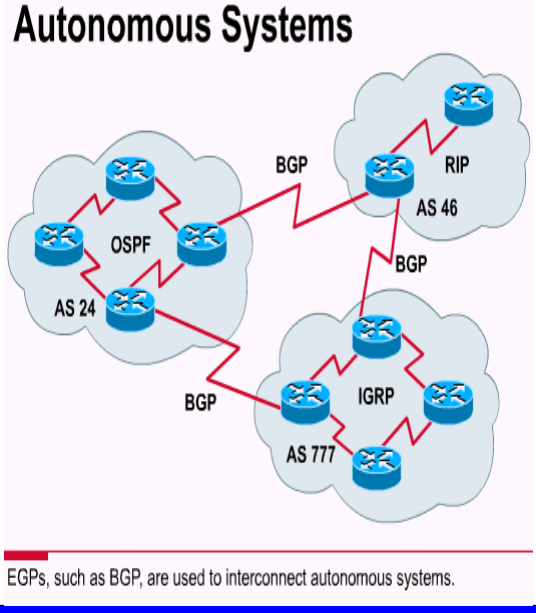
Ch.8 BGP

Lee, HoonJae
hjlee@dongseo.ac.kr
http://kowon.dongseo.ac.kr/~hjlee
http://crypto.dongseo.ac.kr

8.1 Autonomous System

8.1 Autonomous System



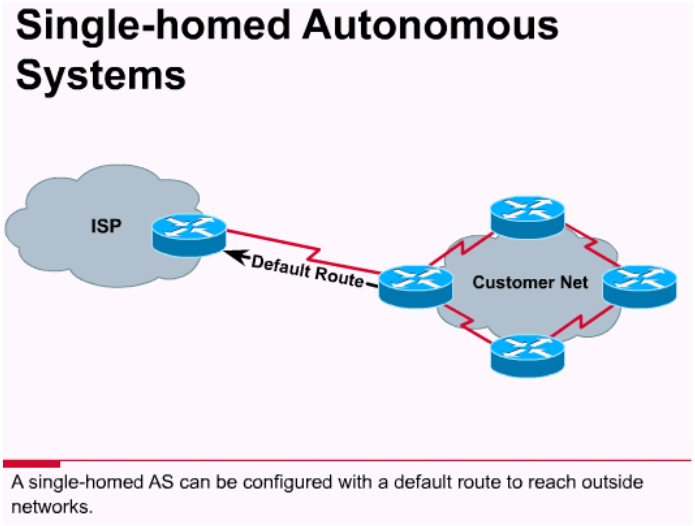
EGPs, such as BGP, are used to interconnect autonomous systems.

◆ AS is any set of routers that share similar routing policies and operate within a single administrative domain.

◆ Each AS has an identifying number, assigned by an Internet registry or a service provider, between 1 and 65,535. AS numbers within the range, 64,512 through 65,535 are reserved for private use (similar to RFC 1918 IP addresses).

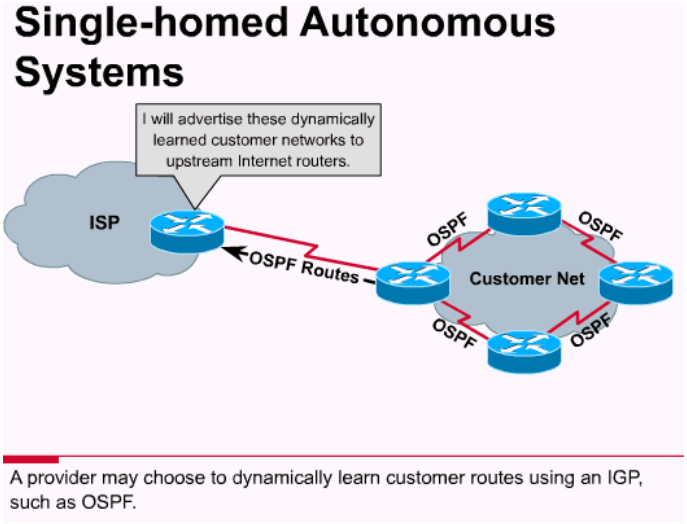
◆ Because of the finite number of available AS numbers, an organization must present justification of its need before it will be assigned an AS number.

8.1 Autonomous System

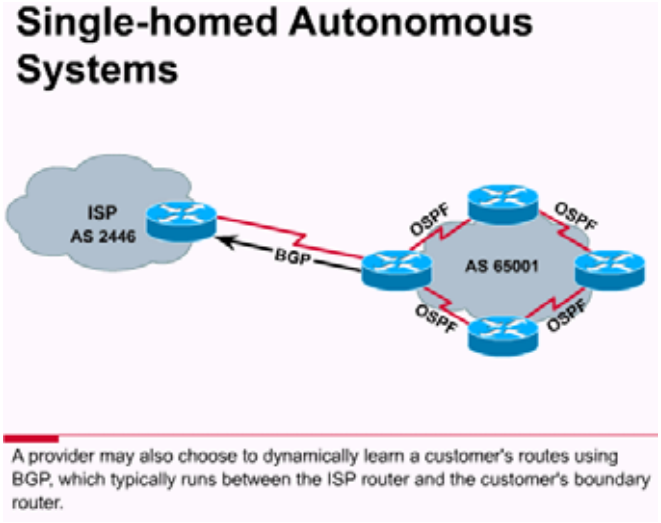


A single-homed AS can be configured with a default route to reach outside networks.

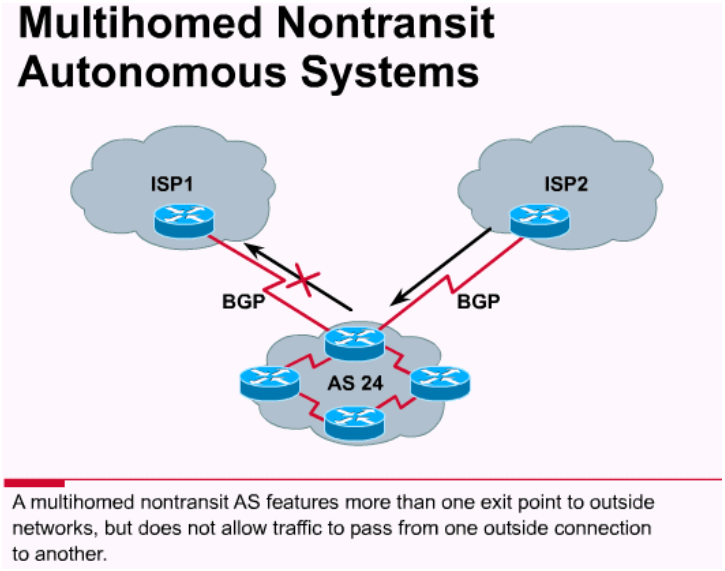
8.1 Autonomous System



8.1 Autonomous System

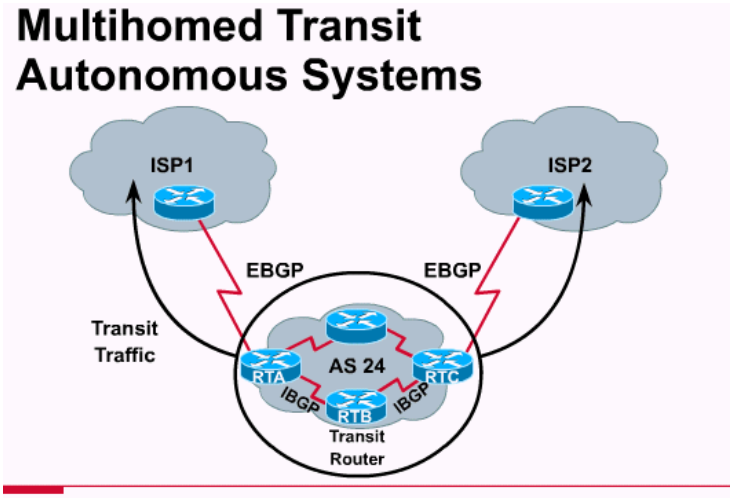


8.1 Autonomous System



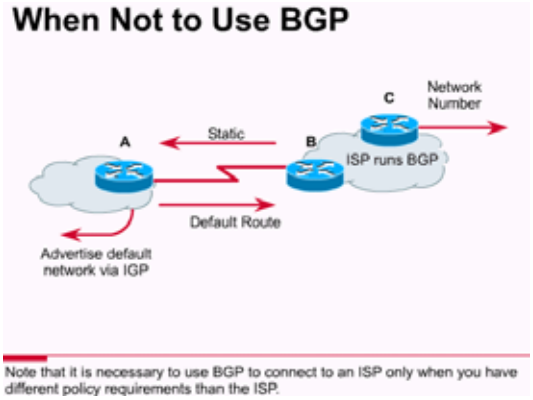
A multihomed nontransit AS features more than one exit point to outside networks, but does not allow traffic to pass from one outside connection to another.

8.1 Autonomous System



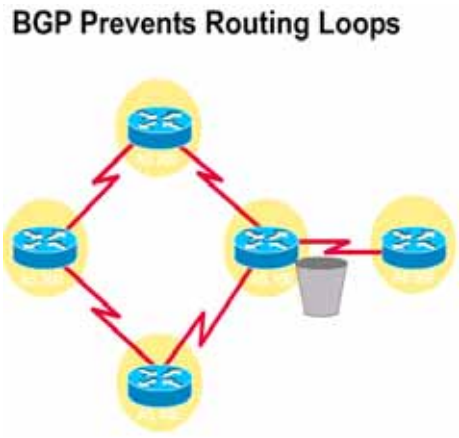
A multihomed transit system can be used for transit traffic by other autonomous systems.

8.1 Autonomous System



router A is advertising a default network into the AS through a local IGP, such as RIP. A static route affords connectivity through router B to the ISP's AS. The ISP is running BGP and is recognized by other BGP routers in the Internet.

8.2 Basic BGP Operation

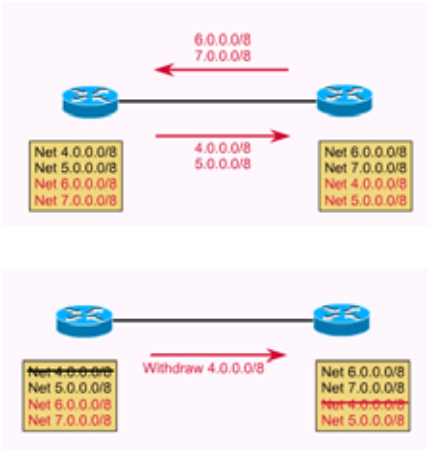


◆ BGP updates are carried using TCP on port 179. In contrast, RIP updates use UDP port 520, while OSPF does not use a Layer 4 protocol. Because BGP requires TCP, IP connectivity must exist between BGP peers, and TCP connections must be negotiated between them before updates can be exchanged. Thus, BGP inherits TCP's reliable, connection-oriented properties.

◆ To guarantee loop-free path selection, BGP constructs a graph of autonomous systems based on the information exchanged between BGP neighbors. As far as BGP is concerned, the whole internetwork is a graph, or tree, of autonomous systems.

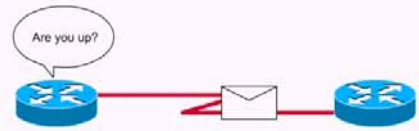
8.2 Basic BGP Operation

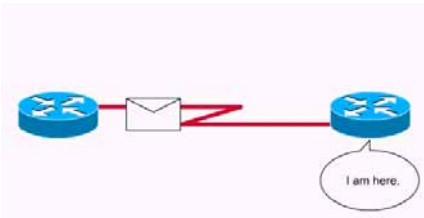
How BGP Works



8.2 Basic BGP Operation

Keepalive Messages

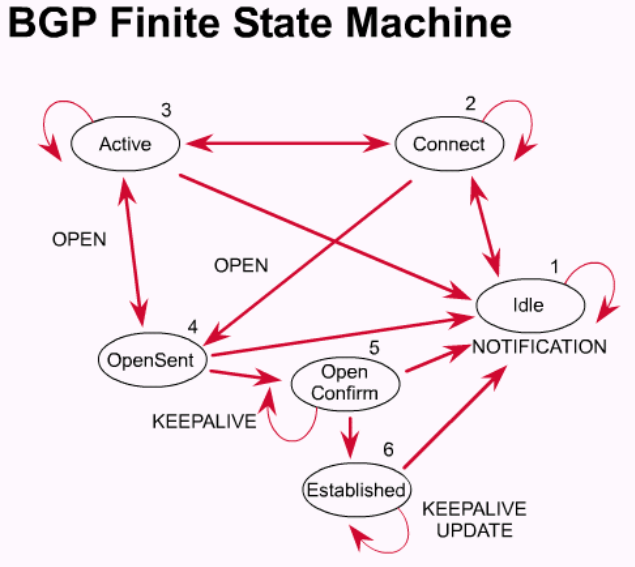




Possible BGP Error Codes

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 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

8.2 Basic BGP Operation



8.2 Basic BGP Operation

The six states of the BGP FSM are described below:

◆ **Idle** - Idle is the first state of a BGP connection. BGP is waiting for a start event, which is normally initiated by an administrator or a network event. At the start event, BGP initializes its resources, resets a connect retry timer, and starts listening for a TCP connection that may be initiated by a remote peer. BGP then transitions to a Connect state. Note that BGP can transition back to Idle from any other state in case of errors.

◆ **Connect** - In the connect state; BGP is waiting for the TCP connection to be completed. If the TCP connection is successful, the state transitions to OpenSent. If the TCP connection fails, the state transitions to the Active state, and the router tries to connect again. If the connect retry timer expires, the state remains in the Connect state, the timer is reset, and a TCP connection is initiated. In case of any other event (initiated by system or administrator), the state returns to Idle.

◆ **Active** - In the Active state, BGP is trying to acquire a peer by initiating a TCP connection. If it is successful, it transitions to OpenSent. If the connect retry timer expires, BGP restarts the connect timer and falls back to the Connect state. While active, BGP is still listening for a connection that may be initiated from another peer. The state may go back to Idle in case of other events, such as a stop event initiated by the system or the operator.

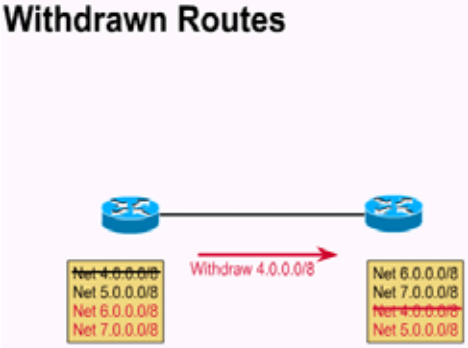
8.2 Basic BGP Operation

◆ **OpenSent** - In the OpenSent state, BGP is waiting for an open message from its peer. The open message is checked for correctness. In case of errors, such as an incompatible version number or an unacceptable AS, the system sends an error notification message and goes back to idle. If there are no errors, BGP starts sending keepalive messages and resets the keepalive timer. At this stage, the hold time is negotiated and the smaller value is taken. If the negotiated hold time is 0, the hold timer and the keepalive timer are not restarted.

◆ **OpenConfirm** - While in OpenConfirm state, BGP is waiting for a keepalive or notification message. If a keepalive message is received, the state goes to the Established state, and the neighbor negotiation is complete. If the system receives an update or keepalive message, it restarts the hold timer (assuming that the negotiated hold time is not 0). If a notification message is received, the state falls back to Idle. The system sends periodic keepalive messages at the rate set by the keepalive timer. In the case of any TCP disconnect or in response to any stop event (initiated by the system or the administrator), the state falls back to Idle. In response to any other event, the system sends a notification message with an FSM error code and returns to the idle state.

◆ **Established** - Established is the final state in the neighbor negotiation; BGP starts exchanging update packets with its peers. If it is nonzero, the hold timer is restarted at the receipt of an update or keepalive message.

8.2 Basic BGP Operation



Withdrawn Routes: Withdrawn routes provide a list of routing updates that are no longer reachable and that need to be withdrawn (removed) from the BGP routing table. Withdrawn routes have the same format as NLRI.

An update message that has no NLRI or path attribute information is used to advertise only routes to be withdrawn from service.

8.2 Basic BGP Operation

BGP Attribute Codes and Their Respective Types

Attribute Code	Type
1 -- ORIGIN	Well-known mandatory
2 -- AS_PATH	Well-known mandatory
3 -- NEXT_HOP	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. It must be recognized by all BGP implementations. If a well-known attribute is missing, a notification error will be generated. This ensures that all BGP implementations agree on a standard set of attributes. An example of a well-known mandatory attribute is the AS_Path attribute.

Well-known discretionary - An attribute that is recognized by all BGP implementations, but may or may not be sent in the BGP update message. An example of a well-known discretionary attribute is the LOCAL_PREF attribute.

Optional transitive - An attribute that may or may not be recognized by all BGP implementations (thus, optional). 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. Whether or not the receiving BGP router recognizes the attribute, it is nontransitive and is not passed along to other BGP peers.

8.2 Basic BGP Operation

◆TextBook

- BGP message header → Marker, Length, Type(Fig. 8-9)
- BGP Open Message → Fig 8-10
- BGP Notification Message → Fig 8-11, Table 8-1
- BGP Keepalive Message
- BGP Update Message → Fig.8-12

8.3 Configuring BGP

Simple BGP Configuration

Configuration for A

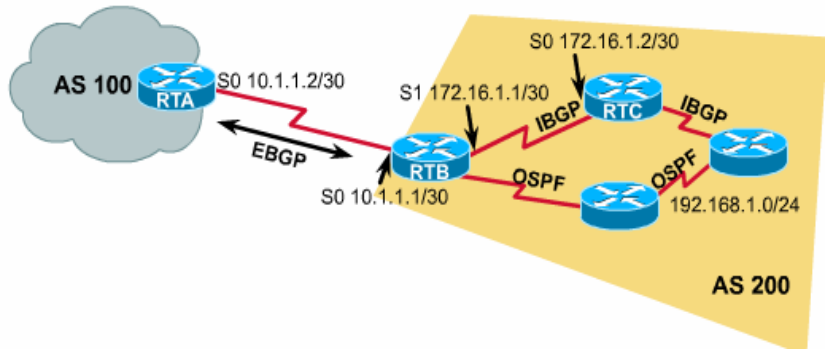
```
router bgp 100
network 19.0.0.0
neighbor 15.1.1.2 remote-as 200
```

Configuration for B

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

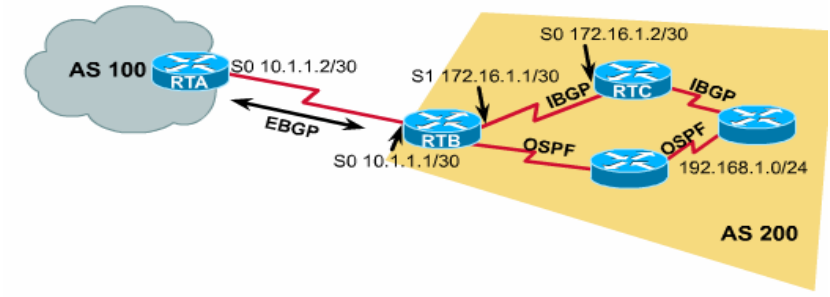

8.3 Configuring BGP

EBGP and IBGP



8.3 Configuring BGP

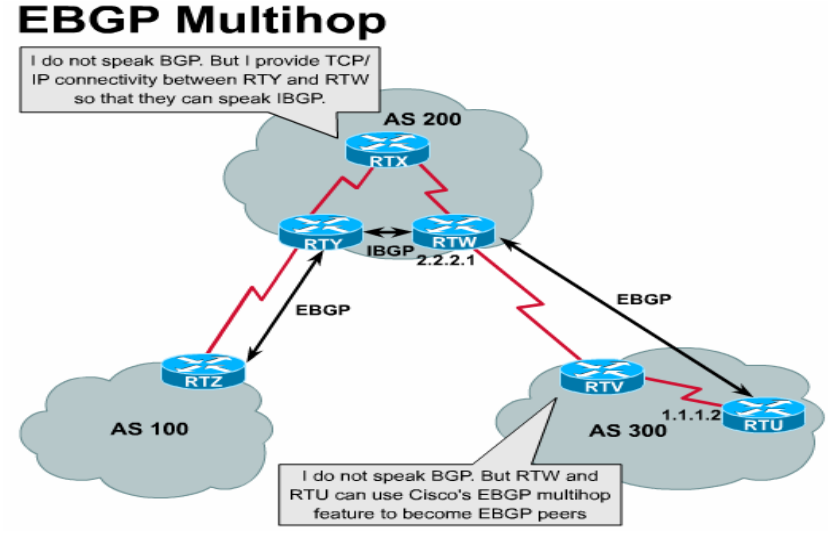
EBGP and IBGP Configuration Example



In this example network, RTB speaks EBGP to RTA, and IBGP to RTC.

8.3 Configuring BGP

EBGP Multihop



8.3 Configuring BGP

Clearing the BGP Table

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


8.3 Configuring BGP

Building Peering Sessions