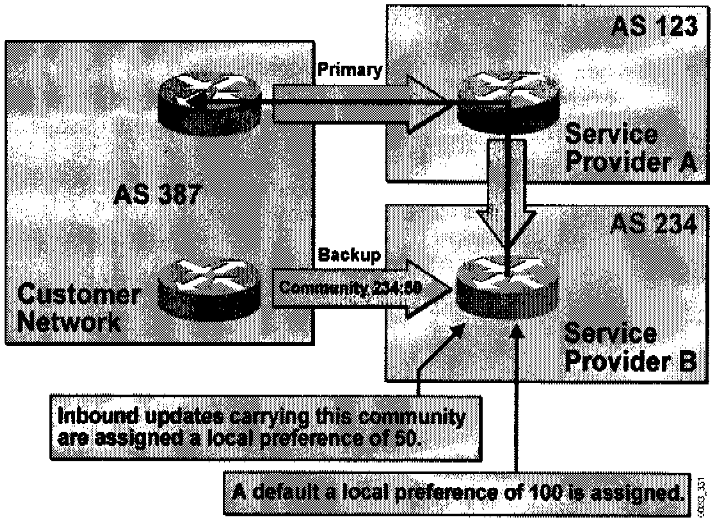
In order to remove incoming traffic from the backup link, the customer must influence route selection in the backup AS. The backup ISP must be forced to prefer the primary path to reach the customer network, although this choice means selecting a route with a longer AS path.

One way to influence route selection is to use local preference in the network of the backup ISP. Using local preference creates an administrative scalability issue if each customer requires its use, because the ISP must maintain the configuration.

One scalable way of setting local preference in an ISP network is to use communities. The customer sets a well-known community value on the routes sent to the backup ISP. The ISP recognizes the community and sets the local preference for these routes. This solution is available only if the ISP has implemented and announced the use of communities. If communities and a local preference setting are used, route selection occurs only if there are alternative routes to compare.

Another way of influencing route selection in the backup ISP is to do AS-path prepending before sending the advertisement to the backup ISP. When the customer sends routes over the backup link, multiple copies of its own AS number are prepended to the AS path of each route. The backup ISP receives the routes and makes normal route selection decisions. No special weight or local preference settings are used; the BGP route selection is based exclusively on the AS-path length. No special configuration is required in the service provider network.

BGP Incoming Link Selection Using BGP Communities

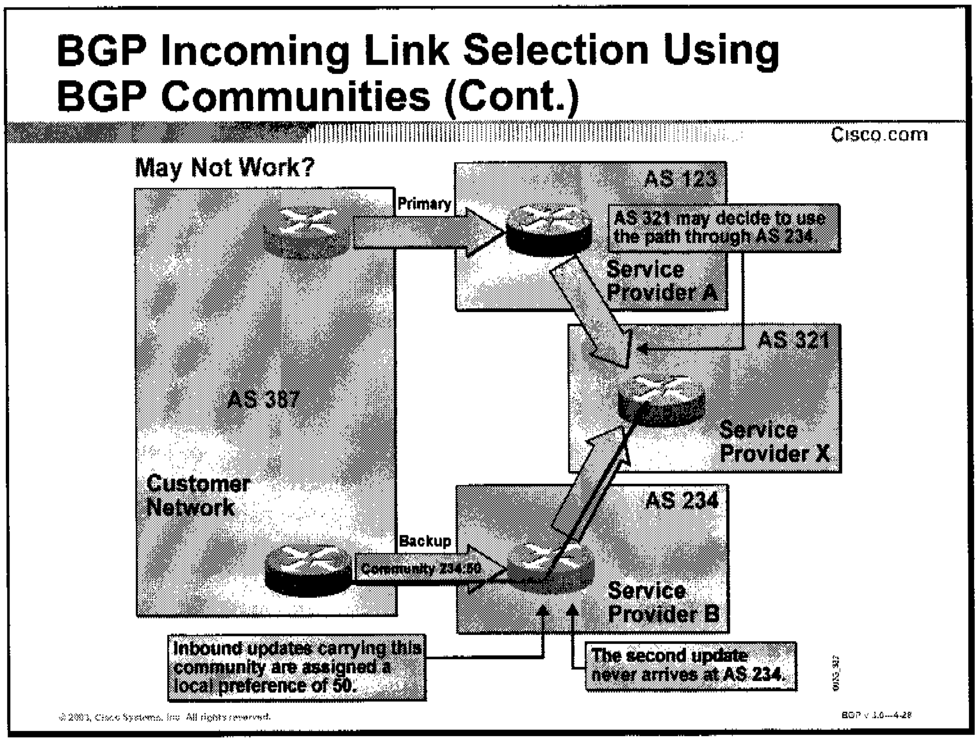


The backup service provider B (AS 234) has defined the meaning of community 234:50. When AS 234 receives routes with this community, the local preference is set to 50.

The customer AS 387 is advertising the route over the primary link without any communities. It is received by AS 123 and propagated to AS 234. When AS 234 receives the route via AS 123, there is no community set. AS 234 therefore assigns the default local preference value of 100.

The customer is also advertising the route over the backup link. However, in this case, the route has the community 234:50 set. When AS 234 receives this route, it recognizes the community, and the local preference value is set to 50.

Route selection is now performed in AS 234. The route received via AS 123 is preferred based on the local preference values.



Although the use of communities is correctly configured, the desired load distribution may not always be achieved. As this example shows, AS 234 does not always receive the primary route although nothing is wrong in the network.

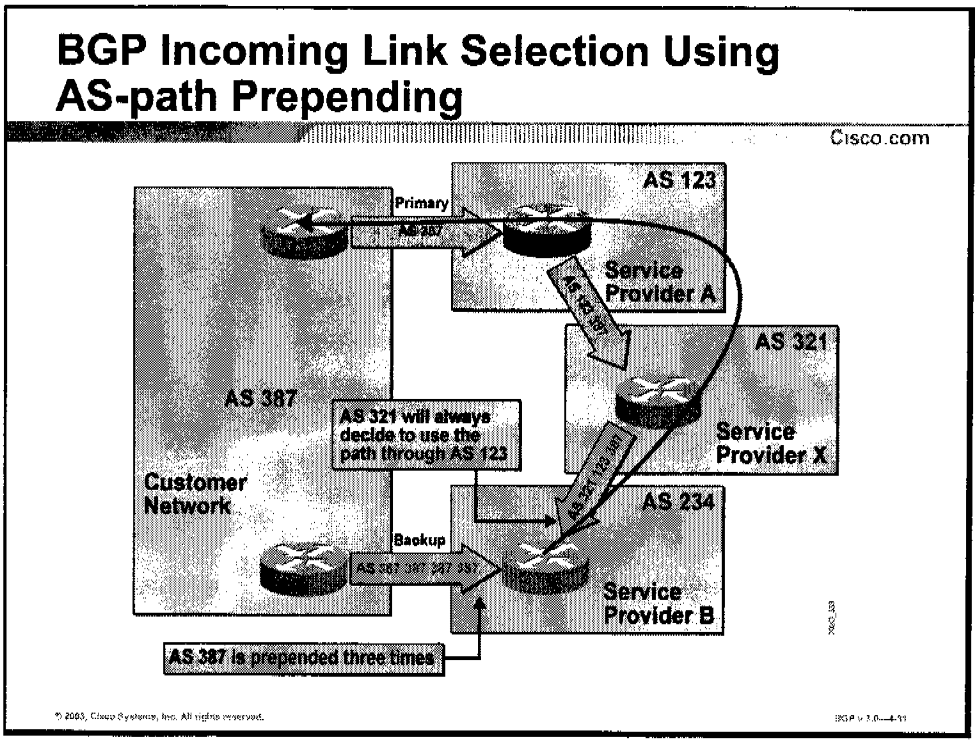
The customer AS 387 sends routes with community 234:50 over the backup link to AS 234. AS 234 receives the routes and sets the local preference to 50. If AS 234 over some period of time selects the directly connected path to AS387as the best, it propagates the route to AS 321. As the route is propagated over the EBGP session between AS 234 and AS 321, the local preference value used within AS 234 is lost.

AS 321 does not have any use for the community 234:50 because this community is defined and implemented only within AS 234. Potentially, the community value can also be stripped off during BGP route propagation.

Customer AS 387 also sends the routes over the primary link to AS 123. The routes are propagated to AS 321, which now sees two alternative routes to the destination networks within AS 387. Neither weight nor local preference is used by the routers in AS 321as criteria for reaching AS 387 for those routes within AS 321. Both alternatives have equal AS-path length.

The route selection decision that will be made in AS 321 is hard to predict, but the outcome definitely influences the route selection decision made in AS 234. If AS 321 prefers the route to the customer network via AS 234 for any reason, then the second-best alternative via AS 123 and the primary link is never propagated to AS 234.

In this case, AS 234 never sees the primary path and has to stick to the backup link and announce the route to AS 321. The network has reached a steady state when the traffic uses the backup link although the primary link is available.



In this example, the customer AS 387 is performing AS-path prepending on the backup link. Three copies of the customer AS number (387) are prepended to the AS path. As the route goes out over the EBGP session, BGP prepends the local AS number to the AS-path attribute. AS 234 receives routes from AS 387 over the backup link with an AS-path length of four (the original AS 387 plus three prepended copies that the customer edge router applied to the AS-path attribute).

The customer advertises networks without AS-path prepending over the primary link. AS 123 receives routes with an AS-path length of one and propagates these routes to AS 321, which then receives them with an AS-path length of two.

If, for a short period of time, AS 321 received the customer routes via AS 234, the AS-path length of those routes would have been five. In that case, AS 321 selects the route from AS 123 as the best and propagates it to AS 234.

AS 234 now sees both alternatives. The customer routes received directly from the customer have an AS-path length of four. The routes received via AS 321 have an AS-path length of three. Because no weight or local preference is configured in this example, AS 234 selects the route via AS 321 as the best.

The desired result, to have all traffic enter the customer network via the primary link, is now achieved.

Note	If the backup ISP is implementing incoming AS-path filters for this customer with the length of the AS path equal to one, the ISP has to change the configuration of the AS-path filter for the customer. The ISP can either create a new filter, allowing multiple copies of the customer AS number only for this customer, or create a common filter for all customers belonging to one peer group, using regular expression variables.

Practice

Q1) What are three benefits of using AS-path prepending instead of BGP communities when you are influencing incoming path selection? (Choose three.)

A) The customer is not dependent upon the service provider for configuration changes and maintenance.

B) Using BGP communities to set local preference may not work in all network scenarios.

C) AS-path prepending requires only that the backup provider support multiple instances of the same AS number in the AS path to function correctly.

D) Using BGP communities to set local preference cannot be scaled in a service provider network.

Load Sharing with Multiple Providers

This topic describes how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment.

Load Sharing with Multiple Providers

Cisco.com

Load sharing for outgoing traffic:

- You can use the same solution as with multihomed customers connected to one service provider

Load sharing for incoming traffic:

- The only load-sharing option that you can use in this setup is to separate address space into two or more smaller address blocks
- Some traffic analysis is needed to fine-tune address space separation according to link bandwidths
- You should use AS-path prepending to ensure symmetric routing as well as backup for noncontiguous address blocks

Load sharing over links to two different ISPs can be compared to doing load sharing over two parallel links to a single ISP. The only difference is that there is only one option available to control incoming traffic. Controlling load distribution of the outgoing traffic is configured in exactly the same way as when a multihomed customer connects to a single service provider.

The customer can control the load distribution of incoming traffic based on traffic destination. The customer divides its address space into several announcements. One announcement is sent to one of the ISPs. Another announcement is sent to the other ISP. For backup purposes, the customer announces the entire address space to both ISPs. The ISPs now use the most explicit route rule, and as long as both links are up, traffic with destinations within one part of the customer address space is routed over one of the links and traffic to the other part is routed over the other link.

It is very difficult to predict the volume of traffic that will be directed to one part of the customer address space and the volume that will be directed to the other part. You should monitor the results of changing route updates by watching the load on the links before and after implementing the change. If the load distribution is not satisfactory, you can further modify the division of the address space. You must then check the load on the links again and further fine-tune the configuration.

A customer may decide to use both the division of address space into several advertisements and AS-path prepending together. Some part of the customer's address space may be advertised by the customer network with a longer AS path over one of the links to fine-tune the load. Also, there may be cases where there are noncontiguous subnets that cannot be divided because the prefixes would be too long. These subnets are evenly distributed between the links in a primary/backup configuration.

Practice

Q1) In what two ways is the load-sharing design different from the one used for multihomed customers connected to a single ISP? (Choose two.)

A) There is no difference in the design for outgoing traffic.

B) There is no difference in the design for incoming traffic.

C) You can use AS-path prepending.

D) The weight, MED and local preference attributes all have to be altered.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- Customers requiring the maximum redundancy in their network design would implement a configuration that is multihomed to multiple service providers.
- A customer that is multihomed to multiple BGP service providers must advertise its address space to both ISPs as well as take care not to transmit any routing information between the two ISPs.
- The internal addresses of the customer must be advertised to both ISPs. Depending on the addressing scheme being used by the customer, NAT may be required.
- Customers connected to only one ISP do not require a public AS number, while customers connected to multiple ISPs must use an AS number that all ISPs agree to.

 BGP v3.0—4-33

Summary (Cont.)

Cisco.com

- You can use AS number translation to append a different AS number to the AS path, allowing the customer to use a single private AS number in the network.
- Outgoing route selection in primary/backup connectivity is achieved using local preference. Incoming route selection should be implemented using either BGP communities to tag customer routes or AS-path prepending.
- Load-sharing configurations for outgoing traffic are the same as those used in the scenario in which the customer is multihomed to a single provider. You can perform load sharing of incoming traffic when multihomed to multiple providers only if separate address spaces are advertised to each provider. You can also use AS-path prepending this configuration for fine-tuning purposes.

 BGP v3.0—4-34

Next Steps

After completing this lesson, go to:

- BGP Transit Autonomous System module

References

For additional information, refer to these resources:

- For more information on the BGP local AS feature, refer to "Configuring the BGP Local-AS Feature" at the following URL: http://www.cisco.com/warp/public/459/39.html

- For more information on load sharing, refer to "Sample Configurations for Load Sharing with BGP in Single and Multihomed Environments" at the following URL: http://www.cisco.com/warp/public/459/40.html

Quiz: Connecting a Multihomed Customer to Multiple Service Providers

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe BGP configuration characteristics used to establish routing between a multihomed customer and multiple service providers
- Describe addressing strategies available to a multihomed customer connected to multiple service providers
- Describe AS numbering strategies available to a multihomed customer connected to multiple service providers
- Describe the operation of AS number translation
- Describe how you can implement a typical backup setup between a multihomed customer and multiple service providers in a BGP environment
- Describe the use of BGP attributes to influence inbound link selection in customer networks multihomed to multiple service providers
- Describe how you can implement load sharing between a multihomed customer and multiple service providers in a BGP environment

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) A multihomed customer is using AS number 65500 internally. The customer is connected to two different providers. Provider 1 (in AS 222) has assigned the customer an AS number of AS 65101. Provider 2 (in AS 333) has assigned the customer an AS number of AS 65201. Given that the customer will use AS number translation for its internal AS, what is the AS-path attribute, attached to routes originated in the customer network, when displayed on a router in the network of Provider 2?

A) 65550 i

B) 65201 i

C) 65201 65550 i

D) 333 65201 i

Q2) What three methods can you use to provide load sharing over network links between a multihomed customer and multiple providers? (Choose three.)

A) advertising of split addressing space to the provider

B) configuring of multiple static routes pointing to the provider

C) use of the BGP **maximum-paths** command to perform load sharing over parallel links

D) AS-path prepending to fine-tune the load-sharing configuration

Q3) What are three BGP configuration characteristics of a multihomed customer connected to multiple providers? (Choose three.)

A) The customer announces assigned addressing to its providers through BGP.

B) The customer announces a default route to its network through BGP.

C) The provider announces a default route, local routes, or full Internet routing to the customer via BGP.

D) The customer configures outbound filters to prevent its network from becoming a transit area.

Q4) A multihomed customer is using AS number 1024 and is connected to two different providers (Provider 1: AS 222 and Provider 2: AS 333). The customer has configured MED to ensure a proper return path so that Provider 1 is the primary provider and Provider 2 is the backup provider. Unfortunately, return traffic continues to use the backup link. What is a possible cause of this problem?

A) The backup provider is ignoring the MED attribute on received routes.

B) The MED attribute cannot be sent to the backup provider because it is local to AS 1024 only.

C) The customer has not set the proper BGP communities to allow the primary and backup providers to correctly set the MED attribute.

D) MED cannot be used in this scenario, because it will not be advertised to providers upstream of Provider 2.

Q5) What are three important considerations for customers wishing to connect to multiple providers? (Choose three.)

A) The customer has to consider whether to use PA or PI address space.

B) The customer has to decide whether to use static routes or BGP to connect to upstream providers.

C) The customer has to decide whether to use a public AS number or a private AS number scheme.

D) The customer has to decide whether to perform load sharing or use a primary/backup implementation over redundant links.

Q6) Which AS number selection is the best possible choice for a customer multihomed to multiple providers?

A) a single public AS number

B) a single private AS number

C) two private AS numbers used in conjunction with AS number translation

D) multiple private AS numbers, one used internally by the customer and the others used in conjunction with AS number translation for each provider

Q7) Given the following router command output, what two methods have been configured to influence return traffic in a primary/backup link for this multihomed customer? (Choose two.)

```
Provider# show ip bgp

BGP table version is 5, local router ID is 10.0.33.34
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
*  10.10.20.0/24    192.168.63.3          2000              0 100 100
100 i
*>                  192.168.64.4                            0 300 100
i
*> 30.30.30.0/24    192.168.63.3             0              0 300 100
i
*> 40.40.40.0/24    192.168.64.4             0              0 100 i
```

A) MED

B) local preference

C) split address advertisement

D) AS-path prepending

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

5

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 issues in transit autonomous systems, ranging from synchronization between an Interior Gateway Protocol (IGP) and BGP to Internal Border Gateway Protocol (IBGP) full-mesh and next-hop requirements.

Upon completing this module, you will be able to:

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

Outline

The module contains these lessons:

- Transit Autonomous System Functions
- IBGP and EBGP Interaction in a Transit Autonomous System
- Packet Forwarding in Transit Autonomous Systems
- Configuring a Transit Autonomous System
- Monitoring and Troubleshooting IBGP in Transit Autonomous Systems

Transit Autonomous System Functions

Overview

The topology of the Internet can be viewed as a series of connections between stub networks, multihomed networks, and transit autonomous systems. A multihomed autonomous system (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 Border Gateway Protocol (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.

Importance

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

Objectives

Upon completing this lesson, you will be able to:

- 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

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

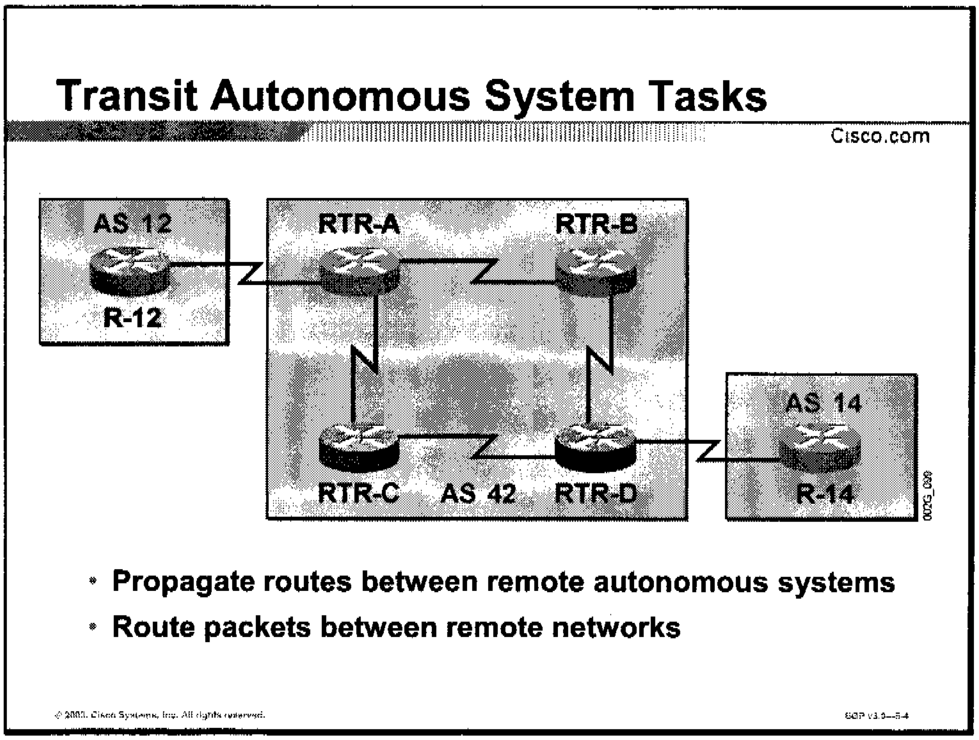
Outline

This lesson includes these topics:

- Overview
- Transit Autonomous System Tasks
- External Route Propagation
- Internal Route Propagation
- Packet Forwarding in an Autonomous System
- Core Router IBGP Requirements in a Transit Autonomous System
- Summary
- Assessment (Quiz): Transit Autonomous System Functions

Transit Autonomous System Tasks

This topic lists 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 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 received in the first task.

Practice

Q1) What are the two main functions of a transit AS? (Choose two.)

A) to propagate routes between remote autonomous systems

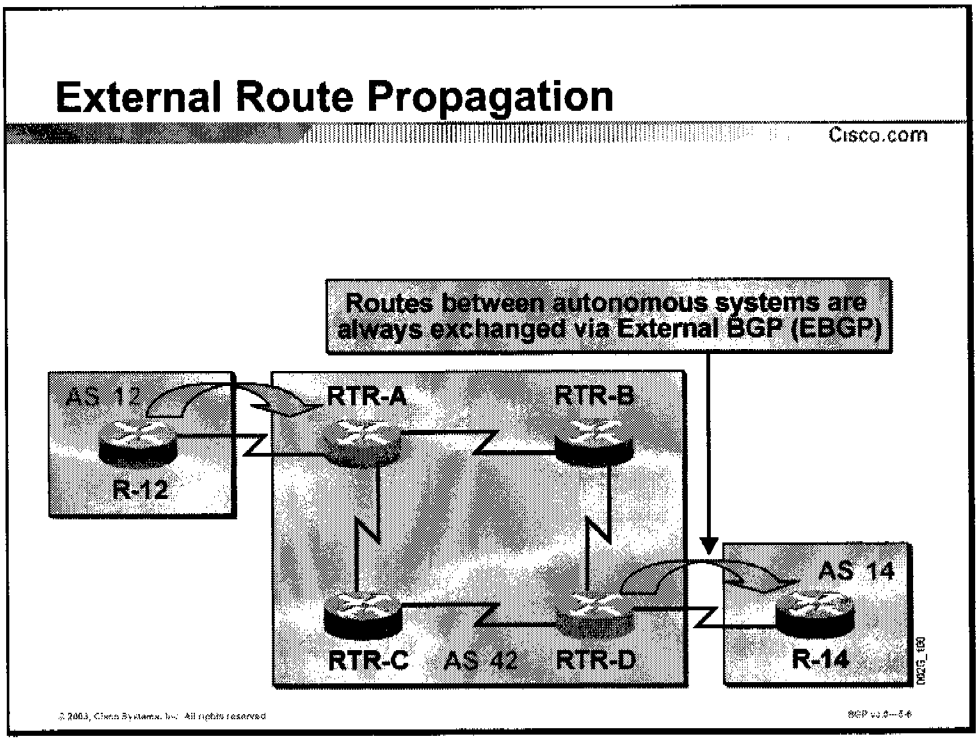
B) to filter noncustomer routes from transiting the AS

C) to route packets between remote networks

D) to connect customer networks to Internet service providers

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 example here, 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 thus need no additional routing information in order to establish the session.

Practice

Q1) How are external routes exchanged between autonomous systems?

A) through EBGP

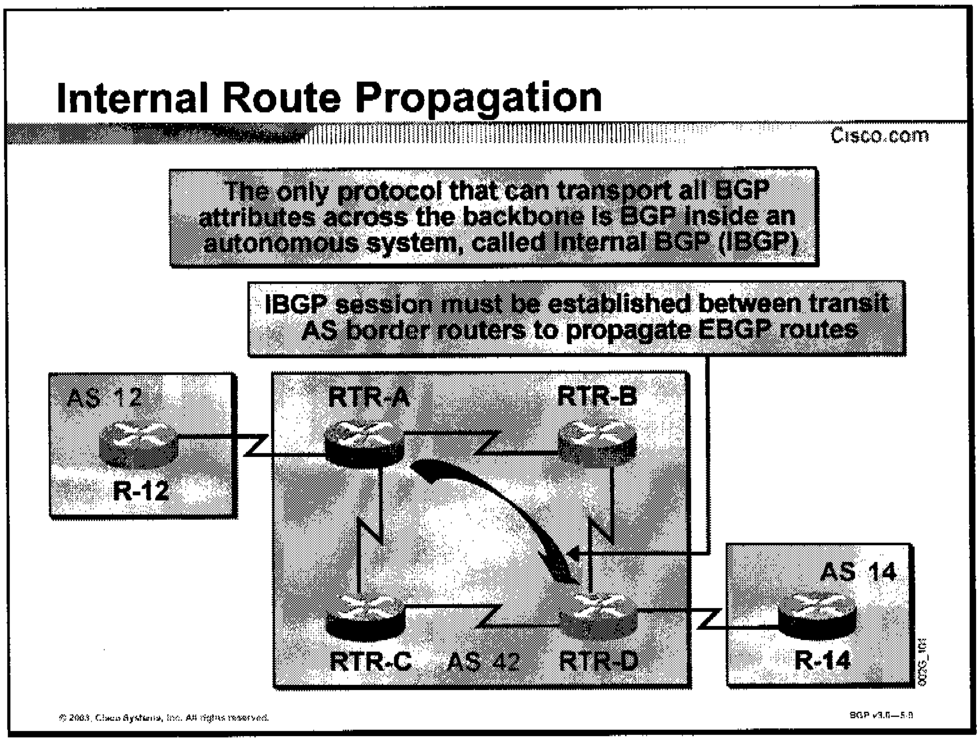
B) through IBGP

C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

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 in order 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.

Practice

Q1) How are BGP routes propagated across an AS?

A) through EBGP

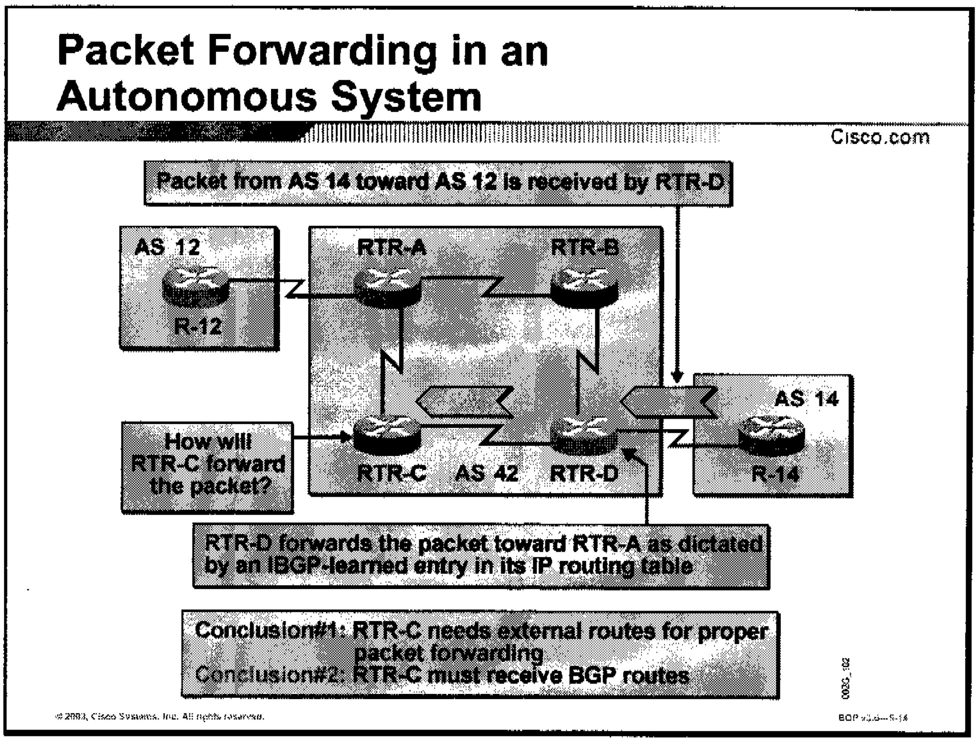
B) through IBGP

C) through route redistribution into an IGP

D) through local advertisement of routes at the edge router

Packet Forwarding in an Autonomous System

This topic explains 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 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 is, of course, an unacceptable situation. To prevent dropped packets due to 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.

Practice

Q1) Why do you have to run BGP on all core routers in a transit AS?

A) to eliminate the possibility of routing loops

B) to optimize the routing across the AS

C) to be able to forward packets to all external destinations

D) to ensure that only one exit point exists for the transit backbone

Core Router IBGP Requirements in a Transit Autonomous System

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

Core Router IBGP Requirements in a Transit Autonomous System

Cisco.com

All core routers must have all external routes

Core routers must receive BGP routes

- **Redistribution of BGP routes into IGP is not scalable**
- **Default routing is not applicable in transit AS core**

Within a transit AS, all routers, which are in a theoretical transit path between external destinations, should have information about all external routes 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 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 when redistributing the 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.

Practice

Q1) Why is the redistribution of BGP routes into IGP not advisable?

A) Redistributing full Internet routing into any IGP would result in the loss of the BGP attributes needed to ensure optimal routing within an AS.

B) Redistributing full Internet routing is not possible, because BGP policies cannot be enforced by IGPs.

C) IGPs are not capable of handling the number of routes currently present on the Internet.

D) The increased convergence times of IGPs as compared to BGP would cause too many flaps, rendering BGP unstable for Internet use.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- Routers in a transit AS receive routing information updates from neighboring autonomous systems, propagate the information through their own AS, and send it to another neighboring AS.
- Two autonomous systems usually exchange routing information over an EBGP session.
- A BGP session between two routers in the same AS is called an Internal BGP (IBGP) session.
- In order for packets to be properly forwarded in a transit AS, all routers must have external routing information.
- The only feasible method of distributing external routing information to all routers in the transit AS is through IBGP.

Next Steps

After completing this lesson, go to:

- IBGP and EBGP Interaction in a Transit Autonomous System lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to "BGP Case Studies Section 1" at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html
- For more information on the attributes in transit autonomous systems, refer to "Using the Border Gateway Protocol for Interdomain Routing" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Quiz: Transit Autonomous System Functions

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- 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

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Why is IBGP a mandatory component of 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?

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 it not recommended to redistribute BGP routes into an IGP for use in a transit backbone?

A) Redistribution removes all BGP attributes 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 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.)

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?

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

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

IBGP and EBGP Interaction in a Transit Autonomous System

Overview

A transit autonomous system (AS), by definition, allows traffic not originating in that AS to travel through it. This activity requires interaction between External Border Gateway Protocol (EBGP) and Internal Border Gateway Protocol (IBGP) in the transit AS. This lesson introduces the requirements of IBGP and 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.

Importance

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

Objectives

Upon completing this lesson, you will be able to:

- Describe AS-path processing in IBGP
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers, and its implications
- Explain 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 routers to announce themselves as the next hop in IBGP updates
- Explain the differences between EBGP and IBGP sessions

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

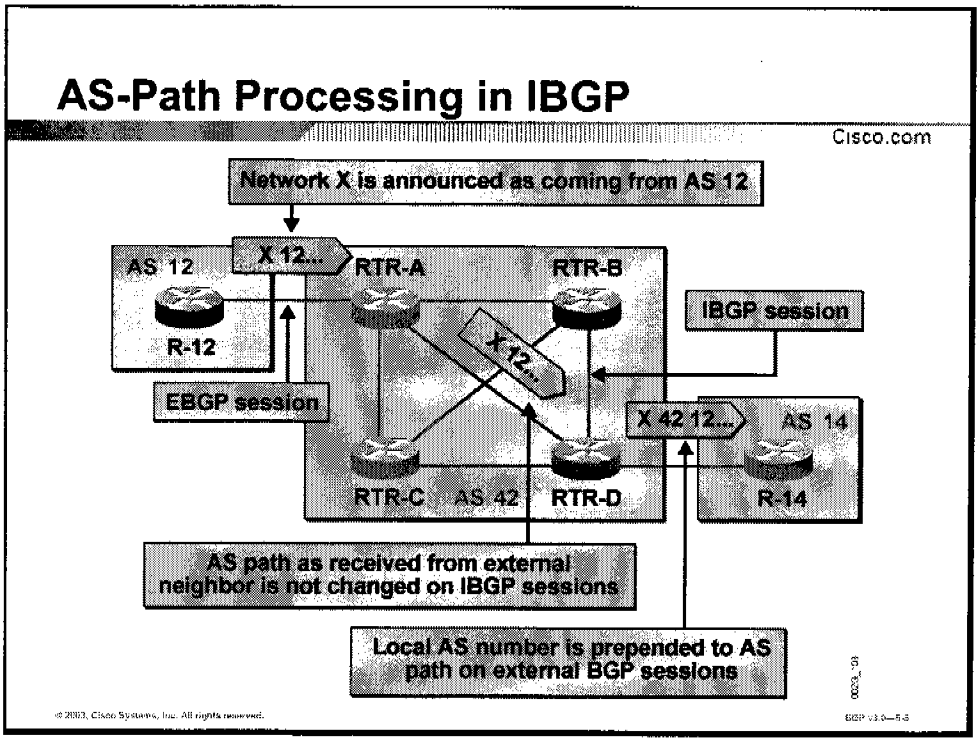
Outline

This lesson includes these topics:

- Overview
- AS-Path Processing in IBGP
- BGP Split Horizon
- IBGP Full Mesh
- IBGP Neighbors
- IBGP Next-Hop Processing
- Transit Network Using External Next Hops
- Transit Network Using Edge Routers as Next Hops
- Differences Between EBGP and IBGP Sessions
- Summary
- Assessment (Quiz): IBGP and EBGP Interaction in a Transit Autonomous System

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 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 example, 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, it 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."

Practice

Q1) How is the AS-path attribute modified across a transit AS?

A) The AS path of each router is prepended to the AS-path attribute.

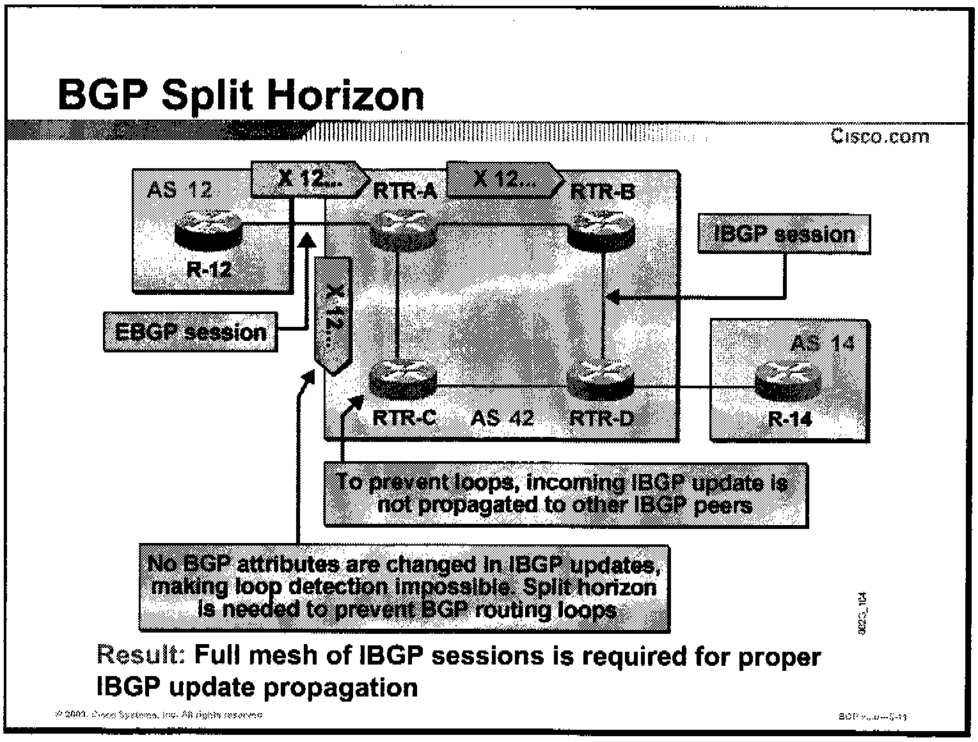
B) The AS path is not changed when the BGP prefix is propagated across IBGP sessions.

C) The AS number of the router sending the information is appended to the AS path.

D) The value of the AS path is reset to the AS path of the transit backbone.

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 received through an IBGP session is never forwarded to another IBGP neighbor, only toward EBGP neighbors. Due to 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.

Practice

Q1) What is the purpose of the BGP split-horizon rule?

A) to propagate IBGP updates to other IBGP peers

B) to prevent BGP routing loops within an AS

C) to send routing information to an EBGP neighbor

D) to prevent BGP routes from being sent back to the advertising router

Q2) What is the impact of BGP split horizon on an AS?

A) IBGP-speaking routers forward BGP updates to all other IBGP-speaking routers within the transit backbone.

B) The IBGP full-mesh requirement within the transit AS is no longer required because BGP routing loops are eliminated.

C) Convergence time within the AS is reduced because fewer routes have to be processed by each IBGP neighbor router.

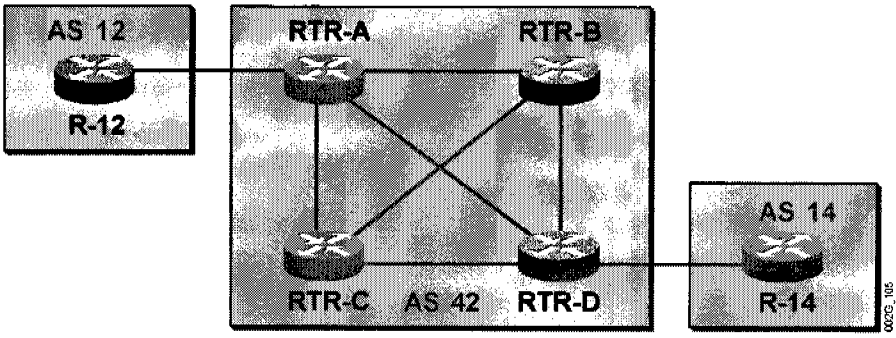
D) All BGP routers in an AS must be updated directly by the border router receiving the update via EBGP.

IBGP Full Mesh

This topic explains the need for a full-mesh topology between IBGP routers, and its implications.

IBGP Full Mesh

Cisco.com



- Full mesh of IBGP sessions has to be established between all BGP-speaking routers in the AS for proper IBGP route propagation
- The IBGP full mesh is a logical mesh of TCP sessions only; physical full mesh is not required

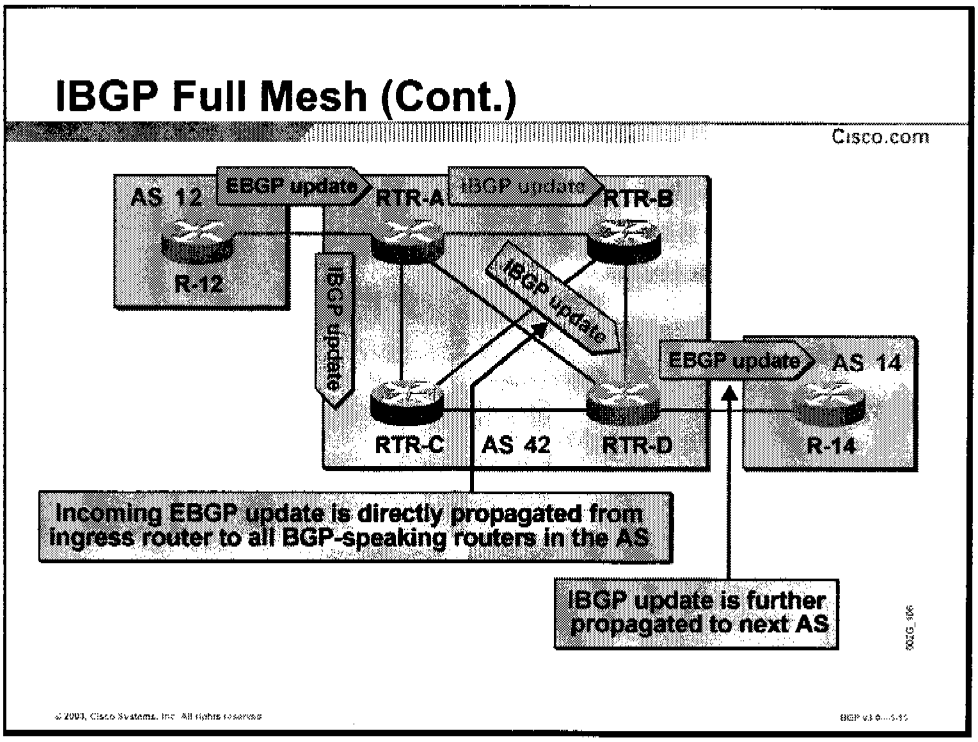
Because every router on the transit path within the AS must have routing information about all external networks received by any of the border routers, all routers must have IBGP sessions to all border routers. But this is not enough, because any of the internal routers could also create new BGP routing 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 sample network, RTR-A must have IBGP sessions with RTR-B, RTR-C, and RTR-D in order to propagate routes 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 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



The figure here illustrates IBGP split-horizon and IBGP full-mesh principles in the sample network. R-12 is sending an update to RTR-A over an EBGP session. Updates 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 due to BGP split horizon.

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

Practice

Q1) What are two reasons why a fully meshed logical topology is required for all IBGP neighbors in an AS? (Choose two.)

A) to receive routing information about external networks from border routers

B) to facilitate BGP split-horizon rules

C) to allow IBGP-speaking routers to inject routes into BGP

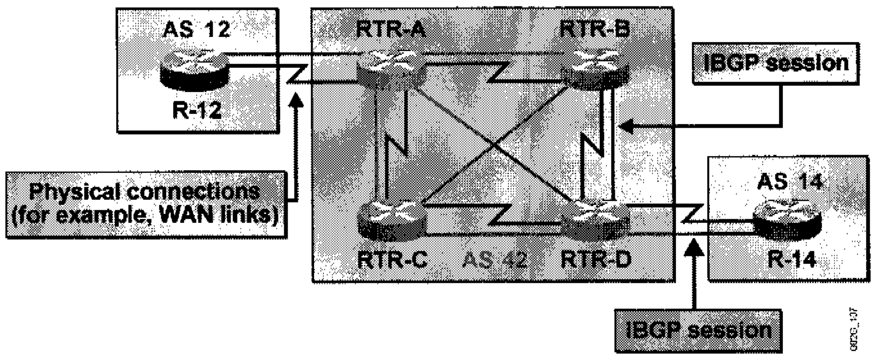
D) to ensure that each IBGP router has the most current BGP table and attributes

IBGP Neighbors

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

IBGP Neighbors

Cisco.com



- **Due to IBGP full-mesh requirements, IBGP neighbors are usually not directly connected**
- **Which interfaces should be chosen as the source and destination addresses of IBGP TCP sessions?**

In this example, 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 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 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 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 as long as the two BGP routers have any usable path between them.

IBGP Neighbors (Cont.)

Cisco.com

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 session**
- **There is no BGP recovery after a failure inside the transit autonomous system**
  - **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.

In order 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 assigned to each loopback interface so that they are reachable by all BGP routers in the AS.

The IBGP sessions 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, BGP sessions 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.

Practice

Q1) What is the recommended way to run IBGP sessions?

A) Establish IBGP sessions using IP addresses of physical interfaces to allow detection of link failures.

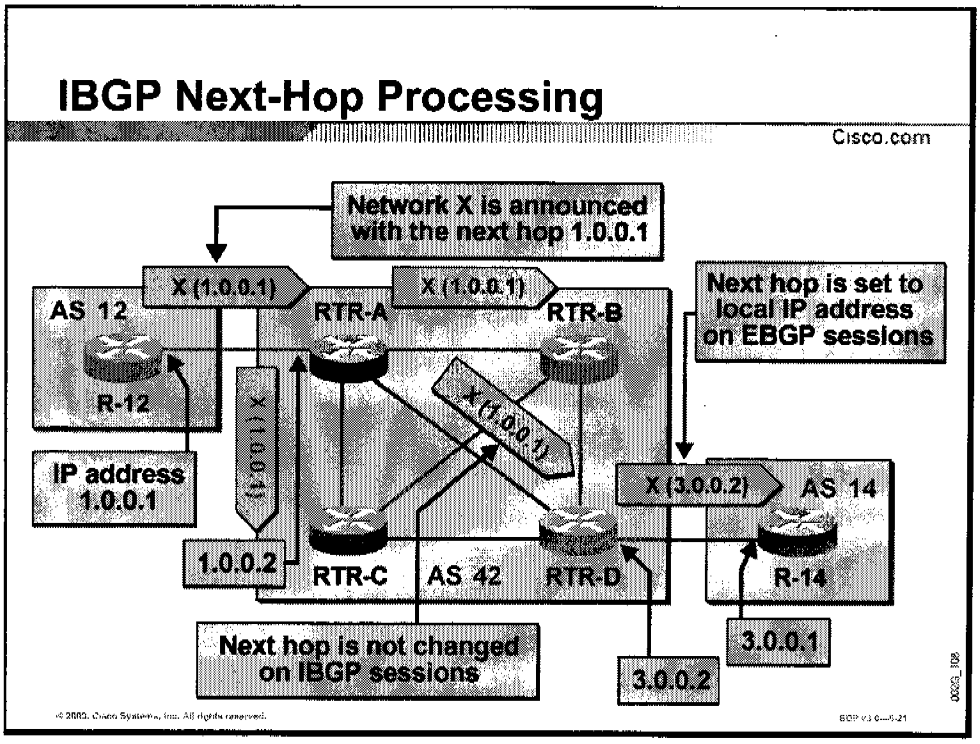
B) Establish an IBGP session using IP addresses of loopback interfaces to prevent IBGP sessions from failing if a physical interface fails.

C) Establish IBGP sessions using the IP addresses of the fastest interface to make sure that packets are forwarded out the fastest interface.

D) There is no recommendation. Any IP address on a router can be used for an IBGP session without any penalties.

IBGP Next-Hop Processing

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



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 forwarding 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 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 will 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 forwarding 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 they 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 will set 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.

Practice

Q1) How is the BGP next hop changed inside an AS?

A) The next-hop address is set to the router-ID of the forwarding router.

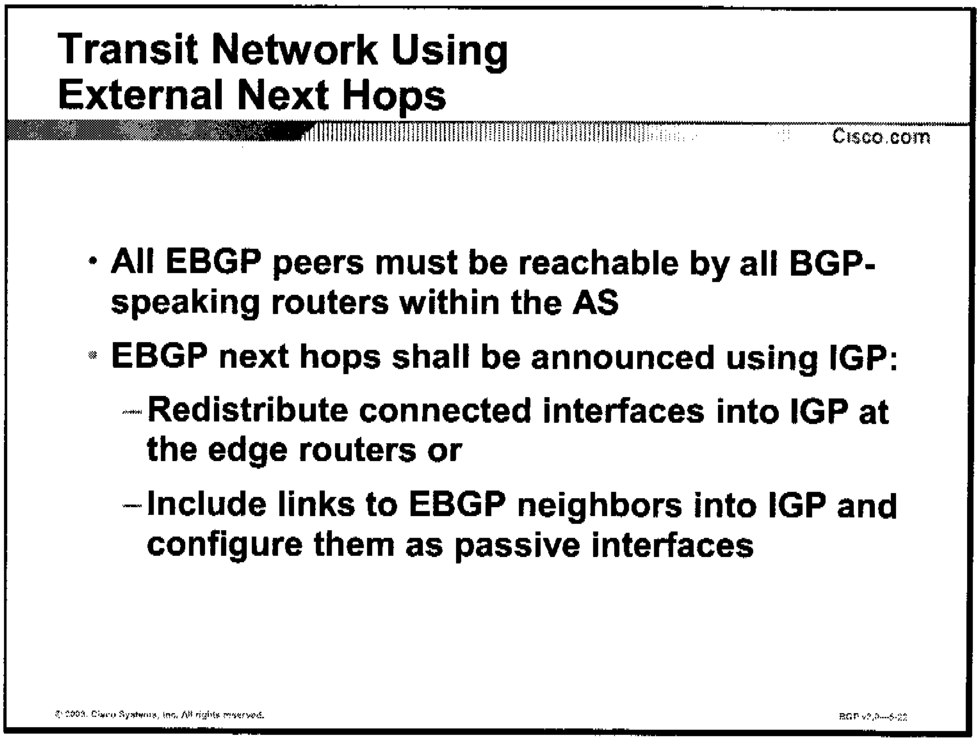
B) The next-hop address is not modified.

C) The next-hop address is set to the address of the receiver.

D) The next-hop attribute is set to the configured default gateway.

Transit Network Using External Next Hops

This topic explains 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 forwarding 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 used to reach the next-hop IP address.

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

One way of making the IGP carry the information 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 redistributed into the IGP might appear as external IGP routes, depending on what IGP is configured within the AS. There exist several scalability issues 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 different autonomous systems to exchange routes using the IGP. In that case, all benefits of having separate autonomous systems will be lost.

Practice

Q1) What are the implications of IBGP next-hop processing on the network design?

A) The network must contain a physical full mesh between all BGP-speaking routers.

B) All BGP external routes must be redistributed into the IGP by the border router.

C) All BGP speakers must have IGP reachability to all external neighbors to be able to perform a recursive lookup to resolve next-hop addresses.

D) You must configure static routes on all IBGP-speaking routers pointing to the border routers of the transit AS.

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.

Transit Network Using Edge Routers as Next Hops

Cisco.com

Alternate design: Next-hop processing is modified at the edge routers

- **Edge routers announce themselves as the next hop in IBGP updates**
- **No redistribution of external subnets is necessary**
- **This design might result in suboptimal routing if multiple paths to a neighboring AS exist**

Use default next-hop processing if at all possible

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 received on the EBGP sessions will be forwarded on the IBGP sessions and the next-hop attribute will be set to the IP address 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 the normal way. But the next-hop address used will be the IP address of the far end of the IBGP session, because the border router changed it. The IP address of the far-end IBGP peer is always known in the forwarding 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.

Transit Network Using Edge Routers as Next Hops (Cont.)

Cisco.com

```
router(config-router)#
neighbor ip-address next-hop-self
```

- Changes next-hop processing at edge routers
- Bypass the BGP next-hop processing and announce the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor
- Has to be set on all IBGP neighbors to fully bypass IBGP next-hop processing

BGP v3.0—5-24

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**

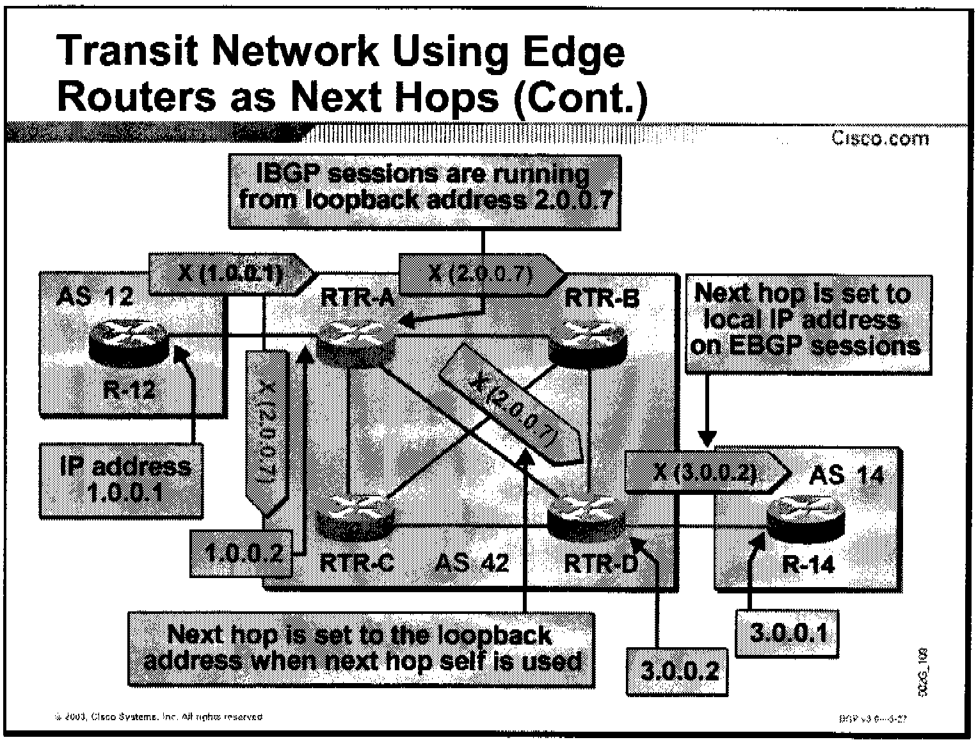
To disable this feature, use the **no** form of this command.

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

Syntax Description

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

Example



The next-hop attribute is normally not changed on IBGP updates. But in the figure, the **next-hop-self** configuration has been used on all IBGP sessions. 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 used as the source address of the IBGP session.

The receiver of IBGP updates will see next-hop information indicating a destination, which might not be directly connected. To resolve this problem, it will check its forwarding table and see if and how the next-hop address can be reached. Then it will 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 the destination address equal to the IP address stated 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 be 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 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 forwarding 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 as they would 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 used on that EBGP session, 3.0.0.2.

Practice

Q1) How can you influence IBGP next-hop processing?

A) by configuring the router to set the next-hop attribute to its router-ID

B) by configuring the router to set the next-hop attribute to the gateway of last resort

C) by configuring the router to ignore the next-hop attribute

D) by configuring the router to set the next-hop attribute to its source address of internal BGP sessions

Differences Between EBGP and IBGP Sessions

This topic explains the differences between EBGP and IBGP sessions.

Differences Between EBGP and IBGP Sessions

Cisco.com

- No BGP attributes are changed in IBGP updates
- Due to BGP split horizon, routes learned from IBGP peer are not advertised to other IBGP peers
- Local preference and MED attributes are propagated only over IBGP sessions
- EBGP peers are directly connected; IBGP peers are usually distant
- Route selection rules slightly prefer EBGP routes

Both EBGP and IBGP sessions forward BGP updates. But they do it in a slightly different manner:

- 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 received on an IBGP session to another IBGP session.

- IBGP border routers will 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.

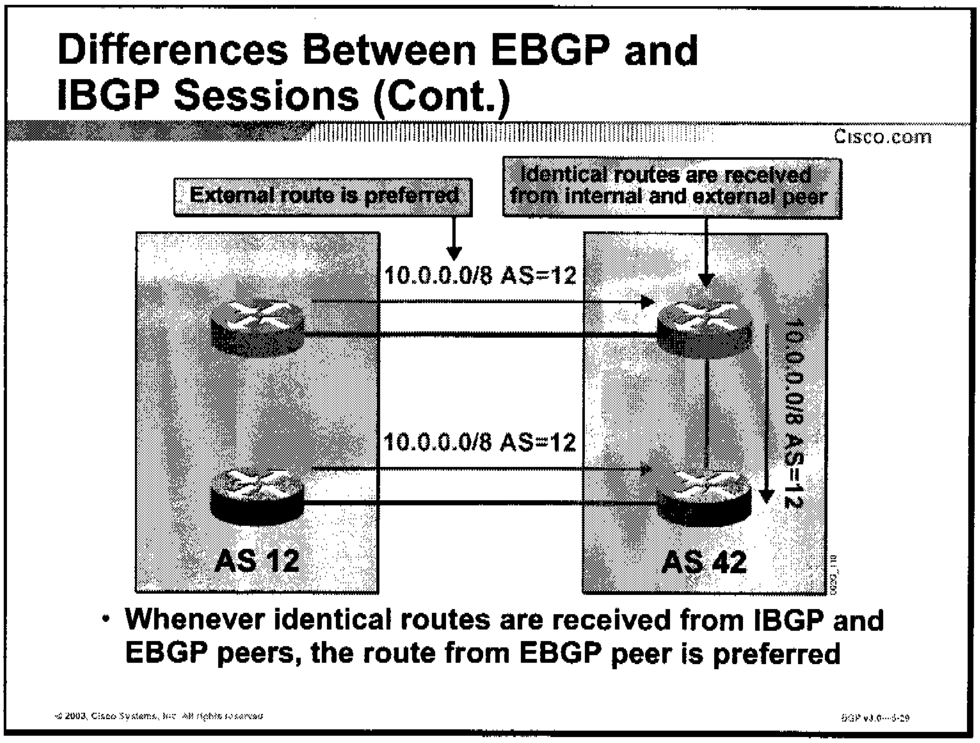
- The multi-exit discriminator (MED) attribute is received from a neighboring AS and distributed within the local AS. But when the update is about to be forwarded over another EBGP session to a third AS, the information is stripped off.

- 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. Since 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



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 received through the EBGP session.

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

In the figure, the lower 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 upper router in AS 42. All essential attributes are the same, so route selection cannot be made easily.

The lower 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 received on the EBGP session is preferred over the update received on the IBGP session.

Practice

Q1) What are the three major differences that exist between EBGP and IBGP? (Choose three.)

A) No BGP attributes are changed in IBGP updates.

B) Route selection rules slightly prefer IBGP routes.

C) Local preference and MED are propagated only over IBGP sessions.

D) Routes learned from IBGP are not advertised to other IBGP peers.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- All BGP routing updates carry the mandatory well-known attribute AS-path, which lists the autonomous systems that the routing update has already crossed. AS-path is not changed when the BGP prefix is propagated across IBGP sessions.
- Routing information loops within the AS are prevented by IBGP split horizon—routing information 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.
- For stability, the best choice when you are configuring IBGP sessions is to establish the session between loopback interfaces of BGP routers.

Summary (Cont.)

Cisco.com

- 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.
- In order to get the recursive routing to work, a router must resolve all possible next-hop references using information in the forwarding table. The IGP used in the AS must carry this information.
- 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.
- BGP attributes are not changed when an update is sent across an IBGP session, unless next-hop-self is configured.

Next Steps

After completing this lesson, go to:

- Packet Forwarding in Transit Autonomous Systems lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to "BGP Case Studies Section 1" at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html

- For more information on the next-hop attribute, refer to "Using the Border Gateway Protocol for Interdomain Routing" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm

Quiz: IBGP and EBGP Interaction in a Transit Autonomous System

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Describe AS-path processing in IBGP
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers, and its implications
- Explain 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 routers to announce themselves as the next hop in IBGP updates
- Explain the differences between EBGP and IBGP sessions

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) Which two statements are true regarding the AS-path attribute as it relates to IBGP? (Choose two.)

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.

Q2) Why is it recommended that loopback interfaces be used to form IBGP neighbor sessions?

A) Using loopback interfaces reduces router memory resource requirements.

B) Using loopback interfaces reduces router CPU resource requirements.

C) Using loopback interfaces ensures IBGP session stability.

D) Using loopback interfaces is more secure than using the physical interface.

Q3) How is the BGP next-hop attribute processed over an IBGP connection?

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 configured on the router.

Q4) Which two statements are true of the full-mesh requirement in IBGP? (Choose two.)

A) The IBGP mesh requires a logical full mesh.

B) A physical full mesh must be maintained within the IBGP AS.

C) Due to 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.

Q5) What three statements are true regarding the next-hop-self configuration in BGP? (Choose three.)

A) Changing the next-hop attribute might cause suboptimal routing.

B) This configuration changes how the next-hop attribute is processed at edge routers.

C) This configuration announces the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor.

D) This configuration removes the requirement for the IGP to carry reachability information for intra-AS destinations.

Q6) What are three differences between IBGP and EBGP sessions? (Choose three.)

A) Route selection rules slightly prefer IBGP routes.

B) Routes learned from IBGP peer 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.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Packet Forwarding in Transit Autonomous Systems

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.

Importance

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

Objectives

Upon completing this lesson, you will be able to:

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

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

This lesson includes these topics:

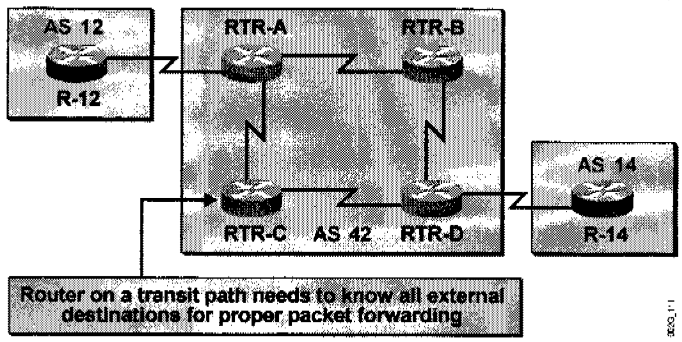
- Overview
- Packet Forwarding in a Transit Autonomous System
- Recursive Lookup in Cisco IOS Software
- Routing Protocols in a Transit Autonomous System
- BGP and IGP Interaction
- Problems with BGP and IGP Interaction
- Summary
- Assessment (Quiz): Packet Forwarding in Transit Autonomous Systems

Packet Forwarding in a Transit Autonomous System

This topic describes packet forwarding in a transit AS.

Packet Forwarding in a Transit Autonomous System

Cisco.com



All core routers must receive external routes—they must run BGP

- **The only scalable design, route redistribution into IGP, is not scalable**

BGP v3.0—5-4

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

Router R-14 will forward to RTR-D IP packets with the destination address matching a network in AS 12. RTR-D will check its routing table and find 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. In order 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 in from the other direction and destined to 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 received by RTR-A and RTR-D, respectively, 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 due to information overload, causing a total network meltdown in the AS. The volume of routing information carried by BGP in the contemporary Internet has long ago passed the limits for what it is possible to carry in any IGP.

Packet Forwarding in a Transit Autonomous System (Cont.)

Cisco.com

- **Routes learned via BGP do not have outgoing interface associated with them in the routing table**
- **Recursive lookup is performed to forward IP packets toward external destinations**

```
wg3pe2# show ip route 128.51.0.0
Routing entry for 128.51.0.0/16
  Known via "bgp 5", distance 200, metric 0
  Tag 99, type internal
  Last update from 192.168.5.1 01:21:25 ago
  Routing Descriptor Blocks:
  * 192.168.5.1, from 192.168.5.1, 01:21:25 ago
      Route metric is 0, traffic share count is 1
      AS Hops 5
wg3pe2# show ip route 192.168.5.1
Routing entry for 192.168.5.1/32
  Known via "ospf 1", distance 110, metric 1563, type intra area
  Redistributing via ospf 1, rip
  Advertised by rip metric 3
  Last update from 192.168.5.17 on Serial1/0.121, 02:09:15 ago
  Routing Descriptor Blocks:
  * 192.168.5.17, from 192.168.5.1, 02:09:15 ago, via Serial1/0.121
      Route metric is 1563, traffic share count is 1
```

No outgoing interface

Route toward BGP next hop

BGP next hop

Outgoing interface

A BGP route is installed in a router's IP routing table 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 is dependent upon the IP switching mechanism used by the router. At the latest, the router performs the recursive lookup when an IP packet with a destination address matching 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.

Practice

Q1) Why do you need to run IBGP on all core routers?

A) to ensure that a full mesh exists between all BGP routers in the AS

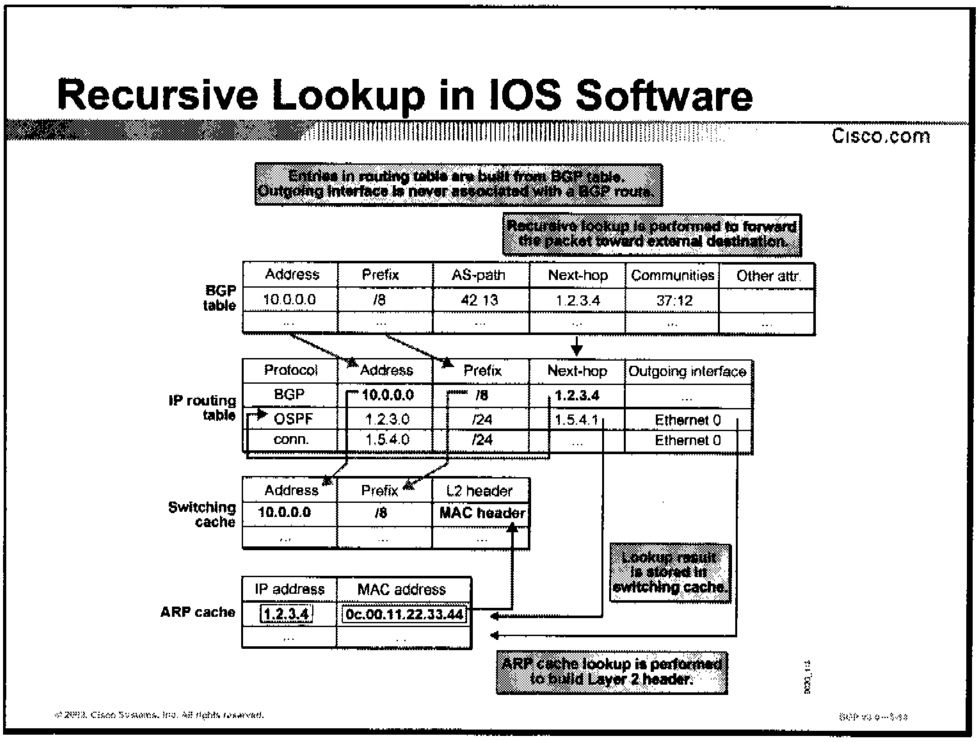
B) to allow routers to properly resolve the next-hop address

C) to allow routers to forward packets toward all external destinations

D) to ensure correct propagation of the gateway of last resort

Recursive Lookup in Cisco IOS Software

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



The figure here 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 will now find 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 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	This example illustrates the recursive lookup performed when the router uses cache-based IP switching mechanisms; for example, fast switching or optimum switching. See the next figure for more information on the differences between Cisco Express Forwarding (CEF) and the cache-based switching mechanisms.

Recursive Lookup in Cisco IOS Software (Cont.)

Cisco.com

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

 BGP v3.0—5-15

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 being done, including recursive lookup in the case of a BGP route. The result of the lookup is being cached for later use when successive packets to the same destination arrive. The process is repeated for every specific destination.

CEF prebuilds a complete IP forwarding table (called the Forwarding Information Base [FIB]) 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.

Practice

Q1) What is recursive lookup in Cisco IOS software?

A) verifying that the route exists both in the BGP table and the IP routing table

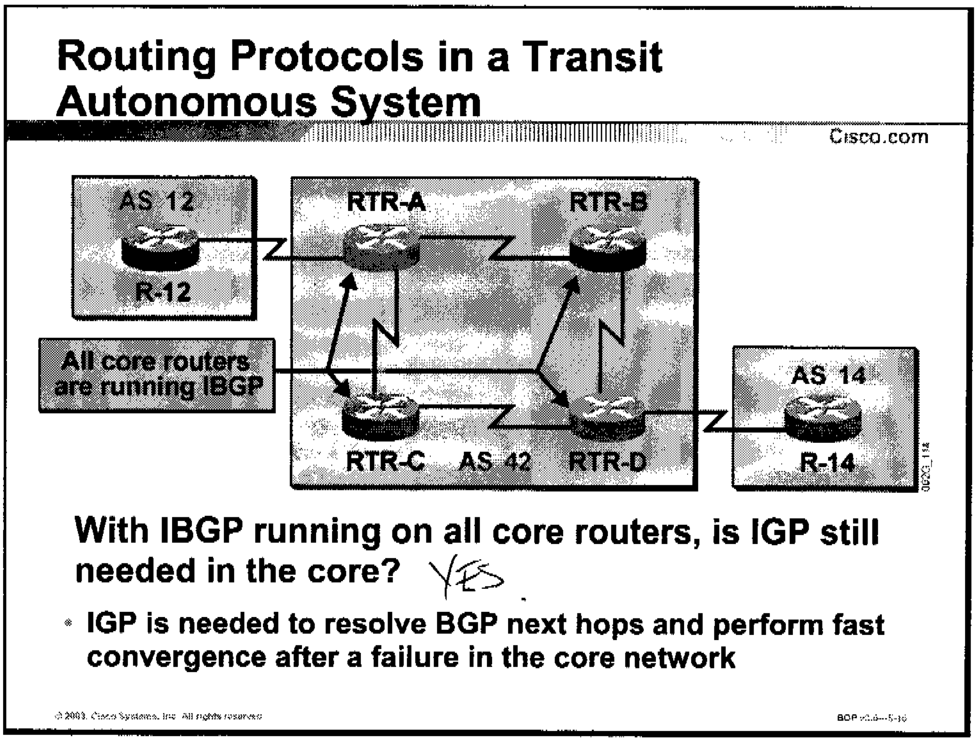
B) resolving the MAC address of the route to build the MAC header

C) searching the routing table to find a path to the next-hop address

D) the process used to build the IP fast cache on the Cisco IOS router

Routing Protocols in a Transit Autonomous System

This topic explains the need for an IGP in a transit backbone 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 needed to establish the IBGP sessions
- To resolve next-hop references (recursive routing)

For example, when RTR-D in the example here receives an IP packet with the destination in AS 12, it will do 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 will determine this fact. The IGP route to the next-hop network will be changed by the router due to newly received IGP route information, and all cache entries relying on the old information will be invalidated. The next recursive lookup that RTR-D performs will now indicate a different outgoing interface than before the problem occurred.

During the IGP convergence process, the BGP routing is not affected. The only routing update exchanged during the transit AS convergence would be 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 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 Autonomous System (Cont.)

Cisco.com

- Core routers need to run BGP and IGP
- BGP shall carry all external routes
- IGP shall propagate BGP next hops and other core subnets only
- All customer routes shall also be 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 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 in order to reduce the amount of information in the IGP and increase IGP stability.

Practice

Q1) What are two reasons why you need an IGP in a transit AS? (Choose two.)

A) to provide the gateway of last resort to BGP speakers

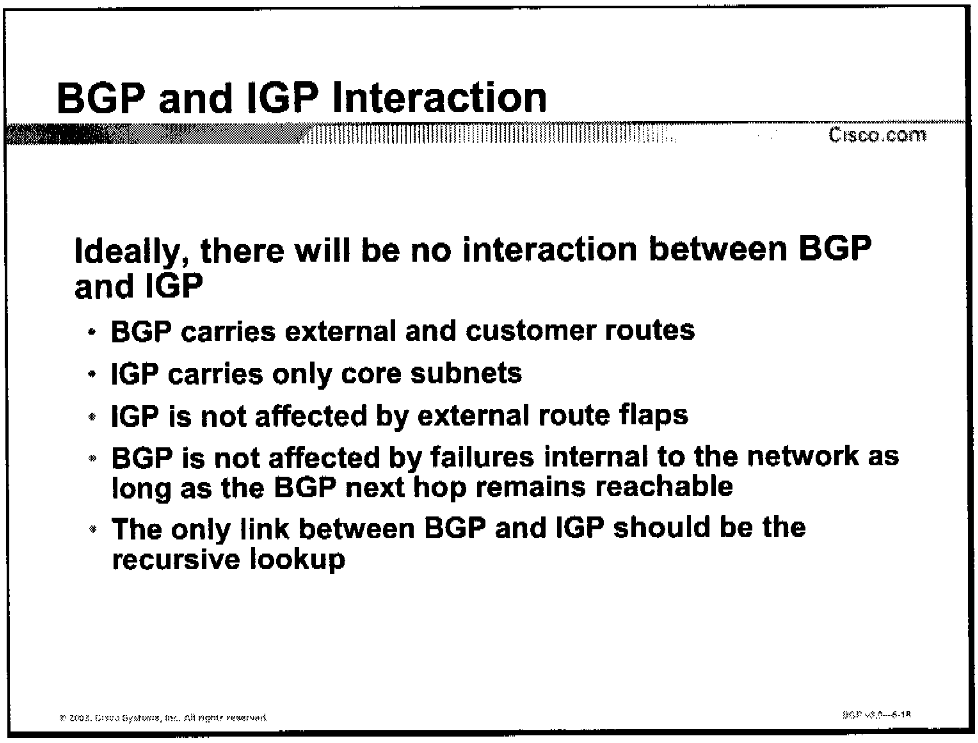
B) to perform the recursive lookup for external networks

C) to carry BGP routes through the transit AS

D) to establish IBGP sessions between nondirectly connected routers

BGP and IGP Interaction

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



Ideally, BGP and the IGP carry two different sets of routing information. BGP carries those routes received from other autonomous systems and those 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 next-hop addresses.

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 due to internal problems within the AS.

BGP and IGP Interaction (Cont.)

Cisco.com

Sometimes, BGP and IGP will propagate the same route

- **Usually due to bad network design**
- **In this case, routes are determinedd in EBGP/IGP/IBGP order based on administrative distances of the routes**

Routing protocol	Default administrative distance
EBGP	20
IGP	90 - 170
IBGP	200

BGP v3.0—5-59

Sometimes the interaction between BGP and the IGP is not ideal due to a number of reasons, including bad network design. In the worst case, the same networks might be carried in both the IGP and in BGP. For example, the subnets connecting the AS with neighboring autonomous systems have to be announced via 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 will install 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 learned via EBGP have a high likelihood of being installed in the forwarding table.

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

Practice

Q1) Why should you transport your customer routes in BGP and not in an IGP?

A) to ensure that the gateway of last resort propagates across the AS

B) to correctly resolve addresses in the next-hop attribute

C) to ensure full reachability to external network destinations

D) to protect the IGP from carrying too many routes

Q2) How should BGP react to a failure inside a transit AS that has redundant paths?

A) BGP sessions across the failed link will terminate.

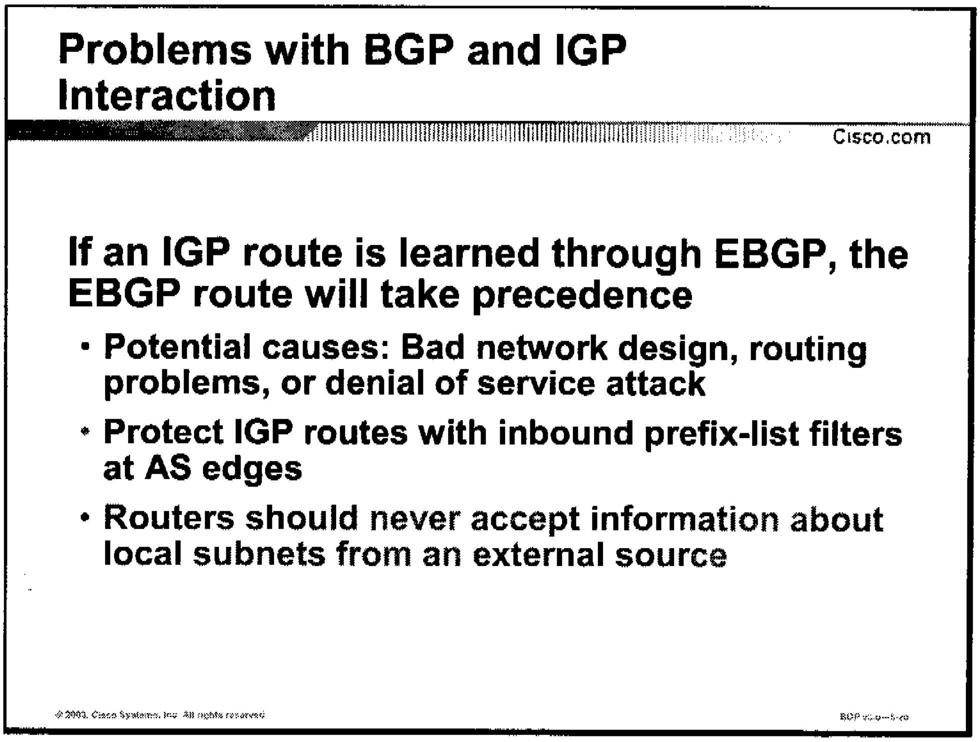
B) BGP does nothing because the IGP finds an alternate path to the neighbors.

C) Failure within the AS causes the BGP to recalculate new paths.

D) BGP will terminate the link between itself and IGP.

Problems with BGP and IGP Interaction

This topic explains 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 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.

Practice

Q1) What happens when the same route is learned via BGP and an IGP?

A) BGP will install both an EBGP and an IBGP route for the same destination.

B) If the same route is learned via both the IGP and EBGP, the EBGP route is preferred.

C) The route learned through the IGP will be installed in the routing table.

D) If the same route is learned via both the IGP and IBGP, the IBGP route is preferred.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- The only scalable way to ensure external destinations are reachable from inside the transit AS is through IBGP.
- A recursive lookup is performed in BGP to resolve the forwarding path reference of the next-hop attribute.
- Packet forwarding to external destinations benefits from the high-speed convergence offered by an IGP; therefore, an IGP is still needed inside a transit AS.
- The IGP should provide reachability toward BGP next-hop addresses only if they are not disturbed by external updates from other autonomous systems (those are handled by BGP).
- IP packets could be erroneously forwarded out of the local AS if an external AS accidentally (or by intent: DoS) feeds the local AS with EBGP routes that should be local.

 BGP v3.0—5-21

Next Steps

After completing this lesson, go to:

- Configuring a Transit Autonomous System lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to "BGP Case Studies Section 1" at the following URL: http://www.cisco.com/warp/public/459/13.html

Quiz: Packet Forwarding in Transit Autonomous Systems

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

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

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) What are two reasons why you must run IBGP on all routers within a transit backbone? (Choose two.)

A) so routers can properly forward packets toward all external destinations

B) to ensure that a full mesh exists between 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

Q2) 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.)

A) to provide routing information needed to establish the IBGP sessions

B) to resolve next-hop references 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

Q3) What is the AD of the following protocols? (Fill in the blanks.)

A) IBGP _______

B) EBGP _______

C) OSPF _______

D) IS-IS _______

E) RIP _______

Q4) What are two reasons why the AD is an important consideration for BGP network design? (Choose two.)

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) AD is not a large concern to BGP design, because the router will always choose the route advertised by the protocol best suited to reach the destination.

Q5) 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.)

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 prior to any BGP packets arriving at the router.

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Configuring a Transit Autonomous System

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.

Importance

Configuring a BGP network in a transit services configuration requires special care to ensure consistency of routing information throughout the AS. Understanding the configuration requirements for a transit backbone is crucial to successfully implementing the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS® commands required to configure IBGP neighbors
- Identify the Cisco IOS commands required to configure IBGP sessions between loopback interfaces
- Identify the Cisco IOS commands required to configure BGP synchronization to ensure successful IBGP operation of the transit AS

- Identify the Cisco IOS commands required to change the AD of BGP routes
- Identify the scalability limitations of IBGP-based backbones

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

Outline

This lesson includes these topics:

- Overview
- Configuring IBGP Neighbors
- Configuring IBGP Sessions Between Loopback Interfaces
- Configuring BGP Synchronization
- Changing the Administrative Distance of BGP Routes
- Scalability Limitations of IBGP-Based Transit Backbones
- Summary
- Assessment (Quiz): Configuring a Transit Autonomous System

Configuring IBGP Neighbors

This topic lists the Cisco IOS commands required to configure IBGP neighbors in an AS.

Configuring IBGP Neighbors

Cisco.com

router(config-router)#

```
neighbor ip-address remote-as as-number
```

- **Configures BGP neighbor**
- **The AS number configured determines whether the session is an EBGP session (neighbor AS is different from local AS) or IBGP session (same AS number)**

router(config-router)#

```
neighbor ip-address description text
```

- **Attaches optional description to a neighbor**

Configuring an IBGP neighbor is a simple task:

- Configure a BGP neighbor, specifying the same AS number.
- Optionally, attach a description to the neighbor to help in documentation and troubleshooting efforts.

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

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

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

Practice

Q1) Which Cisco IOS command do you use to configure a description on a BGP session?

A) **neighbor** *ip-address* **remote-as** *number* **description** *text*

B) **description** *text*

C) **neighbor** *ip-address* **description** *text*

D) **ip bgp description** *text*

Configuring IBGP Sessions Between Loopback Interfaces

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

Configuring IBGP Sessions Between Loopback Interfaces

Cisco.com

```
router(config-router)#
neighbor ip-address update-source interface
```

- **Configures the source interface for the TCP session that carries BGP traffic**
- **For IBGP sessions, the source interface shall be a loopback address**
- **Source address configured on one peering router must match the destination address configured on the other—BGP session will not start otherwise**
- **Make sure that your loopback interfaces are announced in the backbone IGP**

BGP v3.0—5-5

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 will be brought down so only one remains. The TCP packets will have the destination IP address as configured with the **neighbor** command. But they must also have a source IP address assigned. If no update source is configured, the router will set 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 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 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 **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

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

Practice

Q1) Which Cisco IOS command do you use to configure a BGP session between loopback interfaces?

A) **neighbor** *ip-address* **remote-as** *number* **update-source** *interface*

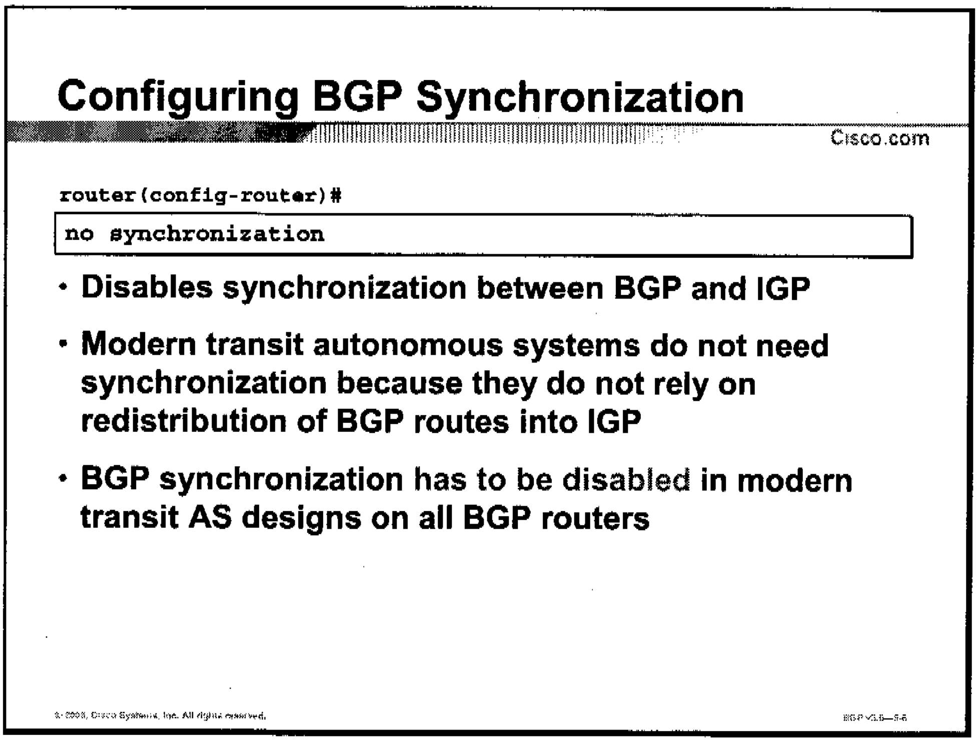
B) **ip bgp source-interface** *interface*

C) **neighbor** *ip-address* **update-source** *interface*

D) **ip bgp update-source** *interface*

Configuring BGP Synchronization

This topic presents the Cisco IOS command 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 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 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 (disable synchronization between BGP and your IGP), use the **no** form of the **synchronization** command.

no synchronization

Syntax Description

This command has no arguments or keywords.

Practice

Q1) What BGP parameter do you need to disable for proper IBGP operation?

A) redistribution

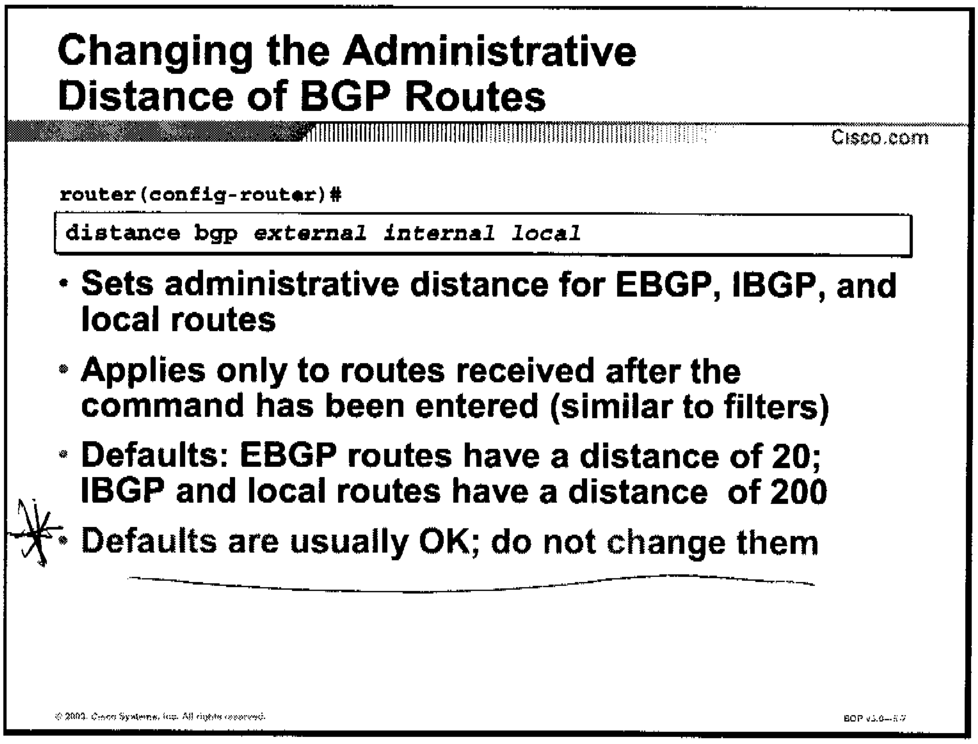
B) AS-path propagation

C) synchronization

D) split horizon

Changing the Administrative Distance of BGP Routes

This topic presents the Cisco IOS command 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

Syntax 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 those 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 those networks listed with a network router configuration command, often as backdoors (BGP backdoor 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

Practice

Q1) How can you change the AD of BGP routes?

A) through AS-path prepending

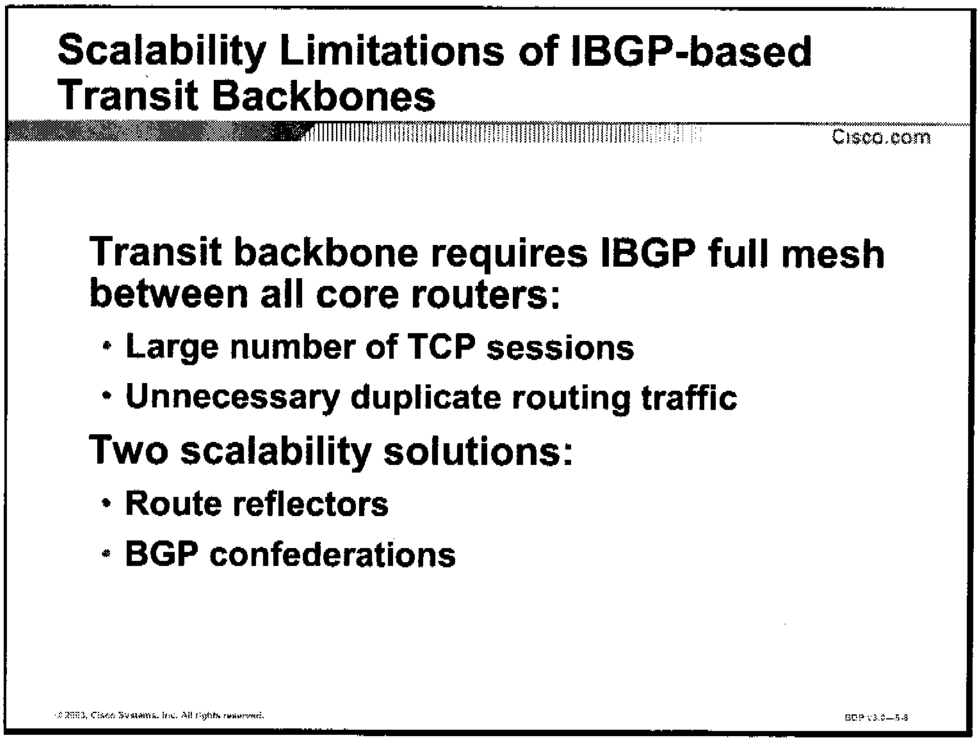
B) with the **distance bgp** router configuration command

C) by redistributing BGP routes into an IGP

D) by disabling BGP synchronization on the router

Scalability Limitations of IBGP-Based Transit Backbones

This topic identifies 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 session between "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.

Since 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 different 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 between themselves using intraconfederation EBGP sessions.

Practice

Q1) What are the scalability limitations of an IBGP-based transit AS?

A) The transit backbone can have only two exit points.

B) All EBGP speakers must have a BGP session to all BGP-speaking routers in the transit AS.

C) The transit area is limited to a maximum distance of 100 routers.

D) A full mesh of IBGP sessions is required due to BGP split-horizon rules.

Q2) What two BGP tools can you use to overcome IBGP scalability issues in a transit AS? (Choose two.)

A) disabling synchronization

B) route reflectors

C) BGP confederations

D) BGP peer groups

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- To configure an IBGP neighbor, use the neighbor command, specifying a remote AS number matching the AS number of the local router.
- When you configure IBGP sessions between loopback interfaces, the interfaces must be announced in the backbone IGP.
- You should disable BGP synchronization in all modern transit AS designs on all BGP routers.
- Although you can change the administrative distances of BGP routes, you typically should not change the default settings for EBGP (20) and IBGP (200).
- The full-mesh IBGP requirement in the transit AS creates scalability issues in the number of TCP sessions and unnecessary, duplicate routing traffic. IBGP scalability solutions to these issues exist.

Next Steps

After completing this lesson, go to:

- Monitoring and Troubleshooting IBGP in Transit Autonomous Systems lesson

References

For additional information, refer to these resources:

- For more information on transit autonomous systems, refer to "BGP Case Studies Section 1" at the following URL: http://www.cisco.com/warp/public/459/13.html
- For more information on BGP solutions for scaling IBGP, refer to "BGP Case Studies Section 4" at the following URL: http://www.cisco.com/warp/public/459/bgp-toc.html#case4

Quiz: Configuring a Transit Autonomous System

Complete the quiz to assess what you have learned in this lesson.

Objectives

This quiz tests your knowledge on how to:

- Identify the Cisco IOS commands required to configure IBGP neighbors
- Identify the Cisco IOS commands required to configure IBGP sessions between loopback interfaces
- Identify the Cisco IOS commands required to configure BGP synchronization to ensure successful IBGP operation of the transit AS
- Identify the Cisco IOS commands required to change the AD of BGP routes
- Identify the scalability limitations of IBGP-based backbones

Instructions

Complete these steps:

Step 1 Answer all questions in this quiz by selecting the best answer(s) to each question.

Step 2 Verify your results against the answer key located in the course appendices.

Step 3 Review the topics in this lesson matching questions with an incorrect answer choice.

Q1) When you are configuring the BGP neighbor session, what differentiates an EBGP neighbor from an IBGP neighbor?

A) The keyword **internal** at the end of the **neighbor** command.

B) IBGP neighbors will have the same AS number specified.

C) A description for the neighbor must be attached with the **neighbor description** command.

D) Directly connected neighbors will automatically form an EBGP session.

Q2) What two steps are required to use a loopback interface for IBGP peering sessions? (Choose two.)

A) Ensure that the loopback interfaces are reachable through an IGP.

B) Ensure that the two neighbors must be directly attached.

C) Verify that each router has multiple physically redundant paths.

D) Configure a **neighbor** statement with the **update-source** command.

Q3) Why is it important to disable BGP synchronization in a transit backbone?

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 in order for BGP routing across the transit backbone to function properly, creating too much overhead.

Q4) What are two negative ramifications of the full-mesh requirement imposed by IBGP? (Choose two.)

A) administratively difficult to apply an AS-wide routing policy

B) requires the use of next-hop-self for proper routing to external destinations

C) large number of TCP sessions

D) unnecessary duplication of routing traffic

Q5) What are two scalability tools that you can use to overcome the full-mesh requirement for IBGP sessions? (Choose two.)

A) confederations

B) floating static routes

C) route reflectors

D) disabling BGP synchronization

Scoring

You have successfully completed the quiz for this lesson when you earn a score of 80 percent or better.

Monitoring and Troubleshooting IBGP in Transit Autonomous Systems

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 encountered in a transit backbone. Common problems with IBGP, as discussed in this lesson, occur when IBGP sessions do not reach the established state, routing information received via IBGP is never selected, and the best BGP route is never installed in the forwarding table.

Importance

Configuring a BGP network in a transit services configuration requires special care to ensure consistency of routing information throughout the autonomous system (AS). Understanding the tools and techniques to monitor and troubleshoot problems in the transit backbone is crucial to successfully implementing and maintaining the transit autonomous network.

Objectives

Upon completing this lesson, you will be able to:

- Identify the Cisco IOS® commands required to monitor IBGP operation
- Identify common IBGP configuration problems
- Troubleshoot IBGP session startup issues

- Troubleshoot IBGP route selection issues
- Troubleshoot IBGP synchronization issues

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- BGP Overview module

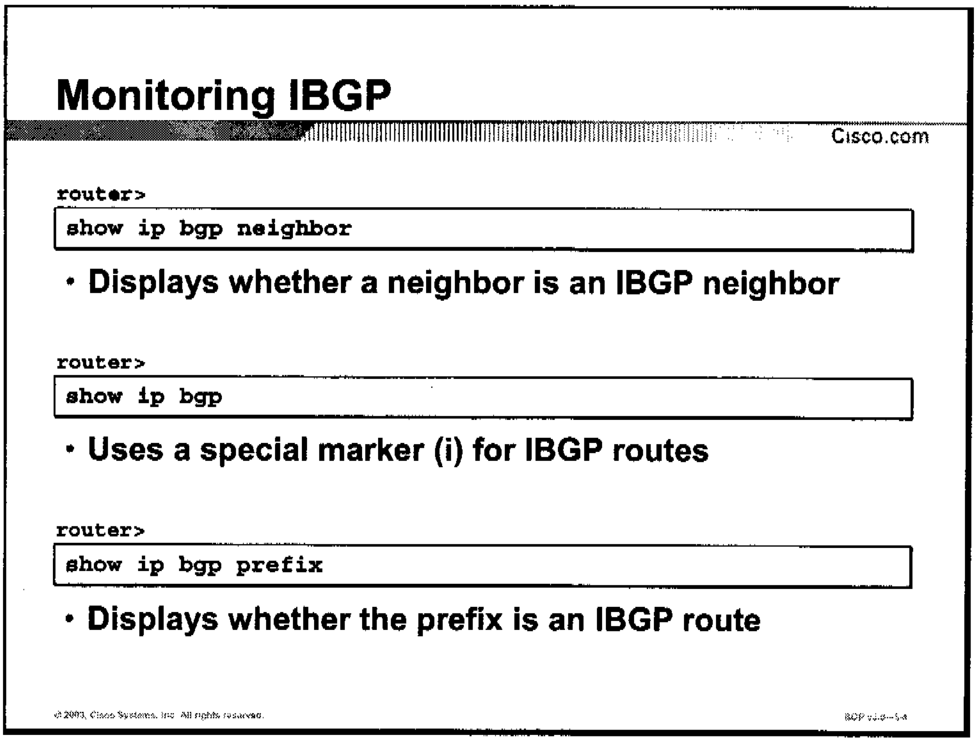
Outline

This lesson includes these topics:

- Overview
- Monitoring IBGP
- Common IBGP Problems
- Troubleshooting IBGP Session Startup Issues
- Troubleshooting IBGP Route Selection Issues
- Troubleshooting IBGP Synchronization Issues
- Summary
- Assessment (Lab): BGP Transit Autonomous Systems

Monitoring IBGP

This topic lists the Cisco IOS commands 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

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 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 received.

dampened-routes	(Optional) Displays the dampened routes to the neighbor at the IP address 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

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

Example

Monitoring IBGP (Cont.)

Cisco.com

```
router# show ip bgp neighbor 192.168.3.101
BGP neighbor is 192.168.3.101,  remote AS 3, internal link
  BGP version 4, remote router ID 192.168.3.101
  BGP state = Established, up for 00:56:08
  Last read 00:00:08, hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received
    Address family IPv4 Unicast: advertised and received
  Received 82 messages, 0 notifications, 0 in queue
  Sent 97 messages, 0 notifications, 0 in queue
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 5 seconds
```


The **show ip bgp neighbors** command displays whether a router is running an IBGP or EBGP session with a BGP neighbor. The indication is given by the "internal link" phrase (marked yellow in the second line of the figure).

Monitoring IBGP (Cont.)

Cisco.com

```
router# show ip bgp 197.99.1.0
BGP routing table entry for 197.99.1.0/24, version 3
Paths: (1 available, best #1)
  Advertised to non peer-group peers:
  192.168.3.103
  99
    192.168.21.99 (metric 20) from 192.168.3.101 (192.168.3.101)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```


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

Practice

Q1) Which Cisco IOS **show** command indicates that a BGP route is an IBGP route?

A) **show ip route**

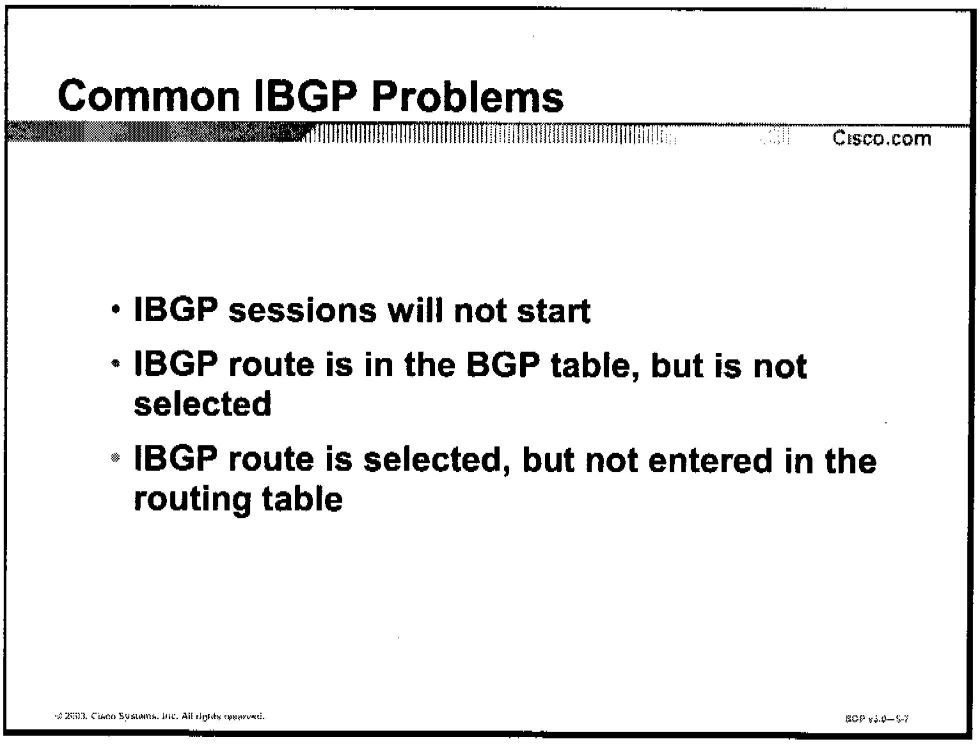
B) **show ip route bgp**

C) **show ip bgp**

D) **show ip bgp internal**

Common IBGP Problems

This topic identifies configuration problems 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. The next three topics give you troubleshooting advice to the most common problems:

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

Practice

Q1) What are the three common IBGP problems in transit backbones? (Choose three.)

A) IBGP sessions will not become established.

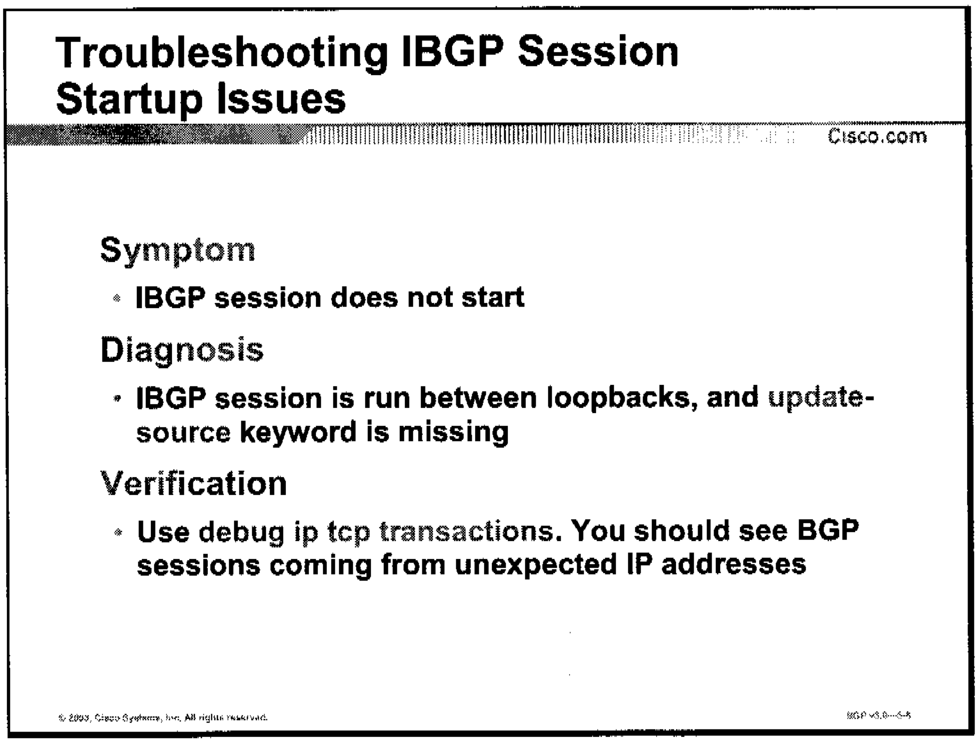
B) IBGP routes are never selected.

C) IBGP routes enter the "stuck in active state."

D) BGP routes are not installed into the IP 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 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 situation is 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.)

Cisco.com

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 internal gateway protocol (IGP) carries the subnets assigned to each of the loopback interfaces.

When verifying the 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.)

Cisco.com

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**

 BGP v3.0—5-10

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 succeed. But for better performance during recovery from network failure, both attempts should have the possibility to succeed. If both attempts do succeed, one of the connections will be brought down.

Practice

Q1) What are three common situations that prevent IBGP sessions from starting? (Choose three.)

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 control 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.

Troubleshooting IBGP Route Selection Issues

This topic describes how to troubleshoot IBGP route selection issues.

Troubleshooting IBGP Route Selection Issues

Cisco.com

Symptom

- **An IBGP route is in the BGP table but is never selected as the best route**

Diagnosis

- **BGP next hop is not reachable**

Verification

- Use **show ip bgp prefix** to find the BGP next hop
- Use **show ip route** to verify next-hop reachability

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, which refers to a next hop that is currently not reachable according to the forwarding table, will be saved in the BGP table, but it cannot be installed by the router into its forwarding 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 using the **show ip bgp** *prefix* command. The next hop is referred to as "inaccessible," if it is not currently reachable according to the forwarding table.

A common mistake is to forget to let the IGP announce the reachability of subnets that physically connect the local AS with a neighboring AS. These subnets are used by the router to establish the EBGP session, and the next hop 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.

Prevent this problem by including the subnet linking the transit AS to neighboring autonomous systems in the IGP using either the **redistribute connected** or the **network and passive-interface** configuration commands.

Practice

Q1) What would prevent IBGP routes from being selected as the best route in the BGP table?

A) Failure to disable BGP synchronization.

B) Failure to disable BGP split horizon.

C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

Troubleshooting IBGP Synchronization Issues

This topic describes how to troubleshoot IBGP synchronization issues.

Troubleshooting IBGP Synchronization Issues

Cisco.com

Symptom

- **An IBGP route is selected as the best route but not entered into the IP routing table**

Diagnosis

- **BGP synchronization is not disabled**

Verification

- **Disable BGP synchronization, clear the BGP sessions, and re-examine the IP routing table after the BGP table becomes stable**

BGP v3.0—5-12

In old BGP designs, redistribution between BGP and 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 the synchronization has to be turned off. However, the default value is to have synchronization enabled.

Routers with BGP synchronization enabled will not install IBGP routes in the forwarding table nor propagate them to other EBGP neighbors.

Fix this problem by configuring **no synchronization** in the router BGP configuration.

Practice

Q1) What common issue could prevent IBGP best routes from being inserted into the IP routing table?

A) Failure to disable BGP synchronization.

B) Failure to disable BGP split horizon.

C) The IGP has no route to the BGP next hop.

D) A default route has not been injected into the IGP.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- You can use the show ip bgp neighbor and show ip bgp prefix commands to monitor IBGP.
- Common IBGP configuration problems include a session not in the established state and issues with injecting routes into the IP routing table.
- One common session startup issue is to use the loopback as an IBGP peer without issuing the update-source configuration command. Another common session startup issue is the presence of a filter.
- 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.
- Routers with BGP synchronization enabled will not install IBGP routes in the forwarding table nor propagate them to other EBGP neighbors.

Next Steps

After completing this lesson, go to:

- Scaling Service Provider Networks module

References

For additional information, refer to these resources:

- For more information on troubleshooting IBGP, refer to "Troubleshooting IP Connectivity and Routing Problems" at the following URL: http://www.cisco.com/univercd/cc/td/doc/cisintwk/itg_v1/tr1907.htm#xtocid27
- For more information on configuring and monitoring IBGP, refer to "Configuring BGP" at the following URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm

Laboratory Exercise: BGP Transit Autonomous Systems

Complete the laboratory exercise to practice what you have learned in this module.

Required Resources

These are the resources and equipment required to complete this exercise:

Your workgroup requires the following components:

- Four Cisco 2610 routers with a WIC-1T and BGP-capable operating system software installed.
- Four CAB-X21FC + CAB-X21MT DTE-DCE serial cable combinations. The DCE side of the cable is connected to the Cisco 3660.
- Two Ethernet 10BASE-T patch cables.
- IBM PC (or compatible) with Windows 95/98 and an installed Ethernet adapter.

The lab backbone requires the following components (supporting up to eight workgroups):

- One Cisco 2610 router with a WIC-1T and BGP-capable operating system software installed
- Two Cisco 2610 routers with BGP-capable operating system software installed
- One Cisco 3640 router with an installed NM-8A/S
- Two Catalyst 2924M-XL Ethernet switches
- Three Ethernet 10BASE-T patch cables

Exercise Objective

In this exercise, you will enable the provider network to behave as a transit AS, given a typical service provider network with multiple BGP connections to other autonomous systems.

After completing this exercise, you will be able to:

- Plan the migration of an existing backbone toward a fully meshed IBGP backbone designed for transit traffic
- Configure IBGP sessions between loopback interfaces
- Configure BGP synchronization to ensure successful IBGP operation of the transit AS
- Monitor IBGP operation

Command List

The commands used in this exercise are described in the table here.

Table 1: Exercise Commands

Command	Description
router bgp *as-number*	Enter BGP configuration mode.
neighbor *ip-address* **remote-as** *as-number*	Establish an IBGP session by using your workgroup number as the AS number.
neighbor *ip-address* **update-source** *interface*	Use IP address of the specified interface' as the source address for the BGP session.
show ip bgp	Inspect the contents of the BGP table.
show ip bgp regexp *regexp*	Use a regular expression to filter the output of the **show ip bgp** command.
no synchronization	Disable synchronization of IGP and BGP routes.

Job Aids

These job aids are available to help you complete the laboratory exercise:

- With the rapid growth of the Internet, you decide to become an Internet service provider (ISP), and you already have your first customer. Unfortunately, the customer is willing to pay you only for connectivity toward your own network and toward AS 37.
- In this exercise, you will transform your network into a transit AS running BGP on all core routers.
- Start a BGP process on all routers in your workgroup. Configure a full mesh of IBGP sessions between all routers in your AS. Establish these BGP sessions between Loopback0 interfaces.

- Propagate only your own networks and networks originating in AS 37 to router "Client."
- Make sure that you accept only the networks originating in AS 99 from router "Client."
- AS 99 should receive only prefixes originating in your AS "x" and AS 37. You should accept only prefixes originating in AS 99 from router "Client."
- Figure 1 displays the required BGP connectivity within your AS as well as the BGP sessions with your customer and your upstream ISPs.

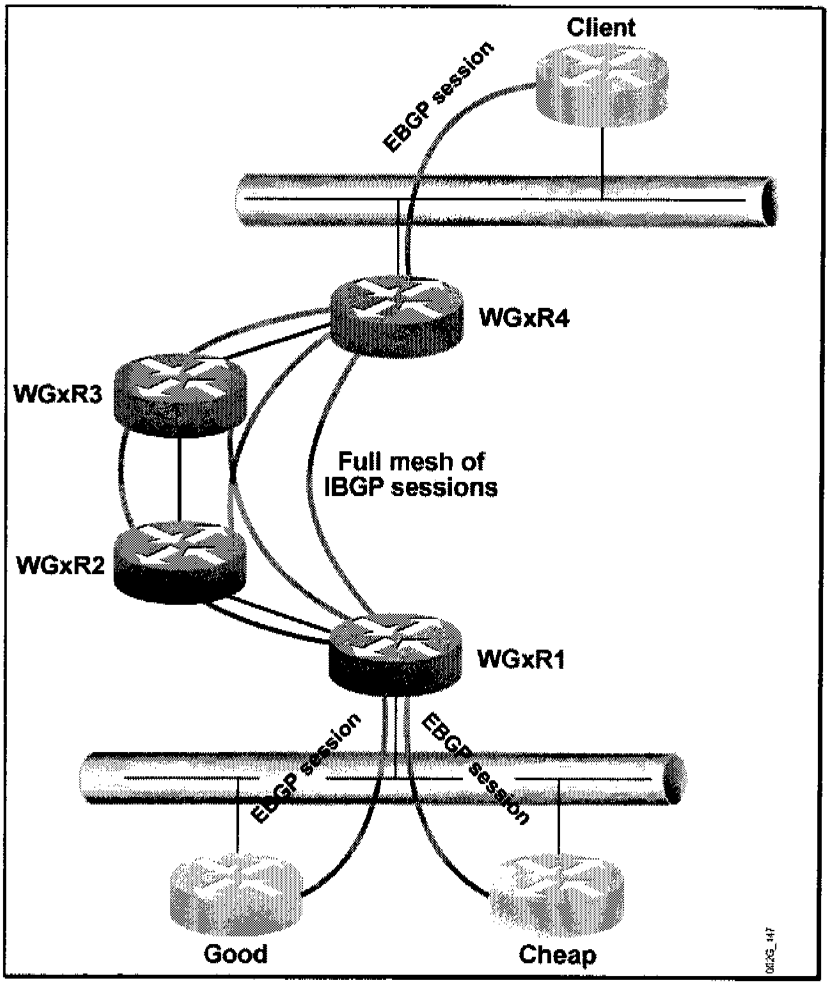


Figure 1: Creating a full mesh of IBGP sessions

Task 1: Configure the BGP Transit Autonomous System

In this task, you will configure your network backbone as a fully meshed IBGP backbone acting as a transit AS.

Exercise Procedure

Complete these steps:

Step 1 To clean up your BGP configuration, remove the BGP process and the default route from IGP on WGxR1.

Step 2 Remove the BGP process from WGxR2.

Step 3 Start the BGP process on all routers in your workgroup.

Step 4 Advertise your prefixes (197.x.0.0/16 and 192.168.x.0/24) in BGP on routers WGxR1 and WGxR4.

Step 5 Re-establish neighbor relationships to routers "Good" and "Cheap" without any filters on router WGxR1. Use the parameters from the following table:

Router	AS number	IP address
Good	20	192.168.20.20
Cheap	22	192.168.20.22

Step 6 Establish a BGP session with router "Client" on router WGxR4. Use the parameters from the following table:

Parameter	Value
Client IP address	192.168.21.99
Client AS number	99

Step 7 Configure all routers in your workgroup as IBGP neighbors (IBGP full mesh). Use loopback interfaces to establish these IBGP sessions.

Exercise Verification

You have completed this exercise when you attain these results:

- Check BGP on all core routers and the router "Client" and ensure that they have established the correct sessions with their peers.

```
WG1R1#sh ip bgp summary
...
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.20.20   4    20    1189    1200       52    0    0 13:34:45
23
192.168.20.22   4    22    1195    1183       52    0    0 13:34:46
23
197.1.2.1       4     1    1174    1196       52    0    0 13:34:58
0
```



```
197.1.4.1       4    1   1170   1188     52    0    0 13:35:17
0
197.1.6.1       4    1   1171   1183     52    0    0 13:34:58
13
```

```
WG1R2#sh ip bgp summary
...
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.1.1       4    1   1200   1179    125    0    0 13:38:09
24
197.1.4.1       4    1   1173   1173    125    0    0 13:38:31
0
197.1.6.1       4    1   1176   1170    125    0    0 13:37:58
13
```

```
WG1R3#sh ip bgp sum
...
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
197.1.1.1       4    1   1193   1175     78    0    0 13:40:33
24
197.1.2.1       4    1   1175   1175     78    0    0 13:40:37
0
197.1.6.1       4    1   1183   1175     78    0    0 13:40:04
13
```

```
WG1R4#sh ip bgp sum
...
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.21.99   4   99   1191   1192     14    0    0 13:40:38
11
197.1.1.1       4    1   1190   1178     14    0    0 13:41:04
24
197.1.2.1       4    1   1173   1179     14    0    0 13:40:54
0
197.1.4.1       4    1   1175   1183     14    0    0 13:40:54
0
```

```
Client#sh ip bgp sum
...
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.21.X    4    1   1147   1146     18    0    0 13:42:01
2
```

■ Check the BGP table on router "Client" and verify that it is correctly receiving BGP routes.

```
Client#sh ip bgp
BGP table version is 119, local router ID is 197.99.111.1
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
*> 10.0.0.0         192.168.21.1                           0 1 20 i
*> 99.0.0.0         0.0.0.0                  0         32768 i
*> 128.20.0.0       192.168.21.1                           0 1 22 i
*> 128.20.12.0/24   192.168.21.1                           0 1 20 i
*> 128.22.0.0       192.168.21.1                           0 1 22 i
*> 128.22.12.0/24   192.168.21.1                           0 1 22 i
*> 128.26.0.0       192.168.21.1                           0 1 22 26
i
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 128.42.0.0       192.168.21.1                           0 1 20 42
i
*> 128.51.0.0       192.168.21.1                           0 1 22 26
51 i
*> 128.213.0.0      192.168.21.1                           0 1 20 213
i
*> 128.214.0.0      192.168.21.1                           0 1 22 214
i
*> 192.20.11.0      192.168.21.1                           0 1 22 i
*> 192.20.12.0/30   192.168.21.1                           0 1 20 i
*> 192.22.11.0      192.168.21.1                           0 1 22 i
*> 192.22.12.0/30   192.168.21.1                           0 1 22 i
*> 192.26.11.0      192.168.21.1                           0 1 22 26
i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.42.11.0      192.168.21.1                           0 1 20 42
i
*> 192.51.11.0      192.168.21.1                           0 1 22 26
51 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 192.213.11.0     192.168.21.1                           0 1 20 213
i
*> 192.214.11.0     192.168.21.1                           0 1 22 214
i
*> 197.1.0.0/16     192.168.21.1             0             0 1 i
*> 197.99.1.0       0.0.0.0                  0         32768 i
*> 197.99.11.0      0.0.0.0                  0         32768 i
*> 197.99.12.0      0.0.0.0                  0         32768 i
*> 197.99.13.0      0.0.0.0                  0         32768 i
*> 197.99.20.0      0.0.0.0                  0         32768 i
*> 197.99.22.0      0.0.0.0                  0         32768 i
*> 197.99.80.0      0.0.0.0                  0         32768 i
*> 197.99.111.1/32  0.0.0.0                  0         32768 i
```



```
*> 197.99.120.0      0.0.0.0                  0         32768 i
*> 197.99.128.0/20   0.0.0.0                  0         32768 i
*> 200.20.0.0/16     192.168.21.1                           0 1 20 i
*> 200.22.0.0/16     192.168.21.1                           0 1 22 i
```

- Use traceroute from router "Client" to the loopback interface on router WGxR1 (197.x.1.1). You should see a path similar to the one below:

```
Client# traceroute 197.1.1.1
Type escape sequence to abort.
Tracing the route to 197.1.1.1

  1 192.168.21.1 4 msec 4 msec 4 msec
  2 192.168.1.9 [AS 1] 20 msec 16 msec 16 msec
  3 192.168.1.5 [AS 1] 32 msec 32 msec 28 msec
  4 192.168.1.1 [AS 1] 44 msec *  44 msec
```

- Use traceroute from router WGxR1 to the loopback interface on router "Client" (197.99.1.1). You should see a path similar to the one below:

```
WG1R1# traceroute 197.99.1.1
Type escape sequence to abort.
Tracing the route to 197.99.1.1

  1 192.168.1.2 16 msec 16 msec 17 msec
  2 192.168.1.6 32 msec 32 msec 28 msec
  3 192.168.1.10 44 msec 40 msec 40 msec
  4 192.168.21.99 48 msec *  44 msec
```

Answer these questions:

Q1) Check the BGP table on router "Client." How many prefixes coming from your AS are in that BGP table? ____________

Q2) Is there any other way of discovering how many prefixes you have advertised to the router "Client"? ________________________________

Task 2: Configure Filters in the BGP Transit Autonomous System

As the last steps in this exercise, you need to establish route filters toward your customer on WGxR4.

Exercise Procedure

Complete these steps:

Step 1 Create an AS-path filter to permit your own networks and networks originating in AS 37.

Step 2 Create an AS-path filter to permit networks originating in AS 99.

Step 3 Apply the AS-path filters to your customer.

Exercise Verification

You have completed this exercise when you attain these results:

- Check to see if router "Client" is receiving only your networks and those originating in AS 37.

```
Client#sh ip bgp
BGP table version is 206, local router ID is 197.99.111.1
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
*> 99.0.0.0         0.0.0.0                  0         32768 i
*> 128.37.0.0       192.168.21.1                           0 1 20 42
37 i
*> 192.37.11.0      192.168.21.1                           0 1 20 42
37 i
*> 192.168.1.0      192.168.21.1             0             0 1 i
*> 197.1.0.0/16     192.168.21.1             0             0 1 i
*> 197.99.1.0       0.0.0.0                  0         32768 i
*> 197.99.11.0      0.0.0.0                  0         32768 i
*> 197.99.12.0      0.0.0.0                  0         32768 i
*> 197.99.13.0      0.0.0.0                  0         32768 i
*> 197.99.20.0      0.0.0.0                  0         32768 i
*> 197.99.22.0      0.0.0.0                  0         32768 i
*> 197.99.80.0      0.0.0.0                  0         32768 i
*> 197.99.111.1/32  0.0.0.0                  0         32768 i
*> 197.99.120.0     0.0.0.0                  0         32768 i
*> 197.99.128.0/20  0.0.0.0                  0         32768 i
```


- Use the **show ip route ospf** command to make sure that your IGP carries only your internal networks.

```
WG1R4#sh ip route ospf
     197.1.8.0/32 is subnetted, 1 subnets
O       197.1.8.1 [110/193] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.1.0/32 is subnetted, 1 subnets
O       197.1.1.1 [110/193] via 192.168.1.9, 1d02h, Serial0/0.1
O    192.168.20.0/24 [110/202] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.3.0/32 is subnetted, 1 subnets
O       197.1.3.1 [110/129] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.2.0/32 is subnetted, 1 subnets
O       197.1.2.1 [110/129] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.5.0/32 is subnetted, 1 subnets
O       197.1.5.1 [110/65] via 192.168.1.9, 1d02h, Serial0/0.1
     197.1.4.0/32 is subnetted, 1 subnets
O       197.1.4.1 [110/65] via 192.168.1.9, 1d02h, Serial0/0.1
     192.168.1.0/24 is variably subnetted, 4 subnets, 2 masks
O       192.168.1.0/30 [110/192] via 192.168.1.9, 1d02h, Serial0/0.1
O       192.168.1.4/30 [110/128] via 192.168.1.9, 1d02h, Serial0/0.1
```

Answer these questions:

Q1) Why did you have to disable synchronization?

Q2) Why did you have to establish a full mesh of IBGP sessions?