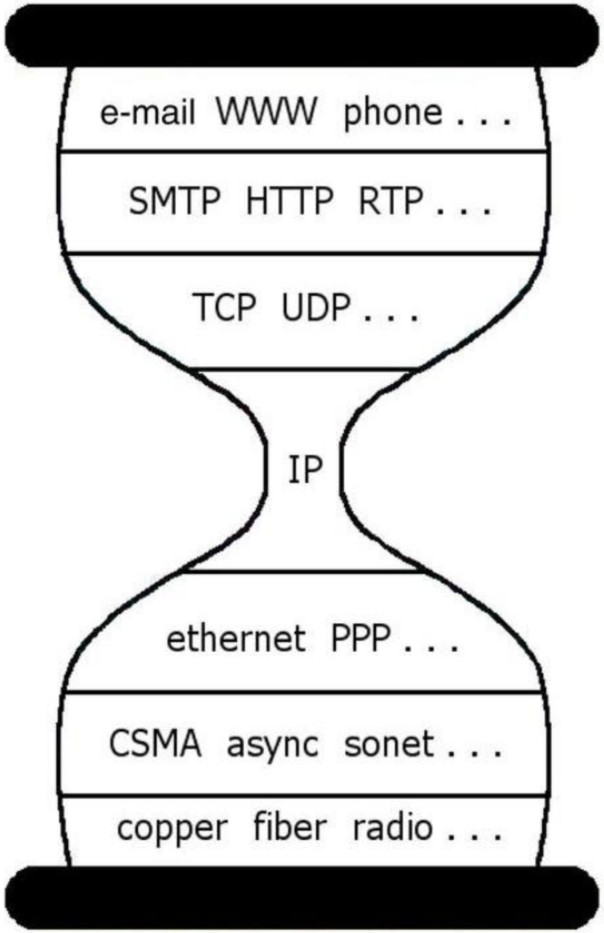
ECE 697J – Advanced Topics in Computer Networks

Networking Introduction
9/9/03

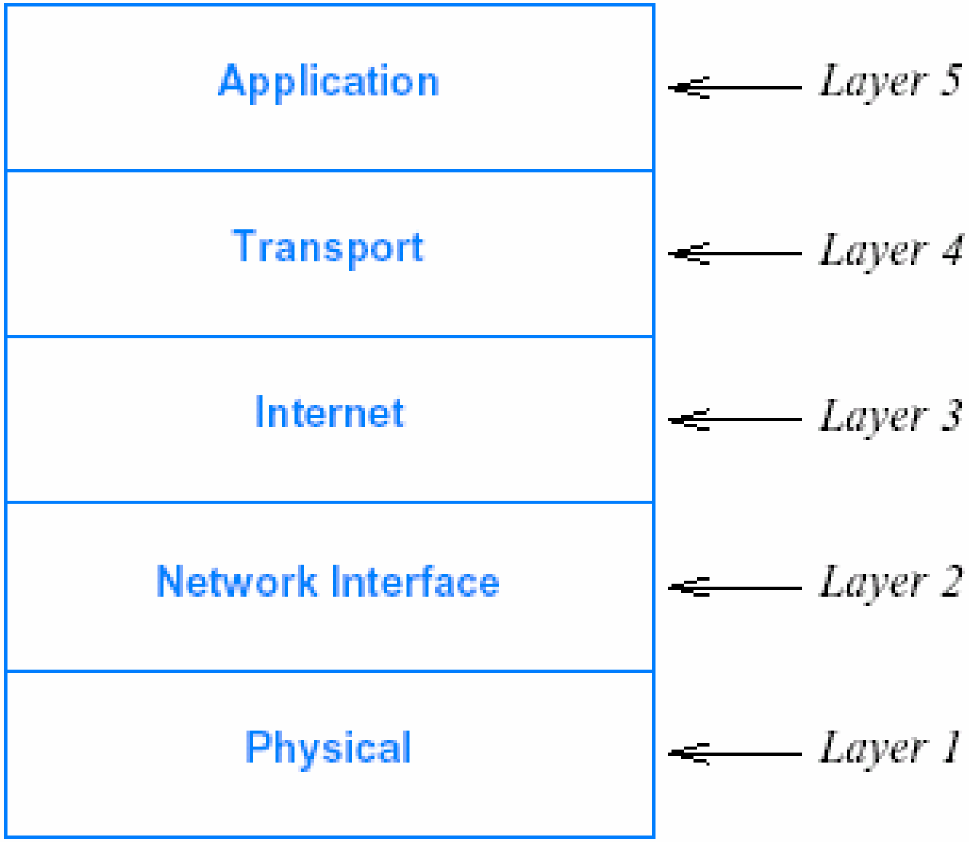
Network Protocols

- Why are several network protocols necessary?
  - Different protocols solve different problems
  - Need a mix of functionality depending on application
- How is interoperability ensured?
  - Common protocol (hourglass model)
  - In the Internet: IP
- What is a protocol suite?
  - A coordinated set of protocols
  - E.g.: HTTP over TCP over IP over Ethernet
- What is a protocol stack?
  - The software that implements a layered protocol suite



Internet Reference Model

- 5-layer reference model:
  - Derived from 7-layer OSI (Open System Interconnect) layer model
- Layer 2 is also known as link layer or data link layer
- We are not discussing physical layer
- Protocol data units:
  - Layer 2: “frame”
  - Layer 3: “packet”
  - Layer 4: “segment” or “user datagram”



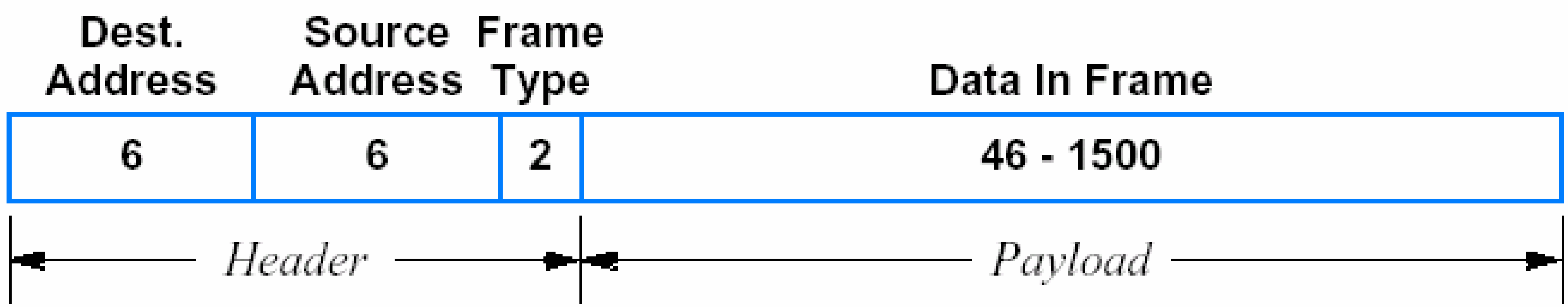
Data Link Layer

- Communication between to adjacent systems
- Point-to-point or shared media communication
  - Specification of media access (e.g., CSMA)
- Unicast, multicast, broadcast communication
  - Source and destination addresses

Type	Meaning
unicast	frame destined for a single station
multicast	frame destined for a subset of computers in the broadcast domain
broadcast	frame destined for all computers in the broadcast domain

Example: Ethernet

- Shared medium: CSMA/CD
- Frame format:



- 48-bit addresses (globally unique)
- 16-bit frame type: indicates next layer protocol
  - 0800 = IP, 0806 = ARP
- Additionally: 64-bit preamble, 32-bit CRC trailer
- Data size:
  - What are the limitations?

Network Layer: Internet Protocol

- Internet packet header:

0	4	8	16	19	24	31
VERS	HLEN	SERVICE	TOTAL LENGTH			
ID			FLAGS	F. OFFSET		
TTL		TYPE	HDR CHECKSUM			
SOURCE						
DESTINATION						
IP OPTIONS (MAY BE OMITTED)					PADDING	
BEGINNING OF PAYLOAD ⋮						

- Provides end-to-end connectivity
  - Global addressing
- Allows for several options:
  - Source routing, route recording, etc.

Internet Protocol Header

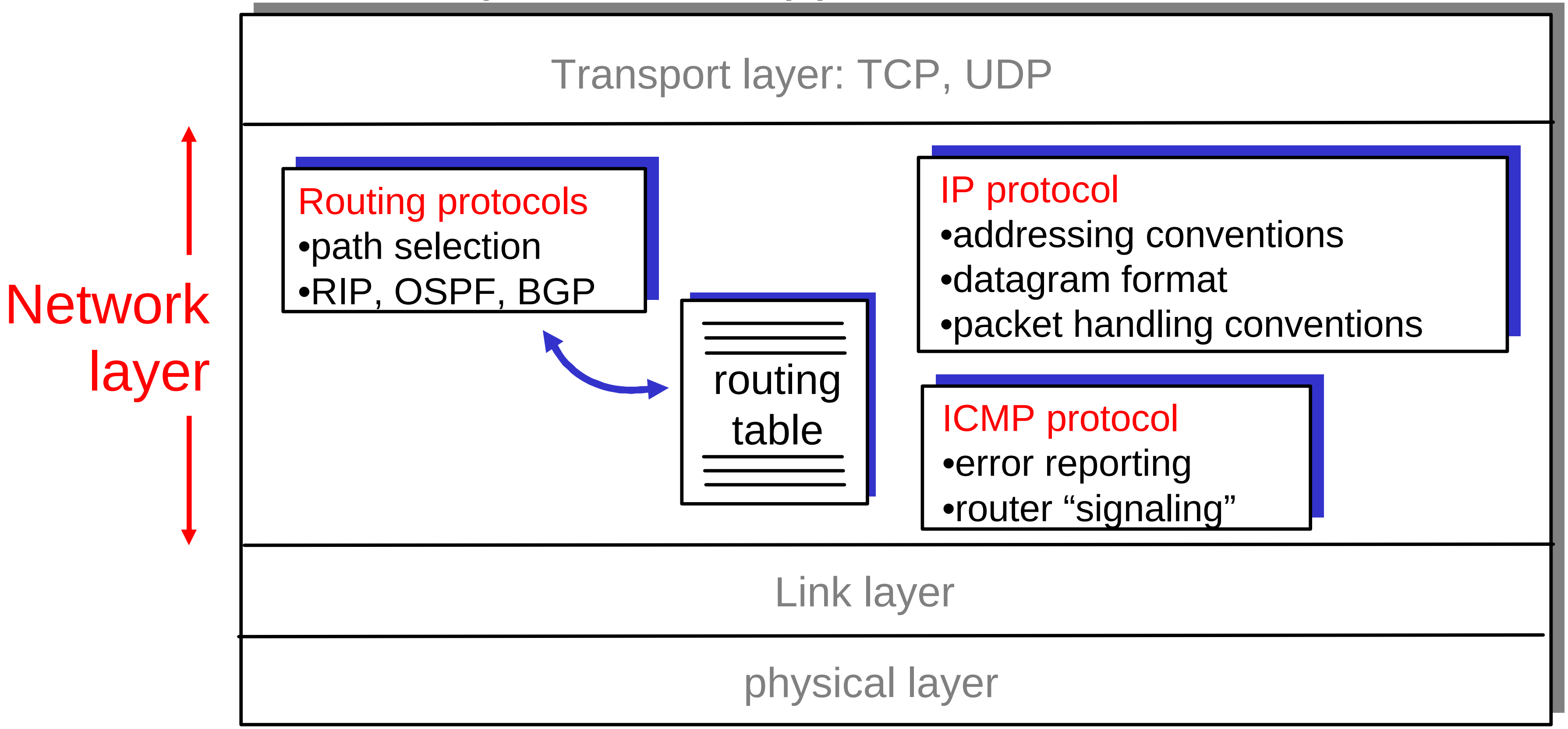
- Header fields:

Field	Meaning
VERS	Version number of IP being used (4)
HLEN	Header length measured in 32-bit units
SERVICE	Level of service desired
TOTAL LENGTH	Datagram length in octets including header
ID	Unique value for this datagram
FLAGS	Bits to control fragmentation
F. OFFSET	Position of fragment in original datagram
TTL	Time to live (hop countdown)
TYPE	Contents of payload area
HDR CHECKSUM	One's-complement checksum over header
SOURCE	IP address of original sender
DESTINATION	IP address of ultimate destination
IP OPTIONS	Special handling parameters
PADDING	To make options a 32-bit multiple

- Which get changed between hops?

IP-Related Protocols

- Several other protocols support IP:



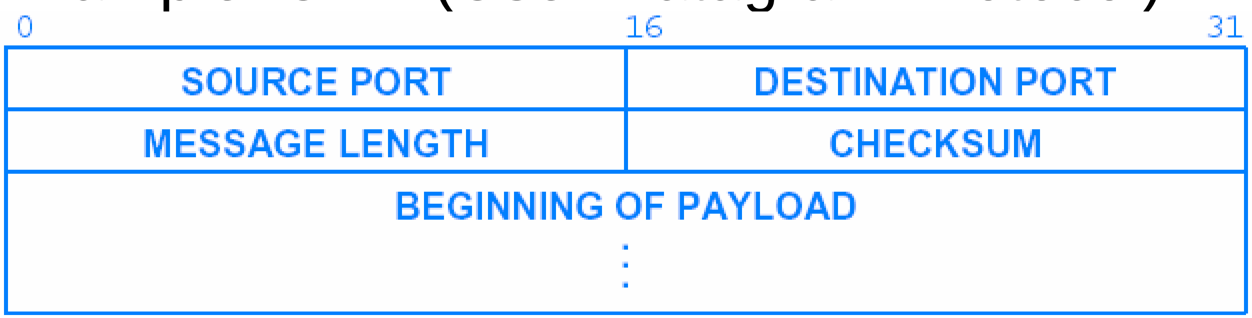
- We'll cover routing tables with packet processing

IP Addressing

- Hosts within a subnet have same address prefix
  - Example: ECE 128.119.86.*, UMass 128.119.*
- Prefixes are assigned by Internet Assigned Numbers Authority (IANA) to Regional Internet Registries (RIR)
- Classless Interdomain Routing (CIDR)
  - Subnet addresses given by prefix and prefix-length
  - Example: 128.119.86.0/24 or 128.119.0.0/16
  - Any prefix length allowed – no more address classes(!)
- Are addresses globally unique?
  - No, Network Address Translators (NAT) allow private subnets
  - Private subnet address space: 192.168.0.0/16, 172.16.0.0/12, 10.0.0.0/8

Transport Layer

- End-to-end protocols for communication between applications
  - Transport layer is not changed inside the network
- How are applications identified?
  - Port numbers used for demultiplexing
- Example: UDP (User Datagram Protocol)

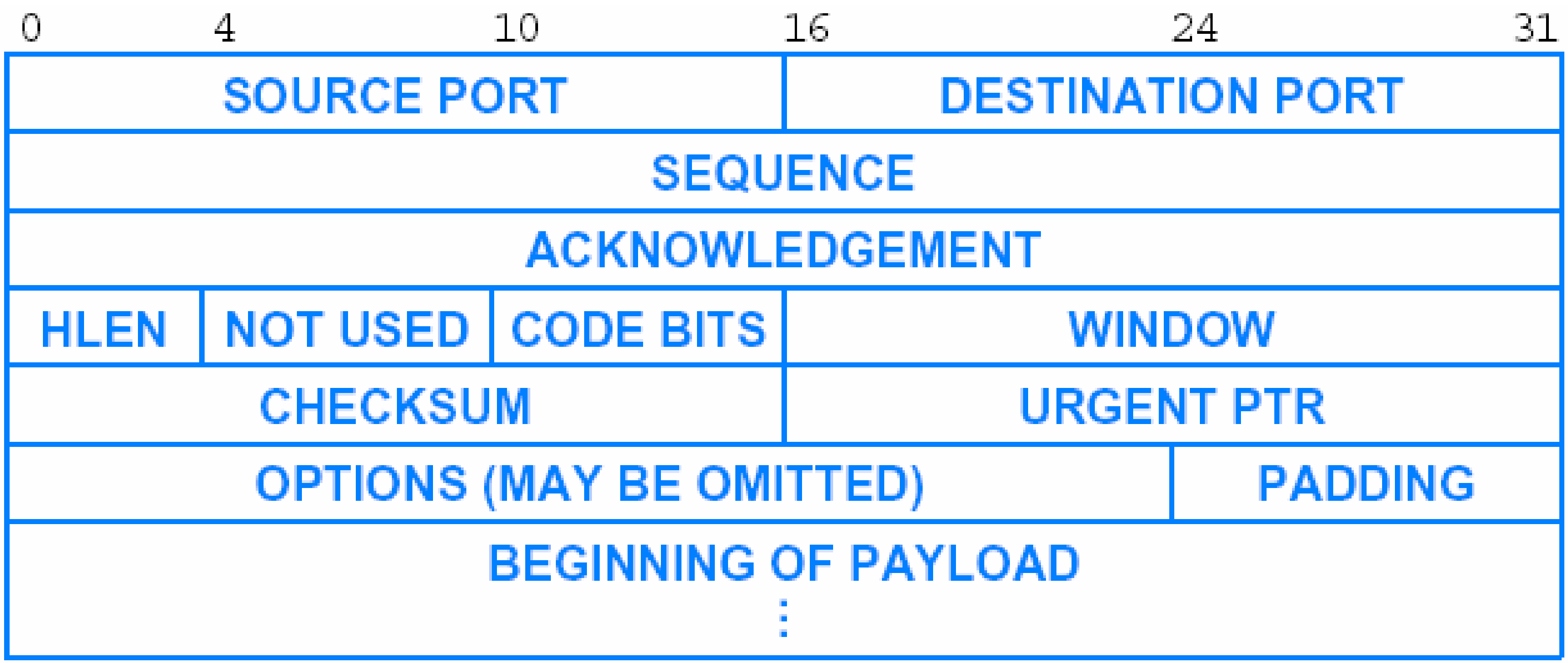


Field	Meaning
SOURCE PORT	ID of sending application
DESTINATION PORT	ID of receiving application
MESSAGE LENGTH	Length of datagram including the header
CHECKSUM	One's-complement checksum over entire datagram

Transport Control Protocol

- UDP shortcomings:
  - Unreliable (packet loss)
  - Packet reordering
  - No congestion control
  - No flow control
- TCP addresses these problems:
  - Sequence numbers
  - Acknowledgements and retransmission timers
  - Congestion and flow control windows

TCP Header



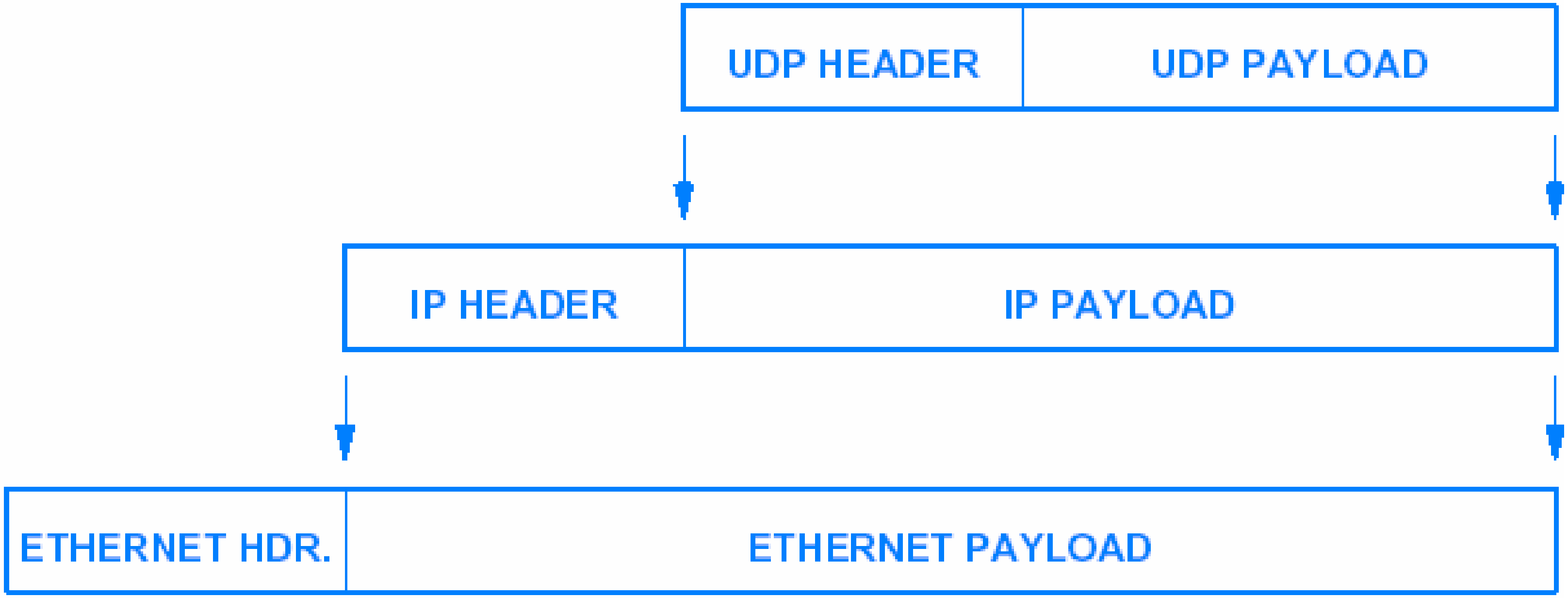
- Checksum, source and destination ports similar to UDP
- Sequence and Ack number is byte count (not packet)

TCP Questions

- How is a connection between two hosts established?
  - Three-way handshake
- How is reliability ensured when packet is lost?
  - Packet is not acknowledged, timeout occurs, retransmission
- How is congestion control achieved?
  - Congestion window is continually increased to use available bandwidth
  - Congestion window is reduced when packet loss occurs

Protocol Encapsulation

- Protocols are encapsulated according to the protocol suites used
- Example:



Real-World Example

- Packet captured with Ethereal:

```
⊞ Frame 304 (1514 bytes on wire, 1514 bytes captured)
⊟ Ethernet II, Src: 00:05:5d:d9:c7:4c, Dst: 00:09:6b:e0:51:d8
    Destination: 00:09:6b:e0:51:d8 (IBM_e0:51:d8)
    Source: 00:05:5d:d9:c7:4c (D-Link_d9:c7:4c)
    Type: IP (0x0800)
⊟ Internet Protocol, Src Addr: 206.65.183.58 (206.65.183.58), Dst Addr: 192.168.0.100 (192.168.0.100)
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
    Total Length: 1500
    Identification: 0xd185
  ⊞ Flags: 0x04
    Fragment offset: 0
    Time to live: 50
    Protocol: TCP (0x06)
    Header checksum: 0x2b0e (correct)
    Source: 206.65.183.58 (206.65.183.58)
    Destination: 192.168.0.100 (192.168.0.100)
⊟ Transmission Control Protocol, Src Port: http (80), Dst Port: 3342 (3342), Seq: 3111424399, Ack: 2138476516, Len: 1460
    Source port: http (80)
    Destination port: 3342 (3342)
    Sequence number: 3111424399
    Next sequence number: 3111425859
    Acknowledgement number: 2138476516
    Header length: 20 bytes
  ⊞ Flags: 0x0018 (PSH, ACK)
    Window size: 2920
    Checksum: 0x3611 (correct)
⊟ Hypertext Transfer Protocol
    Data (1460 bytes)
```

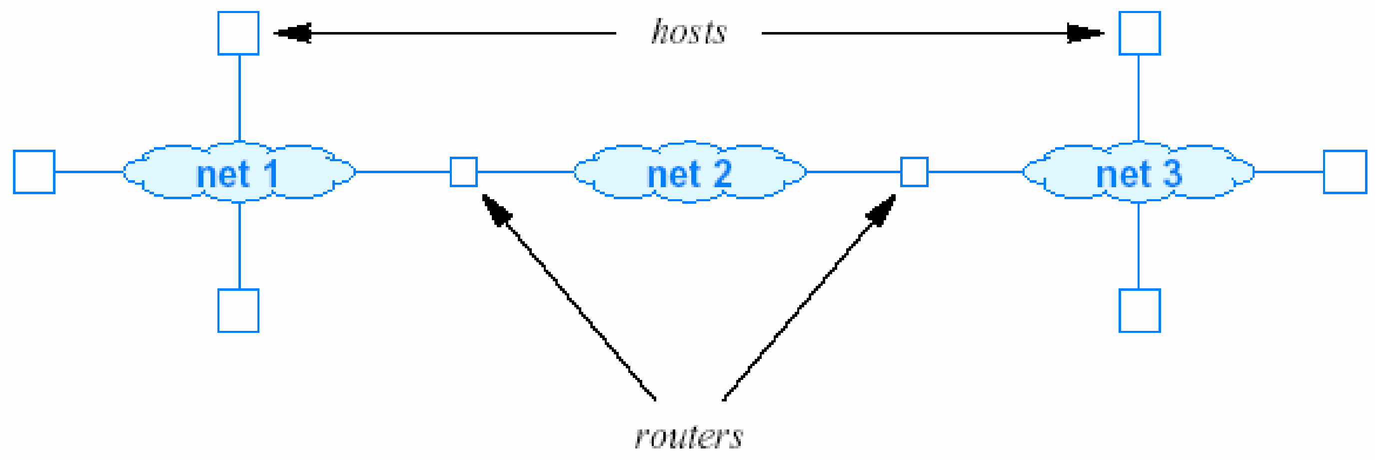
Address Resolution Protocol

- Host needs to know Ethernet address of given IP interface
- Address Resolution Protocol is layer 2 protocol
  - Directly encapsulated in Ethernet frame
- ARP header:

0	8	16	24 31
ETHERNET ADDRESS TYPE (1)		IP ADDRESS TYPE (0800)	
ETH ADDR LEN (6)	IP ADDR LEN (4)	OPERATION	
SENDER'S ETH ADDR (first 4 octets)			
SENDER'S ETH ADDR (last 2 octets)		SENDER'S IP ADDR (first 2 octets)	
SENDER'S IP ADDR (last 2 octets)		TARGET'S ETH ADDR (first 2 octets)	
TARGET'S ETH ADDR (last 4 octets)			
TARGET'S IP ADDR (all 4 octets)			

Network Systems

- The obvious: hosts and routers



- Hosts can be variety of devices:
  - Workstations, servers, wireless PDAs, cell phones, etc.
- But there is more on different layers

Layer 2 Devices

- Bridges:
  - Connection between two networks on data link level
  - Isolation of Ethernet collision domains
- Layer 2 switch:
  - Similar to bridge
  - Often with point-to-point connections on each port
  - High-throughput
- VLAN switch:
  - Supports several Virtual LANs
  - Layer 2 switch that emulates several smaller switches

Layer 3 & 4 Devices

- IP Router
  - Packet forwarding
  - IP destination address lookup, simple packet header processing
- Firewall
  - Blocks packets to certain internal addresses and ports
  - Maintains list of currently active connections
- Network Address Translator (NAT)
  - "Hides" subnet behind single external IP address
  - Rewrites packets to change IP address and port numbers
- Load Balancer
  - Distributes web requests to server farm
  - Uses Layer 4+ (or Layer 7) classification and TCP splicing
- Set-Top Box
  - Decrypts content for service subscribers
- Other devices: Monitor, Policer, Shaper, Analyzer

Homework

- Install tcpdump or Ethereal on your machine
  - Capture packets from a web transfer
  - Telnet into your ecs account and see that your password is sent in cleartext!
  - Ssh into your account and look at encryption
- Read Chapter 4
- Look at web page to decide papers that you are interested in
  - Who wants to do IP lookup?

Papers

- **IP lookup:** Marcel Waldvogel, George Varghese, Jon Turner, Bernhard Plattner. Scalable High Speed IP Lookups. In Proc. of ACM SIGCOMM 97, pages 25-36, Cannes, France, September 1997.
- **Router design:** S. Keshav and Rosen Sharma. Issues and Trends in Router Design. IEEE Communications Magazine, 36(5):144-151, May 1998.
- **Network applications (1):** George Apostolopoulos, David Aubespin, Vinod Peris, Prashant Pradhan,
Debanjan Saha. Design, Implementation and Performance of a Content-Based Switch. In Proc. of IEEE INFOCOM 2000, pages 1117-1126, Tel Aviv, Israel, March 2000.
- **Network applications (2):** Li-wei Lehman, Stephen J. Garland, and David L. Tennenhouse. Active reliable multicast. In Proc. of IEEE INFOCOM 98, pages 581-589, San Francisco, CA, April 1998.
- **Active networking:** David L. Tennenhouse and David J. Wetherall. Towards an active network architecture. Computer Communication Review, 26(2):5-18, April 1996.
- **Scheduling:** M. Shreedhar and George Varghese. Efficient fair queuing using deficit round-robin. IEEE/ACM Transactions on Networking, 4(3): 375-385, June 1996.