Lecture material:

- 3.3.3.1 Users desire universal interconnection
   Implications
   - Don't want to mandate any interconnect topology - organization can add LAN whenever it wants. Backbone graph need not be a star, ring, tree, etc.
   - Need 1 set of names/addresses used throughout internet
   - Network level operations should be independent of physical network topology. (Ethernet vs. radio net vs. T1)

- 3.5 How are networks interconnected to form an internet?

<table>
<thead>
<tr>
<th>Network Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeater</td>
<td>Copy electrical signal</td>
</tr>
<tr>
<td>Bridge</td>
<td>Store, forward packet</td>
</tr>
<tr>
<td>IP Router/Gateway</td>
<td>Store, forward packet</td>
</tr>
</tbody>
</table>

- The internet model implies:
  - Your network may carry an extra load due to traffic on 2 networks that your network connects:

- Why TCP/IP?
  - Vendor-independent: a novelty in 1981
  - Single protocol runs on:
    - micros through CRAYs
    - both LANs and WANs
    - both big (900,000 host Internet) and small (2 machines in my home) networks
  - Lots of performance optimizations over the years
  - Biggest internet in the world runs it
  - Becoming defacto standard in U.S.: .edu, .com, and .gov use it

- Definitions
  - "Internet": the big one running TCP/IP (900,000 hosts)
  - "Internet": any old connection of networks possibly using non-TCP/IP protocol

- TCP/IP protocol suite: (Fig. 5.1)
  - TCP
  - UDP
  - ICMP (Internet Control Message Protocol):
    - When/how to report errors between gateways and hosts
  - IP (and SLIP)
  - ARP (Address resolution protocol): Maps IP address to layer 2 address
  - RARP (Reverse address resolution protocol): Maps layer 2 address into an Internet address - used by diskless workstation upon booting

- RARP client protocol in diskless workstation's ROM
  - Upon booting, diskless host broadcasts RARP packet
  - A satellite link between machines
  - The above vary by a large range of delay, throughput, #nodes/network

- History:
  - 1970's: ARPANET was early experiment in networking
  - Early 1980's: TCP and IP protocols replace earlier ARPANET protocols; emphasis switches to internetworking
  - 1984: ARPANET + MILNET split
  - 1987: NSFNet links supercomputer centers, funds regional networks, allows ubiquitous network service

- IP addresses (Comer Ch. 4)
  - 2 sets of addresses:
    - Human readable: vtopus.cs.vt.edu
    - "Dotted decimal": 128.173.40.1, where:
      - 128.173 = .vt.edu
      - .40,.41 = .cs (.88-ee)
      - .1 = host id 1 (255 possible hosts in .cs.vt.edu)
  - Cat "inet/hosts" for more examples
  - IP address = (netid, hostid)
  - Every LAN has a unique netid; it is 128.173 at .vt.edu
  - Gateways connect netid's together; therefore gateways know about netid's but not about hostid's
  - 5 forms of addresses (4 in Stevens, Fig. 5.2)
    - Each organization (.vt.edu) is generally assigned one netid
    - Q. How many Internet hosts are possible?
      - A. About (see Fig. 4.3) $2^{29} + 2^{29} + 2^{29} + 2^{29} = 3.84x10^{19}$. However, ranges of IP addresses are allocated to organizations. So we may run out of IP addresses much sooner.
      - Class A for lots of hosts on a few networks in an organization
- Class C for a few hosts on each of many networks in an organization

Q. Which class is .vt.edu?
A. 7 bits < 128.173=.vt.edu < 21 bits, so it must be class B

- Routing:
  - If (source netid=destination netid) then direct else indirect delivery
  - More later

- Special addresses (Comer Vol I, Fig. 4.3):
  - IP = (netid=0,hostid) means this network with specified host id
  - IP = all 1's means broadcast on the local network; a host could learn its netid this way from another host
  - Loopback: 127.0.0.1 at .vt.edu

Q. If my VPI Unix host is sold to another university, can it retain its IP address?
A. No: The IP address encodes the netid, which is organization specific. Even if host moves from CS to EE, its IP address must change: 128.173.40.x to 128.88.5.x

- Weaknesses of IP addressing:
  - If host moves from one network to another, its IP address must change
  - Because routing is based on destination netid of IP address, packets may take different routes to a multihomed host depending on which address is used. (Example: willow.cs.vt.edu and locust.cs.vt.edu are connected by both fddi and ethernet; IP address used in socket interface determines which network is used)

- Who assigns IP addresses?
  - netids: Network Information Center (NIC)
  - hostids: your organization decides

- IP protocol (Comer Ch. 7)

  - (Arpanet pioneers to ISO: "Boy, have we got a protocol for you")
  - Connectionless:
    - each datagram from host A to B can follow different path
    - fragments of a single datagram may follow different paths
  - Best-error delivery:
    - IP is unreliable only when network resources (e.g., gateway buffers, network bandwidth) are exhausted or network fails

- IP Functions
  - Defines unit of transfer in internet:
    - 2^16 byte limit on a packet
    - packet header format
  - Performs routing
  - Protocol for unreliable, connectionless delivery:
    - What does host/gateway do when datagram arrives
    - When can datagram be discarded?
    - When/how to report errors (ICMP - Internet Control Message Protocol)
  - Fragmentation:
    - IP datagram has a maximum length of 2**16 bytes. Of the 2**16 bytes, normally 20 bytes are an IP header.
    - A network is said to have a maximum transfer unit (MTU), in units of bytes:
      - Ethernet: MTU=1500
      - FDDI: MTU=4532 (from RFC 1188)
    - IP says that a host or gateway must accept datagrams of length 576 bytes or the maximum of the MTU's of the networks to which the machine is attached, whichever is larger
    - In addition, a transport protocol - such as TCP - may try to match the segment size it gives IP to the underlying network. We will see this later. (TCP will use the MTU of the network, if the source and destination are on the same physical network; otherwise it uses 576-20-20 (the 20's are for the TCP and IP headers) as the maximum segment size (MSS).)
  - Problem: What happens if a datagram must pass through multiple networks with different MTU's (Comer Fig. 7.6)?
    - Solution: Fragmentation (Comer Fig. 7.7)
      - As a datagram traverses the network, it is continually broken into smaller pieces (down to the minimum IP datagram size of 576 bytes).
      - The destination host must then reassemble the original datagram, and pass it to its transport layer.

Q. Does G2 or HostB reassemble fragments (Comer Fig. 7.6)?
A. HostB does, for simplicity; gateways need not store fragments.

A fragment has almost the same header as the original datagram, except for bits indicating if it is a fragment, if it is the last fragment, and the offset.

Note that fragments may arrive out of order at the destination, because they are all routed independently.

Note that the destination must have enough buffer space to fully assemble one IP datagram (of max length 2**16), or reassembly deadlock will occur.

The destination starts a reassembly timer when the first fragment arrives. The timer is used discard fragments if a fragment is lost.
- **IP datagram header format** (Comer Fig. 7.3)
  - VERS = 4
  - HLEN >= 5 (unit is 32-bit words; minimum header length is 5 words)
  - **SERVICE TYPE**:
    - 3 bits for one of 7 priority levels
    - High reliability desirable (on/off bit) (don't use radio network!)
    - High throughput desirable (on/off bit) (use high bandwidth link)
  - **TOTAL LENGTH**: in bytes, for this IP datagram (thus for a fragment TOTAL LENGTH is the fragment length)
  - 3 fields for fragmentation/reassembly:
    - IDENTIFICATION: Unique id assigned to each datagram send by the source host (typically the value of a global counter incremented each time an IP packet is formed). Allows destination to tell which fragments belong to which datagrams.
    - FLAGS:
      - Do not fragment this datagram (useful for network testing)
      - More fragments (after this one)?
    - FRAGMENT OFFSET: offset of data in data field of original datagram in multiples of 8 bytes
  - **TIME TO LIVE**: each gateway dec by 1; discard if 0; detects loops in routing tables
  - **PROTOCOL**: 6 for TCP. This way, IP knows whether to send encapsulated message to UDP or TCP.
  - **HEADER CHECKSUM**: No check on corruption of user data
  - **SOURCE, DESTINATION IP ADDRESS**: Gateways only route based on destination address
  - **OPTIONS**:
    - Record route: record IP address of each gateway as datagram traverses internet
    - Record timestamps: records timestamps (and optionally IP addresses) of each gateway visited
    - Source routing: data field specifies IP addresses to use in routing; ignore routing tables; for tests
      - **strict**: datagram must follow the exact list of IP addresses
      - **loose**: datagram must pass through listed addresses, but can also go through others
  - **IP Routing** (Comer, Ch 8)

- **Routing**: choosing a path over which to send packets
- **Router**: computer choosing path
  - Characteristics of IP routing:
    - Network layer (layer 3)

- Every host and gateway in internet does routing (though gateways have routing tables that hosts normally do not use)
- Routing could be based on:
  - network load
  - datagram length
  - can you think of others?
- Internet routing today: based on estimated shortest path
- **Direct delivery**:
  - Used to deliver datagram from one to another machine attached to same network (same netid)
  - Used to deliver datagrams between any two .vt.edu hosts, since all share same netid
  - Exception: subnetting (described later)
- **Indirect delivery**:
  - Delivery via one or more gateways
  - How to tell which to use:
    - netid's equal: use direct delivery
    - netid's different: use indirect delivery
- Direct delivery algorithm:
  - To transfer IP datagram from host A to host B:
    1. A encapsulates DG in physical frame (layer 2)
    2. A maps B's IP address to B's physical address
    3. A uses network hardware (Ethernet, token ring) to directly deliver DG to B
  - Q: How is step 2 done?
    - A. using ARP
- **Address Resolution Protocol (ARP)** (Comer Ch. 5)
  - Q. How do you think a machine maps an Ip address to a physical address?
    - A1. Give each computer a table mapping IP to physical addr. Disadvantage: Inserting, removing a computer on network requires updating every host table!
    - A2. Encode physical addr into host field of IP address. Disadvantages:
802.3, 802.5 use 6 byte physical address (802.5)
can use 2 byte, however

Physical address on adaptor card must be
customer selectable -- not true for Ethernet.

A3. Dynamic binding (ARP) algorithm:
   a. Host A wants to map IP address I_B of host B to a
      phys addr. So it broadcasts on its network a
      special frame (ARP frame) with IP addr I_B.
   b. Host I_B responds with another ARP frame
      containing I_A's phys addr.

Q. Does ARP get executed every time a host sends a
datagram?
A. No:
   (1) Each host caches IP/phys addr pairs
   (2) The requesting ARP frame from A also contains
       I_A's IP and phy addr so that other hosts can add it
to their cache.

- Indirect delivery details (See Comer Fig. 8.2.)
  - IP routing is done by a table in each gateway
  - Table entry: (N,G)
  - If DG is destination to IP address with netid N,
    then directly deliver it to gateway with IP address G
  - Implies that all G's in table are directly attached to
    host/gateway storing table
  - Table size is proportional to number of networks, not
    number of hosts
  - Table often does not have an entry for every netid; in
    this case gateway has default gateway address to use
  - Hosts usually have an empty table, in which case they
    always use default address

- Complete algorithm: (based on Comer Fig. 8.3)
  Route_IP_Datagram(DG, RT) /* RT=routing table */
  D := destination IP address of DG
  DN := netid of D
  If DN matches any directly connected network then use
direct delivery to D
  else if DN matches N in any table entry (N,G) in RT
  then use direct delivery to G
  else use direct delivery to default gateway address

- Implications:
  - Routing is load independent
  - All datagrams with the same ip destination address
    take the same route unless the routing table is updated.
  - Routing does not share outgoing traffic over multiple
    links (unless routing table is updated)
  - Later in course we'll discuss how routing tables are
    updated. (There are core gateways that run a shortest
    path algorithm to update tables.)

- Subnetting (Comer, Ch. 16.6)
  - Q. Ethernet direct delivery requires broadcast. A class B
    address has 2^16 hosts on one network. Isn't this too
    many hosts to do a broadcast to? If I send a message
    from host X to Y at .vt.edu, would all the other .vt.edu hosts
    "see" the message?
    A. No. Bridges would limit physical broadcast range, yet
    preserve ability to directly broadcast to all campus
    machines according to the IP routing protocol.
    A. No. Subnetting allows .vt.edu to be partitioned into a
    little internet of its own!
    - .vt.edu uses a 6 bit subnetid and a 10 bit host id
  - See Stevens Fig. 5.3, Comer Fig. 16.3