

Maintaining Cisco Service Provider Routing Protocols

Volume 2

Version 1.0

Student Guide



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.

 CCDE, CCENT, CCSI, Cisco Eos, Cisco Explorer, Cisco HealthPresence, Cisco IronPort, the Cisco logo, Cisco Nurse Connect, Cisco Pulse, Cisco SensorBase, Cisco StackPower, Cisco StadiumVision, Cisco TelePresence, Cisco TrustSec, Cisco Unified Computing System, Cisco WebEx, DCE, Flip Channels, Flip for Good, Flip Mino, Flipshare (Design), Flip Ultra, Flip Video, Flip Video (Design), Instant Broadband, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn, Cisco Capital, Cisco Capital (Design), Cisco Financial (Stylized), Cisco Store, Flip Gift Card, and One Million Acts of Green are service marks; and Access Registrar, Aironet, AllTouch, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Lumin, Cisco Nexus, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, Continuum, EtherFast, EtherSwitch, Event Center, Explorer, Follow Me Browsing, GainMaker, iLUNIX, IOS, iPhone, IronPort, the IronPort logo, Laser Link, LightStream, Linksys, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, PCNow, PIX, PowerKEY, PowerPanels, PowerTV, PowerTV (Design), PowerVu, Prisma, ProConnect, ROSA, SenderBase, SMARTnet, Spectrum Expert, StackWise, WebEx, and the WebEx logo are registered trademarks of Cisco and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website 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. (1002R)

DISCLAIMER WARRANTY: THIS CONTENT IS BEING PROVIDED “AS IS.” 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.

Table of Contents

Volume 2

<i>Change Management for Routing Protocols</i>	2-1
Overview	2-1
Module Objectives	2-1
Understanding Routing Protocol Requirements	2-3
Overview	2-3
Objectives	2-3
Overview of Routing Protocols	2-4
Core IGP	2-19
Internal BGP	2-27
Static and Connected Routing	2-46
External BGP	2-56
Summary	2-73
Lesson Self-Check	2-74
Lesson Self-Check Answer Key	2-75
Adding Routers to the Core Network	2-77
Overview	2-77
Objectives	2-77
Change Management for Adding New Core Routers	2-78
Enabling and Tuning OSPF	2-95
Enabling and Tuning IS-IS	2-108
Enabling and Tuning IBGP	2-122
Summary	2-128
Lesson Self-Check	2-129
Lesson Self-Check Answer Key	2-130
Modifying Core Routing Parameters	2-131
Overview	2-131
Objectives	2-131
Modifying Routing Parameters	2-132
Adding New Prefixes	2-141
Maintaining OSPF and IS-IS Filters	2-149
Tuning Routing Protocol Parameters	2-157
Summary	2-173
Lesson Self-Check	2-174
Lesson Self-Check Answer Key	2-175
Connecting Single-Homed Customers to the Network	2-177
Overview	2-177
Objectives	2-177
Single-Homed Customer Requirements	2-178
Customer Access Link	2-184
Customer Addressing and Routing	2-191
Customer Reachability	2-203
Summary	2-210
Lesson Self-Check	2-211
Lesson Self-Check Answer Key	2-212
Connecting Multihomed Customers to the Network	2-213
Overview	2-213
Objectives	2-213
Multihomed Customer Requirements	2-214
Customer Reachability	2-222
Routing Policies	2-234
Route Filtering	2-254
Summary	2-263
Lesson Self-Check	2-264
Lesson Self-Check Answer Key	2-265

Modifying Routing Paths	2-267
Overview	2-267
Objectives	2-267
Influencing Internal Routing Using OSPF	2-268
Influencing Internal Routing Using IS-IS	2-279
Influencing External Routing Using BGP	2-287
Summary	2-299
Lesson Self-Check	2-300
Lesson Self-Check Answer Key	2-301
Module Summary	2-303

Change Management for Routing Protocols

Overview

In this module, you will identify the routing requirements in service provider environments. This module will present a common service provider network and explain the routing options for routing protocols in the core service provider network and on the segment connecting to single-homed and multihomed customers. This module describes different modification and tuning processes for service provider routing protocols. In some cases, default selection of routing paths in a network does not satisfy your requirements. The module will present changes with internal routing protocols and with external routing protocols. In order to efficiently modify routing paths, this module presents the path selection process for each routing protocol and how to implement and configure changes in the path selection.

Module Objectives

Upon completing this module, you will be able to perform changes, in the routing configuration, that are based on designs and implementation templates, as well as Information Technology Infrastructure Library® (ITIL)-based change management processes and procedures. This ability includes being able to meet these objectives:

- Identify the routing requirements in service provider environments
- Identify the tasks that are required to join a new router into the core routing domain
- Identify the tasks that are required to modify core routing parameters in the core routing domain and implement the changes according to the change request
- Identify the tasks that are required to add a single-homed customer to the network and implement the changes according to the change request
- Identify the tasks that are required to add a multihomed customer to the network and implement the changes according to the change request
- Identify the tasks that are required to modify the routing paths in the network using IGPs and BGP and implement the changes according to the change request

Understanding Routing Protocol Requirements

Overview

In this lesson, you will identify the routing requirements in service provider environments. The lesson will present a common service provider network and explain the need for routing protocols in the core service provider network. Routing options between the service provider network and different types of customers will be explained and discussed.

Objectives

Upon completing this lesson, you will be able to identify the routing requirements in service provider environments. This ability includes being able to meet these objectives:

- Describe the characteristics and requirements for routing protocols in service provider environments
- Describe IGP requirements in service provider environments
- Describe internal routing requirements in service provider environments
- Describe static and connected routing requirements in service provider environments
- Describe external routing requirements in service provider environments

Overview of Routing Protocols

This topic describes the characteristics and requirements for routing protocols in service provider environments.

Common Service Provider Network

- Runs BGP or static routing with customer
- Exchanges routes with other service providers via BGP
- Runs IBGP between its own BGP speakers
- Runs one instance of IGP (OSPF or IS-IS):
 - IGP used for internal routes only

© 2010 Cisco Systems, Inc. All rights reserved.

MSRPv1.0-2.3

The common service provider network runs External Border Gateway Protocol (EBGP) or static routing with customers. EBGP is always used as the routing protocol between different service providers.

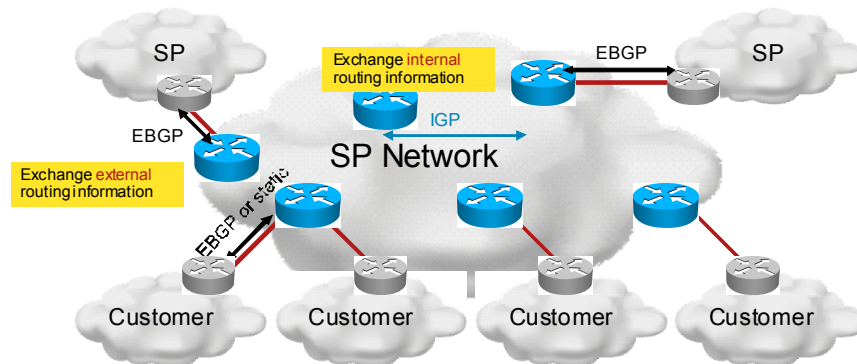
Internal Border Gateway Protocol (IBGP) is required in the provider network because all EBGP-speaking routers in an autonomous system (AS) must exchange external routes via IBGP. Also, non-EBGP speakers are required to take part in the IBGP exchange if they are in a transit path and forward packets that are based on destination IP addresses.

The service provider network also runs an IGP. The protocols of choice are Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS). The IGP is used for two purposes:

- It provides IP connectivity between all IBGP speakers so that TCP sessions for IBGP can be established between Border Gateway Protocol (BGP)-speaking routers.
- It provides optimal routing to the BGP next-hop address.

A single IGP should be used within the entire AS. This setup facilitates effective packet forwarding from the ingress router to egress routers. The IGP is configured to carry internal routes only, including internal links and loopback addresses of the routers. For performance and scalability reasons, no customer routes or external routes should be injected into the IGP.

Common Service Provider Network (Cont.)



© 2010 Cisco Systems, Inc. All rights reserved.

MSRPv1.0-2-4

The typical service provider network consists of a network core that connects various edge devices. Some of the edge devices connect customers; others connect to other service providers.

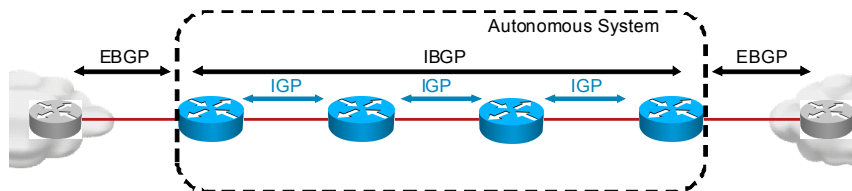
The edge devices that connect to other service providers use EBGP to exchange routing information. The edge devices that connect customers use either static routing or EBGP.

Unless Multiprotocol Label Switching (MPLS) is configured on the service provider backbone, routers in a transit path are also required to have complete routing information. Therefore, these routers take part in the IBGP routing exchange.

An IGP is also required within the service provider network. The IGP is used to carry internal routes, including the loopback interface addresses of IBGP-speaking routers. The IGP provides reachability information to establish IBGP sessions and to perform the recursive routing lookup for the BGP next hop.

Common Service Provider Network (Cont.)

- The provider core IGP is a single instance of IS-IS or OSPF.
- The core IGP is used only within the service provider backbone.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRPv1.0-2-5

The service provider backbone typically uses a single instance of either IS-IS or OSPF as its IGP. The IGP is used within the provider backbone only. The provider backbone exchanges no IGP routing information with customer routers or with routers in other autonomous systems.

Route Propagation in Service Provider Networks

- **BGP route propagation:**
 - BGP carries customer routes.
 - BGP carries other provider routes.
- **IGP route propagation:**
 - IGP is responsible only for the next hop.
- **Do not redistribute BGP into IGP:**
 - IGP performance and convergence time suffer if a large number of routes are carried.
 - No IGP is capable of carrying full Internet routes.
 - A full Internet routing table is huge.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRPv1.0-2-6

It is important to avoid sending any unnecessary routing information in the IGP. The IGP performs best if it carries as few routes as possible. Optimally, the IGP should contain only information about BGP next hops and routes that are internal to the service provider network, enabling the establishment of IBGP sessions.

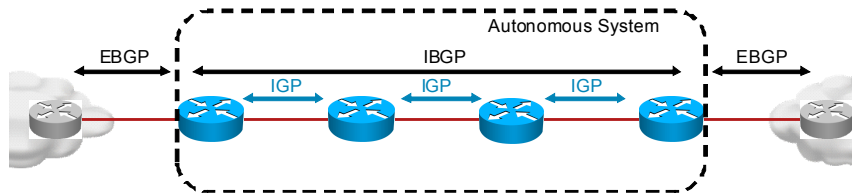
All other routing information should be carried in BGP, which is designed to scale for large volumes of routing information. Customer routes and the routes from other service providers should be carried in BGP. These routes should not be propagated from BGP into the provider IGP.

IGP performance and convergence time suffer if the IGP carries a larger number of routes. The design goal should be to minimize the volume of routing information that the IGP carries. Naturally, the number of route flaps is also reduced when the number of routes is reduced.

BGP scales to a much larger volume of routing information because of inherent BGP design qualities. Potentially, the BGP routers of the service provider can receive the complete Internet routing table, which is huge. You should therefore never redistribute the routing information that has been received by BGP into the IGP, because no IGP is capable of carrying several tens of thousands of routes.

Routing Information Exchange with Other Service Providers

- BGP is used to exchange routing information between Internet service providers.



© 2010 Cisco Systems, Inc. All rights reserved.

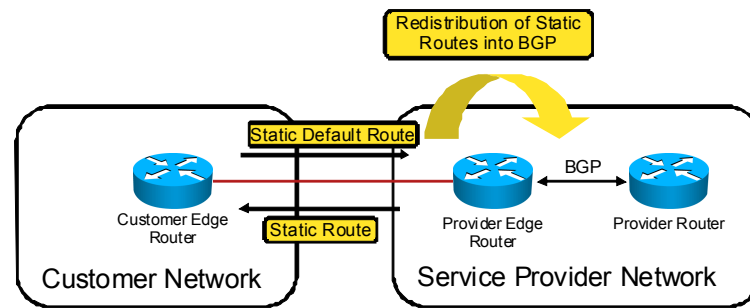
MSPRPv1.0-2.7

Provider edge routers use BGP to exchange routing information with other service provider networks for redundancy and scalability reasons.

Static routing with other service providers is generally not a viable solution due to the dynamic routing requirements of the service provider environment. Routing information is received at provider edge routers using EBGP and then propagated using IBGP to the rest of the service provider network. At another edge router, the routing information is further propagated to a different service provider using EBGP with other autonomous systems.

Routing Information Exchange with Customers

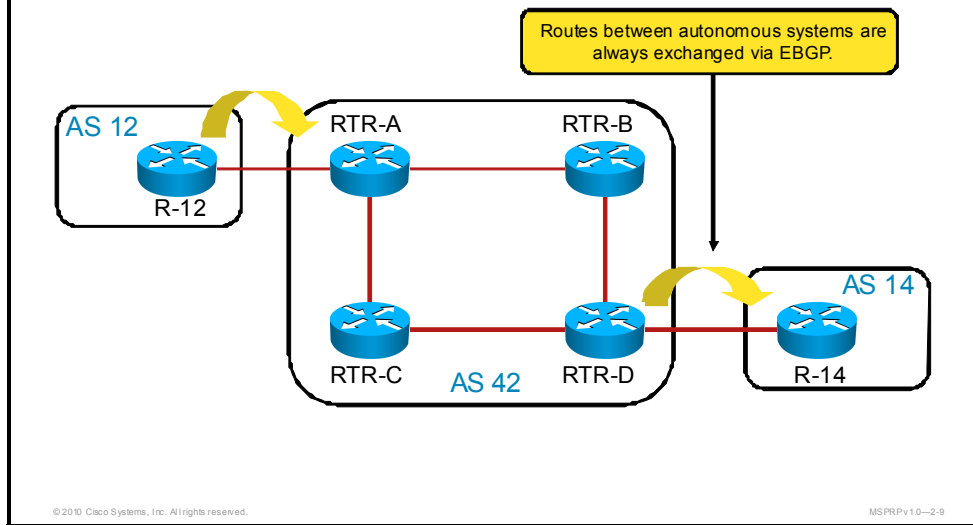
- The provider edge router redistributes static customer routes into BGP.
- BGP carries customer routes.



The provider edge router typically uses static routing to reach customer networks. In this case, the customer typically configures a static default route that points to the edge router of the service provider.

The provider edge router redistributes customer static routes into BGP. The service provider network then uses BGP to propagate the information to the rest of the service provider network using IBGP. The service provider also advertises customer routing information to other autonomous systems using BGP.

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 R-12 and RTR-A is called an EBGP session because R-12 and RTR-A are in different autonomous systems.

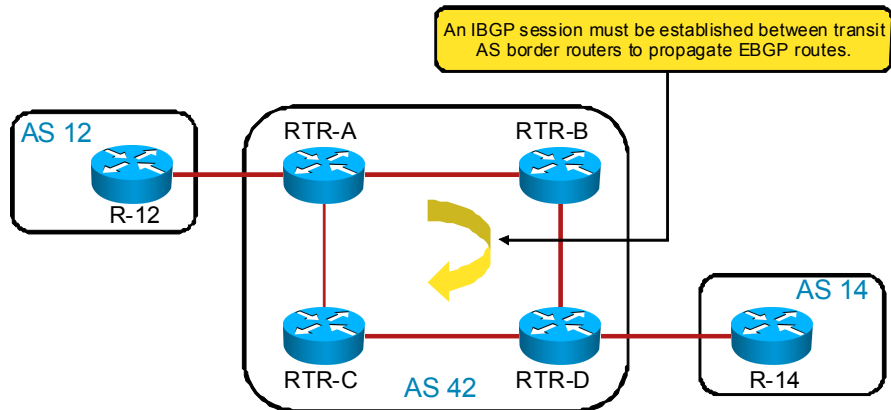
BGP routing information updates consist of the network address, subnet mask, and any number of BGP attributes. No other routing protocol provides the same richness of route attributes as BGP. Translating BGP route attribute information into any other protocol would likely cause a loss of information.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

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

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

Internal Route Propagation

- The only protocol that can transport all BGP attributes across the backbone is IBGP.

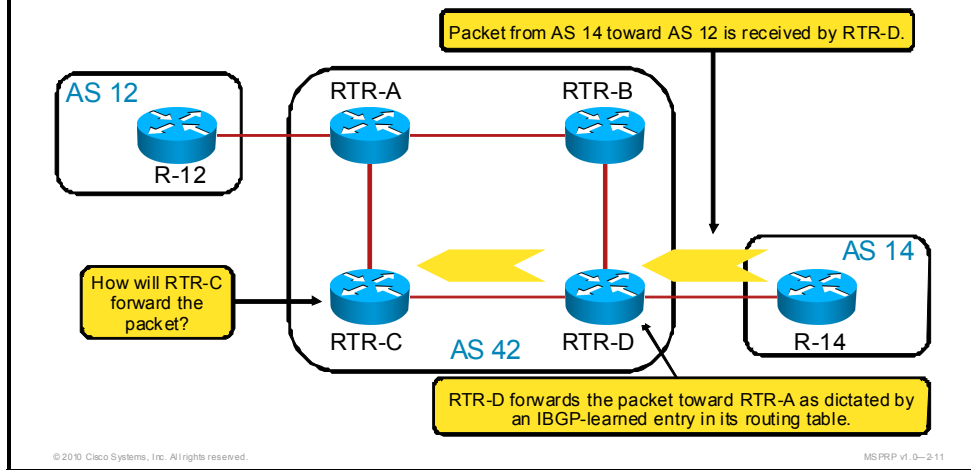


In this example, the BGP session between RTR-A and RTR-D, which are both in the same AS, is an IBGP session.

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

Packet Forwarding in an AS

- RTR-C needs external routes for proper packet forwarding.
- RTR-C must receive BGP routes.



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

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

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

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

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

Core Router IBGP Requirements in a Service Provider AS

- All core routers must have all external routes.
- Core routers must receive BGP routes:
 - Redistribution of BGP routes into IGP is not scalable.
 - Default routing is not applicable in service provider AS core.
 - The only feasible way is by using IBGP.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0--2.12

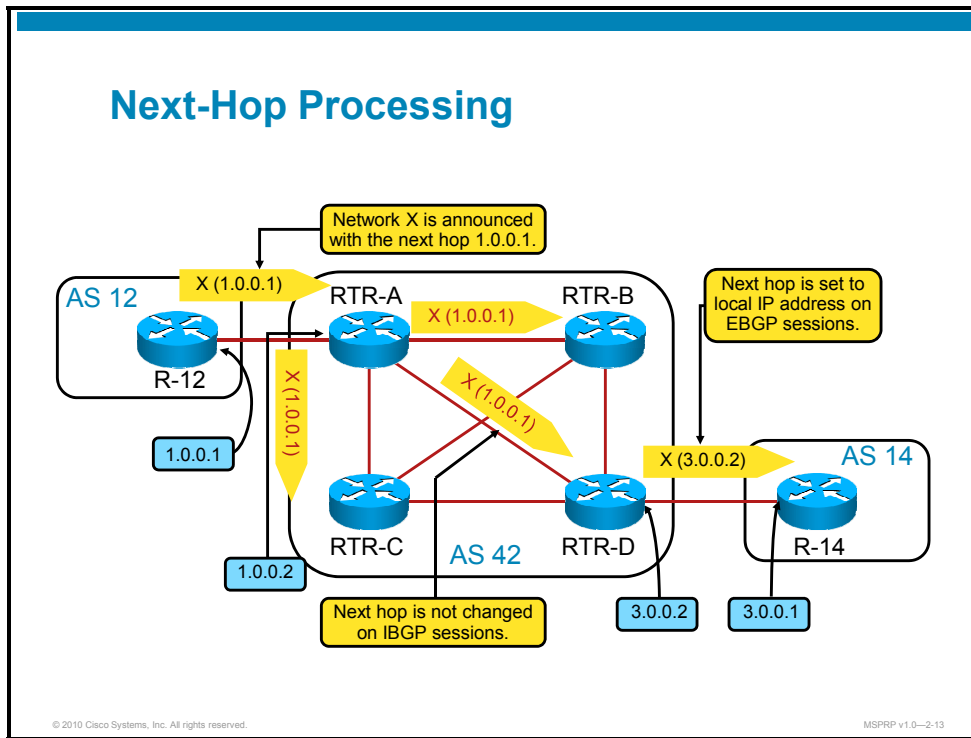
Within a service provider AS, all routers that are in a theoretical transit path between external destinations should have information about all external routes that are received from any neighboring AS. If a single router on a transit path does not have this information, it is possible that an IP packet that is received from a neighboring AS cannot 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.

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

Next-Hop Processing



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

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

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

RTR-A also forwards the BGP update over all its IBGP sessions. It does not change the next-hop attribute, so RTR-B, RTR-C, and RTR-D get information that they can reach network X by forwarding packets to 1.0.0.1. But that IP address is not directly connected, so the routers must look in their routing tables to see if and how they can reach 1.0.0.1. If the recursive route lookup is successful, each router can then route packets to network X in the same direction as it would route packets to 1.0.0.1.

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

Service Provider 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.

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0--2-14

All BGP-speaking routers within the AS receive information about external networks with the next-hop attribute. This attribute is set to the far end of the EBGp sessions reaching the border routers of the AS.

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

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

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

External subnets that are redistributed into the IGP might appear as external IGP routes, depending on what IGP is configured within the AS. There are several scalability issues that are associated with external routes in some routing protocols. For example, Open Shortest Path First (OSPF) carries each external subnet in a separate link-state advertisement (LSA) object. If route redistribution is not desirable for any reason, an alternative method is to include the subnet on which the EBGp session is running in the IGP configuration using the **network** command. To prevent the border router from exchanging IGP routing with the border router of the other AS, you must configure the interface as a passive interface. Failure to do so could cause the two autonomous systems to exchange routes using the IGP. In that case, all benefits of having separate autonomous systems would be lost.

Service Provider 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.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.15

An IGP peer usually does not modify the next-hop attribute when the BGP update is propagated across IBGP sessions. However, you can configure the BGP router to have a different behavior. It can 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. The next-hop attribute will be set to the IP address that is used on the local side of the IBGP session. The original next hop, set by the far end of the EBGp session, will be lost.

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

Service Provider Using Edge Routers as Next Hops (Cont.)

```
Router(config-router)# Cisco IOS Software  
neighbor ip-address next-hop-self
```

```
RP/0/RP0/CPU0:Router(config-bgp-nbr-af)# Cisco IOS XR Software  
next-hop-self
```

- Changes next-hop processing at edge routers
- Bypasses the BGP next-hop processing and announces the local IP address as the BGP next hop in outgoing updates sent to the specified neighbor
- Has to be set on all IBGP neighbors to fully bypass IBGP next-hop processing

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0--2-16

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

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

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

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

Syntax Description

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

To disable next-hop calculation on a Cisco IOS XR router and insert your own address in the next-hop field of BGP updates, use the **next-hop-self** command in an appropriate configuration mode.

- **next-hop-self [disable]**

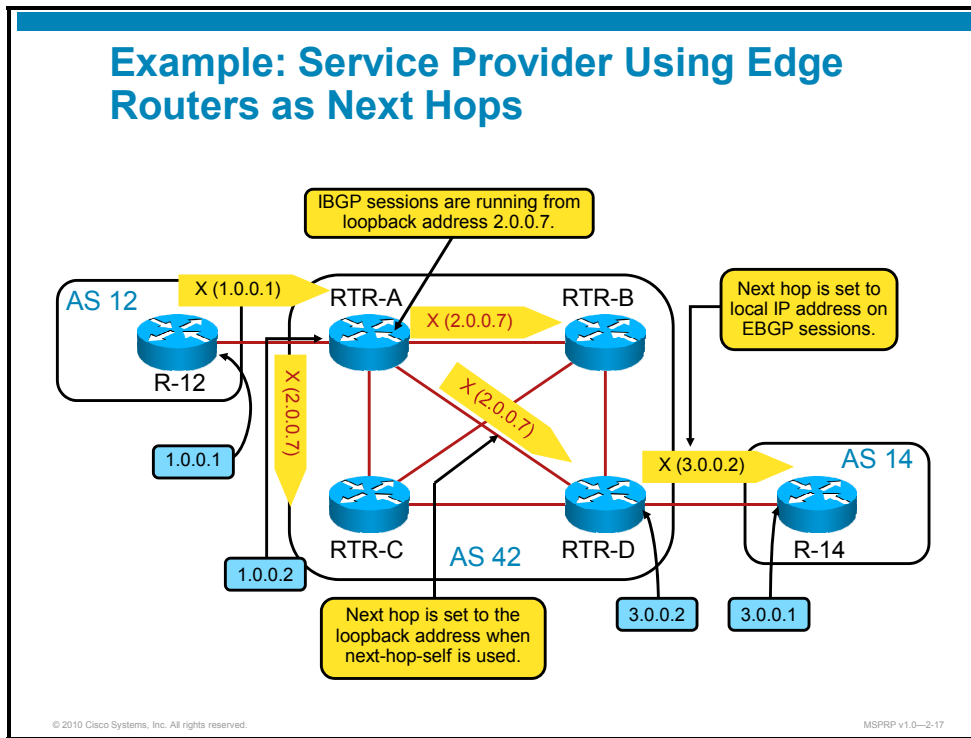
To disable this feature, use the **no** form of this command.

- **no next-hop-self [disable]**

Syntax Description

Parameter	Description
<i>disable</i>	(Optional) Allows a next-hop calculation override when this feature may be inherited from a neighbor group or address family group

Example: Service Provider Using Edge Routers as Next Hops



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 that 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 otherwise, the IBGP session would not have been established.

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

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

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

Core IGP

This topic describes the typical scenarios where IGP routing protocols are used in service provider environments.

Core IGP Requirements

- IGP requirements:
 - Exchange local routes
 - Ensure reachability of BGP next-hop addresses
- User data traffic is forwarded based on FIB table contents:
 - Derived from routing table using recursive lookup.
 - All recursive lookups are performed while the FIB table is built.
 - BGP contains the prefix mapped to a non-adjacent next-hop address.
 - IGP provides routing to BGP next-hop addresses.
- Any IGP can be used in combination with BGP.

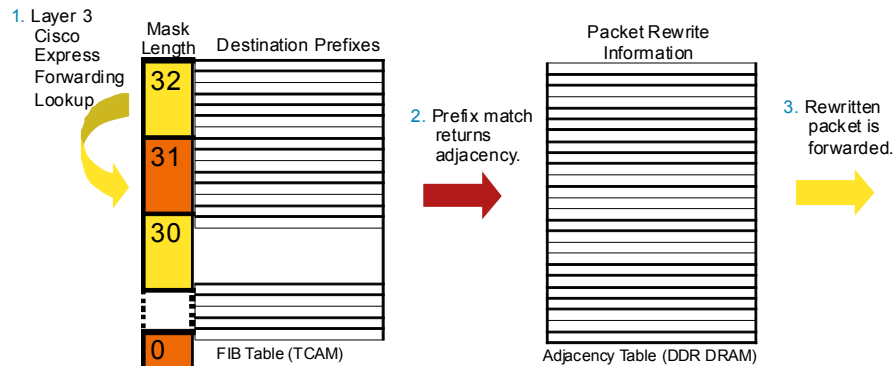
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0--2.18

Basic requirements for core IGP routing protocol is to exchange local routes and ensure reachability of the BGP next-hop IP address. Data traffic is forwarded based on Forwarding Information Base (FIB) table contents, which are derived from a routing table using recursive lookup. Recursive lookup is performed while the FIB table is built. IGP should provide routing to the BGP next-hop IP address. Any IGP can be used in combination with BGP, though normally it is IS-IS or OSPF.

Example: Cisco Express Forwarding on Distributed Cisco 7600 Series Routers

- Layer 3 packets initiate TCAM lookup.
- The longest match returns adjacency with rewrite information.
- The packet is rewritten per adjacency information and forwarded.



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 performed, 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 prebuilds a complete IP forwarding table, called the FIB, which 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. 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.

The following sequence of events describes the operation of Cisco Express Forwarding on a distributed platform (Cisco 7600 Series Routers or Cisco Catalyst 6500 Series Switches).

- The FIB is derived from the IP routing table and is arranged for maximum lookup throughput.
- The adjacency table is derived from the Address Resolution Protocol (ARP) table, and it contains Layer 2 rewrite (MAC rewrite) information for the next hop.
- IP destination prefixes are stored in the ternary content addressable memory (TCAM) table, from the most specific to the least specific entry.
- When the TCAM table is complete, a wildcard entry redirects frames to the Layer 3 engine.
- When the adjacency table is complete, a TCAM table entry points to the Layer 3 engine to redirect the adjacency.
- The FIB lookup is based on the Layer 3 destination address prefix (longest match).

The FIB table is updated when these events occur:

- An ARP entry for the destination next hop changes, ages out, or is removed.
- The routing table entry for a prefix changes.
- The routing table entry for the next hop changes.

These six basic steps form the procedure for initially populating the adjacency table:

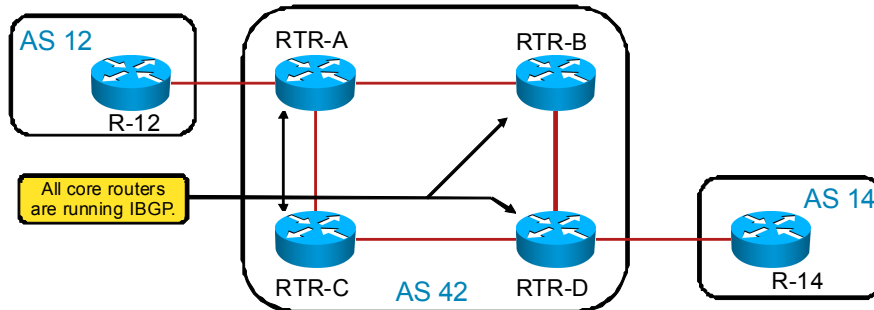
- Step 1** The Layer 3 engine queries the switch for a physical MAC address.
- Step 2** The switch selects a MAC address from the chassis MAC range and assigns it to the Layer 3 engine. The Layer 3 engine assigns this MAC address as a burned-in address for all VLANs and is used by the switch to initiate Layer 3 packet lookups.
- Step 3** The switch installs wildcard FIB entries, which point to drop adjacencies (for processing FIB table lookup misses).
- Step 4** The Layer 3 engine informs the switch of its interfaces participating in MLS (MAC address and associated VLAN). The switch creates the (MAC, VLAN) Layer 2 CAM entry for the Layer 3 engine.
- Step 5** The Layer 3 engine informs the switch about features for interfaces participating in MLS.
- Step 6** The Layer 3 engine informs the switch about all FIB entries that are related to its interfaces and connected networks. The switch populates the FIB entries and points them to Layer 3 engine redirect adjacencies.

Routers using Cisco IOS XR Software use similar distribution of routing tasks between the control plane (route processor [RP]) and the data plane (intelligent line cards). Unlike Cisco IOS Software in which routing mechanisms can be chosen and tuned, Cisco IOS XR Software only supports one routing mechanism, and that is Cisco Express Forwarding.

Routing Protocols in a Service Provider AS

With IBGP running on all core routers, is an IGP still needed in the core?

- An IGP is needed to resolve BGP next hops and to perform fast convergence after a failure in the core network.



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

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

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

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

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

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

Routing Protocols in a Service Provider 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

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2.22

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

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

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

BGP and IGP Interaction

Ideally, there will be no interaction between BGP and the IGP:

- BGP carries external and customer routes.
- The IGP carries only core subnets.
- The IGP is not affected by external route flaps.
- BGP is not affected by failures internal to the network as long as the BGP next hop remains reachable.
- The only link between BGP and the IGP should be the recursive lookup when building the FIB table.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-223

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

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

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

OSPF Requirements

- Traditional design guidelines encourage usage of areas:
 - Summarization to:
 - Limit influence of flaps
 - Minimize routing table size
 - Filtering capabilities on ABRs
 - Contain flooding of LSAs to a smaller number of routers
- Modern networks use the opposite—single-area designs:
 - Support for simplified MPLS TE implementation
 - Optimal routing because each router sees full topology
 - Easier to achieve today:
 - More stable links
 - More CPU power and memory in routers

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0—224

Traditional design guidelines encouraged usage of OSPF areas. By using multiarea design, you can limit the influence of link flaps and minimize routing table size. You can also take advantage of filtering capabilities on Area Border Router (ABR) and contain flooding of link-state advertisements (LSAs) to a smaller number of routers.

On the other hand, modern networks use the opposite approach called single-area design. You can take advantage of optimal routing because each router sees complete topology, which is not the case in single-area design. Nowadays unstable links, CPU power, and memory resource is not such a concern as it was in the past. Single-area OSPF design provides support for simplified Cisco MPLS Traffic Engineering (MPLS TE) implementation.

IS-IS Requirements

- Traditional design guidelines encourage usage of both levels (similar reasons as with OSPF).
- Modern networks tend to use Level 2 only for the same reasons as with OSPF.
- IS-IS is believed to scale slightly better than OSPF.
- The scalability of link-state protocol is primarily affected by the number of routers.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.25

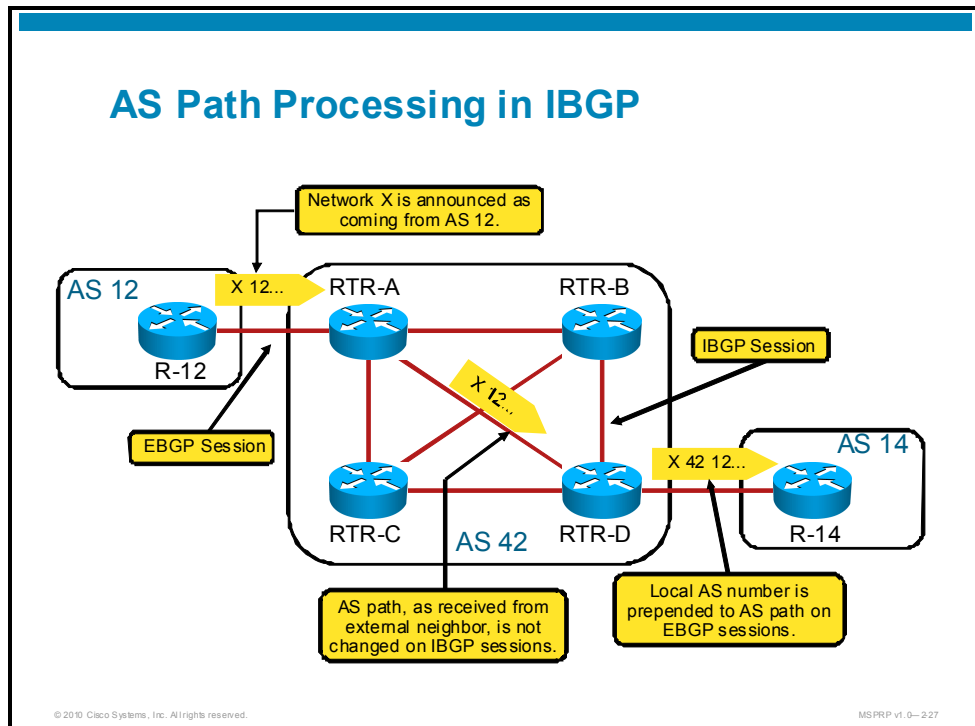
Traditional design guidelines encouraged use of both IS-IS levels. By using design guidelines for both levels, you can limit the influence of link flaps and minimize routing table size.

On the other hand, modern networks take the opposite approach using Level 2-only routers, with which you can take advantage of optimal routing because each router sees complete topology.

Because of its design, IS-IS is believed to scale slightly better than OSPF. IS-IS and OSPF scalability is primarily affected by the number of routers in the network.

Internal BGP

This topic describes the typical scenarios in which internal BGP is used in service provider environments.



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

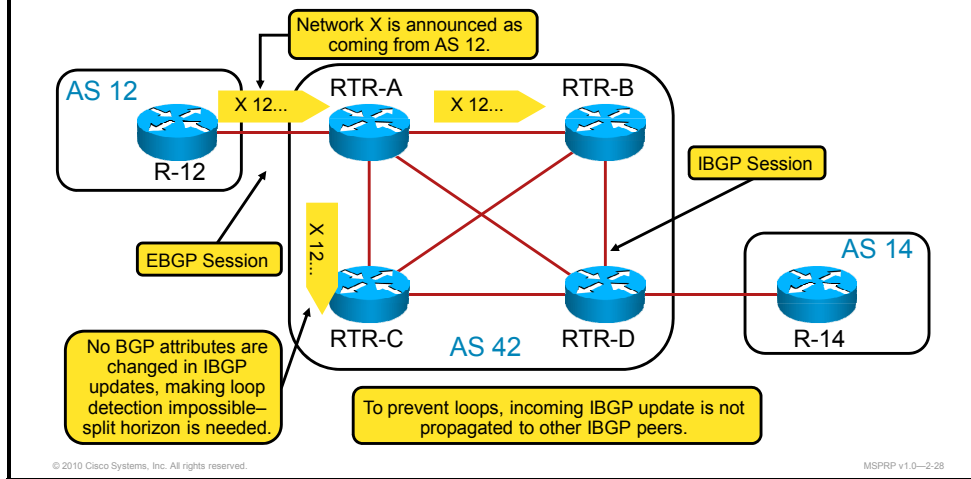
When a router originates a BGP prefix (network X in this example), the AS path is empty. Whenever a BGP prefix is announced over an EBGP session, the AS number of the router that is sending the information is prepended to the AS path. In the example, R-12 inserts "12" in the AS path before forwarding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

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

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

BGP Split Horizon

- Result: Full mesh of IBGP sessions is required for proper IBGP update propagation

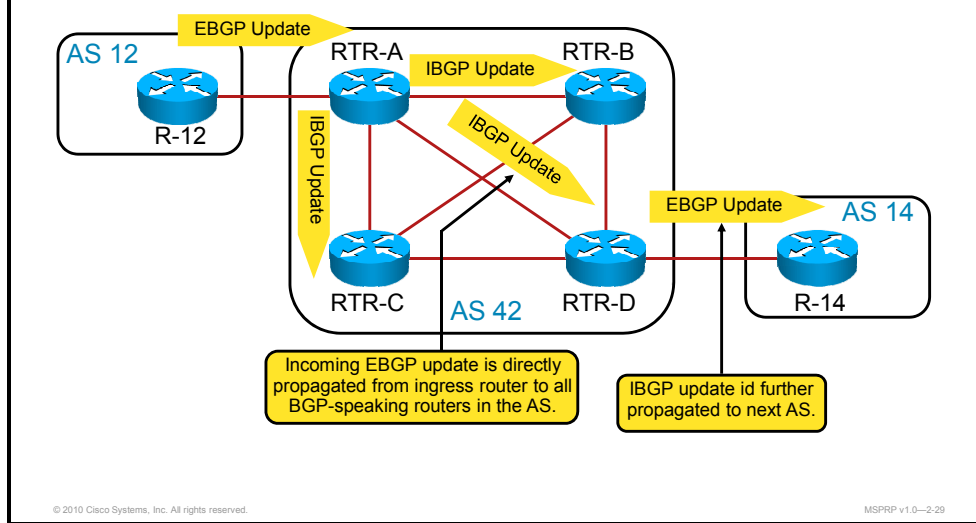


All routers within an AS must make routing decisions in a consistent way. They must have access to the same routing information with the same attributes in order to come to the same conclusion about which exit point of the AS to use. In other words, the BGP attributes should not be changed within the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



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

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

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

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

IBGP Neighbors

- Because of IBGP full-mesh requirements, IBGP neighbors are usually not directly connected.
- 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.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0–2.30

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 that is placed on IGP is that the network must converge before the BGP keepalive timer expires.

IBGP Scalability Issues

- IBGP requires a full mesh between all BGP-speaking routers:
 - Large number of TCP sessions
 - Unnecessary duplicate routing traffic
- Solutions:
 - Route reflectors modify IBGP split-horizon rules.
 - BGP confederations modify IBGP AS path processing.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-231

Classic IBGP split-horizon rules specify that updates that are received on an EBGP session should be forwarded on all IBGP and EBGP sessions. However, updates that are received on an IBGP session should be forwarded only to all EBGP sessions. This rule requires that a BGP boundary router be able to send routing updates to all other BGP-speaking routers in its own AS directly through a separate IBGP session to each of them.

The primary reason for the IBGP split-horizon rule is to avoid routing information loops within the AS. If the information that is received through an IBGP session is forwarded on other IBGP sessions, the information might come back to the originator and be forwarded again in a never-ending loop. The originator would not detect the loop because no BGP attributes are changed on IBGP sessions.

The general design rule in classic IBGP is to have a full mesh of IBGP sessions. But a full mesh of IBGP sessions between n number of routers would require $(n * (n - 1)) / 2$ IBGP sessions. For example, a router with an AS that contains 10 routers would require $(10 * (10 - 1)) / 2 = 45$ IBGP sessions. Imagine the number of sessions (and the associated router configuration) that would be required for a single AS containing 500 routers.

Every IBGP session uses a single TCP session to another IBGP peer. An update that must be sent to all IBGP peers must be sent on each of the individual 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 requirement results in multiplication of the update over the single link.

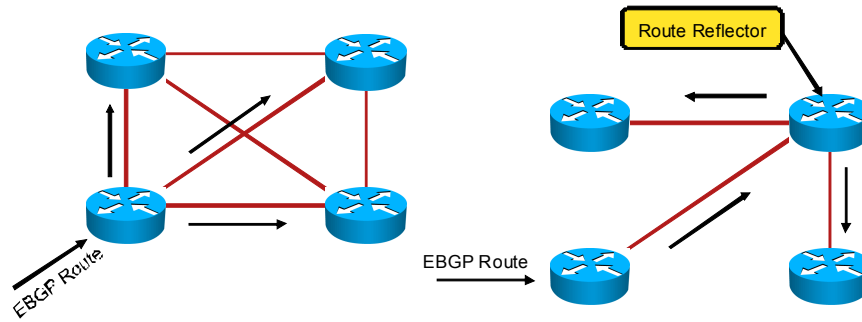
Route reflectors are BGP scalability mechanisms that enable routing information to be redistributed to all routers within an AS while eliminating the need for a fully meshed topology within the AS. This feature reduces the number of TCP sessions that must be maintained, lowering network overhead and CPU and memory resource requirements.

Two different solutions are available to achieve greater scalability when you are faced with the full-mesh rules of IBGP autonomous systems.

- Route reflectors modify the classic IBGP split-horizon rule and allow a particular router to forward incoming IBGP updates to an outgoing IBGP session under certain conditions. This router becomes a concentration router, or a route reflector.
- BGP confederations introduce the concept of a number of smaller autonomous systems within the original AS. The small autonomous systems exchange BGP updates between them using intra-confederation EBGP sessions.

Route Reflector Split-Horizon Rules

- Classic IBGP: IBGP routes are not propagated to other IBGP peers.
- Full mesh of IBGP peers is therefore required.
- Route reflector can propagate IBGP routes to other IBGP peers.
- Full mesh of IBGP peers is no longer required.



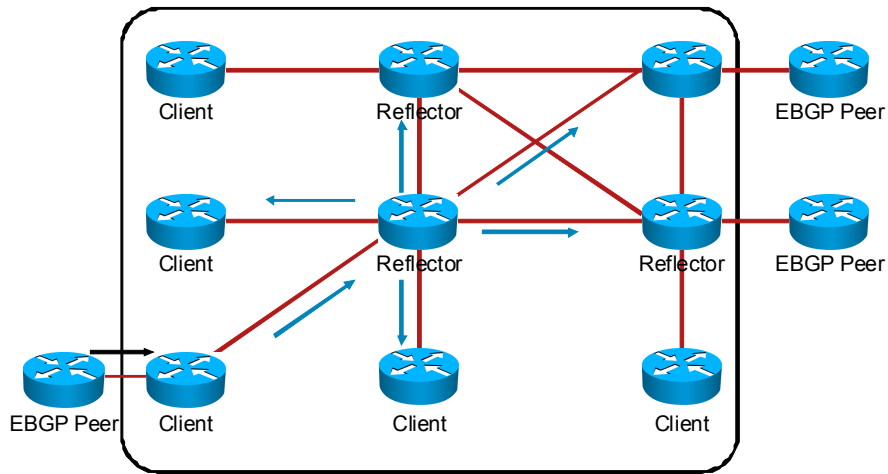
In classic IBGP, the BGP boundary router needs to forward the route that is received from an EBGP peer to every other router within its own AS using a dedicated IBGP session for each one. Also, the BGP boundary router forwards routes that are sourced by a router in the same way. To allow every router to update every other router, a full mesh of IBGP sessions is required.

The IBGP route reflector design relaxes the need for a full mesh. The router that is configured as a route reflector, under certain conditions, will relay updates that are received through an IBGP session to another IBGP session. This capability requires modifications of the classic IBGP split-horizon rules.

The route reflector concept introduces processing overhead on the concentration router and, if it is configured incorrectly, can cause routing loops and instability.

Route Reflector Split-Horizon Rules (Cont.)

- Routes received on reflector from a client are propagated to all other peers.



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.33

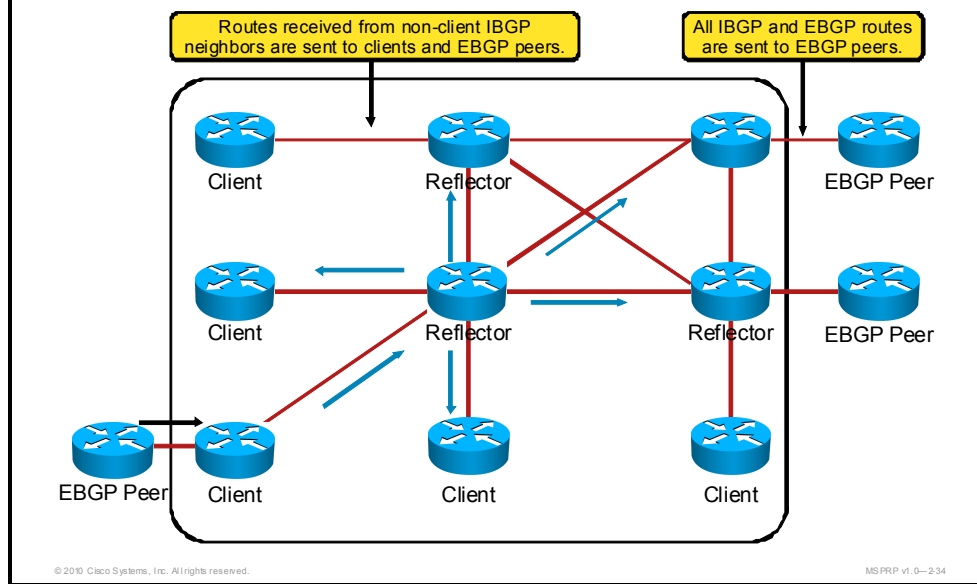
When you implement a route-reflector-based IBGP network, the BGP routers are divided into route reflectors (which implement modified split-horizon rules) and clients (which behave like traditional IBGP routers).

Route reflector clients are excluded from the full mesh. They can have any number of EBGP sessions but may have only one IBGP session, the session with their route reflector. Clients conform to the classic IBGP split-horizon rules and forward a received route from EBGP on their IBGP neighbor sessions. But the route reflector conforms to the route reflector split-horizon rules and recognizes that it has an IBGP session to a client. When the IBGP update is received from the client, the route reflector forwards the update to other IBGP neighbors, therefore alleviating the IBGP full-mesh requirement for its clients.

Similarly, when the route reflector receives an IBGP update from a neighbor that is not its client, it forwards the update to all of its clients.

Forwarding an IBGP update in a route reflector does not change the next-hop attribute or any other common BGP attribute. This feature means that the client will use the optimum route by means of recursive routing, regardless of how it received the BGP route.

Route Reflector Split-Horizon Rules (Cont.)



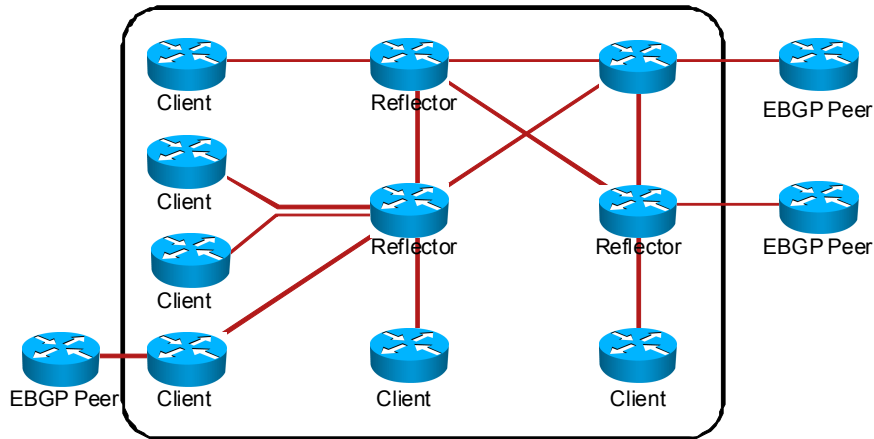
The figure shows how an AS with nine routers running BGP reduces the required number of IBGP TCP sessions from 36 to 11 by using route reflectors.

The table presents detailed IBGP split-horizon rules that are modified by the introduction of BGP route reflectors. For purposes of definition, a route reflector is a BGP speaker that can advertise IBGP learned routes to another IBGP peer and, hence, can reflect routes. IBGP peers of the route reflector fall under two categories: clients and nonclients. The route reflector and its clients form a “cluster.” All IBGP peers of the route reflector that are not part of the cluster are nonclients. A classic IBGP router is a router that does not support route reflector functionality.

Type of Router	Incoming Update From	Is Forwarded To
Classic	EBGP peer	All peers (IBGP and EBGP)
	IBGP peer	EBGP peers
Route reflector	EBGP peer	All peers (IBGP and EBGP)
	Nonclient IBGP peer	EBGP peers and clients
	Client IBGP peer	All peers but the sender
Client	EBGP peer	All peers (IBGP and EBGP)
	IBGP peer	EBGP peers

Redundant Route Reflectors

- The reflector could be a single point of failure.
- Route reflectors must be redundant.



© 2010 Cisco Systems, Inc. All rights reserved.

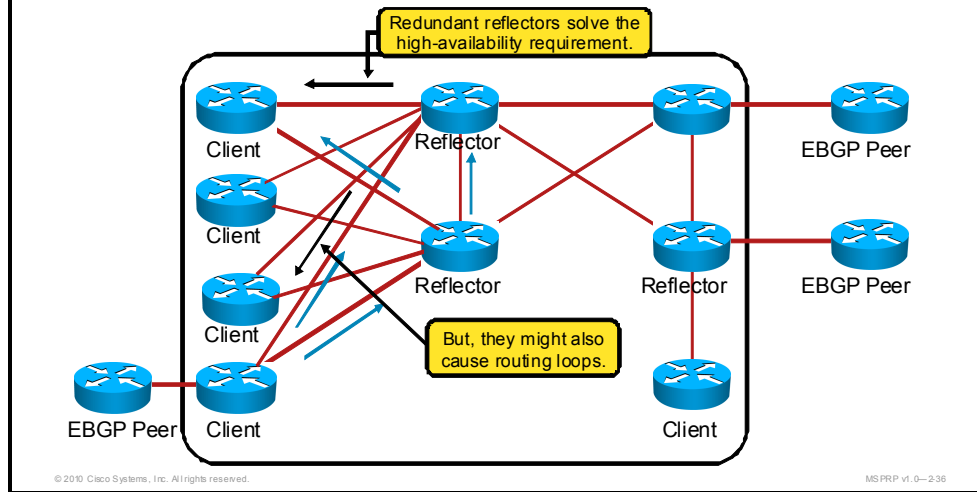
MSPRP v1.0-2.35

Clients may have any number of EBGP peers but may have IBGP sessions only with their route reflector or reflectors. If the reflector fails, its client can no longer send BGP updates to, or receive them from, the rest of the AS. The route reflector is, therefore, a single point of failure.

To avoid introducing a single point of failure into the network, route reflector functionality must be as redundant as the physical network. If a client will still be physically attached to the network after its route reflector has failed, the client should have a redundant route reflector. Thus, in all highly available networks, route reflectors must be redundant.

Redundant Route Reflectors (Cont.)

- The concept of **clusters** is introduced to prevent IBGP routing loops between route reflectors.



A client may have IBGP sessions to more than one route reflector to avoid a single point of failure. Each client will receive the same route from both of its reflectors. Both route reflectors will receive the same IBGP update from their client, and they will both reflect the update to the rest of the clients. Additionally, both route reflectors will get updated from the full mesh and reflect those updates to their clients. As a result, each client will get two copies of all routes. Under certain circumstances (particularly when you use weights on IBGP sessions to influence BGP route selection), improper route reflection can result in an IBGP routing loop that is impossible to detect. Additional BGP attributes are thus necessary to prevent these routing loops.

Route Reflector Clusters

- A group of redundant route reflectors and their clients form a cluster.
- Each cluster must have a unique cluster ID.
- Each time a route is reflected, the cluster ID is added to the cluster-list BGP attribute.
- The route that already contains the local cluster ID in the cluster list is not reflected.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.37

A router that is acting as a route reflector client does not require any specific configuration. It simply has fewer IBGP sessions than it would have if it were part of the full mesh. But improperly configuring the client to also be a reflector could easily cause a loop. An IBGP route coming in from one of the real reflectors to the client could be forwarded by the client, erroneously acting as a reflector, to the other reflector.

Route reflector clusters prevent IBGP routing loops in redundant route reflector designs.

The role of the network designer is to properly identify which route reflectors and their clients will form a cluster. The designer assigns to the cluster a cluster ID number that is unique within the AS.

Note The cluster ID number must be configured in the route reflectors. The clients should not be configured with this information.

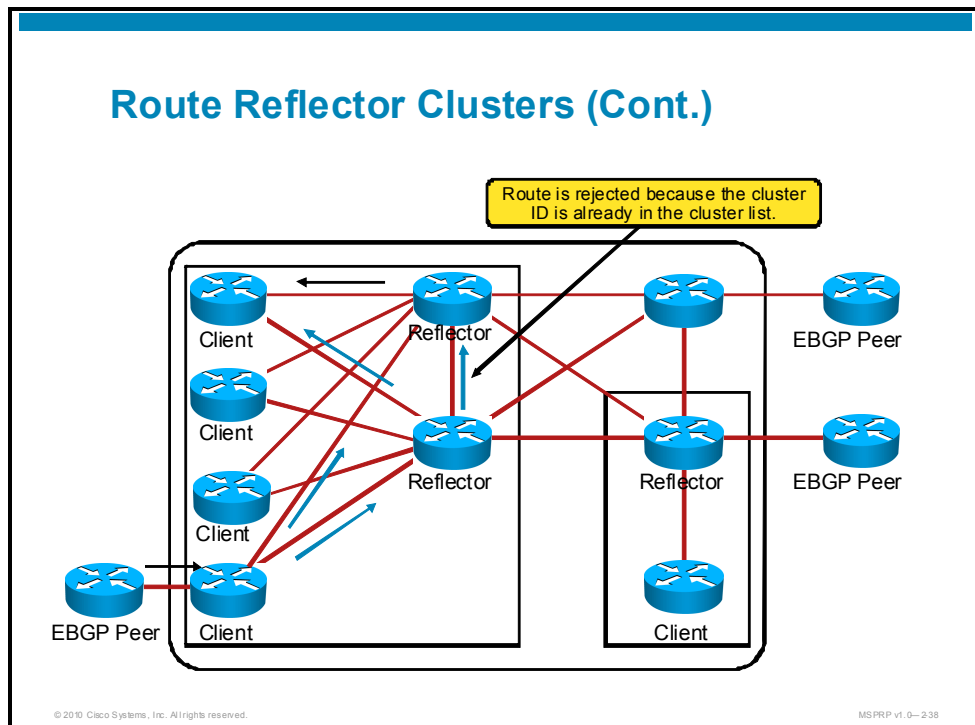
A route reflector router can reflect routes only within a single cluster. A route reflector can, however, participate in another cluster but only as a client. A client can function as a client only to a route reflector belonging to the same cluster.

When a route is reflected, the reflector creates the cluster-list attribute and attaches it to the route if it does not exist. It then sets its cluster ID number in the cluster list or adds its cluster ID number to an already existing cluster-list attribute. If the route, for any reason, is ever reflected back to the same reflector, it will recognize its cluster ID number in the cluster list and not forward it again. The first route reflector that reflects the route also sets an additional BGP attribute, called “originator-ID,” and adds it to the BGP router ID of its client.

Note The cluster-list and originator-ID attributes are nontransitive optional BGP attributes, allowing routers that do not support route reflector functionality to coexist with route reflectors and their clients in the same AS.

Based on cluster-list and originator-ID attributes, routers can implement two loop-prevention mechanisms.

- Any router that receives an IBGP update with the originator-ID attribute set to its own BGP router ID will ignore that update.
- Any route reflector that receives an IBGP update with its cluster-ID attribute already in the cluster list will ignore that update.



The client in the cluster forwards the received EBGP update to both reflectors. The route reflectors forward the update into the IBGP full mesh. This behavior means that they send the update to each other as well. But when a route reflector receives a BGP update from another route reflector, it recognizes their common cluster ID number in the cluster-list attribute. Therefore, the newly received route update is ignored.

Network Design with Route Reflectors

Route reflector rules:

- Route reflector rules divide a transit AS into smaller areas (called clusters).
- Each cluster contains route reflectors and route reflector clients.
- Routers that do not support route reflector functionality act as a one-router cluster or as a route reflector client.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.39

The physical topology of the network could serve as a guide to route reflector design.

Implementing route reflectors within the transit AS will create smaller areas (or groups) of routers. These smaller groupings of routers are called clusters. A cluster consists of route reflector routers, either redundant or nonredundant, and the client routers that are connected to them.

In designing the implementation of route reflectors within a transit AS, identify a group of peripheral routers that are physically connected to the same backbone router or routers. Consider the peripheral routers as clients and the backbone routers as route reflectors. Then, consider that this group of routers forms a cluster.

Note	Additional design examples and rules are available in the module “BGP Transit Autonomous Systems.”

Only the routers that are configured as route reflectors require a Cisco IOS Software version with route reflector functionality. A router lacking this functionality in its installed Cisco IOS Software can function as a client or be a part of the full mesh. Normally, this situation is not a concern because route reflector functionality has been incorporated in Cisco IOS Software since Release 11.1.

Network Design with Route Reflectors (Cont.)

IBGP session rules:

- All clients in a cluster must establish IBGP sessions with, and only with, all route reflectors in the cluster.
- An IBGP full mesh between all route reflectors within the AS is required.
- Routers that are not route reflectors can participate in the IBGP full mesh or be route reflector clients.

© 2010 Cisco Systems, Inc. All rights reserved.

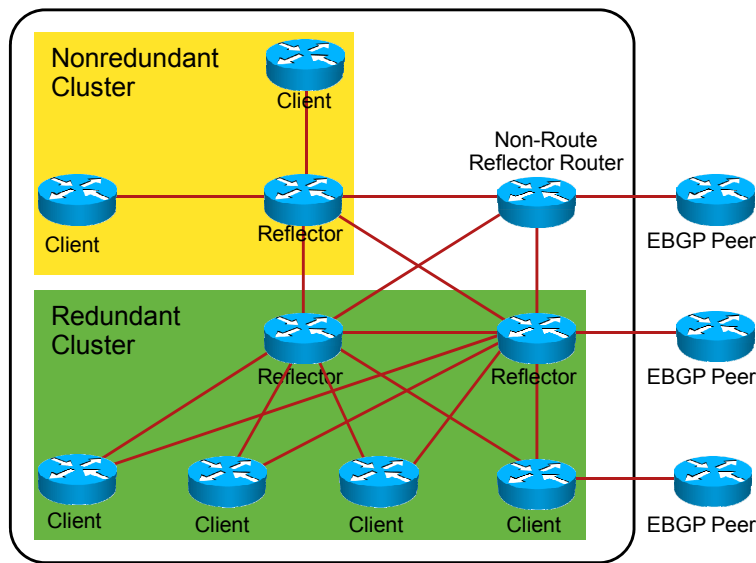
MSRP v1.0—240

The principal goal for designing networks with BGP route reflectors is to reduce the size of the full mesh of IBGP sessions by excluding some routers from the mesh. The routers that are excluded from the full mesh, the clients, have to send their IBGP information to, and receive it from, at least one router that belongs to the full mesh, the route reflector. Thus, the full mesh is still there, but it is smaller, and all route reflectors have to be part of it.

All clients in a cluster should have IBGP sessions with all their route reflectors and their route reflectors only. If a client does not have sessions with all the reflectors in the cluster, the redundancy is violated. If a client has IBGP sessions to routers other than the route reflectors, unnecessary routing traffic is generated.

Both clients and other routers—those routers that are not route reflectors—obey the classic IBGP split-horizon rules. Thus, non-route-reflector routers are either clients to a reflector or are participating directly in the full mesh.

Example: Network Design with Route Reflectors



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-41

In the area that is called the “redundant cluster,” the four client routers and the two route reflector routers make up the cluster. Each of the four client routers has an IBGP session with the two route reflectors and only with those two route reflectors.

In the nonredundant area, each of the two client routers has a single physical connection to a route reflector router. These three routers form a nonredundant cluster. The router that is designated the route reflector in the cluster is already a single point of failure in this physical design. A failure of this router will prevent the clients in the cluster from reaching the rest of the network. Therefore, there is no new single point of failure that is introduced when the router is configured as the only route reflector in this cluster. Each of the two clients has a single IBGP session to the route reflector.

The other router that is shown is not configured as a route reflector nor is it a client to any other route reflector. This other router serves as an example of when a non-route-reflector router participates in the full mesh.

Potential Network Issues

Potential problems that can occur when you deviate from the route reflector network design rules:

Issue:

- Clients do not have sessions with all reflectors in a cluster.
- Clients have sessions with reflectors in several clusters.
- Clients have IBGP sessions with other clients.

Result:

- Clients will not receive all IBGP routes.
- Clients will receive duplicate copies of the same route.
- Clients will receive duplicate copies of the same route.

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0-242

Two nontransitive optional BGP attributes, originator-ID and cluster-list, are both used to prevent fatal loops of information. The use of these two attributes makes a network fairly insensitive to poor configuration. However, for optimal performance, you must have an optimal configuration.

Several problems could occur if you deviate from route reflector network design rules.

- If route reflectors are not connected with IBGP sessions in a full mesh, some clusters will not have all the routes.
- If a client has IBGP sessions with some route reflectors in a cluster, but not with all of them, the client might miss some BGP routes.
- If a client has IBGP sessions to route reflectors that belong to different clusters, the BGP update from the client will be forwarded by the client into the full mesh with different cluster IDs in the cluster-list attribute. When the BGP update enters the mesh, it will reach the other route reflector, which will, unnecessarily, accept the route as valid and forward it into its cluster. This situation, in turn, causes unnecessary duplication of updates to the clients.
- If a client has IBGP sessions to other clients in the same cluster, those clients will receive unnecessary update duplications.

Hierarchical Route Reflectors

Problem:

- In very large networks, a single layer of route reflectors might not be enough.

Solution:

- A hierarchy of route reflectors can be established:
 - A route reflector can be a client of another route reflector.
 - The hierarchy can be as deep as is needed.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-243

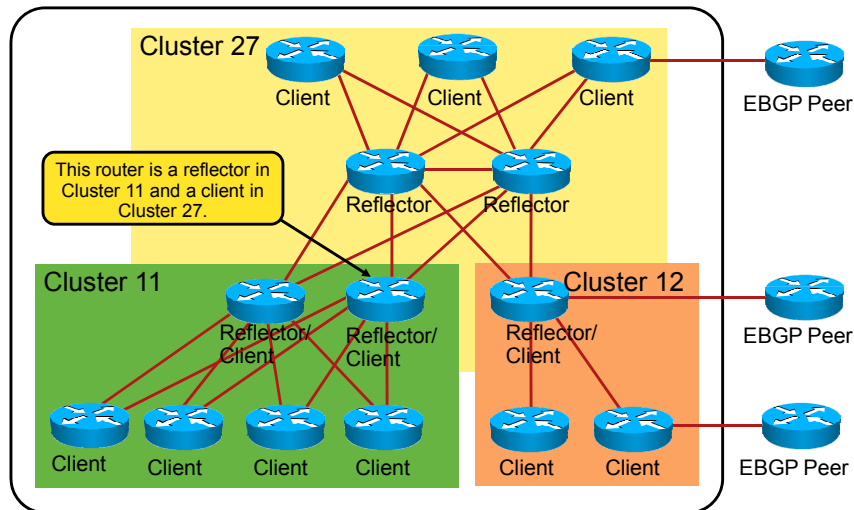
Network designers can build route reflector clusters in hierarchies. With hierarchies, a router serving as a route reflector in one cluster can act as a client in another cluster.

Clients are not configured to be route reflector clients; they simply have fewer IBGP sessions. However, a network designer must configure a route reflector. In configuring an IBGP session on a route reflector, the designer must configure the session to reach a client in order for the route reflector IBGP split-horizon rules to start working. All other IBGP sessions that are configured on the route reflector are a part of the full mesh. Also, the designer must configure the cluster ID on the route reflector.

A router that is configured to be a route reflector will still have ordinary IBGP sessions that are part of the full mesh. If these sessions are reduced in number and only a few remain—and the remaining ones reach a second level of route reflectors—a hierarchy of route reflectors is created.

When a designer builds a first level of clusters, the remaining full mesh is smaller than when all routers belonged to it. But if it is large enough, the designer can build an additional level of route reflectors.

Hierarchical Route Reflectors (Cont.)



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-44

This first step reduced the original full mesh of 14 routers to a full mesh of 8 routers.

A second level of route reflector clusters was built by creating Cluster 27. This second step further reduced the full mesh of eight routers to a full mesh consisting of only two routers. Only the two route reflectors in Cluster 27 should be connected in a full mesh.

When a client in the lowest level receives an EBGP update, it will forward it on all configured IBGP sessions to a route reflector. The route reflector recognizes BGP updates that are received from configured clients and will forward these updates to all other clients that use normal IBGP sessions. The update, sent on a normal IBGP session, will be a second-level client update to the second-level route reflector. The second-level route reflector will recognize that the update was received from a client and will forward it to all other clients and into the full mesh.

Static and Connected Routing

This topic describes the typical scenarios in which static and connected routing is used in service provider environments.

Customer Connectivity Types

- Internet customers have a wide range of connectivity and redundancy requirements:
 - Single permanent connection to the Internet
 - Multiple permanent connections to a single provider in primary and backup configuration
 - Multiple permanent connections to a single provider used for load sharing of traffic
 - Connections to multiple service providers for maximum redundancy

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-246

Service provider customers have different requirements for their Internet connectivity. These different requirements result in different solutions:

- A single permanent connection to one ISP. This solution meets the requirements for the vast majority of customers.
- Multiple permanent connections in which one of the lines is primary and the other line is used for backup only. This setup also provides redundancy on the links. Compared to a dial-up backup, a permanent backup link is preferred for various reasons, such as the severe bandwidth limitations on dial-up lines and the time that is required to establish a dial-up connection.
- Multiple permanent connections to one ISP, which is used for load sharing of traffic. This solution gives redundancy on the links but also provides additional bandwidth.
- Permanent connections to more than one ISP. This solution provides the highest level of redundancy, because it not only can cope with link-level failures but also with failures within the network of a service provider.

Customer to Provider Routing Schemes

- Static or dynamic routing can be used between an Internet customer and an ISP.
- BGP is the only acceptable dynamic routing protocol.
- Because of its lower complexity, static routing is preferred.

© 2010 Cisco Systems, Inc. All rights reserved.

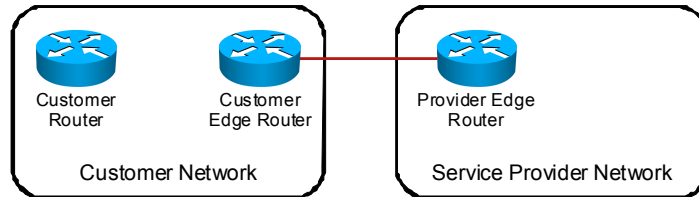
MSRP v1.0--247

Different solutions for connecting a customer network to the network of an ISP require different methods of routing information exchange:

- **Static routing** is preferred because of its lower complexity. In a normal case, the customer network must have a default route to the ISP network, and the ISP network must have a route to the IP prefixes that the customer has in its network. As always, static routing provides very low, if any, redundancy.
- **Dynamic routing** provides redundancy. The customer and the ISP networks must be configured to exchange a common routing protocol. BGP is the only choice because of the large volumes of routing information, the inherent security mechanisms of BGP, and the ability of BGP to process routing policies.

Single Permanent Connection to the Internet

- The simplest setup: a single link between the customer network and the Internet.
- No redundancy for link or equipment failure.
- Static routing is usually adequate.
- BGP should not be used in this setup.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-248

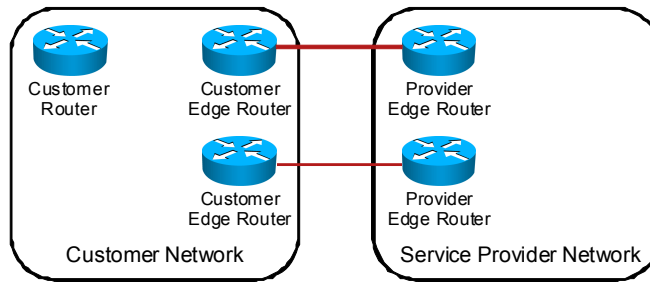
A single permanent connection to one ISP is the most common setup. This setup is also the simplest to implement.

The customer network has an edge router. This router is connected to one of the edge routers of the ISP. The connection is permanent, and there is no redundancy in this solution. Any failure on the permanent link or either of the two edge routers causes a complete outage of the service. Serious failures within the ISP network that affect all customers of this ISP also affect the customer in this example.

When the customer has a single permanent connection to the Internet, static routing is usually adequate. The physical topology does not provide any redundancy, and it is therefore unnecessary to add the complexity of dynamic routing. Keep the network simple by avoiding the use of BGP in this case.

Multiple Permanent Connections— Providing Redundancy

- Customers wanting increased redundancy install several physical links to the Internet.
- Redundant links are used in primary and backup setup or for load sharing.
- Redundancy is for link or equipment failure.
- There is no redundancy for service provider failure.
- You can still use static routing if link and remote equipment can be detected reliably.
- BGP between the customer and the service provider is usually used in this setup.



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0--249

In this setup, one customer edge router connects to one ISP edge router. A different customer edge router is used to connect to another ISP edge router. If one of these routers fails, only one of the connections breaks down; the other connection is still available.

In some cases, the two links may be implemented between the customer and the provider for load sharing and, in other cases, strictly for backup purposes. When load sharing between both links is a desired network characteristic, the distribution of the load over the links is more complicated than when both links terminate in the same router.

Again, because the customer is connected to a single ISP, serious ISP network failures that affect all customers of this ISP will also affect the customer in this scenario, regardless of the backup link.

You can also use static routing for multiple permanent connections between two different routers on the customer network to two different routers on the service provider network, if the link-level procedures can detect failures. When one of the connections is lost, the link-level procedure detects this loss and places the interface in a down state. Because the interface is in the down state, the static route that points out of the down interface becomes invalid. As a result, the router stops the redistribution of the static route into BGP.

However, customers that require the use of multiple connections and multiple routers very often do not rely on the link-level procedures. These customers require a routing protocol such as BGP to detect the failures. Because BGP uses handshaking and reliable transfer, it always detects a failed link or a failed remote router.

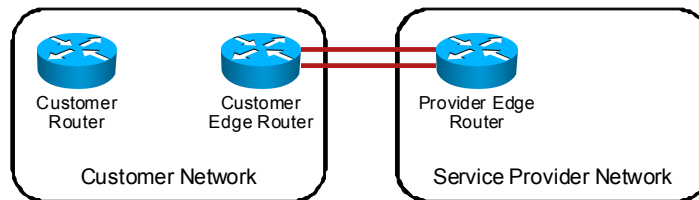
Multiple permanent connections to more than one ISP always require the use of dynamic routing with BGP. The customers that require this type of connection do not only want to protect the network connectivity from link failures or remote router failures. They also want to protect their network connectivity from serious problems in the network of an ISP.

Monitoring the link status cannot detect a problem inside one of the ISP networks. If the link and the ISP edge router are still operating, the link-level procedures do not indicate any problems. However, the ISP network may suffer from severe problems. An ISP network can be partitioned or disconnected from the rest of the Internet without having any problems with the edge router and the access line to the customer network.

The only way to detect this situation is to use BGP with both ISPs and receive complete Internet routing from both of them. When one of the ISPs has problems, the edge router, being the BGP neighbor of the customer, withdraws the routes that it can no longer reach. This action means that the customer routers know which Internet routes that each ISP can reach at the moment.

Multiple Permanent Connections— Providing Load Sharing

- Customers wanting to increase their access speed can install several physical links between a pair of routers.
- There is redundancy for link failure.
- There is no redundancy for equipment failure.
- Load sharing in this setup is optimal.
- Static routing is preferred if physical link failure can be detected.
- Traffic will enter a “black hole” if the physical link failure is not detected.



© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0--2.50

In the example, a single router in the customer network is connected to a single router in the ISP network. The redundancy is limited to the link level because router failures are not covered. Using two parallel links between two routers allows for an optimal distribution of load over the links.

Depending on the switching path that is used in the customer and the ISP routers, load sharing can be performed based on the destination address only (fast switching), on source-destination address pairs (default behavior for Cisco Express Forwarding), or on a packet-by-packet basis (process switching or Cisco Express Forwarding).

As in the previous examples, serious ISP network failures that affect all customers of this ISP will also affect this customer, regardless of the link backup.

Multiple permanent connections between a single router on the customer network and a single router on the service provider network should be configured with static routing, provided that link-level procedures can detect link failure.

With this type of connection, two static routes are configured on each network, pointing to both links between the customer and the ISP. If either of the links fails, the link-level procedures should detect this failure and place the interface in a down state. In this case, the static route is invalid and is not used for forwarding packets. The router will then forward all packets over the remaining link.

If the link-level procedures cannot detect a link failure, the static route pointing out over the failed link is still valid. The router continues to use this static route to send out some of the traffic on the failed interface. This situation effectively creates a “black hole” for some of the traffic.

Characteristics of Static Routing

- The customer network must announce a default route:
 - Redistribute default route into customer IGP if the customer is running EIGRP.
 - Use **default-information originate** if the customer is running OSPF or RIP.
- Customer routes should be carried in BGP, not core IGP:
 - Redistribute static routes into BGP, not IGP.
- Routes to subnets of the provider address block should not be propagated to other autonomous systems:
 - Mark redistributed routes with no-export community.
 - Use static route tags for consistent tagging.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.51

When static routing is implemented between the customer network and the ISP network, the edge router of the customer must announce itself as a default gateway or a gateway of last resort. This procedure must be completed using the IGP within the customer network, because different routers within the customer network must be able to select the best route to the exit point of the network.

Different IGPs use different methods of announcing a router as a gateway of last resort. Enhanced Interior Gateway Routing Protocol (EIGRP) uses the concept of default network, while OSPF and Routing Information Protocol (RIP) send reachability information about network 0.0.0.0/0. In either case, the network operators of the customer network are responsible for configuring their network to use the customer edge router as a gateway of last resort.

When static routing is used between the customer and the provider, the edge router of the provider must propagate a static route that points to the customer network, to all other routers within the ISP network, and also to the rest of the Internet. The network operators in the ISP network propagate the route using a configuration command to start redistributing the routes into BGP.

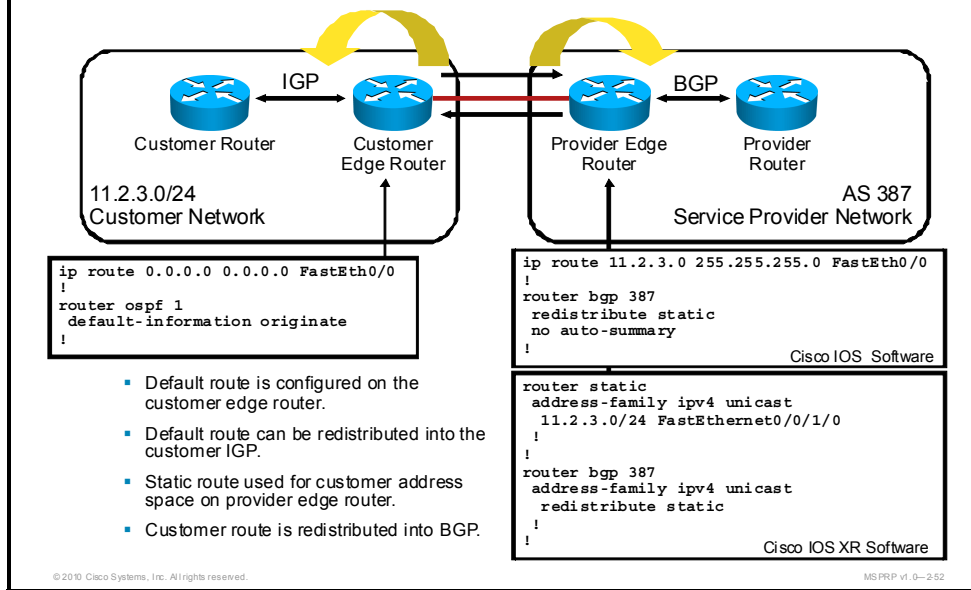
Customer routes should not be redistributed into the IGP of the ISP network. Care should be taken that the IGP of the ISP network does not carry too many routes. Redistributing customer routes into the IGP could potentially cause poor performance and might eventually cause a complete shutdown of IGP routing at the service provider.

If a customer uses provider-assigned addresses (PA addresses) and the ISP announces a large block of addresses for which the network of this customer is only a small portion of the block, then the routes of this customer should not be propagated by the service provider to the rest of the Internet. Instead, the rest of the Internet should receive only an announcement containing the larger block of addresses.

An easy way of achieving this setup is to use communities within the ISP network. Any customer route that should not be announced to the rest of the Internet is marked using the no-export community. To ensure that the BGP communities are propagated, at least over all IBGP sessions, the network operators of the ISP network must configure a send-community option for all IBGP neighbors. The edge routers of the ISP network then see the no-export community and filter those routes out before sending the update to EBGp neighbors.

Communities are set using route maps. A route map can select routes that are based on various attributes. One of these attributes is the route tag. Through configuration, the router can assign a route tag to specific static routes. This option means that the network operators of the ISP network can invent a scheme of tagging in which all static routes that should not be propagated to other autonomous systems are assigned a specific tag. Then a route map can select all routes with that tag and assign them the no-export community.

Characteristics of Static Routing (Cont.)

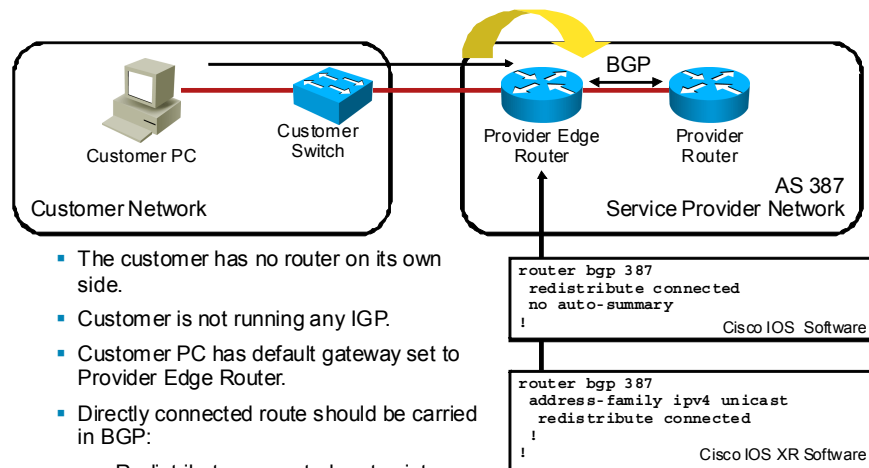


In this case, a routing protocol does not add any redundancy and would only add complexity.

The customer edge router has a static default route pointing to the interface Fast Ethernet 0/0. If the serial interface goes down, the route becomes invalid. The **default-information originate** command is configured in the OSPF process on the customer router. Therefore, the router announces a default route into OSPF as long as it has a valid default route itself.

The service provider edge router also has a static route, declaring the customer IP network number as reachable over the Fast Ethernet 0/0 interface. It also becomes invalid if the interface goes into the down state. The ISP edge router must forward this information to all other ISP routers and to the rest of the Internet. This action is accomplished by redistributing the static route into BGP. As long as the static route is valid, BGP announces it. To the rest of the Internet, the customer network appears as reachable within the AS of the ISP. As far as the rest of the Internet is concerned, the customer is a part of the service provider AS.

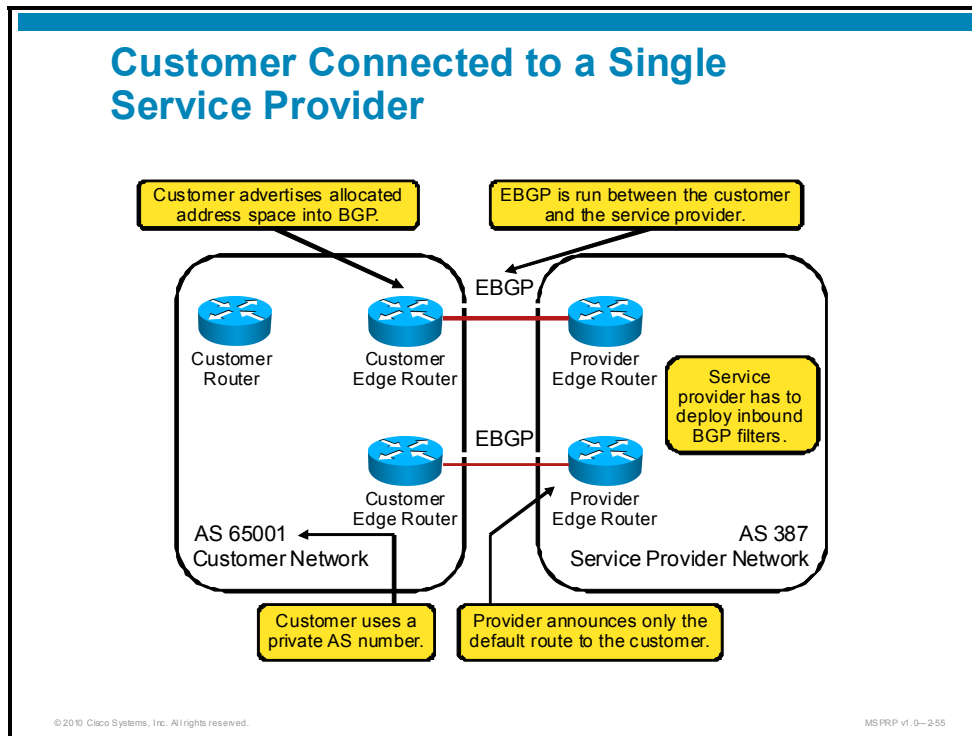
Characteristics of Connected Routing



In the case of connected routing, the customer has no router on its own side and the customer is not running any IGP protocol. Usually, the customer has a switch on its side to which end devices are connected. End devices have a default gateway that is set to the IP of the Provider Edge (PE) router interfaces pointing towards the customer. Directly connected routes should be redistributed into the BGP as shown in the example.

External BGP

This topic describes the typical scenarios in which external BGP is used in service provider environments.



In the example, the customer network is connected to a service provider network using multiple permanent links. BGP is used to exchange routing information between the customer and the provider.

Selecting BGP as the routing protocol between the customer and the provider network ensures that a link failure or the failure of a remote router is detected. In this scenario, the customer does not require the use of a public AS number or complete Internet routing. Instead, a private AS number is assigned to the customer network, and the ISP sends a default route to the customer through BGP.

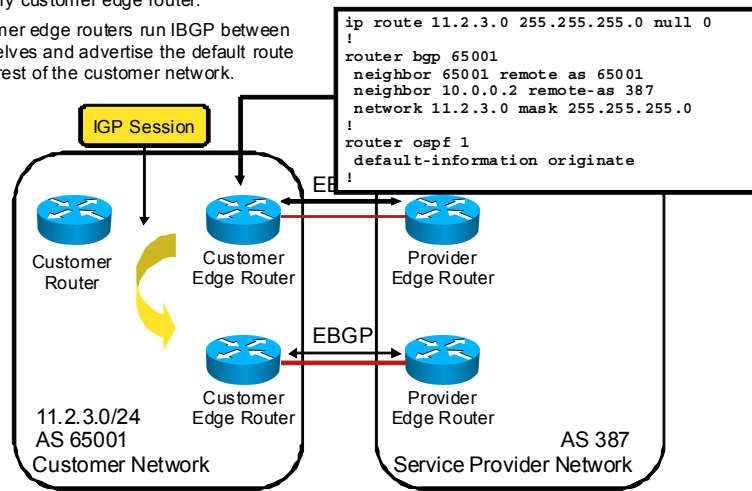
The big difference in this case, as compared to a network scenario in which static routes and redistribution are used, is that routers within the private AS of the customer now advertise the customer routes via BGP. Thus, the customer is responsible for announcing its own address space. The ISP receives routes from the customer and conditionally propagates them (similar to static routing). If the customer uses PA address space and the ISP can summarize the address space, it will not propagate the explicit routes from the customer to the Internet. The private AS number in the AS path attribute must be removed before the ISP can propagate any of the customer routes.

Because the customer is now creating BGP routes that the ISP receives, any error that is made by the customer can influence routing operation within the ISP network and, if propagated, within the Internet as a whole. Announcing a route to a network to which the customer has not been assigned may cause routing problems. There is always a risk that such routing problems can occur in a service provider network. However, the risk is much greater when the customer, whose network administrators usually have less experience with BGP, enters the configuration.

To reduce the risk of erroneous route advertising, the ISP should always filter any BGP information that it has received from the customer network. The ISP should reject routes to networks that are not expected to be in the customer AS. Routes that contain an AS path with unexpected AS numbers should also be rejected.

Customer Connected to a Single Service Provider (Cont.)

- The customer address space is advertised on every customer edge router.
- Customer edge routers run IBGP between themselves and advertise the default route to the rest of the customer network.



Both customer edge routers are configured to run BGP and should advertise all of the customer networks with the **network** command. If only one router advertises the network, a single point of failure has been introduced. The two customer edge routers must also run IBGP between them to make common decisions regarding BGP routing information.

Each customer edge router has an EBGP session with the ISP router on the other side of the link. Over that EBGP session, the ISP announces only a default route to the customer AS. When EBGP receives the default route, it installs it in the routing table and redistributes it into the IGP (in this case, OSPF) of the customer.

Conditional Advertising for Customer Connected to a Single Service Provider

- Customer edge routers should announce the whole customer address space into BGP.
- The static route covering the whole customer address should point to the core of the customer network, not to null 0.
- The customer edge router revokes the BGP announcement of customer address space if the edge router loses connectivity with the customer core.

© 2010 Cisco Systems, Inc. All rights reserved.

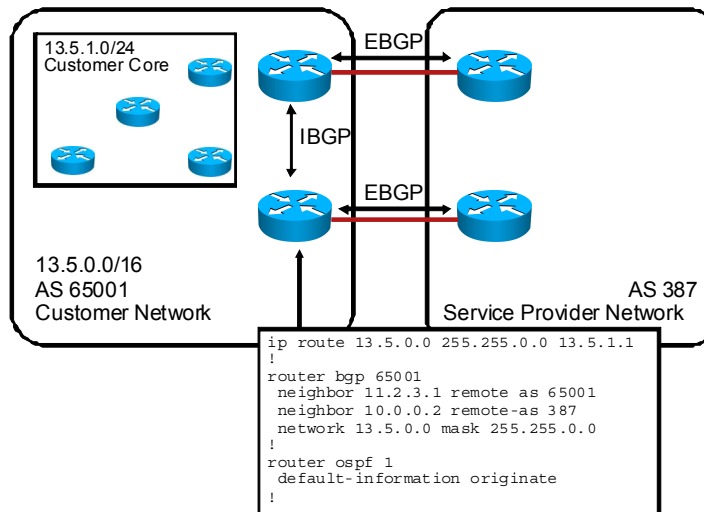
MSRP v1.0--257

As a general principle, the customer should announce addresses as large as possible (the larger the address space that can be aggregated, the better). The BGP advertisement is configured on the customer edge routers using the **network** command. Route advertisement is conditioned by the appearance of a corresponding network or subnet in the routing table of the edge router. If the network or subnet is manually entered into the routing table by configuring a static route to null 0, the condition is always true. The static route is always there, and the BGP advertisement is always performed.

If the customer edge router loses connectivity to the rest of the customer network but is still connected to the ISP network, BGP advertisement must cease. In this case, BGP advertisement can be stopped if BGP advertisements are bound to the reachability status of a specific subnet in the core of the customer network, according to the customer IGP.

The problem with using a static route to null 0 is that it conditions the network statement in the BGP configuration so that BGP always advertises the route. If the customer edge router loses connectivity with the rest of the customer network, the router continues to advertise the entire customer address space. The ISP network receives a valid route from the customer edge router. Traffic is sent to this router, but because the router has lost connectivity with the rest of the network, the traffic is dropped (routed to the null 0 interface using the static route).

Conditional Advertising for Customer Connected to a Single Service Provider (Cont.)



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.58

The address space is further subnetted at the customer site. One of the subnets (subnet 13.5.1.0/24) is identified as being a central part of the customer core network.

The customer edge routers participate in the IGP routing of the customer. This participation means that these routers have information about which of the subnets within the address space 13.5.0.0/16 is currently reachable. If these subnets are available, there is an explicit route to each of them. If any of the subnets go down, or if the path toward them goes down, the route to that subnet is removed from the routing table.

The BGP advertisement in each of the customer edge routers is configured to advertise the complete address space that the customer uses. When the customer edge routers advertise this route, the ISP network, and thus the rest of the Internet, sees the complete address space of the customer as one single route, 13.5.0.0/16.

Advertisement of the customer address space by BGP is conditioned by the appearance of the static route, IP route 13.5.0.0 255.255.0.0 13.5.1.1. If the static route is valid, then the BGP route 13.5.0.0/16 is advertised. The static route is a recursive route, which means that the router takes another look in the routing table for the address 13.5.1.1 before determining what to do with the static route. The idea is that 13.5.1.1 is reachable via the IGP. The IGP announces the subnet 13.5.1.0/24. If this subnet is reachable by the edge router, then the static route to 13.5.0.0/16 is valid. If there is no route to 13.5.1.1, then the static route is invalid.

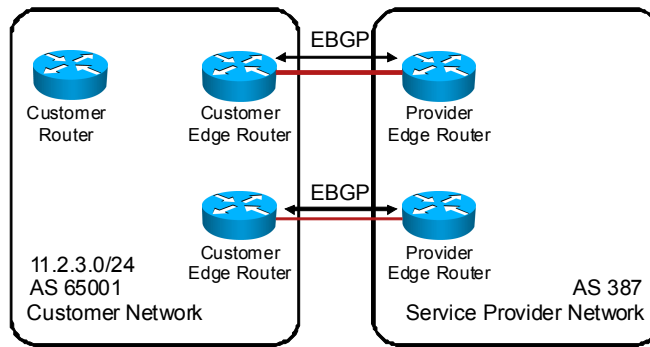
Note The condition, whether to advertise the entire customer address space 13.5.0.0/16, is controlled by the IGP reachability of a single subnet, 13.5.1.0/24.

The IGP configuration also includes origination of the default route by both edge routers.

Configuring BGP on Service Provider Routers

The service provider must:

- Advertise a default route to the customer through BGP
- Filter incoming BGP updates with a prefix list to verify that the customer announces only the assigned address space
- Filter incoming BGP updates with an AS path filter list to verify that the customer uses only its own AS number



In the ISP network, the two edge routers must have BGP sessions that are configured for the customer. There is no point in feeding the complete Internet routing table to the customer. The table contains the same set of routes for both links and the customer always uses the ISP for all traffic toward the Internet. Injection of a default route in the customer network would accomplish the same task.

The customer is responsible for its own advertisements. Because customers are much less likely to be experienced in BGP configuration than the ISP, they are more likely to make errors. Therefore, the ISP must protect itself and the rest of the Internet from those errors.

The service provider should use a prefix list that allows only customer-assigned routes and denies any other route. Using a prefix list ensures that private address space or any other illegal networks that are erroneously announced by the customer never reach the ISP BGP table. Filtering based on the AS path also provides some protection from customer configuration errors. Only routes that originated within the customer AS are allowed. A filter-list performs this check.

If the customer address space is a small part of a larger block that is announced by the ISP, the explicit BGP routes that are received from the customer need not be advertised to the rest of the Internet. The ISP can announce the large block, attracting any traffic toward any subnet within the block. After the traffic enters the ISP network, the more explicit routes to the customer network are available and used. In this case, the provider edge router can tag the BGP routes that are received from the customer with the no-export well-known community, restricting them from being sent by the ISP to any other AS.

Configuring BGP on Service Provider Routers (Cont.)

Router(config-router) # Cisco IOS Software

```
neighbor ip-address default-originate
```

RP/0/0/CPU0:Router(config-bgp-nbr-af) # Cisco IOS XR Software

```
default-originate
```

- By default, the default route (0.0.0.0/0) is not advertised in outgoing BGP updates.
- The **default-originate** command advertises the default route to a BGP neighbor, even if the default route is not present in the router.
- Note: The default route is filtered by the outbound BGP filters (prefix list, filter list, or route map).

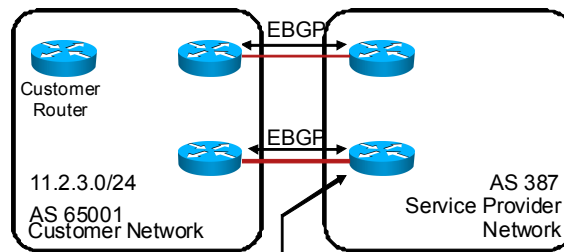
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.60

The default route, 0.0.0.0/0, is not advertised in outgoing BGP updates unless it is explicitly configured. The **neighbor default-originate** router configuration command is used to initiate the advertisement of the default route to a neighbor.

The BGP does not check before the default route is advertised. The default route does not need to be present in the BGP table before it is advertised using this command. The default route is also sent without being filtered by any outgoing prefix lists, filter lists, or route maps.

Configuring BGP on Service Provider— Cisco IOS Routers



Cisco IOS Software

```

router bgp 387
 neighbor 10.0.0.1 remote-as 65001
 neighbor 10.0.0.1 default-originate
 neighbor 10.0.0.1 prefix-list DfltOnly out
 neighbor 10.0.0.1 prefix-list CustA in
 neighbor 10.0.0.1 filter-list 15 in
 neighbor 10.0.0.1 route-map AllCustsIn in
!
ip as-path access-list 15 permit ^65001(_65001)*$
!
ip prefix-list DefaultOnly permit 0.0.0.0/0
ip prefix-list CustA permit 11.2.3.0/24 le 32
ip prefix-list Provider permit 11.2.0.0/16 le 32
!
route-map AllCustsIn permit 10
 match ip prefix-list Provider
 set community no-export additive
!
route-map AllCustsIn permit 9999
!

```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-261

The customer is assigned the private AS number 65001. The BGP session is opened with the customer IP address, 10.0.0.1.

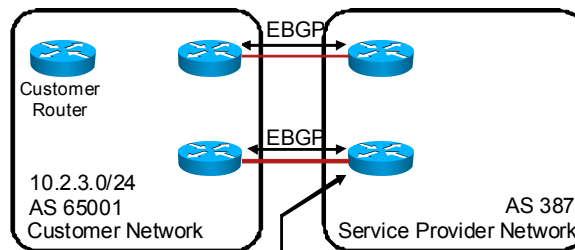
The ISP sends the default route only to the customer. This route is configured using first the **default-originate** command and then the prefix list, DefaultOnly.

Received routes from the customer must first pass the prefix list, CustomerA. There is one dedicated prefix list for each individual customer; the prefix list permits only the routes that the customer is allowed to announce. If the prefix list allows the routes, they must also pass the filter list named “15 in.” In this case, the filter allows the private AS of the customer in any number of repetitions, as long as it is the only AS number in the path. This filter list allows for AS path prepending configurations on the customer side. If both the prefix list and the filter list allow the received route, then the route map, AllCustomersIn, is applied.

The route map is a general route map that is used for all customers. It checks every route that is received, via the prefix-list, Provider. If the route is within the large block of PA address space that the ISP announces to the rest of the Internet, the customer route is marked with the no-export community. This mark means that the route is used within the ISP autonomous system only and is not sent to the rest of the Internet.

Routes that are received from the customer and are allowed by the prefix list and filter list, but do not fall within the PA address space, are allowed by the route map and are not changed in any way. The ISP propagates these routes to the rest of the Internet.

Configuring BGP on Service Provider— Cisco IOS XR Routers



Cisco IOS XR Software

```

router bgp 387
neighbor 10.0.0.1
remote-as 65001
address-family ipv4 unicast
route-policy CustAIn in
route-policy DfltOnly out
default-originate
!
!
!
route-policy DfltOnly
if (destination in (0.0.0.0/0)) then pass else drop endif
end-policy
!
route-policy CustAIn
if (not as-path in (ios-regex '^65001(_65001)*$')) then
drop
endif
if (destination in 10.2.3.0/24 le 32) then
set community (no-export) additive
else
drop
endif
end-policy

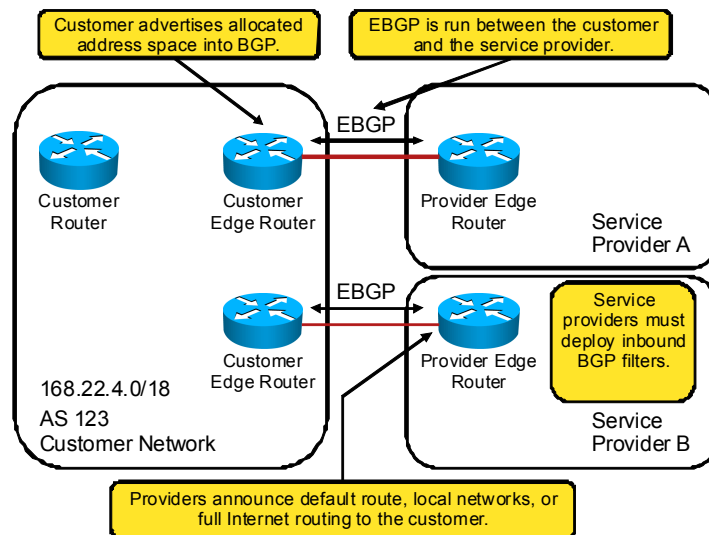
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.62

The sample Cisco IOS XR configuration achieves the same goal as the previous Cisco IOS configuration. The main difference is that the inbound and outbound policies and filters are combined in one pair of routing policies. The CustAIn routing policy first filters out any update that does not originate in the customer's AS—that is not from the approved address range—and then assigns the no-export community to permitted routes.

Customer Connected to a Multiple Service Providers



The highest level of resilience to network failures is achieved in network designs that connect the customer network to two different service provider networks. Customers use this option when the requirement for resilient Internet connectivity is very high. This requirement also involves duplication of equipment to make the customer network fully redundant.

BGP must be used between the customer and both service providers, because static routing will not work in this type of network. It is not enough to detect link failures or a failure in the remote router by link-level procedures. Failures that occur beyond the directly connected router must also be detected, and the only means of detecting these failures is by using a routing protocol. The only routing protocol that is suited for the Internet environment is BGP. Correctly configured, BGP takes care of rerouting in the following situations:

- Link failure between the customer network and the network of one of the ISPs.
- Edge router failure on either the customer or the ISP side.
- Link failure or router failure within the customer network that causes the customer edge router to lose connectivity with the customer network core. This situation requires correct configuration of route advertisement as described in an earlier lesson.
- Link failure or router failure within the ISP network that causes the ISP edge router to lose connectivity with the rest of the Internet

Multihomed customers have multiple permanent links to different ISPs. The links should terminate in different edge routers in the customer network. Otherwise, one of the major advantages, resilience to router failure, is lost.

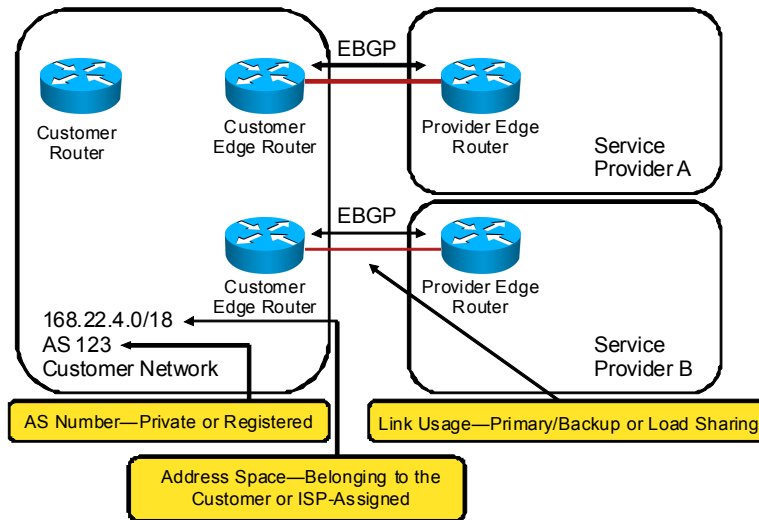
Multihomed customers should use BGP with both ISPs. The customer should advertise its address space to both providers. Route advertisement should be configured in both customer edge routers. The advertisement should be conditioned at the edge routers by the appropriate route policies leading toward the core of the customer network. This setup is analogous to the setup configured when you are connecting a multihomed customer network to a single provider.

The customer should take care not to move any routing information between the two ISPs. It must use outgoing filters to prevent any route that is received from one of the ISPs from being propagated to the other. Otherwise, the customer network appears as a transit network between the two ISPs.

Both ISPs must apply filters on the incoming BGP information from the customer to protect themselves and the rest of the Internet from errors in the BGP configuration of the customer. Each of the service providers must accept routes from the customer that indicate networks within the customer address space only. AS path filter lists should be implemented on the provider edge routers to allow incoming routes only if they have the correct AS path attribute value. If the incoming filters on the ISP edge router accept customer routes, then the service provider should propagate those routes to the rest of the Internet.

Both ISPs must provide the customer with at least some BGP routes. Depending on customer requirements, the volume of BGP routes that the ISP provides could range anywhere from the default route only to the complete Internet routing table.

Customer Connected to Multiple Service Providers (Cont.)



Before configuring the multihomed network, consider the following questions:

- Should any of the links be used as primary and the others as backup?
- Should both links share the load?
- What address space is the customer using? Is the customer address space a PA address or a provider-independent address (PI address)?
- What AS number is the customer using? (Is the customer using a public or a private AS number?)

Multihomed Customer Address Space Selection

Provider-independent address space:

- If the customer owns the address space, there should be no limitations regarding announcing it to both service providers.

Provider-assigned address space:

- If the customer uses ISP-assigned small address blocks, then there is no purpose in using BGP to provide redundant connectivity. NAT is easier to implement and solves the problem of reverse path.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.65

If the customer has its own address space, it should announce it to both service providers. Both providers are responsible for propagating the customer routes to the rest of the Internet without doing any summarization.

However, if the customer uses a small block of addresses that is assigned by one of the ISPs, an alternative design, not involving BGP, is to use two different PA address spaces and do Network Address Translation (NAT). With NAT, the router translates traffic going out over one of its connections to one of the PA addresses. If traffic goes out the other way, the addresses are translated to an address from the address space of the other provider.

Multihomed Customer AS Number Selection

Registered, public AS number (recommended):

- Preferred option, but difficult to obtain
- Does not require ISPs to assign a private AS number
- Consistent routing information in the Internet

Private AS number (discouraged):

- Easier to obtain (even easier with AS translation):
 - One private AS number: The customer has to be able to use the same private AS number with multiple providers.
 - Multiple private AS numbers: The customer obtains a private AS number assigned by each provider and uses one of them internally; the others have to be translated.
- Causes inconsistent routing information

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0--266

Using BGP requires an AS number. The preferred option is to use a registered, public AS number. However, registered AS numbers are assigned only to those customers that really need them because public AS numbers are a scarce resource. A customer with BGP sessions to multiple ISPs must use a registered, public AS number. A customer that is connected to only one ISP does not require a public AS number. In that case, a private AS number in the range 64512 to 65535 is sufficient.

Whenever the customer has a public AS number that is assigned, there are no conflicts in the BGP setup because the number is guaranteed to be unique within the Internet. Route announcements are made by both the customer and the service provider without tampering with the AS path. As a result, the service provider propagates consistent AS path information to the rest of the Internet.

When the customer does not have a public AS number, it must use a private AS number. Because private AS numbers are not propagated to the Internet, several network administrators can, independently of each other, make this assignment. In this case, AS numbers are reused; this conserves AS number space. A service provider normally assigns private AS numbers to its customers. This arrangement ensures that unique, private AS numbers are used among the customers of a single ISP.

When a customer is going to be multihomed and the private AS number that has already been assigned by one of the ISPs comes in conflict with an AS number that has been assigned by the other ISP, the customer needs to consider renumbering the customer AS. If the two service providers can reach a common agreement on which private AS number that the multihomed customer should use, renumbering is a solution. If no common agreement can be made or if renumbering, for some reason, is not an option, AS translation must be configured on the customer network.

No router should ever propagate private AS numbers to the rest of the Internet. An ISP can keep track of which private AS numbers it has assigned to its customers and avoid reuse or conflicts within that scope. However, as soon as the scope is widened to include other ISPs, conflicts will occur. Each ISP, therefore, removes private AS numbers from the AS path before sending routes outside its own AS.

When the routes with the private AS numbers removed are propagated to the rest of the Internet, the AS path looks as though the routes were originated within the public AS of the ISP. All information about the private AS lying behind the public AS is lost. In the case of a multihomed customer, the customer routes are, in the first step, propagated into each of its ISPs' autonomous systems. In the next step, the routes have the private AS number removed as the routes are propagated to the rest of the Internet. Now the customer routes appear to be originating in the autonomous systems of both ISPs. To an outside observer, there is now an AS path inconsistency because it looks as though the same route belongs to different autonomous systems.

Upstream ISP

- Tier 2 and Tier 3 ISPs depend on upstream ISPs to deliver complete Internet routing table.
- Only local routes and leaf customer routes can be sent to upstream ISPs.
- Strict filtering should be used to prevent route leakage:
 - Prevent complete Internet routing from one upstream ISP to be sent to another upstream ISP
 - Prevent routes learned from peers at exchange points to be advertised upstream

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0--267

ISPs use upstream ISPs for Internet access. An upstream ISP usually has a larger network than the contracting ISP and is able to provide the contracting ISP with access to parts of the Internet the contracting ISP by itself has no access to.

Tier 2 and Tier 3 ISPs depend on upstream ISPs to deliver a complete Internet routing table. By definition, a Tier 2 ISP is a network that peers with some networks, but still purchases IP transit or pays settlements to reach at least some portion of the Internet. Tier 3 is, by definition, a network that solely purchases transit from other networks to reach the Internet.

Only local routes and customer routes should be sent to upstream ISPs. Use strict filtering to prevent route leakage. When communicating with upstream ISPs, prevent the sending of complete Internet routing to the upstream ISP and the sending of routes learned from peer routes that are learned points.

Exchange Points

- Multiple service providers can peer in exchange points.
- Exchange points are designed to cover specific geographic areas:
 - Countries
 - Regions
- Exchange points can be used to interconnect Tier 1 and Tier 2 service providers.
- Agreements between pairs of providers determine what routes can be sent and received.
- Strict inbound and outbound BGP filters are required to enforce the agreement.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.68

Exchange points give multiple service providers an opportunity to peer in so-called exchange points. The primary purpose of an exchange point is to allow networks to interconnect directly via the exchange point, rather than through one or more third-party networks. The advantages of the direct interconnection are numerous, but the primary reasons are cost, latency, and bandwidth. Exchange points are designed to cover specific geographic areas like countries and regions.

Exchange points can be used to interconnect Tier 1 and Tier 2 service providers. Agreements between pairs of providers determine what routes can be sent and received. You can achieve that goal by setting strict inbound and outbound BGP filters to enforce the agreement.

Summary

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

Summary

- Common service provider network:
 - Runs BGP or static routing with the customer
 - Exchanges routes with other service providers via EBGp
 - Runs IBGP between its own BGP speakers
 - Runs one instance of IGP (OSPF or IS-IS) for internal routes
- An IGP is needed to resolve BGP next hops and to perform fast convergence after a failure in the core network.
- A full mesh of IBGP sessions is required for proper IBGP update propagation.
- A full mesh of IBGP peers is no longer required for scalability when using route reflectors.
- Static routing is adequate, if there is a single link between the customer network and the Internet.
- EBGp is used for routing among service providers or between the customer and multiple service providers.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2.69

Lesson Self-Check

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

- Q1) When connecting the service provider to a single-homed customer, which kind of routing is recommended? (Choose two) (Source: Overview of Routing Protocols)
- A) BGP on service provider edge router to single-homed customer
 - B) static route on service provider edge router to single-homed customer
 - C) static default route on single-homed customer to service provider edge router
 - D) BGP on single-homed customer to service provider edge router
- Q2) In which two ways can routes be exchanged inside autonomous systems and between different autonomous systems? (Choose two) (Source: Overview of Routing Protocols)
- A) inside AS via IBGP
 - B) inside AS via EBGp
 - C) between AS via IBGP
 - D) between AS via EBGp
- Q3) Why is IGP (OSPF or IS-IS) still needed in the core service provider network? (Choose two) (Source: Core IGP)
- A) to resolve BGP next hops
 - B) to ensure reachability of the customer network
 - C) to assure routing protocol redundancy if BGP does not work properly
 - D) to perform fast convergence after failure in the core network
- Q4) In which manner should IBGP sessions be established? (Source: Internal BGP)
- A) partial mesh
 - B) star topology
 - C) full mesh
 - D) hub-and-spoke topology
- Q5) Which routing protocols should core routers run? (Source: Core IGP)
- A) BGP and IGP
 - B) only BGP
 - C) only IGP
- Q6) Which two are benefits of route reflectors? (Choose two.) (Source: Internal BGP)
- A) route reflectors can propagate IBGP routes to other IBGP peers
 - B) full mesh of IBGP peers is not required
 - C) full mesh of IBGP peers is required
 - D) route cannot propagate IBGP routes to other IBGP peers
- Q7) Which routing approach should the service provider use when connecting a multihomed customer to its network? (Source: External BGP)
- A) IBGP between the service provider and a multihomed customer
 - B) static route to a multihomed customer
 - C) OSPF or IS-IS
 - D) EBGp between the service provider and a multihomed customer

Lesson Self-Check Answer Key

- Q1) B, C
- Q2) A, D
- Q3) A, D
- Q4) C
- Q5) A
- Q6) A, B
- Q7) D

Adding Routers to the Core Network

Overview

In this lesson, you will identify the tasks that are required to join a new router to the core routing domain. The lesson focuses on the tasks that are required to join a new router to an Open Shortest Path First (OSPF) domain, an Intermediate System-to-Intermediate System (IS-IS) domain, and to a Border Gateway Protocol (BGP) autonomous system. The lesson will present the required steps to configure a newly added router to the core network.

Objectives

Upon completing this lesson, you will be able to identify the tasks required to join a new router into the core routing domain. This ability includes being able to meet these objectives:

- Identify the tasks that are required to join a new router into the core routing domain
- Identify the tasks that are required to join a new router into an OSPF domain
- Identify the tasks that are required to join a new router into an IS-IS domain
- Identify the tasks that are required to join a new router into a BGP AS

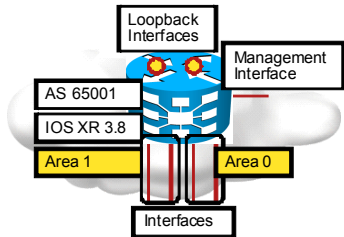
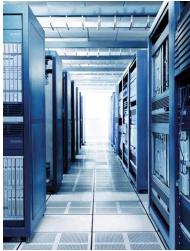
Change Management for Adding New Core Routers

This topic identifies the tasks that are required to join a new router into the core routing domain.

Deployment Methodology

Requirements

- **Determine physical requirements:**
 - Device type and modules
 - Rack space
 - Power
 - Cooling
 - Cabling
- **Determine logical requirements:**
 - Software
 - Network design or reference design
 - Router roles (such as BGP, OSPF, or IS-IS)
 - Implementation templates
 - Parameter pools (such as addresses, areas, levels, interfaces)



© 2010 Cisco Systems, Inc. All rights reserved.MSRP v1.0—2.3

The process of adding a new router to the network includes many tasks that depend on multiple factors:

- Routing design
- Management and provisioning options (manual configuration or provisioning using an element management system)
- Change management procedures (for example, company policies, security policy, ticketing system)

In general, the process always starts by determining the physical and logical requirements for the newly deployed router.

Physical requirements can be derived from the router's documentation, which details the device's physical characteristics:

- Required rack space, typically measured in rack units (RU)
- Power consumption (kW), voltage (100V–240V), frequency (50Hz or 60Hz), power type (AC or DC), number of power supplies and power cords
- Cooling requirements (often measured in BTU)
- Cabling requirements for data (types of cables), in which the number and length are usually determined by the network design

Logical requirements determine the operation of the device in the IP network:

- The Cisco IOS Software or Cisco IOS XR Software version determines the device's capabilities, which are often also determined by the design (at least the minimum version) but should follow the current Cisco recommended version.
- The router role determines the features that must be deployed for the router to perform the desired functions as defined by the network design.
- Parameter pools are used to assign available parameters to the new device, such as IP addresses (for example, physical interfaces, loopback interfaces), OSPF areas or IS-IS areas, and network service access point (NSAP) addresses. New tables can be created in the database to keep track of used interfaces on the new device. Provisioning software would typically automate many of these tasks.

Deployment Methodology (Cont.)

Preparation and Planning

- Create configuration:
 - From templates
 - From parameter databases (data pools)
- Update parameter databases
- Determine deployment impact:
 - Disruptive—requires maintenance window
 - Nondisruptive
- Create implementation plan:
 - Implementation steps
 - Verification steps
 - Rollback procedures (if applicable)

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-4

Once the roles are determined and the appropriate parameters are selected, configurations for the new device can be created. Any assigned parameters must be documented in the appropriate databases as required by the company's change management procedures.

The configuration, in combination with existing network design and network implementation plans, can be used to determine the impact of the deployment on the existing network. In general, the addition of a router is nondisruptive unless some major physical reworking is required (for example, installing new power units, replacing patch panels). If the installation will result in downtime, a regular or special maintenance window should be used to minimize the impact of the installation on end users. Verification steps should be designed to confirm the operation of the router. Rollback procedures are required, if the implementation is disruptive, to be able to quickly revert to the previous solution if the implementation fails.

Deployment Methodology (Cont.)

Deployment Execution

- **Staging (optional):**
 - Assemble and configure in advance.
 - Verify basic operation.
- **Final deployment:**
 - Catalog and inventory the system components.
 - Install equipment in data racks.
 - Complete cabling and other physical connectivity.
 - Verify that all units power up correctly.
 - Capture installation-specific information.
 - **Configure.**
 - **Verify.**
- **Document each step.**

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0—2-5

Network staging is an optional intermediary phase when the assembly, software installation, and configuration can be performed in an isolated environment. Some basic tests can be performed before the final deployment.

Finally, the deployment can occur and the prepared configurations can be deployed and verified to confirm the operation of the device.

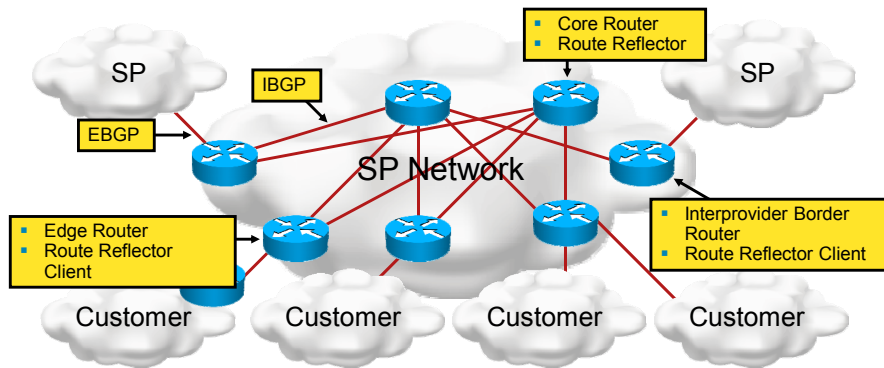
BGP Router Roles

EBGP roles:

- Edge router
- Interprovider border router

IBGP roles:

- Route reflector
- Route reflector client



© 2010 Cisco Systems, Inc. All rights reserved.

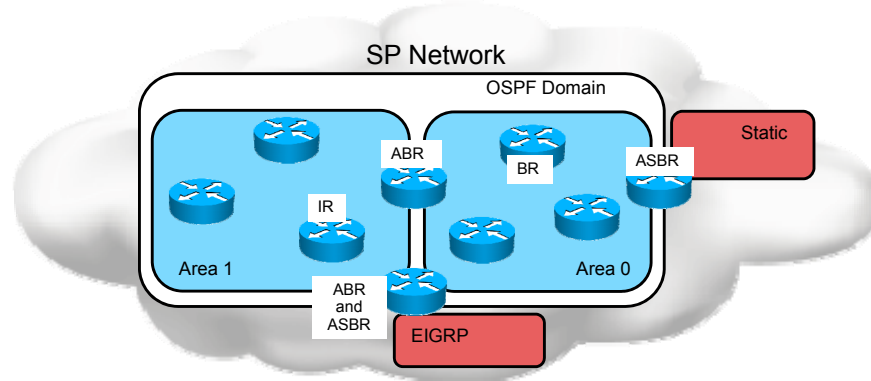
MSPRP v1.0-2.6

The figure illustrates an ISP network where a number of routers are running BGP, which performs different functions:

- Edge routers run External Border Gateway Protocol (EBGP) with customers and may have a number of requirements that are defined for this role (for example, inbound and outbound filtering, and routing policies).
- Interprovider routers also run EBGP but with other service providers. This role is typically much more complicated because it may require more complex filters and policies that may also depend on interprovider agreements.
- All edge routers are typically also clients of route reflectors.
- Route reflectors have specific requirements. Top-level route reflectors are often dedicated BGP routers that are not used to forward any traffic. In multilevel route reflector designs, intermediary route reflectors will be both clients and reflectors.

OSPF Router Roles

- Area border router (ABR)
- Autonomous system boundary router (ASBR)
- Internal router
- Backbone router



The figure illustrates a generic OSPF design with multiple areas. The figure does not show the additional roles where an Area Border Router (ABR) may need to address different types of nonbackbone areas (for example, stub area, not-so-stubby area [NSSA], totally stubby area).

The figure lists the following major roles:

- An ABR is a router that connects one or more OSPF areas to the main backbone network. It is considered a member of all areas it is connected to. An ABR keeps multiple copies of the link-state database in memory, one for each area to which that router is connected.
- An Autonomous System Boundary Router (ASBR) is a router that is connected to more than one AS and that exchanges routing information with routers in other autonomous systems. ASBRs typically also run a non-interior routing protocol (IGP), routing protocol (for example, BGP), or use static routes, or both. An ASBR is used to distribute routes that are received from other autonomous systems throughout its own AS.
- An internal router is a router that only has OSPF neighbor relationships with interfaces in the same area.
- The backbone router is a router that is part of the OSPF backbone. By definition, this backbone router includes all ABRs, since those routers pass routing information between areas. However, a backbone router may also be a router that connects only to other backbone (or area border) routers and is, therefore, not part of any area (other than Area 0).

Note An ABR is always a backbone router, but a backbone router is not necessarily an ABR.

The router types are attributes of an OSPF process. A given physical router may have one or more OSPF processes. For example, a router that is connected to more than one area and receives routes from a BGP process connected to another AS, is both an ABR and an ASBR.

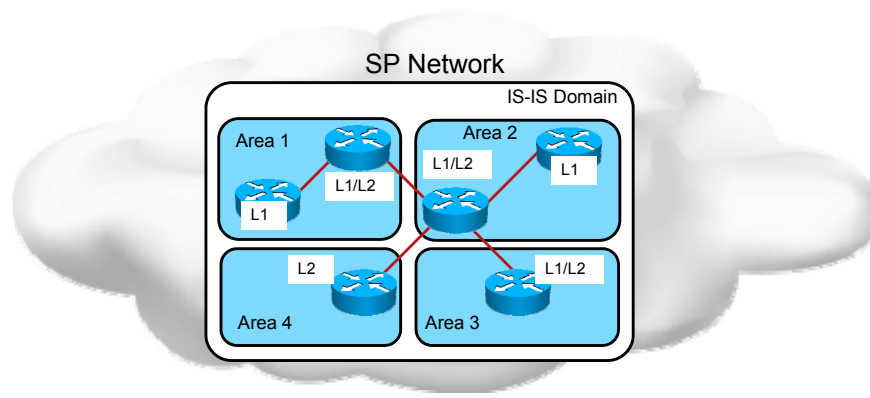
Each router has an identifier, customarily written in the dotted decimal format (for example 1.2.3.4) of an IP address. This ID must be configured to an OSPF instance to be considered the ID. If not explicitly configured, the highest logical IP address will assume the router ID role.

The router ID does not have to be a part of any routable subnet in the network and often is not to avoid confusion.

Note	Do not confuse router types with a designated router (DR) or a backup designated router (BDR), which is an attribute of a router interface, not the router itself.

IS-IS Router Roles

- Level 1 (intra-area)
- Level 2 (interarea)
- Level 1/2 (both)



© 2010 Cisco Systems, Inc. All rights reserved.

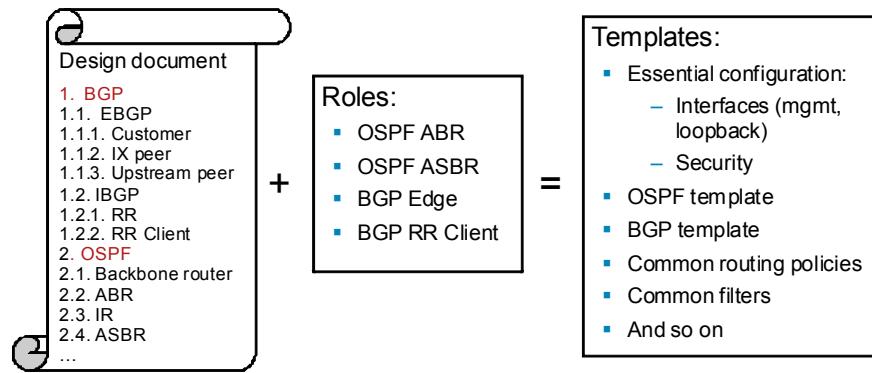
MSRP v1.0—2-8

The figure illustrates various IS-IS routers with different roles in the service provider network:

- **Level 1** routers are intra-area routers. Routing information is exchanged between Level 1 routers and other Level 1 routers in the same area.
- **Level 2** routers are interarea routers that can only form relationships with other Level 2 routers. Level 2 routers only exchange information with other Level 2 routers.
- **Level 1/2** routers exchange information with both levels and are used to connect the interarea routers with the intra-area routers.

Example: Deployment Templates

- Select the appropriate templates based on defined router roles.
- Combine design templates into the configuration template for the deployed router.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.9

The figure illustrates a process of using an existing design to create an implementation plan for a new device not specifically covered by the design or existing implementation plans. The sample device was assigned four roles, which are used to extract the template configuration from the design documents. Once the template is created, the variables need to be filled in to create a workable configuration.

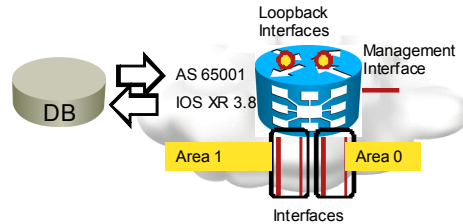
Required Parameters for Deployment

- Fixed parameters are already in the design templates:

- Timers
- Common filters
- Loopback interface numbers
- Cluster ID
- AS number
- Shared passwords

- Variable parameters are stored in a database:

- Interfaces and VLANs
- Addresses (e.g., loopback interfaces, management interfaces, core interfaces, edge interfaces, NSAP)
- Passwords—May be generated on demand and stored in the database



To finalize the configuration, the variables must be retrieved or generated to convert the template into a configuration for the new device:

- Many parameters can be included in the design documents if they are of a fixed nature (for example, BGP AS number, BGP Cluster IDs).
- Most variables have to be determined when the configuration is being created. Examples of this are IP addresses for interfaces, NSAP addresses for IS-IS processes, and routing protocol authentication passwords.

Use the company procedures to allocate available numbers to the new router, and document which resources have been allocated to prevent future duplication. Some parameters require the creation of a password, which may need to be recorded in a database. (For example, create a cryptographically strong password to authenticate new routing protocol adjacencies and record the password in a secured database.)

Parameter Database Maintenance

- Assign parameters from parameter pools
- Update parameter databases

Router	HW	Loopback0	Peers	IGP	IGP Parameters	BGP
Core 1	CRS-1	1.0.0.1	Core 2	IS-IS L2	Area 0001	RR Level 1 CID 1
Core 2	CRS-1	1.0.0.2	Core 1	IS-IS L2	Area 0001	RR Level 1 CID 1
Core 3	CRS-1	1.0.0.3	Core 1	IS-IS L2	Area 0001	RR Level 2 CID 12
Pop 1	CRS-1	1.0.1.1	Core 1 Core 2	IS-IS L1/L2	Area 0002	RRC Level 3 CID 132
Pop 2	CRS-1	1.0.1.2	Core 2 Core 3	IS-IS L1/L2	Area 0003	RRC Level 3 CID 151

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—211

When deploying a new router into a service provider network, refer to the database to obtain new router parameters.

This figure illustrates the deployment of a new Cisco Carrier Routing System (CRS-1) router:

- An available IP address was taken from the pool that is designated for loopback addresses. The address was recorded in the database to prevent any overlapping addressing in the future.
- Peering was defined and the IS-IS operational type was determined based on the role of the router.
- The router is placed into area 2, which influences the NSAP addressing.
- Such tables would contain many more parameters, which are needed for the configuration of the router.

Example: Deployment Templates

BGP—Edge Router

Cisco IOS Software

```
router bgp 23456
  bgp router-id Router-ID
  !
  ! IBGP to Route Reflectors
  neighbor 10.1.1.1 remote-as 23456
  neighbor 10.1.1.1 description Route reflector 1
  neighbor 10.1.1.1 send-community
  neighbor 10.1.1.1 password PASSWORD
  neighbor 10.1.1.2 remote-as 23456
  neighbor 10.1.1.2 description Route reflector 2
  neighbor 10.1.1.2 send-community
  neighbor 10.1.1.2 password PASSWORD
  !
  ! BGP to Customers
  neighbor CUSTOMER_IP remote-as CUSTOMER_AS
  neighbor CUSTOMER_IP description DESCRIPTION
  neighbor CUSTOMER_IP password PASSWORD
  neighbor CUSTOMER_IP prefix-list CUSTOMER_IN in
  neighbor CUSTOMER_IP prefix-list CUSTOMER_OUT out
  neighbor CUSTOMER_IP filter-list 3 in
  neighbor CUSTOMER_IP filter-list 2 out
  !
  ! Upstream ISPs
  neighbor ISP_IP remote-as ISP_AS
  neighbor ISP_IP description DESCRIPTION
  neighbor ISP_IP password PASSWORD
  neighbor ISP_IP prefix-list Upstream_In in
  neighbor ISP_IP prefix-list Upstream_Out out
  neighbor ISP_IP filter-list 2 in
  neighbor ISP_IP filter-list 4 out
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-212

The figure illustrates a portion of a created template, which was combined to include IBGP neighbors, EBGP neighbors, and common filters. The highlighted parameters must be replaced with unique values taken from various sources:

- The router ID is fixed to a specific loopback IP address.
- To maximize security, each BGP session is authenticated using a unique password that needs to be generated for both Internal Border Gateway Protocol (IBGP) peers (route reflectors) and any EBGP peer.
- Each BGP customer is establishing an authenticated BGP session, in which common or dedicated prefix lists and AS path filters are used to prevent accidental or deliberate attacks on BGP.

Example: Deployment Templates

BGP—Edge Router (Cont.)

Cisco IOS Software

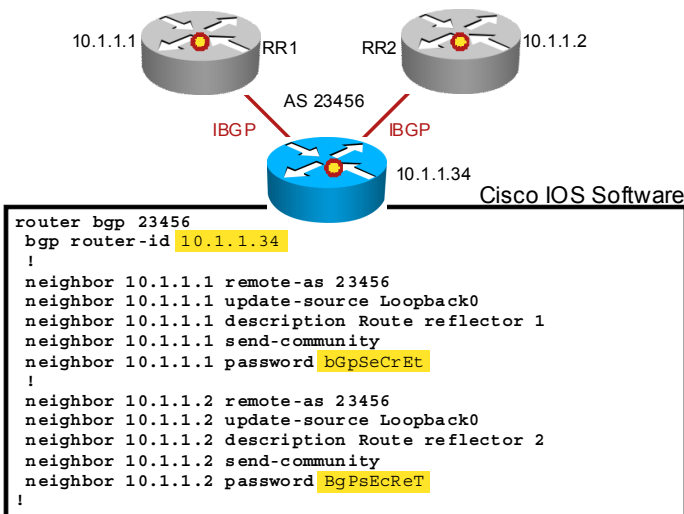
```
! Default Customer Inbound Prefix Filter
ip prefix-list CUSTOMER_IN permit CUSTOMER_PREFIX le ALLOWED_SUBNETTING
!
! Default Customer Outbound Prefix Filter
ip prefix-list CUSTOMER_OUT permit 0.0.0.0/0
!
! Default Upstream ISP Prefix Filters
ip prefix-list UpstreamIn deny 0.0.0.0/0
ip prefix-list UpstreamIn deny 10.0.0.0/8 le 32
ip prefix-list UpstreamIn deny 172.16.0.0/12 le 32
ip prefix-list UpstreamIn deny 192.168.0.0/16 le 32
ip prefix-list UpstreamIn deny OWNED_PREFIXES
ip prefix-list UpstreamIn deny OWNED_PREFIXES
...
ip prefix-list UpstreamIn permit 0.0.0.0/0 le 32
!
ip prefix-list UpstreamOut permit OWNED_PREFIXES
ip prefix-list UpstreamOut permit CUSTOMER_PREFIXES
...
!
! Permit local routes
ip as-path access-list 1 permit ^$
! Permit all routes
ip as-path access-list 2 permit .*
! Default customer inbound filter allowing AS Path prepending
ip as-path access-list 3 permit ^([0-9]+)(\1)$
! Filter to upstream ISPs
ip as-path access-list 4 permit ^$
ip as-path access-list 4 permit ^(Customer1_AS)(\1)*$
ip as-path access-list 4 permit ^(Customer2_AS)(\1)*$
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-213

- A per-customer prefix list is used to accept only the prefixes that the customer owns and to only allow some subnetting. (For example, the customer owns a /24 subnet and is allowed to subnet up to /25).
- A filter that is used in the inbound direction from upstream ISPs is designed to filter out the default route (optional), private addresses, and the ISP's own addresses.
- The ISP's own addresses and addresses belonging to customers are permitted in the outbound direction.
- AS path-based filters provide an additional layer of protection or may be used to simplify filtering when maintaining specific prefix lists is not desired. There are four common filters to match local routes (“^\$”), match any route (“.*”), match the customer's routes with optional AS path prepending (“^([0-9]+)(\1)\$”), and to match local routes and all customer routes.

Example: Deployment Templates



This figure illustrates a configuration of IBGP towards route reflectors when only the password is parameter-based; all other options are static and defined by the design. Additionally, the loopback IP address is used to fix the BGP router ID.

Example: Deployment Templates

BGP—Edge Router

```
router bgp 23456
  bgp router-id ROUTER_ID
  neighbor 10.1.1.1
    remote-as 23456
    password encrypted PASSWORD
    description Route reflector 1
    address-family ipv4 unicast
      send-community
    !
  neighbor 10.1.1.2
    remote-as 23456
    password encrypted PASSWORD
    description Route reflector 2
    address-family ipv4 unicast
      send-community-ebgp
    !
  neighbor CUSTOMER_IP
    remote-as CUSTOMER_X_AS
    password encrypted PASSWORD
    description DESCRIPTION
    address-family ipv4 unicast
      route-policy CUSTOMER_X_IN in
      route-policy CUSTOMER_X_OUT out
    !
  neighbor ISP_IP
    remote-as ISP_AS
    password encrypted PASSWORD
    description DESCRIPTION
    address-family ipv4 unicast
      route-policy ISP_IN in
      route-policy ISP_OUT out
    !
```

© 2010 Cisco Systems, Inc. All rights reserved.

Cisco IOS XR Software

```
prefix-set CUSTOMER_X_IN
  CUSTOMER_PREFIX le SUBNETTING
end-set
!
prefix-set CUSTOMER_X_OUT
  0.0.0.0/0
end-set
!
prefix-set ANY
  0.0.0.0/0 le 32
end-set
!
prefix-set DEFAULT
  0.0.0.0/0
end-set
!
prefix-set RFC1918
  10.0.0.0/8 le 32,
  172.16.0.0/12 le 32,
  192.168.0.0/16 le 32
end-set
!
prefix-set OWNED_PREFIXES
  OWNED_PREFIX,
  OWNED_PREFIX
end-set
!
```

MSPRP v1.0-315

The sample Cisco IOS XR configuration shows a similar implementation of BGP on an edge router to support IBGP sessions to a pair of route reflectors, to upstream ISPs, and to customers. When designing BGP configurations for Cisco IOS XR Software, you must properly translate Cisco IOS commands into Cisco IOS XR commands. You can use the following high-level guidelines:

- Translate prefix lists into prefix sets. Note that prefix sets do not have permit-and-deny statements. You have to group permitted and denied lines into separate prefix sets and reference them later in routing policies.
- Translate route maps, prefix-based filters, and AS path filters into routing policies.

In the example, the configuration on the left illustrates the configuration of BGP sessions. The right side illustrates the configuration of objects to use in routing policies.

Example: Deployment Templates

BGP—Edge Router (Cont.)

```
as-path-set CUSTOMER_IN
ios-regex '^[0-9]+(_\1)*$'
end-set
!
prefix-set CUSTOMER_PREFIXES
CUSTOMER_X_PREFIXES,
CUSTOMER_Y_PREFIXES,
...
end-set
!
route-policy UPSTREAM_IN
if ((destination in RFC1918) or
(destination in DEFAULT) or
(destination in OWNED_PREFIXES)) then
drop
else
pass
endif
end-policy
!
route-policy UPSTREAM_OUT
if ((destination in OWNED_PREFIXES)
or (destination in CUSTOMER_PREFIXES))
then
pass
else
drop
endif
end-policy
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

Cisco IOS XR Software

```
route-policy CUSTOMER_X_IN
if ((destination in CUSTOMER_X_IN)
and (as-path in CUSTOMER_IN)) then
pass
else
drop
endif
end-policy
!
route-policy CUSTOMER_OUT
if (destination in ANY) then
pass
endif
end-policy
!
```

MSRP v1.0—216

You can approach a two-step conversion of route maps and other filtering mechanisms to routing policies:

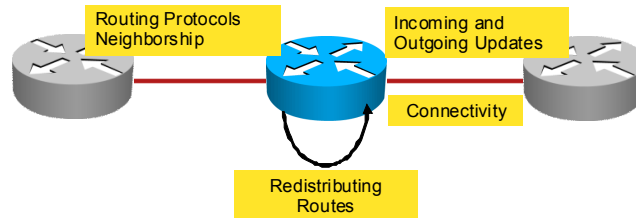
- Translate route maps, prefix lists, and AS path filters into basic routing policies.
- Optimize routing policies by combining if-statements and introducing parameterization.

Refer to online conversion guidelines to convert Cisco IOS configuration to Cisco IOS XR Software:

http://www.ciscosystems.com/en/US/docs/ios_xr_sw/iosxr_r3.5/xr12000_conversion/reference/guide/cn35main.html

Verification Steps

- Connectivity
- Routing protocol adjacencies
- Routing table
- IGP topology database
- BGP table
- Receiving and sending of routing updates (check filters, attributes)
- Redistribution



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-17

The last part is the creation of verification steps to ensure that the newly deployed and configured router is functioning properly. The verification can include:

- Connectivity tests using ping and traceroute
- Adjacency checking to ensure all IGP and BGP adjacencies are established and exchanging routing information
- Routing table verification for some key entries to ensure the reception of route
- IGP topology database to verify the type of routes received

Enabling and Tuning OSPF

This topic describes the tasks that are required to join a new router into an OSPF domain.

OSPF Implementation Tasks

- Enable basic OSPF:
 - Networks and areas
- Tune OSPF:
 - Router ID
 - Default passive interface
 - Adjacency modes
 - Interfaces
 - Hello timers
 - Authentication

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0–2.19

Basic configuration of OSPF does not require much effort (that is, only a single **network** command is needed to start OSPF on all interfaces). For security and stability reasons, however, tune OSPF to meet the requirements.

Adding OSPF Routers

OSPF implementation scenarios:

- Adding backbone routers
- Adding ABRs
- Adding nonbackbone routers

OSPF templates:

- Common OSPF configuration
- Role-specific configuration
- Router-specific configuration

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.20

To add a new router into the OSPF domain, the following scenarios are possible:

- Adding a router into the core network. All router interfaces are in the backbone area (Area 0).
- Adding a router into OSPF as ABR. Router interfaces are set into the backbone area or nonbackbone area.
- Adding a router into the OSPF nonbackbone area.

You will use different templates for adding a new router into the OSPF domain.

Common OSPF Configuration

- Router ID
- Default passive interface
- Area authentication modes and keys
- **network** command—may be area-specific
- Auto cost
- Timers (LSA, SPF)
- Incremental SPF

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-221

For an OSPF routing process to start successfully, it must be able to determine an OSPF router ID. The OSPF database uses the OSPF router ID to uniquely describe each router in the network. Remember that every router keeps a complete topology database of all routers and links in an area (or network). Therefore each router should have a unique router ID.

The OSPF routing process chooses a router ID for itself when it starts up. The router ID is a unique IP address that can be assigned in the following ways:

- By default, the highest IP address of any active physical interface when OSPF starts is chosen as the router ID. The interface does not have to be part of the OSPF process, but it has to be operating. There must be at least one operating IP interface on the router for OSPF to use as the router ID.
- If there is a loopback interface, its address will always be preferred as the router ID instead of a physical interface address—a loopback interface never goes down. If there is more than one loopback interface, then the highest IP address on any active loopback interface becomes the router ID.
- Using the **router-id** command is the preferred procedure to set the router ID and is always used in preference to the other two procedures.

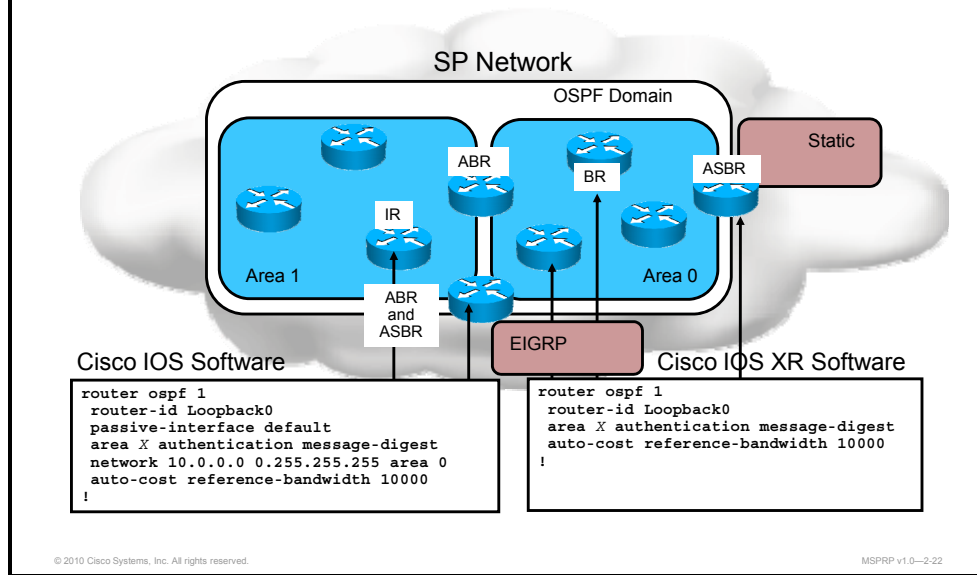
Once the OSPF router ID is set, it does not change, even if the interface that the router is using for the router ID goes down. The OSPF router ID changes only if the router reloads or if the OSPF routing process restarts.

There are a number of other configuration options that influence the behavior of OSPF, such as:

- Making all interfaces passive and then selectively enabling the interfaces that should be active (that is, OSPF should find adjacencies by sending out hello packets). This type of behavior must be configured in Cisco IOS Software, but is the default behavior in Cisco IOS XR Software. This approach is recommended for security reasons in order not to establish adjacencies on unwanted links (for example, customer links).
- Authentication is also recommended for security reasons and can be established on a per-link basis or for an entire area.

- Using the network command is the most basic way of enabling OSPF in Cisco IOS Software.
- OSPF calculates the link cost from the bandwidth value that a router applies to each interface. The best cost of 1 is assigned to Gigabit interfaces. To properly distinguish between Gigabit and 10-Gb interfaces, the auto-cost should be modified to 10000 (Megabits) or more.
- A number of other timers exist to control the sending of hello packets and to determine dead peers (for example, hello and dead intervals), the sending and processing of link-state advertisement (LSA), and so on.
- Incremental Shortest Path First (SPF) calculations can be used to optimize the Dijkstra algorithm and thus minimize the impact of topology changes on a router's CPU.

Common OSPF Configuration (Cont.)



The figure illustrates a simple, common (partial) configuration of OSPF for Cisco IOS Software and Cisco IOS XR Software. A router ID has been fixed to a loopback address and area authentication has been enabled (no keys are shown in the example). The auto-cost is configured to properly weigh interfaces up to the speed of 10 Gb/s.

Cisco IOS Software uses the **network** command to add interfaces into an OSPF process. In the Cisco IOS XR Software, the **interface** commands are used within the OSPF area configuration mode to add interfaces into an OSPF process

Role-Specific OSPF Configuration

ABR

- Filter interarea routes using:
 - Prefix lists
 - Route maps or route policies
- Summarize prefixes
- Originate a default route
- Area types (stub, NSSA)

© 2010 Cisco Systems, Inc. All rights reserved.

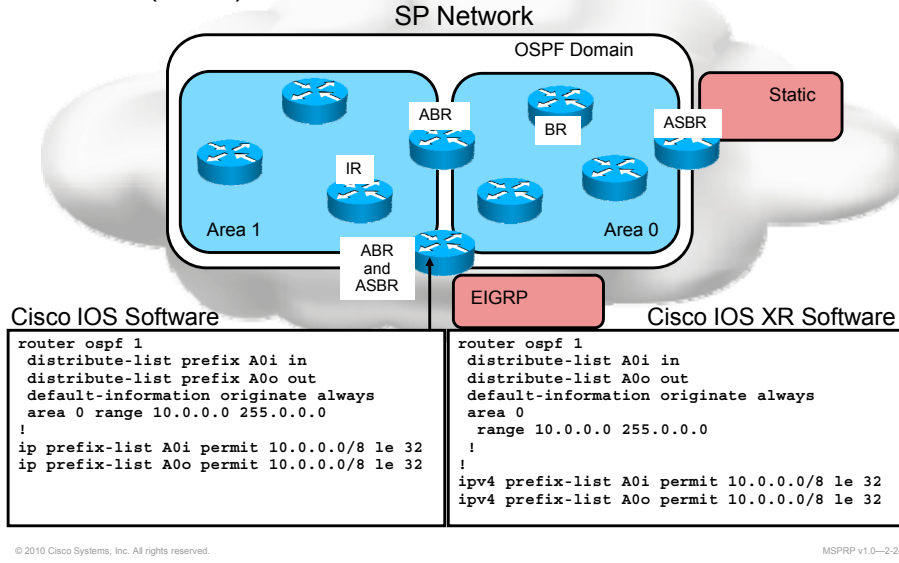
MSRP v1.0-2.23

ABRs are OSPF routers that are part of more than one area. They can perform additional functions such as:

- Filter updates passing between areas using the rules depending on the type of nonbackbone areas (for example, totally stubby area would only receive a default route)
- Filter updates using prefix lists
- Filter updates using route maps (Cisco IOS Software) or routing policies (Cisco IOS XR Software)
- Summarize prefixes to minimize the size of LSAs and of routing tables
- Originate a default route into nonbackbone areas

Role-Specific OSPF Configuration

ABR (Cont.)



The sample configurations illustrate the configuration of an ABR with the following characteristics:

- Filter incoming (into Area 0) and outgoing (from Area 0) updates to only permit prefixes in the private range of 10.0.0.0/8.
- Originate a default route into nonbackbone areas.
- Summarize all prefixes from the private range of 10.0.0.0/8 into the classful update using the **area range** command.

Role-Specific OSPF Configuration

ASBR

- Filter redistributed prefixes using:
 - Route map or route policy
- Summarize redistributed prefixes

© 2010 Cisco Systems, Inc. All rights reserved.

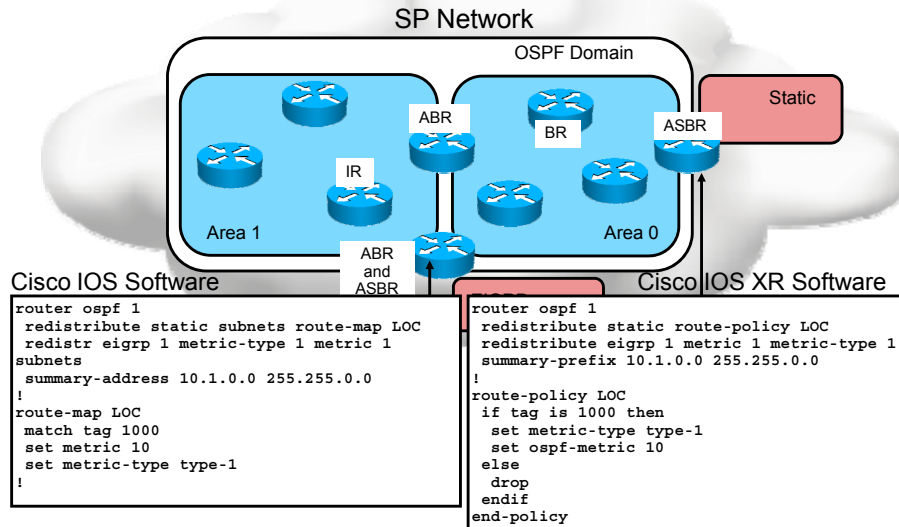
MSPRP v1.0-2.25

ASBRs can perform an additional function of interfacing with other routing protocols through redistribution. ASBRs can also perform filtering upon redistribution using route maps (Cisco IOS Software) or routing policies (Cisco IOS XR Software).

Summarization can also be completed upon redistribution.

Role-Specific OSPF Configuration

ASBR (Cont.)



The figure illustrates the configuration of redistribution from static routes that are tagged with the appropriate value. This value is matched by the route map or routing policy to set the required OSPF attributes. The sample shows that all redistributed routes are set to be of OSPF external type-1 (O E1). An external cost of 10 is set for those routes. Additionally, summarization is configured to minimize the impact of the redistributed routes on the OSPF database.

Router-Specific OSPF Configuration

- Interface options:
 - Cost configuration
 - Adjacency mode
 - Ignore MTU
 - Hello and dead intervals
 - Authentication mode and key
- OSPF process options:
 - Activate interface if default-passive used

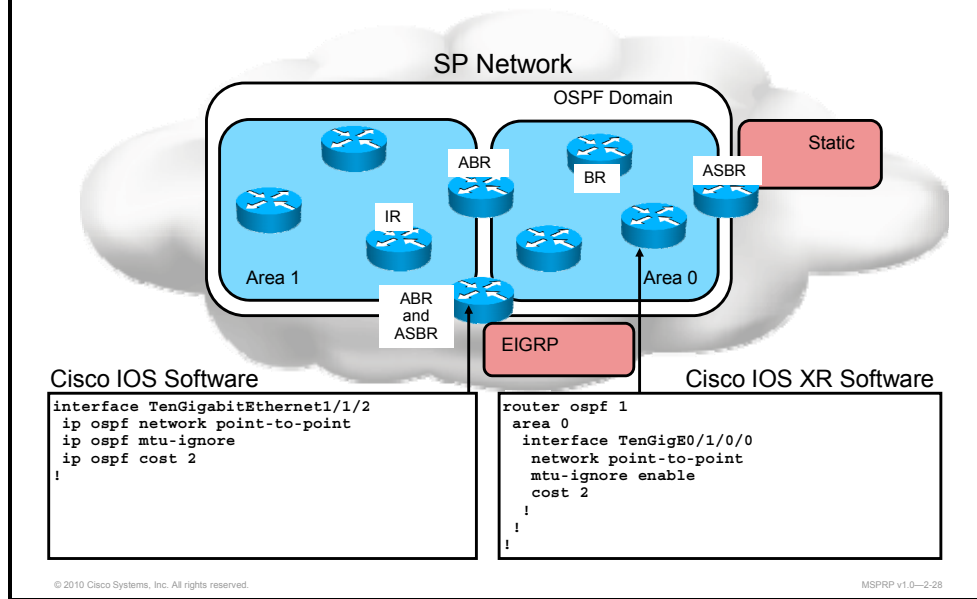
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0–227

Interface-related commands depend on the placement of the router in the network topology. Interfaces can be configured with:

- Specific cost (optional, if the auto-cost assigns the appropriate value).
- Adjacency mode, which is typically configured when using a multiaccess technology such as Ethernet on a point-to-point link.
- MTU ignore feature when MTUs configured on adjacent routers do not match.
- Hello and dead intervals can be modified to improve the time to detect failures.
- Authentication mode or keys, or both, to authenticate OSPF adjacencies.
- Activation of interfaces is required in Cisco IOS XR Software and is also used to configure the previously listed options within the OSPF configuration mode. Cisco IOS Software only requires this activation if the default passive mode has been configured.

Role-Specific OSPF Configuration



The figure shows that the 10-Gb/s interface has been manually configured with the OSPF cost of 2 (presumably to make it less preferable to some other 10-Gb/s interface). The MTU is ignored in the database descriptor packets, and the adjacency mode is set to point-to-point.

Note the difference between Cisco IOS Software, in which all per-interface options are configured in the interface configuration mode using the **ip ospf** command. In Cisco IOS XR Software, all the options are configured in the interface configuration mode within the OSPF area configuration mode.

Verifying OSPF Operation

```
RP/0/RP0/CPU0:CRS#show route ospf
```

```
Router#show ip route ospf
```

```
10.0.0.0/8 is variably subnetted, 3 subnets, 2 masks  
O IA 10.2.1.0/24 [110/782] via 10.64.0.2, 00:03:05, FastEthernet0/0  
Router#
```

```
RP/0/RP0/CPU0:CRS#show ospf interface TenGigE0/1/4/0
```

```
Router#show ip ospf interface TenGigE0/1/4/0
```

```
TenGigE0/1/4/0 is up, line protocol is up  
Internet Address 10.64.0.1/24, Area 0  
Process ID 1, Router ID 10.64.0.1, Network Type BROADCAST, Cost: 1  
Transmit Delay is 1 sec, State DROTHER, Priority 0  
Designated Router (ID) 10.64.0.2, Interface address 10.64.0.2  
No backup designated router on this network  
Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5  
oob-resync timeout 40  
Hello due in 00:00:04  
Supports Link-local Signaling (LLS)  
Index 1/1, flood queue length 0  
Next 0x0(0)/0x0(0)  
Last flood scan length is 1, maximum is 4  
Last flood scan time is 0 msec, maximum is 4 msec  
Neighbor Count is 1, Adjacent neighbor count is 1  
Adjacent with neighbor 10.64.0.2 (Designated Router)
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-229

To verify that OSPF has been properly configured, use the following **show** commands:

- The **show ip route ospf** command in Cisco IOS Software or **show route ospf** command in Cisco IOS XR Software display the OSPF routes known to the router. These commands are one of the best ways to determine connectivity between the local router and the rest of the network. This command also has optional parameters so that you can further specify the information to display, including the OSPF process ID.
- The **show ip ospf interface** (Cisco IOS Software) or **show ospf interface** (Cisco IOS XR Software) commands verify that interfaces are configured in the intended areas. In addition, these commands display other OSPF parameters such as the timer intervals, established adjacencies, and statistics.

Verifying OSPF Operation (Cont.)

```
RP/0/RP0/CPU0:CRS#show ospf neighbor
```

```
Router#show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.64.0.1	0	FULL/DROTHER	00:00:30	10.64.0.1	FastEthernet0/0
10.2.1.1	0	FULL/-	00:00:34	10.2.1.1	Serial0/0/1

```
RP/0/RP0/CPU0:CRS#show ospf neighbor detail
```

```
Router#show ip ospf neighbor detail
```

```
Neighbor 10.64.0.1, interface address 10.64.0.1
  In the area 0 via interface FastEthernet0/0
  Neighbor priority is 0, State is FULL, 16 state changes
  DR is 10.64.0.2 BDR is 0.0.0.0
<output omitted>
```

```
Neighbor 10.2.1.1, interface address 10.2.1.1
  In the area 1 via interface Serial0/0/1
  Neighbor priority is 0, State is FULL, 6 state changes
  DR is 0.0.0.0 BDR is 0.0.0.0
<output omitted>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-230

The **show ip ospf neighbor** (Cisco IOS Software) and the **show ospf neighbor** (Cisco IOS XR Software) commands display a list of neighbors. This list of neighbors includes their router ID, priority, adjacency state (for example, init, exstart, or full), and the dead timer.

Enabling and Tuning IS-IS

Identify the tasks that are required to join a new router into an IS-IS domain.

Integrated IS-IS Configuration Steps

1. Define areas, prepare addressing plan (NETs) for routers, and determine interfaces.
2. Enable IS-IS on the router.
3. Configure the NET.
4. Enable Integrated IS-IS on the appropriate interfaces. Do not forget interfaces to stub IP networks, such as loopback interfaces (although there are no CLNS neighbors there).

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-232

Four steps are required for the basic setup of IS-IS. Additional commands are available for fine-tuning the configuration. Before you configure Integrated IS-IS, you must map out the areas and plan the addressing.

Defining Area and Addressing

Area determined by NET prefix:

- Assign to support two-level hierarchy

Addressing:

- **IP**: Plan to support summarization.
- **CLNS**: Prefix denotes area. System ID must be unique.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-233

Recall that all intra-area traffic in IS-IS must traverse the Level 2 backbone area. Thus, Connectionless Network Service (CLNS) addresses must be planned to execute a two-level hierarchy.

You must decide which routers will be backbone (Level 2) routers, which routers will be Level 1/2, and which will be internal area (Level 1) routers. If some routers must do both Level 1 and Level 2 routing, then you should identify the specific interfaces that will participate in each type of routing.

Remember that the CLNS address of a router is called the NET, and it consists of three main parts:

- The prefix, which identifies the area that the router is a part of
- The system ID, which uniquely identifies each device
- The NSAP selector (NSEL), which must be 0

It is not enough to plan the IS-IS area addressing. You must also plan IP addressing to have a scalable network, and the IP addresses must be planned to allow for summarization of addresses.

Route summarization is the key idea that enables all the benefits of the hierarchical addressing design. Route summarization minimizes routing update traffic and resource utilization.

Be particularly careful when you configure the IP addressing on the router, because it is more difficult to troubleshoot IP address misconfigurations with IS-IS. The IS-IS neighbor relationships are established over Open Systems Interconnection (OSI) CLNS, not over IP. Because of this approach, two ends of a CLNS adjacency can have IP addresses on different subnets, with no impact to the operation of IS-IS.

Changing IS-IS Router Level

Router(config-router)#

```
is-type {level-1 | level-1-2 | level-2-only}
```

- Configures the IS-IS level globally on a router; the default is Level 1/2
- Same command for Cisco IOS XR Software

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—234

Optimizing IS-IS facilitates its smooth functioning and maximizes its efficiency. Remember that the default configuration of IS-IS results in the router having an intermediate system type of Level 1/2. Although this configuration allows all routers to learn of each other and pass routes without too much administrative oversight, it is not the most efficient way to build an IS-IS network.

Routers with the default configuration send out both Level 1 and Level 2 hellos and maintain both Level 1 and Level 2 link-state databases (LSDB). Each router should be configured to support the minimum level of routing required, which does the following:

- **Saves memory:** If a router does not need the LSDB for one of the levels, it will not maintain one.
- **Saves bandwidth:** Hellos and link-state packets (LSP) will be sent only for the necessary level.

If a router is to operate only as an internal area router or a backbone router, then specify this configuration by entering the **is-type** router configuration command. To specify that the router act as an internal area (Level 1)-only router, use the **is-type level-1** command. To specify that the router act as a backbone (Level 2)-only router, use the **is-type level-2-only** command.

If the level type has been changed from the default, and the router needs to return to acting as a Level 1/2 router, then use the **is-type level-1-2** command.

Changing IS-IS Interface Level

Router(config-if)#

```
isis circuit-type {level-1 | level-1-2 | level-2-only}
```

- Cisco IOS syntax
- Configures the type of adjacency on an interface; the default is Level 1/2

RP/0/RP0/CPU0:CRS(config)#

```
router isis instance-id  
interface {type interface-path-id}  
circuit-type {level-1 | level-1-2 | level-2-only}
```

- Cisco IOS XR syntax

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-235

Although the router can be a Level 1/2 router, establishing both types of adjacencies over all interfaces may not be required.

If a particular interface has only Level 1 routers that are connected to it, there is no need for the router to send Level 2 hellos out that interface. Similarly, if an interface has only Level 2 routers that are connected to it, there is no need for the router to send Level 1 hellos out that interface. Doing so would waste bandwidth and router resources trying to establish adjacencies that do not exist.

To make IS-IS more efficient in these types of situations, configure the interface to send only the needed type of hellos by using the interface command **isis circuit-type** and specifying either the **level-1** or **level-2-only** keyword.

If the circuit type is not configured, by default Cisco IOS Software attempts to establish both types of adjacencies over the interface (Level 1/2).

To configure **is-type** from the router configuration mode in Cisco IOS XR Software, you must enter interface subconfiguration mode.

Changing IS-IS Metric

Router(config-if)#

```
isis metric metric [delay-metric [expense-metric [error-metric]]] {level-1 | level-2}
```

- Configures the metric for an interface; the default is 10
- Metric value is from 1 to 63

Router(config-router)#

```
metric default-value {level-1 | level-2}
```

- Alternately, configures the metric globally for all interfaces
- The same command syntax for Cisco IOS Software and Cisco IOS XR Software

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.36

Unlike most other IP protocols, IS-IS on a Cisco router does not take into account line speed or bandwidth when it sets its link metrics. All interfaces are assigned a metric value of 10 by default. In a network with links of varying types and speeds, this assignment can result in suboptimal routing.

To change the metric value, use the **isis metric metric {level-1 | level-2}** interface command. The metric can have different values for Level 1 and Level 2 over the same interface.

Note

The IS-IS specification defines four different types of metrics. Cost, being the default metric, is supported by all routers. Delay, expense, and error are optional metrics. The delay metric measures transit delay, the expense metric measures the monetary cost of link utilization, and the error metric measures the residual error probability that is associated with a link. The default Cisco implementation uses cost only. However, Cisco IOS Software does now allow all four metrics to be set with the optional parameters in the **isis metric** command.

If the metric value for all interfaces needs to be changed from the default value of 10, then the change needs to be performed one by one on all IS-IS interfaces. This change can be time-consuming and error-prone, especially for routers with many IS-IS interfaces.

The **metric** command changes the metric value for all IS-IS interfaces. If the keyword **level-1** or **level-2** is not entered, the metric will be applied to both Level 1 and Level 2 IS-IS interfaces. This command is available only in Cisco IOS Software Release 12.3(4)T and later; it supports only the cost metric.

In this topic, you will identify three different scenarios when adding a new IS-IS router into a service provider domain:

- Adding a router into a single area as Level 1
- Adding a router into the single area and backbone area as Level 1 and Level 2
- Adding a router into the backbone only as Level 2

IS-IS Level 1 Router

Level 1 routers are referred to as station routers:

- Level 1 routers constitute an area.
- Level 1 routers keep one copy of the link-state database (its own area “picture”; intra-area information only).
- They enable “stations” to communicate.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—238

Level 1 routers are also referred to as station routers because they enable stations (end systems) to communicate with each other and the rest of the network.

A contiguous group of Level 1 routers defines an area. The Level 1 routers maintain the Level 1 database, which defines the picture of the area itself and its exit points to neighboring areas.

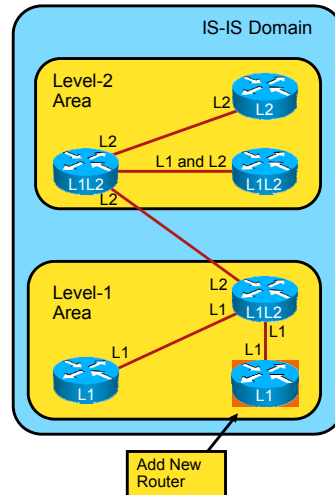
Adding a New IS-IS Router

Level 1 only for Cisco IOS Software

Cisco IOS Software

```
key chain ISIS
key 1
key-string cisco
!
interface Loopback0
isis circuit-type level-1
!
router isis
net 49.0001.1720.3125.4002.00
is-type level-1
authentication mode md5
authentication key-chain ISIS
metric-style wide
passive-interface Loopback0

interface TenGigabitEthernet1/0/0
ip router isis
isis circuit-type level-1
isis network point-to-point
isis metric 20 level-1
isis authentication mode md5
isis authentication key-chain ISIS
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-39

The figure shows a new IS-IS router configuration when added to the Level 1 area as Level 1 router. The configuration contains the following elements:

- Router IS-IS is enabled and is-type is set to Level 1 only.
- One interface is enabled into IS-IS. The interface will be able to set only Level 1 OSPF adjacency.
- A router generates and accepts only new-style type, length, and values (TLVs). Therefore, the router will use less memory and other resources than it would if it generated both old-style and new-style TLVs.

Note This style is appropriate for enabling MPLS traffic engineering across an entire network.

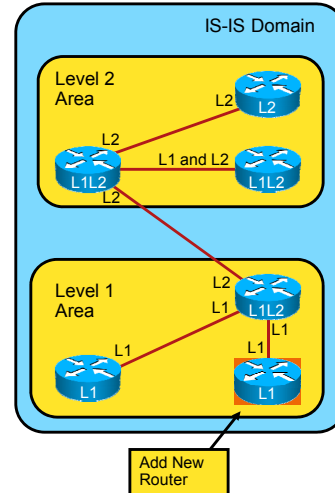
- IS-IS metric is changed for Level 1 only.
- Router will not send IS-IS hello packets to the loopback interface, which additionally minimizes resource usage.
- Authentication is enabled.

Adding a New IS-IS Router

Level 1 only for Cisco IOS XR Software

Cisco IOS XR Software

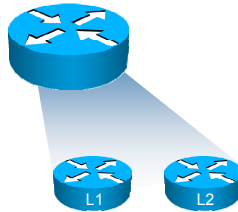
```
router isis XR
net 49.0001.1720.3125.5002.00
is-type level-1
!
address-family ipv4 unicast
metric-style wide
!
!
interface Loopback0
passive
circuit-type level-1
address-family ipv4 unicast
!
!
interface TenGigE0/0/0/0
circuit-type level-1
hello-password hmac-md5 cisco
address-family ipv4 unicast
metric 20
!
```



The sample Cisco IOS XR configuration shows the IS-IS router configuration when added into Level 1. The IS-IS routing protocol is only enabled for Level 1 globally and on all interfaces.

IS-IS Level 1/2 Router

- Level 1/2 routers keep two separate copies of the link-state databases:
 - For Level 1 and Level 2
 - Inform Level 1 routers about an exit point
- Level 1 area is a collection of Level 1 and Level 1/2 routers.
- Backbone area (Level 2) is a set of Level 1/2 and Level 2 routers and must be contiguous.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-41

Level 1/2 routers act as if they were two IS-IS routers:

- They support a Level 1 function to communicate with the other Level 1 routers in their area and maintain the Level 1 LSP information in a Level 1 topology database. They inform other Level 1 routers that they are an exit point from the area.
- They support a Level 2 function to communicate with the rest of the backbone and maintain a Level 2 topology database separately from their Level 1 database.

Note An IS-IS router can be a member of only one area—it can maintain only one Level 1 topology database.

IS-IS does not share the concept of an Area 0 with OSPF. Rather, it can appear as a set of distinct areas that are interconnected by a chain of Level 2 routers weaving their way through and between the Level 1 areas.

Adding a New IS-IS Router

Level 1/2

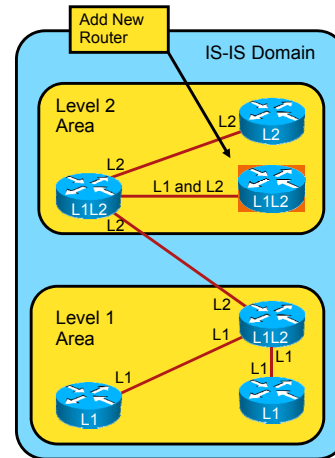
```
router isis
net 49.0001.1720.3125.4002.00
metric-style wide
is-type level-1-2
passive-interface Loopback0

interface TenGigabitEthernet1/0/0
ip router isis
isis circuit-type level-1-2
isis network point-to-point
isis metric 20 level-1
isis metric 25 level-2
```

Cisco IOS
Software

```
router isis XR
net 49.0001.1720.3125.5002.00
is-type level-1-2
!
address-family ipv4 unicast
metric-style wide
!
interface Loopback0
passive
address-family ipv4 unicast
!
!
interface TenGigE0/0/0/0
circuit-type level-1-2
address-family ipv4 unicast
metric 20 level 1
metric 25 level 2
```

Cisco IOS XR
Software



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2-42

The figure shows a Cisco IOS and Cisco IOS XR IS-IS router configuration when added to the Level 1 and Level 2 areas. The configuration contains the following elements:

- Router IS-IS is enabled. The is-type is not changed, since Level 1/2 is default.
- One interface is enabled into IS-IS. The interface will be able to set both Level 1 and Level 2 OSPF adjacencies.
- A router generates and accepts only new-style TLVs.
- The IS-IS metric is changed for Level 1 and for Level 2 separately.
- The router will not send IS-IS hello packets to the loopback interface.

IS-IS Level 2 Router

Level 2 routers are referred to as area routers:

- They store inter-area information.
- They interconnect areas.

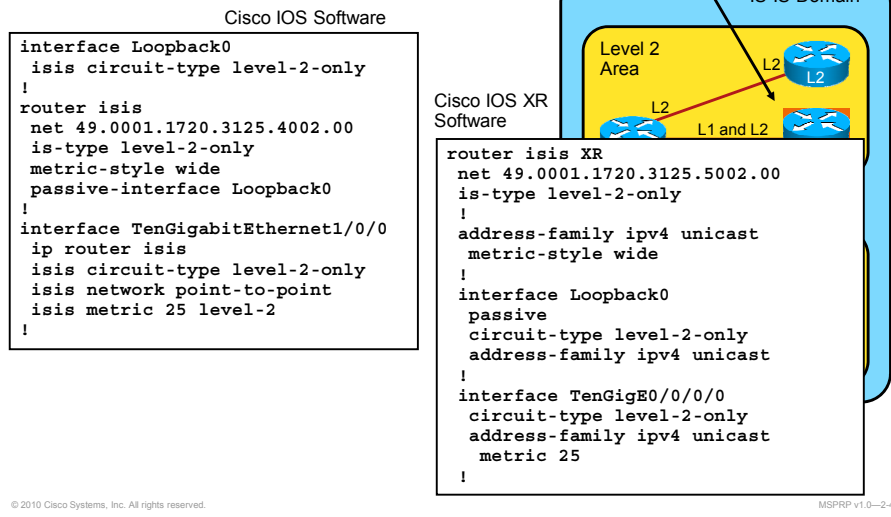
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-243

Level 2 routers are also referred to as area routers because they interconnect the Level 1 areas. Level 2 routers store a separate database, which contains only the interarea topology information.

Adding a New IS-IS Router

Level 2 only



This example shows a new IS-IS router configuration when added to the Level 2 area as Level 2 router only. The configuration contains the following elements:

- The router IS-IS is enabled and is-type is set to Level 2 only.
- One interface is enabled for IS-IS and is only enabled for Level 2 IS-IS adjacencies.
- A router generates and accepts only new-style TLVs. Therefore, the router will use less memory and other resources than it would if it generated both old-style and new-style TLVs.

Note This style is appropriate for enabling MPLS traffic engineering across an entire network.

- The IS-IS metric is changed for Level 2 only.
- The router will not send IS-IS hello packets to the loopback interface, which additionally minimizes resource usage.
- Authentication is enabled.

Verifying IS-IS Operation

show isis topology

Cisco IOS and Cisco IOS XR Software

```
R2#show isis topology
IS-IS paths to level-1 routers
System Id      Metric      Next-Hop      Interface      SNPA
R1              10          R1            Fa0/0          0016.4610.fdb0
R2              --
IS-IS paths to level-2 routers
System Id      Metric      Next-Hop      Interface      SNPA
R1              **
R2              --
R3              35          R3            Se0/0/1        *HDLC*
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-245

The **show isis topology** (Cisco IOS Software and Cisco IOS XR Software) command displays the Level 1 and Level 2 topology tables, which show the least-cost IS-IS paths to the intermediate systems.

Verifying IS-IS Operation

show clns protocol

Cisco IOS Software

```
R2#show clns protocol

IS-IS Router: <Null Tag>
System Id: 0000.0000.0002.00 IS-Type: level-1-2
Manual area address(es):
  49.0001
Routing for area address(es):
  49.0001
Interfaces supported by IS-IS:
  Loopback0 - IP
  Serial0/0/1 - IP
  FastEthernet0/0 - IP
Redistribute:
  static (on by default)
Distance for L2 CLNS routes: 110
RRR level: none
Generate narrow metrics: level-1-2
Accept narrow metrics: level-1-2
Generate wide metrics: none
Accept wide metrics: none
```

Cisco IOS XR Software

```
RP/0/RP0/CPU0:CRS#show isis protocol
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-246

The **show clns protocol** (**show isis protocol** for Cisco IOS XR Software) command displays information for the specific IS-IS processes in the router.

Verifying IS-IS Operation

show clns neighbors and show clns interface

Cisco IOS

```
R2#show clns neighbors
System Id      Interface  SNPA           State  Holdtime  Type  Protocol
R3             Se0/0/1    *HDLC*         Up     28        L2    IS-IS
R1             Fa0/0      0016.4610.fdb0 Up     23        L1    IS-IS
```

Cisco IOS XR

```
RP/0/RP0/CPU0:CRS#show isis neighbors
```

Cisco IOS

```
R2#show clns interface s0/0/1
Serial0/0/1 is up, line protocol is up
Checksums enabled, MTU 1500, Encapsulation HDLC
ERPDUs enabled, min. interval 10 msec.
CLNS fast switching enabled
CLNS SSE switching disabled
DEC compatibility mode OFF for this interface
Next ESH/ISH in 45 seconds
Routing Protocol: IS-IS
  Circuit Type: level-2
  Interface number 0x1, local circuit ID 0x100
  Neighbor System-ID: R3
  Level-2 Metric: 35, Priority: 64, Circuit ID: R2.00
  Level-2 IPv6 Metric: 10
  Number of active level-2 adjacencies: 1
  Next IS-IS Hello in 5 seconds
  if state UP
```

Cisco IOS XR

```
RP/0/RP0/CPU0:CRS#show isis interface
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-247

The **show clns neighbors** (**show isis neighbors** for Cisco IOS XR Software) command displays IS and end-system neighbors, if there are any. The neighbors are the routers with which this router has IS-IS adjacencies.

You can reduce the list of neighbors that are displayed to those neighbors across a particular interface if you specify the interface type and number in the command. The optional keyword **detail** displays the area addresses advertised by the neighbor in the hello messages.

The **show clns interface** (**show isis interface** for Cisco IOS XR Software) command displays information about the interfaces that currently run CLNS.

Enabling and Tuning IBGP

This topic identifies the tasks that are required to join a new router into a BGP autonomous system.

Adding Router to the IBGP

- Route reflector:
 - Configure cluster ID.
 - Configure route reflector clients as BGP neighbors.
 - Configure other IBGP neighbors (full mesh) as BGP neighbors.
- Route reflector client:
 - Configure route reflectors in the cluster as BGP neighbors
- IBGP peer (not RR, not RR client):
 - Configure route reflectors and other IBGP peers (full mesh) as BGP neighbors

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0–249

As part of the planning and preparation that is necessary to add a new router to the AS with a route reflector design, consider the following options:

- Add a new route reflector
- Add a new route reflector client
- Add a new internal BGP peer

Configuring Route Reflectors

```
router(config-router)#
```

```
bgp cluster-id cluster-id
```

- Optionally assigns a cluster ID to the route reflector (default value is router-ID)
- Required only for clusters with redundant reflectors
- Cluster ID cannot be changed after the first client is configured
- Same command for Cisco IOS XR Software

```
router(config-router)#
```

```
neighbor ip-address route-reflector-client
```

```
RP/0/RP0/CPU0:CRS(config)#
```

```
router bgp as-number  
address-family ipv4 unicast  
neighbor peer-IP-address  
remote-as peer-as-number  
address-family ipv4 unicast  
route-reflector-client
```

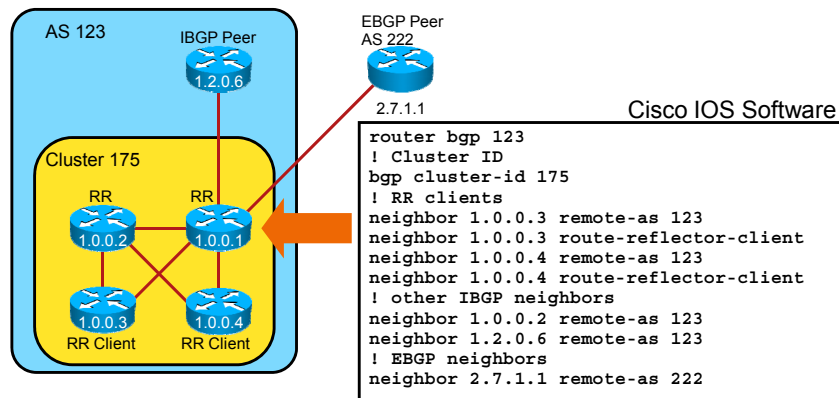
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-250

Use the **bgp cluster-id** command to configure the cluster-ID if the BGP cluster has redundant route reflectors.

The **neighbor ip-address route-reflector-client** command in the Cisco IOS Software and **route-reflector-client** command in the Cisco IOS XR Software are used to configure the router as a BGP route reflector and configure the specified neighbor as its client. When all the clients are disabled, the local router is no longer a route reflector.

Configuring Route Reflectors (Cont.)



In this example, AS 123 has been with route reflectors and has been grouped into route reflector clusters to reduce the full-mesh IBGP requirements and thus optimize the operation of IBGP.

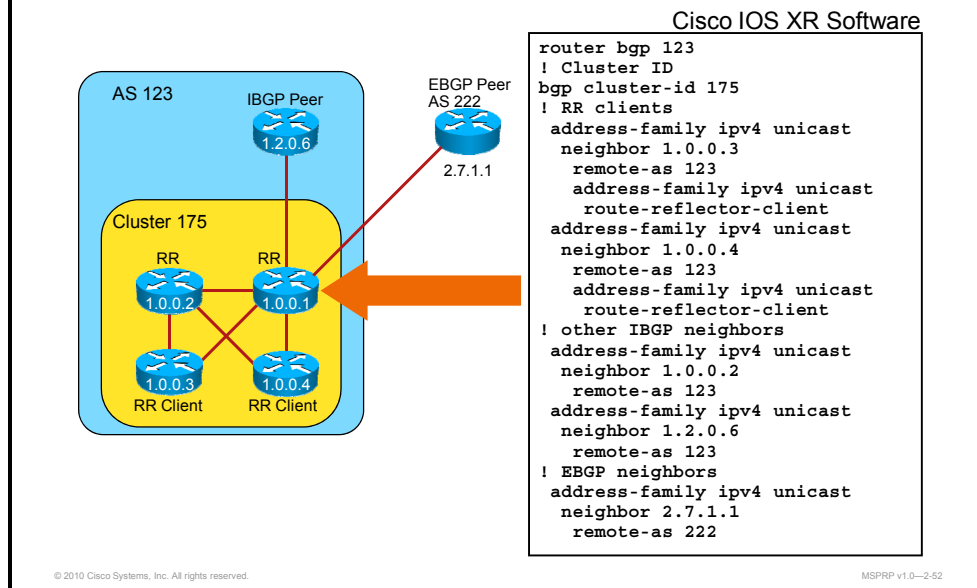
The routers with router-ID 1.0.0.1 and 1.0.0.2 are route reflectors in a cluster that has been assigned cluster-ID 175. The routers with router-ID 1.0.0.3 and 1.0.0.4 are clients to these two route reflectors.

The figure shows a portion of the configuration in router 1.0.0.1. The cluster-ID is assigned to the router under the **router bgp** process definition of the router configuration. After the router has been assigned, the route reflector client configuration is added under the **router bgp** process for the two neighbors that identify the two sessions reaching clients.

The 1.0.0.1 router is also configured for:

- IBGP peerings with router 1.0.0.2, the other route reflector within the cluster, and with router 1.2.0.6
- EBGP peering with router 2.7.1.1

Configuring Route Reflectors (Cont.)



The sample Cisco IOS XR Software configuration for route reflector configuration mainly differs from Cisco IOS Software configuration in that it is more modular (for example, neighbor subconfiguration mode, address family configuration mode for BGP, and for individual neighbors).

To optimize the configuration, multiple neighbors can be configured to inherit common parameters from an address family, a session, or a neighbor group, similar to a Cisco IOS Software session and neighbor templates.

The **af-group** command is used to group address family-specific neighbor commands within an IPv4 or IPv6 address family. Neighbors that have the same address family configuration are able to use the address family group name for their address family-specific configuration. A neighbor inherits the configuration from an address family group by way of the use command. If a neighbor is configured to use an address family group, the neighbor will (by default) inherit the entire configuration from the address family group. However, a neighbor will not inherit the entire configuration from the address family group if items are explicitly configured for the neighbor.

The **session-group** command allows you to create a session group from which neighbors can inherit address family-independent configuration. A neighbor inherits the configuration from a session group by way of the use command. If a neighbor is configured to use a session group, the neighbor (by default) inherits the session group's entire configuration. A neighbor does not inherit the entire configuration from a session group if a configuration is applied directly to that neighbor.

The **neighbor-group** command helps you apply the same configuration to one or more neighbors. Neighbor groups can include session groups and address family groups. This additional flexibility can create a complete configuration for a neighbor. Once a neighbor group is configured, each neighbor can inherit the configuration through the use command. If a neighbor is configured to use a neighbor group, the neighbor (by default) inherits the neighbor group's entire BGP configuration.

Inherited parameters can be overridden on a per neighbor basis.

Monitoring Route Reflectors

Cisco IOS Software

```
Router#show ip bgp neighbors 1.0.0.1
BGP neighbor is 1.0.0.1, remote AS 213, internal link
  Index 1, Offset 0, Mask 0x2
    Route-reflector Client
      BGP version 4, remote router ID 11.0.0.1
      BGP state = Established, table version = 5, up for 01:33:24
      Last read 00:00:24, hold time is 180, keepalive interval is 60 seconds
< text omitted >
```

Cisco IOS XR Software

```
RP/0/RP0/CPU0:CRS#show bgp neighbors 1.0.0.1
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—253

To display information about the TCP and BGP connections to neighbors, use the **show ip bgp neighbors EXEC** command. Use the **show bgp neighbors** command in the Cisco IOS XR Software.

In this case, the **show ip bgp neighbors** command is used on the router, not to see routes or paths that have been received, but to see the status of the neighbor session. Therefore, no other qualifiers than the optional IP address are given.

The **show ip bgp neighbors** command, issued on the route reflector router, indicates that the neighbor is a route reflector client.

Monitoring Route Reflectors (Cont.)

Cisco IOS Software

```
Router#show ip bgp 11.0.0.0
BGP routing table entry for 11.0.0.0/8, version 3
Paths: (1 available, best #1, advertised over IBGP)
  Local, (Received from a RR-client)
    1.0.0.1 (metric 40640000) from 1.0.0.1 (11.0.0.1)
      Original IGP, metric 0, localpref 100, valid, internal, best
```

Cisco IOS XR Software

```
RP/0/RP0/CPU0:CRS#show bgp 11.0.0.1
```

- Routes received from the client, as seen on the reflector

Cisco IOS Software

```
Router#show ip bgp 14.0.0.0
BGP routing table entry for 14.0.0.0/8, version 30
Paths: (1 available, best #1)
  Not advertised to any peer
    Local
      1.0.0.3 (metric 41152000) from 1.0.0.2 (14.1.2.3)
        Original IGP, metric 0, localpref 100, valid, internal, best
        Originator: 14.1.2.3, Cluster list: 0.0.0.175
```

Cisco IOS XR Software

```
RP/0/RP0/CPU0:CRS#show bgp 14.0.0.0
```

- Reflected routes, as seen on the client

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-254

To display entries in the BGP routing table, use the **show ip bgp** (**show bgp** for Cisco IOS XR Software) EXEC command.

When details are displayed for a specific route entry in the BGP table, the cluster-list and originator-ID attributes are also shown.

Note Note that the values in the cluster list are displayed in the IP format even though they were configured as a single number. The configured cluster ID 567 is converted to 0.0.2.55 (2 x 256 + 55).

The first example of a **show ip bgp** command is issued on a route reflector router. It shows that this particular entry in the BGP table was received from a route reflector client.

The second example shows an entry in the BGP table that at some point was reflected from a route reflector. The reflecting router has added the originator-ID and cluster-list attributes to the route.

Summary

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

Summary

- When adding a new router to the service provider network, these are the deployment steps: plan, design, develop, configure, install, and verify.
- RFC 4098 defines the following types of BGP-speaking routers: edge router, subscriber edge router, interprovider border router, and core router.
- The OSPF router types in the service provider network are: ABR, ASBR, IR, and BR.
- The IS-IS router types are: Level 1, Level 2, and Level 1/2 routers.
- IBGP routers in the route reflector scenario are: route reflector, route reflector client, and internal BGP peer.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-255

Lesson Self-Check

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

- Q1) What is the last deployment step when adding a new router to the service provider network? (Source: Change Management for Adding New Core Routers)
- A) plan
 - B) verify
 - C) design
 - D) configure
- Q2) In OSPF, which router connects one or more OSPF areas to the main backbone network? (Source: Change Management for Adding New Core Routers)
- A) Autonomous system border router (ASBR)
 - B) Area Border Router (ABR)
 - C) Internal router (IR)
 - D) Backbone Router (BR)
- Q3) In IS-IS, which routers exchange routing information only with routers in the same area? (Source: Change Management for Adding New Core Routers)
- A) Level 1
 - B) Level 2
 - C) Level 1/2
- Q4) In OSPF, what is the preferred procedure to set the router ID and is used in preference to the other two procedures? (Source: Enabling and Tuning OSPF)
- A) setting highest IP address on physical interface
 - B) setting IP address on loopback interface
 - C) using router-id command
- Q5) In IS-IS, how is metric assigned to interfaces? (Source: Enabling and Tuning IS-IS)
- A) based on bandwidth and delay of an interface
 - B) based on bandwidth of an interface
 - C) metric on all interfaces is by default 100
 - D) metric on all interfaces is by default 10

Lesson Self-Check Answer Key

Q1)	B
Q2)	B
Q3)	A
Q4)	C
Q5)	D

Modifying Core Routing Parameters

Overview

In this lesson, you will identify the tasks that are required to modify core routing parameters in the core routing domain and implement the changes according to the change request. This lesson describes typical routing design changes and describes how a new prefix is added to the routing domain. The lesson also describes different modification and tuning processes for service provider routing protocols.

Objectives

Upon completing this lesson, you will be able to identify the tasks required to modify core routing parameters in the core routing domain and implement the changes according to the change request. This ability includes being able to meet these objectives:

- Describe typical routing design changes
- Add a new prefix to the routing domain
- Modify route filters

Modifying Routing Parameters

This topic describes typical routing design changes.

Modifying Routing Parameters

- **Modification:**
 - Adding
 - Changing
 - Removing
- **Reasons:**
 - Redesign
 - Optimization
 - Fault mitigation



© 2010 Cisco Systems, Inc. All rights reserved. MSPRP v1.9-2.3

To change routing protocol parameters, there may be different modifications:

- Adding a routing protocol on the router or adding an interface to the existing routing protocol
- Changing routing protocol filters, redistribution parameters, timers, and metric
- Removing the routing protocol from the router or removing an interface from the routing protocol

The following are reasons to modify routing parameters:

- Redesigning a network to meet new business requirements
- Optimizing a network to use the same resources with better performance
- Mitigating faults with redundancy and tuned timers

Modifying Routing Parameters (Cont.)

- **Scope:**

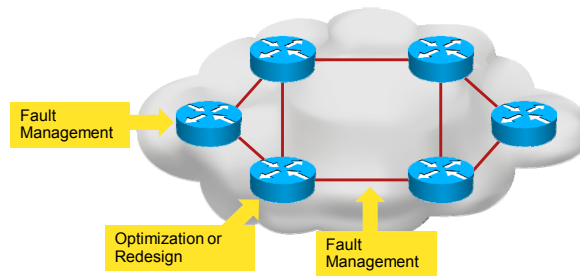
- Network
- Router
- Link

- **Impact:**

- Disruptive
- Nondisruptive

- **Additional actions:**

- Clear process or neighbor relationships
- Remove and add configuration
- Reload router



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9—2-4

Modifying routing parameters may be scoped to the whole service provider network, to the router, or to the link.

The impact to the network when changing routing parameters may be disruptive or non-disruptive. The impact is disruptive when users or applications lose network service for many seconds, minutes, or even hours. Non-disruptive impact is when users or applications do not lose network service or will lose it for a few seconds or less than a second.

The following are additional actions in maintaining routing protocols:

- Clearing process (interior gateway protocol [IGP]) or neighbor relationship (Border Gateway Protocol [BGP])
- Removing and adding the same or a changed configuration
- Reloading the router after a major configuration change

Modifying Routing Parameters (Cont.)

- **Disruptive modifications:**
 - Longer outage
 - Reload required
 - Reconfiguration required
 - Disruptive networkwide modifications
- **Nondisruptive modifications:**
 - No outage or very short outage (seconds)
 - No additional actions required after modification
 - Example: **clear ip route ***
 - Typically very short outage (<< 1 second)

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9–2.5

Disruptive modifications result in the following:

- Longer outage
- Reload required
- Reconfiguration required
- Disruptive networkwide modifications

Non-disruptive modifications result in the following:

- No outage or very short outage (seconds)
- No additional actions are required after modification
- Example: “**clear ip route ***”
- Typically very short outage (less than one second)

OSPF Modifications

- Adding links to an area
- Adding networks
- Modifying summarization
- Router ID
- Filter (prefix list, route map, or policy)
- Incremental SPF
- Timers (LSA, pacing, throttle, hello, dead interval)
- Authentication keys
- Authentication method

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-2-6

The following are possible modifications in the Open Shortest Path First (OSPF):

- Adding links to an area.
- Adding networks to the OSPF process.
- **Modifying summarization:** This modification is normally implemented at the boundaries of Area Border Routers (ABRs). Although summarization could be configured between any two areas, it is better to summarize in the direction of the backbone. This way the backbone receives all the aggregate addresses and in turn injects them, already summarized, into other areas. There are two types of summarization: interarea route summarization and external route summarization.
- **Changing the OSPF router ID:** Once an OSPF router ID is selected, it is not changed, even if the interface that was used to select it changes its operational state or its IP address. To change the OSPF router ID, you have to reset the OSPF process with the **clear ip ospf process** command (even when the new router ID was requested with the **router-id** router configuration command).
- **Filtering:** There are many filtering methods available in the OSPF routing protocol (prefix list, route map, or policy).
- **Incremental Shortest Path First (SPF):** Incremental SPF is more efficient than the complete SPF algorithm, allowing OSPF to converge faster on a new routing topology in reaction to a network event.
- **Timers:** You can perform the following modifications to the OSPF timers: modify the hello interval, control the link-state advertisement (LSA) retransmission interval, modify the router dead interval, and specify the transit delay.
- Changing authentication keys.
- **Changing authentication method:** You can use simple password authentication or Message Digest 5 (MD5) authentication.

OSPF Modifications (Cont.)

- Interface MTUs
- MTU handling
- Link's area membership
- Reference bandwidth
- Interface cost
- Default metric for redistributed routes
- Network type
- Priority

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-2-7

The following are additional modifications in the OSPF:

- Interface maximum transmission units (MTUs).
- **MTU handling:** OSPF checks whether neighbors are using the same MTU on a common interface. This check is performed when neighbors exchange database description (DBD) packets. If the receiving MTU in the DBD packet is higher than the IP MTU configured on the incoming interface, OSPF adjacency will not be established. To disable OSPF MTU mismatch detection on receiving DBD packets, use the **ip ospf mtu-ignore** command in interface configuration mode.
- Link's area membership.
- **Reference bandwidth:** You can change the reference bandwidth under **router ospf**. By default, the reference bandwidth is 100 Mb/s.
- **Interface cost:** By default, the cost of an interface is calculated based on the bandwidth, but you can change the cost of an interface.
- **Default metric for redistributed routes:** Redistributing routes into OSPF from other routing protocols or from static will cause these routes to become OSPF external routes. If no metric is specified, OSPF puts a default value of 20 when redistributing routes from all protocols except BGP routes, which get a metric of 1.

- **Network type:** The type of network to which the OSPF interface is connected determines OSPF adjacency. One of the following OSPF network types must be selected when configuring an interface on an OSPF router: Broadcast, point-to-point, nonbroadcast multiaccess (NBMA).
- **Priority:** The router with the highest OSPF priority on a segment will become the designated router (DR) for that segment. The same process is repeated for the backup designated router (BDR). In a tie, the router with the highest router ID will win. The default for the interface OSPF priority is one. Remember that the DR and BDR concepts are per multiaccess segment. A priority value of zero indicates an interface, which should not be elected as DR or BDR. The state of the interface with priority zero will be DROTHER, a router that is neither a DR nor a BDR.

IS-IS Modifications

- Authentication keys and methods
- Interface MTUs
- Hello padding
- Level or area membership (NSAP address change)
- Incremental SPF
- LSP suppression
- LSP maximum size
- LSP timers (regeneration, refresh, lifetime)
- SPF timers
- Metric style (narrow, wide)
- Attach and overload bits

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-2.8

A number of modifications are possible in the Intermediate System-to-Intermediate System (IS-IS):

- Authentication keys and methods.
- Interface MTUs.
- IS-IS hellos are padded to the complete MTU size. The benefit of padding IS-IS hellos to the complete MTU is that it allows for early detection of errors that result from transmission problems with large frames or errors that result from mismatched MTUs on adjacent interfaces. You can disable hello padding in order to avoid wasting network bandwidth in case the MTU of both interfaces is the same or, in the case of translational bridging. While hello padding is disabled, Cisco routers still send the first five IS-IS hellos padded to the complete MTU size, in order to maintain the benefits of discovering MTU mismatches.
- Level or area membership (network service access point [NSAP] address change).
- IS-IS can be configured to use an incremental SPF algorithm for calculating the SPF routes. Incremental SPF is more efficient than the complete SPF algorithm, allowing IS-IS to converge faster on a new routing topology in reaction to a network event.
- Link-state packet (LSP) suppression.
- LSP maximum size.
- LSP timers (regeneration, refresh, lifetime).
- SPF, partial route calculation (PRC) and LSP generation timers need to be tuned according to the level of stability of the network and the stability that is required in the routing domain. For instance, setting low values will trigger a fast convergence with a potential risk of high resource utilization if there are flapping routes that cause network turnover. Setting high values will keep the network stable with slower convergence.
- Metric style (narrow, wide).
- Attach and overload bits.

BGP Modifications

- BGP timers (scan)
- Maximum paths
- Networks
- Summarization (aggregation)
- Redistribution (route map or policy)
- Route selection (best path)
- Dampening

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.9—2-9

Possible modifications in the BGP:

- BGP timers (scan)
- Maximum paths
- Networks
- Summarization (aggregation)
- Redistribution (route map or policy)
- Route selection (best path)
- Dampening

BGP Modifications (Cont.)

Neighbor parameters:

- Affecting **sessions**:
 - Remote AS
 - Neighbor's IP address
 - Hold time and keepalive
 - Authentication key
- Affecting **update processing**:
 - Route reflector client
 - Advertisement interval
 - Filter (prefix list, AS path filter, route map, or policy)
 - Next-hop processing
 - Community propagation
 - Soft reconfiguration

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9—2-10

Neighbor parameter modifications that affect BGP sessions and BGP update processing include:

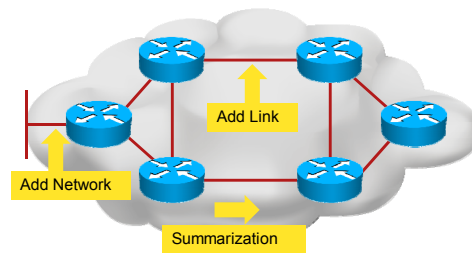
- Affecting BGP sessions:
 - Remote autonomous system (AS)
 - Neighbor's IP address
 - Holdtime and keepalive
 - Authentication key
- Affecting BGP update processing:
 - Route reflector client
 - Advertisement interval
 - Filter (prefix list, AS path filter, route map, or policy)
 - Next-hop processing
 - Community propagation
 - Soft reconfiguration

Adding New Prefixes

This topic describes how a new prefix is added to the routing domain.

Adding New Prefixes

- Adding new prefixes into routing domain:
 - Link to an area
 - Network
 - Summarization
- The scope is area or network.
- The impact is nondisruptive.
- Deployment options:
 - Gradual deployment
 - Overnight migration (that is, requires a maintenance window)



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9--2.12

To add a new prefix into a routing domain, these are the following scenarios:

- Add link to an area
- Add network into a routing domain
- Summarization on the ABRs or Autonomous System Boundary Routers (ASBRs)

The scope of changes when adding a new prefix into a routing domain is area or network. The impact of many of those changes is non-disruptive.

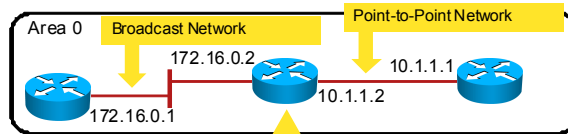
These are the deployment options:

- Gradual deployment with some deadlines and time limitations
- Overnight migration with a maintenance window required

Adding a Prefix to an OSPF Domain

Link and Network

- Adding new prefixes into an OSPF routing domain:
 - Link to an area
 - Network
 - Summarization



Cisco IOS Software

```
router ospf 1
! adding network
network 172.16.0.0 0.0.0.255 area 0
! adding link
network 10.1.1.2 0.0.0.0 area 0
!
```

Cisco IOS XR Software

```
router ospf xr
area 0
interface TenGigE0/1/4/0
network point-to-point
!
interface TenGigE0/1/4/1
network broadcast
!
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-213

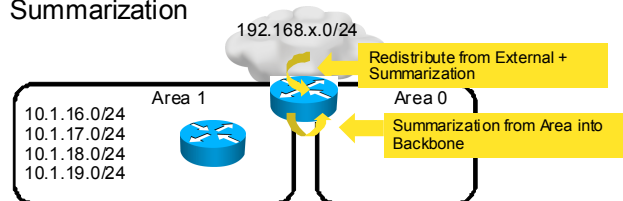
The figure shows Cisco IOS and Cisco IOS XR Software examples of how to add a link or network to the OSPF area. In the Cisco IOS XR configuration example, the OSPF network type on the 10.1.1.2 link is also set to point-to-point. The OSPF network type on the 172.16.0.2 link is set to broadcast.

Adding a Prefix to OSPF Domain

Summarization

Adding new prefixes into an OSPF routing domain:

- Link to an area
- Network
- Summarization



Cisco IOS Software

```
router ospf 1
! Inter-Area Route Summarization
area 1 range 10.1.16.0 255.255.252.0
! External Route Summarization
summary-address 192.168.0.0
255.255.0.0
!
```

Cisco IOS XR Software

```
router ospf 1
area 1
range 10.1.16.0 255.255.252.0
!
summary-prefix 192.168.0.0
255.255.0.0
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-214

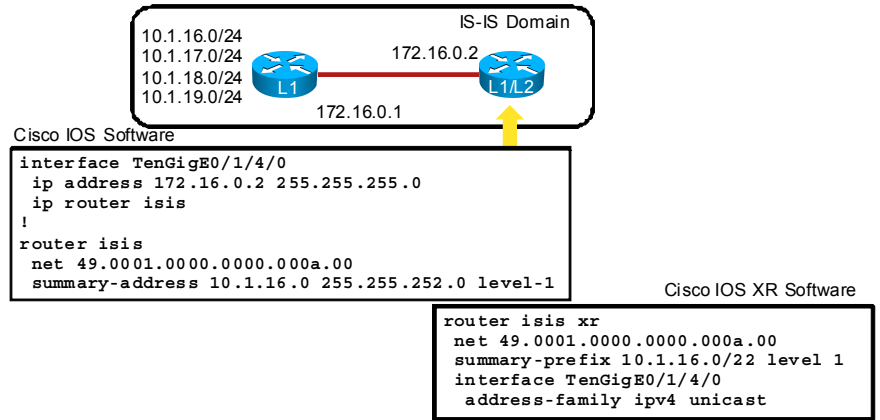
This figure shows Cisco IOS and Cisco IOS XR Software examples of how to add a prefix to an OSPF domain by summarization. The example shows summarization when external routes are redistributed and summarized into an OSPF domain as 192.168.0.0/16 and summarization from the nonbackbone area (Area 1) into the backbone area (Area 0) as 10.1.16.0/22.

Adding a Prefix to IS-IS

Link and Summarization

Adding new prefixes into IS-IS routing domain:

- Link to IS-IS
- Summarization



This figure shows Cisco IOS and Cisco IOS XR Software examples of how to add a prefix to an IS-IS domain. The example shows adding a link to an IS-IS domain and summarization of networks from the L1 into L2 area as 10.1.16.0/22.

Announcing Prefixes 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.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9–2.16

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

There are two ways to do this configuration:

- List the network numbers that are candidates to be advertised using the network configuration command. If any of the listed networks are reachable by the local router according to its routing table, then the network is injected as a route into the BGP table.
- Redistribute routing information that has been learned by other routing protocols into the BGP table. You can apply 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.

Adding a Prefix to BGP

network Command

- The addition of a **route map** (Cisco IOS Software) or a **route policy** (Cisco IOS XR Software) allows BGP parameters and attributes to be defined.
- The route map or route policy option can be used for the following:
 - Changing the **weight** 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

Cisco IOS Software

```
router bgp 1
network 10.1.1.0 mask 255.255.255.0 route-map SetAttr
!
```

Cisco IOS XR Software

```
router bgp 1
address-family ipv4 unicast
network 10.1.1.0/24 route-policy SetAttr
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-217

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 (Cisco IOS Software) or policy (Cisco IOS XR Software) option in the basic network command.

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

- **Weight (default value = 32768):** The weight attribute that is only local to a router 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. A weight can be a number from 0 to 65,535. Paths that the router originates have a weight of 32,768 by default, and other paths have a weight of 0.
- **Community (default value = nonexistent):** Used for tagging routes at their source.
- **Local preference (default value = 100):** Unlike the weight attribute, the local preference attribute is propagated throughout the local AS. The local preference is used for AS-wide BGP best-path selection.
- **Multi-exit discriminator (MED) (default value = 0):** Used for return-path selection in topologies where multiple exit points to the same neighbor AS exist.

Adding a Prefix to BGP

Redistributing

- 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
- Redistribution using route maps (Cisco IOS Software) or route policy (Cisco IOS XR Software)
 - Origin can be set to IGP
 - Other BGP path attributes can also be set:
 - Metric
 - Next-hop
 - Community

Cisco IOS Software

```
router bgp 1
 redistribute ospf 1 route-map OSPF_into_BGP
!
```

Cisco IOS XR Software

```
router bgp 1
 address-family ipv4 unicast
  redistribute ospf 1 route-policy OSPF_into_BGP
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-218

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 complete control over networks that can 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 can be difficult to maintain.

If there are many 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."

Route maps (Cisco IOS Software) or policy (Cisco IOS XR Software) can be configured on the router to filter updates and modify various attributes. A configured route map or policy can be applied to routes being redistributed from the IGP.

Adding a Prefix to BGP

Aggregating

- Summarization is called “aggregation” in BGP.
- Aggregation creates summary routes from networks already in BGP table.
- Individual networks can be announced or suppressed.
- Benefits:
 - Smaller BGP routing tables
 - More stable networks (contained route flapping)
- Drawback:
 - Potential issues with multihomed customers

Cisco IOS Software

```
router bgp 1
 aggregate-address 10.1.0.0 255.255.0.0 summary-only
!
```

Cisco IOS XR Software

```
router bgp 1
 address-family ipv4 unicast
 aggregate-address 10.1.0.0/16 summary-only
!
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–219

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.

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 the summary indicates, then the summary route is also injected into the BGP table on the route and advertised to other routers. This action creates more information in the BGP table. To get any benefits from the aggregation, you must suppress the more specific routes that are covered by the route summary. This suppression is an option to the **aggregate** configuration command.

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

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

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 a single route summary advertises. 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.

Maintaining OSPF and IS-IS Filters

This topic describes how routing protocol filters are maintained.

Maintaining Routing Protocol Filters

- Simple prefix-based filtering using prefix lists
- Complex filtering using route maps or route policies
- OSPF filters on ABRs and ASBRs (redistribution)
- IS-IS filtering between levels or upon redistribution
- BGP filtering using prefix lists, AS path access lists, route maps or route policies, or even route refresh with outbound route filtering

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9—221

Routing protocols can make use of various filters at different locations. OSPF can use filtering on ABRs between areas or on ASBRs upon redistribution. Simple prefix-based filtering can be implemented using prefix lists. More complex filtering requires the usage of route maps (Cisco IOS Software) or routing policies (Cisco IOS XR Software).

OSPF Route-Map Based Filtering

- Filtering happens at the moment when OSPF is installing the route in the routing table.
- Route map could be based on the following match options:
 - Interface
 - IP address
 - IP next-hop
 - IP route-source
 - Metric
 - Route type
 - Tag
- In the Cisco IOS XR Software, route policy must be used instead route map.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9–222

Users can define a route map to prevent OSPF routes from being added to the routing table. This filtering happens at the moment when OSPF is installing the route in the routing table. This feature has no effect on LSA flooding. In the route map, the user can match on any attribute of the OSPF route. That is, the route map could be based on the following match options:

- Interface
- IP address
- IP next hop
- IP route source
- Metric
- Route type
- Tag

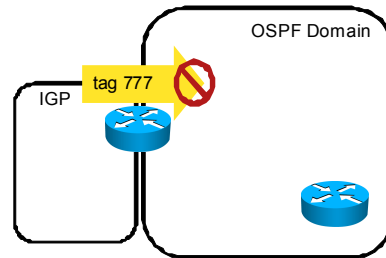
In the Cisco IOS XR Software, route policies are used instead of route maps.

Configuring OSPF Route Map-Based Filtering

- Useful during redistribution
- Filtering based on:
 - Route tag
 - Route type
 - Route source
 - Interface
 - Next hop

Cisco IOS Software

```
route-map tag-filter deny 10
 match tag 777
!
route-map tag-filter permit 20
!
router ospf 1
 distribute-list route-map tag-filter in
!
```



© 2010 Cisco Systems, Inc. All rights reserved.

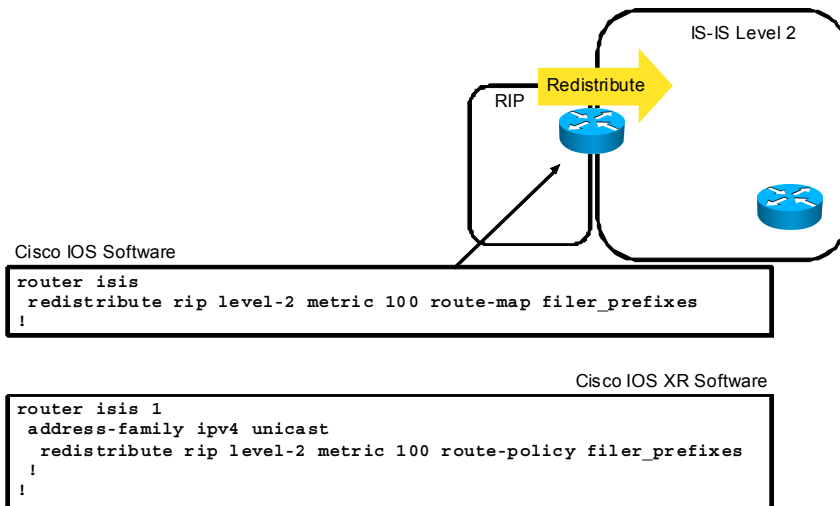
MSRP v1.9-223

This feature can be useful during redistribution if the user tags the prefixes when they get redistributed on ASBRs and later uses the tag to filter the prefixes from being installed in the routing table on other routers.

- **Filtering based on route tag:** Users can assign tags to external routes when they are redistributed to OSPF. Then the user can deny or permit those routes in the OSPF domain by identifying that tag in the route-map and distribute-list in commands.
- **Filtering based on route type:** In OSPF, the external routes could be Type 1 or Type 2. Users can create route maps to match either Type 1 or Type 2 and then use the distribute-list in command to filter certain prefixes. Also, route maps can identify internal routes (interarea and intra-area), and then those routes can be filtered.
- **Filtering based on route source:** When a match is done on the route source, the route source represents the OSPF Router ID of the LSA originator of the LSA in which the prefix is advertised.
- **Filtering based on interface:** When a match is done on the interface, the interface represents the outgoing interface for the route that OSPF is trying to install in the routing table.
- **Filtering based on next-hop:** When a match is done on the next hop, the next hop represents the next hop for the route that OSPF is trying to install in the routing table.

In this Cisco IOS example, OSPF external LSAs have a tag. The value of the tag is examined before the prefix is installed in the routing table. All OSPF external prefixes that have the tag value of 777 are filtered (prevented from being installed in the routing table). The permit statement with sequence number 20 has no match conditions. There are no other route map statements after sequence number 20, so all other conditions are permitted.

IS-IS Filtering



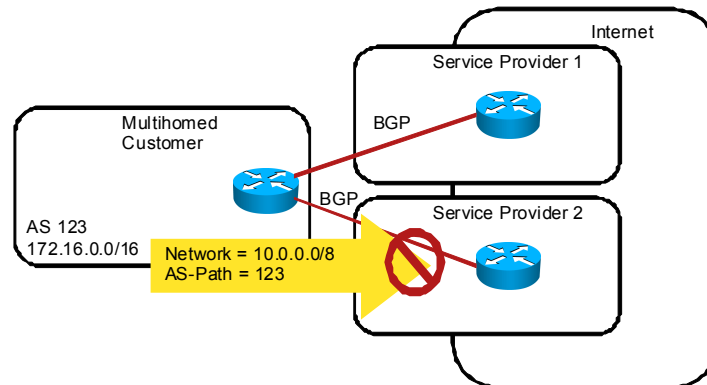
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-224

The example shows a route map-based (Cisco IOS Software) and policy-based (Cisco IOS XR Software) filtering when redistributing Routing Information Protocol (RIP) routes into an IS-IS routing protocol. Routes are filtered with a route map or a policy into IS-IS Level 2 with metric 100.

Requirements for Prefix-Based Filters

- Service providers must filter customer updates to ensure that the customers announce only their assigned address space.



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, which the customer may use for address links, loopback interfaces, or other devices that should never access the Internet.

To avoid problems, the service provider can apply an IP prefix filter on the incoming information from the customer. The service provider will accept only network numbers that are permitted by an access list or prefix list.

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, and so on).
- The traditional way of accomplishing this function is to clear the BGP session.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9—226

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.

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)

Route Refresh (Cont.)

- Inbound soft reconfiguration consumes memory on the receiving router.
- It is needed only because there is no mechanism in standard BGP to request retransmission of BGP routes.
- BGP route refresh is an optional BGP capability that allows a BGP router to request retransmission of BGP routes from a neighbor.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–227

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.

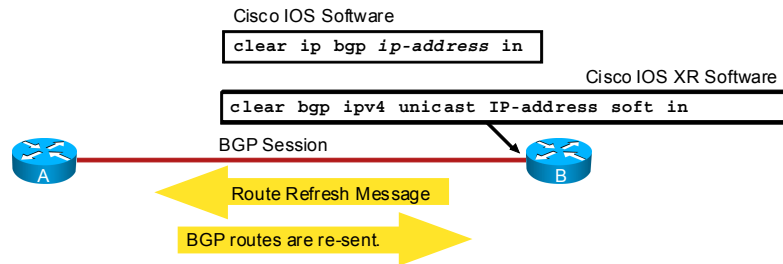
The only reason for making the extra copy is to be able to replay the data through the new routing policy without tearing down the session and re-establishing it.

What is needed is a mechanism to ask the neighbor router to issue a “clear soft outbound.” If this mechanism was 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.

Route Refresh (Cont.)

- Step 1: Route refresh is negotiated when the BGP session is established.
- Step 2: Inbound routing policy is changed on router B.
- Step 3: Operator requests inbound route refresh.
- Step 4: Router B sends route refresh message to router A.
- Step 5: Router A resends all BGP routes to router B.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-228

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, a network administrator can issue the **clear ip bgp ip-address soft in** (Cisco IOS Software) or **clear bgp ipv4 unicast ip-address soft in** (Cisco IOS XR Software) command in the local router. The router checks whether the route refresh capability is available, and if it is, requests resending of the neighbor's BGP table instead of replaying its own stored copy.

Tuning Routing Protocol Parameters

This topic describes how routing protocol parameters are tuned.

Tuning OSPF and IS-IS Parameters

- Goal of IGP tuning:
 - Improve **convergence** (failure detection and rerouting)
 - Improve **stability**
 - Improve **security** (authentication of adjacencies)
- IGP tuning scope:
 - Isolated to a single link Standard Change Minor Change Major Change
 - Area- and level-wide Major Change
 - Routing domain wide Major Change
- Regardless of scope, IGP tuning should be covered by the network design, and it should be tested for convergence and stability.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–230

Tuning routing protocols should preferably have positive results in terms of performance, security, and stability. Most changes that encompass more than one link or one adjacency are regarded as major changes. These changes should follow the proper path in which the change is thoroughly investigated, designed, and tested in a lab before being implemented networkwide.

Modification Results

- Desired:
 - Improvement in performance, security, and stability
- Undesired:
 - Temporary downtime (for example, IGP adjacency flap)
 - High CPU utilization
 - Occasional IGP adjacency flaps (for example, due to high CPU utilization and aggressive timers)

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9–231

Care should be taken when implementing changes since any changes in the routing protocol configuration can cause temporary downtime due to adjacencies being reestablished. In a much worse case, it may cause permanent instability due to overtuning.

Tuning Routing Protocol Parameters

Configuring OSPF or IS-IS timers:

- Hello and dead intervals
- LSA timers
- SPF timers

Configuring OSPF or IS-IS security:

- Authentication
- TTL security (OSPF only)

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-232

Several parameters can be used for routing protocol parameters:

- OSPF timers
- IS-IS timers
- OSPF authentication types
- IS-IS authentication
- Authentication for BGP

Hello and Dead Intervals

- Dead interval determines worst-case failure detection time:
 - OSPF default: 40 seconds
 - IS-IS default: 30 seconds
- Improving failure detection:
 - Tuning hello and dead intervals
 - Enabling fast hellos
 - Use of Bidirectional Forwarding Detection

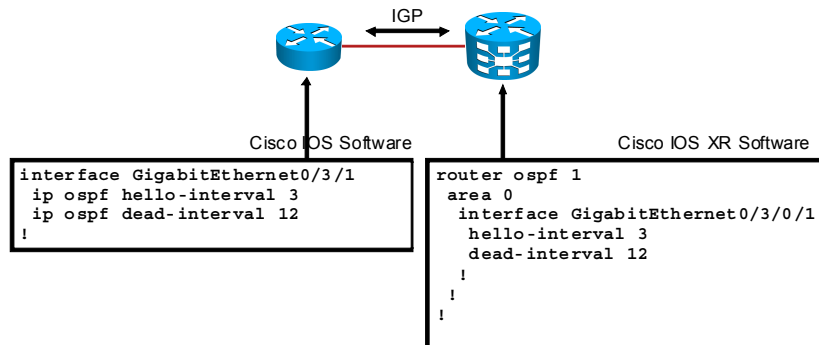
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0–233

OSPF and IS-IS have default timers that allow them to converge within approximately one minute. Various options exist that allow the two protocols to converge faster (even in less than one second).

OSPF Hello and Dead Intervals

- Changing hello and dead intervals can result in mismatched configuration between an adjacent pair of routers (lost adjacency):
 - Hello interval in seconds (default: 10)
 - Dead interval in seconds (default: 40)



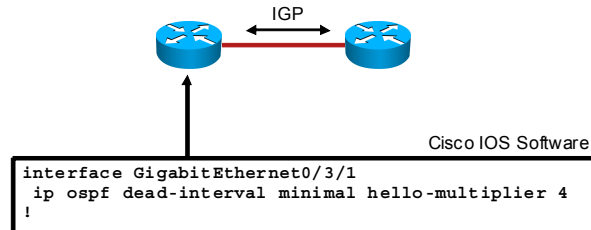
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0–234

The sample configurations show how to improve the convergence of OSPF using Cisco IOS Software or Cisco IOS XR Software.

OSPF Fast Hellos

- Setting dead interval to minimal results in as much as one second to detect failures.
- Hello multiplier determines how many hello messages are sent within one second.
- **Warning:** Can cause high CPU utilization and lost adjacencies on routers with many OSPF peers.



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-235

Fast hellos can be used with OSPF to enable OSPF to converge within one second. Setting the dead interval to “minimal” results in as much as one second to detect failures. The hello multiplier determines how many hello messages are sent within one second.

Configuring OSPF Timers

- Flood packet pacing timer in milliseconds (default: 33)
- Group packet pacing timer in seconds (default: 240)
- Impact of changing these parameters:
 - Does not affect established OSPF adjacencies
 - May have occasional negative effects in large environments if set to too aggressive values

Cisco IOS Software

```
router ospf 1
 timers pacing flood 10
 timers pacing lsa-group 60
!
```

Cisco IOS XR Software

```
router ospf 1
 timers pacing flood 10
 timers lsa group-pacing 60
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–236

In rare situations, you might need to change OSPF packet-pacing default timers to mitigate CPU or buffer utilization issues that are associated with flooding very large numbers of LSAs. The OSPF Update Packet-Pacing Configurable Timers feature allows you to configure the rate at which OSPF LSA flood pacing, retransmission pacing, and group pacing updates occur.

Configuring OSPF flood pacing timers allows you to control interpacket spacing between consecutive link-state update packets in the OSPF transmission queue. Configuring OSPF retransmission pacing timers allows you to control interpacket spacing between consecutive link-state update packets in the OSPF retransmission queue. Cisco IOS Software groups the periodic refresh of LSAs to improve the LSA packing density for refreshes in large topologies. The group timer controls the interval that is used for group LSA refreshment. However, this timer does not change the frequency at which individual LSAs are refreshed (the default refresh occurs every 30 minutes).

Note The default settings for OSPF packet-pacing timers are suitable for the majority of OSPF deployments. You should change the default timers only as a last resort.

The OSPF Update Packet-Pacing Configurable Timers feature provides the administrator with a mechanism to control the rate at which LSA updates occur in order to reduce high CPU or buffer utilization that can occur when an area is flooded with a very large number of LSAs.

Do not change the packet pacing timers unless all other options to meet OSPF packet flooding requirements have been exhausted. Specifically, network operators should prefer summarization, stub area usage, queue tuning, and buffer tuning before changing the default timers. Furthermore, there are no guidelines for changing timer values; each OSPF deployment is unique and should be considered on a case-by-case basis. The network operator assumes risks that are associated with changing the default timer values.

Example: Configuring OSPF Timers

Cisco IOS XR Software

- How often self-originated LSAs should be refreshed:
- Limits the frequency that new processes of any particular OSPF Version 2 LSA can be accepted during flooding:
- Changes the interval at which OSPF link-state LSAs are collected into a group for flooding:

```
router ospf 1
timers lsa refresh seconds
```

```
router ospf 1
timers lsa min-arrival seconds
```

```
router ospf 1
timers lsa group-pacing seconds
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-237

The figure shows OSPF timers.

- The **timers lsa refresh** sets how often self-originated LSAs should be refreshed, in seconds. The default is 1800 seconds for both OSPF and OSPFv3.
- The **timers lsa min-arrival** limits the frequency that new processes of any particular OSPF Version 2 LSA can be accepted during flooding. The default is one second.
- The **timers lsa group-pacing** changes the interval at which OSPF link-state LSAs are collected into a group for flooding. The default is 240 seconds.

OSPF Shortest Path First Throttling

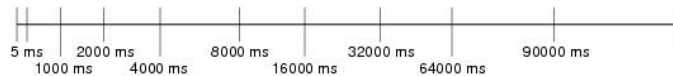
- SPF scheduling and delay SPF calculations during network instability.
- SPF is scheduled to calculate the SPT when there is a change in topology.
- One SPF run may include multiple topology change events.
- The interval is chosen dynamically.
- Within the boundary.

Cisco IOS Software

```
router ospf 1
 timers throttle spf 5 1000 90000
!
```

Cisco IOS XR Software

```
router ospf 1
 timers throttle spf 5 1000 90000
!
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-238

The OSPF Throttling feature makes it possible to configure SPF scheduling in millisecond intervals and to potentially delay SPF calculations during network instability. SPF is scheduled to calculate the shortest path tree (SPT) when there is a change in topology. One SPF run may include multiple topology change events.

The interval at which the SPF calculations occur is chosen dynamically and is based on the frequency of topology changes in the network. The chosen interval is within the boundary of the user-specified value ranges. If network topology is unstable, SPF throttling calculates SPF scheduling intervals to be longer until topology becomes stable.

SPF calculations occur at an interval that is set by the **timers throttle spf** command. The wait interval indicates the amount of time to wait until the next SPF calculation occurs. Each wait interval after that calculation is twice as long as the previous one until the wait interval reaches the maximum wait time specified.

The SPF timing can be better explained using an example. In this example, the start interval is set at 5 ms, the wait interval at 1000 ms, and the maximum wait time is set at 90,000 ms.

OSPF LSA Throttling

- Dynamic mechanism to slow down LSA updates in OSPF during times of network instability
- Allows faster OSPF convergence by providing LSA rate limiting
- Enabled by default and allows faster OSPF convergence
- Controls the generation of LSAs:
- Controls the minimum interval for accepting the same LSA:

```
router ospf 1
timers throttle lsa all
!
```

```
router ospf 1
timers lsa arrival
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9-239

The OSPF LSA Throttling feature provides a dynamic mechanism to slow down LSA updates in OSPF during times of network instability. It also allows faster OSPF convergence by providing LSA rate limiting in ms.

Before the OSPF LSA Throttling feature was available, LSA generation was rate-limited for five seconds. That meant that changes in an LSA could not be propagated in milliseconds, so the OSPF network could not achieve millisecond convergence.

The OSPF LSA Throttling feature is enabled by default and allows faster OSPF convergence (in milliseconds). This feature can be customized. One command controls the generation of LSAs and another command controls the receiving interval. This feature also provides a dynamic mechanism to slow down the frequency of LSA updates in OSPF during times of network instability.

The **timers throttle lsa all** command controls the generation of LSAs. The first LSA is always generated immediately upon an OSPF topology change, and the next LSA generated is controlled by the minimum start interval. The subsequent LSAs generated for the same LSA are rate-limited until the maximum interval is reached. The "same LSA" is defined as an LSA instance that contains the same LSA ID number, LSA type, and advertising router ID.

The **timers lsa arrival** command controls the minimum interval for accepting the same LSA. If an instance of the same LSA arrives sooner than the interval that is set, the LSA is dropped. It is recommended that the arrival interval be less than or equal to the holdtime interval of the **timers throttle lsa all** command.

Configuring IS-IS Timers

- Customizes IS-IS throttling of SPF calculations

```
router isis 1
spf-interval [level-1 | level-2] spf-max-wait [spf-initial-wait
spf-second-wait]
```

- Customizes IS-IS throttling of PRC calculations

```
router isis 1
prc-interval prc-max-wait [prc-initial-wait prc-second-wait]
```

- Sets the minimum interval at which link-state PDUs (LSPs) are generated

```
router isis 1
lsp-gen-interval [level-1 | level-2] lsp-max-wait [lsp-
initial-wait lsp-second-wait]
```

- Sets the maximum time for which link-state PDUs (LSPs) persist without being refreshed

```
router isis 1
max-lsp-lifetime [hours] value
```

- Sets the minimum interval at which LSPs are refreshed

```
router isis 1
lsp-refresh-interval seconds
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-240

SPF, PRC, and LSP generation timers need to be tuned according to the level of stability of the network and stability that is required in the routing domain. For instance, setting low values will trigger a fast convergence with a potential risk of high resource utilization if there are flapping routes that cause network turnover. Setting high values will keep the network stable with slower convergence.

It is recommended to leave the default value for the LSP generation interval at 5 seconds and also to increase the maximum lifetime for LSPs to 65,535 seconds, in order to conserve CPU usage for generating and refreshing LSPs.

- The **spf-interval** command customizes IS-IS throttling of SPF calculations. The recommended values for the **spf-max-wait**, **spf-initial-wait**, and **spf-second-wait** arguments are 5, 1, and 20, respectively.
- The **prc-interval** command customizes IS-IS throttling of PRC calculations. The recommended values for the **prc-max-wait**, **prc-initial-wait**, and **prc-second-wait** arguments are 5, 1, and 20, respectively.
- The **lsp-gen-interval** command sets the minimum interval at which link-state PDUs (LSPs) are generated. The recommended values for the **lsp-max-wait**, **lsp-initial-wait**, and **lsp-second-wait** arguments are 5, 1, and 20, respectively.
- The **max-lsp-lifetime** command sets the maximum time for which link-state PDUs (LSPs) persist without being refreshed. To reduce network resources that are used for LSP generation, increase the LSP maximum lifetime value of 65,535.
- The **lsp-refresh-interval** command sets the minimum interval at which LSPs are refreshed. To reduce network resources that are used for LSP refresh, increase the value to the LSP refresh interval to the maximum value of 65,000 seconds.

OSPF Authentication Types

- OSPF supports two types of authentication:
 - Simple password (or plain text) authentication
 - MD5 authentication
- Router generates and checks every OSPF packet. Router authenticates the source of each routing update packet that it receives.
- Configure a “key” (password); each participating neighbor must have the same key configured.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–241

This figure describes two types of authentication that are used in OSPF:

- Simple password or plain-text authentication
- MD5 authentication

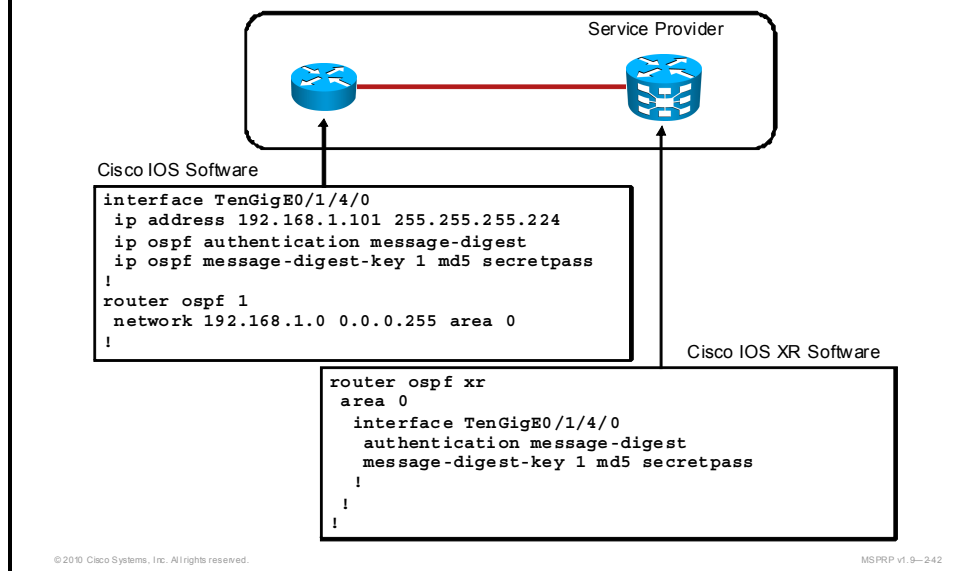
OSPF neighbor authentication (also called neighbor router authentication or route authentication) can be configured such that routers can participate in routing based on predefined passwords.

Recall that when neighbor authentication has been configured on a router, the router authenticates the source of each routing update packet that it receives. This authentication is accomplished by the exchange of an authenticating key (sometimes referred to as a password) that is known to both the sending and the receiving router.

By default, OSPF uses null authentication, which means that routing exchanges over a network are not authenticated. OSPF supports two other authentication methods: simple password authentication (also called plain-text authentication) and MD5 authentication.

OSPF MD5 authentication includes a nondecreasing sequence number in each OSPF packet to protect against replay attacks.

Example: MD5 Authentication Configuration



The figure shows a network that is illustrating the configuration of MD5 authentication. The configuration of the right router is shown in this figure.

MD5 authentication is configured on interface TenGigabit Ethernet 0/1/4/0 with the **ip ospf authentication message-digest** (Cisco IOS Software) command. The interface is configured with an authentication key number 1 set to “secretpass.”

The figure also shows the configuration on the Cisco IOS XR Software.

IS-IS Authentication

- IS-IS authentication can cause disruptions:
 - First-time deployment
 - Modification of authentication password
- Classic deployment or password change should be performed during maintenance window.
- IS-IS supports **smooth transition**:
 - Step 1: Enable send-only mode during deployment/modification
 - Step 2: Configure authentication modes and passwords
 - Step 3: Disable send-only mode to activate authentication

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9–243

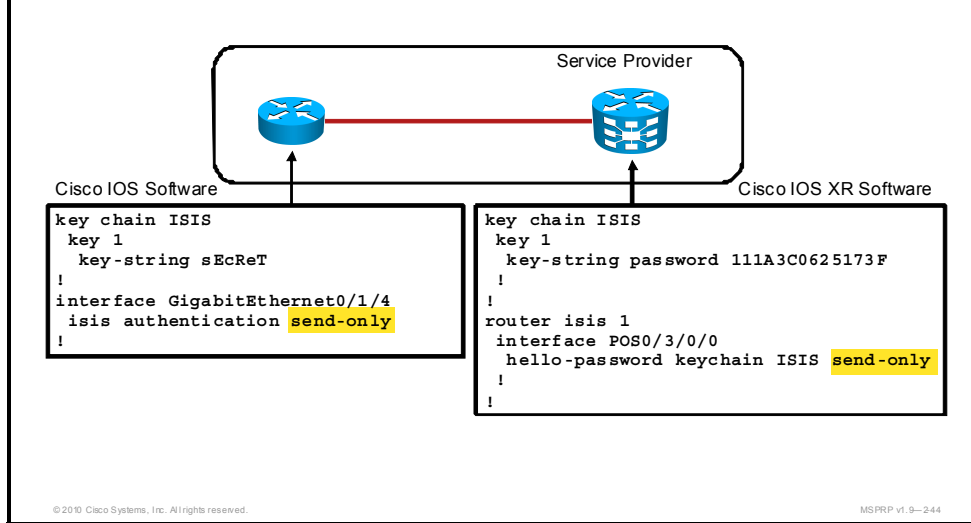
IS-IS allows for the configuration of a password for a specified link, an area, or a domain. Routers that want to become neighbors must exchange the same password for their configured level of authentication. A router not in possession of the appropriate password is prohibited from participating in the corresponding function. That is, it may not initialize a link, be a member of an area, or be a member of a Level 2 domain, respectively.

Cisco IOS Software allows three types of IS-IS authentication to be configured.

- **IS-IS authentication:** For a long time, this method was the only way to configure authentication for IS-IS.
- **IS-IS Hashed Message Authentication Code-Message Digest 5 (HMAC-MD5) authentication:** This feature adds an HMAC-MD5 digest to each IS-IS protocol data unit (PDU). It was introduced in Cisco IOS Software Release 12.2(13)T and is only supported on a limited number of platforms.
- **Enhanced Cleartext Authentication:** With this new feature, cleartext authentication can be configured using new commands that allow passwords to be encrypted when the software configuration is displayed. It also makes passwords easier to manage and change.

IS-IS Authentication with Smooth Transition

Step 1



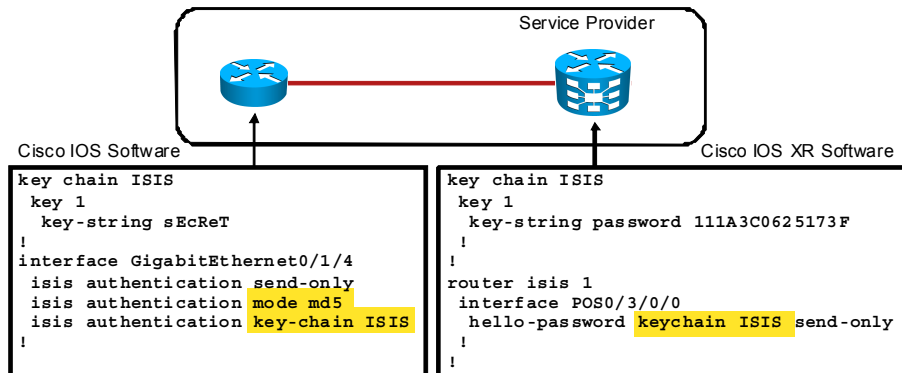
The figure explains how to configure authentication for IS-IS in Cisco IOS Software and Cisco IOS XR Software.

Authentication is available to limit the establishment of adjacencies via the **hello-password** configuration and to limit the exchange of LSPs via the LSP password.

The implementation is split into three steps to ensure minimum interruption to the network. In the first step, you configure authentication using the **send-only** keyword, which adds authentication information to the LSAs without routers checking authentication upon reception of LSAs.

IS-IS Authentication with Smooth Transition

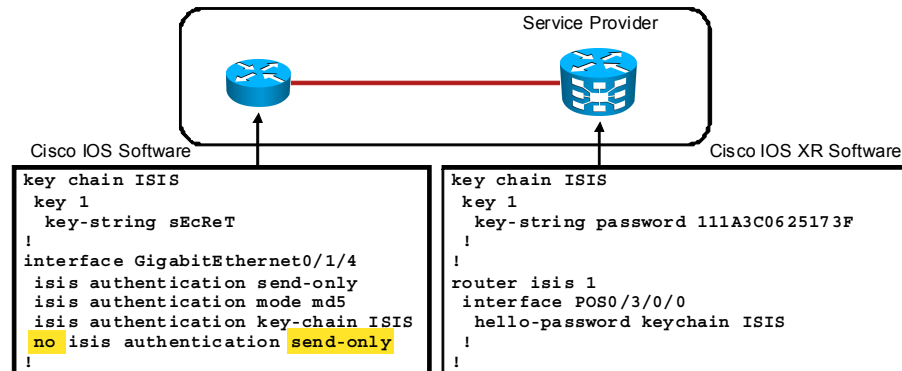
Step 2



Passwords are configured in the second step when adding authentication information to LSAs.

IS-IS Authentication with Smooth Transition

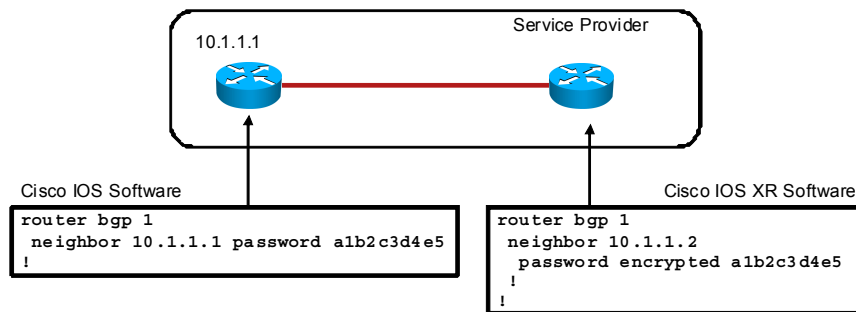
Step 3



When all routers have been configured with passwords, you can start removing the **send-only** keyword to fully activate authentication.

Authentication for BGP

- Between two BGP peers.
- Each TCP segment of the BGP session is authenticated and its integrity is checked.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.9-247

You can configure MD5 authentication between two BGP peers, meaning that each segment sent on the TCP connection between the peers is verified. MD5 authentication must be configured with the same password on both BGP peers; otherwise, the connection between them will not be made. Configuring MD5 authentication causes the Cisco IOS Software to generate and check the MD5 digest of every segment that is sent on the TCP connection. If authentication is invoked and a segment fails authentication, then an error message will display in the console.

The example enables the authentication feature between this router and the BGP neighbor at 10.1.1.1. The password that must also be configured for the neighbor is a1b2c3d4e5. The remote peer must be configured before the hold-down timer expires.

Summary

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

Summary

- Modifications to IGPs:
 - Test in a lab.
 - Evaluate for risks and impact on running network.
- Perform disruptive changes during maintenance windows:
 - Change authentication passwords.
 - Change timers.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.9—248

Lesson Self-Check

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

- Q1) What is a default OSPF reference bandwidth? (Source: Modifying Routing Parameters)
- A) 10 Mb/s
 - B) 100 Mb/s
 - C) 1 Gb/s
 - D) 10 Gb/s
- Q2) Which two options will change OSPF router ID? (Choose two.) (Source: Modifying Routing Parameters)
- A) configuring loopback IP address and resetting OSPF process
 - B) configuring router ID using **router-id** command
 - C) configuring loopback IP address
 - D) configuring router ID using **router-id** command and resetting OSPF process
- Q3) What are two alternatives for injecting local routes into the BGP table? (Choose two.) (Source: Adding New Prefixes)
- A) list them using the **network** command in BGP configuration mode
 - B) redistribute them into BGP
 - C) redistribute BGP routes into IGP (OSPF or IS-IS)
 - D) list them using the **network** command in OSPF configuration mode
- Q4) Which two types of authentication supports OSPF? (Choose two.) (Source: Tuning Routing Protocol Parameters)
- A) CHAP authentication
 - B) MD5 authentication
 - C) simple password authentication
 - D) RADIUS authentication
- Q5) Which two are benefits of BGP aggregation? (Choose two.) (Source: Adding New Prefixes)
- A) more stable internetworks because of less route flapping
 - B) bigger routing tables
 - C) smaller routing tables
 - D) more route flapping

Lesson Self-Check Answer Key

- Q1) B
- Q2) A, D
- Q3) A, B
- Q4) B, C
- Q5) A, C

Connecting Single-Homed Customers to the Network

Overview

There are multiple ways to connect customer networks to the Internet. When connecting single-homed customer networks to the ISP network, the customer uses either a single connection to the ISP or multiple connections to the same ISP. In this case, Border Gateway Protocol (BGP) is generally not used, because static routing is the simplest way to implement connectivity between the customer and the ISP. This lesson discusses solutions for connecting single-homed customer networks to ISPs with static routing and its redistribution into internal BGP on service provider edge routers. Additionally the lesson discusses tasks to perform on the customer edge router.

Objectives

Upon completing this lesson, you will be able to identify the routing and addressing requirements when adding a new single-homed customer to the service provider network. This ability includes being able to meet these objectives:

- Describe the typical routing requirements when adding a new single-homed customer to the service provider network
- Provide addressing for the access link of a new single-homed customer
- Provide addressing for a single-homed customer
- Ensure reachability for a single-homed customer

Single-Homed Customer Requirements

This topic describes the typical routing requirements when adding a new single-homed customer.

Single-Homed Customer Requirements

- Identify customer **connectivity requirements**:
 - Outbound sessions only
 - Bidirectional sessions
- **Public addressing requirements**:
 - Single IP address (for example, for outbound sessions only)
 - Multiple IP addresses (that is, provider-assigned subnet required)
- **Availability requirements**:
 - No special requirements
 - Simple backup required (for example, dial, DSL)
- Use designed and tested templates to connect single-homed customers. Standard Change

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.3

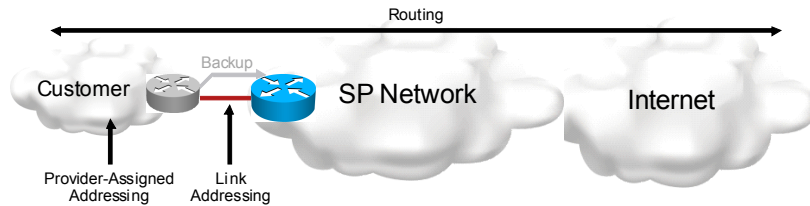
When adding a new single-homed customer, the service provider should first identify customer connectivity, addressing, and availability requirements. Based on those requirements, appropriate actions are taken.

When considering connectivity requirements, the service provider should identify whether the customer needs connectivity only for outbound sessions or for bidirectional sessions. For outbound sessions, only a single IP public address for the customer is needed while for bidirectional sessions, multiple public IP addresses may be needed. The service provider should also be aware of the customer's availability requirements.

When connecting a new customer, use the connectivity profile that the customer selected. It should be one of the available designed and tested solutions for which there are implementation templates and procedures that you can follow to successfully connect a new single-homed customer. Although deploying access for a new single-homed customer is regarded as a standard change, consider the customer's service level agreement (SLA) when modifying the configuration for an existing customer. A change in connectivity (for example, migrating the customer's link to a new router, changing the connectivity type) may cause downtime to the customer. Given that the customer has a tight SLA, make changes during a regular maintenance window that is acceptable to the customer.

Single-Homed Customer Requirements (Cont.)

- Addressing requirements:
 - Link addressing
 - Provider-assigned public IP subnet
- Routing requirements:
 - Static routing (typical)
 - BGP can be used to detect failures and reroute onto backup link (such as dial, DSL); private AS number can be assigned by ISP.



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.4

Customers that are connected to a single ISP usually get their address space assigned by the ISP. An ISP is usually assigned a large address space to delegate to its customers. Because all customers of one ISP obtain their addresses from one address space or a few address spaces, the ISP can probably aggregate the customer addresses before sending them to the rest of the Internet.

Customers that are connected to a single ISP are using provider-assigned addresses. If the customer changes its service provider, the customer must return its provider-assigned addresses to the old ISP and receive a new assignment of provider-assigned addresses from the new ISP. Otherwise, the ISPs are no longer able to perform efficient address aggregation.

In the single-homed environment, static routing is usually adequate because of its lower complexity. Normally, the customer network must have a route to the IP prefixes that the customer has in its network.

If there is a backup link from the customer to the service provider, the customer requires a routing protocol such as BGP to detect the failures. Because BGP uses handshaking and reliable transfer, it always detects a failed link or a failed remote router.

Customers should use public autonomous system (AS) numbers only when they are required. A customer that uses BGP to exchange routing information with only one ISP does not require a public AS number. This customer can use a private AS number.

An ISP network that is running BGP with some of its customers must determine whether a public or a private AS number is required for each customer. When the customer can use a private AS number, the ISP must allocate one from the range of private AS numbers (64512 to 65535). The ISP must make sure not to assign any of the private AS numbers to more than one customer.

Addressing and Routing Requirements

- **Only link addressing:**
 - Customer using Port Address Translation (PAT)
 - Simplified routing
 - Requires public addressing on access link
 - Limited support for inbound sessions (static PAT)
- **Link addressing and a PA subnet:**
 - Routing for PA subnet
 - May use private addressing on access link
 - Supports servers at customer premises for inbound sessions

© 2010 Cisco Systems, Inc. All rights reserved.

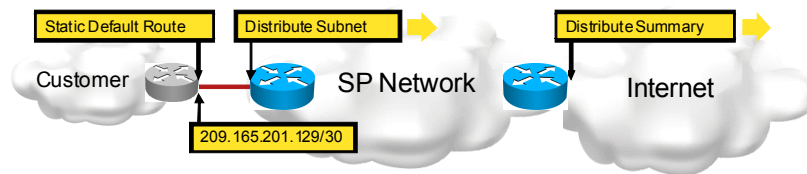
MSRP v1.0—2-5

If the service provider uses only link addressing, then the customer uses Port Address Translation (PAT). Using PAT means that private customer IP addresses are translated into the public IP address of the customer's edge interface pointing to the service provider. Routing is simplified; only a static default route on the customer's edge router is needed. The service provider should distribute the appropriate subnet. When using only the link addressing option, the public IP address on the access link is required. There is limited support for inbound sessions; static PAT should be used.

If the service provider uses link addressing and a provider-assigned subnet, it should provide routing for the provider-assigned subnet. A static route for the provider-assigned subnet is needed on the provider edge router. Another difference with only link addressing is that a private IP address on the access link can be used. This second option of addressing also supports servers at customer premises for inbound sessions.

Only Link Addressing

- **Addressing:**
 - Access link
 - Public address required (such as /30 or /31 subnet)
 - Dynamic PAT for outbound sessions
 - Optionally, static PAT for inbound sessions to servers
- **Routing options:**
 - Include access link subnet in IGP
 - Include access link subnet in BGP
- **Note!** Do not run IGP with customer.



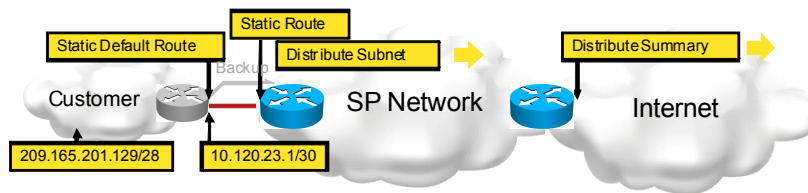
The figure shows an example of only link addressing. The service provider gives the public address only to the access link, normally with a mask of /30. The customer uses private IP addresses in its own network and uses dynamic PAT for outbound sessions. Optionally, for inbound sessions to publicly available servers, static PAT is used.

A static default route on the customer's edge router is needed, and the service provider should distribute the access link subnet. There are two options for doing that. The service provider could include the access link subnet in the interior gateway protocol (IGP), (Open Shortest Path First [OSPF] or Intermediate System-to-Intermediate System [IS-IS]), or BGP. It is very important that the service provider does not run IGP with the customer.

ISP should aggregate the customer addresses before sending the routes to the rest of the Internet.

Provider-Assigned Address Space

- **Addressing:**
 - Access link can use private addresses.
 - Public subnet assigned to customer from provider's pool.
- **Access link routing options:**
 - Include access link subnet in IGP (native, redistribution).
 - Include access link subnet in BGP (native, redistribution).
 - **Note!** Do not propagate private addresses to other autonomous systems (for example, tag with no-export community).
- **PA subnet routing:**
 - Use static route on edge router.
 - Distribute subnet using BGP.
 - Summarize PA subnets when sending to other ISPs.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2.7

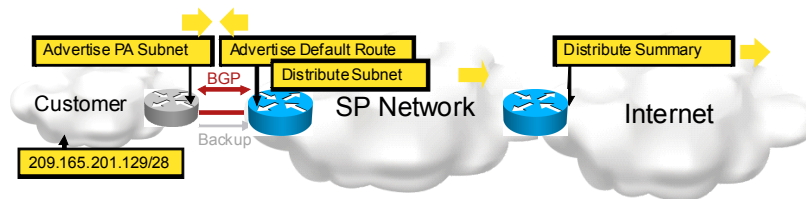
The figure shows an example of a PA address space. The service provider assigns an IP address, which could be private, to the access link. The customer receives a public subnet from the provider's pool. PAT could be still used if there are not enough public IP addresses for the customer network.

The service provider could include the access link subnet in IGP (OSPF or IS-IS) or BGP. In both cases, it could be done using native route propagation or the redistribution of connected routes. The service provider should not propagate private addresses to other autonomous systems.

The static default route on the customer edge router should be configured and the service provider should propagate the provider-assigned subnet to the rest of the Internet. A static route on the provider edge router to the customer's provider-assigned subnet is configured. Usually this static route is redistributed into BGP, ensuring that all other ISP routers and the rest of the Internet will receive information about the provider-assigned subnet of the customer. ISP should summarize the provider-assigned subnets before sending the routes to the rest of the Internet.

BGP for Backup Link

- **Addressing:** Access link only or PA subnet
- **Routing options:**
 - BGP between service provider and customer
 - Originate default route to customer
 - Only accept assigned subnet
 - Include access link subnet into IGP (next-hop reachability)
 - Backup routing can be implemented using RADIUS-assigned routes



The figure shows an example in which the customer has a backup connection to the ISP. If there is a backup link from the customer to the ISP, the customer requires a routing protocol such as BGP to detect the failures. If a primary link goes down, BGP will detect the failure and redirect traffic to the backup link.

Advertising the default route to the customer is recommended. Include the access link subnet into IGP (OSPF or IS-IS) to solve issues with next-hop reachability. ISP should filter routing updates that are received from the customer accepting only those updates from an assigned subnet. One option for implementing backup routing is to use RADIUS-assigned routes.

Customer Access Link

This topic describes addressing for the access link of a new single-homed customer.

Customer Access Link

- **Addressing:**
 - Private
 - Public
- **Core routing** for access link subnet:
 - BGP
 - IGP
- Changing access link addressing for an existing customer will cause temporary downtime for the customer:
 - Consult the customer SLA
 - Coordinate actions with the customer
 - Deploy changes to meet customer SLA

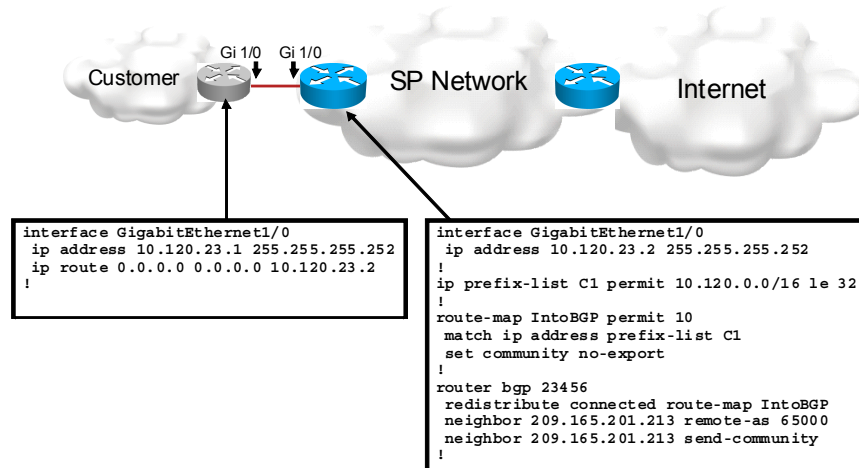
© 2010 Cisco Systems, Inc. All rights reserved. MSPRP v1.0–2.10

Different examples of propagating a customer access link subnet into the service provider network will be described:

- Redistributing a connected route using the no-export community
- Redistributing a connected route and aggregating the customer's addresses before sending them to other ISPs
- Using interior gateway protocols (OSPF and IS-IS)

Example: Customer Access Link

Private Addressing and BGP—Cisco IOS Software

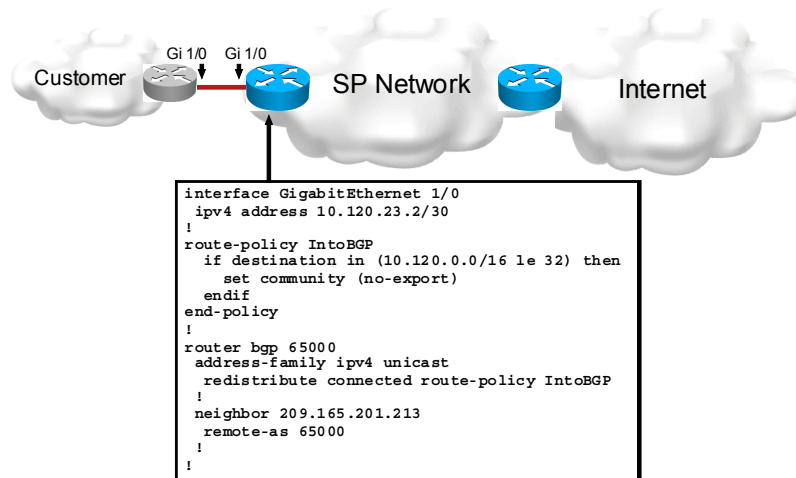


The example shows a solution for Cisco IOS Software in which private addressing on a customer access link is used and directly connected routes are redistributed into BGP on the ISP edge router. This connected route is marked with a predefined community called **no-export**. If a router in the service provider network receives an update carrying this community, it will not propagate it to any external BGP neighbors. This means that the update is not propagated to other ISPs. To ensure that the BGP communities are propagated, the network operators of the ISP network must configure a **send-community** option for all IBGP neighbors. The edge routers of the ISP network then see the no-export community and filter out those routes before sending the update to External Border Gateway Protocol (EBGP) neighbors. Communities are set using route maps. A route map can select routes that are based on various attributes.

A static default route is configured on the customer edge router, and a connected route is redistributed into BGP on the ISP edge router. An IBGP neighbor relationship is established between the service provider edge router and the router that has connection to other ISPs.

Example: Customer Access Link

Private Addressing and BGP—Cisco IOS XR Software



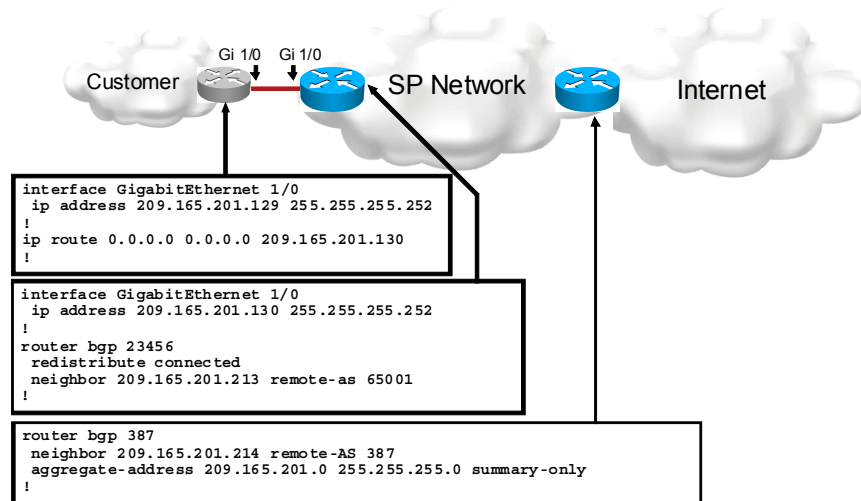
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-3.12

The example shows a solution for Cisco IOS XR Software in which private addressing on the customer access link is used, and directly connected routes are redistributed into BGP on the service provider edge router. Again, a static default route is configured on the customer edge router, and the connected route is redistributed into BGP on the service provider edge router. BGP community and extended-community filtering cannot be configured for Internal Border Gateway Protocol (IBGP) neighbors in Cisco IOS XR Software. Communities and extended communities are always sent to IBGP neighbors.

Example: Customer Access Link

Public Addressing and BGP—Cisco IOS Software

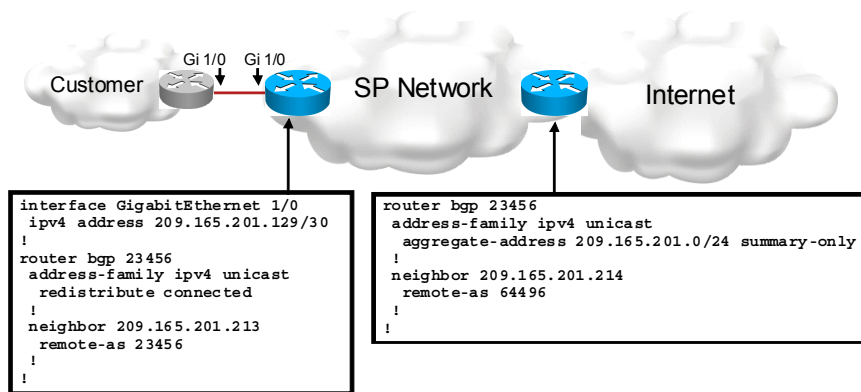


The sample configuration shows where public addressing for the customer access link is used. A static default route is configured on the customer edge router, and the connected route is redistributed into BGP on the ISP edge router. An IBGP neighbor relationship is established between the ISP edge router and the router that has connection to other ISPs.

Because ISP should aggregate the customer addresses before sending the routes to the rest of the Internet, the **aggregate-address** command with **summary-only** option is used. Before BGP advertises these customer routes to the rest of the Internet, an aggregation of the subnets into a larger announcement would be appropriate. When the **summary-only** option is used, only the route summary will be advertised, not the more specific routes.

Example: Customer Access Link

Public Addressing and BGP—Cisco IOS XR Software



© 2010 Cisco Systems, Inc. All rights reserved.

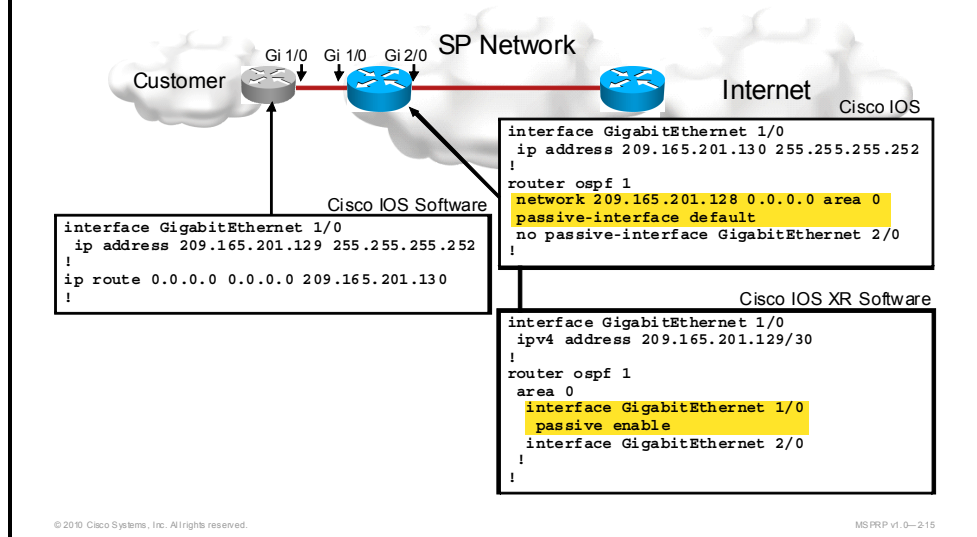
MSRP v1.0—214

The figure shows configuration for Cisco IOS XR Software in which public addressing for the customer access link is used. A static default route is configured on the customer edge router, and the connected route is redistributed into BGP on the ISP edge router. An IBGP neighbor relationship is established between the ISP edge router and the router that has connection to other ISPs.

ISP should aggregate the customer addresses before sending the routes to the rest of the Internet.

Example: Customer Access Link

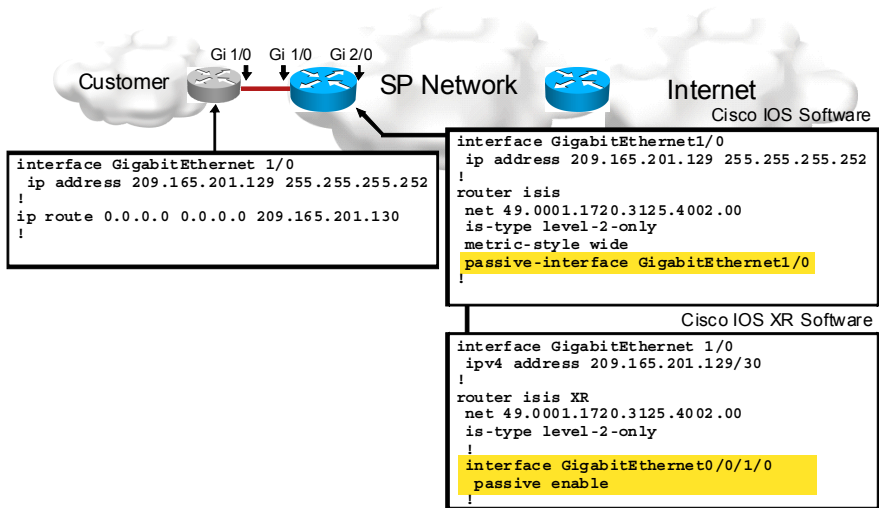
OSPF



Another option for propagating the customer access link subnet into the ISP network is using an interior gateway protocol, for instance OSPF, as shown in the configuration for Cisco IOS Software. The customer access link subnet is announced into the service provider network using the **network** command under the OSPF process. The service provider should run OSPF only on interfaces pointing to other routers inside the service provider network and not on interfaces pointing to the customers. The easiest way to accomplish this is to use the **passive-interface default** command, which disables the sending of routing updates on all interfaces. Then the sending of routing updates on interfaces pointing to other routers in the service provider network is enabled using the **no passive-interface** command.

Example: Customer Access Link

IS-IS



Similarly to OSPF, IS-IS can be enabled in the core and the customer access links are passively added to the database. The sample configuration shows how to passively add the links to IS-IS in Cisco IOS and to the Cisco IOS XR command-line interface.

Customer Addressing and Routing

This topic describes addressing for a single-homed customer.

Customer Addressing and Routing

- **Addressing:**
 - Public provider-assigned addresses
- **Routing:**
 - Static
 - BGP
- Changing addressing, routing, or filtering for an existing customer may cause temporary down time for the customer:
 - Consult the customer SLA
 - Coordinate actions with the customer
 - Deploy changes to meet customer's SLA
 - Use migration procedures that cause no or minimum disruption to customer traffic

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0—218

Customers that are connected to a single ISP usually receive provider-assigned address space from the ISP. The customer has to renumber its network when it changes its ISP.

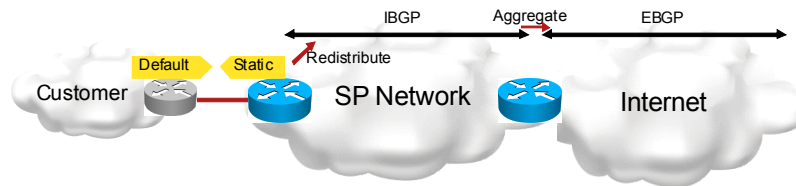
Two examples will be discussed:

- Static routing for a single-homed customer with provider-assigned addresses
- BGP routing for a single-homed customer that uses a backup link to the ISP

Customer Addressing and Routing

Static Routing

- Default static route from customer to service provider
- Static route from service provider to the customer
- Redistribution of static route into IBGP
- Aggregation of customer routes before sending them to other service providers



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.19

As the figure shows, the customer network is connected to the Internet by using a single permanent connection to a single ISP. A routing protocol does not add any redundancy and would only add complexity, which is why the static route from the ISP to the customer is used. This static route should be redistributed into the IBGP of the ISP. It is recommended that customer routes are aggregated before sending them to other ISPs.

Customer Addressing and Routing

Static Routing (Cont.)

- Identify all possible combinations of services offered to a customer, including QoS services.
- Assign a tag to each combination of services.
- Configure a route map that matches defined tags and sets BGP communities or other BGP attributes.
- Redistribute static routes into BGP through a route map.
- For each customer, configure a static route toward the customer with the proper tag.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-220

You can easily extend the principle of using tags when you configure static routes, and of assigning different communities that are based on those tags, to implement a more complex routing policy. To propagate static routes in an ISP network, complete these steps:

- Step 1** Identify all the different SLAs that are offered to customers and then all the different combinations of these SLAs.
- Step 2** Assign each combination its own tag value and its own community.
- Step 3** Configure a route map, which selects routes with each of the assigned tags and sets the corresponding community value. Because the processing of a route map stops when the match clause of a statement is met, each route should be assigned a single combination of communities only. Therefore, take great care to assign a tag and a community combination to each combination of services that is provided.

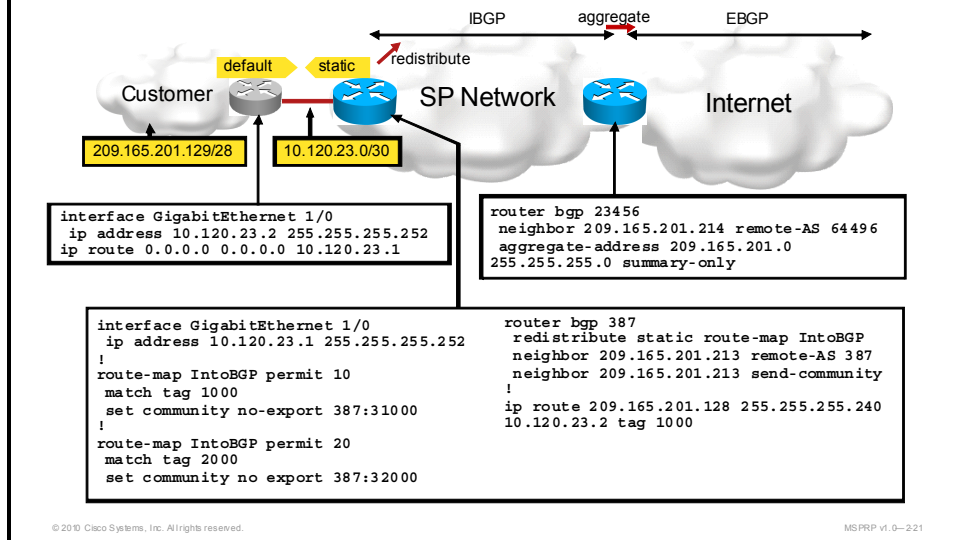
Note When the provider edge routers redistribute static routes into BGP, these routes must pass through the route map. BGP assigns the correct community depending on the tag values that are given on the configuration line for each of the static routes.

- Step 4** Finally, configure static routes. Before you configure a static route for a specific customer, identify the combination of the services that are provided to this customer. Then look up the corresponding tag value. After you have configured the route, assign the tag.

With this routing policy, every static route to a customer network is assigned a tag, and the redistributed BGP route is assigned a corresponding community. The BGP communities that are attached to the routes signal to other routers in the ISP network the particular service combination you should use.

Example: Customer Addressing and Routing

Static Routing Configuration—Cisco IOS Software



The figure shows a Cisco IOS configuration in which the ISP offers two different SLAs to its customers. Each service is tagged with its own tag value and its own community combination. Route maps are configured in the ISP edge router that has the static routes to the customer network. Redistribution of the configured customer static routes into BGP is also performed at the ISP edge router.

Because the route map has no “permit any” statement at the end, the static routes that are not assigned any of the tags being used are not redistributed. The route map filters out these routes, forcing the network operators to make a tag assignment to all customer routes.

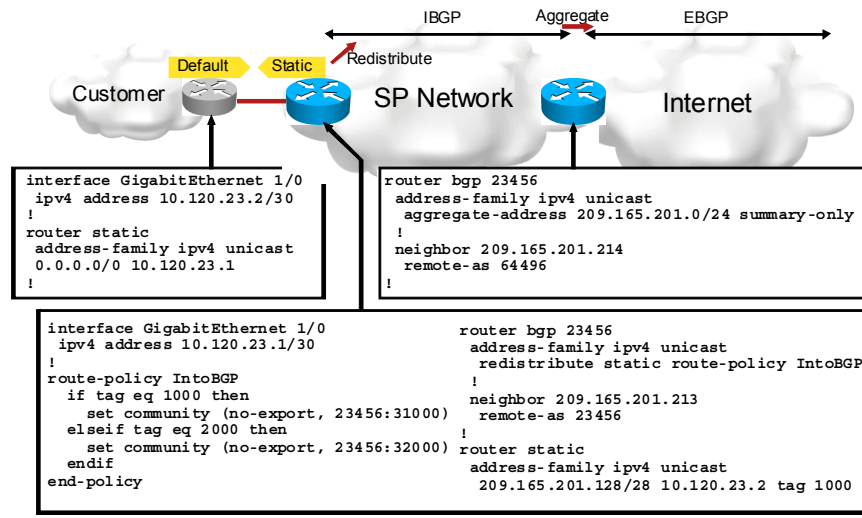
The figure shows how the ISP edge router uses the route map named “IntoBGP” when redistributing static routes into BGP. Because the route map assigns community values that other routers will use within the ISP network, network operators must configure all IBGP neighbors with the **send-community** qualifier.

When connecting customers, the network operators identify which service combination to use for this particular customer. The customers who have a static route that is tagged with tag 1000 receive one service level. The customers who have static routes that are tagged with tag 2000 receive another service level. Because provider-assigned addressing is used, **no-export** community is set, meaning that provider-assigned blocks are not propagated to upstream ISPs.

A static route to the customer is configured and assigned the appropriate tag value of 1000, which represents the specified services that are assigned to the customer. ISP aggregates the customer addresses before sending the routes to the rest of the Internet.

Example: Customer Addressing and Routing

Static Routing Configuration—Cisco IOS XR Software

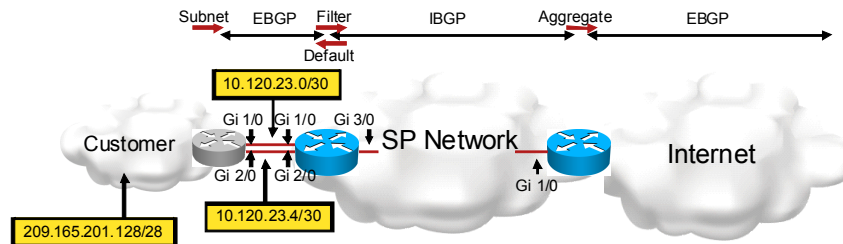


This figure shows a Cisco IOS XR configuration in which the ISP offers two different service levels to its customers. All statements that were described in the previous example are also true for the example shown in this figure.

Example: Customer Addressing and Routing

Backup Link with BGP

- BGP between customer and service provider.
- In case of failure on primary link, traffic is rerouted to the backup link.



© 2010 Cisco Systems, Inc. All rights reserved.

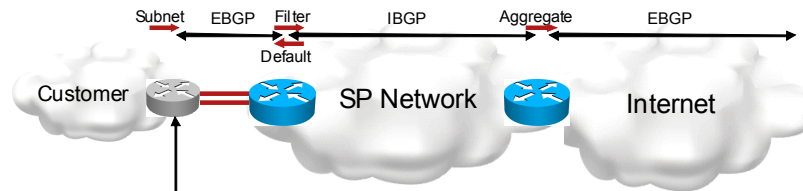
MSPRP v1.0-223

The figure shows an example of where the backup link between the customer and the ISP is used. If there is a failure on the primary link, traffic is rerouted to the backup link. The EBGP between the customer and the ISP edge router is responsible for that link. Among the routers in the ISP network, IBGP is established.

The customer's provider-assigned subnet is advertised through BGP. ISP is sending a default route to the customer and filtering incoming routing updates, allowing only those routes from the assigned subnet to the customer. Routes should be aggregated before sending them to other ISPs.

Example: Customer Addressing and Routing

Backup Link with BGP—Cisco IOS Software



```

interface GigabitEthernet1/0
 ip address 10.120.23.1 255.255.255.252
!
interface GigabitEthernet2/0
 ip address 10.120.23.5 255.255.255.252
!
ip prefix-list Default permit 0.0.0.0/0
!
route-map MED50 permit 10
 set metric 50
!
route-map MED10 permit 10
 set metric 10
!
route-map LocalPref permit 10
 set local-preference 200
!

router bgp 65001
 network 209.165.201.128 mask 255.255.255.240
!
neighbor 10.120.23.2 remote-as 23456
neighbor 10.120.23.2 prefix-list Default in
neighbor 10.120.23.2 route-map LocalPref in
neighbor 10.120.23.2 route-map MED10 out
!
neighbor 10.120.23.6 remote-as 23456
neighbor 10.120.23.6 prefix-list Default in
neighbor 10.120.23.6 route-map MED50 out
!

```

© 2010 Cisco Systems, Inc. All rights reserved.

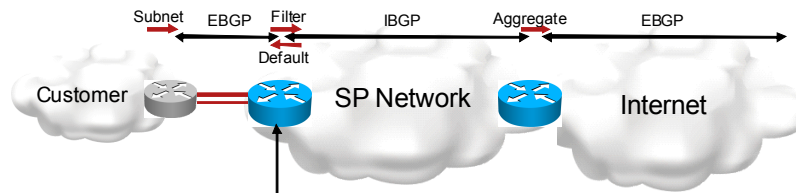
MSRP v1.0-224

The figure shows a Cisco IOS Software configuration of the customer edge router, which includes IP addressing, routing, and influencing the link selection for traffic flow between the customer and the ISP. Two BGP neighbors, one for each link between the customer and the ISP are configured. The customer's provider-assigned subnet 209.165.201.128/28 is advertised through BGP using the **network** command.

Prefix list “Default” permits only the receipt of default routes from the ISP. By configuring route maps “MED” and “LocalPref,” you assure that the upper link is selected as the primary path in both directions.

Example: Customer Addressing and Routing

Backup Link with BGP—Cisco IOS Software (Cont.)



```
interface GigabitEthernet1/0
ip address 10.120.23.2 255.255.255.252
!
interface GigabitEthernet2/0
ip address 10.120.23.6 255.255.255.252
!
interface GigabitEthernet3/0
ip address 10.120.24.1 255.255.255.252
!
ip route 0.0.0.0 0.0.0.0 Null0
!
ip prefix-list C1 permit 209.165.201.128/28
!
ip prefix-list Default permit 0.0.0.0/0
!

router bgp 23456
 redistribute static
 neighbor 10.120.23.1 remote-as 65001
 neighbor 10.120.23.1 default-originate
 neighbor 10.120.23.1 prefix-list C1 in
 neighbor 10.120.23.1 prefix-list Def out
 neighbor 10.120.23.5 remote-as 65001
 neighbor 10.120.23.5 default-originate
 neighbor 10.120.23.5 prefix-list C1 in
 neighbor 10.120.23.5 prefix-list Def out
 neighbor 198.18.2.3 remote-as 23456
 neighbor 198.18.2.3 next-hop-self
```

© 2010 Cisco Systems, Inc. All rights reserved.

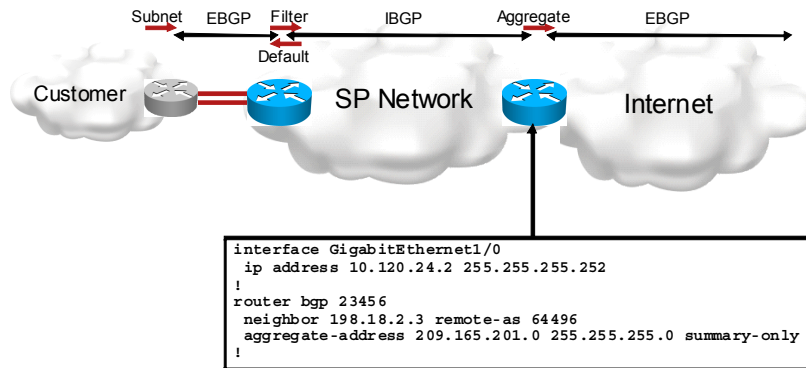
MSPRP v1.0-225

The figure shows a Cisco IOS Software configuration of the ISP edge router, which includes IP addressing, routing, and filtering. Three BGP neighbors are configured—two for redundant links to the customer and one for internal BGP sessions with other ISP routers. The default route is propagated to the customer using the **default-originate option** command. The next-hop reachability issue is solved using the **next-hop-self** command.

ISP is filtering incoming routing updates from the customer using prefix-list CustNet, allowing only the provider-assigned subnet 209.165.201.128/28 to be advertised. ISP uses prefix list “Default” to make sure that only the default route is sent to the customer.

Example: Customer Addressing and Routing

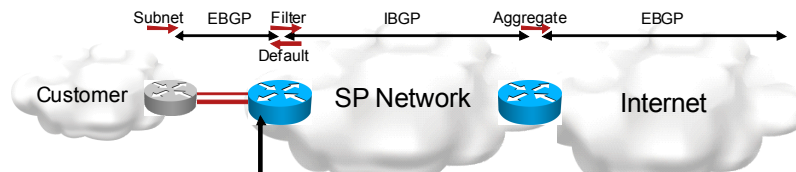
Backup Link with BGP—Cisco IOS Software (Cont.)



The figure shows a Cisco IOS Software configuration of the ISP edge router connecting to other ISPs. The aggregation of customer routes is configured before sending them to other ISPs.

Example: Customer Addressing and Routing

Backup Link with BGP—Cisco IOS XR Software



```
interface GigabitEthernet 1/0
  ipv4 address 10.120.23.1/30
!
interface GigabitEthernet 2/0
  ipv4 address 10.120.23.5/30
!
route-policy Default
  if destination in (0.0.0.0/0) then
    pass
  endif
end-policy
!
route-policy MED
  set med 50
end-policy
!
route-policy LocalPref
  set local-preference 200
end-policy

router bgp 23456
  address-family ipv4 unicast
    network 209.165.201.128/28
  !
  neighbor 10.120.23.2
    remote-as 65001
    address-family ipv4 unicast
    route-policy LocalPref in
  !
  neighbor 10.120.23.6
    remote-as 65001
    address-family ipv4 unicast
    route-policy Default in
    route-policy MED out
  !
!
```

© 2010 Cisco Systems, Inc. All rights reserved.

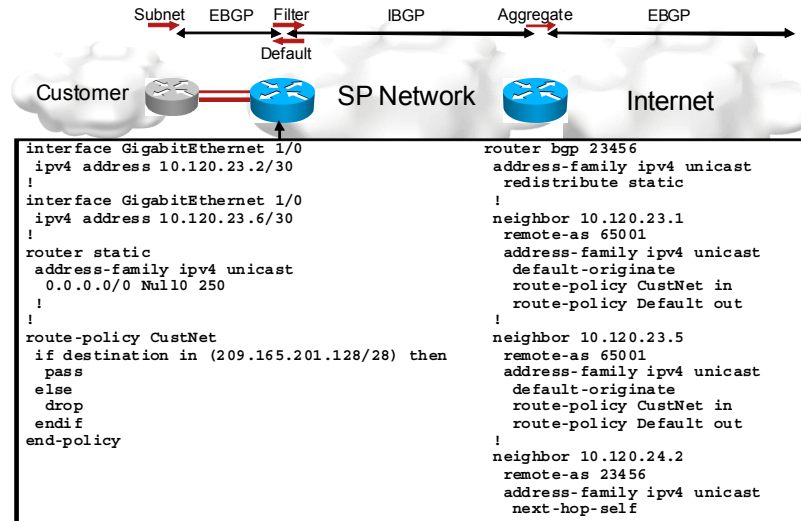
MSRP v1.0-227

The figure shows the Cisco IOS XR Software configuration of the customer edge router, which includes IP addressing, routing, and influencing the link selection for the traffic flow between the customer and the ISP. Two BGP neighbors, one for each link between the customer and the ISP, are configured. The customer provider-assigned subnet 209.165.201.128/28 is advertised through BGP using the **network** command.

The route policy “Default” permits only the receipt of default routes from the ISP. By configuring route policies “MED” and “LocalPref,” you assure that the upper link is selected as the primary path in both directions.

Example: Customer Addressing and Routing

Backup Link with BGP—Cisco IOS XR Software (Cont.)

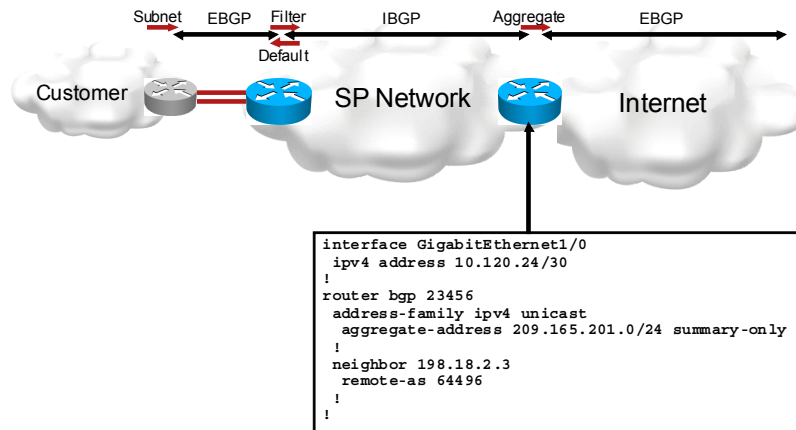


The figure shows Cisco IOS XR Software configuration of the ISP edge router, which includes IP addressing, routing, and filtering. Three BGP neighbors are configured—two for redundant links to the customer and one for an internal BGP session with the other ISP router. The default route is propagated to the customer using the **default-originate** command. The next-hop reachability issue is solved using the **next-hop-self** command.

ISP is filtering incoming routing updates from the customer using route policy CustNet, allowing only the provider-assigned subnet 209.165.201.128/28 to be advertised. The provider uses the route policy “Default” to make sure that only the default route is sent to the customer.

Example: Customer Addressing and Routing

Backup Link with BGP—Cisco IOS XR Software (Cont.)



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.29

The figure shows Cisco IOS XR Software configuration of the ISP edge router connecting to the other ISP. The aggregation of customer routes is configured before sending them to other ISPs.

Customer Reachability

This topic describes how to ensure reachability for a single-homed customer.

Customer Reachability

- **Configuration:**
 - Provider-assigned subnet in the customer network
 - Public addressing on customer access link
 - Default static route on customer edge router to the service provider
 - Static route on service provider edge router to the customer subnet
 - Redistribution of static route into IBGP of service provider
 - Aggregation of customer's routes before sending to other ISPs
- **Verification:**
 - Testing reachability of customer edge provider-assigned subnet and access link from Internet-facing provider edge router using **show** commands.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-231

The generic example of connecting a single-homed customer to the service provider will be shown, and the verification of configuration will be done using some **show** commands. The configuration will include the following:

- Provider-assigned subnet in the customer network
- Public addressing on the customer access link
- Default static route on the customer edge router to the ISP
- Static route on the ISP edge router to the customer subnet
- Redistribution of the static route into the IBGP of the service provider
- Aggregation of the customer routes before sending to other ISPs

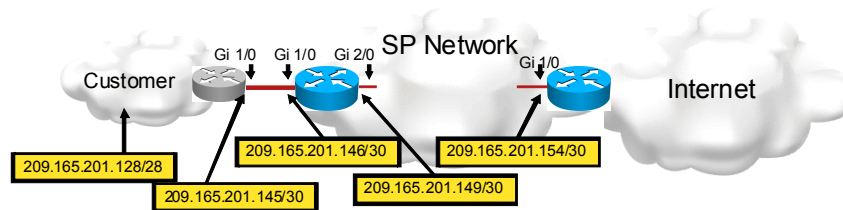
The verification will include the following:

- Testing reachability of the customer edge provider-assigned subnet and the access link from the Internet-facing provider edge router using **show** commands

Customer Reachability

Addressing

- Provider-assigned subnet in the customer network (209.165.201.128/28)
- Public addressing on customer access link (209.165.201.144/30)



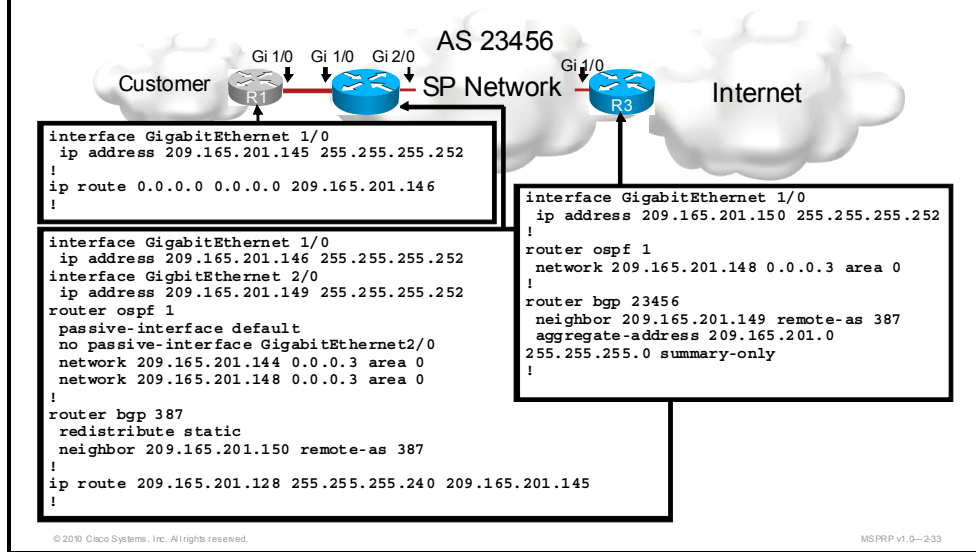
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-232

The addressing used in the example is shown in the figure. The customer received provider-assigned subnet 209.165.201.128/28 for its own network. The customer access link uses public subnet 209.165.201.144/30, while the link between ISP routers uses the subnet 209.165.201.148/30.

Example: Customer Reachability

Cisco IOS Software



The figure shows Cisco IOS configurations for the customer edge router, the provider edge router facing the customer, and the ISP edge router facing other ISPs.

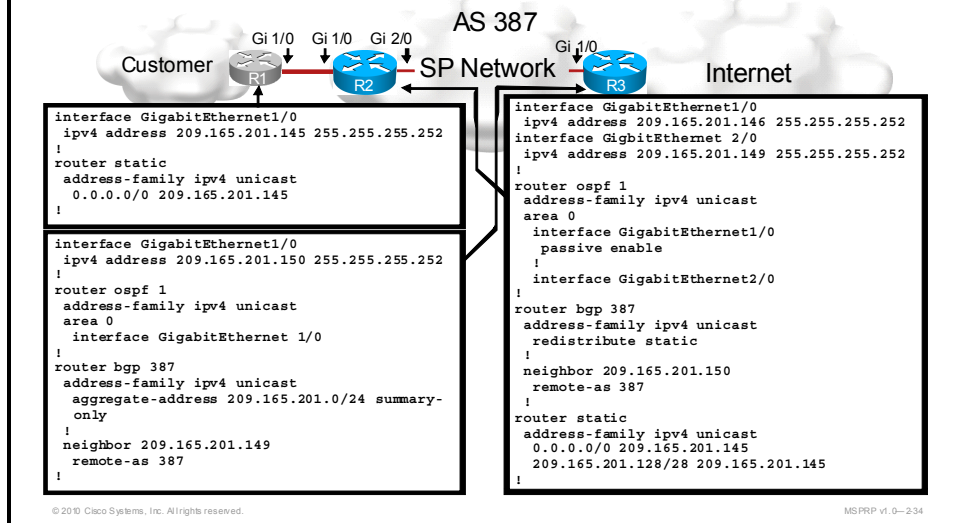
On customer edge router R1, the default static route to ISP is configured, pointing to the Gigabit Ethernet 1/0 interface of the service provider edge router R2.

On router R2, the customer access subnet 209.165.201.144/30 is propagated to the ISP network through an OSPF protocol. The interface pointing to the customer is configured as passive to prevent sending OSPF updates to the customer. The static route to the provider-assigned public subnet of the customer is also configured on router R2. This static route is redistributed into IBGP, which is running between service provider routers R2 and R3.

The aggregation of customer routes before sending them to other ISPs is configured on R3. This means that R3 has specific routes to all customers, but other ISPs have only one aggregate route to all customers in AS 387.

Example: Customer Reachability

Cisco IOS XR Software



The figure shows Cisco IOS XR configurations for the customer edge router, the provider edge router facing the customer, and the ISP edge router facing other ISPs.

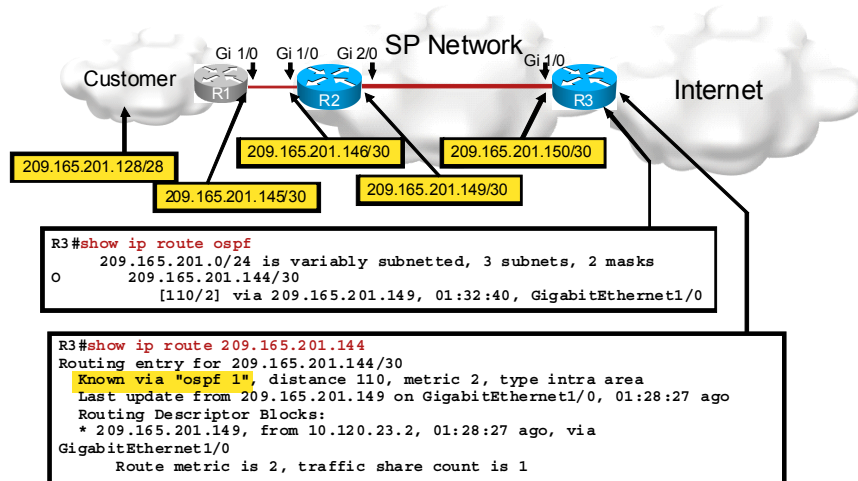
On the customer edge router R1, the default static route to ISP is configured, pointing to the Gigabit Ethernet 1/0 interface of ISP edge router R2.

On router R2, the customer access subnet 209.165.201.144/30 is propagated to the ISP network through the OSPF protocol. The interface pointing to the customer is configured as passive to prevent sending OSPF updates to the customer. The static route to the provider-assigned public subnet of the customer is also configured on router R2. This static route is redistributed into IBGP, which is running between ISP routers R2 and R3.

The aggregation of customer routes before sending them to other ISPs is configured on R3. This means that R3 has specific routes to all customers, but other ISPs have only one aggregate route to all customers in AS 387.

Example: Customer Reachability

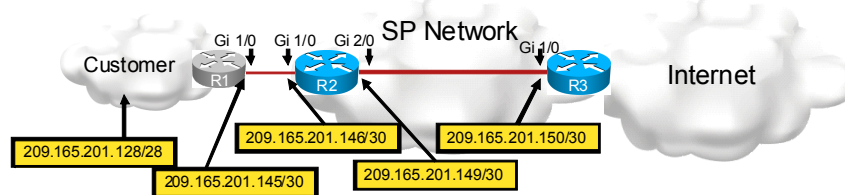
Verification



Outputs in the figure verify reachability of the customer access link from ISP router R3. Both outputs are indicating that the route from ISP router R3 to the customer access subnet 209.165.201.144/30 is learned through OSPF protocol.

Example: Customer Reachability

Verification (Cont.)



```
R3# show ip route bgp
209.165.201.0/24 is variably subnetted, 3 subnets, 2 masks
B    209.165.201.128/28 [200/0] via 209.165.201.145, 02:12:32
```

```
R3# show ip route ospf
209.165.201.0/24 is variably subnetted, 3 subnets, 2 masks
O    209.165.201.144/30
      [110/2] via 209.165.201.149, 02:21:02, GigabitEthernet1/0
```

```
R3# show ip bgp
BGP table version is 2, local router ID is 209.165.201.150
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop           Metric LocPrf Weight Path
*>i 209.165.201.128/28 209.165.201.145             0    100    0 ?
```

© 2010 Cisco Systems, Inc. All rights reserved.

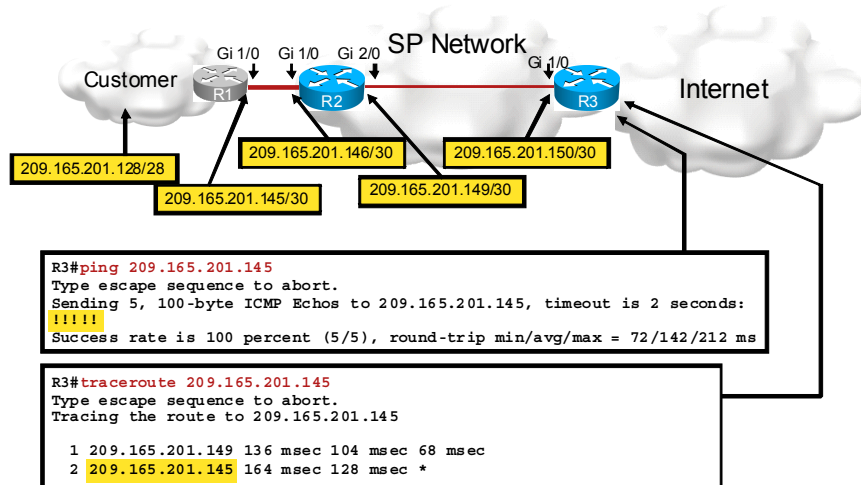
MSRP v1.0-336

Outputs in the figure verify reachability of the provider-assigned customer subnet from ISP router R3. Outputs are indicating that the route from the ISP router R3 to the customer subnet 209.165.201.128/30 is learned through BGP protocol.

For customer subnet 209.165.201.128/30, the next-hop address is 209.165.201.145, which is learned through OSPF protocol.

Example: Customer Reachability

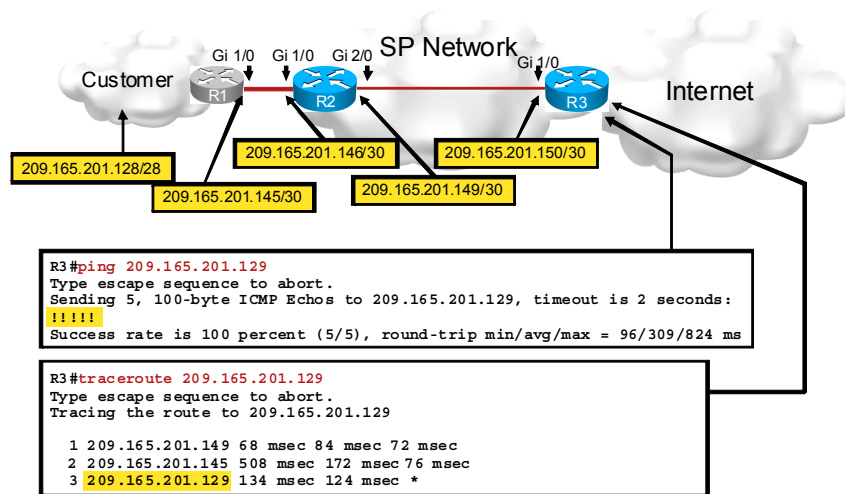
Verification (Cont.)



The figure shows two outputs, both checking reachability of the customer's edge router R1 interface Gigabit Ethernet 1/0 with IP address 209.165.201.145. Based on output using the **ping** command, the connection is working fine. Using the **traceroute** command, you can check the route that was selected for packets that are destined for the customer access link.

Example: Customer Reachability

Verification (Cont.)



The figure shows two outputs, both checking reachability of the provider-assigned subnet that the customer received. Reachability of the IP address 209.165.201.129 inside this subnet is checked. Based on output using the **ping** command, the connection is working fine. Using the **traceroute** command, you can check the route that was selected for packets that are destined for the customer subnet.

Summary

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

Summary

- You must identify single-homed customer requirements:
 - Addressing requirements
 - Availability requirements
 - Routing requirements
- Static routing is adequate for the single-homed customer.
- BGP is adequate if the backup connection to ISP exists to detect failures and reroute onto the backup link .
- Customer has static default route configured to ISP.
- ISP has static route pointing to customer subnet configured.
- Customer access link could be propagated into the ISP network using redistribution into IBGP or using IGP (OSPF or IS-IS).
- Static route to the customer subnet is usually redistributed into IBGP of ISP.
- Before sending customer routes to other ISPs, routes should be aggregated.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-239

Lesson Self-Check

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

- Q1) Which three requirements should be identified when connecting a new single-homed customer to an ISP? (Choose three.) (Source: Single-homed Customer Requirements)
- A) switching
 - B) addressing
 - C) routing
 - D) availability
- Q2) Which kind of routing is adequate in a single-homed environment (without backup connection to ISP)? (Source: Customer Addressing and Routing)
- A) dynamic
 - B) static
- Q3) What should be configured on a single-homed customer edge router? (Source: Customer Addressing and Routing)
- A) dynamic routing protocol
 - B) static default route
 - C) static route
 - D) aggregation of routes
- Q4) What should be configured on an ISP edge router pointing to the single-homed customer? (Source: Customer Addressing and Routing)
- A) dynamic routing protocol
 - B) static default route
 - C) static route
 - D) aggregation of routes
- Q5) What should be configured on an ISP edge router pointing to other ISPs to decrease routing table size? (Source: Customer Addressing and Routing)
- A) dynamic routing protocol
 - B) static default route
 - C) static route
 - D) aggregation of routes

Lesson Self-Check Answer Key

- Q1) B, C, D
- Q2) B
- Q3) B
- Q4) C
- Q5) D

Connecting Multihomed Customers to the Network

Overview

In this lesson, you will identify the mechanisms to connect a customer to two service providers for maximum redundancy. The lesson focuses on the mechanisms that are available to connect a customer to two service providers for load balancing, as well as maximum redundancy. The lesson also discusses the Border Gateway Protocol (BGP) characteristics that are used to configure customer and service provider routers to accommodate multiple connections between them.

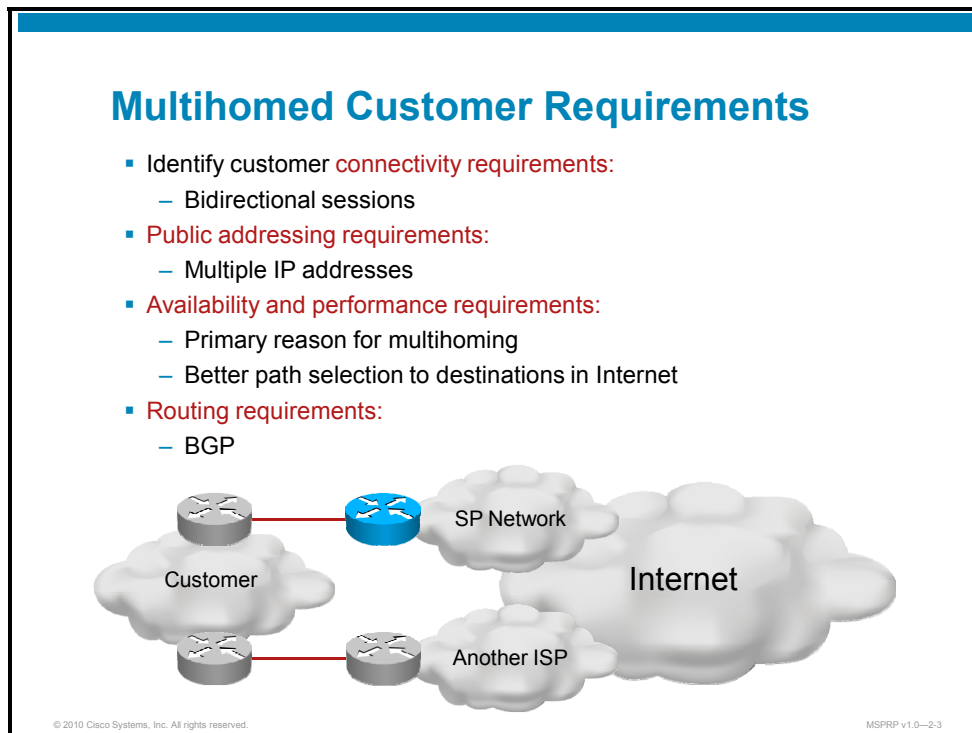
Objectives

Upon completing this lesson, you will be able to identify the mechanisms to connect a customer to two service providers for maximum redundancy. This ability includes being able to meet these objectives:

- Describe the typical routing requirements when adding a new multihomed customer
- Ensure reachability for a multihomed customer
- Provide the required routing policy to a multihomed customer
- Use strict route filtering for routes coming from a multihomed customer

Multihomed Customer Requirements

This topic describes customer requirements for multihomed connectivity.



Some customers that provide services on publicly available servers (e-commerce, e-learning) require high availability and bidirectional load balancing. The highest level of availability is achieved in multihomed networks that connect to two different service providers. Multihomed users have links to different service providers, which should terminate in different edge routers in the customer network.

The customer would usually be assigned with several IP addresses (subnet) from the provider-independent address space. This address space has to be advertised from the customer to both service providers and from them further into the Internet. Both service providers should advertise at least some BGP routes. Depending on customer requirements, the volume of BGP routes that the ISP provides can range from a default route only to the complete Internet routing table.

This setup should address rerouting in the following situations:

- Link failure between the customer and one ISP
- Edge router failure on either the customer or the ISP side
- Link failure in the customer network, where inside routers connect to one of the edge routers
- Link or router failure in one of the ISPs

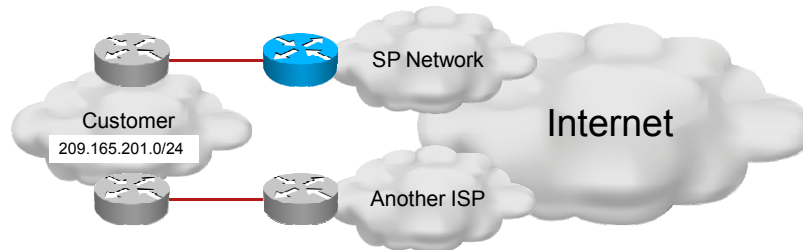
The only routing mechanism that can be used in this case is BGP.

Another reason for multihoming can be performance requirements. In multihomed networks, path manipulation can be performed using BGP. Traffic going to destinations that are closer to one service provider should take a path through that provider. Traffic going to a destination closer to the second service provider should go through the second provider.

Multihomed Customer Requirements

Addressing

- Addressing requirements:
 - Provider-independent address space
 - Obtain provider-assigned address from local ISP



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-4

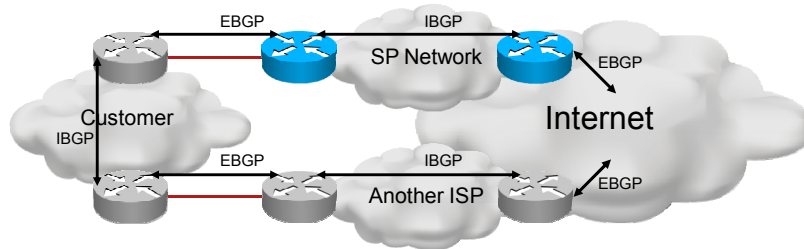
Customers can obtain address space in two different ways:

- The customer can use provider-independent address space, when a Regional Internet Registry assigns IP addresses to end users without going through a service provider. The benefit is independence from ISP. In this case, the customer has to advertise this address space to the ISP using BGP.
- The customer can use provider-assigned address space that is assigned by an ISP. In this case, the ISP will advertise the address space to the Internet. However, in high availability design, the other ISP will not advertise addressing space belonging to the other ISP. In this case, there is no purpose in using BGP to provide redundant connectivity.

Multihomed Customer Requirements

Routing

- Routing requirements:
 - The only acceptable routing protocol is BGP.
 - Public autonomous system number (AS number) is required.
 - Obtain public AS number from local regional Internet registry.
- Routing policy requirements:
 - Primary and backup
 - Load balancing



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.5

BGP is the only routing protocol that is usually supported by ISPs. This is also the only protocol that can be used to advertise the huge number of networks that are available today on the Internet. Since BGP uses a rich metric (also called attributes), a path selection can also be configured on the customer and the ISP side.

The use of BGP requires an autonomous system (AS) number. A customer can use a private AS number that is assigned by the ISP or can obtain a public AS number from a Regional Internet Registry. The latter is recommended since the routing over the Internet will be consistent. When using private AS numbers, the ISP has to remove private AS numbers from the AS path of customer networks before advertising them further to the Internet. Those customer routes then appear as originating in the ISP AS.

The customer can decide to use the high availability setup in **primary/backup** or **load balancing** configuration. In primary/backup configuration, one link will mainly be used, and the backup will be used only if there is a failure on the primary link. For outgoing traffic, this configuration can be easily achieved on the customer side by increasing a local preference for networks advertised by the primary ISP. For incoming traffic, AS prepending can be used for routes that are advertised to the backup ISP, thus lengthening the AS path. In this case, routers in the Internet will prefer the path with the shorter AS path, which is through the primary ISP.

Another configuration is load balancing, in which both links will be used to balance the traffic going to and from the Internet.

Multihomed Customer Requirements

- **ISP-provided information:**
 - Access link address and other parameters
 - Available services based on BGP communities
- **Customer-provided information:**
 - Autonomous system number
 - Provider-independent address space
 - Routing table requirements:
 - Default route
 - Partial Internet routing table
 - Full Internet routing table
 - Other, depending on ISP services

© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0-2.6

The ISP should provide the following information, which is needed to implement high availability setup.

- IP addresses for the access link, AS number, and other parameters.
- BGP communities that can be used for different services. The customer can mark networks with communities specified by the ISP before advertising them to the ISP. The ISP can then assign different attributes to these routes (change local preference or configure AS path prepending).

The customer must provide some parameters:

- AS number
- Provider-independent address space

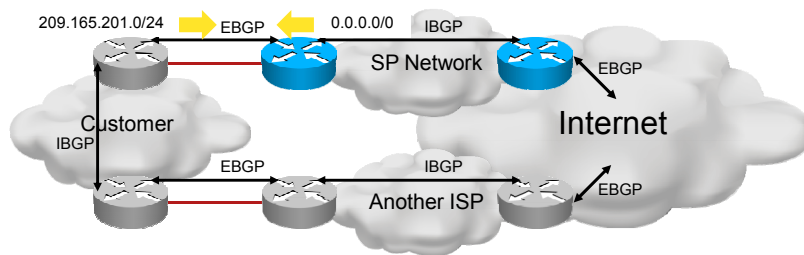
When a customer decides to use multihoming with BGP, there are three different ways to implement it:

- Each ISP passes only a default route to the customer.
- Each ISP passes a default route and local routes (routes belonging to ISP and customers that are connected to this ISP).
- Each ISP passes a complete Internet routing table.

Multihomed Customer Requirements

Default Route Only

- Send only default route to customer:
 - Outbound filter permitting only default route
- Accept only customer prefix:
 - Inbound AS path access list permitting customer-originated prefixes
 - Inbound prefix list permitting customer prefixes



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.7

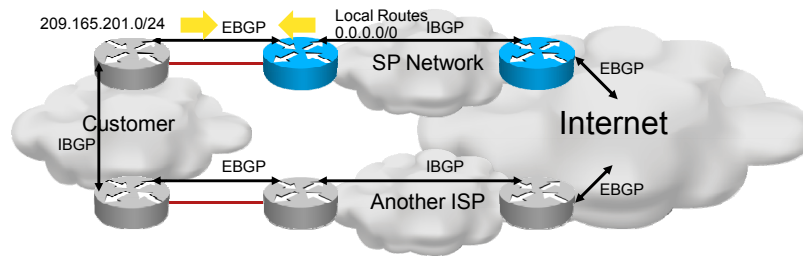
In the first implementation option, presented in the figure, each ISP sends only the default route to the customer. Load balancing cannot be achieved using this implementation. One ISP will act as primary and another ISP will act as secondary.

The ISP edge routers should be configured with inbound and outbound filters to filter routing updates. The edge router should pass only the default route to the customer, and it should receive only routes belonging to the customer, thus preventing the customer from becoming a transit AS. Outbound and inbound filtering can be implemented using prefix lists. However, it can be cumbersome for an ISP to identify all customer routes. Another option is to filter routes using an AS path access list.

Multihomed Customer Requirements

Routing: Local Routes

- Send local routes, routes from direct customers, and default route:
 - Outbound filter permitting local routes (ISP's address space and prefixes of direct customers) and default route
- Accept only customer prefix:
 - Inbound AS path access list permitting customer-originated prefixes
 - Inbound prefix list permitting customer prefixes



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-8

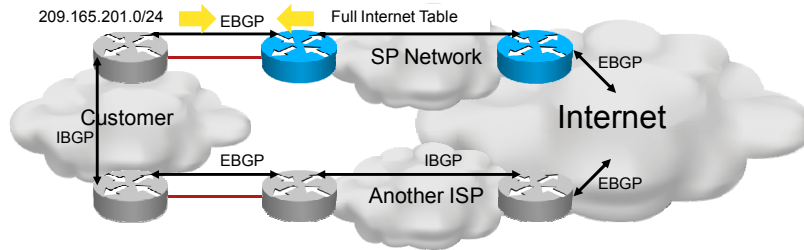
As depicted in the figure, each ISP passes a default route and local routes in the second implementation option. These local routes belong to the ISP and other customers that are connected directly to the ISP. Traffic going to the backup ISPs, or its directly connected customer, will take a path through the backup ISP. Traffic going to the Internet will use a default route that is advertised by the primary ISP. The benefit of this setup is more optimal routing; the disadvantage is more complex configuration.

Inbound and outbound filtering has to be configured on ISP edge routers. The outbound filter permits only the default route and local routes to be advertised. This filtering can be implemented using prefix lists and an AS path access list. Inbound filtering permits only customer routes to be advertised to the ISP. It can be implemented using prefix lists or AS path access lists.

Multihomed Customer Requirements

Routing: Full Internet Routing Table

- Send full Internet routing table:
 - No outbound filter in Cisco IOS Software
 - Outbound routing policy permitting everything in Cisco IOS XR Software
- Accept only customer prefix:
 - Inbound AS path access list permitting customer-originated prefixes
 - Inbound prefix list permitting customer prefixes



In the third implementation option, each ISP passes a complete Internet routing table to the customer. This solution offers several benefits:

- Optimal routing, because traffic to destinations in the Internet will take the shortest (preferred) path.
- Load balancing, which can be achieved by manipulating path attributes.
- Simple configuration, because no filters are needed on Cisco IOS routers, and a filter permitting everything is needed on Cisco IOS XR routers.

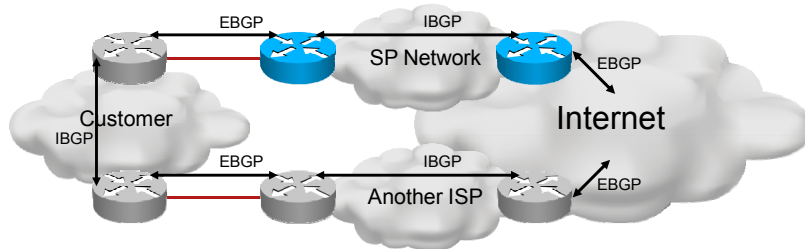
However, a disadvantage is huge routing tables on the customer side, where routers with enough memory must be used.

Inbound filtering is still needed, in this case, to permit only customer routes to be advertised to the ISPs, thus preventing a customer from becoming a transit AS.

Multihomed Customer Requirements

Routing: BGP Policy

- BGP path selection on customer side:
 - Change local preference
- BGP path selection on the Internet:
 - AS path prepend
 - Translate BGP communities to local preference



© 2010 Cisco Systems, Inc. All rights reserved.

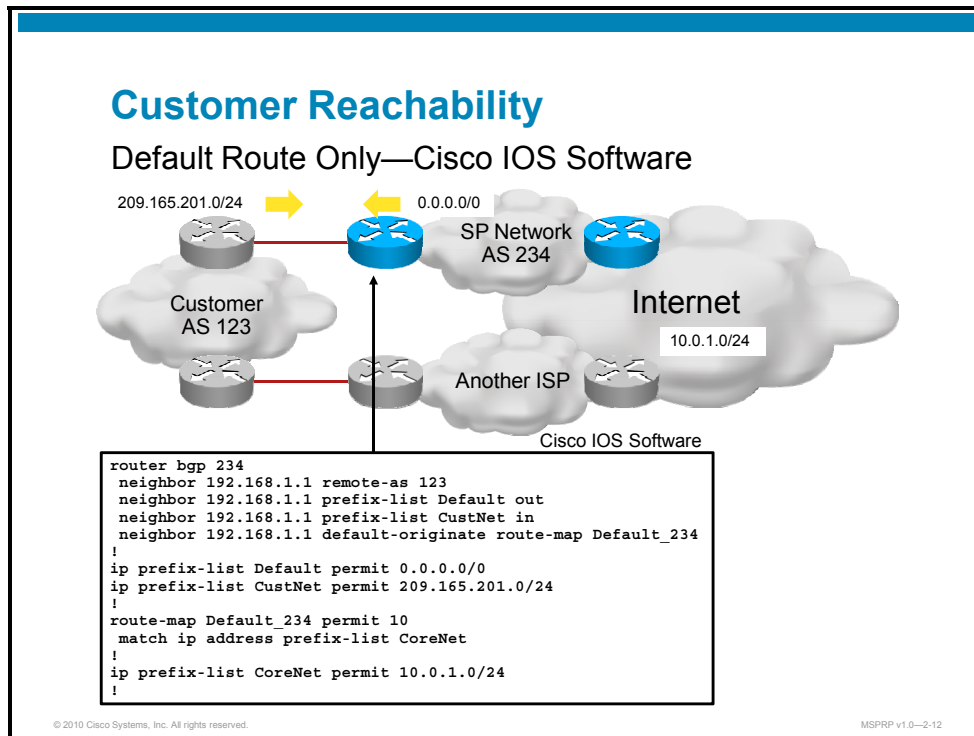
MSPRP v1.0-2-10

In addition to routing filters, routing policies should be implemented to influence path selection for incoming and outgoing traffic. On the customer side, the local preference for routes that are received from the primary ISP should be increased for outgoing traffic to prefer that path.

For traffic originating in the ISP and beyond, path manipulation can be configured either on the customer side or the ISP side. The customer can configure AS path prepending to lengthen an AS path of routes advertised to the backup ISP. Another solution is to signal ISP preference using BGP communities that are defined by the ISP. Those communities can be translated to some other BGP attributes inside the ISP network, such as local preference or AS path prepend.

Customer Reachability

This topic describes Cisco IOS Software and IOS XR Software configuration scenarios for multihoming with BGP.



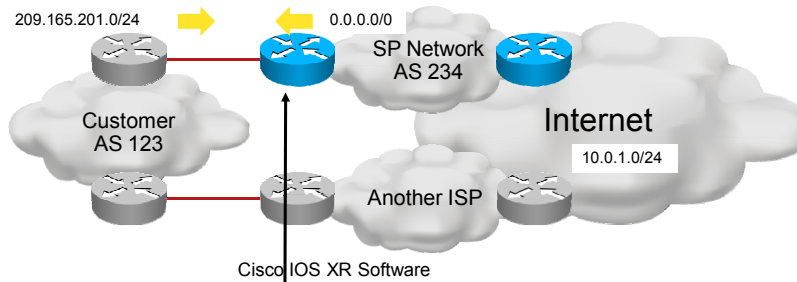
In the first scenario, both ISPs advertise only a default route to the customer. Since the unconditional BGP default route advertisement could lead to traffic black holes, advertising default routes to the BGP neighbor is conditional. If the primary ISP loses connectivity to the network core, but still advertises the default route to the customer, the primary ISP will attract the traffic from AS 123. It will then drop it due to failed uplink.

The figure shows a configuration on a Cisco IOS edge router on the ISP side. For filtering routing updates, prefix lists are configured and applied to the BGP neighbor. For outbound filtering, a prefix list *Default*, which permits only a default route, is configured. This prefix list is applied to the BGP neighbor (customer edge router) in the outbound direction. For inbound filtering, a prefix list *CustNet*, which permits only the customer network, is configured. This prefix list is applied to the same BGP neighbor, but in the inbound direction.

Conditional default route advertisement is also configured. The route map, which matches the core network, is applied to the **default-originate** command. The presence of the core network (10.0.1.0/24) in the routing table indicates that the core network is available. Only in this case, a default route will be advertised to the customer router.

Customer Reachability

Default Route Only—Cisco IOS XR Software



```
router bgp 234
 neighbor 192.168.1.1
 remote-as 123
 address-family ipv4 unicast
 default-originate route-policy
Default_234
 route-policy Default out
 route-policy CustNet in
!
route-policy Default
 if destination in (0.0.0.0/0) then
  pass
 endif
end-policy
```

```
route-policy CustNet
 if destination in (209.165.201.0/24) then
  pass
 endif
end-policy
!
route-policy Default_234
 if rib-has-route in (10.0.1.0/24) then
  pass
 endif
end-policy
```

© 2010 Cisco Systems, Inc. All rights reserved.

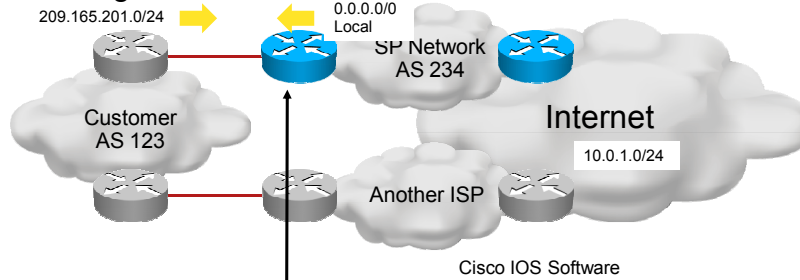
MSPRP v1.0-2-13

In the Cisco IOS XR Software, route policy language is used to filter routing updates. The *Default_234* route policy is configured for conditional default route advertisement. This route policy verifies whether the core network is available in the routing table. Only in this case, the default route is advertised to the customer router.

Two policies are configured for route filtering. Route policy *Default* allows only a default route to be propagated to the customer. It is applied to the BGP neighbor in the outbound direction. Route policy *CustNet* allows only the customer network to be advertised to the ISP router. The route policy is applied to the neighbor in the inbound direction.

Customer Reachability

Routing: Local Routes—Cisco IOS Software



```
router bgp 234
 neighbor 192.168.1.1 remote-as 123
 neighbor 192.168.1.1 filter-list 1 out
 neighbor 192.168.1.1 prefix-list CustNet in
 neighbor 192.168.1.1 default-originate route-map Default_234
!
ip as-path access-list 1 permit ^[0-9]*$
ip prefix-list CustNet permit 209.165.201.0/24
!
route-map Default_234 permit 10
 match ip address prefix-list CoreNet
!
ip prefix-list CoreNet permit 10.0.1.0/24
!
```

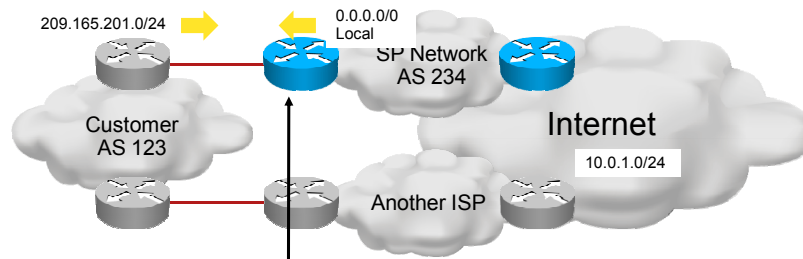
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-14

In the second scenario, each ISP sends a default route and local routes to a customer. The figure depicts a configuration on a Cisco IOS router. Conditional default route advertisement is configured. Inbound filtering is configured to allow only the customer network to be advertised to the ISP, thus preventing a customer from becoming a transit AS. Outbound filtering is configured using the AS path access list, which permits only networks originating in the AS 234 or in other autonomous systems connected directly to the AS 234. The regular expression that is used to specify the permitted AS path matches the AS path without any AS numbers or matches an AS path with one AS number.

Customer Reachability

Routing: Local Routes—Cisco IOS XR Software



```
router bgp 234
 neighbor 192.168.1.1
 remote-as 123
 address-family ipv4 unicast
 default-originate route-policy Default_234
 route-policy SPNet out
 route-policy CustNet in
!
route-policy SPNet
 if as-path length le 1 then
  pass
 endif
end-policy
```

```
route-policy CustNet
 if destination in (209.165.201.0/24) then
  pass
 endif
end-policy
!
route-policy Default_234
 if rib-has-route in (10.0.1.0/24) then
  pass
 endif
end-policy
```

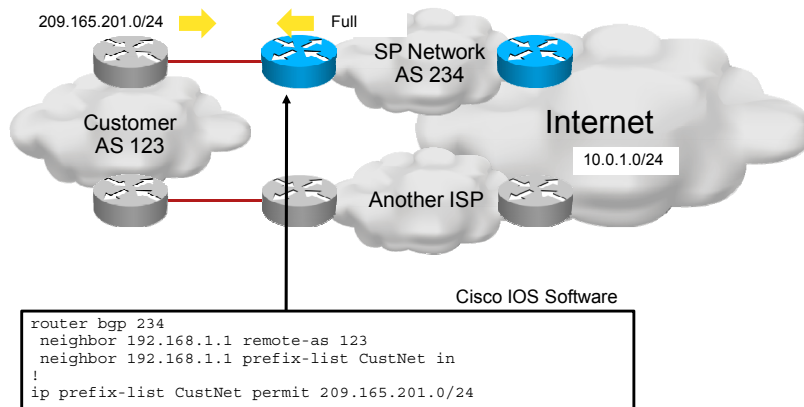
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-15

In Cisco IOS XR Software, route policy language is used to filter routing updates. The *Default_234* route policy is used for conditional default route advertisement. The *CustNet* policy is used to filter incoming updates and allows only customer routes. The *SPNet* policy is used to filter outgoing updates and permits routes with the AS path of length 1 or less. This policy matches routes originating from the AS 234 and from autonomous systems connected directly to the AS 234.

Customer Reachability

Routing: Full Internet Routing Table—Cisco IOS Software



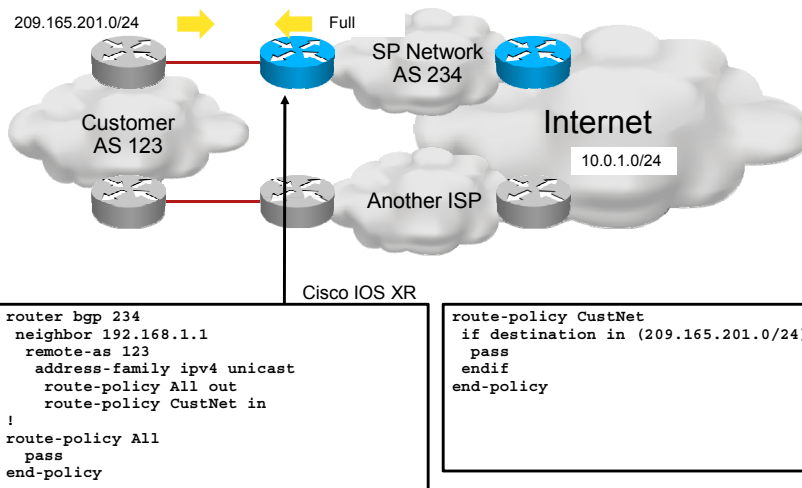
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-16

In the third scenario, each ISP sends a complete Internet routing table to the customer. The figure presents a configuration on a Cisco IOS router. Outbound filtering in this case is not needed, since the complete routing table is sent to the customer. Inbound filtering is still configured to receive only the customer route from the customer. Default route advertisement in this case is not needed.

Customer Reachability

Routing: Full Internet Routing Table—Cisco IOS XR Software



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-17

The figure presents a configuration on a Cisco IOS XR router. The *All* route policy is configured, which allows all routes to be propagated to the customer. The *CustNet* route policy is configured to allow only the customer route to be advertised to the ISP.

Customer Reachability

Verification

- Verify route propagation towards:
 - Customer (default route, partial or full routing table)
 - Internet (customer prefixes)
- Verify path selection on the customer side:
 - BGP commands
 - Ping and traceroute
- Verify path selection on the Internet side—use looking glass:
 - BGP commands
 - Ping and traceroute

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0–218

Verification of routes on the customer and the Internet side can be done by examining routing and BGP tables.

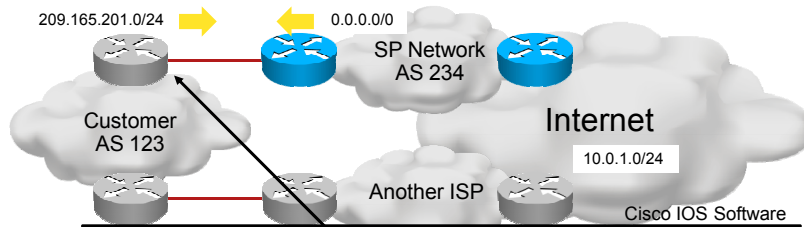
Path selection on the customer side can be verified using various BGP commands. In Cisco IOS Software, the **show ip bgp** command can be used to examine the BGP table. The **show ip bgp neighbors *neighbor_ip* advertised-routes** command can be used to verify which networks are advertised to a BGP neighbor. In Cisco IOS XR Software, the **show bgp** command can be used to examine the BGP table. The **show bgp neighbors *x.x.x.x* advertised-routes** command can be used to verify the networks that are advertised to a BGP neighbor.

Ping and **traceroute** can be used to verify connectivity and path selection from the customer to the Internet.

To verify path selection for incoming traffic originating on the Internet, BGP looking glasses—where BGP commands, **ping** and **traceroute**—can be used.

Customer Reachability Verification

Default Route Only



```
Customer# show ip bgp
BGP table version is 19, local router ID is 192.168.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop           Metric LocPrf Weight Path
*> 0.0.0.0         192.168.1.2             0         0 234 i
r>i209.165.201.0   172.16.0.1              0        100    0 i

Customer# show ip route
<...output omitted...>
Gateway of last resort is 192.168.1.2 to network 0.0.0.0

    172.16.0.0/24 is subnetted, 1 subnets
C      172.16.0.0 is directly connected, GigabitEthernet2/0
D    209.165.201.0/24 [90/130816] via 172.16.0.1, 00:10:57, GigabitEthernet2/0
C    192.168.1.0/24 is directly connected, GigabitEthernet1/0
B*    0.0.0.0/0 [20/0] via 192.168.1.2, 00:05:11
```

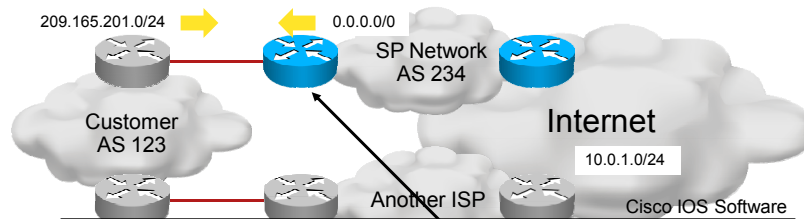
© 2010 Cisco Systems, Inc.

P v1.0—2-19

In the figure, the **show ip bgp** command is used on the customer side to verify that the customer received only the default route from the ISP. This default route was put in the routing table. The customer route (209.165.201.0/24) is marked with Routing Information Base (RIB) failure, because the network has been learned also through the Enhanced Interior Gateway Routing Protocol (EIGRP).

Customer Reachability Verification

Default Route Only (Cont.)



```
ISP# show ip bgp
BGP table version is 6, local router ID is 192.168.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop           Metric LocPrf Weight Path
*>i10.0.1.0/24    192.168.2.2             0      100      0  I
*> 10.10.10.0/24  0.0.0.0                 0           32768  i
*> 209.165.201.0 192.168.1.1             0           0 123  I

ISP# show ip bgp neighbors 192.168.1.1 advertised-routes
<...output omitted...>

Originating default network 0.0.0.0

   Network        Next Hop           Metric LocPrf Weight Path
Total number of prefixes 0
```

© 2010 Cisco Systems, Inc. All rights reserved.

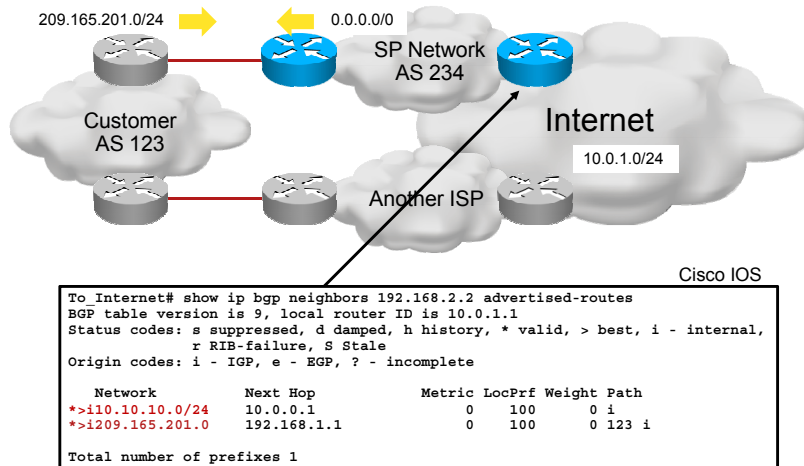
MSRP v1.0-2-20

In the figure, the **show ip bgp** command is used on the ISP side to verify that three routes are in the BGP table. Route 209.165.201.0/24 is advertised from the customer, 10.10.10.0/24 originates from the AS 234, and 10.0.1.0/24 is advertised from the Internet. Since the core network is available in the routing table, the default route will be advertised to the customer.

In the printout, the **show ip bgp neighbor_ip advertised-routes** command is used to verify that no routes are advertised to the customer edge router except the default route.

Customer Reachability Verification

Default Route Only (Cont.)



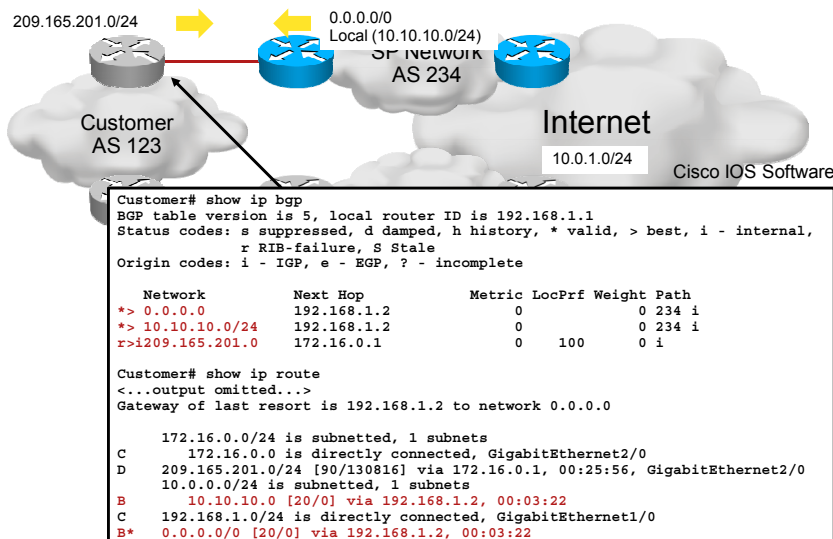
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-21

In the figure, the **show ip bgp neighbor ip advertised-routes** command is used to verify advertised routes from the AS 234 to the Internet router (not shown in the figure). It can be seen that the ISP advertises the customer route (209.165.201.0/24) and local routes (10.10.10.0/24) to the Internet.

Customer Reachability Verification

Routing: Local Routes



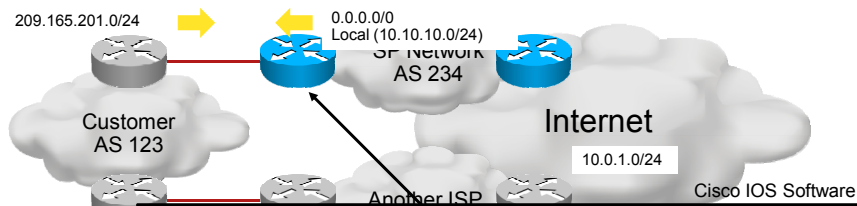
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-22

In the figure, the **show ip bgp** command is used to verify that the customer router receives a default route and a route belonging to the ISP. The **show ip route** command can be used to verify that both are put into the routing table. The customer router would use specific routes in the routing table to reach destinations in the ISP. To reach destinations beyond the AS 234, the customer router would use a default route.

Customer Reachability Verification

Routing: Local Routes (Cont.)



```
ISP# show ip bgp
BGP table version is 6, local router ID is 192.168.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop           Metric LocPrf Weight Path
*>i10.0.1.0/24      192.168.2.2             0   100      0 1 i
*> 10.10.10.0/24    0.0.0.0                 0           32768 i
*> 209.165.201.0   192.168.1.1             0           0 123 i

ISP# show ip bgp neighbors 192.168.1.1 advertised-routes
<...output omitted...>

Originating default network 0.0.0.0

   Network        Next Hop           Metric LocPrf Weight Path
*> 10.10.10.0/24    0.0.0.0                 0           32768 i

Total number of prefixes 1
```

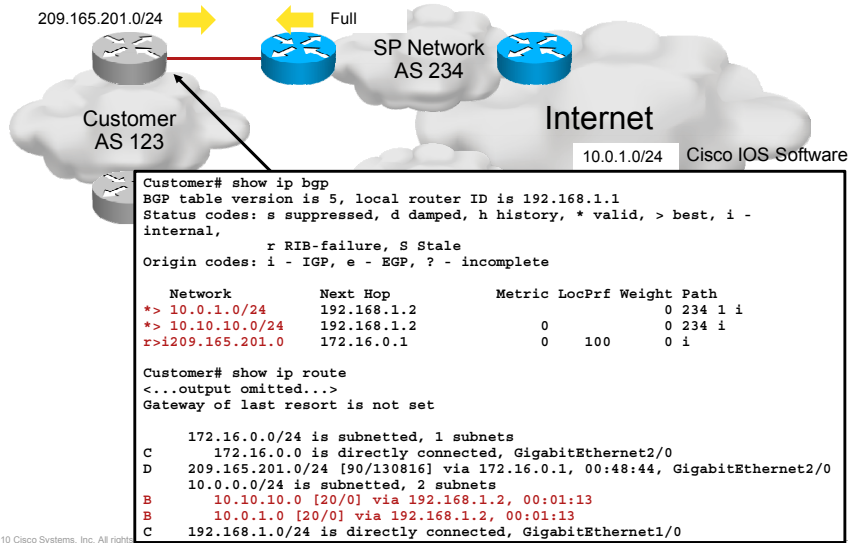
In the printout, the **show ip bgp** command is used on the ISP router to verify that three routes are in the BGP table: route 209.165.201.0/24 is advertised from the customer, 10.10.10.0/24 originates from the AS 234, and 10.0.1.0/24 is advertised from the Internet. Since the core network is available in the routing table, the default route will be advertised to the customer.

The **show ip bgp neighbor ip advertised-routes** command is used to verify that only the default route and the local ISP's routes (10.10.10.0/24) are advertised to the customer router.

The printout on the router connecting to the Internet, in this case, does not differ from the printout in the example with the default route only.

Customer Reachability Verification

Routing: Full Internet Routing Table

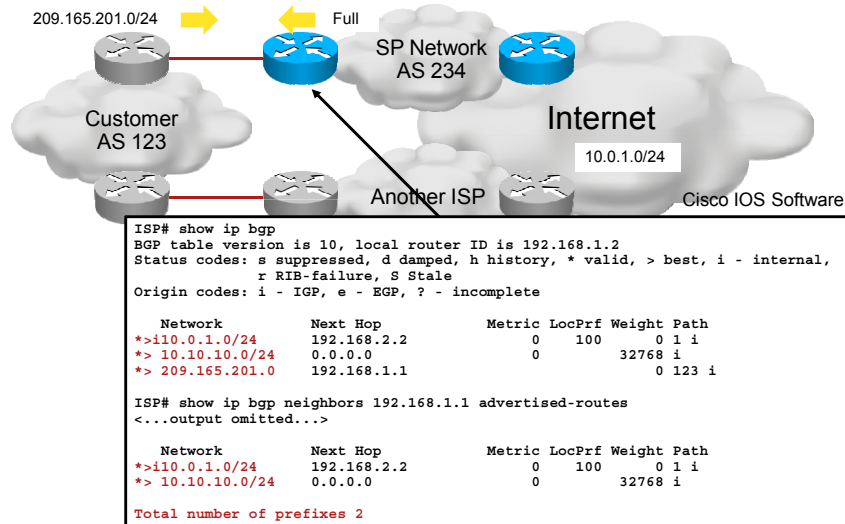


In the printout, the **show ip bgp** and **show ip route** commands are used to verify that a complete routing table is sent to the customer. The customer sees networks belonging to the AS 234 (10.10.10.0/24) and networks from the Internet (10.0.1.0/24). The **show ip route** command is used to verify that all networks that are learned through BGP are put into the routing table except the customer route, which is learned also through EIGRP.

The AS path for the Internet route indicates that the route originated in the AS 1 (Internet) and went through the AS 234 (ISP) to reach the customer.

Customer Reachability Verification

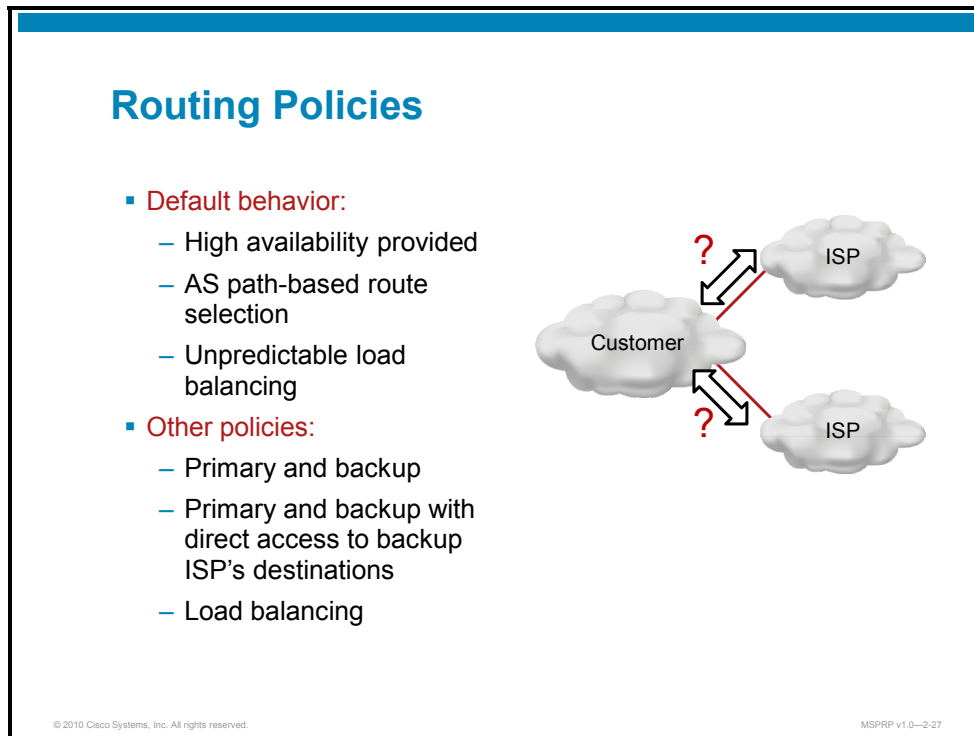
Routing: Full Internet Routing Table (Cont.)



In the printout on the ISP side, the **show ip bgp** command and **show ip bgp neighbors neighbor_ip advertised-routes** command are used to verify that all networks (local and from the Internet) are advertised to the customer router.

Routing Policies

This topic describes mechanisms and configurations on Cisco IOS and Cisco IOS XR platforms for implementing routing policies to influence path selection.



If each ISP advertises only a default route to the customer, one default route will be used as primary and another as secondary. Without any path manipulation, the oldest path will be used as primary for outgoing traffic destined for Internet destinations. Decisions for incoming traffic will be based on the AS path, and a path with the shortest AS path will be used. The result will probably be an asymmetric routing.

If each ISP advertises a complete routing table to the customer and there is no path manipulation that is configured, an unpredictable load balancing will occur. Outgoing traffic, which is destined for Internet destinations, will take a path with the shortest AS path. Decisions for incoming traffic will be based on the AS path, and a path with the shortest AS path will be used. The result will be unpredictable load balancing and asymmetric routing.

In multihomed BGP environments, different policies are possible:

- Primary and backup configuration
- Primary and backup configuration with direct access to the backup ISP's destinations
- Load balancing configuration

For the first scenario, the following configuration is recommended: configure a primary default route by increasing a local preference on the customer router for the default route that is received from the primary ISP. This way, routers inside the customer network will prefer the path through the primary ISP. For incoming traffic, the customer can configure AS path prepending for routes going to the secondary ISP, thus lengthening the AS path and making the path less preferred.

In the second scenario, the customer router is configured the same way as in the first scenario. The only difference is that the backup ISP will send a partial routing table (default route and local routes) instead of only the default route. For traffic going to the networks in the backup ISP, specific routes on the customer edge router would be used, and they would point to the backup ISP. For traffic going to all other destinations, a primary default route would be used.

In the third configuration, both ISPs have to send complete routing tables to the customer. The customer will increase local preference for half of the networks that are received by the primary ISP and the other half of networks that are received by the second ISP. This way, load balancing will be configured for incoming traffic, and one ISP will act as a backup for the other ISP. Load balancing for outgoing traffic can be achieved by subnetting the customer network and advertising one subnet to one ISP and another one to the other ISP (for example, splitting one /24 subnet into two /25 subnets). Note that it is very likely that some upstream ISPs will filter out prefixes longer than /24; this action will prevent this solution from working properly.

Routing Policies (Cont.)

- **Implemented by customer:**
 - Local preference usage for outbound path
 - AS path prepending for return path
 - Alternatively, conditional advertising for return path (slower convergence)
- **Optionally aided by ISP:**
 - Manually:
 - Local preference and/or AS path prepending for return path
 - Through signaling:
 - Translating BGP communities to local preference and/or AS path prepending

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-228

Routing policies can be implemented on the customer side. Optionally, the ISP can also implement policies.

The customer can implement routing policies:

- Changing local preference for the outbound path to influence the path selection for outbound traffic. The customer increases the local preference for routes that are received from the primary ISP to prefer paths through the primary ISP.
- AS path prepending to influence path selection for incoming traffic. The customer prepends an AS number to the routes advertised to the backup ISP to lengthen the AS path, thus making the path through the backup ISP less preferred.
- BGP conditional advertisement of customer routes to influence path selection for incoming traffic. In this case, the customer tracks a route in the primary ISP network. If the tracked route is available in the customer's BGP table, the customer advertises routes only to the primary ISP. If the tracked route is not present in the BGP table (this indicates that the connection to the primary ISP is down), the customer advertises its routes also to the backup ISP.

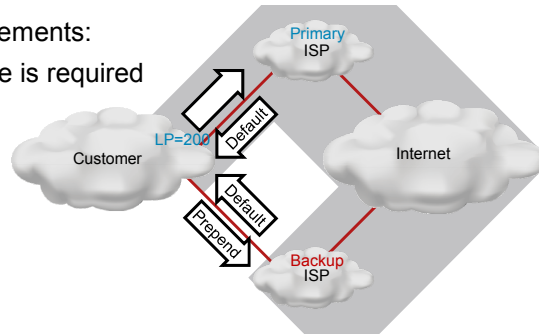
Routing policies can be implemented also on the ISP side. The ISP can set local preference for incoming routing updates and AS path prepending for outgoing updates (to upstream autonomous systems) in two ways:

- **Manually:** ISP changes local preference and configures AS path prepending manually for each customer route.
- **Through signaling:** The customer signals its needs using BGP communities. Each customer route is marked with BGP communities defined by the ISP. The ISP then matches those communities and sets specific local preference and AS path prepending.

Customer-Implemented Policies

Primary and Backup

- Primary and backup:
 - **Outbound path:** Higher local preference for routes from preferred ISP
 - **Return path:** AS path prepending of routes sent to backup ISP
- Routing table requirements:
 - Only default route is required



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-29

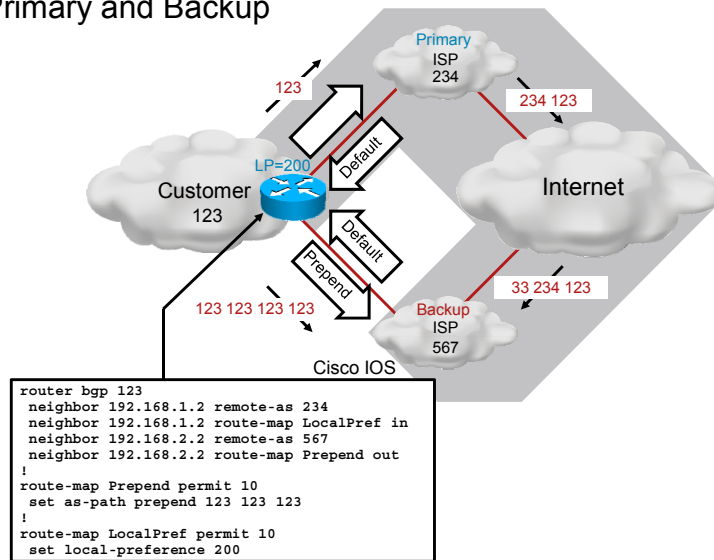
The figure depicts an example, in which one ISP is used as primary and another one as backup. In this case, only the default route is required from each ISP. The customer implements routing policies.

The customer selects the primary ISP for outgoing traffic by setting a higher local preference (200) for routes that are received from the preferred ISP. Since the default local preference is set to 100, the path through the primary ISP will be chosen.

The customer configures AS path prepending of routes that are sent to the backup ISP to lengthen the AS path. Returning traffic originating from the Internet will thus choose a shorter path through the primary ISP.

Customer-Implemented Policies

Primary and Backup



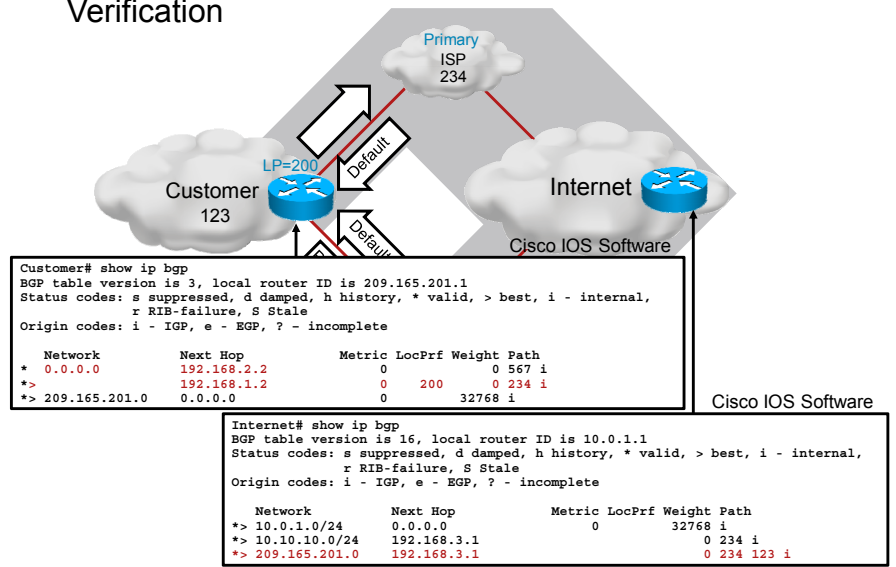
The sample configuration will result in AS 567 receiving a route to the customer network with an AS path containing the AS number 123 four times (three times because of prepending and once because of the External Border Gateway Protocol [EBGP] session). On the customer router, the default route that is received from the primary ISP will have a local preference of 200. The default route received from the backup ISP will have a local preference set to 100, which is a default value.

In the first step, two route maps are configured. The *LocalPref* route map is used to set local preference for received routes to 200. The *Prepend* route map is used to prepend the AS number to the advertised customer routes three times.

In the second step, the *LocalPref* route map is applied to the primary ISP neighbor in the inbound direction. The *Prepend* route map is applied to the backup ISP neighbor in the outbound direction.

Customer-Implemented Policies

Verification



In the figure, the **show ip bgp** command is used to verify the path selection. On the customer site, two default routes are present in the BGP table, but only one with the higher local preference is chosen and put into a routing table. The next hop of the chosen default route points to the BGP neighbor in the primary ISP (192.168.1.2).

On the Internet router, only one path to the network 209.265.201.0/24 is available. Two paths should be available, but the backup ISP does not advertise the route to the Internet because it received a route with the shortest AS path (AS 123, AS 234, AS Internet) from the Internet. Recall that the customer advertised a route to the backup ISP with the AS path prepending, where AS number 123 was prepended three times. The customer edge router also added the fourth copy of AS 123 because the route was sent across the EBGP session. The AS path of the customer route that was received by the customer was longer than the AS path of the customer route that was received from the Internet.

However, if the link through the primary ISP fails, the primary ISP will stop advertising the customer route to the Internet. In that case, the backup ISP would advertise the route to the Internet.

Customer-Implemented Policies

Primary and Backup with Direct Traffic

Primary and backup with direct traffic to backup ISP's destinations:

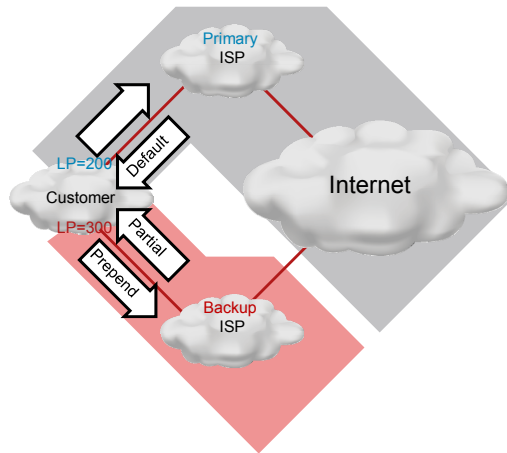
- Outbound path:
 - Higher local preference for routes from preferred ISP
 - Even higher local preference for routes connected to backup ISP's AS (optional)
- Return path: AS path prepending of routes sent to backup ISP

Routing requirements:

- Default route from primary ISP
- Partial routing table and default route from backup ISP

Issues:

- Identifying prefixes more closely through backup ISP
- How many times to prepend?



In some cases, it is desirable for traffic going to the destinations in backup ISP to take the shorter path through the backup ISP. In that case, the backup ISP should advertise default and local routes, while the primary ISP advertises only the default route. Local preference of the default route that is received from the primary ISP should be set to 200 to prefer that default route. The customer router will have specific routes to reach destinations in the backup ISP. These routes will point to the backup ISP default route, which will point to the primary ISP to reach all other destinations. If the primary ISP fails, the default route, which points to the secondary ISP, would be used.

However, to make sure that the secondary ISP would be used to reach networks in that ISP, the customer can set local preference for local routes that are received from the backup ISP to an even higher value, for example 300. This local preference is useful if the customer is not sure if the filtering of routes is properly configured on the primary ISP.

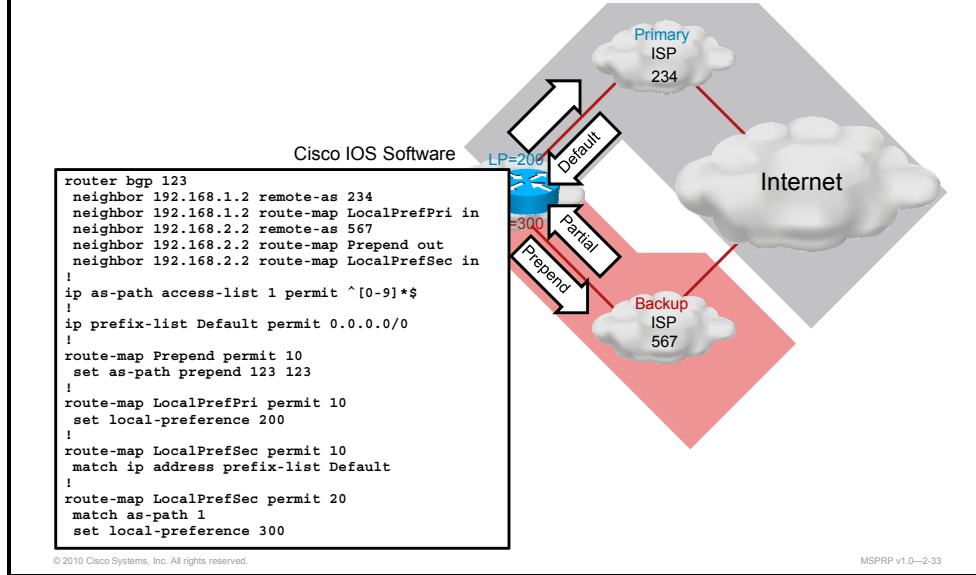
For the returning traffic, an AS path prepending would be configured for customer routes advertised to the backup ISP. In this case, the customer has to make sure not to prepend the AS number too many times. Traffic originating in the backup ISP should take the shortest path to the customer.

There are two issues arising in this scenario:

- How to identify prefixes that are closer through the second ISP?
- How many times to prepend the AS number, so that the traffic originating in the backup ISP goes directly to the customer and other traffic goes to the customer through the primary ISP?

Customer-Implemented Policies

Primary Backup with Direct Traffic (Cont.)



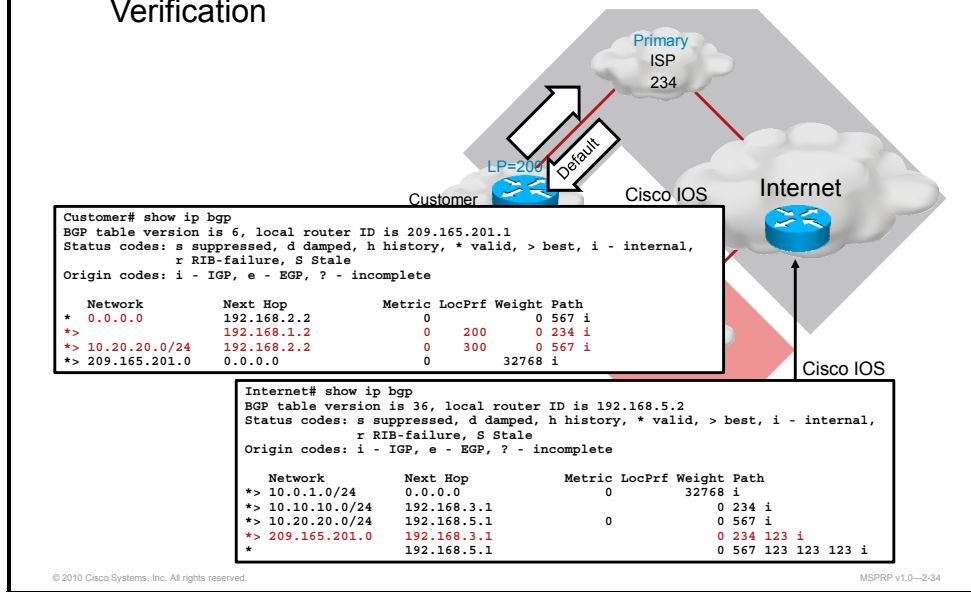
The sample configuration will result in the customer preferring the default route through the primary ISP. Returning traffic originating in the Internet will also prefer the path through the primary ISP. Traffic going to the backup ISP's networks and returning traffic will, on the other hand, take the shorter path through the backup ISP.

In the first step, three route maps are configured. The *Prepend* route map is used to prepend the AS number twice to routes going to the backup ISP. The *LocalPrefPri* route map is used to set a local preference for a default route that is received from the primary ISP to 200. The *LocalPrefSec* route map is used to match routes with AS-path length less or equal to one (routes originating in the backup ISP and autonomous systems directly connected to the backup ISP). Local preference of those routes is set to 300. Because the default route would also match the AS-path access list, another statement is configured at the beginning of that route map (sequence number 10), which matches the default route and just permits it without modifying any attributes (keeps the default local preference 100).

In the second step, these route maps are applied to BGP neighbors. The *LocalPrefPri* route map is applied to the primary ISP neighbor in the inbound direction. The *LocalPrefSec* route map is applied to the backup ISP neighbor in the inbound direction. The *Prepend* route map is applied to the backup ISP neighbor in the outbound direction.

Customer-Implemented Policies

Verification



In the figure, the **show ip bgp** command is used to verify the path selection. To the customer router, two default routes are advertised. Local preference of the one received from the primary ISP (192.168.1.2) is set to 200 to prefer it. Local routes from the backup ISP are also advertised, and local preference of those routes is set to 300.

On the Internet router, the primary and secondary ISPs advertise the customer network. The path through the primary ISP (192.168.3.1) will be preferred since the AS path is shorter.

Customer-Implemented Policies

Load Balancing

Outbound path:

- Higher local preference for half of routes from one ISP
- Higher local preference for the other half of routes from the other ISP

Return path:

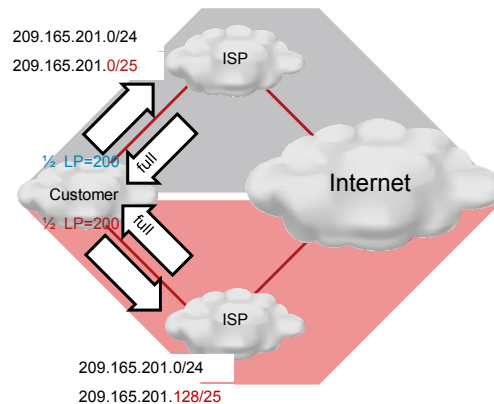
- Split prefix into two halves.
- Advertise one half to each ISP.
- Advertise full prefix to both ISPs.

Routing requirements:

- Full routing table from both ISPs

Issues:

- How to split full Internet routing table into two halves?
- Will load balancing (inbound and outbound) perform as expected?



The figure illustrates an example with a load-balancing configuration on the customer edge router. Both ISPs will serve as backups for each other. In this scenario, each ISP should advertise complete routing tables. Load balancing for outgoing traffic can be achieved by setting a local preference to a higher value for half of the routes that are received from one ISP and for the other half that are received from the other ISP. Traffic statistics should be gathered to identify traffic patterns (how much traffic goes to specific destinations in the Internet). Based on the traffic statistics, the customer can decide which routes' local preference should be increased on which ISP.

For incoming traffic, load balancing can be configured by splitting the customer network into two subnets and advertising one subnet to one ISP and the other subnet to another ISP.

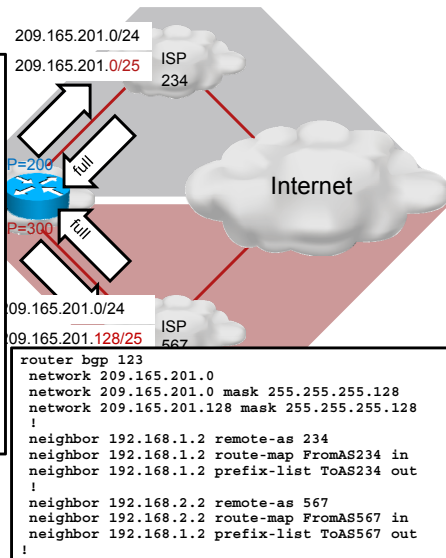
Two issues arise in this scenario:

- How to equally split the complete Internet routing table into two halves?
- Will load balancing perform as expected? Will there be any asymmetric routing?

Customer-Implemented Policies

Load Balancing (Cont.)

```
! Match routes originating in even ASNs
ip as-path access-list 1 permit [02468]$
!
! Match routes originating in odd ASNs
ip as-path access-list 2 permit [13579]$
!
route-map FromAS234 permit 10
match as-path 1
set local-preference 200
!
route-map FromAS234 permit 20
!
route-map FromAS567 permit 10
match as-path 2
set local-preference 200
!
route-map FromAS567 permit 20
!
ip prefix-list ToAS234 permit 209.165.201.0/0
ip prefix-list ToAS234 permit 209.165.201.0/25
!
ip prefix-list ToAS567 permit 209.165.201.0/0
ip prefix-list ToAS567 permit 209.165.201.128/25
!
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-38

The sample configuration will result in the customer router increasing the local preference to 200 for half of the networks that are received from one ISP and for the other half of networks that are received from the other ISP. The example illustrates one of the ways in which the complete Internet routing table can be split into two parts (not necessarily equal halves in the number of prefixes or in the resulting amount of load-balanced traffic). In the example, the prefixes that are received from one ISP will be assigned a local preference of 200 if they originate in an even-numbered AS (the AS-path attribute ends with 0, 2, 4, 6, or 8). The prefixes that are received from the other ISP will be assigned a local preference of 200 if they originate in an odd-numbered AS (the AS path attribute ends with 1, 3, 5, 7, or 9). Other prefixes will be assigned the default local preference of 100.

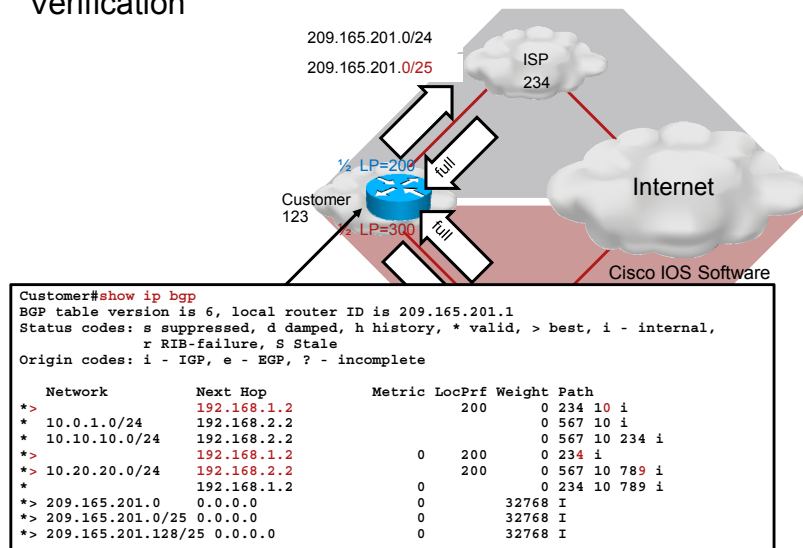
In the first step, two route maps are configured. Each matches one half of networks and sets a local preference to a value of 200. Because of the implicit deny at the end of the route map, another statement is configured that permits everything else without modifying any of the attributes.

In the second step, route maps are applied to BGP neighbors. One route is applied to the AS 234 neighbor and another to the AS 567. Both are applied in the inbound direction.

The load-balancing configuration for incoming traffic is not shown in the example. The customer should split its address space into two halves and advertise one subnet to one ISP and other one to the other ISP. Advertisement could be done by redistribution from an interior routing protocol (IGP) into BGP on the edge routes or by advertising the subnets through an Internal Border Gateway Protocol (IBGP) session from the customer's inside routers.

Customer-Implemented Policies

Verification



In the figure, the **show ip bgp** command is used to verify load balancing for outgoing traffic. Two networks are advertised from the Internet, 10.0.1.0/24 and 172.16.1.0/24. The customer router receives routes from both ISPs, but it changes the local preference to 200 for one route that is received from each ISP, respectively. This results in the customer choosing AS 234 for traffic going to the 10.0.1.0/24 network and AS 567 for traffic going to the 172.16.1.0/24 network.

ISP-Aided Routing Policy

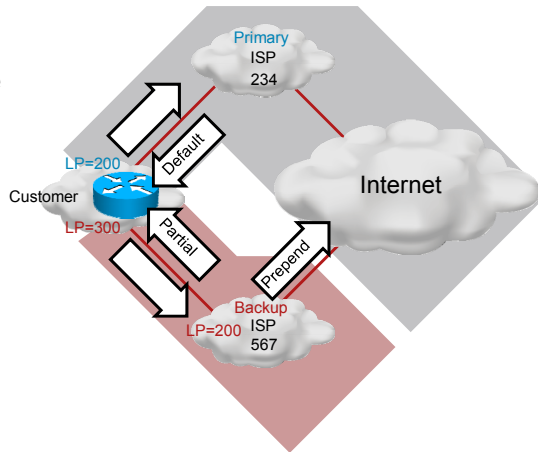
Primary and Backup with Direct Traffic

Manual configuration:

- Backup ISP sets higher local preference on customer routes.
- Backup ISP does prepending towards the Internet.

Issues:

- Per-customer configurations; difficult to maintain



Primary and backup scenarios, with direct traffic routed to the backup ISP, can also be optionally implemented on the ISP side. Routing requirements are the same as in the scenario without ISP-aided routing policy. The primary ISP sends only the default route, and the backup ISP sends the default route and local routes. The customer increases the local preference for the default route that is received from the primary ISP to 200 and sets local preference for local routes from the backup ISP to 300.

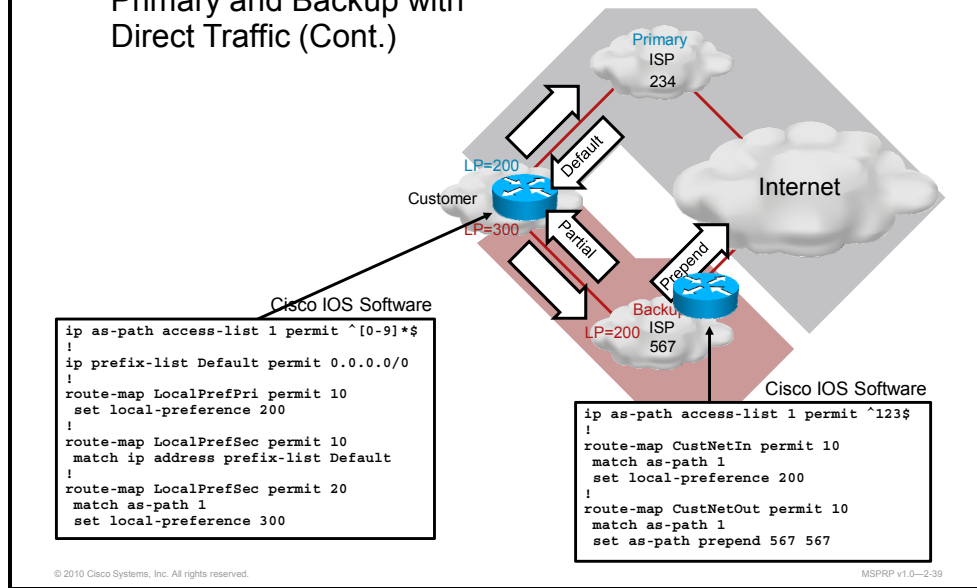
The difference from the scenario without ISP-aided routing policy is that the backup ISP router does the path manipulation for the incoming traffic, not the customer. The backup ISP will set the local preference of received customer routes to 200 on the ingress router. All traffic originating in the backup ISP, or traffic from the directly attached other customers, will take the direct path to the customer.

On the egress router connecting the backup ISP to the Internet, an AS-path prepending will be configured. The customer routes will be advertised from the backup ISP with the longer AS path, thus making them less desirable for traffic originating on the Internet. All traffic originating on the Internet and going to the customer will then take a path through the primary ISP, unless the connection through the primary fails.

However, this solution is not scalable. The backup ISP must identify customer routes and implement route manipulation separately for each customer.

ISP-Aided Routing Policy

Primary and Backup with Direct Traffic (Cont.)



The sample configuration will result in the customer router preferring the default route that is received from the primary ISP. Local preference for local routes that are received from the backup ISP is changed to 300. AS path prepending is not configured on the customer side, but on the backup ISP side instead. Local preference of the customer routes is also increased on the backup ISP to prefer the direct path for traffic originating in the AS 567 and autonomous systems connected directly to it. When the backup ISP advertises customer routes to the Internet on the egress router, AS number 567 will be prepended twice.

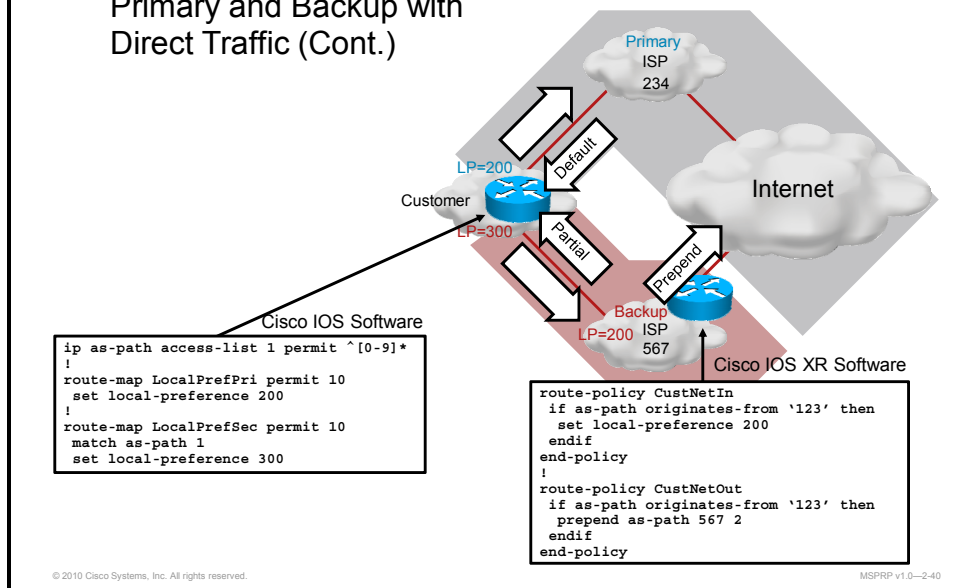
The configuration on the customer router does not differ from the example without ISP-aided routing policy, except the outbound policy is not configured on the router.

The configuration on the backup ISP Cisco IOS router is done in two steps. In the first step, two route maps are configured. The *CustNetIn* route map matches all routes originating in the AS 123 and sets a local preference for received routes to 200. The *CustNetOut* route map matches all routes originating in the AS 123 and sets the AS path prepending for customer routes when those routes are advertised further to the Internet.

In the second step, which is not shown in the configuration example, route maps should be applied to neighbors. *CustNetIn* should be applied to the BGP neighbor in the customer AS in the inbound direction, and *CustNetOut* should be applied to the BGP neighbor in the upstream AS in the outbound direction. Usually, this configuration is not done on the same router. The *CustNetIn* policy map would be configured on the ISP's router connecting to the customer, and the *CustNetOut* policy map would be configured on the ISP's router connecting to the Internet.

ISP-Aided Routing Policy

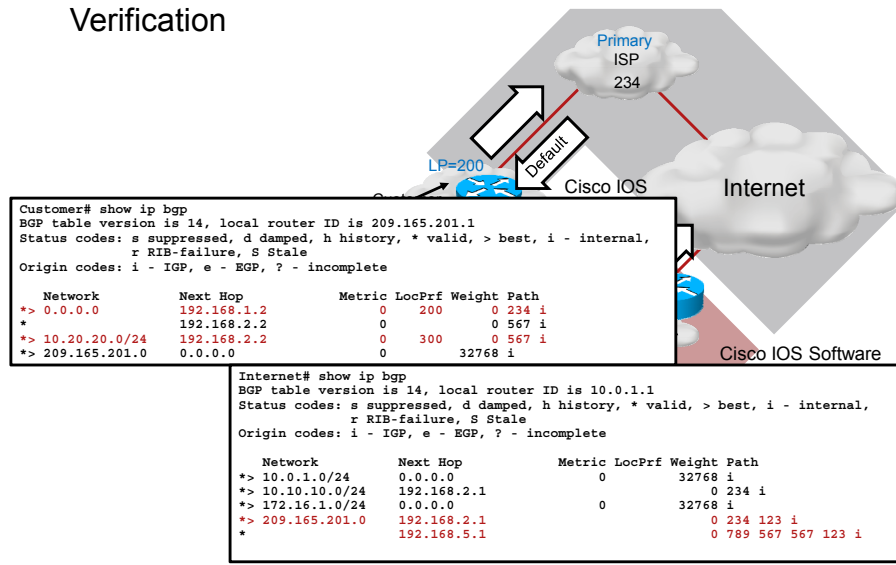
Primary and Backup with
Direct Traffic (Cont.)



The figure illustrates a sample configuration on a Cisco IOS XR router. Two route policies are configured that both match routes originating in the AS 567. The *CustNetIn* route policy sets the local preference to 200 for routes received from the customer. The *CustNetOut* route policy prepends the AS 123 two times to customer routes that are advertised further to the Internet. The *CustNetIn* route policy should be applied to the BGP neighbor in the customer ISP in the inbound direction. The *CustNetOut* should be applied to the BGP neighbor on the Internet in the outbound direction.

ISP-Aided Routing Policy

Verification



In the figure, the **show ip bgp** command is used to verify path manipulation on the customer and Internet routers. To the customer router, two default routes are advertised. Local preference of the one received from the primary ISP is set to 200 to prefer it. Local routes from the backup ISP are also advertised, and local preference of those routes is set to 300.

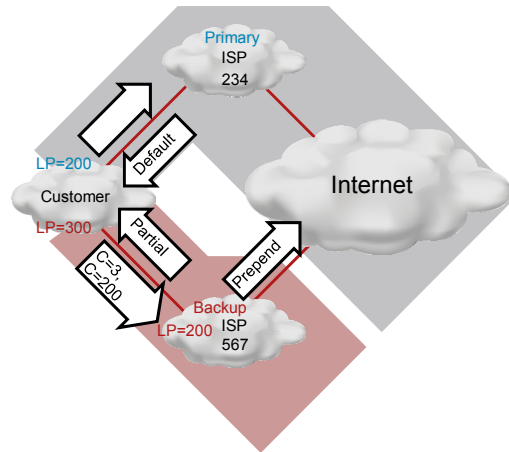
To the Internet router, both ISPs advertise the customer route. Because an AS-path prepending is configured on the backup ISP router connecting to the Internet, the one advertised by the primary ISP has the shorter AS path and is preferred.

ISP-Aided Routing Policy

Primary Backup with Direct Traffic

Signaling:

- Customer signals “backup” service using BGP community.
- Backup ISP sets higher local preference on customer routes based on BGP community.
- Backup ISP does prepending towards the Internet based on BGP community.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-42

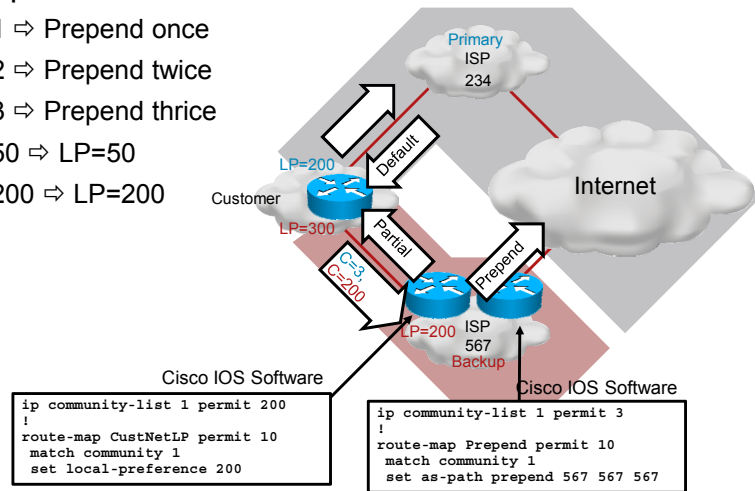
The figure illustrates an example when ISP aids with routing policy for direct traffic. In this example, the customer signals backup service using BGP communities, which the backup ISP defines. These BGP communities are attached to the customer routes. The backup ISP then sets higher local preference for customer routes and does AS path prepending of routes that are advertised towards the Internet based on matched BGP communities.

ISP-Aided Routing Policy

Primary and Backup with Direct Traffic (Cont.)

Backup ISP's services:

- C=1 ⇒ Prepend once
- C=2 ⇒ Prepend twice
- C=3 ⇒ Prepend thrice
- C=50 ⇒ LP=50
- C=200 ⇒ LP=200



A customer can signal needed backup services to the backup ISP using BGP communities, which are defined by the ISP in a similar manner as in the example. If the customer appends BGP community with value 1 to an advertised route, the ISP will prepend its own AS to the customer route when advertising it further to the Internet. If the customer appends the BGP community with the value 50 to an advertised route, the ISP will set local preference to 50 when advertising the customer route inside the AS 567.

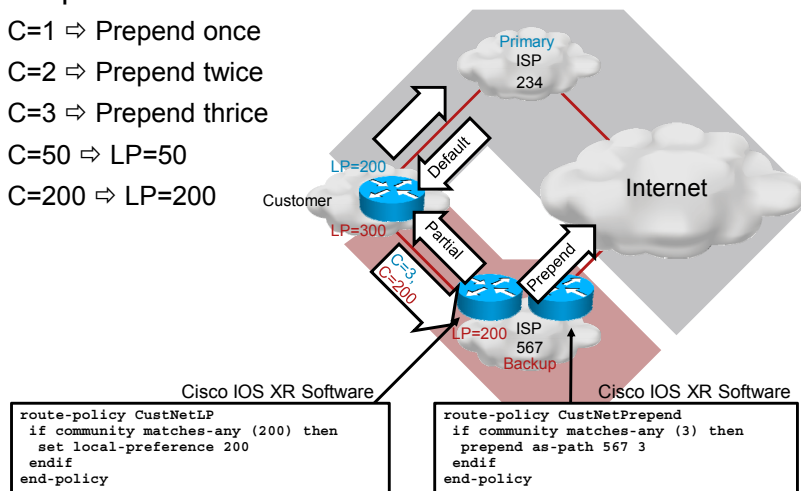
The configuration in the example will result in the backup ISP ingress router matching routes that are marked with BGP community 200 and setting local preference for those routes to 200 when advertising them to the internal ISP's routers. The configuration on the egress backup ISP's router will match BGP community 3 and prepend the AS 567 three times to routes advertised to the Internet.

ISP-Aided Routing Policy

Primary and Backup with Direct Traffic (Cont.)

Backup ISP's services:

- C=1 ⇒ Prepend once
- C=2 ⇒ Prepend twice
- C=3 ⇒ Prepend thrice
- C=50 ⇒ LP=50
- C=200 ⇒ LP=200



The figure shows the configuration example on the Cisco IOS XR routers. On the router connecting to the customer, the *CustNetLP* route policy is configured that matches routes with BGP community 200. The router then sets local preference 200 to routes advertised to the IBGP neighbor. On the router connecting to the Internet, the *CustNetPrepend* route policy is configured that matches routes with BGP community 3. It prepends AS number 567 three times to routes advertised to the Internet.

ISP-Aided Routing Policy

Verification

```
ISP# show ip bgp
BGP table version is 45, local router ID is 10.10.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop        Metric LocPrf Weight Path
*> 10.0.1.0/24     192.168.5.2          0   100      0 1 i
*> 10.20.20.0/24   0.0.0.0              0           32768 i
*> 10.10.10.0/24   192.168.5.2          0   100      0 1 567 i
*> 172.16.1.0/24   192.168.5.2          0   100      0 1 i
*> 209.165.201.0   192.168.3.1          0   200      0 123 i

ISP# show ip bgp 209.165.201.0
BGP routing table entry for 209.165.201.0/24, version 45
Paths: (1 available, best #1, table Default-IP-Routing-Table)
  Advertised to update-groups:
    2
  123
    192.168.1.1 from 192.168.1.1 (209.165.201.1)
      Origin IGP, metric 0, localpref 200, valid, external, best
      Community: 3 200
```

```
Internet# show ip bgp
BGP table version is 55, local router ID is 192.168.5.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop        Metric LocPrf Weight Path
*> 10.0.1.0/24     0.0.0.0              0           32768 i
*> 10.20.20.0/24   192.168.2.1          0           0 567 i
*> 172.16.1.0/24   0.0.0.0              0           32768 i
* 209.165.201.0    192.168.5.1          0   567 567 567 567 123 i
*>                 192.168.2.1          0   234 123 i
```

In the figure the **show ip bgp** and **show ip bgp prefix** are used to verify the path manipulation. On the ISP router connecting to the customer, the customer network (209.165.201.0/24) is present in the BGP table. Local preference of this route is set to 200. It can also be seen that two BGP communities (3 and 200) are appended to the customer route.

On the Internet router, both ISPs advertise the customer route. Because an AS path prepending is configured on the backup ISP router connecting to the Internet, the one advertised by the primary ISP has the shorter AS path and is preferred.

Route Filtering

This topic describes mechanisms and configurations on Cisco IOS and Cisco IOS XR platforms for implementing route filtering when advertising routes from one AS to another.

Route Filtering

- Route filtering is required to ensure compliance with peering agreements.
- Ideally, both parties will use filters.
- Filtering is combined with routing policy implementation.

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-247

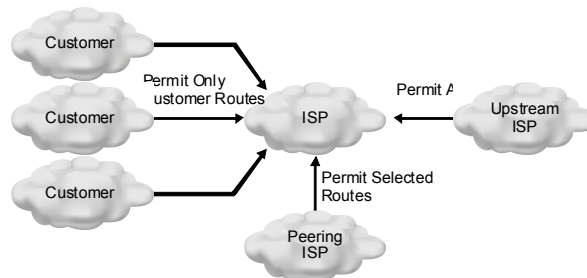
Route filtering is required on routers to ensure compliance with peering agreements. If the customer wants to receive only the default route, the ISP has to make sure that only default routes will be sent to the customer. Ideally, the ISP and the customer will both use filters (one in inbound, other in outbound direction) to make sure that filtering really is implemented.

Filtering is often combined with routing policies to achieve the desired path selection.

Route Filtering

Inbound

- Typical filters:
 - AS path filters
 - Prefix lists (more difficult to maintain)



© 2010 Cisco Systems, Inc. All rights reserved.

MS PRP v1.0-248

Inbound filtering on the ISP side should be configured for different parties:

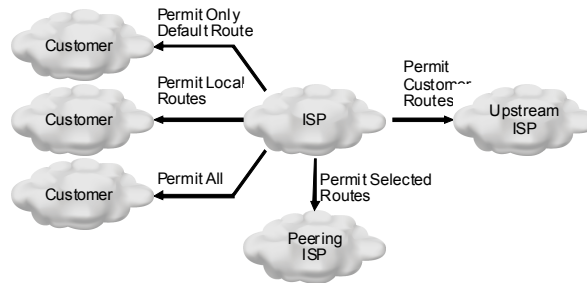
- **Customers:** Customer routes should only be permitted in the ISP AS.
- **Peering ISP:** ISPs should agree on which routes will be advertised from one to another.
- **Upstream ISP:** All routes should be permitted from the upstream ISP.

Typical filters that are used are AS-path filters and prefix lists that can be applied directly to a BGP neighbor configuration in the inbound direction. Recall that filtering is often used with route policy implementation, in which some BGP attributes must also be set (local preference). In this case, route maps (Cisco IOS Software) or route policies (Cisco IOS XR Software) must be used. However, AS-path filters and prefix lists can still be used to identify routes in a route map or policy map.

Route Filtering

Outbound

- Typical filters:
 - AS path filters
 - Prefix lists (more difficult to maintain)



© 2010 Cisco Systems, Inc. All rights reserved.

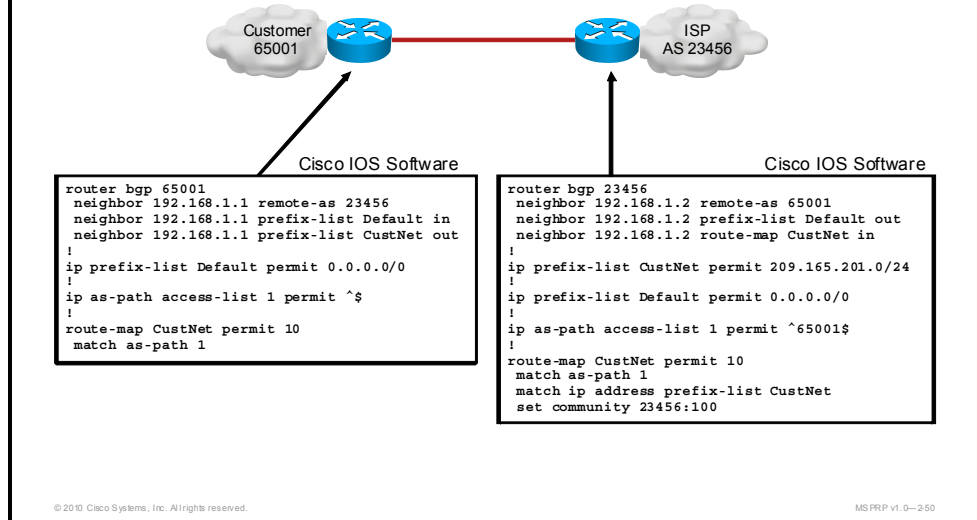
MSRP v1.0-249

Also, outbound filtering on the ISP side should be configured for different parties:

- **Customers:** The ISP can, based on customer requirements, advertise either default route only, default route and local routes, or the complete Internet routing table to a customer.
- **Peering ISP:** ISPs should agree on which routes will be advertised from one to another.
- **Upstream ISP:** All customer routes should be advertised to the upstream ISP.

Typical filters that are used are AS path filters and prefix lists that can be applied directly to a BGP neighbor configuration in the outbound direction. Recall that filtering is often used together with route policy implementation, when some BGP attributes must also be set (AS path prepending). In this case, route maps (Cisco IOS Software) or route policies (Cisco IOS XR Software) must be used. However, AS path filters and prefix lists can still be used to identify routes in a route map or route policy.

Example: Customer Filter

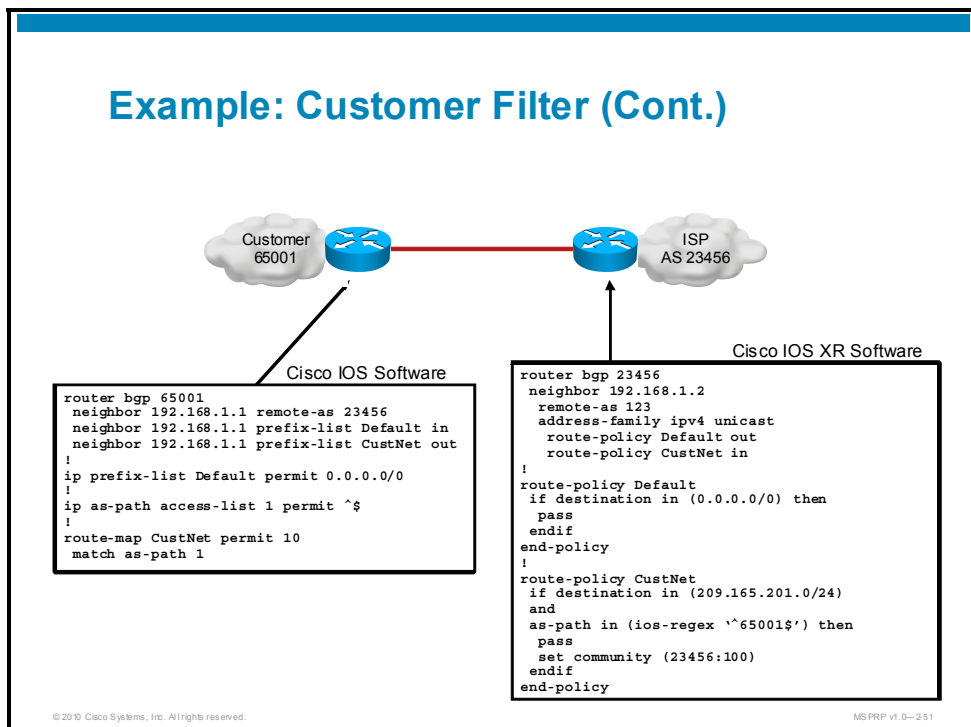


The example shows the configuration for route filtering on the customer side and the ISP side in inbound and outbound directions.

On the customer router, the *CustNet* route map is configured that matches all routes originating in the AS 65001 (locally originating routes will have an empty AS path). This route map is applied to the EBGP neighbor in the ISP AS in the outbound direction. For inbound filtering, the *Default* prefix list is configured that permits only the default route. This prefix list is applied to the EBGP neighbor in the ISP AS in the inbound direction.

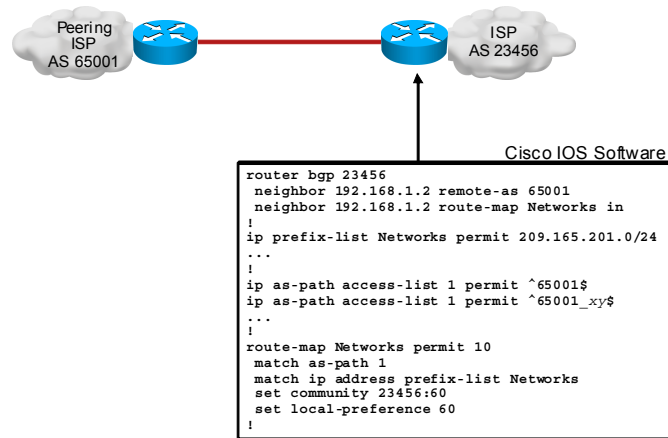
On the ISP router, the *Default* prefix list is configured that permits only the default route to be advertised to the customer. This prefix list is applied to the EBGP neighbor in the customer AS in the outbound direction. To filter incoming routes, the *CustNet* route map is configured that matches customer prefixes with the prefix list and all other routes originating in the AS 65001. This route map is applied to the EBGP neighbor in the customer AS in the inbound direction. Additionally, ISP tags customer routes with BGP community 23456:100.

Example: Customer Filter (Cont.)



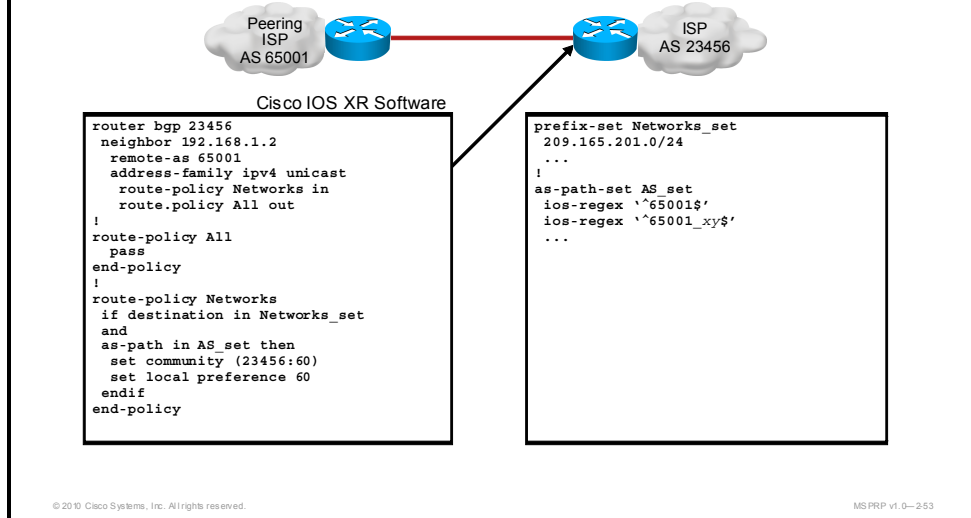
The figure shows the configuration on the ISP Cisco IOS XR router. Two route policies are configured. The *Default* route policy permits only the default route and is applied to the EBGP neighbor in the outbound direction. The *CustNet* route policy permits only customer prefixes and routes originating in the AS 65001. The route policy also tags customer routes with the BGP community. This route policy is applied to the BGP neighbor in the inbound direction.

Example: Peering Filters



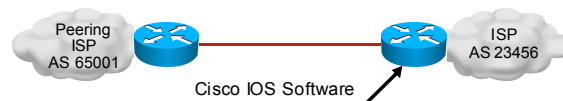
The figure illustrates an example of a peering filter, in which all outgoing updates are permitted (no route map is configured in the outbound direction). In inbound direction peering, ISP routes—and routes originating in AS 65001 and in directly connected autonomous systems — are permitted. Routes that are received from peering autonomous systems are also tagged with BGP community 23456:60. Local preference is set to 60.

Example: Peering Filters (Cont.)



The figure shows the configuration of a peering filter on the Cisco IOS XR router. Route policy *All* is configured to permit all routes to be advertised to the peering ISP. The *Networks* route policy is configured to receive only routes that are specified in the *Network_set* prefix set and *AS_set* AS path set (peering ISP routes and routes originating in the AS 65001 and in directly connected autonomous systems) from the peering ISP. Those networks are marked with BGP community 23456:60. Local preference for those networks is set to 60.

Example: Upstream Filters



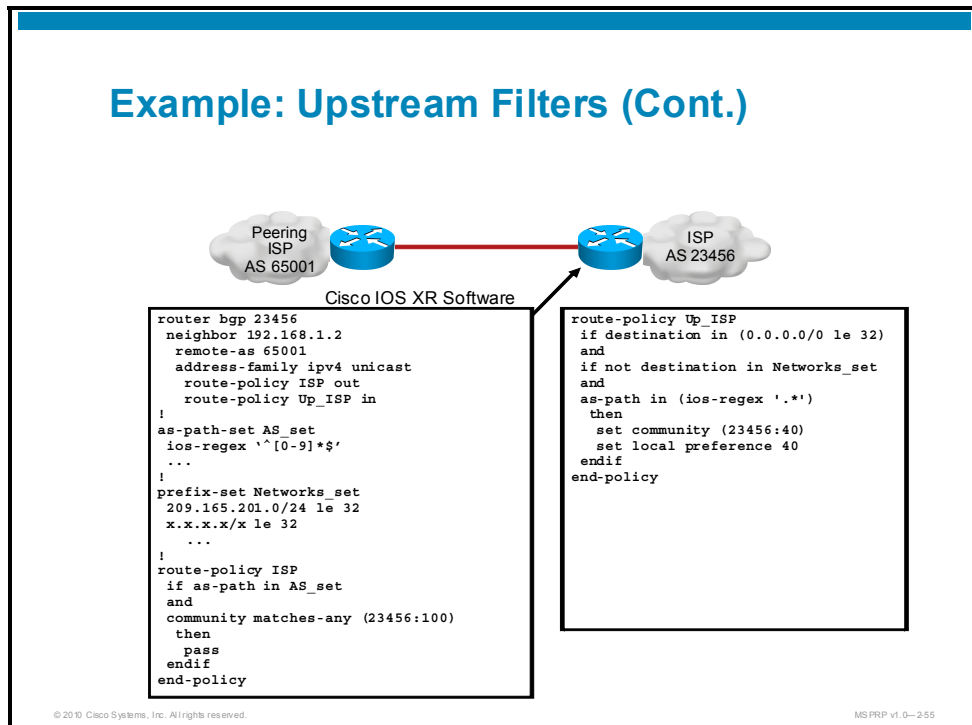
```
router bgp 23456
 neighbor 192.168.1.2 remote-as 65001
 neighbor 192.168.1.2 route-map Up_ISP in
 neighbor 192.168.1.2 route-map ISP out
 !
 ip prefix-list Up_ISP_Net deny 209.165.201.0/24 le 32
 ip prefix-list Up_ISP_Net deny x.x.x.x/x le 32
 ...
 ip prefix-list Up_ISP_Net permit 0.0.0.0/0 le 32
 !
 ip community-list 1 permit 23456:100
 !
 ip as-path access-list 1 permit ^[0-9]*$
 !
 ip as-path access-list 2 permit .*
 !
 route-map Up_ISP permit 10
 match ip address prefix-list Up_ISP_Net
 match as-path 2
 set community 23456:40
 set local-preference 40
 !
 route-map ISP permit 10
 match community 1
 match as-path 1
 !
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-254

The following example of the upstream filter will filter out local routes from the upstream ISP. The default route and routes from other autonomous systems will be permitted to the AS 23456. Advertised routes will also be tagged with BGP community 23456:40; local preference will be set to 40. Outbound filtering of routes that are sent to the upstream ISP will permit tagged customer routes with BGP community 23456:100, routes originating in AS 23456, and other directly connected autonomous systems.

Example: Upstream Filters (Cont.)



The figure presents a sample configuration to configure upstream filters on a Cisco IOS XR router. Two route policies are configured. The *ISP* route policy matches any customer route that is tagged with BGP community 23456:100, routes originating in the AS 23456, and in autonomous systems directly attached to it. Those routes are advertised to the upstream ISP. The *Up_ISP* route policy filters incoming updates from the upstream ISP and permits the default route and all other routes except local routes (defined in the *Networks_set* prefix set). This route policy also tags received routes with BGP community 23456:40 and sets the local preference to 40.

Summary

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

Summary

- A multihomed connection using BGP offers the highest level of resiliency and redundancy.
- A customer that is multihomed to two ISPs must advertise its own networks to both ISPs and take care not to advertise any other networks.
- Each ISP can, based on customer requirements, advertise only the default route, default route and local routes, or full Internet routing table to a customer.
- Routing policies should be implemented to meet the customer's required path selection. On Cisco IOS routers, route maps can be used. On Cisco IOS XR routers, route policies can be used.
- Routing filters should be implemented to make sure that only required networks are advertised from one party to another.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-256

Lesson Self-Check

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

- Q1) Which routing protocol can be used in the multihomed connectivity scenario? (Source: Multihomed Customer Requirements)
- A) OSPF
 - B) IS-IS
 - C) BGP
 - D) RIPv2
- Q2) Which three pieces of information should a multihomed customer provide when connecting to an ISP? (Choose three.) (Source: Multihomed Customer Requirements)
- A) routing table requirements
 - B) AS number
 - C) address space
 - D) access link addressing
- Q3) Which two are the default behaviors of multihomed connectivity? (Choose two.) (Source: Routing Policies)
- A) high availability
 - B) symmetric routing
 - C) unpredictable load balancing
- Q4) Which two policies can be implemented in a multihomed scenario? (Choose two.) (Source: Routing Policies)
- A) one ISP primary, other backup
 - B) primary and backup ISPs with direct access to primary ISP's destinations
 - C) load balancing
 - D) symmetric routing
- Q5) How can a customer influence path selection for outgoing traffic? (Source: Routing Policies)
- A) by changing MED for routes that are received from the ISP
 - B) by changing the local preference for routes that are received from the ISP
 - C) by prepending an AS number to routes received from the ISP
 - D) by changing the weight for routes that are received from the ISP
- Q6) How can a customer influence path selection for incoming traffic? (Source: Routing Policies)
- A) by changing MED for routes that are sent to the ISP
 - B) by changing local preference for routes that are sent to the ISP
 - C) by prepending an AS number to routes sent to the ISP
 - D) by changing the weight for routes that are sent to the ISP
- Q7) Which tools can be used to implement route filtering on a Cisco IOS router? (Source: Route Filtering)
- A) route maps
 - B) route policies
 - C) AS path filters
 - D) prefix lists

Lesson Self-Check Answer Key

- Q1) C
- Q2) A, B, D
- Q3) A, C
- Q4) A, C
- Q5) B
- Q6) C
- Q7) A

Modifying Routing Paths

Overview

In some cases, the default selection of routing paths in a network does not satisfy your requirements. Changes are required with internal routing protocols, such as Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS), and with the external routing protocol, which is Border Gateway Protocol (BGP). In order to efficiently modify routing paths, you should understand the path selection process for each routing protocol and how to implement and configure changes in the path selection.

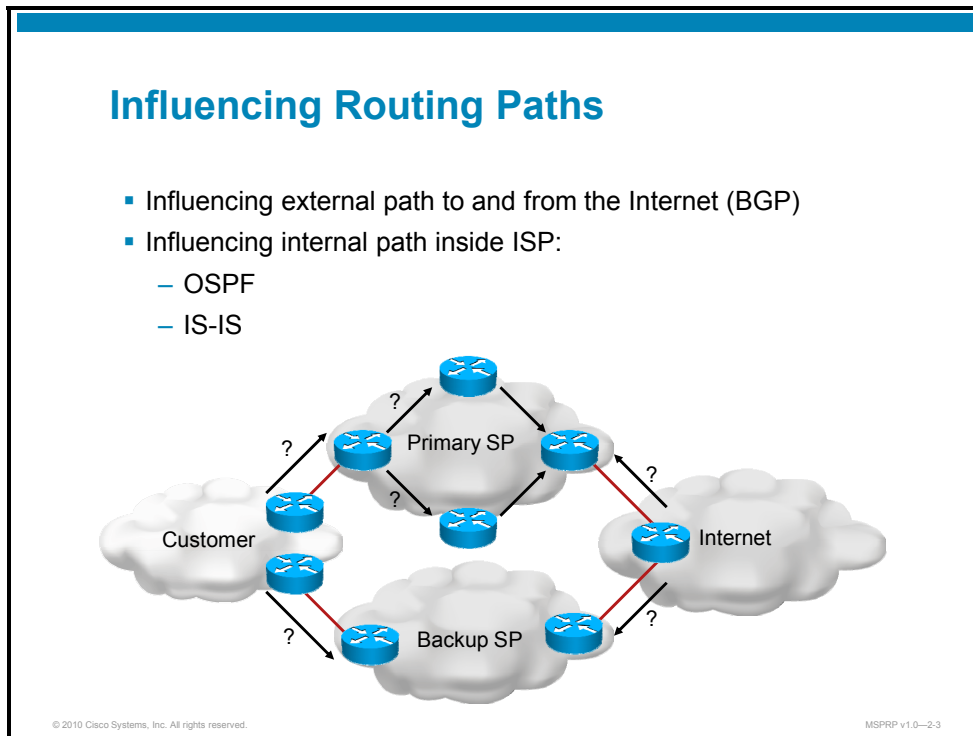
Objectives

Upon completing this lesson, you will be able to modify the routing path selection for internal and external routing protocols. This ability includes being able to meet these objectives:

- Influence internal routing using OSPF
- Influence internal routing using IS-IS
- Influence external routing using BGP

Influencing Internal Routing Using OSPF

This topic describes default OSPF path selection and default metric calculation. Mechanisms to influence path selection and OSPF load balancing are also described.



In some cases, default path selection results in suboptimal routing. As shown in the figure, you—as network administrators—can influence path selection in several places.

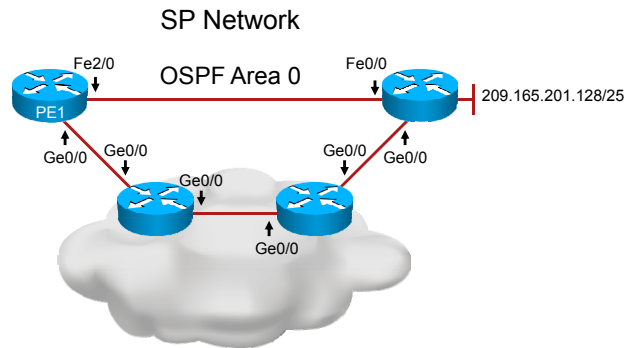
- The path that is taken from the customer network to the Internet through one of the service providers (external routing).
- The path that is taken from the Internet to the customer through one of the service providers (external routing).
- The path that is taken inside the service provider network (internal routing).

However, the first two path selection scenarios are valid only when a customer is multihomed to two service providers. In this case, BGP is used to advertise routes from a customer to the service providers and vice versa. Therefore, to influence path selection from a customer to the Internet and vice versa, you must change BGP attributes for paths that are learned from both service providers.

Usually the default path selection inside the autonomous system (AS), where internal gateway protocols are used, does not reflect your requirements. In this case, you can influence path selection by changing the metrics that the internal routing protocols (OSPF and IS-IS) use.

Default OSPF Path Selection

- OSPF uses cost as a metric.
- OSPF calculates path cost by adding all link costs along a path.
- Cost is, by default, calculated as the inverse of interface bandwidth:
 - $\text{Cost} = \text{BW}_{\text{ref}}/\text{BW}$, where $\text{BW}_{\text{ref}} = 108 \text{ b/s}$ by default
- In this case, worse path will be used.



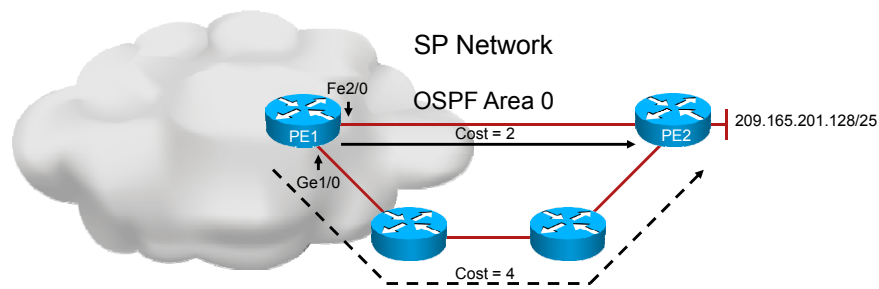
OSPF uses cost as a metric. The cost is an indication of the overhead to send packets over a specific path. Path cost is calculated by summing all link (interface) costs along a path. Interface cost is computed automatically for each interface that is enabled for OSPF using the following formula:

$\text{Cost} = \text{BW}_{\text{ref}}/\text{BW}$, where BW_{ref} is 10^8 b/s , and BW is a bandwidth of an interface that is configured by the **bandwidth** interface command.

OSPF cost is a 16-b number, which ranges from 1 to 65,535, in which paths with lower costs are more desirable. For example, an Ethernet link gets a cost of 10, and a Fast Ethernet link gets a cost of 1. The problem arises with high-bandwidth links (more than 100 Mb/s), in which the default cost calculation does not work anymore. All links faster than 100 Mb/s would be assigned the same cost, which is 1. Without OSPF costs that are set manually on each interface, the chosen path does not necessarily reflect the actual best path through a network.

Note Note that the OSPF metric is relevant only for outbound traffic.

Default OSPF Path Selection (Cont.)



```
PE1#show ip route
<...output omitted...>
Gateway of last resort is not set

 209.165.201.0/25 is subnetted, 1 subnets
O    209.165.201.128 [110/2] via 10.0.0.2, 00:00:41, FastEthernet2/0
 10.0.0.0/30 is subnetted, 4 subnets
O    10.0.0.8 [110/2] via 10.0.0.14, 00:04:04, GigabitEthernet1/0
C    10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0--2.5

As shown in the example, two paths exist from the PE1 router to the 209.165.201.128/25 network. Although the lower path should be preferred since it uses links with Gigabit Ethernet interfaces, the upper path will be used because the cost calculated by the OSPF is lower.

The cost to reach the 209.165.201.128/25 network from the PE2 router is 1 because the network is directly connected. The cost to reach the same network from the PE1 router via the upper path is 2 because the cost to reach PE2 from the PE1 is 1.

The cost to reach the same network from the PE1 router via the lower path is 4 because each link is assigned the cost of 1. Although Gigabit Ethernet interfaces should be preferred instead of Fast Ethernet, they all have the same cost because of the default cost calculation.

In this case, set OSPF costs manually for packets to take the lower, faster path. In general, change the OSPF cost when you use interfaces faster than 100 Mb/s.

Note	Recall that the OSPF cost of links with interfaces equal to or faster than the Fast Ethernet is 1 by default.

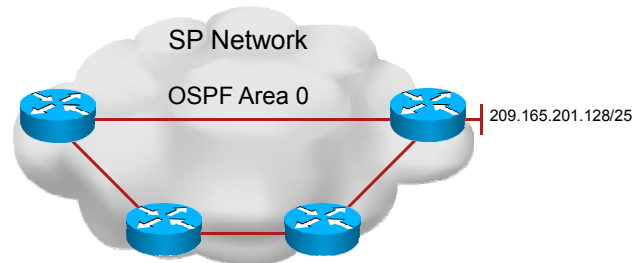
Influencing OSPF Path Selection

- Change the **OSPF cost** on a specific interface.
- Change the **reference bandwidth** used to compute default OSPF costs.

Standard Change

Minor Change

Major Change



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2/6

You can influence OSPF path selection in two ways.

- You can override the automatically calculated OSPF cost by setting the *cost manually on each interface* as needed.
- You can change the *reference bandwidth*, which is used to compute OSPF cost. In this case you must be sure that you change the reference bandwidth on all routers in a network to have the same cost calculation over the whole network.

If you increase the reference bandwidth, the computed OSPF cost will reflect the actual overhead of an interface, including interfaces faster than 100 Mb/s.

When changing the OSPF cost for a selected link, it is important to remember that the costs should be configured equally on both ends of the link (on the adjacent routers). A change would likely result in the rerouting of traffic. If the change had been thoroughly planned, it should not have any negative effect on the running traffic; the rerouting will happen without losing a single packet. These types of changes (if covered by service provider standards and procedures) can therefore be regarded as standard—or minor—changes.

The impact of changing the reference bandwidth, unlike changing a cost for a single link, is much greater. The reference bandwidth influences how OSPF calculates the cost for links that have not been specifically configured with a static cost. To assure that all routers have the same view of the topology and the costs of individual links, it is important to have **all** routers that are configured with the same reference bandwidth. This change is rarely implemented today because it is typically part of a service provider network initial design; all service providers already use links that are faster than 100 Mb/s. The change is most commonly implemented when a service provider first introduces a new link that is of a greater bandwidth than the links previously used. For example, when a service provider introduced a 10-Gb/s link, it would have to modify the reference bandwidth on all routers, unless the reference bandwidth was already set to support this speed. In the future, service providers may need to change the reference bandwidth again when they introduce faster interfaces (for example, adding 40-Gb/s or 100-Gb/s links).

Influencing OSPF Path Selection

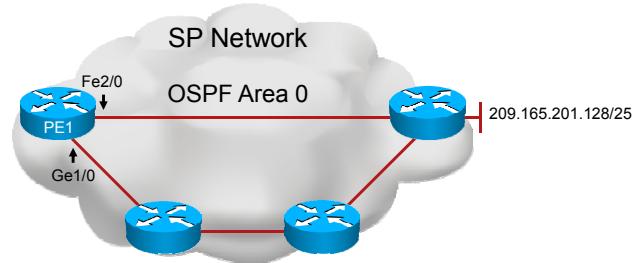
Cisco IOS Software

- Change the OSPF cost on the interface, or

```
interface FastEthernet2/0
ip ospf cost 10
interface GigabitEthernet1/0
ip ospf cost 1
```

- Change the OSPF reference bandwidth on all routers

```
router ospf 1
auto-cost reference-bandwidth 1000
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0--2.7

The sample configuration shows how to influence OSPF path selection in Cisco IOS Software.

The first option is to change the OSPF cost manually on a specific interface using the **ip ospf cost** command in the interface configuration mode. In the example, the cost for Fast Ethernet 2/0 has been changed to 10, and the cost for Gigabit Ethernet 1/0 has been changed to 1. This way, the PE1 router will prefer the path through the lower path.

The second option is to change the reference bandwidth using the **auto-cost reference bandwidth** command in the **router ospf** configuration mode. In the example, the reference bandwidth is set to 1000 Mb/s. Consequently, the cost of the Fast Ethernet link is changed to 10, and the cost of the Gigabit Ethernet is changed to 1. Thus the links are differentiated. In this case you have to make sure to change the reference bandwidth on all routers to have unified cost calculation over the network.

Influencing OSPF Path Selection

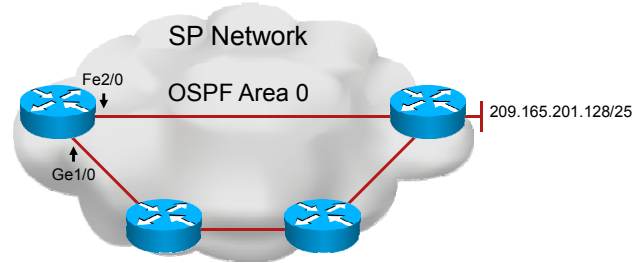
Cisco IOS XR Software

- Change the OSPF cost on the interface, or

```
router ospf 1
 area 0
  interface FastEthernet2/0
   cost 10
  interface GigabitEthernet1/0
   cost 1
```

- Change the OSPF reference bandwidth on all routers

```
router ospf 1
 auto-cost reference-bandwidth 1000
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-8

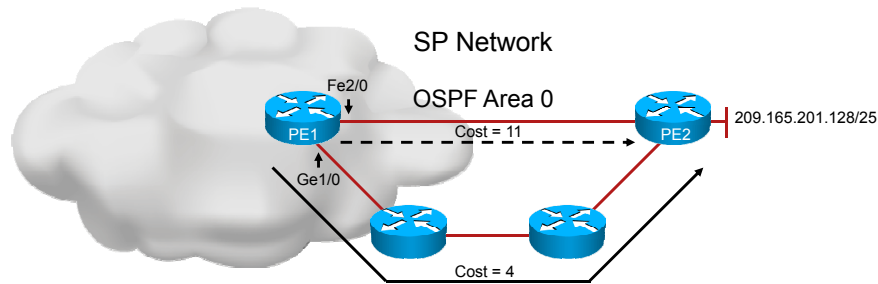
The sample configuration shows how to influence OSPF path selection in Cisco IOS XR Software.

The interface cost is changed in the **router ospf** configuration mode under the interface submode using the **cost** command.

The reference bandwidth is changed under the **router ospf** configuration mode using the **auto-cost reference bandwidth** command.

Changing OSPF Cost

Verification



```
PE1#show ip route
<...output omitted...>

Gateway of last resort is not set

  209.165.201.0/25 is subnetted, 1 subnets
O    209.165.201.128 [110/4] via 10.0.0.14, 00:02:59, GigabitEthernet1/0
  10.0.0.0/30 is subnetted, 4 subnets
O    10.0.0.8 [110/2] via 10.0.0.14, 00:04:04, GigabitEthernet1/0
C    10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2.0

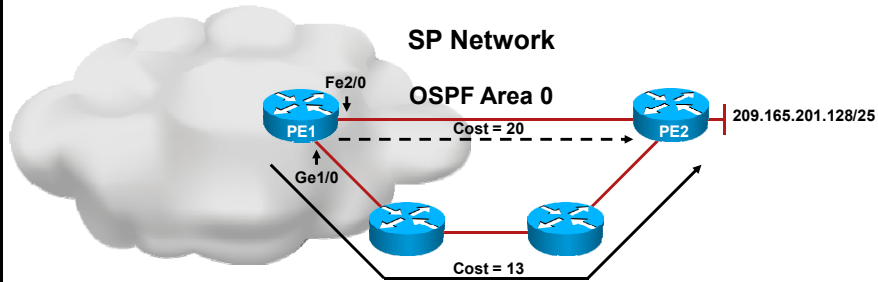
The sample printout shows the path selection after changing the OSPF cost on interfaces. The cost to reach the 209.165.201.128/25 network from the PE1 router over the upper path is now 11, because the cost of the Fast Ethernet 2/0 interface has been changed to 10. Cost from the PE2 router to reach the network is still 1 because the interface cost of that router has not been changed.

The cost of the lower path remained 4 because the cost of the Gigabit Ethernet 1/0 has been assigned a value of 1. Costs of links further to the destination remained the same (calculated automatically), because cost has not been changed on other routers.

This way, the lower path will be preferred because of the lower overall cost to the destination.

Changing OSPF Reference Bandwidth

Verification



```
PE1#show ip route
<...output omitted...>

Gateway of last resort is not set

  209.165.201.0/25 is subnetted, 1 subnets
O    209.165.201.128 [110/13] via 10.0.0.14, 00:00:08, GigabitEthernet1/0
  10.0.0.0/30 is subnetted, 4 subnets
O    10.0.0.8 [110/2] via 10.0.0.14, 00:02:16, GigabitEthernet1/0
C    10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-10

The sample printout shows the path selection after changing the reference bandwidth to 1000 Mb/s on all routers.

The cost to reach the destination network from the PE2 router is now 10, because the automatically calculated cost of Fast Ethernet interfaces is now 10. The overall cost to reach the destination network from the PE1 router over the upper path is 20.

The path cost to reach the destination network from the PE1 router over the lower path is now 13, because the calculated cost of Gigabit Ethernet interfaces is 1. The initial cost of 10, to reach the network from the PE2 router, is added to the cost of Gigabit Ethernet three times because of three links between the PE1 and PE2 routers.

- Up to 16 equal-cost paths can be used by OSPF for load balancing.
- Only 4 equal-cost paths are used by default.
- Per-destination load balancing (Cisco Express Forwarding):
 - Hash is computed from the source-destination pair of the packet.
 - Hash points to an adjacency table entry.



MSPRP v1.0—2-1

The number of paths that the load balancing uses is limited by the number of entries the routing protocol puts in the routing table. Four entries is the default for OSPF, but as many as 16 equal-cost paths can be used.

- **Per-destination (fast switching):** The router distributes the packets that are based on the destination address. Packets going to one destination will take one path, and packets going to another destination will take another path. If there are not many destinations, equal load across links is not guaranteed. However, in service provider networks in which packets are going to many different destinations, equal load across links will probably be achieved.
- **Per-packet (process switching):** The router sends one packet for one destination over the first path and the second packet for the same destination over the second path, and so on. Per-packet load balancing guarantees equal load across all links. However, there is potential that the packets may arrive at the destination out of order because differential delays may exist within the network. The second issue is that every packet will be processed in CPU.

2-276 Maintaining Cisco Service Provider Routing Protocols (MSPRP) v1.0

Cisco Express Forwarding has two main components:

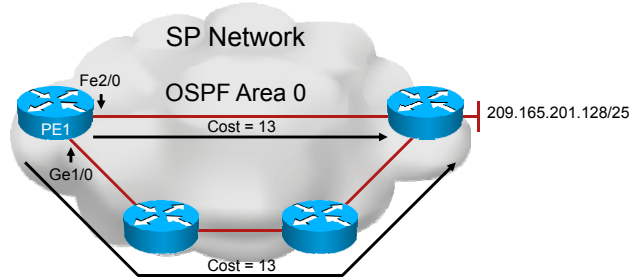
- **Forwarding Information Base (FIB):** It maintains a mirror image of the forwarding information that is contained in the routing table. If there are multiple equal-cost paths to one destination, all will be present.
- **Adjacency Tables:** Maintain Layer 2 next-hop addresses for all FIB entries.

For per-destination load balancing, a hash is computed out of the source and destination IP address. This hash points to one of the adjacency entries in the adjacency table, which provides the Layer 2 information to encapsulate a packet.

Note	Cisco Express Forwarding also supports per-packet load balancing. However, it is not recommended; the packets may arrive at the destination out of order because differential delay may exist within the network.

OSPF Load Balancing

Verification



```
PE1#show ip route
<...output omitted...>
Gateway of last resort is not set
  209.165.201.0/25 is subnetted, 1 subnets
O       209.165.201.128 [110/13] via 10.0.0.14, 00:00:12, GigabitEthernet1/0
       [110/13] via 10.0.0.2, 00:01:07, FastEthernet2/0
  10.0.0.0/30 is subnetted, 4 subnets
O       10.0.0.8 [110/2] via 10.0.0.14, 00:02:16, GigabitEthernet1/0
C       10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

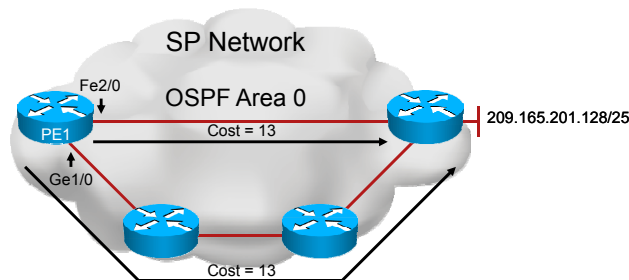
MSPRP v1.0—2-12

In the example, the OSPF cost has been changed on Fast Ethernet 2/0 and Gigabit Ethernet 1/0 interfaces to have equal-cost paths to the destination network.

In the printout presenting the **show ip route** command, two equal-cost paths with a cost of 13 are available to reach the 209.165.201.128/25 network.

OSPF Load Balancing

Verification (Cont.)



```
PE1#show ip cef 209.165.201.128
209.165.201.128/25, version 33, epoch 0, per-destination sharing
0 packets, 0 bytes
  via 10.0.0.14, GigabitEthernet1/0, 0 dependencies
    traffic share 1
    next hop 10.0.0.14, GigabitEthernet1/0
    valid adjacency
  via 10.0.0.2, FastEthernet2/0, 0 dependencies
    traffic share 1
    next hop 10.0.0.2, FastEthernet2/0
    valid adjacency
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-13

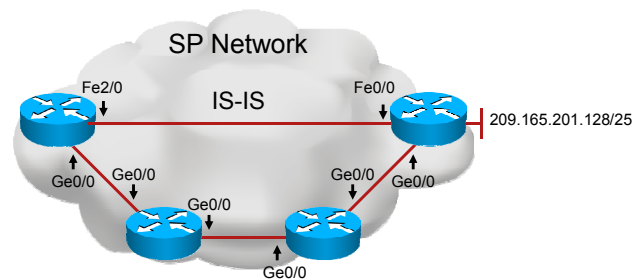
In the printout, the **show ip cef** command is used to verify Cisco Express Forwarding load balancing. Per-destination load balancing is configured (by default), and two paths are available to reach the 209.165.201.128/25 network.

Influencing Internal Routing Using IS-IS

This topic describes default IS-IS path selection and default metric calculation. Mechanisms to influence path selection and IS-IS load balancing are also described.

Default IS-IS Path Selection

- IS-IS calculates path metric by adding all link metrics along a path.
- IS-IS by default uses 6-b long metric (narrow-style metric).
- All IS-IS enabled interfaces are assigned a metric of 10.
- In this case, worse path will be used.



© 2010 Cisco Systems, Inc. All rights reserved.

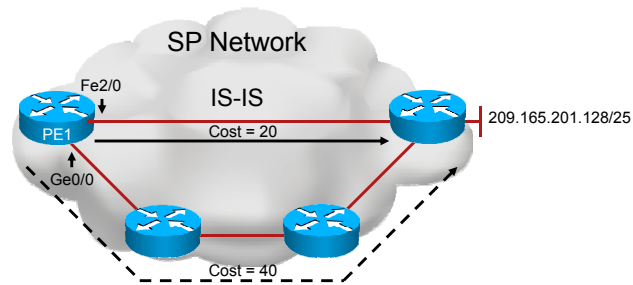
MSPRP v1.0-2-15

IS-IS calculates the path metric by adding all link metrics along a path. The path with the lower metric will be preferred.

IS-IS on a Cisco router does not take into account the bandwidth of an interface when it sets its metric. By default, all IS-IS enabled interfaces will have a metric of 10. In a network with different links, this default assignment could result in suboptimal routing.

By default, the IS-IS metric uses a 6-b long number, which ranges from 1 to 63 (narrow-style metric).

Default IS-IS Path Selection (Cont.)



```
PE1#show ip route
<...output omitted...>

Gateway of last resort is not set

  209.165.201.0/25 is subnetted, 1 subnets
i L1  209.165.201.128 [115/20] via 10.0.0.2, FastEthernet2/0
  10.0.0.0/30 is subnetted, 4 subnets
i L1  10.0.0.8 [115/20] via 10.0.0.14, GigabitEthernet1/0
C     10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-16

As shown in the example, two paths exist from the PE1 router to the 209.165.201.128/25 network. Although the lower path should be preferred since it uses links with Gigabit Ethernet interfaces, the upper path will be used because the metric of that path is lower. The metric of the upper path is 20 because there are two segments, each with a metric of 10. The metric of the lower path is 40 because there are four segments, each with a metric of 10.

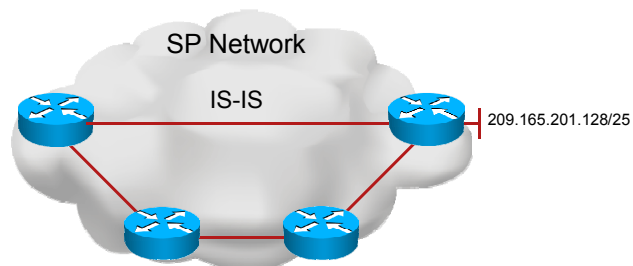
Influencing IS-IS Path Selection

- Change the **IS-IS metric** on a specific interface.
- Change **metric globally** for all interfaces.
- Use **wide-style** metric (24-b long).

Standard Change

Minor Change

Major Change



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-17

You can influence the IS-IS path selection only by changing the metric on a specific interface. You can also change the IS-IS metric globally for all interfaces (not valid for Cisco IOS XR Software).

Optionally, you can also configure the IS-IS to use a wide-style metric, which is 24-b long and ranges from 1 to 16777214. This way, you can set the metrics of links in a network in a more granular fashion.

Influencing IS-IS Path Selection

Cisco IOS Software

- Change the IS-IS metric on the interface

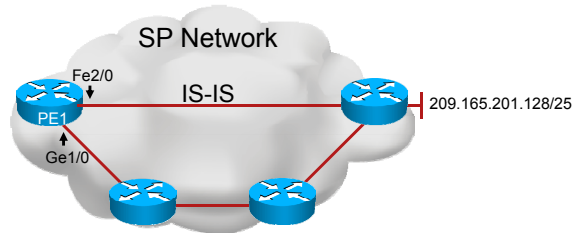
```
interface FastEthernet2/0
isis metric 30
interface GigEthernet1/0
isis metric 5
```

- Change the IS-IS metric globally (optional)

```
router isis
metric 30
```

- Change metric style (optional)

```
router isis
metric-style wide
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-18

The example shows how to change the IS-IS metric in Cisco IOS Software. The metric on the Fast Ethernet 2/0 has been changed to 30, and the metric on the Gigabit Ethernet 1/0 has been changed to 5 using the **isis metric** command in the interface configuration mode.

Optionally, you can also change the IS-IS metric globally for all interfaces using the **metric** command in the **router isis** configuration mode. You can also configure IS-IS to use wide-style metrics (24-b long) for more granular control of path selection using the **metric-style wide** command in **router isis** configuration mode.

Influencing IS-IS Path Selection

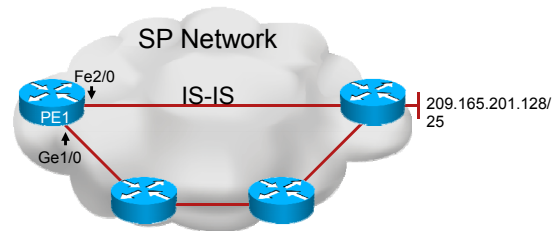
Cisco IOS XR Software

- Change the IS-IS metric on the interface

```
router isis 1
interface FastEthernet2/0
  address-family ipv4 unicast
  metric 30
interface GigaEthernet1/0
  address-family ipv4 unicast
  metric 5
```

- Change metric style (optional)

```
router isis 1
  address-family ipv4 unicast
  metric-style wide
```



The example shows how to change the IS-IS metric in Cisco IOS XR Software. The metric on the Fast Ethernet 2/0 has been changed to 30. The metric on the Gigabit Ethernet 1/0 has been changed to 5 using the **metric** command in the **router isis** configuration mode under the interface submenu.

Optionally, you can also configure the IS-IS to use wide-style metric using the **metric-style wide** command in the **router isis** configuration mode.

Influencing IS-IS Path Selection

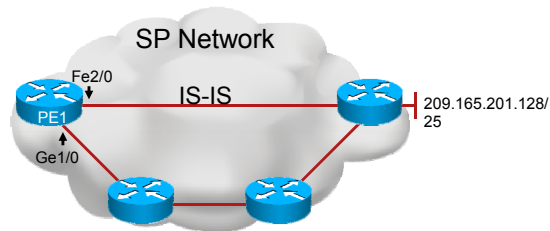
Cisco IOS XR Software

- Change the IS-IS metric on the interface

```
router isis 1
interface FastEthernet2/0
 address-family ipv4 unicast
  metric 30
interface GigaEthernet1/0
 address-family ipv4 unicast
  metric 5
```

- Change metric style (optional)

```
router isis 1
 address-family ipv4 unicast
  metric-style wide
```



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-19

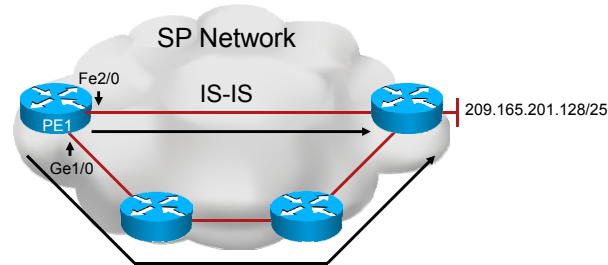
The sample printout shows the path selection after changing the IS-IS metric. The metric of the Fast Ethernet 2/0 interface has been changed to 30. The metric to reach the 209.165.201.128/25 network from the PE1 router over the upper path is now 40. (One unchanged interface adds 10 to the overall metric.)

The metric to reach the same network over the lower path is now 35 because the metric of the Gigabit Ethernet has been changed to 5. (Three unchanged interfaces add 30 to the overall metric.)

The best path is over the lower link now because of the lower metric.

IS-IS Load Balancing

- Up to 16 equal-cost paths can be used by IS-IS for load balancing.
- 4 equal-cost paths are used by default.
- Per-destination load balancing (Cisco Express Forwarding):
 - Hash is computed from the source/destination pair of the packet.
 - Hash points to an adjacency table entry.



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0-2-21

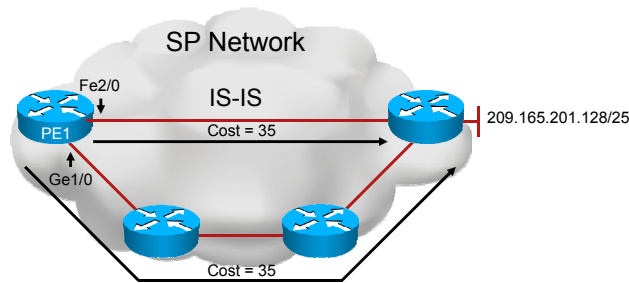
In cases where the router receives and installs multiple paths with the same administrative distance and cost to a destination, load balancing can occur. IS-IS supports only equal-cost load balancing, which means that all paths must have the same cost to reach a destination.

The number of entries the routing protocol puts in the routing table limits the number of paths that the load balancing uses. Four entries is the default for IS-IS, but as many as 16 equal-cost paths can be used.

A routing protocol (IS-IS, in this case) builds a routing table with multiple entries to a destination, and Cisco Express Forwarding load-balances packets across multiple paths. Load balancing is per-destination by default.

IS-IS Load Balancing

Verification



```
PE1#show ip route
<...output omitted...>

Gateway of last resort is not set

209.165.201.0/25 is subnetted, 1 subnets
i L1 209.165.201.128 [115/35] via 10.0.0.14, GigabitEthernet1/0
    [115/35] via 10.0.0.2, FastEthernet2/0

10.0.0.0/30 is subnetted, 4 subnets
i L1 10.0.0.8 [115/15] via 10.0.0.14, GigabitEthernet1/0
C 10.0.0.12 is directly connected, GigabitEthernet1/0
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

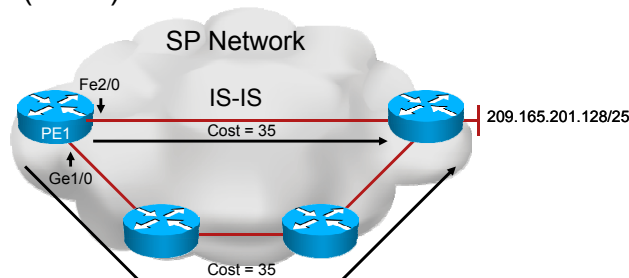
MSRP v1.0—2-22

In the example, the IS-IS metric has been changed on Fast Ethernet 2/0 and Gigabit Ethernet 1/0 interfaces to have equal-cost paths to the destination network.

In the printout presenting the **show ip route** command, two equal-cost paths with a metric of 35 are available to reach the 209.165.201.128/25 network.

IS-IS Load Balancing

Verification (Cont.)



```
PE1#show ip cef 209.165.201.128
209.165.201.128/25, version 45, epoch 0, per-destination sharing
0 packets, 0 bytes
via 10.0.0.14, GigabitEthernet1/0, 0 dependencies
  traffic share 1
  next hop 10.0.0.14, GigabitEthernet1/0
  valid adjacency
via 10.0.0.2, FastEthernet2/0, 0 dependencies
  traffic share 1
  next hop 10.0.0.2, FastEthernet2/0
  valid adjacency
<...output omitted...>
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0—2-23

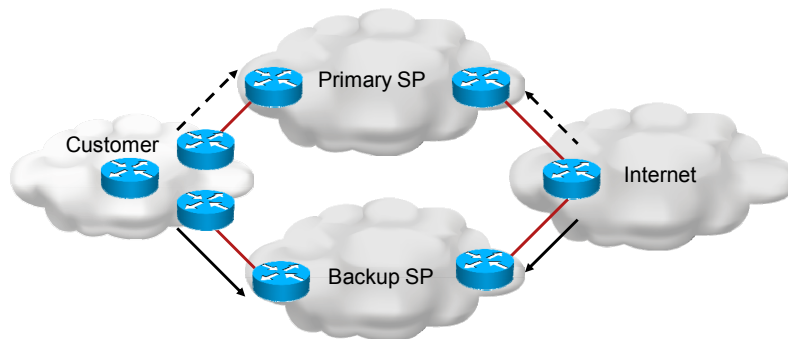
In the printout, the **show ip cef** command is used to verify Cisco Express Forwarding load balancing. Per-destination load balancing is configured (by default) and two paths are available to reach the 209.165.201.128/25 network.

Influencing External Routing Using BGP

This topic describes default BGP path selection. Mechanisms and configuration are described to influence BGP path selection.

Default BGP Path Selection

- BGP uses rich metrics (BGP attributes).
- Default behavior without path manipulation:
 - Path with the shortest AS path will be used.
 - Next criteria is path age.



When a customer is multihomed to two service providers, the BGP is used as the routing protocol between the customer and the service providers. BGP uses rich metrics, which are called BGP attributes. If no path manipulation is configured, the most common reason for path selection is the preference for the shortest AS path. If the AS path length of two paths is the same, the next criteria, path age, is used.

Three different scenarios are possible when multihoming using BGP.

- Each ISP passes only the *default route* to the customer.
- Each ISP passes the *default route and local routes* (routes belonging to the ISP and customers that are connected to this ISP).
- Each ISP passes a complete Internet routing table.

In any case, the customer might not be satisfied with the default BGP path selection.

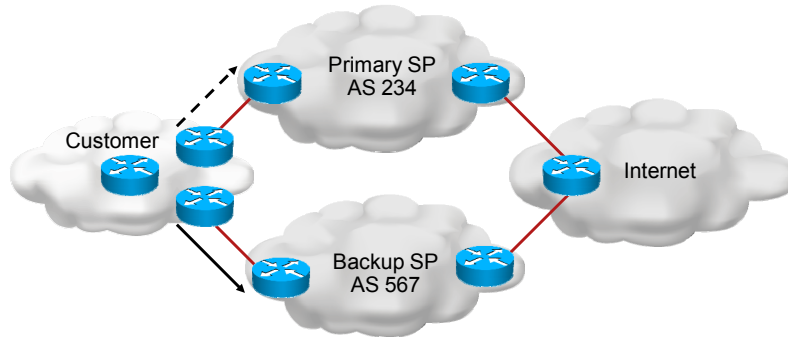
In the first scenario, the customer receives two default routes. Without any path manipulation, both default routes have the same BGP attributes (default values), and the AS path is of the same length. In this case, the oldest path (which is probably more stable) is preferred.

In the second scenario, the problem remains the same. Only traffic going to the specific destinations in the backup ISP goes through the backup ISP; there are specific routes that are received for each destination from that ISP. For any other traffic, the oldest default route would be used.

In the third scenario, the customer has a complete Internet routing table with entries to all destinations. In this case, a path to a specific destination over one ISP might have a different AS path length than a path over the other ISP. Unpredictable load balancing of traffic going to different destinations can occur.

For returning traffic originating on the Internet, the path selection is done based on the AS path length. Without any path manipulation, traffic takes the shortest path over one of the service providers. The result is unpredictable traffic patterns and asymmetric routing.

Default BGP Path Selection (Cont.)



```

Customer#show ip bgp
<...output omitted...>

Network          Next Hop        Metric LocPrf Weight Path
* i0.0.0.0        192.168.1.2      0      100      0 234 i
*>i               192.168.3.2      0      100      0 567 i
*> 209.165.201.0  0.0.0.0          0                      32768 i

```

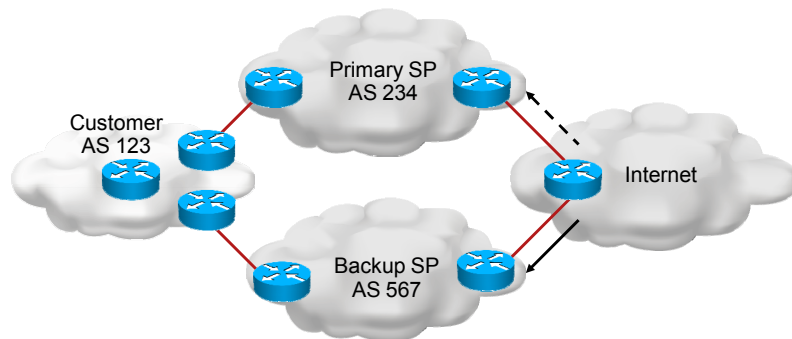
© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-26

In the example, the customer is connected to the Internet through two service providers. Each customer edge router receives a default route that is advertised further to the customer network.

The printout shows using the **show ip bgp** command, a situation in which the customer router receives two default routes with default attributes and an AS path length of 1. Since all BGP attributes are the same, the oldest path, which points to the backup ISP, is used (path that is marked with the “>”).

Default BGP Path Selection (Cont.)



```
Internet#show ip bgp
<...output omitted...>
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.1.0/24	0.0.0.0	0		32768	i
* 209.165.201.0	192.168.2.1			0	234 123 i
*>	192.168.5.1			0	567 123 i

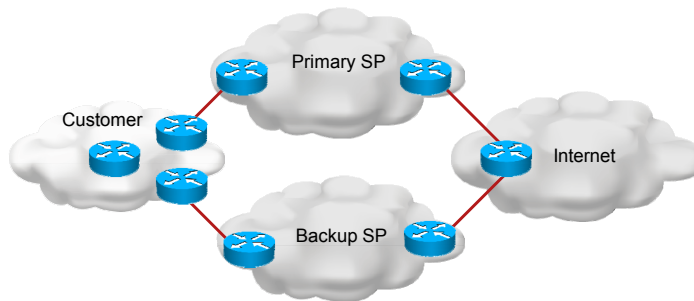
© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-2.27

The figure shows an output of the **show ip bgp** command on the Internet router. The customer network 209.165.201.0/24 is accessible over two paths. Because the AS path length of both paths is the same, again the oldest path, which points to the backup ISP, is used.

Influencing BGP Path Selection

- Customer side:
 - Outbound traffic: Increase local preference for routes received from preferred ISP.
 - Inbound traffic: Configure AS path prepending for routes sent to backup ISP.
- ISP side:
 - Manual for each customer
 - Through signalization using BGP communities



© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-28

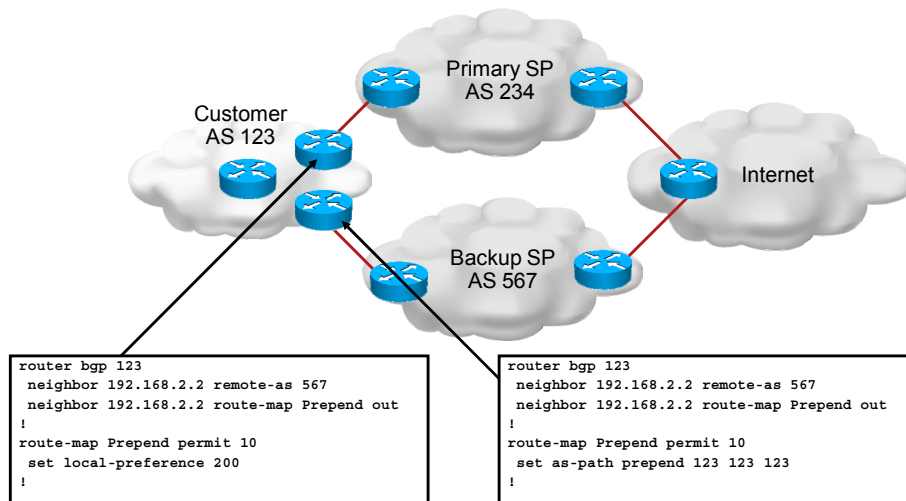
BGP path selection can be configured on the customer side. Outbound traffic can be influenced on the customer side by increasing local preference of routes that are received from the primary ISP. For inbound traffic, path selection can be influenced on the customer side by configuring AS path prepend for customer routes that are sent to the backup ISP. AS path prepending lengthens the AS path, thus making the path through the backup ISP less preferred.

BGP path selection can be influenced also on the ISP side. Instead of configuring the AS path prepend on the customer side, you can configure it on a backup ISP side. ISP can do this configuration manually for each customer (for example, they can match all routes originating in the customer AS), or it can be done through signalization using BGP communities. In this case, the customer appends BGP community (which an ISP specifies) to routes sent to a service provider. The ISP then matches this BGP community and, based on the community value, configures AS path prepending when customer routes are advertised further to the Internet.

In any case, BGP path selection for outbound traffic has to be done on the customer side by increasing local preference of routes that are received from the primary ISP.

Influencing BGP Path Selection

Cisco IOS Software



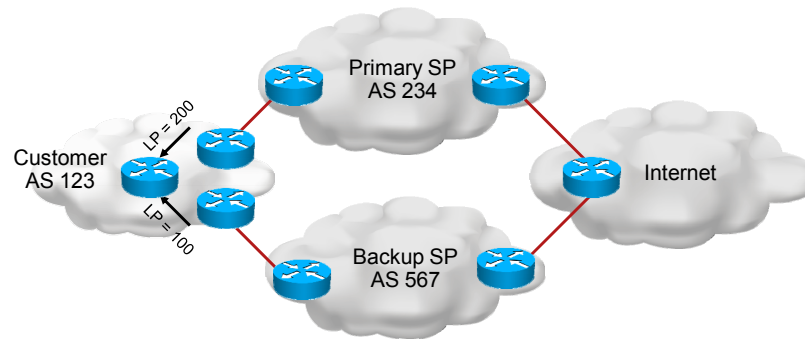
The first configuration example shows how to influence BGP path selection in Cisco IOS Software. Path manipulation is done on the customer side in this case.

On the router that is connected to the primary ISP, a route map is configured that sets local preference of routes that are received from the primary ISP to 200. This route map is applied to the BGP neighbor in the inbound direction.

On the router that is connected to the backup ISP, a route map is configured that prepends the customer AS number three times to routes sent to the backup ISP. This route map is applied to the BGP neighbor in the outbound direction. Local preference of routes that are received from the backup ISP is not affected by this route map and is set to the default value, which is 100.

Influencing BGP Path Selection

Verification



```
Customer#show ip bgp
<...output omitted...>
      Network        Next Hop           Metric LocPrf Weight Path
* i0.0.0.0           192.168.3.2             0    100      0 567 i
*>i                   192.168.1.2             0    200      0 234 i
*> 209.165.201.0     0.0.0.0                 0                 32768 i
```

© 2010 Cisco Systems, Inc. All rights reserved.

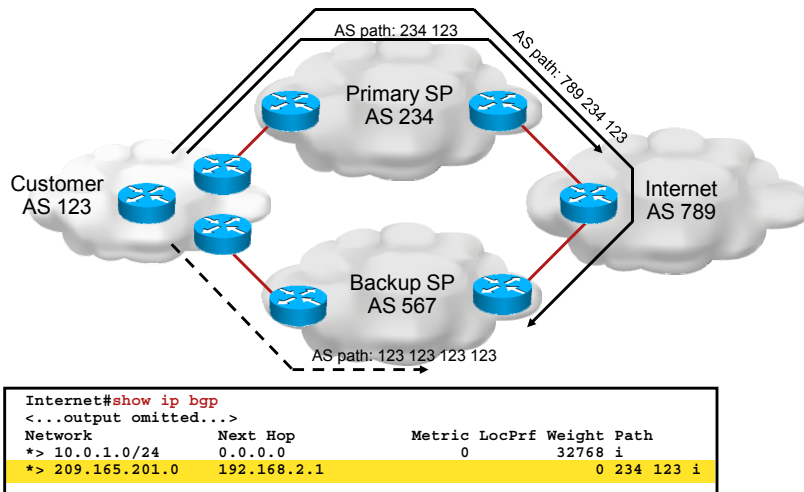
MSPRP v1.0—2-30

The sample printout shows the path selection on the customer side after influencing the path selection. Each customer edge router receives a default route, which is advertised further to the customer network. In this case, the customer edge router that is connected to the primary ISP sets local preference of the default route that is received from the primary ISP to 200. Recall that the local preference attribute is advertised further through IBGP sessions.

The customer router then receives two default routes, one with local preference set to 100 and one with local preference set to 200. A path with higher local preference is preferred, and this path points to the primary ISP (selected path is marked with “>”).

Influencing BGP Path Selection

Verification (Cont.)



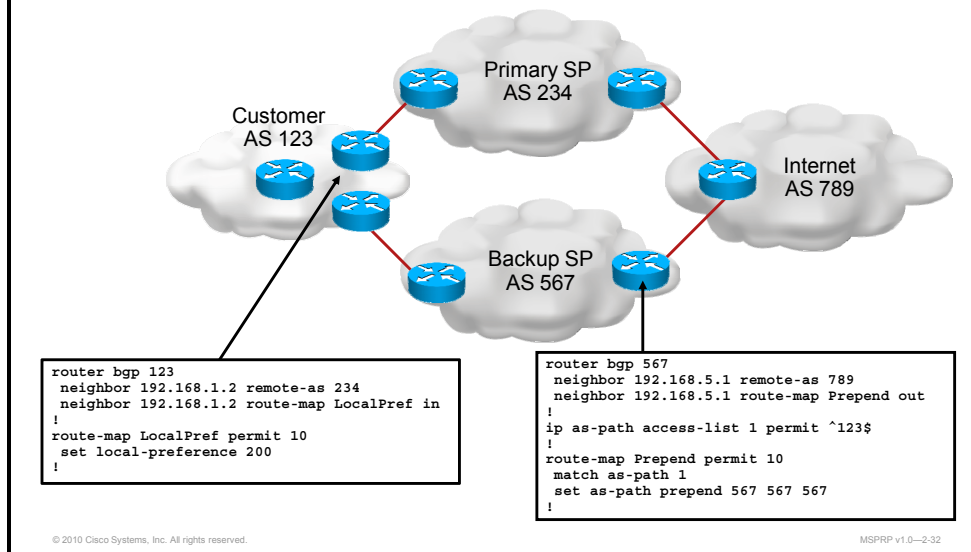
The sample printout shows the path selection on the Internet side after influencing the path selection.

You could expect that there would be two paths present on the Internet router to reach the customer network. However, the **show ip bgp** command reveals that there is only one present in the BGP table. The reason is that the backup ISP receives two BGP updates. One is received from the customer directly with an AS path length of 4. Recall that AS path prepending has been configured, which prepends the AS number three times. The fourth copy of the AS number is prepended because of the EBGP session. Another update about the customer network is received from the Internet with an AS path length of 3. Traffic originating in the backup ISP and going to the customer would prefer the path through the Internet and primary ISP because of the shorter AS path. If the backup ISP advertised the customer network to the Internet, there would be a violation of the split-horizon rule.

If the link between the customer and the primary ISP fails, the backup ISP would advertise the customer network to the Internet and connectivity would be possible over the backup ISP.

Influencing BGP Path Selection

Cisco IOS Software



The second configuration example shows how to influence BGP path selection on the customer and ISP side. The customer influences path selection for incoming traffic by increasing local preference of the default route that is received from the primary ISP. This influence is achieved by configuring a route map and applying this route map to the BGP neighbor in the inbound direction.

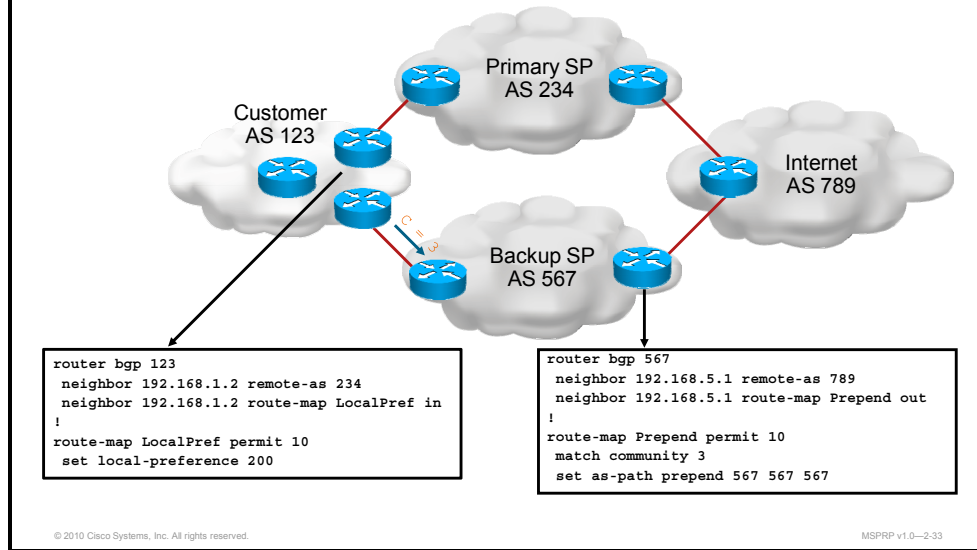
Manipulation of the path selection process for incoming traffic is done on the ISP side by configuring AS path prepending for customer routes, when those routes are advertised further to the Internet. On the ISP router, which is connected to the Internet, a route map is configured that matches all routes originating in the customer AS (123). AS path prepending is configured to prepend the AS number three times. This route map is applied to the BGP neighbor in the Internet in the outbound direction.

This configuration is useful in the second multihoming scenario, in which both ISPs advertise the default route, and the backup ISP also advertises local routes. Traffic originating in the backup ISP and going to the customer network goes to the customer AS directly in this case, because the direct path would have a shorter AS path.

The overhead of this configuration is that the backup ISP must manually match customer routes that are based on prefixes or AS numbers.

Influencing BGP Path Selection

Cisco IOS Software (Cont.)



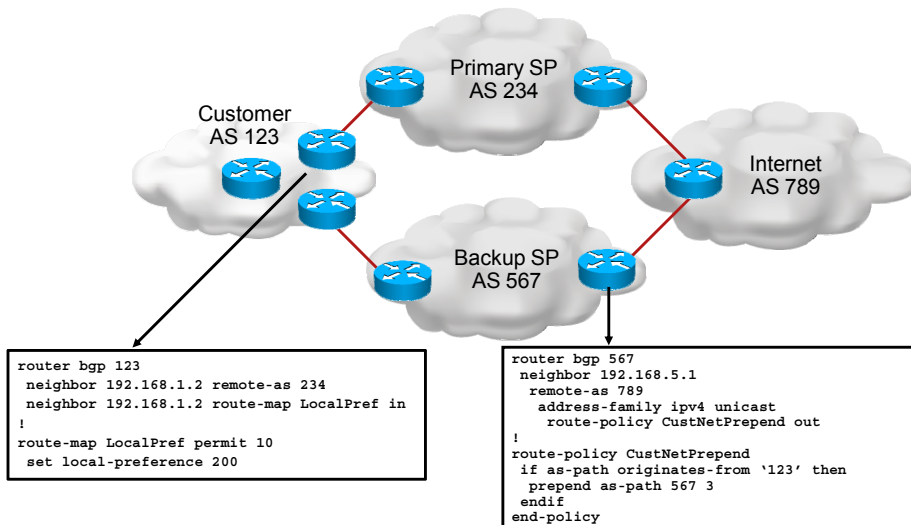
The third configuration example differs from the second one only in matching the customer routes. In this case, BGP communities are used for the customer to signal services to the ISP. When the customer advertises the customer network to the backup ISP, a BGP community (specified by the ISP) is appended to the network update (configuration is not shown in the example). ISP then matches communities and based on the values of the community, AS path prepending is configured. For example, if the customer uses the community value of 3, this means that the ISP should append the AS number three times.

A route map is configured on the ISP router connecting to the Internet, which matches routes with the BGP community set to 3 and prepends the AS number three times.

The benefit of this configuration is that the backup ISP can match routes that are based on communities specified by the ISP. The customer's job is to use the right communities.

Influencing BGP Path Selection

Cisco IOS XR Software



© 2010 Cisco Systems, Inc. All rights reserved.

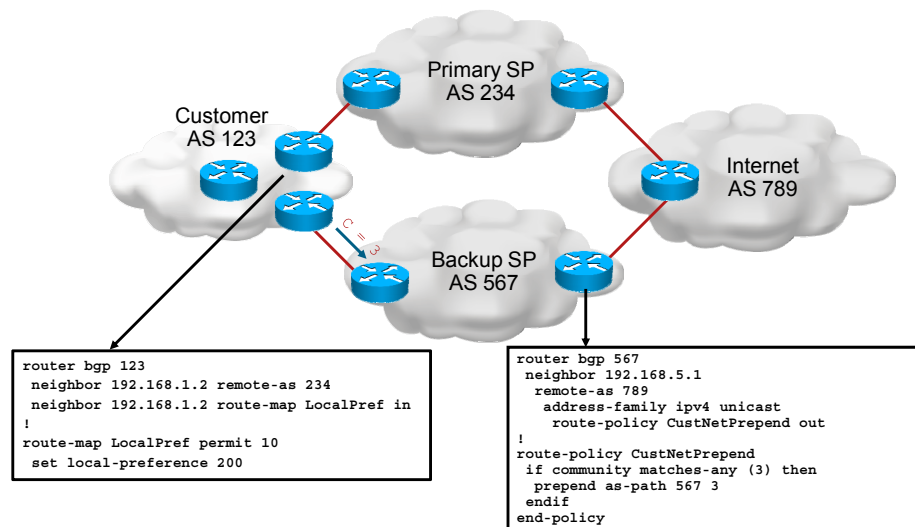
MSPRP v1.0—2-34

The sample configuration shows how to configure the BGP path selection, on the service provider side, manually on the Cisco IOS XR Software.

A route policy is configured that matches routes originating in the AS 123 and prepends the AS number 567 three times. This route policy is applied to the BGP neighbor in the outbound direction.

Influencing BGP Path Selection

Cisco IOS XR Software (Cont.)

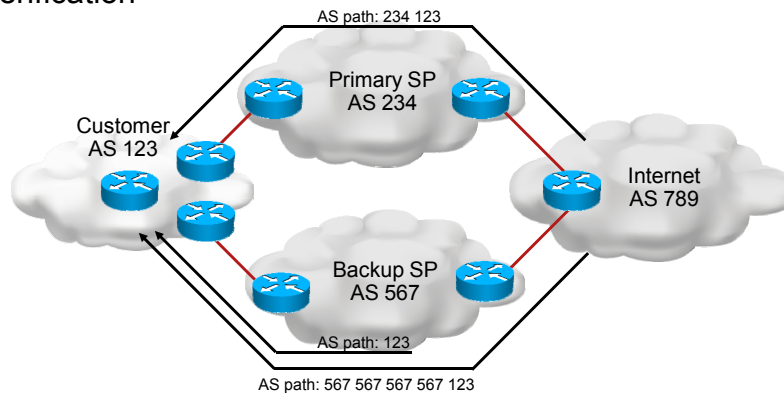


The sample configuration shows how to configure the BGP path selection on the service provider side using signaling on the Cisco IOS XR Software.

A route policy is configured that matches routes with the BGP community set to 3 and prepends the AS number 567 three times. This route policy is applied to the BGP neighbor in the outbound direction.

Influencing BGP Path Selection

Verification



```
Internet#show ip bgp
<...output omitted...>
Network      Next Hop      Metric LocPrf Weight Path
*> 10.0.1.0/24 0.0.0.0        0       32768 i
*> 209.165.201.0 192.168.2.1    0       0 234 123 i
                  192.168.5.1    0       0 567 567 567 567 123 i
```

© 2010 Cisco Systems, Inc. All rights reserved.

MSPRP v1.0—2-36

The sample printout shows the path selection, after influencing, on the Internet side. The AS path prepending has been configured on the backup ISP router. AS path prepending can be done manually or through signalization using BGP communities, as mentioned previously. This configuration is appropriate for this scenario, where the backup ISP also advertises local routes to the customer. In this case, traffic to and from the backup ISP should go directly through the backup ISP.

The **show ip bgp** command shows that there are two paths available on the Internet router to reach the customer network. One is over the primary ISP (AS path is 234 123), and another one is over the backup ISP (AS path is 567 567 567 567 123). The shortest path over the primary ISP is preferred. However, traffic originating in the backup ISP and going to the customer will take a direct path to the customer.

Summary

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

Summary

- OSPF cost is calculated based on the interface bandwidth and reference bandwidth.
- All IS-IS enabled interfaces are assigned a metric of 10.
- Up to 16 equal-cost paths can be used by OSPF and IS-IS for load balancing, but only 4 equal-cost paths are used by default.
- BGP uses rich metrics. Without route manipulation, the most common criteria for path selection is AS path length.
- BGP path manipulation can be configured on the customer side, the ISP side, or both.

© 2010 Cisco Systems, Inc. All rights reserved.

MSRP v1.0-237

Lesson Self-Check

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

- Q1) In which two ways can you influence OSPF path selection? (Choose two.) (Source: Influencing Internal Routing Using OSPF)
- A) by changing OSPF cost on interface
 - B) by changing reference bandwidth
 - C) by changing the metric style to wide
 - D) by changing administrative distance
- Q2) How can you influence IS-IS path selection? (Source: Influencing Internal Routing Using IS-IS)
- A) by changing reference bandwidth
 - B) by changing the metric style to wide
 - C) by changing IS-IS metric on interface
 - D) by changing IS-IS metric globally
- Q3) What is the most common criteria for BGP path selection when there is no path manipulation? (Source: Influencing External Routing using BGP)
- A) local preference
 - B) AS path length
 - C) path age
 - D) MED
- Q4) How can you influence BGP path selection for inbound traffic? (Source: Influencing External Routing Using BGP)
- A) by changing local preference of advertised routes
 - B) by configuring AS path prepending of received routes
 - C) by changing local preference of received routes
 - D) by configuring AS path prepending of advertised routes
- Q5) How can you influence BGP path selection for outbound traffic? (Source: Influencing External Routing Using BGP)
- A) by changing weight of advertised routes
 - B) by configuring AS path prepending of received routes
 - C) by changing local preference of received routes
 - D) by changing MED of advertised routes

Lesson Self-Check Answer Key

- Q1) A, B
- Q2) C
- Q3) B
- Q4) D
- Q5) C

Module Summary

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

Module Summary

- Common service provider network:
 - Runs BGP or static routing with the customer
 - Exchanges routes with other service providers via EBGp
 - Runs IBGP between its own BGP speakers
 - Runs one instance of IGP (OSPF or IS-IS) for internal routes
- When adding a new router to the service provider network, follow these deployment steps: plan, design, develop, configure, install, and verify.
- Reasons to modify routing parameters are redesigning, optimizing, and mitigating faults.
- Consider the impact and risks associated with making changes in the routing configuration.
- Also consider the customer SLA to determine the impact on customers affected by the changes in the routing configuration.
- Static routing is adequate for the single-homed customer.
- A multihomed connection to two ISPs using BGP offers customers high levels of resiliency and redundancy.

© 2010 Cisco Systems, Inc. All rights reserved.

MS RRP v1.0--2.1

A common service provider network runs static routing with single-homed customers and Border Gateway Protocol (BGP) with multihomed customers. Service providers exchange routes with other service providers via External Border Gateway Protocol (EBGP), while running Internal Border Gateway Protocol (IBGP) between their own BGP speakers. Service providers must also run Open Shortest Path First (OSPF) or Intermediate System-to-Intermediate System (IS-IS) for the needs of internal routes.

When you add a new router to the service provider network, follow six deployment steps:

- Plan
- Design
- Develop
- Configure
- Install
- Verify

At times, you may need to modify routing parameters; possible reasons include redesigning or optimizing a network and mitigating faults. When connecting a single-homed customer, static routing is adequate. On the other hand, a multihomed connection to two service providers using BGP offers high levels of resiliency and redundancy. Route manipulation for BGP is recommended because BGP uses rich metrics. Without route manipulation, the most common criterion for path selection is an autonomous system path length.

