An Introduction to BGP

Andrew Parnell – andrew@parnell.ca

Some basics

- BGP – Border Gateway Protocol
- Current Version is BGP4 – RFC4271
- Previous versions do not support CIDR
- Uses TCP 179
- Exchanges routing information (prefixes) between different networks (Autonomous Systems)
- Not concerned with link state or internal topology, this is what the IGP is for – eg OSPF, IS-IS, OLSR, …
- Neighbours must be configured manually

Some more basics…

- Primarily used for multi-homing between different networks – eg the internet
- If not multi-homed, BGP offers little value – why not just use a default route?
- Offers very flexible policies for route selection:
  - perhaps you have two upstream providers, prefer the one who charges you the least
- Very limited *inbound* routing control – almost useless!
  - There are ways to influence inbound traffic (eg prepending, deaggregation)
  - but ultimately you have no control over another AS's routing policy

Autonomous Systems

- All BGP speaking routers belong to only one AS
- Identified by its AS Number
- Most ASNs on the internet are a 2 byte value
  - eg 701, 49835, 4
- Assigned by RIRs (eg RIPE, ARIN) similarly as IP addreses
- Recently extended to 4-byte ASNs due to exhaustion
  - Actually backwards compatible!

Prefixes

- Routers exchange prefixes with each other in order to build the network topology
- A prefix is a network range, eg 109.69.8.0/23, 2001:470::/32
- Currently ~335000 IPv4 prefixes, ~4500 IPv6 prefixes on the internet, announced by ~37000 ASNs

AS Paths

- As prefixes are propagated, each AS appends its own ASN to form an AS path
- Example – guifi.net announces their IPv4 prefix 109.69.8.0/23 from AS49835
- Their upstream providers add this route into their tables, propagates to their peers
- Example AS path: 3356 174 49835 i
  - AS3356 leans the prefix from AS174, who has learned the prefix from AS49835

A simple internet

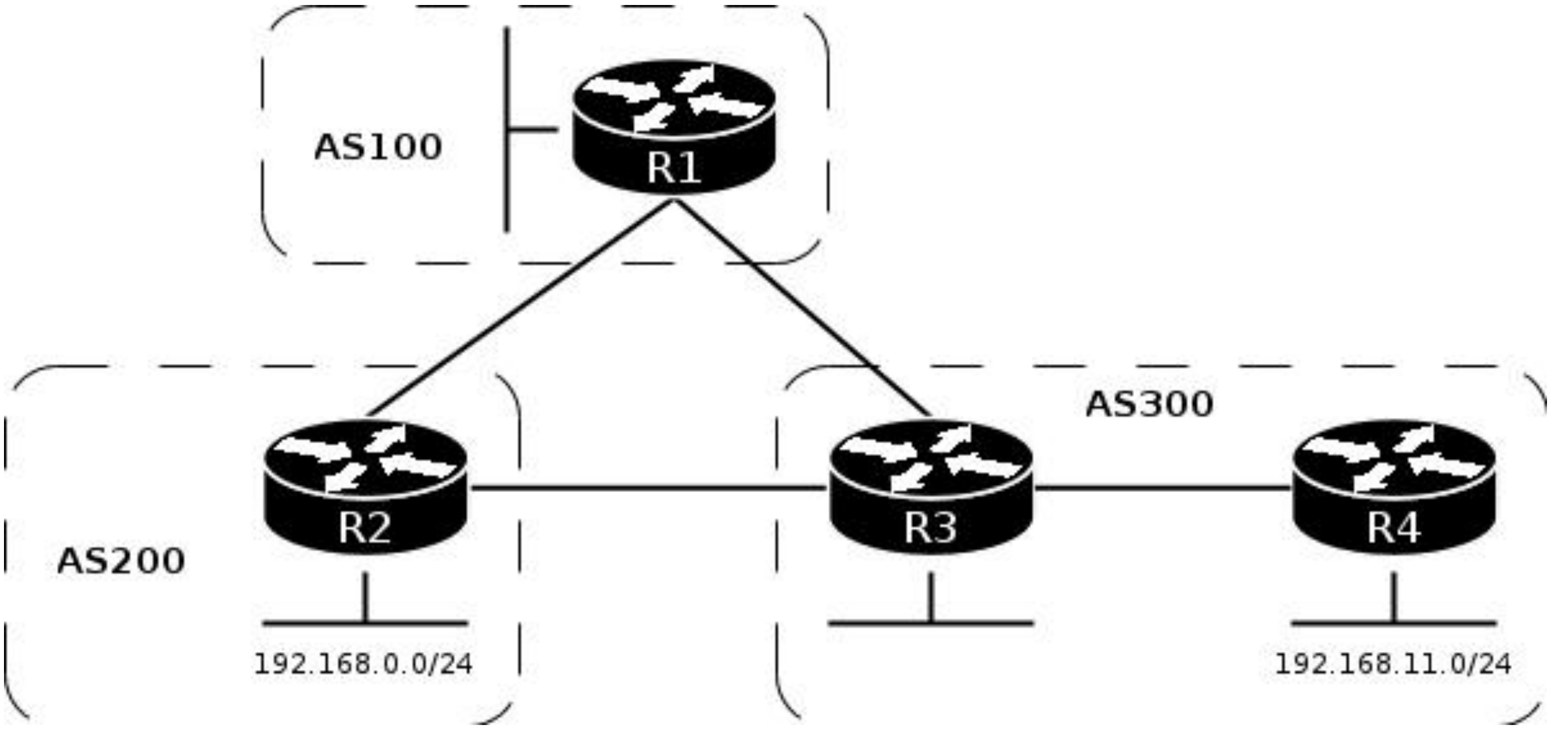

Image source: wiki.mikrotik.com

Hijacking - A problem!

- BGP is trust-based
- Anybody may announce any prefix, regardless of whether it is theirs
- Some noteworthy examples of this:
  - Pakistan wishes to censor youtube, inadvertently propagates a false route to the internet. Youtube becomes mostly unreachable globally (2008)
  - China Telecom originated 37,000 prefixes not belonging to them in 15 minutes, causing massive outage of services globally (2010)
- Routing Registries intend to solve this, however not yet widely implemented

Further Reading

- BGP information is not a secret, as such there are many tools to see what is going on
  - http://bgp.he.net
  - http://robtex.com
- Many networks operate a Looking Glass to view their perspective of the internet
  - telnet://route-server.east.allstream.com
  - http://www.lookinglass.org/
-

The End.

- Lots of questions?
- Hungry? :)