

Lecture 9: Internetworking

CSE 123: Computer Networks
Alex C. Snoeren



Lecture 9 Overview



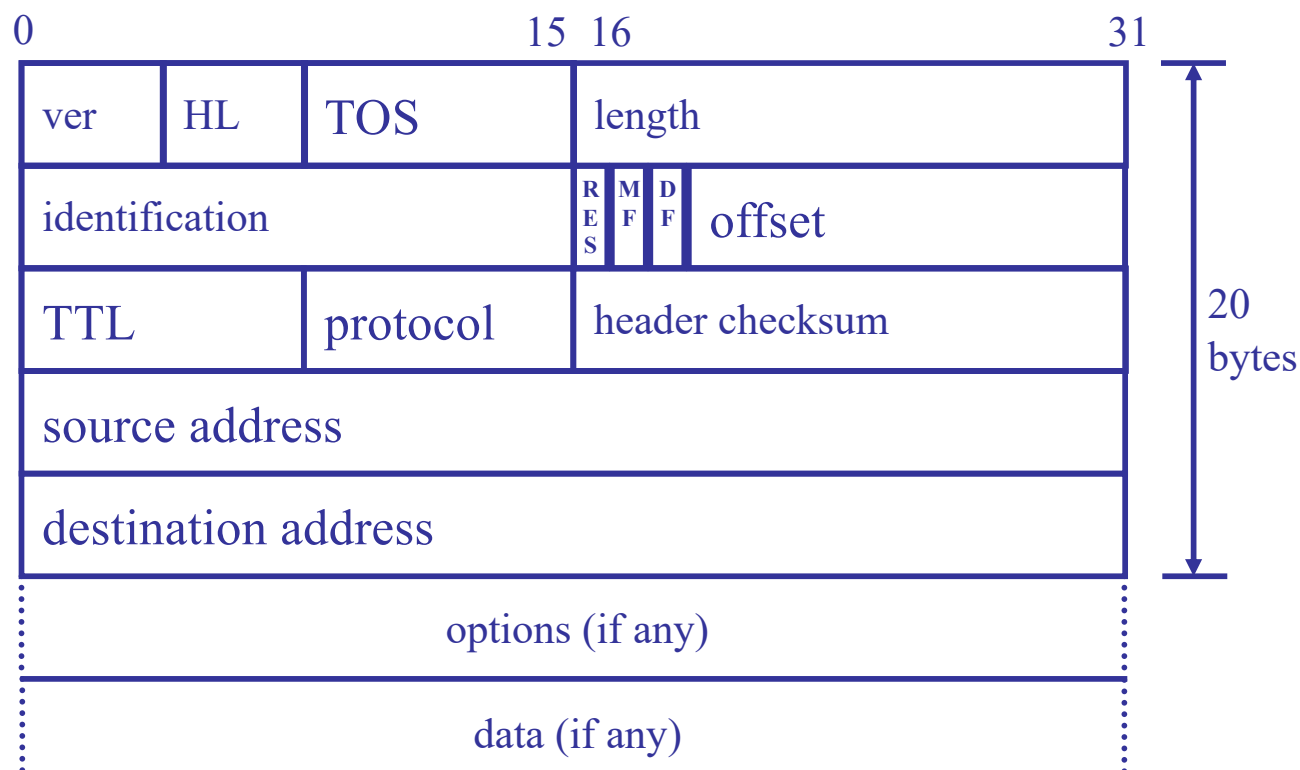
- IP Fragmentation
- ICMP, the *other* network-layer protocol
- IP Addresses
 - Class-based addressing

IP Philosophy



- Impose few demands on network
 - ◆ Make few assumptions about what network can do
 - ◆ No QoS, no reliability, no ordering, no large packets
 - ◆ No persistent state about communications; no connections
- Manage heterogeneity at hosts (not in network)
 - ◆ Adapt to underlying network heterogeneity
 - ◆ Re-order packets, detect errors, retransmit lost messages...
 - ◆ Persistent network state only kept in hosts (fate-sharing)
- Service model: **best effort**, a.k.a. *send and pray*

IP Packet Header



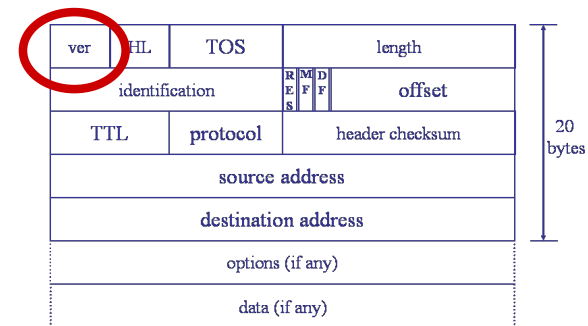
Version field



- Which version of IP is this?
 - ◆ Plan for change
 - ◆ Very important!

What versions of IP are in use today?

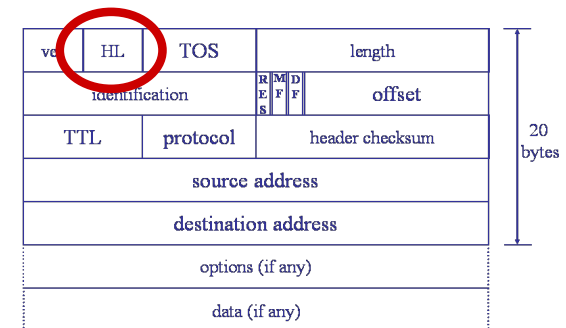
- A. IPv4
- B. IPv6
- C. Both v4 and v6
- D. More than two



Header length



- How big is IP header?
 - ♦ Counted in 32-bit words
 - ♦ Variable length
 - » Options
 - ♦ Engineering consequences of variable length...



- Most IP packet headers are 20 bytes long

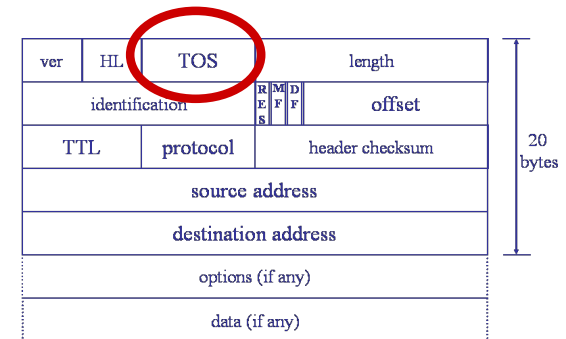
What is the content of the HL field for most IP packets?

- A. 20
- B. 5
- C. Something else

Type-of-Service



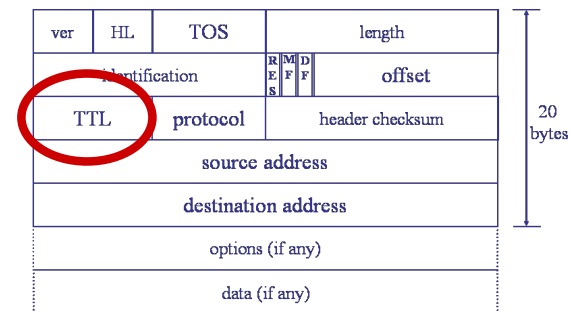
- How should this packet be treated?
 - ◆ Care/don't care for delay, throughput, reliability, cost
 - ◆ How to interpret, how to apply on underlying net?
 - ◆ Largely unused until 2000 (hijacked for new purposes, ECN & Diffserv)



TTL (Time-to-Live)



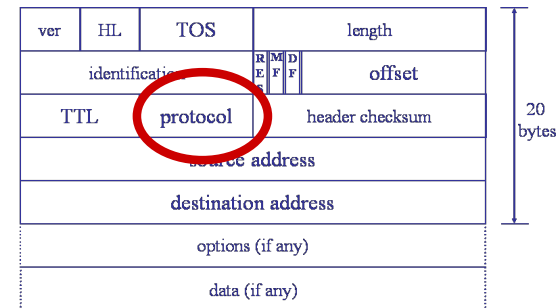
- How many more routers can this packet pass through?
 - ◆ Designed to limit packet from looping forever
- Each router decrements TTL field
- If TTL is 0 then router discards packet



Protocol



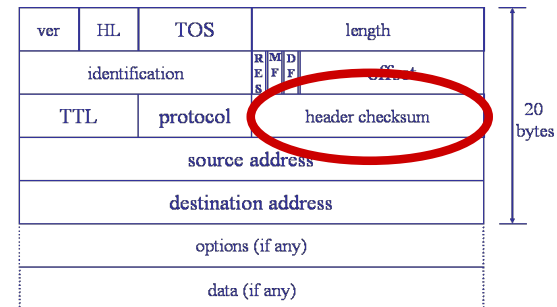
- Which transport protocol is the data using?
 - ◆ i.e. how should a host interpret the data
- TCP = 6
- UDP = 17



IP Checksum



- Header contains simple checksum
 - ◆ Validates content of header *only*
- Recalculated at each hop
 - ◆ Routers need to update TTL
 - ◆ Hence straightforward to modify
- Ensures *correct* destination receives packet





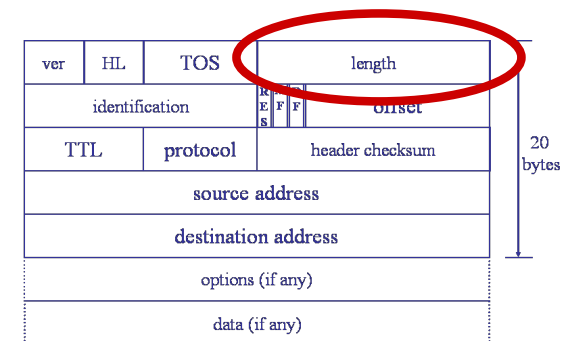
So what does IP do?

- Addressing
- Fragmentation
 - ♦ E.g. FDDI's maximum packet is 4500 bytes while Ethernet is 1500 bytes, how to manage this?
- *Some* error detection
- Routers only forward packets to next hop. They do *not*:
 - » Detect packet loss, packet duplication
 - » Reassemble or retransmit packet

Length



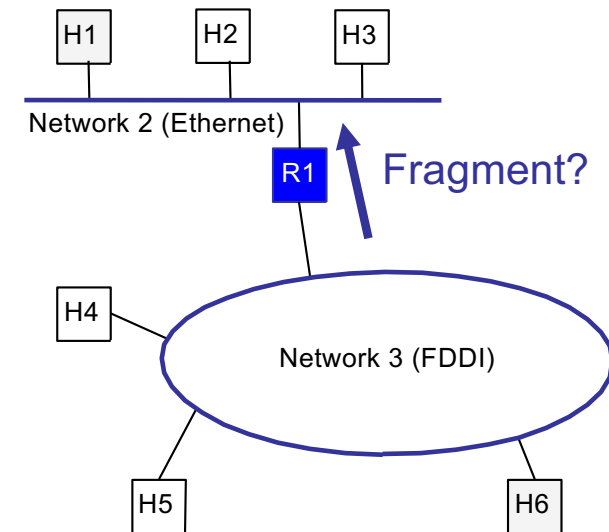
- How long is whole packet in bytes?
 - ◆ Includes header
 - ◆ Limits total packet to 64K
 - ◆ Redundant?



Fragmentation



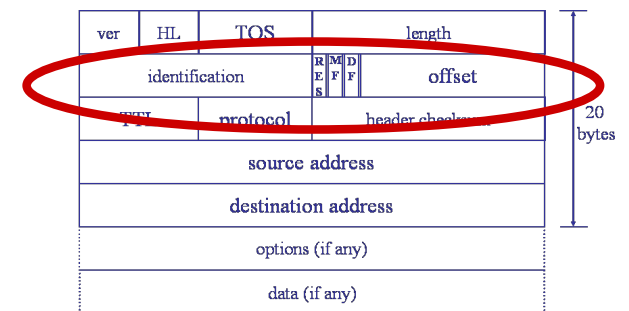
- Different networks may have different maximum frame sizes
 - ◆ Maximum Transmission Unit (MTUs)
 - ◆ Ethernet 1.5K, FDDI 4.5K
- Router breaks up single IP packet into two or more smaller IP packets
 - ◆ Each **fragment** is labeled so it can be correctly **reassembled**
 - ◆ *End host* reassembles them into original packet



IP ID and Bitflags



- Source inserts unique value in identification field
 - ◆ Also known as the IPID
 - ◆ If packet is fragmented, the router copies this value into any fragments
- Offset field indicates position of current fragment (in bytes/8)
 - ◆ Zero for non-fragmented packet
- Bitflags provide additional information
 - ◆ More Fragments bit helps identify last fragment
 - ◆ Don't Fragment bit prohibits (further) fragmentation
 - ◆ Note recursive fragmentation easily supported—just requires care with More Fragments bit





Fragmentation Example

What is the length of the final fragment?

- A. 1500
- B. 4000
- C. 1000
- D. 1040

	length	ID	MF	offset	
	=4000	=x	=0	=0	

One large datagram becomes several smaller datagrams

(Offset actually encoded as bytes/8)

	length	ID	MF	offset		0
	=1500	=x	=1	=0		
	length	ID	MF	offset		185
	=1500	=x	=1	=1480		
	length	ID	MF	offset		370
	=????	=x	=0	=2960		



Costs of Fragmentation

- Interplay between fragmentation and retransmission
 - ◆ A single lost fragment may trigger retransmission
 - ◆ Any retransmission will be of entire packet (why?)
- Packet must be completely reassembled before it can be consumed on the receiving host
 - ◆ Takes up buffer space in the mean time
 - ◆ When can it be garbage collected?
- Why not reassemble at each router?



Path MTU Discovery

- Path MTU is the smallest MTU along path
 - ◆ Packets less than this size don't get fragmented
- Fragmentation is a burden for routers
 - ◆ We already avoid reassembling at routers
 - ◆ Avoid fragmentation too by having hosts learn path MTUs
- Hosts send packets, routers return error if too large
 - ◆ Hosts can set "don't fragment" flag
 - ◆ Hosts discover limits, can size packets at source

Aside: ICMP



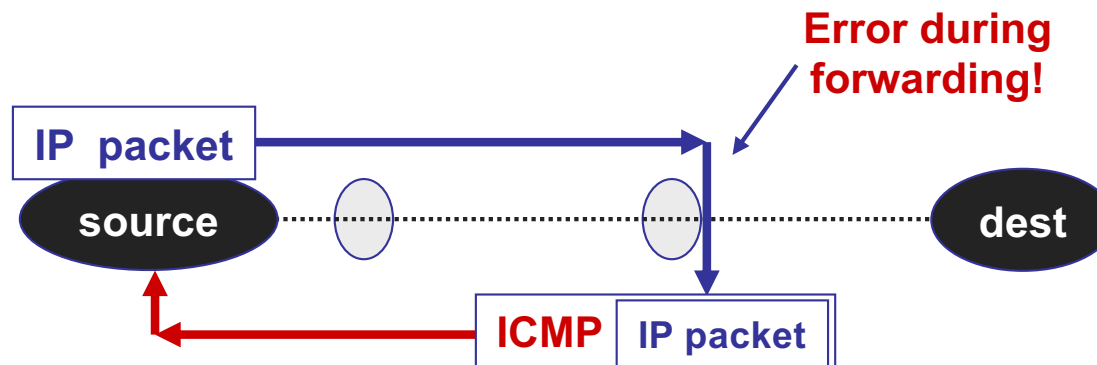
- What happens when things go wrong?
 - ◆ Need a way to test/debug a large, widely distributed system
- ICMP = Internet Control Message Protocol (RFC792)
 - ◆ Companion to IP – required functionality
 - ◆ Technically a network protocol, but implemented “above” IP
- Used for error and information reporting:
 - ◆ Errors that occur during IP forwarding
 - ◆ Queries about the status of the network

Who needs to know about packet forwarding failures?

- A. The sender**
- B. The receiver**
- C. Previous routers**
- D. All of the above**



ICMP Error Message Generation



- ICMP messages include portion of IP packet that triggered the error (if applicable) in their payload



Common ICMP Messages

- Destination unreachable
 - ◆ “Destination” can be host, network, port, or protocol – and *fragmentation needed*
- Redirect
 - ◆ To shortcut circuitous routing
- TTL Expired
 - ◆ Used by the “traceroute” program
 - » traceroute traces packet routes through Internet
- Echo request/reply
 - ◆ Used by the “ping” program
 - » ping just tests for host liveness

For Next Time



- Keep reading 3.3 in P&D
- HW 2 due before class Friday
- Project 1b and HW questions covered in discussion tomorrow