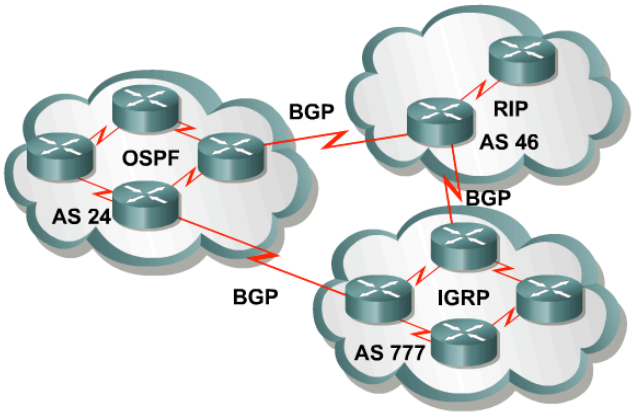
The Network Layer

Routing Algorithms

Exterior Gateway Protocols

- Protocols that run outside an enterprise, or between autonomous systems (ASs), are called exterior gateway protocols (EGPs).
- Typically, EGPs are used to exchange routing information between Internet Service Providers (ISPs), or in some cases between a customer's AS and the provider's network.
- Border Gateway Protocol, version 4 (BGP4), is the most common EGP and is considered the Internet standard.

Exterior Gateway Protocols

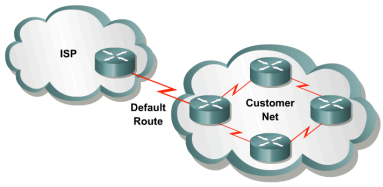


EGPs, such as BGP, are used to interconnect autonomous systems

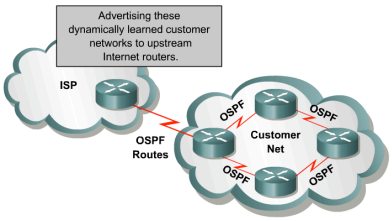
Autonomous Systems

- Because the routing and security policies of one organization may conflict with the policies of another, internetworks are divided into domains, or autonomous systems (AS).
- Each AS typically represents an independent organization and applies its own unique routing and security policies.
- EGPs facilitate the sharing of routing information between autonomous systems
- Each AS has an identifying number that is assigned by an Internet registry or a service provider. This number is between one (1) and 65,535.
- AS numbers within the range of 64,512 through 65,535 are reserved for private use.

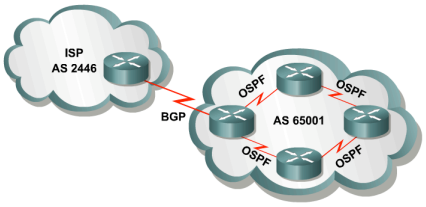
Single Homed ASs



Using Static Routing with summarization to announce customer routes to the outside world

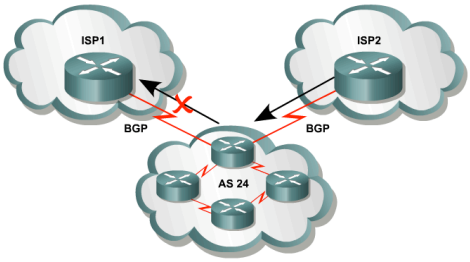


Using a IGP to learn the customer routes



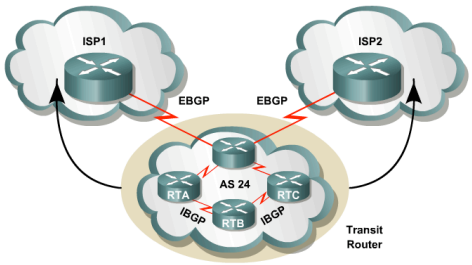
Using a EGP, such as BGP, to learn and announce the customer routes

Multihomed Nontransit AS



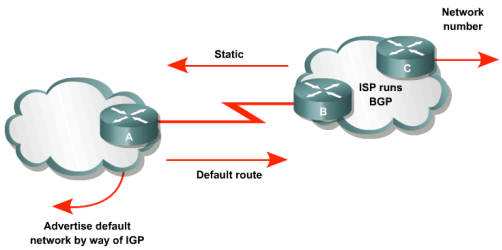
- An AS is a multihomed system if it has more than one exit point to outside networks
- An AS connected to the Internet can be multihomed to a single provider or multiple providers
- A nontransit AS does not allow transit traffic to pass through it. Transit traffic is any traffic that has a source and destination outside the AS
- A nontransit AS would advertise only its own routes to both providers to which it connects. **It would not advertise routes that it learned from one provider to another**.

Multihomed Transit ASs



- A multihomed transit system has more than one connection to the outside world and can be used for transit traffic by other autonomous systems.
- The multihomed AS views transit traffic as any traffic originating from outside sources bound for outside destinations
- A transit AS can route transit traffic by running BGP internally so that multiple border routers in the same AS can share BGP information
- 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)

When NOT to use BGP

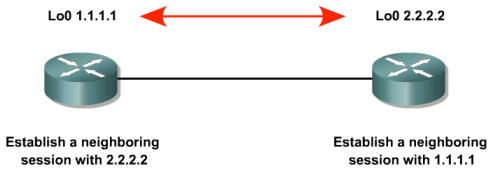


Do not use BGP within the AS in the following situations:
•There is only a single connection to the Internet or other AS.
•Internet routing policy and route selection are not of concern to the AS.
•The BGP routers have insufficient RAM or processor power to handle constant updates.
•There is limited understanding of route filtering and the BGP path selection process.
•Low bandwidth link between autonomous systems

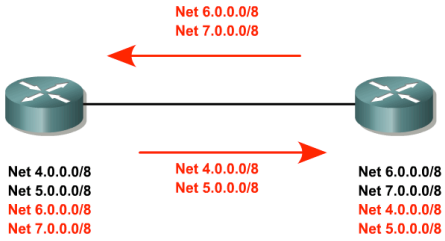
Basic BGP Operation

- The function of BGP is to exchange routing information between autonomous systems and guarantee the selection of a loop free path
- Common IGPs such as RIP and OSPF use technical metrics. BGP does not use technical metrics. Instead, BGP makes routing decisions based on network policies, or rules.
- BGP updates are carried using TCP on port 179
- To guarantee loop free path selection, BGP constructs a graph of autonomous systems based on the information exchanged between BGP neighbors
- 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.

Basic BGP Operation

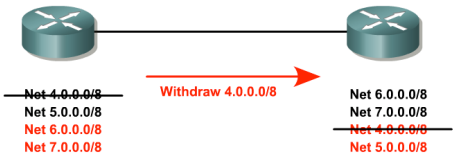


When two routers establish a TCP enabled BGP connection, they are called neighbors or peers.
Each router running BGP is called a BGP speaker.



When BGP neighbors first establish a connection, they exchange all candidate BGP routes.

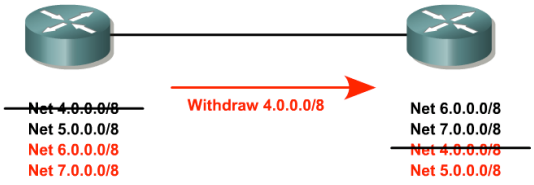
Basic BGP Operation



After this initial exchange, incremental updates are sent as network information changes.

- Incremental updates are more efficient than complete table updates. This is especially true with BGP routers, which may contain the **complete Internet routing table**.
- Peers advertise destinations that are reachable through them by using update messages. These messages contain **route prefix, AS path, path attributes** such as the degree of preference for a particular route, and other properties.
- If there are no routing changes to transmit to a peer, a BGP speaker will periodically send keepalive messages to maintain the connection. These **19-byte keepalive packets are sent every 60 seconds by default**.

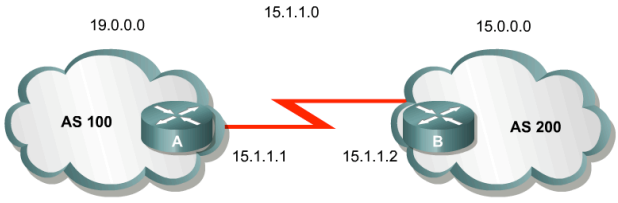
Basic BGP Operation



Rather than advertise reachable destinations as a network and a subnet mask, BGP advertises them using network-layer reachability information (NLRI), which consists of prefixes and prefix lengths.

NET= 10.1.0.0 MASK= 255.255.255.0 advertised as 10.1.0.0/24

A simple BGP Configuration



Configuration for A

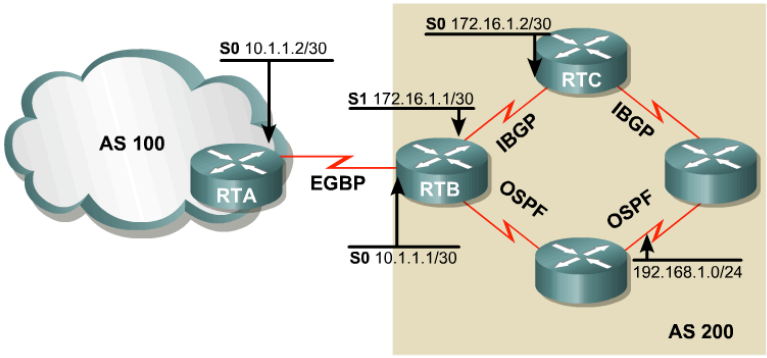
```
router bgp 100
network 19.0.0.0
neighbor 15.1.1.2 remote-as 200
```

Configuration for B

```
router bgp 200
network 15.0.0.0
neighbor 15.1.1.1 remote-as 100
```

Command	Description
`show ip bgp`	Displays entries in the BGP routing table. A network can be specified to get more detailed information about a particular prefix. Use the subnets keyword to get information about a particular prefix and all its subnets.
`show ip bgp summary`	Displays a summary of all BGP connections.
`show ip bgp neighbors`	Displays detailed information for each BGP connection.
`show ip bgp paths`	Displays all the BGP paths in the database.

EBGP and IBGP



EBGP session – Occurs between routers in two different autonomous systems. These routers are usually adjacent to one another, sharing the same medium and a subnet (RTA)

IBGP session – Occurs between routers in the same AS, and is used to coordinate and synchronize routing policy within the AS. Neighbors may be located anywhere in the AS, even several hops away from one another. An IBGP session typically occurs between routers in the same AS in an ISP

BGP Message Types

- **Open Message –** This message is used to establish connections with peers and includes fields for the BGP version number, the AS number, hold time, and Router ID.
- **Keepalive Message –** This message type is sent periodically between peers to maintain connections and verify paths held by the router sending the keepalive. The recommended keepalive interval is one third of the hold time interval.
- **Notification Message –** This message type is used to inform the receiving router of errors. This message includes a field for error codes that can be used to troubleshoot BGP connections.
- **Update Message –** The update messages contain all the information BGP uses to construct a loop free picture of the internetwork. There are three basic components of an update message. They are network-layer reachability information (NLRI), path attributes, and withdrawn routes. These three elements are described briefly in the following sections.

BGP update packet 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	—

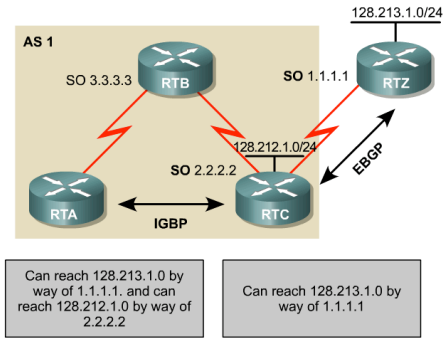
Well-known mandatory – An attribute that must exist in the BGP update packet.
Well-known discretionary – An attribute that is recognized by all BGP implementations, but may or may not be sent in the BGP update message.
Optional transitive – An attribute that may or may not be recognized by all BGP implementations. Because the attribute is transitive, BGP should accept and advertise the attribute even if it is not recognized.
Optional nontransitive – An attribute that may or may not be recognized by all BGP implementations and is not passed along to other BGP peers.

BGP Attributes: Next Hop

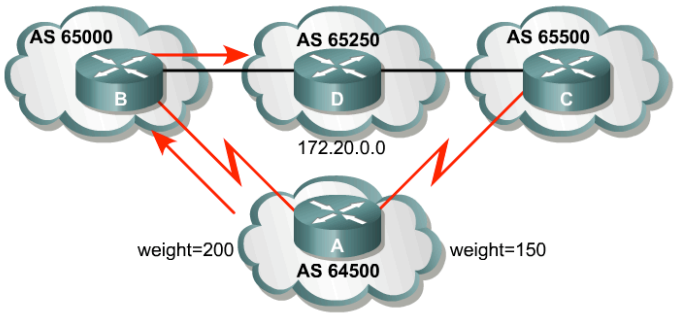
For EBGP sessions the "next hop" to reach a route is the IP address of the router that has announced the route.

For IBGP sessions, where 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 by way of EBGP, the next hop learned from EBGP is carried unaltered into IBGP.

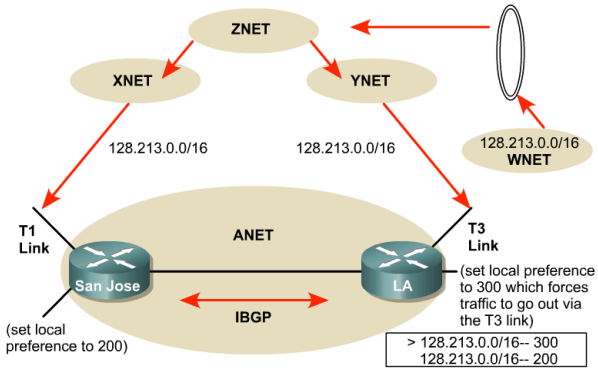


BGP Attributes: Weight



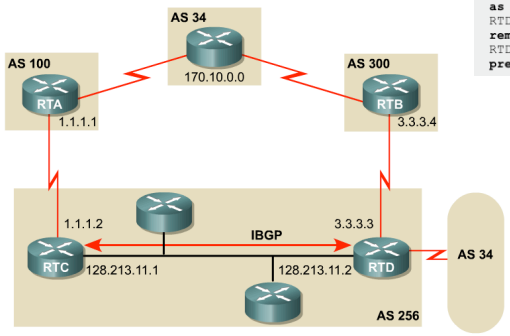
- The Weight attribute gives higher preference to the route that has a higher weight (similar to the Local Preference attribute but not exchanged between routers).
- The weight parameter has a higher precedence than any other attribute.
- It is the most important attribute when determining route preference. The Weight attribute is a Cisco proprietary attribute.

BGP Attributes: Local Preference



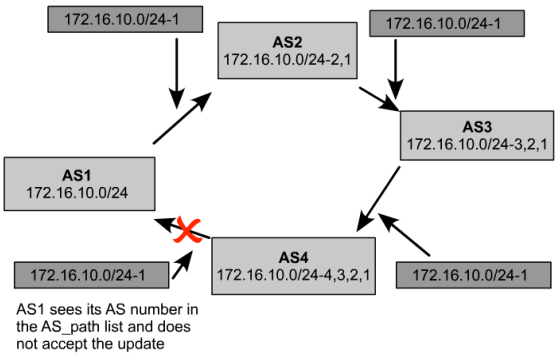
- The Local Preference attribute is a degree of preference given to a route for comparison with other routes for the same destination.
- Higher Local Preference values are preferred.
- Local Preference, as indicated by the name, is local to the AS and is exchanged between IBGP peers only.
- Local Preference is not advertised to EBGP peers.

BGP Attributes: Local Preference



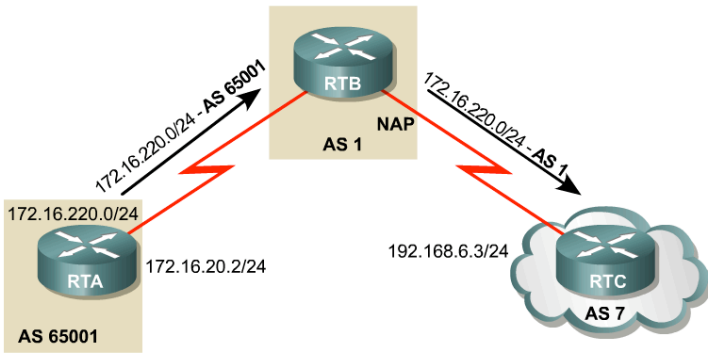

```
RTC(config)#router bgp 256
RTC(config-router)#neighbor 1.1.1.1 remote-
as 100
RTC(config-router)#neighbor 128.213.11.2
remote-as 256
RTC(config-router)#bgp default local-
preference 150

RTD(config)#router bgp 256
RTD(config-router)#neighbor 3.3.3.4 remote-
as 300
RTD(config-router)#neighbor 128.213.11.1
remote-as 256
RTD(config-router)#bgp default local-
preference 200
```

BGP Attributes: AS_Path



- AS_Path attribute is the sequence of AS numbers that a route has traversed to reach a destination.
- AS_Path information is one of the attributes that BGP looks at to determine the best route to reach a destination.
- In comparing two or more different routes, given that all other attributes are identical, a shorter path is always preferred. In case of a tie in AS_Path length, other attributes are used to make the decision.

AS_Path and Private AS Numbers

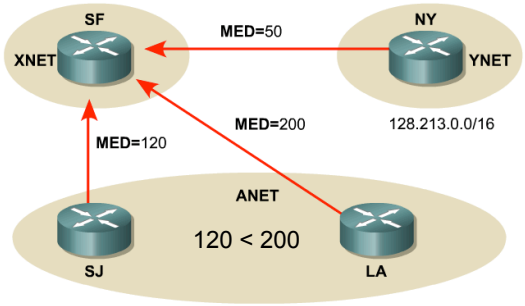


- If a customer is single-homed or multihomed to the same provider, the provider generally requests that the customer use an AS number taken from the private pool, 64,512 to 65,535.
- Private AS numbers must be stripped off before leaving the provider's system

BGP Attributes: Origin

- The Origin attribute indicates the origin of the routing update.
- BGP allows the following three types of origins:
  - **IGP –** The prefix is internal to the originating AS.
  - **EGP –** The prefix was learned by way of 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.

BGP Attributes: Multiple Exit Discriminator



- MED informs external neighbors about the preferred path into an AS that has multiple entry points.
- A lower MED is preferred over a higher MED.
- By default the router compares MED attributes for paths **from external neighbors that are in the same AS**
- When BGP forwards the routing update to another AS, the MED is reset to zero. This is true unless the outgoing MED is set to a specific value.

Overview of the BGP routing Process

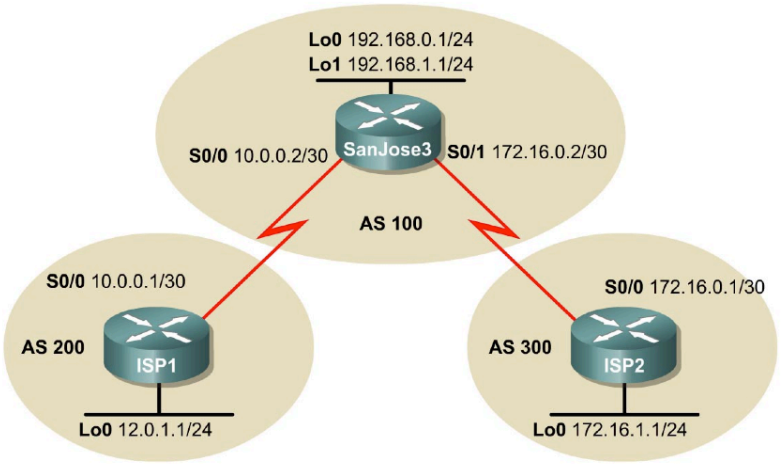
- The Cisco implementation of BGP keeps track of all BGP updates in a BGP table separate from the IP routing table.
- In case multiple routes to the same destination exist, BGP picks only the best route and sends it to the peers.
- BGP router may originate routing updates to advertise networks that belong to its own AS.
- Valid local routes originated in the system and the best routes learned from BGP peers are then installed in the IP routing table. The IP routing table is used for the final routing decision.

BGP Decision Process

Decision Process

1. If the next hop is inaccessible, the route is ignored.
2. The BGP router will prefer the path with the largest weight.
3. If the weights are the same, the BGP router will prefer the route with the largest local preference.
4. If the routes have the same local preference, the BGP router will prefer the route that was locally originated by this router.
5. If the local preference is the same, the BGP router will prefer the route with the shortest AS_Path.
6. If the AS_Path length is the same, the BGP router will prefer the route with the lowest origin type.
7. If the origin type is the same, the BGP router will prefer the route with the lowest MED.
8. If the routes have the same MED, the BGP router will prefer EBGP first, then confederation external, and IBGP last.
9. If all the preceding scenarios are identical, the BGP router will prefer the route that can be reached by way of the closest IGP neighbor.
10. If the internal path is the same, the BGP router ID will be a tiebreaker. The BGP router will prefer the route coming from the BGP router with the lowest router ID.

BGP Lab


Lo0 192.168.0.1/24
Lo1 192.168.1.1/24
SanJose3
S0/0 10.0.0.2/30
S0/1 172.16.0.2/30
AS 100
S0/0 10.0.0.1/30
AS 200
ISP1
Lo0 12.0.1.1/24
S0/0 172.16.0.1/30
AS 300
ISP2
Lo0 172.16.1.1/24