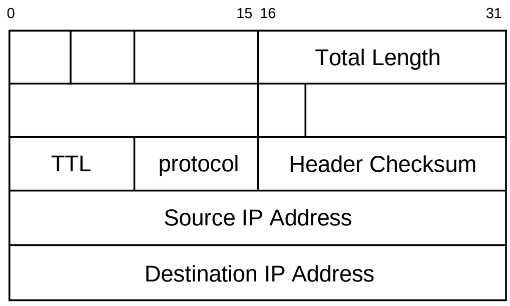
COMP312-09A Communications and Systems Software

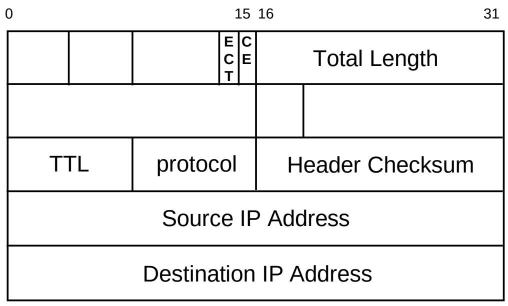
Lecture 10 – IP

Matthew Luckie
mluckie@cs.waikato.ac.nz

IPv4 Header learned in 202

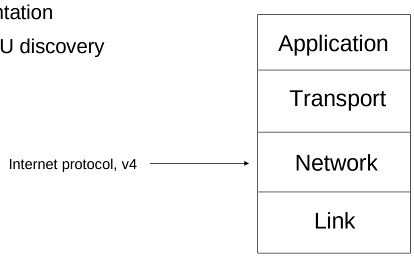


IPv4 Header learned in 312 so far



Overview

- Internet addresses
- Fragmentation
- Path MTU discovery



Internet addresses

- 130.217.250.39

- 32 bit globally unique IPv4 address
- dotted decimal presentation format
- network id prefix, host id suffix

IPv4 Binary

```
   130    .    217    .    250    .     39
1000 0010  1101 1001  1111 1010  0010 0111
```

In the dark ages

- IPv4 addresses fell into one of five classes
  - Class A: 8 bit network id (/8)
    - 24 bit host id. 2^{24} = 16777216 addresses
  - Class B: 16 bit network id (/16)
    - 16 bit host id. 2^{16} = 65536 addresses
  - Class C: 24 bit network id (/24)
    - 8 bit host id. 2^{8} = 256 addresses
  - Class D: Multicast
  - Class E: Future use
- These days, we don't talk about classes very much, except when reflecting on what a poor idea it was.

IP address classes

- Class A addresses begin with 0
  - 0.0.0.0/8 to 127.0.0.0/8
- Class B addresses begin with 10
  - 128.0.0.0/16 to 191.0.0.0/16
- Class C addresses begin with 110
  - 192.0.0.0/24 to 223.0.0.0/24
- Class D addresses begin with 1110
  - 224.0.0.0 to 239.0.0.0
- Class E addresses begin with 1111
  - 240.0.0.0 to 255.0.0.0

IPv4 Binary

```
   130   .   217   .   250   .    39
1000 0010  1101 1001  1111 1010  0010 0111
```

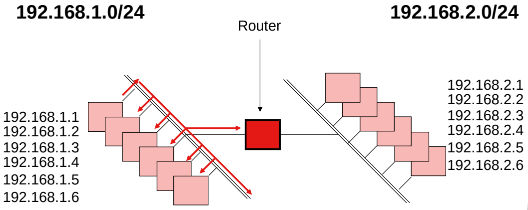
- Class B address, as the IP address begins 10

Special use addresses

- Each IPv4 network has two reserved addresses
  - Network address: host ID portion all zeros
  - Broadcast address: host ID portion all ones
- Network address: used to identify "this network" in routing tables
- Broadcast address: used to address all hosts on this IPv4 subnet.
  - E.g. windows networking broadcasting the existence of a windows computer on the network
  - Broadcast traffic is not forwarded outside of the subnet

Broadcast

192.168.1.0/24 **192.168.2.0/24**



192.168.1.1 broadcasts to 192.168.1.255
The router(s) on this subnet will receive the broadcast frame, but not forward it

Special use addresses: 130.217.250.39/16

```
   130   .   217   .   250   .    39
1000 0010  1101 1001  1111 1010  0010 0111

Network address:
   130   .   217   .    0    .    0
1000 0010  1101 1001  0000 0000  0000 0000

Broadcast address:
   130   .   217   .   255   .   255
1000 0010  1101 1001  1111 1111  1111 1111
```

Exercise

- What are the network and broadcast addresses for
  - 192.168.2.10 / 22
  - 192.168.254.238 / 28

Special use addresses: 192.168.2.10/22

```
   192    .    168    .      2    .     10
1100 0000   1010 1000   0000 0010   0000 1010
```


Special use addresses: 192.168.2.10/22

```
   192    .    168    .      2    .     10
1100 0000   1010 1000   0000 0010   0000 1010

Network address
   192    .    168    .      0    .     0
1100 0000   1010 1000   0000 0000   0000 0000

Broadcast address
   192    .    168    .      3    .    255
1100 0000   1010 1000   0000 0011   1111 1111
```


Special use addresses: 192.168.254.238/28

```
   192    .    168    .    254    .    238
1100 0000   1010 1000   1111 1110   1110 1110
```


Special use addresses: 192.168.254.238/28

```
   192    .    168    .    254    .    238
1100 0000   1010 1000   1111 1110   1110 1110

Network address
   192    .    168    .    254    .    224
1100 0000   1010 1000   1111 1110   1110 0000

Broadcast address
   192    .    168    .    254    .    239
1100 0000   1010 1000   1111 1110   1110 1111
```


Usable addresses

- A usable address is one in a subnet that can be used to identify a host.
- Different network prefix lengths have different numbers of usable addresses.
- In most scenarios, the smallest usable prefix is a /30
  - 2 bits host-id suffix gives 4 addresses
  - Two of the four are used for broadcast and network addresses
  - Two of the four are used to number hosts in the network
- $2^{(32\text{-networklen})} - 2$

0 15 16 31

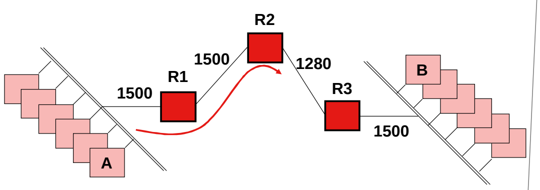
Ver	HLEN	DSCP	ECT	CE	Total Length
TTL		protocol			Header Checksum
Source IP Address					
Destination IP Address					

Ver: specifies IP version, in IPv4 this is always 4.
HLEN: specifies IP header length in 32-bit words; is at least 5, can be more if IP options are included
DSCP: Differentiated Services Code Point; used to provide Quality of Service

IP fragmentation

- IP networks can be made up of many different link types
  - Ethernet
  - X.25
  - Tunnels and VPNs
- Each of these link types have a different maximum transmission units (MTUs)
  - i.e the maximum sized packet which they can transmit

IP fragmentation



Host A sends a 1500 byte IP packet to host B. Router R1 forwards the packet to R2, but R2 cannot forward it to R3 as the packet is.

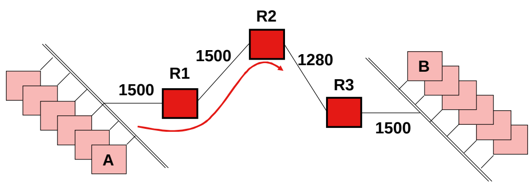
IP fragmentation

0 15 16 31

Ver	HLEN	DSCP	ECT	CE			Total Length
Identifier					DF	MF	Fragment offset
TTL		protocol					Header Checksum
Source IP Address							
Destination IP Address							

- The second 32 bit word in the IPv4 header is used for fragmentation
- 16 bit identifier
  - Assigned by the sender, identifies a single unique packet in a flow
  - Usually implemented as a simple counter, but could be random
- DF bit: Don't Fragment
  - A router should not fragment this packet
- MF bit: More Fragments
  - A router which fragments a packet sets this bit on all fragments but the last
- Fragment offset (13 bits)
  - count of 64-bit words which identify where this fragment goes in relation to other frames.
  - used to reassemble fragments at the end host

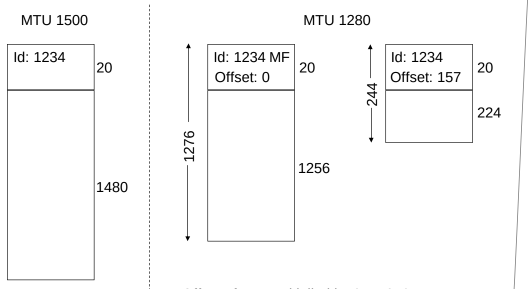
IP fragmentation



R2, provided DF bit is not set, will take a 1500 byte IP packet and break it up.

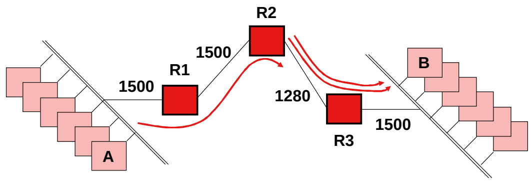
1500 – 20 = 1480 byte of IP data.

IP fragmentation at R2



Offset of 157 multiplied by 8 = 1256
1256 + 224 = 1480

IP fragmentation

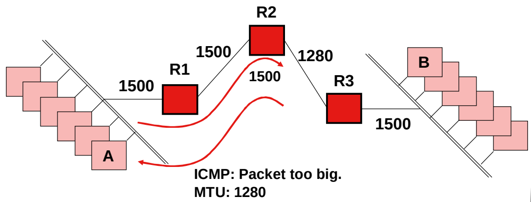


R2 creates two fragments and forwards them to B

IP fragmentation

- Inefficient
- Causes routers to have to break the packet into pieces
  - Much harder to optimise for, compared to simply making a forwarding decision
  - Will incur processing delay at the router
- Causes more IP packets to enter the network
  - If any fragment is lost, sender has to discard all fragments
- Alternative: Path MTU Discovery

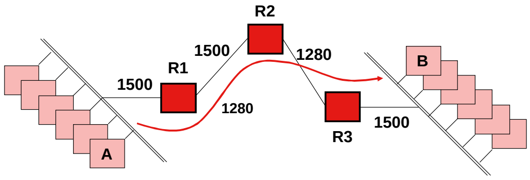
Path MTU Discovery



R2 sees DF bit in IP header is set, does not fragment.
It discards the frame, and sends an ICMP message to A.

The message is a "Packet Too Big" message, and it usually includes the MTU of the next-hop

Path MTU Discovery



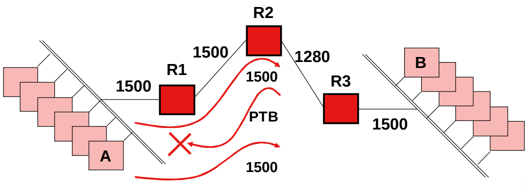
A sends packets up to 1280 bytes in size from then on.

Actually, periodically it will "forget" the 1280 value just to see if it is possible to send larger frames.

Issues with Path MTU Discovery

- Relies on ICMP, which is often blocked by firewalls.
- If you get a job in networking, please do not firewall ICMP packet too big messages. You're hurting the internet if you do.

Path MTU Discovery



Router R1 is configured to discard (firewall) ICMP messages. It discards the “Packet too big” message, too.

A never receives the ICMP message, keeps trying to send 1500 byte packets.

Summary

- IP addresses were historically given out in a class-based method.
- Fragmentation
- Path MTU Discovery