

Lecture 11:

Fragmentation & Addressing

CSE 123: Computer Networks

Stefan Savage





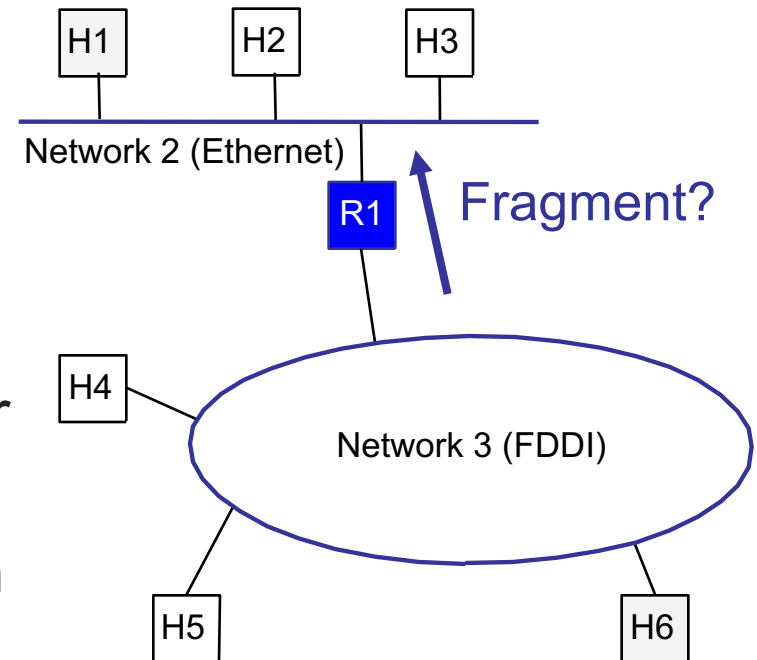
So what *does* IP do?

- 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 packets
- Today we'll talk about fragmentation and addressing



Fragmentation

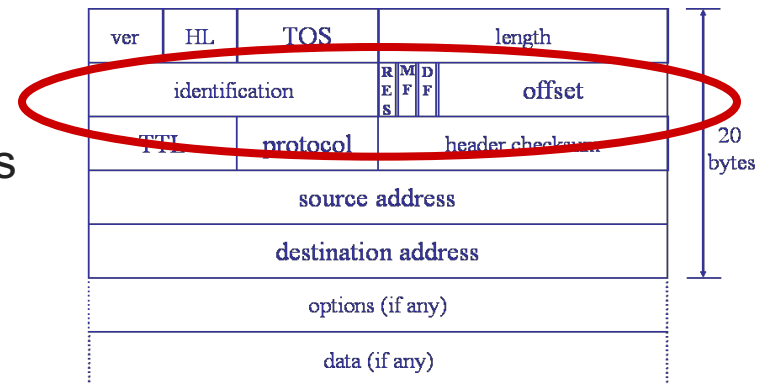
- Different networks may have different maximum frame sizes
 - ◆ Maximum Transmission Unit (MTUs)
 - ◆ Ethernet 1.5K, FDDI 4.5K
- Router breaks up single IP packet into two or more smaller IP packets
 - ◆ Each **fragment** is labeled so it can be correctly **reassembled**
 - ◆ *End host* reassembles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

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

One large datagram becomes
several smaller datagrams

(Offset actually encoded as bytes/8)

	length =1500	ID =x	MF =1	offset =0	
--	-----------------	----------	----------	--------------	--

0

	length =1500	ID =x	MF =1	offset =1480	
--	-----------------	----------	----------	-----------------	--

185

	length =1040	ID =x	MF =0	offset =2960	
--	-----------------	----------	----------	-----------------	--

370



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
 - » ICMP protocol: special IP packet format for sending error msgs
 - ◆ Reassembly at destination as before

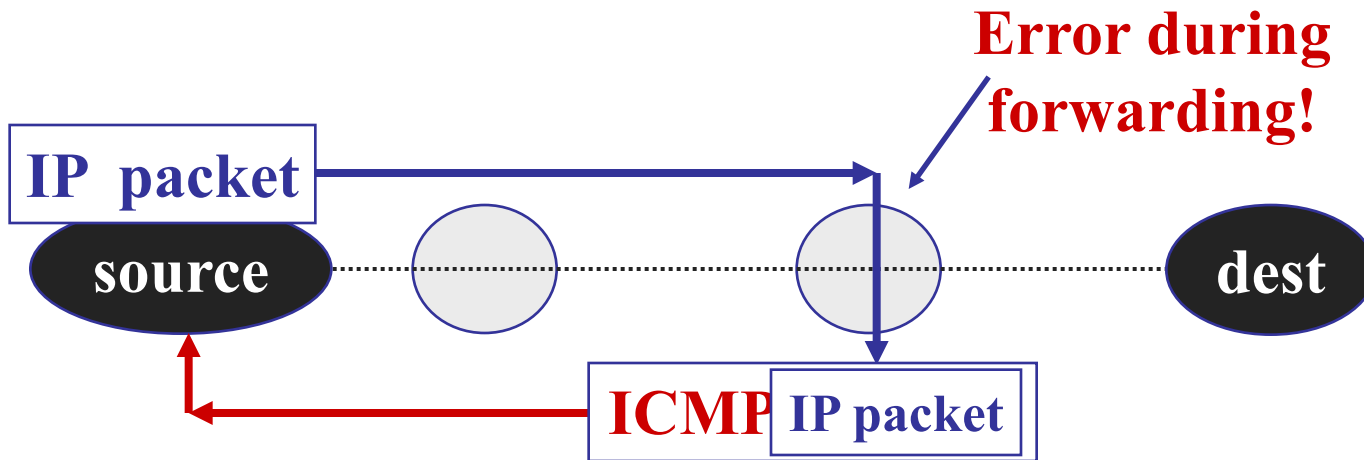
Aside: ICMP



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



ICMP Error Message Generation





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
 - » traceroute traces packet routes through Internet
- Echo request/reply
 - ◆ Used by the “ping” program
 - » ping just tests for host liveness
- ICMP messages include portion of IP packet that triggered the error (if applicable) in their payload



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
 - ◆ Don't waste valuable bandwidth sending tons of ICMP messages

Addressing Considerations



- Fixed length or variable length addresses?
- Issues:
 - ◆ Flexibility
 - ◆ Processing costs
 - ◆ Header size
- Engineering choice: IP uses fixed length addresses

Addressing Considerations (2)



- Hierarchical vs flat
 - ◆ How much does each router need to know?
- Original DARPAnet IP addressing (24 bits)
 - ◆ Global inter-network address (8 bits)
 - ◆ Local network-specific address (16 bits)



- Very successful, but now obsolete... what assumption do you think was problematic?



IP Addresses

- 32-bits in an IPv4 address
 - ◆ Dotted decimal format a.b.c.d
 - ◆ Each represent 8 bits of address
- Hierarchical: Network part and host part
 - ◆ E.g. IP address 128.54.70.238
 - ◆ 128.54 refers to the UCSD campus network
 - ◆ 70.238 refers to the host ieng6.ucsd.edu
- Which part is network vs. host?



Class-based Addressing

- Most significant bits determines “class” of address

Class A

0	Network	Host
---	---------	------

127 nets, 16M hosts

Class B

		14	16
1	0	Network	Host

16K nets, 64K hosts

Class C

21			8	
1	1	0	Network	Host

2M nets, 254 hosts

- Special addresses
 - ◆ Class D (1110) for multicast, Class E (1111) experimental
 - ◆ 127.0.0.1: local host (a.k.a. the loopback address)
 - ◆ Host bits all set to 0: network address
 - ◆ Host bits all set to 1: broadcast address



IP Forwarding Tables

- Router needs to know where to forward a packet
- Forwarding table contains:
 - ◆ List of network names and **next hop** routers
 - ◆ Local networks have entries specifying which interface
 - » Link-local hosts can be delivered with Layer-2 forwarding
- E.g. `www.ucsd.edu` address is `132.239.180.101`
 - ◆ Class B address – class + network is `132.239`
 - ◆ Lookup `132.239` in forwarding table
 - ◆ Prefix – part of address that really matters for routing



For Next Time

- Read 3.2.5-6, 9.3.1 in P&D
- HW2, Project 1