Border Gateway Protocol (BGP)

- BGPv4 is the standard Inter-AS routing protocol for the Internet
- RFC 1771 (+ see 1772, 1773)

- A BGP route is based on reachability, not (necessarily) cost
BGP protocol

- Based on DV concepts
- A “Path Vector” protocol
  - Routers advertise complete routes to destination AS’s, not costs (normally)
  - Being able to control which AS your route goes through is very important to the government
    (ensure traffic from Vandenberg to the White House is not routed through another country)

Path vector problems?

- Resolves the Count-to-Infinity problem with Split Horizon algorithm (a router receiving a path vector can check to see if it is on the path)
Terminology

- **AS 1**: RIP
- **AS 2**: OSPF
- **BGP Peers**: Between AS 1 and AS 2
- **“Stub” AS**: IP packet that is sent from or to this AS: “local” traffic
- **“Transit” AS**: Any others: “transit” traffic

Speakers and gateways

- Admin of each AS declares one or more BGP routers to be “BGP speaker:
  - Establish BGP sessions to other AS’s
  - Determine routes to advertise
- “Gateways” are BGP routers through which traffic enters and leaves the AS
  - Speaker is not necessarily a gateway
Routing

- BGP routers exchange route info
  - Inter-AS – exchange done directly
  - Intra-AS – exchange via internal AS paths

- Note that BGP routes to networks, not individual destinations or routers

Route Advertisement

- Sent between BGP routers
- Only when something changes
- Consists of:
  - Network address in CIDR format (e.g. 127.04.114/24)
  - Attributes (we’ll talk about those later)
BGP route advertising

- Administrator is free to set own policies for advertising routes
  - For example:
    - Particular AS’s preferred
    - Particular AS’s avoided

Processing route advertisements

- Advert treated as a “contract” from the peer router to forward traffic to the ntwk
- BGP router can ignore adverts:
  - If own AS number is in the path
  - If forbidden AS is in the path
  - etc.
Choosing between paths

- A BGP router may receive several paths to a single destination, but uses only one
- Major selection criteria:
  - Preferences
    - set locally by administrator (metric, filters)
    - suggested by neighboring routers (local_pref, MED)
  - Minimum number of AS’s crossed
  - Route Filters
    - Prefix List, AS-Path, Community
- Routes not selected are saved for backup

Observations

- Other AS’s have to trust that a speaker will advertise the best route (and “best” may mean something different to each)
- A speaker does not have to advertise a route, even if it knows one
  - Can refuse to provide transit
  - Can try to block access to other AS’s
How are BGP messages transmitted?

- TCP
- Port 179

BGP message types

- OPEN
  - Establishes link to a BGP peer, Authenticates
- UPDATE
  - Contains a route being advertised (no more than one)
  - Optionally, provides the address of a network for which the sender wants to withdraw a previously-advertised route (multiple)
- KEEPALIVE
  - ACK or "I’m still alive" (sent to neighbors on a regular basis: typ 30sec)
- NOTIFICATION
  - Error or other BGP control message
Some Attributes (in UPDATE messages)

- ORIGIN: Source of the route information
- AS-PATH: List of AS’s to traverse on the way to the network
- NEXT_HOP: IP addr of router to start AS_PATH
- UNREACHABLE: Previously-advertised route has become unreachable

BGP variants

- eBGP
  - The common form of BGP. What we’ve been talking about.

- iBGP
  - A modified “internal use” variation of BGP.
  - Used among edge routers of the same AS to synchronize external routes.
BGP AS-Path Padding

- Append the local AS multiple times upon advertisement
- “Scruffy” fix to emulate Distance-Vector
- Big corporations (WorldCom, Sprint, etc) use this in their Internet routing tables for ISPs

Ex:

Original AS Path = 100 300 500
..with Padding = 100 100 100 300 500

AS Path 100 now has a distance of 3.