BGP

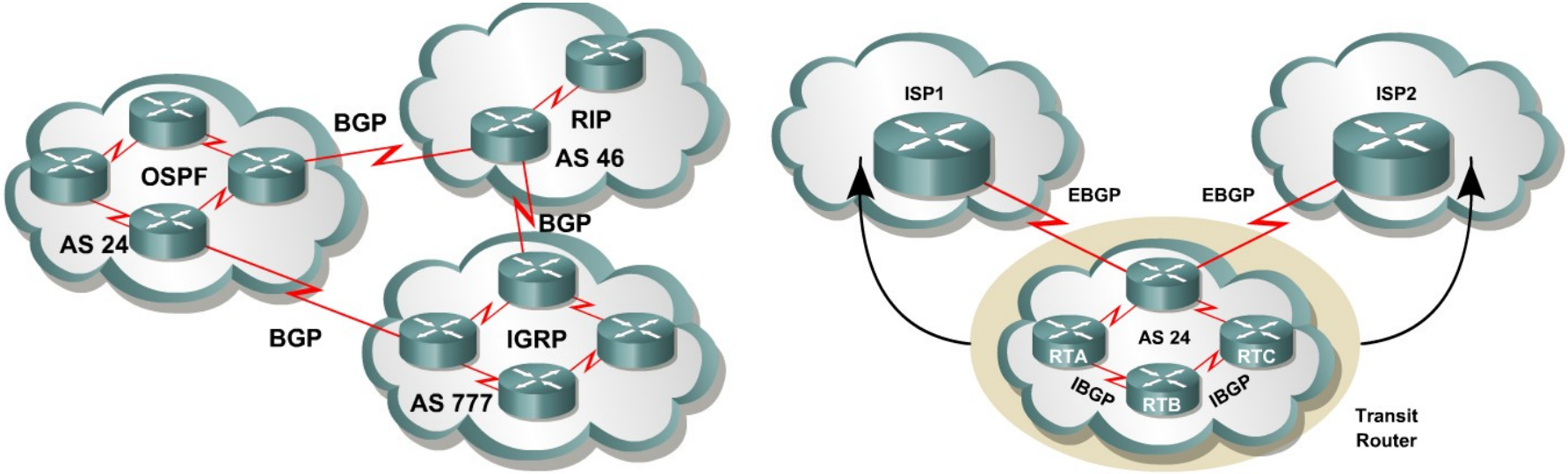
Cabrillo College

Module 4

Terms

- **IGP** (Interior Gateway Protocol) - RIP, IGRP, EIGRP, OSPF = Routing protocol used to exchange routing information within an autonomous system.
- **EGP** (Exterior Gateway Protocol) - BGP = Routing protocol used to exchange routing information between autonomous systems.
- **Autonomous System** = (From RFC 1771) “A set of routers under the single technical administration, using an IGP and common metrics to route packets within the AS, and using an EGP to route packets to other AS’s.”
- **BGP** is a path vector or an advanced distance vector routing protocol.

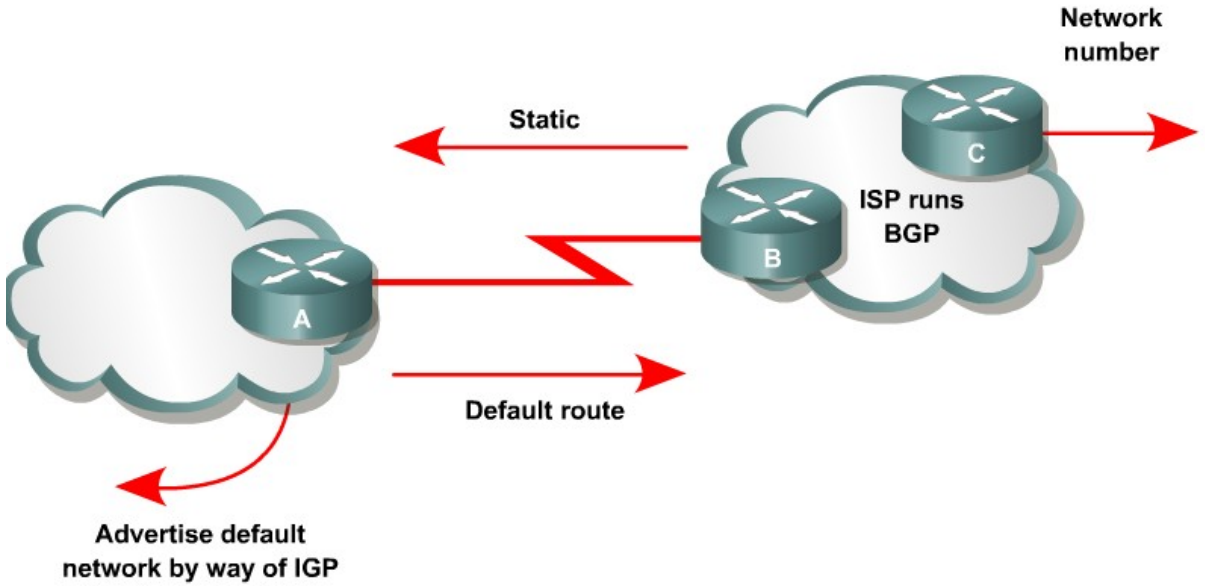
When to use BGP and when not to use BGP – Cisco CCO



Use BGP when the effects of BGP are well understood and one of the following conditions exist:

- The AS allows packets to transit through it to reach another AS (transit AS).
- The AS has multiple connections to other AS's.
- The flow of traffic entering or exiting the AS must be manipulated. This is policy based routing and based on **attributes**.

When to use BGP and when not to use BGP

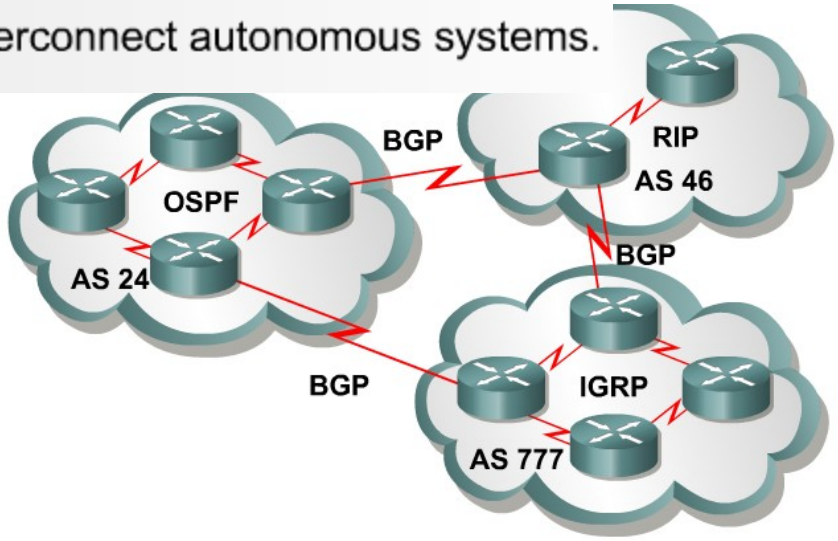


Do not use BGP if you have one or more of the following conditions:

- A single connection to the Internet or another AS
- No concern for routing policy or routing selection
- A lack of memory or processing power on your routers to handle constant BGP updates
- A limited understanding of route filtering and BGP path selection process
- Low bandwidth between AS's

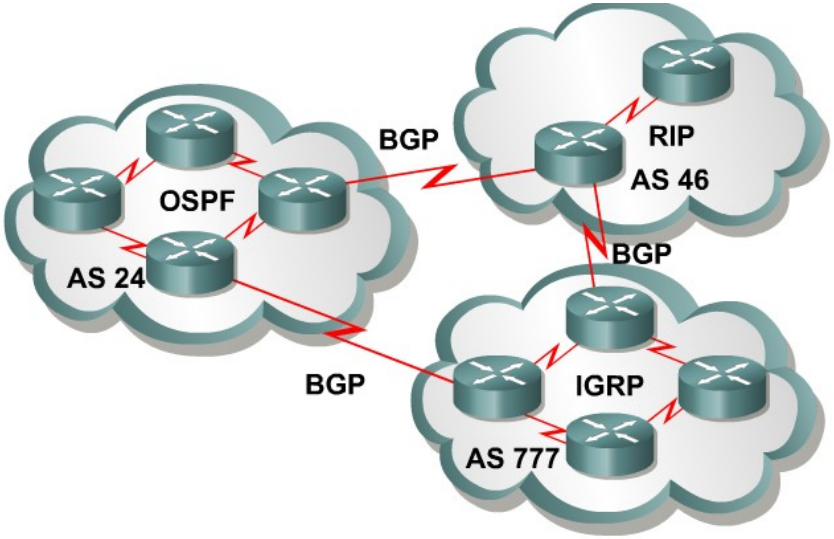
Overview of autonomous systems

EGPs, such as BGP, are used to interconnect autonomous systems.



- An AS is a group of routers that share similar routing policies and operate within a single administrative domain.
- An AS can be a collection of routers running a single IGP, or it can be a collection of routers running different protocols all belonging to one organization.
- In either case, the outside world views the entire Autonomous System as a single entity.

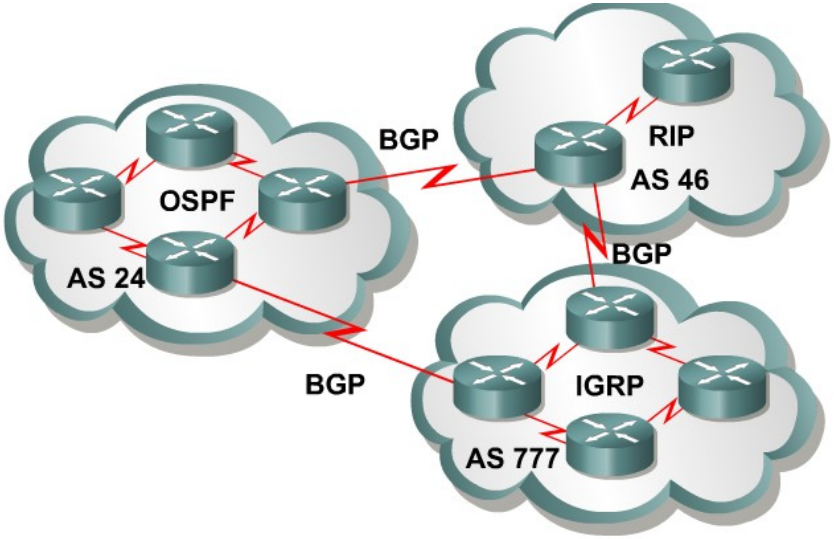
Overview of autonomous systems



AS Numbers

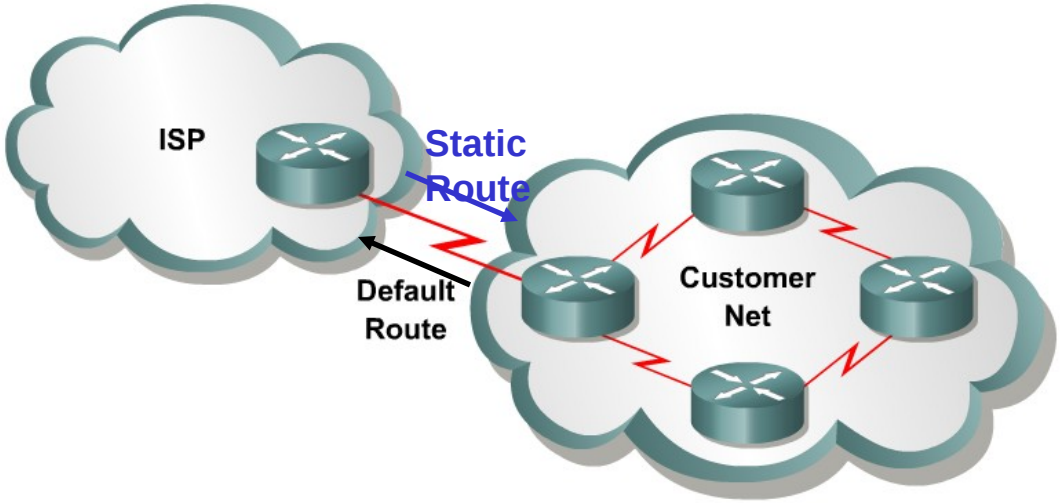
- Each AS has an identifying number that is assigned by an Internet registry or a service provider.
- This number is between **1 and 65,535**.
- AS numbers within the range of **64,512 through 65,535** are reserved for **private** use.
- This is similar to RFC 1918 IP addresses.
- Because of the finite number of available AS numbers, an organization must present justification of its need before it will be assigned an AS number.

Overview of autonomous systems



- Today, the Internet Assigned Numbers Authority (IANA) is enforcing a policy whereby organizations that connect to a single provider and share the provider's routing policies use an AS number from the private pool, 64,512 to 65,535.

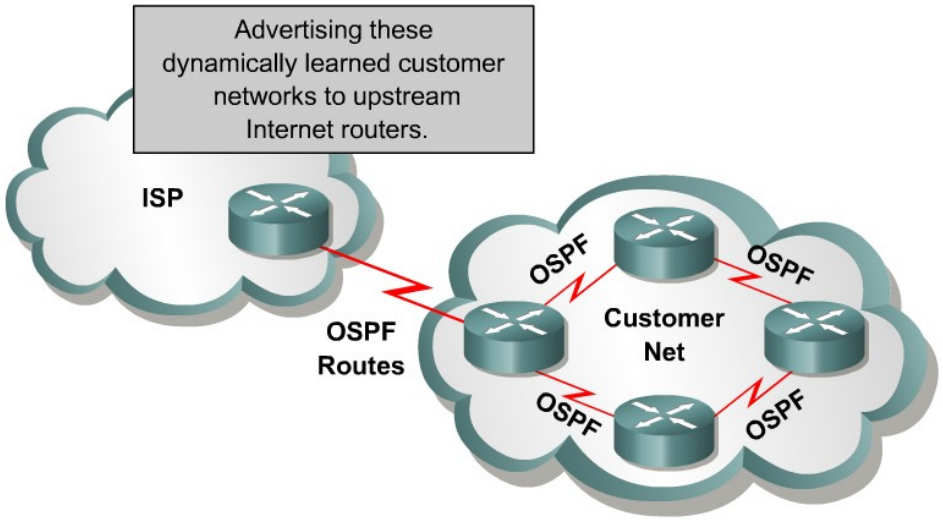
Single-homed autonomous systems



A single-homed AS can be configured with a default route to reach outside networks.

- If an AS has only **one exit point** to outside networks, it is considered a **single-homed system**.
- Single-homed autonomous systems are often referred to as **stub** networks or stubs.
- Stubs can rely on a **default route** to handle all traffic destined for non-local networks.
- BGP is **not** normally needed in this situation.

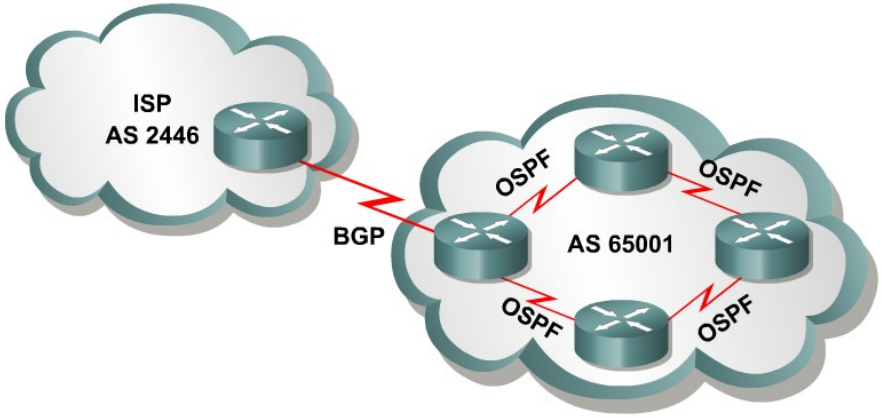
Single-homed autonomous systems



A provider may choose to dynamically learn customer routes using an IGP, such as OSPF.

- Use an IGP – Both the provider and the customer use an **IGP** to share information regarding the customer's networks.
- This provides the benefits associated with dynamic routing.
- BGP is **not** normally needed in this situation.

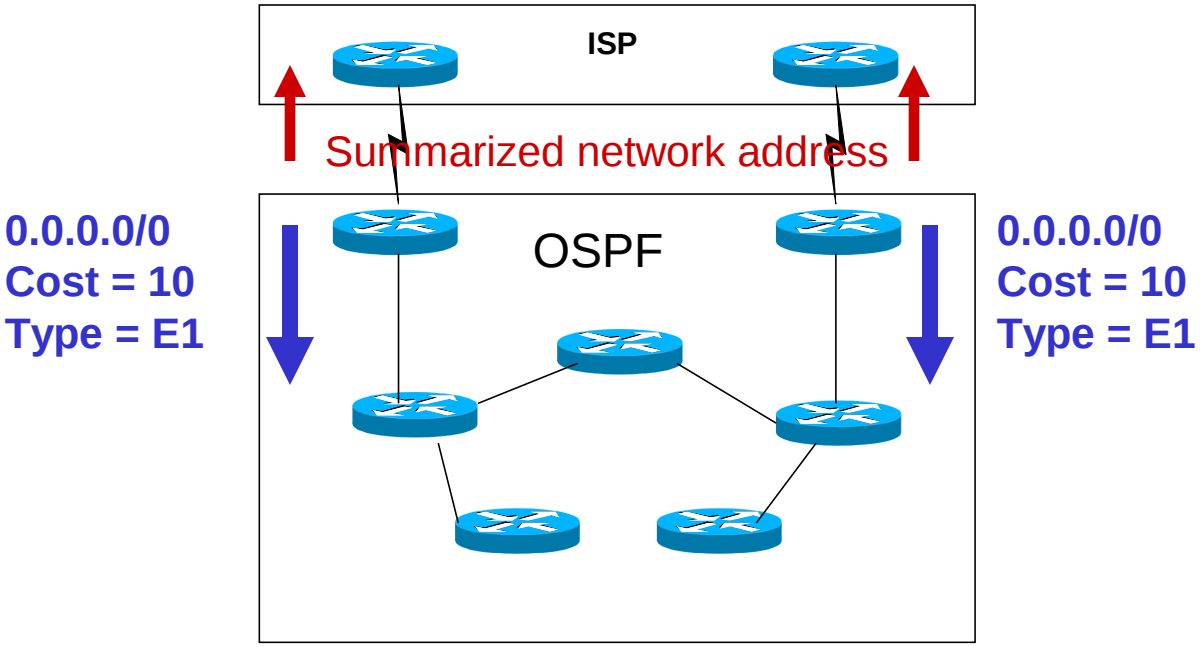
Single-homed autonomous systems



A provider may also choose to dynamically learn a customer's routes using BGP, which typically runs between the ISP router and the customer's boundary router.

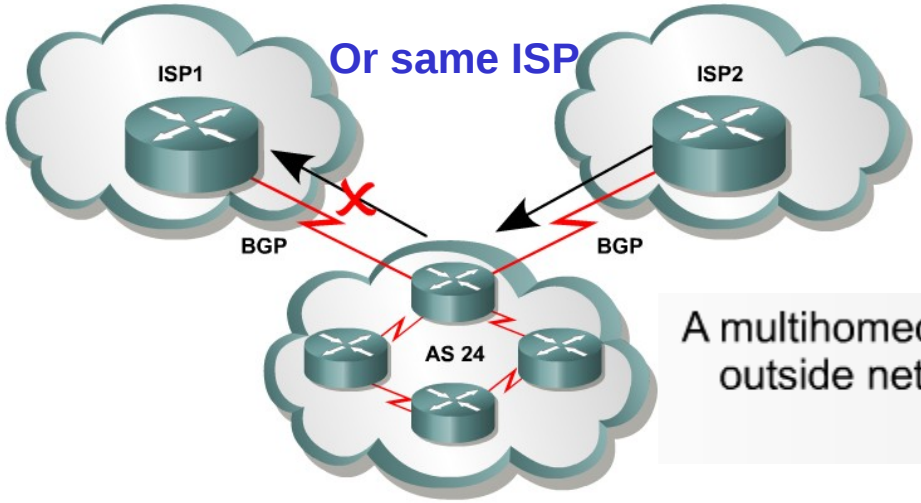
- **Use an EGP** – The third method by which the ISP can learn and advertise the customer's routes is to use an EGP such as BGP.
- In a **single-homed autonomous system** the *customer's routing policies are an extension of the policies of the provider*.
  - For this reason the Internet number registries are unlikely to assign an AS number.
  - Instead, the provider can give the customer an AS number from the **private** pool of AS numbers, 64,512 to 65,535.
  - The provider will strip off these numbers when advertising the customer's routes towards the core of the Internet.

Multi-homed to a Single Autonomous Systems



- This is an improved topology over Single-Home AS, providing for **redundancy**.
- One option may be to use one link as the **primary** link and the other as a **backup** link.
- A better design would be to **use both paths**, with each one providing backup for the other in the event of link or router failure.
- In most cases this will be sufficient for good internetwork performance.

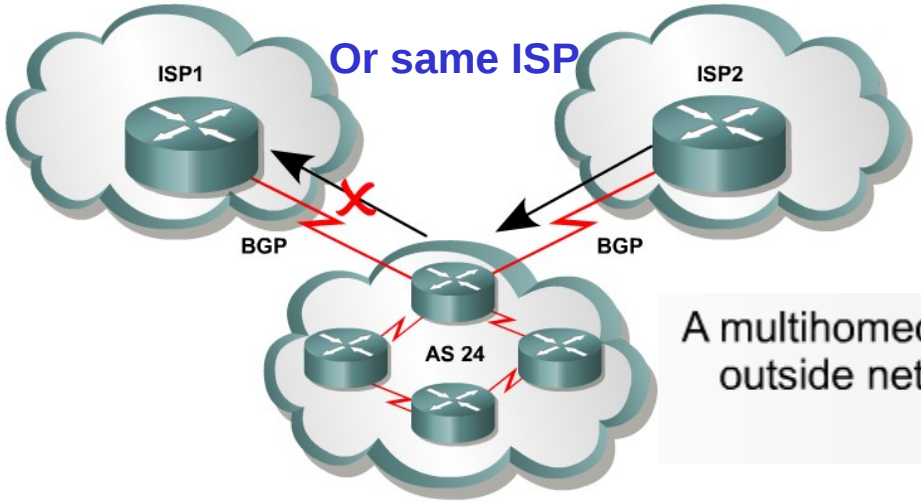
Multihomed nontransit autonomous systems



A multihomed nontransit AS features more than one exit point to outside networks, but does not allow traffic to pass from one outside connection to another.

- An AS is a **multihomed system** if it has *more than one exit point* to outside networks.
- A **non-transit AS** does not allow transit traffic-that is, any traffic that has a source and destination outside the AS—to pass through it.
- A **non-transit AS** would advertise only its *own* routes to both the providers it connects to—it would not advertise routes it learned from one provider to another.
- This makes certain that ISP1 will not use AS 24 to reach destinations that belong to ISP2, and ISP2 would not use AS 24 to reach destinations that belong to ISP1.

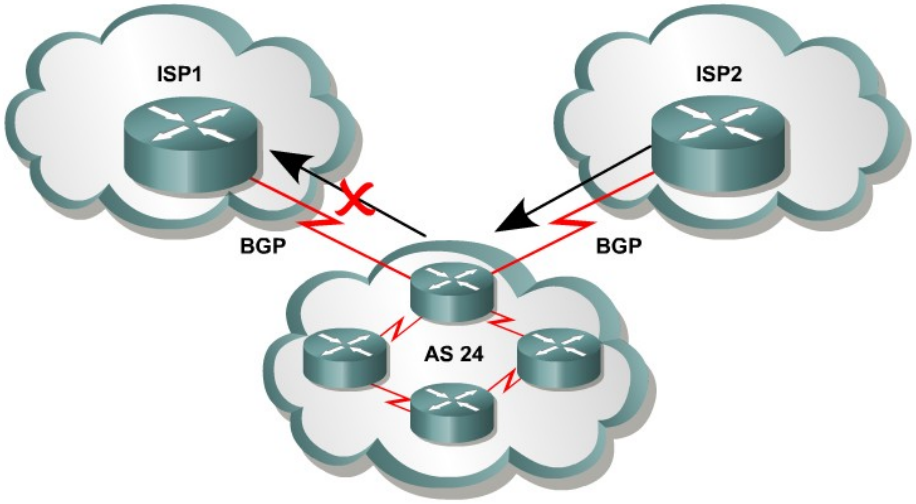
Multihomed nontransit autonomous systems



A multihomed nontransit AS features more than one exit point to outside networks, but does not allow traffic to pass from one outside connection to another.

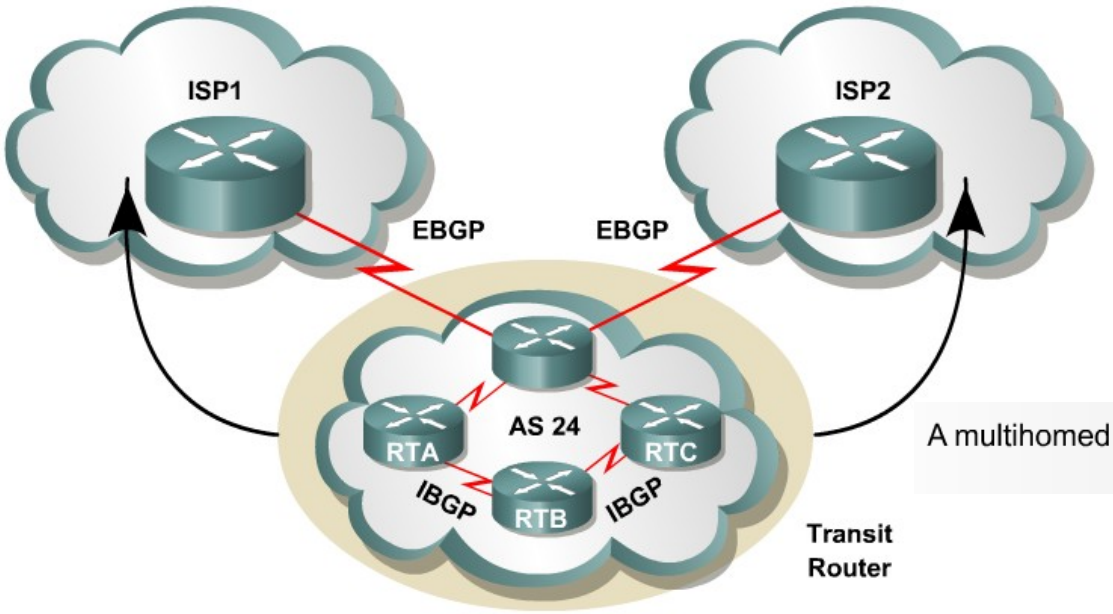
- Multihomed nontransit autonomous systems do not really need to run BGP4 with their providers.
- It is usually **recommended** and **often required** by ISPs.
- As it will be seen later in this module, BGP4 offers numerous advantages, including increased control of route propagation and filtering.

Multihomed nontransit autonomous systems



- ***Incoming route advertisements influence your outgoing traffic, and outgoing advertisements influence your incoming traffic.***
- If the provider advertises routes into your AS via BGP, your internal routers have more accurate information about external destinations.
  - BGP also provides tools for setting routing policies for external destinations.
- If your internal routes are advertised to the provider via BGP, you have influence over which routes are advertised at which exit point.
  - BGP also provides tools for your influencing (to some degree) the choices the provider makes when sending traffic into your AS.

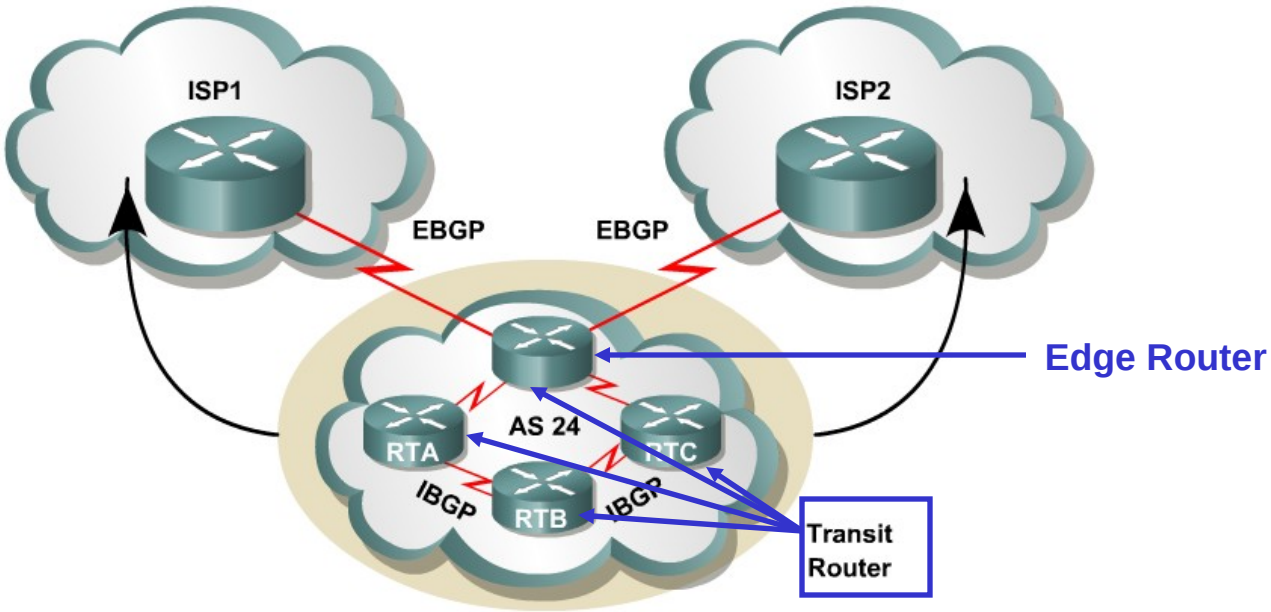
Multi-homed Transit Autonomous Systems



A multihomed transit system can be used for transit traffic by other autonomous systems.

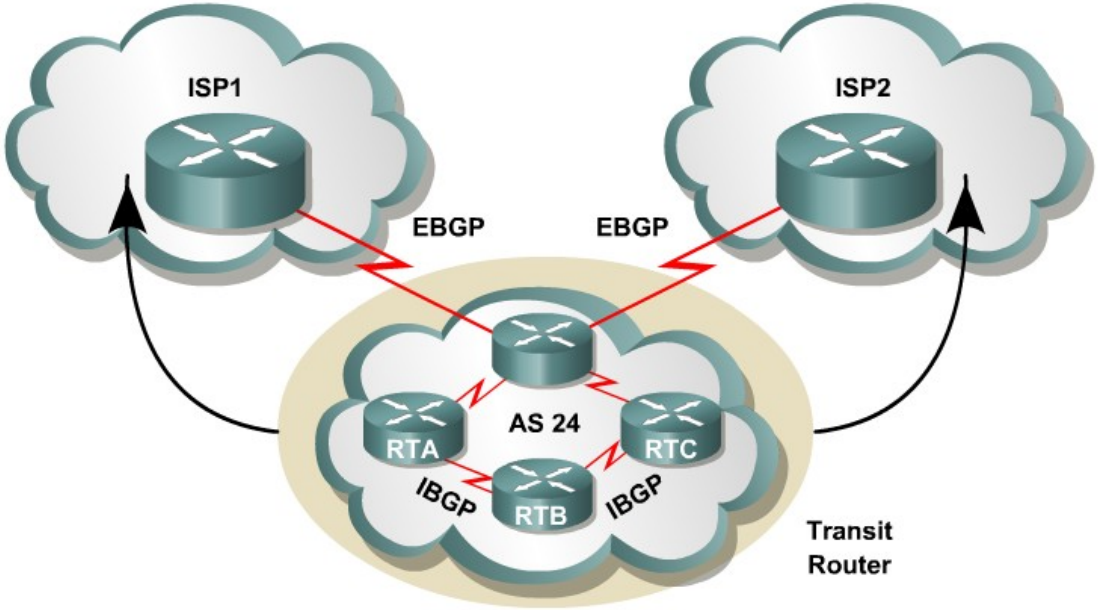
- A multi-homed **transit system** has more than one connection to the outside world and can be used for transit traffic by other autonomous systems.
  - From the point of view of the multi-homed AS, transit traffic is any traffic originating from outside sources bound for outside destinations

Multi-homed Transit Autonomous Systems



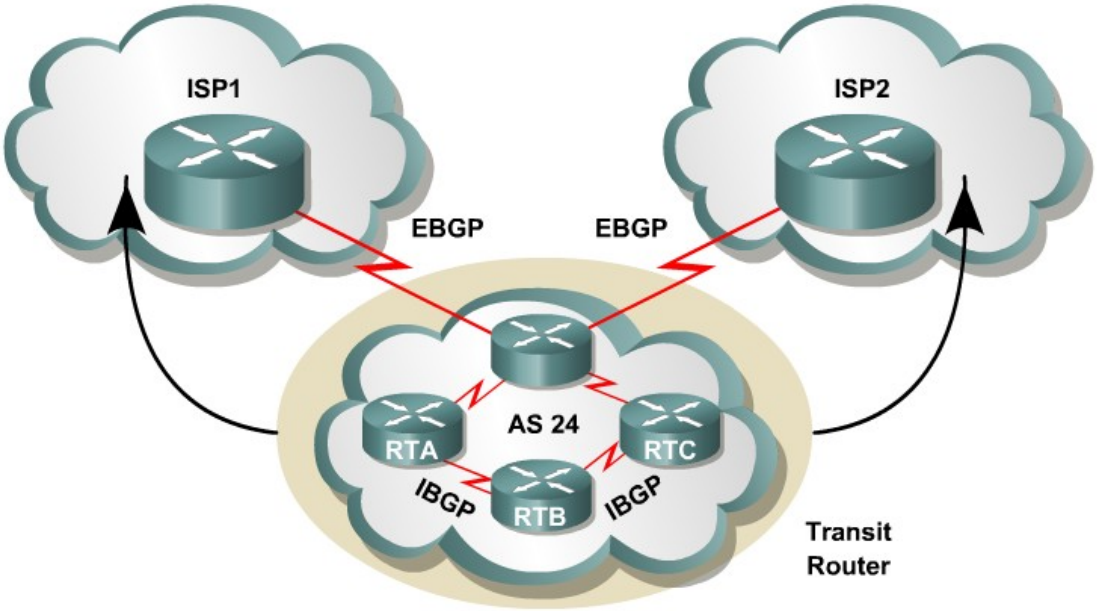
- When BGP is running inside an AS, it is referred to as **Internal BGP (IBGP).**
- When BGP runs between autonomous systems, it is called **External BGP (EBGP).**
- If the role of a BGP router is to route IBGP traffic, it is called a **transit router.**
- Routers that sit on the boundary of an AS and that use EBGP to exchange information with the ISP are called **border or edge routers**.

BGP Hazards – Doyle, Routing TCP/IP



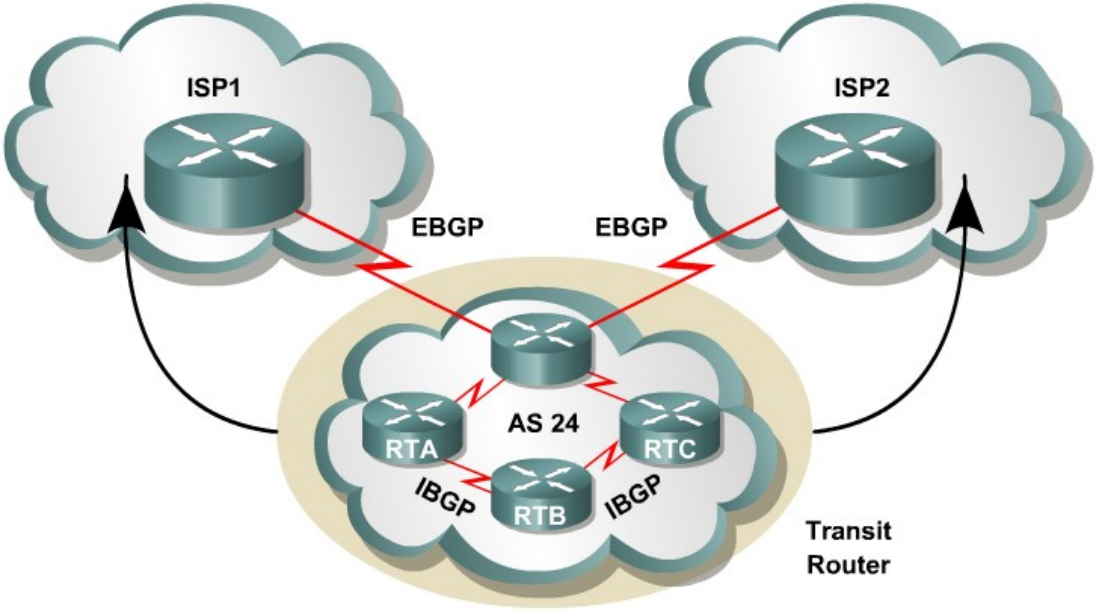
- Creating a BGP “peering” relationship involves an interesting combination of trust and mistrust.
- You must trust the network administrator on that end to know what they are doing.
- At the same time, if you are smart, you will take every practical measure to protect yourself in the event that a mistake is made on the other end.
- “Paranoia is your friend.”

BGP Hazards – Inadvertent Transit Domain



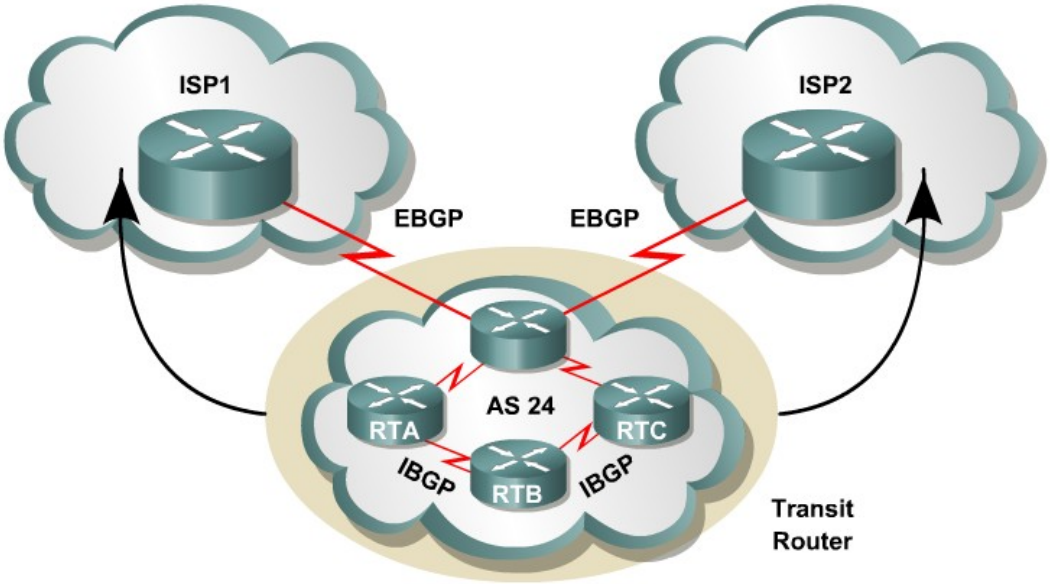
- We inadvertently advertise routes learned from ISP2 to ISP1.
- ISP1 customers will see our network as the best path to ISP2 customers.
- We have become a transit domain for packets from ISP1 to ISP2.

BGP Basics



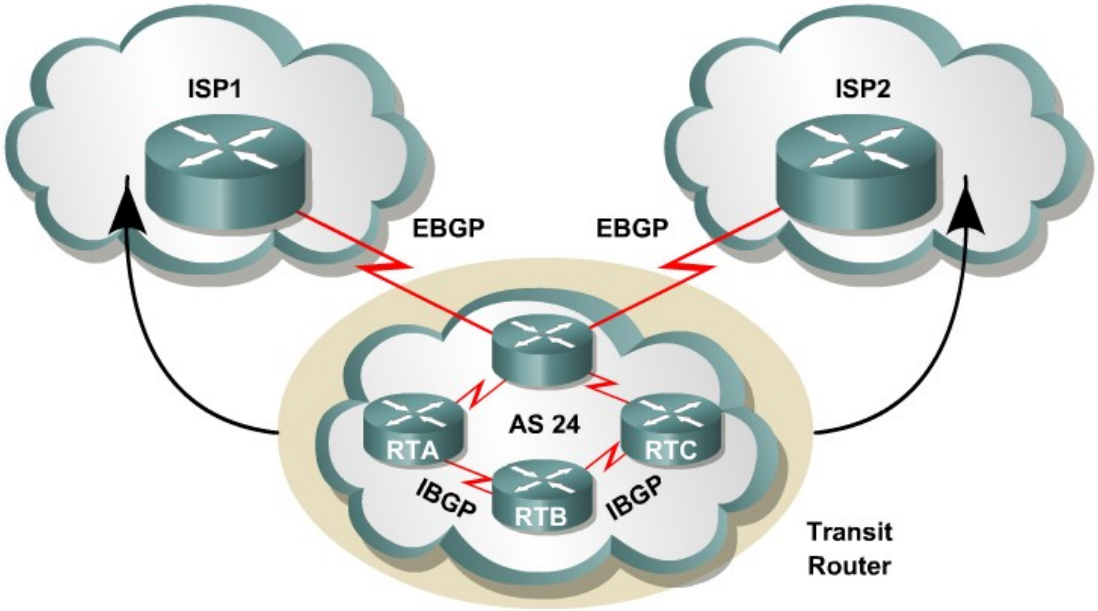
- BGP is a path vector routing protocol.
- Defined in RFC 1772
- BGP is a distance vector routing protocol, in that it relies on downstream neighbors to pass along routes from their routing table.
- BGP uses a list of AS numbers through which a packet must pass to reach a destination.

BGP Basics



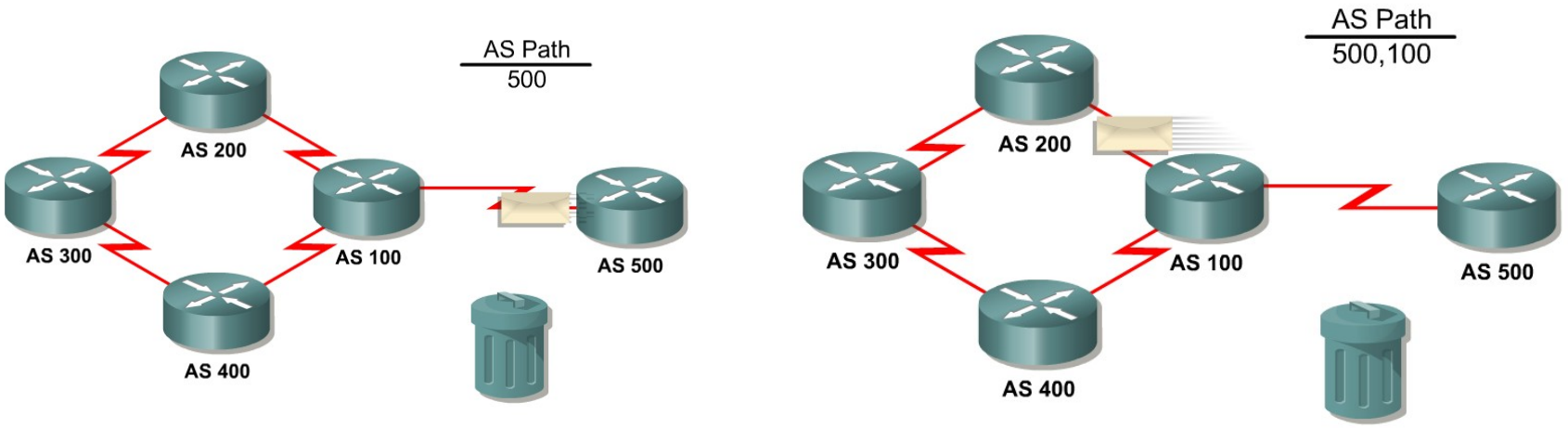
- The function of BGP is to:
  - Exchange routing information between autonomous systems
  - Guarantee the selection of a loop free path.
- BGP4 is the first version of BGP that supports CIDR and route aggregation.
- Common IGPs such as RIP, OSPF, and EIGRP use technical metrics.
  - BGP does not use technical metrics.
- BGP makes routing decisions based on network policies, or rules (later)
- BGP does not show the details of topologies within each AS.
- BGP sees only a tree of autonomous systems.
- Cisco routers maintain a separate routing table to hold BGP routes – **show ip bgp** – later.

BGP Basics

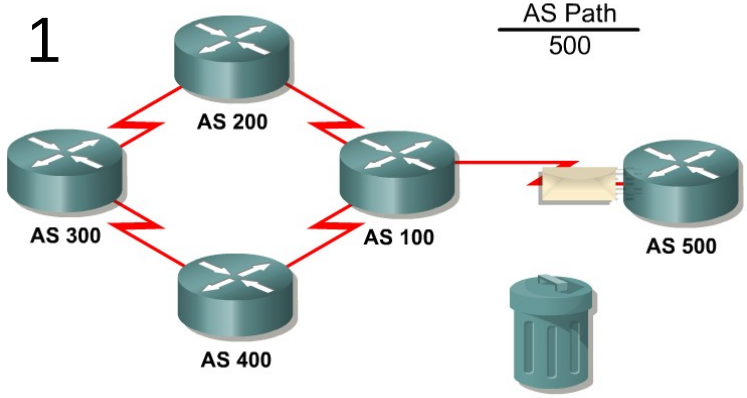


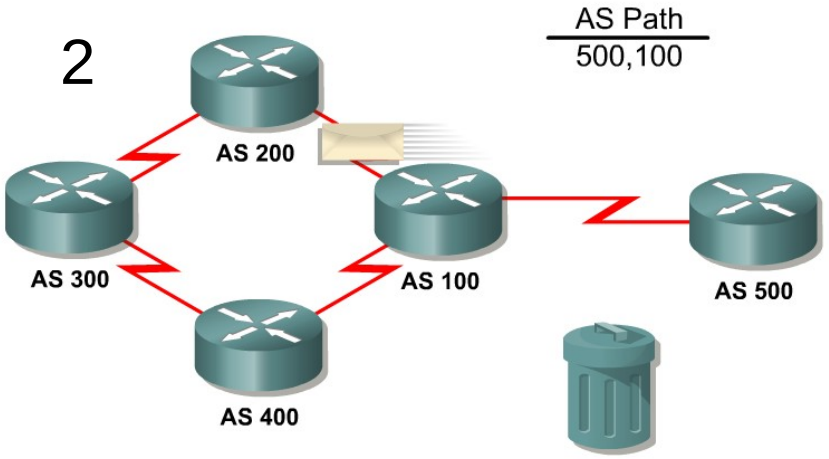
- BGP updates are carried using TCP on port 179.
  - In contrast, RIP updates use UDP port 520
  - OSPF, IGRP, EIGRP does not use a Layer 4 protocol
- Because BGP requires TCP, IP connectivity must exist between BGP peers.
- TCP connections must also be negotiated between them before updates can be exchanged.
- Therefore, BGP inherits those reliable, connection-oriented properties from TCP.

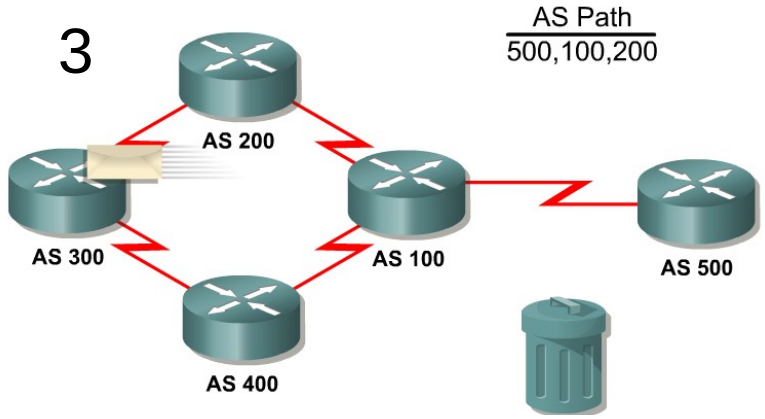
Loop Free Path

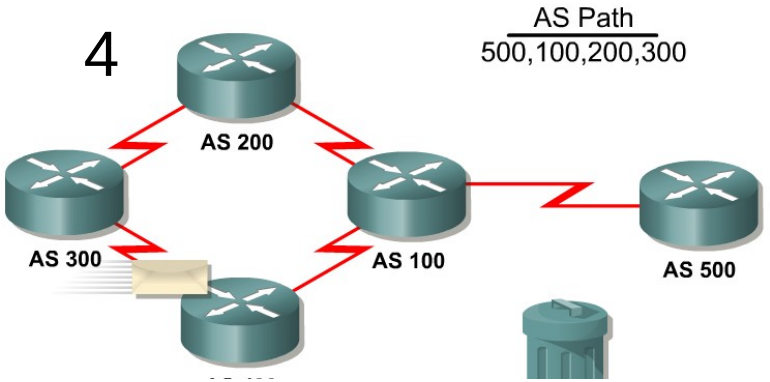


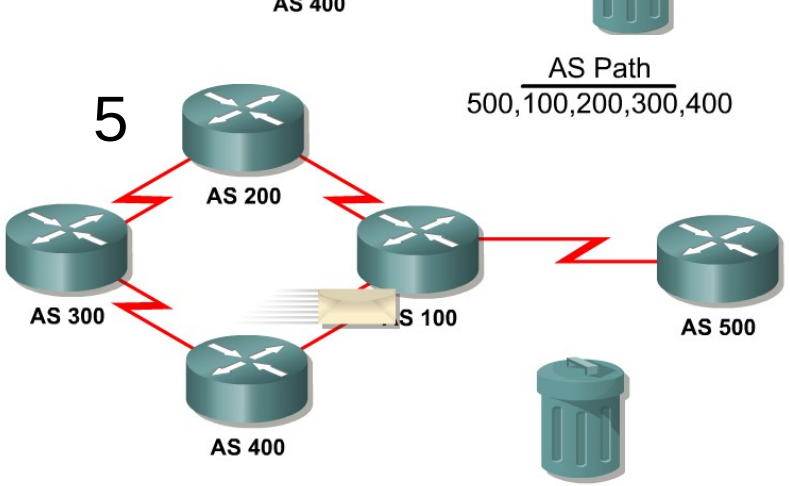
- To guarantee loop free path selection, BGP constructs a graph of autonomous systems based on the information exchanged between BGP neighbors.
- BGP views the whole internetwork as a graph, or tree, of autonomous systems.
- The connection between any two systems forms a path.
- The collection of path information is expressed as a sequence of AS numbers called the AS Path.
- This sequence forms a route to reach a specific destination

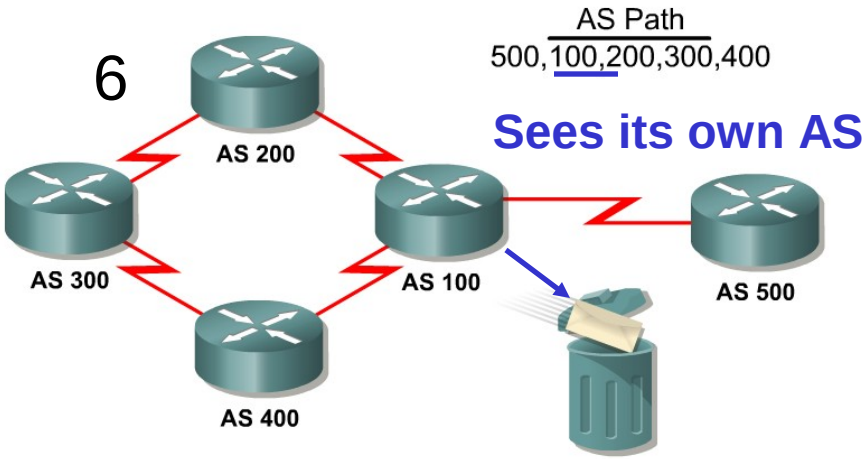

1
AS Path
500
AS 200
AS 300
AS 100
AS 500
AS 400


2
AS Path
500,100
AS 200
AS 300
AS 100
AS 500
AS 400

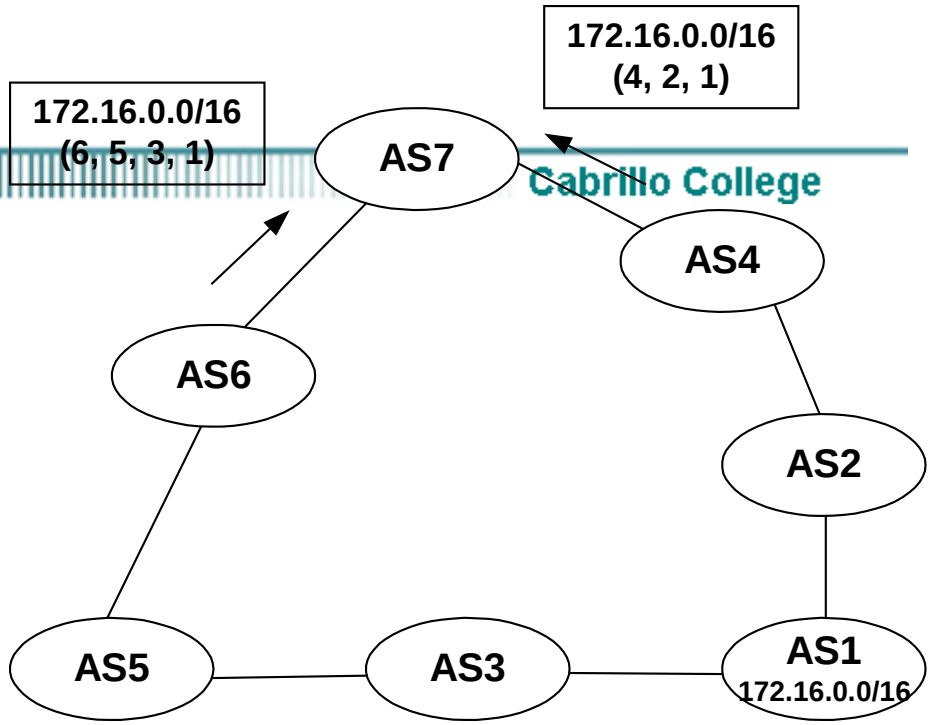

3
AS Path
500,100,200
AS 200
AS 300
AS 100
AS 500
AS 400


4
AS Path
500,100,200,300
AS 200
AS 300
AS 100
AS 500


5
AS Path
500,100,200,300,400
AS 200
AS 300
AS 500
AS 400


6
AS Path
500,100,200,300,400
Sees its own AS
AS 200
AS 300
AS 100
AS 500
AS 400

Loop Free Path

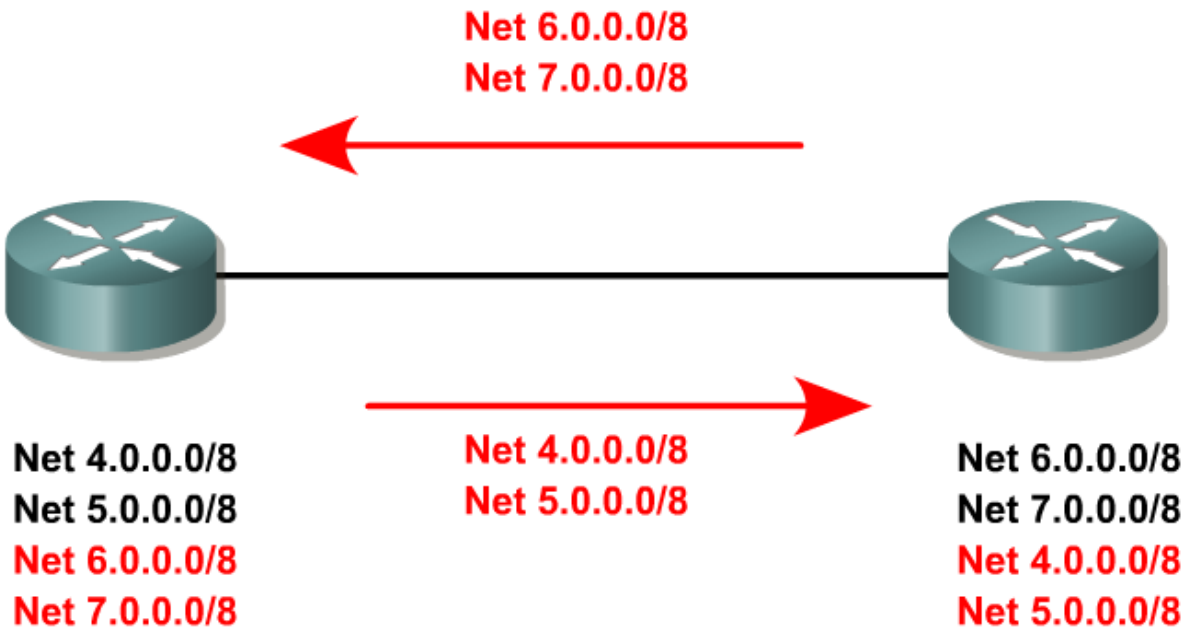


- The list of AS numbers associated with a BGP route is called the **AS_PATH** and is one of several path attributes associated with each route.
- Path attributes will be discussed in much more detail later.
- The shortest inter-AS path is very simply determined by the least number of AS numbers.
- All things being equal, BGP prefers routes with shorter AS paths.
- In this example, AS7 will choose the shortest path (4, 2, 1).
- We will see later what happens with equal cost paths.

BGP Operation

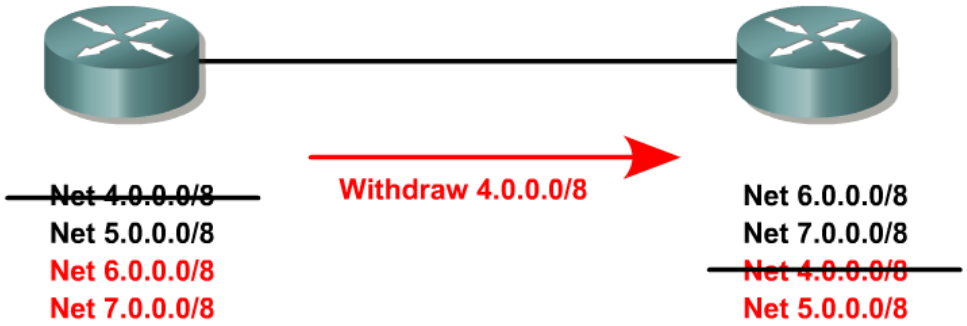
- When two routers establish a TCP-enabled BGP connection between each other, they are called ***neighbors*** or ***peers***.
- Each router running BGP is called a ***BGP speaker***.

Initial Exchange



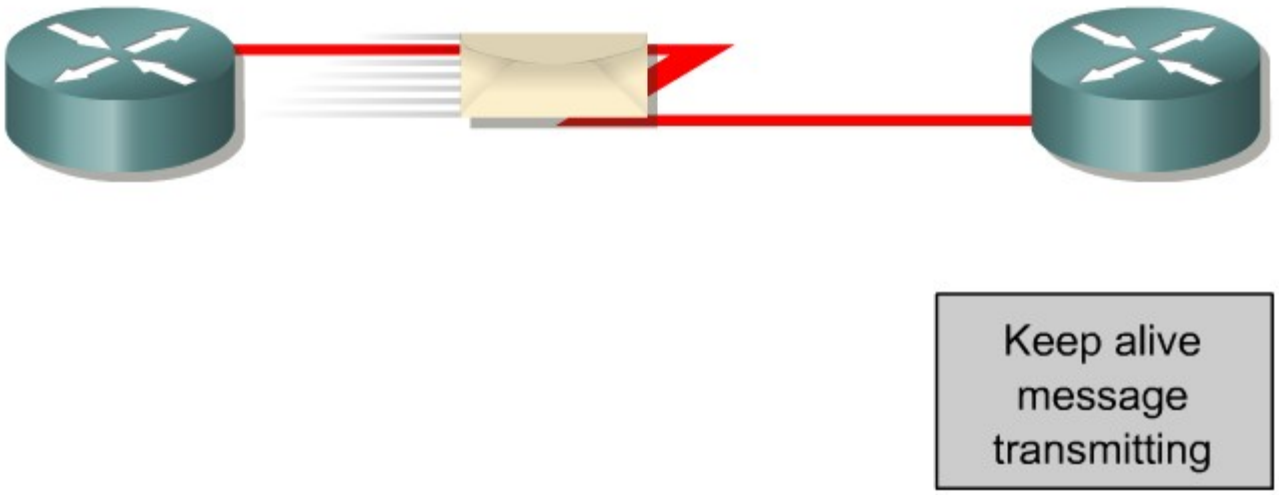
- When BGP neighbors first establish a connection, they exchange all candidate BGP routes.
- After this initial exchange, incremental updates are sent as network information changes.

Withdrawn Routes



- The information for network reachability can change, such as when a route becomes unreachable or a better path becomes available.
- BGP informs its neighbors of this by withdrawing the invalid routes and injecting the new routing information.
- Withdrawn routes are part of the update message. BGP routers keep a table version number that tracks the version of the BGP routing table received from each peer.
- If the table changes, BGP increments the table version number.
- A rapidly incrementing table version is usually an indication of instabilities in the network, or a misconfiguration.

BGP Keepalives



- Peers exchange **keepalive messages** to ensure the connection is maintained.
- The Cisco default keepalive interval is **60 seconds** (RFC 1771 does not specify a standard time).
- If three keepalive intervals (**180 seconds**) pass the peer declares its neighbor down.
- These can be modified with timers **bgp command**.

BGP Message Types

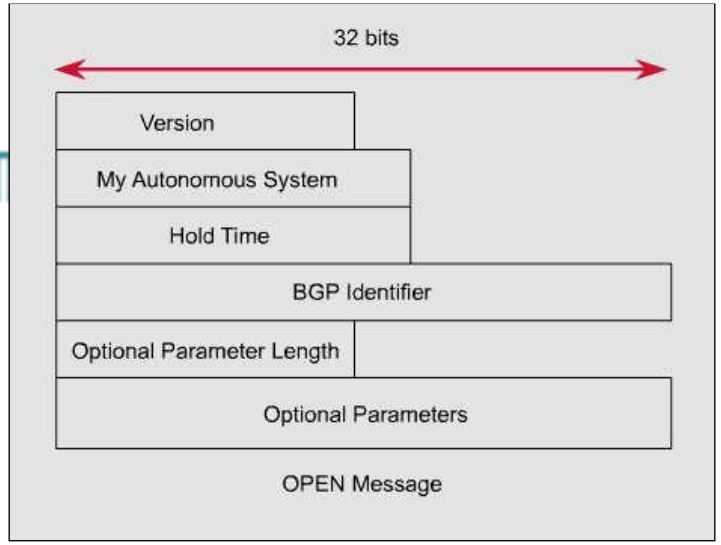
- Before establishing a BGP peer connection the two neighbors must perform the standard TCP three-way handshake and open a TCP connection to port 179.
- After the TCP session is established, BGP peers exchanges several messages to open and confirm connection parameters and to send BGP routing information.
- All BGP messages are unicast to the one neighbor over the TCP connection.
- There are four BGP message types:
  - **Type 1: OPEN**
  - **Type 2: KEEPALIVE**
  - **Type 3: UPDATE**
  - **Type 4: NOTIFICATION**

BGP Message Types

Each BGP Message contains the following header:

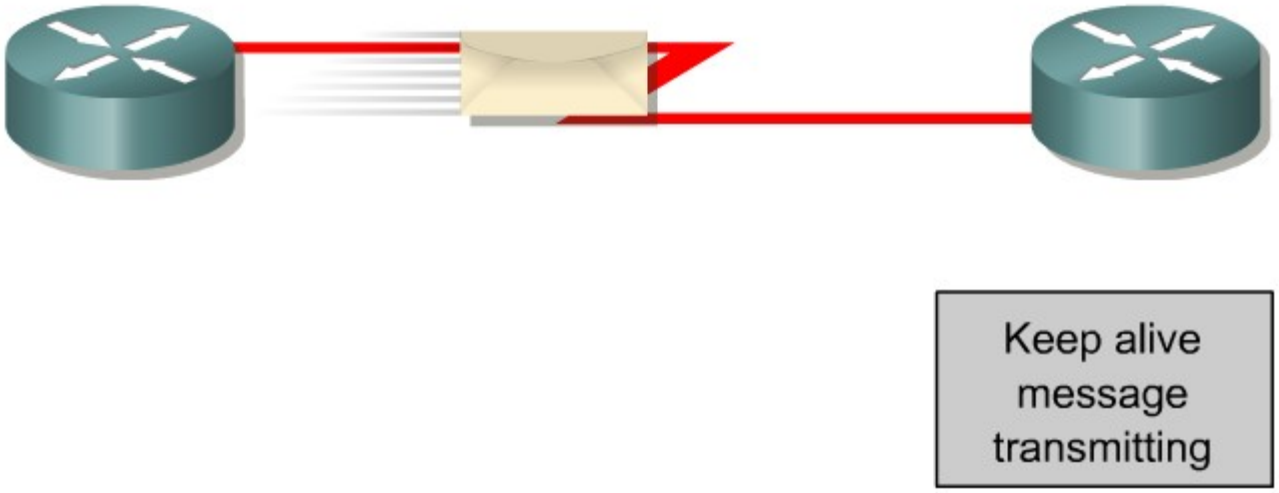
- **Marker**: The marker field is used to either authenticate incoming BGP messages or to detect loss of synchronization between two BGP peers.
- **Length**: The length field indicates the total BGP message length, including the header.

Type 1:
BGP Open Message



- After the TCP session is established, both neighbors send Open messages.
- This message is used to establish connections with peers.
- Each neighbor uses this message to identify itself and to specify its BGP operational parameters including:
  - **BGP version number** (defaults to version 4)
  - **AS number**: AS number of the originating router, determines if BGP session is EBGP or IBGP.
  - **BGP identifier**: IP address that identifies the neighbor using the same method as OSPF router ID.
  - **Optional parameter**: authentication, multiprotocol support and route refresh.

Type 2: BGP Keepalive Message



- This message type is sent periodically between peers to maintain connections and verify paths held by the router sending the keepalive.
- If a router accepts the parameters specified in its neighbor's Open message, it responds with a Keepalive.
- Subsequent Keepalives are **sent every 60 seconds** by Cisco default or equal to one-third the agreed-upon hold time (180 seconds).
- If the periodic timer is set to a value of zero (0), no keepalives are sent.

Type 3: BGP Update Message

- The UPDATE messages contain all the information BGP uses to construct a loop-free picture of the internetwork.
- Update messages advertises feasible routes, withdrawn routes, or both.
- The three basic components of an UPDATE message are:
  - **Network-Layer Reachability Information (NLRI)**
  - **Path Attributes**
  - **Withdrawn Routes**

Type 3: BGP Update Message

Network-Layer Reachability Information (NLRI)

- This is one or more (Length, Prefix) tuples that advertise IP address prefixes and their lengths.
- 192.168.160.0/19
  - Prefix = 192.168.160.0
  - Prefix Length = 19

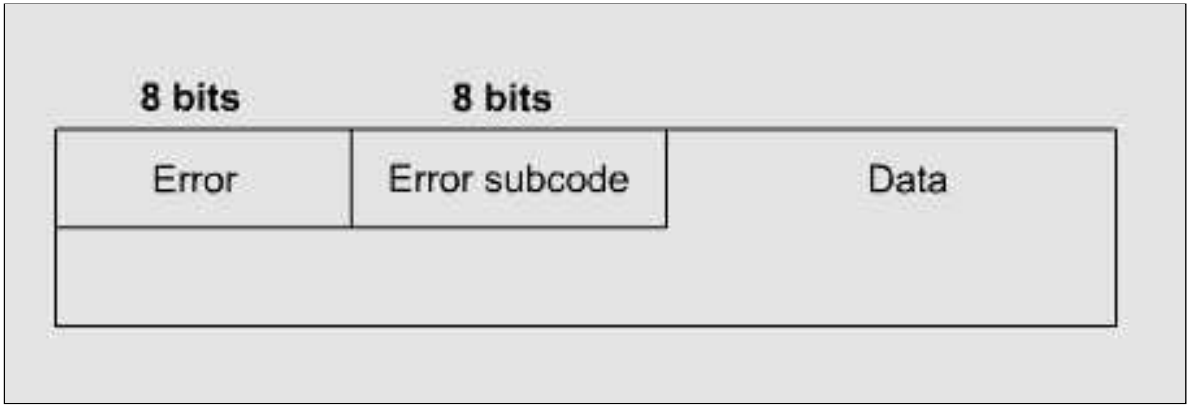
Path Attributes

- This is described later, providing the information that allows BGP to choose a shortest path, detect routing loops, and determine routing policy.

Withdrawn Routes

- These are (Length, Prefix) tuples describing destination that have become unreachable and are being withdrawn from service.
- An update message that has no NLRI or path attribute information is used to advertise only routes to be withdrawn from service.

Type 4: BGP Notification Message



- A NOTIFICATION message is sent whenever an **error** is detected and always causes the BGP connection to close.
- The NOTIFICATION message is composed of the Error Code (8 bits), Error Subcode (8 bits), and a Data fields (variable length).

Path Attributes

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

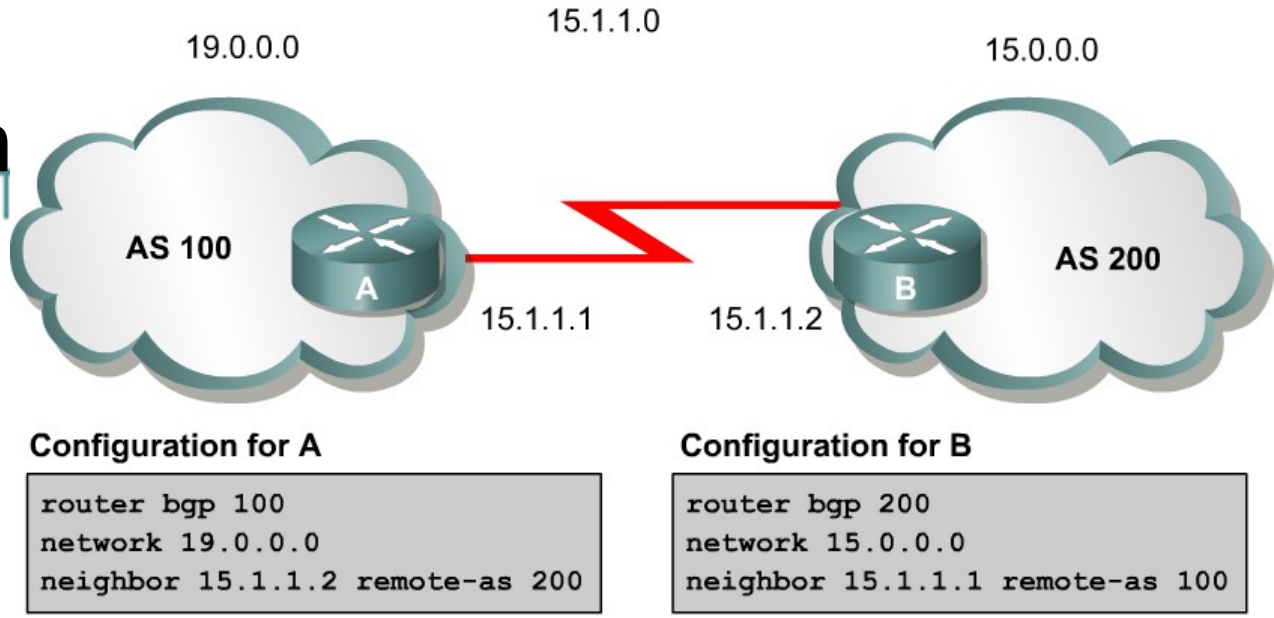
- Much of the work you will do configuring BGP focuses on **path attributes**.
- Each route has its own set of defined attributes, which can include path information, route preference, next-hop, and aggregation information.
- Administrators use these values to enforce **routing policy**.
- Based on attribute values, you can configure BGP to filter routing information, prefer certain paths, or otherwise customize its behavior.
- Every **UPDATE message** has a variable-length sequence of path attributes in the form <attribute type, attribute length, attribute value>.

Path Attributes

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

- Since you will use path attributes extensively when configuring routing policy, you should note that not all vendor implementations of BGP recognize the same attributes.
- In fact, path attributes come in four different types:
  - **Well-known mandatory**
  - **Well-known discretionary**
  - **Optional transitive**
  - **Optional non-transitive**

BGP Configuration

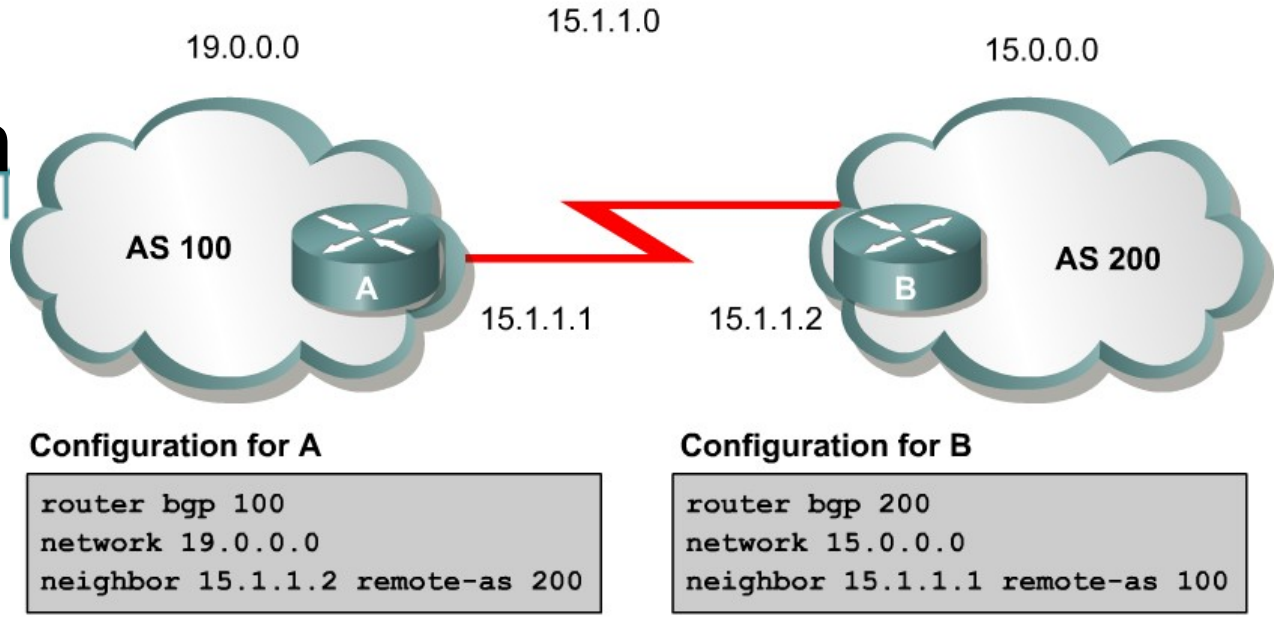


- To begin configuring a BGP process, issue the following familiar command:

  `Router(config)#`**`router bgp`** ***`AS-number`***

- BGP configuration commands appear on the surface to mirror the syntax of familiar IGP (for example, RIP, OSPF) commands.
- Although the syntax is similar, the function of these commands is significantly different.
- **Note**: Cisco IOS permits only one BGP process to run at a time, thus, **a router cannot belong to more than one AS.**

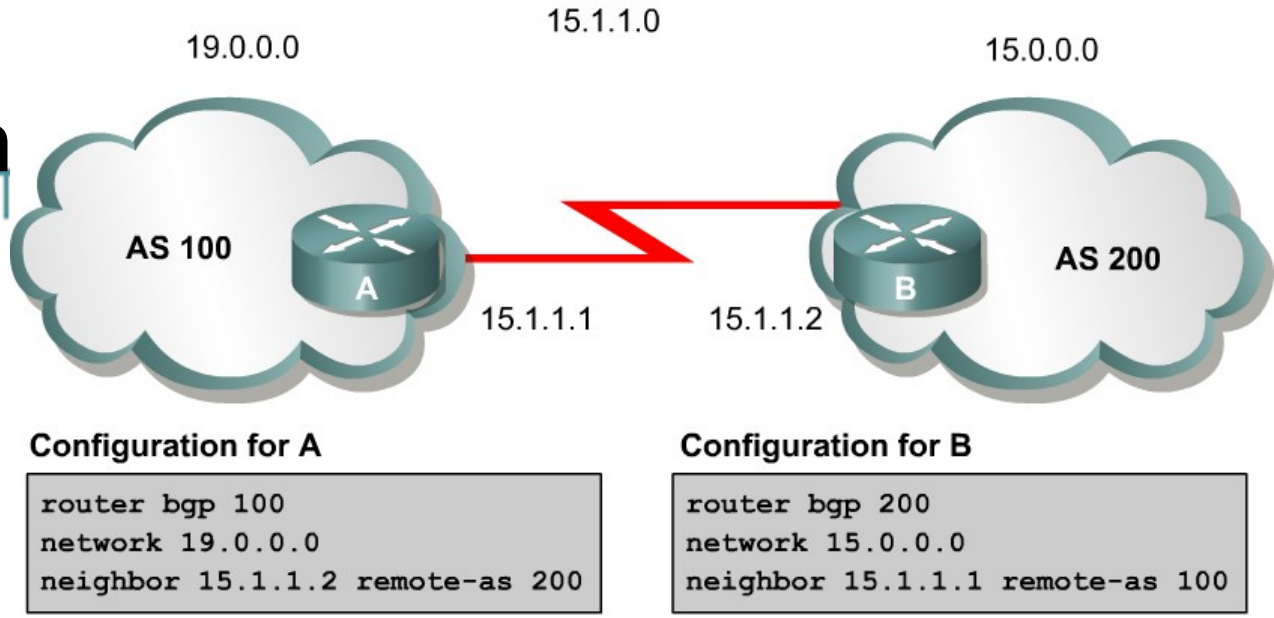
BGP Configuration



`Router(config-router)#`**`network`** ***`network-number`*** **`[mask`** ***`network-mask`*** **`]`**

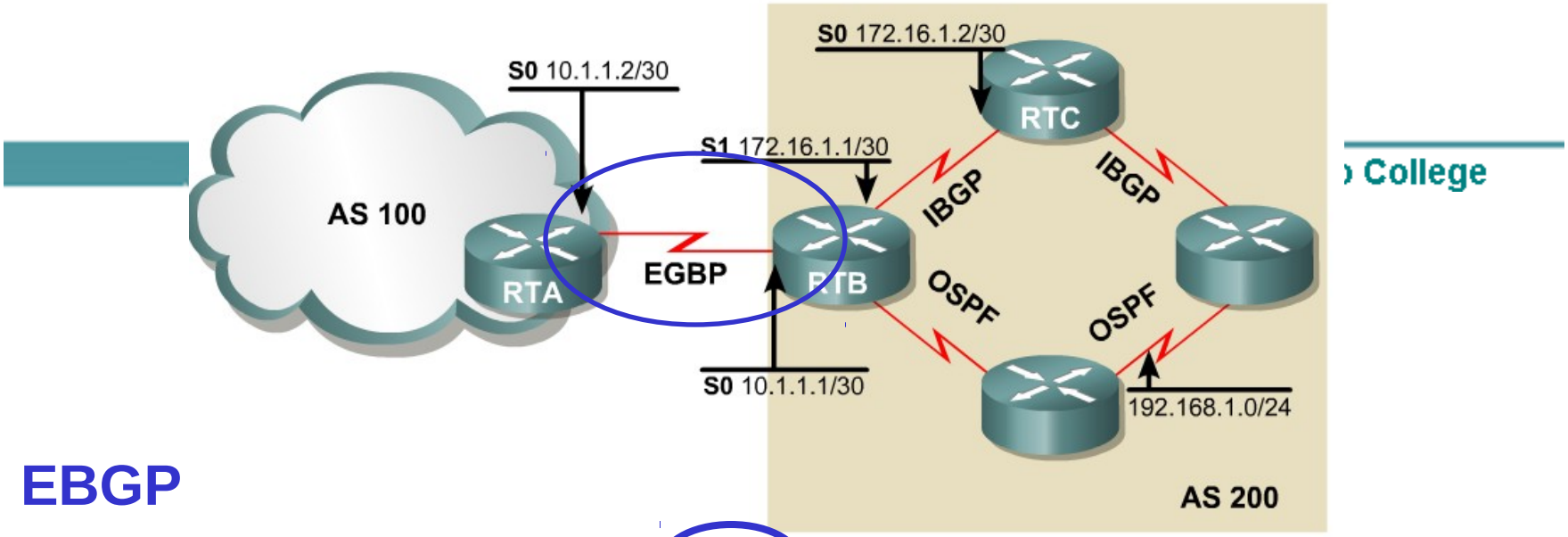
- The **network** command is used with **IGPs**, such as RIP, to determine the interfaces on which to send and receive updates, as well as which directly connected networks to advertise.
- However, when configuring **BGP**, the **network** command does **not** affect what interfaces BGP runs on.
- In BGP, the **network** command tells the BGP process **what locally learned networks to advertise**.
- The networks can be **connected routes, static routes, or routes learned via a dynamic routing protocol**, such as RIP.
  - Thus, configuring just a **network** statement will **not** establish a BGP neighbor relationship. This is a major difference between BGP and IGPs.

BGP Configuration



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

- In order for a BGP router to **establish a neighbor relationship with another BGP router**, you must issue the this configuration command.
- This command serves to identify a peer router with which the local router will establish a session.
- The ***AS-number*** argument determines whether the neighbor router is an EBGP or an IBGP neighbor.



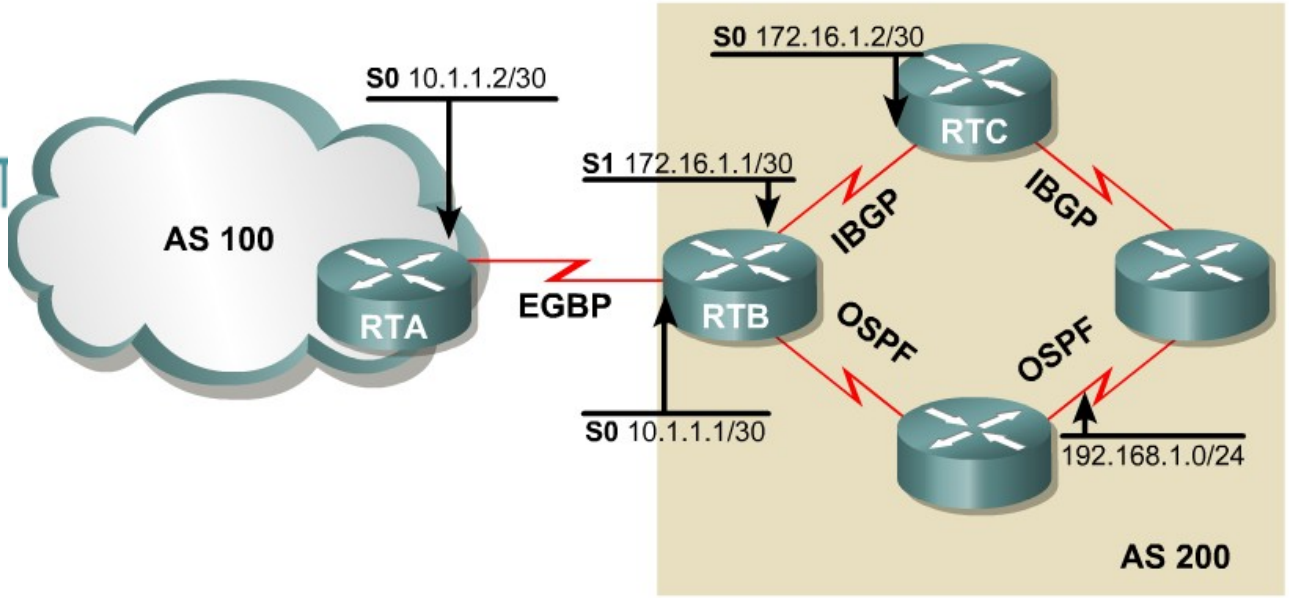
EBGP

```
RTA(config)#router bgp 100
RTA(config-router)#neighbor 10.1.1.1 remote-as 200

RTB(config)#router bgp 200
RTB(config-router)#neighbor 10.1.1.2 remote-as 100
```

- RTB: Note that the **neighbor** command's **remote-as** value, 100, is different from the AS number specified by the **router bgp** command (200).
- Because the two AS numbers are different, BGP will start an **EBGP** connection with RTA.
- Communication will occur between autonomous systems.

BGP Configuration



- Finally, whenever you are configuring BGP, you will notice that changes you make to an existing configuration may not appear immediately.
- To force BGP to clear its table and reset BGP sessions, use the **clear ip bgp** command. The easiest way to enter this command is as follows:

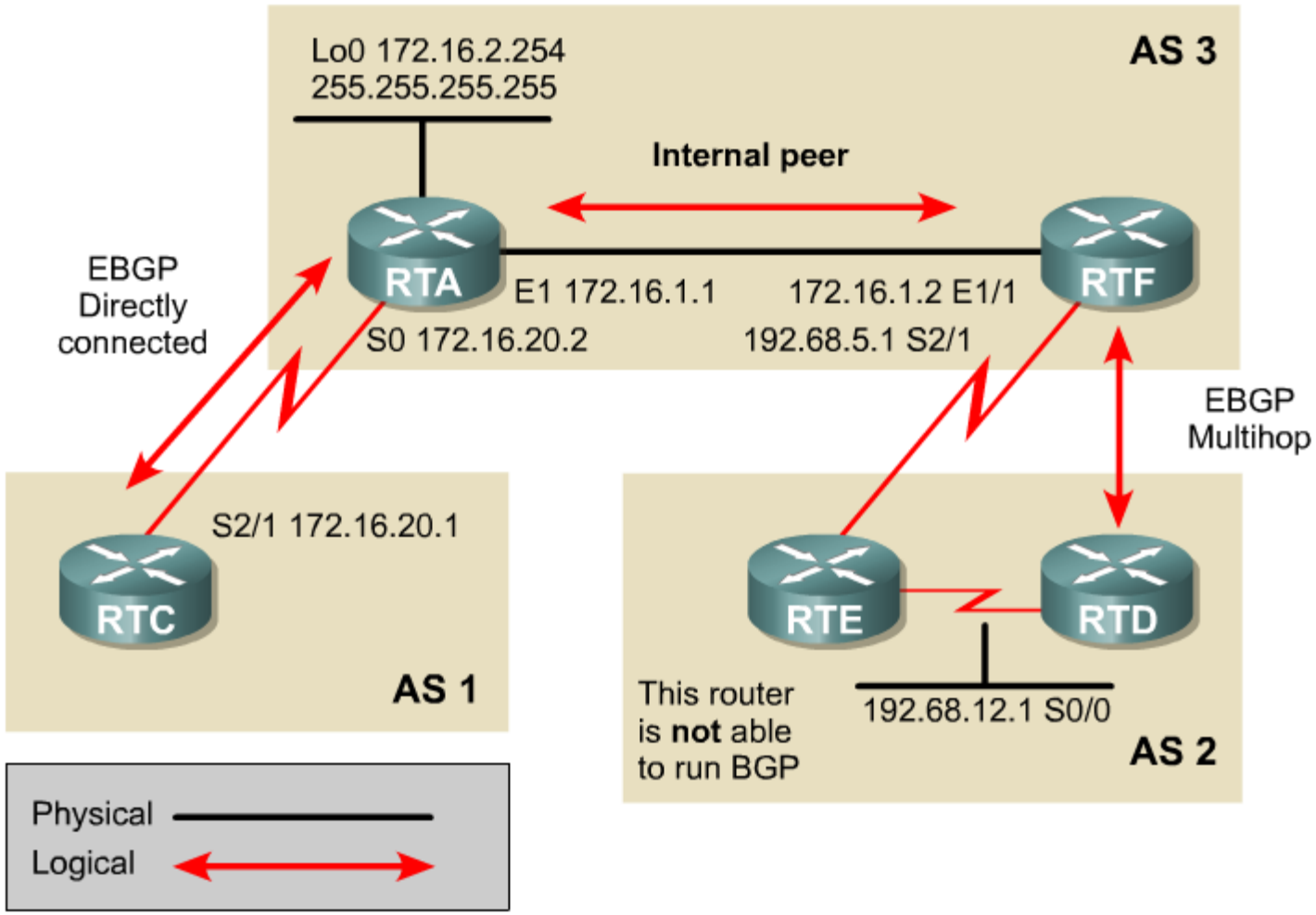
```
Router#clear ip bgp *
Router#clear ip bgp 10.0.0.0
```

Use this command with CAUTION, better yet, not at all, in a production network.

From the net…

- "clear ip bgp * OOPS! Not me but a colleague who was an employee of a large ISP with a 3 letter title. Got back from a Cisco routing course and thought they would try out some commands on the core network. It took 45 minutes for the core to reconverge. P45 followed"

Example


Lo0 172.16.2.254
255.255.255.255
AS 3
Internal peer
EBGP
Directly
connected
RTA
E1 172.16.1.1
172.16.1.2 E1/1
RTF
S0 172.16.20.2
192.68.5.1 S2/1
EBGP
Multihop
S2/1 172.16.20.1
RTC
AS 1
RTE
RTD
This router
is not able
to run BGP
192.68.12.1 S0/0
AS 2
Physical
Logical

Verifying BGP Configuration

- If the router has not installed the BGP routes you expect, you can use the **show ip bgp** command to verify that BGP has learned these routes.
- ***More later…***

```
RTA#show ip bgp
BGP table version is 3, local router ID is 10.2.2.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i1.0.0.0         192.168.1.6              0    100      0 200 400 e
*>i10.1.1.1/32     10.1.1.1                 0    100      0 i
*>i172.16.1.0/24   10.1.1.1                 0    100      0 i
* i192.168.1.32/27 192.168.1.6              0    100      0 200 i
```

Verifying BGP Configuration

- If an expected BGP route does not appear in the BGP table, you can use the **show ip bgp neighbors** command to verify that your router has established a BGP connection with its neighbors.

```
RTA#show ip bgp neighbors
BGP neighbor is 172.24.1.18,  remote AS 200, external link
  BGP version 4, remote router ID 172.16.1.1
  BGP state = Established, up for 00:03:25
  Last read 00:00:25, hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received
    Address family IPv4 Unicast: advertised and received
  Received 7 messages, 0 notifications, 0 in queue
  Sent 8 messages, 0 notifications, 0 in queue
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 30 seconds
<output omitted>
```

Implementing Policy

Common BGP Attributes

- Next Hop
- AS_Path
- Atomic Aggregate
- Aggregator
- Local Preference
- Weight
- Multiple Exit Discriminator (MED)
- Origin

- Traffic inside and outside an AS always flows according to the road map laid out by routes.
- Altering the routes changes traffic behavior.
  - How do I prevent my private networks from being advertised?
  - How do I filter routing updates coming from a particular neighbor?
  - How do I make sure that I use this link or this provider rather than another one?

BGP Attributes

- ORIGIN
- NEXT_HOP
- AS_PATH
- LOCAL_PREF
- Weight
- MULTI_EXIT_DISC (MED)

The ORIGIN attribute

- Well-known mandatory attribute (type code 1)
- Indicates the origin of the routing update
  - **IGP:** The prefix is internal to the originating AS.
  - **EGP**: The prefix was learned via some EGP, such as BGP.
  - **INCOMPLETE**: The prefix was learned by some other means, probably redistribution.
- BGP considers the ORIGIN attribute in its decision-making process to establish a preference ranking among multiple routes.
- Specifically, BGP prefers the path with the lowest origin type, where
  - IGP is lower than EGP
  - and EGP is lower than INCOMPLETE.

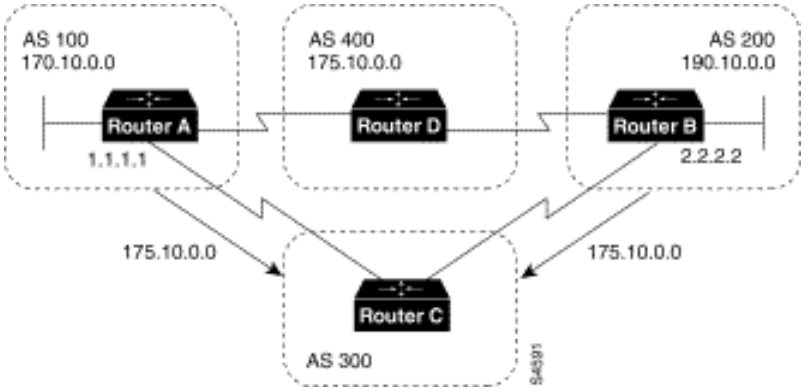
NEXT_HOP

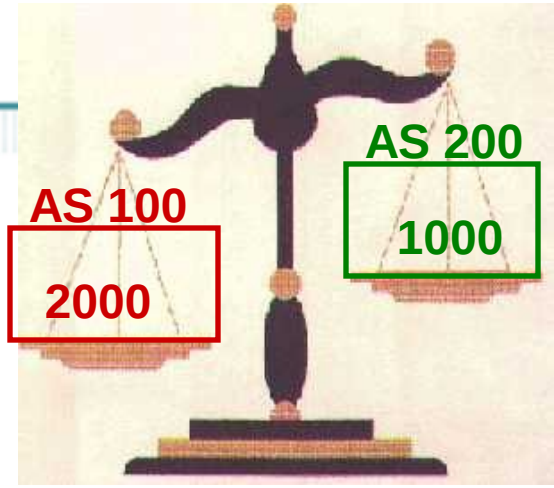
- The **NEXT_HOP** attribute is a well-known mandatory attribute (type code 3).
- In terms of an **IGP**, such as RIP, the “next hop” to reach a route is the IP address of the router that has announced the route.
  - Note: The abbreviation **IGP** (Interior Gateway Protocol) will always be in **green**, so not to get it confused with **IBGP** (Interior BGP)
- The **NEXT_HOP** concept with BGP is slightly more elaborate.
- For **EBGP** sessions, the next hop is **the IP address of the neighbor that announced the route**
- For **IBGP** sessions, **for routes originated inside the AS, the next-hop is the IP address of the neighbor that announced the route**.
- **For routes injected into the AS via EBGP, the next hop learned from EBGP is carried unaltered into IBGP**.
  - The next hop is the IP address of the EBGP neighbor from which the route was learned.

AS_PATH

- An **AS_PATH** attribute is a well-known mandatory attribute (type code 2).
- It is the **sequence of AS numbers a route has traversed to reach a destination**.
- The AS that originates the route adds its own AS number when sending the route to its external BGP peers.
- Thereafter, each AS that receives the route and passes it on to other BGP peers will prepend its own AS number to the list.
- **Prepending** is the act of adding the AS number to the beginning of the list.
- The **final list** represents all the AS numbers that a route has traversed with the AS number of the AS that originated the route all the way at the end of the list.
- This type of **AS_PATH** list is called an AS_SEQUENCE, because all the AS numbers are ordered sequentially.

The WEIGHT attribute





- The weight attribute is a **special Cisco attribute** that is used in the path selection process **when there is more than one route to the same destination**.
- The weight attribute is **local to the router on which it is assigned**, and it is **not propagated** in routing updates.
- By **default**, the weight attribute is **32768** for paths that the router originates and **zero** for other paths.
- Routes with a **higher weight are preferred** when there are multiple routes to the same destination.

The LOCAL_PREF Attribute

- Well-known discretionary attribute (type code 5).
- Degree of preference given to a route to compare it with other routes for the same destination
  - Higher LOCAL_PREF values are preferred
- **Local to the AS**
  - Exchanged between IBGP peers only
  - It is not advertised to EBGP peers
- **Routers within a multi-homed AS may learn that they can reach the same destination network via neighbors in two (or more) different autonomous systems**.
  - There could be two or more exit points from the local AS to any given destination.
- You can use the **LOCAL_PREF** attribute **to force your BGP routers to prefer one exit point over another** when routing to a particular destination network.

The MED attribute

- The MULTI_EXIT_DISC (Multi-Exit Discriminator) attribute is an optional non-transitive attribute (type code 4).
- **Informs external neighbors about the preferred path** into an AS that has multiple entry points.
- A lower MULTI_EXIT_DISC (or MED) is preferred over a higher MED.

The MED attribute

Multi-Exit Discriminator Attribute

- The multi-exit discriminator (MED) attribute is a ***hint*** **to external neighbors** about the preferred path **into an AS** when there are **multiple entry points** into the AS.
- A **lower MED value is preferred** over a higher MED value.
- The **default** value of the **MED** attribute is **0**.
- Unlike local preference, the **MED attribute is exchanged between AS's**, but a **MED attribute that comes into an AS does not leave the AS.**
- When an update enters the AS with a certain MED value, that value is used for decision making within the AS.
- When BGP sends that update to another AS, the MED is reset to 0.
- Unless otherwise specified, the router compares MED attributes for paths from external neighbors that are in the same AS.
- **If you want MED attributes from neighbors in other ASes to be compared**, you must configure the **bgp always-compare-med** command.

Summary of the BGP Path Selection Process

- BGP selects only one path as the best path.
- When the path is selected, BGP puts the selected path in its routing table and propagates the path to its neighbors.
- BGP uses the following criteria, in the order presented, to select a path for a destination:

1. If the path specifies a next hop that is inaccessible, drop the update.
2. Prefer the path with the **largest weight**.
3. If the weights are the same, prefer the path with the **largest local preference**.
4. If the local preferences are the same, prefer the **path that was originated by BGP** running on this router.
5. If no route was originated, prefer the route that has the **shortest AS_path**.
6. If all paths have the same AS_path length, prefer the path with the **lowest origin** type (where IGP is lower than EGP, and EGP is lower than Incomplete).
7. If the origin codes are the same, prefer the path with the **lowest MED attribute**.
8. If the paths have the same MED, prefer the **external path** over the internal path.
9. If the paths are still the same, prefer the path through the **closest IGP neighbor**.
10. Prefer the path with the lowest IP address, as specified by the BGP **router ID**.