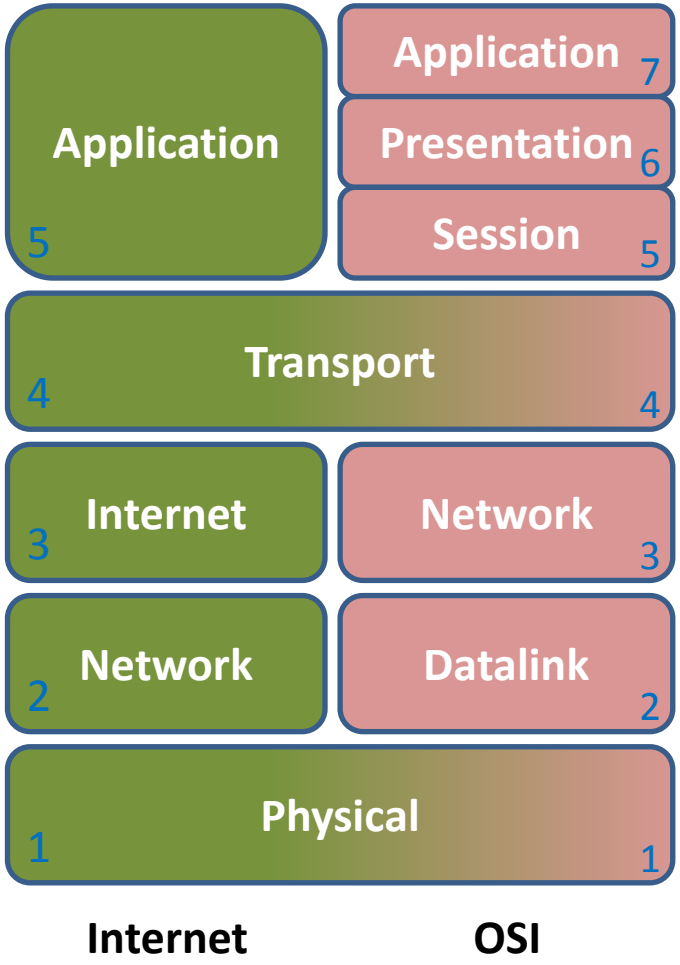
Internetworking

G54ACC – IP and Up

Lecture 1

Recap

- IP sits at layer 3
  - OSI *network layer*
  - TCP/IP *Internet layer*
- Approximates a *fully-connected network* abstraction
- IP packets are *hop-by-hop* routed by *store-and-forward* devices



Contents

- Overview
- Addressing
- Routing *vs*. Forwarding
- Address Management
- Control
- Evolution
- Alternatives & Summary

Contents

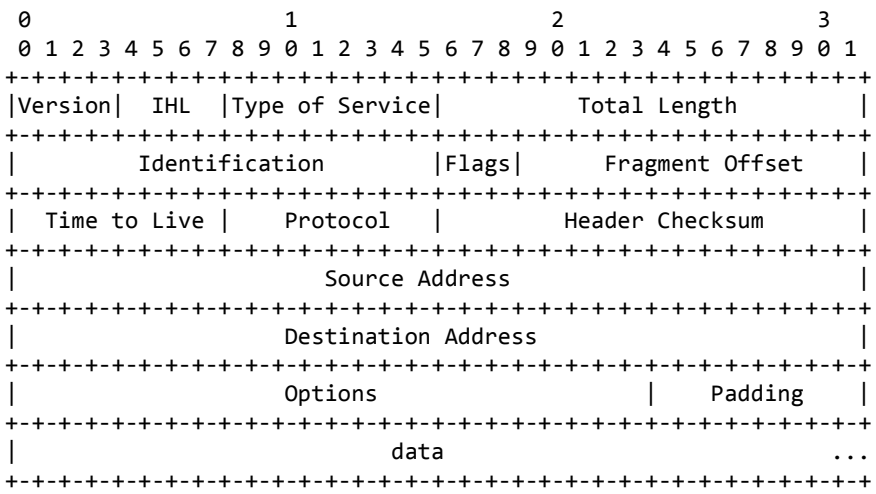
- Overview
  - Encapsulation
- Addressing
- Routing *vs*. Forwarding
- Address Management
- Control
- Evolution
- Alternatives & Summary

Overview

- Hosts are allocated addresses
- Generated data is encapsulated
- Packets passed to datalink for transmission
  - Commonly Ethernet in some form
  - *Historic interest: Token-Ring, ATM*
- Routers receive packets, examine, forward
  - May have non-IP devices between routers
  - Details of datalink are abstracted

Encapsulation

- 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



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

Contents

- Overview
- Addressing
  - Address Aggregation
  - Classless Inter-Domain Routing
- Routing *vs*. Forwarding
- Address Management
- Control
- Evolution
- Alternatives & Summary

Addressing

- Host addresses in IPv4 are 32 bits long
  - Intended to be globally unique
  - Really it's *interfaces* that get allocated addresses
  - 32 bits is no longer enough
    - E.g., virtualization, mobile telephony
- Commonly depicted in 4 parts
  - 32 bits == 4 bytes
  - Dotted quad notation (msb first): `a.b.c.d`
  - E.g., `128.243.35.39` == `0x80f32327` == `2163417895`

Address Aggregation

- Need to determine the *route* from the *address*
  - How do we get where we want to go?
  - Group addresses for efficiency: 2^{32} bits not enough but still a lot!
- Divide address into *network* and *host* parts
  - Hierarchical addressing
- Originally fixed: first 8 bits for network, remaining for hosts
  - What to do when >254 networks?
- Next, moved to *classful*: A, B, C, (D/multicast, E/reserved)
  - Indicate length of network part by high-order bits
  - But many wasted addresses as Internet grew

Classless Inter-Domain Routing

- Subnetting: make *netmask* explicit
  - Cluster addresses internally
- CIDR, RFC 1338
  - Gives rise to *prefixes* with explicit *prefix lengths*
- E.g., 172.16/12 == (172.16.0.0, 255.240.0.0)
  - A = x/8, B = x/16, C = x/24 but required routing protocol changes

Contents

- Overview
- Addressing
- Routing *vs*. Forwarding
  - Forwarding
  - Other Forwarding Actions
- Address Management
- Control
- Evolution
- Alternatives & Summary

Routing *vs.* Forwarding

- Router receives an IP packet: what to do?
  - Drop or forward?
  - If forward, which port to transmit it on?
- Making this decision is **forwarding**
  - Which port gets it closer to the destination address?
  - IP bases this decision (almost) solely on `dstip`
- 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 port
- Routing generates (prefix, port) 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

Other Forwarding Actions

- Decrement the *time-to-live* (TTL) field
  - Ensures that looping packets eventually die
- Perform any *fragmentation* necessary
  - IP abstracts datalinks with different max PDUs
  - Generate *N* packets from the first, setting fragment offset
- ...unless Don't Fragment is set:
  - Path MTU Discovery

Contents

- Overview
- Addressing
- Routing *vs*. Forwarding
- Address Management
  - Macro – Internet-wide
  - Micro – Network-local
- Control
- Evolution
- Alternatives & Summary

Address Management: Macro

- Internet Assigned Numbers Authority (IANA)
  - Co-ordinates number spaces
- Delegates *netblocks* to Regional Internet Registries (RIRs)
  - Africa, Asia/Pacific, North America, Latin America, EMEA
  - ...and down to National and Local registries (NIRs, LIRs)
- Approach appropriate registry for an allocation
  - Must provide suitable justification
  - Much harder to get a large allocation (== short prefix)
- Deeply manual process involving much politics
  - http://www.iana.org/numbers/ and http://www.iana.org/

Address Management: Micro

- Used to also be very manual
  - A big text file containing (IP, Ethernet) address map
  - Also needed to maintain subnet allocations for routing protocols
- Dynamic Host Configuration Protocol, DHCP
  - RFC 2131
  - “Can anyone give me an address?”
  - Usually provides a pile of other configuration information
- Address Resolution Protocol, ARP
  - “Does anyone know who’s at <ipaddr>?”
- Both utilise IP broadcast address, 255.255.255.255

Contents

- Overview
- Addressing
- Routing *vs*. Forwarding
- Address Management
- Control: ICMP
- Evolution
- Alternatives & Summary

Internet Control Message Protocol

- ICMP: RFC 792
  - In-band control: uses IP messages
- Most common operations:
  - ECHO REQUEST/RESPONSE (ping – “are you there?”)
  - UNREACHABLE (host, network, gateway)
  - TIME EXCEEDED (TTL was decremented to zero)
- Implementation of *traceroute*
  - How to determine route your packets take?
  - Transmit packet to destination with increasing TTL
  - Receive ICMP time exceeded from each enabled gateway

Contents

- Overview
- Addressing
- Routing *vs*. Forwarding
- Address Management
- Control
- Evolution
  - Multicast
  - Mobile
  - IPv6
- Alternatives & Summary

Multicast

- Have considered *unicast* (1:1) and *broadcast* (1:all)
- IP natively supports *multicast* (M:N)
  - Efficiency gains if the routers do the replication
  - Note multicast support exists at some lower layers too...
- Everything becomes more complex
  - Need to manage *group membership*
  - Need to manage routing between groups
- For example,
  - IGMP controls local group membership
  - PIM, MOSPF handle routing between local groups
  - MBGP permits different routing between networks

Mobile

- IP can natively support mobility; RFC3344
  - Again, so do some lower layers...
- Assign *Care-of-Address* in addition to home address
  - *Home-agent* and *foreign-agent* manage nodes
- Again, adds complexity: protocols for
  - Authenticated registration to inform HA of CoA
  - Extension to ICMP for agent discovery
  - Routing protocols, typically tunnelling
- What about multicast? (ARP? Broadcast? ...)

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
- Addressing
- Routing *vs*. Forwarding
- Address Management
- Control
- Evolution
- Alternatives & Summary

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

Summary

- IP provides encapsulation and packetization of data from higher layers
- Packets are forwarded through the network toward the specified destination by routers
- Building up information to forward is routing
  - *Topic of next lecture*
- Several control protocols, processes and organizations provide IP infrastructure