

CSE/EE 461 – Lecture 8

IP/ICMP and the Network Layer

David Wetherall
djw@cs.washington.edu

Last Time

- Focus:
 - What to do when one shared LAN isn't big enough?
- Interconnecting LANs
 - Bridges and LAN switches
 - But there are limits ...

Application
Presentation
Session
Transport
Network
Data Link
Physical

This Lecture

- Focus:
 - How do we build large networks?
- Introduction to the Network layer
 - Internetworks
 - Service models
 - IP, ICMP

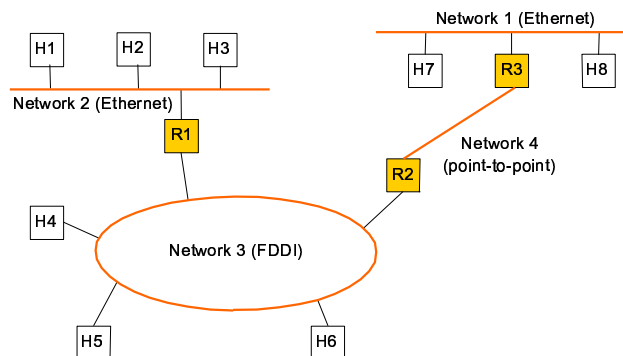
Application
Presentation
Session
Transport
Network
Data Link
Physical

djw // CSE/EE 461, Winter 2001

L8.3

Internetworks

- Set of interconnected networks, e.g., the Internet
 - Scale and heterogeneity



djw // CSE/EE 461, Winter 2001

L8.4

The Network Layer

- Job is to provide end-to-end data delivery between hosts on an internetwork
- Provides a higher layer of addressing

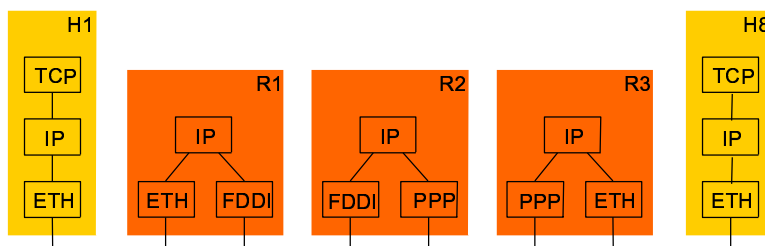
Application
Presentation
Session
Transport
Network
Data Link
Physical

djw // CSE/EE 461, Winter 2001

L8.5

In terms of protocol stacks

- IP is the network layer protocol used in the Internet
- Routers are network level gateways
- Packet is the term for network layer PDUs

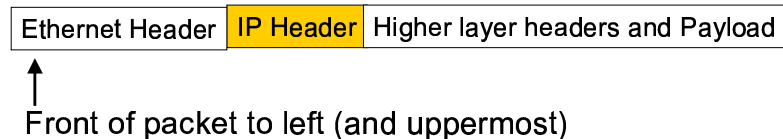


djw // CSE/EE 461, Winter 2001

L8.6

In terms of packet formats

- View of a packet on the wire on network 1 or 2
- Routers work with IP header, not higher
 - Higher would be a “layer violation”
- Routers strip and add link layer headers



djw // CSE/EE 461, Winter 2001

L8.7

Network Service Models

- Datagram delivery: postal service
 - Also connectionless, best-effort or unreliable service
 - Network can't guarantee delivery of the packet
 - Each packet from a host is routed independently
 - Example: IP
- Virtual circuit models: telephone
 - Also connection-oriented service
 - Signaling: connection establishment, data transfer, teardown
 - All packets from a host are routed the same way (router state)
 - Example: ATM, Frame Relay, X.25

djw // CSE/EE 461, Winter 2001

L8.8

Datagrams or Virtual Circuits?

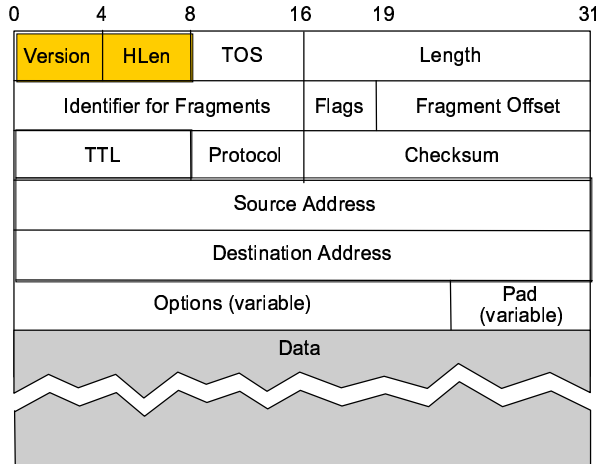
- Pros and Cons?
 - Simplicity/robustness versus stronger resource allocation
- We return to these tradeoffs later
 - Quality of Service (QOS)
 - These issues at the heart of current Internet evolution
 - Intserv (connection oriented) vs Diffserv (“connectionless”)

Internet Protocol (IP)

- IP (RFC791) defines a “best effort” service
 - May be loss, reordering, duplication, and errors!
 - Currently IPv4 (IP version 4), IPv6 on the way
- Routers forward packets using predetermined routes
 - Routing protocols (RIP, OSPF, BGP) run between routers to maintain routes (routing table, forwarding information base)
- Global, hierarchical addresses, not flat addresses
 - 32 bits in IPv4 address; 128 bits in IPv6 address
 - ARP (Address Resolution Protocol) maps IP to MAC addresses

IPv4 Packet Format

- Version is 4
- Header length is number of 32 bit words
- Limits size of options

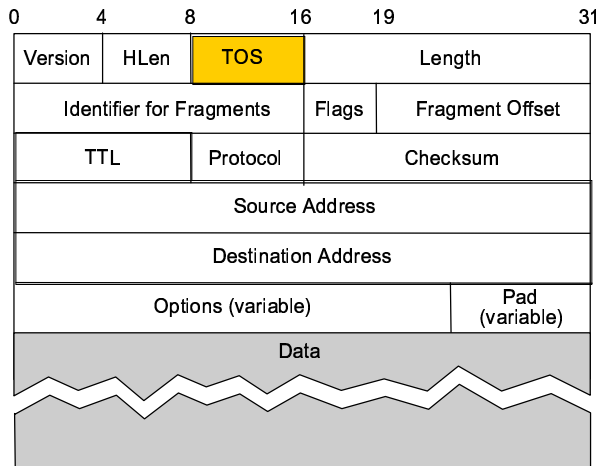


djw // CSE/EE 461, Winter 2001

L8.11

IPv4 Header Fields ...

- Type of Service
- Abstract notion, never really worked out
 - Routers ignored
- But now being redefined for Diffserv

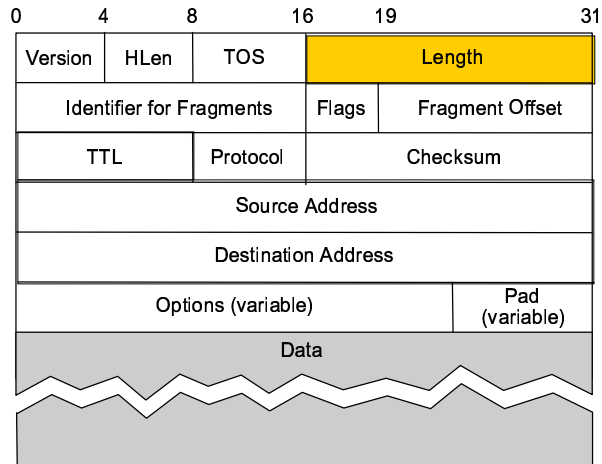


djw // CSE/EE 461, Winter 2001

L8.12

IPv4 Header Fields ...

- Length of packet
- Min 20 bytes, max 65K bytes (limit to packet size)

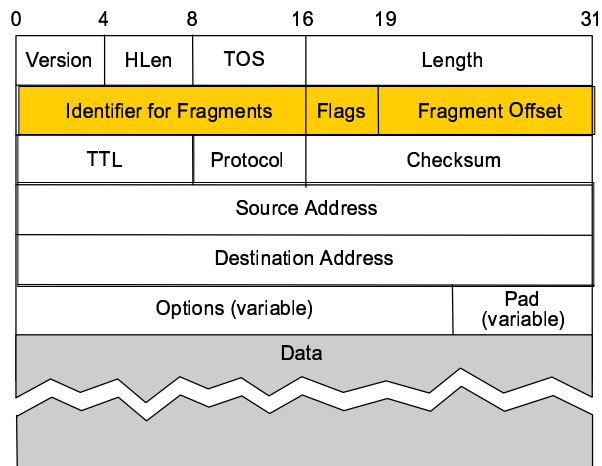


djw // CSE/EE 461, Winter 2001

L8.13

IPv4 Header Fields ...

- Fragment fields
- Different LANs have different frame size limits
- May need to break large packet into smaller fragments

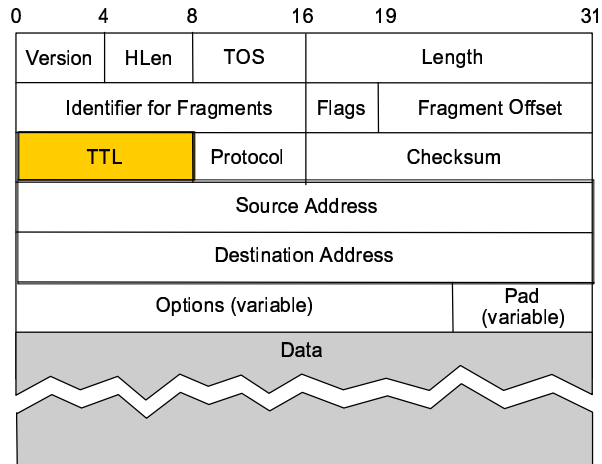


djw // CSE/EE 461, Winter 2001

L8.14

IPv4 Header Fields ...

- Time To Live
- Decremented by router and packet discarded if = 0
- Prevents immortal packets

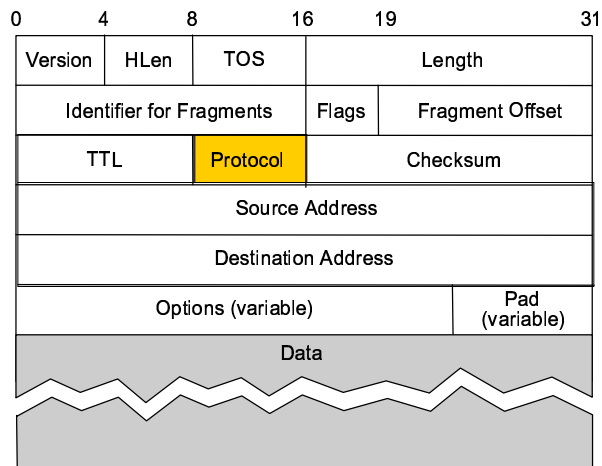


djw // CSE/EE 461, Winter 2001

L8.15

IPv4 Header Fields ...

- Identifies higher layer protocol
 - E.g., TCP, UDP

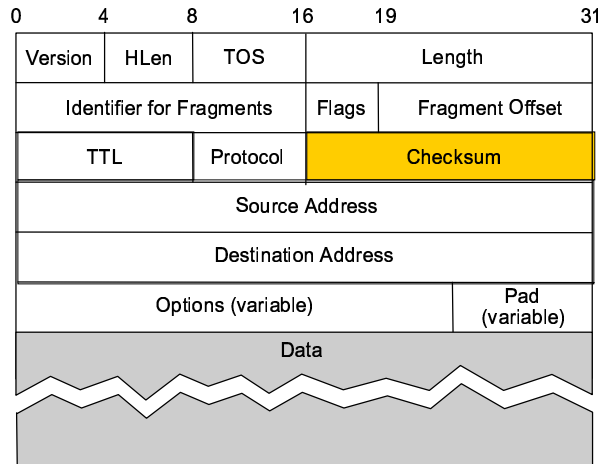


djw // CSE/EE 461, Winter 2001

L8.16

IPv4 Header Fields ...

- Header checksum
- Recalculated by routers (TTL drops)
- Doesn't cover data
- Disappears for IPv6

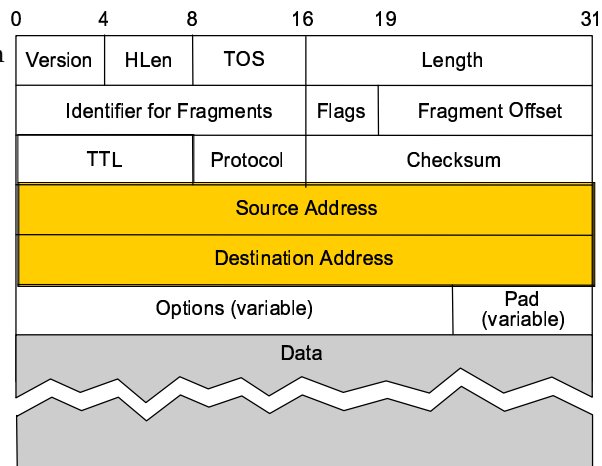


djw // CSE/EE 461, Winter 2001

L8.17

IPv4 Header Fields ...

- Source/destination addresses
 - Not Ethernet
- Unchanged by routers
- Not authenticated by default

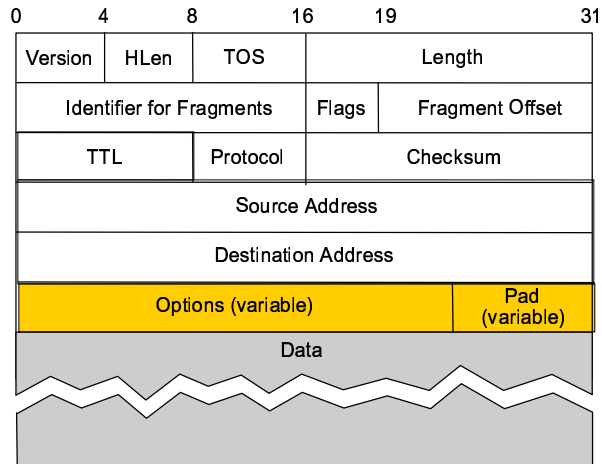


djw // CSE/EE 461, Winter 2001

L8.18

IPv4 Header Fields ...

- IP options indicate special handling
 - Timestamps
 - "Source" routes
- Rarely used ...

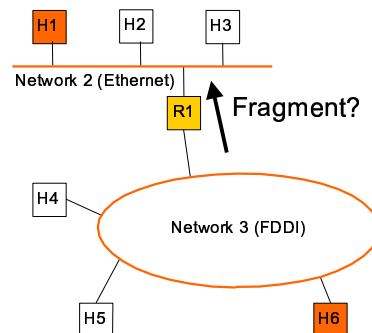


djw // CSE/EE 461, Winter 2001

L8.19

Fragmentation Issue

- Different networks may have different frame limits (MTUs)
 - Ethernet 1.5K, FDDI 4.5K
- Don't know if packet will be too big for path beforehand
 - IPv4: fragment on demand and reassemble at destination
 - IPv6: network returns error message so host can learn limit

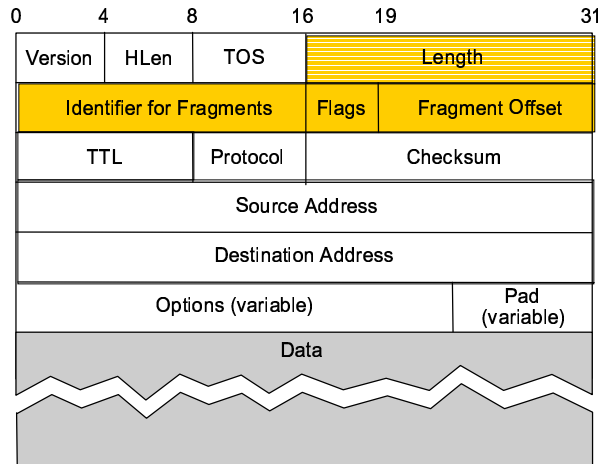


djw // CSE/EE 461, Winter 2001

L8.20

Fragment Fields

- Fragments of one packet identified by (source, dest, frag id) triple
 - Make unique
- Offset gives start, length changed
- Flags are More Fragments (MF) Don't Fragment (DF)



djw // CSE/EE 461, Winter 2001

L8.21

Fragment Considerations

- Relating fragments to original datagram provides:
 - Tolerance of loss, reordering and duplication
 - Ability to fragment fragments
- Consequences of fragmentation:
 - Loss of any fragments causes loss of entire packet
 - Need to time-out reassembly when any fragments lost

djw // CSE/EE 461, Winter 2001

L8.22

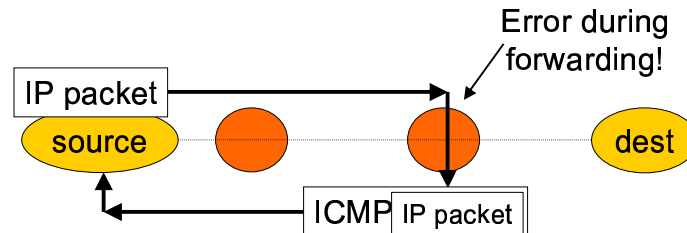
Path MTU Discovery

- 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 discover limits, can fragment at source
 - Reassembly at destination as before
- Learned lesson from IPv4, streamlined in IPv6

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
- Used for error and information reporting:
 - Errors that occur during IP forwarding
 - Queries about the status of the network

ICMP Generation



djw // CSE/EE 461, Winter 2001

L8.25

Common ICMP Messages

- Destination unreachable
 - “Destination” can be host, network, port or protocol
- Redirect
 - To shortcut circuitous routing
- TTL Expired
 - Used by the “traceroute” program
- Echo request/reply
 - Used by the “ping” program
- ICMP messages include portion of IP packet that triggered the error (if applicable) in their payload

djw // CSE/EE 461, Winter 2001

L8.26

ICMP Restrictions

- The generation of error messages is limited to avoid cascades ... error causes error that causes error!
- Don't generate ICMP error in response to:
 - An ICMP error
 - Broadcast/multicast messages (link or IP level)
 - IP header that is corrupt or has bogus source address
 - Fragments, except the first
- ICMP messages are often rate-limited too.

Key Concepts

- Network layer provides end-to-end data delivery across an internetwork, not just a LAN
 - Issues of scale and heterogeneity
 - Datagram and virtual circuit service models
 - IP/ICMP is the network layer protocol of the Internet
- Up next: More detailed look at routing and addressing