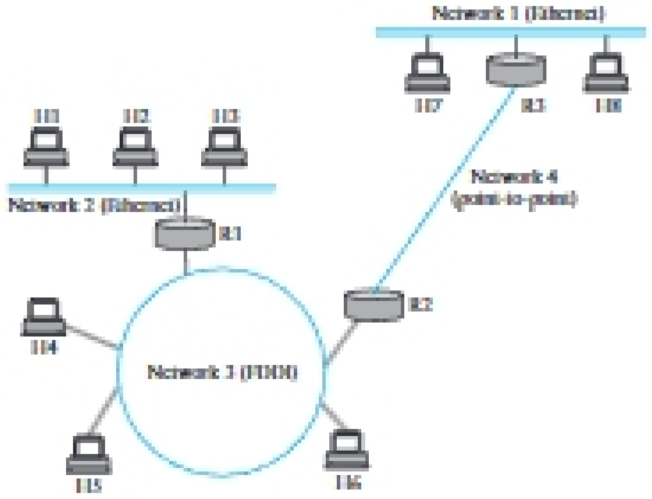
Overview

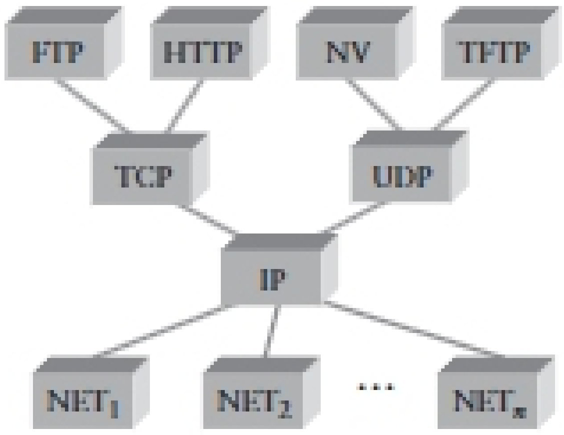
- Internet Protocol
  - What it provides and its header
  - Fragmentation and assembly
  - IP address assignment
- Address mapping and allocation
- Forwarding: switching, circuits, and source routing

Internet Protocol Goal

- Glue lower-level networks together



The Hourglass, Revisited



Internet Protocol

- Connectionless (datagram-based)
- Best-effort delivery (unreliable service)
  - packets are lost
  - packets are delivered out of order
  - duplicate copies of a packet are delivered
  - packets can be delayed for a long time

IPv4 packet format

```
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
```

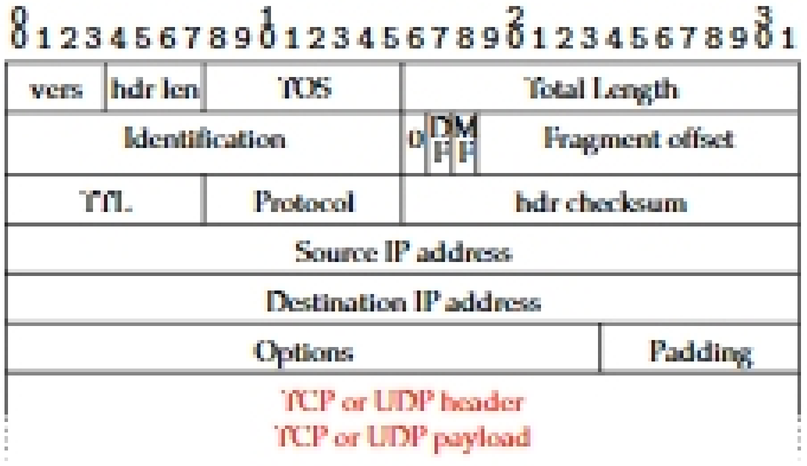
vers	hdr len	TOS	Total Length			
Identification			0	DF	MF	Fragment offset
TTL		Protocol	hdr checksum			
Source IP address						
Destination IP address						
Options						Padding
Data						

IP header details

- Routing is based on destination address
- TTL (time to live) decremented at each hop (avoids loops)
  - TTL mostly saves from routing loops
  - But other cool uses…
- Fragmentation possible for large packets
  - Fragmented in network if crosses link w. small frame size
  - MF bit means more fragments for this IP packet
  - DF bit says "don't fragment" (returns error to sender)
- Following IP header is "payload" data
  - Typically beginning with TCP or UDP header

Example Encapsulation

IPv4 packet format



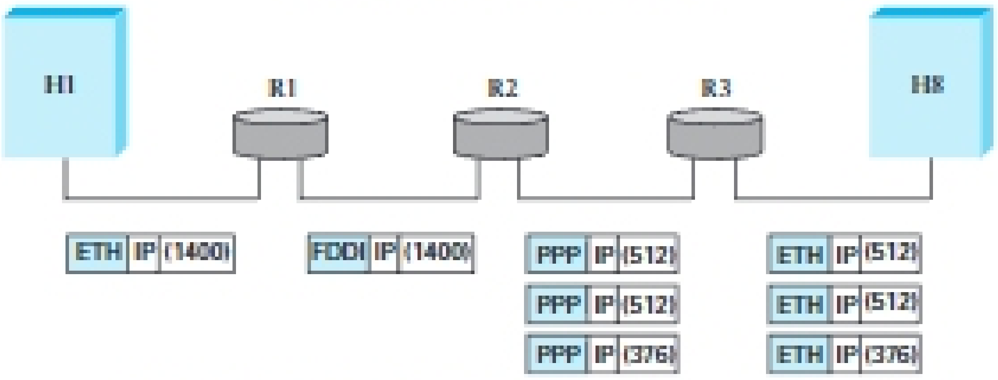
Other IP Fields

- Version: 4 (IPv4) for most packets, there's also IPv6 (lecture 9)
- Header length (in case of options)
- Type of Service (diffserv, we won't go into this)
- Protocol identifier (UDP: 17, TCP: 6, ICMP:1, why is TCP earlier?)
- Checksum over the header
- Let's look at a packet with wireshark

Fragmentation & Reassembly

- Each network has some maximum transmission unit (MTU)
- Strategy
  - Fragment when necessary (MTU < size of Datagram)
  - Source host tries to avoid fragmentation
    When fragment is lost, whole packet must be retransmitted!
  - Re-fragmentation is possible
  - Fragments are self-contained datagrams
  - Delay reassembly until destination host
  - Do not recover from lost fragments

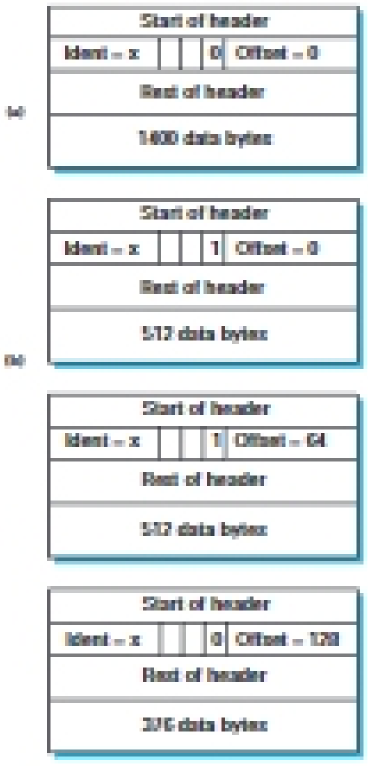
Fragmentation example



- Ethernet MTU is 1,500 bytes
- PPP MTU is 576 bytes
  - R2 Must fragment IP packets to forward them

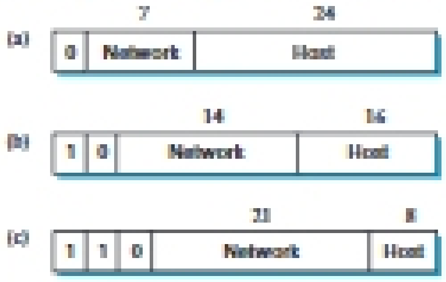
Fragmentation example (continued)

- IP addresses plus ident field identify fragments belonging to same packet
- MF (more fragments) bit is 1 in all but last fragment
- Fragment size multiple of 8 bytes
  - Multiply offset field by 8 to get fragment position within original packet



Format of IP addresses

- Globally unique (or made to seem that way)
- Hierarchical: network + host
  - Aggregating addresses saves memory in routers, simplifies routing (as we will see next lecture)
- Originally, routing prefix embedded in address:



(Still hear "class A," "class B," "class C")

- Now, routing info on "CIDR" blocks, addr+prefix-len
  - E.g., 171.67.0.0/16

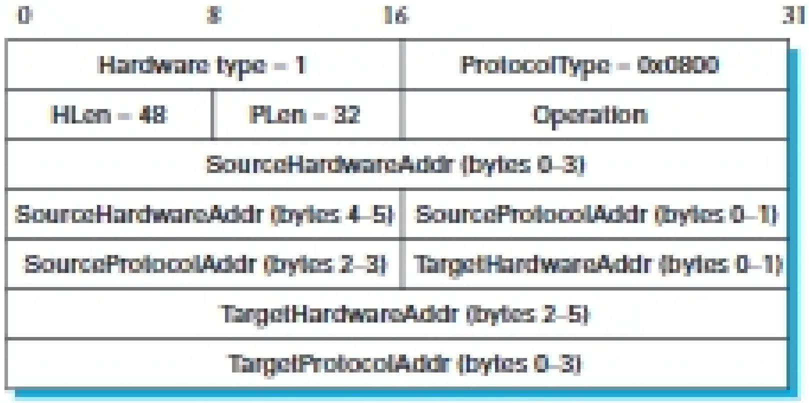
Translating IP to lower-level addresses

- Map IP addresses into physical addresses
  - E.g., Ethernet address of destination host
  - Or Ethernet address of next hop router
- Techniques
  - Encode link layer address in host part of IP address (option is available, but only in IPv6)
  - Each network node maintains a lookup table (link→IP)
- ARP – *address resolution protocol*
  - Table of IP to link layer address bindings
  - Broadcast request if IP address not in table
  - Everybody learns physical address of requesting node (broadcast)
  - Target machine responds with its link layer address
  - Table entries are discarded if not refreshed

Need for Address Translation

- Layer 2 (link) address names a hardware interface
  - E.g., my wireless ethernet 00:13:42:d2:93:11 [digits changed to protect the weak MAC-layer filtering]
- Layer 3 (network) address names a host
  - E.g., www6.stanford.edu is 171.67.22.48
- Details:
  - A single host can have multiple hardware interfaces, so multiple link layer addresses for a single network address
  - A node is asked to forward a packet to another IP address: out which hardware interface does it send the packet?

Arp Ethernet packet format

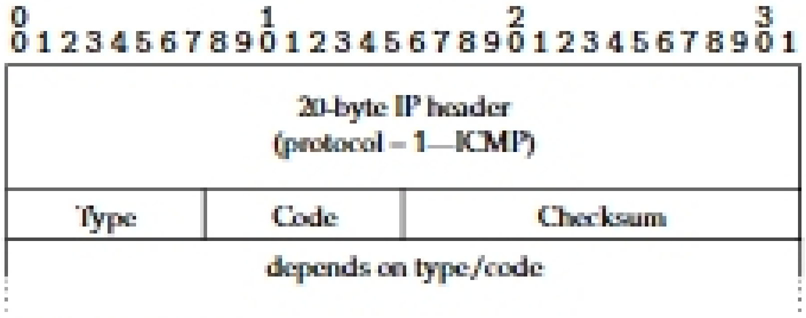


Hardware type = 1		ProtocolType = 0x0800
HLen = 48	PLen = 32	Operation
SourceHardwareAddr (bytes 0–3)		
SourceHardwareAddr (bytes 4–5)		SourceProtocolAddr (bytes 0–1)
SourceProtocolAddr (bytes 2–3)		TargetHardwareAddr (bytes 0–1)
TargetHardwareAddr (bytes 2–5)		
TargetProtocolAddr (bytes 0–3)		

Internet Control Message Protocol (ICMP)

- Echo (ping)
- Redirect (from router to source host)
- Destination unreachable (protocol, port, or host)
- TTL exceeded (so datagrams don't cycle forever)
- Checksum failed
- Reassembly failed
- Cannot fragment
- **Many ICMP messages include part of packet that triggered them**
  - Example: Traceroute

ICMP message format



20-byte IP header (protocol = 1—ICMP)		
Type	Code	Checksum
depends on type/code		

- Types include:
  - echo, echo reply, destination unreachable, time exceeded, …
  - See http://www.iana.org/assignments/icmp-parameters