Border Gateway Protocol

- Routing Protocol used to exchange routing information between networks
  - exterior gateway protocol
- RFC1771
  - work in progress to update
  - `draft-ietf-idr-bgp4-10.txt`
- Currently Version 4
- Runs over TCP

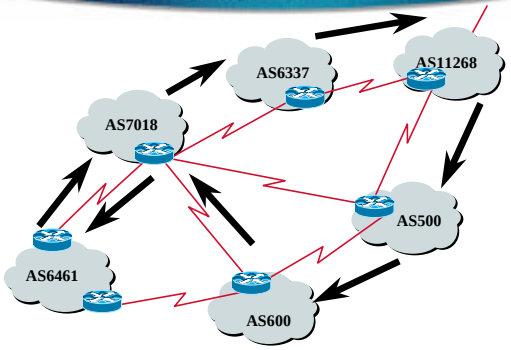
BGP

- Path Vector Protocol
- Incremental Updates
- Many options for policy enforcement
- Classless Inter Domain Routing (CIDR)
- Widely used for Internet backbone
- Autonomous systems

Path Vector Protocol

- BGP is classified as a ***path vector*** routing protocol (see RFC 1322)
  - A path vector protocol defines a route as a pairing between a destination and the attributes of the path to that destination.

Path Vector Protocol



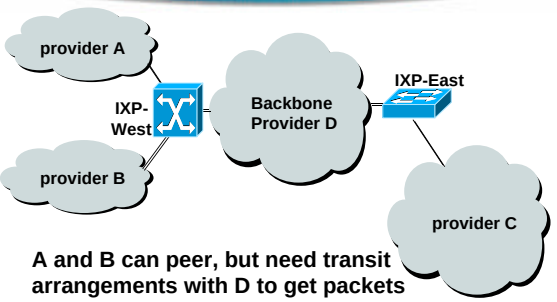
Definitions

- Transit - carrying traffic across a network, usually for a fee
- Peering - exchanging routing information and traffic
- Default - where to send traffic when there is no explicit match is in the routing table

Default Free Zone

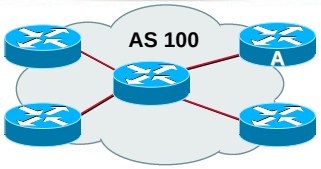
The default free zone is made up of Internet routers which have explicit routing information about the rest of the Internet, and therefore do not need to use a default route.

Peering and Transit example



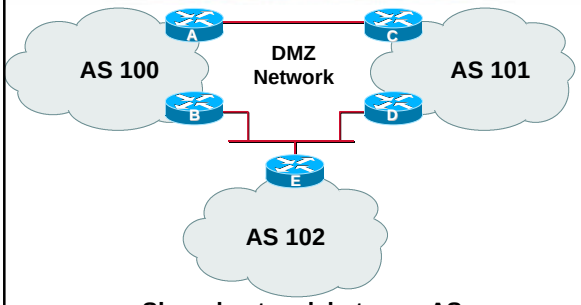
A and B can peer, but need transit arrangements with D to get packets to/from C

Autonomous System (AS)



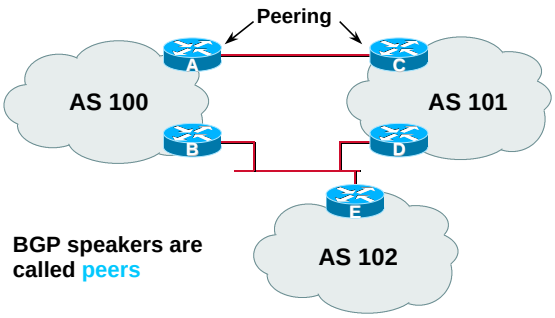
- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control

Demarcation Zone (DMZ)



- Shared network between ASes

BGP Basics



BGP speakers are called peers

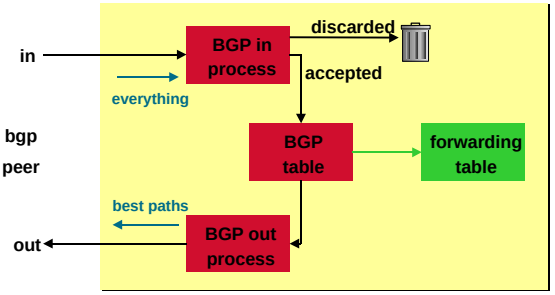
BGP General Operation

- Learns multiple paths via internal and external BGP speakers
- Picks the best path and installs in the forwarding table
- Policies applied by influencing the best path selection

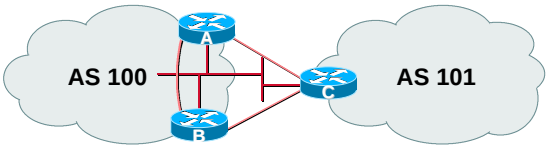
Constructing the Forwarding Table

- BGP "in" process
  - receives path information from peers
  - results of BGP path selection placed in the BGP table
  - "best path" flagged
- BGP "out" process
  - announces "best path" information to peers
- Best paths installed in forwarding table if:
  - prefix and prefix length are unique
  - lowest "protocol distance"

Constructing the Forwarding Table



External BGP Peering (eBGP)



- Between BGP speakers in different AS
- Should be directly connected
- Do not run an IGP between eBGP peers

Configuring External BGP

Router A in AS100

```
interface ethernet 5/0
ip address 222.222.10.2 255.255.255.240
router bgp 100
 network 220.220.8.0 mask 255.255.252.0
 neighbor 222.222.10.1 remote-as 101
 neighbor 222.222.10.1 prefix-list RouterC in
 neighbor 222.222.10.1 prefix-list RouterC out
```

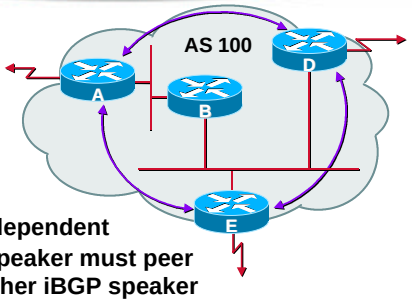
Router C in AS101

```
interface ethernet 1/0/0
ip address 222.222.10.1 255.255.255.240
router bgp 101
 network 220.220.16.0 mask 255.255.240.0
 neighbor 222.222.10.2 remote-as 100
 neighbor 222.222.10.2 prefix-list RouterA in
 neighbor 222.222.10.2 prefix-list RouterA out
```


Internal BGP (iBGP)

- BGP peer within the same AS
- Not required to be directly connected
- iBGP speakers need to be fully meshed
  - they originate connected networks
  - they do not pass on prefixes learned from other iBGP speakers

Internal BGP Peering (iBGP)

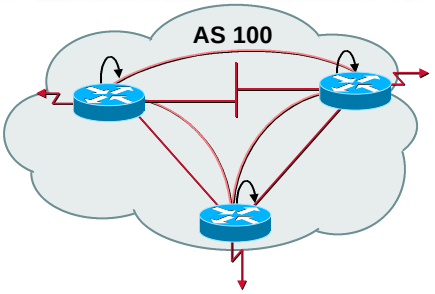


- Topology independent
- Each iBGP speaker must peer with every other iBGP speaker in the AS

Stable iBGP Peering

- Peer with loop-back address
- iBGP session is not dependent on state of a single interface
- iBGP session is not dependent on physical topology
- Loop-back interface does not go down - ever!

Peering to Loop-Back Address



Configuring Internal BGP

Router A

```
interface loopback 0
ip address 215.10.7.1 255.255.255.255
router bgp 100
  network 220.220.1.0
  neighbor 215.10.7.2 remote-as 100
  neighbor 215.10.7.2 update-source loopback0
  neighbor 215.10.7.3 remote-as 100
  neighbor 215.10.7.3 update-source loopback0
```

Router B

```
interface loopback 0
ip address 215.10.7.2 255.255.255.255
router bgp 100
  network 220.220.5.0
  neighbor 215.10.7.1 remote-as 100
  neighbor 215.10.7.1 update-source loopback0
  neighbor 215.10.7.3 remote-as 100
  neighbor 215.10.7.3 update-source loopback0
```


Inserting prefixes into BGP

- Two ways to insert prefixes into BGP
  - redistribute static
  - network command

Inserting prefixes into BGP - redistribute static

- Configuration Example:

```
router bgp 109
 redistribute static
ip route 198.10.4.0 255.255.254.0 serial0
```

- Static route must exist before redistribute command will work
- Forces origin to be “incomplete”
- Care required!

Inserting prefixes into BGP - redistribute static

- Care required with redistribute!
  - redistribute <routing-protocol> means everything in the <routing-protocol> will be transferred into the current routing protocol
  - Will not scale if uncontrolled
  - Best avoided if at all possible
  - redistribute normally used with “route-maps” and under tight administrative control

Inserting prefixes into BGP - network command

- Configuration Example

```
router bgp 109
 network 198.10.4.0 mask 255.255.254.0
ip route 198.10.0.0 255.255.254.0 serial0
```

- A matching route must exist in the routing table before the network is announced
- Forces origin to be "IGP"

Configuring Aggregation

- Three ways to configure route aggregation
  - redistribute static
  - aggregate-address
  - network command

Configuring Aggregation

- Configuration Example:

```
router bgp 109
 redistribute static
ip route 198.10.0.0 255.255.0.0 null0 250
```

- static route to "null0" is called a pull up route
  - packets only sent here if there is no more specific match in the routing table
  - distance of 250 ensures this is last resort static
  - care required - see previously!

Configuration Aggregation - Network Command

- Configuration Example

```
router bgp 109
 network 198.10.0.0 mask 255.255.0.0
ip route 198.10.0.0 255.255.0.0 null0 250
```

- A matching route must exist in the routing table before the network is announced
- Easiest and best way of generating an aggregate

Configuring Aggregation - aggregate-address command

- Configuration Example

```
router bgp 109
 network 198.10.4.0 mask 255.255.252.0
 aggregate-address 198.10.0.0 255.255.0.0 [ summary-only ]
```

- Requires more specific prefix in routing table before aggregate is announced
- {summary-only} keyword
  - optional keyword which ensures that only the summary is announced if a more specific prefix exists in the routing table

Auto Summarisation

- Automatically summarises subprefixes to the classful network.
  - Example:

```
61.10.8.0/22 --> 61.0.0.0/8
```

- Must be turned off for any Internet connected site using BGP.

```
router bgp 109
 no auto-summary
```

Synchronisation

- **In Cisco IOS, BGP does not advertise a route before all routers in the AS have learned it via an IGP**
- **Disable synchronisation if:**
  - AS doesn't pass traffic from one AS to another, or
  - All transit routers in AS run BGP, or
  - iBGP is used across backbone

```
router bgp 109
 no synchronization
```


Summary

- **BGP4 - distance vector protocol**
- **iBGP versus eBGP**
- **stable iBGP - peer with loopbacks**
- **announcing prefixes & aggregates**
- **no synchronization & no auto-summary**