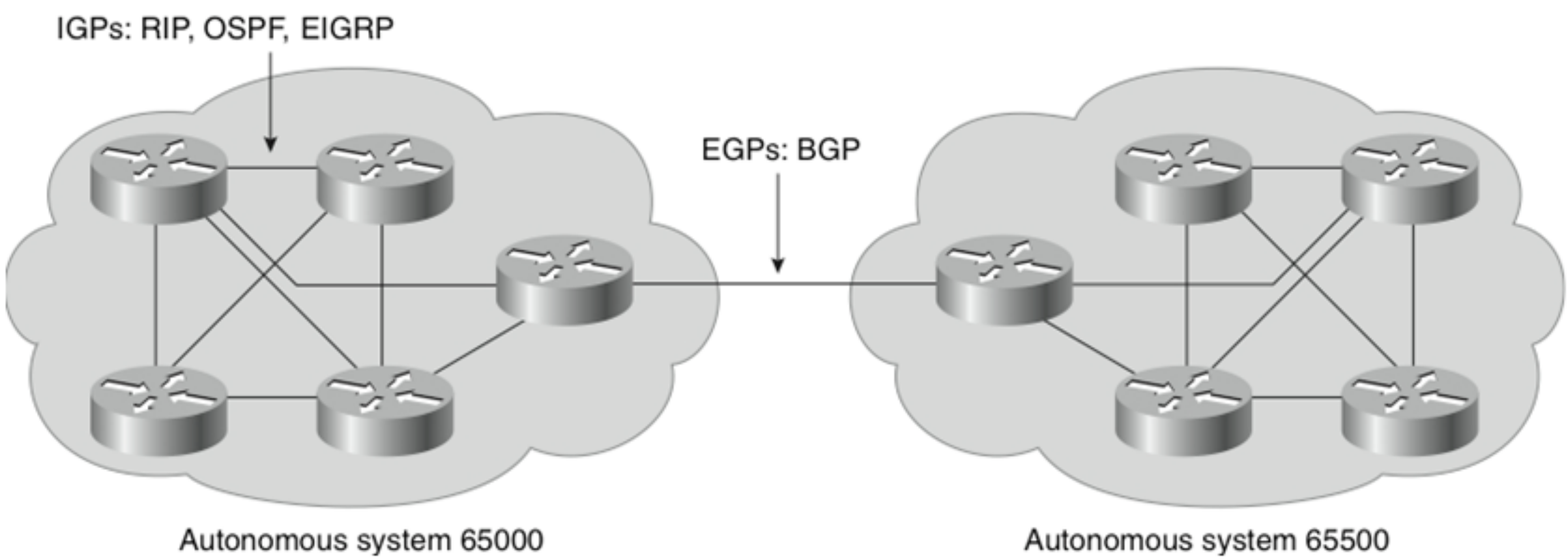
BGP Terminology, Concepts, and Operation

IGP versus EGP

- **Interior gateway protocol (IGP)**
  - A routing protocol operating within an Autonomous System (AS).
  - RIP, OSPF, and EIGRP are IGPs.
- **Exterior gateway protocol (EGP)**
  - A routing protocol operating between different AS.
  - BGP is an interdomain routing protocol (IDRP) and is an EGP.



Autonomous Systems (AS)

- An AS is a group of routers that share similar routing policies and operate within a single administrative domain.

- An AS typically belongs to one organization.
- If an AS connects to the public Internet using an exterior gateway protocol such as BGP, then it must be assigned a unique AS number which is managed by the Internet Assigned Numbers Authority (IANA).

- Any organization can acquire an ASN. This requires a contractual agreement with a sponsoring LIR (Local Internet Registry) or with the RIPE NCC.
- To do this, you need to become a member of the RIPE NCC(Regional Internet Registries) or find a sponsoring LIR.

AS Numbers

- AS numbers can be between **1** to **65,535**.
  - RIRs manage the AS numbers between **1** and **64,512**.
  - The **64,512 - 65,535** numbers are reserved for private use (similar to IP Private addresses).
  - The IANA is enforcing a policy whereby organizations that connect to a single provider use an AS number from the private pool.

AS Numbers

- **Note**:
  - On Friday 14 September, 2012, the RIPE NCC, the Regional Internet Registry (RIR) for Europe, the Middle East and parts of Central Asia, distributed the last blocks of IPv4 address space from the available pool. For this reason, the IETF has released RFC 4893 and RFC 5398.
  - These RFCs describe BGP extensions to increase the AS number from the two-octet (16-bit) field to a four-octet (32-bits) field, increasing the pool size from **65,536** to **4,294,967,296** values.

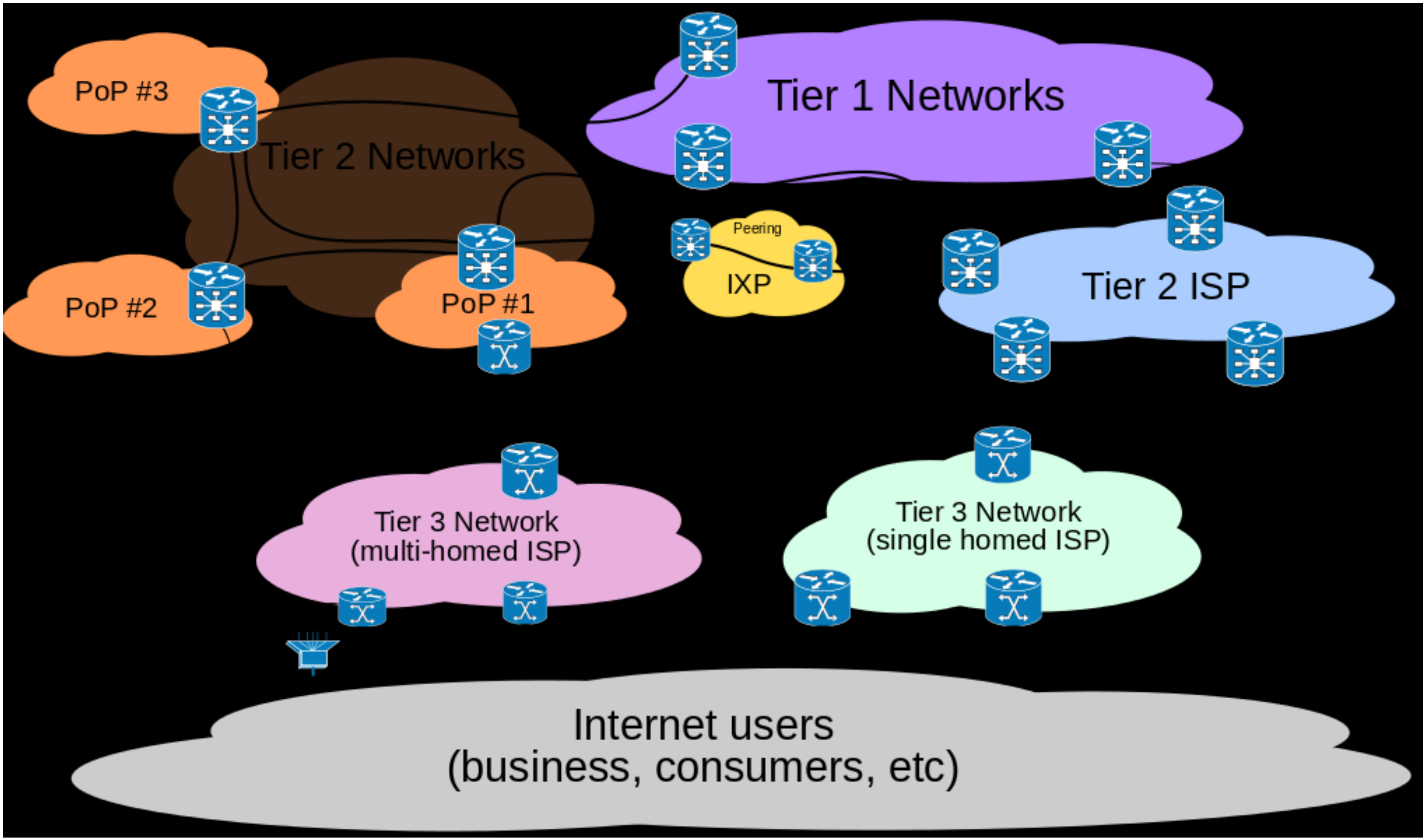
IANA- Internet Assigned Numbers Authority

- The IANA is responsible for allocating AS numbers through five Regional Internet Registries (RIRs).

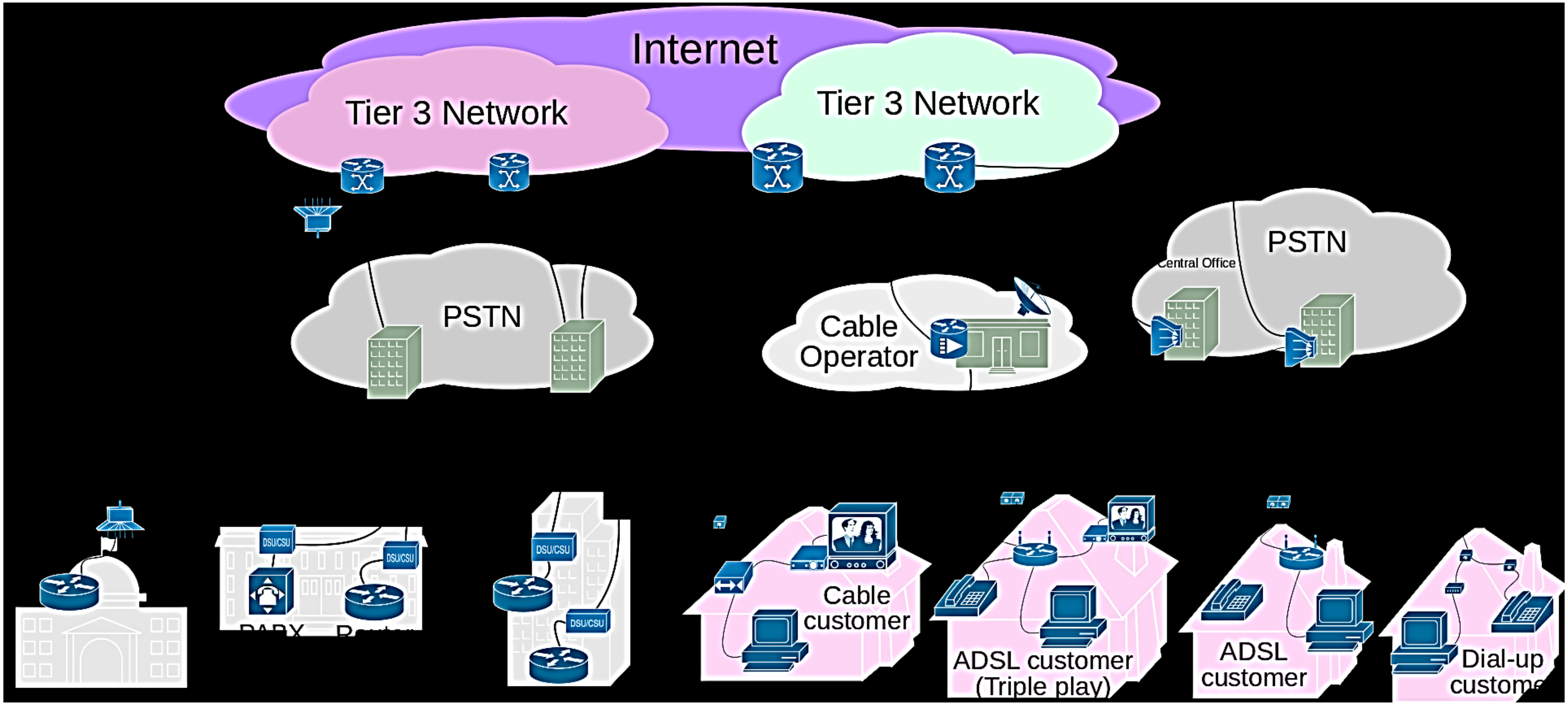
Regional Internet Registries (RIRs)

RIR Name	Geographic Coverage	Link
AfriNIC	Continent of Africa	www.afrinic.net
APNIC (Asia Pacific Network Information Centre)	Asia Pacific region	www.apnic.net
ARIN (American Registry for Internet Numbers)	Canada, the United States, and several islands in the Caribbean Sea and North Atlantic Ocean	www.arin.net
LACNIC (Latin America and Caribbean Internet Addresses Registry)	Central and South America and portions of the Caribbean	www.lacnic.net
RIPE (Réseaux IP Européens)	Europe, the Middle East, and Central Asia	www.ripe.net

https://en.wikipedia.org/wiki/Internet_service_provider#/media/File:Internet_Connectivity_Distribution_%26_Core.svg



https://en.wikipedia.org/wiki/Internet_service_provider#/media/File:Internet_Connectivity_Access_layer.svg



BGP Basics

- The Internet is a collection of autonomous systems that are interconnected to allow communication among them.
  - BGP provides the routing between these autonomous systems.
- BGP is a path vector protocol.
- It is the only routing protocol to use TCP (reliable).
  - OSPF and EIGRP operate directly over IP. IS-IS is at the network layer.
  - RIP uses the User Datagram Protocol (UDP) for its transport layer.

BGP Basics

- BGP version 4 (BGP-4) is the latest version of BGP.
  - Defined in RFC 4271.
  - Supports supernetting, CIDR and VLSM .

- BGP4 and CIDR prevent the Internet routing table from becoming too large.
  - Without CIDR, the Internet would have 2,000,000 + entries.
  - With CIDR, Internet core routers manage around 500,000 entries (2014)
  - http://bgp.potaroo.net/

of Current BGP Routes

As of August 30, 2010, there were 332,145 routes in the routing tables of the Internet core routers.

http://bgpupdates.potaroo.net/instability/bgpupd.html

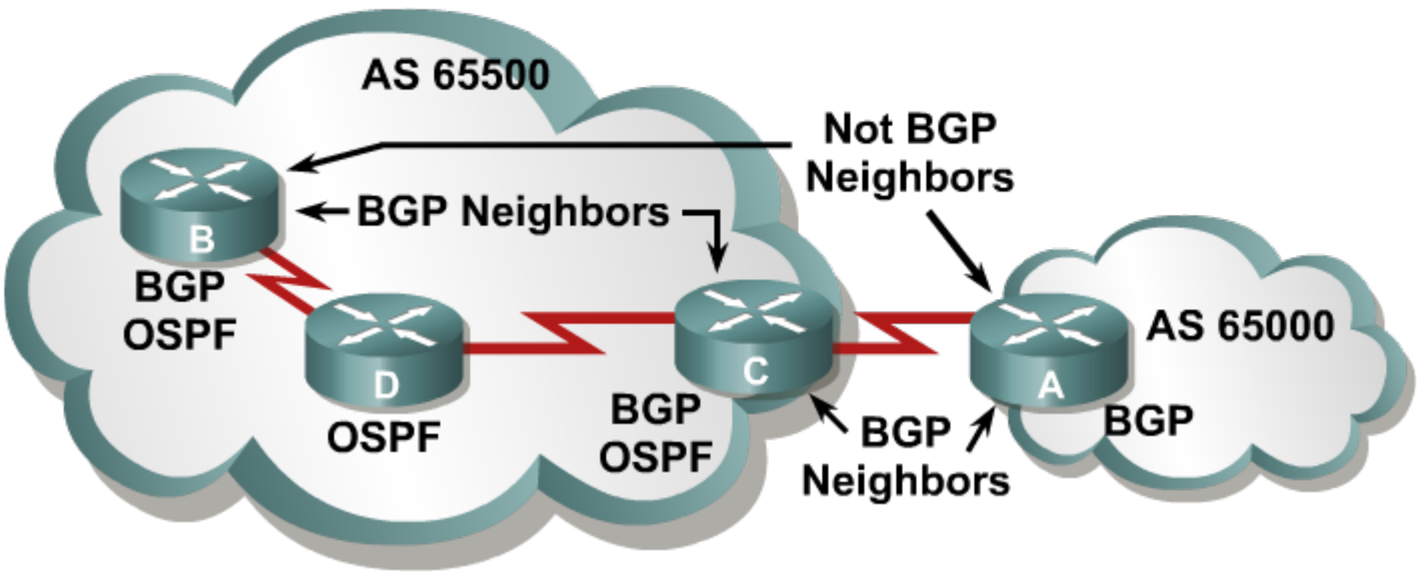
```
7 Day BGP Profile: 24-August-2010 00:00 - 30-August-2010 23:59 (UTC+1000)

Number of BGP Update Messages:          1195261
Number of Prefix Updates:               2787149
Number of Prefix Withdrawals:            490070
Average Prefixes per BGP Update:           2.74
Average BGP Update Messages per second:    1.73
Average Prefix Updates per second:         4.74
Peak BGP Update Message Rate per second: 3848  (19:25:51 Mon, 30-Aug-2010)
Peak Prefix Update Rate per second:      66398 (07:07:37 Mon, 30-Aug-2010)
Peak Prefix Withdraw Rate per second:    16512 (19:26:14 Mon, 30-Aug-2010)
Prefix Count:                            342962
Updated Prefix Count:                    332145
Stable Prefix Count:                      10817
Origin AS Count:                          35292
Updated Origin AS Count:                  34786
Stable Origin AS Count:                     506
Unique Path Count:                       215660
Updated Path Count:                      195814
Stable Path Count:                        19846
```

Peers = Neighbors

- A “BGP peer,” also known as a “BGP neighbor,” is a specific term that is used for BGP speakers that have established a neighbor relationship.
- Any two routers that have formed a TCP connection to exchange BGP routing information are called BGP peers or BGP neighbors.

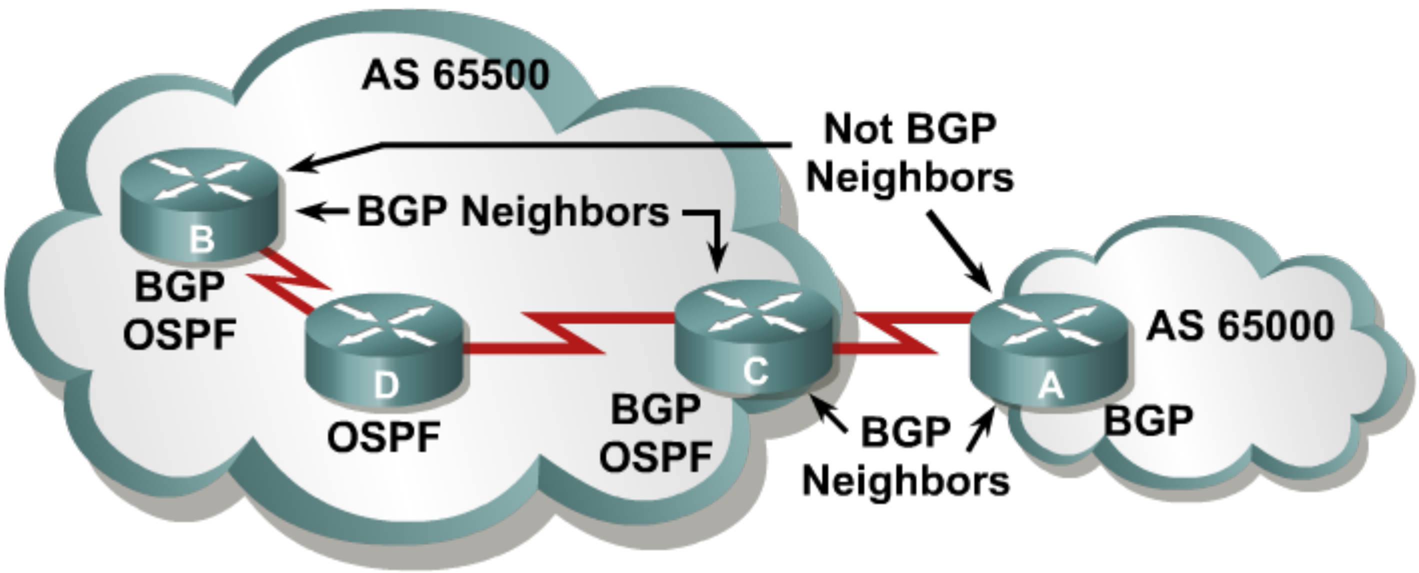
BGP Operational Overview

- When two routers establish a TCP enabled BGP connection, they are called **neighbors** or **peers**.
  - Peer routers exchange multiple connection messages.
- Each router running BGP is called a **BGP speaker**.



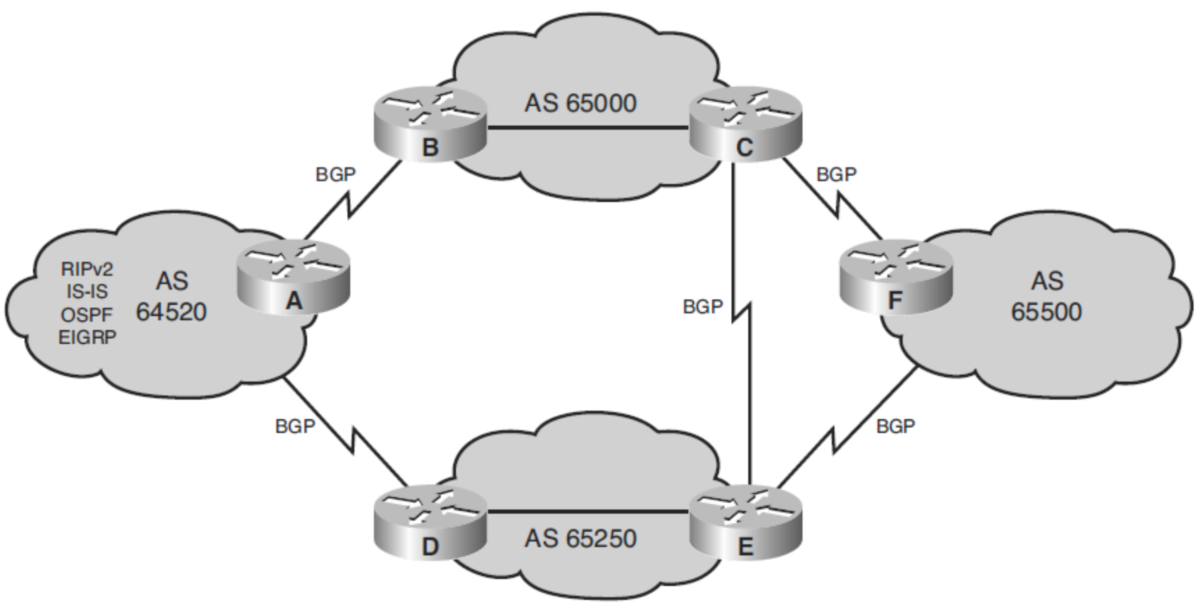
BGP Operational Overview

- When BGP neighbors first establish a connection, they exchange all candidate BGP routes.
  - After this initial exchange, incremental updates are sent as network information changes.



BGP Use Between AS

- BGP provides an interdomain routing system that guarantees the loop-free exchange of routing information between autonomous systems.



Comparison BGP with IGPs

- BGP works differently than IGPs because it does not make routing decisions based on best path metrics.
  - Instead, BGP is a policy-based routing protocol that allows an AS to control traffic flow using multiple BGP attributes.
- Routers running BGP exchange network attributes including a list of the full path of BGP AS numbers that a router should take to reach a destination network.
- BGP allows an organization to fully use all of its bandwidth by manipulating these path attributes.

Comparing IGPs with BGP

Protocol	Interior or Exterior	Type	Hierarchy Required?	Metric
RIP	Interior	Distance vector	No	Hop count
OSPF	Interior	Link state	Yes	Cost
IS-IS	Interior	Link state	Yes	Metric
EIGRP	Interior	Advanced distance vector	No	Composite
BGP	**Exterior**	**Path vector**	**No**	**Path vectors (attributes)**

Connecting Enterprise Networks to an ISP

- Modern corporate IP networks connect to the global Internet.

- Requirements that must be determined for connecting an enterprise to an ISP include the following:
  - Public IP address space
  - Enterprise-to-ISP connection link type and bandwidth
  - Connection redundancy
  - Routing protocol

Public IP Address Space

- Public IP addresses are used:
  - By internal enterprise clients to access the Internet using NAT.
  - To make enterprise servers accessible from the Internet using static NAT.

- Public IP addresses are available from ISPs and RIRs.
  - Most enterprises acquire their IP addresses and AS number from ISPs.
  - Large enterprises may want to acquire IP addresses and AS number from a RIR.

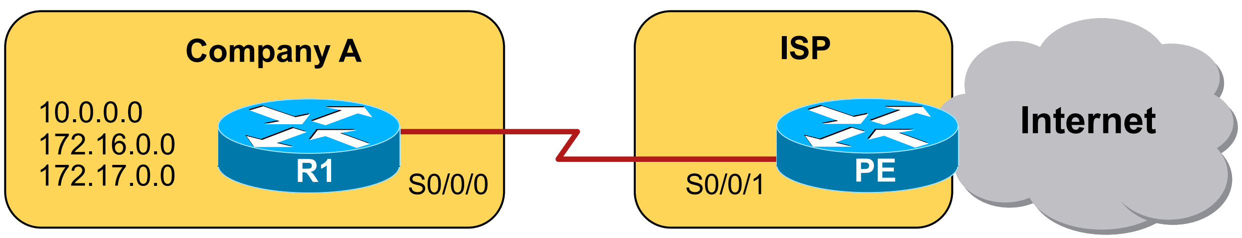
Connection and Routing Questions

- Which connection options does the ISP offer?
- Which routing options does the ISP offer?
- Will the enterprise network be connected to multiple ISPs?
- Does the routing need to support one link to an ISP or multiple links, to one or multiple ISPs?
- Is traffic load balancing over multiple links required?
- How much routing information needs to be exchanged with the ISP?
- Does the routing need to respond to the changes in the network topology, such as when a link goes down?

Using Static Routes Example

- Static routes are the simplest way to implement routing with an ISP.
  - Typically a customer has a single connection to an ISP and the customer uses a default route toward the ISP while the ISP deploys static routes toward the customer.

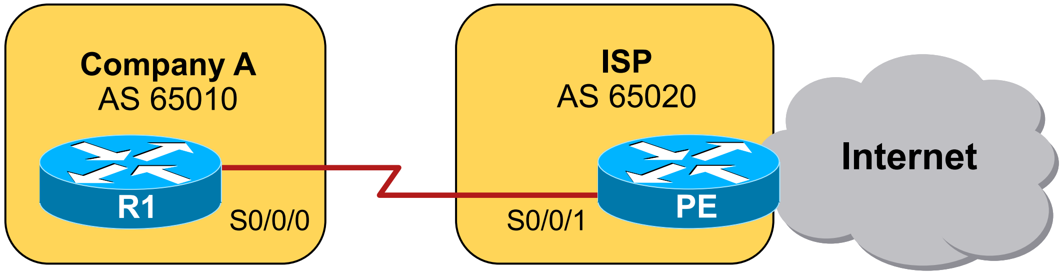
```
R1(config)# router eigrp 110
R1(config-router)# network 10.0.0.0
R1(config-router)# exit
R1(config)# ip default-network 0.0.0.0
R1(config)# ip route 0.0.0.0 0.0.0.0 serial 0/0/0
```




```
PE(config)# ip route 10.0.0.0 255.0.0.0 serial 0/0/1
PE(config)# ip route 172.16.0.0 255.255.0.0 serial 0/0/1
PE(config)# ip route 172.17.0.0 255.255.0.0 serial 0/0/1
```

Using BGP

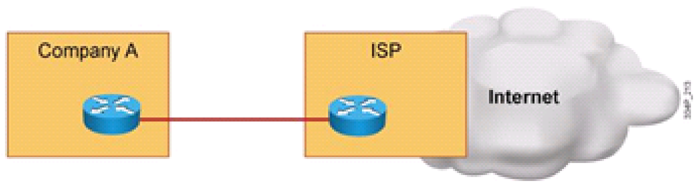
- BGP can be used to dynamically exchange routing information.
- BGP can also be configured to react to topology changes beyond a customer-to-ISP link.



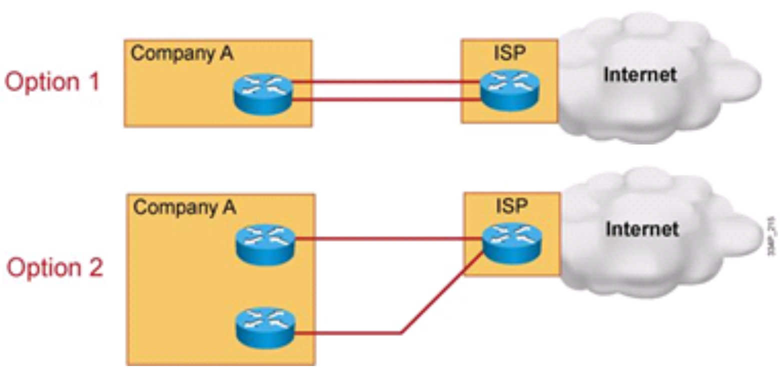
Connection Redundancy

Connecting to One ISP

Single-homed

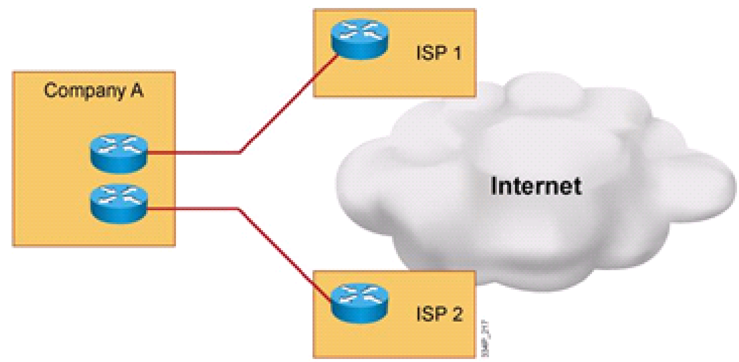


Dual-homed

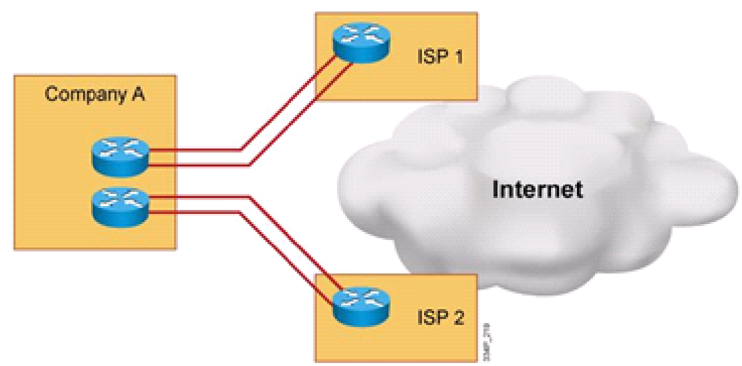


Connecting to Two or more ISPs

Multihomed

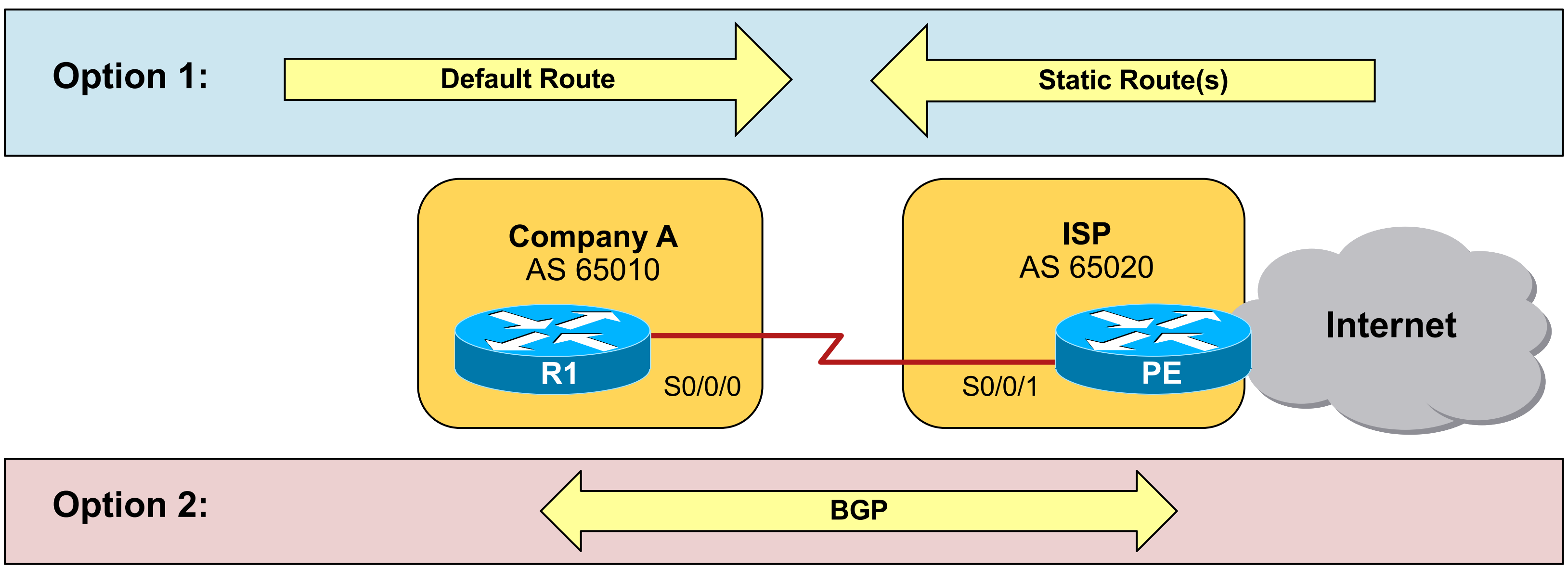


Dual-multihomed



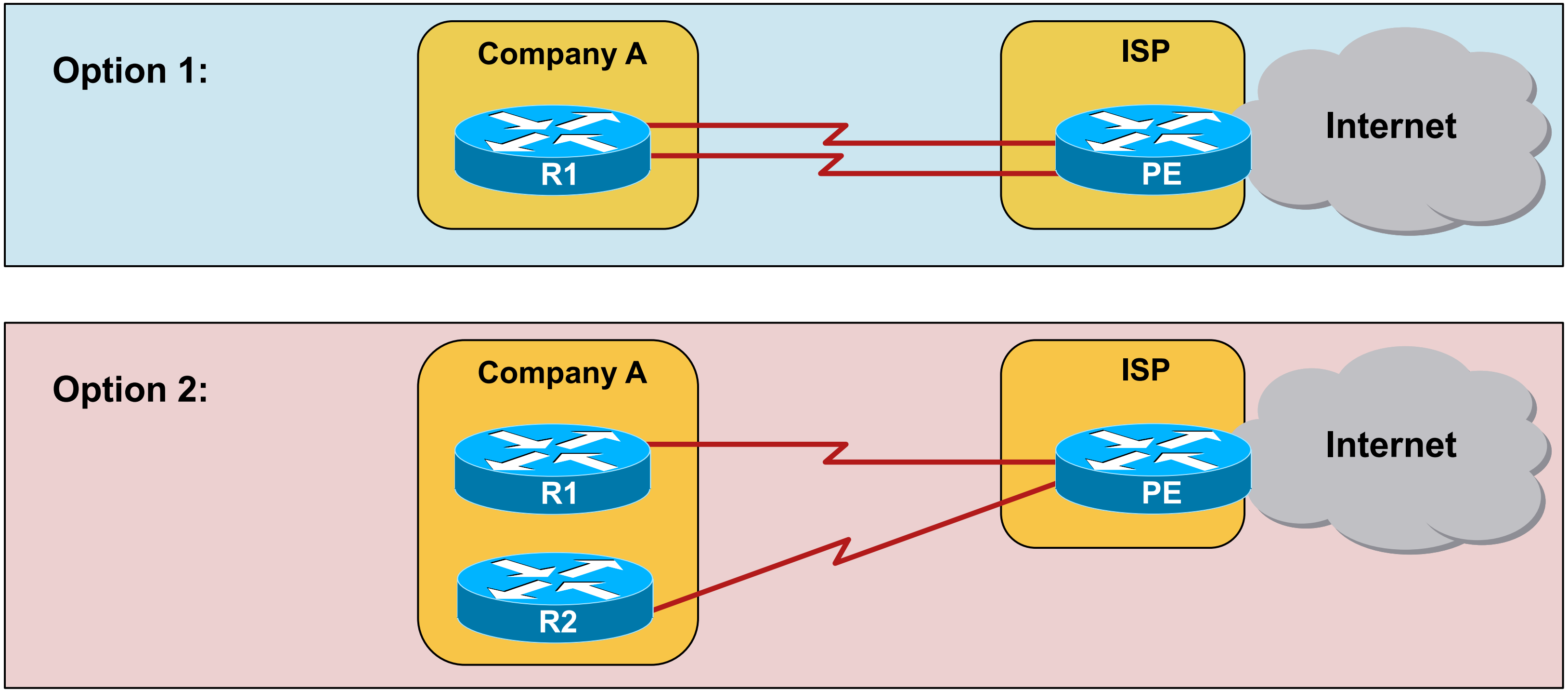
Connecting to One ISP: Single-Homed

- The connection type depends on the ISP offering (e.g., leased line, xDSL, Ethernet) and link failure results in a no Internet connectivity.
- The figure displays two options:
  - **Option 1**: Static routes are typically used with a static default route from the customer to the ISP, and static routes from the ISP toward customer networks.
  - **Option 2**: When BGP is used, the customer dynamically advertises its public networks and the ISP propagates a default route to the customer.



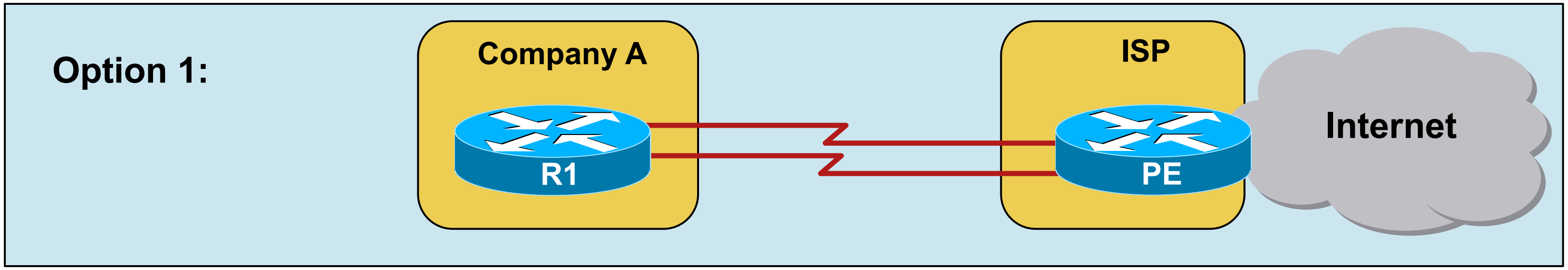
Connecting to One ISP: Dual-Homed

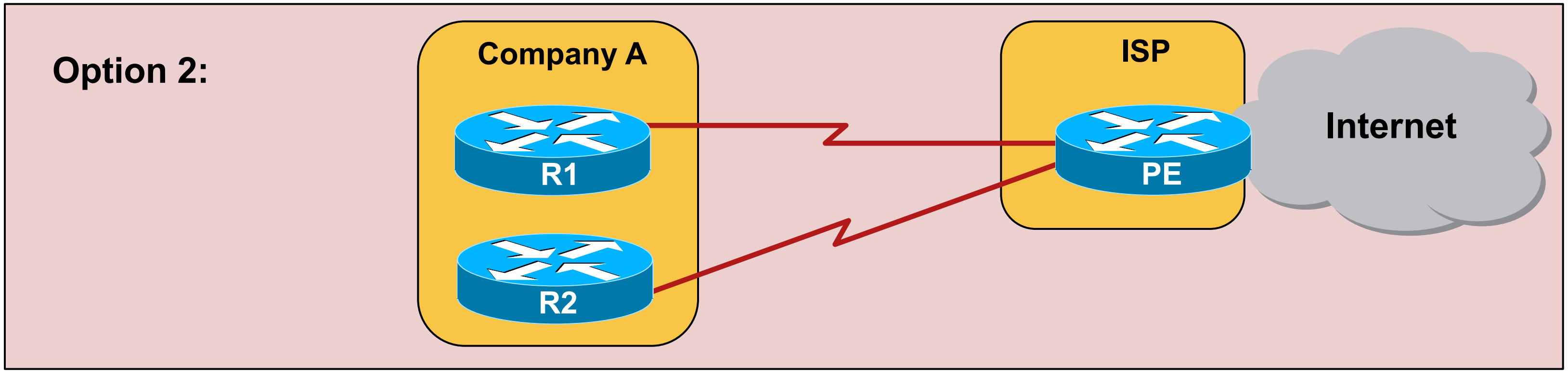
- The figure displays two dual-homed options:
  - **Option 1**: Both links can be connected to one customer router.
  - **Option 2**: To enhance resiliency, the two links can terminate at separate routers in the customer's network.



Connecting to One ISP: Dual-Homed

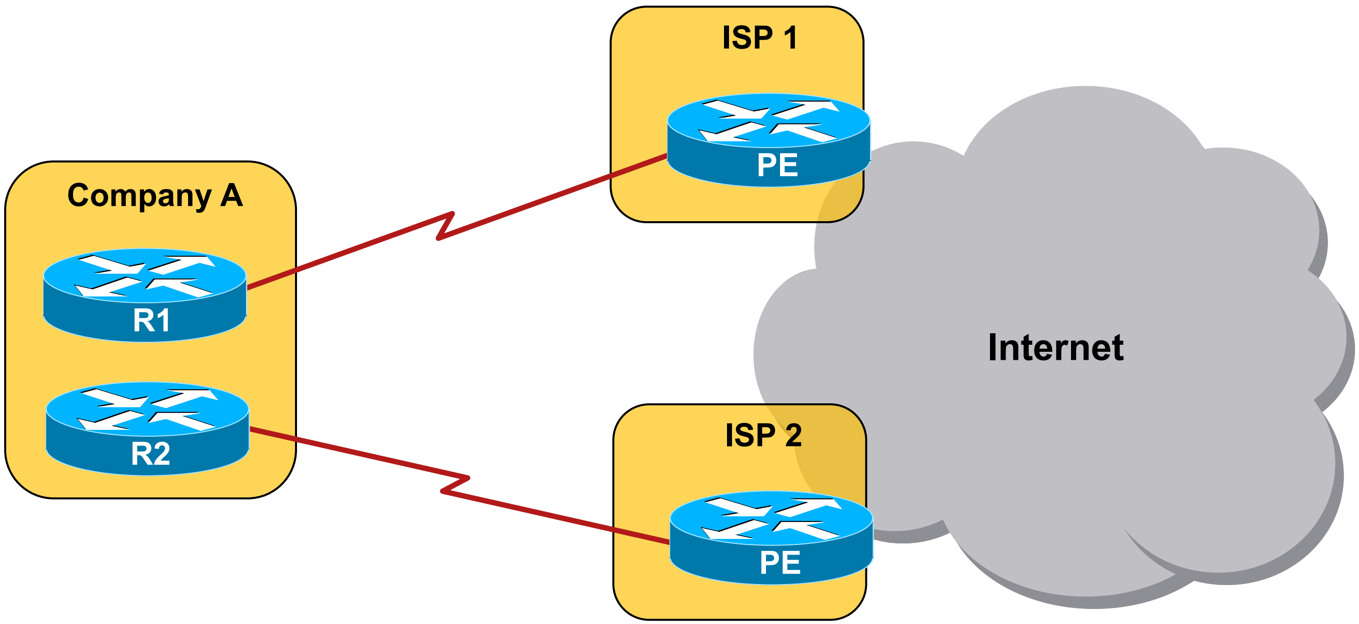
- Routing deployment options include:
  - Primary and backup link functionality in case the primary link fails.
  - Load sharing using Cisco Express Forwarding (CEF).
- Regardless, routing can be either static or dynamic (BGP).





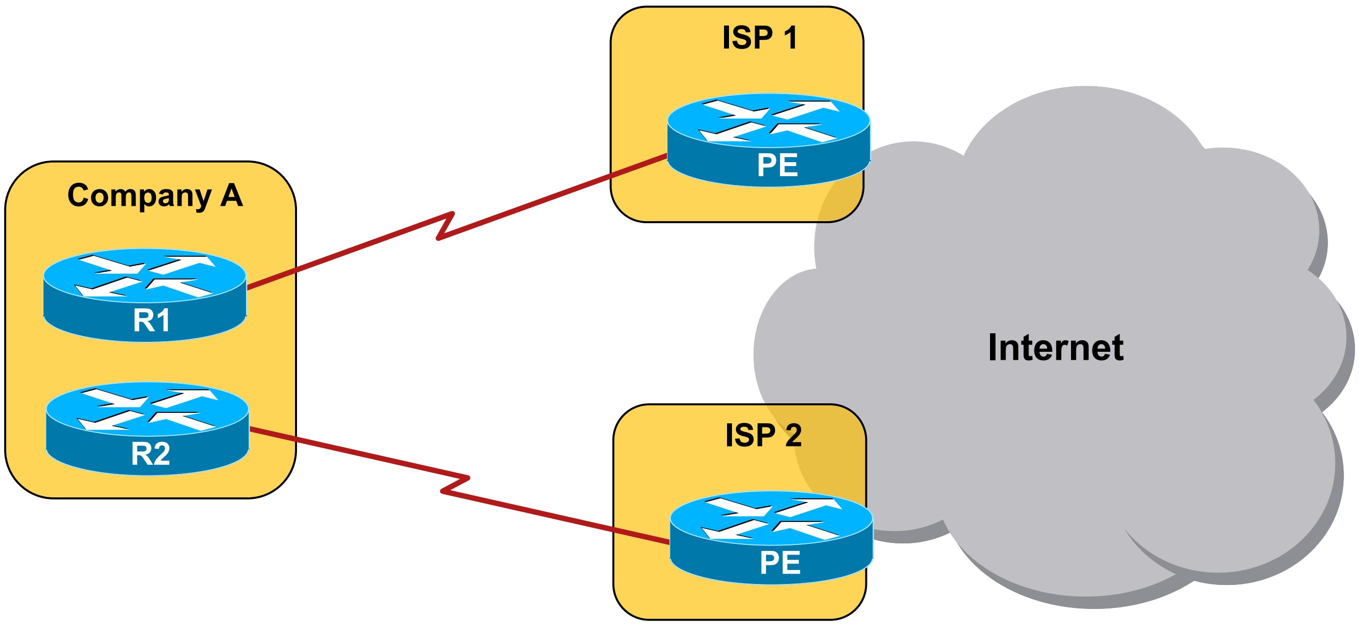
Connecting to Multiple ISPs: Multihomed

- Connections from different ISPs can terminate on the same router, or on different routers to further enhance the resiliency.
- Routing must be capable of reacting to dynamic changes therefore BGP is typically used.



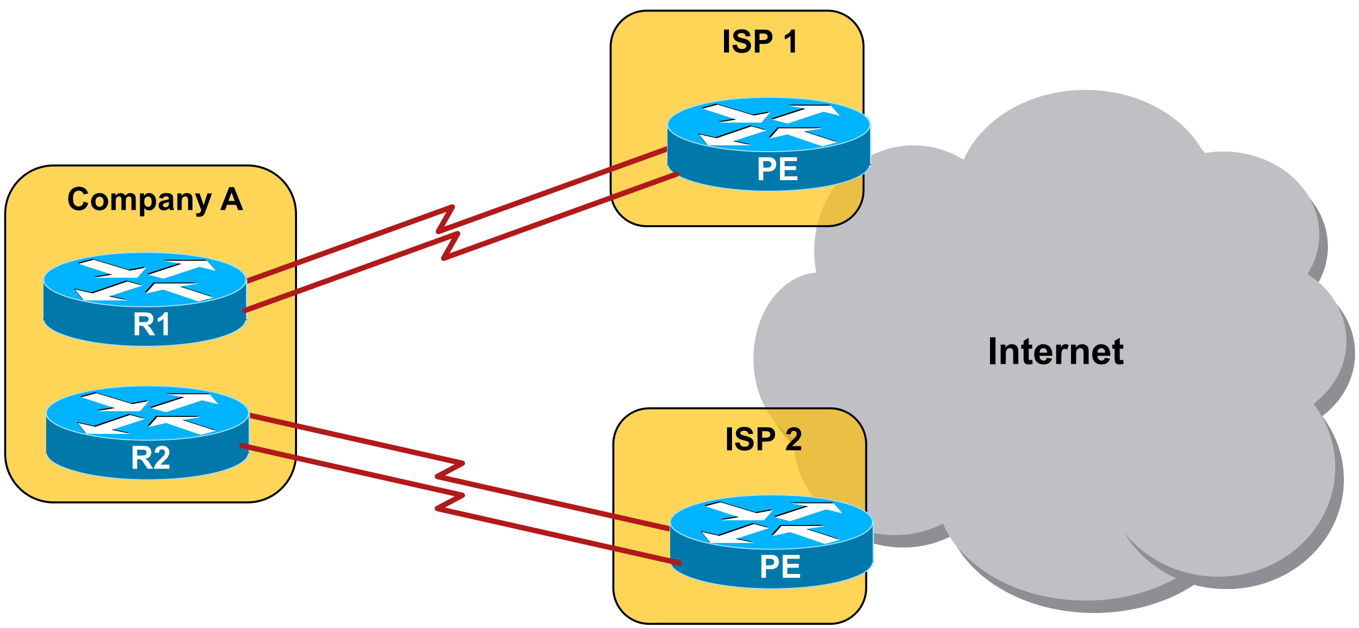
Connecting to Multiple ISPs: Multihomed

- Multihomed benefits include:
  - Achieving an ISP-independent solution.
  - Scalability of the solution, beyond two ISPs.
  - Resistance to a failure to a single ISP.
  - Load sharing for different destination networks between ISPs.



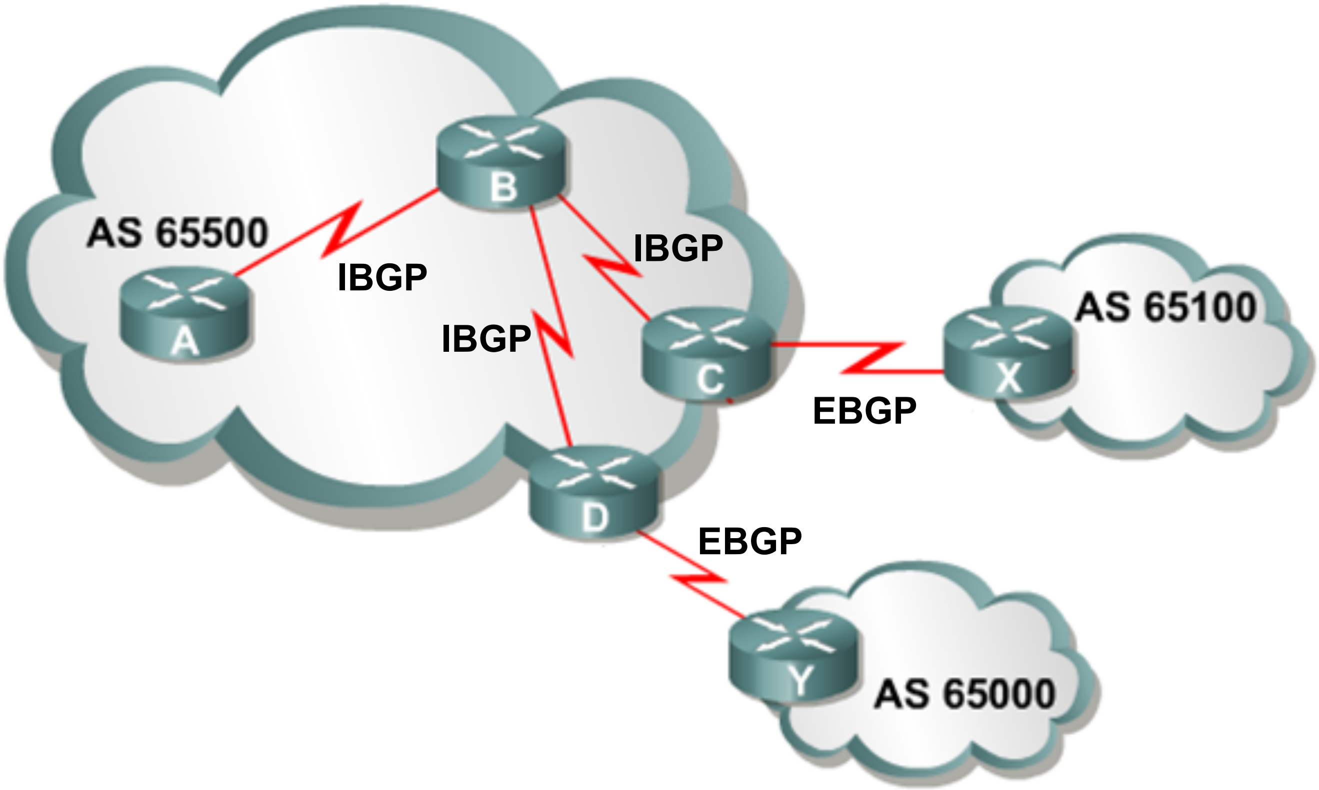
Connecting Multiple ISPs: Dual-Multihomed

- Dual multihomed includes all the benefits of multihomed connectivity, with enhanced resiliency.
- The configuration typically has multiple edge routers, one per ISP, and uses BGP.



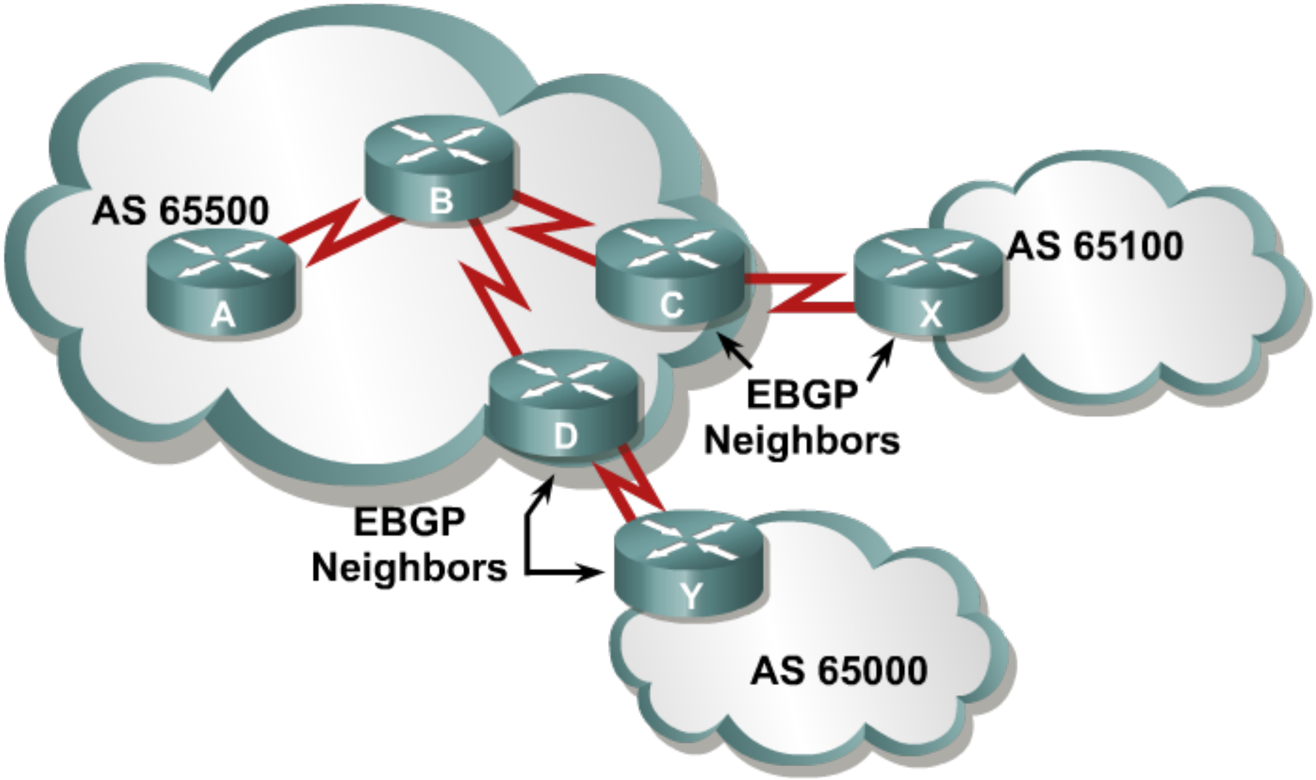
Using BGP in an Enterprise Network

- When BGP is running between routers in different AS, it is called **External BGP (EBGP)**.
- When BGP is running between routers in the same AS, it is called **Internal BGP (IBGP)**.



External BGP

- EBGP neighbors are in different autonomous systems.
  - EBGP neighbors need to be directly connected.

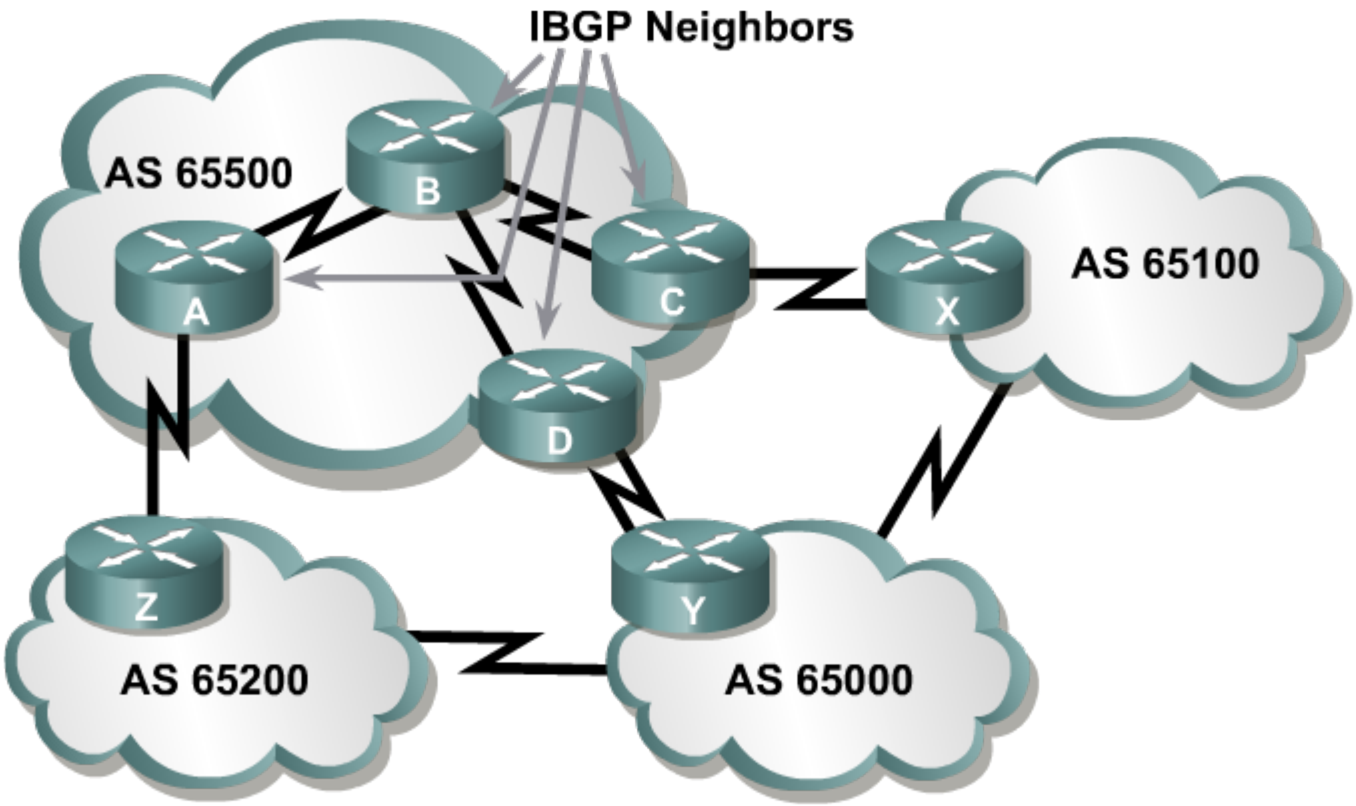


EBGP Neighbor Relationship Requirements

- Define neighbors:
  - A TCP session (three-way handshake) must be established before starting BGP routing update exchanges.
- Reachability:
  - EBGP neighbors are usually directly connected.
- Different AS number:
  - EBGP neighbors must have different AS numbers.

Internal BGP

- IBGP neighbors are in the same autonomous systems.
  - IBGP neighbors do not need to be directly connected.

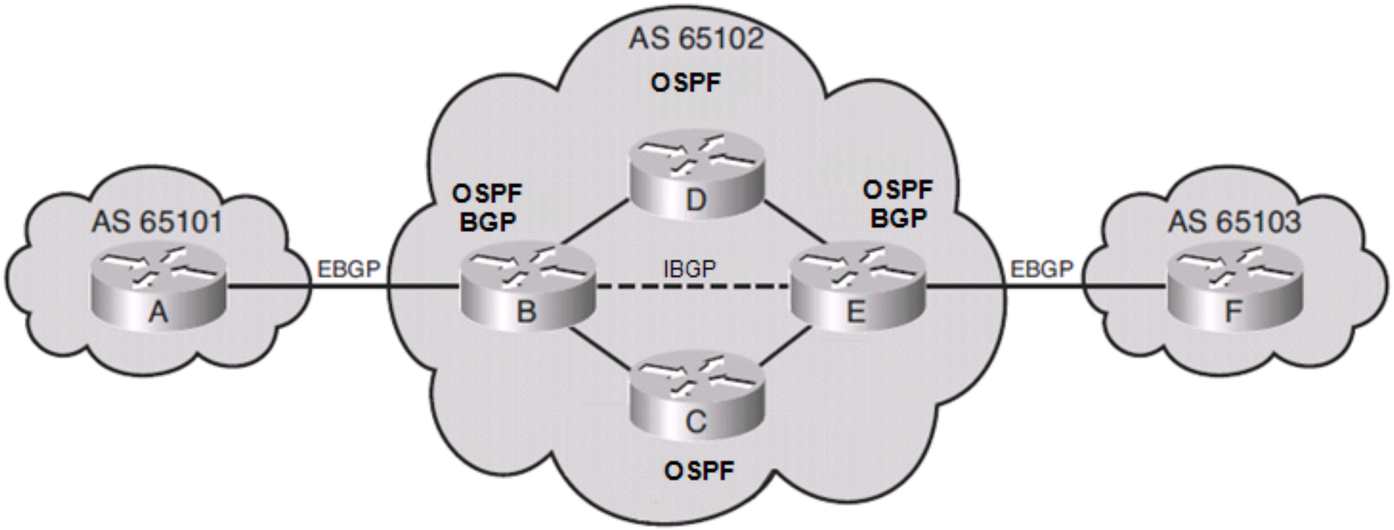


IBGP Neighbor Relationship Requirements

- **Define neighbors:**
  - A TCP session (three-way handshake) must be established before starting BGP routing update exchanges.
- **Reachability:**
  - IBGP neighbors must be reachable usually by using an IGP.
  - Loopback IP addresses are typically used to identify IBGP neighbors.
- **Same AS number:**
  - IBGP neighbors must have the same AS number.

IBGP in a Transit AS

- A transit AS is an AS that routes traffic from one external AS to another external AS.
- In this example, AS 65102 is a service provider network.
  - Only the two edge routers (router B and E) are running BGP and have established an IBGP neighbor relationship using OSPF.
  - Although the EBGP routes could be redistributed into OSPF, the potential number of BGP routes may overwhelm OSPF and is therefore not recommended.

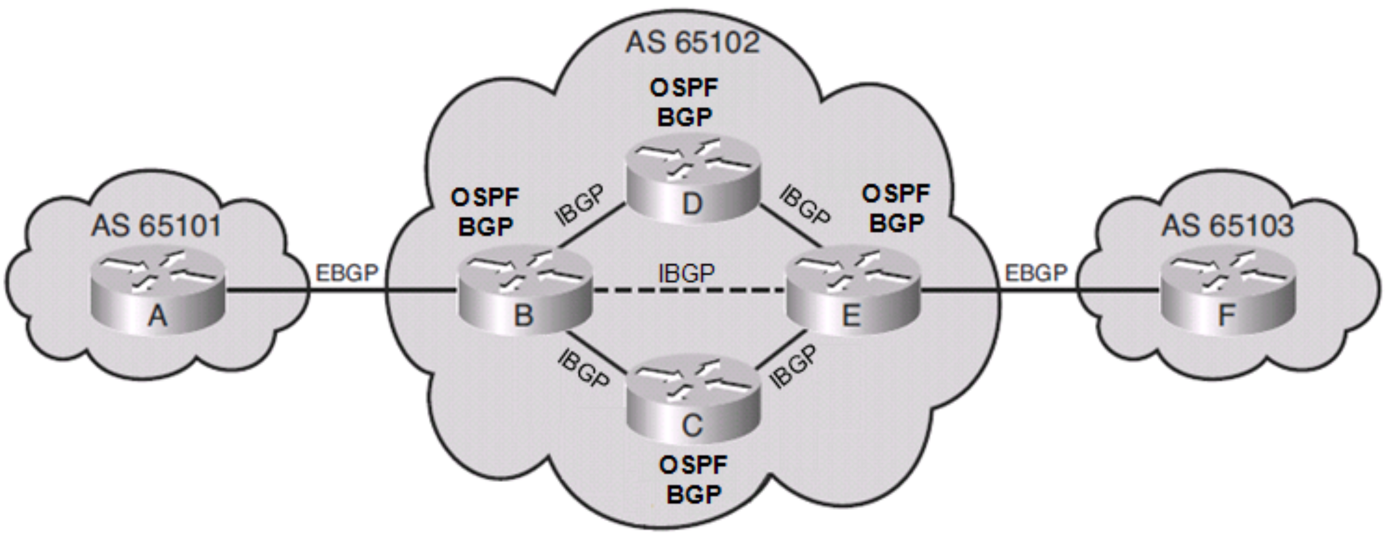


BGP Synchronization

- The BGP synchronization rule states that:
  - *“A BGP router should not use, or advertise to an external neighbor, a route learned by IBGP, unless that route is local or is learned from the IGP.”*
  - If synchronization is enabled, a router learning a route via IBGP waits until the IGP has propagated the route within the autonomous system and then advertises it to external peers.
  - With the default of synchronization disabled, BGP can use and advertise to external BGP neighbors routes learned from an IBGP neighbor that are not present in the local routing table.
- BGP synchronization is disabled by default in Cisco IOS Software Release 12.2(8)T and later.
  - It was on by default in earlier Cisco IOS Software releases.

IBGP in a Transit AS

- A better solution for a provider network would be to have a fully meshed BGP internetwork.
  - BGP runs on all internal routers and all routers establish IBGP sessions.
  - IBGP routers have complete knowledge of external routes.



IBGP in a Nontransit AS

- A nontransit AS is an AS that does not route traffic from one external AS to another external AS.
  - Nontransit AS networks are typically enterprise networks.

- All routers in a nontransit AS must still have complete knowledge of external routes.

- To avoid routing loops within an AS, BGP specifies that routes learned through IBGP are never propagated to other IBGP peers.
  - It is assumed that the sending IBGP neighbor is fully meshed with all other IBGP speakers and has sent each IBGP neighbor the update.

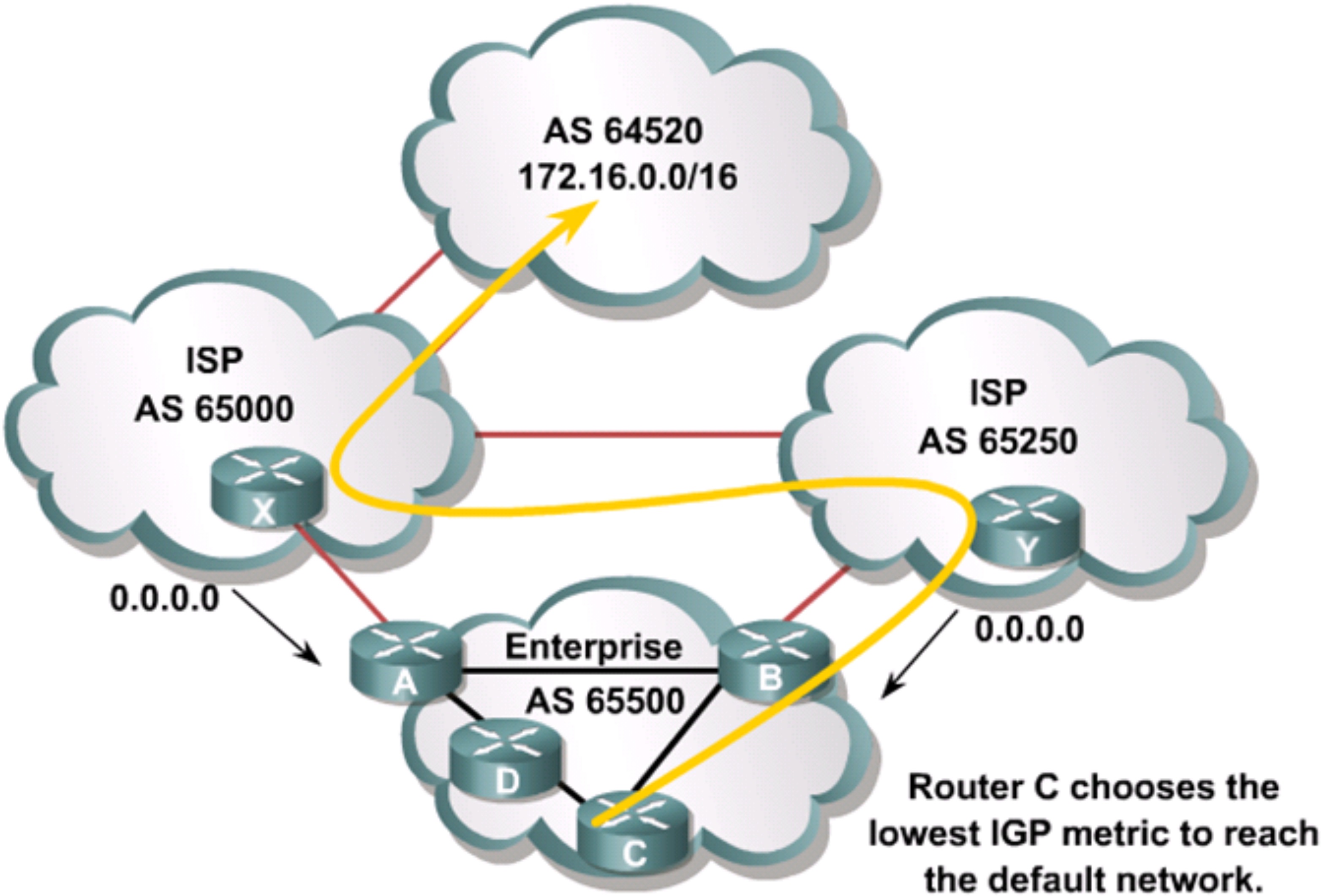
Three Multihoming Connection Options

1. Each ISP passes only a default route to the AS.
   - The default route is passed on to internal routers.

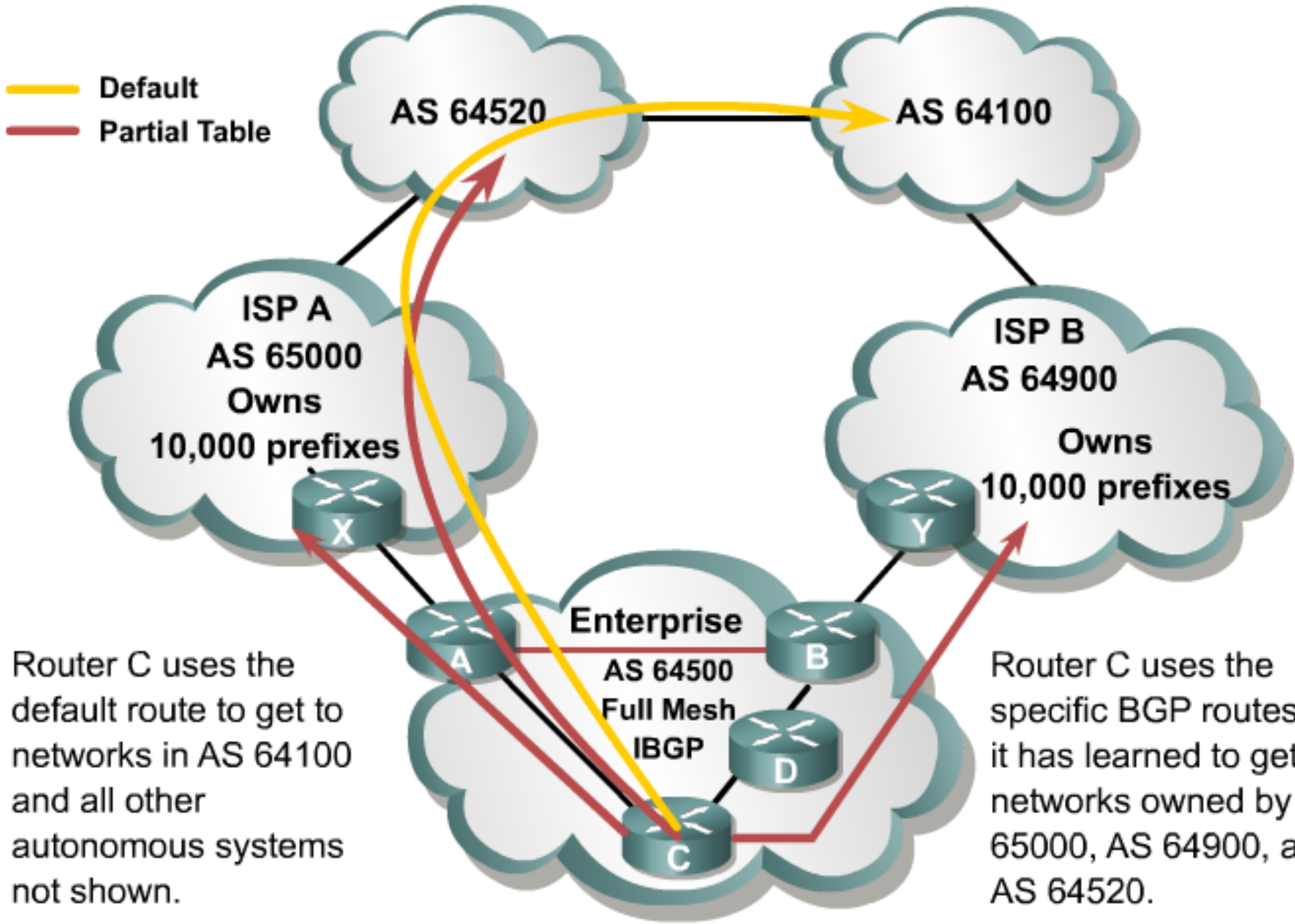
2. Each ISP passes only a default route and provider-owned specific routes to the AS.
   - These routes may be propagated to internal routers, or all internal routers in the transit path can run BGP to exchange these routes.

3. Each ISP passes all routes to the AS.
   - All internal routers in the transit path run BGP to exchange these routes.

Default Routes from All Providers



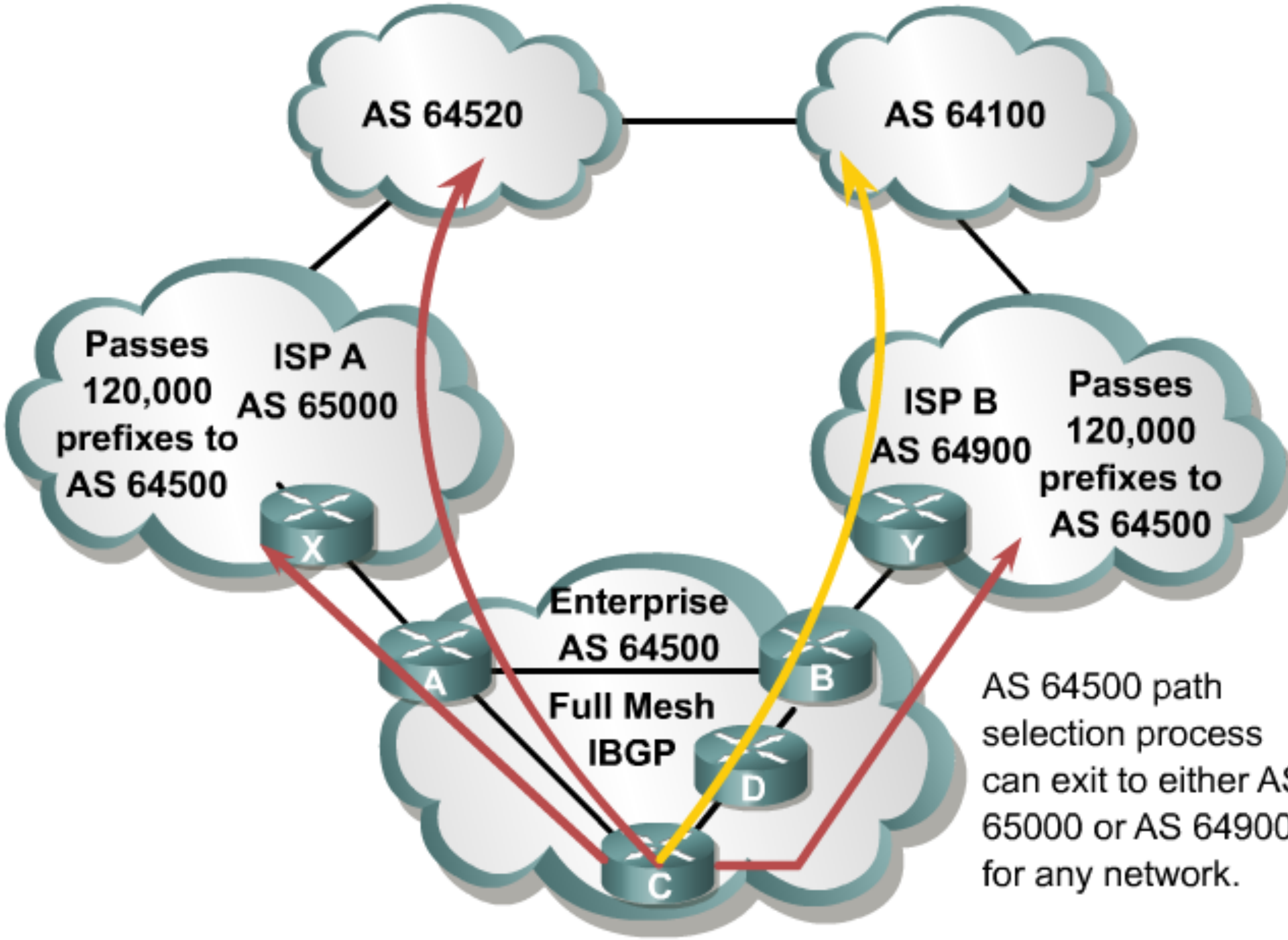
Default Routes and Partial Updates



Router C uses the default route to get to networks in AS 64100 and all other autonomous systems not shown.

Router C uses the specific BGP routes that it has learned to get to networks owned by AS 65000, AS 64900, and AS 64520.

Full Routes from All Providers



AS 64500 path selection process can exit to either AS 65000 or AS 64900 for any network.

When to Use BGP

- Most appropriate when the effects of BGP are well-understood and at least one of the following conditions exists:
  - The AS has multiple connections to other autonomous systems.
  - The AS allows packets to transit through it to reach other autonomous systems (eg, it is a service provider).
  - Routing policy and route selection for traffic entering and leaving the AS must be manipulated.

When Not to Use BGP

- Do not use BGP if one or more of the following conditions exist:
  - A single connection to the Internet or another AS.
  - Lack of memory or processor power on edge routers to handle constant BGP updates.
  - You have a limited understanding of route filtering and the BGP path-selection process.
- In these cases, use static or default routes instead.

BGP Table

- BGP keeps its own table for storing BGP information received from and sent to BGP neighbors.

  - This table is also known as the BGP table, BGP topology table, BGP topology database, BGP routing table, and the BGP forwarding database.

- The router offers the best routes from the BGP table to the IP routing table.

BGP Tables

- Neighbor table
  - List of BGP neighbors
- BGP table (forwarding database)
  - List of all networks learned from each neighbor
  - Can contain multiple paths to destination networks
  - Contains BGP attributes for each path
- IP routing table
  - List of best paths to destination networks

BGP Message Types

- There are four different BGP message types:

Open Message

Octets	16	2	1	1	2	2	4	1	7
	Marker	Length	Type	Version	AS	Hold Time	BGP ID	Optional Length	Optional

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes length	Withdrawn Routes	Attribute Length	Attributes	NLRI

Notification Message

Octets	16	2	1	1	1	Variable
	Marker	Length	Type	Error Code	Error Sub-code	Diagnostic Data

Keepalive Message

Octets	16	2	1
	Marker	Length	Type

Open Message

- Once a TCP connection has been established, the Open message is sent and includes a set of parameters that have to be agreed upon before a full BGP adjacency can be established.
- Once both BGP peers have agreed upon mutual capabilities, they can start exchanging routing information by means of BGP Update messages.

Open Message

Octets	16	2	1	1	2	2	4	1	7
	Marker	Length	Type	Version	AS	Hold Time	BGP ID	Optional Length	Optional

Update Message

- Update messages contain all the information BGP uses to construct a loop-free picture of the internetwork.
- A BGP update message has information on one path only; multiple paths require multiple update messages.
  - All the attributes in the update message refer to that path, and the networks are those that can be reached through it.

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes Length	Withdrawn Routes	Attribute Length	Path Attributes	NLRI

Update Message

- An update message includes the following information:
  - Unreachable routes information
  - Path attribute information
  - Network-layer reachability information (NLRI)
    - This field contains a list of IP address prefixes that are reachable by this path.

Update Message

				Unreachable Routes Information		Path Attributes Information		NLRI Information
Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes Length	Withdrawn Routes	Attribute Length	Path Attributes	NLRI

NLRI (Network-layer reachability information) format

- The NLRI is a list of <**length**, **prefix**> tuples.
  - One tuple for each reachable destination.
  - The **prefix** represents the reachable destination
  - The prefix **length** represents the # of bits set in the subnet mask.

IP Address Subnet Mask	NLRI
10.1.1.0 255.255.255.0	**24, 10.1.1.0**
192.24.160.0 255.255.224.0	**19, 192.24.160.0**

Notification Message

- A BGP notification message is sent when an error condition is detected.
  - The BGP connection is closed immediately after this is sent.
- Notification messages include an error code, an error subcode, and data related to the error.

Notification Message

Octets	16	2	1	1	1	Variable
	Marker	Length	Type	Error Code	Error Sub-code	Diagnostic Data

Notification Message

- Sample error codes and their associated subcodes.

Error Code	Error Subcode
1--Message Header Error	1--Connection Not Synchronized 2--Bad Message Length 3--Bad Message Type
2--OPEN Message Error	1--Unsupported Version Number 2--Bad Peer AS 3--Bad BGP Identifier 4--Unsupported Optional Parameter 5--Authentication Failure 6--Unacceptable Hold Time
3--UPDATE Message Error	1--Malformed Attribute List 2--Unrecognized Well-Known Attribute 3--Missing Well-Known Attribute 4--Attribute Flags Error 5--Attribute Length Error 6--Invalid Origin Attribute 7--AS Routing Loop 8--Invalid NEXT_HOP Attribute 9--Optional Attribute Error 10--Invalid Network Field 11--Malformed AS_path
4--Hold Timer Expired	NOT applicable
5--Finite State Machine Error (for errors detected by the FSM)	NOT applicable
6--Cease (for fatal errors besides the ones already listed)	NOT applicable

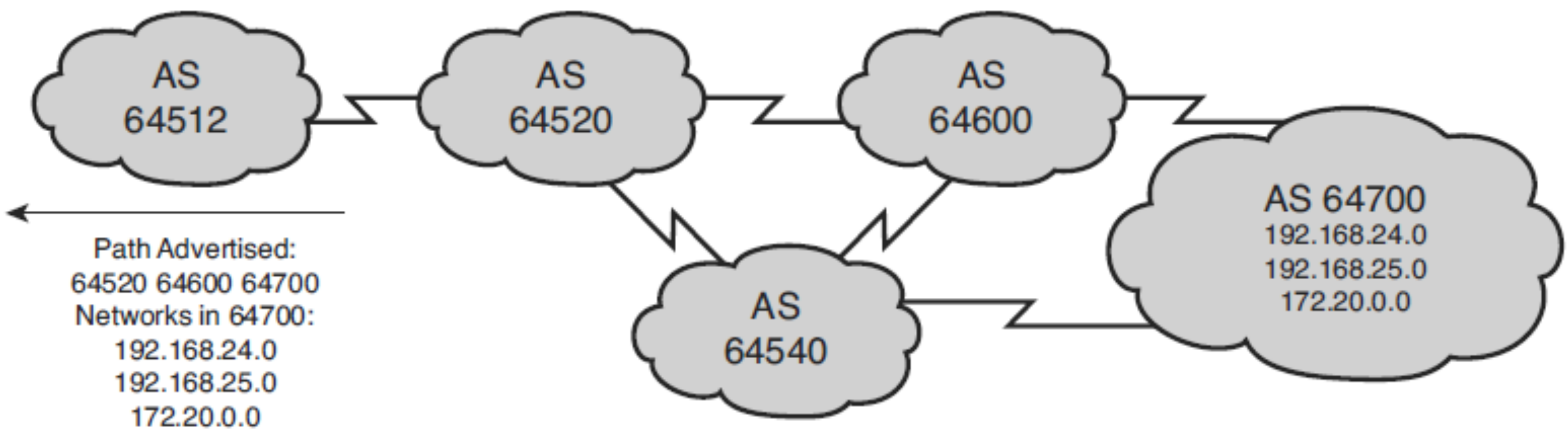
BGP Path Vector Characteristics

- Internal routing protocols announce a list of networks and the metrics to get to each network.

- In contrast, BGP routers exchange network reachability information, called **path vectors**, made up of **path attributes**.

BGP Path Vector Characteristics

- The path vector information includes:
  - A list of the full path of BGP AS numbers (hop by hop) necessary to reach a destination network.
  - Other attributes including the IP address to get to the next AS (the next-hop attribute) and how the networks at the end of the path were introduced into BGP (the origin code attribute).



Path Attributes

- Path attributes are a set of BGP metrics describing the path to a network (route).
  - BGP uses the path attributes to determine the best path to the networks.
  - Some attributes are mandatory and automatically included in update messages while others are manually configurable.

- BGP attributes can be used to enforce a routing policy.

- Configuring BGP attributes provides administrators with many more path control options.
  - E.g., filter routing information, prefer certain paths, customize BGP’s behavior.

Path Attributes

- A BGP update message includes a variable-length sequence of path attributes describing the route.
- A path attribute consists of three fields:
  - Attribute type
  - Attribute length
  - Attribute value

BGP Attribute Type

- Type code 1 ORIGIN
- Type code 2 AS_PATH
- Type code 3 NEXT_HOP
- Type code 4 MULTI_EXIT_DISC
- Type code 5 LOCAL_PREF
- Type code 6 ATOMIC_AGGREGATE
- Type code 7 AGGREGATOR
- Type code 8 Community (Cisco-defined)
- Type code 9 Originator-ID (Cisco-defined)
- Type code 10 Cluster list (Cisco-defined)

Path Attributes Information

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes Length	Withdrawn Routes	Attribute Length	Path Attributes	NLRI

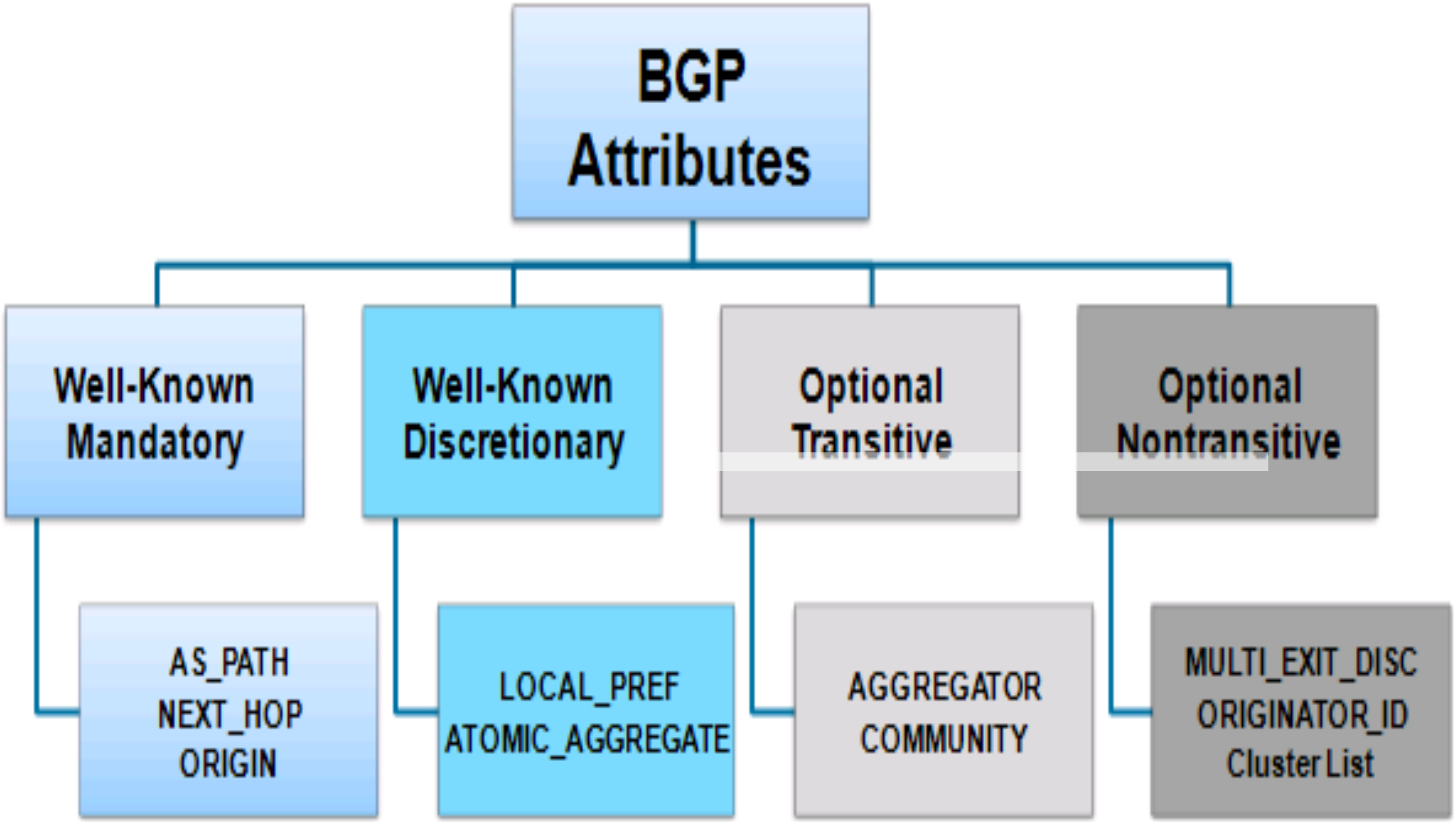
Attributes

- Some attributes are mandatory and automatically included in update messages while others are manually configurable.

Attribute	EBGP	IBGP	
AS_PATH	Well-known Mandatory	Well-known Mandatory	Automatically included in update message
NEXT_HOP	Well-known Mandatory	Well-known Mandatory	
ORIGIN	Well-known Mandatory	Well-known Mandatory	
LOCAL_PREF	Not allowed	Well-known Discretionary	Can be configured to help provide path control.
ATOMIC_AGGREGATE	Well-known Discretionary	Well-known Discretionary	
AGGREGATOR	Optional Transitive	Optional Transitive	
COMMUNITY	Optional Transitive	Optional Transitive	
MULTI_EXIT_DISC	Optional Nontransitive	Optional Nontransitive	

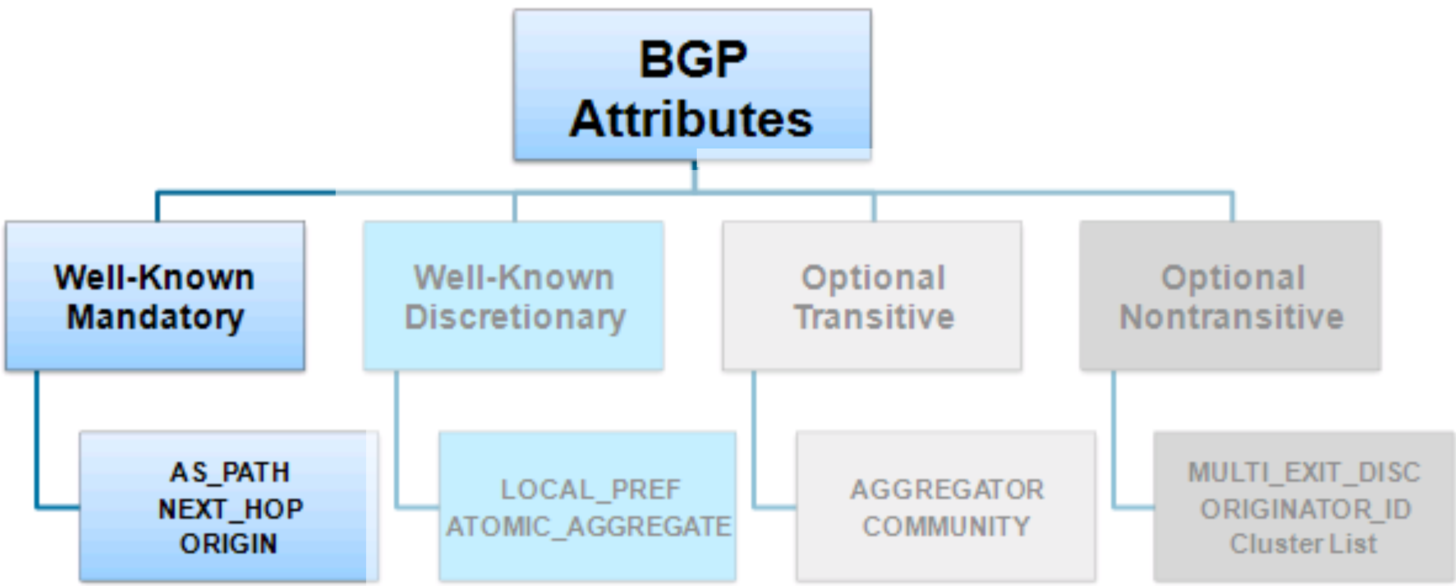
Path Attributes

- There are four different attribute types.
  - Not all vendors recognize the same BGP attributes.



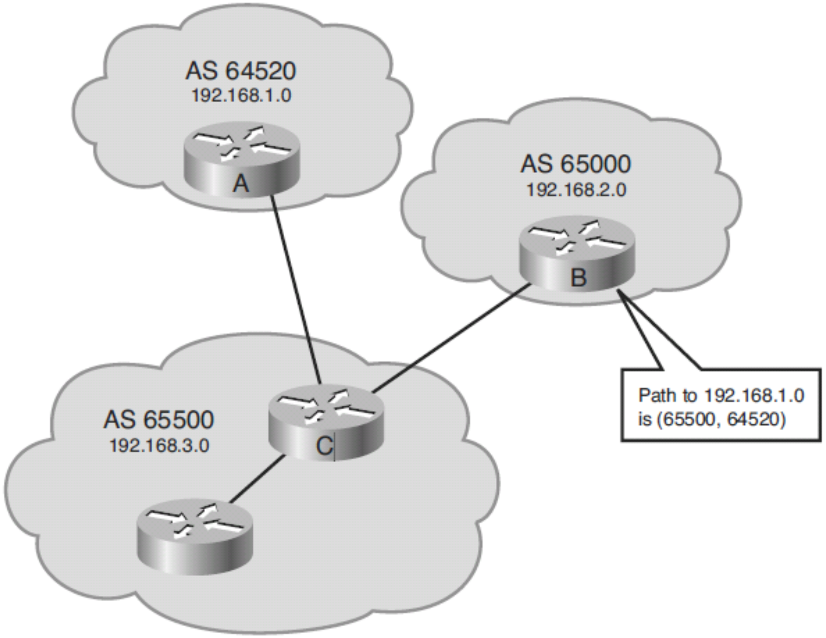
Well-Known Mandatory

- Attribute is recognized by all implementations of BGP and must appear in a BGP update message.
  - If missing, a notification error will be generated.
- Well-known mandatory attributes ensures that all BGP implementations agree on a standard set of attributes.



Well-Known Mandatory: AS_PATH

- The AS_PATH attribute contains a list of AS numbers to reach a route.
- Whenever a route update passes through an AS, the AS number is added to the beginning of the AS_PATH attribute before it is advertised to the next EBGP neighbor.



Well-Known Mandatory: NEXT_HOP

- The NEXT_HOP attribute indicates the IP address that is to be used to reach a destination.

- The IP address is the entry point of the next AS along the path to that destination network.

  - Therefore, for EBGP, the next-hop address is the IP address of the neighbor that sent the update.

Well-Known Mandatory: ORIGIN

- The ORIGIN attribute defines the origin of the path which could be:
  - **IGP**:
    - The route is interior to the originating AS and normally occurs when a `network` command is used to advertise the route via BGP.
    - An origin of IGP is indicated with an “**i**” in the BGP table.
  - **EGP**:
    - (Obsolete) The route is learned via EGP which is considered a historic routing protocol and is not supported on the Internet.
    - An origin of EGP is indicated with an “**e**” in the BGP table.
  - **Incomplete**:
    - The route’s origin is unknown or is learned via some other means and usually occurs when a route is redistributed into BGP.
    - An incomplete origin is indicated with a “**?**” in the BGP table.

Well-Known Mandatory: ORIGIN

```
R1# show ip bgp
BGP table version is 24, local router ID is 172.16.1.2
Status codes: s suppressed, d damped, h history, * valid, > b
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric  LocPrf  Weight  Path
*> 192.208.10.0     192.208.10.5             0               0  300 i
*> 172.16.1.0       0.0.0.0                  0           32768 i
<output omitted>
```

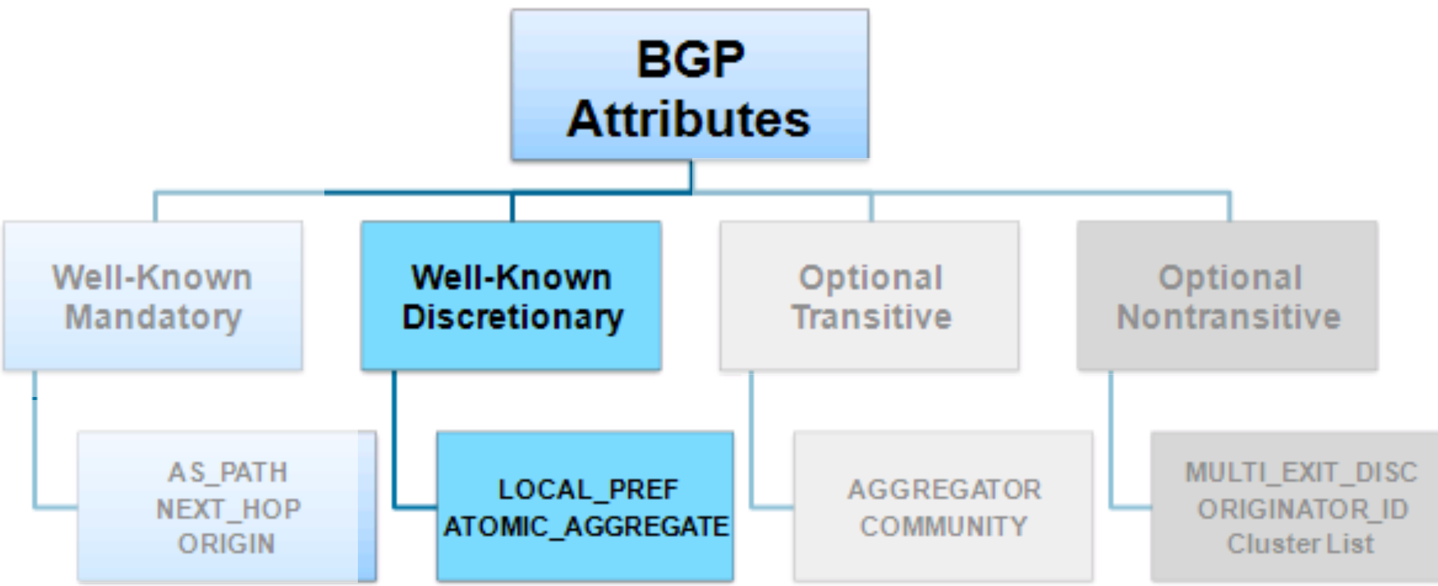
i = Route generated by the **network** command.

```
R1# show ip bgp
<output omitted>
   Network          Next Hop              Metric  LocPrf  Weight  Path
*> 10.1.1.0/24       0.0.0.0                   0           32768  ?
*> 192.168.1.0/24   10.1.1.2                  84           32768  ?
*> 192.168.2.0/24   10.1.1.2                  74           32768  ?
<output omitted>
```

? = Route generated by unknown method (usually redistributed).

Well-Known Discretionary

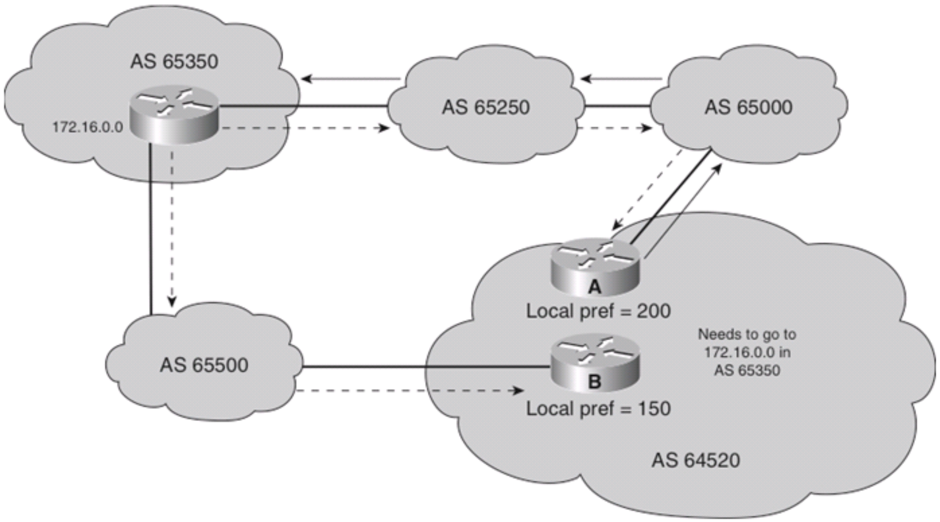
- Attribute is recognized by all implementations of BGP but may not be sent in the BGP update message.



Well-Known Discretionary: LOCAL_PREF

- The Local Preference attribute provides an indication to the “local” routers in the AS about which path is preferred to exit the AS.
  - A path with a higher local preference is preferred.
  - The default value for local preference on a Cisco router is 100.
- It is configured on a router and exchanged between IBGP routers.
  - It is not passed to EBGP peers.

Well-Known Discretionary: LOCAL_PREF



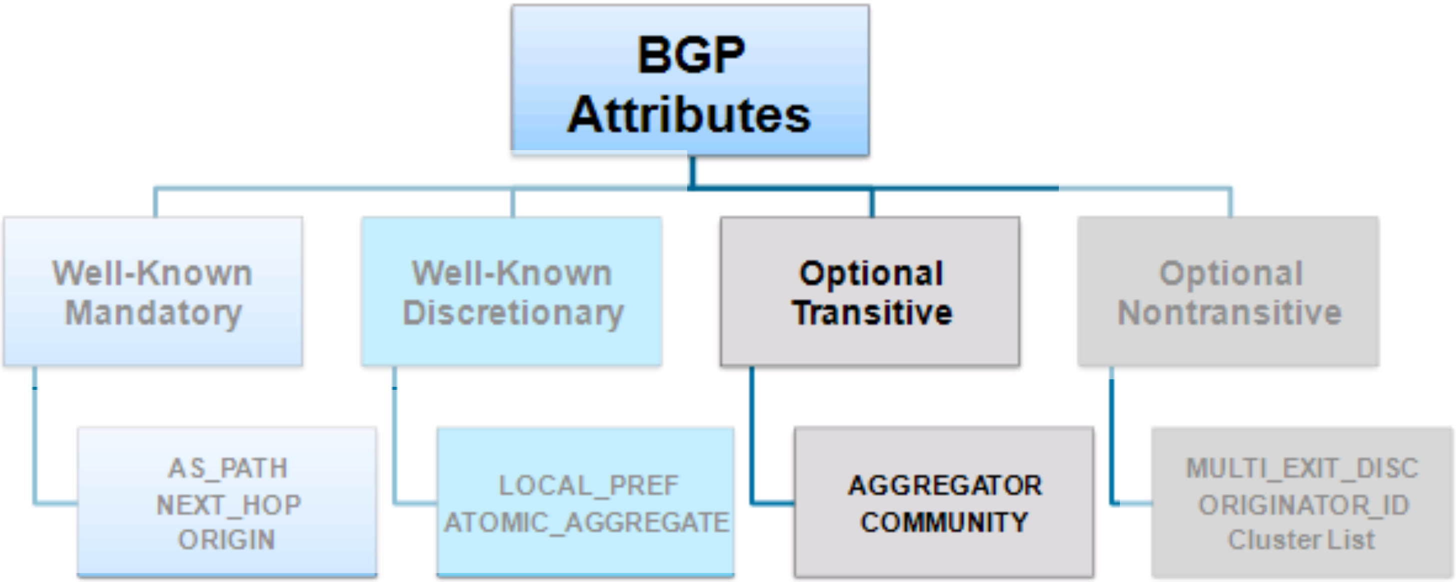
- Routers A and B are IBGP neighbors in AS 64520 and both receive updates about network 172.16.0.0 from different directions.
  - The local preference on router A is set to 200.
  - The local preference on router B is set to 150.
- Because the local preference for router A is higher, it is selected as the preferred exit point from AS 64520.

Well-Known Discretionary: ATOMIC_AGGREGATE

- The Atomic Aggregate attribute is used to indicate that routes have been summarized.
  - Attribute warns that the received information may not necessarily be the most complete route information available.
- Attribute is set to either True or False with “true” alerting other BGP routers that multiple destinations have been grouped into a single update.
  - Router update includes its router ID and AS number along with the supernet route enabling administrators to determine which BGP router is responsible for a particular instance of aggregation.
  - Tracing a supernet to its original "aggregator" may be necessary for troubleshooting purposes.

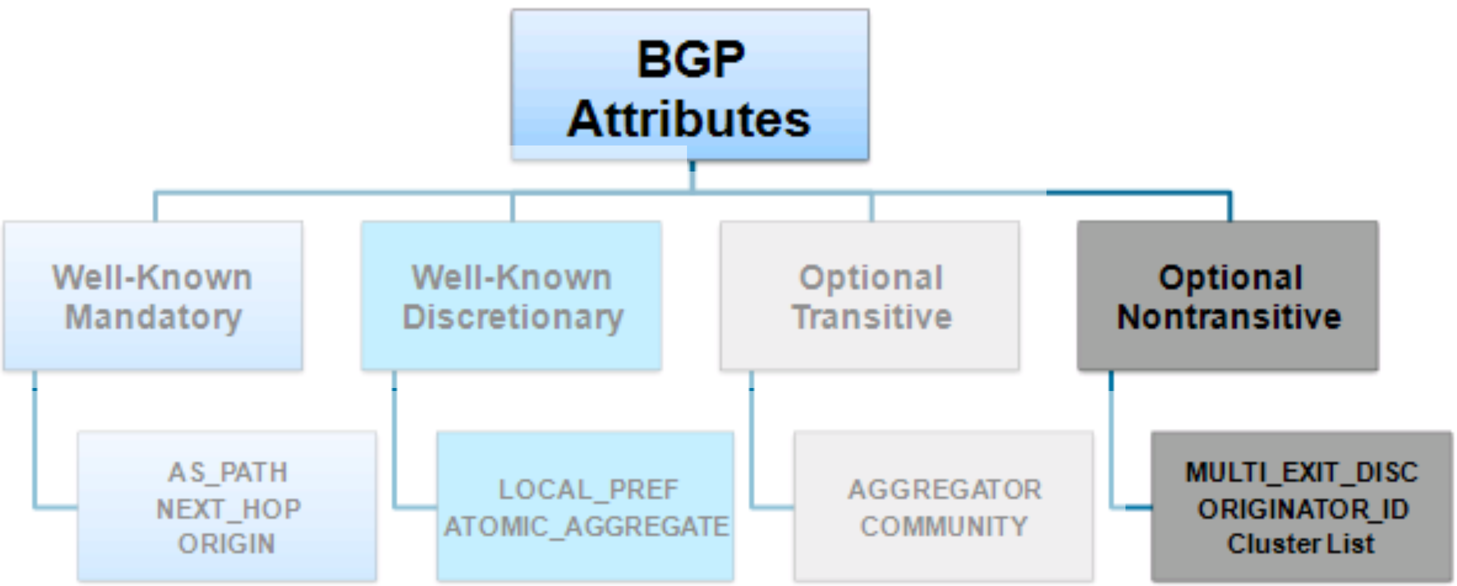
Optional Transitive

- Attribute may or may not be recognized by all BGP implementations.
- Because the attribute is transitive, BGP accepts and advertises the attribute even if it is not recognized.



Optional Nontransitive

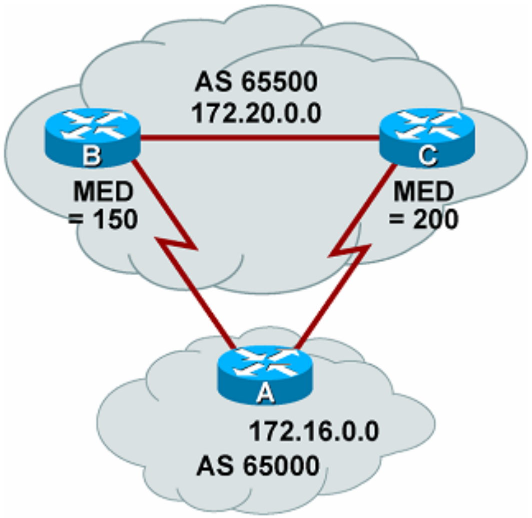
- Attribute that may or may not be recognized by all BGP implementations.
- Whether or not the receiving BGP router recognizes the attribute, it is nontransitive and is not passed along to other BGP peers.



Optional Nontransitive: MED

- The Multiple Exit Discriminator (MED) attribute, also called the *metric,* provides a hint to external neighbors about the preferred path into an AS that has multiple entry points.
  - Lower MED is preferred over a higher MED!
- The MED is sent to EBGP peers and those routers propagate the MED within their AS.
  - The routers within the AS use the MED, but do not pass it on to the next AS.
  - When the same update is passed on to another AS, the metric will be set back to the default of 0.
- By using the MED attribute, BGP is the only protocol that can affect how routes are sent into an AS.

Optional Nontransitive: MED

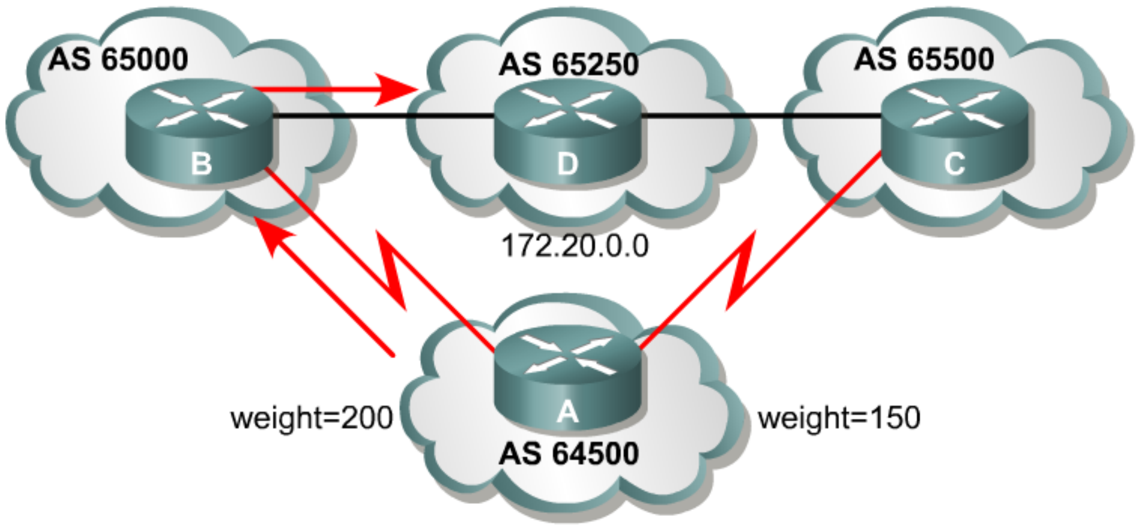


- Routers B and C include a MED attribute in the updates to router A.
  - Router B MED attribute is set to 150.
  - Router C MED attribute is set to 200.
- When A receives updates from B and C, it picks router B as the best next hop because of the lower MED.

Cisco Weight Attribute

- The Weight attribute is a Cisco proprietary attribute.
- Similar in function to the local preference, the weight attribute applies when 1 router has multiple exit points.
  - Local preference is used when 2+ routers provide multiple exit points.
- It is configured locally on a router and is not propagated to any other routers.
  - Routes with a higher weight are preferred when multiple routes exist to the same destination.
- The weight can have a value from 0 to 65535.
  - Paths that the router originates have a weight of 32768 by default, and other paths have a weight of 0 by default.

Cisco Weight Attribute



- Routers B and C learn about network 172.20.0.0 from AS 65250 and propagate the update to router A.
  - Therefore Router A has two ways to reach 172.20.0.0.
- Router A sets the weight of updates as follows:
  - Updates coming from router B are set to 200
  - Updates coming from router C are set to 150.
- Router A uses router B because of the higher weight.

BGP Route Selection Process

1. Prefer the route with highest weight.
2. Prefer the route with highest LOCAL_PREF.
3. Prefer the locally generated route (network or aggregate routes).
4. Prefer the route with the shortest AS-PATH.
5. Prefer the route with the lowest ORIGIN (IGP<EGP<incomplete)
6. Prefer the route with the lowest MED.
7. Prefer the EBGP route over IBGP route.
8. Prefer the route through the closest IGP neighbor.
9. Prefer the oldest EBGP route.
10. Prefer the route with the lowest neighbor BGP router ID value.
11. Prefer the route with the lowest neighbor IP address.

Configuring BGP

Implementing Basic BGP

- The information necessary to implement BGP routing includes the following:
  - The AS numbers of enterprise and service provider.
  - The IP addresses of all the neighbors (peers) involved.
  - The networks that are to be advertised into BGP

- In the implementation plan, basic BGP tasks include the following:
  - Define the BGP process
  - Establish the neighbor relationships
  - Advertise the networks into BGP

Verifying BGP

- After implementing BGP, verification should confirm proper deployment on each router.
- Verification tasks include verifying:
  - That the appropriate BGP neighbor relationships and adjacencies are established.
  - That the BGP table is populated with the necessary information.
  - That IP routing table is populated with the necessary information.
  - That there is connectivity in the network between routers and to other devices.
  - That BGP behaves as expected in a case of a topology change, by testing link failure and router failure events.

Enable BGP Routing

- Define BGP as the IP routing protocol.

```
Router(config)#
router bgp autonomous-system
```

- The *autonomous-system* value is either an internally generated number (if not connecting to a provider network) or obtained from an ISP or RIR.
  - It is a required parameter.
  - It can be any positive integer in the range from 1 to 65535.
- Only one instance of BGP can be configured on the router at a single time.

Defining BGP Neighbors

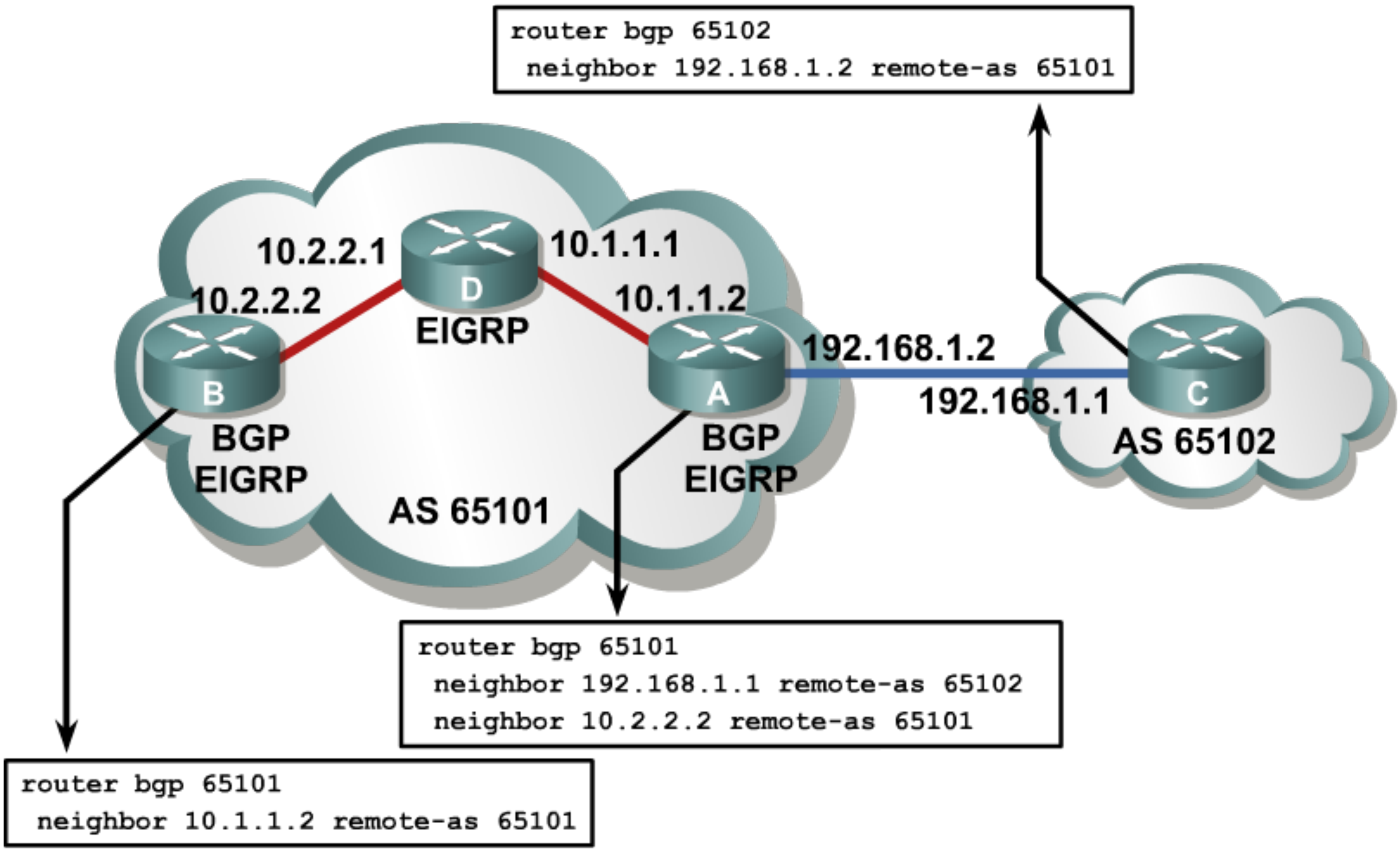
- Identify peer router with which to establish a BGP session.

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

```
neighbor {ip-address | peer-group-name} remote-as
  autonomous-system
```

- The *ip-address* is the destination address of the BGP peer.
  - The address must be reachable before attempting to establish the BGP relationship.
- The *autonomous-system* value is used to identify if the session is with internal BGP (IBGP) peers or with external BGP (EBGP) peers.
  - If the value is the same as the router's AS, then an IBGP session is attempted.
  - If the value is not the same as the router's AS, then an EBGP session is attempted.

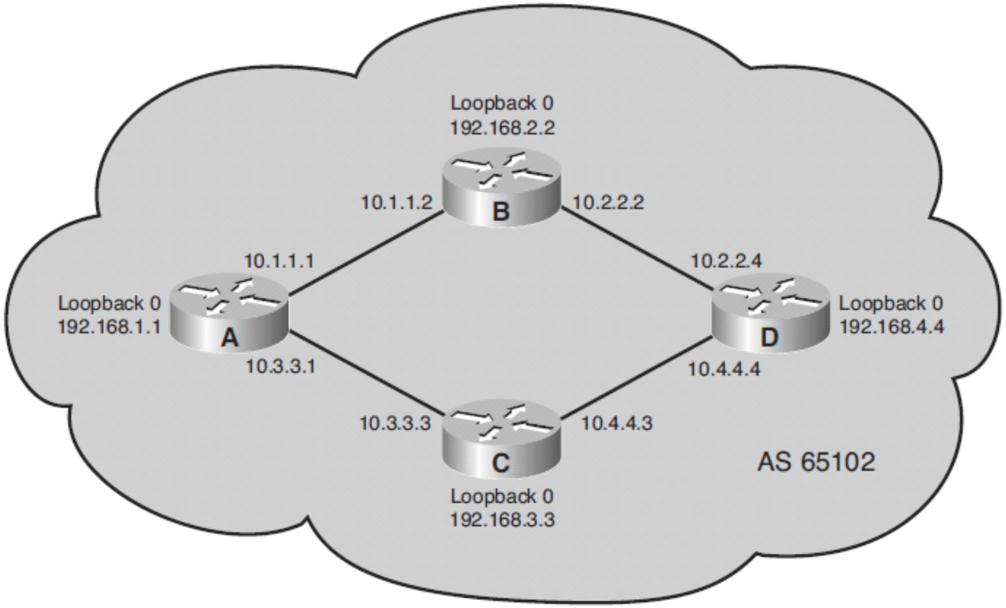
Example: BGP `neighbor` Command



IBGP Source IP Address Problem

- BGP does not accept unsolicited updates.
  - It must be aware of every neighboring router and have a **`neighbor`** statement for it.
- For example, when a router creates and forwards a packet, the IP address of the outbound interface is used as that packet's source address by default.
  - For BGP packets, this source IP address must match the address in the corresponding **`neighbor`** statement on the other router or the routers will not establish the BGP session.
  - This is not a problem for EBGP neighbors as they are typically directly connected.

IBGP Source IP Address Problem



- When multiple paths exist between IBGP neighbors, the BGP source address can cause problems:
  - Router D uses the **`neighbor 10.3.3.1 remote-as 65102`** command to establish a relationship with A.
  - However, router A is sending BGP packets to D via B therefore the source IP address of the packets is 10.1.1.1.
  - The IBGP session between A and D cannot be established because D does not recognize 10.1.1.1 as a BGP neighbor.

IBGP Source IP Address Solution

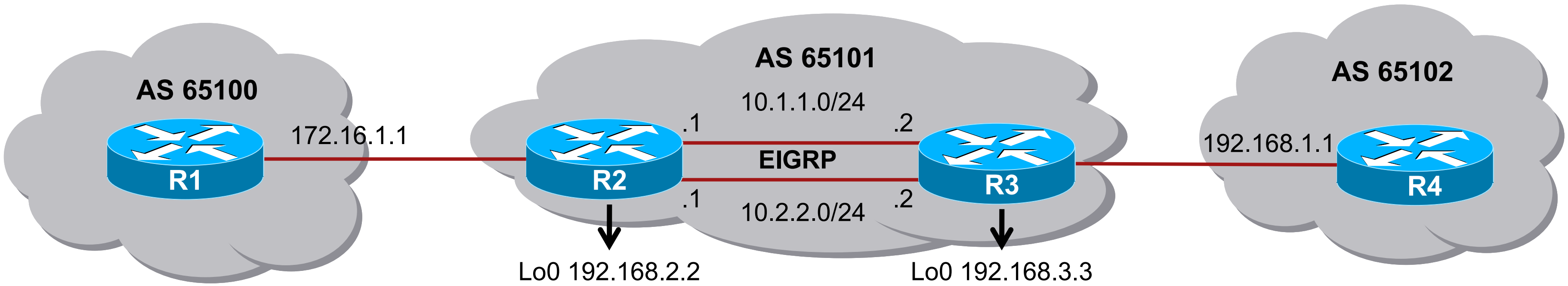
- Establish the IBGP session using a loopback interface.

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

> **neighbor** {*ip-address* | *peer-group-name*} **update-source**
> *loopback interface-number*

- Informs the router to use a loopback interface address for all BGP packets.
- Overrides the default source IP address for BGP packets.
- Typically only used with IBGP sessions.
- As an added bonus, physical interfaces can go down for any number of reasons but loopbacks never fail.

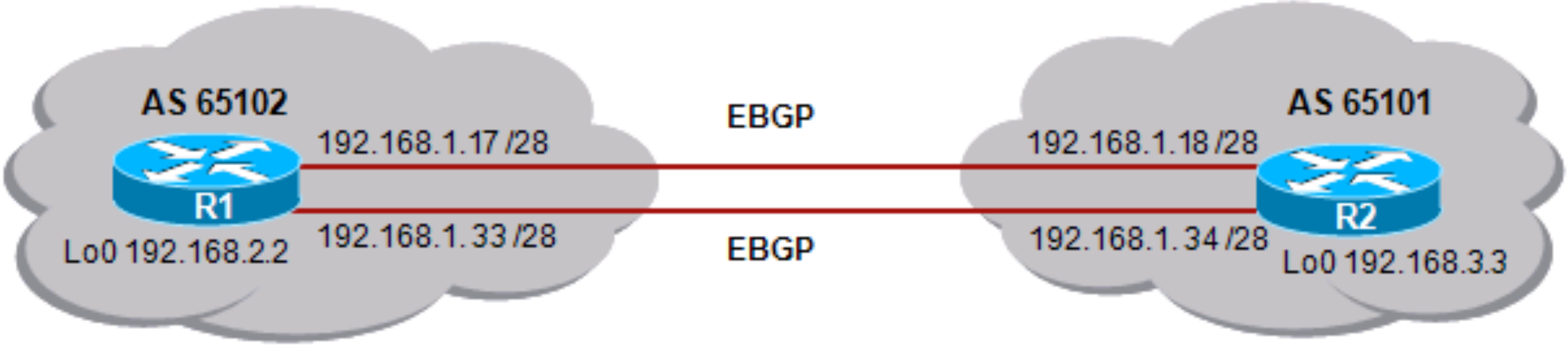
IBGP Source IP Address Example




```
R2(config)# router bgp 65101
R2(config-router)# neighbor 172.16.1.1 remote-as 65100
R2(config-router)# neighbor 192.168.3.3 remote-as 65101
R2(config-router)# neighbor 192.168.3.3 update-source loopback0
R2(config-router)# exit
R2(config)# router eigrp 1
R2(config-router)# network 10.0.0.0
R2(config-router)# network 192.168.2.0
R2(config-router)#
```

```
R3(config)# router bgp 65101
R3(config-router)# neighbor 192.168.1.1 remote-as 65102
R3(config-router)# neighbor 192.168.2.2 remote-as 65101
R3(config-router)# neighbor 192.168.2.2 update-source loopback0
R3(config-router)# exit
R3(config)# router eigrp 1
R3(config-router)# network 10.0.0.0
R3(config-router)# network 192.168.3.0
R3(config-router)#
```

EBGP Dual-Homed Problem



- R1 in AS 65102 is dual-homed with R2 in AS 65101.
- A problem can occur if R1 only uses a single **`neighbor`** statement pointing to 192.168.1.18 on R2 .
  - If that link fails, the BGP session between these AS is lost, and no packets pass from one autonomous system to the next, even though another link exists.
- A solution is configuring two **`neighbor`** statements on R1 pointing to 192.168.1.18 and 192.168.1.34.
  - However, this doubles the BGP updates from R1 to R2.

EBGP Dual-Homed Solution

- The ideal solution is to:
  - Use loopback addresses.
  - Configure static routes to reach the loopback address of the other router.
  - Configure the **`neighbor ebgp-multihop`** command to inform the BGP process that this neighbor is more than one hop away.

Enable Multihop EBGP

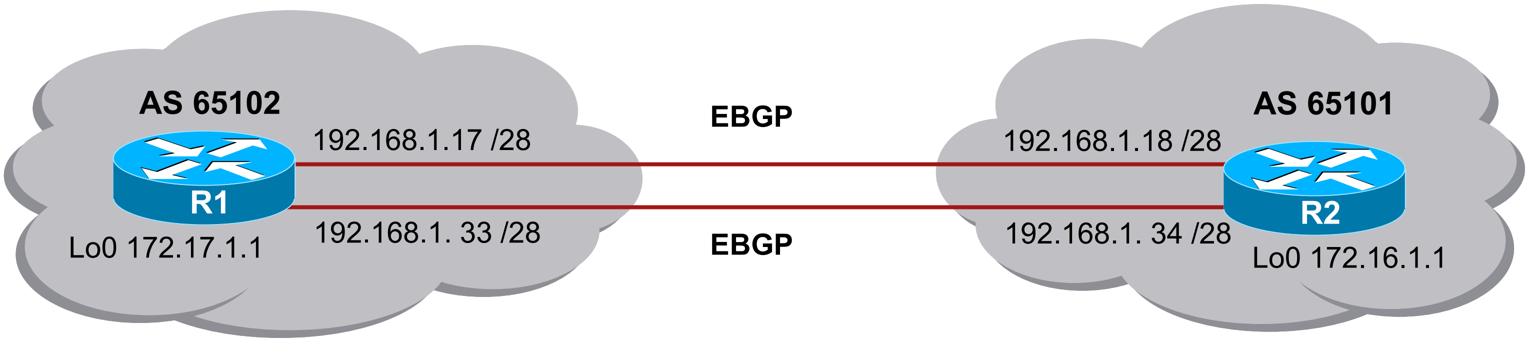
- Increase the time-to-live (TTL) for EBGP connections.

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

```
neighbor {ip-address | peer-group-name} ebgp-multihop
  [ttl]
```

- This command is of value when redundant paths exist between EBGP neighbors.
- The default *ttl* is 1, therefore BGP peers must be directly connected.
  - The range is from 1 to 255 hops.
- Increasing the *ttl* enables BGP to establish EBGP connections beyond one hop and also enables BGP to perform load balancing.

Multihop EBGP Example



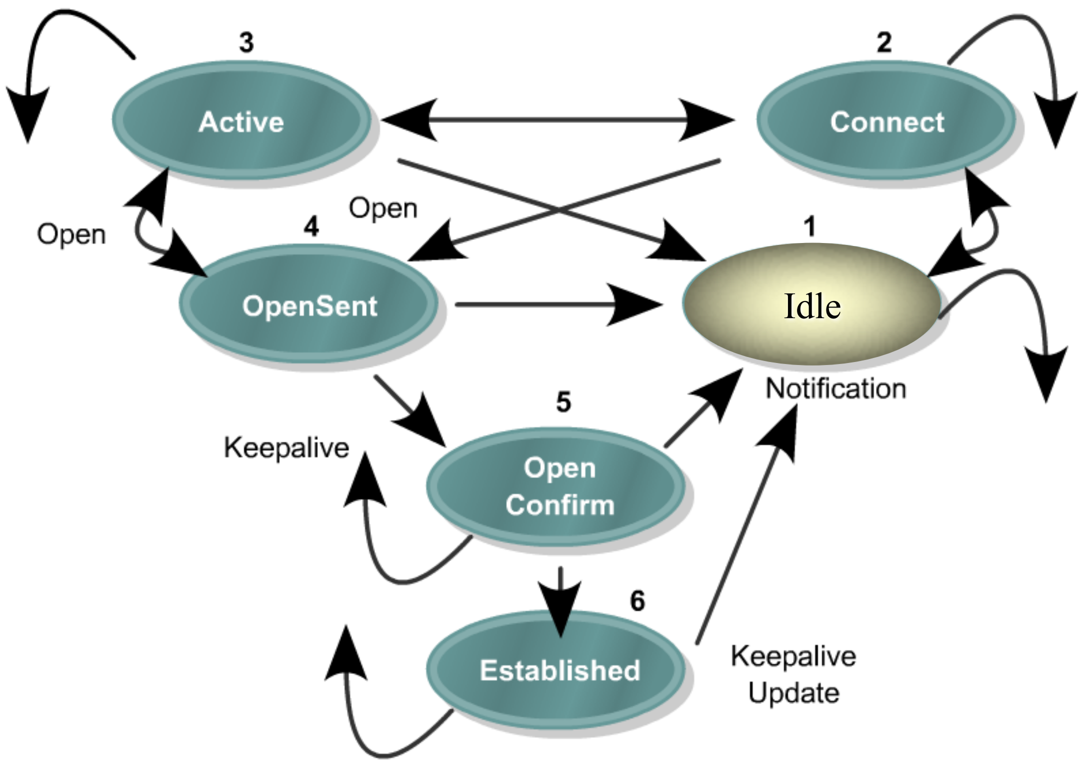

```
R1(config)# router bgp 65102
R1(config-router)# neighbor 172.16.1.1 remote-as 65101
R1(config-router)# neighbor 172.16.1.1 update-source loopback0
R1(config-router)# neighbor 172.16.1.1 ebgp-multihop 2
R1(config-router)#  exit
R1(config)# ip route 172.16.1.1 255.255.255.255 192.168.1.18
R1(config)# ip route 172.16.1.1 255.255.255.255 192.168.1.34
R1(config)#
```

```
R2(config)# router bgp 65101
R2(config-router)# neighbor 172.17.1.1 remote-as 65102
R2(config-router)# neighbor 172.17.1.1 update-source loopback0
R2(config-router)# neighbor 172.17.1.1 ebgp-multihop 2
R2(config-router)#  exit
R2(config)# ip route 172.17.1.1 255.255.255.255 192.168.1.17
R2(config)# ip route 172.17.1.1 255.255.255.255 192.168.1.33
R2(config)#
```

BGP States

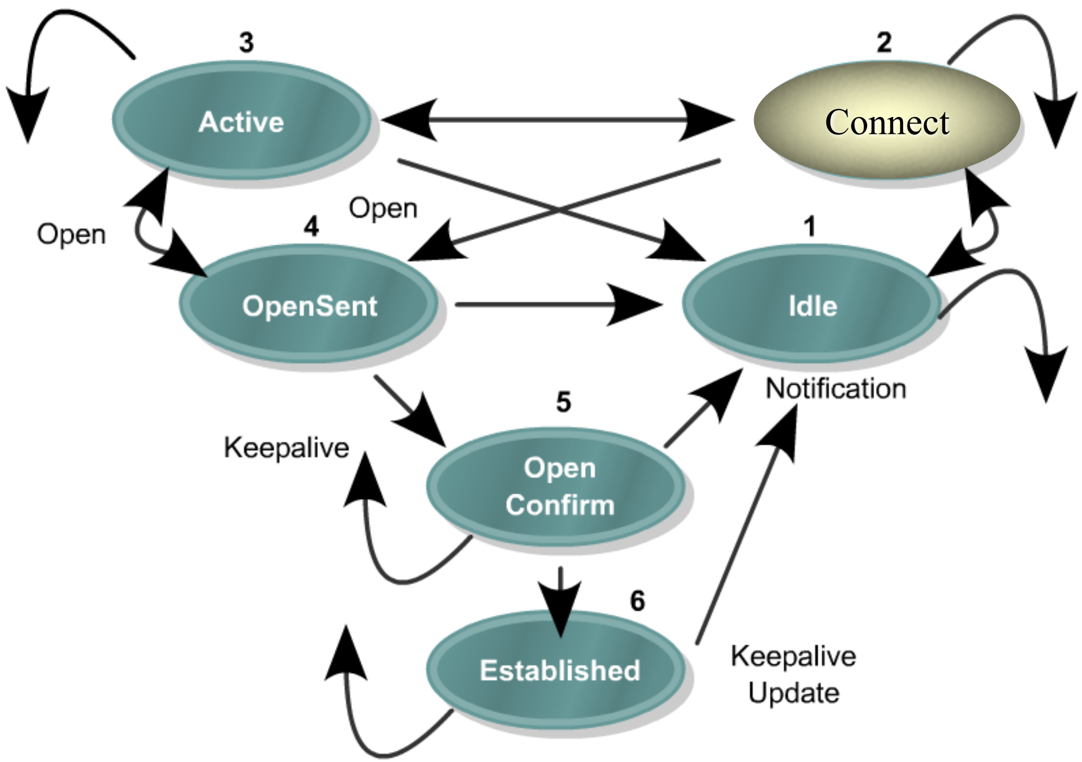
- BGP is a state machine that takes a router through the following states with its neighbors:
  - **Idle**
  - **Connect**
  - **Open sent**
  - **Open confirm**
  - **Established**
- The **Idle** state begins once the **`neighbor`** command is configured.

Idle State



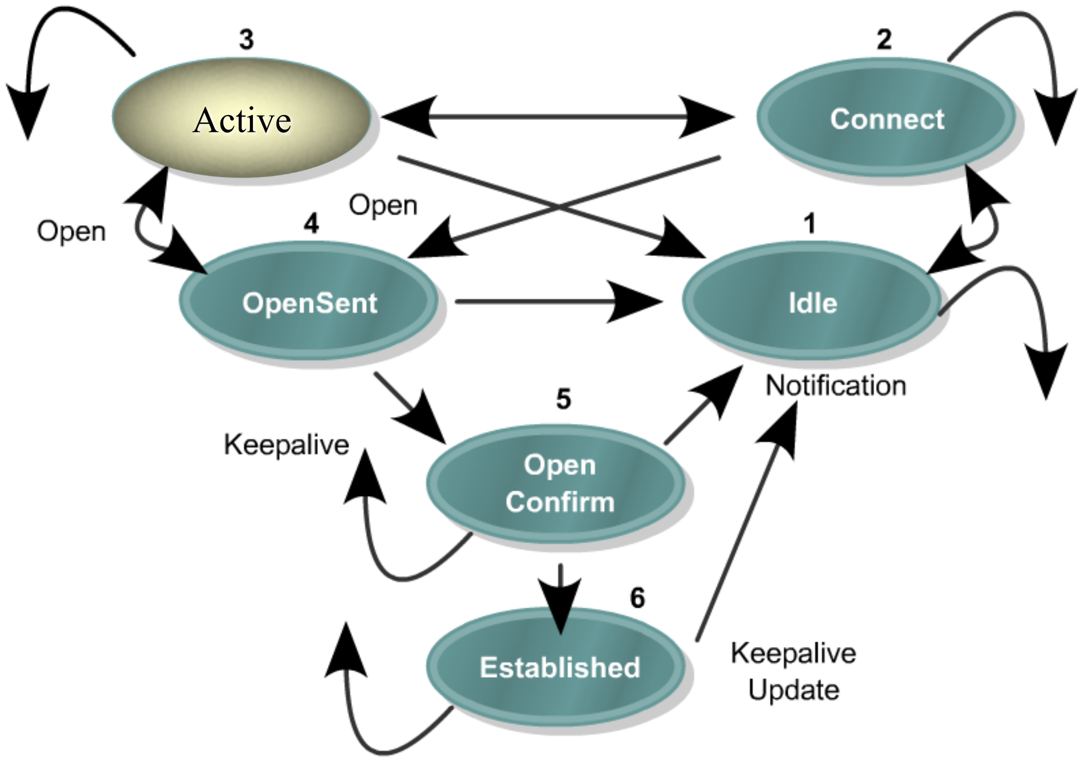
- The router is searching the routing table to see whether a route exists to reach the neighbor.
- If a router remains in this state then the router is:
  - Waiting for a static route to that IP address or network to be configured.
  - Waiting for the IGP to learn about this network from another router.

Connect State



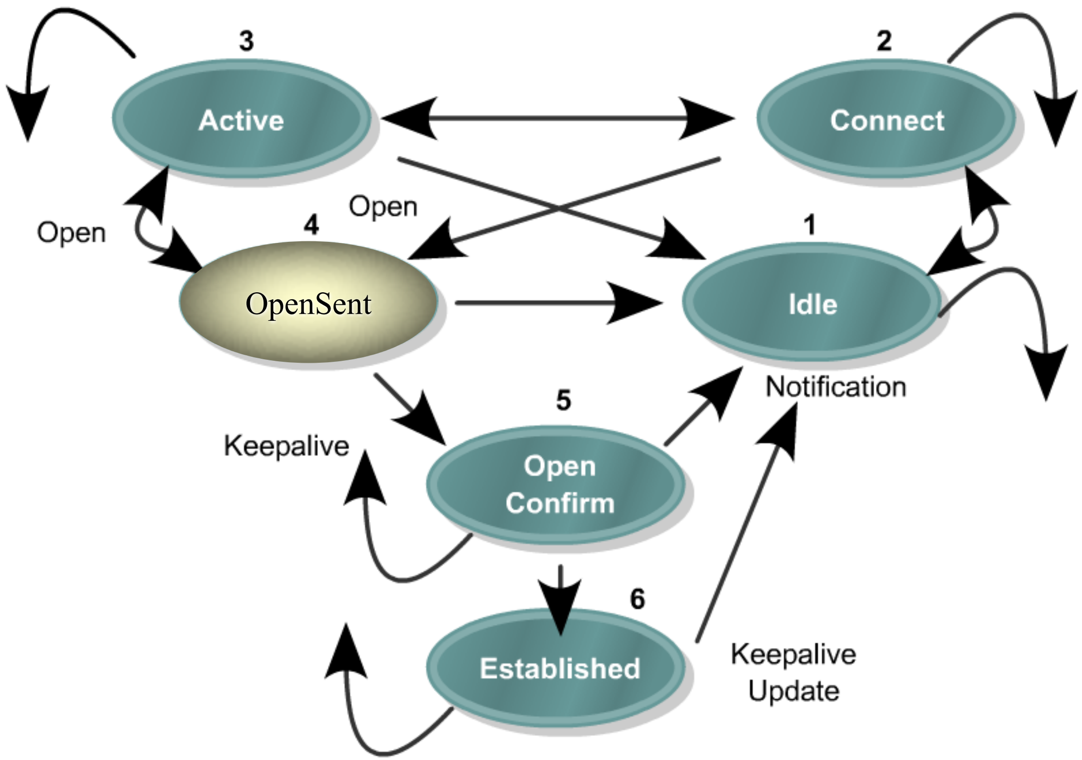
- The router found a route to the neighbor and has completed the three-way TCP handshake.

Active State



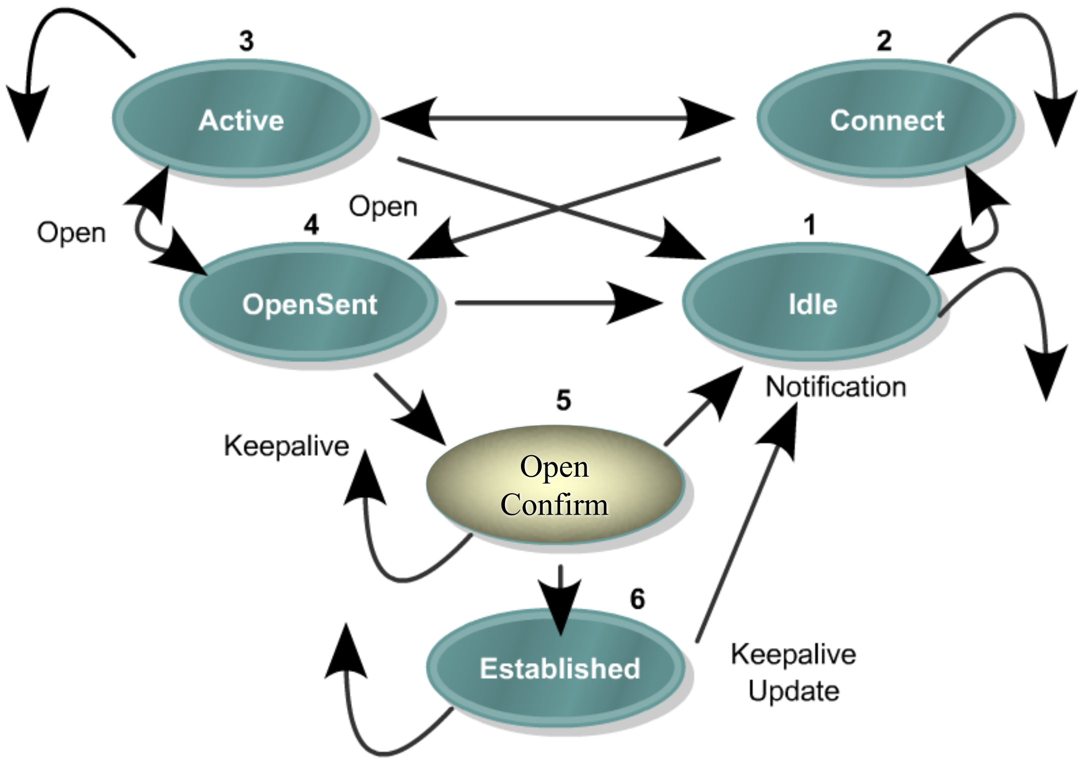
- BGP is trying to acquire a peer by initiating a TCP connection.
- If it is successful, it transitions to OpenSent otherwise the state returns to Idle.
- If the router remains is this state it means that the router has not received a response (open confirm packet) back from the neighbor.
  - Reasons for this include missing neighbor statement or incorrect AS number.

Open Sent State



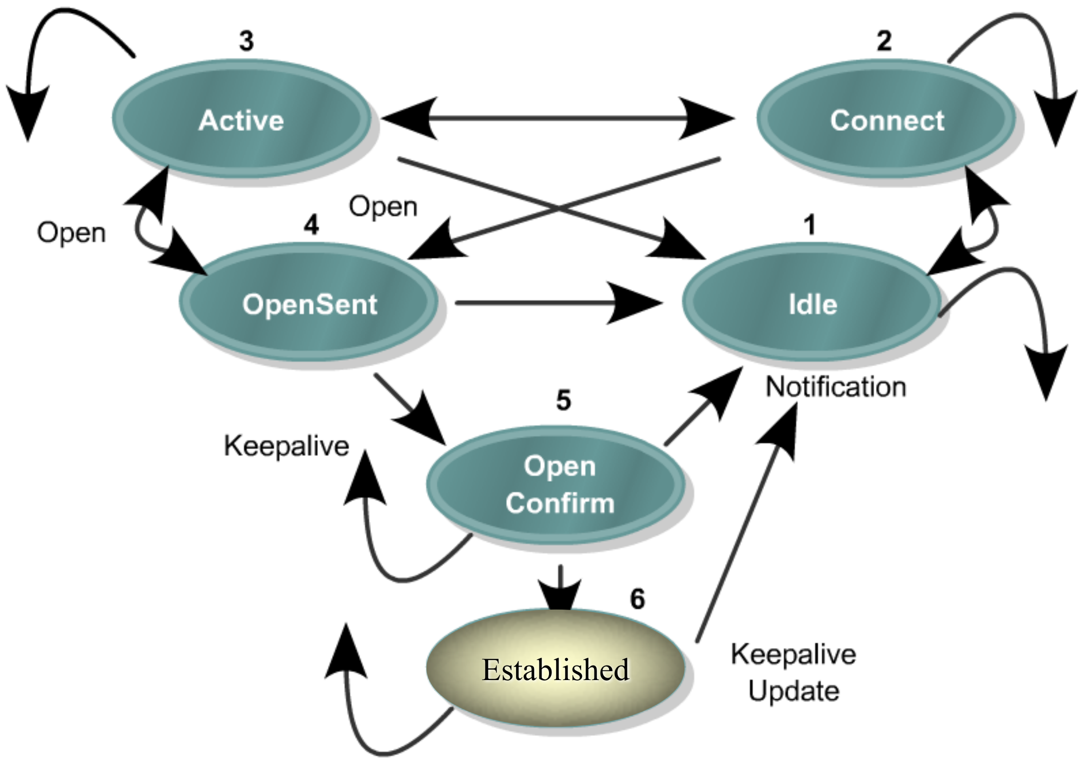
- An open message was sent, with the parameters for the BGP session.

Open Confirm



- The router received agreement on the parameters for establishing a session.

Established State



- This is the desired state for a neighbor relationship.
- It means peering is established and routing begins.

Verifying BGP: `show ip bgp neighbors`

Verify the BGP neighbor relationship.

```
R1# show ip bgp neighbors
BGP neighbor is 172.31.1.3,  remote AS 64998, external link
  BGP version 4, remote router ID 172.31.2.3
  BGP state = Established, up for 00:19:10
  Last read 00:00:10, last write 00:00:10, hold time is 180, keepalive
interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(old & new)
    Address family IPv4 Unicast: advertised and received
  Message statistics:
    InQ depth is 0
    OutQ depth is 0
                         Sent       Rcvd
    Opens:                  7          7
    Notifications:          0          0
    Updates:               13         38
<output omitted>
```

BGP video

Yakov Rekhter

http://www.youtube.com/watch?v=HAOVNYSnL7k