



C O N C E P T S

AS (Autonomous System)—A network under a single administrative control, like a single company or a single ISP.

ASN (AS Number)—used by BGP both in choosing a best path to a destination and in preventing routing loops. Note that I will occasionally abbreviate the plural "Autonomous Systems" as "ASNs" to avoid an unfortunate acronym while only slightly mangling the literal meaning.

IGP (Internal Gateway Protocol)—Adaptively share and choose best routes *within* an AS. Examples include RIP, EIGRP, and OSPF.

EGP (Exterior Gateway Protocol)—Handle routing *between* autonomous systems, for example routing a packet from one AS to the next until it reaches the AS of its destination.

BGP (Border Gateway Protocol)—The only EGP in use today. It concentrates on advertising all reachable destinations in the internet. As of February 2017, this includes almost 700,000 entries.

eBGP (external BGP)—BGP used as an EGP to advertise prefixes between autonomous systems—the only type we care about for CCNA.

iBGP (internal BGP)—advertises prefixes between routers within an AS—not part of CCNA. Perhaps useful when an AS has more than one connection to the world and those edge routers need to tell each other about their access. In that case, it could also help internal routers choose which edge router is best for a given external destination.

PA (Path Attribute)—more complex than just a simple metric. Used by BGP to choose paths. Beyond the scope of CCNA, but simpler than OSPF, so scalable. When an update arrives, a router only has to decide if the new route is better than the current one.

B G P A T T H E E N T E R P R I S E E D G E

Single-homed—your enterprise has one connection to one ISP. This is all we care about for CCNA.

The first word refers to connection redundancy and the second refers to ISP redundancy:

	SINGLE LINK PER ISP	REDUNDANT LINKS PER ISP
SINGLE ISP	Single Homed	Dual Homed
MULTIPLE ISPS	Single Multihomed	Dual Multihomed

eBGP Advertisements

- The enterprise would advertise to the ISP a single route to the enterprise's entire public address block. No internal subnetting details would be useful
- The ISP would advertise a single default route. It could advertise all the routes in the internet, but they'd all have the next same-hop, so what's the point of overtaxing your routers

HOW BGP WORKS

BGP differs from the IGP's covered in CCNA.

- The network command serves a different purpose
- BGP has no concept of being enabled on an interface
- BGP doesn't dynamically discover neighbors; they're manually defined

Routes that are advertised can be learned in three ways:

- The BGP network command
- Redistribution—we don't know about this yet.
- From a neighbor

BGP messages use TCP port 179.

NLRI—network prefix and length information contained in update messages

Peers—Once two BGP routers pass all parameter checks, they

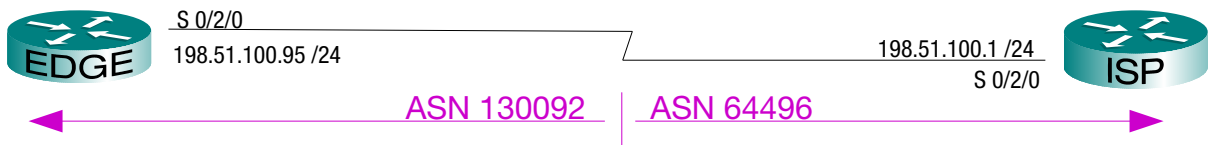
- Become peers and open a TCP connection.
- They add one or more NLRI + PA entries to their BGP tables based on their static configurations, like "network" statements
- Each peer advertises its best route for each NLRI in its table
- Each adds received networks to its own table if they're better or new

CONFIGURATION

At the enterprise edge, you'd start BGP, defining your own ASN and give it the address of a BGP router at the ISP to peer with.

```
R1(config)# router bgp <own ASN>
R1(config-router)# neighbor <ip address> remote-as <ASN>
```

For example:



```
EDGE(config)# router bgp 130092
EDGE(config-router)# neighbor 198.51.100.1 remote-as 64496
```

```
ISP(config)# router bgp 64496
ISP(config-router)# neighbor 198.51.100.95 remote-as 130092
Jan 2 13:59:04.207: %BGP-5-ADJCHANGE: neighbor 198.51.100.95 Up
The routers are now neighbors, but have no routes to share with each other
```

Verifying Neighbors

The IOS command "show tcp brief" displays all TCP connections, including our BGP connection to our ISP. We can tell that that connection [line 3] belongs to BGP because it is using port 179. TCP connections are bidirectional, so there will only be one. Whichever peer created the connection (random chance; us in this case) will use an ephemeral (temporary) port number (56662 in our case) and connect to the BGP port (179) on the other end.

```

1 EDGE# show tcp brief
2 TCB          Local Address          Foreign Address      (state)
3 48B34098    198.51.100.95.56662    198.51.100.1.179    ESTAB
4
5 EDGE# show ip bgp summary
6 BGP router identifier 198.51.100.95, local AS number 130092
7 BGP table version is 1, main routing table version 1
8
9 Neighbor      V      AS  MsgRcvd  MsgSent  TblVer  InQ  OutQ  Up/Down  State/PfxRcd
10 198.51.100.1  4      64496  53       52      1     0    0 00:44:22  0

```

The command "show ip bgp summary" provides a table with one line per peer. We can tell that our only peer [line 10] is an eBGP peer because its ASN is different from our own [line 6]. With iBGP, they'd be the same.

If the neighbor state is "established," then the far right column "State/PfxRcd" shows the number of prefixes received from the peer (in our case 0). For any other state, that column would tell you which non-established state the peer is in.

STATE	MEANING
Established	Success—the routers are peers and can exchange routes
Idle	The neighbor is administratively shutdown by the "neighbor shutdown" command or the router is waiting to retry
Connect	TCP is in the process of making a connection
Active	TCP connection complete; no BGP messages have been sent yet
Opensent	The first BGP message (an open message) has been sent
Openconfirm	We've received an open message from the other router; but the neighbor relationship can still fail (perhaps parameter testing still needed?)

The BGP router identifier [line 6] is chosen just like with other protocols, in order of preference:

- Explicit Configuration
EDGE(config)# **router bgp 130092**
EDGE(config-router)# **bgp router-id 2.2.2.2**
Note that BGP is a little more picky than OSPF or EIGRP; the router-id must be a valid IP address. For example, you can't use 0.0.0.2, like I did with the other protocols.
- Highest IPv4 address among the "up" loopback interfaces
- Highest IPv4 address among the "up" real interfaces

Neighbor Shut Down (Admin Disable)

```

1 EDGE(config)# router bgp 130092
2 EDGE(config-router)# neighbor 198.51.100.1 shutdown
3                               To revive the relationship, use "no neighbor 198.51.100.1 shutdown"
4 EDGE(config-router)# do show ip bgp summary
5 BGP router identifier 2.2.2.2, local AS number 130092
6 BGP table version is 1, main routing table version 1
7
8 Neighbor      V      AS  MsgRcvd  MsgSent  TblVer  InQ  OutQ  Up/Down  State/PfxRcd
9 198.51.100.1  4      64496  0       0      1     0    0 00:00:42  Idle (Admin)

```

Unlike the IGPs on the CCNA exam, BGP's network command explicitly adds a network to its table of routes so that it will be advertised. This is the only way to add a route to the table that is covered.

Injecting and Advertising a Classful Net

If your enterprise has been given all of a classful network, you would want it advertised to the ISP so they can advertise to the rest of the world that packets to those addresses should be sent to you. You would never do this with any private addresses that you are using behind NAT because those addresses belong to everyone, not just you.

First, make sure the address range you want to inject and advertise is in your local routing table. If it's not, BGP won't advertise it.

```

1 EDGE(config)# interface gi0/0
2 EDGE(config-if)# description World-Routable DMZ
3 EDGE(config-if)# ip address 203.0.113.1 255.255.255.0
4 EDGE(config-if)# no shutdown
5 EDGE(config-if)# do show ip route
6
7 Legend omitted
8
9 198.51.100.0/24 is variably subnetted, 2 subnets, 2 masks
10 C 198.51.100.0/24 is directly connected, Serial0/2/0
11 L 198.51.100.95/32 is directly connected, Serial0/2/0
12 C 203.0.113.0/24 is directly connected, GigabitEthernet0/0
13 L 203.0.113.1/32 is directly connected, GigabitEthernet0/0

```

Next, use the network command to add the prefix and length (address & mask) to the BGP table. BGP will check the routing table, and if it finds an exact match (prefix and length), it will add that network to its own table.

```

1 EDGE(config)# router bgp 130092
2 EDGE(config-router)# network 203.0.113.0 mask 255.255.255.0
3
4 EDGE(config-router)# do show ip bgp
5 BGP table version is 2, local router ID is 2.2.2.2
6 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
7 r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
8 f RT-Filter
9 Origin codes: i - IGP, e - EGP, ? - incomplete
10
11 Network Next Hop Metric LocPrf Weight Path
12 *> 203.0.113.0 0.0.0.0 0 32768 i
13
14 Next hop is 0.0.0.0 because the net is directly connected. Otherwise, it would be the next-hop address within our network, just like with EIGRP, OSPF, or RIP

```

Now, BGP will advertise the best route for each network in its table, and we can see our net appear in the ISP's BGP table and routing table, ready for further advertisement to the rest of the world.

Notice that the BGP path is our enterprise's ASN. Think of it as a next-hop ASN, representing the big picture of a packet bouncing from one ISP to the next with little concern over individual routers within those ASNs. There was no path ASN in the edge router's BGP table because the net was in the local ASN.

Also, the next hop IP address is set to the address of the eBGP neighbor, by default.

In the routing table, the administrative distance is 20 (for eBGP). iBGP uses an AD of 200.

```

1 ISP# show ip bgp
2 BGP table version is 2, local router ID is 198.51.100.1
3 Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
4               r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
5 f RT-Filter
6 Origin codes: i - IGP, e - EGP, ? - incomplete
7
8   Network          Next Hop          Metric LocPrf Weight Path
9 *> 203.0.113.0     198.51.100.95    0         0 130092 i
10
    The * and > in the left column mean that the route is valid and best (see legend above)
ISP# show ip route
    Lots omitted
    198.51.100.0/24 is variably subnetted, 2 subnets, 2 masks
C    198.51.100.0/24 is directly connected, Serial0/2/0
L    198.51.100.1/32 is directly connected, Serial0/2/0
B    203.0.113.0/24 [20/0] via 198.51.100.95, 00:04:42

```

Advertising a Subnet

Advertising a subnet is no different; just use the correct mask in the network statement. Best practice is to give your ISP a single route to your entire address space. Your ISP really doesn't care how you subdivided it or how subnets are routed after the ISP hands off your traffic to you.

A Static Discard Route

So far, we know

- Your ISP really wants a single route to your entire address space, no matter how much subnetting you do
- BGP checks your routing table for a route to a network before advertising it
- If you do subnet, you may not have a route in your edge router's routing table for the entire address space. If that happens, BGP won't advertise it to the ISP

For example, we control 203.0.113.0 /24 and have subdivided it.

```

1 EDGE# show ip route
2
    Lots omitted
3    203.0.113.0/24 is variably subnetted, 6 subnets, 4 masks
4 C    203.0.113.0/25 is directly connected, GigabitEthernet0/0.10
5 L    203.0.113.1/32 is directly connected, GigabitEthernet0/0.10
6 C    203.0.113.128/26 is directly connected, GigabitEthernet0/0.20
7 L    203.0.113.129/32 is directly connected, GigabitEthernet0/0.20
8 C    203.0.113.192/28 is directly connected, GigabitEthernet0/0.30
9 L    203.0.113.193/32 is directly connected, GigabitEthernet0/0.30

```

Since the only match of both the address 203.0.113.0 and the mask of /24 is a heading and not a route, BGP will refuse to add it to its table and advertise it.

```

1 EDGE(config)# router bgp 130092
2 EDGE(config-router)# network 203.0.113.0 mask 255.255.255.0
3 EDGE(config-router)# do show ip bgp
4 EDGE(config-router)#
5
    No response means that the BGP table is empty

```

But you can create a phony "static discard route" matching the entire address space. It's just a static route with a null destination. After you do that, you'll see your address space in the BGP table.

```

1 EDGE(config)# ip route 203.0.113.0 255.255.255.0 null0
2 EDGE(config)# end
3 EDGE# show ip route
4
5           Lots Omitted
6           203.0.113.0/24 is variably subnetted, 7 subnets, 5 masks
7 S       203.0.113.0/24 is directly connected, Null0
8 C       203.0.113.0/25 is directly connected, GigabitEthernet0/0.10
9 L       203.0.113.1/32 is directly connected, GigabitEthernet0/0.10
10 C      203.0.113.128/26 is directly connected, GigabitEthernet0/0.20
11 L      203.0.113.129/32 is directly connected, GigabitEthernet0/0.20
12 C      203.0.113.192/28 is directly connected, GigabitEthernet0/0.30
13 L      203.0.113.193/32 is directly connected, GigabitEthernet0/0.30
EDGE# show ip bgp
BGP table version is 2, local router ID is 203.0.113.193
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
f RT-Filter
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop           Metric LocPrf Weight Path
* > 203.0.113.0     0.0.0.0             0         32768 i

```

Thanks to the way routing works, static discard routes are harmless.

- Routing always uses the longest netmask match (the most precisely matching route)
- Packets destined for a subnet that lacks a route are discarded
- The static discard route acts as a "none of the above" catch-all route for the classful network and explicitly discards packets to routeless (unused) subnets that would have died anyway.

LEARNING A DEFAULT ROUTE FROM THE ISP

Your ISP has almost certainly configured BGP on their end to tell your router to set the ISP as your default-gateway. Once you have BGP working, this would automatically show up as a 0.0.0.0 route in BGP and a default gateway setting in your routing table.

```

1 EDGE# show ip bgp
2
3           Stuff Omitted
4   Network          Next Hop           Metric LocPrf Weight Path
5 * > 0.0.0.0         198.51.100.1       0         0 64496 i
6 * > 203.0.113.0     0.0.0.0             0         32768 i
EDGE# show ip route
7
8           Legend Omitted
9           Gateway of last resort is 198.51.100.1 to network 0.0.0.0
10 B* 0.0.0.0/0 [20/0] via 198.51.100.1, 00:04:45
11           The B means the route came from BGP, the star means that it's a default route candidate

```