The IP Network Layer

G54ACC

Lecture 2

richard.mortier@nottingham.ac.uk

Contents

- Overview
- ARP
- Fragmentation
- Routing *vs*. Forwarding
- Other issues and IPv6

Contents

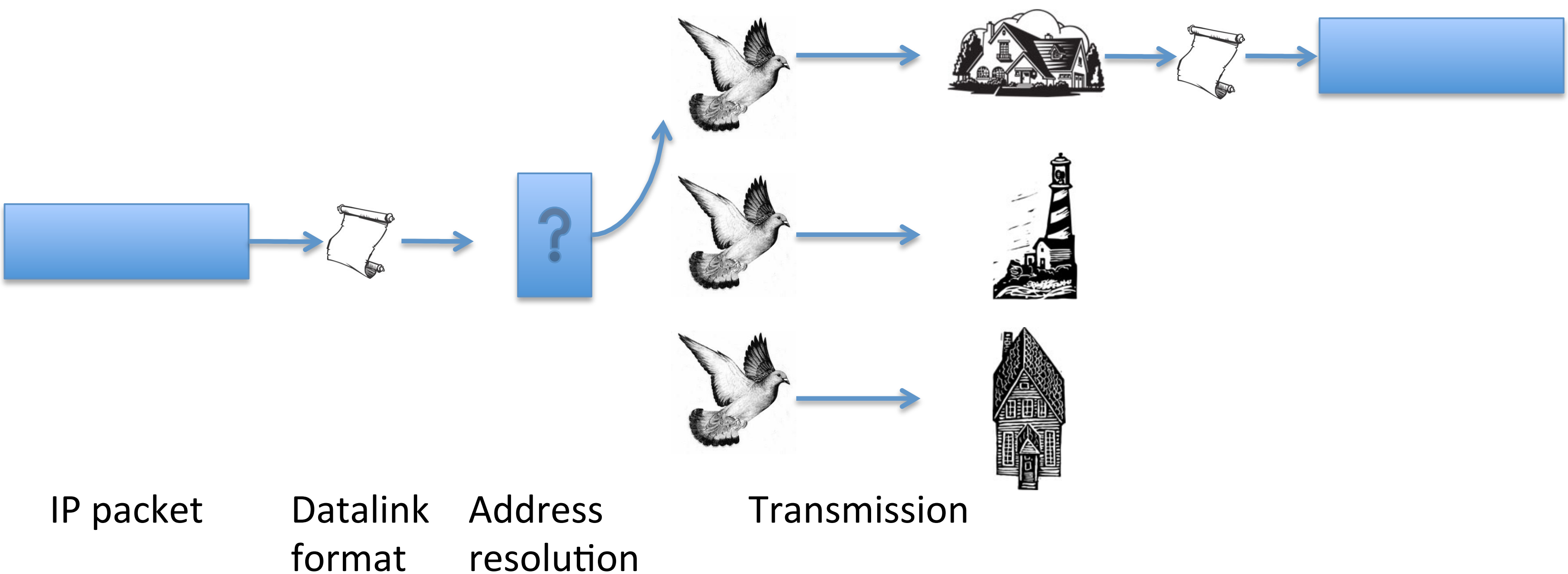
- Overview
  - Encapsulation
  - Mapping IP
  - Ethernet
- ARP
- Fragmentation
- Routing *vs*. Forwarding
- Other issues and IPv6

Overview

- Hosts are allocated addresses
  - In fact, really *interfaces* that are addressed
- Generated data is encapsulated
- Packets passed to datalink for transmission
  - We'll focus on Ethernet, but there are others
- Routers receive packets, examine, forward
  - May have non-IP devices between routers
  - Details of datalink are abstracted

Avian Carriers

- RFC 1149 (Request For Comments)
  - All Internet standards start this way



Encapsulation (reminder)

- Data comes down from above in chunks
  - Sockets API imposes buffers even on streams
  - These chunks may be too large for the datalink
  - Packetization generates suitable sized chunks
- IP header is prepended

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version|  IHL  |Type of Service|          Total Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Identification        |Flags|      Fragment Offset    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Time to Live |    Protocol   |         Header Checksum       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Destination Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           data                              ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RFC 791, Internet Protocol
http://www.rfc-editor.org/rfc/rfc791.txt
IEN 46, Addressing and Naming
http://www.postel.org/ien/txt/ien46.txt

Mapping IP

- Think of common scenarios
  - Laptop – Ethernet & WiFi
  - Smartphone/iPad – WiFi & 3G & Bluetooth
  - ADSL router – Ethernet & ADSL
- Two common problems:
  - Need to map IP address to “subnetwork” address
  - Need to deal with different packet sizes:
    - MTU – “Maximum Transmission Unit”

Ethernet (reminder)

- 48 bit static addresses
  - Traditionally in NIC firmware
- "All ones" is broadcast address
  - `ff:ff:ff:ff:ff:ff`
- *Type* is the next layer protocol
- Ethernet (IEEE 802.3) is one of a family of 802 networks:
  - 802.11 WiFi
  - 802.16 WiMax …

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      62 bits of preamble                    ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...               used for bit synchronization              |SOF|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      48 bits Ethernet                       ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...    Destination Address       |        48 bits Ethernet    ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...                         Source Address                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Length or Type          |         data               ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...             between 46 and 1500 bytes in total              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             32 bits Frame Check Sequence (FCS)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

SOF: start-of-frame delimiter

name	description	note
IEEE 802.1	Bridging (networking) and Network Management	
IEEE 802.2	Logical link control	inactive
IEEE 802.3	Ethernet	
IEEE 802.4	Token bus	disbanded
IEEE 802.5	Defines the MAC layer for a Token Ring	inactive
IEEE 802.6	Metropolitan Area Networks	disbanded
IEEE 802.7	Broadband LAN using Coaxial Cable	disbanded
IEEE 802.8	Fiber Optic TAG	disbanded
IEEE 802.9	Integrated Services LAN	disbanded
IEEE 802.10	Interoperable LAN Security	disbanded
IEEE 802.11 a/b/g/n	Wireless LAN (WLAN) & Mesh (Wi-Fi certification)	
IEEE 802.12	demand priority	disbanded
IEEE 802.13	Used for 100BASE-X Ethernet	
IEEE 802.14	Cable modems	disbanded
IEEE 802.15	Wireless PAN	
IEEE 802.15.1	Bluetooth certification	
IEEE 802.15.2	IEEE 802.15 and IEEE 802.11 coexistence	
IEEE 802.15.3	High-Rate WPAN certification	
IEEE 802.15.4	Low-rate WPAN certification	
IEEE 802.15.5	Mesh networking for WPAN	
IEEE 802.16	Broadband Wireless Access (WiMAX certification)	
IEEE 802.16.1	Local Multipoint Distribution Service	
IEEE 802.17	Resilient packet ring	
IEEE 802.18	Radio Regulatory TAG	
IEEE 802.19	Coexistence TAG	
IEEE 802.20	Mobile Broadband Wireless Access	
IEEE 802.21	Media Independent Handoff	
IEEE 802.22	Wireless Regional Area Network	
IEEE 802.23	Emergency Services Working Group	New (March, 2010)

For completeness…

(not examinable ☺)

Contents

Address Resolution Protocol (ARP)

- Ethernet header
  - Sent to broadcast address
  - From source Ethernet address
  - Type = `0x0806`
- ARP header
  - Htype = `0x01`
  - Ptype = `0x0800`
  - Hlen = 6, Plen = 4
- Operation = 1 (request) broadcasts:
  - Sender Ethernet address
  - Sender IP address
  - Target IP address
- Operation = 2 (reply) from target
  - Insert target's Ethernet address

Internet Protocol (IPv4) over Ethernet ARP packet		
bit offset	**0 – 7**	**8 – 15**
0	Hardware type (HTYPE)	
16	Protocol type (PTYPE)	
32	Hardware address length (HLEN)	Protocol address length (PLEN)
48	Operation (OPER)	
64	Sender hardware address (SHA) (first 16 bits)	
80	(next 16 bits)	
96	(last 16 bits)	
112	Sender protocol address (SPA) (first 16 bits)	
128	(last 16 bits)	
144	Target hardware address (THA) (first 16 bits)	
160	(next 16 bits)	
176	(last 16 bits)	
192	Target protocol address (TPA) (first 16 bits)	
208	(last 16 bits)	

ARP cache

- IP hosts use ARP to find other IP hosts
- Cache entries, e.g.,

```
ppshorizon305:G54ACC drm$ arp –an
? (128.243.35.1) at 0:18:74:1e:bd:40 on en0 ifscope [ethernet]
? (128.243.35.255) at ff:ff:ff:ff:ff:ff on en0 ifscope [ethernet]
```

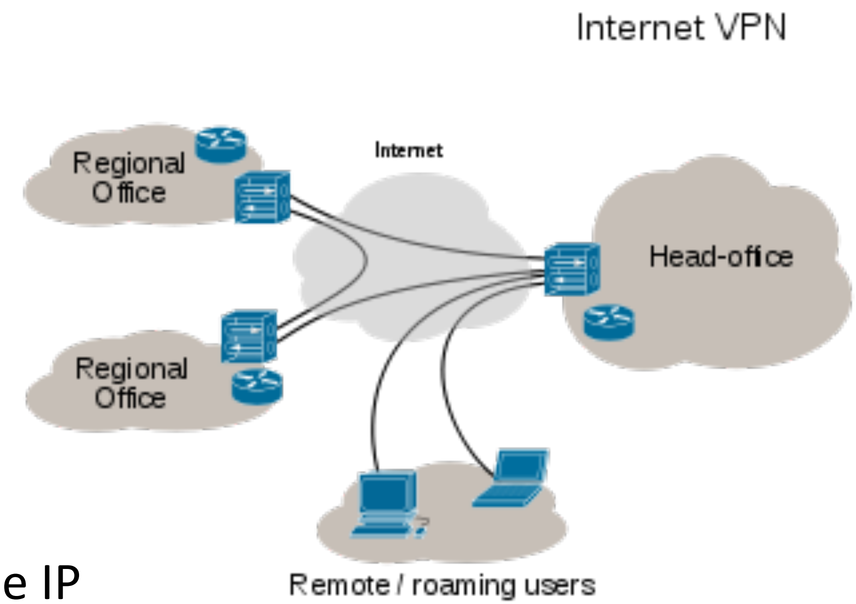
- But be warned, data structure may not be small!

```
marian$ /usr/sbin/arp -a | wc
     563    2810   40483
```

- Perhaps cache with TCP / UDP information
  – see other lectures …
- What's the cache eviction policy?
  – "Soft" – unused entries evicted after some time…

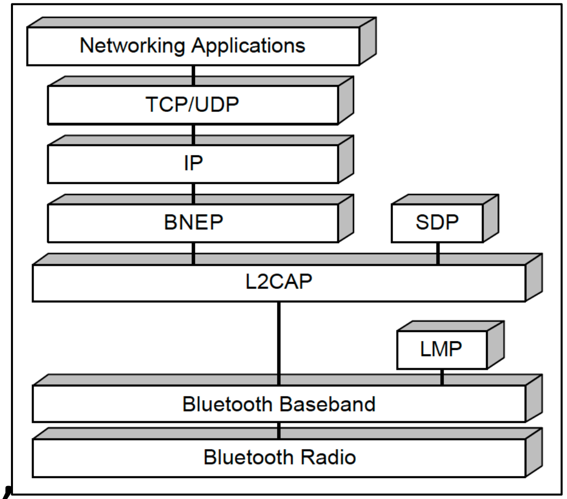
IP as a Subnetwork!

- May want to treat the Internet as a subnetwork
- Why?
  Virtual Private Network



- Examples:
  - IP over IP: RFC 2005 e.g., for Mobile IP
  - PPTP: RFC 2637 – PPP over GRE
  - L2TP: RFC 2661 – PPP inside UDP packets
  - IPSEC: RFC 2401 – Encryption per IP packet
  - SSTP: MSFT – PPP or L2TP over SSL
- Add encryption if we want it private…

...mind if I call you Bruce(*)



- Bluetooth
  - v1.0 was IP/PPP/RFCOM/L2CAP
  - v1.1 changed to BNEP
    - Make Bluetooth look like Ethernet
- ATM
  - Invented LANE - "LAN Emulation"
    - Make ATM look like Ethernet
- Why?
  - Many companies built their protocols directly on Ethernet
  - And it "keeps things simple"

(*) http://orangecow.org/pythonet/sketches/bruces.htm & YouTube

Alternatives

- Packet systems: CLNP, CLNS
  - Actually still used, a little! IS-IS/CLNP/802.3
- Circuit switching
  - A real actual copper circuit (old telephones)
  - Digital telephony implies framing (SONET, &c.)
  - Starting to move onto IP network anyway
- Statistical multiplexing with circuits
  - Virtual circuits (ATM)
  - Flow switching (Ipsilon)
- Compare circuit setup *vs*. Per-packet routing

Contents

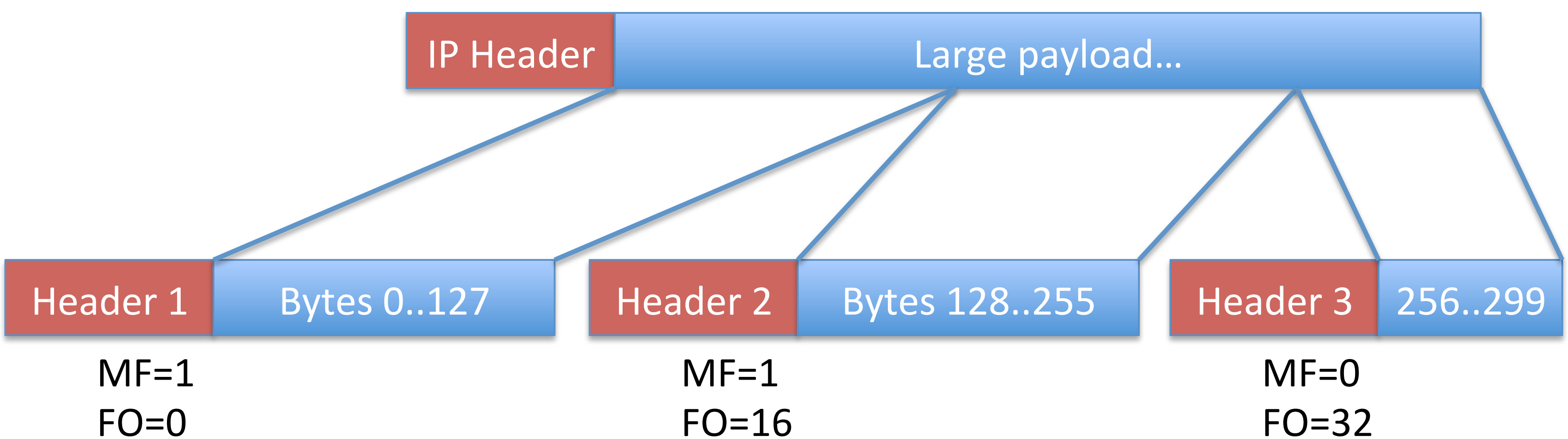
- Overview
- ARP
- Fragmentation
  - Reassembly
  - Loss
- Routing *vs*. Forwarding
- Other issues and IPv6

Fragmentation

- IP allows datagram sizes up to 64 kB
- Physical networks often only support smaller frame types (Maximum Transmission Unit, MTU)
  - e.g., Ethernet 1500 bytes, dialup PPP ~256 bytes
- So, a single IP datagram may need to be divided into “fragments” for transmission…
- See IP header details:
  - MF/DF (More Fragments/Don’t Fragment) bits
  - ID (Identity)
  - FO (Fragment Offset)

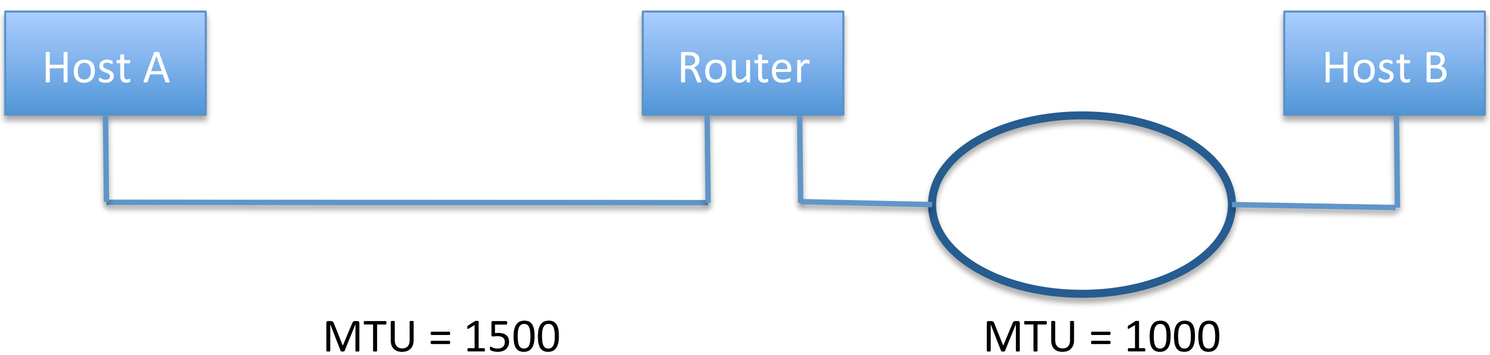
IP Fragmentation

- Each fragment is a (new) IP packet
  - Has full IP header containing the original source & destination
- Identification field is the same for each fragment
- Fragment offset identifies where in original packet
  - NB. Fragment offset in units of 8 octets, i.e., byte offset is (8 * fragment offset)
- “More Fragments” flag set in all but last fragment
- For example, given 300 byte IP data on MTU of 128 bytes



Fragmenting Packets

- May be
  - done by sending host
  - done by intermediate router:
  - prevented with IP “Don’t Fragment” flag
  - avoided by Path MTU discovery (see ICMP…)



Reassembling Fragments

- **ONLY** done by the ultimate destination of the packet
  - ...after checking header checksum, destination, and ID
  - ...but before any other processing
- Receiver must maintain a pool of fragments
  - All discarded after a timeout
- After all fragments of a datagram are received, the datagram is reassembled and processed

Handling Loss

- Lose one fragment and you lose the whole datagram
  - Bad if fragment loss is likely or the number of fragments is large
- E.g., NFS v2 used UDP, v.3 uses TCP
  - The underlying block size increased, 8kB to 32kB
  - This generated many more segments!
  - Thus UDP gave a higher effective packet loss rate, wasting more bandwidth

Contents

- Overview
- ARP
- Fragmentation
- Routing *vs*. Forwarding
  - Other forwarding actions
  - Longest Prefix Matching
- Other issues and IPv6

Routing *vs.* Forwarding

- Router receives an IP packet: what to do?
  - Drop or forward?
  - If forward, which interface to transmit it on?
- Making this decision is **forwarding**
  - Which interface is closest to the destination?
  - IP bases this decision (almost) solely on the destination IP address
  - Also decrement *time-to-live* (TTL) field to ensure looping packets eventually die
- Building up the information to do so is **routing**
  - Where are all the addresses at the moment?

Forwarding

- Need to map packet's destination to interface
  - Group addresses for efficiency
  - Use *prefixes* with explicit *prefix lengths*
  - E.g., 172.16/12; 10/8; 192.168/16; 128.243/16
- Routing generates (prefix, interface) mapping
  - The *forwarding table*
- Map *address* to *prefix* via *longest prefix match*
  - Means the *most specific* entry is used
  - If no match, then use *default route*; else drop
- How to structure forwarding table?
  - Minimise lookup latency
  - Handle variable and substantial update dynamics

Longest Prefix Matching

192	168	10	12	**/32** – Host
1100 0000	1010 1000	0000 1010	0000 1100	
192	168	0	0	**/16**
1100 0000	1010 1000	0000 0000	0000 0000	
192	168	8	0	**/21**
1100 0000	1010 1000	0000 1000	0000 0000	
192	168	10	0	**/23**
1100 0000	1010 1000	0000 1010	0000 0000	
192	168	10	0	**/24**
1100 0000	1010 1000	0000 1010	0000 0000	
192	168	4	0	**/24**
1100 0000	1010 1000	0000 0100	0000 0000	

Contents

- Overview
- ARP
- Fragmentation
- Routing *vs*. Forwarding
- Other issues and IPv6

Other Issues with IPv4

- Multicast
  - Considered *unicast* (1:1) and *broadcast* (1:all)
  - IP natively supports *multicast* (M:N)
  - Need to manage *group membership*, *routing*
- Mobility (RFC3344)
  - Assign *Care-of-Address* in addition to home address
  - *Home-agent* and *foreign-agent* manage nodes
  - What about multicast? (ARP? Broadcast? ...!)
- Everything becomes more complex...

IPv6

- IPv4 is dead! Long live IPv6!
  - Well, sort of: in transition since 1998
- Primary benefit: larger address space (128b)
  - But what about NAT
- Other benefits:
  - Integrated security, back-ported as IPSec
  - Various auto-configuration mechanisms
  - Mobility support: avoid *triangular routing*
  - Jumbograms, extended options space, ...
- Costs: all the other protocols need to change too!
  - *Network layer is **much more** than just an encapsulation*

Contents

- Overview
  - Encapsulation
  - Mapping IP
  - Ethernet
- ARP
  - ARP caching
  - Alternatives
- Fragmentation
  - Reassembly
  - Loss
- Routing *vs*. Forwarding
  - Other forwarding actions
  - Longest Prefix Matching
- Other issues and IPv6

Summary

- IP interconnects networks, enabling adoption of new technologies
  - Encapsulates and packetizes data from higher layers
- Two key issues at the interface
  - Mapping addresses
  - Dealing with MTUs
- Packets forwarded through the network toward specified destination by routers
  - Building up information to forward is *routing*
- Numerous control protocols, processes and organizations required to provide IP infrastructure

Quiz (1)

1. What function does ARP perform? Why?
2. ARP uses subnet broadcast to do what it does. Suggest any alternative mechanism.
3. Two hosts, A (IP address 192.168.0.1, Ethernet address aa:aa:aa:aa:aa:aa) and B (IP address 192.168.0.2, Ethernet address bb:bb:bb:bb:bb:bb) are connected to an Ethernet but have never previously sent or received anything. Write down what has to happen for A to send an IP packet to B.
4. In what situations is fragmentation required?
5. Name two IP header fields that will be altered when an IP router forwards an IP packet, and say why.
6. The two hosts above, A and B, now connect via a gateway (an IP router). What happens in the same situation now?

Quiz (2)

7. List three different IP options, and say what they do.
8. Why does the IP header contain a Protocol field, and how is it used?
9. The IP header has the source address first, then the destination address; but the Ethernet header has the destination address first, then the source address. Say why both source and destination addresses are there in each case, and also why the Ethernet header puts the destination first.
10. Assuming losses are independent, explain the example on sl.21 concerning the higher effective loss rate with NFS v3 vs. NFS v2.