

Configuring BGP on Cisco Routers

Student Guide

Volume 1

Version 4.0

Part Number: 97-3539-01



Americas Headquarters

Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters

Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters

Cisco Systems International BV
Amsterdam,
The Netherlands

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

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

DISCLAIMER WARRANTY: THIS CONTENT IS BEING PROVIDED "AS IS" AND AS SUCH MAY INCLUDE TYPOGRAPHICAL, GRAPHICS, OR FORMATTING ERRORS. CISCO MAKES AND YOU RECEIVE NO WARRANTIES IN CONNECTION WITH THE CONTENT PROVIDED HEREUNDER, EXPRESS, IMPLIED, STATUTORY OR IN ANY OTHER PROVISION OF THIS CONTENT OR COMMUNICATION BETWEEN CISCO AND YOU. CISCO SPECIFICALLY DISCLAIMS ALL IMPLIED WARRANTIES, INCLUDING WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE, OR ARISING FROM A COURSE OF DEALING, USAGE OR TRADE PRACTICE. This learning product may contain early release content, and while Cisco believes it to be accurate, it falls subject to the disclaimer above.

Welcome Students

Students, this letter describes important course evaluation access information!

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

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

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

Sincerely,
Cisco Systems Learning

© 2015 Cisco Systems, Inc.

Table of Contents

Course Introduction	1
Overview	1
Course Goal and Objectives	3
Course Flow	4
Your Training Curriculum	5
Additional References	8
Module 1: BGP Overview	9
<u>Lesson 1: Introducing BGP</u>	<u>11</u>
Interdomain Routing	12
Why External Routing Protocols?	14
BGP Characteristics	16
BGP Development Considerations	19
Single-Homed Customers	20
Multihomed Customers	22
Transit Autonomous Systems	24
BGP Limitations	25
Summary	26
<u>Lesson 2: Understanding BGP Path Attributes</u>	<u>27</u>
BGP Path Attributes	28
Well-Known BGP Attributes	29
Optional BGP Attributes	32
AS-Path Attribute	34
Next-Hop Attribute	36
Summary	40
<u>Lesson 3: Establishing BGP Sessions</u>	<u>41</u>
BGP Neighbor Discovery	42
Establishing a BGP Session	45
BGP Keepalives	48
MD5 Authentication	50
Summary	51
<u>Lesson 4: Processing BGP Routes</u>	<u>53</u>
Receiving Routing Updates	54
Building BGP Table	56
BGP Route Selection Criteria	57
BGP Route Propagation	59
Building IP Routing Table	60
Advertising Local Networks	61
Automatic Summarization	64
Summary	68
<u>Lesson 5: Configuring Basic BGP</u>	<u>69</u>
BGP Routing Process	70
Configuring External Neighbors	71
Discovery 1: Configure Basic BGP	72
Announcing Networks in BGP	81

Redistributing Routes into BGP	85
Discovery 2: Announcing Networks in BGP	86
BGP Conditional Route Injection.....	99
BGP Support for TTL Security Check.....	101
Discovery 3: Implement BGP TTL Security Check	102
Multihomed Customer Problem.....	106
Summary.....	108
Lesson 6: Monitoring and Troubleshooting BGP.....	109
Monitoring Overall BGP Routing	110
Monitoring BGP Neighbors	112
Monitoring BGP Table	113
Debugging BGP	115
BGP Session Startup Problems.....	118
BGP Neighbor Not Reachable	119
BGP Neighbor Not Configured.....	122
BGP AS Number Mismatch	124
Summary.....	126
Lesson 7: Module Summary	127
References.....	127
Lesson 8: Module Self-Check	129
Module Self-Check Answers.....	141
Answer Key.....	141
Module 2: BGP Transit Autonomous Systems.....	145
Lesson 1: Working with Transit AS	147
Transit AS Tasks.....	148
External Route Propagation.....	149
Internal Route Propagation.....	150
Packet Forwarding in AS	151
Core Router IBGP Requirements in Transit AS.....	152
Discovery 4: BGP Route Propagation	153
Summary.....	160
Lesson 2: Interacting with IBGP and EBGp in Transit AS	161
AS-Path Processing in IBGP	162
Multipath Load Sharing in BGP	163
BGP Split Horizon	165
IBGP Full Mesh.....	166
IBGP Neighbors.....	168
IBGP Next-Hop Processing.....	170
Discovery 5: IBGP Full Mesh	172
Transit Network Using Edge Routers as Next Hops Example	187
Differences Between EBGp and IBGP Sessions.....	188
Scalability Limitations of IBGP-Based Transit Backbones	190
Summary.....	191
Lesson 3: Forwarding Packets in Transit AS.....	193
Packet Forwarding in Transit AS	194
Recursive Lookup in Cisco IOS Software.....	197
Routing Protocols in Transit AS	199
BGP and IGP Interaction	201
Discovery 6: BGP Administrative Distance	203

Problems with BGP and IGP Interaction	213
Summary.....	214
<u>Lesson 4: Monitoring and Troubleshooting IBGP in Transit AS</u>	<u>215</u>
Monitoring IBGP	216
Common IBGP Problems.....	218
Troubleshooting IBGP Session Startup Issues	219
Troubleshooting IBGP Route Selection Issues	222
Troubleshooting IBGP Synchronization Issues.....	223
Summary.....	224
<u>Lesson 5: Module Summary</u>	<u>225</u>
References.....	225
<u>Lesson 6: Module Self-Check</u>	<u>227</u>
<u>Module Self-Check Answers.....</u>	<u>235</u>
Answer Key	235
<u>Module 3: Route Selection Using Policy Controls.....</u>	<u>237</u>
<u>Lesson 1: Using Multihomed BGP Networks</u>	<u>239</u>
Business Requirements for Multihomed BGP Networks	240
Technical Requirements for Multihomed BGP Networks	241
BGP Route Selection Without BGP Policies	242
Multihomed Customer Routing Policies	244
Influencing BGP Route Selection	245
Transit Traffic Issue	247
Routing Update Reliability Issue	248
Return Traffic Issue.....	249
Summary.....	250
<u>Lesson 2: Employing AS Path Filters.....</u>	<u>251</u>
AS Path Filtering Scenarios	252
AS Path Regular Expressions	254
String Matching	255
Applying AS Path Filters	264
Configuring BGP AS Path Filters	265
Discovery 7: Configure Non-Transit Autonomous System	268
Monitoring AS Path Filters	275
Summary.....	278
<u>Lesson 3: Filtering with Prefix Lists</u>	<u>279</u>
Requirements for Prefix-Based Filters.....	280
Prefix Lists vs. IP Access Lists.....	281
Configuring Prefix Lists	284
BGP Filters Implementation.....	288
Implementing Prefix Lists in the BGP Process	289
Discovery 8: Filtering Customer Prefixes	293
Modifying Prefix Lists.....	300
Monitoring Prefix Lists.....	301
Summary.....	304
<u>Lesson 4: Using Outbound Route Filtering.....</u>	<u>305</u>
Outbound Route Filtering	306
Inbound vs. Outbound Filtering Example	307
BGP Prefix-Based Outbound Route Filtering.....	308
Outbound Route Filter Message	309

Configuring Outbound Route Filtering	312
Using Outbound Route Filtering	315
Discovery 9: Prefix-Based Outbound Route Filtering	316
Summary	324
Lesson 5: Applying Route Maps as BGP Filters	325
Route Map Overview	326
BGP Route-Map Policy List Support	330
BGP Route Map Continue.....	334
Prefix List Use in Route Maps.....	338
BGP Filters	340
Using Route Maps as BGP Filters.....	341
Discovery 10: Configure Route Maps as BGP Filters.....	343
Summary	354
Lesson 6: Implementing Changes in BGP Policy	355
Traditional Filtering Limitations.....	356
BGP Soft Reset Enhancement	358
Route Refresh.....	360
Configuring Route Refresh	363
Monitoring Route Refresh.....	365
Summary.....	367
Lesson 7: Module Summary	369
References.....	370
Lesson 8: Module Self-Check	373
Module Self-Check Answers.....	383
Answer Key.....	383
Glossary.....	385

Course Introduction

Overview

In this course you will learn how to design, implement, and verify a BGP in the customer and service provider networks.

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

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

Learner Skills and Knowledge

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

Learner Skills and Knowledge

- Intermediate to advanced knowledge of Cisco IOS Software configuration
- Configuring and troubleshooting RIP, EIGRP, OSPF and IS-IS
- Skills and knowledge equivalent to those learned in:
 - *Interconnecting Cisco Networking Devices v2.0, Part 1 (ICND1 v2.0) and Part 2 (ICND2 v2.0), or*
 - *Interconnecting Cisco Networking Devices: Accelerated Version 2.0 (CCNAX v2.0)*
 - *Implementing Cisco IP Routing (ROUTE v2.0)*
 - *Building Cisco Service Provider Next-Generation Networks Part 1 (SPNGN1) v1.2*
 - *Building Cisco Service Provider Next-Generation Networks Part 2 (SPNGN2) v1.2*

Course Goal and Objectives

This topic describes the course goal and objectives.

Course Goal

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

- Describe how to configure, monitor, and troubleshoot basic BGP to enable interdomain routing in a network scenario with multiple domains
- Describe how to use BGP policy controls to influence the BGP route selection process in a network scenario in which you must support connections to multiple ISPs
- Describe how to use BGP attributes to influence the route selection process in a network scenario where you must support multiple connections.
- Describe how to successfully connect the customer network to the Internet in a network scenario in which multiple connections must be implemented

© 2015 Cisco and/or its affiliates. All rights reserved.

2

Course Goal (Cont.)

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

- Describe how to configure the service provider network to behave as a transit AS in a typical implementation with multiple BGP connections to other autonomous systems.
- Enable route reflection as possible solution to BGP scaling issues in a typical service provider network with multiple BGP connections to other autonomous systems
- Describe the available BGP tools and features to optimize the scalability of the BGP routing protocol in a typical BGP network

© 2015 Cisco and/or its affiliates. All rights reserved.

3

The goal of the course is to provide learners with in-depth knowledge of BGP.

Course Flow

This topic presents the suggested flow of the course materials.

Course Flow		
	AM	PM
Day 1	Course Intro Module 1: BGP Overview	Module 1 (Cont.)
Day 2	Module 2: BGP Transit Autonomous Systems	Module 3: Route Selection Using Policy Controls
Day 3	Module 3 (Cont.)	Module 4: Route Selection Using Attributes
Day 4	Module 4 (Cont.) Module 5: Customer-to-Provider Connectivity with BGP	Module 5 (Cont.)
Day 5	Module 6: Scaling Service Provider Networks Module 7: Optimizing BGP Scalability	Module 7 (Cont.)

© 2015 Cisco and/or its affiliates. All rights reserved.

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

Your Training Curriculum

This topic presents the training curriculum for this course.



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

Training Curriculum

This subtopic presents the learning center information.



Additional information is available at:

- https://learningnetwork.cisco.com/community/learning_center
- <https://learningnetwork.cisco.com/community/certifications/ccna>

Learner Introductions

- Your name
- Your company
- Job responsibilities
- Skills and knowledge
- Brief history
- Objective

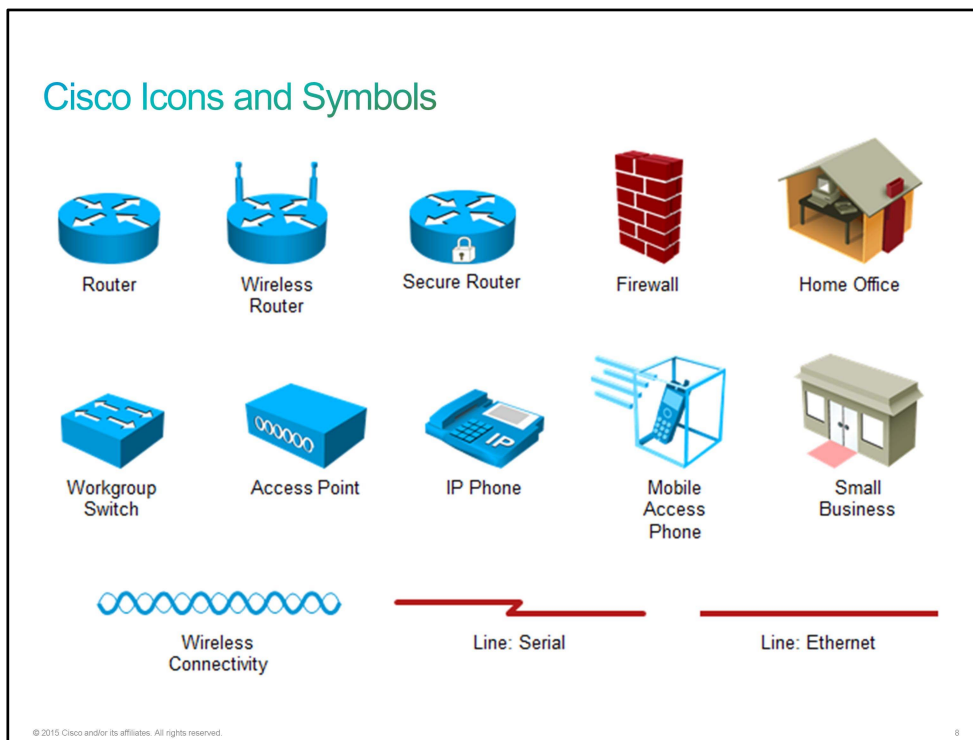


© 2015 Cisco and/or its affiliates. All rights reserved.

7

Additional References

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



Cisco Glossary of Terms

For additional information on Cisco terminology, refer to the Cisco Internetworking Terms and Acronyms glossary of terms at [http://docwiki.cisco.com/wiki/Category:Internetworking_Terms_and_Acronyms_\(ITA\)](http://docwiki.cisco.com/wiki/Category:Internetworking_Terms_and_Acronyms_(ITA)).

Module 1: BGP Overview

Introduction

[BGP](#) is an interdomain (interautonomous system) routing protocol that is used to exchange routing information for the Internet. BGP, by design, is a very robust and scalable routing protocol. Because BGP is deployed as an interdomain routing protocol, it has many rich features that allow you to implement a variety of routing policies.

In this module you will learn the basic BGP technology, details about BGP session establishment and routing information exchange. You will also learn the basic Cisco IOS BGP configuration and troubleshooting tasks.

Upon completing this module, you will be able to:

- Defines how to identify appropriate BGP use and limitations.
- Describe BGP path attributes and the functionality of each attribute.
- Describe BGP neighbors and neighbor session establishment procedures.
- Describe BGP route processing and BGP routing updates
- Describe how to configure basic BGP
- Describe how to perform the steps to correct basic BGP configuration and session errors

Lesson 1: Introducing BGP

Overview

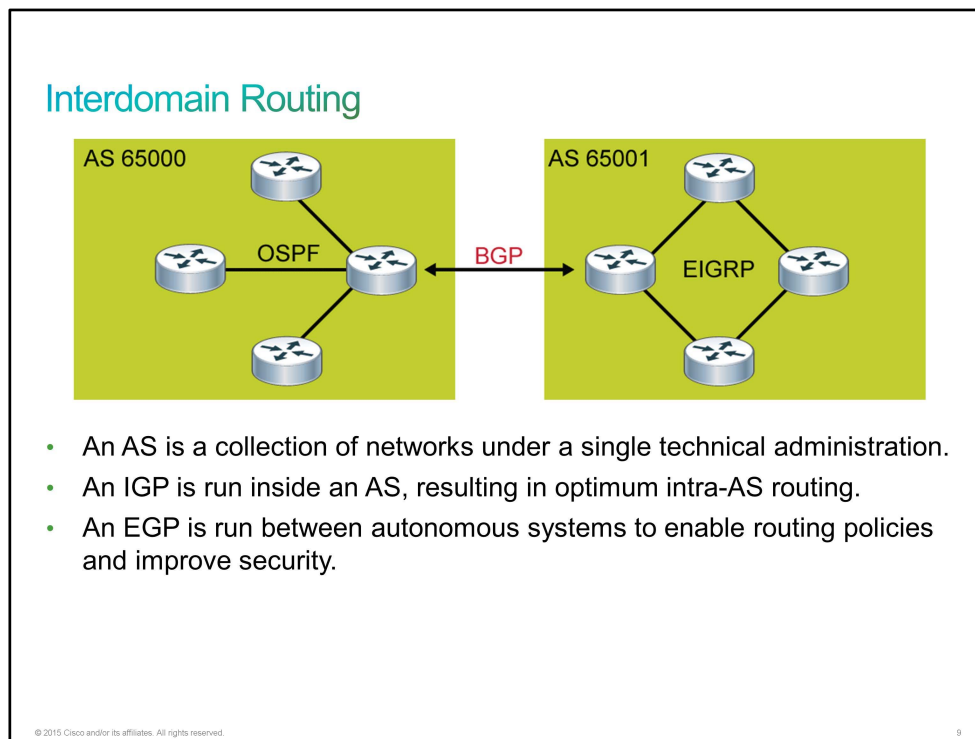
The [BGP](#) is a very robust and scalable routing protocol. You can see that from the fact that it is the routing protocol that is used on the Internet. Service providers and customer networks, such as universities and corporations, usually use an [IGP](#) such as [EIGRP](#), [OSPF](#), or [IS-IS](#) for the exchange of routing information within their networks. Any communication between these IGP's and the Internet or between service providers will be accomplished through BGP. This lesson introduces basic BGP characteristics and features.

Upon completing this lesson, you will be able to:

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

Interdomain Routing

When talking to people who are involved with Internet routing, network administrators commonly use the terms "autonomous system," "interdomain routing," "interior routing protocol," and "exterior routing protocol."



These terms, which may be confusing for a novice, are defined as follows:

- An [AS](#) is a collection of networks under a single technical administration. Other definitions refer to a collection of routers or [IP](#) prefixes, but in the end they are all essentially the same thing. The important principle is the technical administration, which means sharing the same routing protocol and routing policy. Legal and administrative ownership of the routers does not matter with autonomous systems. Autonomous systems are identified by AS numbers. AS numbers are 16-bit, unsigned integers ranging from 1 to 65,535. Public AS numbers (1 to 64,511) are assigned and managed by an Internet registry. A range of private AS numbers (64,512 to 65,535) has been reserved for customers that need an AS number to run [BGP](#) in their private networks. New 32-bit AS numbers were created when the AS number pool from [IANA](#) approached exhaustion.
- Interdomain routing is routing between autonomous systems. It is usually based on a set of policies, not just the technical characteristics of the underlying infrastructure.
- Exterior routing protocols (BGP being the only exterior routing protocol that is used today) are protocols that have the right set of functions to support various interdomain routing policies. Such protocols are contrary to interior routing protocols (for example, [OSPF](#), [IS-IS](#), or [EIGRP](#)). Interior routing protocols focus only on finding the optimum (usually fastest) route between two points, without respect to routing policies.

Design Goals for Interdomain Routing

Design Goals for Interdomain Routing

Scalability

- The Internet has more than 600,000 aggregated routes and is still growing.

Secure routing information exchange

- Routers from another AS cannot be trusted.
- Tight filters are required; authentication is desirable.

Support for routing policies

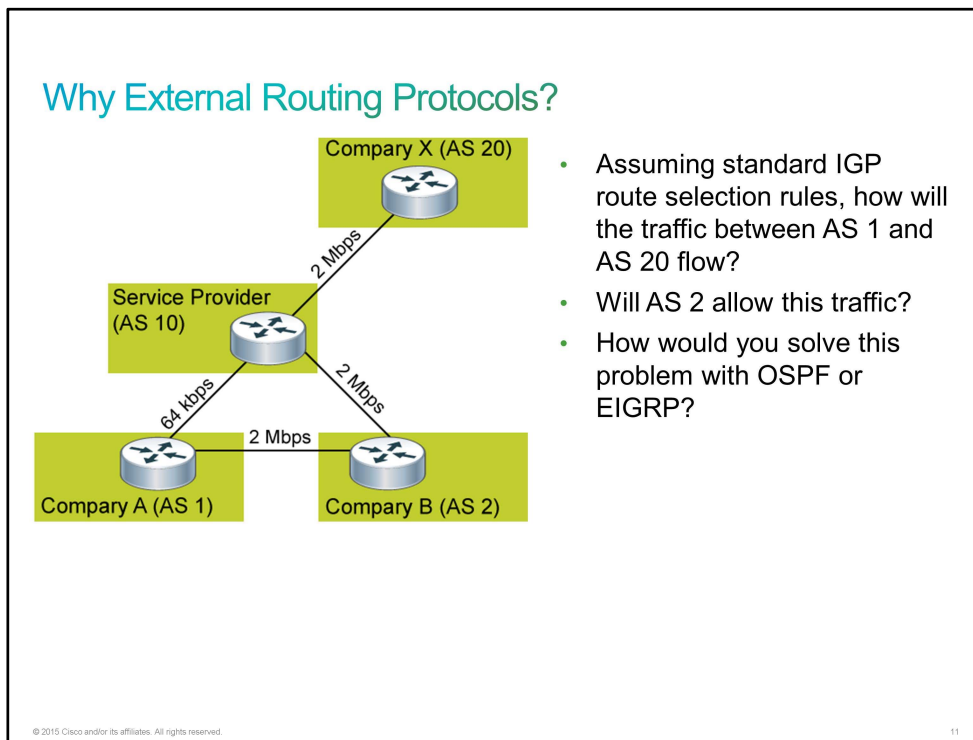
- Routing between autonomous systems might not always follow the optimum path.
- Exterior routing protocols have to support a wide range of customer requirements.

© 2015 Cisco and/or its affiliates. All rights reserved.

10

Why External Routing Protocols?

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



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

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

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

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

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

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

BGP Characteristics

[BGP](#) is a distance vector protocol. So, BGP announces to its neighbors those [IP](#) networks that it can reach itself. The receivers of that information say, "If that [AS](#) can reach those networks, then I can reach them via the AS."

BGP Characteristics

- BGP is a distance vector protocol with enhancements:
 - Reliable updates
 - Triggered updates only
 - Rich metrics (called path attributes)
- BGP is designed to scale to huge internetworks

© 2015 Cisco and/or its affiliates. All rights reserved.

12

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

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

BGP enables reliable information exchange and is capable of batching the routing updates. These two characteristics allow BGP to scale to large, Internet-sized networks.

BGP Characteristics (Cont.)

Reliable updates:

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

© 2015 Cisco and/or its affiliates. All rights reserved.

13

The reliable transport mechanism that BGP uses is standard [TCP](#). BGP is an application protocol that uses both the TCP and IP protocols for reliable connections.

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

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

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

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

BGP Characteristics (Cont.)

Common BGP uses:

- Customers that are connected to more than one service provider
- Service provider networks (transit autonomous systems)
- Service providers exchanging traffic at an exchange point (CIX, GIX, NAP, ...).
 - NAP defines the exchange points between region and core.
 - CIX or GIX points define international exchange points.
- Network cores of large-enterprise customers

NAP—Network Access Point, CIX—Commercial Internet eXchange, GIX—Global Internet eXchange

© 2015 Cisco and/or its affiliates. All rights reserved.

14

BGP Development Considerations

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

BGP Development Considerations

- BGP was designed to perform well in the following areas:
 - Interdomain routing applications
 - Huge internetworks with large routing tables
 - Environments that require complex routing policies
- Some design tradeoffs were made:
 - BGP uses TCP for reliable transport—CPU-intensive
 - Scalability is the top priority—slower convergence

© 2015 Cisco and/or its affiliates. All rights reserved.

15

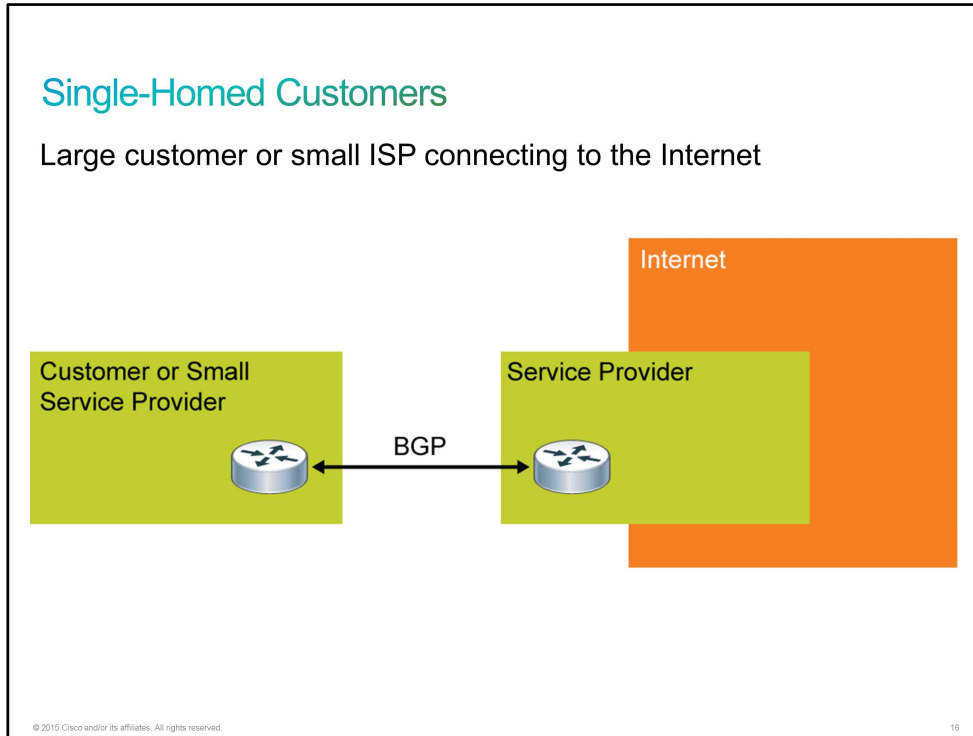
TCP, a well-known and widely proven protocol, was chosen as the transport mechanism. That decision kept BGP simple, but it increased the CPU resource requirements for routers running BGP. The point-to-point nature of TCP also introduces a slight increase in network traffic. This increase occurs because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

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

Note	BGP convergence times can be modified with the configuration of nondefault values for BGP scan and advertisement timers.
-------------	--

Single-Homed Customers

Normal Internet access to a single [ISP](#) does not require [BGP](#); static routes are more commonly used to handle this situation. Small ISPs that buy Internet connectivity from other ISPs use this type of connectivity more often. Especially if they want to start their business the proper way—by using their own [AS](#) number and having their own address space.



The figure shows a customer network that is connected to the Internet using a single ISP, but such a scenario is generally not the case when BGP is used.

Use Guidelines—Single-Homed Customers

Use Guidelines—Single-Homed Customers

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

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

Use static routes in all other cases:

- Static routes always simpler than BGP

© 2015 Cisco and/or its affiliates. All rights reserved.

17

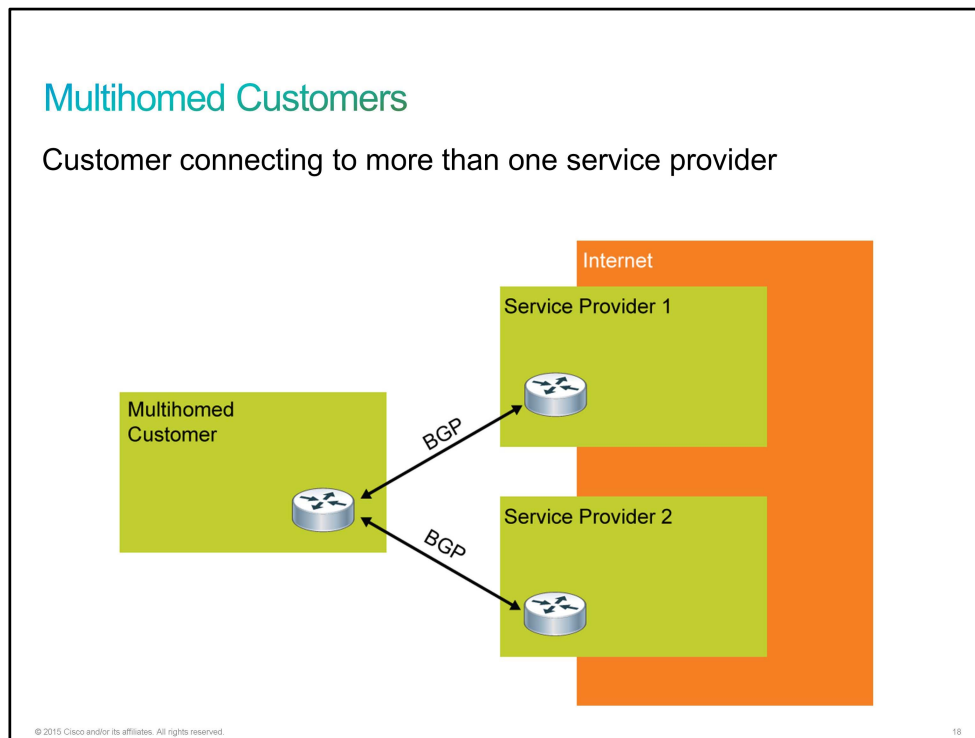
Under certain conditions, BGP must be configured between the customer and the service provider. For example, BGP is needed when customers are multihomed to the same service provider, that is, the customer networks have multiple links connecting them with the service provider network. Such customers require dynamic routing protocol interaction with the service provider to detect link failures. Private AS numbers (AS numbers between 64,512 and 65,534 or between 4,200,000,000—4,294,967,294) are usually implemented in BGP configurations for these customers.

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

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

Multihomed Customers

The [BGP](#) is appropriate for the multihomed customer. The customer must have its own officially assigned [AS](#) number. The customer is also responsible for announcing its own [IP](#) networks to both [ISPs](#). Both ISPs forward all routes that they receive from the Internet to the customer network. The customer should avoid forwarding any routing information that they receive from one ISP to the other. Otherwise, the customer becomes a transit provider between the two ISPs. Most customers like to avoid this situation because it creates a resource drain on routers and network links.



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

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

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

User Guidelines—Multihomed Customers

User Guidelines—Multihomed Customers

- BGP is almost mandatory for multihomed customers.
- Multihomed customers have to use public AS numbers.
- Multihomed customers should use a provider-independent address space.

© 2015 Cisco and/or its affiliates. All rights reserved.

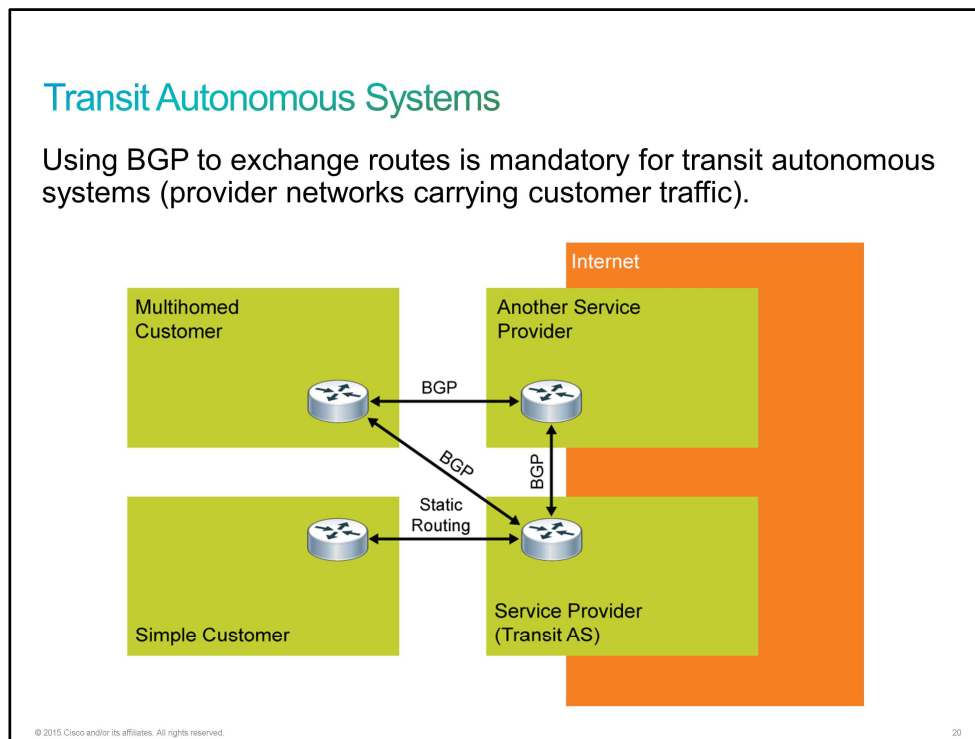
19

The following use guidelines apply to multihomed customers:

- Although there are designs where BGP could be avoided, most multihomed customers need to use BGP with their service providers.
- Multihomed customers must have their own AS numbers. It is recommended to use a public AS number (between 1 and 23.455, between 23.457 and 64.534, or between 131.072 and 4.199.999.999).
- Multihomed customers should use a provider-independent address space, which is allocated to them directly by an Internet registry.

Transit Autonomous Systems

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



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

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

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

BGP Limitations

BGP-enabled routers make forwarding decisions based on the destination [IP address](#) only. The source IP address does not affect the BGP decision.

BGP Limitations

BGP and associated tools cannot express all routing policies.

- You cannot influence the routing policies of downstream autonomous systems.

“BGP does not enable one AS to send traffic to a neighbor AS intending that the traffic take a different route from that taken by traffic originating in the neighbor AS.” (RFC 1771)

© 2015 Cisco and/or its affiliates. All rights reserved.

21

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- BGP is an enhanced distance vector protocol with reliable transport.
- Customers that plan to connect to more than one ISP, and small ISPs that plan to have multiple Internet connections in the future, usually use BGP with their service provider. Most multihomed customers use BGP with their service providers.
- A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information that is received from one AS to another AS.
- In the BGP, you cannot influence the routing policies of downstream autonomous systems.

Lesson 2: Understanding BGP Path Attributes

Overview

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

Upon completing this lesson, you will be able to:

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

BGP Path Attributes

Each [BGP](#) update consists of one or more [IP](#) subnets and a set of attributes that are attached to them.

BGP Path Attributes

- BGP metrics are called path attributes.
- BGP attributes are categorized as "well-known" and "optional."
- All compliant implementations must recognize well-known attributes.
- You should not expect that all implementations can recognize optional attributes. Only some implementations (could be private) are able to recognize optional attributes.

© 2015 Cisco and/or its affiliates. All rights reserved.

23

All BGP implementations are required to recognize certain attributes. Those attributes are called "well-known BGP attributes."

Attributes that are not well-known are called "optional." Optional attributes are either specified in a later extension of BGP or even in private vendor extensions that are not documented in a standard document.

Well-Known BGP Attributes

Well-known attributes are divided into mandatory and discretionary.

Well-Known BGP Attributes

- All well-known attributes are propagated to other neighbors.
- The difference between mandatory and discretionary well-known attributes:
 - **Mandatory** well-known attributes must be present in all update messages.
 - **Discretionary** well-known attributes are optional; they could be present in update messages.

© 2015 Cisco and/or its affiliates. All rights reserved.

24

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

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

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

Mandatory Well-Known BGP Attributes

Mandatory Well-Known BGP Attributes

- Origin
 - The origin of a BGP route

i	Route originated in an IGP
e	Route originated in EGP
?	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

© 2015 Cisco and/or its affiliates. All rights reserved.

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

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

Discretionary Well-Known BGP Attributes

Discretionary Well-Known BGP Attributes

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

© 2015 Cisco and/or its affiliates. All rights reserved.

28

All BGP implementations must support discretionary well-known attributes. However, discretionary well-known attributes do not have to be present in all BGP updates. Routers use discretionary well-known attributes only when those functions are required.

The following are descriptions of these two attributes:

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

Optional BGP Attributes

Optional [BGP](#) attributes are transitive or nontransitive.

Optional BGP Attributes

- The difference between transitive or nontransitive optional BGP attributes:
 - **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).

© 2015 Cisco and/or its affiliates. All rights reserved.

27

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

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

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

Optional BGP Attributes (Cont.)

Nontransitive attributes

- MED:
 - Used to discriminate between multiple entry points to a single AS

Transitive attributes

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

© 2015 Cisco and/or its affiliates. All rights reserved.

28

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

Transitive optional attributes include the following:

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

AS-Path Attribute

An edge router modifies the [AS](#)-path attribute every time information about a particular [IP](#) subnet passes over an AS border. When a router first originates a route in [BGP](#), the AS-path attribute is empty. The local AS number is prepended to the AS path each time that the route crosses an AS boundary.

AS-Path Attribute

- 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 attribute 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.

© 2015 Cisco and/or its affiliates. All rights reserved.

29

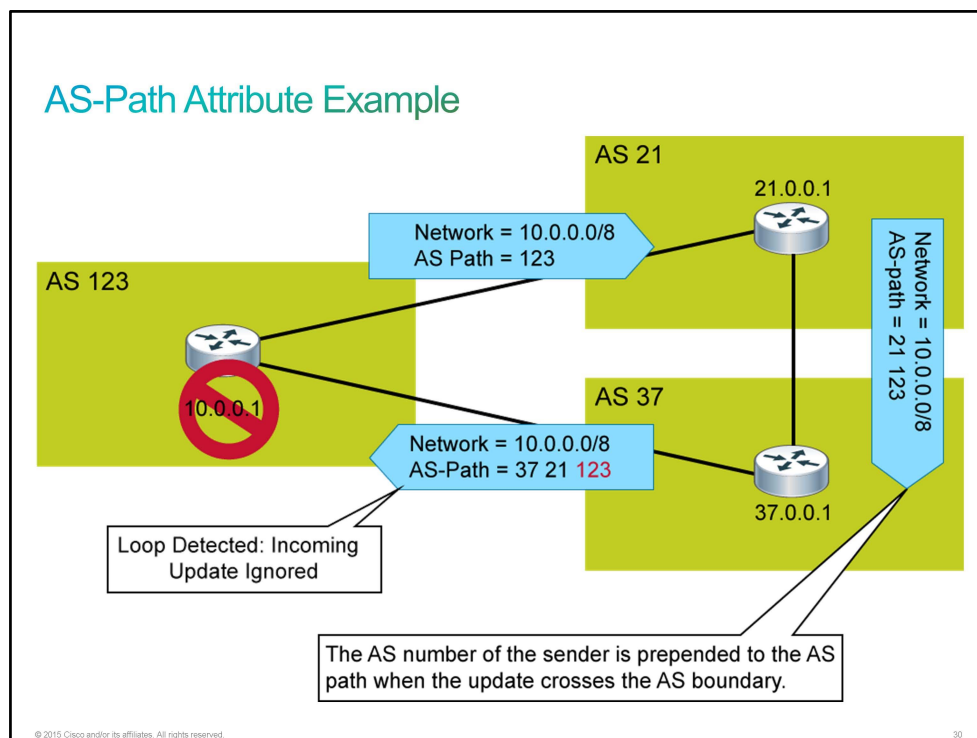
There are several consequences of AS-path attribute behavior:

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

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

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

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

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

Next-Hop Attribute

The [BGP](#) next-hop attribute identifies the [IP address](#) that a router should use to forward packets toward the destination that is announced in a BGP routing update.

Next-Hop Attribute

- The next-hop attribute indicates the next-hop IP address that is used for packet forwarding.
- It is usually set to the IP address of the sending EBGP router.
- Can be set to a third-party IP address to optimize routing.

© 2015 Cisco and/or its affiliates. All rights reserved.

31

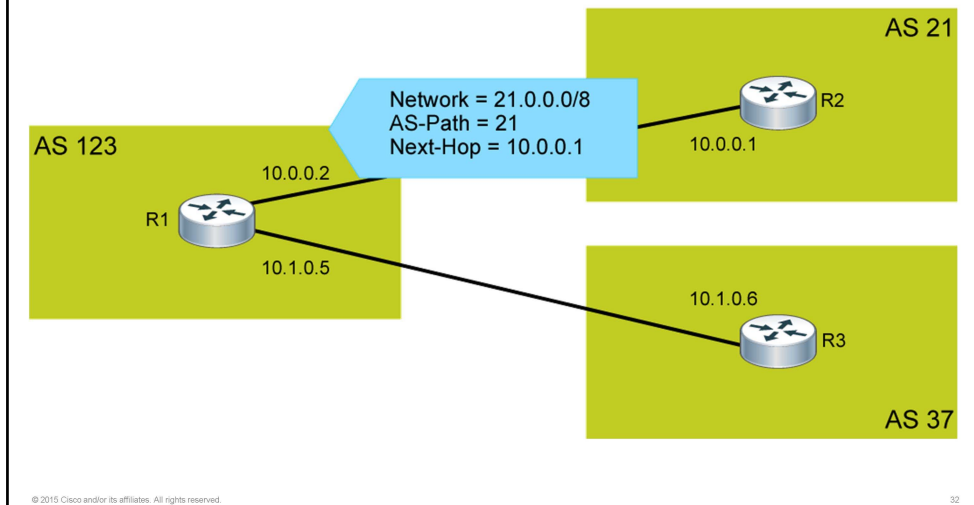
In most cases, the sending router sets the next-hop attribute to its own IP address. There are cases, however, where the next-hop IP address points to a third router.

Next-Hop Processing

The figure shows the usual next-hop processing.

Next-Hop Processing

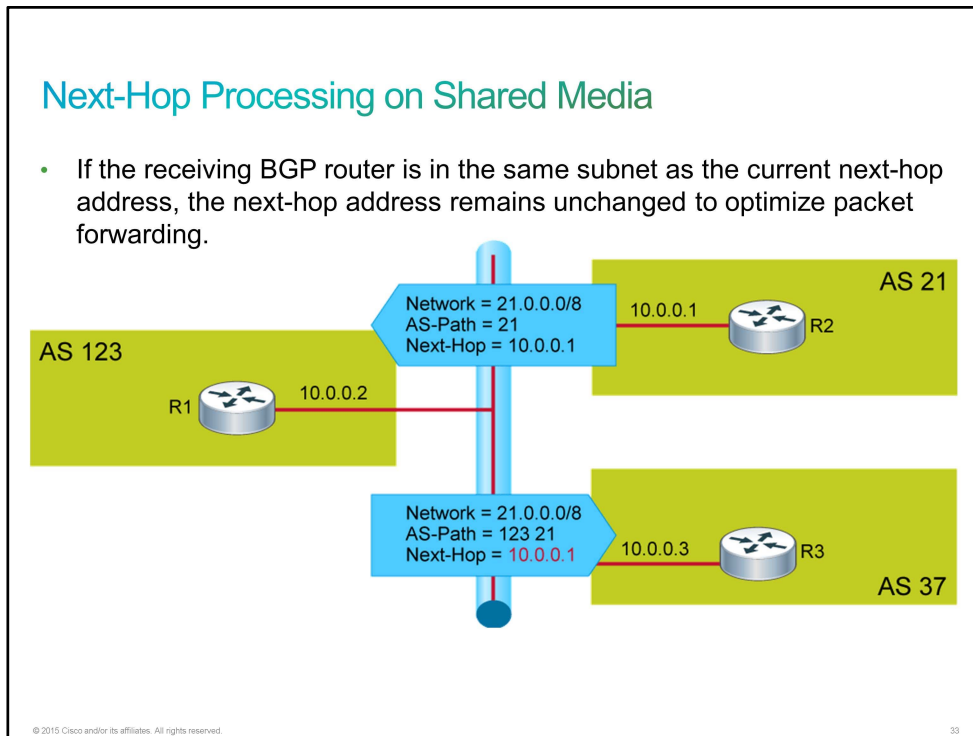
- Next-hop attribute is usually set to the IP address of the sending router.



The processing is as follows:

- R2 announces network 21.0.0.0/8 to R1. The outgoing IP address of R2 (the address that is used to establish the BGP [TCP](#) session) is used as the BGP next hop.
- R1 receives the routing update and installs it in its BGP table and routing table. Should R1 need to forward packets toward network 21.0.0.0/8, it would send those packets toward the IP address 10.0.0.1 (R2).
- When R1 propagates the information about 21.0.0.0/8 to R3, it sets the BGP next-hop attribute to its own IP address.

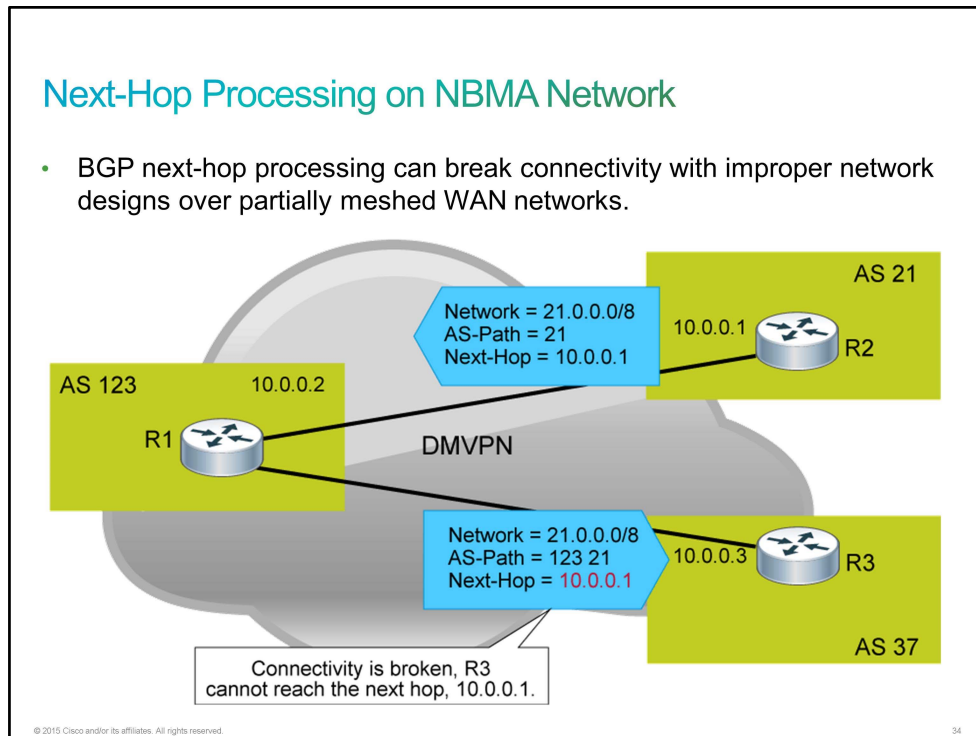
Next-Hop Processing on Shared Media



The next-hop processing changes if the BGP routers connect to a shared subnet. In the figure here, if R1 announces the network 21.0.0.0/8 to R3 with the BGP next-hop address set to R1, the packets from AS 37 toward network 21.0.0.0/8 will have to cross the shared LAN twice. R1 thus sends the routing update toward R3 with the BGP next-hop address unchanged (still pointing toward R2), 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 Processing on NBMA Network



BGP next-hop processing results in optimum data transfer over shared media (for example, a LAN subnet). In partially meshed networks (such as [DMVPN](#)), BGP next-hop processing can break IP connectivity.

Consider, for example, the network diagram in the figure. R1 sends a routing update about network 21.0.0.0/8 to R3 with R2 set to the next-hop address (as they are all in the same subnet). Because there is no direct connection (virtual circuit) between R3 and R2 but R3 still tries to send packets directly toward R2, the connectivity between AS 37 and AS 21 is broken.

There are two ways to solve this connectivity loss:

- Use the subinterfaces on R1 to make sure that R2 and R3 are in different subnets. BGP next-hop processing will then ensure that R1 is the BGP next-hop in the outgoing BGP updates.
- Disable the BGP next-hop processing on R1 in an existing multipoint DMVPN design that shares a common subnet. (This option is strongly discouraged in normal BGP design because routing problems should be solved with a proper network design of point-to-point subinterfaces.)

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- BGP metrics that are attached to a BGP route are called "path attributes." Some path attributes are well-known. Every BGP implementation should be able to recognize the well-known attributes.
- Some of the well-known attributes are mandatory and have to be present in every BGP update. Mandatory well-known attributes are the 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).
- The AS-path attribute lists the autonomous systems that the routing update has already crossed. This attribute is used for BGP loop detection and BGP route selection.
- The next-hop attribute specifies the IP address that should be used for packet forwarding. The next hop is usually set to the IP address of the BGP router sending the update.

© 2015 Cisco and/or its affiliates. All rights reserved.

35

Lesson 3: Establishing BGP Sessions

Overview

Understanding the [BGP](#) neighbor session establishment process is a key component to understanding the fundamental operation of the BGP protocol.

BGP is an [EGP](#) that has been designed for scalability and policy control. As a result, BGP requires neighboring routers to be explicitly configured before BGP routing updates can be sent between them. This situation differs from [IGPs](#) such as [EIGRP](#), [IS-IS](#), and [OSPF](#), that discover neighbors by using a broadcast packet or a hello protocol. In this lesson, BGP neighbor session establishment procedures are discussed.

Upon completing this lesson, you will be able to:

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

BGP Neighbor Discovery

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

BGP Neighbor Discovery

- BGP neighbors are not discovered; they must be configured manually.
- Configuration must be done on both sides of the connection.
- Both routers will attempt to connect to the other with a TCP session on port number 179.
- Only the session with the higher router-ID remains after the connection attempt.
- The source IP address of incoming connection attempts is verified against a list of configured neighbors.

© 2015 Cisco and/or its affiliates. All rights reserved.

36

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

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

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

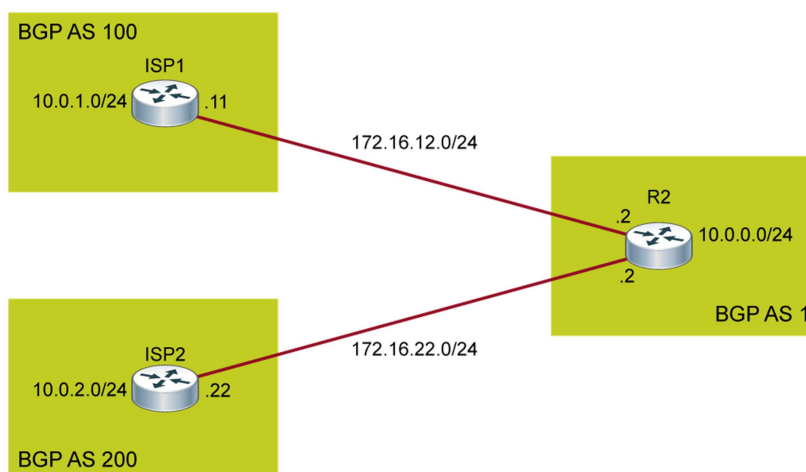
A router may not open a BGP session to itself. If the configured neighbor IP address is, in fact, an IP address of the local router, the router recognizes the problem and tears down the session. The router-ID is also used for this verification.

BGP Neighbor Discovery Example

This example illustrates a small BGP network.

BGP Neighbor Discovery Example

Small BGP Network



© 2015 Cisco and/or its affiliates. All rights reserved.

37

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

BGP Neighbor Discovery Example (Cont.)

Initially, all BGP sessions to the neighbors are idle.

```
R2# show ip bgp summary
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 31, main routing table version 31

Neighbor      V    AS  MsgRcvd MsgSent  TblVer  InQ  OutQ  Up/Down  State/PfxRcd
172.16.12.11  4   100      0       0        1    0    0  00:00:37  Idle
172.16.22.22  4   200      0       0        1    0    0  00:00:14  Idle
```

© 2015 Cisco and/or its affiliates. All rights reserved.

38

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

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

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

Establishing a BGP Session

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

Establishing a BGP Session

- A TCP session is established when the neighbor becomes reachable.
- BGP Open messages are exchanged.

```
R2# debug ip tcp transactions
R2# debug ip bgp
R2# debug ip bgp events
BGP: 172.16.12.11 active went from Idle to Active
BGP: 172.16.12.11 open active, local address 172.16.12.2
TCP0: state was LISTEN -> SYNRCVD [179 -> 172.16.12.11(11374)]
TCP0: state was SYNRCVD -> ESTAB [179 -> 172.16.12.11(11374)]
TCBEFF01278 accepting EC3FFBE8 from 172.16.12.11.11374
BGP: 172.16.12.11 passive open to 172.16.12.2
BGP: 172.16.12.11 passive went from Idle to Connect
BGP: 172.16.12.11 passive rcv message type 1, length (excl. header) 38
BGP: ses global 172.16.12.11 (0xEFF0F1B8:0) pas Receive OPEN
BGP: 172.16.12.11 passive rcvd OPEN w/ remote AS 100, 4-byte remote AS 100
BGP: 172.16.12.11 passive went from Connect to OpenSent
BGP: 172.16.12.11 passive went from OpenSent to OpenConfirm
BGP: 172.16.12.11 passive went from OpenConfirm to Established
BGP(0): 172.16.12.11 was the first peer to be established for IPv4 Unicast
%BGP-5-ADJCHANGE: neighbor 172.16.12.11 Up
```

© 2015 Cisco and/or its affiliates. All rights reserved.

38

The debug output shows how the router creates a socket data structure and binds it to its local IP address 172.16.12.2 and the well-known port 179. Then the router sends a [TCP SYN](#) packet to the configured peer router IP address of 172.16.12.11 and a high port number 11374. The connection attempt succeeds, and the TCP session is then ready to transfer the BGP information.

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

Establishing a BGP Session (Cont.)

The BGP Open message contains the following:

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

© 2015 Cisco and/or its affiliates. All rights reserved.

40

The parameters in the BGP Open message are as follows:

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

Establishing a BGP Session (Cont.)

BGP neighbors—steady state

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

```
R2# show ip bgp summary
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 16, main routing table version 16
15 network entries using 2220 bytes of memory
15 path entries using 960 bytes of memory
2/2 BGP path/bestpath attribute entries using 272 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3500 total bytes of memory
BGP activity 78/63 prefixes, 84/69 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	18	15	16	0	0	00:11:46	6
172.16.22.22	4	200	5	8	16	0	0	00:00:27	9

© 2015 Cisco and/or its affiliates. All rights reserved.

41

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

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

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

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

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

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

BGP Keepalives

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

BGP Keepalives

- A TCP-based BGP session does not provide any means of verifying BGP neighbor presence:
 - Except when sending BGP traffic
- BGP needs an extra mechanism:
 - Keepalive BGP messages provide verification of neighbor existence.
 - Keepalive messages are sent every 60 seconds.

© 2015 Cisco and/or its affiliates. All rights reserved.

42

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

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

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

BGP Keepalives (Cont.)

- Keepalive interval value is not communicated in the BGP Open message.
- Keepalive value is selected as follows:
 - Configured value, if local holdtime is used
 - Configured value, if holdtime of neighbor is used and $\text{keepalive} < (\text{holdtime} / 3)$
 - Smaller integer in relation to $(\text{holdtime} / 3)$, if holdtime of neighbor is used and $\text{keepalive} > (\text{holdtime} / 3)$

© 2015 Cisco and/or its affiliates. All rights reserved.

43

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

The BGP process selects the keepalive interval value according to these conditions:

1. If the locally configured value of holdtime is selected (being the lower of two), the peers use the locally used keepalive interval.
2. If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is less than a third of the holdtime interval, the peers use the locally configured keepalive.
3. If the holdtime interval of the neighbor is selected, and the locally configured keepalive interval is more than a third of the holdtime interval, the peers use the smaller integer value in relation to $(\text{holdtime} / 3)$.

Keepalive Value Example

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

MD5 Authentication

Authentication between [BGP](#) neighbors can be negotiated between BGP-speaking routers using optional parameters in the Open message.

MD5 Authentication

- BGP peers may optionally use MD5 TCP authentication using a shared secret.
- Both routers must be configured with the same password (MD5 shared secret).
- Each TCP segment is verified.

© 2015 Cisco and/or its affiliates. All rights reserved.

44

If you are using [MD5](#) authentication, every [TCP](#) segment on the BGP session will be transmitted to the configured neighbor along with a checksum. The checksum is calculated together with a secret known by the two routers using the MD5 algorithm. The common secret is never transmitted on the network. If the receiver, which is using the same common secret, calculates the same checksum from the TCP segment, then the receiver can be pretty sure that the information is transmitted from the correct source and the information has not been altered.

Authentication of BGP sessions is a vital tool to avoid [DoS](#) attacks.

Note	Cisco routers support key chain to configure BGP authentication. BGP supports only HMAC-MD5 and HMAC-SHA1-12 cryptographic algorithms.
-------------	--

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- With interior routing protocols, adjacent routers are usually discovered through a dedicated hello protocol. In BGP, neighbors must be manually configured to increase routing protocol security.
- BGP neighbors, once configured, establish a TCP session and exchange the BGP Open message, which contains the parameters that each BGP router proposes to use.
- The router uses BGP keepalives to provide a verification of the existence of a configured BGP neighbor.
- MD5 authentication can be configured on a BGP session to help prevent spoofing, DoS attacks, or man-in-the-middle attacks.

© 2015 Cisco and/or its affiliates. All rights reserved.

45

Lesson 4: Processing BGP Routes

Overview

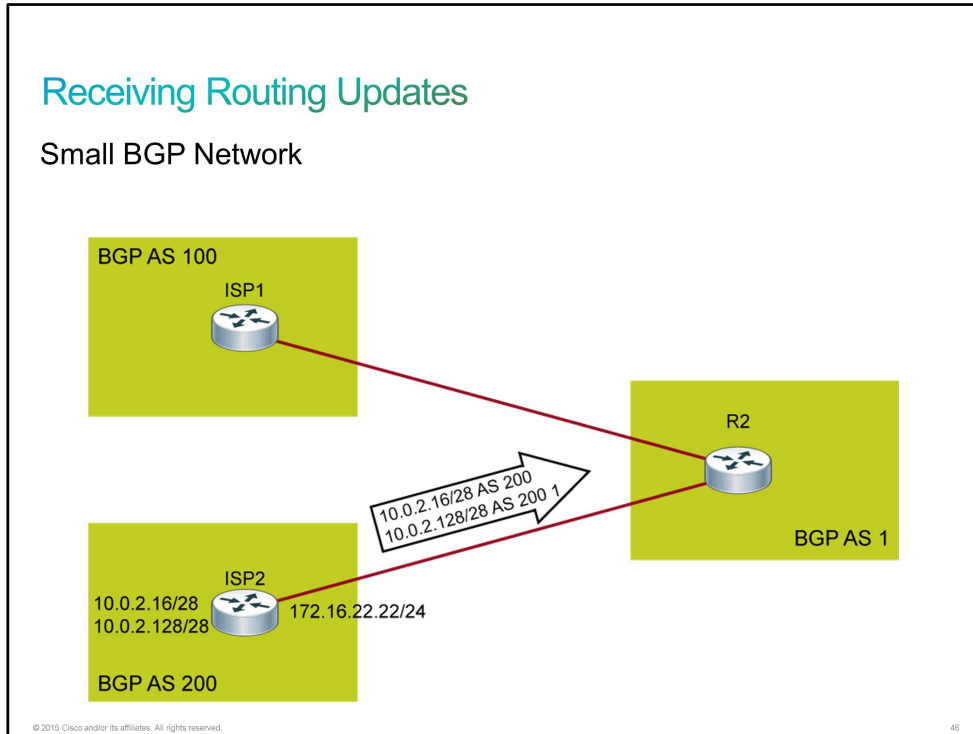
Route processing is fundamental to the operation of [BGP](#). Knowledge of the BGP route selection process, route propagation, and how the BGP and IP routing tables are built is key to properly configuring BGP and troubleshooting BGP routing issues.

Upon completing this lesson, you will be able to:

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

Receiving Routing Updates

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



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

Receiving Routing Updates (Cont.)

Information from the BGP tables is exchanged after adjacency establishment.

```
R2# debug ip bgp updates
%BGP-5-ADJCHANGE: neighbor 172.16.22.22 Up
BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22, origin i,
metric 0, merged path 200, AS_PATH
BGP(0): 172.16.22.22 rcv UPDATE about 10.0.2.128/28 -- DENIED due to: AS-PATH
contains our own AS;
BGP(0): 172.16.22.22 rcvd 10.0.2.16/28
BGP(0): Revise route installing 1 of 1 routes for 10.0.2.16/28 ->
172.16.22.22(global) to main IP table
```

© 2015 Cisco and/or its affiliates. All rights reserved.

47

The debug output shows how information about network 10.0.2.16/28 is received from neighbor 172.16.22.22. The neighbor indicates that [IP](#) packets to destination IP addresses in network 10.0.2.16/28 can be forwarded to the next-hop address 172.16.22.22. The AS path 200 indicates that the final destination is in AS 200. The metric is the [MED](#) value.

Network 10.0.2.128/28 is denied. The reason is AS path loop detection. The receiving router detects its own AS number in the AS path and silently discards (denies) the route.

Building BGP Table

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

Building BGP Table

All inbound updates are placed into the BGP table.

```
R2# show ip bgp
BGP table version is 11, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	1.0.0.0/24	172.16.22.22	0		0 200	99 i
*>		172.16.12.11	0		0 100	99 i
*>	10.0.1.0/24	172.16.12.11	0		0 100	i
*>	10.0.2.0/28	172.16.22.22	0		0 200	i
*>	10.0.2.16/28	172.16.22.22	0		0 200	i
*>	10.0.2.32/28	172.16.22.22	0		0 200	i
*>	10.0.2.48/28	172.16.22.22	0		0 200	i
*>	10.0.2.64/28	172.16.22.22	0		0 200	i
*>	10.0.2.80/28	172.16.22.22	0		0 200	i
*>	10.0.2.96/28	172.16.22.22	0		0 200	i
*>	10.0.2.112/28	172.16.22.22	0		0 200	i

© 2015 Cisco and/or its affiliates. All rights reserved.

46

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

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

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

BGP Route Selection Criteria

When a router has more than one alternative route to reach the same [IP](#) subnet (network and mask), the router has to select one of the routes as best in its default mode of operation. To make this selection, the router uses the [BGP](#) attributes that are attached to the various updates.

BGP Route Selection Criteria

- Exclude routes with inaccessible next hop
- Prefer highest weight (local to router)
- Prefer highest local preference (global within AS)
- Prefer routes that the router originated
- Prefer shortest AS path (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 lowest BGP router-ID

© 2015 Cisco and/or its affiliates. All rights reserved.

49

The selection criteria are checked in the order that is indicated in the following steps. The first check that indicates a difference is used for route selection, and no further testing is done.

1. The router checks whether the next-hop attribute indicates an [IP address](#) that is reachable according to the current routing table. It is not necessary to have a direct connection to the next hop. It can very well be several router hops away and the route to it learned by the [IGP](#). If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best.
2. The router prefers the route with the higher weight. The weight is not carried with the updates; it is a value that is assigned to the route by the local router and considered only within the router itself.
3. If the local preference attributes are different, the route with the highest value is selected as best.
4. If one of the routes is injected into the BGP table by the local router, the local router prefers it to any routes that it receives from other BGP routers.
5. At this point, the lengths of the [AS](#) paths are compared (the content is not checked; only the number of autonomous systems in each AS path is counted). The route with the shortest length is selected.
6. If the AS-path lengths are the same, the origin code is checked. BGP will prefer the path with the lowest origin type: IGP is lower than [EGP](#), and EGP is lower than Incomplete.
7. The router next compares [MED](#) values but only if it receives the updates from the same neighboring AS. Routes with a lower MED are preferred.

8. At this point, it is clear that the destination network is outside the local AS and that there is not much difference among the alternatives. Because the IP packets to the destination network must leave the AS, it is better that they do so as quickly as possible. If any of the alternatives are received from a BGP peer in another AS, that alternative is preferred.
9. If the router receives all alternatives from peer routers in the local AS, each of them will indicate an exit point, and the closest exit is used. Distance to the exit point is calculated by comparing the IGP costs against the BGP next hops, as indicated in the routing table.
10. If the router receives all alternatives from [EBGP](#) neighbors, the most stable path (the oldest path) is preferred.
11. If the router still cannot differentiate among the routes, it nevertheless has to make a decision and select the best route. It checks the BGP sessions on which it received the updates and chooses the route that was received on the session for which the peer router has the lowest BGP router-ID.

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

BGP Route Propagation

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

BGP Route Propagation

The best BGP routes are propagated to BGP neighbors.

```
R2# debug ip bgp updates
BGP(0): (base) 172.16.12.11 send UPDATE (format) 10.0.2.0/28, next 172.16.12.2,
metric 0, path 200
BGP(0): 172.16.12.11 NEXT_HOP is on same subnet as the bgp peer and set to
172.16.12.11 for net 10.0.1.0/24, flags 200, sb: AC100C00, mask: FFFFFFF0
BGP(0): (base) 172.16.12.11 send UPDATE (format) 10.0.1.0/24, next
172.16.12.11, metric 0, path 100
BGP(0): (base) 172.16.12.11 send UPDATE (format) 10.0.0.0/28, next 172.16.12.2,
metric 0, path Local
```

© 2015 Cisco and/or its affiliates. All rights reserved.

50

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

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

Building IP Routing Table

The route in the [BGP](#) table that BGP selects as the best is a candidate for installation in the [IP](#) forwarding table or routing table.

Building IP Routing Table

The best BGP routes are copied into the IP routing table based on administrative distance.

```
R2# show ip route
<... output omitted ...>

Gateway of last resort is not set

      1.0.0.0/24 is subnetted, 1 subnets
B       1.0.0.0 [20/0] via 172.16.22.22, 00:03:07
      10.0.0.0/8 is variably subnetted, 20 subnets, 3 masks
<... output omitted ...>
B       10.0.1.0/24 [20/0] via 172.16.12.11, 00:03:06
B       10.0.2.0/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.16/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.32/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.48/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.64/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.80/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.96/28 [20/0] via 172.16.22.22, 00:06:46
B       10.0.2.112/28 [20/0] via 172.16.22.22, 00:06:46
<... output omitted ...>
```

© 2015 Cisco and/or its affiliates. All rights reserved.

51

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

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

In this example, network 1.0.0.0/24 is reachable via 172.16.22.22. After the router has installed the route in the routing table, user data traffic starts to be forwarded.

Advertising Local Networks

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

Advertising Local Networks

- The BGP router process keeps a list of local networks (defined with the **network** command or through redistribution).
- The BGP process periodically scans the IP routing table and inserts or revokes routes from the BGP routing table based on their presence in the IP routing table.

© 2015 Cisco and/or its affiliates. All rights reserved.

52

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

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

Advertising Local Networks Example

In this example, network 10.0.0.0/28 is directly connected to interface Loopback1.

Advertising Local Networks Example

The BGP route is revoked after the network is removed from the routing table.

```
R2# debug ip routing
R2# debug ip bgp updates
is_up: Loopback1 0 state: 6 sub state: 1 line: 0
RT: interface Loopback1 removed from routing table
RT: del 10.0.0.0 via 0.0.0.0, connected metric [0/0]
RT: delete subnet route to 10.0.0.0/28
RT: del 10.0.0.1 via 0.0.0.0, connected metric [0/0]
RT: delete subnet route to 10.0.0.1/32
BGP(0): route 10.0.0.0/28 down
BGP(0): no valid path for 10.0.0.0/28
BGP: topo global:IPv4 Unicast:base Remove_fwdroute for 10.0.0.0/28
BGP(0): (base) 172.16.12.11 send unreachable (format) 10.0.0.0/28
```

© 2015 Cisco and/or its affiliates. All rights reserved.

53

The route to 10.0.0.0/28 was previously installed in the BGP table because it was listed with a network statement and it was in the routing table as directly connected. When the Loopback1 interface goes down, the router removes the directly connected route from its routing table. Because the route no longer exists in the routing table, it must also be removed from the BGP table.

Because there has been a change in the BGP table, the BGP neighbors must be informed. The router sends a BGP update message to its neighbor indicating that network 10.0.0.0/28 is now unreachable.

Advertising Local Networks Example (Cont.)

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

```
RT: add 10.0.0.0/28 via 0.0.0.0, connected metric [0/0]
RT: interface Loopback1 added to routing table
RT: updating connected 10.0.0.1/32 (0x0):
    via 0.0.0.0 Lo1
RT: add 10.0.0.1/32 via 0.0.0.0, connected metric [0/0]
BGP(0): route 10.0.0.0/28 up
BGP: topo global:IPv4 Unicast:base Remove_fwdroute for 10.0.0.0/28
BGP(0): redistributedlocal route 10.0.0.0/28 modified
BGP(0): (base) 172.16.12.11 send UPDATE (format) 10.0.0.0/28, next 172.16.12.2,
metric 0, path Local
```

© 2015 Cisco and/or its affiliates. All rights reserved.

54

In this example, network 10.0.0.0/28 is listed with a network statement in the BGP process. However, the network was not in the routing table of the router, so the network was not injected into its BGP table.

Later, the Loopback1 interface comes back up again. This reappearance means that the network 10.0.0.0/28 is now in the routing table as a directly connected route. As a result, the router once again injects the 10.0.0.0/28 network into its BGP table and then updates its configured neighbor.

Automatic Summarization

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

Verify if BGP automatic summarization is disabled or enabled. The following examples will have BGP automatic summarization enabled.

Automatic Summarization

- Verify BGP automatic summarization.
- Enable automatic summarization when:
 - Summarization of IGP-to-BGP redistributed routes to major network boundary required
 - Using classful **network** command to summarize subnets to a major network boundary
- Disable automatic summarization when:
 - Summarization on IGP-to-BGP redistribution not desired
 - Using classless variant of the **network** command

© 2015 Cisco and/or its affiliates. All rights reserved.

55

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

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

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

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

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

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

Automatic Summarization Example

In this example, five subnets of the major class A network 10.0.0.0/8 (10.0.0.0/28, 10.0.0.16/28, 10.0.0.32/28, 10.0.0.48/28, and 10.0.0.64/28) exist in the routing table.

Automatic Summarization Example

```
R2# show ip route
      10.0.0.0/8 is variably subnetted, 11 subnets, 3 masks
<... output omitted ...>
C       10.0.0.0/28 is directly connected, Loopback1
C       10.0.0.16/28 is directly connected, Loopback2
C       10.0.0.32/28 is directly connected, Loopback3
C       10.0.0.48/28 is directly connected, Loopback4
C       10.0.0.64/28 is directly connected, Loopback5
<... output omitted ...>
```

- Five subnets for 10.0.0.0 exist in the routing table.
- Automatic summarization is enabled for BGP.
- BGP has been configured to locally announce 10.0.0.0.

```
router bgp 1
 network 10.0.0.0
 auto-summary
```

© 2015 Cisco and/or its affiliates. All rights reserved.55

Automatic Summarization Example (Cont.)

```
R2# show ip bgp
BGP table version is 12, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop        Metric LocPrf Weight Path
*>  1.0.0.0/24       172.16.12.11      0             0 100 99 i
*>  10.0.0.0         0.0.0.0           0             32768 i
*>  10.0.1.0/24      172.16.12.11      0             0 100 i
```

- Classful network summary is inserted into BGP table.

© 2015 Cisco and/or its affiliates. All rights reserved.

57

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

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

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

Subnets 10.0.0.0/28, 10.0.0.16/28, 10.0.0.32/28, 10.0.0.48/28, and 10.0.0.64/28 were summarized during insertion into the BGP table to the classful network 10.0.0.0/8. This action occurred because a classful **network** command and automatic summarization were configured on the router. If automatic summarization is disabled, no insertion into the BGP table occurs at all.

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

Automatic Summarization Example (Cont.)

```
R2# show ip route
<... output omitted ...>
    10.0.0.0/8 is variably subnetted, 11 subnets, 3 masks
C       10.0.0.0/28 is directly connected, Loopback1
C       10.0.0.16/28 is directly connected, Loopback2
C       10.0.0.32/28 is directly connected, Loopback3
C       10.0.0.48/28 is directly connected, Loopback4
C       10.0.0.64/28 is directly connected, Loopback5
<... output omitted ...>
```

- Five subnets 10.0.0.0 exist in the routing table.
- Automatic summarization is enabled for BGP.
- BGP has been configured to redistribute connected routes into BGP.

```
router bgp 1
 redistribute connected
 auto-summary
```

© 2015 Cisco and/or its affiliates. All rights reserved.

58

Automatic Summarization Example (Cont.)

```
R2# show ip bgp
BGP table version is 15, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	1.0.0.0/24	172.16.12.11	0		0 100 99	i
*>	10.0.0.0	0.0.0.0	0		32768	?
*>	10.0.1.0/24	172.16.12.11	0		0 100	i
*>	172.16.0.0	0.0.0.0	0		32768	?

- Classful network summary is inserted into BGP table.

© 2015 Cisco and/or its affiliates. All rights reserved.

59

In this example, automatic summarization is enabled, resulting in the summarization of redistributed subnets to their classful boundaries. Subnets 10.0.0.0/28, 10.0.0.16/28, 10.0.0.32/28, 10.0.0.48/28, and 10.0.0.64/28 will be summarized into the single class A network 10.0.0.0/8. The network 10.0.0.0/8 is a locally sourced summary with all the attributes of a locally sourced BGP route (next hop = 0.0.0.0, weight = 32768, empty AS-path list). The origin of the route is marked as incomplete because the route is sourced through redistribution.

If automatic summarization was disabled, more specific routes would be present in the BGP table instead of the summary prefix 10.0.0.0/8.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

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

© 2015 Cisco and/or its affiliates. All rights reserved.

60

Summary (Cont.)

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

© 2015 Cisco and/or its affiliates. All rights reserved.

61

Lesson 5: Configuring Basic BGP

Overview

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

Upon completing this lesson, you will be able to:

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

BGP Routing Process

You will use the **router bgp** command to start the [BGP](#) routing process in the router. There can be, at most, one BGP process in a router. It must be assigned the local [AS](#) number.

BGP Routing Process

```
router(config)# router bgp as-number
```

- Starts BGP routing.
- Get your AS number from American Registry for Internet Numbers (<http://www.arin.net>) or Réseaux IP Européens (<http://www.ripe.net>).
- Use private AS numbers (between 64,512 and 65,534 or between 4,200,000,000 and 4,294,967,294) if you run BGP in a private network.
- Only one BGP routing process per router is allowed.

© 2015 Cisco and/or its affiliates. All rights reserved.62

The AS number is a 16-bit or 32-bit integer. New 32-bit AS numbers were created when the AS number pool from [IANA](#) approached exhaustion. The AS number pool was extended from 16 to 32 bits. It must uniquely identify the AS among all routers that are exchanging BGP routing information, either directly or indirectly. This requirement means that the AS numbers must be unique when BGP information is exchanged with the Internet.

The AS number can be a public AS number or a private AS number. Public AS numbers range from 1 to 64,534 or 131,072 to 4,199,999,999. An Internet registry ([ARIN](#): <http://www.arin.net> or [RIPE](#): <http://www.ripe.net>) assigns public AS numbers. Private AS numbers range from 64,512 to 65,534 or 4,200,000,000— 4,294,967,294. Private AS numbers are never propagated onto the public Internet.

Configuring External Neighbors

[BGP](#) does not automatically discover neighbors. You will need to explicitly configure BGP neighbors. The local router will try to connect to the indicated [IP address](#) and also accept incoming connection attempts from the indicated IP address.

Configuring External Neighbors

```
router(config-router)# neighbor ip-address remote-as as-number
```

- Defines an external neighbor
- External neighbor has to be reachable over directly connected subnet

```
router(config-router)# neighbor ip-address description neighbor description
```

- Assigns a description to an external neighbor.

```
router(config-router)# neighbor ip-address shutdown
```

- Disables a BGP neighbor

© 2015 Cisco and/or its affiliates. All rights reserved.63

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

You may optionally configure other attributes with the neighbor on successive configuration lines, referring to the same neighbor IP address but indicating different attributes.

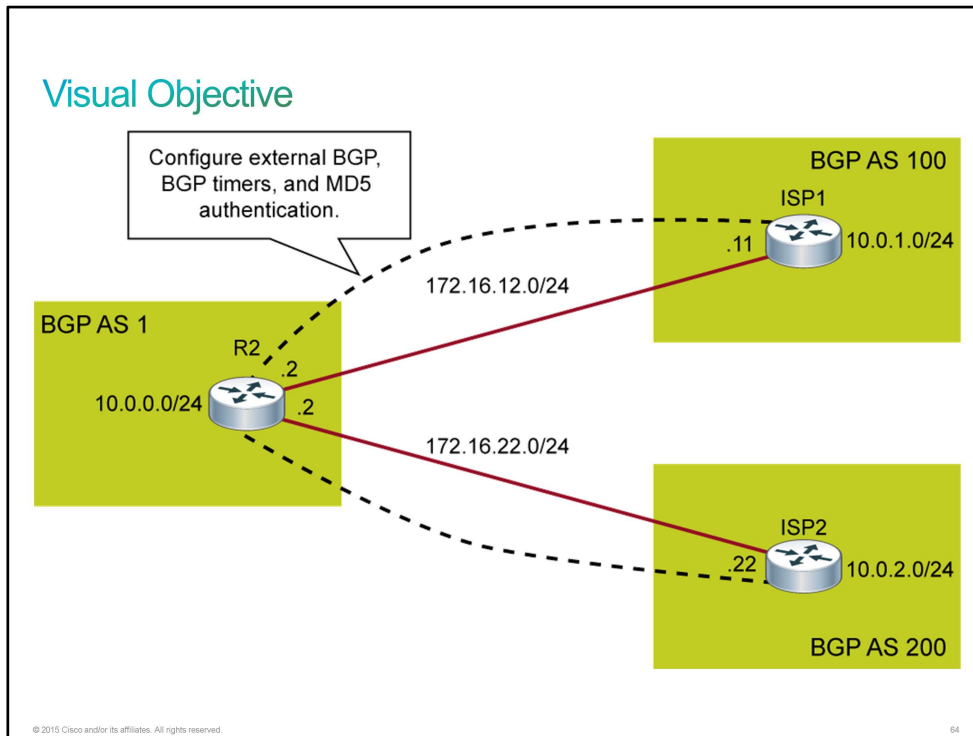
With the **neighbor description** command, a description (text string) can be entered that describes the neighbor.

When you are debugging or troubleshooting, you can use the **neighbor shutdown** command to temporarily disable BGP neighbor without deleting BGP neighbor configuration. You can shut down the BGP neighbor during extensive modification of routing policies to prevent inconsistent routing data.

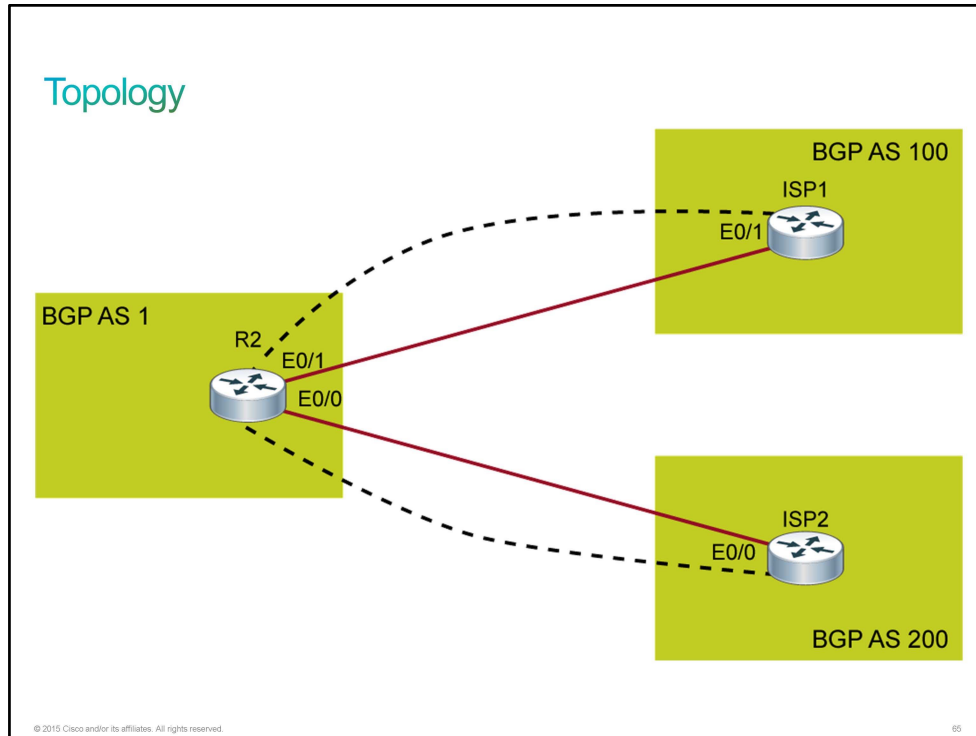
Discovery 1: Configure Basic BGP

Overview

Through this discovery, you will learn how to configure external BGP, BGP timers, and MD5 authentication.



Topology

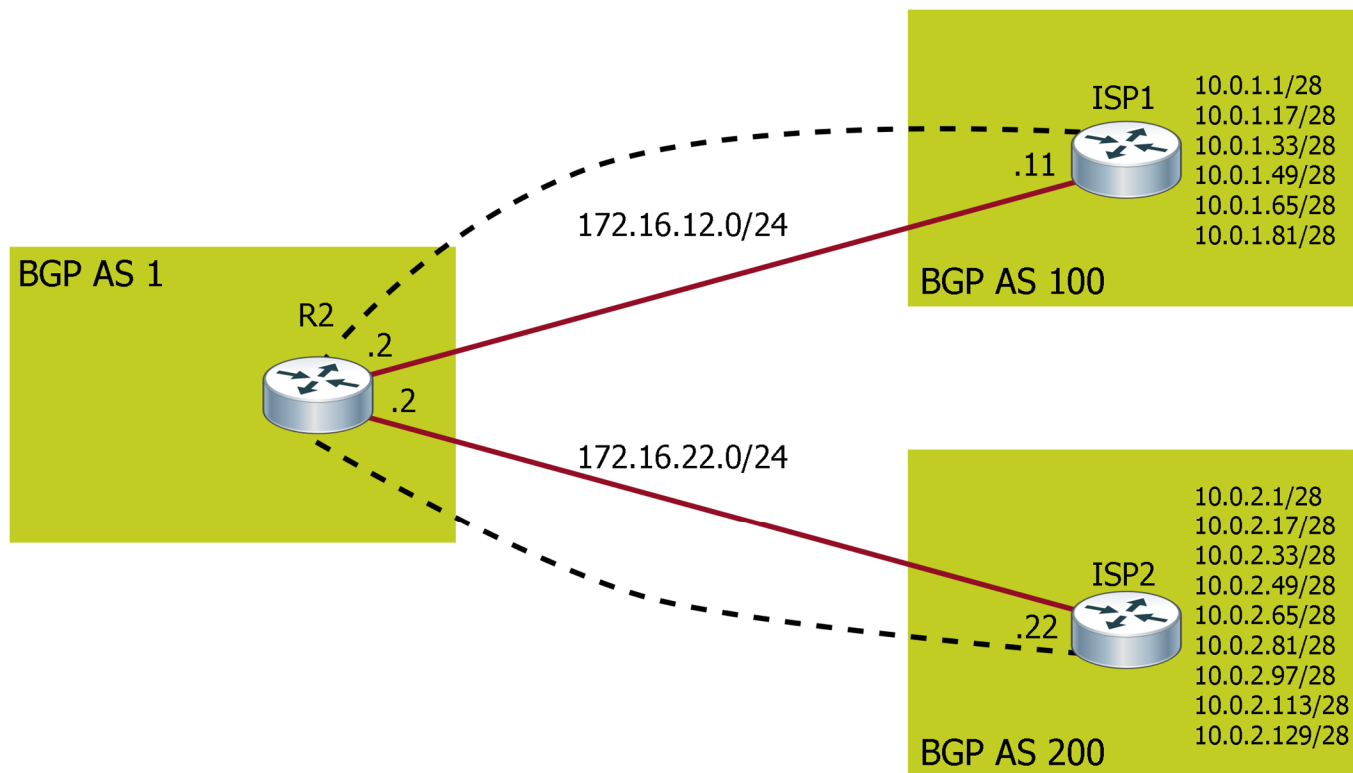


Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks



Configure Basic BGP

Discovery Steps

Step 1 You will configure an external BGP session between R2 and ISP1 routers. On the R2 router, configure BGP external neighbor ISP1.

ISP1 has IP address 172.16.12.11 and BGP AS 100. R2 router is already configured in the BGP AS 1.

```
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 remote-as 100
```

Step 2 On the R2 router, verify the state of the BGP session.

To verify the BGP session, use the **show ip bgp summary** command.

```
R2# show ip bgp summary
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 34, main routing table version 34
9 network entries using 1332 bytes of memory
9 path entries using 576 bytes of memory
1/1 BGP path/bestpath attribute entries using 136 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 2068 total bytes of memory
BGP activity 15/6 prefixes, 21/12 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	0	0	1	0	0	never	Idle
172.16.22.22	4	200	1359	1357	34	0	0	20:29:01	9

You should see "Idle" state next to the neighbor 172.16.12.11.

Step 3 On the ISP1 router, configure BGP external neighbor R2.

R2 has IP address 172.16.12.2 and BGP AS 1. ISP1 router is already configured in the BGP AS 100.

```
ISP1(config)# router bgp 100
ISP1(config-router)# neighbor 172.16.12.2 remote-as 1
```

Shortly, after configuring external BGP neighbor on the ISP1 router, you will see the following message on the ISP1 router's console:

```
*Feb  5 01:13:18.858: %BGP-5-ADJCHANGE: neighbor 172.16.12.2 Up
```

External BGP session between R2 and ISP1 is successfully established.

Step 4 On the R2 router, verify the state of the BGP session.

To verify the BGP session, use the **show ip bgp summary** command.

```
R2# show ip bgp summary
```

```
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 40, main routing table version 40
15 network entries using 2220 bytes of memory
15 path entries using 960 bytes of memory
2/2 BGP path/bestpath attribute entries using 272 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3500 total bytes of memory
BGP activity 21/6 prefixes, 27/12 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	32	36	40	0	0	00:25:36	6
172.16.22.22	4	200	1398	1398	40	0	0	21:04:31	9

Instead of state information, you should see the number of prefixes router received from its neighbor. Now you know that the state is Established and that router R2 received 6 prefixes from the ISP1 router.

Step 5 On the R2 router confirm that BGP state is established.

To confirm BGP state, use the **show ip bgp neighbor** command.

```
R2# show ip bgp neighbors 172.16.12.11
```

```
BGP neighbor is 172.16.12.11, remote AS 100, external link
```

```
BGP version 4, remote router ID 10.0.1.81
```

```
BGP state = Established, up for 00:32:50
```

```
Last read 00:00:46, last write 00:00:11, hold time is 180, keepalive interval is 60 seconds
```

```
Neighbor sessions:
```

```
1 active, is not multisession capable (disabled)
```

```
Neighbor capabilities:
```

```
Route refresh: advertised and received(new)
```

```
Four-octets ASN Capability: advertised and received
```

```
Address family IPv4 Unicast: advertised and received
```

```
Enhanced Refresh Capability: advertised and received
```

```
Multisession Capability:
```

```
Stateful switchover support enabled: NO for session 1
```

From the output, also observe the default hold time and keepalive interval.

Configuring BGP Timers

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

Configuring BGP Timers

```
router(config-router)# timers bgp keepalive holdtime
```

- Changes the default values of BGP timers per BGP process.
- Only the holdtime value is communicated in the BGP Open message.
- Both peers use the smallest configured holdtime value on BGP peers.

```
router(config-router)# neighbor ip-address timers keepalive holdtime
```

- Changes the default values of BGP timers per specific neighbor.
- Overrides the **bgp** settings of the timers.

© 2015 Cisco and/or its affiliates. All rights reserved.

66

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

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

Step 6 You will change BGP timers for session between R2 and ISP1 routers. On the R2 router, change the default hold time and keepalive interval for ISP1 neighbor only.

Use a hold time of 30 seconds and keepalive interval 10 seconds.

```
R2(config)# router bgp 1  
R2(config-router)# neighbor 172.16.12.11 timers 10 30
```

Step 7 On the ISP1 router, change the default hold time and keepalive interval for BGP process.

Use hold time of 30 seconds and keepalive interval 10 seconds.

```
ISP1(config)# router bgp 100  
ISP1(config-router)# timers bgp 10 30
```

Timers are negotiated at BGP neighbor establishment. On the ISP1 router, clear BGP sessions by using the following command:

```
ISP1# clear ip bgp *
*Feb  5 02:41:15.321: %BGP-5-ADJCHANGE: neighbor 172.16.12.2 Down User reset
*Feb  5 02:41:15.321: %BGP_SESSION-5-ADJCHANGE: neighbor 172.16.12.2 IPv4 Unicast
topology base removed from session User reset
*Feb  5 02:41:15.382: %BGP-5-ADJCHANGE: neighbor 172.16.12.2 Up
```

Step 8 On the R2 router, verify the BGP hold time and keepalive interval.

BGP timers have changed for ISP1 neighbor:

```
R2# show ip bgp neighbors 172.16.12.11
BGP neighbor is 172.16.12.11, remote AS 100, external link
  BGP version 4, remote router ID 10.0.1.81
  BGP state = Established, up for 00:01:03
  Last read 00:00:08, last write 00:00:00, hold time is 30, keepalive interval is 10
seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
```

BGP timers are default for ISP2 neighbor:

```
R2# show ip bgp neighbors 172.16.22.22
BGP neighbor is 172.16.22.22, remote AS 200, external link
  BGP version 4, remote router ID 10.0.2.129
  BGP state = Established, up for 00:11:01
  Last read 00:00:57, last write 00:00:06, hold time is 180, keepalive interval is 60
seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
```

Configuring MD5 Authentication

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

Configuring MD5 Authentication

```
router(config-router)# neighbor ip-address password string
```

- Enables MD5 authentication on a specific BGP session.
- Password string on both routers must match.

© 2015 Cisco and/or its affiliates. All rights reserved.

67

Step 9 You will enable MD5 authentication for BGP session between R2 and ISP1 routers. On the R2 router, enable MD5 authentication on the BGP session to the ISP1 neighbor.

Use MD5 authentication password **cisco**.

```
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 password cisco
```

Step 10 On the ISP1 router, enable MD5 authentication on the BGP session to the R2 neighbor.

Use MD5 authentication password **cisco**.

```
ISP1(config)# router bgp 100
ISP1(config-router)# neighbor 172.16.12.2 password cisco
```

Make sure the password that is used on ISP1 is same as password used on R2. Space character after password will become part of the password.

If password is misconfigured, you will see the following message on the router console:

```
%TCP-6-BADAUTH: Invalid MD5 digest from 172.16.12.2(60123) to 172.16.12.11(179) (RST)
tableid - 0
```

Step 11 On the ISP1 router, verify that BGP session to the R2 router is established.

```
ISP1# show ip bgp neighbors 172.16.12.2
BGP neighbor is 172.16.12.2, remote AS 1, external link
  BGP version 4, remote router ID 172.16.22.2
  BGP state = Established, up for 00:08:19
  Last read 00:00:02, last write 00:00:02, hold time is 30, keepalive interval is 10
seconds
  Configured hold time is 30, keepalive interval is 10 seconds
  Minimum holdtime from neighbor is 0 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
```

Step 12 On the R2 router, verify that BGP prefixes are received from ISP1 router.

You should see six BGP prefixes received from ISP1 router.

```
R2# show ip bgp summary
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 28, main routing table version 28
15 network entries using 2220 bytes of memory
15 path entries using 960 bytes of memory
2/2 BGP path/bestpath attribute entries using 272 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3500 total bytes of memory
BGP activity 63/48 prefixes, 75/60 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	69	73	28	0	0	00:10:15	6
172.16.22.22	4	200	19	24	28	0	0	00:13:06	9

Announcing Networks in BGP

Before you inject any local routing information into [BGP](#) table for advertising to other BGP routers, some basic configuration is required.

Announcing Networks in BGP

Only administratively defined networks are announced in BGP.

- Manually configure networks to be announced.
- Use redistribution from IGP.
- Use aggregation to announce summary prefixes.

© 2015 Cisco and/or its affiliates. All rights reserved.

88

There are two ways to do this configuration:

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

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

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

Announcing Networks in BGP (Cont.)

```
router(config-router)# (no) auto-summary
```

- Enables or disables summarization of networks before insertion into the BGP table:
 - Locally inserted networks (using the **network** command)
 - Redistributed routes
- Enabled by default

© 2015 Cisco and/or its affiliates. All rights reserved.

69

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

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

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

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

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

Announcing Networks in BGP (Cont.)

```
router(config-router)# network major-network-number
```

To manually define a major network

- Allows advertising of major networks into BGP.
- At least one of the subnets must be present in the routing table.
- Behavior depends on the presence of the **auto-summary** command.
- The meaning of the **network** command in BGP is completely different from any other routing protocol.

© 2015 Cisco and/or its affiliates. All rights reserved.

70

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

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

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

Announcing Networks in BGP (Cont.)

```
router(config-router)# network major-network-number route-map route-map-name
```

- The addition of the **route-map** option allows network parameters to be modified before you enter them into the BGP table.
- The **route-map** option can be used for the following:
 - Changing the weight value of a locally sourced route
 - Tagging sourced routes with BGP communities
 - Setting the local preference for a specific network
 - Changing the value of the MED for a specific network

© 2015 Cisco and/or its affiliates. All rights reserved.

71

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

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

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

Announcing Networks in BGP Example

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

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

Redistributing Routes into BGP

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

Redistributing Routes into BGP

- Easier than listing networks in BGP process in large networks.
- Redistributed routes carry origin attribute "incomplete."
- Always filter redistributed routes to prevent route leaking.
- Avoid in service provider environments.

© 2015 Cisco and/or its affiliates. All rights reserved.

72

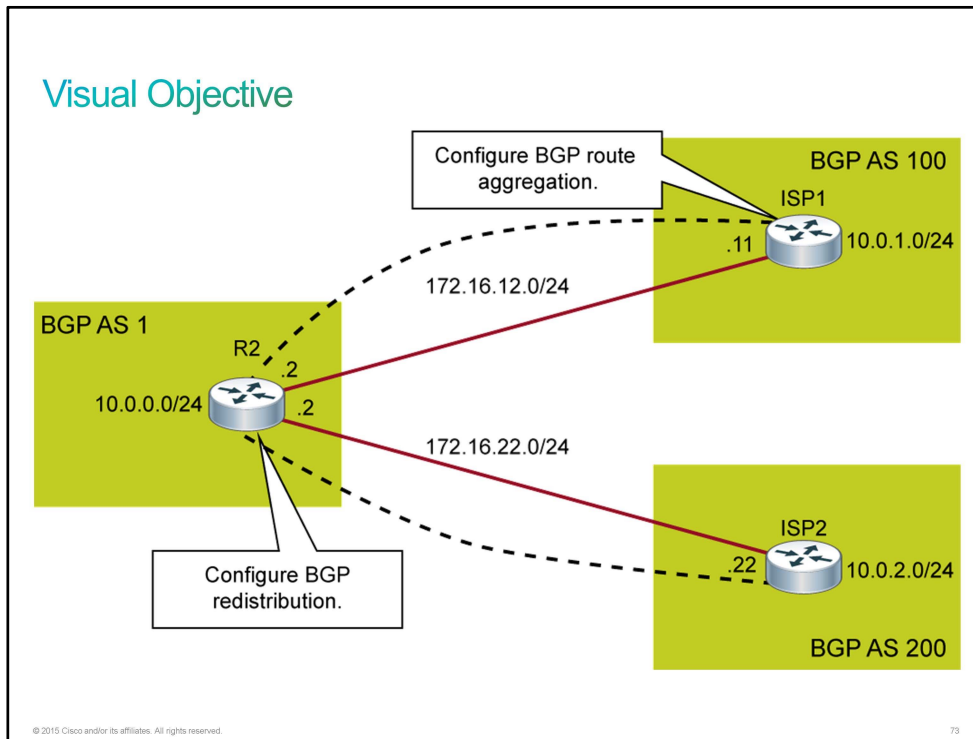
If there are a lot of networks to be advertised, and BGP is used primarily to achieve scalability, not routing security (for example, in enterprise networks), it could be easier to let the local [IGP](#) find the routes and then redistribute them into BGP. However, this approach introduces the risk that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

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

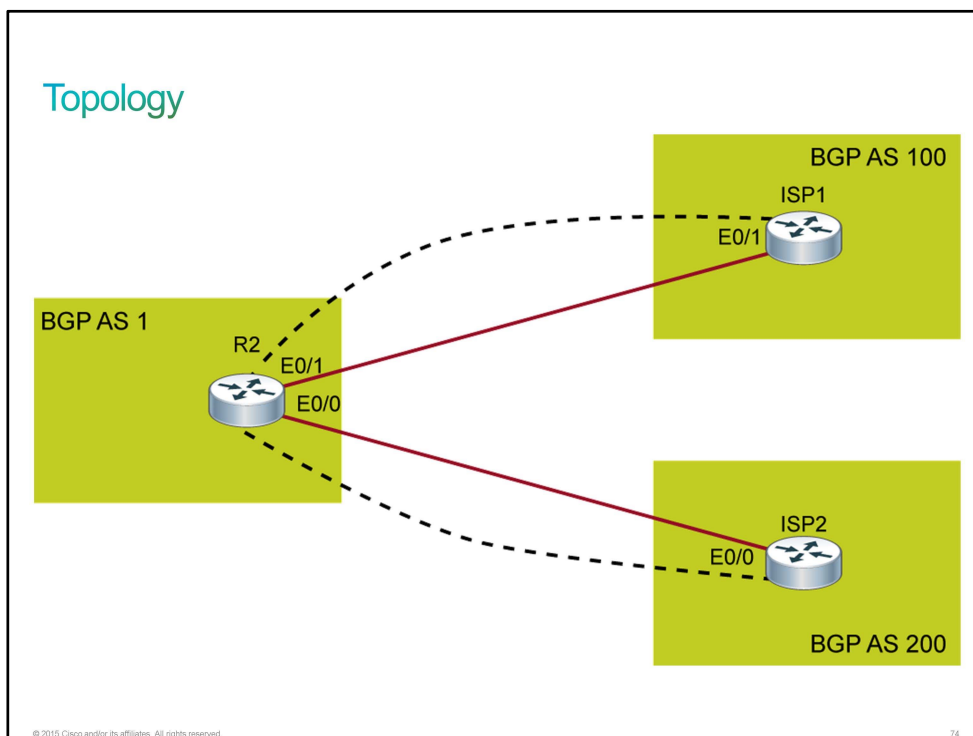
Discovery 2: Announcing Networks in BGP

Overview

Through this discovery, you will learn how to configure [BGP](#) redistribution and BGP route aggregation.



Topology

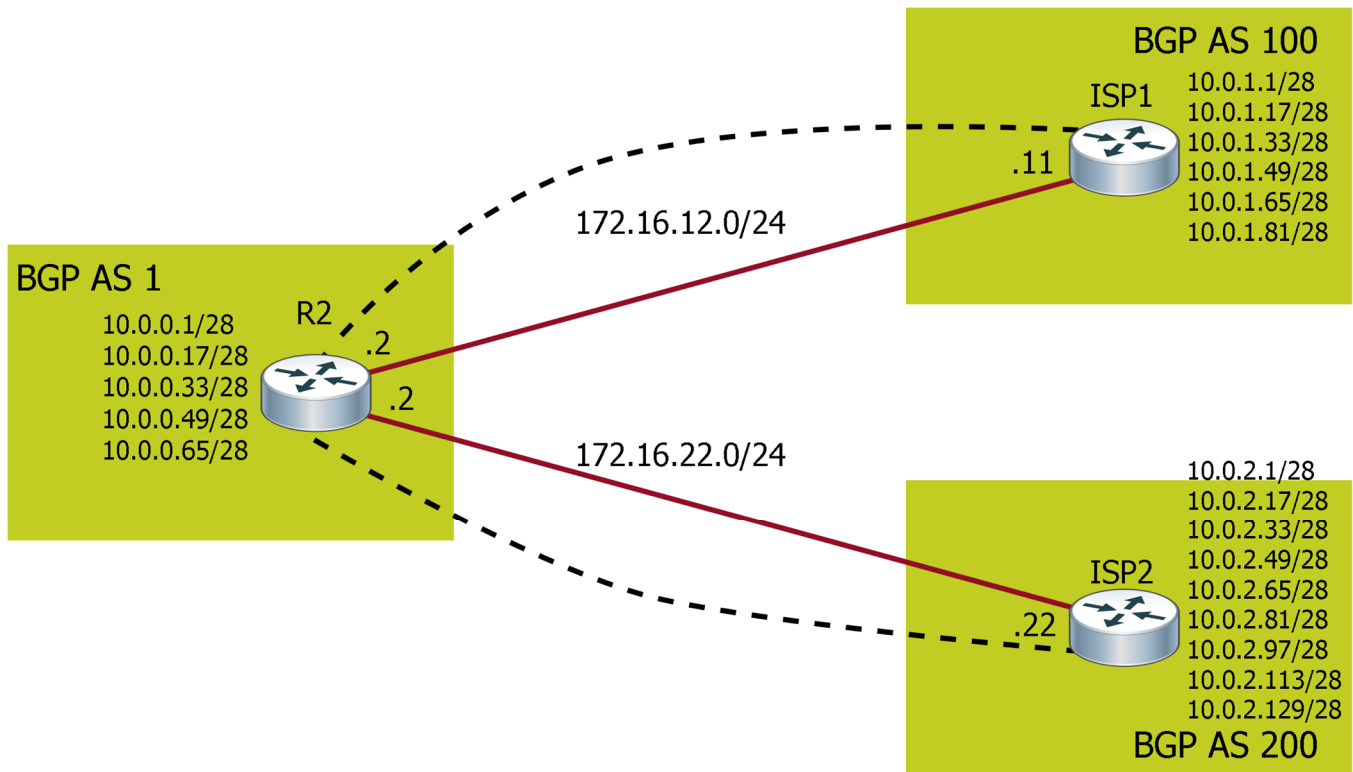


Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks



Announcing Networks in BGP

Discovery Steps

Step 1 You will redistribute connected routes into BGP. R2 router is preconfigured with five Loopback interfaces and is already running BGP in the [AS 1](#).

Initially, verify which Loopback interfaces are on the R2 router. Read Loopback interface's [IP...](#) addresses and subnet masks. You will need this information when configuring access list.

```
R2# show ip interface | include Loopback|Internet address
< text omitted >
Loopback1 is up, line protocol is up
  Internet address is 10.0.0.1/28
Loopback2 is up, line protocol is up
  Internet address is 10.0.0.17/28
Loopback3 is up, line protocol is up
  Internet address is 10.0.0.33/28
Loopback4 is up, line protocol is up
  Internet address is 10.0.0.49/28
Loopback5 is up, line protocol is up
  Internet address is 10.0.0.65/28
```

On the R2 router, redistribute connected routes into BGP:

```
R2(config)# router bgp 1
R2(config-router)# redistribute connected
```

This way, you will include all connected interfaces, including Loopback interfaces, into the BGP

routing table.

Step 2 On the R2 router, configure access list that permits even-numbered Loopback interfaces (Loopback 2 and Loopback 4).

Use access list number 1.

```
R2(config)# access-list 1 permit 10.0.0.16
R2(config)# access-list 1 permit 10.0.0.48
```

Step 3 On the R2 router, make sure that only even-numbered Loopback interfaces are redistributed into BGP.

```
R2(config)# router bgp 1
R2(config-router)# distribute-list 1 out connected
```

Step 4 Clear all BGP sessions on the R2 router.

```
R2# clear ip bgp *
```

Step 5 On the R2 router, verify that two routes are added into BGP routing table.

```
R2# show ip bgp
BGP table version is 18, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.16/28	0.0.0.0	0		32768	?
*>	10.0.0.48/28	0.0.0.0	0		32768	?
*>	10.0.1.0/28	172.16.12.11	0		0	100 i
*>	10.0.1.16/28	172.16.12.11	0		0	100 i
*>	10.0.1.32/28	172.16.12.11	0		0	100 i
*>	10.0.1.48/28	172.16.12.11	0		0	100 i
*>	10.0.1.64/28	172.16.12.11	0		0	100 i
*>	10.0.1.80/28	172.16.12.11	0		0	100 i
*>	10.0.2.0/28	172.16.22.22	0		0	200 i
*>	10.0.2.16/28	172.16.22.22	0		0	200 i
*>	10.0.2.32/28	172.16.22.22	0		0	200 i
*>	10.0.2.48/28	172.16.22.22	0		0	200 i
*>	10.0.2.64/28	172.16.22.22	0		0	200 i
*>	10.0.2.80/28	172.16.22.22	0		0	200 i
*>	10.0.2.96/28	172.16.22.22	0		0	200 i
*>	10.0.2.112/28	172.16.22.22	0		0	200 i
*>	10.0.2.128/28	172.16.22.22	0		0	200 i

When you see next hop is 0.0.0.0, you know that these routes are locally inserted into BGP routing.

Also observe that origin code is incomplete for routes that are inserted by R2 router.

Redistribution Using Route Maps

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

Redistribution Using Route Maps

- Origin can be set to "IGP" with a route-map.
- Other BGP path attributes can also be set:
 - Metric
 - Next-hop
 - Community

© 2015 Cisco and/or its affiliates. All rights reserved.

75

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

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

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

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

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

Step 6 On the R2 router, configure route map, where you will match based on previously configured access list and set origin code to IGP.

```
R2(config)# route-map evenLoop permit 10
R2(config-route-map)# match ip address 1
R2(config-route-map)# set origin igp
```

Step 7 On the R2 router, redistribute connected routes into BGP by using previously configured route map.

```
R2(config)# router bgp 1
R2(config-router)# redistribute connected route-map evenLoop
```

Step 8 On the R2 router, verify that origin code of two added routes in the BGP routing table has changed into IGP.

```
R2# show ip bgp | include 0.0.0.0
*> 10.0.0.16/28      0.0.0.0          0          32768 i
*> 10.0.0.48/28     0.0.0.0          0          32768 i
```

Configuring Classless BGP

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

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

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

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

Configuring Classless BGP

- BGP supports CIDR.
- Any BGP router can advertise individual networks or supernets (prefixes).
- Prefix notation is used with BGP instead of subnet masks.
 - 192.168.0.0/16 = 192.168.0.0 255.255.0.0

© 2015 Cisco and/or its affiliates. All rights reserved.

76

Step 9 On the R2 router, configure Loopback 1 interface subnet to be advertised into BGP. The Loopback 1 interface has a classless prefix preconfigured (10.0.0.0/28).

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

```
R2(config)# router bgp 1
R2(config-router)# network 10.0.0.0 mask 255.255.255.240
```

Step 10 On the R2 router, verify that Loopback 1 interface subnet is present in the BGP routing table.

```

R2# show ip bgp
BGP table version is 21, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.12.11	0		0 100	i
*>	10.0.1.16/28	172.16.12.11	0		0 100	i
*>	10.0.1.32/28	172.16.12.11	0		0 100	i
*>	10.0.1.48/28	172.16.12.11	0		0 100	i
*>	10.0.1.64/28	172.16.12.11	0		0 100	i
*>	10.0.1.80/28	172.16.12.11	0		0 100	i
*>	10.0.2.0/28	172.16.22.22	0		0 200	i
*>	10.0.2.16/28	172.16.22.22	0		0 200	i
*>	10.0.2.32/28	172.16.22.22	0		0 200	i
*>	10.0.2.48/28	172.16.22.22	0		0 200	i
*>	10.0.2.64/28	172.16.22.22	0		0 200	i
*>	10.0.2.80/28	172.16.22.22	0		0 200	i
*>	10.0.2.96/28	172.16.22.22	0		0 200	i
*>	10.0.2.112/28	172.16.22.22	0		0 200	i
*>	10.0.2.128/28	172.16.22.22	0		0 200	i

```

R2# show ip bgp 10.0.0.0/28
BGP routing table entry for 10.0.0.0/28, version 21
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    3
  Refresh Epoch 1
  Local
    0.0.0.0 from 0.0.0.0 (10.0.0.65)
      Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best

```

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

Step 11 On the R2 router, advertise subnet 10.0.0.0/24 into BGP. Subnet 10.0.0.0/24 is representing all Loopback interfaces that are configured on R2 router.

```

R2(config)# router bgp 1
R2(config-router)# network 10.0.0.0 mask 255.255.255.0

```

Step 12 On the R2 router, verify that subnet 10.0.0.0/24 is present in the BGP routing table.

```

R2# show ip bgp 10.0.0.0/24
% Network not in table

```

Route is not present in the BGP routing table.

The **network** command with the **mask** option tells BGP that 10.0.0.0/24 is a candidate for being advertised. The mask keyword and the mask 255.255.255.0 are required because the mask is not

the natural one. However, before the candidate route is actually advertised, the router checks the routing table for an exact match (both network number and mask). It will always be found because there is a static route for it. This static route points to the null interface, which is always available.

The conclusion is that 10.0.0.0/24 will always be advertised by this router. All other BGP routers will use this information and forward any IP packets with the destination IP address in the interval 10.0.0.0 to 10.0.0.255 (inclusive) in the direction of this router. When those packets arrive, the router, in this example, must have more explicit routes to the different parts of the 10.0.0.0/24 address range. IGP could answer this need but it is not shown in this configuration example.

Step 13 On the R2 router, configure explicit static route for 10.0.0.0/24 pointing to the null interface.

```
R2(config)# ip route 10.0.0.0 255.255.255.0 Null 0
```

If, however, an IP packet arrives with a destination address to which this router does not have a more explicit route, the static route will route the packet to the null interface, where it is dropped. This routing is a safety precaution that will prevent a routing loop, which might occur when route summaries are used in combination with default routing. Suppose that a packet arrives from the Internet to a subnet of 10.0.0.0/24, which is currently not reachable. In this case, the packet might follow the default route toward the Internet because there is no more explicit route. Of course, the packet would immediately be routed back again, and a routing loop would occur.

Step 14 On the R2 router, verify that subnet 10.0.0.0/24 is now present in the BGP routing table.

```
R2# show ip bgp 10.0.0.0/24
BGP routing table entry for 10.0.0.0/24, version 22
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    3
  Refresh Epoch 1
  Local
    0.0.0.0 from 0.0.0.0 (10.0.0.65)
      Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
```

Route 10.0.0.0/24 is present in the BGP routing table.

Aggregating BGP Networks

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

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

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

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

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

Aggregating BGP Networks

Summarization is called "aggregation" in BGP.

- Aggregation creates summary routes (called "aggregates") from networks already in BGP table.
- Individual networks can be announced or suppressed.

© 2015 Cisco and/or its affiliates. All rights reserved.

77

Step 15 On the ISP1 router, configure BGP aggregation on the subnet 10.0.1.0/24.

Do not use **summary-only** keyword, to observe both the route summary, and the more specific routes that will be advertised.

```
ISP1(config)# router bgp 100
ISP1(config-router)# aggregate-address 10.0.1.0 255.255.255.0
```

Step 16 On the ISP1 router, verify presence of aggregated route 10.0.1.0/24 in the BGP routing table.

ISP1# show ip bgp

BGP table version is 73, local router ID is 10.0.1.81

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	172.16.12.2	0		0	1 i
*>	10.0.0.0/24	172.16.12.2	0		0	1 i
*>	10.0.0.16/28	172.16.12.2	0		0	1 i
*>	10.0.0.48/28	172.16.12.2	0		0	1 i
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.0/24	0.0.0.0			32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.2.0/28	172.16.12.2			0	1 200 i
*>	10.0.2.16/28	172.16.12.2			0	1 200 i
*>	10.0.2.32/28	172.16.12.2			0	1 200 i
*>	10.0.2.48/28	172.16.12.2			0	1 200 i
*>	10.0.2.64/28	172.16.12.2			0	1 200 i
*>	10.0.2.80/28	172.16.12.2			0	1 200 i
*>	10.0.2.96/28	172.16.12.2			0	1 200 i
*>	10.0.2.112/28	172.16.12.2			0	1 200 i
*>	10.0.2.128/28	172.16.12.2			0	1 200 i

Step 17 On the R2 router, verify presence of aggregated route 10.0.1.0/24 and more specific routes in the BGP routing table.

R2# show ip bgp

BGP table version is 37, local router ID is 10.0.0.65

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.12.11	0		0	100 i
*>	10.0.1.0/24	172.16.12.11	0		0	100 i
*>	10.0.1.16/28	172.16.12.11	0		0	100 i
*>	10.0.1.32/28	172.16.12.11	0		0	100 i
*>	10.0.1.48/28	172.16.12.11	0		0	100 i
*>	10.0.1.64/28	172.16.12.11	0		0	100 i
*>	10.0.1.80/28	172.16.12.11	0		0	100 i
*>	10.0.2.0/28	172.16.22.22	0		0	200 i
*>	10.0.2.16/28	172.16.22.22	0		0	200 i
*>	10.0.2.32/28	172.16.22.22	0		0	200 i
*>	10.0.2.48/28	172.16.22.22	0		0	200 i
*>	10.0.2.64/28	172.16.22.22	0		0	200 i
*>	10.0.2.80/28	172.16.22.22	0		0	200 i
*>	10.0.2.96/28	172.16.22.22	0		0	200 i
*>	10.0.2.112/28	172.16.22.22	0		0	200 i
*>	10.0.2.128/28	172.16.22.22	0		0	200 i

This approach is generally not desirable. Therefore, suppression of individual routes, described next, is used in most cases.

Step 18 On the ISP1 router, configure BGP aggregation on the subnet 10.0.1.0/24 by using keyword **summary-only**.

```
ISP1(config)# router bgp 100
ISP1(config-router)# aggregate-address 10.0.1.0 255.255.255.0 summary-only
```

Step 19 On the ISP1 router, verify presence of aggregated route 10.0.1.0/24 in the BGP routing table.

```
ISP1# show ip bgp
BGP table version is 79, local router ID is 10.0.1.81
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	172.16.12.2	0		0 1	i
*>	10.0.0.0/24	172.16.12.2	0		0 1	i
*>	10.0.0.16/28	172.16.12.2	0		0 1	i
*>	10.0.0.48/28	172.16.12.2	0		0 1	i
s>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.0/24	0.0.0.0			32768	i
s>	10.0.1.16/28	0.0.0.0	0		32768	i
s>	10.0.1.32/28	0.0.0.0	0		32768	i
s>	10.0.1.48/28	0.0.0.0	0		32768	i
s>	10.0.1.64/28	0.0.0.0	0		32768	i
s>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.2.0/28	172.16.12.2			0 1 200	i
*>	10.0.2.16/28	172.16.12.2			0 1 200	i
*>	10.0.2.32/28	172.16.12.2			0 1 200	i
*>	10.0.2.48/28	172.16.12.2			0 1 200	i
*>	10.0.2.64/28	172.16.12.2			0 1 200	i
*>	10.0.2.80/28	172.16.12.2			0 1 200	i
*>	10.0.2.96/28	172.16.12.2			0 1 200	i
*>	10.0.2.112/28	172.16.12.2			0 1 200	i
*>	10.0.2.128/28	172.16.12.2			0 1 200	i

The prefix 10.0.1.0/24 is injected because there is at least one more specific route within the summary range. In this case 10.0.1.0/28, 10.0.1.16/28, 10.0.1.32/28, 10.0.1.48/28, 10.0.1.64/28, and 10.0.1.80/28 are within the range. The more specific routes are marked as suppressed using the lowercase letter "s." The "s" means that they are still present and available in the BGP table of the router, but they are not advertised on any BGP session.

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

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

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

them.

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

Step 20 On the R2 router, verify that only BGP aggregated route is received from ISP1 router.

You will not see more specific routes (10.0.1.0/28, 10.0.1.16/28, 10.0.1.32/28, 10.0.1.48/28, 10.0.1.64/28, or 10.0.1.80/28) advertised from ISP1 router.

R2# **show ip bgp**

BGP table version is 29, local router ID is 10.0.0.65

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.1.0/24	172.16.12.11	0		0 100	i
*>	10.0.2.0/28	172.16.22.22	0		0 200	i
*>	10.0.2.16/28	172.16.22.22	0		0 200	i
*>	10.0.2.32/28	172.16.22.22	0		0 200	i
*>	10.0.2.48/28	172.16.22.22	0		0 200	i
*>	10.0.2.64/28	172.16.22.22	0		0 200	i
*>	10.0.2.80/28	172.16.22.22	0		0 200	i
*>	10.0.2.96/28	172.16.22.22	0		0 200	i
*>	10.0.2.112/28	172.16.22.22	0		0 200	i
*>	10.0.2.128/28	172.16.22.22	0		0 200	i

BGP Conditional Route Injection

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

BGP Conditional Route Injection

The BGP Conditional Route Injection feature has the following functions:

- It provides means to originate a prefix into a BGP routing table without the corresponding match.
- It allows more specific routes to be generated based on administrative policy or traffic engineering information to provide more specific control over the forwarding of packets to these more specific routes, which are injected into the BGP routing table only if the configured conditions are met.
- It improves accuracy of common route aggregation by conditionally injecting or replacing less specific prefixes with more specific prefixes.

© 2015 Cisco and/or its affiliates. All rights reserved.78

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

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

BGP Conditional Route Injection Example

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

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

BGP Support for TTL Security Check

The BGP Support for TTL Security Check feature introduces a lightweight security mechanism to protect EBGp peering sessions from CPU utilization-based attacks. These types of attacks are typically brute-force DoS attacks that attempt to disable the network by flooding the network with IP packets that contain forged source and destination IP addresses. This feature protects the EBGp peering session by comparing the value in the TTL field of received IP packets against a hop count. The hop count is configured locally for each EBGp peering session. If the value in the TTL field of the incoming IP packet is greater than or equal to the locally configured value, the IP packet is accepted and processed normally. If the TTL value in the IP packet is less than the locally configured value, the packet is silently discarded and no ICMP message is generated. Not responding to a forged packet is part of a designed behavior that helps prevent CPU utilization-based attacks.

BGP Support for TTL Security Check

BGP Support for TTL Security Check feature has the following characteristics:

- Lightweight security mechanism to protect EBGp peering sessions from CPU utilization-based attacks.
- Protects the EBGp peering session by comparing the value in the TTL field of received IP packets against a hop count that is configured locally for each EBGp peering session.
- Supports both directly connected peering sessions and multihop EBGp peering sessions.

© 2015 Cisco and/or its affiliates. All rights reserved.

79

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

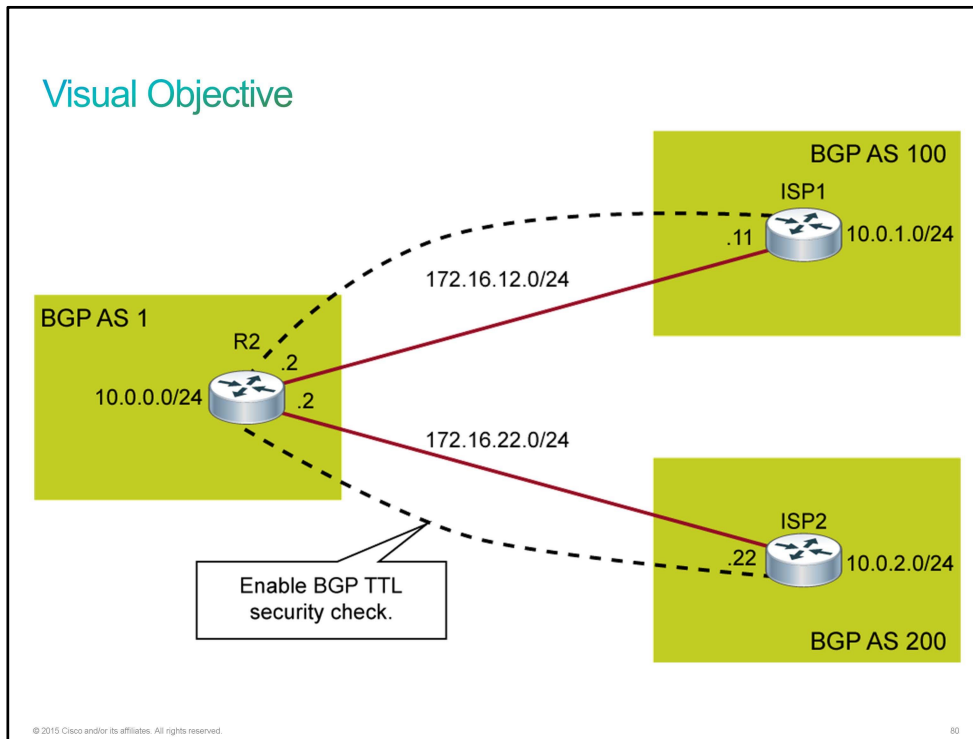
This feature supports both directly connected peering sessions and multihop EBGp peering sessions. The incoming packets that contain invalid TTL values do not affect the BGP peering session. The BGP peering session remains open, and the router silently discards the invalid packet. The BGP session, however,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

The BGP Support for TTL Security Check feature should be configured on each participating router. It provides an effective and easy-to-deploy solution to protect EBGp peering sessions from CPU utilization-based attacks. When this feature is enabled, a host cannot attack a BGP session if the host is not a member of the local or remote BGP network or if the host is not directly connected to a network segment between the local and remote BGP networks. This solution greatly reduces the effectiveness of DoS attacks against a BGP AS.

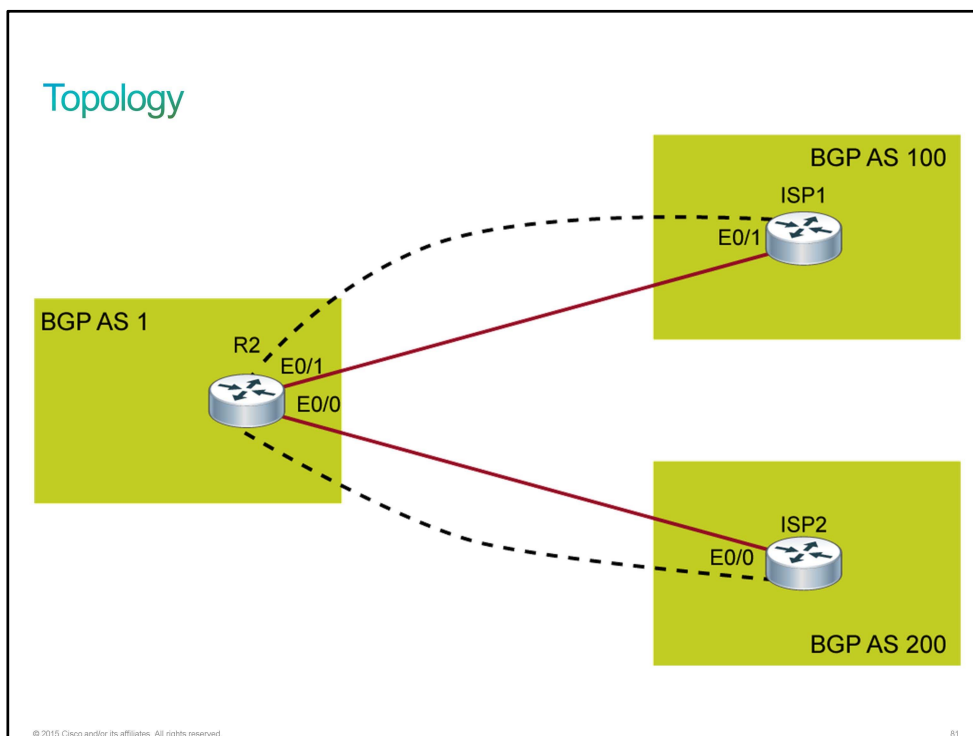
Discovery 3: Implement BGP TTL Security Check

Overview

Through this discovery, you will learn how to enable [BGP TTL](#) security check.

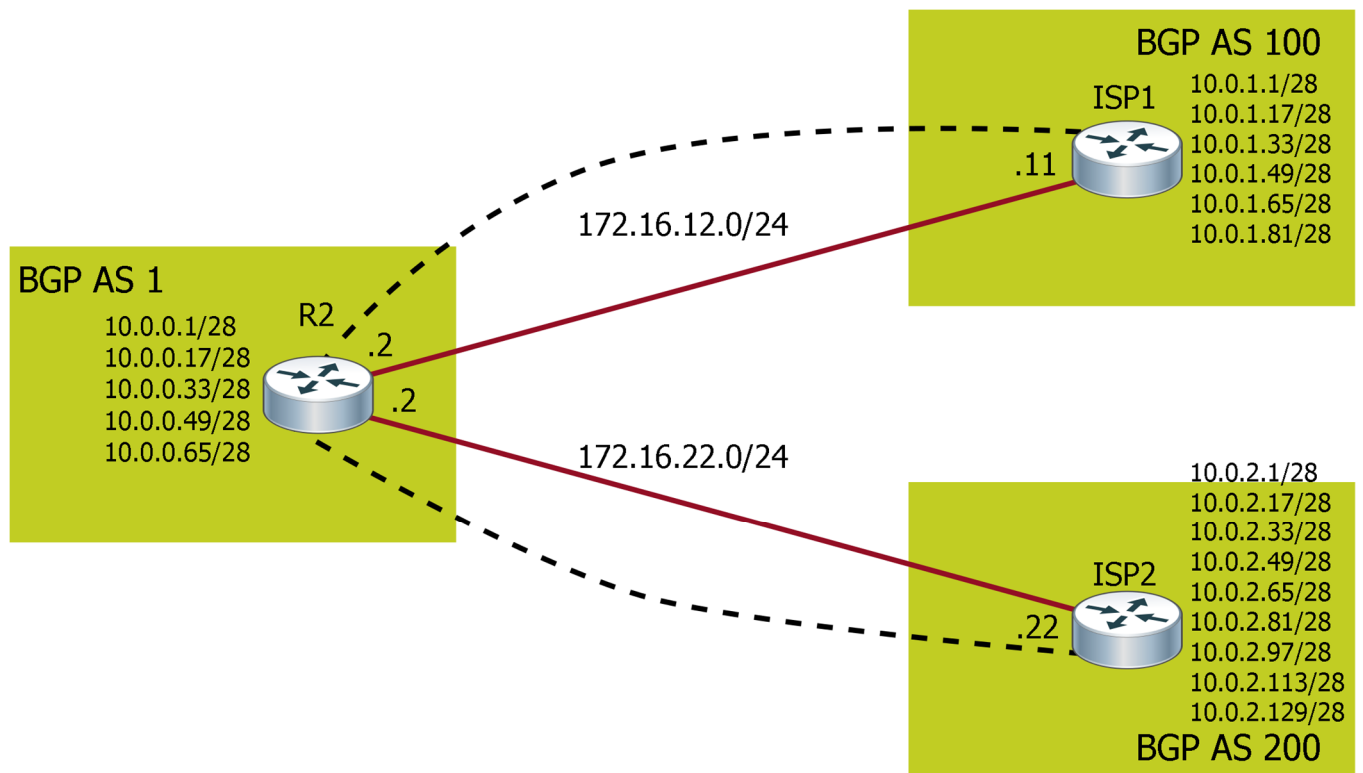


Topology



Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."



Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks

Implement BGP TTL Security Check

Discovery Steps

Step 1 To secure a BGP peering session between R2 and ISP2 routers configure BGP TTL security check.

On the R2 router, configure BGP TTL security check to the ISP2 with hop count 10.

```
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.22.22 ttl-security hops 10
```

BGP session between R2 and ISP2 routers will go down.

Router R2 will start sending BGP packets with TTL 255 and will accept BGP packets with 245 or more.

Step 2 On the ISP2 router, configure BGP TTL security check to the R2 with hop count 10.

```
ISP2(config)# router bgp 200
ISP2(config-router)# neighbor 172.16.22.2 ttl-security hops 10
```

BGP session between R2 and ISP2 routers will go up.

```
%BGP-5-ADJCHANGE: neighbor 172.16.22.2 Up
```

Step 3 On the ISP2 router, verify what TTL values are used for BGP neighbor.

```
ISP2# show ip bgp neighbors 172.16.22.2 | include TTL
Connection is ECN Disabled, Minimum incoming TTL 245, Outgoing TTL 255
```

You can see BGP packets are sent with TTL number 255 and are accepted with TTL value 245 or more.

Step 4 On the R2 router, verify what TTL values are used for BGP neighbor.

```
R2# show ip bgp neighbors 172.16.22.22 | include TTL
Connection is ECN Disabled, Minimum incoming TTL 245, Outgoing TTL 255
```

You can see BGP packets are sent with TTL number 255 and are accepted with TTL value 245 or more.

Step 5 On the R2 router, verify that BGP session to the ISP2 is established.

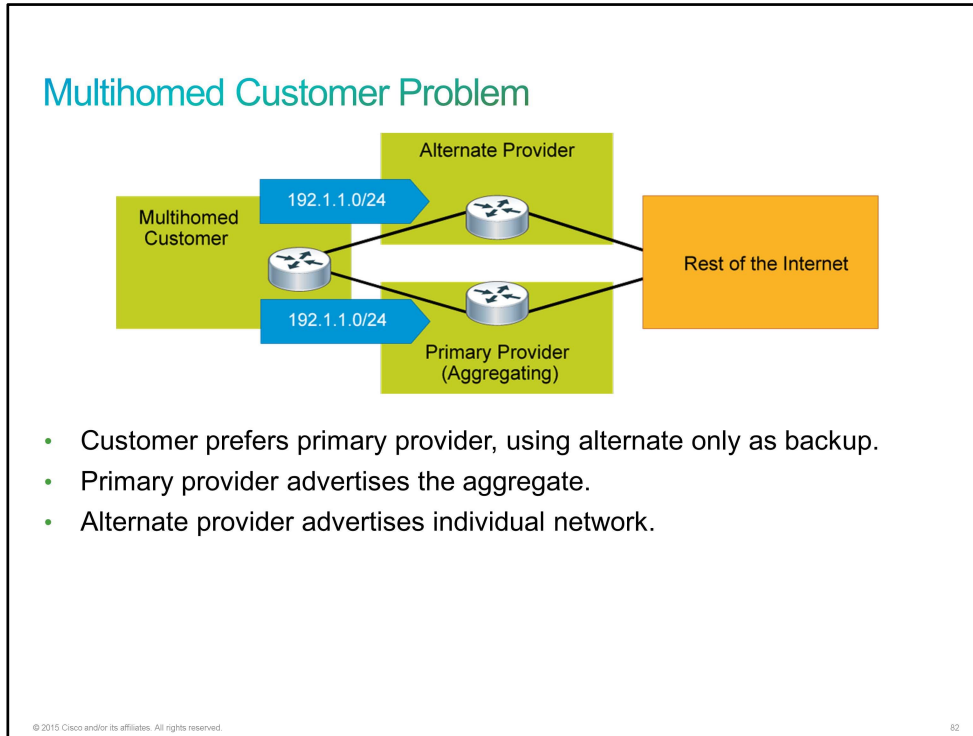
```
R2# show ip bgp summary
BGP router identifier 10.0.0.65, local AS number 1
BGP table version is 63, main routing table version 63
14 network entries using 2072 bytes of memory
14 path entries using 896 bytes of memory
3/3 BGP path/bestpath attribute entries using 408 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3424 total bytes of memory
BGP activity 38/24 prefixes, 38/24 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	241	247	63	0	0	00:37:51	1
172.16.22.22	4	200	19	23	63	0	0	00:12:47	9

BGP session between R2 and ISP2 routers is secured with BGP TTL security check.

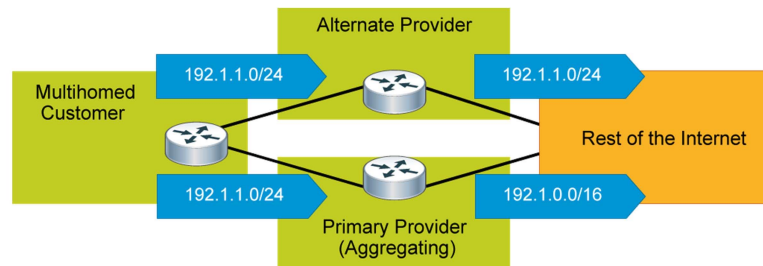
Multihomed Customer Problem

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



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

Multihomed Customer Problem (Cont.)



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

© 2015 Cisco and/or its affiliates. All rights reserved.

83

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

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

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

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- The BGP process in a Cisco router is started with the **router bgp** command and BGP neighbors are added with **neighbor** command.
- MD5 authentication can be used to secure a connection between two BGP neighbors. BGP also supports TTL security check feature, which protects EBGP sessions from CPU utilization-based attacks.
- Local networks are announced in BGP by listing them with the **network** command or by redistributing them.
- BGP supports aggregation but BGP route aggregation is not appropriate in multihomed topologies.

© 2015 Cisco and/or its affiliates. All rights reserved.

84

Lesson 6: Monitoring and Troubleshooting BGP

Overview

To ensure that basic [BGP](#) configurations are operating correctly, you need to be familiar with BGP monitoring commands. If basic BGP configurations are not functioning as expected, BGP troubleshooting skills are critical to successful problem resolution.

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

Upon completing this lesson, you will be able to:

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

Monitoring Overall BGP Routing

To display the status of all [BGP](#) connections, use the **show ip bgp summary** EXEC command.

This command has no arguments or keywords.

Monitoring Overall BGP Routing

```
router> show ip bgp summary
```

- Displays BGP memory use, and displays BGP neighbors and the state of communication with them

```
R2# show ip bgp summary
BGP router identifier 172.16.22.2, local AS number 1
BGP table version is 10, main routing table version 10
9 network entries using 1332 bytes of memory
9 path entries using 576 bytes of memory
1/1 BGP path/bestpath attribute entries using 136 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 2068 total bytes of memory
BGP activity 9/0 prefixes, 9/0 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
172.16.12.11	4	100	0	0	1	0	0	never	Idle
172.16.22.22	4	200	63	62	10	0	0	00:54:24	9

© 2015 Cisco and/or its affiliates. All rights reserved. 85

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

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

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

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

- IP address of the neighbor as configured in the local router
- BGP version number that the router uses when communicating with the neighbor
- [AS](#) number of the remote neighbor
- Number of messages and updates that have been received from the neighbor since the session was established
- Number of messages and updates that have been sent to the neighbor since the session was established

- Version number of the local BGP table that has been included in the most recent update to the neighbor
- Number of messages that are waiting to be processed in the incoming queue from this neighbor
- Number of messages that are waiting in the outgoing queue for transmission to the neighbor
- How long the neighbor has been in the current state and the name of the current state (the state "Established" is not printed out, so no state name indicates "Established")

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

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

Many messages in the incoming queue indicate a lack of CPU resources in the local router. Many messages in the outgoing queue indicate a lack of bandwidth to the remote router or a lack of CPU resources in the remote router.

Monitoring BGP Neighbors

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

Monitoring BGP Neighbors

```
router> show ip bgp neighbors ip-address
```

- Displays detailed neighbor information

```
R2# show ip bgp neighbors 172.16.12.11
BGP neighbor is 172.16.12.11, remote AS 100, external link
  BGP version 4, remote router ID 10.0.1.81
  BGP state = Established, up for 00:03:10
  Last read 00:00:26, last write 00:00:31, hold time is 180, keepalive interval
  is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
    Multisession Capability:
    Stateful switchover support enabled: NO for session 1
```

© 2015 Cisco and/or its affiliates. All rights reserved. 86

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

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

Monitoring BGP Table

To display entries in the [BGP](#) routing table, use the **show ip bgp** EXEC command.

Monitoring BGP Table

```
router> show ip bgp
```

- Displays all routes in the BGP table in summary format

```
R2# show ip bgp
```

```
BGP table version is 18, local router ID is 172.16.22.2
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,  
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,  
               x best-external, a additional-path, c RIB-compressed,
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	1.0.0.0/24	172.16.22.22	0		0	200 i
*		172.16.12.11	0		0	100 99 i
*>	10.0.1.0/28	172.16.12.11	0		0	100 i
*>	10.0.1.16/28	172.16.12.11	0		0	100 i
*>	10.0.1.32/28	172.16.12.11	0		0	100 i
*>	10.0.1.48/28	172.16.12.11	0		0	100 i
*>	10.0.1.64/28	172.16.12.11	0		0	100 i
*>	10.0.1.80/28	172.16.12.11	0		0	100 i
*>	10.0.2.0/28	172.16.22.22	0		0	200 i

<... output omitted ...>

© 2015 Cisco and/or its affiliates. All rights reserved.

87

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

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

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

Monitoring the BGP Table (Cont.)

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

- Displays detailed information about all paths for a single prefix

```
R2# show ip bgp 1.0.0.1
BGP routing table entry for 1.0.0.0/24, version 18
Paths: (2 available, best #1, table default)
  Advertised to update-groups:
    1
  Refresh Epoch 1
  200
    172.16.22.22 from 172.16.22.22 (10.0.2.129)
      Origin IGP, metric 0, localpref 100, valid, external, best
  Refresh Epoch 1
  100 99
    172.16.12.11 from 172.16.12.11 (10.0.1.81)
      Origin IGP, metric 0, localpref 100, valid, external
```

- BGP attributes are listed for each route: AS-Path, Next-Hop, advertising router IP address, advertising router Router-ID.
- The bottom row lists other BGP attributes.

© 2015 Cisco and/or its affiliates. All rights reserved.

88

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

In this example, the information about network 1.0.0.1 is displayed. There are two routes to 1.0.0.1. One is received from neighbor 172.16.12.11 and the other from 172.16.22.22.

The BGP route selection process has selected this route as the best. So, BGP will try to install this route into the routing table. Installation of routes in the routing table is made based on the [AD](#).

Debugging BGP

There are several Cisco IOS commands that you can use to perform debugging of basic [BGP](#) configurations.

Debugging BGP

```
router# debug ip tcp transactions
```

- Displays all TCP transactions (start of session, session errors, etc.)

```
router# debug ip bgp events
```

- Displays significant BGP events (neighbor state transitions, update runs)

© 2015 Cisco and/or its affiliates. All rights reserved. 89

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

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

Debugging BGP (Cont.)

```
router# debug ip bgp keepalives
```

- Debugs BGP keepalive packets

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

- Displays all incoming or outgoing BGP updates
- *Use with caution*

© 2015 Cisco and/or its affiliates. All rights reserved.90

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

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

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

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

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

Debugging BGP (Cont.)

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

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

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

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

© 2015 Cisco and/or its affiliates. All rights reserved.

91

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

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

BGP Session Startup Problems

There are several common session startup issues that you can experience when configuring basic [BGP](#).

BGP Session Startup Problems

Common BGP session startup symptoms:

- BGP neighbors do not become active.
- BGP neighbor is active, but the session is never established.
- BGP neighbor oscillates between idle and active.

BGP Neighbor Not Reachable

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

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

BGP Neighbor Not Reachable

Symptom:

- BGP neighbors do not become active.
 - **show ip bgp neighbors** displays the neighbor state as Idle for several minutes.

Diagnosis:

- Neighbor is not directly connected.

Verification:

- Verify with **show ip route**.

© 2015 Cisco and/or its affiliates. All rights reserved.93

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

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

BGP Neighbor Not Reachable (Cont.)

Symptom:

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

Diagnosis:

- Neighbor is not reachable.

Verification:

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

© 2015 Cisco and/or its affiliates. All rights reserved.

94

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

BGP Neighbor Not Reachable Example

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

BGP Neighbor Not Reachable Example

```
ISP1# debug ip tcp transactions
TCP special event debugging is on
<... output omitted ...>
*Feb 6 05:04:01: TCBEFE6AF18 created
*Feb 6 05:04:01: TCBEFE6AF18 setting property TCP_VRFTABLEID (20) EC3C8C6C
<... output omitted ...>
*Feb 6 05:04:01: TCP: Random local port generated 43645, network 1
*Feb 6 05:04:01: TCBEFE6AF18 bound to 172.16.12.11.43645
*Feb 6 05:04:01: Reserved port 43645 in Transport Port Agent for TCP IP type 1
*Feb 6 05:04:01: TCP: sending SYN, seq 2161553223, ack 0
*Feb 6 05:04:01: TCP0: Connection to 172.16.12.22:179, advertising MSS 1460
*Feb 6 05:04:01: TCP0: state was CLOSED -> SYNSENT [43645 ->
172.16.12.22(179)]
<... output omitted ...>
*Feb 6 05:04:31: TCP0: state was SYNSENT -> CLOSED [43645 ->
172.16.12.22(179)]
*Feb 6 05:04:31: TCB 0xEFE6AF18 destroyed
```

- SYN packet is sent.
- SYN-ACK reply never comes back. TCP session is closed.

© 2015 Cisco and/or its affiliates. All rights reserved.

95

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

BGP Neighbor Not Configured

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

BGP Neighbor Not Configured

Symptom:

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

Diagnosis:

- This router is not configured as the BGP neighbor on the neighboring router.

Verification:

- Check IP addresses of BGP neighbors with **show ip bgp summary** on the neighboring router.

© 2015 Cisco and/or its affiliates. All rights reserved.

96

BGP Neighbor Not Configured Example

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

BGP Neighbor Not Configured Example

```
ISP2# debug ip tcp transactions
*Feb  9 00:35:12.675: TCBEFA28EF0 created
<... output omitted ...>
*Feb  9 00:35:12: TCBEFA28EF0 bound to 172.16.22.22.45207
*Feb  9 00:35:12: Reserved port 45207 in Transport Port Agent for TCP IP type 1
*Feb  9 00:35:12: TCP: sending SYN, seq 4270019552, ack 0
*Feb  9 00:35:12: TCP0: Connection to 172.16.22.2:179, advertising MSS 1460
*Feb  9 00:35:12: TCP0: state was CLOSED -> SYNSENT [45207 -> 172.16.22.2(179)]
*Feb  9 00:35:12: Released port 45207 in Transport Port Agent for TCP IP type 1
delay 240000
*Feb  9 00:35:12: TCP0: state was SYNSENT -> CLOSED [45207 -> 172.16.22.2(179)]
*Feb  9 00:35:12: TCP0: bad seg from 172.16.22.2 -- closing connection: port
45207 seq 0 ack 4270019553 rcvnxt 0 rcvwnd 0 len 0
*Feb  9 00:35:12: TCP0: connection closed - remote sent RST
*Feb  9 00:35:12: TCB 0xEFA28EF0 destroyed
```

- SYN packet is sent.
- Neighbor replies with an RST packet. TCP session is closed.

© 2015 Cisco and/or its affiliates. All rights reserved.

97

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

BGP AS Number Mismatch

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

BGP AS Number Mismatch

Symptom:

- BGP neighbor oscillates between Active and Idle.
 - **debug ip tcp transactions** displays the TCP session being established and torn down immediately.

Diagnosis:

- There is an AS number mismatch between BGP neighbors.

Verification:

- Verify the AS numbers that are configured for neighboring routers using the **show ip bgp summary** on both routers.

© 2015 Cisco and/or its affiliates. All rights reserved.

98

BGP AS Number Mismatch Example

This example illustrates a mismatch in an AS number.

BGP AS Number Mismatch Example

```
R2# debug ip tcp transaction
R2# debug ip bgp event
```

```
*Feb 6 06:37:44: TCBEE9C5A38 created
*Feb 6 06:37:44: TCP0: state was LISTEN -> SYNRCVD [179 ->
172.16.12.11(60291)]
*Feb 6 06:37:44: TCP: tcb EE9C5A38 connection to 172.16.12.11:60291, peer MSS
1460, MSS is 516
*Feb 6 06:37:44: TCP: sending SYN, seq 2355663235, ack 879693158
*Feb 6 06:37:44: TCP0: Connection to 172.16.12.11:60291, advertising MSS 1460
*Feb 6 06:37:44: TCP0: state was SYNRCVD -> ESTAB [179 -> 172.16.12.11(60291)]
*Feb 6 06:37:44: TCBEE9C4B20 accepting EE9C5A38 from 172.16.12.11.60291
*<... output omitted ...>
*Feb 6 06:37:44: %BGP-3-NOTIFICATION: sent to neighbor 172.16.12.11 passive
2/2 (peer in wrong AS) 2 bytes 0064
*<... output omitted ...>
*Feb 6 06:37:44: TCP0: FIN processed
*Feb 6 06:37:44: TCP0: state was ESTAB -> CLOSEWAIT [179 ->
172.16.12.11(60291)]
*<... output omitted ...>
*Feb 6 06:37:49: TCP0: sending FIN
```

- First, TCP session is established.
- BGP notification is sent because of AS number mismatch.

© 2015 Cisco and/or its affiliates. All rights reserved.

99

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

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

© 2015 Cisco and/or its affiliates. All rights reserved.

100

Lesson 7: Module Summary

Overview

This topic summarizes the key points that were discussed in this module.

Module Summary

- BGP has reliable transport provided by TCP, a rich set of metrics called BGP path attributes, and scalability features such as batched updates that make it suitable for very large networks.
- Configured BGP neighbors establish a TCP session and exchange the BGP Open message, which contains the parameters that each BGP router proposes to use.
- Some path attributes are well-known and should be recognized by every BGP implementation. Some of the well-known attributes, such as AS-path, next-hop, and origin, are mandatory and have to be present in every BGP update.
- The route selection process takes into account various BGP attributes that are attached to the route, as well as local decisions.
- When you are configuring BGP neighbors, you will enable the BGP routing protocol process, establish neighbors, and advertise local routes.

© 2015 Cisco and/or its affiliates. All rights reserved.

101

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Border Gateway Protocol*.
http://docwiki.cisco.com/wiki/Border_Gateway_Protocol

- Cisco Systems, Inc. *Configuring BGP*.
http://www.cisco.com/c/en/us/td/docs/ios/12_2/ip/configuration/guide/fipr_c/1cfbgp.html
- Cisco Systems, Inc. *Using the Border Gateway Protocol for Interdomain Routing*
http://docwiki.cisco.com/wiki/Internetworking_Case_Studies

Lesson 8: Module Self-Check

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

1. Which statement about the AS is true? (Source: "Introducing BGP")
 - A. The AS is a collection of networks under a single administrative domain.
 - B. The AS is a collection of networks that belong to different enterprise networks.
 - C. The AS requires IGP protocol to exchange routing information between autonomous systems.
 - D. The AS number can be any number between 0 and 255.

2. Which routing method best describes BGP? (Source: "Introducing BGP")
 - A. distance vector
 - B. link state
 - C. path vector
 - D. hybrid of link state and distance vector

3. A transit AS is an AS that exchanges BGP routing information with other autonomous systems and forwards information that is received from one AS to another AS. True or false? (Source: "Introducing BGP")
 - A. true
 - B. false

4. Which three items are BGP enhancements to traditional distance vector routing protocols? (Choose three.) (Source: "Introducing BGP")
- A. reliable updates
 - B. use of triggered updates only
 - C. enhanced security
 - D. rich metrics
 - E. route summarization
 - F. snapshot updates
5. Which protocol facilitates reliable update capabilities in BGP? (Source: "Introducing BGP")
- A. TCP
 - B. UDP
 - C. HSRP
 - D. ICMP
6. What are three characteristics of an AS? (Choose three.) (Source: "Introducing BGP")
- A. AS uses IGP for intradomain routing.
 - B. AS uses EGP for interdomain routing.
 - C. AS is a collection of networks under a common administrative authority.
 - D. AS consists of a group of network domains.
 - E. AS automatically summarizes addresses
 - F. AS is regulated by the IETF.
7. Which three scenarios are common scenarios where BGP is used? (Choose three.) (Source: "Introducing BGP")
- A. a customer with a connection to multiple service providers
 - B. service provider networks acting as transit systems and forwarding external traffic through their network
 - C. a single-site customer intranet with complex administrative policies between departments
 - D. as the core routing protocol in very large enterprise networks
 - E. as the routing protocol in an IS-IS backbone area
 - F. as the core routing protocol in an SNA network
8. What are three recommended BGP use guidelines for multihomed customer networks? (Choose three.) (Source: "Introducing BGP")
- A. Most multihomed customers should use BGP with their service providers.
 - B. Most multihomed customers should forward routing information that is received from one provider to the other provider.
 - C. The multihomed customer must have its own public AS number.
 - D. Multihomed customers should use a provider-independent, public address space.
 - E. The multihomed customer may use and advertise RFC 1918 addresses.
 - F. Multihomed customers should use the AS number of their primary ISP.

9. What is a limitation of the BGP routing protocol? (Source: "Introducing BGP")
- A. You cannot use BGP to implement hop-by-hop routing policy controls.
 - B. You cannot use BGP to influence the routing policy in a downstream AS.
 - C. BGP cannot control forwarding of packets based on their destination address.
 - D. BGP cannot scale to very large networks with more than 110,000 routes.
10. Which three statements are true of BGP mandatory well-known attributes? (Choose three.) (Source: "Understanding BGP Path Attributes")
- A. They must be present in all BGP updates.
 - B. All BGP-compliant implementations must recognize them.
 - C. All BGP-compliant routers must adhere to policies specified in mandatory attributes.
 - D. All well-known attributes are propagated to other neighbors.
 - E. They must be present in some BGP updates.
11. Which three attributes are BGP mandatory well-known attributes? (Choose three.) (Source: "Understanding BGP Path Attributes")
- A. next-hop
 - B. weight
 - C. AS-path
 - D. origin
 - E. MED
 - F. local preference
12. Which three possible values are assigned to the BGP origin attribute? (Choose three.) (Source: "Understanding BGP Path Attributes")
- A. IGP
 - B. EGP
 - C. unknown
 - D. internal
 - E. external
 - F. MED
13. Which nontransitive optional BGP attribute is useful in assisting with the route selection process when multiple links to another AS exist? (Source: "Understanding BGP Path Attributes")
- A. next-hop
 - B. local preference
 - C. MED
 - D. AS-path

14. In which two ways can the BGP next-hop attribute be modified? (Choose two.) (Source: "Understanding BGP Path Attributes")
- A. If the next-hop attribute is in the same IP subnet as the receiving router, the attribute is unchanged; otherwise, it is set to the IP address of the sending router.
 - B. The next-hop attribute is always set to the IP address of the sending router.
 - C. The next-hop attribute is modified only when BGP packets exit an AS.
 - D. The BGP next-hop attribute is modified only when BGP packets traverse point-to-point links.
15. Which three statements regarding the BGP AS-path attribute are true? (Choose three.) (Source: "Understanding BGP Path Attributes")
- A. The local AS number is prepended to the AS path each time that the route crosses an AS boundary.
 - B. The AS that originally injected the route into BGP is always found at the rightmost end of the AS path.
 - C. The AS-path attribute can be used to avoid routing loops.
 - D. BGP routes with an empty AS path were injected into BGP from outside the local AS.
 - E. The local AS number is appended to the end of the AS path each time that the route crosses an AS boundary.
 - F. The AS that originally injected the route into BGP is always found at the leftmost end of the AS path.
16. Which description applies to the local preference attribute? (Source: "Understanding BGP Path Attributes")
- A. well-known mandatory
 - B. well-known discretionary
 - C. optional transitive
 - D. optional nontransitive
17. Which two attributes are BGP optional transitive attributes? (Choose two.) (Source: "Understanding BGP Path Attributes")
- A. MED
 - B. weight
 - C. aggregator
 - D. atomic aggregate
 - E. community
18. What is indicated by a state of "Idle" in the output of the **show ip bgp summary** command? (Source: "Establishing BGP Sessions")
- A. The router is currently not attempting to establish a connection with a neighbor.
 - B. The connection to the configured neighbor has timed out.
 - C. The connection to a BGP neighbor has been established, and no errors have been received on the connection.
 - D. The connection to a BGP neighbor has been established, and no packets have been sent.

19. How are neighbors detected in BGP? (Source: "Establishing BGP Sessions")
- A. They are discovered automatically
 - B. They must be configured manually.
20. Router-ID is a number that uniquely identifies the router. How is it selected? (Source: "Establishing BGP Sessions")
- A. Router-ID is the highest IP address of any loopback interface. If there is no loopback interface, the router-ID must be configured manually.
 - B. Router-ID must be configured manually.
 - C. Router-ID is the highest IP address of any loopback interface. If there is no loopback interface, the router uses the highest IP address of any interface.
 - D. Router-ID is the lowest IP address of any loopback interface. If there is no loopback interface, the router uses the lowest IP address of any interface.
21. What happens if two TCP connection attempts between configured BGP neighbors succeed? (Source: "Establishing BGP Sessions")
- A. Both connections will be terminated, and the neighbors will re-establish a neighbor relationship.
 - B. One connection will be maintained as primary and the other as backup.
 - C. One of the two connections will be torn down
 - D. The router with the lower router-ID will determine if the second connection is torn down or used as a backup TCP connection.
22. Given the following BGP session states, in which state is the router while waiting for the other router to respond to the BGP Open message? (Source: "Establishing BGP Sessions")
- A. OpenConfirm
 - B. Established
 - C. Idle
 - D. OpenSent
 - E. Active
23. Given the following BGP session states, what is their order of progression during the creation of a successful neighbor session? (Source: "Establishing BGP Sessions")

OpenConfirm	1
Established	2
Active	3
Idle	4
OpenSent	5

24. BGP sessions are established over TCP using port number 179. True or false? (Source: "Establishing BGP Sessions")
- A. true
 - B. false
25. What does the field "TblVer" indicate in the output of the **show ip bgp summary** command? (Source: "Establishing BGP Sessions")
- A. the current version of BGP in use by the router
 - B. the number of route prefixes that are contained in the BGP update of the router
 - C. BGP messages that have been received from that neighbor
 - D. the last version of the BGP database that was sent to that neighbor
26. What occurs when you use MD5 between two BGP neighbors? (Source: "Establishing BGP Sessions")
- A. Every packet is encrypted with MD5.
 - B. The IP header is encrypted using MD5.
 - C. An MD5 checksum is calculated and sent with each packet so that its source can be verified.
 - D. A username and password are embedded in an IP datagram that is matched to a username and password on the remote neighbor.
27. What does a router that is running BGP do with a BGP update that contains its own AS path? (Source: "Processing BGP Routes")
- A. The router checks to see whether the information that is contained in the update is better than its current information. If it is, it will update its BGP table.
 - B. The router accepts the route update.
 - C. The router silently discards (denies) the route.
 - D. The router returns an error to the router that sent the update.
28. Which three attributes must always be present in the routing update? (Choose three.) (Source: "Processing BGP Routes")
- A. MED
 - B. AS-path
 - C. local-preference
 - D. next-hop
 - E. origin
29. When a router has more than one alternative route to reach the same destination, local preference is checked before the weight attribute. True or false? (Source: "Processing BGP Routes")
- A. true
 - B. false

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

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

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

The router checks whether the next-hop attribute indicates an IP address that is reachable according to the current routing table. If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best. Step 7

The router compares MED values, but only if it receives the updates from the same neighboring AS. Routes with a lower MED are preferred. Step 1

If the local preference attributes are different, the route with the highest value is selected as best. Step 2

If one of the routes is injected into the BGP table by the local router, the local router prefers it to any routes that it receives from other BGP routers. Step 3

The router prefers the route with the higher weight. Step 5

If the AS-path lengths are the same, the origin code is checked. BGP prefers the path with the lowest origin type: IGP is lower than EGP, and EGP is lower than Incomplete. Step 4

The lengths of the AS paths are compared (the content is not checked; only the number of autonomous systems in each AS path is counted). The route with the shortest length is selected. Step 6

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

- A. automatically, after a BGP neighbor session is established
- B. manually, with the **network** command
- C. through redistribution into the BGP process
- D. by advertising them to the BGP table on the router after Cisco Discovery Protocol discovers connected networks

33. What are two situations when is it appropriate to disable automatic summarization in BGP? (Choose two.) (Source: "Processing BGP Routes")
- A. when BGP neighbors are not configured to advertise aggregate routes to upstream providers
 - B. when the classless variant of the **network** command is used
 - C. when you are using a classless IGP in the AS
 - D. when the effects of automatic summarization of IGP-to-BGP redistribution are not desired
34. What is the AD of BGP routes in the IP routing table that were learned from BGP neighbors in a different AS? (Source: "Processing BGP Routes")
- A. 1
 - B. 20
 - C. 90
 - D. 120
35. Which three BGP attributes are displayed for each route in the BGP table when you are using the **show ip bgp** command? (Choose three.) (Source: "Processing BGP Routes")
- A. weight
 - B. communities
 - C. origin
 - D. AS-path
36. What is the valid AS number range for a BGP process on a Cisco router? (Source: "Configuring Basic BGP")
- A. 1 to 256
 - B. 1 to 32768
 - C. 1 to 65535
 - D. 1 to 131072
37. Which two parameters must you configure with the **neighbor** command to establish a BGP session with an external neighbor? (Choose two.) (Source: "Configuring Basic BGP")
- A. neighbor IP address
 - B. subnet mask of the IP network
 - C. remote AS number
 - D. local AS number
 - E. description of the neighbor
38. What is the best method to temporarily disable a BGP neighbor session? (Source: "Configuring Basic BGP")
- A. Remove the **neighbor** command from the BGP router process.
 - B. Remove the BGP router process from the configuration.
 - C. Terminate the neighbor connection with the **neighbor shutdown** command
 - D. Disconnect the neighbor by initiating a router reload.

39. Which two of the following statements about configuring BGP timers are accurate? (Choose two.)
(Source: "Configuring Basic BGP")
- A. Changing the BGP default holdtime and keepalive timers is usually not recommended.
 - B. The **neighbor timers** command sets the timers for a specific BGP peer or peer group.
 - C. The **timers bgp** command sets the timers for a specific BGP peer or peer group.
 - D. Holdtime indicates the frequency (in seconds) with which the Cisco IOS software sends messages to its peer.
40. Which two of the following are characteristics of the string component of the **neighbor {ip-address | peer-group-name} password string** command? (Choose two.) (Source: "Configuring Basic BGP")
- A. can contain any alphanumeric characters, including spaces
 - B. case-sensitive password of up to 100 characters
 - C. first character can be a number
 - D. cannot specify a password in the format "number-space-anything"
41. Which three steps must you complete to advertise a classless prefix into BGP? (Choose three.) (Source: "Configuring Basic BGP")
- A. Configure the prefix with the **network** command.
 - B. Specify the **mask** keyword with the locally advertised route.
 - C. Configure the **redistribute connected** command under the BGP router process.
 - D. Use a static route pointing to null 0 that matches the prefix.
42. Which origin code is carried with routes that are redistributed into BGP? (Source: "Configuring Basic BGP")
- A. internal
 - B. external
 - C. unknown
 - D. incomplete
43. Which two of the following statements about the classless behavior of BGP are correct? (Choose two.)
(Source: "Configuring Basic BGP")
- A. When an exact match is not found in the IP routing table a matching prefix is automatically configured on the router.
 - B. In the **network ip-prefix-address mask subnet-mask** command, the prefix does not have to match an entry in the IP routing table.
 - C. To advertise classless networks into BGP (a subnet or a supernet), you can use the **network** command with the **mask** keyword and the subnet mask specified.
 - D. If the keyword **mask** and the subnet mask are omitted, the network is assumed to have its natural mask according to the network class.

44. What are two benefits of using route aggregation in BGP? (Choose two.) (Source: "Configuring Basic BGP")
- A. It ensures that even if aggregate networks are down, the aggregate is advertised, which eliminates black holes.
 - B. It reduces the amount of memory that is used in the router to store the BGP table.
 - C. It reduces route flapping and its effects on router CPU resources.
 - D. BGP attribute granularity is maintained, which ensures optimal path selection.
45. Which two of the following are characteristics of the BGP Conditional Route Injection feature? (Choose two.) (Source: "Configuring Basic BGP")
- A. allows you to originate a prefix into a BGP routing table without the corresponding match
 - B. enabled with the **bgp inject-map exist-map** command
 - C. allows conditional injecting or replacing more specific prefixes with less specific prefixes
 - D. allows origination of a prefix into a BGP routing table only with the corresponding match
46. Which two of the following are characteristics of the BGP Support for TTL Security Check feature? (Choose two.) (Source: "Configuring Basic BGP")
- A. should be configured on only one participating router
 - B. prevents BGP sessions from expiring even if keepalive packets are not received before the session timer expires
 - C. protects the EBGP peering session by comparing the value in the TTL field of received IP packets against a hop count that is configured locally for each EBGP peering session
 - D. supports both directly connected peering sessions and multihop EBGP peering sessions
47. Which two of the following are functions of the **show ip bgp summary** command? (Choose two.)
- A. displays BGP memory use
 - B. displays BGP neighbors and status of communication with them
 - C. locates problems in BGP sessions that are up and established
 - D. displays the BGP routing table
48. Which command do you use to display detailed BGP neighbor information? (Source: "Monitoring and Troubleshooting BGP")
- A. **show ip bgp summary**
 - B. **show ip bgp**
 - C. **show ip bgp neighbors address**
 - D. **show ip bgp detail**
49. To get information about the TCP session and BGP parameters of the session, you can issue **show ip bgp summary** command. True or false? (Source: "Monitoring and Troubleshooting BGP")
- A. true
 - B. false

50. Which two of the following statements about the **show ip bgp** command that is used to monitor the BGP routing table are accurate? (Choose two.) (Source: "Monitoring and Troubleshooting BGP")
- A. The **show ip bgp** command shows an abbreviated list of information about each route, displaying one line per prefix.
 - B. The **show ip bgp** command shows a full list of information about each route, displaying one line per prefix.
 - C. All of the BGP attributes that are associated with the route are displayed on the line.
 - D. If the BGP table contains more than one route to the same network, the routes are displayed on successive lines of the command output.
51. Which debug command should you enable to troubleshoot BGP session startup issues where the TCP connection never succeeds? (Source: "Monitoring and Troubleshooting BGP")
- A. **ip bgp updates**
 - B. **ip packets**
 - C. **ip bgp keepalives**
 - D. **ip tcp transactions**
52. When debugging an exchange of updates, you should use **debug ip bgp keepalive** command. True or False?
- A. true
 - B. false
53. What are the three most common session startup issues that you can experience when configuring basic BGP? (Choose three.) (Source: "Monitoring and Troubleshooting BGP")
- A. BGP neighbors do not become active.
 - B. BGP routing loops cause black holes.
 - C. A BGP neighbor is active, but the BGP session is not established.
 - D. The BGP neighbor state oscillates between Idle and Active.
 - E. The BGP session is active, but the neighbor cannot be reached.
 - F. BGP keepalives experience intermittent failures.
54. What is the most common reason for a BGP session not leaving the Idle state? (Source: "Monitoring and Troubleshooting BGP")
- A. The TCP port for the connection is not configured.
 - B. The external neighbor is not directly connected.
 - C. The TCP SYN packet is answered with an RST packet.
 - D. The neighbors have been configured with the same AS number.
55. What will result from attempting to open a BGP connection with a neighbor that has not been properly configured for BGP? (Source: "Monitoring and Troubleshooting BGP")
- A. The BGP session will remain in the Idle state.
 - B. The neighbor session will be established, and the session startup parameters will be negotiated over the TCP session.
 - C. The BGP session will be immediately terminated with a TCP RST packet.
 - D. The BGP session will become "stuck in Active state."

56. When a BGP neighbor oscillates between Active and Idle, what is the likely diagnosis? (Source: "Monitoring and Troubleshooting BGP")
- A. There are mismatched keepalive intervals.
 - B. There is an AS number mismatch between BGP neighbors.
 - C. One router is not configured as the BGP neighbor on the neighboring router.
 - D. The BGP neighbor is not reachable.

Module Self-Check Answers

Answer Key

1. A, B
2. A
3. A
4. A, B, D
5. A
6. A, B, C
7. A, B, D
8. A, C, D
9. B
10. A, B, D
11. A, C, D
12. A, B, C
13. C
14. A, B
15. A, B, C
16. B
17. C, E
18. A
19. B
20. C
21. C
22. D
23. Idle — 1

Active — 2
OpenSent — 3
OpenConfirm — 4
Established — 5

24. A

25. D

26. C

27. D

28. B, D, E

29. B

30. D

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

The router checks whether the next-hop attribute indicates an IP address that is reachable according to the current routing table. If the next hop is not reachable, the router does not consider the BGP route as a candidate to become selected as the best. — Step 1

The router prefers the route with the higher weight. — Step 2

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

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

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

If the AS-path lengths are the same, the origin code is checked. BGP prefers the path with the lowest origin type: IGP is lower than EGP, and EGP is lower than Incomplete. — Step 6

32. B, C

33. B, D

34. B

35. A, C, D

36. C

37. A, C

38. C

39. A, B

40. A, D

41. A, B, D

42. D

43. C, D

44. B, C

45. A, B

- 46. C, D
- 47. A, B
- 48. C
- 49. B
- 50. B, D
- 51. D
- 52. B
- 53. A, C, D
- 54. B
- 55. C
- 56. B

Module 2: BGP Transit Autonomous Systems

Introduction

This module is one of the focal points of the [BGP](#) curriculum: a discussion of BGP issues in a transit [AS](#). In this module you will learn basic BGP transit AS issues, ranging from synchronization between an [IGP](#), and BGP to [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 in a transit AS between EBGP and IBGP
- Describe the function of an IGP in forwarding packets through an AS
- Verify proper operation of a configured BGP transit network by performing the steps necessary to correct basic IBGP configuration errors

Lesson 1: Working with Transit AS

Overview

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

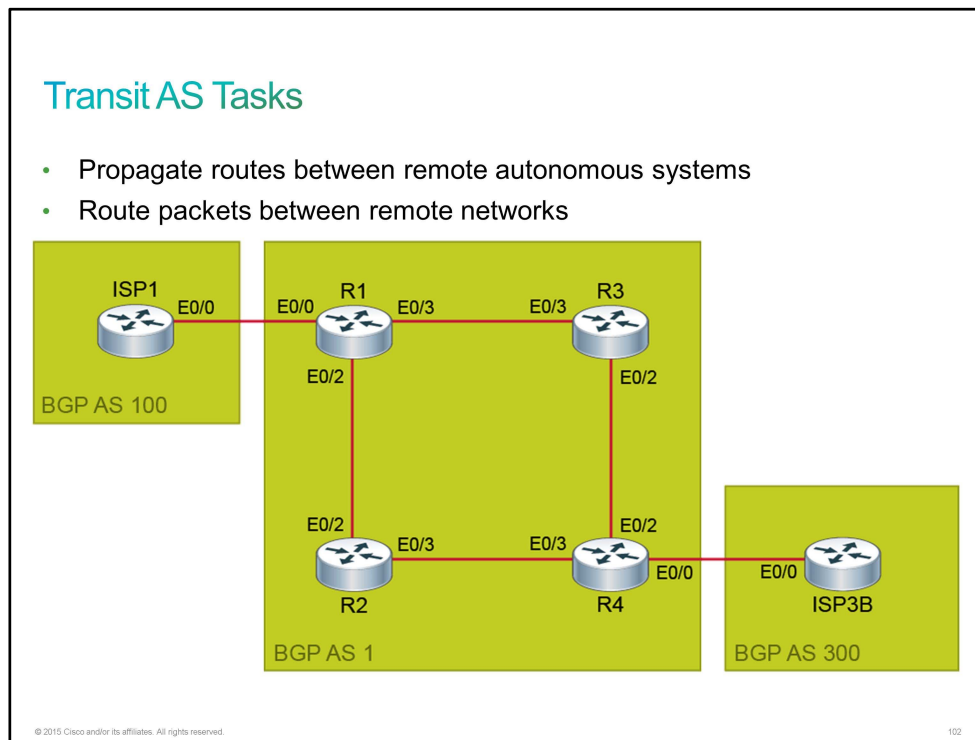
The topology of the Internet can be viewed as a series of connections between stub networks, multihomed networks, and transit autonomous systems. A multihomed AS containing more than one connection to the outside world and allowing traffic not originating in that AS to travel through it is a transit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

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

Transit AS Tasks

A transit [AS](#) is required to carry traffic that originates from or is destined for locations outside of that AS. A transit AS has two basic functions.

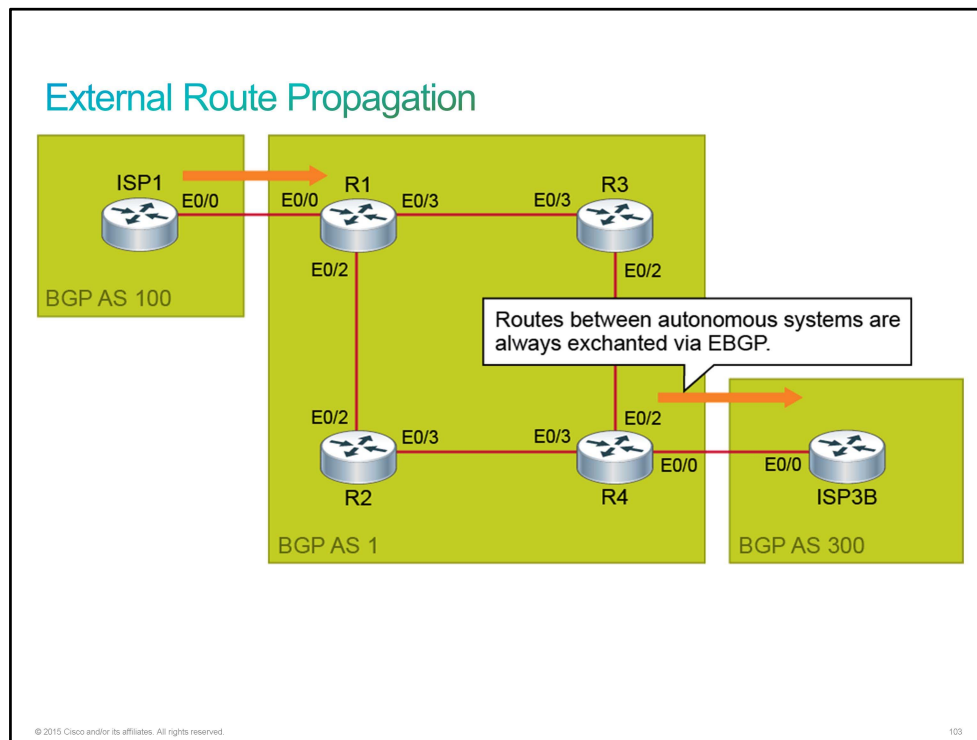


Routers in a transit AS have to perform two tasks:

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

External Route Propagation

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



In the figure, the BGP session between ISP1 and R1 is called an [EBGP](#) session because ISP1 and R1 are in different autonomous systems.

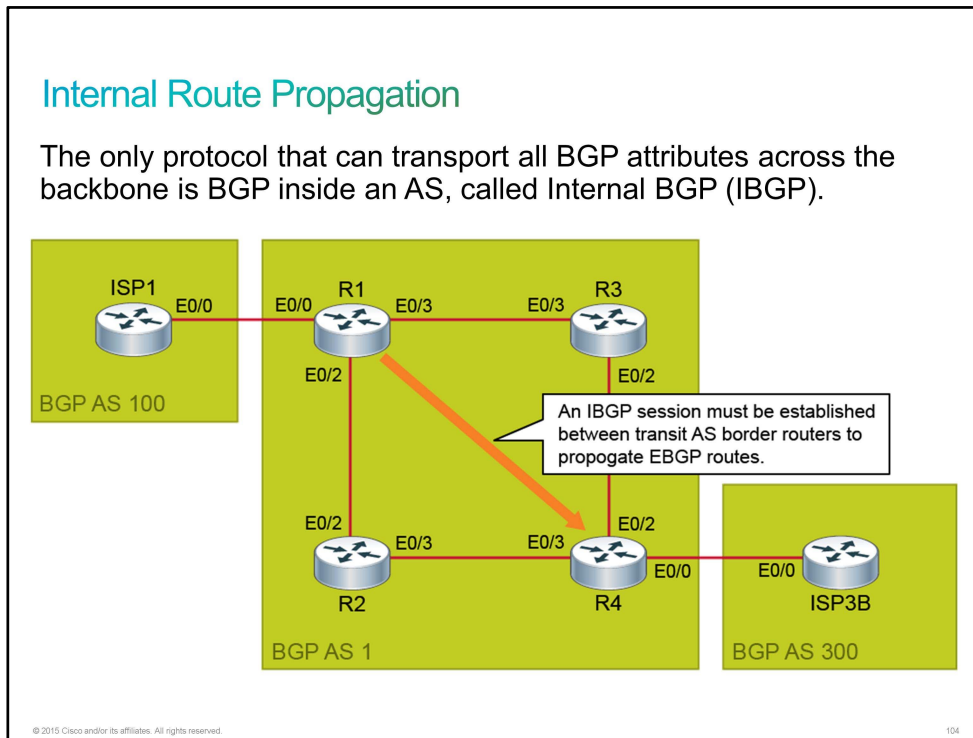
BGP routing information updates consist of the network address, subnet mask, and any number of BGP attributes. No other routing protocol provides the same richness of route attributes as BGP. Translating BGP route attribute information into any other protocol would likely cause a loss of information. Therefore, the EBGP information that R1 receives is not translated; it is just forwarded to other BGP-speaking routers (R2, R3, and R4 in the figure) within the [AS](#).

Likewise, R4 has BGP information and can propagate it to ISP3B in AS 300 over the EBGP session.

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

Internal Route Propagation

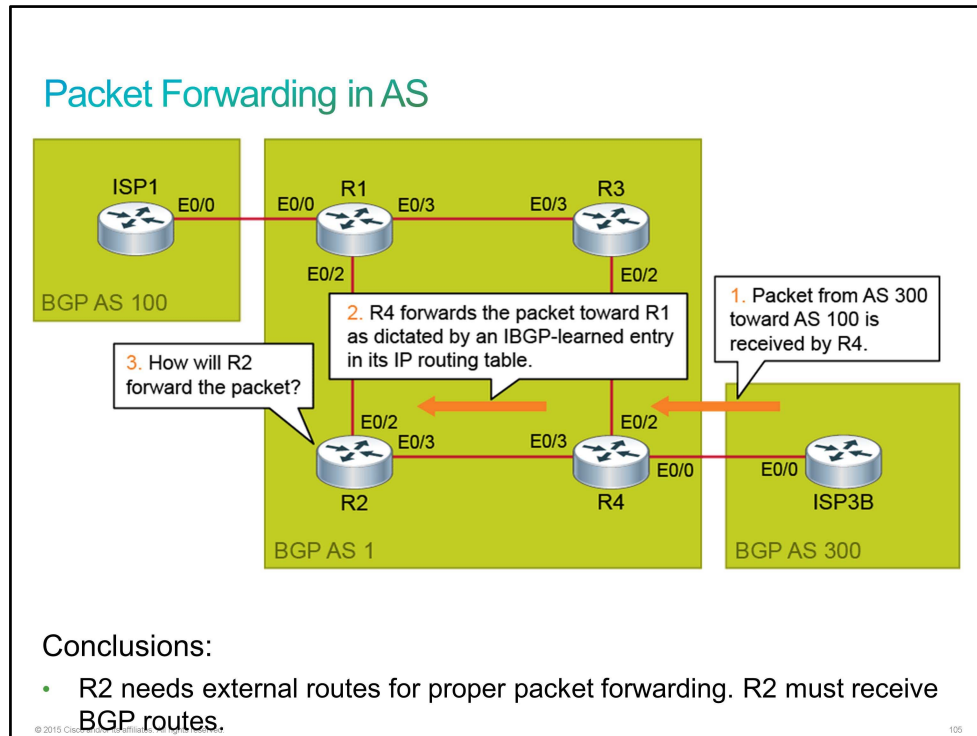
In this example, the [BGP](#) session between R1 and R4, which are both in the same [AS](#), is an [IBGP](#) session.



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

Packet Forwarding in AS

In this example, after [AS 300](#) has received the routing information about reachable networks inside [AS 100](#), [IP](#) packets can start to flow. In the figure, IP packets flow from [AS 300](#) toward [AS 100](#). ISP3B, the egress router in [AS 300](#), forwards IP packets with destinations in [AS 100](#) toward R4, according to information received through [EBGP](#).



R4 now uses the [IBGP](#) information that it received from R1 and forwards the packets in the direction of R1, which in this case means via R2.

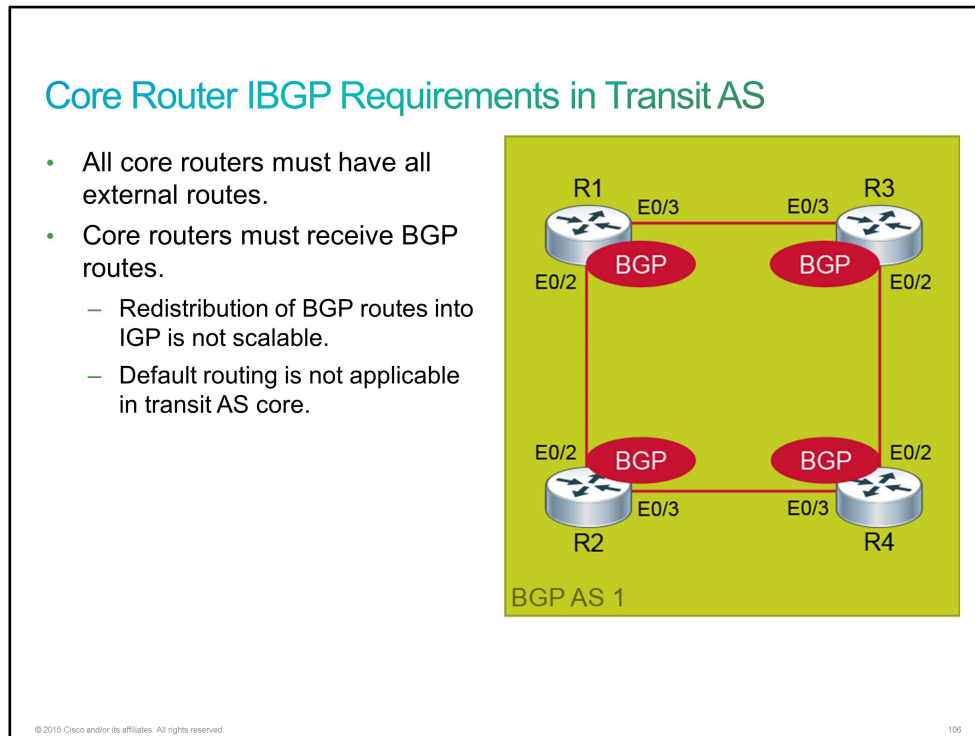
When the IP packets reach R2, the router checks its routing table for a matching entry, but it fails to find one. The packet is dropped because the destination network is unreachable from the perspective of R2.

This situation is, of course, unacceptable. To prevent dropped packets resulting from unreachable networks, R2 must also have routing information about the networks reachable inside [AS 100](#). The same information that R4 received from R1 over the IBGP session must be propagated to R2.

Note R3 has the same network reachability requirements as R2, because R4 could forward the packets via R3 as well as via R2.

Core Router IBGP Requirements in Transit AS

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



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

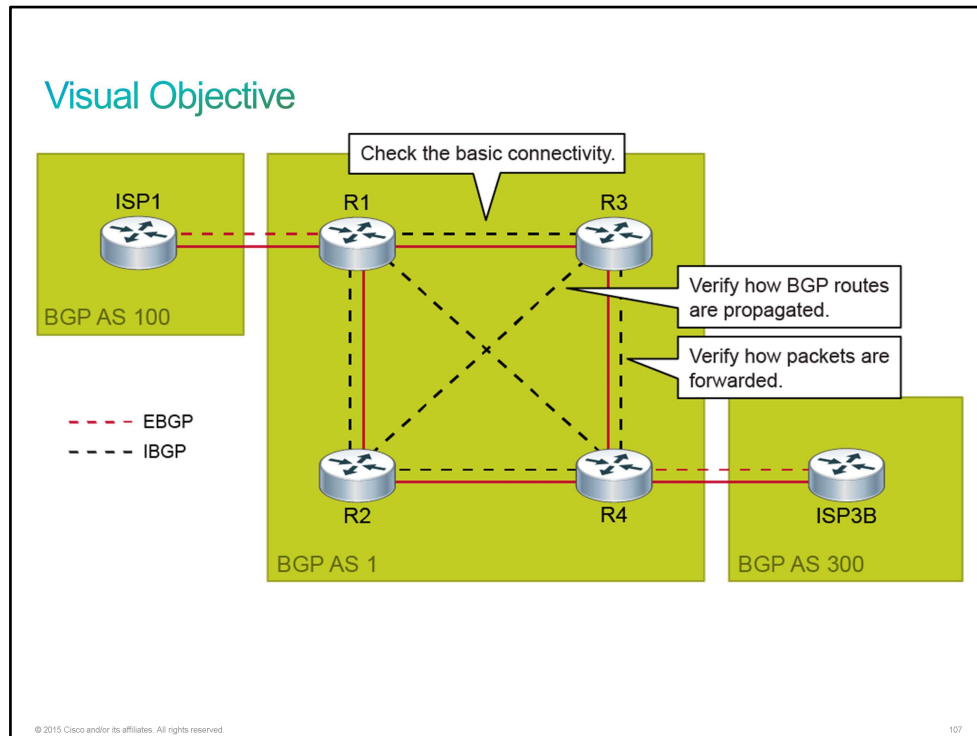
Note The risk of losing information during redistribution of EBGP routes into an IGP is not the reason why BGP is used to update intermediate routers in the transit path instead of an IGP. Redistribution into an IGP is not used because of the scalability issues that would arise from doing so.

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

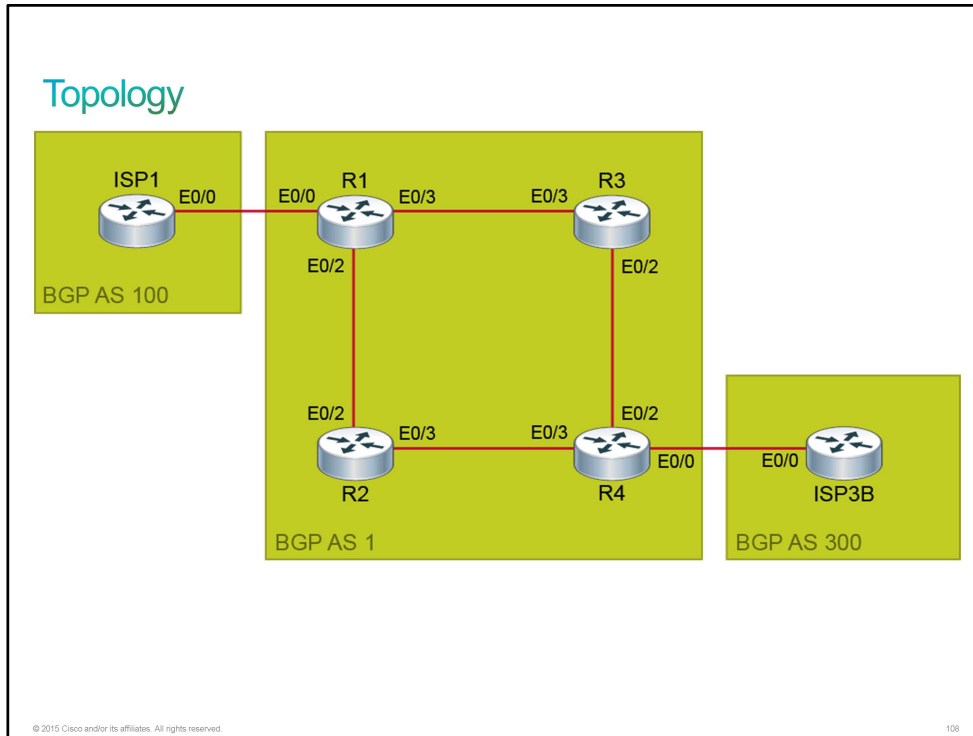
Discovery 4: BGP Route Propagation

Overview

Through this discovery, you will learn how the [BGP](#) routes are propagated between different autonomous systems that are connected directly and need for full mesh within an [AS](#).



Topology



Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

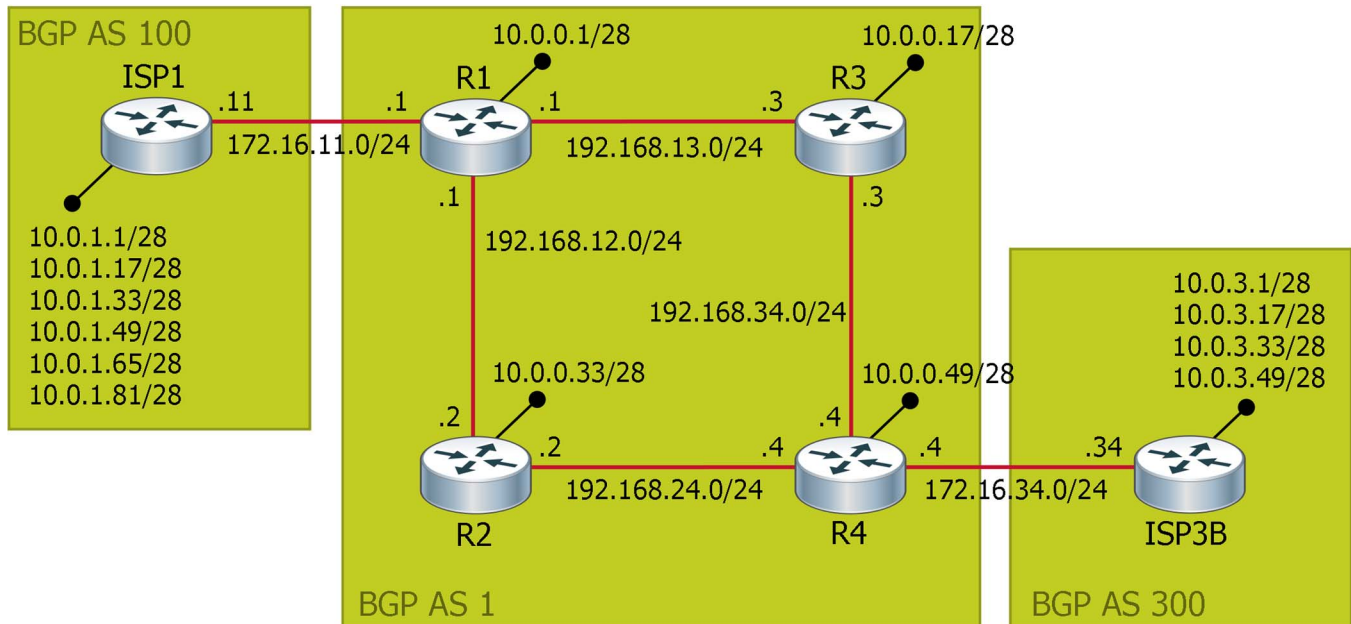
Device	Interface	IP address	Description
ISP1	Ethernet 0/0	172.16.11.11/24	Connection to R1
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
R1	Ethernet 0/0	172.16.11.1/24	Connection to ISP1
R1	Ethernet 0/2	192.168.12.1/24	Connection to R2
R1	Ethernet 0/3	192.168.13.1/24	Connection to R3
R1	Loopback 1	10.0.0.1/28	Loopback simulates LAN network
R2	Ethernet 0/2	192.168.12.2/24	Connection to R1
R2	Ethernet 0/3	192.168.24.2/24	Connection to R4
R2	Loopback 1	10.0.0.33/28	Loopback simulates LAN network
R3	Ethernet 0/2	192.168.34.3/24	Connection to R4
R3	Ethernet 0/3	192.168.13.3/24	Connection to R1
R3	Loopback 1	10.0.0.17/28	Loopback simulates LAN network
R4	Ethernet 0/0	172.16.34.4/24	Connection to ISP3B
R4	Ethernet 0/2	192.168.34.4/24	Connection to R3
R4	Ethernet 0/3	192.168.24.4/24	Connection to R2
R4	Loopback 1	10.0.0.49/28	Loopback simulates LAN network
ISP3B	Ethernet 0/0	172.16.34.34/24	Connection to R4
ISP3B	Loopback 1 Loopback 2 Loopback 3 Loopback 4	10.0.3.1/28 10.0.3.17/28 10.0.3.33/28 10.0.3.49/28	Loopbacks simulate LAN networks

BGP Route Propagation

Discovery Steps

Step 1 You will verify route 10.0.1.0/28 is propagated from AS 100 into AS 300.

[IP addresses](#), [IGP](#) and BGP are preconfigured as shown in the topology below:



On the ISP1 router, verify route 10.0.1.0/28 is present in the IP routing table.

```
ISP1# show ip route
< output omitted >
Gateway of last resort is not set

10.0.0.0/8 is variably subnetted, 20 subnets, 2 masks
B    10.0.0.0/28 [20/0] via 172.16.11.1, 01:49:30
B    10.0.0.16/28 [20/0] via 172.16.11.1, 01:49:30
B    10.0.0.32/28 [20/0] via 172.16.11.1, 01:49:30
B    10.0.0.48/28 [20/0] via 172.16.11.1, 01:49:30
C    10.0.1.0/28 is directly connected, Loopback1
L    10.0.1.1/32 is directly connected, Loopback1
C    10.0.1.16/28 is directly connected, Loopback2
L    10.0.1.17/32 is directly connected, Loopback2
C    10.0.1.32/28 is directly connected, Loopback3
<... output omitted ...>
```

IP address 10.0.1.1/28 is configured on the Loopback 1 interface.

Step 2 On the ISP1 router, verify route 10.0.1.0/28 is present in the BGP routing table as locally originated.

```
ISP1# show ip bgp
< output omitted >
      Network                Next Hop                Metric LocPrf Weight Path
*> 10.0.0.0/28                172.16.11.1                0
*> 10.0.0.16/28               172.16.11.1
*> 10.0.0.32/28               172.16.11.1
*> 10.0.0.48/28               172.16.11.1
*> 10.0.1.0/28                0.0.0.0                    0 32768 i
*> 10.0.1.16/28               0.0.0.0                    0 32768 i
*> 10.0.1.32/28               0.0.0.0                    0 32768 i
*> 10.0.1.48/28               0.0.0.0                    0 32768 i
*> 10.0.1.64/28               0.0.0.0                    0 32768 i
*> 10.0.1.80/28               0.0.0.0                    0 32768 i
*> 10.0.3.0/28                172.16.11.1                0 1 300 i
*> 10.0.3.16/28               172.16.11.1                0 1 300 i
*> 10.0.3.32/28               172.16.11.1                0 1 300 i
*> 10.0.3.48/28               172.16.11.1                0 1 300 i
```

The route 10.0.1.0/28 is locally originated, because next hop IP address is 0.0.0.0 and AS path attribute is empty.

Step 3 On the R1 router, verify route 10.0.1.0/28 is present in the BGP routing table.

```
R1# show ip bgp
< output omitted >
      Network                Next Hop                Metric LocPrf Weight Path
*> 10.0.0.0/28                0.0.0.0                    0 32768 i
*>i 10.0.0.16/28               10.0.0.17                  0 100 0 i
*>i 10.0.0.32/28               10.0.0.33                  0 100 0 i
*>i 10.0.0.48/28               10.0.0.49                  0 100 0 i
*> 10.0.1.0/28                172.16.11.11               0 0 100 i
*> 10.0.1.16/28               172.16.11.11               0 0 100 i
*> 10.0.1.32/28               172.16.11.11               0 0 100 i
*> 10.0.1.48/28               172.16.11.11               0 0 100 i
*> 10.0.1.64/28               172.16.11.11               0 0 100 i
*> 10.0.1.80/28               172.16.11.11               0 0 100 i
*>i 10.0.3.0/28                172.16.34.34               0 100 0 300 i
*>i 10.0.3.16/28               172.16.34.34               0 100 0 300 i
*>i 10.0.3.32/28               172.16.34.34               0 100 0 300 i
*>i 10.0.3.48/28               172.16.34.34               0 100 0 300 i
```

The route 10.0.1.0/28 has next hop IP address 172.16.11.11 and AS path attribute 100.

Step 4 On the R4 router, verify route 10.0.1.0/28 is present in the BGP routing table.

```
R4# show ip bgp
< text omitted >
      Network          Next Hop          Metric LocPrf Weight Path
*>i 10.0.0.0/28         10.0.0.1              0      100      0 i
*>i 10.0.0.16/28        10.0.0.17             0      100      0 i
*>i 10.0.0.32/28        10.0.0.33             0      100      0 i
*> 10.0.0.48/28         0.0.0.0               0              32768 i
*>i 10.0.1.0/28         172.16.11.11          0      100      0 100 i
*>i 10.0.1.16/28        172.16.11.11          0      100      0 100 i
*>i 10.0.1.32/28        172.16.11.11          0      100      0 100 i
*>i 10.0.1.48/28        172.16.11.11          0      100      0 100 i
*>i 10.0.1.64/28        172.16.11.11          0      100      0 100 i
*>i 10.0.1.80/28        172.16.11.11          0      100      0 100 i
*> 10.0.3.0/28          172.16.34.34          0              0 300 i
*> 10.0.3.16/28         172.16.34.34          0              0 300 i
*> 10.0.3.32/28         172.16.34.34          0              0 300 i
*> 10.0.3.48/28         172.16.34.34          0              0 300 i
```

The route 10.0.1.0/28 has next hop IP address 172.16.11.11 and AS path attribute 100. You can see route is learned via internal BGP.

Step 5 On the R4 router, verify what is the best path to reach route next hop IP address 172.16.11.11.

```
R4# show ip route
< output omitted >
      172.16.0.0/16 is variably subnetted, 3 subnets, 2 masks
O      172.16.11.0/24 [110/30] via 192.168.34.3, 02:13:15, Ethernet0/2
                                [110/30] via 192.168.24.2, 02:13:15, Ethernet0/3
<... output omitted ...>
```

Output shows that next hop IP address 172.16.11.11 is reachable via two [OSPF](#) routes. Router R4 will load share between two OSPF routes. Some packets will be sent to the R2 router and some to the R3 router.

So, R2 and R3 routers need IP routing information to reach next hop 172.16.11.11.

Step 6 On the ISP3B router, verify route 10.0.1.0/28 is present in the BGP routing table.

```
ISP3B# show ip bgp
< output omitted >
      Network          Next Hop          Metric LocPrf Weight Path
*> 10.0.0.0/28          172.16.34.4              0 1 i
*> 10.0.0.16/28          172.16.34.4              0 1 i
*> 10.0.0.32/28          172.16.34.4              0 1 i
*> 10.0.0.48/28          172.16.34.4              0 1 i
*> 10.0.1.0/28          172.16.34.4              0 1 100 i
*> 10.0.1.16/28          172.16.34.4              0 1 100 i
*> 10.0.1.32/28          172.16.34.4              0 1 100 i
*> 10.0.1.48/28          172.16.34.4              0 1 100 i
*> 10.0.1.64/28          172.16.34.4              0 1 100 i
*> 10.0.1.80/28          172.16.34.4              0 1 100 i
*> 10.0.3.0/28           0.0.0.0                  0 32768 i
*> 10.0.3.16/28          0.0.0.0                  0 32768 i
*> 10.0.3.32/28          0.0.0.0                  0 32768 i
*> 10.0.3.48/28          0.0.0.0                  0 32768 i
```

The route 10.0.1.0/28 has next hop IP address 172.16.34.4 and AS path attribute 1 100.

Step 7 You will verify, how packets are forwarded from AS 300 to the IP 10.0.1.1 in the AS 100. On the ISP3B router, ping next hop IP address of the BGP route 10.0.1.0/24.

In the previous step, you have learned next hop IP address is 172.16.34.4. Source ping from the ISP3B router's Loopback 1 interface.

```
ISP3B# ping 172.16.34.4 source Loopback 1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.34.4, timeout is 2 seconds:
Packet sent with a source address of 10.0.3.1
!!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

From the ISP3B router, you can successfully reach next hop 172.16.34.4 (router R4).

Step 8 From the R4 router, to reach IP address 10.0.1.1, next hop IP address is 172.16.11.11. From the R4 router, ping next hop IP address from the BGP route 10.0.1.0/24.

Source ping from the R4 router's Loopback 1 interface.

```
R4# ping 172.16.11.11 source Loopback 1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.11.11, timeout is 2 seconds:
Packet sent with a source address of 10.0.0.49
!!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

From the R4 router, you can successfully reach next hop 172.16.11.11 (ISP1 router).

Step 9 From the ISP3B router, ping directly to the IP address 10.0.1.1. Source ping from the ISP3B router's Loopback 1 interface.

Ping should be successful. The successful ping demonstrates IP connectivity between AS 100 and AS 300 via AS 1.

```
ISP3B# ping 10.0.1.1 source Loopback 1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.0.1.1, timeout is 2 seconds:
Packet sent with a source address of 10.0.3.1
!!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Routers in a transit AS receive routing information updates from neighboring autonomous systems, propagate the information through their own AS, and send it to other neighboring autonomous systems.
- Two autonomous systems usually exchange routing information over an EBGp session.
- A BGP session between two routers in the same AS is called an IBGP session.
- 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.

Lesson 2: Interacting with IBGP and EBGP in Transit AS

Overview

Configuring a [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.

Upon completing this lesson, you will be able to:

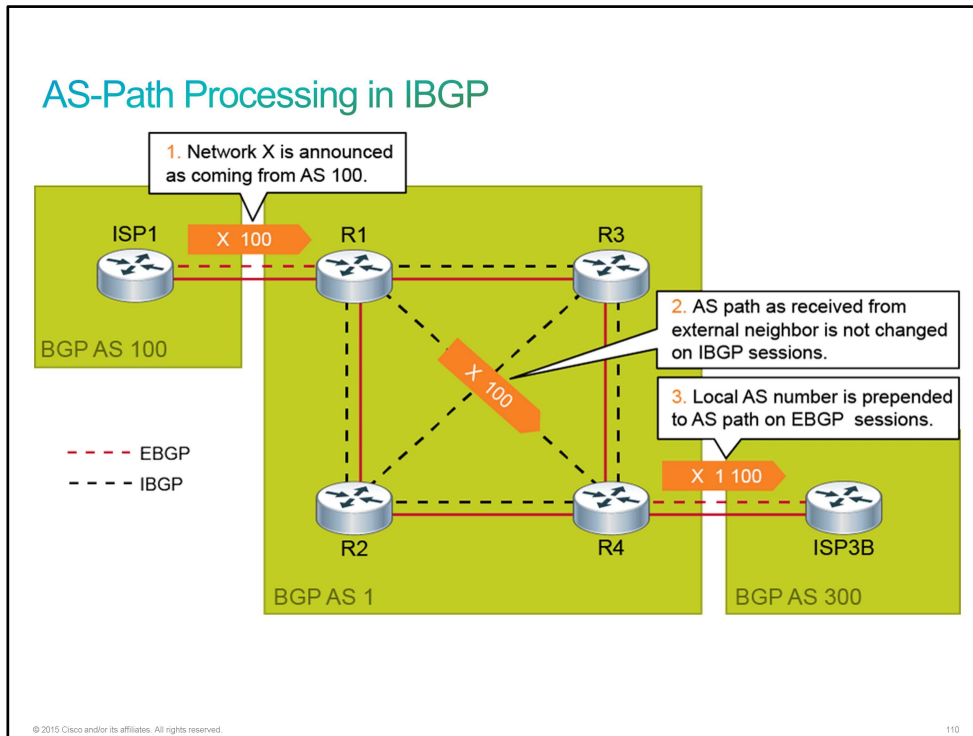
- Describe AS-path processing in IBGP
- Describe BGP multipath load sharing
- Explain the need for BGP split horizon
- Explain the need for a full-mesh topology between IBGP routers and the implications of that need
- List the benefits of establishing IBGP neighbor sessions using loopback interfaces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

AS-Path Processing in IBGP

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

When a router originates a BGP prefix (network X in this example), the AS path is empty. Whenever a BGP prefix is announced over an [EBGP](#) session, the AS number of the router that is sending the information is prepended to the AS path. In the example, ISP1 inserts "100" in the AS path before forwarding the routing update to R1.

The AS path is not changed when the BGP prefix is propagated across [IBGP](#) sessions because the routing update has not crossed an AS boundary.

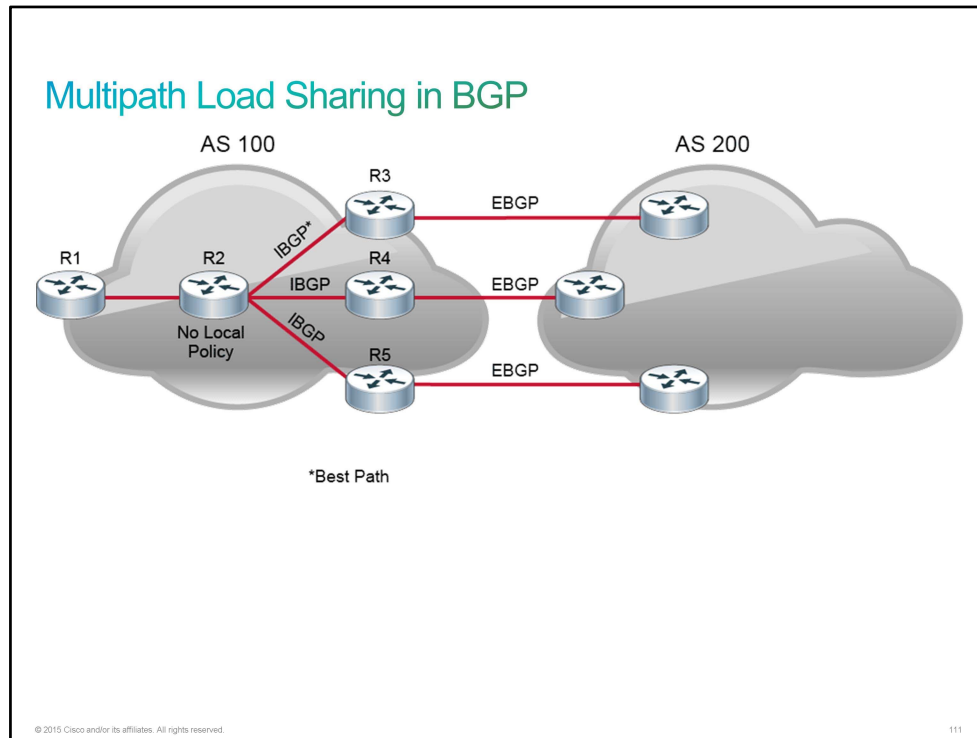


In the figure, R1 forwards the information over an IBGP session to R4 with the AS-path unchanged. The AS-path information about network X will be the same in all routers within AS 1, because all the routers are updated using IBGP sessions from R1.

When R4 forwards the information about network X to ISP3B, R4 prepends its own AS number (1) to the AS path. Thus, ISP3B receives the routing information about network X with an AS-path attribute of "1 100."

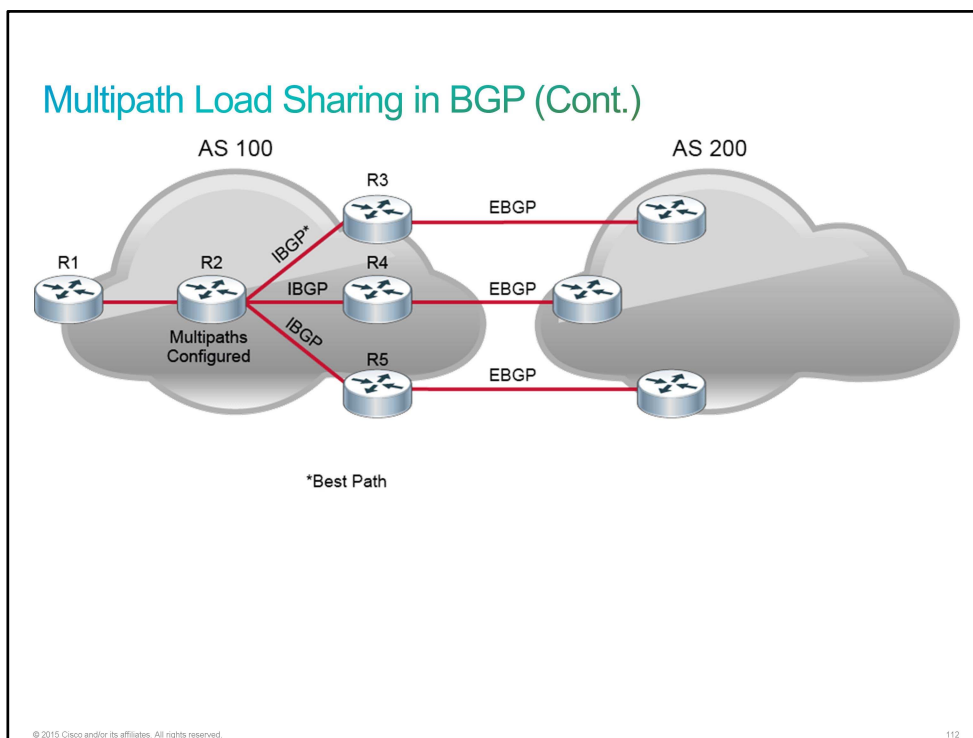
Multipath Load Sharing in BGP

When a BGP-speaking router with no local policy configured receives NLRI from multiple IBGP sources for the same destination, the router chooses one IBGP path as the best path. The best path is then installed in the IP routing table of the router.



The figure illustrates that with three paths to AS 200, the R2 router determines that one of the paths to AS 200 is the best path. So, it uses only this path to reach AS 200.

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



On the R2 router in the figure, the paths to the R3, R4, and R5 routers are configured as multipaths and can be used to reach AS 200. The load to AS 200 is equally shared between the paths.

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

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

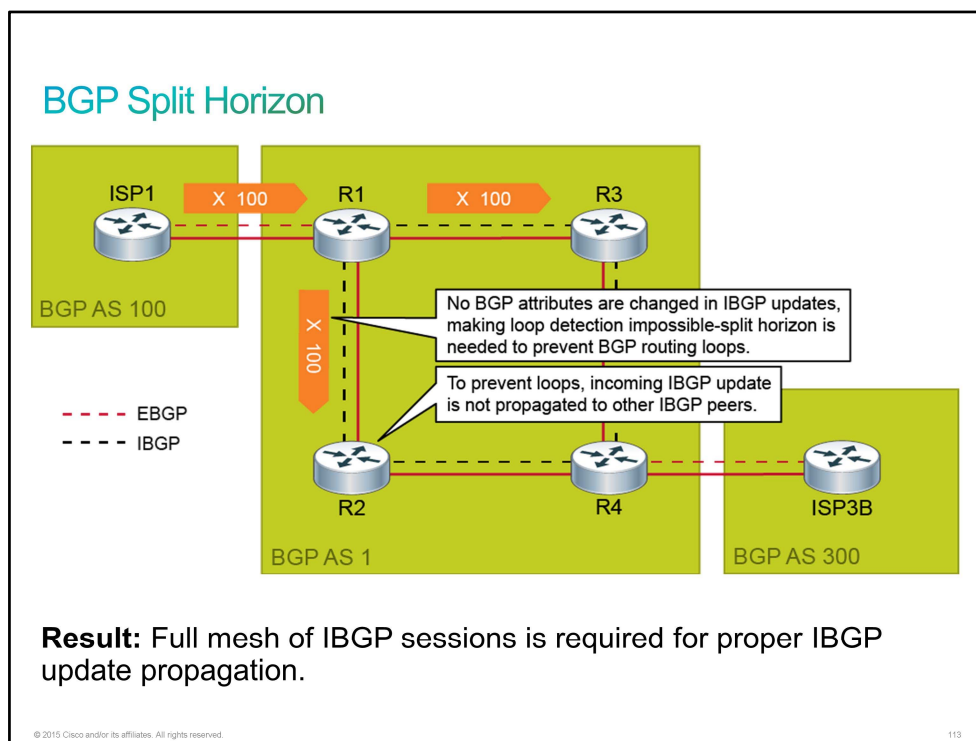
Even if the criteria are met and multiple paths are considered multipaths, the BGP-speaking router still designates one of the multipaths as the best path and advertises this best path to its neighbors.

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

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

BGP Split Horizon

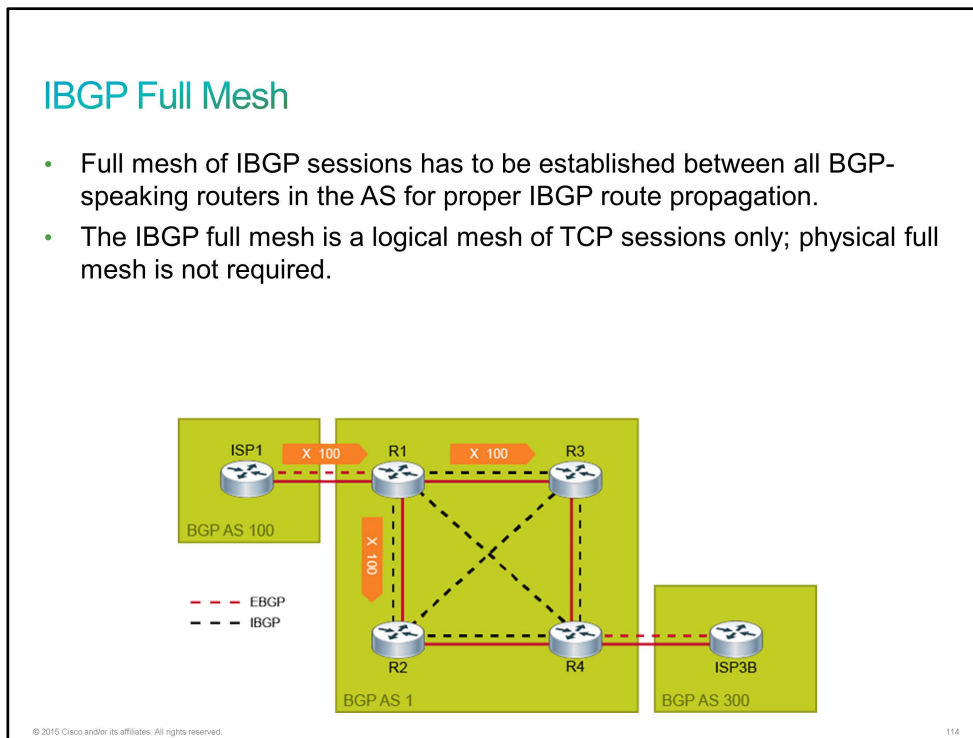
All routers within an [AS](#) must make routing decisions in a consistent way. They must have access to the same routing information with the same attributes in order to come to the same conclusion about which exit point of the AS to use. In other words, the [BGP](#) attributes should not be changed 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 BGP-speaking routers modify AS path only on [EBGP](#) sessions, this loop-preventing mechanism is only useful between autonomous systems, not within them.

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

IBGP Full Mesh



Every router on the transit path within the [AS](#) must have routing information about all external networks that are received by any of the border routers. So, R2 and R3 must have [IBGP](#) sessions to all border routers.

This level of communication is not enough, though, because any of the internal routers could also create new [BGP](#) routing information (for example, originate a customer network). These updates must also reach all the routers within the AS. The conclusion is that all BGP routers within an AS must have IBGP sessions with every other BGP router in the AS. This requirement results in a full mesh of BGP sessions between BGP-speaking routers in an AS.

In the network in the figure, R1 must have IBGP sessions with R2, R3, and R4 to propagate routes that are received from AS 100 to all routers within AS 1. Similarly, R4 must have IBGP sessions with R1, R2, and R3 to be able to propagate routes that are received from AS 300 to all routers within AS 1.

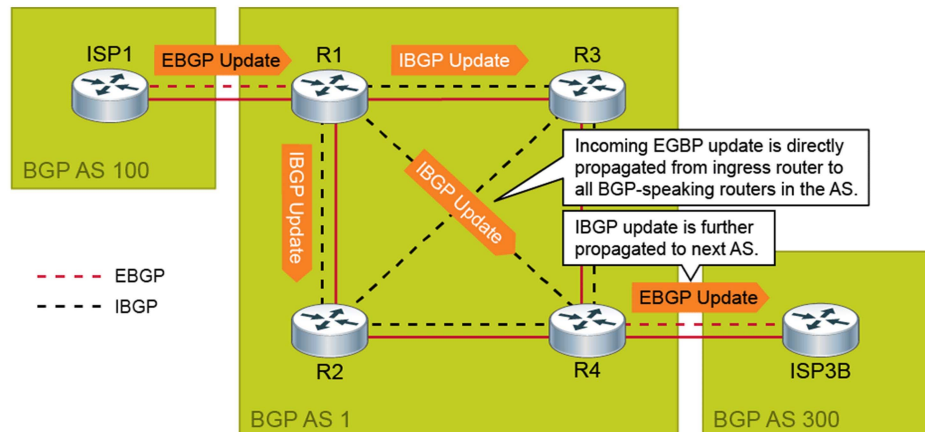
Note The IBGP session between R2 and R3 is not strictly necessary for proper forwarding of [IP](#) packets between external destinations. It does become mandatory if R2 or R3 starts to originate BGP networks. To prevent potential future connectivity issues, it is a good practice to establish a full mesh of IBGP sessions regardless of whether they are needed at the time of network deployment or not.

The [IGP](#) that runs within AS 1 provides enough information to any BGP router within AS 1 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.

IBGP Full Mesh Example

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

IBGP Full Mesh Example



© 2015 Cisco and/or its affiliates. All rights reserved.

115

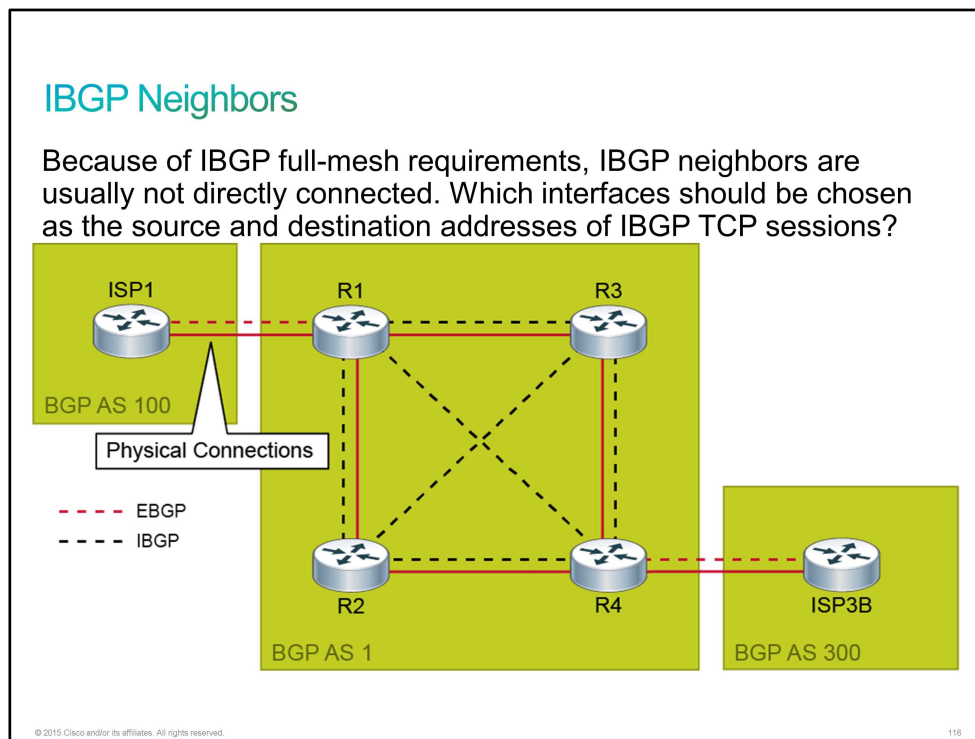
ISP1 is sending an update to R1 over an EBGP session. Updates that are received on an [EBGP](#) session should be forwarded on all other IBGP sessions, so R1 updates R2, R3, and R4. All routers within AS 1 are updated directly by R1.

R2 and R3 are prevented from forwarding the update that they received from R1 because of BGP split horizon.

R4, which received the information on an IBGP session, is prevented from updating R2 and R3 because of the same split-horizon rule. But R4 will update ISP3B over an EBGP session.

IBGP Neighbors

In the figure, the transit [AS 1](#) has a redundant physical topology. The [IGP](#) provides reachability information for all routers and networks within AS 1. This way, IGP enables all routers in the AS to establish [IBGP](#) sessions to all other routers, even if the routers are not directly connected.



If the IBGP session between R1 and R4 was established using [IP addresses](#) that belong to the physical interfaces, the IBGP session would go down if either of the physical interfaces went down. The IP address of an interface that is in the down state is invalid. As a result, the router would tear down the [TCP](#) session that is used for [BGP](#) between the routers. Then, all IP packets that are received with a destination address pointing to that interface will also be dropped.

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

IBGP Neighbors (Cont.)

Always run IBGP sessions between loopback interfaces.

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

© 2015 Cisco and/or its affiliates. All rights reserved.

117

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

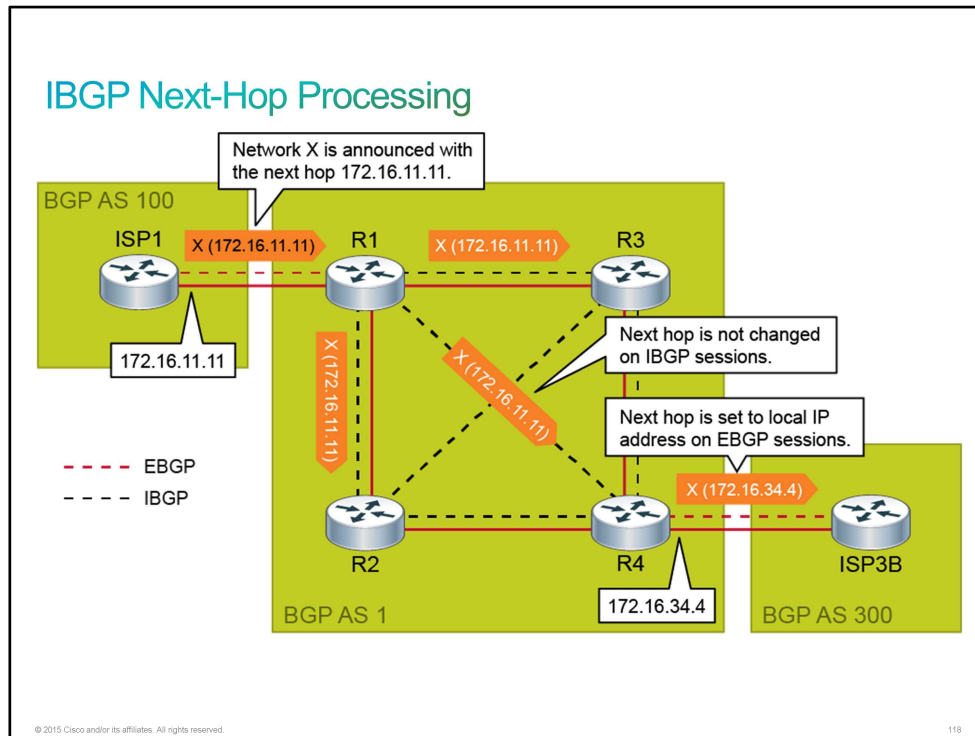
To establish BGP connectivity between the loopback interfaces, the IP addresses of these interfaces have to be reachable by both routers. It is important that the [IGP](#) carry information about the subnets that are assigned to each loopback interface so that the interfaces are reachable by all BGP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

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

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

IBGP Next-Hop Processing

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

In the figure, ISP1 sends a BGP update about network X. Because it is sending this update over an EBGP session to R1, the next-hop attribute is set to the IP address that is used at the ISP1 side of the EBGP session, 172.16.11.11.

R1 can use this information and route packets to network X by forwarding them to ISP1.

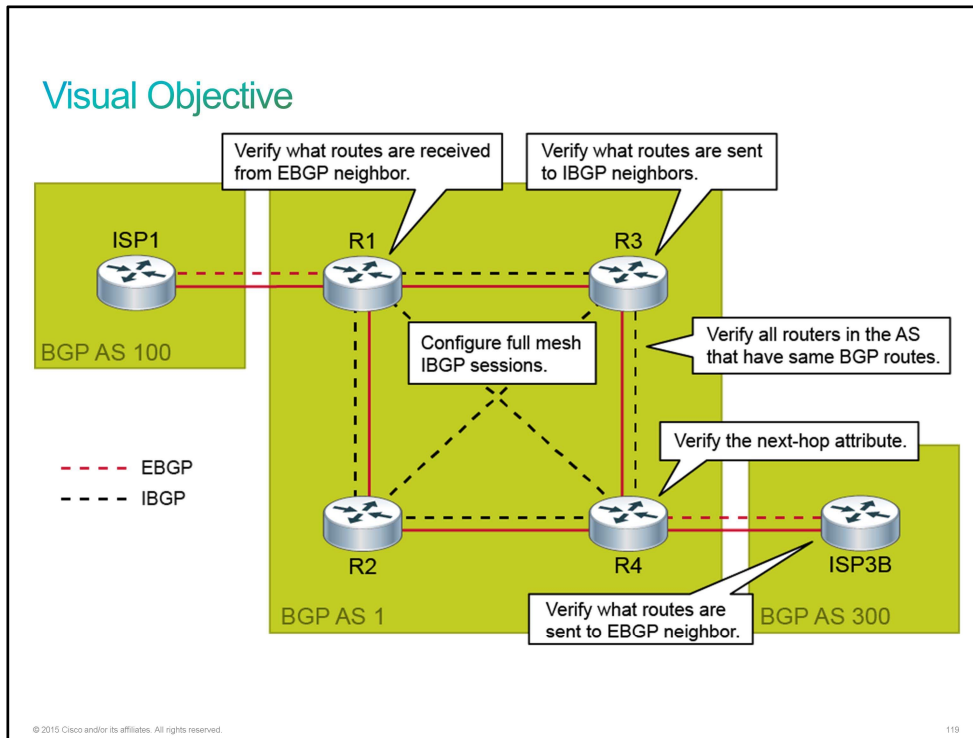
R1 also forwards the BGP update over all its IBGP sessions. It does not change the next-hop attribute, so R2, R3, and R4 get information that they can reach network X by forwarding packets to 172.16.11.11. But that IP address is not directly connected, so the routers must look in their routing tables to see if and how they can reach 172.16.11.11. If the recursive route lookup is successful, each router can then route packets to network X in the same direction as it would route packets to 172.16.11.11.

R4 also forwards the BGP update about network X to ISP3B. The connection between these routers is an EBGP session. So, R4 sets the next-hop attribute to its own IP address, 172.16.34.4, which is used by R4 on the EBGP session toward ISP3B.

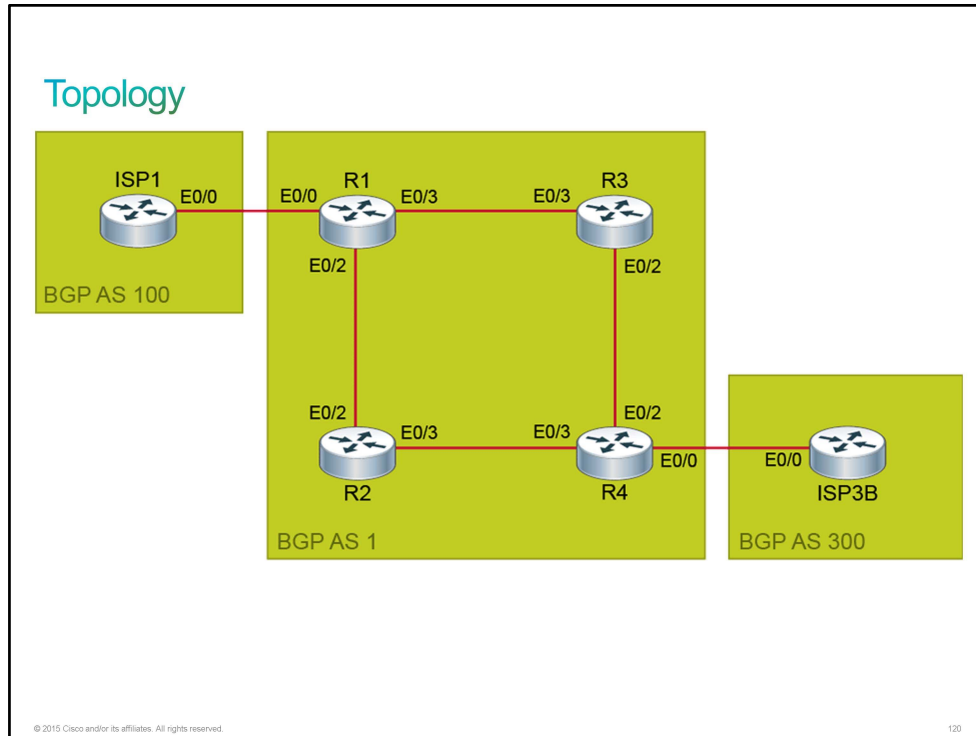
Discovery 5: IBGP Full Mesh

Overview

Through this discovery, you will learn how the [BGP](#) split horizon works and how it impacts the routing table. You will learn why full mesh [IBGP](#) sessions within an [AS](#) are required and how next-hop is processed during route propagation within an AS.



Topology



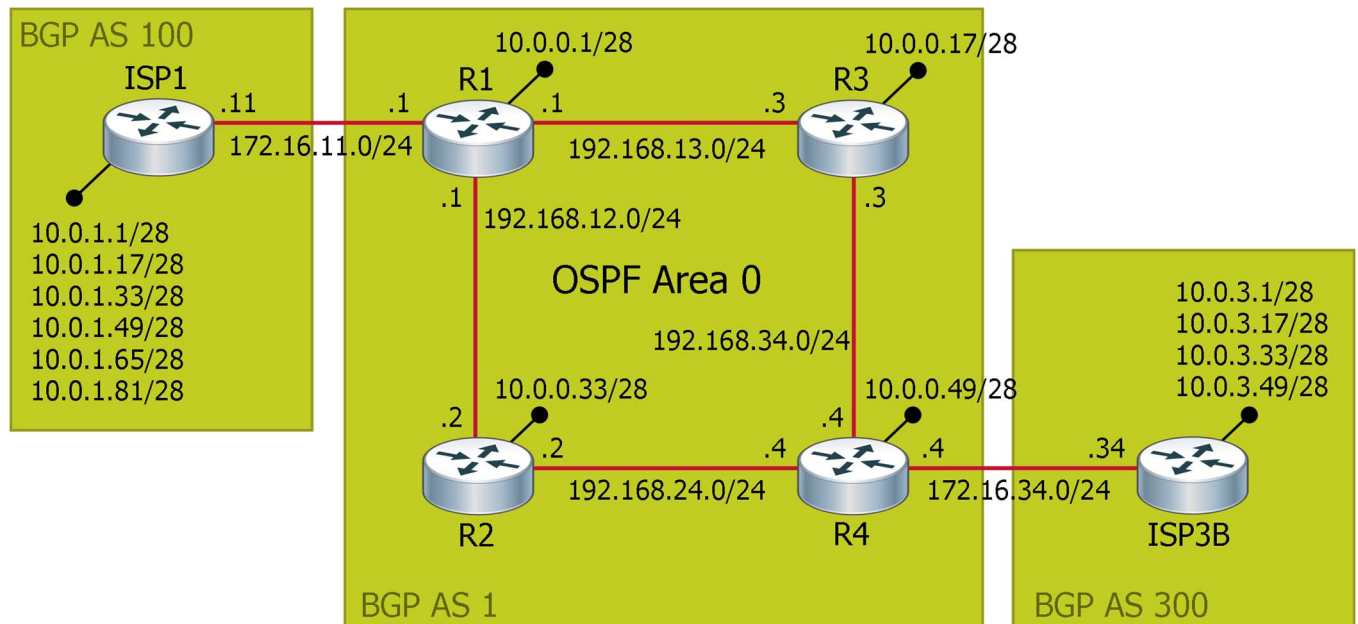
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
ISP1	Ethernet 0/0	172.16.11.11/24	Connection to R1
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
R1	Ethernet 0/0	172.16.11.1/24	Connection to ISP1
R1	Ethernet 0/2	192.168.12.1/24	Connection to R2
R1	Ethernet 0/3	192.168.13.1/24	Connection to R3
R1	Loopback 1	10.0.0.1/28	Loopback simulates LAN network
R2	Ethernet 0/2	192.168.12.2/24	Connection to R1
R2	Ethernet 0/3	192.168.24.2/24	Connection to R4
R2	Loopback 1	10.0.0.33/28	Loopback simulates LAN network
R3	Ethernet 0/2	192.168.34.3/24	Connection to R4
R3	Ethernet 0/3	192.168.13.3/24	Connection to R1
R3	Loopback 1	10.0.0.17/28	Loopback simulates LAN network
R4	Ethernet 0/0	172.16.34.4/24	Connection to ISP3B
R4	Ethernet 0/2	192.168.34.4/24	Connection to R3
R4	Ethernet 0/3	192.168.24.4/24	Connection to R2
R4	Loopback 1	10.0.0.49/28	Loopback simulates LAN network
ISP3B	Ethernet 0/0	172.16.34.34/24	Connection to R4
ISP3B	Loopback 1 Loopback 2 Loopback 3 Loopback 4	10.0.3.1/28 10.0.3.17/28 10.0.3.33/28 10.0.3.49/28	Loopbacks simulate LAN networks

[IP addresses](#) and [IGP](#) are preconfigured as shown in the topology below:



BGP is also configured as [EBGP](#) (ISP1 to R1 and ISP3B to R4) and as [IBGP](#) (R1 to R2, R1 to R3, R3 to R4, and R2 to R4). IBGP is not configured between R1 to R4 or between R2 to R3.

IBGP Full Mesh

Discovery Steps

Step 1 On the R1 router, verify what BGP routes are received from ISP1 router.

An external BGP session is established between R1 and ISP1 routers.

```
R1# show ip bgp
BGP table version is 10, local router ID is 10.0.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.0.0/28	0.0.0.0	0		32768	i
*>i 10.0.0.16/28	10.0.0.17	0	100	0	i
*>i 10.0.0.32/28	10.0.0.33	0	100	0	i
*> 10.0.1.0/28	172.16.11.11	0		0 100	i
*> 10.0.1.16/28	172.16.11.11	0		0 100	i
*> 10.0.1.32/28	172.16.11.11	0		0 100	i
*> 10.0.1.48/28	172.16.11.11	0		0 100	i
*> 10.0.1.64/28	172.16.11.11	0		0 100	i
*> 10.0.1.80/28	172.16.11.11	0		0 100	i

The R1 router received six BGP prefixes from the ISP1 router. For these prefixes, the next hop IP address is 172.16.11.11.

Step 2 On the R1 router, verify what BGP prefixes are sent to internal BGP peer R2 (10.0.0.33).

R1# **show ip bgp neighbors 10.0.0.33 advertised-routes**

BGP table version is 10, local router ID is 10.0.0.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.11.11	0		0	100 i
*>	10.0.1.16/28	172.16.11.11	0		0	100 i
*>	10.0.1.32/28	172.16.11.11	0		0	100 i
*>	10.0.1.48/28	172.16.11.11	0		0	100 i
*>	10.0.1.64/28	172.16.11.11	0		0	100 i
*>	10.0.1.80/28	172.16.11.11	0		0	100 i

Total number of prefixes 7

You can see that all the prefixes that are received from the external BGP peer are sent to the internal BGP peer.

Step 3 On the R2 router, verify if all six prefixes that are originated in AS 100 are received from R1 router.

R2# **show ip bgp**

BGP table version is 16, local router ID is 10.0.0.33

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>i	10.0.0.0/28	10.0.0.1	0	100	0	i
*>	10.0.0.32/28	0.0.0.0	0		32768	i
*>i	10.0.0.48/28	10.0.0.49	0	100	0	i
* i	10.0.1.0/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.16/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.32/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.48/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.64/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.80/28	172.16.11.11	0	100	0	100 i
* i	10.0.3.0/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.16/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.32/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.48/28	172.16.34.34	0	100	0	300 i

On the R2 router, you should see all six prefixes from AS 100 in the BGP routing table.

Step 4 On the R1 router, verify what BGP prefixes are sent to internal BGP peer R3 (10.0.0.17).

```

R1# show ip bgp neighbors 10.0.0.17 advertised-routes
BGP table version is 10, local router ID is 10.0.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.11.11	0		0	100 i
*>	10.0.1.16/28	172.16.11.11	0		0	100 i
*>	10.0.1.32/28	172.16.11.11	0		0	100 i
*>	10.0.1.48/28	172.16.11.11	0		0	100 i
*>	10.0.1.64/28	172.16.11.11	0		0	100 i
*>	10.0.1.80/28	172.16.11.11	0		0	100 i

Total number of prefixes 7

You can see all prefixes received from external BGP peer are sent to internal BGP peer.

Step 5 On the R3 router, verify if all six prefixes that are originated in AS 100 are received from R1 router.

```

R3# show ip bgp
BGP table version is 16, local router ID is 10.0.0.17
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>i	10.0.0.0/28	10.0.0.1	0	100	0	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>i	10.0.0.48/28	10.0.0.49	0	100	0	i
* i	10.0.1.0/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.16/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.32/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.48/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.64/28	172.16.11.11	0	100	0	100 i
* i	10.0.1.80/28	172.16.11.11	0	100	0	100 i
* i	10.0.3.0/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.16/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.32/28	172.16.34.34	0	100	0	300 i
* i	10.0.3.48/28	172.16.34.34	0	100	0	300 i

On the R3 router, you should see all six prefixes from AS 100 in the BGP routing table.

Step 6 On the R2 router, verify what BGP prefixes are sent to internal BGP peer R4 (10.0.0.49).

```

R2# show ip bgp neighbors 10.0.0.49 advertised-routes
BGP table version is 16, local router ID is 10.0.0.33
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop              Metric LocPrf Weight Path
*>  10.0.0.32/28     0.0.0.0                      0         32768 i

Total number of prefixes 1

```

You can see that no prefixes are received from internal BGP peer and sent to internal BGP peer. There is a single prefix, locally originated, that is sent to internal BGP peer.

Step 7 On the R3 router, verify what BGP prefixes are sent to internal BGP peer R4 (10.0.0.49).

```

R3# show ip bgp neighbors 10.0.0.49 advertised-routes
BGP table version is 16, local router ID is 10.0.0.17
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop              Metric LocPrf Weight Path
*>  10.0.0.16/28     0.0.0.0                      0         32768 i

Total number of prefixes 1

```

You can see that no prefixes are received from internal BGP peer and sent to internal BGP peer. There is a single prefix, locally originated, that is sent to internal BGP peer.

R4 router has no internal BGP session to the R1 router.

Step 8 On the R4 router, you will not find any prefixes from AS 100.

```

R4# show ip bgp
BGP table version is 8, local router ID is 10.0.0.49
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop              Metric LocPrf Weight Path
*>i 10.0.0.16/28     10.0.0.17                  0        100      0 i
*>i 10.0.0.32/28     10.0.0.33                  0        100      0 i
*>  10.0.0.48/28     0.0.0.0                    0          32768 i
*>  10.0.3.0/28      172.16.34.34               0          0 300 i
*>  10.0.3.16/28     172.16.34.34               0          0 300 i
*>  10.0.3.32/28     172.16.34.34               0          0 300 i
*>  10.0.3.48/28     172.16.34.34               0          0 300 i

```

Step 9 Configure BGP session between R1 and R4 routers.

BGP session is sourced from Loopback 1 interface. Router R1 Loopback 1 interface has IP address 10.0.0.1 and router R4 Loopback 1 interface has IP address 10.0.0.49. Both routers are in the BGP AS 1.

```
R1(config)# router bgp 1
R1(config-router)# neighbor 10.0.0.49 remote-as 1
R1(config-router)# neighbor 10.0.0.49 update-source Loopback 1
```

```
R4(config)# router bgp 1
R4(config-router)# neighbor 10.0.0.1 remote-as 1
R4(config-router)# neighbor 10.0.0.1 update-source Loopback 1
```

Step 10 On the R1 router, verify what BGP prefixes are sent to internal BGP peer R4 (10.0.0.49).

```
R1# show ip bgp neighbors 10.0.0.49 advertised-routes
BGP table version is 11, local router ID is 10.0.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.11.11	0		0	100 i
*>	10.0.1.16/28	172.16.11.11	0		0	100 i
*>	10.0.1.32/28	172.16.11.11	0		0	100 i
*>	10.0.1.48/28	172.16.11.11	0		0	100 i
*>	10.0.1.64/28	172.16.11.11	0		0	100 i
*>	10.0.1.80/28	172.16.11.11	0		0	100 i

Total number of prefixes 7

You can see all prefixes received from external BGP peer are sent to internal BGP peer.

Step 11 On the R4 router, verify if all six prefixes that are originated in AS 100, are received from R1 router.

On the R4 router, you should see all six prefixes from AS 100 in the BGP routing table.

```
R4# show ip bgp
BGP table version is 9, local router ID is 10.0.0.49
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i 10.0.0.0/28	10.0.0.1	0	100	0	i
*>i 10.0.0.16/28	10.0.0.17	0	100	0	i
*>i 10.0.0.32/28	10.0.0.33	0	100	0	i
*> 10.0.0.48/28	0.0.0.0	0		32768	i
* i 10.0.1.0/28	172.16.11.11	0	100	0	100 i
* i 10.0.1.16/28	172.16.11.11	0	100	0	100 i
* i 10.0.1.32/28	172.16.11.11	0	100	0	100 i
* i 10.0.1.48/28	172.16.11.11	0	100	0	100 i
* i 10.0.1.64/28	172.16.11.11	0	100	0	100 i
* i 10.0.1.80/28	172.16.11.11	0	100	0	100 i
*> 10.0.3.0/28	172.16.34.34	0		0	300 i
*> 10.0.3.16/28	172.16.34.34	0		0	300 i
*> 10.0.3.32/28	172.16.34.34	0		0	300 i
*> 10.0.3.48/28	172.16.34.34	0		0	300 i

Now all routers in the AS 1 have same BGP routes.

Step 12 BGP session between R2 and R3 routers is not needed in this lab setup, because R2 and R3 have no external BGP sessions. But in order to be consistent and prepare BGP to accommodate future expansions, configure BGP session between R2 and R3 routers.

BGP session is sourced from Loopback 1 interface. Router R2 Loopback 1 interface has IP address 10.0.0.33 and router R3 Loopback 1 interface has IP address 10.0.0.17. Both routers are in the BGP AS 1.

```
R2(config)# router bgp 1
R2(config-router)# neighbor 10.0.0.17 remote-as 1
R2(config-router)# neighbor 10.0.0.17 update-source Loopback 1
```

```
R3(config)# router bgp 1
R3(config-router)# neighbor 10.0.0.33 remote-as 1
R3(config-router)# neighbor 10.0.0.33 update-source Loopback 1
```

Step 13 On the R4 router, verify what BGP prefixes are sent to external BGP peer ISP3B (172.16.34.34).

```
R4# show ip bgp neighbors 172.16.34.34 advertised-routes
BGP table version is 9, local router ID is 10.0.0.49
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>i	10.0.0.0/28	10.0.0.1	0	100	0	i
*>i	10.0.0.16/28	10.0.0.17	0	100	0	i
*>i	10.0.0.32/28	10.0.0.33	0	100	0	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i

Total number of prefixes 4

You can see there are no prefixes from AS 100.

Step 14 On the R4 router, verify why routes from AS 100 are not sent to external BGP peer ISP3B (172.16.34.34).

Examine single BGP route from AS 100.

```
R4# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 0
Paths: (1 available, no best path)
  Not advertised to any peer
  Refresh Epoch 2
  100
    172.16.11.11 (inaccessible) from 10.0.0.1 (10.0.0.1)
      Origin IGP, metric 0, localpref 100, valid, internal
```

The output shows that route 10.0.1.0/28 has inaccessible next hop 172.16.11.11. This inaccessibility is a reason that the route is not marked as a good route and is not advertised toward external BGP peer.

```
R4# show ip route 172.16.11.11
% Subnet not in table
```

On the R4 router, there is no route to the next hop 172.16.11.11 in the IP routing table.

Step 15 On the R1 router, verify what is next hop IP address of the BGP route 10.0.1.0/28.

```
R1# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 3
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    2
  Refresh Epoch 1
  100
    172.16.11.11 from 172.16.11.11 (10.0.1.81)
      Origin IGP, metric 0, localpref 100, valid, external, best
```

Output shows that route 10.0.1.0/28 has valid next hop 172.16.11.11.

Transit Network Using External Next Hops

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

Transit Network Using External Next Hops

- All EBGP peers must be reachable by all BGP-speaking routers within the AS.
- EBGP next hops shall be announced using the IGP.
 - Redistribute connected interfaces into the IGP at the edge routers.
 - Include links to EBGP neighbors into the IGP and configure them as passive interfaces.

© 2015 Cisco and/or its affiliates. All rights reserved.

121

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

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

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

External subnets that are redistributed into the IGP might appear as external IGP routes, depending on what IGP is configured within the AS. There are several scalability issues that are associated with external routes in some routing protocols. For example, [OSPF](#) carries each external subnet in a separate [LSA](#) object. If route redistribution is not desirable for any reason, an alternative method is to include the subnet on which the EBGp session is running in the IGP configuration using the **network** command. To prevent the border router from exchanging IGP routing with the border router of the other AS, you must configure the interface as a passive interface. Failure to do so could cause the two autonomous systems to exchange routes using the IGP. In that case, all benefits of having separate autonomous systems would be lost.

Step 16 On the R1 router, redistribute subnet from the [Ethernet](#) 0/0, facing to the ISP1 router, into OSPF routing protocol.

Between routers in the AS 1 (R1, R2, R3, and R4) there is OSPF with process ID 1 in the area 0 preconfigured.

```
R1(config)# route-map toISP1 permit 10
R1(config-route-map)# match interface Ethernet 0/0
R1(config-route-map)# exit
R1(config)# router ospf 1
R1(config-router)# redistribute connected subnets route-map toISP1
```

You need to use **subnets** keyword, because subnet of the network 172.16.0.0/16 is used between R1 and ISP1 routers.

Step 17 On the R4 router, examine BGP route 10.0.1.0/28 from AS 100 again.

```
R4# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 10
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    1
  Refresh Epoch 3
  100
    172.16.11.11 (metric 20) from 10.0.0.1 (10.0.0.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

Output shows that route 10.0.1.0/28 has valid next hop 172.16.11.11 and is selected as best route.

```
R4# show ip route 172.16.11.11
Routing entry for 172.16.11.0/24
  Known via "ospf 1", distance 110, metric 20, type extern 2, forward metric 20
  Last update from 192.168.34.3 on Ethernet0/2, 00:10:23 ago
  Routing Descriptor Blocks:
    192.168.34.3, from 10.0.0.1, 00:10:23 ago, via Ethernet0/2
      Route metric is 20, traffic share count is 1
    * 192.168.24.2, from 10.0.0.1, 00:10:23 ago, via Ethernet0/3
      Route metric is 20, traffic share count is 1
```

On the R4 router, there is also routing information to reach the next hop 172.16.11.11.

Step 18 On the R4 router, verify what BGP prefixes are sent to external BGP peer ISP3B (172.16.34.34).

```
R4# show ip bgp neighbors 172.16.34.34 advertised-routes
BGP table version is 15, local router ID is 10.0.0.49
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i 10.0.0.0/28	10.0.0.1	0	100	0	i
*>i 10.0.0.16/28	10.0.0.17	0	100	0	i
*>i 10.0.0.32/28	10.0.0.33	0	100	0	i
*> 10.0.0.48/28	0.0.0.0	0		32768	i
*>i 10.0.1.0/28	172.16.11.11	0	100	0	100 i
*>i 10.0.1.16/28	172.16.11.11	0	100	0	100 i
*>i 10.0.1.32/28	172.16.11.11	0	100	0	100 i
*>i 10.0.1.48/28	172.16.11.11	0	100	0	100 i
*>i 10.0.1.64/28	172.16.11.11	0	100	0	100 i
*>i 10.0.1.80/28	172.16.11.11	0	100	0	100 i

Total number of prefixes 10

You can see there are also prefixes from AS 100 advertised to external BGP peer ISP3B.

Step 19 On the ISP3B router, verify what is next hop IP address of the BGP route 10.0.1.0/28.

```
ISP3B# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 10
Paths: (1 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  1 100
    172.16.34.4 from 172.16.34.4 (10.0.0.49)
      Origin IGP, localpref 100, valid, external, best
```

Output shows that route 10.0.1.0/28 has different next hop IP address (172.16.34.4), as seen on the R4 router (172.16.11.11). When BGP route is sent to external BGP peer, the next-hop IP address changes.

Transit Network Using Edge Routers as Next Hops

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

Transit Network Using Edge Routers as Next Hops

- **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.

© 2015 Cisco and/or its affiliates. All rights reserved.

122

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

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

Step 20 On the R1 router, remove the previously configured redistribution.

```
R1(config)# router ospf 1
R1(config-router)# no redistribute connected
```

On the R4 router, the BGP route (10.0.1.0/28) next hop IP address 172.16.11.11 will again become inaccessible.

```
R4# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 16
Paths: (1 available, no best path)
  Not advertised to any peer
  Refresh Epoch 3
  100
    172.16.11.11 (inaccessible) from 10.0.0.1 (10.0.0.1)
      Origin IGP, metric 0, localpref 100, valid, internal
```

Step 21 Configure R1 router to be next hop for all IBGP speaking routers.

The IBGP speaking routers of the R1 router are R2 (10.0.0.33), R3 (10.0.0.17), and R4 (10.0.0.49).

```
R1(config)# router bgp 1
R1(config-router)# neighbor 10.0.0.17 next-hop-self
R1(config-router)# neighbor 10.0.0.33 next-hop-self
R1(config-router)# neighbor 10.0.0.49 next-hop-self
```

Step 22 On the R4 router, examine BGP route 10.0.1.0/28 from AS 100 again.

```
R4# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 22
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    1
  Refresh Epoch 3
  100
    10.0.0.1 (metric 21) from 10.0.0.1 (10.0.0.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

Output shows that route 10.0.1.0/28 has valid next hop 10.0.0.1 and is selected as best route. Next hop 10.0.0.1 is R1 Loopback 1 interface.

```
R4# show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 21, type intra area
  Last update from 192.168.34.3 on Ethernet0/2, 00:46:42 ago
  Routing Descriptor Blocks:
    192.168.34.3, from 10.0.0.1, 00:46:42 ago, via Ethernet0/2
      Route metric is 21, traffic share count is 1
  * 192.168.24.2, from 10.0.0.1, 00:46:42 ago, via Ethernet0/3
      Route metric is 21, traffic share count is 1
```

On the R4 router, there is routing information to reach the next hop 10.0.0.1.

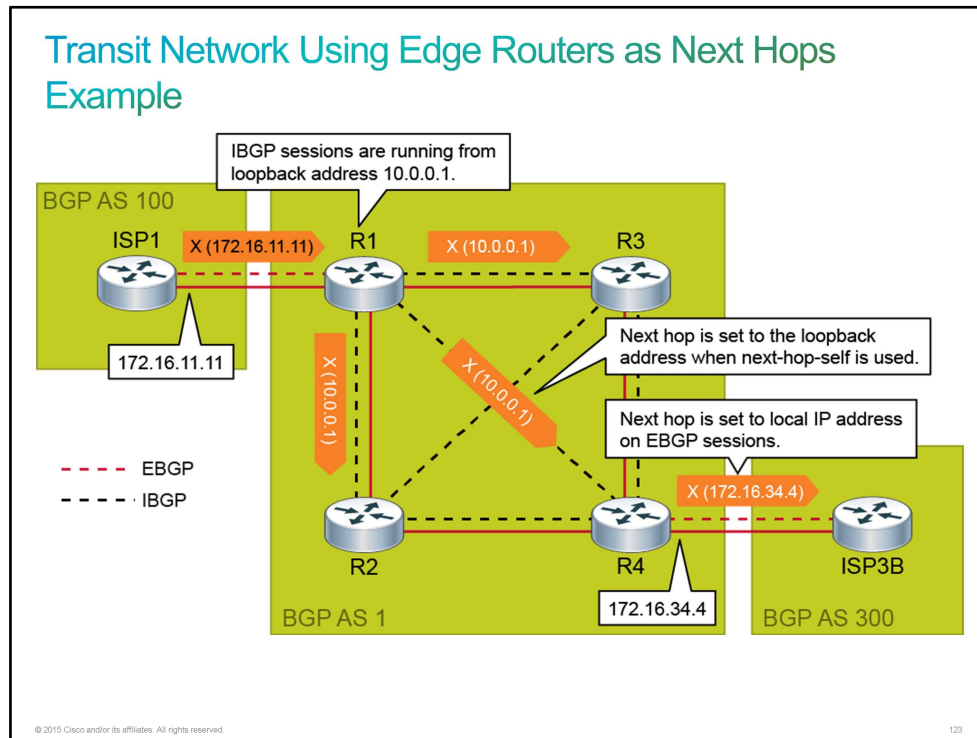
Step 23 On the ISP3B router, verify what is next hop IP address of the BGP route 10.0.1.0/28.

```
ISP3B# show ip bgp 10.0.1.0/28
BGP routing table entry for 10.0.1.0/28, version 22
Paths: (1 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  1 100
    172.16.34.4 from 172.16.34.4 (10.0.0.49)
      Origin IGP, localpref 100, valid, external, best
```

The output shows that route 10.0.1.0/28 has different next hop IP address (172.16.34.4), as seen on the R4 router (10.0.0.1).

Transit Network Using Edge Routers as Next Hops Example

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



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

In the figure, ISP1 sends a BGP update about network X. Because it is sending a BGP update over an EBGP session to R1, the next-hop attribute is set to the IP address that is used at the ISP1 side of the EBGP session, 172.16.11.11. R1 can use this information and route packets to network X by forwarding them to ISP1.

R1 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 R2, R3, and R4 will get information that they can reach network X by forwarding packets to 10.0.0.1. But that address is not directly connected. The routers will inspect the routing table to see if and how they can reach 10.0.0.1. They can then route packets to network X in the same direction that they would use to route packets to 10.0.0.1.

R4 also forwards the BGP update about network X to ISP3B. This session is an EBGP session, which means that R4 will set the next-hop attribute to its own IP address that is used on that EBGP session, 172.16.34.4.

Differences Between EBGP and IBGP Sessions

Both [EBGP](#) and [IBGP](#) sessions forward [BGP](#) updates; however, they do it in slightly different ways.

Differences Between EBGP and IBGP Sessions

- No BGP attributes are changed in IBGP updates.
- Because of BGP split horizon, routes that are learned from an IBGP peer are not advertised to other IBGP peers.
- Local preference is propagated only over IBGP sessions.
- EBGP peers are directly connected; IBGP peers are usually distant.
- Route selection rules slightly prefer EBGP routes.

© 2015 Cisco and/or its affiliates. All rights reserved.

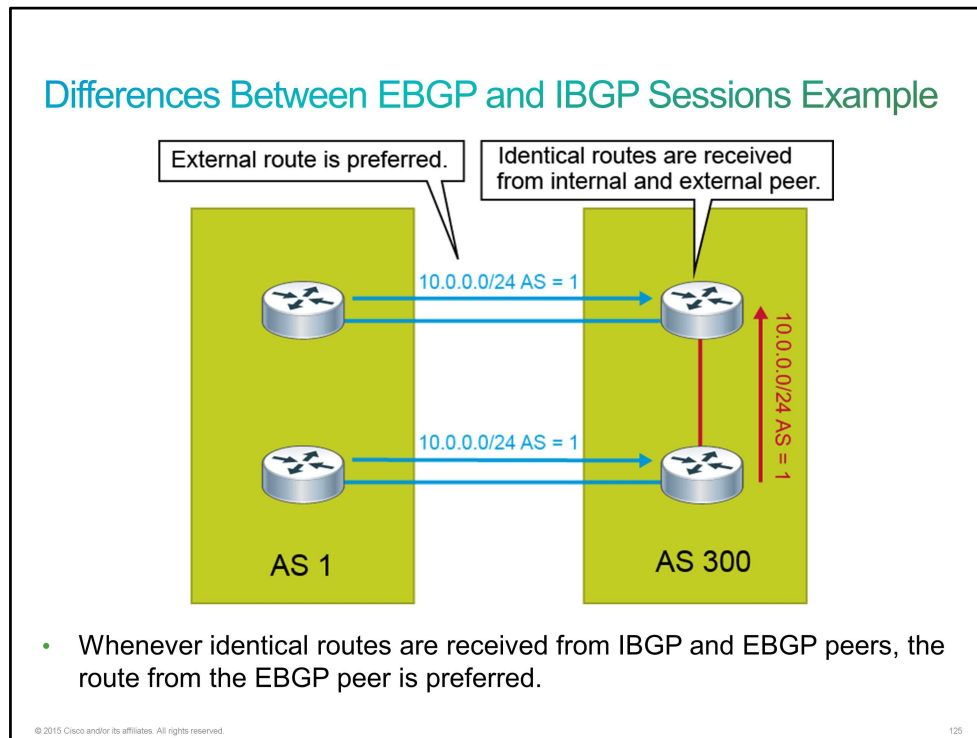
124

Here are the differences between EBGP and IBGP sessions in forwarding BGP updates:

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

Differences Between EBG and IBGP Sessions Example

This example illustrates the preference of the EBG route.



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

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

In the figure, the upper router in AS 300 receives BGP updates about network 10.0.0.0/24 over two different paths. One update is received over the EBG session to AS 1. The other update is received over the IBGP session to the lower router in AS 300. All essential attributes are the same, so route selection cannot be made easily.

The upper router in AS 300 realizes that IP packets with destination addresses within network 10.0.0.0/24 should sooner rather than later leave AS 300. It is better to make them leave the AS right away. So the update that was received on the EBG session is preferred over the update that was received on the IBGP session.

Scalability Limitations of IBGP-Based Transit Backbones

The requirements that are associated with [IBGP](#)-based transit backbones can result in limited scalability.

Scalability Limitations of IBGP-Based Transit Backbones

Transit backbone requires IBGP full mesh between all core routers.

- Large number of TCP sessions
- Unnecessary, duplicate routing traffic

There are two scalability solutions:

- Route reflectors
- BGP confederations

© 2015 Cisco and/or its affiliates. All rights reserved. 125

IBGP split-horizon rules mandate an IBGP connection between every border router and every other [BGP](#) router in an [AS](#).

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

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

Two solutions are available:

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- All BGP routing updates carry the mandatory well-known attribute AS-path. The AS-path attribute is not changed when the BGP prefix is propagated across IBGP sessions.
- The IBGP multipath load sharing feature enables the BGP speaking router to select multiple IBGP paths as the best paths to a destination.
- IBGP split horizon prevents routing information loops within the AS.
- All BGP routers within an AS must have IBGP sessions with every other BGP router in the AS.

© 2015 Cisco and/or its affiliates. All rights reserved.

127

Summary (Cont.)

- For recursive routing to work, a router must resolve all possible next-hop references that use information in the routing table.
- 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.
- BGP attributes are not changed when an update is sent across an IBGP session unless next-hop-self is configured.
- 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.

© 2015 Cisco and/or its affiliates. All rights reserved.

128

Lesson 3: Forwarding Packets in Transit AS

Overview

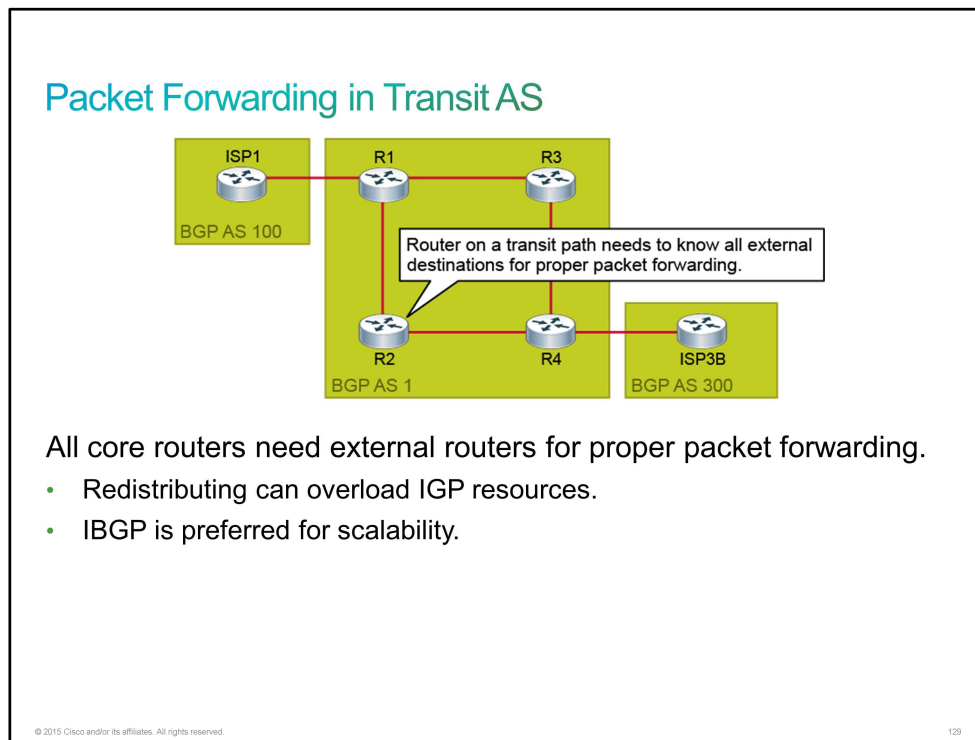
A transit [AS](#) requires interaction between [EBGP](#) and [IBGP](#) and between IBGP and an [IGP](#) in the transit AS. This lesson describes packet forwarding through a transit AS and discusses the requirements for successful packet forwarding, such as recursive route lookup and an IGP in the transit AS. This lesson concludes with a discussion of the interaction between IBGP and an IGP running within the transit AS.

Upon completing this lesson, you will be able to:

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

Packet Forwarding in Transit AS

When [BGP](#) updates have propagated through the transit [AS](#) to all neighboring autonomous systems, the [IP](#) traffic can start to flow.



In the figure, ISP3B router forwards to R4 IP packets with the destination address matching a network in AS 100. R4 checks its routing table and finds that there is a [BGP](#) route for that destination. The BGP route has a next-hop reference, which points to the far end of the [EBGP](#) session between ISP1 and R1. So R4 once again checks the routing table and finds that it should forward the packet to R2 in this case.

Thus, R2 receives the IP packet with a destination address indicating a host within AS 100. To be able to forward this packet, R2 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 R2 could receive another packet, coming from the other direction and destined for AS 300.

The conclusion is that both R2 and R3, to handle all possible cases, must have routing information to all the external networks that R1 and R4 have. The only scalable way of providing routers with this information is to update R2 and R3 with [IBGP](#) from both R1 and R4.

In theory, the external information that R1 and R4 receive could be redistributed by these ingress routers into the [IGP](#) in use within the transit AS. However, no IGP can handle the volume of information that BGP can. So there would always be a risk that the IGP would break because of information overload, causing a total network meltdown in the AS. The volume of routing information that the BGP carries in the contemporary Internet long ago passed the limits of what it is possible to carry in any IGP.

Packet Forwarding in Transit AS (Cont.)

- Routes that are learned via BGP do not have an outgoing interface associated with them in the routing table.
- Recursive lookup is performed to forward IP packets toward external destinations.

```
R2# show ip route 10.0.1.0 255.255.255.240
Routing entry for 10.0.1.0/28
  Known via "bgp 1", distance 200, metric 0
  Tag 100, type internal
  Last update from 10.0.0.1 16:37:59 ago
  Routing Descriptor Blocks:
    * 10.0.0.1, from 10.0.0.1, 16:37:59 ago
      Route metric is 0, traffic share count is 1
      AS Hops 1
      Route tag 100
      MPLS label: none
```

- Observe that the route that is known via BGP has no outgoing interface, but it has a BGP next hop.

© 2015 Cisco and/or its affiliates. All rights reserved.

130

Packet Forwarding in Transit AS (Cont.)

```
R2# show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 11, type intra area
  Last update from 192.168.12.1 on Ethernet0/2, 17:23:10 ago
  Routing Descriptor Blocks:
    * 192.168.12.1, from 10.0.0.1, 17:23:10 ago, via Ethernet0/2
      Route metric is 11, traffic share count is 1
```

- The route toward BGP next hop is known via OSPF and has an outgoing interface.

© 2015 Cisco and/or its affiliates. All rights reserved.

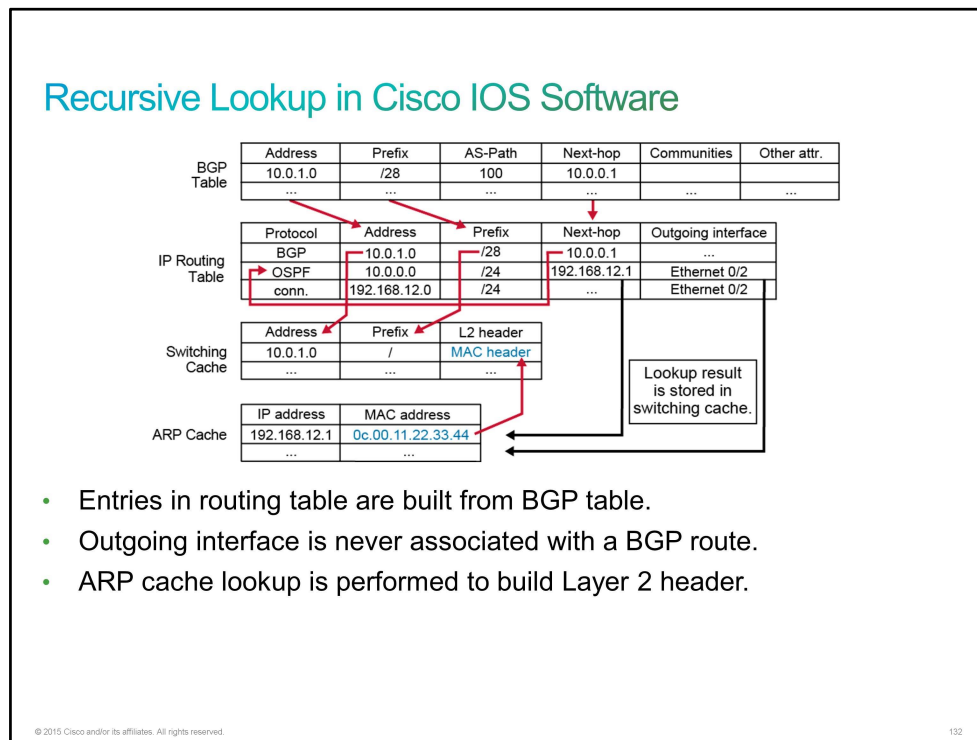
131

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

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

Recursive Lookup in Cisco IOS Software

Recursive lookup refers to the [BGP](#) routes for which another routing lookup is required in order to resolve the BGP route next hop.



The figure presents the steps in the recursive lookup process in Cisco IOS software. The router has received a BGP update about network 10.0.1.0/28. It was associated with an [AS-path](#) attribute set to 100, a next-hop attribute set to the [IP address](#) 10.0.0.1. Some other attributes were also carried with the update.

Because the next-hop address 10.0.0.1 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 10.0.1.0/28 is received, the router searches the routing table and finds the installed BGP route. The router takes the indicated next-hop address 10.0.0.1 and searches the routing table again. It now finds a match with the [OSPF](#) route to subnet 10.0.0.0/24. The 10.0.0.0/24 route has an outgoing interface set to interface [Ethernet](#) 0/2 and a next hop set to 192.168.12.1. So, the packets that are destined for network 10.0.1.0/28 should be forwarded via 192.168.12.1, which is directly reachable over Ethernet 0/2. The [ARP](#) table is used to find the [MAC address](#) for IP address 192.168.12.1. The MAC address is used to forward the IP packet to network 10.0.1.0/28 out the Ethernet 0/2 interface. The [MAC](#) header is stored in the cache for successive packets to network 10.0.1.0/28.

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

Recursive Lookup in Cisco IOS Software (Cont.)

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

© 2015 Cisco and/or its affiliates. All rights reserved.

133

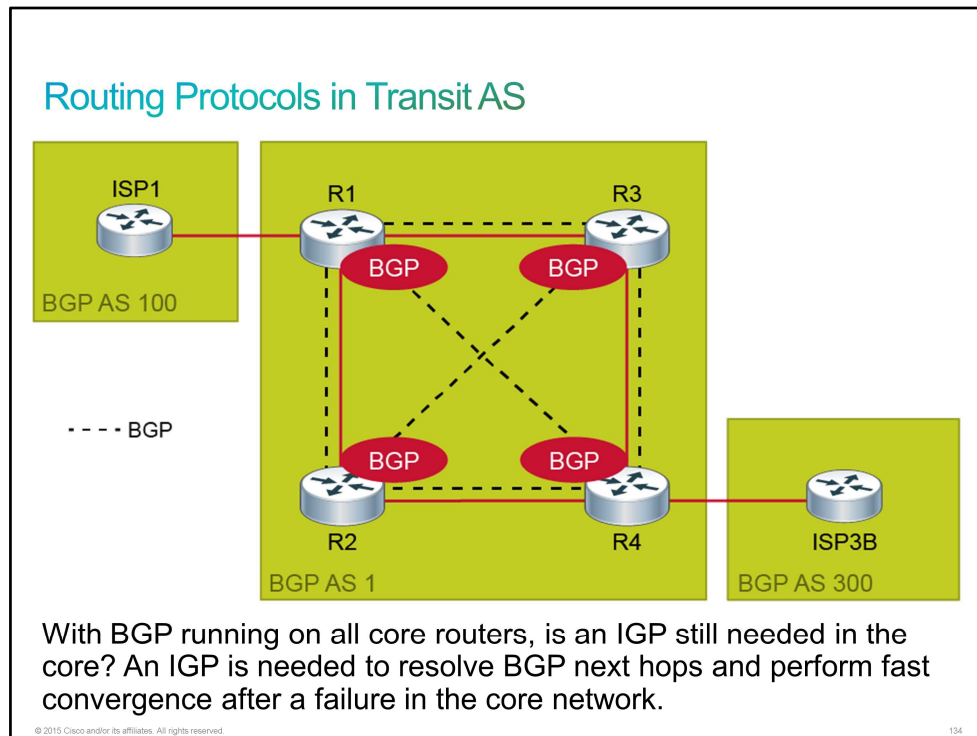
Traditional Cisco IOS switching mechanisms used the traffic-driven, cache-based switching approach. Both fast switching and optimum switching populate the IP switching cache on demand, meaning that before any IP packets are forwarded, the cache is empty. After the first packet to a specific destination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

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

Routing Protocols in Transit AS

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

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



When R4 in the figure receives an [IP](#) packet with the destination in AS 100, it does a recursive lookup to find the outgoing interface to be used for packet forwarding. It performs the recursive lookup that is based on the [IGP](#) information. If there is suddenly an internal problem within AS 1, and the next-hop address is reachable a different way, the IGP determines this fact. The router changes the IGP route to the next-hop network because of the newly received IGP route information, and all cache entries that rely on the old information are invalidated. The next recursive lookup that R4 performs will indicate a different outgoing interface than before the problem occurred.

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

The packet forwarding to external destinations thus benefits from the high-speed convergence that the IGP offers. 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 Transit AS (Cont.)

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

© 2015 Cisco and/or its affiliates. All rights reserved.

135

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, this information includes only the links within the core network, the loopback interfaces, and the external subnets that are used in EBGp sessions with the 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. This approach reduces the amount of information in the IGP and increases IGP stability.

BGP and IGP Interaction

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

BGP and IGP Interaction

Ideally, there will be no interaction between BGP and the IGP.

- BGP carries external and customer routes.
- The IGP carries only core subnets.
- External route flaps do not affect IGP.
- Failures internal to the network do not affect BGP as long as the BGP next hop remains reachable.
- The only link between BGP and the IGP should be the recursive lookup.

© 2015 Cisco and/or its affiliates. All rights reserved.

136

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

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

BGP and IGP Interaction (Cont.)

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

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

Routing Protocol Default Administrative Distance

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

© 2015 Cisco and/or its affiliates. All rights reserved.

137

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

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

BGP considers both EBG/IBGP routes in the BGP selection process. BGP will therefore never try to install both an EBG 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 EBG route as the best route for a given destination network, it will try to install that route with a very low AD. So, the routes that are learned via EBG have a high likelihood of being installed in the routing table.

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

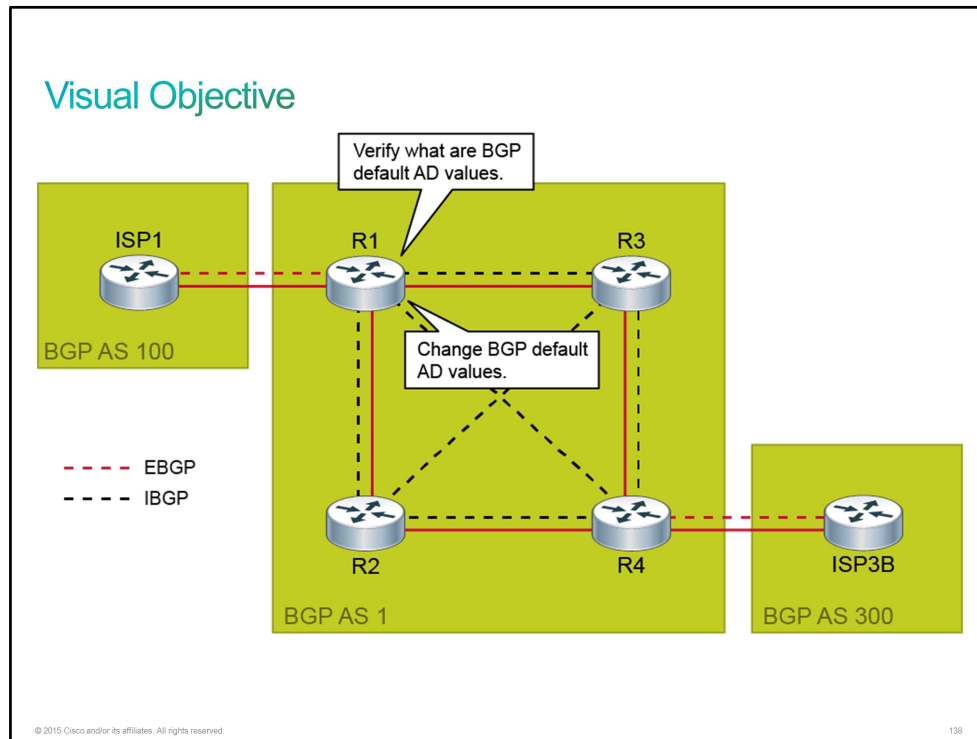
All IGPs, such as [EIGRP](#), [OSPF](#), [IS-IS](#), and so on, have a medium likelihood of being installed. The ADs for IGPs fall between the ADs of EBG and IBGP.

Note The reason for giving EBG a low default AD is because EBG 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.

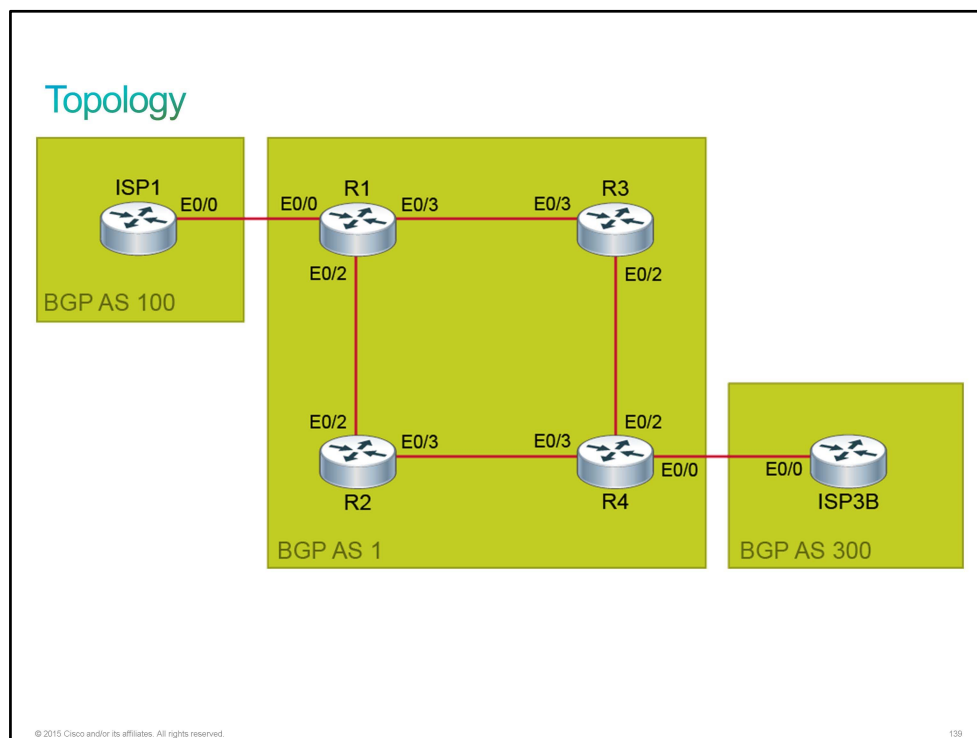
Discovery 6: BGP Administrative Distance

Overview

Through this discovery, you will learn how to change BGP administrative distances.



Topology



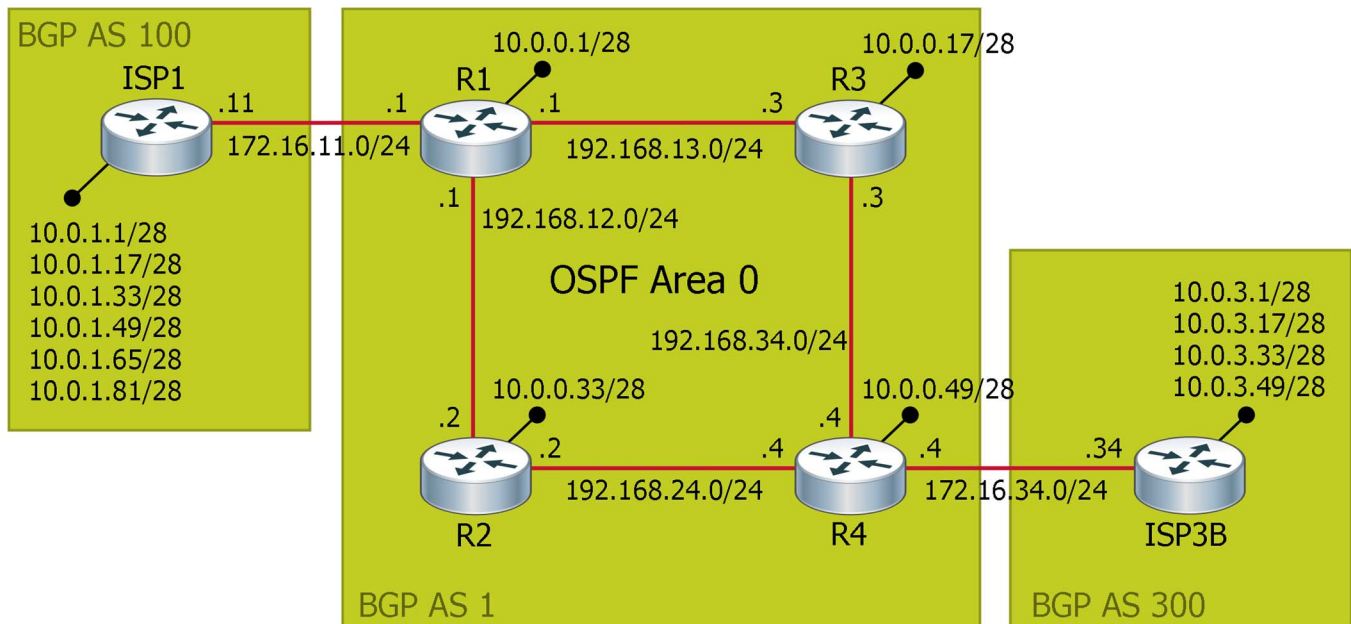
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
ISP1	Ethernet 0/0	172.16.11.11/24	Connection to R1
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
R1	Ethernet 0/0	172.16.11.1/24	Connection to ISP1
R1	Ethernet 0/2	192.168.12.1/24	Connection to R2
R1	Ethernet 0/3	192.168.13.1/24	Connection to R3
R1	Loopback 1	10.0.0.1/28	Loopback simulates LAN network
R2	Ethernet 0/2	192.168.12.2/24	Connection to R1
R2	Ethernet 0/3	192.168.24.2/24	Connection to R4
R2	Loopback 1	10.0.0.33/28	Loopback simulates LAN network
R3	Ethernet 0/2	192.168.34.3/24	Connection to R4
R3	Ethernet 0/3	192.168.13.3/24	Connection to R1
R3	Loopback 1	10.0.0.17/28	Loopback simulates LAN network
R4	Ethernet 0/0	172.16.34.4/24	Connection to ISP3B
R4	Ethernet 0/2	192.168.34.4/24	Connection to R3
R4	Ethernet 0/3	192.168.24.4/24	Connection to R2
R4	Loopback 1	10.0.0.49/28	Loopback simulates LAN network
ISP3B	Ethernet 0/0	172.16.34.34/24	Connection to R4
ISP3B	Loopback 1 Loopback 2 Loopback 3 Loopback 4	10.0.3.1/28 10.0.3.17/28 10.0.3.33/28 10.0.3.49/28	Loopbacks simulate LAN networks

IP addresses and IGP are preconfigured as shown in the topology below:



BGP is also configured as EBGP (ISP1 to R1 and ISP3B to R4) and as IBGP (full mesh between R1, R2, R3, and R4).

BGP Administrative Distance

Discovery Steps

Step 1 On the R1 router, verify what are BGP administrative distances.

```

R1# show ip protocols
*** IP Routing is NSF aware ***

Routing Protocol is "bgp 1"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  IGP synchronization is disabled
  Automatic route summarization is disabled
  Neighbor(s):
    Address                FiltIn FiltOut DistIn DistOut Weight RouteMap
    10.0.0.17
    10.0.0.33
    10.0.0.49
    172.16.11.11
  Maximum path: 1
  Routing Information Sources:
    Gateway                Distance    Last Update
    10.0.0.17                200        00:04:14
    10.0.0.33                200        00:04:33
    10.0.0.49                200        00:04:09
    172.16.11.11            20         00:04:33
  Distance: external 20 internal 200 local 200

Routing Protocol is "ospf 1"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  Router ID 10.0.0.1
  Number of areas in this router is 1. 1 normal 0 stub 0 nssa
  Maximum path: 4
  Routing for Networks:
    10.0.0.0 0.0.0.255 area 0
    192.168.0.0 0.0.255.255 area 0
  Routing Information Sources:
    Gateway                Distance    Last Update
    10.0.0.17                110        00:09:56
    10.0.0.33                110        00:10:11
    10.0.0.49                110        00:09:32
  Distance: (default is 110)

```

The BGP external AD is 20, internal and local AD are 200. You can also learn OSPF AD value.

Changing Administrative Distance of BGP Routes

You can change BGP administrative distances, to allow the use of BGP route, that could be a better route to a node.

Changing Administrative Distance of BGP Routes

```
router(config-router)# distance bgp external-distance internal-distance local-distance
```

- This command sets the AD for EBGp, IBGP, and local routes.
- This change applies only to routes received after the command has been entered (similar to filters).
- Defaults: EBGp routes have a distance of 20; IBGP and local routes have a distance of 200.
- *The defaults are usually correct; do not change them.*

© 2015 Cisco and/or its affiliates. All rights reserved.

140

Step 2 On the R1 router, change EBGp AD into 100 and IBGP AD into 190.

Leave default BGP local AD.

```
R1(config)# router bgp 1  
R1(config-router)# distance bgp 100 190 200
```

Step 3 On the R1 router, verify what are BGP administrative distances again.

```

R1# show ip protocols
*** IP Routing is NSF aware ***

Routing Protocol is "bgp 1"
  Outgoing update filter list for all interfaces is not set
  Incoming update filter list for all interfaces is not set
  IGP synchronization is disabled
  Automatic route summarization is disabled
  Neighbor(s):
    Address           FiltIn FiltOut DistIn DistOut Weight RouteMap
    10.0.0.17
    10.0.0.33
    10.0.0.49
    172.16.11.11
  Maximum path: 1
  Routing Information Sources:
    Gateway           Distance      Last Update
    10.0.0.17          200           00:19:41
    10.0.0.33          200           00:20:00
    10.0.0.49          200           00:19:35
    172.16.11.11       20            00:19:59
  Distance: external 100 internal 190 local 200
<... output omitted ...>

```

The BGP EBGp and IBGP ADs changed.

Step 4 On the R1 router, verify IP routing table.

R1# show ip route

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default, U - per-user static route
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
+ - replicated route, % - next hop override

Gateway of last resort is not set

```
10.0.0.0/8 is variably subnetted, 18 subnets, 2 masks
C    10.0.0.0/28 is directly connected, Loopback1
L    10.0.0.1/32 is directly connected, Loopback1
B    10.0.0.16/28 [200/0] via 10.0.0.17, 01:21:20
O    10.0.0.17/32 [110/11] via 192.168.13.3, 01:22:35, Ethernet0/3
B    10.0.0.32/28 [200/0] via 10.0.0.33, 01:21:39
O    10.0.0.33/32 [110/11] via 192.168.12.2, 01:22:50, Ethernet0/2
B    10.0.0.48/28 [200/0] via 10.0.0.49, 01:21:14
O    10.0.0.49/32 [110/21] via 192.168.13.3, 01:22:11, Ethernet0/3
    [110/21] via 192.168.12.2, 01:22:11, Ethernet0/2
B    10.0.1.0/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.1.16/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.1.32/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.1.48/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.1.64/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.1.80/28 [20/0] via 172.16.11.11, 01:21:39
B    10.0.3.0/28 [200/0] via 10.0.0.49, 01:21:14
B    10.0.3.16/28 [200/0] via 10.0.0.49, 01:21:14
B    10.0.3.32/28 [200/0] via 10.0.0.49, 01:21:14
B    10.0.3.48/28 [200/0] via 10.0.0.49, 01:21:14
172.16.0.0/16 is variably subnetted, 2 subnets, 2 masks
C    172.16.11.0/24 is directly connected, Ethernet0/0
L    172.16.11.1/32 is directly connected, Ethernet0/0
192.168.12.0/24 is variably subnetted, 2 subnets, 2 masks
C    192.168.12.0/24 is directly connected, Ethernet0/2
L    192.168.12.1/32 is directly connected, Ethernet0/2
192.168.13.0/24 is variably subnetted, 2 subnets, 2 masks
C    192.168.13.0/24 is directly connected, Ethernet0/3
L    192.168.13.1/32 is directly connected, Ethernet0/3
O    192.168.24.0/24 [110/20] via 192.168.12.2, 01:22:21, Ethernet0/2
O    192.168.34.0/24 [110/20] via 192.168.13.3, 01:22:21, Ethernet0/3
```

You can still see BGP routes are using default AD values.

Step 5 On the R1 router, clear BGP sessions.


```
R1# clear ip bgp *
R1#
%BGP-5-ADJCHANGE: neighbor 10.0.0.17 Down User reset
%BGP_SESSION-5-ADJCHANGE: neighbor 10.0.0.17 IPv4 Unicast topology base removed from
session User reset
%BGP-5-ADJCHANGE: neighbor 10.0.0.33 Down User reset
%BGP_SESSION-5-ADJCHANGE: neighbor 10.0.0.33 IPv4 Unicast topology base removed from
session User reset
%BGP-5-ADJCHANGE: neighbor 10.0.0.49 Down User reset
%BGP_SESSION-5-ADJCHANGE: neighbor 10.0.0.49 IPv4 Unicast topology base removed from
session User reset
%BGP-5-ADJCHANGE: neighbor 172.16.11.11 Down User reset
%BGP_SESSION-5-ADJCHANGE: neighbor 172.16.11.11 IPv4 Unicast topology base removed from
session User reset
%BGP-5-ADJCHANGE: neighbor 10.0.0.33 Up
%BGP-5-ADJCHANGE: neighbor 10.0.0.17 Up
%BGP-5-ADJCHANGE: neighbor 172.16.11.11 Up
%BGP-5-ADJCHANGE: neighbor 10.0.0.49 Up
```

After all BGP neighbors will come "Up," proceed to the next step.

Step 6 On the R1 router, verify IP routing table again.

R1# show ip route

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default, U - per-user static route
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
+ - replicated route, % - next hop override

Gateway of last resort is not set

```
10.0.0.0/8 is variably subnetted, 18 subnets, 2 masks
C    10.0.0.0/28 is directly connected, Loopback1
L    10.0.0.1/32 is directly connected, Loopback1
B    10.0.0.16/28 [190/0] via 10.0.0.17, 00:01:28
O    10.0.0.17/32 [110/11] via 192.168.13.3, 01:29:04, Ethernet0/3
B    10.0.0.32/28 [190/0] via 10.0.0.33, 00:01:28
O    10.0.0.33/32 [110/11] via 192.168.12.2, 01:29:19, Ethernet0/2
B    10.0.0.48/28 [190/0] via 10.0.0.49, 00:01:28
O    10.0.0.49/32 [110/21] via 192.168.13.3, 01:28:40, Ethernet0/3
    [110/21] via 192.168.12.2, 01:28:40, Ethernet0/2
B    10.0.1.0/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.1.16/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.1.32/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.1.48/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.1.64/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.1.80/28 [100/0] via 172.16.11.11, 00:01:28
B    10.0.3.0/28 [190/0] via 10.0.0.49, 00:01:28
B    10.0.3.16/28 [190/0] via 10.0.0.49, 00:01:28
B    10.0.3.32/28 [190/0] via 10.0.0.49, 00:01:28
B    10.0.3.48/28 [190/0] via 10.0.0.49, 00:01:28
172.16.0.0/16 is variably subnetted, 2 subnets, 2 masks
C    172.16.11.0/24 is directly connected, Ethernet0/0
L    172.16.11.1/32 is directly connected, Ethernet0/0
192.168.12.0/24 is variably subnetted, 2 subnets, 2 masks
C    192.168.12.0/24 is directly connected, Ethernet0/2
L    192.168.12.1/32 is directly connected, Ethernet0/2
192.168.13.0/24 is variably subnetted, 2 subnets, 2 masks
C    192.168.13.0/24 is directly connected, Ethernet0/3
L    192.168.13.1/32 is directly connected, Ethernet0/3
O    192.168.24.0/24 [110/20] via 192.168.12.2, 01:28:50, Ethernet0/2
O    192.168.34.0/24 [110/20] via 192.168.13.3, 01:28:50, Ethernet0/3
```

Now, you can see that BGP routes AD values have changed per configuration.

Problems with BGP and IGP Interaction

There are several potential problems that might arise from [BGP](#) and [IGP](#) interaction.

Problems with BGP and IGP Interaction

If an IGP route is learned through EBGp, the EBGp route will take precedence.

- Potential causes include bad network design, routing problems, or denial-of-service attack.
- Protect IGP routes with inbound prefix-list filters at AS edges.
- Routers should never accept information about local subnets from an external source.

© 2015 Cisco and/or its affiliates. All rights reserved.

141

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

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- All core routers need external routers for proper packet forwarding.
- 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 that IGP offers.
- The IGP should provide reachability toward BGP next-hop addresses only.
- 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.

Lesson 4: Monitoring and Troubleshooting IBGP in Transit AS

Overview

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

Upon completing this lesson, you will be able to:

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

Monitoring IBGP

To display information about the [TCP](#) and [BGP](#) connections to neighbors, use the **show ip bgp neighbors** EXEC command. To display entries in the BGP routing table, use the **show ip bgp** EXEC command.

Monitoring IBGP

```
router> show ip bgp neighbors
```

- Displays whether a neighbor is an IBGP neighbor

```
router> show ip bgp
```

- Uses a special marker (i) for IBGP routes

```
router> show ip bgp prefix
```

- Displays whether the prefix is an IBGP route

© 2015 Cisco and/or its affiliates. All rights reserved. 143

Monitoring IBGP Example

This example shows the commands to monitor [IBGP](#).

Monitoring IBGP Example

```
R1# show ip bgp neighbors 10.0.0.49
BGP neighbor is 10.0.0.49, remote AS 1, internal link
  BGP version 4, remote router ID 10.0.0.49
  BGP state = Established, up for 00:47:57
  Last read 00:00:00, last write 00:00:37, hold time is 180, keepalive interval
  is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  <... output omitted ...>
```

- Internal BGP neighbor

```
R1# show ip bgp neighbors 172.16.11.11
BGP neighbor is 172.16.11.11, remote AS 100, external link
  BGP version 4, remote router ID 10.0.1.81
  BGP state = Established, up for 00:49:24
  Last read 00:00:37, last write 00:00:51, hold time is 180, keepalive interval
  is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  <... output omitted ...>
```

- External BGP neighbor

© 2015 Cisco and/or its affiliates. All rights reserved.

144

The **show ip bgp neighbors** command displays whether a router is running an IBGP (internal) or [EBGP](#) (external) session with a BGP neighbor. The "internal link" or "external link" phrases provide you with this information.

Monitoring IBGP Example (Cont.)

```
R1# show ip bgp 10.0.3.0
BGP routing table entry for 10.0.3.0/28, version 12
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    2
  Refresh Epoch 1
  300
    10.0.0.49 (metric 21) from 10.0.0.49 (10.0.0.49)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

- Route that is received from internal BGP neighbor.

```
R1# show ip bgp 10.0.1.0
BGP routing table entry for 10.0.1.0/28, version 4
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    1
  Refresh Epoch 1
  100
    172.16.11.11 from 172.16.11.11 (10.0.1.81)
      Origin IGP, metric 0, localpref 100, valid, external, best
```

- Route that is received from external BGP neighbor.

© 2015 Cisco and/or its affiliates. All rights reserved.

145

The **show ip bgp prefix** command displays whether a BGP route was received from an IBGP (internal) or EBGP (external) neighbor. The word "internal" or "external" provides you with this information.

Common IBGP Problems

Troubleshooting the [BGP](#) configuration of a transit [AS](#) can be cumbersome, because there are a number of common pitfalls that you might encounter.

Common IBGP Problems

Three of the most common problems are the following:

- IBGP sessions do not start (do not reach the Established state).
- IBGP route is in the BGP table but is not selected.
- IBGP route is selected but is not entered in the routing table.

Troubleshooting IBGP Session Startup Issues

There are several approaches to [IBGP](#) session startup issues troubleshooting.

Troubleshooting IBGP Session Startup Issues

Symptom:

- IBGP session does not start.

Diagnosis:

- IBGP session is run between loopbacks, and **update-source** keyword is missing.

Verification:

- Use **debug ip tcp transactions**. You should see BGP sessions coming from unexpected IP addresses.

© 2015 Cisco and/or its affiliates. All rights reserved.147

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

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

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

Troubleshooting IBGP Session Startup Issues (Cont.)

Symptom:

- IBGP session does not start.

Diagnosis:

- Loopback interfaces are not reachable.

Verification:

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

© 2015 Cisco and/or its affiliates. All rights reserved.

148

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 **IGP** carries the subnets that are assigned to each of the loopback interfaces.

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

Troubleshooting IBGP Session Startup Issues (Cont.)

Symptom:

- IBGP session does not start.

Diagnosis:

- Packet filters prevent establishment of BGP sessions.

Verification:

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

© 2015 Cisco and/or its affiliates. All rights reserved.

149

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

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

Troubleshooting IBGP Route Selection Issues

A router can use a [BGP](#) update to reach network destinations only if the next-hop address specified in the BGP update is reachable. A BGP update that refers to a next hop that is currently not reachable according to the routing table will be saved in the BGP table. However, the router cannot install this update in its routing table. If the next-hop address later becomes reachable, the BGP route will become a candidate route that the router could use for packet forwarding to that destination.

Troubleshooting IBGP Route Selection Issues

Symptom:

- An IBGP route is in the BGP table but is never selected as the best route.

Diagnosis:

- The 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.

© 2015 Cisco and/or its affiliates. All rights reserved.150

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

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

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

Troubleshooting IBGP Synchronization Issues

In old [BGP](#) designs, redistribution between BGP and an [IGP](#) was a common practice, and these protocols had to be synchronized to ensure proper packet forwarding. In modern designs, redistribution is no longer used and synchronization has to be turned off. The default value is disabled synchronization.

Troubleshooting IBGP Synchronization Issues

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.

© 2015 Cisco and/or its affiliates. All rights reserved.151

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

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

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- You can use the **show ip bgp neighbors** and **show ip bgp prefix** commands to monitor IBGP operation.
- There are a number of problems that can occur during IBGP session startup. You can use **debug ip tcp transactions** to see BGP sessions coming from unexpected IP addresses.
- 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 be able to install IBGP routes in the routing table or propagate them to other EBGP neighbors. The BGP synchronization is disabled by default.

Lesson 5: Module Summary

Overview

This topic summarizes the key points that were discussed in this module.

Module Summary

- Because all transit autonomous systems are required to carry traffic originating from or destined to locations outside of that AS, a degree of interaction and coordination between BGP and the IGP is necessary. You should take a special care to ensure consistency of routing information throughout the AS.
- Both EBGP and IBGP sessions forward BGP updates, but they do it in slightly different ways.
- An IGP is still needed inside a transit AS. The high-speed convergence offered by an IGP helps in the packet forwarding to external destinations.
- Configuring in a transit AS involves configuring IBGP neighbors, BGP synchronization, and IBGP sessions between loopback interfaces.
- The three common IBGP configuration problems concern session startup, route selection, and synchronization, and there are specific commands and procedures that you can use to solve those problems.

© 2015 Cisco and/or its affiliates. All rights reserved.

153

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *BGP Case Studies*
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/26634-bgp-toc.html>

- Cisco Systems, Inc. *Using the Border Gateway Protocol for Interdomain Routing*
http://docwiki.cisco.com/wiki/Internetworking_Case_Studies
- Cisco Systems, Inc. *Troubleshooting TCP/IP*. "Troubleshooting IP Connectivity and Routing Problems."
<http://www.cisco.com/en/US/docs/internetworking/troubleshooting/guide/tr1907.html#wp1021148>
- Cisco Systems, Inc. *Configuring BGP*.
http://www.cisco.com/c/en/us/td/docs/ios/12_2/ip/configuration/guide/fipr_c/1cfbgp.html
- Cisco Systems, Inc. *Cisco IOS IP Routing Configuration Guide, Release 15M&T*
http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bgp/configuration/15-mt/irg-15-mt-book.html

Lesson 6: Module Self-Check

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

1. Why is IBGP a mandatory component of a transit AS? (Source: "Working with a Transit AS")
 - A. It is the only feasible way to ensure that all routers in the AS have consistent external routing information.
 - B. It eliminates the scalability issues of running an IGP within the transit AS.
 - C. Running IBGP on all routers is the only way to satisfy the filtering requirements of the transit AS.
 - D. An IGP is not capable of handling the potential routing loops in the transit AS.

2. IBGP neighbors must be directly connected. True or false? (Source: "Working with a Transit AS")
 - A. true
 - B. false

3. Why is redistributing BGP routes into an IGP for use in a transit backbone NOT recommended? (Source: "Working with a Transit AS")
 - A. Redistribution removes all BGP attributes that are needed to ensure optimal routing within the transit AS.
 - B. An IGP cannot enforce complex administrative policies and route selection rules.
 - C. IGP's cannot scale to the demands that are presented by the number of routes on the Internet.
 - D. IGP's are not stable when faced with a flapping network.

4. How is EBGp used in a transit AS? (Source: "Working with a Transit AS")
 - A. as a means of transporting customer routes across the transit backbone
 - B. to exchange routes between different autonomous systems and the transit AS
 - C. to enhance scalability by transporting IGP routes for the transit AS
 - D. as a means of injecting local routes into the transit backbone

5. The received EBGp information has to be translated and then forwarded to other BGP-speaking routers within the transit AS. True or false? (Source: "Working with a Transit AS")
- A. true
 - B. false
6. All routers within a transit AS must have all external routes, received from any neighboring AS. True or false? (Source: "Working with a Transit AS")
- A. true
 - B. false
7. What are the two key functions of a transit AS? (Choose two.) (Source: "Working with a Transit AS")
- A. to filter out routes that do not belong to customers of the service provider
 - B. to provide Internet connectivity to customers of the service provider
 - C. to propagate routes between remote autonomous systems
 - D. to route packets between remote networks
8. How are BGP routes sent across the transit backbone? (Source: "Working with a Transit AS")
- A. by redistributing BGP into an IGP and then back into BGP
 - B. by using 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
9. Which two statements are true regarding the AS-path attribute as it relates to IBGP? (Choose two.) (Source: "Interacting with IBGP and EBGp in a Transit AS")
- A. Each router in the AS appends its AS number to the AS path on outgoing BGP updates.
 - B. The AS path inside an AS will be empty for routes originating inside a neighboring AS.
 - C. The AS-path attribute is not used to detect routing loops inside an AS.
 - D. The AS-path attribute is not modified within the AS.
10. What is the primary function of the IBGP multipath load-sharing feature? (Source: "Interacting with IBGP and EBGp in a Transit AS")
- A. to choose one IBGP path as the best path to a destination
 - B. to choose multiple IBGP paths as the best paths to a destination
 - C. to designate one path as the best path and advertise this best path to its neighbors
 - D. to enable a router to handle all the traffic destined for a particular site
11. Which of the following statements about a BGP split horizon is accurate? (Source: "Interacting with IBGP and EBGp in a Transit AS")
- A. Because of the BGP split horizon, any router can relay IBGP information within the AS.
 - B. BGP split horizon is useful only within autonomous systems.
 - C. Routing information loops within the AS are prevented by an IBGP split horizon.
 - D. With a BGP split horizon, routing information that is received through an IBGP session is forwarded to another IBGP neighbor.

12. Why is it recommended that loopback interfaces be used to form IBGP neighbor sessions? (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. reduces router memory resource requirements
 - B. reduces router CPU resource requirements
 - C. ensures IBGP session stability
 - D. is more secure than using the physical interface
13. How is the BGP next-hop attribute processed over an IBGP connection? (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. The next-hop address is set to the address of the receiving router.
 - B. The next-hop address is not modified over the IBGP session.
 - C. The next-hop address is set to the IP address of the nearest EBGP peer.
 - D. The next-hop attribute is set to the IP address of the nearest EBGP peer; if no external AS connection has been configured, the next hop is set to the default gateway that is configured on the router.
14. Which two statements are true of the full-mesh requirement in IBGP? (Choose two.) (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. The IBGP mesh must be a logical full mesh.
 - B. A physical full mesh must be maintained within the IBGP AS.
 - C. Because of BGP split horizon, no router can relay IBGP information within the AS.
 - D. All routers within the AS must be directly connected to ensure correct delivery of BGP routing information.
15. Which three statements regarding the next-hop-self configuration in BGP are true? (Choose three.) (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. Changing the next-hop attribute might cause suboptimal routing.
 - B. The configuration changes how the next-hop attribute is processed at edge routers.
 - C. The configuration announces the local IP address as the BGP next hop in outgoing updates that are sent to the specified neighbor.
 - D. The configuration removes the requirement for the IGP to carry reachability information for intra-AS destinations.
16. Why must all EBGP peers be reachable by all BGP-speaking routers within the transit AS? (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. BGP-speaking routers in a transit AS use the next-hop-self attribute to find their EBGP neighbors.
 - B. EBGP peers in a transit AS use the length of the AS path to decide which BGP route to install in the routing table.
 - C. When BGP routes are used in the routing table, the router checks how it would have reached the next-hop address, and it installs the BGP route with the same forwarding indication as for the route that is used to reach the next-hop IP address.
 - D. All BGP peers do not need to speak to each other within a transit AS.

17. Which command is used to configure the router as the next hop for a BGP-speaking neighbor or peer group? (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. **neighbor description**
 - B. **neighbor remote as**
 - C. **maximum-paths ibgp**
 - D. **neighbor next-hop-sf**
18. What are three differences between IBGP and EBGP sessions? (Choose three.) (Source: "Interacting with IBGP and EBGP in a Transit AS")
- A. Route selection rules slightly prefer IBGP routes.
 - B. Routes that are learned from IBGP peers are not advertised to other IBGP peers.
 - C. EBGP peers are directly connected, and IBGP peers are usually distant.
 - D. By default, no BGP attributes are changed in IBGP updates.
19. When you are configuring the BGP neighbor session, what differentiates an EBGP neighbor from an IBGP neighbor? (Source: "Interacting with IBGP and EBGP in Transit AS")
- A. The keyword **internal** is at the end of the **neighbor** command.
 - B. IBGP neighbors will have the same AS number specified.
 - C. A description for the internal or external neighbor must be attached with the **neighbor description** command.
 - D. Directly connected neighbors will automatically form an EBGP session.
20. What are two scalability tools that you can use to overcome the full-mesh requirement for IBGP sessions? (Choose two.) (Source: "Interacting with IBGP and EBGP in Transit AS")
- A. confederations
 - B. floating static routes
 - C. route reflectors
 - D. disabling BGP synchronization
21. What are two negative ramifications of the full-mesh requirement that is imposed by IBGP? (Choose two.) (Source: "Interacting with IBGP and EBGP in Transit AS")
- A. administrative difficulty of applying an AS-wide routing policy
 - B. requirement to use next-hop-self for proper routing to external destinations
 - C. large number of TCP sessions
 - D. unnecessary duplication of routing traffic
22. It is not recommended to redistribute the information from external routers into the IGP which is used within the transit AS. True or false? (Source: "Forwarding Packets in a Transit AS")
- A. true
 - B. false

23. What are two reasons why you must run IBGP on all routers within a transit backbone? (Choose two.) (Source: "Forwarding Packets in a Transit AS")
- A. so routers can properly forward packets toward all external destinations
 - B. to ensure that a full mesh exists among all routers in the AS
 - C. to allow routers to properly process the BGP next-hop attribute
 - D. because IGP cannot scale large enough to handle redistribution of BGP routes
24. Routes learned via BGP do not have an outgoing interface associated with them in the routing table, so recursive lookup is performed to forward IP packets toward external destinations. True or false? (Source: "Forwarding Packets in a Transit AS")
- A. true
 - B. false
25. If a transit backbone has IBGP running on all routers, what are two reasons why it is still necessary to use an IGP? (Choose two.) (Source: "Forwarding Packets in a Transit AS")
- A. to provide routing information that is needed to establish the IBGP sessions
 - B. to resolve next-hop references that are used in recursive routing
 - C. so that BGP routes can be properly transported through the AS
 - D. to provide user workstations with a network default gateway
26. Map routing protocol to its function. (Source: "Forwarding Packets in a Transit AS")

IBGP	exchanging information about how to reach internal destinations
IGP	exchanging information about how to reach external destinations

27. What is the AD of the following protocols? (Source: "Forwarding Packets in a Transit AS")

IS-IS	200
RIP	20
OSPF	110
EBGP	115
IBGP	120

28. What are two reasons why the AD is an important consideration for BGP network design? (Choose two.) (Source: "Forwarding Packets in a Transit AS")
- A. The AD affects how routes are selected for use in the IP routing table.
 - B. The AD controls how routing information is entered into the BGP table.
 - C. If a route is advertised by both an IGP and through EBGp, the router will prefer the external route.
 - D. The AD is not a large concern to BGP design, because the router will always choose the route that is advertised by the protocol that is best suited to reach the destination.
29. If a router learns the same information via EBGp and IGP, which one will it use? (Source: "Forwarding Packets in a Transit AS")
- A. the one from IGP
 - B. the one from EBGp
30. With regard to recursive route lookups, what are two ways in which CEF is different from traditional Cisco IOS switching mechanisms such as route caching? (Choose two.) (Source: "Forwarding Packets in a Transit AS")
- A. Traditional Cisco IOS switching mechanisms wait for the first packet to arrive before recursive lookup can take place.
 - B. New entries in the IP routing table will trigger a recursive lookup in traditional Cisco IOS switching mechanisms.
 - C. CEF prebuilds a complete IP forwarding table based on the IP routing table.
 - D. CEF will build a FIB directly from the entries in the BGP table before any BGP packets arrive at the router.
31. Which administrative distance of BGP route cannot be changed? (Source: "Forwarding Packets in a Transit AS")
- A. external distance
 - B. internal distance
 - C. local distance
 - D. remote distance
32. Why is it important to disable BGP synchronization in a transit backbone? (Source: "Monitoring and Troubleshooting IBGP in Transit AS")
- A. IGP's can support the routing requirements of full Internet routing, and hence synchronization is no longer necessary.
 - B. Because BGP redistribution into an IGP is no longer practical, enabling the synchronization feature is no longer applicable.
 - C. Synchronization requires all BGP transit routes to be explicitly mapped to an exit point, creating too much administrative overhead.
 - D. Synchronization requires BGP attributes to be properly mapped to IGP metrics for BGP routing across the transit backbone to function properly, creating too much overhead.

33. Which two steps are required when you use a loopback interface for IBGP peering sessions? (Choose two.) (Source: "Monitoring and Troubleshooting IBGP in Transit AS")
- A. Ensure that the loopback interfaces are reachable through an IGP.
 - B. Ensure that the two neighbors are directly attached.
 - C. Verify that each router has multiple physically redundant paths.
 - D. Configure a neighbor statement with the **update-source** command.
34. Which Cisco IOS **show** command indicates that a BGP route is an IBGP route? (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. **show ip route**
 - B. **show ip route bgp**
 - C. **show ip bgp**
 - D. **show ip bgp internal**
35. Which Cisco IOS **show** command displays the information about the TCP and BGP connection to neighbors? (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. **show ip bgp prefix**
 - B. **show ip bgp neighbors**
 - C. **show ip bgp**
 - D. **show ip bgp internal**
36. BGP runs on the well-known TCP port 169. True or False? (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. true
 - B. false
37. Which three of the following are the most common BGP implementation problems? (Choose three.) (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. IBGP sessions do not reach the Established state.
 - B. TCP window size is set incorrectly.
 - C. Routing information that is received via IBGP is never selected.
 - D. The best BGP route is never installed in the routing table.
38. What are three common situations that prevent IBGP sessions from starting? (Choose three.) (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. The IBGP session has been configured to peer to a loopback interface, but **update-source** has not been configured on the neighbor.
 - B. An access-list filter is blocking access to TCP port 179.
 - C. The IBGP session has been configured to peer to a loopback interface, but the loopback interface has not been administratively enabled with the **no shutdown** command.
 - D. The IBGP session has been configured to peer to a loopback interface, but the interfaces are not reachable via the IGP.

39. Which common issue could prevent IBGP best routes from being inserted into the IP routing table?
(Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. failure to disable BGP synchronization
 - B. failure to disable BGP split horizon
 - C. lack of a route to the BGP next hop for the IGP
 - D. failure to inject a default route into the IGP
40. Which two of the following statements about solving IBGP synchronization problems are accurate?
(Choose two.) (Source: "Monitoring and Troubleshooting IBGP in a Transit AS")
- A. Routers with BGP synchronization enabled will be able to install IBGP routes in the routing table and propagate them to other EBGP neighbors.
 - B. Routers with BGP synchronization enabled will not be able to install IBGP routes in the routing table or propagate them to other EBGP neighbors.
 - C. The default value is to have synchronization enabled.
 - D. The default value is to have synchronization disabled.

Module Self-Check Answers

Answer Key

1. A
2. B
3. C
4. B
5. B
6. A
7. C, D
8. B
9. C, D
10. B
11. C
12. C
13. B
14. A, C
15. A, B, C
16. C
17. D
18. B, C, D
19. B
20. A, C
21. C, D
22. A
23. A, D

- 24. A
- 25. A, B
- 26. IGP — exchanging information about how to reach internal destinations
 - IBGP — exchanging information about how to reach external destinations
- 27. IBGP — 200
 - EBGP — 20
 - OSPF — 110
 - IS-IS — 115
 - RIP — 120
- 28. A, C
- 29. B
- 30. A, C
- 31. D
- 32. B
- 33. A, D
- 34. C
- 35. B
- 36. B
- 37. A, C, D
- 38. A, B, D
- 39. A
- 40. B, C

Module 3: Route Selection Using Policy Controls

Introduction

[BGP](#) enables traffic in Internet backbones to determine an optimal path to its destination across networks comprising more than one [AS](#). Routes that are learned via BGP have properties that are associated with them that aid BGP in determining the best route to a particular destination. There are many instances in which the default BGP route selection does not match administrative or business policies. Likewise, redundant network designs often require enterprises to run BGP when they are connected to more than one ISP. In these situations, full BGP routing tables and default BGP route selection are not desirable.

This module provides information on how to connect Internet customers to multiple service providers. It introduces the need for filtering BGP updates and changing BGP route selection policies. In addition, this module describes different Cisco IOS mechanisms (AS path filters, prefix-lists, route-maps) available for BGP route filtering.

Upon completing this module, you will be able to:

- Influencing BGP route selection in a customer scenario where connections to multiple ISPs must be supported
- Configure BGP to influence route selection using AS path filters
- Define how to successfully configure BGP to influence route selection using prefix-list filters in a customer scenario in which connections to multiple ISPs must be supported
- Defines how to use outbound route filtering to minimize the impact of BGP routing updates on router resources in an operational BGP network
- Describe route-maps and how you can use them for BGP filtering.
- Define how to configure the soft reconfiguration feature to minimize the impact of expediting BGP policy updates in a typical BGP network

Lesson 1: Using Multihomed BGP Networks

Overview

In some circumstances, it is important to have multiple paths to an [ISP](#). There are business and technical reasons to configure a [BGP](#) network in a multihomed configuration. Mission-critical applications often call for redundant network designs. When access to applications is provided over the Internet, enterprises typically use multihomed BGP networks to achieve their goals of high availability.

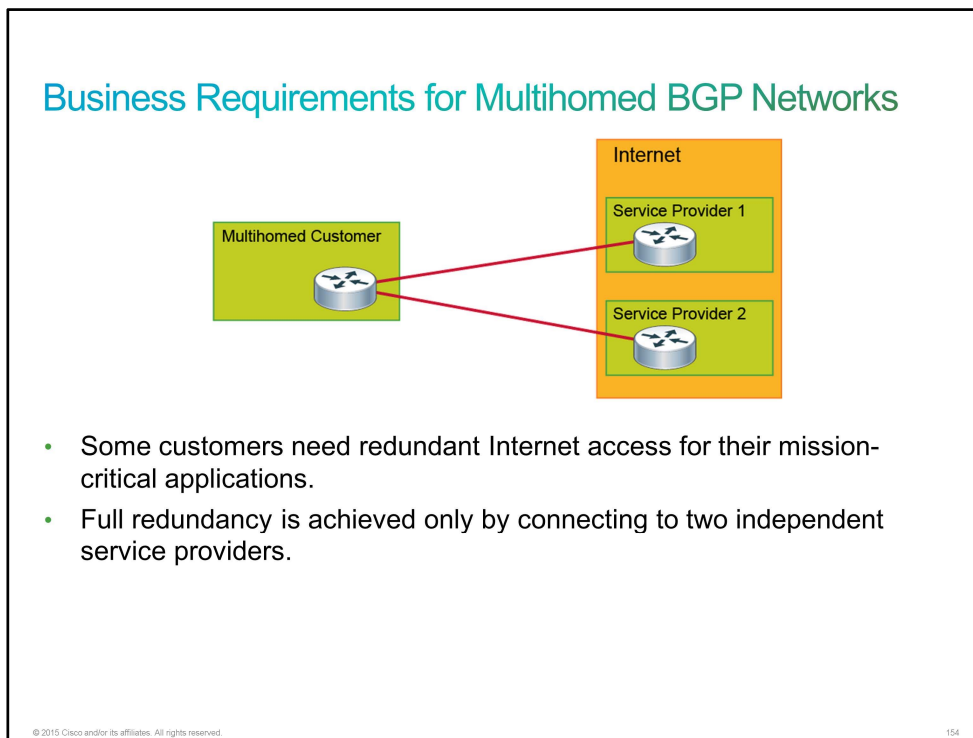
Full BGP routing tables and default BGP route selection might ordinarily be considered desirable characteristics of a network. However, the overhead of full BGP routing tables is not warranted in these situations. Furthermore, the default route selection in BGP often does not match the business and technical requirements for multihomed enterprise networks that use BGP.

Upon completing this lesson, you will be able to:

- List the business requirements for multihomed BGP networks in service provider environments
- Describe the technical requirements for multihomed BGP networks in service provider environments
- Explain the need for BGP policies that influence route selection in a multihomed BGP network
- Describe typical routing policies for multihomed BGP customers
- Explain the need to influence BGP route selection in a service provider environment
- Explain the need for BGP filters in a multihomed customer
- Explain the need for BGP filters in a service provider
- Explain the return traffic issue

Business Requirements for Multihomed BGP Networks

Companies who use web servers (or similar servers) and offer mission-critical business services over the Internet often like to have their networks redundantly connected to the Internet. When the companies calculate the expected loss of business because of an unexpected disconnection from the Internet, they may conclude that having two connections to the Internet is profitable. In such cases, the company may consider being a customer to two different [ISPs](#) or having two separate connections to one ISP.

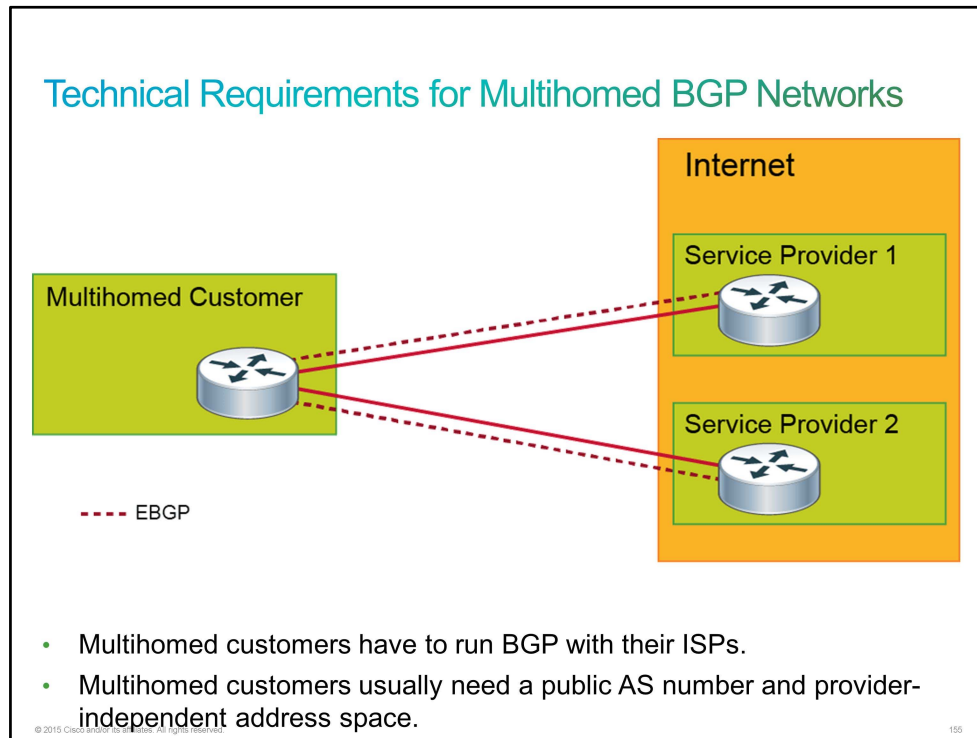


With two connections to one single ISP, [BGP](#) is usually not required. This solution provides a backup for a link failure and router failure. However, this solution does not provide a backup for problems in the network of the ISP or the connection of the ISP to the rest of the Internet.

Full redundancy is achieved only by connecting to two independent ISPs. If one of the ISP networks loses its connection to the rest of the Internet, the customer will still reach the rest of the Internet via the other service provider. At the same time, the customer will still reach those users who are directly connected to the failing ISP via its direct connection.

Technical Requirements for Multihomed BGP Networks

The multihomed customer network must exchange [BGP](#) information with both [ISP](#) networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.



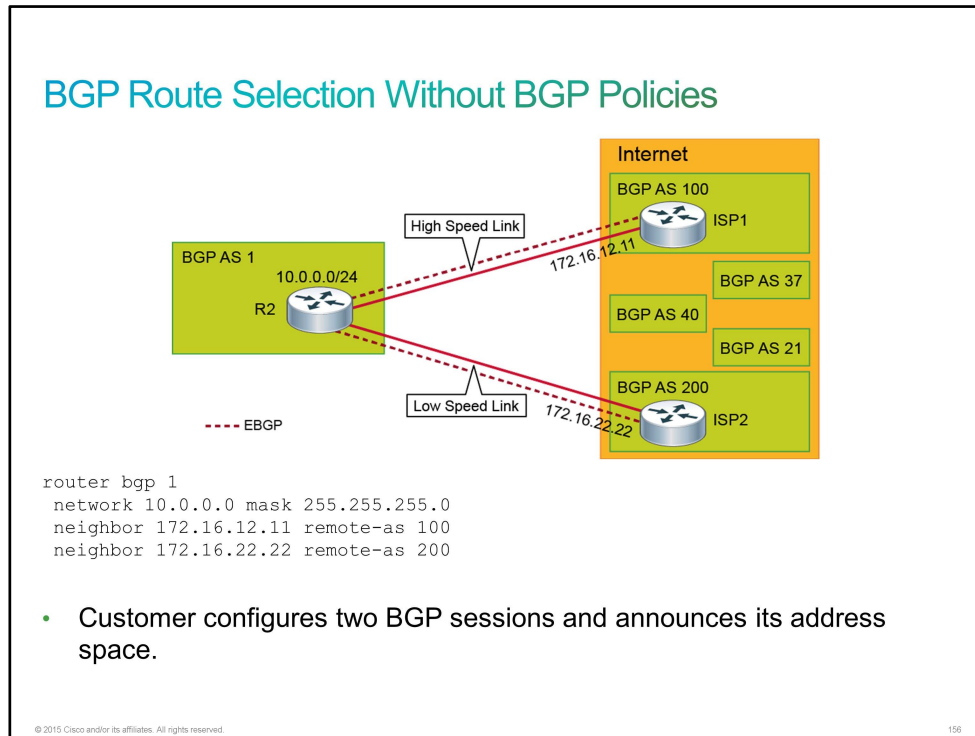
The customer must, in most cases, have its own public [AS](#) number and announce its own IP networks to both ISPs. The ISPs will propagate customer announcements to the rest of the Internet, and the customer will be seen as reachable via both ISP networks. The customer network also receives full Internet routing from both ISPs. This capability gives the customer network the opportunity to choose the best connection at that time to reach any destination on the Internet.

Most customers are not multihomed. They do not exchange BGP with their ISP. Instead, they use default routing to the ISP, and the ISP does static routing to the customer. ISPs use this fact to optimize the number of prefixes that they announce into the Internet. IP network numbers are usually assigned to customers from a range of IP networks that are delegated to the ISP. This situation means, in the ideal case, that all customers that are connected to one single ISP can have their IP networks summarized in a few BGP updates.

In the multihomed scenario, however, the ISP cannot benefit from IP network number assignment from the delegated range. The customer is connected to two different ISPs, and it is not obvious from which provider-assigned address space it should get the [IP addresses](#). The best solution is to do the assignment from a range completely independent of the providers, a provider-independent address space.

BGP Route Selection Without BGP Policies

The simple approach, which is illustrated in the figure, may be the source of many problems. By simply starting [BGP](#) sessions with both [ISPs](#), and announcing the customer networks to both ISPs, the customer could experience difficulties as a result of the default behaviors of BGP. The following example illustrates problems that may occur in this environment.



BGP Route Selection Without BGP Policies Example

BGP Route Selection Without BGP Policies Example

- The BGP routes are selected based on AS-path length.
- The default BGP route selection does not always result in optimum routing.

```
R2# show ip bgp
BGP table version is 30, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

      Network        Next Hop           Metric LocPrf Weight Path
*>  10.0.0.0/24      0.0.0.0             0           32768 i
*>  10.0.1.0/28      172.16.12.11        0           0 100 i
*>  10.0.2.0/28      172.16.22.22        0           0 200 i
*>  10.0.21.0/24     172.16.22.22        0           0 200 21 i
*                   172.16.12.11        0           0 100 37 40 21 i
*  10.0.37.0/24     172.16.22.22        0           0 200 21 40 37 i
*>                   172.16.12.11        0           0 100 37 i
*>  10.0.40.0/24     172.16.22.22        0           0 200 21 40 i
*                   172.16.12.11        0           0 100 37 40 i
```

© 2015 Cisco and/or its affiliates. All rights reserved.

157

The multihomed customer is connected to two ISPs: AS 100 and AS 200. The two ISPs are connected to the upstream ISPs: AS 21, AS 40, or AS 37.

The customer receives all routes from both service providers, giving redundancy. The default BGP route selection prefers the shortest **AS** path. If the AS path lengths are equal, BGP prefers the most stable route, or the route that is received from the peer with the lower router-ID.

In many cases, however, this route is not the most optimal way to reach all destinations. For example, the bandwidth that is available to reach the ISPs has not been taken into consideration. To change the route selection behavior, some BGP parameters must be configured to support more complex routing policies.

Multihomed Customer Routing Policies

Multihomed Customer Routing Policies

Multihomed customers could require a number of routing policies, for example:

- One provider is primary; the other is backup.
- Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary provider.
- All traffic to a particular part of the world goes through one ISP.
- Traffic toward a specific destination goes through only one of the ISPs.

© 2015 Cisco and/or its affiliates. All rights reserved.

158

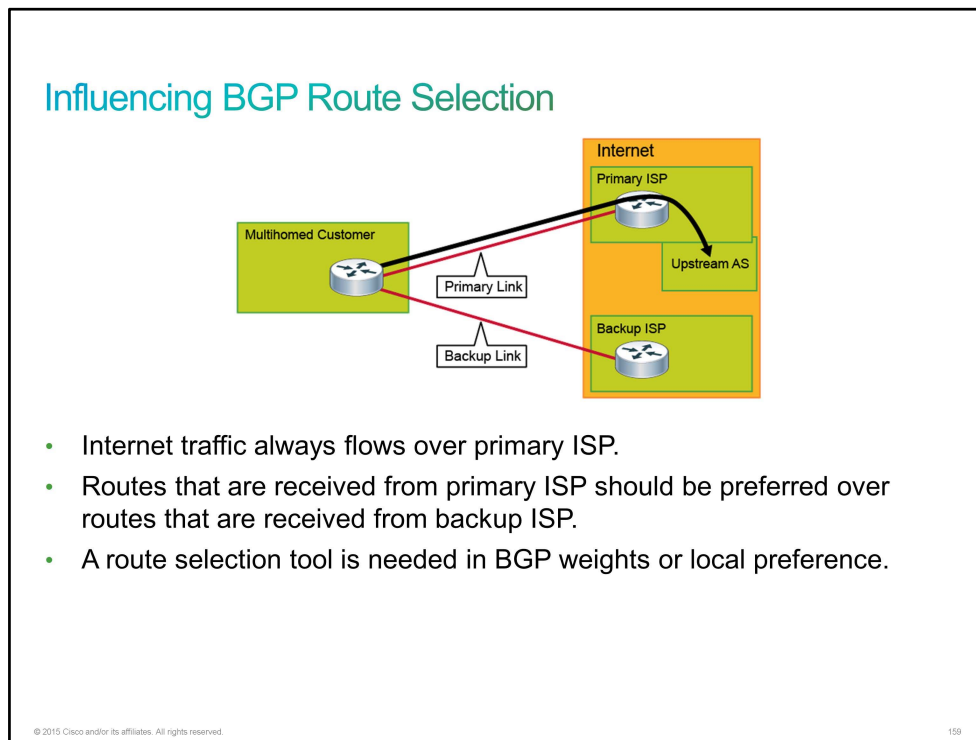
Depending on the circumstances, here are the different policies that a multihomed customer might require:

- One of the two [ISPs](#) can be considered the primary connection. This distinction can be the result of available bandwidth or commercial agreements. However, although one of the ISPs is considered the primary connection, some users may have direct connections to the secondary ISP. Therefore, going via the primary ISP to reach users that are connected to the secondary ISP may be suboptimal.
- Destinations in one part of the world may be reached more optimally via one of the ISPs, rather than via the other. This situation occurs because the two ISPs may have different infrastructures and peering agreements with other ISPs.

It is virtually impossible to establish a routing policy that gives optimal routing to every destination on the Internet. Optimization can be done only with the most common destinations in mind. This situation can result in specific rules on how to reach specific destination networks or the [AS](#).

Influencing BGP Route Selection

When one of the two [ISPs](#) is designated as a primary ISP and the other as a backup, [BGP](#) attributes must be configured to influence BGP route selection.

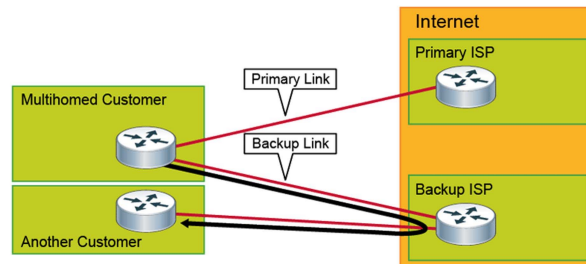


If both ISP connections terminate in one single customer router, all routes that are received from the primary ISP can be assigned a BGP weight. A higher weight indicates a more preferred path.

However, the weight value is local to one router. The weight value is not shared between routers. If one ISP connection terminates in one of the customer routers and the other ISP connection terminates in another, the two customer routers must agree on which link to use. Using local preference instead of weight can cause this situation. All routes that are received from the primary ISP over the primary link are assigned a local preference value, which is higher than the default value of 100. The customer router that receives the routes from the primary ISP completes the assignment and communicates the information to the other routers within the [AS](#) of the customer.

When using either weight or local preference, the customer AS reaches all destinations on the Internet via the primary link as long as it is available. It reaches those destinations within the AS of the secondary ISP. If there is a link failure or failures within the network of the primary ISP, some of the routes, or all of the routes, will no longer be received over the primary link. In that case, the AS of the customer no longer sees those destinations as reachable over the primary link. The only remaining choice is the backup link. Therefore, the customer network uses the backup link only to reach destinations that are not reachable over the primary link.

Influencing BGP Route Selection (Cont.)



- Internet traffic flows over primary ISP; traffic to customers of backup ISP goes direct.
- Route selection has to be performed based on AS numbers in the AS path.

© 2015 Cisco and/or its affiliates. All rights reserved.

160

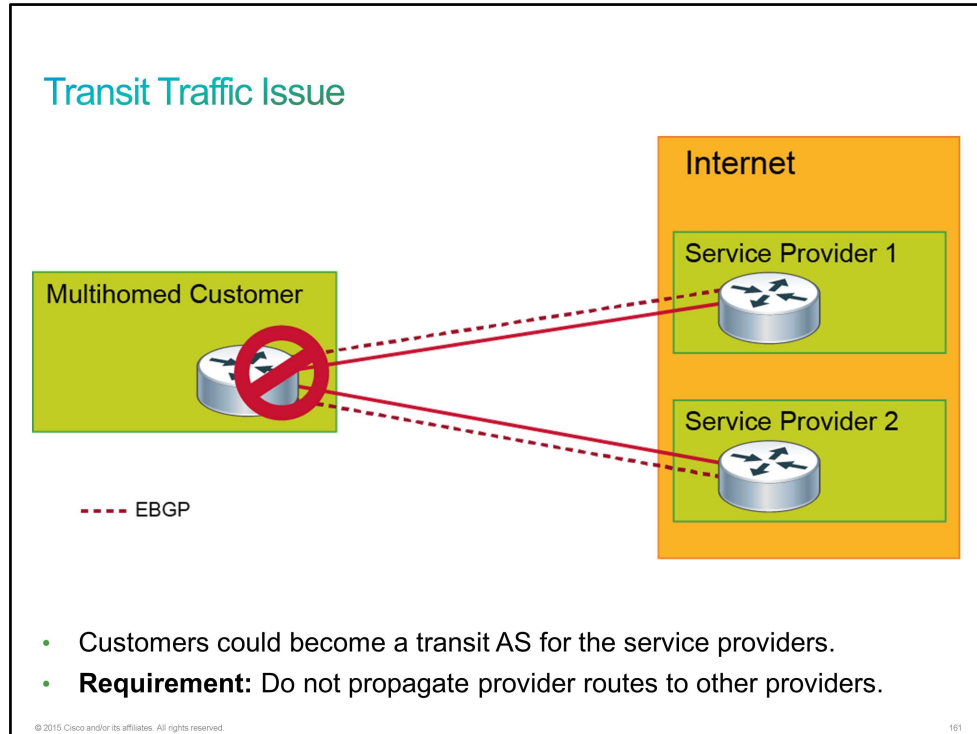
In most cases, it is optimal to reach other customers who are connected to the backup ISP via the backup link, compared with reaching them via the primary link.

The routing policy, where routes are blindly preferred if they are received on the primary link, can easily be modified to use the backup link for destinations in the AS of the backup ISP. On the router, filtering tools can be configured to select routing information that is based on the content in the AS path attribute. Those routes, with an AS path attribute matching specific selection criteria, can be assigned an even higher weight or local preference.

This approach results in a routing policy that gives precedence to reaching destinations within the AS of the primary ISP and within all ASs upstream of the primary ISP over the primary link. Destinations within the AS of the backup ISP receive precedence over the backup link.

Transit Traffic Issue

When [BGP](#) has selected the best path, the router advertises the information to all neighboring autonomous systems, except for the session that the path was received on. This functionality is called "BGP split-horizon," which prevents near-range routing loops. This situation causes the customer [AS](#) to become a transit AS between the two [ISPs](#), and should be avoided.

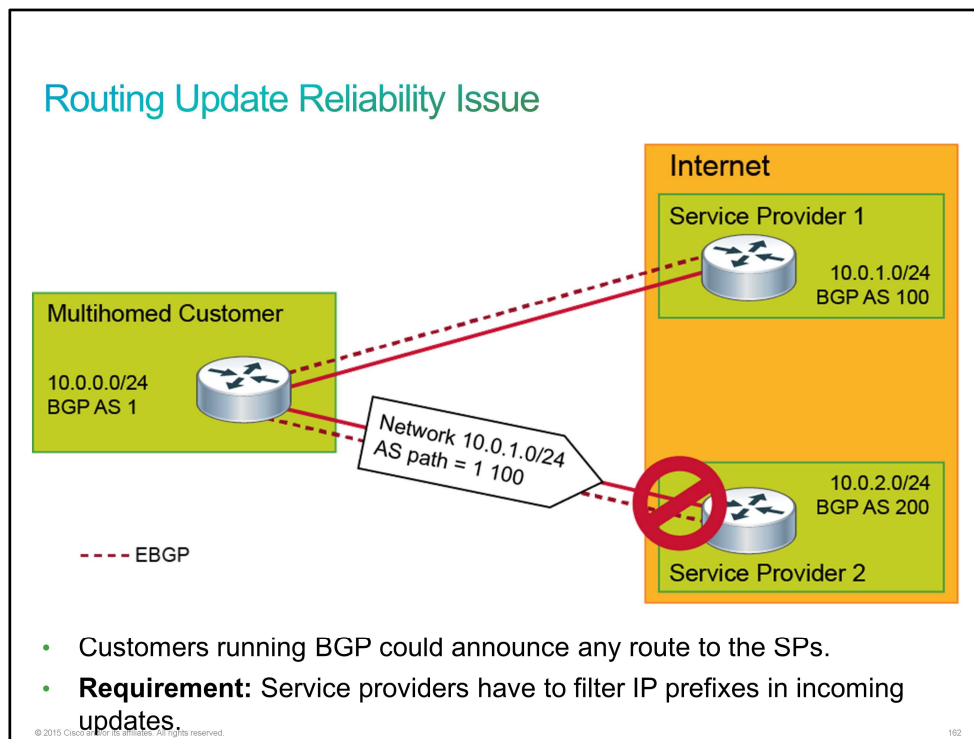


Most customers do not intend to create transit traffic between ISP networks. The access lines to the ISPs are not suited to carry this volume of traffic, and the customer certainly does not want the transit traffic to consume the bandwidth.

The solution to this problem is to filter the outgoing information to both ISPs. Filtering of routing information is performed based on the content of the AS path attribute that is assigned to every BGP route. Only routes having an AS path attribute that indicates that they are sourced by the AS of the customer are allowed to be sent to either of the two ISPs.

Routing Update Reliability Issue

Without some sort of filtering, [BGP](#) routing information that the [AS](#) of the customer creates can potentially be propagated all over the Internet. In this way, the customer can inject erroneous information into the Internet routing tables.



Customers are much less experienced in avoiding these kinds of problems than are service providers. There is much more risk of errors being introduced when a customer is assigned its own AS and uses BGP with the [ISP](#), as compared with the single-homed scenario in which the ISP has sole responsibility to announce BGP routes to the rest of the Internet.

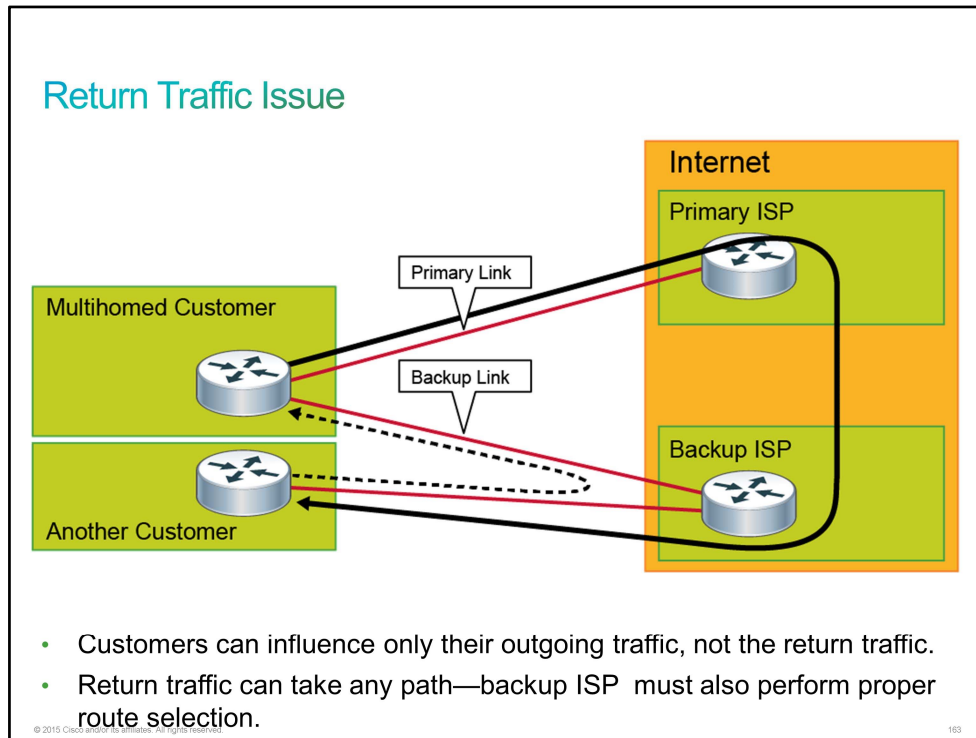
The ISP can stop almost all of the Internet problems that a customer can cause by improperly configuring its BGP. The ISP should also filter all incoming information from the customer and accept only what is supposed to arrive. The ISP should discard anything outside strict limits. In this way, the ISP prevents the propagation of erroneous information to the rest of the Internet.

The ISP can maintain a list of the IP network numbers that the customer is announcing and filter out any other route. If this approach is not possible because of the volume of those lists, the ISP should at least be able to filter out the most obvious erroneous announcements.

Note Private addresses, according to [RFC](#) 1918, should never be announced to the Internet.

Return Traffic Issue

The customer can easily define a policy about how to send outgoing IP packets on the correct link. It is much harder to influence the neighboring [AS](#) about how to direct the IP packets coming into the customer network.



A customer that creates a routing policy in which one of the two [ISPs](#) is always preferred may see the return traffic arriving on the backup link. This situation means that the customer has configured the weight or local preference to make sure that all outgoing traffic is leaving the customer AS over the primary link. However the backup ISP does not have any such configuration. Therefore, return traffic enters the customer AS by using the shortest AS path as its selection criterion.

The best way to solve this problem is for the customer to ask the backup ISP to change its routing policy. The change should cause the backup ISP to prefer reaching the customer AS via the AS of the primary ISP. The backup ISP must implement this change in its own AS.

Note Sometimes the backup ISP administrator might be reluctant to change the configuration for a single customer. In this case, the customer should use another [BGP](#) feature, the AS path prepending tool, to influence the selection of the primary or backup link by lengthening the AS path of routes that are sent to the backup provider.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Some customers need redundant Internet access for their mission-critical applications and address this need by having two separate connections to one ISP or implementing a multihomed configuration (connecting to two different Internet service providers).
- The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.
- An approach to multihoming that is too simple can be a source of problems. Starting BGP sessions and announcing customer networks to multiple ISPs by using the default behavior of BGP may not result in optimal routing.

© 2015 Cisco and/or its affiliates. All rights reserved.

164

Summary (Cont.)

- Depending on the circumstances, a multihomed customer may require different policies. For example, one of the two ISPs being considered the primary connection or reaching destinations in one part of the world more optimally via one of the ISPs. Optimization should be done with the most common destinations in mind, resulting in specific rules on how to reach specific destination networks or the AS.
- In BGP route selection, a routing policy may be created that gives precedence to reaching destinations within the AS of the primary ISP and all upstream ASs over the primary link. All destinations within the AS of the backup ISP can be reached over the backup link.
- When BGP has selected the best path and the information has been propagated to all neighboring autonomous systems, the customer AS may become a transit AS between the two ISPs. The customer must avoid this situation by using BGP filters.

© 2015 Cisco and/or its affiliates. All rights reserved.

165

Lesson 2: Employing AS Path Filters

Overview

In network implementations that require connections to multiple [ISPs](#), you will typically use [AS](#) path filters to influence [BGP](#) route selection. It is important for you as a network administrator to understand the syntax of an AS path regular expression. It is also important to understand how string-matching operators function when they are using AS path regular expressions to match BGP routes.

BGP allows connectivity between multiple ISPs for redundancy and scalability. You can employ AS path filters to remedy the problems that are associated with the various connectivity methods that are used within BGP. This lesson explains the methods that are used to implement BGP AS path filters.

Upon completing this lesson, you will be able to:

- Identify network scenarios in which you must support connections to multiple ISPs and in which AS path filters can be used to influence route selection
- Describe the function of an AS path regular expression
- Explain how string-matching operators function when you are using AS path regular expressions to match BGP routes
- Identify where you can apply an AS path filter when configuring a router to influence route selection
- Identify the Cisco IOS commands that are required to configure AS path filters to influence route selection
- Identify the Cisco IOS commands that are required to monitor the operation of configured AS path filters

AS Path Filtering Scenarios

Several scenarios require filtering and selection of routing information, which is based on the content of the [AS](#) path attribute. Each [BGP](#) route must have an AS path attribute. It is a well-known mandatory attribute and must therefore be present in each BGP update.

AS Path Filtering Scenarios

Several scenarios require BGP route-filtering which is based on the AS path:

- Announce only local routes to the ISP—AS path needs to be empty
- Select routes that are based on a specific AS number in the AS path
- Accept routes for specific AS only from some BGP neighbors

AS path filters use regular expressions.

© 2015 Cisco and/or its affiliates. All rights reserved.166

Using selection criteria that are based on the AS path attribute, a router can identify a set of specific routes from the total set of routes that it receives. Those routes where the AS path contents match the criteria are selected. Routes that do not match the criteria are not selected.

The AS path is a sequence of numbers. Each number indicates an AS. When a route is sourced by a **network** command in a BGP process or redistribution into a BGP process, the AS path attribute is created and left empty. Each time the egress router advertises the route to another AS, it modifies the AS path attribute. Egress router prepends its AS number to the AS path attribute.

While a newly sourced route is still within the AS in which it was created, the AS path is empty. Assume that the AS has a requirement to filter out all but the routes that are local to itself before sending them to a neighboring AS. The AS will permit sending of the routes with the empty AS path and will deny all others.

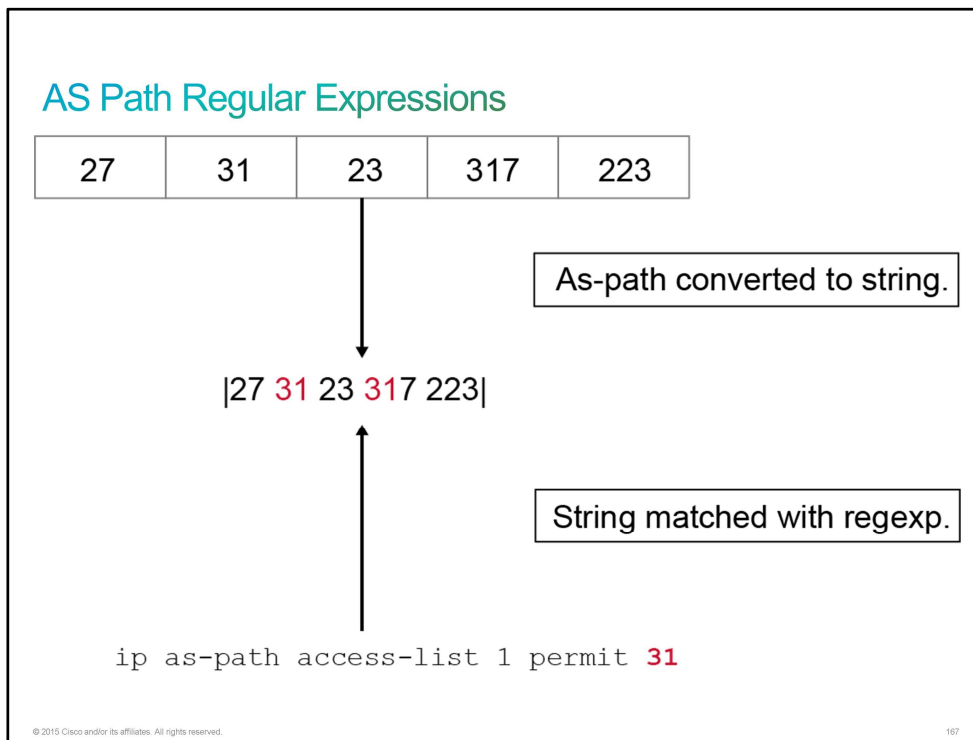
Routers can also filter incoming routes that are based on their AS path attributes. Some destination autonomous systems should not be received from a certain neighbor. Therefore, the routes matching that AS in the AS path can be filtered on the receiving router in case they are accidentally sent.

Selection that is based on the AS path is also a tool that you can use when changing the weight or local preference attributes only for some destination ASs.

When routers filter BGP updates based on the content of the AS path attribute, they use regular expressions. Regular expressions are commonly found in the [UNIX](#) environment and also in some Microsoft Windows-based applications. Regular expressions are a string-matching tool. A regular expression consists of a string of characters. Some of these characters have special meanings, such as functioning as wildcards and operators. Some of these characters simply mean themselves, for example, A to Z, a to z, or 0 to 9. A regular expression is said to match a string if the ordinary characters and the applied meaning of the special operator characters can be translated into the matched string. When a regular expression matches, the selection test is said to be true. If it does not match, the test is false.

AS Path Regular Expressions

The [AS](#) path attribute that is carried with all [BGP](#) routes in a BGP update is a very compact binary encoding of a sequence of integer numbers. It is not a sequence that can be tested by using a regular expression.



Cisco IOS software internally translates the binary encoding into a character string. Each AS number in the sequence is converted into a string using decimal representation. The space character separates each AS number in the AS path attribute. The router applies the regular expression test to this internally created character string.

Characters in a regular expression that are not assigned a specific operation match themselves. The regular expression "31" matches all occurrences of the character "3" followed by the character "1" in the AS path. In this example, "31" matches at two occurrences. One occurrence is sufficient to make the test true. No occurrence means that the test failed.

String Matching

How does string matching functions when you are using AS path regular expressions to match BGP routes?

Regular Expressions

Regular Expressions

A string of characters in a regular expression matches any equivalent substring in the AS path.

How many times does "31" match?

```
|213 317 2316 31|
```

Answer: Three times

```
|213 317 2316 31|
```

© 2015 Cisco and/or its affiliates. All rights reserved.168

The regular expression "31" will match any occurrence of "3" followed by "1" regardless of the characters immediately preceding the "3" and immediately following the "1." Therefore, "31" will match an occurrence of "3" and "1" in the middle of an [AS](#) number.

The regular expression "31" matches the AS path string "213 317 2316 31" three times, because "31" matches a part of "317," "2316," and "31."

Alternatives

Alternatives

Expression *expr1|expr2* matches the string if either subexpression matches the string.

How many times does "21|31" match?

```
|213 317 2316 31|
```

Answer: Four times

```
|213 317 2316 31|
```

© 2015 Cisco and/or its affiliates. All rights reserved.169

The character "|" (vertical bar) has a special meaning. It is an operator that means "or." The regular expression "21|31" matches the sequence of "2" followed by "1" or the sequence of "3" followed by "1." Therefore, this sample regular expression will match a two-character sequence: the "21" or the "31."

The regular expression "21|31" matches the AS path string "213 317 2316 31" four times, because "21" matches a part of "213" and "31" matches a part of both "317" and "2316" and also "31."

Ranges and Wildcard Characters

Ranges and Wildcard Characters

A range of characters matches any single character in the range.

Example: "[1234]" or "[1-4]"

Dot "." matches any single character

How many times does "[1-3].[34]" match?

```
| 213 317 2316 31 |
```

Answer: Twice

```
| 213 317 2316 31 |
```

```
| 213 317 2316 31 |
```

© 2015 Cisco and/or its affiliates. All rights reserved.170

The pair of brackets "[" and "]" has a special meaning. Brackets surround a set of characters of which any one matches. The set of characters is either expressed as the list of the matches (for example, "[1234]") or the sequence with the starting character, a hyphen, and the ending character (for example, "[1-4]"). Both examples match one single character, which must be any one in the set of the four characters "1," "2," "3," and "4."

The character "." (a dot) matches any single character. Small regular expressions can be combined into a larger expression. Such a combination is matching if all of the parts match one after the other. The sample regular expression "[1-3].[34]" matches a sequence of three characters, of which the first must be either "1," "2," or "3." The second character can be any character, and the third must be either "3" or "4."

Note	The space character delimiting two AS numbers is just a character. The dot (".") for example, matches this character.
-------------	---

The regular expression "[1-3].[34]" matches the AS path string "213 317 2316 31" twice. Initially, it matches "213." The leading "[1-3]" matches the leading "2." The dot, which matches any character, matches the "1," and "[34]" matches the trailing "3." The regular expression also matches in "213 317 2316 31." The match is a little harder to see, because the dot (".") matches the space character between "213" and "317."

Matching Delimiters

Matching Delimiters

^	Matches beginning of string
\$	Matches end of string
 	Matches any delimiter (beginning, end, white space, tab, comma)

|213 317 218 31 731|

Does "^21" match?

Answer: Yes

|**2**13 317 218 31 731|

© 2015 Cisco and/or its affiliates. All rights reserved.

171

Matching Delimiters (Cont.)

|213 317 218 31 731|

Does "31\$" match?

Answer: Yes

|213 317 218 31 **731**|

Does "_31_" match?

Answer: Yes

|213 317 218 **31** 731|

© 2015 Cisco and/or its affiliates. All rights reserved.

172

A character string must have a start and an end. The character with the special meaning "^" matches the beginning of a string. Because all strings have a beginning, the "^" character matches all strings. However, the "^" character is used to position the part of the regular expression that follows. The character following the "^" character must be the first character of the string; otherwise, that character would not match the beginning of the string.

The special character "\$" is used analogously, but it means the end of the string. The character preceding the "\$" must be the last character in the string; otherwise, the "\$" does not match the end of the string.

The underscore ("_") matches any delimiter. The space character between two AS numbers is an example of a delimiter. The beginning of the string and the end of the string are also considered delimiters. Other delimiters are the tab and the comma (","). The underscore ("_") is used to ensure that the desired AS number is found in an AS path string but not as part of some other AS number. For example, the regular expression "31" will match the AS number string "317," but the regular expression "_31_" will not. Both "31" and "_31_" will match the AS number string "31."

The regular expression "^21" can match the AS path string "213 317 218 31 731" only one time because there is only one beginning of the string. The regular expression "^21" matches only if the string starts with the sequence "21," which it does.

The regular expression "31\$" can match the AS path string "213 317 218 31 731" only one time because there is only one end of the string. The regular expression "31\$" matches only if the string ends with the sequence "31," which it does.

The regular expression "_31_" can, in theory, match an AS path string several times. However, in this case, when matched against the string "213 317 218 31 731," the regular expression "_31_" matches only the AS number "31" in the AS path.

Grouping

Grouping

Parentheses can be used to group smaller regular expressions into larger expressions.

How many times does "(213|218)_31" match?

```
|213 317 1218 316 31|
```

Answer: Twice

```
|213 317 1218 316 31|
```

© 2015 Cisco and/or its affiliates. All rights reserved.173

Complicated expressions must sometimes be grouped with parentheses, "(" and ")." This feature can be useful in the following example. You are searching for a sequence of two or more AS numbers of which the first can match any of the specific AS (in this example, "213" or "218"). However the last must be a specific AS ("31" in this example). If the parentheses were not used here, the expression would match either the single AS "213" or the sequence of the two (218 31).

The regular expression "(213|218)_31" matches the AS path string "213 317 1218 316 31" twice. The first match is "**213 317** 1218 316 31." The second match is "213 317 **1218 316** 31."

Special Characters

Special Characters

To use the special character as single-character patterns, remove the special meaning by preceding each character with backslash "\"

How do you match "AS 213" in the beginning of the string?

```
| (213 317) 1218 316 31|
```

Answer:

```
^\ (213_
```

© 2015 Cisco and/or its affiliates. All rights reserved.174

Sometimes the target string that you are trying to match with a regular expression contains some of the characters that also have special meanings in the regular expression. To match these characters in the target string, use the backslash "\" together with the character in the regular expression.

Note	This type of regular expression syntax is used for matching AS path strings inside a BGP confederation. A confederation is used to eliminate the scaling problem of full-mesh IBGP by splitting the AS into smaller regional autonomous systems. The example shows that 213 and 317 were part of a confederation by its use of "(" and ")."
-------------	---

Repeating Operators

Repeating Operators

*	Matches zero or more atoms
?	Matches zero or one atom
+	Matches one or more atoms

An atom is a single character or a grouping.

How do you match AS sequences "23 45" and "23 78 45" in a single regular expression?

Answer:

`_23 (_78) ? _45_`

© 2015 Cisco and/or its affiliates. All rights reserved.

175

The special characters, star (asterisk), "*", question mark, "?," and plus, "+," all apply repetition of the expression that immediately precedes them.

The star (asterisk), "*", means that the expression that immediately precedes it is repeated zero or more times. The expression may not be there, but it may also be there any number of times. The expression "1*" will match a sequence of no characters or a sequence of any number of the character "1."

A question mark means, that the expression that immediately precedes it is repeated zero or one time. The expression may not be there, but it may also be there once. The expression "1?" will match a sequence of no characters or the single character "1."

The plus sign ("+") means that the expression that immediately precedes it is repeated one or more times. The expression must be there at least once. The expression "1+" will match a sequence of one or more of the characters "1."

String Matching Example

String Matching Example	
<code>_100_</code>	Going through AS 100
<code>^100\$</code>	Directly connected to AS 100
<code>_100\$</code>	Originated in AS 100
<code>^100_.</code>	Networks behind AS 100
<code>^[0-9]+\$</code>	AS paths one AS long
<code>^([0-9]+)(_1)*\$</code>	Prepending performed in neighboring originating AS
<code>^\$</code>	Networks that originate in local AS
<code>*</code>	Matches everything

© 2015 Cisco and/or its affiliates. All rights reserved. 176

Regular expressions can be arbitrarily complex. However, most searching is accomplished using regular expressions similar to the regular expressions found in the following examples:

- If you are searching for all routes that have AS 100 in their AS paths, the regular expression to use is `"_100_."`
- If you are searching for all routes that are sourced in your directly connected neighboring AS 100, the regular expression to use is `"^100$."` Use an expression with both a carat ("`^`") and dollar sign ("`$`") present when you are searching for an exact match.
- If you are searching for all routes that are sourced in AS 100, which is not necessarily a directly connected neighboring AS, the regular expression to use is `"_100$."` The dollar sign ("`$`") indicates that the AS path must end with AS 100. Meaning that the route was sourced in AS 100. The underscore ("`_`") is used to make sure that it is AS 100 at the end of the string and not, for example, AS 2100.
- If you are searching for all the routes that are reachable behind AS 100, the regular expression to use is `"^100_."` The carat ("`^`") indicates that the AS path must start with AS 100. The underscore ("`_`") is used to make sure that this number is not matching with, for example, AS 1001. The dot ("`.`") is used to indicate that the AS path does not end with AS 100, and that there must be something following AS 100.
- If you are searching for all routes that are sourced in any AS directly neighboring your AS, the regular expression to use is `"^[0-9]+$."` The `"[0-9]"` part means any digit. The plus sign ("`+`") repetition character means one or more times. Therefore, the combination `"[0-9]+"` means a sequence of one or more of digits. The carat ("`^`") and dollar sign ("`$`") mean the beginning and the end of the string. Therefore, the string may consist only of a sequence of one or more digits.

- If you are searching for all routes that are sourced in any AS directly neighboring your AS, and possibly performing AS path prepending (multiplication of a directly connected AS number), the regular expression to use is `"^([0-9]+)(_\1)*$"`. The expression in the first set of parentheses matches any AS number. The parentheses store the value of the matched AS, and the second part of the regular expression, including a variable, then recalls it. The variable `"\1"` is put into parentheses for the purpose of the multiplier operator `"*"`. I meaning that this part can match any number of successive occurrences of the same AS number that the `"[0-9]+"` expression matched. For example, this regular expression matches AS paths `"99 99 99,"` `"200,"` `"101 101,"` or `"5 5 5 5 5,"` but it does not match the AS path `"101 99."`
- The combination `"^$"` means an empty string and is used when you are searching for all routes that are sourced in the local AS.
- Sometimes a search is made to select a few specific routes and do something special with them, while the rest of the routes will be handled in a different way. To search for all routes, regardless of the content of their AS path attribute, use the regular expression `".*"`. The dot (`"."`) matches any single character. The repetition character, star (asterisk), `"*"`, means that the match should be repeated zero or more times. Thus, the combination `".*"` matches any string.

Commonly Used Characters in Expressions Example

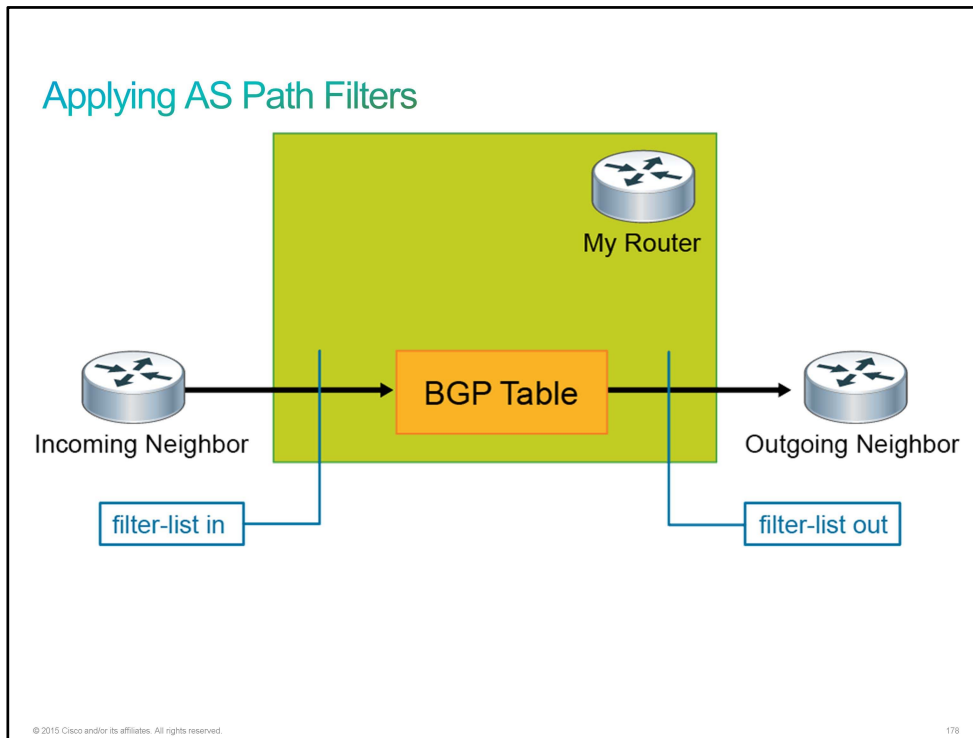
Commonly Used Characters in Expressions Example	
.	Any single character, including a space
*	Zero or more sequences of pattern
+	One or more sequences of pattern
?	Zero or one occurrence of pattern
^	Beginning of string
\$	End of string
_	Match any delimiter (including beginning, end, space, tab, comma)
\	Remove special meaning of character that follows
[]	Match one character in a range
	Logical OR

© 2015 Cisco and/or its affiliates. All rights reserved. 177

The figure lists the most commonly used characters in expressions.

Applying AS Path Filters

AS path filters that are configured on a router select those routes that are allowed.



Routes that are selected behave as described here:

- The selected routes enter the local [BGP](#) table when the selection is applied on the incoming routes from a neighbor; routes that are not selected are silently dropped.
- The selected routes are transmitted to the neighbor when the selection is applied on the outgoing routes to the neighbor. Routes that are not selected are used locally but are never sent to the neighbor.

Configuring BGP AS Path Filters

[AS](#) path access list creates an AS path filter. This access list is applied to a set of routes from which to select a subset. Routes that the access list permits are included in the subset, and those routes that are denied are not included. As in all access lists, the candidate to be permitted or denied membership in the subset is tested against all the lines in the access list. The order of testing is the order in which the list is configured. The first match indicates "permit" or "deny," as specified. If the end of the access list is reached without any explicit match, the candidate is implicitly denied.

Configuring BGP AS Path Filters

```
router(config)# ip as-path access-list number {permit|deny} regex
```

- Configures AS path access-list

```
router(config-router)# neighbor ip-address filter-list as-path-filter {in|out}
```

- Configures inbound or outbound AS path filter for specified BGP neighbor

© 2015 Cisco and/or its affiliates. All rights reserved.

179

The test by the AS path access list is performed by using regular expressions that are applied on the AS path attribute of the route.

The access list can, for example, be applied on the routes received from, or sent to, a specific [BGP](#) neighbor.

ip as-path access-list

To define a BGP AS path access-list, use the **ip as-path access-list** global configuration command:

```
ip as-path access-list access-list-number {permit | deny} as-regular-expression
```

To disable use of the access-list, use the **no** form of this command.

```
no ip as-path access-list access-list-number {permit | deny} as-regular-expression
```

Syntax Description

Parameter	Description
<i>access-list-number</i>	Integer from 1 to 199 that indicates the regular expression access-list number.
permit	Permits access for matching conditions.
deny	Denies access to matching conditions.
<i>as-regular-expression</i>	AS in the access-list using a regular expression (See the "Regular Expressions" appendix in the <i>Cisco IOS Dial Services Command Reference</i> for information about forming regular expressions.)

neighbor filter-list

To set up a BGP filter, use the **neighbor filter-list** router configuration command:

```
neighbor {ip-address | peer-group-name} filter-list access-list-number {in | out}
```

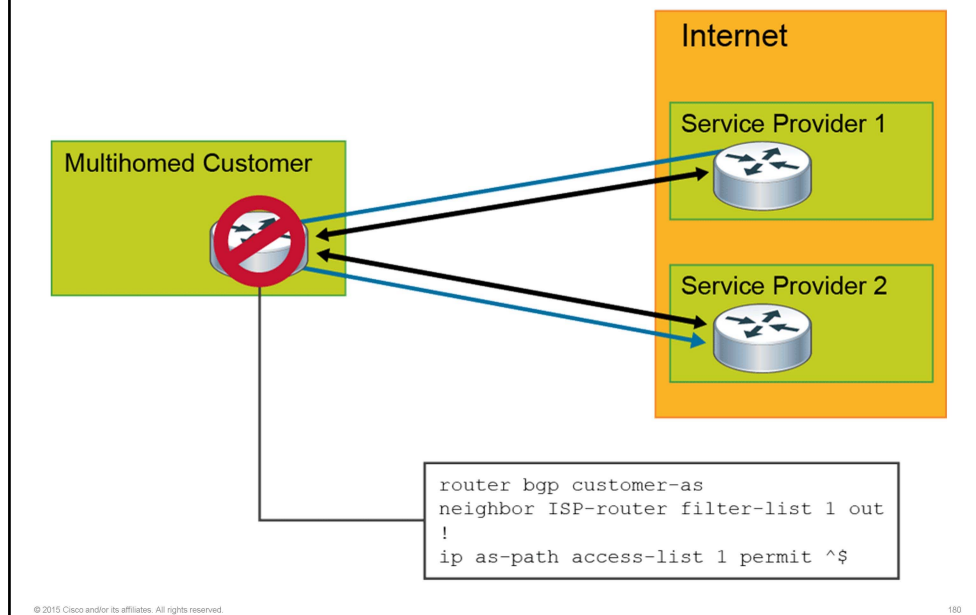
To disable this function, use the **no** form of this command:

```
no neighbor {ip-address | peer-group-name} filter-list access-list-number {in | out}
```

Syntax Description

Parameter	Description
<i>ip-address</i>	IP address of the neighbor.
<i>peer-group-name</i>	Name of a BGP peer group.
<i>access-list-number</i>	Number of an AS path access-list. Define this access-list with the ip as-path access-list command.
in	Access-list to incoming routes.
out	Access-list to outgoing routes.

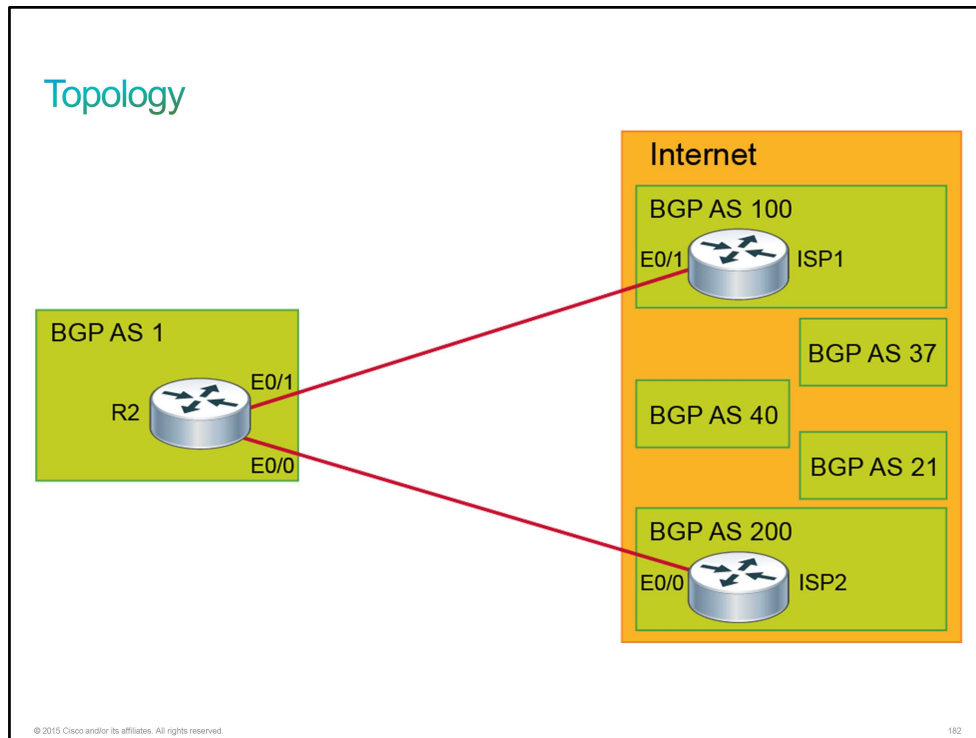
Configuring BGP AS Path Filters (Cont.)



Multihomed customers do not want to act as a transit AS between their service providers. The customer avoids this situation by not transmitting all its routes to its service providers. The service provider sends IP packets to the customer only if the destination addresses match one of the routes that the customer has sent by BGP to the service provider. By making sure that only locally sourced routes are sent, the customer avoids receiving IP packets for destinations outside its own AS.

Within the customer AS, the locally sourced routes have empty AS paths. The regular expression "^\$" matches the empty string. The command **ip as-path access-list 1** permits only the routes that are locally sourced and implicitly denies the rest. By applying this filter-list on outgoing information to all neighbors, the customer will announce local routes only.

Topology



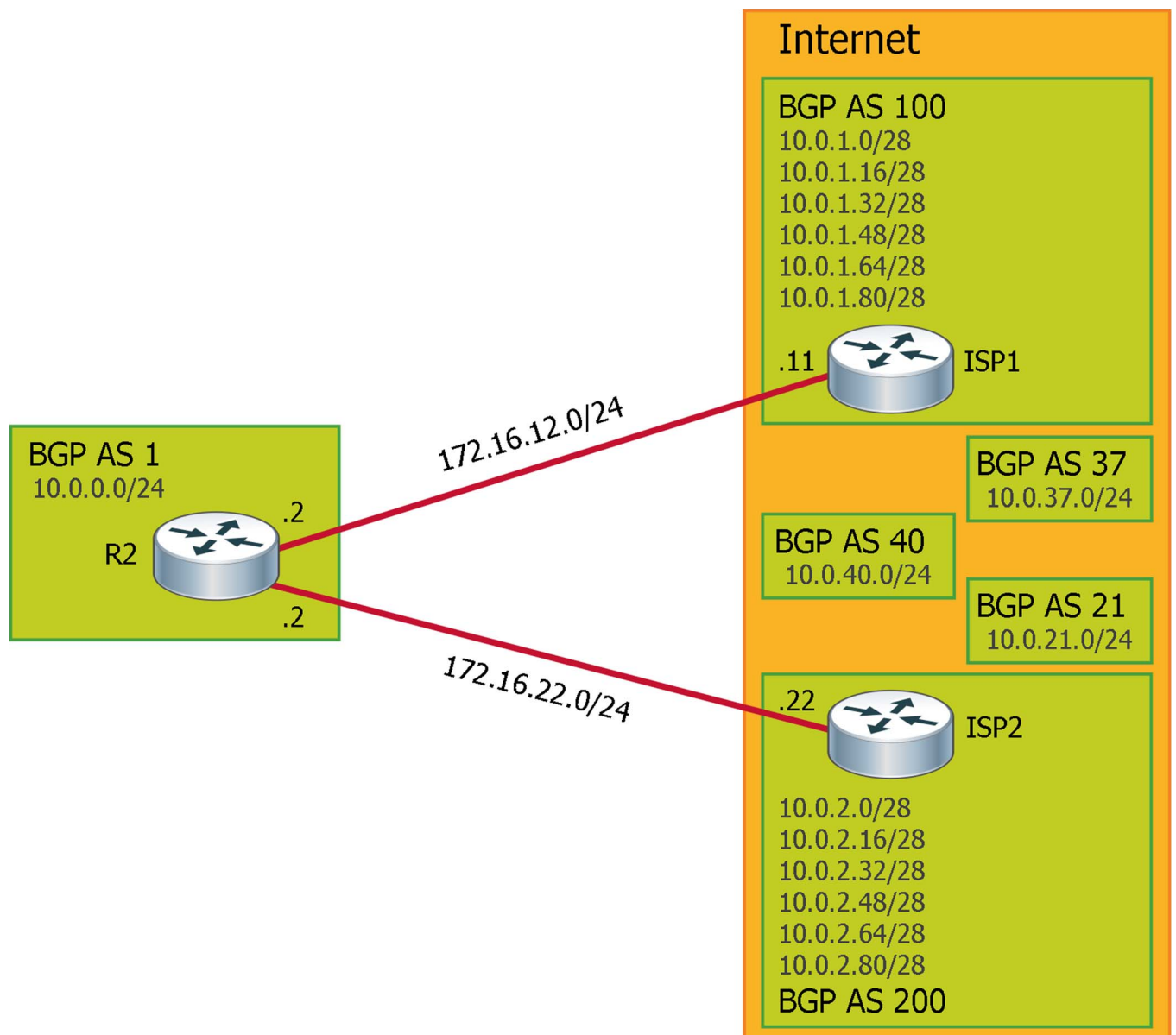
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP1	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate more networks in different autonomous systems.
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks
ISP2	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate more networks in different autonomous systems.

[IP addresses](#) and advertised networks in BGP are preconfigured as shown below:



BGP is also preconfigured as [EBGP](#) (R2 to ISP1 and R2 to ISP2).

Configure AS Path Filters

Discovery Steps

Step 1 On the ISP2 router, verify that you receive routes that the router ISP1 advertises.

```
ISP2# show ip bgp
BGP table version is 32, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	172.16.22.2	0			0 1 i
*>	10.0.1.0/28	172.16.22.2				0 1 100 i
*>	10.0.1.16/28	172.16.22.2				0 1 100 i
*>	10.0.1.32/28	172.16.22.2				0 1 100 i
*>	10.0.1.48/28	172.16.22.2				0 1 100 i
*>	10.0.1.64/28	172.16.22.2				0 1 100 i
*>	10.0.1.80/28	172.16.22.2				0 1 100 i
*>	10.0.2.0/28	0.0.0.0	0		32768	i
*>	10.0.2.16/28	0.0.0.0	0		32768	i
*>	10.0.2.32/28	0.0.0.0	0		32768	i
*>	10.0.2.48/28	0.0.0.0	0		32768	i
*>	10.0.2.64/28	0.0.0.0	0		32768	i
*>	10.0.2.80/28	0.0.0.0	0		32768	i
*>	10.0.2.96/28	0.0.0.0	0		32768	i
	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.2.112/28	0.0.0.0	0		32768	i
*>	10.0.2.128/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*	10.0.37.0/24	172.16.22.2				0 1 100 37 i
*>		0.0.0.0	0		32768	i
*	10.0.40.0/24	172.16.22.2				0 1 100 37 40 i
*>		0.0.0.0	0		32768	i

You can see that ISP1 router in AS 100 either originates or advertises all highlighted routes to the router R2:

- Networks 10.0.1.0/28, 10.0.1.16/28, 10.0.1.32/28, 10.0.1.48/28, 10.0.1.64/28 and 10.0.1.80/28 were originated by ISP1 router in AS 100 and advertised to the router R2.
- Router ISP 1 advertises networks 10.0.37.0/24 and 10.0.40.0/24 to the router R2.
- The customer router R2 originates route 10.0.0.0/24.

All these routes have next hop set to IP address 172.16.22.2, which is the IP address of the R2 router.

Step 2 On the R2 router, verify which routes are being advertised to EBGP neighbor router ISP2.

```
R2# show ip bgp neighbors 172.16.22.22 advertised-routes
BGP table version is 22, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.12.11	0		0	100 i
*>	10.0.1.16/28	172.16.12.11	0		0	100 i
*>	10.0.1.32/28	172.16.12.11	0		0	100 i
*>	10.0.1.48/28	172.16.12.11	0		0	100 i
*>	10.0.1.64/28	172.16.12.11	0		0	100 i
*>	10.0.1.80/28	172.16.12.11	0		0	100 i
*>	10.0.2.0/28	172.16.22.22	0		0	200 i
*>	10.0.2.16/28	172.16.22.22	0		0	200 i
*>	10.0.2.32/28	172.16.22.22	0		0	200 i
*>	10.0.2.48/28	172.16.22.22	0		0	200 i
*>	10.0.2.64/28	172.16.22.22	0		0	200 i
*>	10.0.2.80/28	172.16.22.22	0		0	200 i
*>	10.0.2.96/28	172.16.22.22	0		0	200 i
	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.2.112/28	172.16.22.22	0		0	200 i
*>	10.0.2.128/28	172.16.22.22	0		0	200 i
*>	10.0.21.0/24	172.16.22.22	0		0	200 21 i
*>	10.0.37.0/24	172.16.12.11	0		0	100 37 i
*>	10.0.40.0/24	172.16.12.11	0		0	100 37 40 i

Total number of prefixes 19

You can see that ISP1 router in AS 100 either originates or advertises all highlighted routes to the router R2:

- The customer router R2 advertises the locally originated route 10.0.0.0/24 to the router ISP2.
- The customer router R2 advertises networks 10.0.1.0/28, 10.0.1.16/28, 10.0.1.32/28, 10.0.1.48/28, 10.0.1.64/28 and 10.0.1.80/28 to the router ISP2.
- The customer router R2 advertises networks 10.0.37.0/24 and 10.0.40.0/24 to the router ISP2.

Step 3 Configure AS path filtering on the R2 router, so that only locally originated network will be advertised to both ISP1 and ISP2 routers.

```
R2(config)# ip as-path access-list 1 permit ^$
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 filter-list 1 out
R2(config-router)# neighbor 172.16.22.22 filter-list 1 out
```

Step 4 Clear all BGP sessions on the R2 router.

```
R2# clear ip bgp *
```

Step 5 On the ISP2 router, verify that you do not receive routes that the ISP1 router advertises.

```
ISP2# show ip bgp
BGP table version is 26, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	172.16.22.2	0		0	1 i
*>	10.0.2.0/28	0.0.0.0	0		32768	i
*>	10.0.2.16/28	0.0.0.0	0		32768	i
*>	10.0.2.32/28	0.0.0.0	0		32768	i
*>	10.0.2.48/28	0.0.0.0	0		32768	i
*>	10.0.2.64/28	0.0.0.0	0		32768	i
*>	10.0.2.80/28	0.0.0.0	0		32768	i
*>	10.0.2.96/28	0.0.0.0	0		32768	i
*>	10.0.2.112/28	0.0.0.0	0		32768	i
*>	10.0.2.128/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i

The only route that is received from the router R2 with next hop IP address 172.16.22.2 is 10.0.0.0/24, which is locally originated network on customer's router R2.

Step 6 On the R2 router, verify which routes are being advertised to EBGp neighbor router ISP2.

```
R2# show ip bgp neighbors 172.16.22.22 advertised-routes
BGP table version is 22, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	0.0.0.0	0		32768	i

Total number of prefixes 1

You should only see the network 10.0.0.0/24 being advertised to the EBGp neighbor ISP2. This network is locally originated on the R2 router. You made the customer AS1 a non-transit autonomous system.

Monitoring AS Path Filters

The Cisco IOS commands that are most frequently used to monitor the operation of configured [AS](#) path filters include **show ip as-path-access-list**, **show ip bgp regexp**, and **show ip bgp filter-list**.

Monitoring AS Path Filters

```
router# show ip as-path-access-list [filter list]
```

- Displays all routes in the BGP table matching regular-expression in one or all filter-lists

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

- Displays one or all filter-lists

```
router# show ip bgp filter-list access-list-number
```

- Displays all routes in the BGP table that the specified AS path access-list permits

© 2015 Cisco and/or its affiliates. All rights reserved.

183

show ip as-path-access-list Command

show ip as-path-access-list Command

- Displaying configured filters:

```
R2# show ip as-path-access-list
AS path access list 1
  permit ^$
AS path access list 7
  deny _21_
  permit .*
AS path access list 25
  permit 40$
```

© 2015 Cisco and/or its affiliates. All rights reserved.

184

The **show ip as-path-access-list** command displays a specific access-list or all AS path access-lists in the router.

show ip bgp regexp Command

show ip bgp regexp Command

- Routes that an expression matches:

```
R2# show ip bgp regexp _21_
BGP table version is 20, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop        Metric LocPrf Weight Path
*> 10.0.21.0/24      172.16.22.22         0           0 200 21 i
*   172.16.12.11      172.16.22.22         0           0 100 37 40 21 i
*   10.0.37.0/24      172.16.22.22         0           0 200 21 40 37 i
*   10.0.40.0/24      172.16.22.22         0           0 200 21 40 i
```

© 2015 Cisco and/or its affiliates. All rights reserved.

185

Because regular expressions sometimes get complex, thorough testing of them is required. The **show ip bgp regexp** command displays all routes currently in the [BGP](#) table that have an AS path attribute that matches the typed-in regular expression. Use the **show ip bgp regexp** command to test a regular expression that is typed in on the command line. The result is a printout on the screen of all those routes currently in the BGP table that had an AS path attribute matching the typed-in regular expression.

In the example, you wish to find all BGP routes passing through AS 21.

show ip bgp regexp Command

To display routes that match an AS path regular expression, use the **show ip bgp regexp** privileged EXEC command:

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

Syntax Description

Parameter	Description
<i>regular-expression</i>	Regular expression to match the BGP AS paths

show ip bgp regexp Command

- Routes that a filter-list matches:

```
R2# show ip as-path-access-list 25
AS path access list 25
    permit 40$

R2# show ip bgp filter-list 25
BGP table version is 20, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

      Network          Next Hop           Metric LocPrf Weight Path
*    10.0.40.0/24      172.16.22.22             0         0 200 21 40 i
*>                   172.16.12.11             0         0 100 37 40 i
```

© 2015 Cisco and/or its affiliates. All rights reserved.

186

An AS path access-list is even more complex because it is a combination of several regular expressions. There is one expression on each access-list line. Use the **show ip bgp filter-list** command to test the entire AS path access-list. The result is a printout on the screen of all the routes currently in the BGP table that the access-list permits their AS path attribute.

The AS path access-list number 25 in the example figure consists of one single line. It permits the routes that originate in AS 40. The **show ip bgp filter-list** command displays all the routes currently in the BGP table that the AS path access-list 25 permits.

show ip bgp filter-list Command

To display routes that conform to a specified filter-list, use the **show ip bgp filter-list** privileged EXEC command:

```
show ip bgp filter-list access-list-number
```

Syntax Description

Parameter	Description
<i>access-list-number</i>	Integer from 1 to 199 that indicates the regular expression access-list number

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Several scenarios require BGP route-filtering that is based on AS path, including:
 - Announcing only local routes to the ISP (AS path needs to be empty).
 - Selecting routes that are based on a specific AS number in the AS path.
 - Accepting routes for a specific AS only from some BGP neighbors
- By applying specific selection criteria to the contents of the AS path attribute, routers can select a subset of routes from the total set of routes that are received.
- Cisco IOS software internally translates the AS path encoding, which is carried with all BGP routes into a character string. This string is then tested against the regular expression.
- String matching operates when you are using AS path regular expressions to match BGP routes.

© 2015 Cisco and/or its affiliates. All rights reserved.

187

Summary (Cont.)

- You can use AS path filters to select those routes that will be allowed.
- AS path access-list creates an AS path filter, which is applied to a set of routes from which to select a subset. The **ip as-path access-list** global configuration command defines a BGP AS path access-list, and the **neighbor filter-list** router configuration command sets up a BGP filter.
- There are a number of Cisco IOS commands that are required to monitor the operation of configured AS path filters, including **show ip as-path-access-list**, **show ip bgp regexp**, and **show ip bgp filter-list**.

© 2015 Cisco and/or its affiliates. All rights reserved.

188

Lesson 3: Filtering with Prefix Lists

Overview

Where multiple paths between a customer and an [ISP](#) exist, there is a requirement to filter certain information during [BGP](#) updates. You filter the information to influence the route selection or to enforce an administrative policy. To meet this requirement, you must use filters. Using prefix lists is typically easier than using standard IP access lists and provides performance benefits. It is important for you to understand the commands to apply filtering of inbound or outbound updates with prefix lists and where they should be applied.

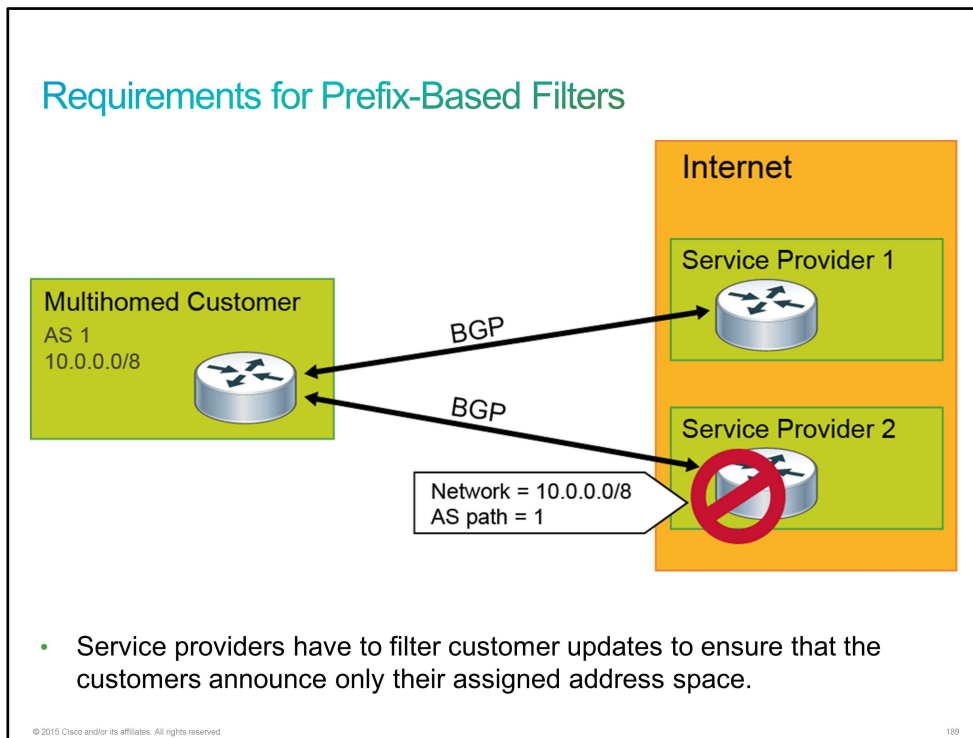
In this lesson, you will learn the requirements for using prefix-based filters in customer implementations where connections to multiple ISPs must be supported. You will also learn the advantages of prefix lists over IP access lists. The commands to apply filtering of inbound or outbound updates with prefix lists and to configure prefix list filters are discussed, and also where network administrators should apply them.

Upon completing this lesson, you will be able to:

- Identify the requirement for prefix-based filters in network implementations where multiple connections between a customer and ISPs exist
- List the advantages of prefix lists versus IP access lists
- Identify the Cisco IOS command that is required to configure prefix list filters
- Describe where you can implement prefix lists in a BGP network
- Identify the Cisco IOS commands that are required to apply filtering of inbound or outbound updates with prefix lists
- Identify the Cisco IOS commands that are required to modify configured prefix-list filters
- Identify the Cisco IOS commands that are required to monitor the operation of configured prefix list filters

Requirements for Prefix-Based Filters

Customers with multihomed networks are responsible for announcing their own networks using [BGP](#). Typically, customers are not as experienced with BGP as service providers, and therefore problems are more likely to occur. A service provider with a multihomed customer must take precautions not to accept, use, or forward any erroneous routing information that is received from the customer.



The customer is assigned a set of IP network numbers that it should announce. If the customer announces any additional networks, something is wrong. The customer may have forgotten not to act as a transit [AS](#) and may have started propagating routes that it has received from the other service provider. Or, the customer may have accidentally started to announce its private address space. The customer may use this private address space for address links, loopback interfaces, or other devices that should never access the Internet.

To avoid problems, the service provider can apply an IP prefix filter on the incoming information from the customer. The service provider will accept only network numbers that the access list or prefix list permits.

Prefix Lists vs. IP Access Lists

Traditionally, the filtering of IP network numbers has been accomplished using an access list. The access list is then bound to either the incoming or outgoing information of a neighbor by using the **neighbor distribute-list** command. A [BGP](#) update about a network number that the access list permits will be accepted. An update about a network number that the access list denies will be dropped.

Prefix Lists vs. IP Access Lists

Traditional prefix filters

- Traditional IP prefix filters were implemented with IP access lists configured with the **distribute-list** command.
- IP access lists used as route filters have several drawbacks:
 - Subnet mask cannot be easily matched.
 - Access lists are evaluated sequentially for every IP prefix in the routing update.
 - Extended access lists can be cumbersome to configure.

© 2015 Cisco and/or its affiliates. All rights reserved.

190

However, standard access lists do not support the testing of the subnet masks. If the access list permits 10.0.0.0/16, it would also permit 10.0.0.0/8.

Extended access lists can do testing on both an IP network number and a subnet mask, but the syntax is cumbersome.

Prefix Lists vs. IP Access Lists (Cont.)

Prefix lists

- Route-filtering mechanism
- Significant performance improvement on long filters
 - Inside Cisco IOS software, the prefix list is a tree structure and is not scanned sequentially.
- Support for incremental updates
 - Individual entries in prefix lists can be inserted or deleted.
- More user-friendly CLI
 - The CLI for using access lists to filter BGP updates is difficult to understand and use, because it uses the packet-filtering format.
- Greater flexibility; can match on subnet masks

© 2015 Cisco and/or its affiliates. All rights reserved.

191

The **ip prefix-list** configuration command has several benefits if you compare it to the **access-list** command. The intended use of prefix lists is limited to route filtering, whereas access lists were initially intended for packet filtering, which was then extended to filter routes.

The prefix list is internally transformed into a tree structure, with each branching of the tree serving as a test. Cisco IOS software determines a verdict of either "permit" or "deny" much faster this way, compared to sequentially interpreting an access list.

The configuration [CLI](#) that you use when configuring the **ip prefix-list** command, allows you to assign a line number to each line of the prefix list. The router uses this number to sort the entries in the prefix list. If the lines are initially assigned line numbers, with some spacing in between them, administrators can later insert extra lines. Individual lines can also be removed without removing the entire list.

Routers match network numbers in a routing update against the prefix list, using as many bits as indicated. For example, a prefix list can be specified to be 10.0.0.0/16, which will match 10.0.0.0 routes but not 10.1.0.0 routes.

Optionally, the prefix list can also specify the size of the subnet mask. In addition, the prefix list can indicate that the subnet mask must be in a specified range.

Prefix Lists vs. IP Access Lists (Cont.)

- Key access list features are preserved.
 - Filtering using "permit" or "deny"
 - Order dependency (first match wins)
 - Security-focused: no match means "deny"
- The matching mechanism has changed.
 - Matches routes in a part of address space with subnet mask longer or shorter than a set number

© 2015 Cisco and/or its affiliates. All rights reserved.

192

The prefix list shares several similarities with the access list. It can consist of any number of lines, each of which indicates a test and a result. The router can interpret the lines in the specified order, although this process is optimized in the Cisco IOS software. When a router evaluates a route against the prefix list, the first line that matches results in either a "permit" or "deny." If none of the lines in the list match, the result is "implicitly deny."

Testing is done using prefixes. The indicated number of bits in the prefix is compared with the same number of bits in the network number in the update. If the bits match, testing continues with an examination of the number of bits set in the subnet mask. The prefix list line can indicate a range within which the number must fall to pass the test. If no range is indicated, the subnet mask must match the prefix size.

Configuring Prefix Lists

You can filter certain information during [BGP](#) updates using prefix list filters to influence route selection or to enforce an administrative policy.

Configuring Prefix Lists

```
router(config)# ip prefix-list list-name [seq seq] {permit|deny} network/len  
[ge value] [le value]
```

- Prefix lists have names and sequence numbers (like route-maps).
- An entry with no **le** or **ge** parameter matches exactly the specified prefix.
- An entry with an **le** or **ge** parameter matches any route within the address space of *address/prefix*. But the prefix must be longer or equal to **ge value** and shorter than or equal to **le value**.

© 2015 Cisco and/or its affiliates. All rights reserved.193

ip prefix-list Configuration Command

To create an entry in a prefix list, use the **ip prefix-list global** configuration command.

```
ip prefix-list list-name [seq seq-value] {permit | deny} network/len [ge ge-value] [le  
le-value]
```

Syntax Description

Parameter	Description
<i>list-name</i>	Name of a prefix
seq	(Optional) Applies the sequence number to the prefix
<i>seq-value</i>	(Optional) Specifies the sequence number for the prefix
permit	Permits access for matching conditions
deny	Denies access to matching conditions
<i>network/len</i>	(Mandatory) The network number and length (in bits) of the subnet mask
ge	(Optional) Applies the <i>ge-value</i> to the range specified
<i>ge-value</i>	(Optional) Specifies the lesser value of a range (the "from" portion of the range description)
le	(Optional) Applies the <i>le-value</i> to the range specified
<i>le-value</i>	(Optional) Specifies the greater value of a range (the "to" portion of the range description)

When multiple entries of a prefix list match a given prefix, the sequence number of a prefix-list entry identifies the entry with the lowest sequence number. In this case, the entry with the smallest sequence number is considered to be the "real" match.

Note	You can specify sequence values for prefix list entries in any increments that you want (the automatically generated numbers are increased in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert more entries into the prefix list. If you choose very large increments, you could run out of sequence values.
-------------	--

You can use the parameters **ge** and **le** to specify the range of the prefix length to be matched for prefixes that are more specific than *network/len*. The exact match is assumed when neither **ge** nor **le** is specified. The range is assumed to be from *ge-value* to 32 only if the **ge** attribute is specified. The range is assumed to be from *len* to *le-value* only if the **le** attribute is specified.

Configuring Prefix Lists (Cont.)

Prefix list matching rules

- Prefix list entries with no **ge** or **le** option match only the specified route.
 - Similar to IP access lists with no wildcard bits
 - Matching also considers the subnet mask

Which of the following routes will be matched by:

```
ip prefix-list MyList permit 192.168.0.0/16?
```

192.168.0.0/16  192.168.0.0/20  192.168.2.0/16 

© 2015 Cisco and/or its affiliates. All rights reserved.

194

Prefix list entries without the **ge** or **le** option match only the route with the specified [IP address](#) and subnet mask. In the example here, the prefix list entry **permit 192.168.0.0/16** will not match the route 192.168.2.0/24 because of the mismatch in the IP address. It will also not match the route 192.168.0.0/20 because of the mismatch in the subnet mask.

Configuring Prefix Lists (Cont.)

- A prefix list entry with **ge** or **le** option matches any prefix within specified address space where the subnet mask falls within the specified limits.

Which of the following routes will be matched by:

```
ip prefix-list MyList permit 192.168.0.0/16 le 20?
```

192.168.0.0/16  192.168.17.0/20  192.168.2.0/24 

```
ip prefix-list MyList permit 192.168.0.0/16 ge 18?
```

192.168.0.0/16  192.168.17.0/20  192.168.2.0/24 

© 2015 Cisco and/or its affiliates. All rights reserved.

195

Prefix list entries with the **ge** or **le** option specified match any prefix within the address space that the *network/len* parameter specifies, under one condition. The subnet mask length of the route must fall within the range that the **le** and **ge** parameters specify.

In the first example in the figure, the route 192.168.2.0/24 is not matched by the prefix list entry **permit 192.168.0.0/16** even though the IP address falls within the specified address range. It is because the subnet mask is too long.

In the second example, the route 192.168.0.0/16 is not matched by prefix list entry **permit 192.168.0.0/18** because the subnet mask is too short.

Configuring Prefix Lists Example

Configuring Prefix Lists Example

What will be matched by:

- a) ip prefix-list A **permit 0.0.0.0/0 ge 32**
 - b) ip prefix-list B **permit 128.0.0.0/2 ge 17**
 - c) ip prefix-list C **permit 0.0.0.0/0 le 32**
 - d) ip prefix-list D **permit 0.0.0.0/0**
 - e) ip prefix-list E **permit 0.0.0.0/1 le 24**
-
- a) All host routes
 - b) Any subnet in class B address space
 - c) All routes
 - d) Just the default route
 - e) Any prefix in class A address space covering at least 256 addresses

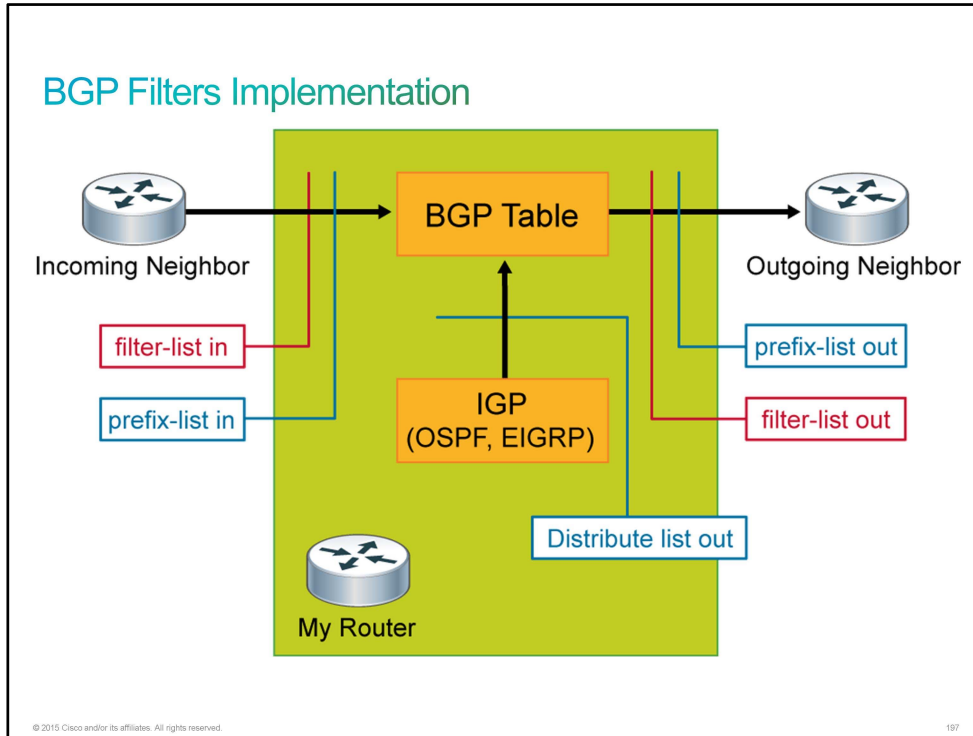
© 2015 Cisco and/or its affiliates. All rights reserved.195

The list contains some commonly used prefix list examples.

In the list, prefix list A **permit 0.0.0.0/0 ge 32** will match all host routes. Prefix list B **permit 128.0.0.0/2 ge 17** will match any subnet in a class B address space. Prefix list C **permit 0.0.0.0/0 le 32** will match all routes, but only the default route will be matched by prefix list D **permit 0.0.0.0/0**. Finally, any prefix in a class A address space that covers at least 256 addresses will be matched to prefix list E **permit 0.0.0.0/1 le 24**.

BGP Filters Implementation

You can optionally apply filter lists and prefix lists on either incoming or outgoing neighbors in any combination. Both the incoming prefix list and the incoming filter list must permit the routes that are received from a neighbor before they are accepted into the [BGP](#) table. Outgoing routes must pass both the outgoing filter list and the outgoing prefix list before being transmitted to the neighbor.



When a router is configured to redistribute routing information from an [IGP](#) into BGP, the following must be true before a route is injected into the BGP table:

- The routes must successfully pass any prefix list or access list that is applied to the redistribution.

Implementing Prefix Lists in the BGP Process

You can use prefix lists to filter incoming or outgoing [BGP](#) updates to neighbors. You can also use prefix lists to filter routes that are being redistributed into the BGP process from other routing protocols.

Implementing Prefix Lists in the BGP Process

```
router(config-router)# neighbor {ip-address|peer-group-name} prefix-list  
prefix-listname {in|out}
```

- Filters inbound or outbound BGP routing updates for a configured neighbor session

```
router(config-router)# distribute-list prefix-list prefix-list out routing-  
process
```

- Filters routes redistributed from specified routing process into BGP

© 2015 Cisco and/or its affiliates. All rights reserved.198

neighbor prefix-list Command

To distribute BGP neighbor information as specified in a prefix list, use the **neighbor prefix-list** router configuration command:

```
neighbor {ip-address | peer-group-name} prefix-list prefix-listname {in | out}
```

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

```
no neighbor {ip-address | peer-group-name} prefix-list prefix-listname {in | out}
```

Syntax Description

Parameter	Description
<i>ip-address</i>	IP address of neighbor
<i>peer-group-name</i>	Name of a BGP peer group
<i>prefix-listname</i>	Name of a prefix list
in	Access list is applied to incoming advertisements to that neighbor
out	Access list is applied to outgoing advertisements from that neighbor

Note	A BGP peer group is a group of BGP neighbors with the same update policies. Route maps, distribute lists, filter lists, and so on usually set update policies. Instead of defining the same policies for each separate neighbor, a peer group name is configured on the router, and these policies are assigned to the peer group.
-------------	--

distribute-list out Command

To suppress networks from being advertised in updates, use the **distribute-list out** router configuration command:

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

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

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

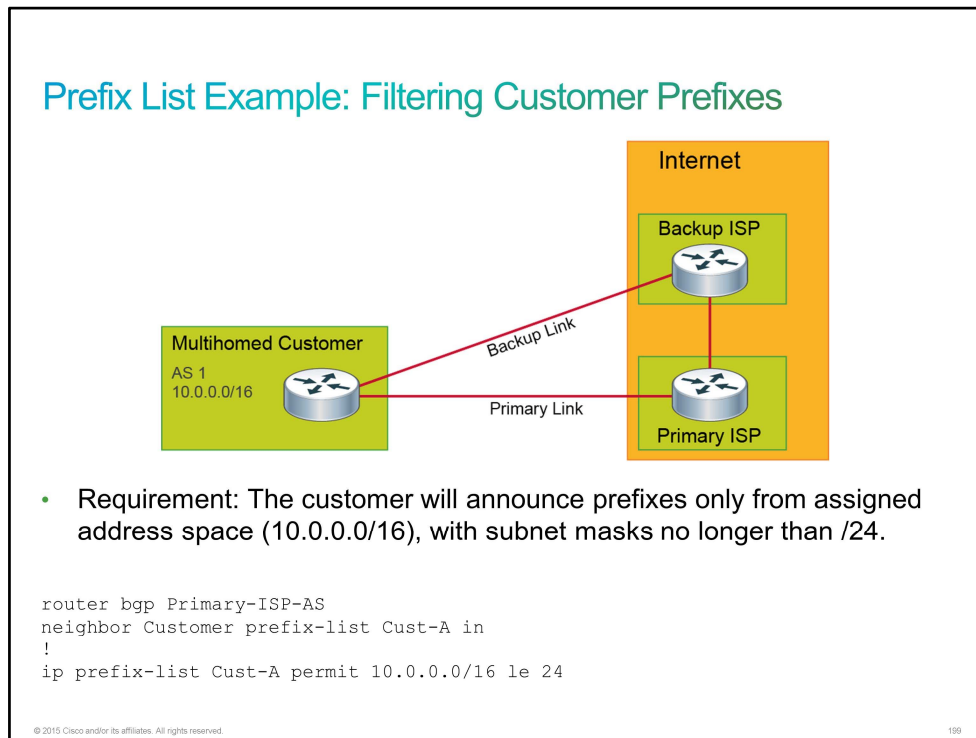
Syntax Description

Parameter	Description
<i>access-list-number</i> <i>name</i>	Standard IP access list number or name. The list defines which networks should be received and which should be suppressed in routing updates.
<i>prefix-listname</i>	Name of a prefix list. The list defines which networks should be received and which should be suppressed in routing updates. This decision is based on matching of the network prefix to the prefixes in the list.
out	Applies the access list to outgoing routing updates.
<i>interface-name</i>	(Optional) Name of a particular interface.
<i>routing-process</i>	(Optional) Name of a particular routing process, or the keyword static or connected .
<i>autonomous-system-number</i>	(Optional) AS number.

Note	Although you can use the neighbor prefix-list router configuration command as an alternative to the neighbor distribute-list command, do not use both the neighbor prefix-list and neighbor distribute-list command filtering for the same neighbor in any given direction. These two commands are mutually exclusive, and only one command (neighbor prefix-list or neighbor distribute-list) can be applied for each inbound or outbound direction.
-------------	---

Prefix List Example: Filtering Customer Prefixes

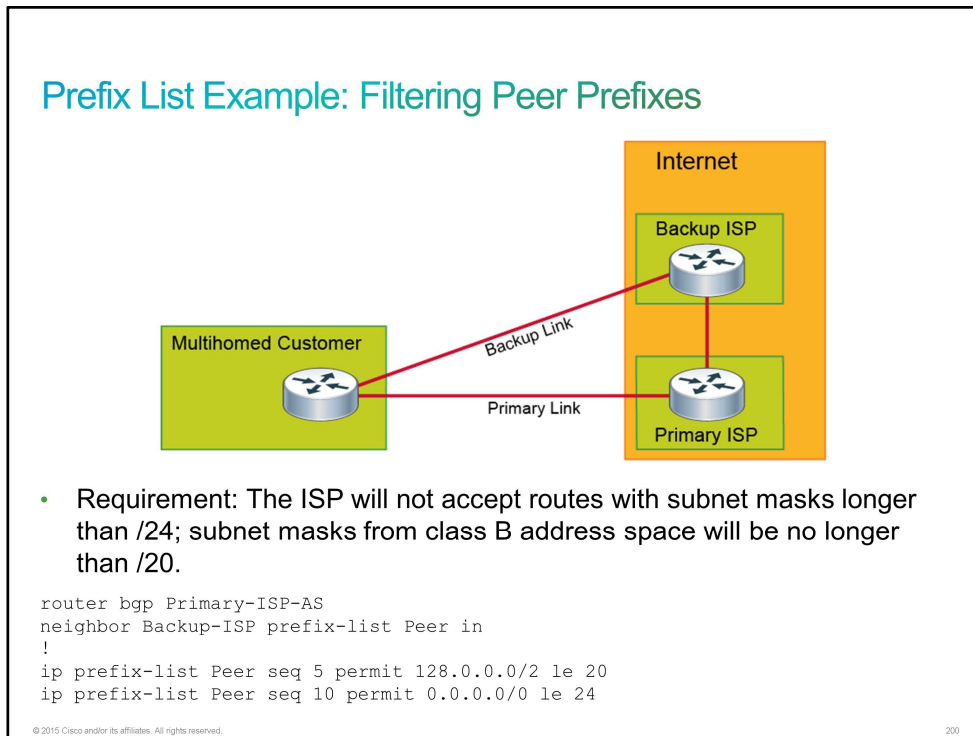
The figure illustrates filtering customer prefixes.



In this example, a multihomed customer has been assigned the address space 10.0.0.0/16. The customer may subnet this address space but may not announce subnets smaller than a subnet mask of 255.255.255.0. Larger subnets are accepted. If the customer has subnetted the network into smaller subnets, it must summarize the routing information about those subnets into at least /24 prefixes before announcing them.

The primary [ISP](#) implements a prefix list with a Cust-A name, to perform the filtering of incoming information from the multihomed customer. The prefix list permits all routes that are received from the customer that have 10.0 in the first 16 bits and have a subnet mask of 24 bits or less. Any other routes from the customer are denied and silently ignored.

Prefix List Example: Filtering Peer Prefixes



In this example, the primary ISP will not accept any route from the customer that indicates a subnet smaller than a 255.255.255.0 subnet mask. The class B network, however, must not be subnetted into subnets smaller than a 255.255.240.0 subnet mask.

The primary ISP implements this route by using a prefix list named Peer. The first line in the prefix list checks whether it is a class B network. Remember that a class B address always has the binary sequence 10 as the first 2 bits in the first byte. The second line matches any prefix.

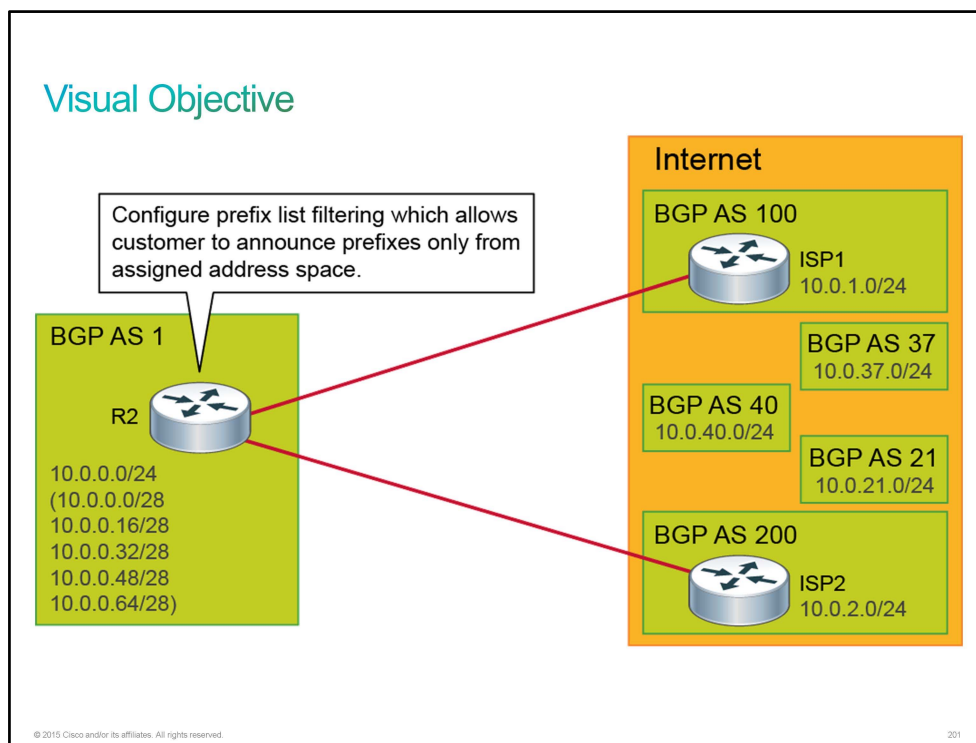
When the primary ISP receives a route from the customer, it compares the route with both lines. If the route is a class B network, both lines match. Testing continues with checking the subnet mask. An upper bound is explicitly indicated, giving a maximum prefix length of 20 bits.

If the received route is not a class B network, only the second line matches. In this case, the subnet mask length must be greater than or equal to 0 and less than or equal to 24. This rule provides a route less explicit than a /24 prefix.

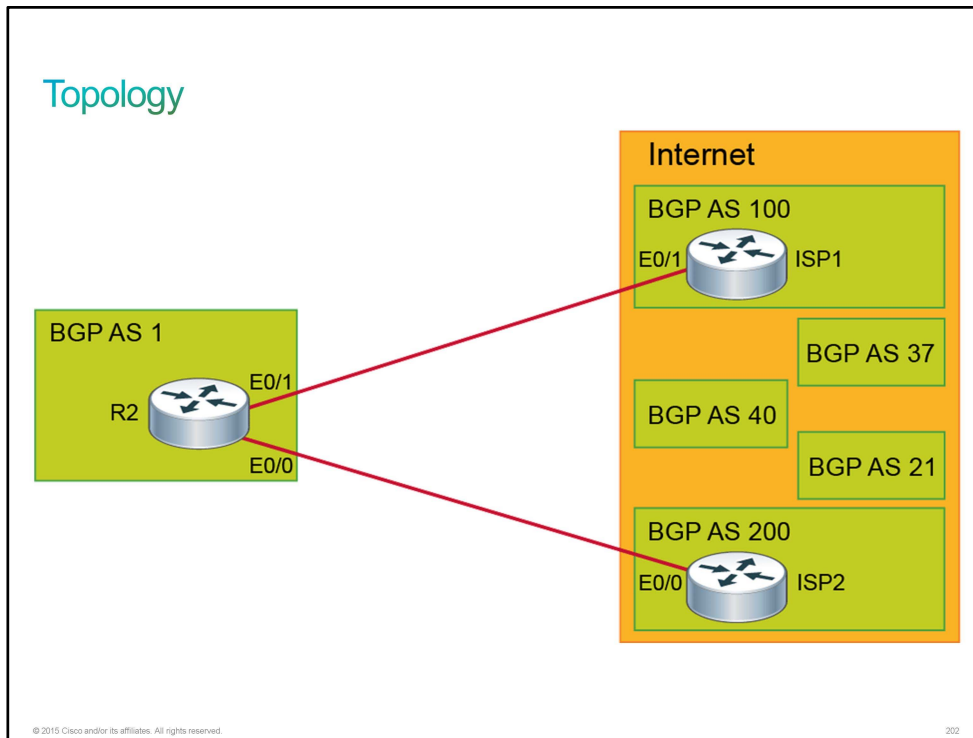
Discovery 8: Filtering Customer Prefixes

Overview

Through this discovery, you will learn how to configure a prefix list to filter customer prefixes. You as an ISP1 service provider will allow the customer on the router R2 to announce prefixes only from the assigned address space (10.0.0.0/24). You will not allow the customer to further subnet the assigned address space, thus subnet mask should be 24.



Topology



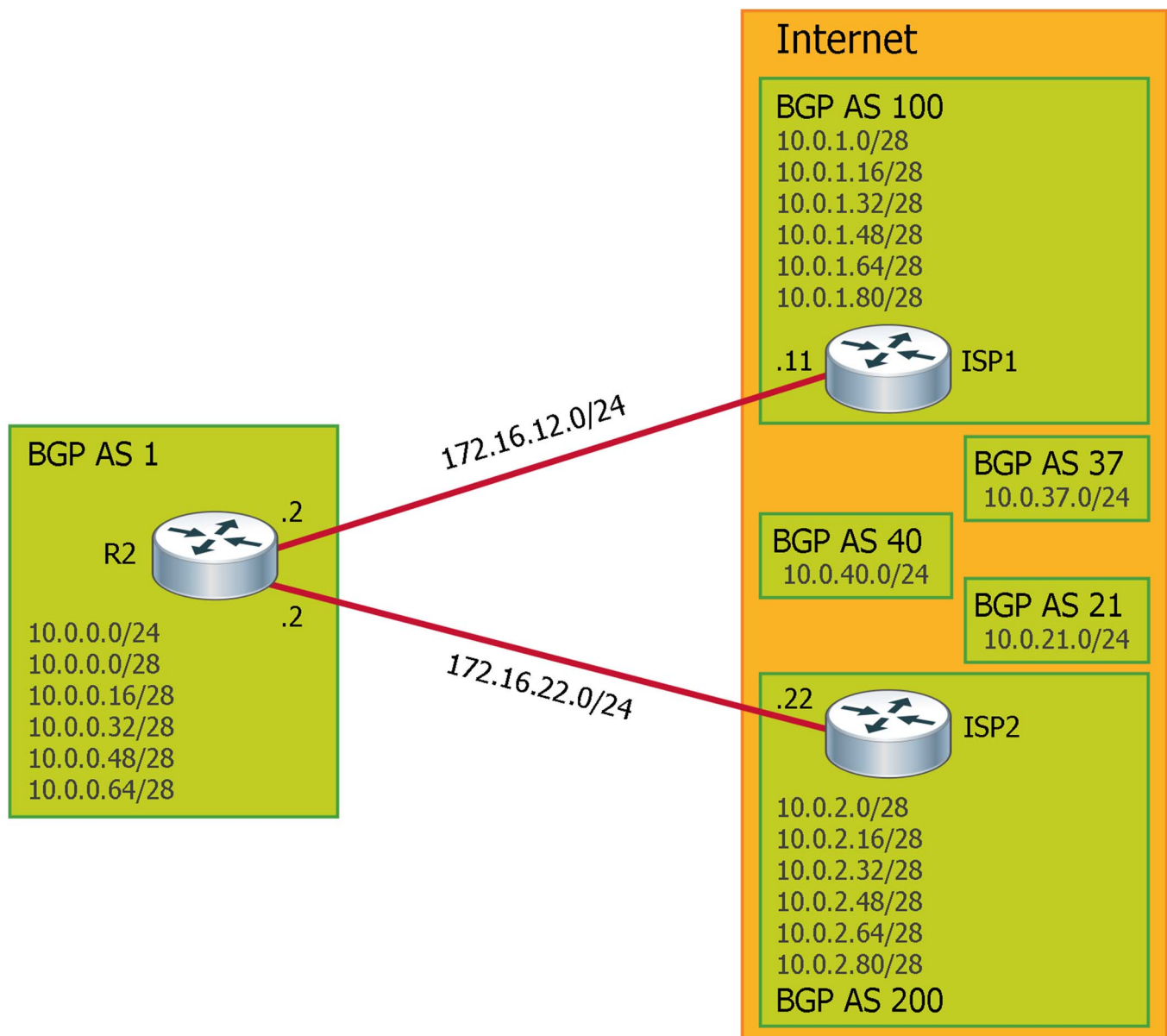
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP1	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate additional networks in different autonomous systems.
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28	Loopbacks simulate LAN networks
ISP2	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate additional networks in different autonomous systems.

[IP addresses](#) and advertised networks in [BGP](#) are preconfigured as shown below:



BGP is also preconfigured as [EBGP](#) (R2 to ISP1 and R2 to ISP2).

The R2 router announces these networks:

- 10.0.0.0/24 (assigned address space to the Customer)
- 10.0.0.16/28
- 10.0.0.32/28
- 10.0.0.48/28
- 10.0.0.64/28

Filtering Customer Prefixes

Discovery Steps

Step 1 On the ISP1 router, verify that you receive routes with mask larger than /24 from the customer router R2.

```
ISP1# show ip bgp
BGP table version is 15, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	172.16.12.2	0		0	1 i
*>	10.0.0.16/28	172.16.12.2	0		0	1 i
*>	10.0.0.32/28	172.16.12.2	0		0	1 i
*>	10.0.0.48/28	172.16.12.2	0		0	1 i
*>	10.0.0.64/28	172.16.12.2	0		0	1 i
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i

You can see that customer announces four networks with mask /28.

Step 2 On the R2 router, verify that customer router R2 originates assigned network 10.0.0.0/24 and four subnets from this assigned address space with mask /28.

```
R2# show ip bgp neighbors 172.16.12.11 advertised-routes
BGP table version is 58, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.32/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.0.64/28	0.0.0.0	0		32768	i

Total number of prefixes 5

Customer assigned address space is 10.0.0.0/24, which is further subnetted to four /28 subnets, which are also announced in BGP updates.

Step 3 Configure prefix list filtering on the ISP1 router to allow the customer on the R2 router to announce prefixes only from the assigned address space (10.0.0.0/24).

You should not allow the customer to further subnet the assigned address space, thus subnet mask should be 24.

```
ISP1(config)# ip prefix-list Customer permit 10.0.0.0/24
ISP1(config)# router bgp 100
ISP1(config-router)# neighbor 172.16.12.2 prefix-list Customer in
```

Step 4 Clear all BGP sessions on the R2 router.

```
R2# clear ip bgp *
```

Step 5 On the ISP1 router, verify that you receive routes with mask /24 only from the customer router R2.

```
ISP1# show ip bgp
BGP table version is 21, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	172.16.12.2	0		0	1 i
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i

You can see that only customer network 10.0.0.0 with mask /24 is put into the BGP table because of the prefix list filtering configured.

Step 6 On the R2 router, verify that customer router R2 still originates assigned network 10.0.0.0/24 and four subnets from this assigned address space with mask /28.

```
R2# show ip bgp neighbors 172.16.12.11 advertised-routes
BGP table version is 58, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.32/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.0.64/28	0.0.0.0	0		32768	i

Total number of prefixes 5

This way, you confirm that filtering of BGP updates occurs on the ISP1 router.

Modifying Prefix Lists

Lines in a prefix list are assigned sequence numbers. These assignments significantly improve the manageability of the list. These sequence number assignments provide the opportunity to remove a specific line. If the spacing between the sequence numbers allows, they also give you the ability to insert a line between two existing lines.

Modifying Prefix Lists

```
router# show ip prefix-list list-name [detail|summary]
```

- Displays the prefix list and the sequence numbers

```
router(config)# no ip prefix-list seq seq condition
```

- Erases the line with the specified sequence number from the prefix list

```
router(config)# ip prefix-list seq seq condition
```

- Inserts the line into the prefix list at the specified point

© 2015 Cisco and/or its affiliates. All rights reserved.

203

To display a currently configured prefix list and its sequence numbers, use the **show ip prefix-list** command with the **detail** keyword.

You can specify sequence values for prefix list entries in any increments that you want (the automatically generated numbers are increased in units of 5). If you specify the sequence values in increments of 1, you will not be able to insert more entries into the prefix list. If you choose very large increments, you could run out of sequence values.

Monitoring Prefix Lists

You can monitor prefix lists by displaying information about a prefix list or by using prefix lists for selective filtering of [BGP](#) table output on Cisco IOS devices.

Monitoring Prefix Lists

```
router# show ip prefix-list [detail | summary] prefix-list-name [network/length]
[seq sequence-number] [longer] [first-match]
```

- To display information about a prefix list or prefix-list entries

```
router# show ip bgp prefix-list prefix-list-name
```

- Displays all routes in the BGP table matching the prefix list.
- Used for easier monitoring of a desired network prefix group in the BGP table.

© 2015 Cisco and/or its affiliates. All rights reserved.204

show ip prefix-list Command

To display information about a prefix list or prefix list entries, use the **show ip prefix-list** EXEC command:

```
show ip prefix-list [detail | summary] name [network/len] [seq seq-num] [longer]
[first-match]
```

Syntax Description

Parameter	Description
detail summary	(Optional) Displays detailed or summarized information about all prefix lists
<i>name</i>	(Optional) The name of a specific prefix list
<i>network/len</i>	(Optional) The network number and length (in bits) of the network mask
seq	(Optional) Applies the sequence number to the prefix-list entry
<i>seq-num</i>	The sequence number of the prefix
longer	Displays all entries of a prefix that are more specific than the given <i>network/len</i>
first-match	Displays the entry of a prefix that matches the given <i>network/len</i>

You can use **show ip bgp prefix-list** command to display selected routes from a BGP routing table based on the contents of a prefix list. Use this command for selective filtering of BGP table output on Cisco IOS devices based on network prefix groups.

To perform prefix list-based BGP table filtering, follow these steps:

1. Configure a prefix list that permits ranges of networks that are meant to be displayed in the BGP table output.
2. Include a reference to a configured prefix list in the **show ip bgp prefix-list** command.

Note	The support for prefix list BGP table filtering was added in Cisco IOS Software Release 12.2(11)T and 12.0(14)ST.
-------------	---

Monitoring Prefix Lists (Cont.)

```
ISP1# show ip prefix-list detail InFilter
ip prefix-list InFilter:
  count: 4, range entries: 3, sequences: 5 - 20, refcount: 2
  seq 5 deny 128.0.0.0/2 le 15 (hit count: 0, refcount: 1)
  seq 10 deny 192.0.0.0/3 ge 25 (hit count: 0, refcount: 2)
  seq 15 deny 193.0.0.0/8 ge 21 (hit count: 0, refcount: 1)
  seq 20 permit 0.0.0.0/0 (hit count: 0, refcount: 1)
```

- Displays details about InFilter prefix list

© 2015 Cisco and/or its affiliates. All rights reserved.

205

In this example, the **show ip prefix-list** command has been issued with the **detail** keyword. The output of the command displays detailed information about configured prefix lists. The sequence numbers, the prefix list entries, and the number of times that the corresponding prefix matched each entry.

Monitoring Prefix Lists (Cont.)

```
ISP1(config)# ip prefix-list MyFilter permit 10.0.1.0/24 ge 28
```

```
ISP1# show ip bgp prefix-list MyFilter
BGP table version is 21, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i

- Displays all entries in BGP table with the first three octets equal to 10.0.1 and with length of a subnet greater or equal to 28

© 2015 Cisco and/or its affiliates. All rights reserved.

206

This example shows a simple prefix list-based filtering of the BGP table. The prefix list filter permits all networks with the first three octets equal to 10.0.1 and with length of a subnet greater or equal to 28. The **show ip bgp prefix-list** command output displays only the networks that the prefix list filter permits.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Customers with multihomed networks are responsible for announcing their own networks using BGP. Service providers with multihomed customers must take precautions not to accept, use, or forward any erroneous routing information that is received from their customers.
- Prefix lists have a number of advantages over access lists, including faster "permit" or "deny" determinations and easier CLI editing.
- Prefix lists are configured using the **ip prefix-list** global configuration command.
- Filter lists and prefix lists can be optionally applied on either incoming or outgoing neighbors in any combination.

© 2015 Cisco and/or its affiliates. All rights reserved.

207

Summary (Cont.)

- Prefix lists can filter incoming or outgoing BGP updates to neighbors and filter routes that are being redistributed into the BGP process from other routing protocols. Use the **neighbor prefix-list router configuration command** to distribute BGP neighbor information as specified in a prefix list.
- Certain Cisco IOS commands (such as the **show ip prefix-list** command) are used to modify configured prefix list filters.
- To display or monitor statistics about a prefix list or prefix list entries, you can use the **show ip prefix-list EXEC** command.

© 2015 Cisco and/or its affiliates. All rights reserved.

208

Lesson 4: Using Outbound Route Filtering

Overview

You can use [ORF](#) as an extra mechanism to minimize the number of updates that are requested from a neighbor. It will reduce the link bandwidth consumption and [CPU](#) use when a router requests a route refresh. An ORF also allows filtering of information that external networks should not receive (such as [RFC 1918](#) information). Understanding how to monitor outbound route filtering capabilities is also important. A [BGP](#) neighbor that supports specific ORF capabilities will report those capabilities to a monitoring neighbor and can then send a filter of the supported type to the neighbor.

This lesson discusses the function of outbound route filtering in a BGP network. The format and function of ORF messages are discussed, as well as the commands that enable ORF negotiations and the activation of an ORF prefix list. The commands that are used to trigger a route refresh are also detailed. Finally, there is a discussion on how to monitor the operations of a configured ORF in a BGP network.

Upon completing this lesson, you will be able to:

- Describe the function of outbound route filtering in a BGP network
- Compare standard inbound filtering and outbound route filtering
- Describe the function of BGP prefix-based outbound route filtering
- Describe the format and function of an ORF message
- Identify the Cisco IOS command that is required to enable ORF negotiations and activate an ORF prefix list
- Identify the Cisco IOS command that is used to trigger a route refresh message
- Identify the Cisco IOS command that is required to monitor the operation of a configured ORF

Outbound Route Filtering

Outbound route filtering is a prefix-based [BGP](#) feature that is enabled through the advertisement of [ORF](#) capabilities to peer routers and is integrated in Cisco IOS Software Release 12.2(4)T. The advertisement of the ORF capability indicates that a BGP-speaking router will accept a prefix list from a neighbor. It will apply the prefix list to locally configured ORFs (if any exist). When this capability is enabled, the BGP speaker can install an inbound prefix list filter to the remote peer as an outbound filter, which reduces unwanted routing updates.

Outbound Route Filtering

- The purpose of outbound route filtering is to reduce the amount of BGP traffic and CPU use needed to process routing updates.
- Routers exchange inbound filter configurations, which are used as outbound filters on neighboring routers.
- Filters are described in ORF entries.
- ORF entries are part of the route refresh message.

© 2015 Cisco and/or its affiliates. All rights reserved.

209

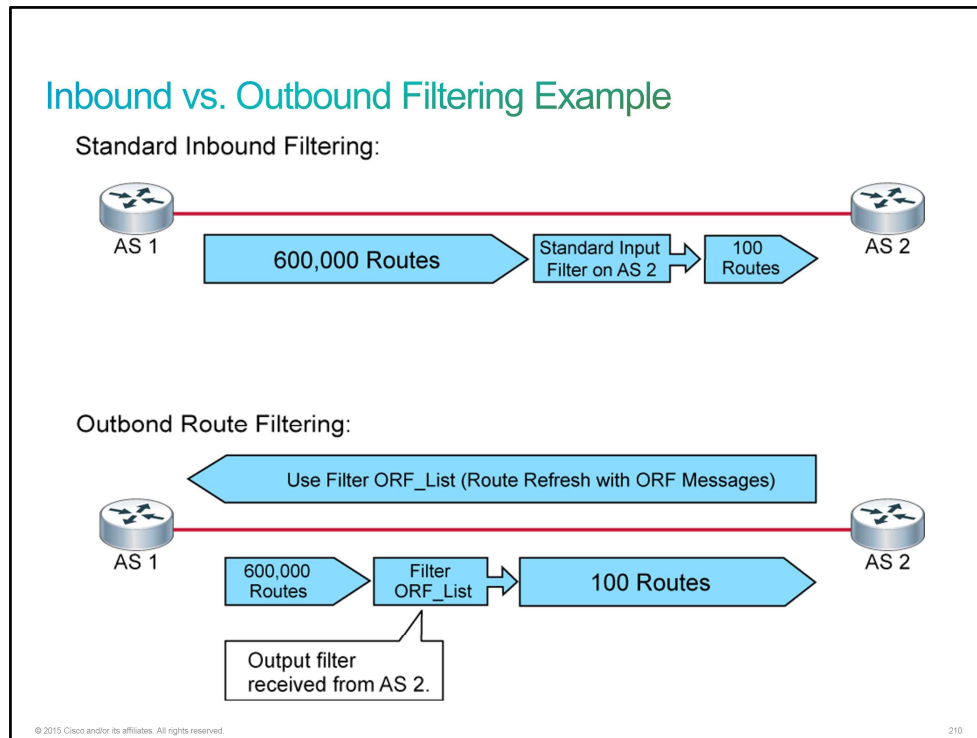
The standard route refresh message contains the [AFI](#) for which the refresh is needed. Outbound route filtering is an extra mechanism that is used to minimize the number of updates that are requested from a neighbor.

This mechanism reduces the link bandwidth consumption and [CPU](#) use when a router requests a route refresh. Filters that the routers should use with the route refresh are described in ORF entries that are part of the route refresh message.

You can configure the ORF feature with send, receive, or send and receive capabilities. The local peer advertises the ORF capability in the send mode. The remote peer receives the ORF capability in receive mode and applies the filter as the outbound policy. The local and remote peers exchange updates to maintain the ORF for each router. Peer routers exchange updates depending on the ORF prefix list capability that is advertised. The remote peer starts sending updates to the local peer after it receives a route refresh request or an ORF prefix list with an immediate status.

Inbound vs. Outbound Filtering Example

The figure illustrates the comparison between standard inbound filtering and outbound route filtering.



Outbound route filtering can limit the number of unwanted routing updates, which will reduce the amount of resources that are required for routing update generation and processing. This feature also reduces the amount of resources that are required to receive and discard routes. The receiving router would otherwise filter out these routes, if the [ORF](#) feature were not available.

The example shows two scenarios:

- The first example shows that 600,000 routes are sent from AS1 to AS2 neighbor, and the input filter on AS2 permits only 100 of these routes.
- The second example shows how a route refresh with a filter is sent from AS2 to the AS1 neighbor. The AS1 neighbor then uses the filter before sending the updates. This way, only 100 updates are sent from AS1 to AS2.

BGP Prefix-Based Outbound Route Filtering

The [BGP](#) prefix-based outbound route filtering feature uses BGP [ORF](#) send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers. Configuring this feature can help reduce the amount of system resources that are required for generating and processing routing updates by filtering out unwanted routing updates at the source.

BGP Prefix-Based Outbound Route Filtering

- Uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers
- Helps to reduce the amount of system resources that are required for generating and processing routing updates by filtering out unwanted routing updates at the source
- Limits the number of unwanted routing updates, which will reduce the amount of resources that are required for routing update generation and processing
- Reduces the amount of resources that are queried to receive and discard routes that would otherwise be filtered out

© 2015 Cisco and/or its affiliates. All rights reserved.

211

The BGP prefix-based outbound route filtering feature is enabled through the advertisement of ORF capabilities to peer routers. The advertisement of the ORF capability indicates that a BGP speaker will accept a prefix list from a neighbor. It will apply the prefix list to locally configured ORFs (if any exist). When this capability is enabled, the BGP speaker can install the inbound prefix list filter to the remote peer as an outbound filter, which reduces unwanted routing updates.

The BGP prefix-based outbound route filtering feature can be configured with send, receive, or send and receive ORF capabilities. The local peer advertises the ORF capability in send mode. The remote peer receives the ORF capability in receive mode and applies the filter as an outbound policy. The local and remote peers exchange updates to maintain the ORF on each router. Updates are exchanged between peer routers by address family depending on the ORF prefix list capability that is advertised. The remote peer starts sending updates to the local peer after a route refresh has been configured with the **clear ip bgp** command. Or after an ORF prefix list with immediate status is processed. The BGP speaker will continue to apply the inbound prefix list to received updates after the speaker pushes the inbound prefix list to the remote peer.

Outbound Route Filter Message

It is important to understand the fields [ORF](#) message.

Outbound Route Filter Message

ORF format

- An ORF message consists of the following fields:
 - AFI/SAFI
 - ORF type
 - When to refresh
 - List of ORF entries
- ORF entries depend on the ORF type.
- The ORF capability needs to be negotiated for every supported ORF type.

© 2015 Cisco and/or its affiliates. All rights reserved.

212

An ORF message contains the following information:

- [AFI](#) and [SAFI](#), for which the filter should be used.
- ORF type, which identifies the type of filter.
- When to refresh (immediate or deferred refresh).
- List of ORF entries where the actual filter is defined.

You can use the AFI/SAFI component of the ORF message to provide a coarse level of granular control by limiting the ORF to only the routes whose [NLRI](#) matches the configured AFI/SAFI component.

The router has to negotiate the ORF capability for each ORF type that is supported in the ORF message.

Outbound Route Filter Message (Cont.)

ORF types:

- NLRI (ORF type = 1)
 - Filters that are based on the prefix
- Communities (ORF type = 2)
 - Filters that are based on standard BGP community attributes
- Extended communities (ORF type = 3)
 - Filters that are based on extended BGP community attributes
- Prefix list (ORF type = 128)
 - Filters that are based on Cisco implementation of prefix filtering

© 2015 Cisco and/or its affiliates. All rights reserved.

215

The value that is contained in the ORF type determines the content that is contained in the ORF message.

Currently, ORF type 0 is reserved, ORF types 1 to 127 are assigned by the [IANA](#), and ORF types 128 to 255 are vendor-specific (and not assigned by the IANA). Commonly used ORF types are as follows:

- ORF type 1 is used to filter based on the NLRI.
- ORF type 2 is used to filter based on standard [BGP](#) community attributes.
- ORF type 3 is used to filter based on extended BGP community attributes.
- ORF type 128 is used to filter based on the Cisco proprietary implementation of prefix filtering (prefix lists).

Outbound Route Filter Message (Cont.)

AFI/SAFI is IPv4 unicast.

ORF type is NLRI:

- **Action:** ADD, DELETE, or DELETE ALL
- **Match:** PERMIT or DENY
- **Scope:** EXACT or REFINE
- **NLRI:** Prefix
- **When:** IMMEDIATE or DEFER

© 2015 Cisco and/or its affiliates. All rights reserved.

214

The ORF-type setting determines the content of the ORF value. An ORF type of NLRI-based filtering (type 1) uses the following actions:

- **ADD:** Adds a line to a prefix list filter on the remote peer
- **DELETE:** Removes a line from a filter that was previously installed on a remote peer
- **DELETE ALL:** Removes all previously installed filters on the remote peer

For each filter entry, there is a match component that specifies either PERMIT or DENY. A PERMIT asks the peer to send updates with routes that match the set of entries as specified in the ORF. DENY specifies that the remote peer should not send updates for the entries matching the entries that are specified in the ORF.

For prefixes specified with a match component of PERMIT, the remote peer is asked to pass a prefix with a scope of EXACT (an exact match) or REFINE (its subnets).

Also contained within the ORF message is the when-to-refresh field. A router can set this field to IMMEDIATE—asking the remote peer to refresh as soon as it has finished processing the ORF message. Or it can set it to DEFER—asking the remote peer to wait until it receives a subsequent route refresh message with the same AFI/SAFI.

Configuring Outbound Route Filtering

You need to enable [ORF](#) negotiations and activate an ORF prefix list when configuring outbound route filtering.

Configuring Outbound Route Filtering

```
router(config-router)# neighbor ip-address capability orf prefix-list  
[receive|send|both]
```

- This command enables negotiation of prefix list ORF capability during session setup.
- The ORF-capable BGP speaker will install ORFs per neighbor.
- **Option:**
 - "Both" allows sending and receiving of prefix lists.
 - "Send" allows only sending of prefix lists.
 - "Receive" allows only receiving of prefix lists.

© 2015 Cisco and/or its affiliates. All rights reserved.215

Cisco routers support the uploading of their prefix lists to a neighbor. You need to use the **neighbor ip-address capability orf prefix-list receive** command to advertise this capability. You need to use **neighbor ip-address capability orf prefix-list send** command to upload the inbound prefix filter to the neighbor. The uploaded filter is then used on the neighboring router after a statically configured outbound prefix list (if it exists) is applied.

The **neighbor ip-address capability orf prefix-list** command enables the negotiation of the prefix-list ORF capability during [BGP](#) session setup. The prefix-list-based ORF (ORF type = 128) is the only ORF type that Cisco IOS software supports.

neighbor orf prefix-list Command

To advertise ORF capabilities to a peer router, use the **neighbor orf prefix-list** command in address family or router configuration mode:

```
neighbor {ip-address} [capability] orf prefix-list [receive | send | both]
```

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

```
no neighbor {ip-address} [capability] orf prefix-list [receive | send | both]
```

Syntax Description

Parameter	Description
<i>ip-address</i>	The IP address of the neighbor router
capability	(Optional) Informs the specified neighbor that this router has ORF capabilities
receive	(Optional) Enables the ORF prefix list capability in receive mode
send	(Optional) Enables the ORF prefix list capability in send mode
both	(Optional) Enables the ORF prefix list capability in both receive and send modes

Configuring Outbound Route Filtering (Cont.)



```
AS1:
router bgp 100
 neighbor 172.16.12.2 capability orf prefix-list receive
AS2:
ip prefix-list FILTER seq 5 permit 10.0.1.16/28
!
router bgp 1
 neighbor 172.16.12.11 capability orf prefix-list send
 neighbor 172.16.12.11 prefix-list FILTER in
!
```

- The command **capability orf prefix-list send** on one router requires **capability orf prefix-list receive** on a neighboring router.

© 2015 Cisco and/or its affiliates. All rights reserved. 216

The example shows the configuration of two routers where one router has uploaded an input prefix list to the neighbor to be used as an output filter.

The following configuration steps are necessary to enable outbound route filtering:

1. Enable negotiation of outbound filtering based on prefix lists.
2. Attach an input prefix list to a neighbor.
3. Enable sending of input prefix list to the neighbor.

In the example, an ORF has been configured on AS2 to advertise the filter to AS1. An IP prefix list with the name FILTER is created to specify the 10.0.1.16/28 subnet for outbound route filtering. The ORF send capability is configured on AS2 so that AS2 can advertise the ORF to AS1.

AS1 is configured to advertise the ORF receive capability to AS2. AS1 will install the ORF, defined in the FILTER prefix list, after the ORF capabilities have been exchanged. An inbound soft reset needs to be initiated on AS1 at the end of this configuration to activate the ORF.

Using Outbound Route Filtering

Identify the Cisco IOS command that is used to trigger a route refresh message.

Using Outbound Route Filtering

```
router# clear ip bgp neighbor in [prefix-filter]
```

- This command triggers a route refresh message.
- This command includes a prefix list in the route refresh message if configured and supported on both ends.
- The prefix list is sent at session setup.
- Use the **prefix-filter** option to refresh the remote filter.

© 2015 Cisco and/or its affiliates. All rights reserved.217

Use the **clear ip bgp neighbor** command with the **prefix-filter** keyword to push out the existing [ORF](#) prefix list so that a new route refresh will be received from a neighbor. The neighbor will use the ORF prefix list that was previously negotiated.

You need to use the **clear ip bgp neighbor** command only when the filter has been modified because the neighbor will store the filter for subsequent route refresh requests. The neighbor will then use the filter on all updates toward the router that originated the filter.

Note	You should enter the in keyword when you are using the clear ip bgp neighbor command because inbound route refresh is desired; only the inbound prefix list filter is pushed to the neighbor and used by the neighbor in the outbound direction.
-------------	--

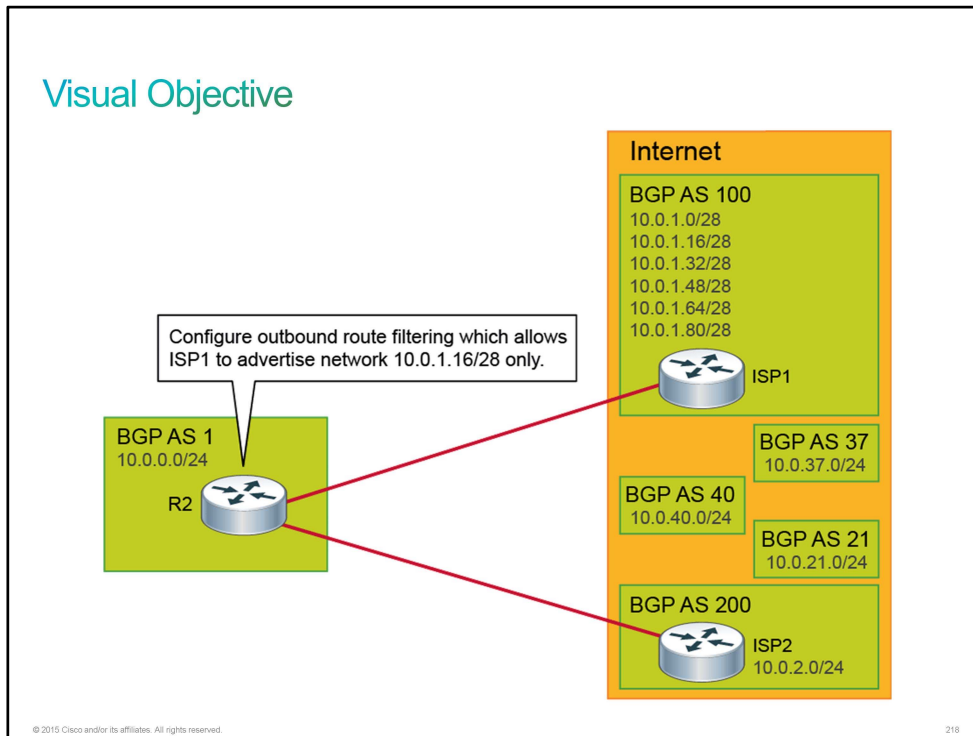
The router will ignore the **prefix-filter** keyword if ORF capability has not been received, or the send capability has not been enabled.

When the **clear ip bgp neighbor** command is used without the **prefix-filter** keyword, a normal route refresh is performed. You should always use the **prefix-filter** keyword when ORF inbound routing policy changes occur.

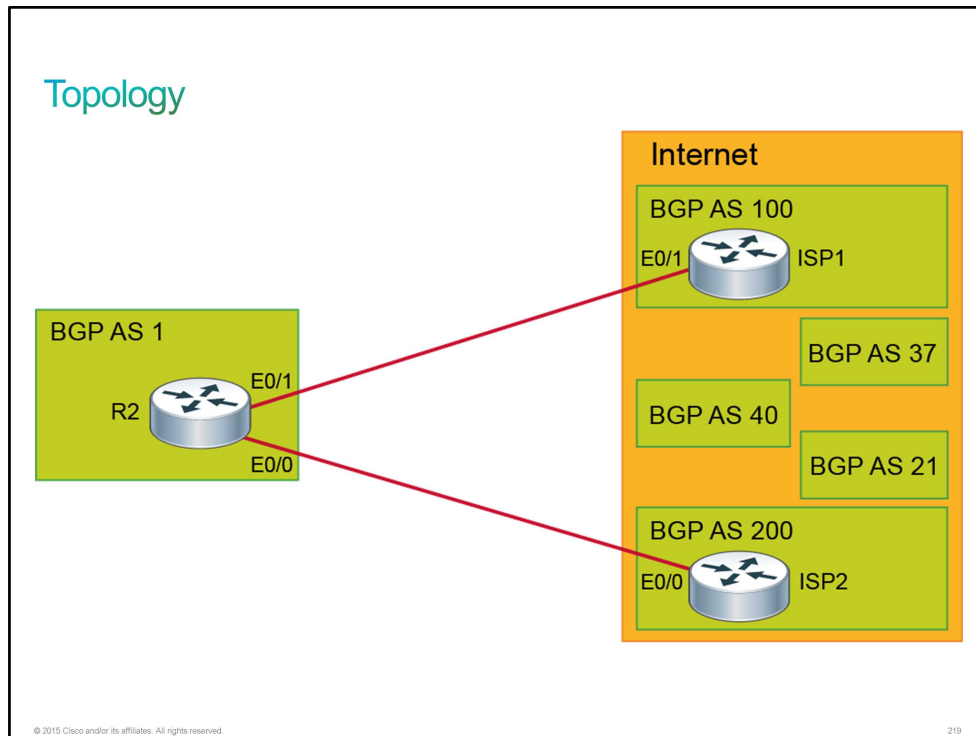
Discovery 9: Prefix-Based Outbound Route Filtering

Overview

Through this discovery, you will learn how to configure prefix-based outbound route filtering. You will create an outbound route filter and configure the R2 router to advertise the filter to the ISP1 router. The filter will allow the ISP1 router to advertise 10.0.1.16/28 network only to the R2 router.



Topology



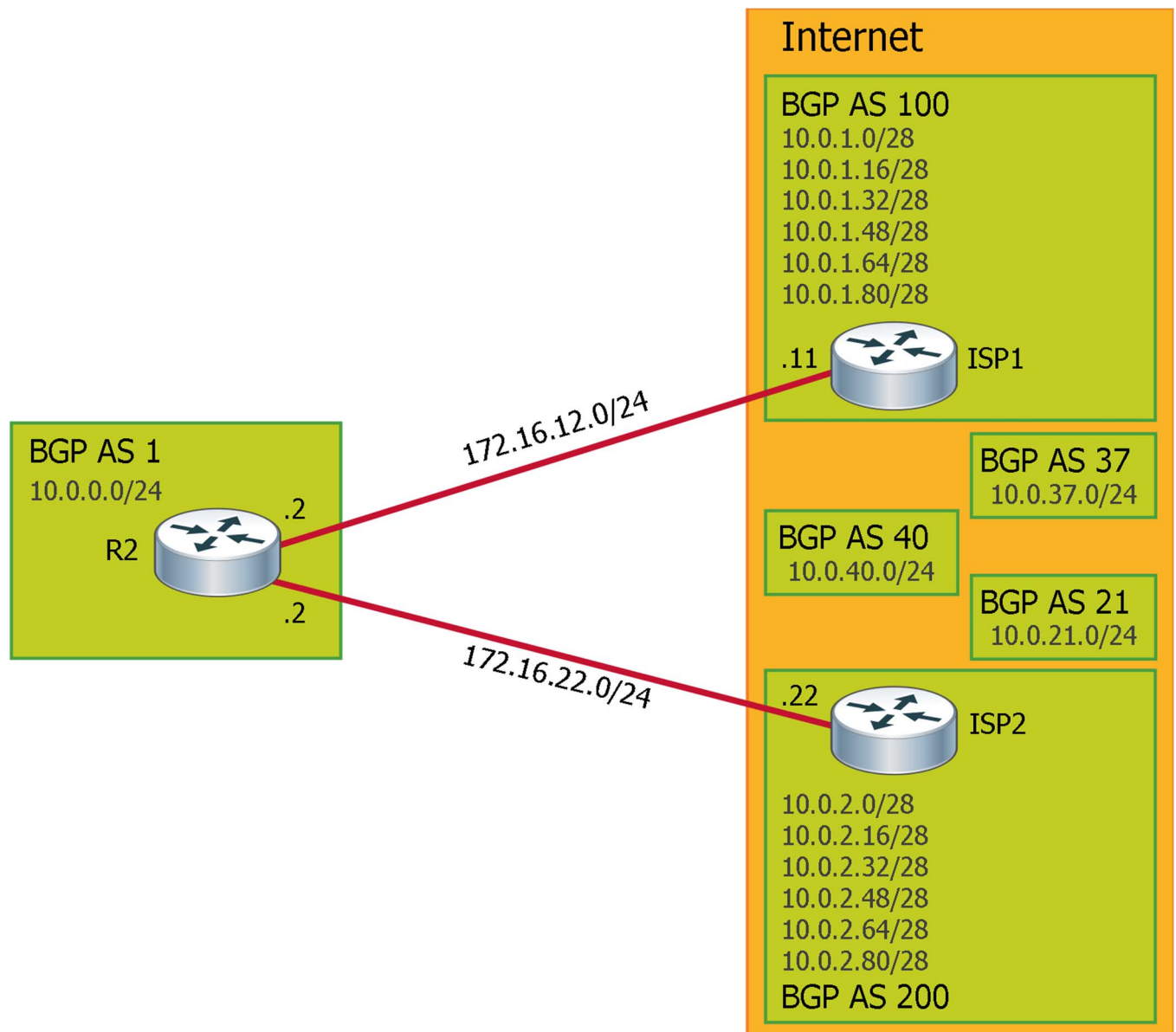
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP1	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate extra networks in different autonomous systems.
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks
ISP2	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate extra networks in different autonomous systems.

IP addresses and advertised networks in BGP are preconfigured as shown below:



BGP is also preconfigured as [EBGP](#) (R2 to ISP1 and R2 to ISP2). The ISP1 router announces these networks:

- 10.0.1.0/28 (assigned address space to the Customer)
- 10.0.1.16/28
- 10.0.1.32/28
- 10.0.1.48/28
- 10.0.1.64/28
- 10.0.1.80/28

Prefix-Based Outbound Route Filtering

Discovery Steps

Step 1 On the ISP1 router, verify which networks are advertised via BGP to the R2 router.

```
ISP1# show ip bgp neighbor 172.16.12.2 advertised-routes
BGP table version is 15, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i

You should see these BGP routes being advertised to the R2 router:

- Any subnets in 10.0.1.0/24 range.
- Networks 10.0.21.0/24, 10.0.37.0/24 and 10.0.40.0/24.

Step 2 On the R2 router, verify which routes are received via BGP from the ISP1 router.

```

R2# show ip bgp
BGP table version is 162, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.32/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.0.64/28	0.0.0.0	0		32768	i
*>	10.0.1.0/28	172.16.12.11	0		0 100	i
*>	10.0.1.16/28	172.16.12.11	0		0 100	i
*>	10.0.1.32/28	172.16.12.11	0		0 100	i
*>	10.0.1.48/28	172.16.12.11	0		0 100	i
*>	10.0.1.64/28	172.16.12.11	0		0 100	i
*>	10.0.1.80/28	172.16.12.11	0		0 100	i
*>	10.0.2.0/28	172.16.22.22	0		0 200	i
*>	10.0.2.16/28	172.16.22.22	0		0 200	i
*>	10.0.2.32/28	172.16.22.22	0		0 200	i
	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.2.48/28	172.16.22.22	0		0 200	i
*>	10.0.2.64/28	172.16.22.22	0		0 200	i
*>	10.0.2.80/28	172.16.22.22	0		0 200	i
*>	10.0.2.96/28	172.16.22.22	0		0 200	i
*>	10.0.2.112/28	172.16.22.22	0		0 200	i
*>	10.0.2.128/28	172.16.22.22	0		0 200	i
*	10.0.21.0/24	172.16.12.11	0		0 100 37 40 21	i
*>		172.16.22.22	0		0 200 21	i
*>	10.0.37.0/24	172.16.12.11	0		0 100 37	i
*		172.16.22.22	0		0 200 21 40 37	i
*>	10.0.40.0/24	172.16.12.11	0		0 100 37 40	i
*		172.16.22.22	0		0 200 21 40	

You should see these BGP routes being received on the R2 router from the ISP1 router:

- Any subnets in 10.0.1.0/24 range.
- Networks 10.0.21.0/24, 10.0.37.0/24 and 10.0.40.0/24.

Step 3 Configure the ISP1 router to advertise [ORF](#) receive capability from the R2 neighbor.

```

ISP1(config)# router bgp 100
ISP1(config-router)# neighbor 172.16.12.2 capability orf prefix-list receive

```

Step 4 Configure the R2 router to advertise ORF send capability to the ISP1 router.

```

R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 capability orf prefix-list send

```

Step 5 On the R2 router, configure the prefix list with the name FILTER, which will only permit 10.0.1.16/28 network. This prefix list will be used for prefix-based outbound route filtering.

```

R2(config)# ip prefix-list FILTER seq 5 permit 10.0.1.16/28

```

Step 6 On the R2 router, apply the "FILTER" prefix list to incoming advertisements from the ISP1 neighbor.

```
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 prefix-list FILTER in
```

Step 7 Clear all BGP sessions on the router ISP1.

```
ISP1# clear ip bgp *
```

Step 8 On the ISP1 router, verify that the ORF capability exchange is enabled.

```
ISP1# show ip bgp neighbors 172.16.12.2
BGP neighbor is 172.16.12.2, remote AS 1, external link
  BGP version 4, remote router ID 10.0.0.65
  BGP state = Established, up for 00:15:44
  Last read 00:00:27, last write 00:00:09, hold time is 180, keepalive interval is 60
seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
    Multisession Capability:
    Stateful switchover support enabled: NO for session 1
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

      Sent      Rcvd
  Opens:          1        1
  Notifications:    0        0
  Updates:         2        2
  Keepalives:      18       19
  Route Refresh:    0        0
  Total:          21       25
  Default minimum time between advertisement runs is 30 seconds

For address family: IPv4 Unicast
  Session: 172.16.12.2
  BGP table version 15, neighbor version 15/0
  Output queue size : 0
  Index 28, Advertise bit 0
  28 update-group member
  AF-dependant capabilities:
    Outbound Route Filter (ORF) type (128) Prefix-list:
    Send-mode: received
    Receive-mode: advertised
<...output omitted...>
```

You should see that the ORF send mode capability was received and the ORF receive mode capability was advertised.

Step 9 On the ISP1 router, verify that it has received an IP prefix list that defines an outbound route filter for the 10.0.1.16/28 subnet.


```
ISP1# show ip bgp neighbors 172.16.12.2 received prefix-filter
Address family: IPv4 Unicast
ip prefix-list 172.16.12.2: 1 entries
seq 5 permit 10.0.1.16/28
```

Step 10 On the ISP1 router, verify that BGP is advertising 10.0.1.16/28 subnet only to the R2 router.

```
ISP1# show ip bgp neighbor 172.16.12.2 advertised-routes
BGP table version is 16, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.1.16/28	0.0.0.0	0		32768	i

Step 11 On the R2 router, verify that only 10.0.1.16/28 was received from the ISP1 router.

```
R2# show ip bgp
BGP table version is 137, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.0.0/24	0.0.0.0	0		32768	i
*> 10.0.0.16/28	0.0.0.0	0		32768	i
*> 10.0.0.32/28	0.0.0.0	0		32768	i
*> 10.0.0.48/28	0.0.0.0	0		32768	i
*> 10.0.0.64/28	0.0.0.0	0		32768	i
*> 10.0.1.16/28	172.16.12.11	0		0	100 i
*> 10.0.2.0/28	172.16.22.22	0		0	200 i
*> 10.0.2.16/28	172.16.22.22	0		0	200 i
*> 10.0.2.32/28	172.16.22.22	0		0	200 i
*> 10.0.2.48/28	172.16.22.22	0		0	200 i
*> 10.0.2.64/28	172.16.22.22	0		0	200 i
*> 10.0.2.80/28	172.16.22.22	0		0	200 i
*> 10.0.2.96/28	172.16.22.22	0		0	200 i
*> 10.0.2.112/28	172.16.22.22	0		0	200 i
Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.2.128/28	172.16.22.22	0		0	200 i
*> 10.0.21.0/24	172.16.22.22	0		0	200 21 i
*> 10.0.37.0/24	172.16.22.22	0		0	200 21 40 37 i
*> 10.0.40.0/24	172.16.22.22	0		0	200 21 40 i

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- The BGP prefix-based outbound route filtering feature uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers.
- Configuring ORF helps you to reduce the amount of system resources that are required for generating and processing routing updates by filtering out unwanted routing updates at the source.
- The ORF message contains the information that is used to determine which updates will be passed.
- The **neighbor ip-address capability orf prefix-list** command with the **send** and **receive** keywords enables ORF negotiations and activates an ORF prefix-list.

© 2015 Cisco and/or its affiliates. All rights reserved.

220

Summary (Cont.)

- Use the **clear ip bgp neighbor** command to trigger a BGP route refresh.
- With the **show ip bgp neighbors** command, neighbor-supported ORF capabilities are displayed as "advertised, received," and a filter of the supported type can be sent to the neighbor.
- You can verify that prefix list was received on ORF receiver using the **show ip bgp neighbors ip-address received prefix** command.

© 2015 Cisco and/or its affiliates. All rights reserved.

221

Lesson 5: Applying Route Maps as BGP Filters

Overview

[BGP](#) is a powerful routing protocol that supports a wide variety of administrative policy controls and route selection features. You cannot achieve many complex filtering goals and administrative policies by using only single-purpose filtering methods or by compounding multiple filtering methods together. Route maps provide a method to perform various compound, complex filtering operations within a single tool. Understanding the operation and use of route maps is a critical component in the successful implementation of any large-scale BGP deployment.

In this lesson, you will learn about route maps and how you can use them for BGP filtering. You will learn the commands that are required to use route maps with prefix lists. You will also learn how to use prefix lists as BGP filters and how to monitor previously configured route maps.

Upon completing this lesson, you will be able to:

- Describe the high-level function of a route-map
- Describe the function of the BGP Route-Map Policy List Support feature
- Describe the function of the BGP Route-Map Continue feature
- Identify the Cisco IOS commands that are required to configure a route map to match against a prefix list
- Identify where you can apply route maps as route filters in a BGP network
- Identify the Cisco IOS command that is required to enable a route map as a BGP route filter
- Identify the Cisco IOS commands that are required to monitor the operation of a configured route map that is used as a BGP filter

Route Map Overview

A route map is a filter, and what the route map denies is dropped. Additionally, you can use the route map to modify attributes of the permitted routes.

Route Map Overview

Route maps are very complex access lists:

- Access lists have lines.
 - Route maps contain statements.
- Access lists use addresses and masks.
 - Route maps use match conditions.
- With access lists, there is an access list number.
 - With route maps, there is a route map name.
- Statements in route maps are numbered.
 - You can insert and delete statements in a route map.
 - You can edit match conditions in a statement.
- Route map statements can modify matched routes with "set" options.

© 2015 Cisco and/or its affiliates. All rights reserved.

222

Route maps are similar to access lists. Both have a set of tests to be performed, and several tests can be done in sequence. The first match produces the result of either "permit" or "deny."

An access list has a number of lines, each indicating a testing condition. The route map is more complex than the access list. The route map consists of several groups of configuration lines; each group is called a statement. The statement has a sequence number that provides the opportunity to remove or modify an explicit statement without removing the entire route map. There is also an opportunity to add a new statement between two existing statements.

Each route map statement starts with a configuration line indicating the name of the route map, the sequence number, and whether the result should be permitted or denied if the testing matches. The statement then continues, following configuration lines with the match clauses. Matching can be done in several ways: testing on the prefix, the AS path, or some other attribute. The statement concludes with optional "set" statements, where attributes may be modified or set.

Route Map Overview (Cont.)

```
route-map name [permit|deny sequence]
match condition
match condition
set parameter
```

- The default statement action is "permit."
- A route that is not matched by any statement is dropped.
- "Permit all" is achieved by specifying "permit" without a "match" clause.
- Match conditions in one statement are ANDed together.
- The first matching statement permits or denies the route.

© 2015 Cisco and/or its affiliates. All rights reserved.

223

A route map consists of several statements. Each statement starts with the route map configuration line, on which the name of the route map must be indicated. A good practice is to always indicate the **permit** or **deny** keyword followed by a sequence number.

The matching clauses for the statements are listed on the match lines following the route map line. There may be several match lines, each referring to a different test to be performed. All tests must be passed for the statement to be matched. If any of the match line tests fails, the next route map statement is tried. Statements are tried in sequence number order. If there are no more statements in the route map, the result is, implicitly, "deny."

If all of the match clauses succeed, there is a match for the statement and the indicated result is used. If the result is to deny, the route is then silently ignored. If the result is to permit, the route is accepted and the set clauses are applied. The set clauses allow one or more attributes to be changed or set to specific values before the route is accepted.

Route Map Overview (Cont.)

- Route map conditions are specified in the match statement.
- Route maps can match on:
 - Network number and subnet mask that are matched with an IP prefix list
 - Route originator
 - BGP next-hop address
 - BGP origin
 - Tag attached to IGP route
 - AS path
 - BGP community that is attached to BGP route
 - IGP route type (internal/external ...)

© 2015 Cisco and/or its affiliates. All rights reserved.

224

Each route map statement can have several match clauses, and each match clause is given its own configuration line. The match clause refers to the tests that will be made on the candidate route. Tests of the candidate route can be based on the following criteria:

- IP network numbers and subnet masks, by referring to a prefix list or access list that will be applied on the route.
- Route originator, by referring to a prefix list or access list that will be applied on the value of the originator [BGP](#) attribute.
- Next hop, by referring to a prefix list or access list that will be applied on the value of the next-hop BGP attribute.
- Origin code, by testing the value of the origin BGP attribute.
- Tag value that is attached to an [IGP](#) route—used only when redistribution from an IGP into BGP occurs.
- [AS](#) path, by referring to an AS path access list that will be applied on the value of the AS path BGP attribute.
- Community, by referring to a community list that will be applied on the value of the Community BGP attribute.
- IGP route type, by testing if the IGP route is internal or external—used only when redistribution from an IGP to BGP occurs.

Route Map Overview (Cont.)

- Route maps can also change the attributes of BGP routes.
- Route maps can set:
 - Origin
 - BGP next-hop
 - Weight
 - BGP community
 - Local preference
 - MED

© 2015 Cisco and/or its affiliates. All rights reserved.

225

Each route map statement may have several set clauses. Each set clause is applied to the route when the route map statement permits the route. With a route map, the following can be set:

- Origin BGP attribute
- Next-hop BGP attribute
- Weight
- Community BGP attribute
- Local preference BGP attribute
- MED BGP attribute, by setting the metric

BGP Route-Map Policy List Support

The BGP Route-Map Policy List Support feature introduces new functionality to [BGP](#) route maps, allowing you to group route map match clauses into named lists—policy lists. A policy list functions like a macro.

BGP Route-Map Policy List Support

- Adds the capability for a network operator to group route map match clauses into named lists that are called policy lists.
- Simplifies the configuration of BGP routing policy in medium-size and large networks. You can preconfigure policy lists with groups of match clauses and then reference these policy lists within different route maps.
- Eliminates need to manually reconfigure each recurring group of match clauses that occur in multiple route map entries.

© 2015 Cisco and/or its affiliates. All rights reserved.

226

When a policy list is referenced in a route map, all of the match clauses are evaluated and processed as if they had been configured directly in the route map. The BGP Route Map Policy List Support feature simplifies the configuration of BGP routing policy in medium-size and large networks. A network operator can preconfigure policy lists with groups of match clauses and then reference these policy lists within different route maps. You no longer need to manually reconfigure each recurring group of match clauses in multiple route map entries.

A policy list is like a route map that contains only match clauses. The policy list is created and then referenced within a route map. There are no changes to match clause semantics and route map functions. Match clauses are configured in policy lists with permit and deny statements. The route map evaluates and processes each match clause and permits or denies routes based on the configuration. AND and OR semantics in the route map function the same way for policy lists as they do for match and set clauses. There are some commands that are related to the BGP Route-Map Policy List Support feature: the **ip policy-list** command, the **match policy-list** command, and the **show ip policy-list** command.

BGP Route-Map Policy-List Support (Cont.)

```
router# ip policy-list policy-list-name {permit | deny}
```

- Creates a BGP policy list

```
router# match policy-list policy-list-name
```

- Configures a route map to evaluate and process a BGP policy list in a route map

```
router# show ip policy-list policy-list-name
```

- Displays one or all filter lists

© 2015 Cisco and/or its affiliates. All rights reserved.

227

To create a BGP policy list, use the **ip policy-list** command in the policy map configuration mode.

```
ip policy-list policy-list-name {permit | deny}
```

To remove a policy-list, use the **no** form of this command

```
no ip policy-list policy-list-name
```

Syntax Description

Parameter	Description
<i>policy-list-name</i>	Name of the configured policy list
permit	Permits access for matching conditions
deny	Denies access to matching conditions

To configure a route map to evaluate and process a BGP policy list in a route map, use the **match policy-list** command in the route map configuration mode.

```
match policy-list policy-list-name
```

To remove a path list entry, use the **no** form of this command.

```
no match policy-list policy-list-name
```

Syntax Description

Parameter	Description
<i>policy-list-name</i>	Name of the configured policy list

To display information about a configured policy list and policy list entries, use the **show ip policy-list** command in the user EXEC mode.

```
show ip policy-list policy-list-name
```

Syntax Description

Parameter	Description
<i>policy-list-name</i>	Name of the configured policy list

Configuring Policy List Examples

The following configuration example creates a BGP policy-list that permits matches on the [AS](#) path and [MED](#) of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-1 permit
Router(config-policy-list)# match as-path 1
Router(config-policy-list)# match metric 10
Router(config-policy-list)# end
```

The following configuration example creates a BGP policy list that permits matches on the specified BGP community and the next hop of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-2 permit
Router(config-policy-list)# match community 20
Router(config-policy-list)# match metric 10
Router(config-policy-list)# ip community-list 20 permit 20:1
Router(config-policy-list)# end
```

The following configuration example creates a BGP policy list that denies matches on the specified BGP community and the next hop of a router:

```
Router(config)# ip policy-list POLICY-LIST-NAME-3 deny
Router(config-policy-list)# match community 20
Router(config-policy-list)# match metric 10
Router(config-policy-list)# end
```

Configuring Route Maps to Reference Policy List Examples

The configuration examples in this section create BGP route maps that reference BGP policy lists with the **route-map** route map configuration command.

The following configuration example creates a route map that references policy lists and separate match and set clauses in the same configuration. This example uses AND semantics between POLICY-LIST-NAME-1 and POLICY-LIST-NAME-2.

```
Router(config)# route-map MAP-NAME-1 10
Router(config-route-map)# match ip-address 1
Router(config-route-map)# match policy-list POLICY-LIST-NAME-1
Router(config-route-map)# match policy-list POLICY-LIST-NAME-2
Router(config-route-map)# set community 10:1
Router(config-route-map)# set local-preference 140
Router(config-route-map)# end
```

The following configuration example creates a route map that references policy lists and separate match and set clauses in the same configuration. This example uses OR semantics between POLICY-LIST-NAME-3 and POLICY-LIST-NAME-4.

```
Router(config)# route-map MAP-NAME-2 10
Router(config-route-map)# match policy-list POLICY-LIST-NAME-3 POLICY-LIST-NAME-4
Router(config-route-map)# set community 10:1
Router(config-route-map)# set local-preference 140
Router(config-route-map)# end
```

Verifying BGP Route-Map Policy List Support

To verify that a policy list has been created, use the **show ip policy-list** command. The output of this command displays the policy list name and configured match clauses. The following sample output is similar to the output that will be displayed:

```
Router# show ip policy-list
policy-list POLICY-LIST-NAME-1 permit
  Match clauses:
    metric 20
policy-list POLICY-LIST-NAME-2 permit
  Match clauses:
    as-path (as-path filter): 1
```

Note	A policy list name can be specified when the show ip policy-list command is entered. This option can be useful for filtering the output of this command and verifying a single policy list
-------------	---

To verify that a route map has been created and a policy list is referenced, use the **show route-map** command. The output of this command displays the route map name and policy-lists that the configured route maps reference. The following sample output is similar to the output that will be displayed:

```
Router# show route-map
route-map ROUTE-MAP-NAME-1, deny, sequence 10
  Match clauses:
  Set clauses:
  Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME-1, permit, sequence 10
  Match clauses:
    IP Policy lists:
      POLICY-LIST-NAME-1
  Set clauses:
  Policy routing matches: 0 packets, 0 bytes
```

BGP Route Map Continue

The BGP Route-Map Continue feature introduces the continue clause to the BGP route map configuration. The continue clause provides more programmable policy configuration and route filtering. It introduces the ability to execute extra entries in a route map after an entry is executed with successful match and set clauses. Continue clauses allow you to configure and organize more modular policy definitions to reduce the number of policy configurations that are repeated within the same route map.

BGP Route-Map Continue

- Introduces the continue clause to the BGP route map configuration, providing more programmable policy configuration and route filtering
- Enables you to execute extra entries in a route map after an entry is executed with successful match and set clauses
- Allows configuration and organization of more modular policy definitions to reduce the number of policy configurations that are repeated within the same route map
- Allows modularization of network policy configuration so that repeated policy definitions can be reduced within the same route map

© 2015 Cisco and/or its affiliates. All rights reserved.

238

Continue clauses provide a programmable method to organize and control the flow of a route map. Route map configuration was linear before this feature was introduced. Continue clauses also allow you to modularize network policy configuration so that repeated policy definitions can be reduced within the same route map.

Route Map Operation Without Continue Clauses

A route map evaluates match clauses until a successful match occurs. After the match occurs, the route map stops evaluating match clauses and starts executing set clauses, in the order in which they were configured. If a successful match does not occur, the route map "falls through" and evaluates the next sequence number of the route map. It repeats until all configured route map entries have been evaluated or a successful match occurs. Each route map sequence is tagged with a sequence number to identify the entry. Route map entries are evaluated in order, starting with the lowest sequence number and ending with the highest sequence number. If the route map contains only set clauses, the set clauses are executed automatically, and the route map does not evaluate any other route map entries.

Route Map Operation with Continue Clauses

When a continue clause is configured, the route map continues to evaluate and execute match clauses in the specified route map entry after a successful match occurs. The continue clause can be configured to go to (or jump to) a specific route map entry by specifying the sequence number. Or, if a sequence number is not specified, to go to the next sequence number. This behavior is called an "implied continue." If a match clause exists, the continue clause is executed only if a match occurs. If no successful matches occur, the continue clause is ignored.

If a match clause does not exist in the route map entry but a continue clause does, the continue clause is automatically executed and goes to the specified route map entry. If a match clause exists in a route map entry, the continue clause is executed only when a successful match occurs. When a successful match occurs and a continue clause exists, the route map executes the set clauses and then goes to the specified route map entry. If the next route map contains a continue clause, the route map executes the continue clause if a successful match occurs. If a continue clause does not exist in the next route map, the route map is evaluated normally. If a continue clause exists in the next route map but a match does not occur, the route map does not continue. It falls through to the next sequence number, if one exists.

Note	A continue clause can be executed without a successful match if a route map entry does not contain a match clause.
-------------	--

BGP Route-Map Continue (Cont.)

```
router# continue sequence-number
```

- Configures a route map to go to a route map entry with a higher sequence number

```
router# show route-map [map-name]
```

- Displays configured route maps

© 2015 Cisco and/or its affiliates. All rights reserved.229

You will use two commands with the BGP Route-Map Continue feature, the **continue** command and the **show route-map** command.

To configure a route map to go to a route map entry with a higher sequence number, use the **continue** command in route map configuration mode.

```
continue sequence-number
```

To remove a continue clause from a route map, use the **no** form of this command.

no continue

Syntax Description

Parameter	Description
<i>sequence-number</i>	(Optional) Route map sequence number. If a route map sequence number is not specified when configuring a continue clause, the continue clause continues to the route map entry with the next sequence number. This behavior is referred to as an "implied continue"

To display the configured route maps, use the **show route-map** command in the EXEC mode.

```
show route-map [map-name]
```

Syntax Description

Parameter	Description
<i>map-name</i>	(Optional) Name of a specific route map

BGP Route-Map Continue Clause Configuration Example

The following example shows the continue clause configuration in a route map sequence.

The first continue clause in route map entry 10 indicates that the route map will go to route map entry 30 if a successful match occurs. If a match does not occur, the route map will fall through to route map entry 20. If a successful match occurs in route map entry 20, the set action will be executed and the route map will not evaluate any additional route map entries.

If a successful match does not occur in route map entry 20, the route map will fall through to route map entry 30. This sequence does not contain a match clause, so the set clause will be automatically executed. The continue clause will go to the next route map entry because a sequence number is not specified.

If there are no successful matches, the route map will fall through to route map entry 30 and execute the set clause, and route map entry 40 will not be evaluated.

```
route-map ROUTE-MAP-NAME permit 10
  match ip address 1
  match metric 10
  set as-path prepend 10
  continue 30
!
route-map ROUTE-MAP-NAME permit 20
  match ip address 2
  match metric 20
  set as-path prepend 10 10
!
route-map ROUTE-MAP-NAME permit 30
  set as-path prepend 10 10 10
  continue
!
route-map ROUTE-MAP-NAME permit 40
  match community 10:1
  set local-preference 104
```

BGP Route-Map Continue Clause Verification Example

To verify the configuration of continue clauses, use the **show route-map** command. The output of this command displays configured route maps, match, set, and continue clauses. The following sample output is similar to the output that will be displayed:

```
Router# show route-map
route-map ROUTE-MAP-NAME, permit, sequence 10
Match clauses:
  ip address (access-lists): 1
  metric 10
Continue: sequence 40
Set clauses:
  as-path prepend 10
Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 20
Match clauses:
  ip address (access-lists): 2
  metric 20
Set clauses:
  as-path prepend 10 10
Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 30
Match clauses:
Continue: to next entry 40
Set clauses:
  as-path prepend 10 10 10
Policy routing matches: 0 packets, 0 bytes
route-map ROUTE-MAP-NAME, permit, sequence 40
Match clauses:
  community (community-list filter): 10:1
Set clauses:
  local-preference 104
Policy routing matches: 0 packets, 0 bytes
route-map LOCAL-POLICY-MAP, permit, sequence 10
Match clauses:
Set clauses:
  community 655370
Policy routing matches: 0 packets, 0 bytes
```

Prefix List Use in Route Maps

You can use prefix lists as match criteria in several matching statements in the route map.

Prefix List Use in Route Maps

```
router(config-route-map)# match ip address prefix-list list-name
```

- Uses prefix list to match routes in route-map match condition

```
router(config-route-map)#match ip next-hop prefix-list list-name
```

- Matches routes where the next hop matches the conditions in the prefix list

```
router(config-route-map)# match ip route-source prefix-list list-name
```

- Matches routes that are received from BGP peer that matches the prefix list

© 2015 Cisco and/or its affiliates. All rights reserved.

230

You can use the **match ip address** command in the route map configuration mode to perform policy routing on packets. You can use the same command to distribute any routes that have a destination network number address that a standard access list, an extended access list, or a prefix list permits.

```
match ip address {access-list-number [access-list-number... | access-list-name...] |  
access-list-name [access-list-number... | access-list-name] | prefix-list prefix-list-  
name [prefix-list-name...]}
```

To remove the **match ip address** entry, use the **no** form of this command.

```
no match ip address {access-list-number [access-list-number... | access-list-name...]  
| access-list-name [access-list-number... | access-list-name] | prefix-list prefix-  
list-name [prefix-list-name...]}
```


Syntax Description

Parameter	Description
<i>access-list-number...</i>	Number of a standard or extended access list. It can be an integer from 1 to 199. The ellipsis indicates that multiple values can be entered.
<i>access-list-name...</i>	Name of a standard or extended access list. It can be an integer from 1 to 199. The ellipsis indicates that multiple values can be entered.
prefix-list	Distributes routes based on a prefix list.
<i>prefix-list-name...</i>	Name of a specific prefix list. The ellipsis indicates that multiple values can be entered.

To redistribute any routes that have a next-hop router address that is passed by one of the specified access lists, use the **match ip next-hop** command in the route map configuration mode.

```
match ip next-hop {access-list-number | access-list-name} [...access-list-number |  
...access-list-name]
```

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

```
no match ip next-hop {access-list-number | access-list-name} [...access-list-number |  
...access-list-name]
```

Syntax Description

Parameter	Description
<i>access-list-number</i>	Number of a standard or extended access-list. It can be an integer from 1 to 199.
<i>access-list-name</i>	Name of a standard or extended access-list. It can be an integer from 1 to 199.

To redistribute routes that have been advertised by routers and access servers at the address that the access lists specify, use the **match ip route-source** command in the route map configuration mode.

```
match ip route-source {access-list-number | access-list-name} [...access-list-number |  
...access-list-name]
```

To remove the route-source entry, use the **no** form of this command.

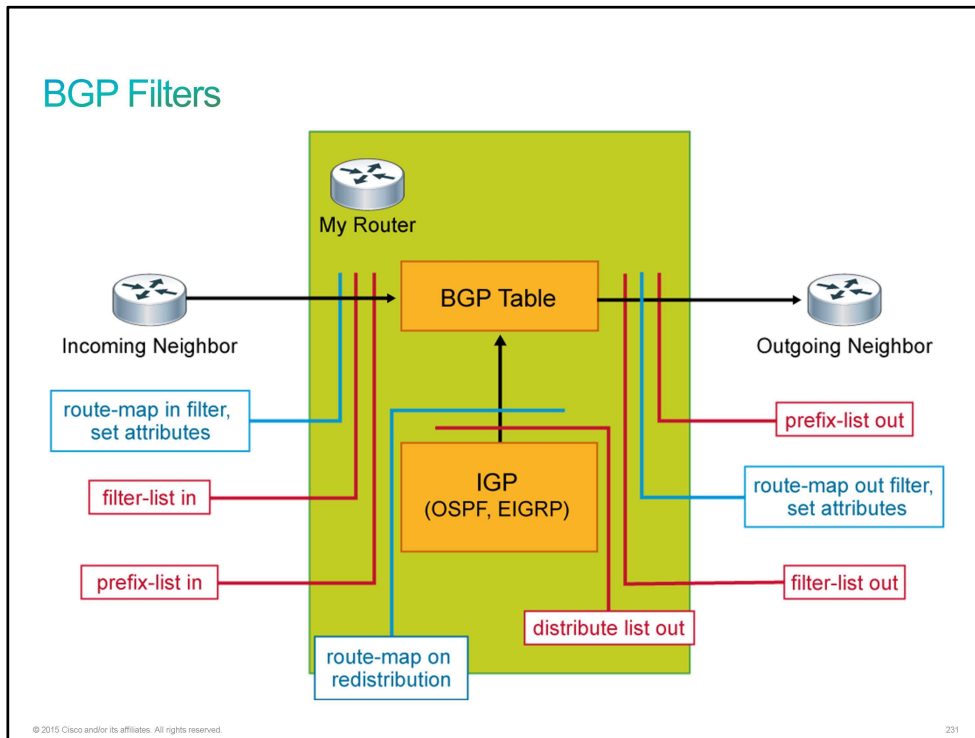
```
no match ip route-source {access-list-number | access-list-name} [...access-list-number  
| ...access-list-name]
```

Syntax Description

Parameter	Description
<i>access-list-number</i>	Number of a standard or extended access list. It can be an integer from 1 to 199.
<i>access-list-name</i>	Name of a standard or extended access list. It can be an integer from 1 to 199.

BGP Filters

You can optionally apply filter lists, prefix lists, and route maps to either incoming or outgoing information or any combination of the two.



The incoming prefix list, the incoming filter list, and the incoming route map must all permit the routes that are received from a neighbor before being accepted into the BGP table. Outgoing routes must pass the outgoing filter list, the outgoing prefix list, and the outgoing route map before being transmitted to the neighbor. When a router is configured to redistribute routing information from an IGP into BGP, the routes must successfully pass any prefix list or route map that is applied to the redistribution. This happens before a route is injected into the BGP table.

Using Route Maps as BGP Filters

You can apply a route map on incoming or outgoing routing information for a neighbor. The route map must permit the routing information in order to be accepted. If there is no statement in the route map explicitly permitting a route, then the route will be implicitly denied and dropped.

Using Route Maps as BGP Filters

```
router(config-router)# neighbor ip-address route-map name [in | out]
```

- This command applies a route map to incoming or outgoing BGP updates.
- Prefixes not permitted by the route map are discarded.
- Route maps can also change BGP attributes in incoming or outgoing updates.
- Route maps, filter lists, and prefix lists are evaluated in sequence (effectively ANDed together).

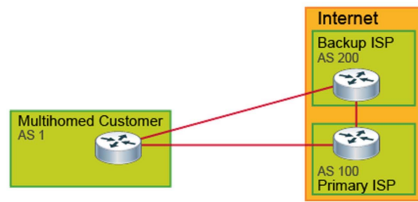
© 2015 Cisco and/or its affiliates. All rights reserved.

232

The permitted routes may have their attributes set or changed by the set clauses in the route map. Setting attributes on routes is useful when you are influencing the route selection. One of the statements in the route map can permit some routes and change their attributes. Another statement in the route map could permit other routes and not alter their attributes. When route selection is performed, the attribute values indicate that one route is preferred over the other.

Using Route Maps as BGP Filters (Cont.)

- **Requirement:** The customer will accept only a default route from both service providers.



```
router bgp 1
 neighbor Primary-ISP remote-as 100
 neighbor Primary-ISP route-map filter in
 neighbor Secondary-ISP remote-as 200
 neighbor Secondary-ISP route-map filter in
!
route-map filter permit 10
 match ip address prefix-list DefaultOnly
!
ip prefix-list DefaultOnly seq 10 permit 0.0.0.0/0
```

© 2015 Cisco and/or its affiliates. All rights reserved.

233

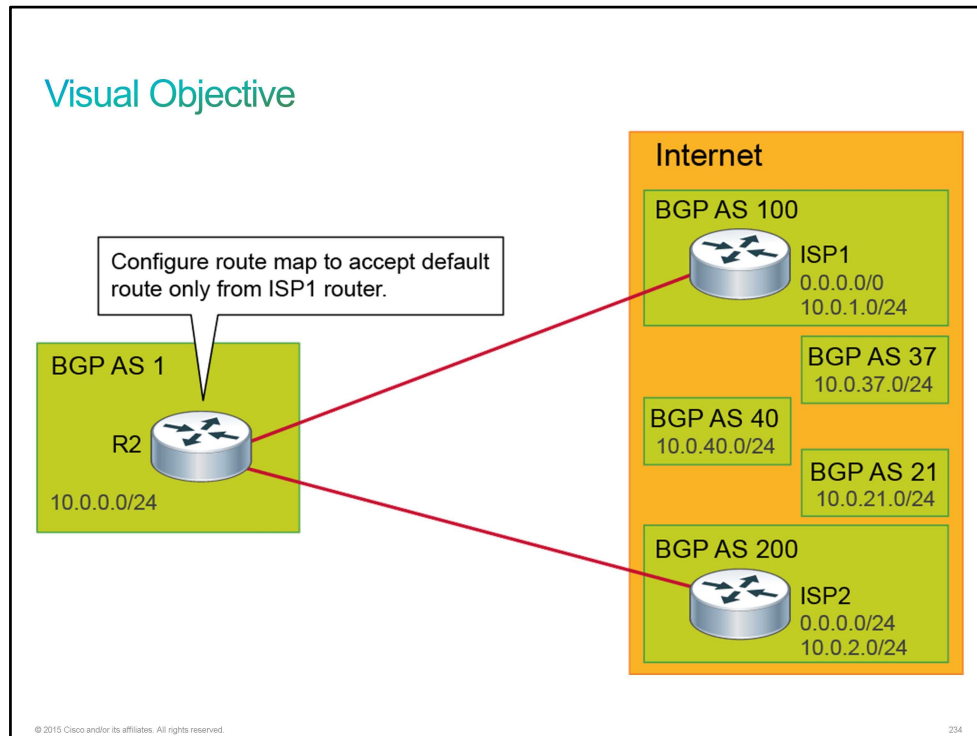
In this example, the customer will accept only a default route from both service providers. The customer does not want to put any other network into the [BGP](#) table than the default route. The default route is announced via BGP to the customer from both service providers.

In the example, the route map with a name *filter* is applied in inbound direction on the customer router for both service provider neighbors. The route map is matching the prefix list, which permits default route only.

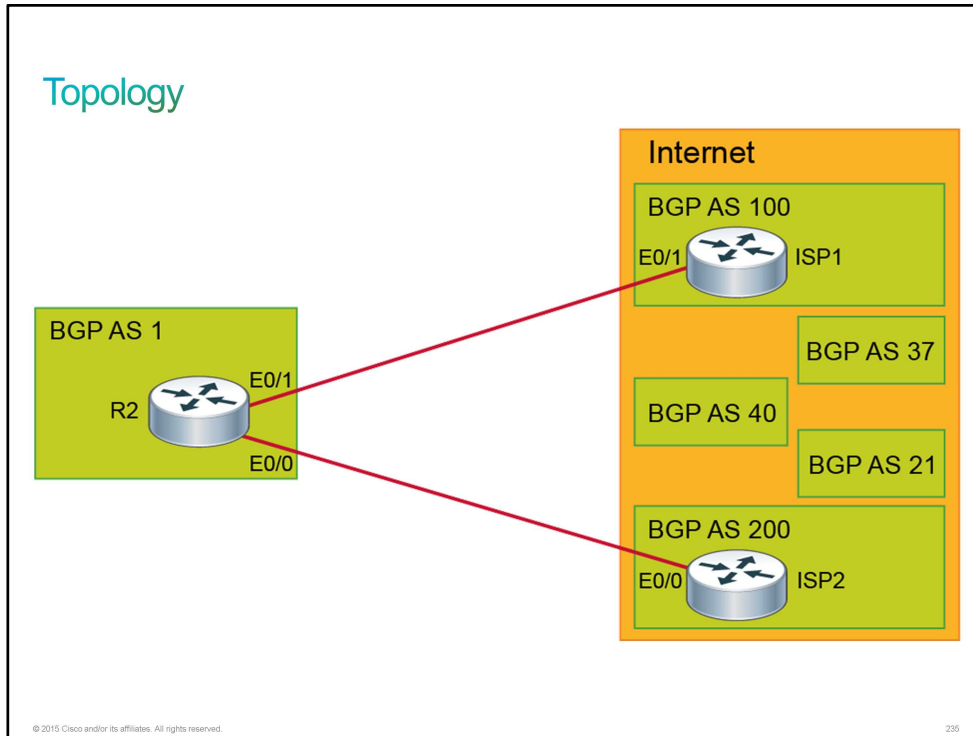
Discovery 10: Configure Route Maps as BGP Filters

Overview

Through this discovery, you will learn how to configure route maps to filter [BGP](#) updates. You as a customer configuring router R2 will accept only the default route that is announced via BGP from the ISP1 router. You will not accept any other networks that are announced via BGP from the service provider.



Topology



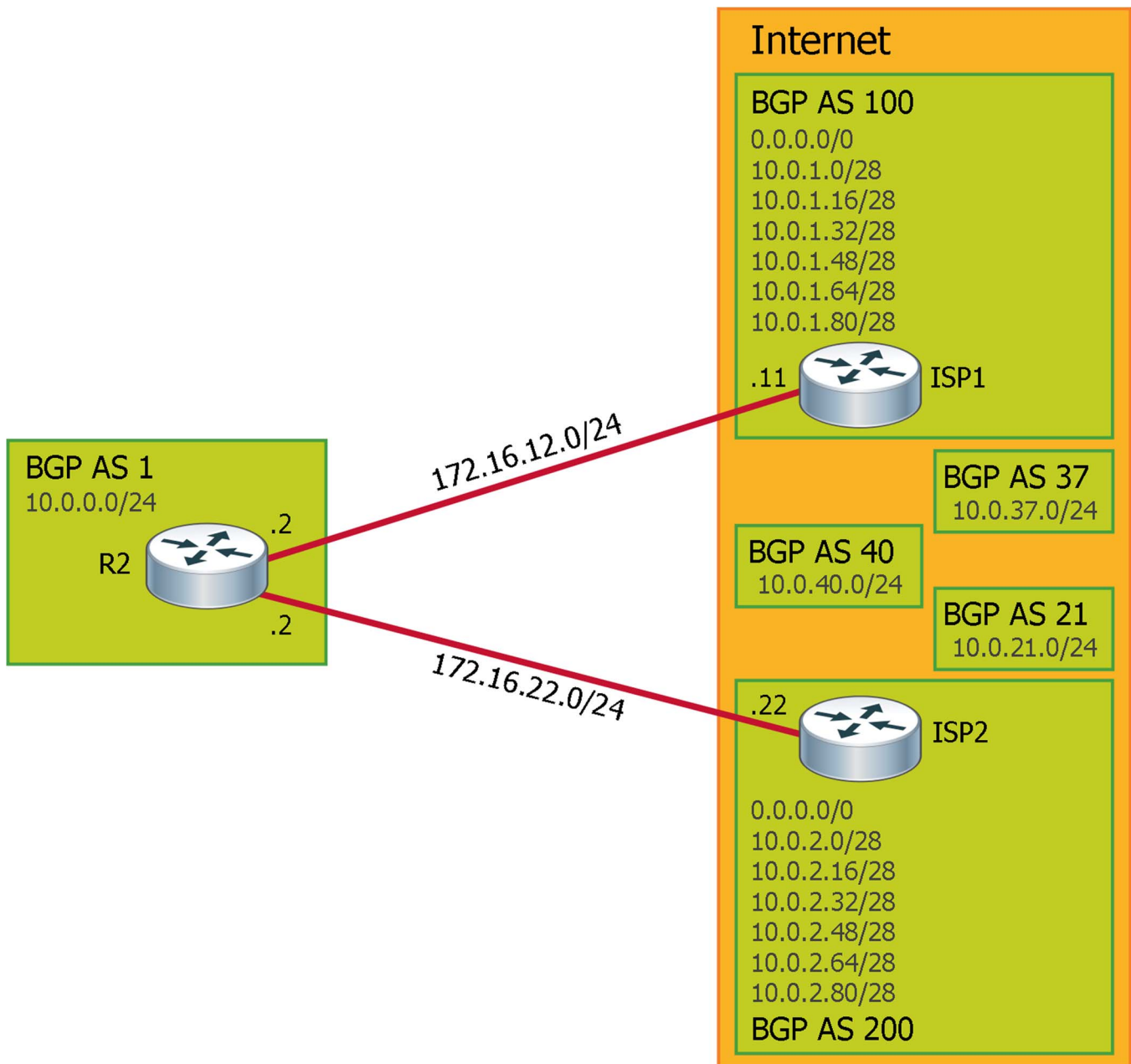
Job Aids

Note If you shut down an interface on a real router or switch, the connected device will see it as "down/down." Due to virtualization specifics, [IOL](#) behavior is slightly different. If you shut down an interface on a router or switch, the connected device will see it as "up/up." In IOL, the status of an interface can only be "up/up" or "administratively down/down."

Device Information

Device	Interface	IP address	Description
R2	Ethernet 0/0	172.16.22.2/24	Connection to ISP2
R2	Ethernet 0/1	172.16.12.2/24	Connection to ISP1
R2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5	10.0.0.1/28 10.0.0.17/28 10.0.0.33/28 10.0.0.49/28 10.0.0.65/28	Loopbacks simulate LAN networks
ISP1	Ethernet 0/1	172.16.12.11/24	Connection to R2
ISP1	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6	10.0.1.1/28 10.0.1.17/28 10.0.1.33/28 10.0.1.49/28 10.0.1.65/28 10.0.1.81/28	Loopbacks simulate LAN networks
ISP1	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate extra networks in different autonomous systems.
ISP2	Ethernet 0/0	172.16.22.22/24	Connection to R2
ISP2	Loopback 1 Loopback 2 Loopback 3 Loopback 4 Loopback 5 Loopback 6 Loopback 7 Loopback 8 Loopback 9	10.0.2.1/28 10.0.2.17/28 10.0.2.33/28 10.0.2.49/28 10.0.2.65/28 10.0.2.81/28 10.0.2.97/28 10.0.2.113/28 10.0.2.129/28	Loopbacks simulate LAN networks
ISP2	Loopback 21 Loopback 37 Loopback 40	10.0.21.1 10.0.37.1 10.0.40.1	Loopbacks simulate extra networks in different autonomous systems.

[IP addresses](#) and advertised networks in BGP are preconfigured as shown below:



BGP is also preconfigured as [EBGP](#) (R2 to ISP1 and R2 to ISP2). The ISP1 router announces these networks:

- Default route 0.0.0.0/0
- 10.0.1.16/28
- 10.0.1.32/28
- 10.0.1.48/28
- 10.0.1.64/28
- 10.0.1.80/28

Configure Route Maps as BGP Filters

Discovery Steps

Step 1 On the ISP1 router, verify that the default route is being announced in an EBGp session to the customer router R2.

```
ISP1# show ip bgp neighbor 172.16.12.2 advertised-routes
BGP table version is 13, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Originating default network 0.0.0.0

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.1.0/28	0.0.0.0	0		32768	i
*>	10.0.1.16/28	0.0.0.0	0		32768	i
*>	10.0.1.32/28	0.0.0.0	0		32768	i
*>	10.0.1.48/28	0.0.0.0	0		32768	i
*>	10.0.1.64/28	0.0.0.0	0		32768	i
*>	10.0.1.80/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i

You can see that the ISP1 router originates the default route.

Step 2 On the ISP2 router, verify that the default route is being announced in an EBGp session to the customer router R2.

```
ISP2# show ip bgp neighbors 172.16.22.2 advertised-routes
BGP table version is 20, local router ID is 10.0.40.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Originating default network 0.0.0.0

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.2.0/28	0.0.0.0	0		32768	i
*>	10.0.2.16/28	0.0.0.0	0		32768	i
*>	10.0.2.32/28	0.0.0.0	0		32768	i
*>	10.0.2.48/28	0.0.0.0	0		32768	i
*>	10.0.2.64/28	0.0.0.0	0		32768	i
*>	10.0.2.80/28	0.0.0.0	0		32768	i
*>	10.0.2.96/28	0.0.0.0	0		32768	i
*>	10.0.2.112/28	0.0.0.0	0		32768	i
*>	10.0.2.128/28	0.0.0.0	0		32768	i
*>	10.0.21.0/24	0.0.0.0	0		32768	i
*>	10.0.37.0/24	0.0.0.0	0		32768	i
*>	10.0.40.0/24	0.0.0.0	0		32768	i
	Network	Next Hop	Metric	LocPrf	Weight	Path

Total number of prefixes 12

You can see that the ISP2 router originates the default route.

Step 3 On the R2 router, verify that the customer router R2 receives default route from both neighbors, the ISP1 and ISP2 routers. You should also make sure that the R2 router receives several other routes from both [ISP](#) routers.

R2# **show ip bgp**

BGP table version is 26, local router ID is 10.0.0.65

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
x best-external, a additional-path, c RIB-compressed,

Origin codes: i - IGP, e - EGP, ? - incomplete

RPKI validation codes: V valid, I invalid, N Not found

Network	Next Hop	Metric	LocPrf	Weight	Path
* 0.0.0.0	172.16.22.22			0 200	i
*>	172.16.12.11			0 100	i
*> 10.0.0.0/24	0.0.0.0	0		32768	i
*> 10.0.0.16/28	0.0.0.0	0		32768	i
*> 10.0.0.32/28	0.0.0.0	0		32768	i
*> 10.0.0.48/28	0.0.0.0	0		32768	i
*> 10.0.0.64/28	0.0.0.0	0		32768	i
*> 10.0.1.0/28	172.16.12.11	0		0 100	i
*> 10.0.1.16/28	172.16.12.11	0		0 100	i
*> 10.0.1.32/28	172.16.12.11	0		0 100	i
*> 10.0.1.48/28	172.16.12.11	0		0 100	i
*> 10.0.1.64/28	172.16.12.11	0		0 100	i
*> 10.0.1.80/28	172.16.12.11	0		0 100	i
*> 10.0.2.0/28	172.16.22.22	0		0 200	i
Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.2.16/28	172.16.22.22	0		0 200	i
*> 10.0.2.32/28	172.16.22.22	0		0 200	i
*> 10.0.2.48/28	172.16.22.22	0		0 200	i
*> 10.0.2.64/28	172.16.22.22	0		0 200	i
*> 10.0.2.80/28	172.16.22.22	0		0 200	i
*> 10.0.2.96/28	172.16.22.22	0		0 200	i
*> 10.0.2.112/28	172.16.22.22	0		0 200	i
*> 10.0.2.128/28	172.16.22.22	0		0 200	i
*> 10.0.21.0/24	172.16.22.22	0		0 200 21	i
*>	172.16.12.11	0		0 100 37 40	21 i
*> 10.0.37.0/24	172.16.22.22	0		0 200 21 40	37 i
*>	172.16.12.11	0		0 100 37	i
*> 10.0.40.0/24	172.16.22.22	0		0 200 21 40	i
*>	172.16.12.11	0		0 100 37 40	i

You can see that the default route on the R2 router was received from two different EBGP neighbors:

- One default route was received from the ISP1 router with the AS path attribute set to 100 and BGP next hop set to 172.16.12.11. This address is the IP address of the ISP1 router.
- Another default route was received from the ISP2 router with the AS path attribute set to 200 and BGP next hop set to 172.16.22.22. This address is the IP address of ISP2 router.

You should receive these BGP routes on the R2 router:

- Any subnets in 10.0.1.0/24 range from the ISP1 router
- Any subnets in 10.0.2.0/24 range from the ISP2 router
- Networks 10.0.21.0/24, 10.0.37.0/24, 10.0.40.0/24 from both the ISP1 and ISP2 routers

Step 4 Configure the route map on the R2 router. The route map should permit only the default route, that is defined in the prefix list, to be accepted from the ISP1 EBGp neighbor.

```
R2(config)# ip prefix-list DefaultOnly permit 0.0.0.0/0
R2(config)# route-map ISP1 permit 10
R2(config-route-map)# match ip address prefix-list DefaultOnly
R2(config-route-map)# exit
R2(config)# router bgp 1
R2(config-router)# neighbor 172.16.12.11 route-map ISP1 in
```

Step 5 Enable debugging of BGP updates on the R2 router.

```
R2# debug ip bgp update
```

Step 6 Clear all BGP sessions on the R2 router.

```
R2# clear ip bgp *
```

Step 7 On the R2 router, observe that all non default route BGP updates received from the ISP1 router were denied because of the configured route map. Only the default route that is received via the EBGp session from the ISP1 router should be installed to the routing table.

```

%BGP-5-ADJCHANGE: neighbor 172.16.12.11 Down User reset
*Mar 6 00:36:02.681: %BGP_SESSION-5-ADJCHANGE: neighbor 172.16.12.11 IPv4 Unicast
topology base removed from session User reset
*Mar 6 00:36:02.681: %BGP-5-ADJCHANGE: neighbor 172.16.22.22 Down User reset
*Mar 6 00:36:02.681: %BGP_SESSION-5-ADJCHANGE: neighbor 172.16.22.22 IPv4 Unicast
topology base removed from session User reset
*Mar 6 00:36:03.107: %BGP-5-ADJCHANGE: neighbor 172.16.12.11 Up
*Mar 6 00:36:03.109: %BGP-5-ADJCHANGE: neighbor 172.16.22.22 Up
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd UPDATE w/ attr: nexthop 172.16.12.11,
origin i, metric 0, merged path 100 37 40, AS_PATH
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.40.0/24 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd UPDATE w/ attr: nexthop 172.16.12.11,
origin i, metric 0, merged path 100 37, AS_PATH
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.37.0/24 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd UPDATE w/ attr: nexthop 172.16.12.11,
origin i, metric 0, merged path 100 37 40 21, AS_PATH
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.21.0/24 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd UPDATE w/ attr: nexthop 172.16.12.11,
origin i, metric 0, merged path 100, AS_PATH
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.0/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.16/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.32/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.48/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.64/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.113: BGP(0): 172.16.12.11 rcvd 10.0.1.80/28 -- DENIED due to: route-
map;
*Mar 6 00:36:03.114: BGP(0): 172.16.12.11 rcvd UPDATE w/ attr: nexthop 172.16.12.11,
origin i, merged path 100, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.12.11 rcvd 0.0.0.0/0
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22,
origin i, metric 0, merged path 200 21 40, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.40.0/24
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22,
origin i, metric 0, merged path 200 21 40 37, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.37.0/24
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22,
origin i, metric 0, merged path 200 21, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.21.0/24
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22,
origin i, metric 0, merged path 200, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.0/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.16/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.32/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.48/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.64/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.80/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.96/28
R2#
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.112/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 10.0.2.128/28
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop 172.16.22.22,
origin i, merged path 200, AS_PATH
*Mar 6 00:36:03.114: BGP(0): 172.16.22.22 rcvd 0.0.0.0/0
*Mar 6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 0.0.0.0/0 ->
172.16.12.11(global) to main IP table
*Mar 6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.0/28 ->
172.16.22.22(global) to main IP table
*Mar 6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.16/28 ->

```

```

172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.32/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.48/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.64/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.80/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.96/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.112/28 ->
172.16.22.22(global) to main IP table
*Mar  6 00:36:04.130: BGP(0): Revise route installing 1 of 1 routes for 10.0.2.128/28 ->
172.16.22.22(global) to main IP table

```

Step 8 On the R2 router, verify the content of BGP table.

```

R2# show ip bgp
BGP table version is 19, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	0.0.0.0	172.16.22.22			0 200	i
*>		172.16.12.11			0 100	i
*>	10.0.0.0/24	0.0.0.0	0		32768	i
*>	10.0.0.16/28	0.0.0.0	0		32768	i
*>	10.0.0.32/28	0.0.0.0	0		32768	i
*>	10.0.0.48/28	0.0.0.0	0		32768	i
*>	10.0.0.64/28	0.0.0.0	0		32768	i
*>	10.0.2.0/28	172.16.22.22	0		0 200	i
*>	10.0.2.16/28	172.16.22.22	0		0 200	i
*>	10.0.2.32/28	172.16.22.22	0		0 200	i
*>	10.0.2.48/28	172.16.22.22	0		0 200	i
*>	10.0.2.64/28	172.16.22.22	0		0 200	i
*>	10.0.2.80/28	172.16.22.22	0		0 200	i
*>	10.0.2.96/28	172.16.22.22	0		0 200	i
	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.2.112/28	172.16.22.22	0		0 200	i
*>	10.0.2.128/28	172.16.22.22	0		0 200	i
*>	10.0.21.0/24	172.16.22.22	0		0 200 21	i
*>	10.0.37.0/24	172.16.22.22	0		0 200 21 40 37	i
*>	10.0.40.0/24	172.16.22.22	0		0 200 21 40	i

You should only see the default route that is installed for BGP updates from the ISP1 router with next hop IP address 172.16.12.11. All other networks were received from the ISP2 router.

Displaying Routes Matching the Route Map



Displaying Routes Matching the Route Map

Monitoring Route-Maps (Cont.)

```
router# show ip bgp route-map route-map-name
```

- Displays all routes in BGP table matching the route map
- Used for filtering the **show ip bgp** output on basis of BGP path attributes:
 - Community
 - Local preference
 - Weight
 - Origin
 - Next-hop
- Can also filter based on prefixes
- Allows powerful combined filtering

© 2015 Cisco and/or its affiliates. All rights reserved.

236

You can also use route maps for selective and powerful filtering of the BGP table. The **show ip bgp route-map** command displays the selected routes from a BGP routing table based on the contents of a route map.

A route map can match the routes based on BGP path attributes (local preference, community, weight, origin, next-hop) or prefix lists and access lists (matching IP prefixes). The power of route map filtering lies in the possibility of combining different filters (for example, filtering on community, prefix, and next-hop values).

Step 9 Configure the route map on the R2 router. The route map should permit only networks passing through the autonomous system 40.

```
R2(config)# ip as-path access-list 2 permit _40_  
R2(config)# route-map PassAS40 permit 10  
R2(config-route-map)# match as-path 2
```

Step 10 On the R2 router, display the content of BGP table matching the configured route map.

```

R2# show ip bgp route-map PassAS40
BGP table version is 19, local router ID is 10.0.0.65
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	10.0.37.0/24	172.16.22.22	0		0 200 21	40 37 i
*>	10.0.40.0/24	172.16.22.22	0		0 200 21	40 i

You should only see the networks containing "40" in the AS path.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- A route map is a filter that has the ability to drop denied routes as well as modify attributes of the permitted routes.
- The BGP Route-Map Policy List Support feature introduces new functionality to BGP route maps. It adds the ability for a network operator to group route map match clauses into named lists that are called policy lists.
- The BGP Route-Map Continue feature introduces the continue clause to BGP route map configuration. Continue clauses provide a programmable method to organize and control the flow of a route map.

© 2015 Cisco and/or its affiliates. All rights reserved.

237

Summary (Cont.)

- You can configure a route map to match against a prefix list by using the **match ip address**, **match ip next-hop**, and **match ip route-source** commands.
- Filter lists, prefix lists, and route maps can optionally all be applied on either incoming or outgoing information in any combination.
- A route map can be applied on incoming or outgoing routing information to or from a neighbor. However, the route map must permit the routing information in order to be accepted.
- Monitoring route maps is possible using the **show ip bgp** and **debug ip bgp update** commands.

© 2015 Cisco and/or its affiliates. All rights reserved.

238

Lesson 6: Implementing Changes in BGP Policy

Overview

Because of the huge volumes of routing information that [BGP](#) is capable of handling, traditional routing update methods are not feasible. Routing policies for a BGP neighbor may include filtering mechanisms such as route maps, distribute lists, prefix lists, and [AS](#) path filter lists. Each of these filters may affect inbound or outbound routing table updates.

Whenever there is an administrative change in a routing policy, the BGP session must be reset before the new policy can take effect. To accomplish this task, you can use two types of reset: hard reset and soft reset. Clearing a BGP session using a hard reset invalidates the cache and results in a negative impact on the operation of networks, because the information in the cache becomes unavailable. A soft reset is recommended because it allows routing tables to be reconfigured and activated without clearing the BGP session.

This lesson discusses routing updates in a BGP environment and the traditional methods of forcing BGP route updates after changes in a filter policy. The function and benefits of soft reconfiguration and route refresh are also discussed. The lesson also presents the commands that are required to perform a soft reconfiguration and route refresh and explains how to monitor and troubleshoot these features.

Upon completing this lesson, you will be able to:

- Identify the limitations of the traditional methods of forcing BGP route updates after changing a filter policy
- Describe the function of the BGP Soft Reset Enhancement feature
- Describe the function and benefits of the route refresh function
- Identify the Cisco IOS command that is required to trigger a route refresh
- Identify the Cisco IOS commands that are required to monitor route refresh operation

Traditional Filtering Limitations

BGP can potentially handle huge volumes of routing information. But it cannot handle it when you change configuration lines in filters or route maps. In this case, the router cannot go through the huge table of BGP information and calculate which entry is no longer valid in the local table. Nor can the router determine which route or routes, already advertised, should be withdrawn from a neighbor. There is an obvious risk that the first configuration change will be immediately followed by a second, which would cause the whole process to start all over again.

Traditional Filtering Limitations

- All filters apply only to new incoming and outgoing updates.
- To change the outbound routing policy, you have to resend BGP updates to your neighbors.
- To change the inbound routing policy, you have to force your neighbor to resend the updates to you.
- The traditional mechanism is to clear BGP sessions.

© 2015 Cisco and/or its affiliates. All rights reserved.

239

To avoid such a problem, Cisco IOS software applies changes only on the updates that are received or transmitted after the configuration change has been performed. This approach means that the new routing policy, which the new filters enforce, is applied only on routes that are received or sent after the change. If network administrators would like to apply the policy change on all routes, they have to force the router to let all routes pass through the new filter.

If the filter is applied to outgoing information, the router has to resend the entire BGP table through the new filter. If the filter is applied to incoming information, the router needs its neighbor to resend its entire BGP table so that it passes through the new filters.

Traditionally, to accomplish these goals, you have torn down the affected BGP sessions after completing a configuration change. After the sessions are down, all information that is received on those sessions is invalidated and removed from the BGP table. Also, the remote neighbor will detect a session down state, and it likewise will invalidate the routes that are received on the session. After a period of 30 to 60 seconds, the sessions are re-established automatically and the entire BGP table is exchanged again, but through the new filters. This process, however, disrupts packet forwarding.

Traditional Limitations of Clearing the BGP Session

Traditional Limitations of Clearing the BGP Session

```
router# clear ip bgp { * | ip-address | peer-group-name }
```

- This command tears down the BGP session with all neighbors, a specific neighbor, or all neighbors in a peer group.
- All BGP routes are lost after the session is torn down; connectivity through the BGP neighbor is lost.
- A new session is re-established within 30 to 60 seconds.
- A full routing update is exchanged once the session is re-established, resulting in enforcement of new routing policy.
- Processing the full Internet routing table can take a long time.
- **Clearing the BGP session is a very disruptive way to implement routing policies.**

© 2015 Cisco and/or its affiliates. All rights reserved.

240

The EXEC command **clear ip bgp** tears down one or several BGP sessions. The BGP sessions are terminated, and the [TCP](#) connections closed. The neighbors go into the Idle state and stay there for approximately 30 seconds. Next, the neighbor session goes into the Active state, and the sessions are re-established.

You can implement the **clear ip bgp** command with the * (asterisk) argument, which applies to all sessions. You can also make a reference to a specific session or group of sessions to tear down.

When the session is down, all routes that are received over the session by both routers are invalidated. When the session is once again in the Established state, all BGP routes have to be resent by both peers. All routes also have to pass through the new filters, which enforces the new policy.

Exchanging the complete Internet routing table takes time, bandwidth, and [CPU](#) resources. IP packet forwarding to and from the neighbor is down for several minutes. Also, revoking and reannouncing the routes will be registered by the rest of the Internet as a flap for each route.

BGP Soft Reset Enhancement

Previously, to perform a soft reset for inbound routing table updates, the **neighbor soft-reconfiguration** command directed the Cisco IOS software in the local [BGP](#) router to store all received (inbound) routing policy updates without modification. This method is memory-intensive and not recommended unless absolutely necessary. (Outbound updates have never required the extra memory and this feature does not affect them.)

BGP Soft Reset Enhancement

- BGP Soft Reset Enhancement provides automatic support for dynamic soft reset of inbound BGP routing table updates that is not dependent upon stored routing table update information.
- Requires no preconfiguration (as with the **neighbor soft-reconfiguration** command).
- Requires much less memory than the previous soft reset method for inbound routing table updates.

© 2015 Cisco and/or its affiliates. All rights reserved.

241

Whenever there is a change in the routing policy, the BGP session must be cleared, or reset, for the new policy to take effect. There are two types of reset, hard reset, and soft reset. Clearing a BGP session using a hard reset invalidates the cache and results in a negative impact on the operation of networks as the information in the cache becomes unavailable. Soft reset is recommended because it allows routing tables to be reconfigured and activated without clearing the BGP session. Soft reset is done on per-neighbor basis. There are two types of soft resets:

- **Dynamic inbound soft reset:** When soft reset is used to generate inbound updates from a neighbor
- **Outbound soft reset:** When soft reset is used to send a new set of updates to a neighbor.

The BGP Soft Reset Enhancement feature, however, provides automatic support for dynamic soft reset of inbound BGP routing table updates that is not dependent on stored routing table update information. The new method requires no preconfiguration (as with the **neighbor soft-reconfiguration** command) and requires much less memory than the previous soft reset method for inbound routing table updates.

There are a number of benefits to the BGP Soft Reset Enhancement feature:

- **Allows dynamic route refresh requests:** This feature provides a way to initiate nondisruptive routing policy changes. It achieves this by allowing the dynamic exchange of route refresh requests between BGP routers, and the subsequent readvertisement of the respective outbound routing tables.

- **Requires no preconfiguration:** Because support for the soft reset using the route refresh capability is included in this release of the Cisco IOS software, no further router configuration is required. You can initiate a soft inbound reset using only the **clear ip bgp in** command.
- **Requires no additional memory resources:** Unlike a soft reset using the stored inbound routing table updates provided by the **neighbor soft-reconfiguration** command, when both BGP peers support the route refresh capability, inbound routing table updates are not stored in the local BGP router. The soft reset requests are exchanged dynamically, and no additional memory is required.
- **Provides flexibility:** There are now two available methods for inbound soft reset. The older method, using stored inbound routing table updates, and the method that this feature provides, using dynamic exchange of update information.

When the routing policy of a BGP neighbor changes, the session must be reset (cleared) for the changes to take effect. Because resetting a BGP session can be disruptive to networks, a soft reset method is recommended for reconfiguring the routing table. Previously, in order to reconfigure the inbound routing table, both the local BGP router and the BGP peer first needed to be configured to store incoming routing policy updates. BGP soft reconfiguration used to be configured using the **neighbor soft-reconfiguration** command. More resources, particularly memory, were required to store the inbound routing table updates. The **clear ip bgp** command could then initiate the soft reset, which generated a new set of inbound routing table updates using the stored information.

The BGP Soft Reset Enhancement feature provides an additional method for soft reset. It allows the dynamic exchange of route refresh requests and routing information between BGP routers and the subsequent readvertisement of the respective outbound routing table. Soft reset using the route refresh capability does not require preconfiguration and consumes no additional memory resources.

To use this new method, both BGP peers must support the soft route refresh capability. This capability is advertised in the Open message that is sent when a peer sends its routing table update. Any router running BGP with this software release automatically supports the route refresh capability. Routers running earlier Cisco IOS software releases do not support the route refresh capability and must use the older soft reset method. If the soft reset fails, you can still clear the BGP session. However, it will have a negative impact on network operations and should be used only as a last resort.

Note	Outbound resets have never required preconfiguration or storing of routing table updates and remain unchanged by the BGP Soft Reset Enhancement feature.
-------------	--

Route Refresh

Route refresh is one of the capabilities of [BGP](#). Routers use the route refresh feature to request a neighbor to resend all the routing information when it is needed.

Route Refresh

- Route refresh is a BGP capability.
- It is used to request a neighbor to resend routing information.
- It is typically used after configuration changes to update the BGP table (route map, distribute list, prefix list, filter list, weight, local preference, MED, etc.).
- The traditional way of accomplishing this function is to clear the BGP session.
- Inbound soft reconfiguration consumes memory on the receiving router.
 - It is needed only because there is no mechanism in standard BGP to request retransmission of BGP routes.
- BGP route refresh is an optional BGP capability that allows a BGP router to request retransmission of BGP routes from a neighbor.

© 2015 Cisco and/or its affiliates. All rights reserved.

242

There are several ways of refreshing the routing information from a neighbor:

- Clearing the neighbor relationship
- Soft-clearing the neighbor relationship (if soft reconfiguration is enabled for this specific neighbor)
- Using route refresh (if the neighbor supports this capability)

The **soft-reconfiguration inbound** feature consumes large volumes of memory in the Internet environment. The number of routes that can be received from a peer router on the Internet is so large that it is not feasible to store an extra copy.

The only reason for making the extra copy is to be able to replay the data through the new routing policy without tearing down the session and re-establishing it.

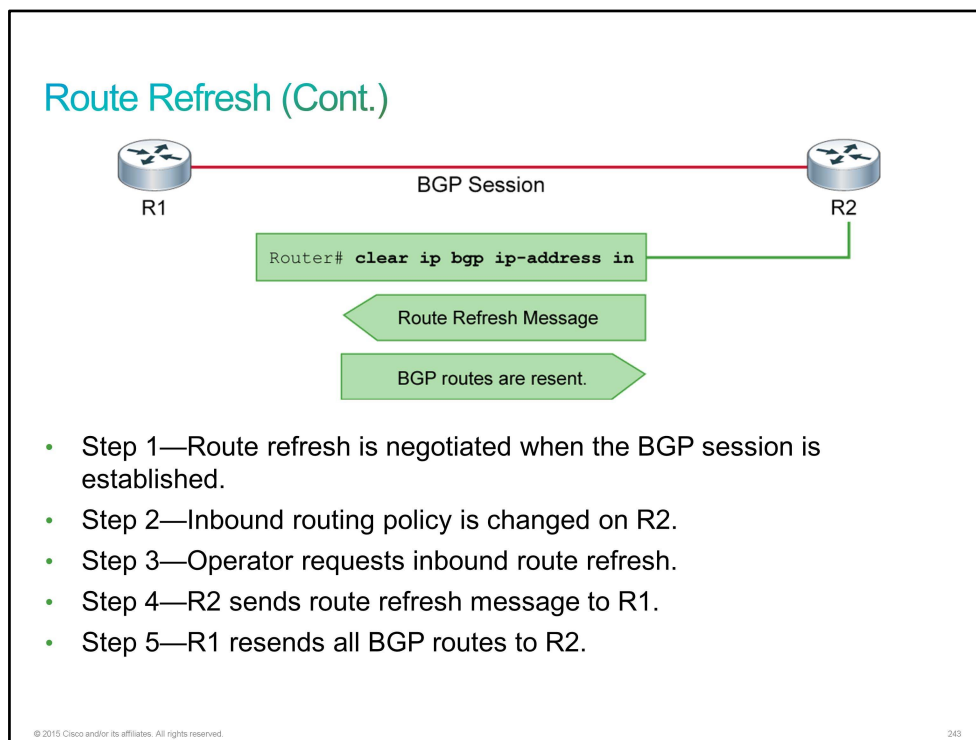
What is needed is a mechanism to ask the neighbor router to do a "clear soft outbound." If this were possible, the extra copy would not be needed. The neighboring router, of course, has its own copy in its BGP table, which it could resend to the local router whenever it is signaled to do so.

There is no such mechanism in standard BGP, but there is an optional BGP capability that allows one router to request a refresh from its neighbor: route refresh.

The table compares the various methods of BGP session reset, stating the advantages and disadvantages of each.

Note To use soft reset without preconfiguration, both BGP peers must support the soft route refresh capability, which is advertised in the Open message that is sent when the peers establish a [TCP](#) session. Routers that run Cisco IOS software releases earlier than Release 12.1 do not support the route refresh capability and must clear the BGP session using the **neighbor soft-reconfiguration** command.

Type of Reset	Advantages	Disadvantages
Hard reset	No memory overhead.	The prefixes in the BGP, IP, and FIB tables that the neighbor provides, are lost. Not recommended.
Outbound soft reset	No configuration, no storing of routing table updates.	Does not reset inbound routing table updates.
Dynamic inbound soft reset	Does not clear the BGP session or cache. Does not require storing of routing table updates, and has no memory overhead.	Both BGP routers must support the route refresh capability (Cisco IOS Software Release 12.1 and later releases).
Configured inbound soft reset (uses the neighbor soft-reconfiguration command)	Can be used when both BGP routers do not support the automatic route refresh capability.	Requires preconfiguration. Stores all received (inbound) routing policy updates without modification, and is thus memory-intensive. Recommended only when absolutely necessary.



The router must negotiate the ability to use the route refresh feature when the BGP session is first established. The local router keeps a record that the capability is available with the neighbor. There is no need to keep a copy of the routing information that is received from the neighbor if it has the ability to refresh.

After reconfiguring the filters and route maps that will implement a new routing policy, you can issue the **clear ip bgp *ip-address* soft in** command in the local router. The router checks whether the route refresh capability is available, and if it is, requests resending of the BGP table of the neighbor instead of replaying its own copy.

Configuring Route Refresh

Use the **clear ip bgp * in** command, to send a route refresh message to all neighbors or **clear ip bgp ip-address in** to send a route refresh message to a specific neighbor.

Configuring Route Refresh

```
router# clear ip bgp { * | ip-address | peer-group-name } in
```

- Sends a route refresh message to the neighbor or neighbors.
- Only works if the neighbor has previously advertised the route refresh capability.

© 2015 Cisco and/or its affiliates. All rights reserved.

244

You need not use the **soft** keyword, because soft reset is automatically assumed when the route refresh capability is supported.

To reset a [BGP](#) connection with BGP soft reconfiguration, use the **clear ip bgp** privileged EXEC command at the system prompt.

```
clear ip bgp { * | ip-address | peer-group-name } [soft [in | out]]
```

Syntax Description

Parameter	Description
*	Resets all current BGP sessions.
<i>ip-address</i>	Resets only the identified BGP neighbor.
<i>peer-group-name</i>	Resets the specified BGP peer group.
soft	(Optional) Soft reset. Does not reset the session
in out	(Optional) Triggers inbound or outbound soft reconfiguration. If the in or out option is not specified, both inbound and outbound soft reset are triggered.

Note When doing **clear ip bgp * in**, if **soft-reconfiguration inbound** is configured for a neighbor, it will cause a soft action. If **soft-reconfiguration inbound** is not configured, it will cause a route refresh.

Monitoring Route Refresh

You need to verify whether the [BGP](#) neighbor supports the route refresh capability.

Monitoring Route Refresh

```
R2# show ip bgp neighbors 172.16.22.22
BGP neighbor is 172.16.22.22, remote AS 200, external link
  BGP version 4, remote router ID 10.0.40.1
  BGP state = Established, up for 02:31:21
  Last read 00:00:23, last write 00:00:49, hold time is 180, keepalive interval
  is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
```

- Verifies the support for route refresh capability.

© 2015 Cisco and/or its affiliates. All rights reserved.

245

Use the **show ip bgp neighbor** command to see whether the neighbor supports the route refresh message. In the example, neighbor 172.16.22.22 supports route refresh capability.

Monitoring Route Refresh (Cont.)

```
R2# debug ip bgp
*Mar 10 05:06:53.243: BGP: ses global 172.16.22.22 (0xECFD0708:0) act Send OPEN
*Mar 10 05:06:53.243: BGP: 172.16.22.22 active sending OPEN, version 4, my as:
1, holdtime 180 seconds, ID A000041
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active rcv message type 1, length
(excl. header) 38
*Mar 10 05:06:53.247: BGP: ses global 172.16.22.22 (0xECFD0708:0) act Receive
OPEN
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active rcv OPEN, version 4, holdtime
180 seconds
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active OPEN has CAPABILITY code: 128,
length 0
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active OPEN has ROUTE-REFRESH
capability(old) for all address-families
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active rcvd OPEN w/ optional parameter
type 2 (Capability) len 2
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active OPEN has CAPABILITY code: 2,
length 0
*Mar 10 05:06:53.247: BGP: 172.16.22.22 active OPEN has ROUTE-REFRESH
capability(new) for all address-families
<... output omitted ...>
*Mar 10 05:06:53.248: BGP: 172.16.22.22 sending REFRESH_REQ(5) for afi/safi:
1/1, refresh code is 1
```

- Debug output after BGP session reset

© 2015 Cisco and/or its affiliates. All rights reserved.

246

Use **debug ip bgp** to display the negotiation of capabilities. Debugging displays the received capabilities.

The example shows a debugging output, after the BGP session was reset, using a hard reconfiguration. You can see, that a neighbor is advertising both old-style and standard (new-style) route refresh. The "old" refers to the old capability code of 128 that the old Cisco implementations of Route Refresh used, before the [RFC 2918](#) was in place. The "new" refers to the RFC 2918 compliant capability code 2 that all recent BGP implementations that support the Route Refresh use, according to RFC 2918. After the session has been established, the router sends an initial standard route refresh message for the address family 1/1 ([IPv4](#) unicast).

Monitoring Route Refresh (Cont.)

```
R2# debug ip bgp
R2# debug ip bgp updates
R2# clear ip bgp 172.16.22.22 in
*Mar 10 05:34:28.458: BGP: 172.16.22.22 sending REFRESH_REQ(5) for afi/safi:
1/1, refresh code is 0
*Mar 10 05:34:28.459: BGP: 172.16.22.22 rcv message type 5, length (excl.
header) 4
*Mar 10 05:34:28.459: BGP: 172.16.22.22 rcvd REFRESH_REQ for afi/safi: 1/1,
refresh code is 1
*Mar 10 05:34:28.459: BGP: nbr_topo global 172.16.22.22 IPv4 Unicast:base
(0xECFD0708:1) rcvd Refresh Start-of-RIB
*Mar 10 05:34:28.459: BGP: nbr_topo global 172.16.22.22 IPv4 Unicast:base
(0xECFD0708:1) refresh_epoch is 3
*Mar 10 05:34:28.460: BGP(0): 172.16.22.22 rcvd UPDATE w/ attr: nexthop
172.16.22.22, origin i, metric 0, merged path 200, AS_PATH
*Mar 10 05:34:28.461: BGP(0): 172.16.22.22 rcvd 10.0.2.0/28
*Mar 10 05:34:28.461: BGP(0): 172.16.22.22 rcvd 10.0.2.16/28
*Mar 10 05:34:28.461: BGP(0): 172.16.22.22 rcvd 10.0.2.32/28
```

- Debug output after route refresh

© 2015 Cisco and/or its affiliates. All rights reserved.

247

Debugging also shows a route refresh message being sent to a neighbor after the network administrator issues the **clear ip bgp ip-address in** command from the privileged EXEC mode. When the neighbor 172.16.22.22 receives the route refresh message, it resends the BGP updates to the router R2. Updates are then processed against the configured BGP inbound filters on the R2 router.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Because of the huge volumes of routing information, BGP cannot use traditional routing update methods
- The BGP Soft Reset Enhancement feature provides automatic support for dynamic soft reset of inbound BGP routing table updates that is not dependent upon stored routing table update information. This method requires no preconfiguration and needs much less memory than the previous soft reset method for inbound routing table updates.
- Route refresh is a BGP capability that is used to request a neighbor to resend routing information after configuration changes.
- The **clear ip bgp** ip-address in command sends a route refresh message to the neighboring router and executes if the neighbor has previously advertised the route refresh capability.
- To verify whether a neighbor supports route refresh, you can use the **show ip bgp neighbor** command. To display the negotiation process, you can use the **debug ip bgp** command.

© 2015 Cisco and/or its affiliates. All rights reserved.

248

Lesson 7: Module Summary

Overview

This topic summarizes the key points that were discussed in this module.

Module Summary

- The multihomed customer network must exchange BGP information with both ISP networks. Dynamic routing is required for full redundancy, and BGP is the only protocol available that can be used in this scenario.
- An AS-path filter is created by an AS-path access-list. The access-list is applied to a set of routes from which to subset can be selected.
- Use prefix-lists to filter incoming or outgoing BGP updates to neighbors and to filter routes that are being redistributed into the BGP process from other routing protocols.

Module Summary (Cont.)

- Outbound route filtering is a mechanism that is used to minimize the number of updates that are requested from a neighbor.
- Route-maps provide a method to perform a variety of compound, complex filtering operations (such as dropping denied routes and modifying attributes of the permitted routes) within a single tool.
- Soft reconfiguration provides the ability to run all routes through filters without tearing down the sessions.

© 2015 Cisco and/or its affiliates. All rights reserved.

250

This module discussed BGP route filtering and BGP route selection policies. The module described multihomed BGP networks and identified the need for BGP route selection. This module also addressed configuring BGP to influence route selection by using AS-path filters, prefix-list filters, and route-maps. Outbound route filtering was also explained. In addition, details about soft reconfiguration and route refresh were provided.

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Sample Configuration for BGP with Two Different Service Providers (Multihoming)*.
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/23675-27.html>
- Cisco Systems, Inc. *Using Regular Expressions in BGP*.
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/13754-26.html>
- Cisco Systems, Inc. *BGP Case Studies "BGP Case Studies 1."*
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/26634-bgp-toc.html#BGPsec1>
- Cisco Systems, Inc. *BGP Case Studies "BGP Case Studies 3."*
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/26634-bgp-toc.html#sec3>
- Cisco Systems, Inc. *Configuring BGP*.
http://www.cisco.com/c/en/us/td/docs/ios/12_2/ip/configuration/guide/fipr_c/1cfbgp.html
- Cisco Systems, Inc. *BGP Prefix-Based Outbound Route Filtering*.
http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bgp/configuration/xr-3e/irg-iproute-bgp-xr-3e-book/irg-oubound-route-filtering.pdf
- Cisco Systems, Inc. *Compatible Systems Setup Guides: BGP Configuration Guide*.
<http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/17612-bgp.html>

- Cisco Systems, Inc. *BGP Soft Reset Enhancement*.
http://www.cisco.com/c/en/us/td/docs/ios/12_0s/feature/guide/s_sftrst.html?mdfid=275960490

Lesson 8: Module Self-Check

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

1. What are two reasons why a customer would want to connect to two ISPs? (Choose two.) (Source: "Using Multihomed BGP Networks")
 - A. To expand capacity for Internet traffic.
 - B. To better protect confidential information as it travels through the Internet.
 - C. To provide redundancy to mission-critical services that are offered over the Internet.
 - D. To efficiently route Internet traffic to two different divisions within the company.

2. What are the two technical requirements for multihomed customers? (Choose two.) (Source: "Using Multihomed BGP Networks")
 - A. The ISPs must assign a range of IP network numbers to the customer.
 - B. The customer network must exchange BGP information with each ISP network.
 - C. In most cases, the customer must have its own public AS number.
 - D. The customer network must not use AS path filters.

3. Which of the following statements best illustrates the importance of BGP policies that influence route selection in a multihomed BGP network? (Source: "Using Multihomed BGP Networks")
 - A. The default BGP route selection does not always result in optimum routing.
 - B. The default BGP route selection always results in optimum routing.
 - C. After the route selection behavior has been set, it cannot be changed.
 - D. The customer receives all routes from both service providers, giving redundancy; therefore, BGP policies are not necessary.

4. There is no need for ISPs and customers to influence the BGP path selection as BGP has built-in mechanisms to find the optimal way to reach the destination. True or false? (Source: "Using Multihomed BGP Networks")
- A. true
 - B. false
5. Which three of the following are potential customer routing policies? (Choose three.) (Source: "Using Multihomed BGP Networks")
- A. One service provider is designated as primary, and the other is a backup.
 - B. Traffic is load-balanced across both ISP networks.
 - C. Traffic toward a specific destination goes through only one of the ISPs.
 - D. Traffic to direct customers of the ISPs goes direct; all other traffic goes through the primary ISP.
 - E. The two ISPs may have similar peering agreements with other ISPs.
 - F. The ISPs use default routing to the customer, and the customer uses static routing to the ISP (or ISPs).
6. Which statement about the need to influence BGP route selection in a service provider environment is accurate? (Source: "Using Multihomed BGP Networks")
- A. If both ISP connections terminate in one single customer router, only some routes that are received from the primary ISP can be assigned a BGP weight.
 - B. In most cases, it is more optimal to reach other customers connected to the backup ISP via the backup link, compared with reaching them via the primary link.
 - C. All routes that are received from the primary ISP over the primary link are assigned a local preference value, which is lower than the default value of 100.
 - D. When one ISP connection terminates in one single customer router, all routes that are received from that ISP are assigned a BGP weight.
7. Customer could become transit AS if no filtering to multihomed ISPs is performed. True or false? (Source: "Using Multihomed BGP Networks")
- A. true
 - B. false
8. Which routes should customer announce to service providers in multihomed scenario? (Source: "Using Multihomed BGP Networks")
- A. Customer should announce only networks originating in customer's AS.
 - B. Customer should announce all networks received from primary ISP to secondary ISP.
 - C. Customer should announce all networks received from secondary ISP to primary ISP.
 - D. Customer should not announce any networks.
9. ISP should allow customer to announce any possible networks via BGP. True or false? (Source: "Using Multihomed BGP Networks")
- A. true
 - B. false

10. Which three goals represent appropriate reasons to apply AS path filters? (Choose three.) (Source: "Employing AS Path Filters")
- A. To ensure that only locally originated routes are announced.
 - B. To limit routes that are advertised from IBGP neighbors.
 - C. To select a subset of all routes based on their originating AS.
 - D. To limit neighbor route updates to specific AS-originated routes.
 - E. To ensure that all destination autonomous systems should be received from a specified neighbor.
 - F. To change the weight or local preference attributes for all destination autonomous systems.
11. Which AS path is matched by the regular expression "72\$"? (Source: "Employing AS Path Filters")
- A. 213 72 218 31 727
 - B. 27 317 271 50 72
 - C. 315 27 723 19 91
 - D. 72 591 368 20 87
12. Which regular expression would you use to match all BGP networks originated in AS 89? (Source: "Employing AS Path Filters")
- A. ^89\$
 - B. ^89_
 - C. _89_
 - D. _89\$
13. Which regular expression would you use to match all BGP networks going through AS 76? (Source: "Employing AS Path Filters")
- A. ^76\$
 - B. ^76_
 - C. _76_
 - D. _76\$
14. Which BGP networks are matched by using regular expression ^54\$? (Source: "Employing AS Path Filters")
- A. All routes that are sourced in your directly connected neighboring AS 54.
 - B. All routes that have AS 54 in their AS paths.
 - C. All routes that are sourced in AS 54.
 - D. All the routes that are reachable behind AS 54.

15. What is the difference between the regular expressions "_100_" and "_100\$"? (Source: "Employing AS Path Filters")
- A. The first expression refers to routes that have the substring "100" in their AS paths; the second expression refers only to routes that are directly connected to AS 100.
 - B. The first expression refers to routes that have the substring "100" in their AS paths; the second expression refers only to routes that originated in AS 100.
 - C. The first expression refers to routes that go through AS 100; the second expression refers to routes that originated in AS 100.
 - D. The first expression refers to routes that are directly connected to AS 100; the second expression refers to routes that originated in AS 100.
16. How do you match AS paths that contain exactly two single-digit AS numbers? (Source: "Employing AS Path Filters")
- A. use the expression "***"
 - B. use the expression ".."
 - C. use the expression "[0-9]_[0-9]"
 - D. use the expression "^[0-9]_[0-9]\$"
17. Which three steps are required to apply a new inbound routing policy to a neighbor? (Choose three.) (Source: "Employing AS Path Filters")
- A. Define an AS path access list.
 - B. Attach the AS path filter to inbound or outbound updates for a specific BGP neighbor.
 - C. Send incoming and outgoing AS path filters to the BGP neighbor.
 - D. Force the updates to go through the new filter.
18. How can you test your regular expression? (Source: "Employing AS Path Filters")
- A. **show ip bgp access-list** command
 - B. **show ip bgp filter** command
 - C. **show ip bgp regexp** command
 - D. **show ip bgp summary** command
19. What are two reasons that a multihomed customer needs prefix-lists? (Choose two.) (Source: "Filtering with Prefix-Lists")
- A. To ensure that only valid IP prefixes are announced to the ISPs.
 - B. To set a limit on the number of prefixes that can be accepted from the ISPs.
 - C. To prevent the customer from receiving its own IP prefixes from the ISP.
 - D. To verify that the customer has received full Internet route tables.

20. Which three of the following choices are advantages of prefix lists over access lists? (Choose three.) (Source: "Filtering with Prefix-Lists")
- A. significant performance improvement on long filters
 - B. support for incremental updates
 - C. ability to consist of any number of lines, each of which indicates a test and a result
 - D. flexibility
 - E. more complex command-line interface
 - F. sequential scanning of prefix-lists within Cisco IOS software
21. When you define prefix lists, what are two reasons to use sequence numbers? (Choose two.) (Source: "Filtering with Prefix-Lists")
- A. To reference the associated access list for the prefix list entry.
 - B. To provide a means of linking an AS path filter list to the prefix list.
 - C. To provide an execution order for prefix-list entries.
 - D. To provide a means of inserting or deleting list entries.
22. Which of the following statements is accurate about the **ge** and **le** parameters in the **ip prefix-list** global configuration command required to configure prefix-list filters? (Source: "Filtering with Prefix-Lists")
- A. No match is assumed when neither **ge** nor **le** is specified.
 - B. The range is assumed to be from *ge-value* to 142 only if the **ge** attribute is specified.
 - C. The range is assumed to be from *len* to 32 only if the **le** attribute is specified.
 - D. An exact match is assumed when neither **ge** nor **le** is specified.
23. Which of the following routes will be matched by the the prefix list **10.0.1.0/24 ge 28**? (Source: "Filtering with Prefix Lists")
- A. 10.0.1.0/26
 - B. 10.1.1.0/28
 - C. 10.0.1.0/29
 - D. 10.0.0.0/28
24. Which of the following statements about implementing prefix lists in a BGP network is accurate? (Source: "Filtering with Prefix Lists")
- A. You can optionally apply filter lists and prefix lists on either incoming or outgoing neighbors in any combination.
 - B. You can optionally apply filter lists and prefix lists only on outgoing neighbors in any combination
 - C. Either the incoming prefix list or the incoming filter list must permit the routes that are received from a neighbor before they are accepted into the BGP table.
 - D. Outgoing routes must pass the outgoing prefix list before being transmitted to the neighbor.
25. Which three steps are required to apply a new inbound routing policy (based on prefix list) to a neighbor? (Choose three.) (Source: "Filtering with Prefix Lists")
- A. Define an IP prefix list.
 - B. Attach the prefix list to inbound or outbound updates for a specific BGP neighbor.
 - C. Send incoming and outgoing AS path filters to the BGP neighbor.
 - D. Force the updates to go through the new filter.

26. How can you apply the same prefix list to multiple BGP neighbors on a router? (Source: "Filtering with Prefix Lists")
- A. By configuring a neighbor prefix-list statement for each BGP peer.
 - B. By configuring a neighbor distribute-list statement for each neighbor.
 - C. By using the BGP *peer-group* option with the **neighbor** statement.
 - D. By configuring the prefix list as a global filter under the BGP routing process.
27. How can you use the **show ip prefix-list** command to display the prefix list entry that matches a specific prefix and length? (Source: "Filtering with Prefix Lists")
- A. not a feature of the **show ip prefix-list** command
 - B. by specifying the **detail** keyword
 - C. with the longer keyword to display all matches except those with more specific entries
 - D. by specifying the **first-match** keyword
28. Which of the following best describes the capabilities of the proprietary ORF type that is supported on Cisco routers? (Source: "Using Outbound Route Filtering")
- A. standard BGP communities filtering
 - B. extended BGP communities filtering
 - C. AS path filtering
 - D. prefix list filtering
29. ORF feature is more effective than standard inbound filtering when it comes to the amount of resources that are required for routing update generation and processing. True or false? (Source: "Using Outbound Route Filtering")
- A. true
 - B. false
30. What are two key benefits to using outbound route filtering? (Choose two.) (Source: "Using Outbound Route Filtering")
- A. conserves CPU cycles
 - B. improves security
 - C. reduces bandwidth that is used by unnecessary routing updates
 - D. increases neighbor availability
31. Which two of the following statements about BGP prefix-based outbound route filtering are accurate? (Choose two.) (Source: "Using Outbound Route Filtering")
- A. Uses BGP ORF send and receive capabilities to minimize the number of BGP updates that are sent between BGP peers.
 - B. Can limit the number of unwanted routing updates.
 - C. Increases the amount of resources required to receive and discard routes that would otherwise be filtered out.
 - D. Can be used to increase the amount of processing on a router that is not accepting full routes from a service provider network.

32. How should you configure the **neighbor capability orf prefix-list** command on a router that is applying a prefix list filter as an outbound route policy? (Source: "Using Outbound Route Filtering")
- A. send
 - B. receive
 - C. both
 - D. prefix filter
33. Which of the following is true regarding ORF capabilities options? (Choose two.) (Source: "Using Outbound Route Filtering")
- A. "both" option allows sending and receiving of prefix lists.
 - B. "send" option allows receiving of prefix lists.
 - C. "receive" option allows receiving of prefix lists.
 - D. "both" option allows only sending of prefix lists.
34. What are two methods of determining that a router has ORF capabilities exchange configured? (Choose two.) (Source: "Using Outbound Route Filtering")
- A. **show running-config | begin bgp** command
 - B. **show ip bgp negotiate** command
 - C. **show ip bgp neighbors** command
 - D. **show ip prefix-list** command
35. Which command should be used to verify that IP prefix-list was received on the ORF receiver? (Source: "Using Outbound Route Filtering")
- A. **show ip bgp neighbors IP_address received prefix-filter**
 - B. **show ip prefix-list**
 - C. **show ip bgp neighbors IP_address send prefix-filter**
 - D. **show orf prefix-list received**
36. What are two prerequisites before you can configure ORF prefix list functionality? (Choose two.) (Source: "Using Outbound Route Filtering")
- A. A route refresh must be sent using the **clear ip bgp** command.
 - B. A BGP peering session between the ORF routers must be up and running.
 - C. ORF capabilities must be enabled on both routers.
 - D. You must configure a prefix list filter on the receiving router.
37. Which of the following statements about the function of a route map is accurate? (Source: "Applying Route Maps as BGP Filters")
- A. A route map cannot be used to modify attributes of the permitted routes.
 - B. A route map is a filter that uses a series of match conditions, and that which is denied by the route map is dropped.
 - C. A route map is less complex than the access list.
 - D. Each route map statement starts with a series of match clauses.

38. Which three of the following statements are accurate about the BGP Route Map Policy List Support feature? (Choose three.) (Source: "Applying Route Maps as BGP Filters")
- A. The BGP Route Map Policy List Support feature allows a network operator to group route-map **match** clauses into named lists called policy-lists.
 - B. The network operator manually reconfigures each recurring group of match clauses that occur in multiple route-map entries.
 - C. The AND and OR semantics in the route map function differently for policy-lists than for **match** and **set** clauses.
 - D. To create a BGP policy-list, use the **ip policy-list** command in policy-map configuration mode.
 - E. To configure a route- map to evaluate and process a BGP policy-list in a route map, use the **match policy-list** command in route-map configuration mode.
 - F. To display information about a configured policy-list and policy-list entries, use the **show ip policy-list** command in route-map configuration mode.
39. Which two of the following are functions of the BGP Route Map Continue feature? (Choose two.) (Source: "Applying Route Maps as BGP Filters")
- A. Provides the ability to pause if a sequence number is not specified.
 - B. Provides the capability to execute additional entries in a route-map after an entry is executed with successful **match** and **set** clauses.
 - C. Allows modularization of network policy configuration so that repeated policy definitions can be expanded within the same route map.
 - D. Allows configuration and organization of more modular policy definitions to reduce the number of policy configurations that are repeated within the same route map.
40. Which of the following commands is used to distribute any routes that have a destination network number address that is permitted by a standard access list, an extended access list, or a prefix list, or to perform policy routing on packets? (Source: "Applying Route Maps as BGP Filters")
- A. **match ip next-hop**
 - B. **match ip route-source**
 - C. **match ip address**
 - D. **show ip bgp route-map**
41. How do you implement a "**permit all**" statement when you are using route maps? (Source: "Applying Route Maps as BGP Filters")
- A. By default, a route- map has an "implicit permit any" statement if no match is found.
 - B. You must configure a route map with a "**permit**" parameter and no match clause.
 - C. You must configure a route map with a "**deny**" parameter and a "deny none" clause.
 - D. You must configure a route map with a "**permit any**" match clause.
42. What happens to incoming BGP updates that do not match any route-map match clauses? (Source: "Applying Route Maps as BGP Filters")
- A. They are entered into the BGP table.
 - B. They are entered into the BGP table and marked with a weight of 32768.
 - C. They are not accepted by the router or entered into the BGP table.
 - D. They are entered into the BGP table if a matching route exists in the IP routing table.

43. Which three BGP attributes can you set using route maps? (Choose three.) (Source: "Applying Route Maps as BGP Filters")
- A. MED
 - B. path origin
 - C. administrative distance
 - D. weight metric
 - E. next-hop
 - F. atomic aggregate
44. What are two reasons for using route map sequence numbers? (Choose two.) (Source: "Applying Route Maps as BGP Filters")
- A. To allow insertion or deletion of route-map entries.
 - B. To order the execution sequence of route-map match clauses.
 - C. To provide an ordered execution sequence for the route map.
 - D. To map between prefix-list statements and route-map match clauses.
45. Which command should be used on R1 (IP address 1.1.1.1) to apply a route map named Filter to incoming BGP updates coming from R2 (IP address 2.2.2.2)? (Source: "Applying Route Maps as BGP Filters")
- A. **neighbor 2.2.2.2 route-map Filter in**
 - B. **neighbor 1.1.1.1. route-map Filter in**
 - C. **neighbor 2.2.2.2 route-map Filter out**
 - D. **neighbor 1.1.1.1. route-map Filter out**
46. Why is clearing a BGP session a disruptive change in routing policy? (Source: "Implementing Changes in BGP Policy")
- A. Clearing a BGP session takes a long time and can disrupt packet forwarding.
 - B. You cannot recover information that is sent while the BGP session is being cleared.
 - C. You cannot automatically re-establish sessions that are torn down during the clearing operation.
 - D. You cannot selectively tear down BGP sessions; you must clear sessions with all neighbors.
47. Using the **clear ip bgp *** command, all BGP sessions will stay up and no BGP routes are lost. True or false? (Source: "Implementing Changes in BGP Policy")
- A. true
 - B. false
48. Which two of the following are functions of the BGP Soft Reset Enhancement feature? (Choose two.) (Source: "Implementing Changes in BGP Policy")
- A. allows dynamic route refresh requests
 - B. requires no preconfiguration
 - C. provides newer method for inbound soft reset that uses stored inbound routing table updates
 - D. uses expanded memory

49. BGP route refresh is BGP capability that allows a BGP router to request retransmission of BGP routes from a neighbor. True or false? (Source: "Implementing Changes in BGP Policy")
- A. true
 - B. false
50. Which of the following statements about the command that is required to perform a route refresh is accurate? (Source: "Implementing Changes in BGP Policy")
- A. You will use the **clear ip bgp * in** command to send a route refresh message to all neighbors.
 - B. You will use the **clear ip bgp ip-address in** command to send a route refresh message to all neighbors.
 - C. You must use the **soft** keyword with the **clear ip bgp** command because soft reset is not automatically assumed when the route refresh capability is supported.
 - D. The **clear ip bgp** command works even if the neighbor has not previously advertised the route refresh capability.
51. Which command would you use on R1 (IP address 1.1.1.1) to request from neighboring R2 (IP address 2.2.2.2) to resend routing information? (Source: "Implementing Changes in BGP Policy")
- A. **clear ip bgp 1.1.1.1 out**
 - B. **clear ip bgp 1.1.1.1 in**
 - C. **clear ip bgp 2.2.2.2 out**
 - D. **clear ip bgp 2.2.2.2 in**
52. Which RFC 2918 compliant capability code 2 used by all recent BGP implementations will tell you that router supports the Route Refresh feature? (Source: "Implementing Changes in BGP Policy")
- A. 1
 - B. 218
 - C. 4
 - D. 2
53. Which command should be used to display the negotiation of BGP capabilities after session reset? (Source: "Implementing Changes in BGP Policy")
- A. **debug ip**
 - B. **debug bgp packets**
 - C. **debug ip bgp updates**
 - D. **debug ip bgp**
54. How do you determine whether a BGP neighbor supports route refresh? (Source: "Implementing Changes in BGP Policy")
- A. A flag in the BGP table indicates the presence of route refresh capability
 - B. The **show ip bgp neighbor** command indicates whether the option is supported.
 - C. Initiate the **debug ip bgp negotiation** command to see whether the router has completed a route refresh capabilities exchange.
 - D. Execute the **clear ip bgp *** command. Command-line BGP status messages will indicate route refresh support capabilities.

Module Self-Check Answers

Answer Key

1. A, C
2. B, C
3. A
4. B
5. A, C, D
6. B
7. A
8. A
9. B
10. A, C, D
11. B
12. D
13. C
14. A
15. C
16. D
17. A, B, D
18. C
19. A, C
20. A, B, D
21. C, D
22. D
23. C

- 24. A
- 25. A, B, D
- 26. C
- 27. D
- 28. D
- 29. A
- 30. A, C
- 31. A, B
- 32. B
- 33. A, C
- 34. A, C
- 35. A
- 36. B, C
- 37. B
- 38. A, D, E
- 39. B, D
- 40. C
- 41. B
- 42. C
- 43. A, D, E
- 44. A, C
- 45. A
- 46. A
- 47. B
- 48. A, B
- 49. A
- 50. A
- 51. D
- 52. D
- 53. D
- 54. B

Glossary

ACK

acknowledgment. Notification sent from one network device to another to acknowledge that some event occurred (for example, the receipt of a message).

AD

administrative distance. Rating of the trustworthiness of a routing information source. Administrative distance often is expressed as a numerical value between 0 and 255. The higher the value, the lower the trustworthiness rating.

AFI

Address Family Identifier.

ARIN

American Registry for Internet Numbers. A nonprofit organization that administers and registers IP numbers for the geographical areas that are currently managed by Network Solutions (InterNIC). Those areas include, but are not limited to, North America, South America, South Africa, and the Caribbean.

ARP

Address Resolution Protocol. Internet protocol that is used to map an IP address to a MAC address. Defined in RFC 826.

AS

autonomous system. A collection of networks under a common administration sharing a common routing strategy. Autonomous systems are subdivided by areas. An autonomous system must be assigned a unique 16-bit number by the IANA.

ATM

Asynchronous Transfer Mode. The international standard for cell relay in which multiple service types (such as voice, video, or data) are conveyed in fixed-length (53-byte) cells. Fixed-length cells allow cell processing to occur in hardware, thereby reducing transit delays. ATM is designed to take advantage of high-speed transmission media, such as E3, SONET, and T3.

BGP

Border Gateway Protocol. Interdomain routing protocol that replaces EGP. BGP exchanges reachability information with other BGP systems. It is defined by RFC 1163.

BGP4

Border Gateway Protocol version 4.

CLI

command-line interface. An interface that allows the user to interact with the operating system by entering commands and optional arguments. The UNIX operating system and DOS provide CLIs.

CPU

central processing unit. The hardware within a computer system or smartphone that carries out the instructions of a computer program by performing the basic arithmetical, logical, and input-output operations of the system.

DMVPN

Dynamic Multipoint VPN.

DoS

denial of service. An intentional or unintentional attack on a device that makes the resource unavailable to perform its normal function.

DPT

Dynamic Packet Transport

EBGP

Exterior Border Gateway Protocol

EGP

Exterior Gateway Protocol. It's the Internet protocol for exchanging routing information between autonomous systems. Documented in RFC 904. This is not to be confused with the general term *exterior gateway protocol*. EGP is an obsolete protocol that was replaced by BGP.

EIGRP

Enhanced Interior Gateway Routing Protocol. It's the advanced version of IGRP developed by Cisco. It provides superior convergence properties and operating efficiency, and it combines the advantages of link-state protocols with those of distance vector protocols.

Ethernet

Baseband LAN specification invented by Xerox Corporation and developed jointly by Xerox, Intel, and Digital Equipment Corporation. Ethernet networks use CSMA/CD and run over a variety of cable types at 10 Mbps. Ethernet is similar to the IEEE 802.3 series of standards. It is the most commonly used LAN technology because its protocol is easy to understand, implement, manage, and maintain. It allows low-cost network implementations, provides extensive topological flexibility for network installation, and guarantees successful interconnection and operation of standards-compliant products, regardless of manufacturer.

Fast Ethernet

Any of a number of 100-Mbps Ethernet specifications. Fast Ethernet offers a speed increase 10 times that of the 10BaseT Ethernet specification while preserving such qualities as frame format, MAC mechanisms, and MTU. Such similarities allow the use of existing 10BaseT applications and network management tools on Fast Ethernet networks. Based on an extension to the IEEE 802.3 specification.

FIB

forwarding information base.

Gigabit Ethernet

Standard for a high-speed Ethernet, approved by the IEEE (Institute of Electrical and Electronics Engineers) 802.3z standards committee in 1996.

IANA

Internet Assigned Numbers Authority. Organization operated under the auspices of the ISOC as a part of the IAB. IANA delegates authority for IP address-space allocation and domain-name assignment to the InterNIC and other organizations. IANA also maintains a database of assigned protocol identifiers that is used in the TCP/IP stack, including autonomous system numbers.

IBGP

Internal Border Gateway Protocol.

ICMP

Internet Control Message Protocol. Network layer Internet protocol that reports errors and provides other information that is relevant to IP packet processing. Documented in RFC 792.

IGP

interior gateway protocol. Internet protocol used to exchange routing information within an autonomous system. Examples of common Internet IGPs include IGRP, OSPF, and RIP.

IOL

Special build of Cisco IOS Software for Linux, created specially for virtualized environments.

IP

Internet Protocol. Network layer protocol in the TCP/IP stack offering a connectionless internetwork service. IP provides features for addressing, type-of-service specification, fragmentation and reassembly, and security. Defined in RFC 791.

IP address

A 32-bit address assigned to hosts using TCP/IP. An IP address belongs to one of five classes (A, B, C, D, or E) and is written as 4 octets separated by periods (dotted decimal format). Each address consists of a network number, an optional subnetwork number, and a host number. The network and subnetwork numbers together are used for routing, and the host number is used to address an individual host within the network or subnetwork. A subnet mask is used to extract network and subnetwork information from the IP address. CIDR provides a new way of representing IP addresses and subnet masks. Also called an Internet address.

IPv4

IP version 4. Internet Protocol version 4 is the fourth version in the development of the IP and the first version of the protocol to be widely deployed. Along with IPv6, IPv4 is at the core of standards-based internetworking methods of the Internet. IPv4 is still used to route most traffic across the Internet. IPv4 is a connectionless protocol for use on packet-switched link layer networks (for example, Ethernet). It operates on a best-effort delivery model in that it does not guarantee delivery and does not assure proper sequencing or avoidance of duplicate delivery.

IS-IS

Intermediate System-to-Intermediate System. OSI link-state hierarchical routing protocol based on DECnet Phase V routing, whereby ISs (routers) exchange routing information based on a single metric to determine network topology.

ISP

Internet service provider. Company that provides Internet access to other companies and individuals.

LAN

local-area network. A high-speed, low-error data network covering a relatively small geographic area (up to a few thousand meters). LANs connect workstations, peripherals, terminals, and other devices in a single building or other geographically limited area. LAN standards specify cabling and signaling at the physical and data link layers of the OSI model. Ethernet, FDDI, and Token Ring are widely used LAN technologies.

LSA

link-state advertisement. A broadcast packet used by link-state protocols that contains information about neighbors and path costs. LSAs are used by the receiving routers to maintain their routing tables. Sometimes called an LSP.

MAC

Media Access Control. The lower of the two sublayers of the data link layer that is defined by the IEEE. The MAC sublayer handles access to shared media, such as whether token passing or contention will be used.

MAC address

a standardized data link layer address that is required for every port or device that connects to a LAN. Other devices in the network use these addresses to locate specific ports in the network and to create and update routing tables and data structures. A MAC address is 6 bytes long and is controlled by the IEEE. It is also known as a hardware address, MAC layer address, and physical address.

MD5

Message Digest 5. A one-way hashing algorithm that produces a 128-bit hash. Both MD5 and Secure Hash Algorithm (SHA) are variations on MD4 and are designed to strengthen the security of the MD4 hashing algorithm. Cisco uses hashes for authentication within the IPsec framework. MD5 is also used for message authentication in SNMP v.2. MD5 verifies the integrity of the communication, authenticates the origin, and checks for timeliness.

MED

multi-exit discriminator.

NLRI

Network Layer Reachability Information.

ORF

outbound route filter.

OSPF

Open Shortest Path First. Link-state, hierarchical IGP routing algorithm proposed as a successor to RIP in the Internet community. OSPF features include least-cost routing, multipath routing, and load balancing. OSPF was derived from an early version of the IS-IS protocol.

RFC

Request for Comments. Document series that is used as the primary means for communicating information about the Internet. Some RFCs are designated by the IAB as Internet standards. Most RFCs document protocol specifications, such as Telnet and FTP, but some RFCs are humorous or historical. RFCs are available online from numerous sources.

RIPE

Reseaux IP Europeens. Group formed to coordinate and promote TCP/IP-based networks in Europe.

RST

reset. A type of message.

SAFI

Subsequent Address Family Identifier.

SYN

synchronization.

SYN-ACK

synchronization-acknowledgment.

TCP

Transmission Control Protocol. Connection-oriented transport layer protocol that provides reliable full-duplex data transmission. TCP is part of the TCP/IP protocol stack.

TCP/IP

Transmission Control Protocol/Internet Protocol. Common name for the suite of protocols developed by the U.S. DoD in the 1970s to support the construction of worldwide internetworks. TCP and IP are the two best-known protocols in the suite.

TLV

type, length, value.

TTL

Time to Live. A mechanism that limits the lifespan or lifetime of data in a computer or network.

UNIX

operating system developed in 1969 at Bell Laboratories. UNIX has gone through several iterations since its inception. These include UNIX 4.3 BSD (Berkeley Standard Distribution), developed at the University of California at Berkeley, and UNIX System V, Release 4.0, developed by AT&T.

