COMP202-08B
Computer Communications

Lecture 10

IP Routing #2

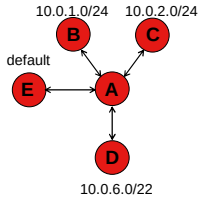
Intro to BGP

THE UNIVERSITY OF WAIKATO

So far

- Internet Protocol v4 (IPv4)
  - 32 bit IP addresses
  - Structured addressing scheme, Internet-wide addressing
- Prefix of the address corresponds to the network where the machine is to be found
  - 130.217.250.39/16 is to be found in the 130.217.0.0/16 network
  - Important scaling property
    - One entry in the routing table for 2^{16} host addresses

IP forwarding

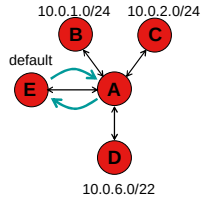


- A has a message to send to 10.0.2.15
- There are two matching routes
  - 10.0.2.0/24, via C
  - Default, via E

A's routing table

Destination	Gateway	Interface
10.0.1.0/24	B	eth0
10.0.2.0/24	C	eth1
10.0.6.0/23	D	eth2
0.0.0.0/0	E	eth3

IP forwarding

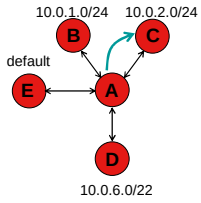


- If it sends it via E
  - E might send it back to A, since E may know to reach C via A.
  - If it does, the message will get stuck in a loop, which is not good.

A's routing table

Destination	Gateway	Interface
10.0.1.0/24	B	eth0
10.0.2.0/24	C	eth1
10.0.6.0/23	D	eth2
0.0.0.0/0	E	eth3

IP forwarding



- Routing tables are ordered from longest network prefix to shortest network prefix
  - /24 is a longer prefix than /22
- The first matching entry in the table is selected
- That is, the most specific destination prefix is chosen

A's routing table

Destination	Gateway	Interface
10.0.1.0/24	B	eth0
10.0.2.0/24	C	eth1
10.0.6.0/23	D	eth2
0.0.0.0/0	E	eth3

IP forwarding

- Order table from most specific to least specific network prefix. Until a match is made:
  - AND the destination address with the subnet mask for the route
  - Compare the result with the network. If they match, return this route
- Example: Select route for 10.0.7.8

Destination	Gateway	Interface
10.0.1.0/24	B	eth0
10.0.2.0/24	C	eth1
10.0.6.0/23	D	eth2
0.0.0.0/0	E	eth3

IP forwarding

- Example: Select route for 10.0.7.8
- Step 1:

	Destination	Gateway	Interface
1.	10.0.1.0/**24**	B	eth0
2.	10.0.2.0/24	C	eth1
3.	10.0.6.0/23	D	eth2
4.	0.0.0.0/0	E	eth3

10.0.1.0/24 00001010 00000000 00000001 00000000

10.0.7.8 00001010 00000000 00000111 00001000
/24 Subnet mask 11111111 11111111 11111111 00000000
00001010 00000000 00000111 00000000

Do not match

IP forwarding

- Example: Select route for 10.0.7.8
- Step 2:

	Destination	Gateway	Interface
1.	10.0.1.0/24	B	eth0
2.	10.0.2.0/**24**	C	eth1
3.	10.0.6.0/23	D	eth2
4.	0.0.0.0/0	E	eth3

10.0.2.0/24 00001010 00000000 00000010 00000000

10.0.7.8 00001010 00000000 00000111 00001000
/24 Subnet mask 11111111 11111111 11111111 00000000
00001010 00000000 00000111 00000000

No Match.

IP forwarding

- Example: Select route for 10.0.7.8
- Step 3:

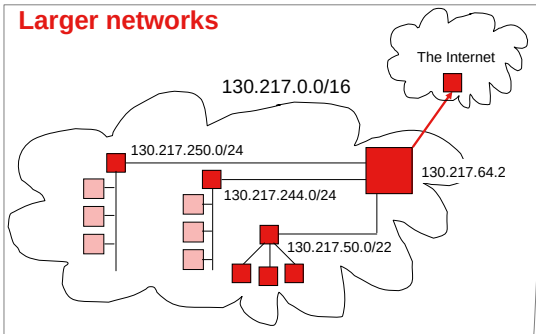
	Destination	Gateway	Interface
1.	10.0.1.0/24	B	eth0
2.	10.0.2.0/24	C	eth1
3.	10.0.6.0/**23**	D	eth2
4.	0.0.0.0/0	E	eth3

10.0.6.0/23 00001010 00000000 00000110 00000000

10.0.7.8 00001010 00000000 00000111 00001000
/23 Subnet mask 11111111 11111111 11111110 00000000
00001010 00000000 00000110 00000000

Match!

Larger networks

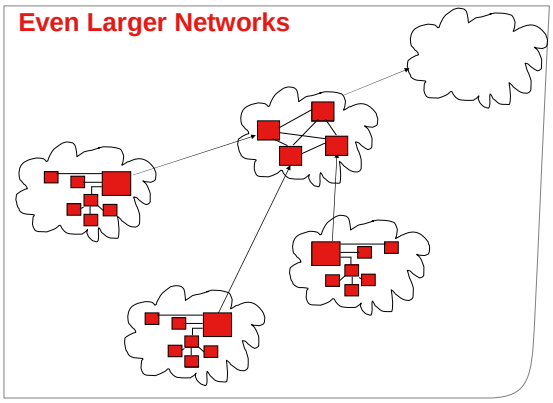


Smaller networks can be aggregated to a single larger network. Useful scaling property, as the upstream router doesn't need to know the complex details of the downstream network.

Larger Networks

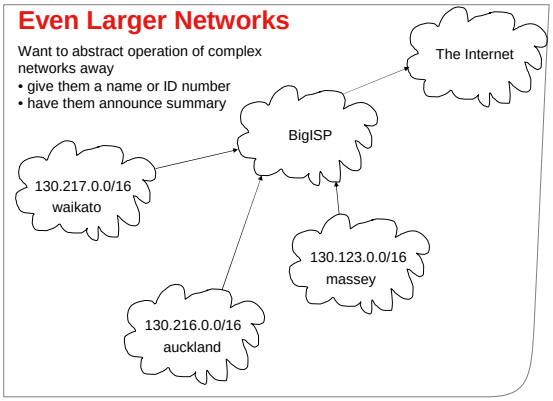
- Still possible with this example to manage routing manually, using, e.g.
  - route command (UNIX systems)
  - route add –net 130.216.244.0/20 130.217.64.2
    - Add a route for 130.216.244.0/20 via 130.217.64.2
- Manual (static) routing becomes more difficult as networks get larger
  - Old networks go away
  - New networks are added
  - Impossible for the Internet

Even Larger Networks



Even Larger Networks

Want to abstract operation of complex networks away
- give them a name or ID number
- have them announce summary



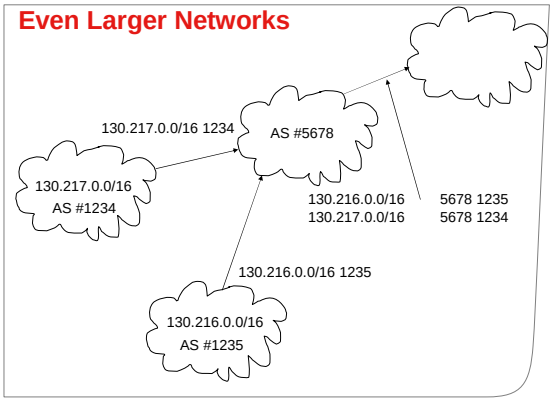
BGP

- Boarder Gateway Protocol (BGP)
- Protocol used to organise Internet-wide routing
- Internet is organised into “Autonomous Systems”
  - ASes
  - Each AS is assigned a number
  - Number space used to 16 bit (1-65535) but now 32 bit
- Connections between ASes are manually specified
  - Peering
  - Route based on policy
  - Usually involving legal contracts, money.
  - Once established, dynamic routing takes place

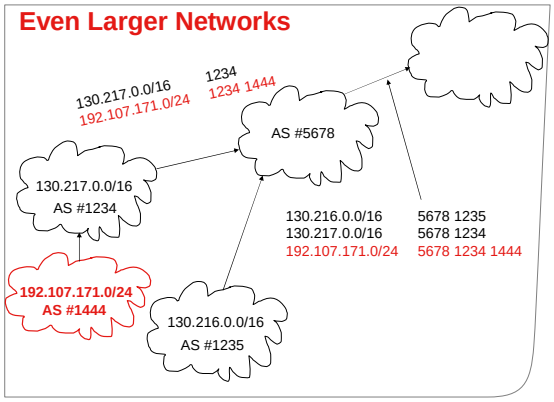
BGP

- Peers announce to each other the network prefixes which can be reached through them
- Peers prepend their AS number to the announcement
  - Create an AS path

Even Larger Networks

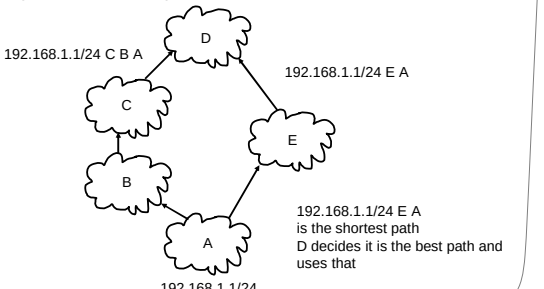


Even Larger Networks



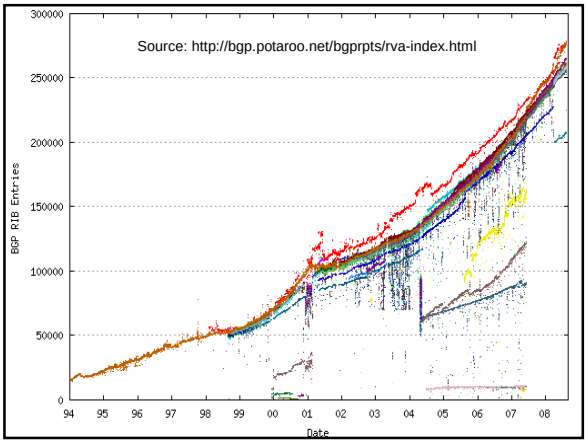
BGP

- Routers may receive an advertisement for the same IP prefix from multiple ASes



In class video: BGP the movie

- History: addresses used to be handed out either in
  - /8 chunks: 16777216 addresses
  - /16 chunks: 65536 addresses
  - /24 chunks: 256 addresses
- Universities such as Waikato are too large for a /24, so asked for (and obtained) /16
  - Inefficient use of address space
  - Currently less than 1/4th of the addresses handed to Waikato are used. (actually less than that, but that's another story)
- Now have the ability to ask for a network prefix of appropriate size



Summary

- IP Routing tables are organised so that the most specific route is chosen to a destination
- Internet routing is organised using the BGP protocol
  - ASes
  - An AS being a network operated by a single organisation