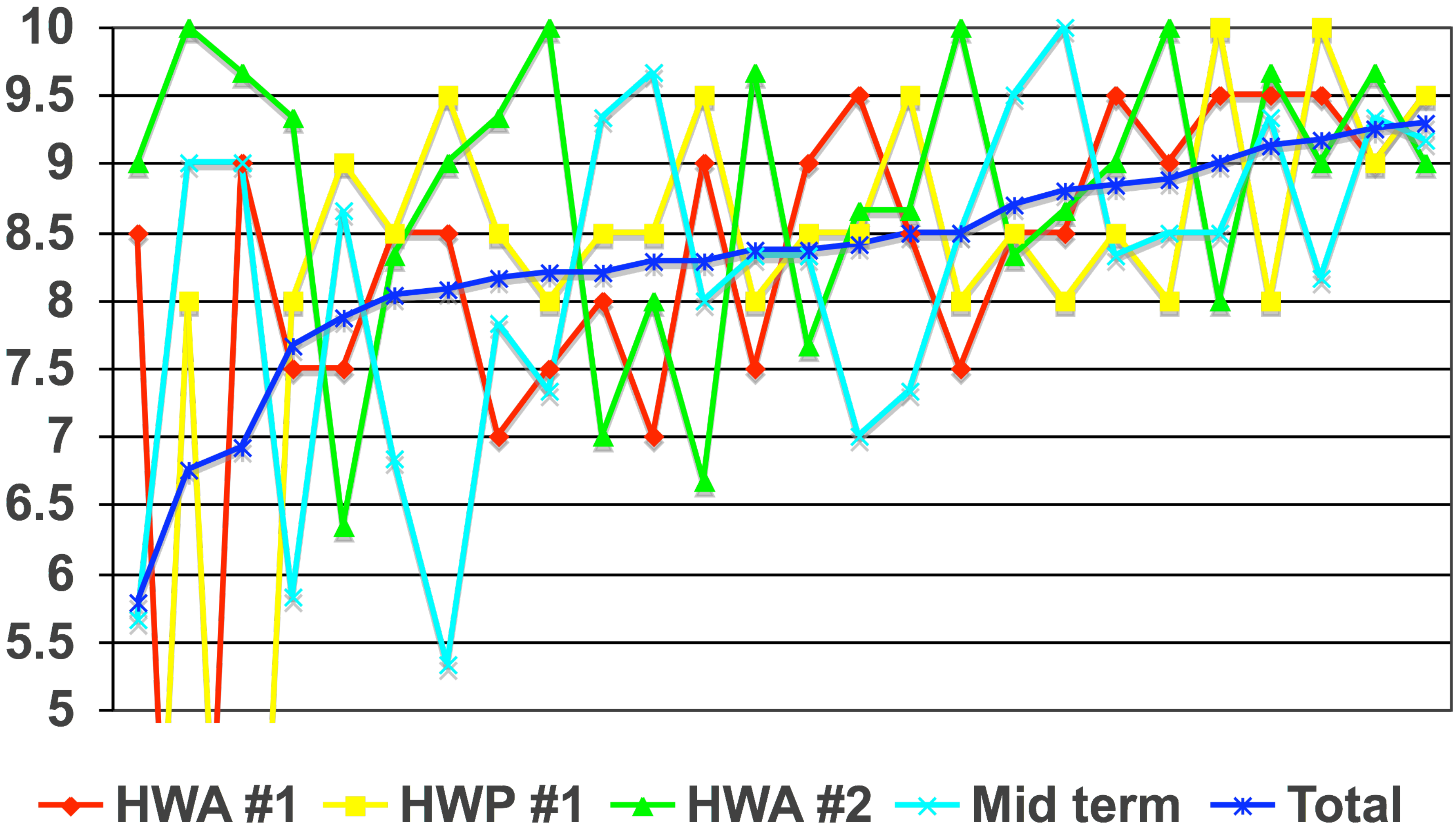
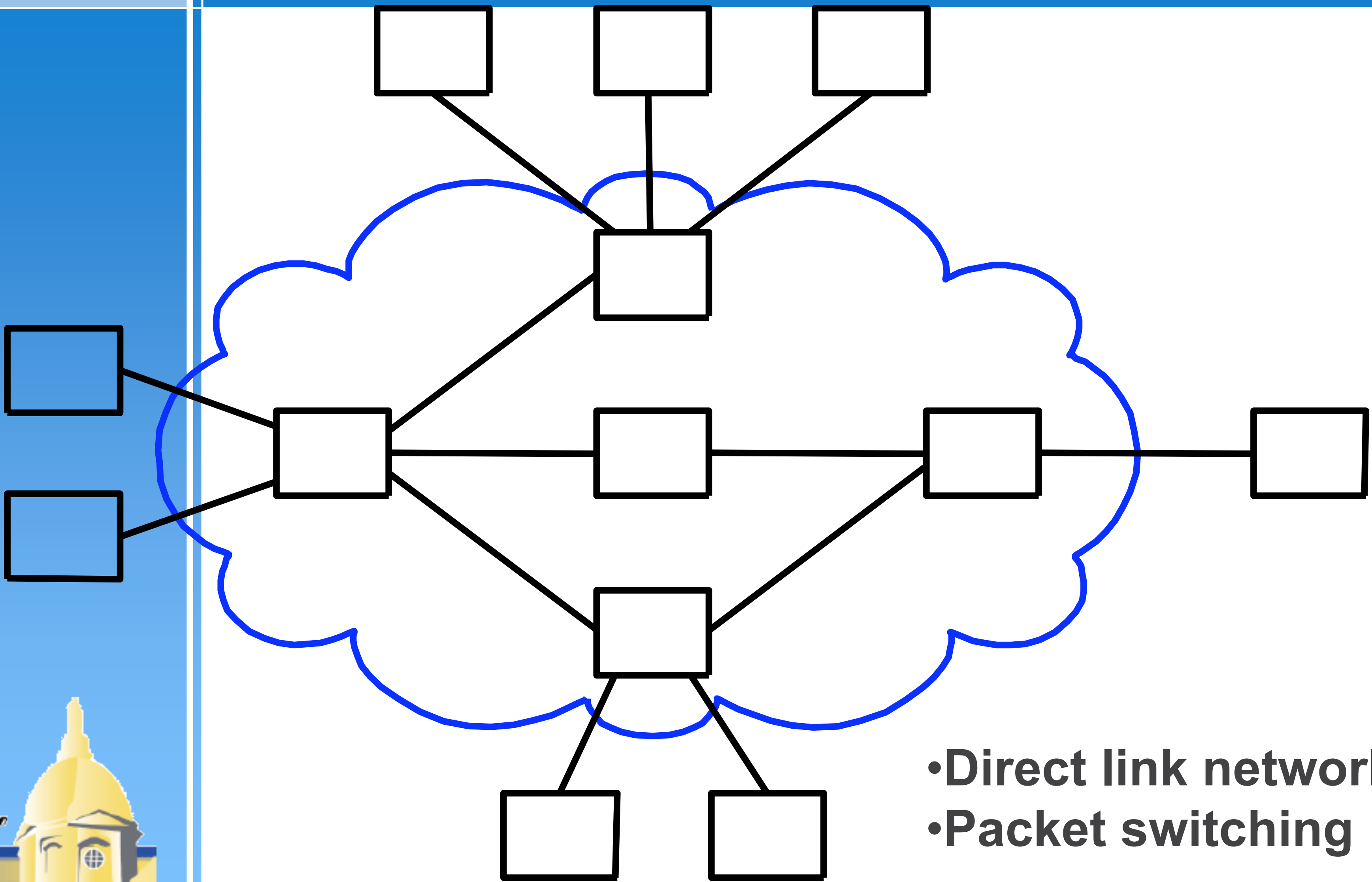
Mid term grade distribution

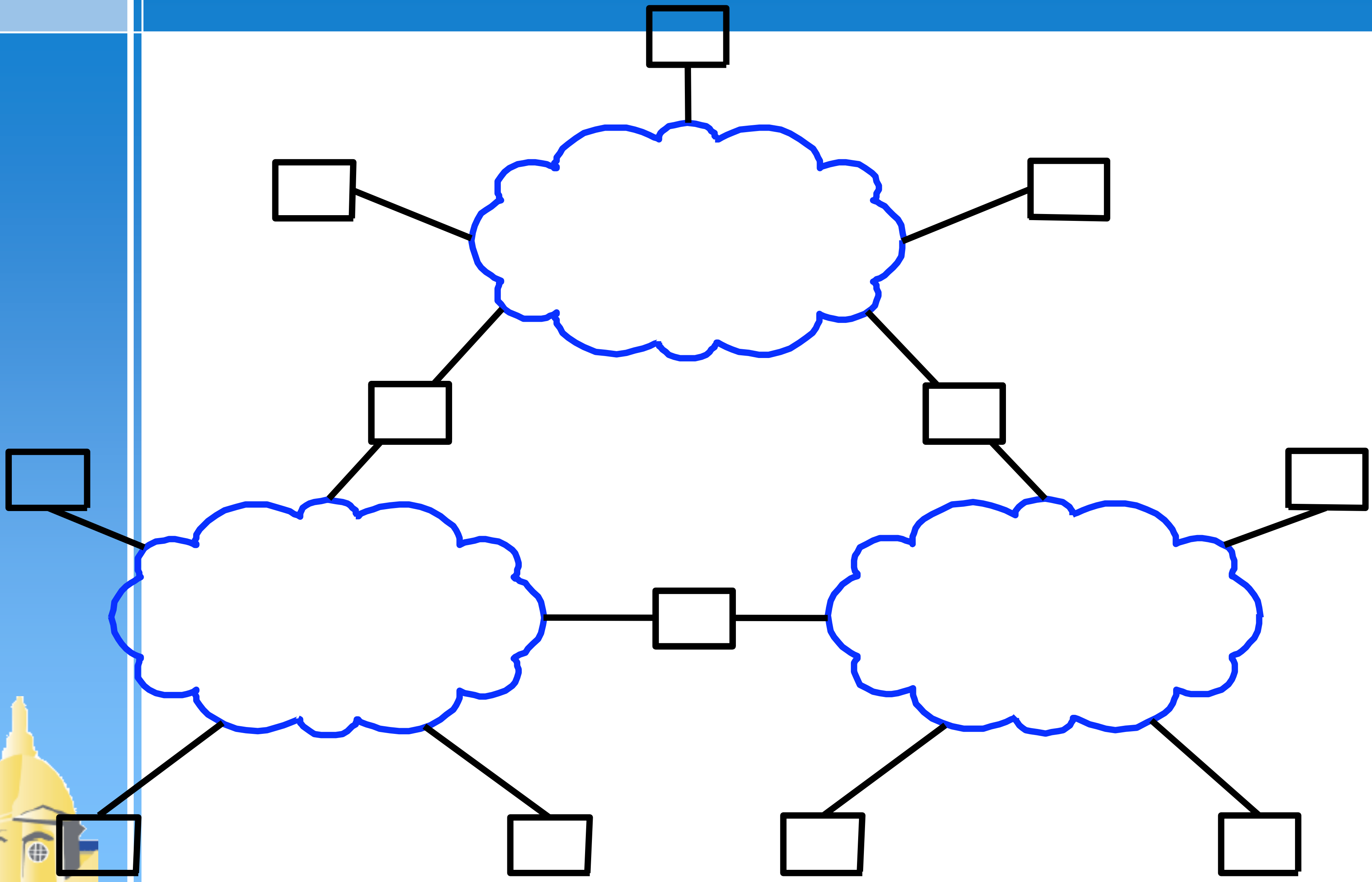


Recap from before spring break



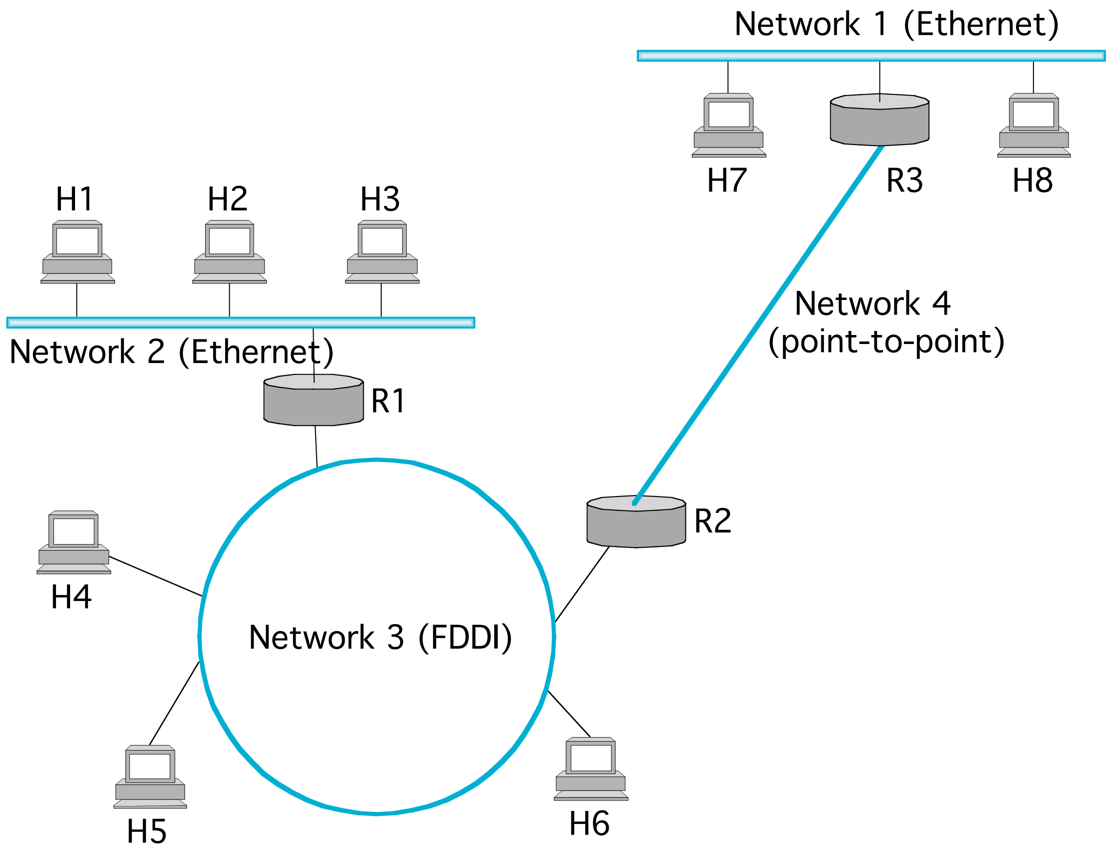
- **Direct link networks**
- **Packet switching**

Internetworking



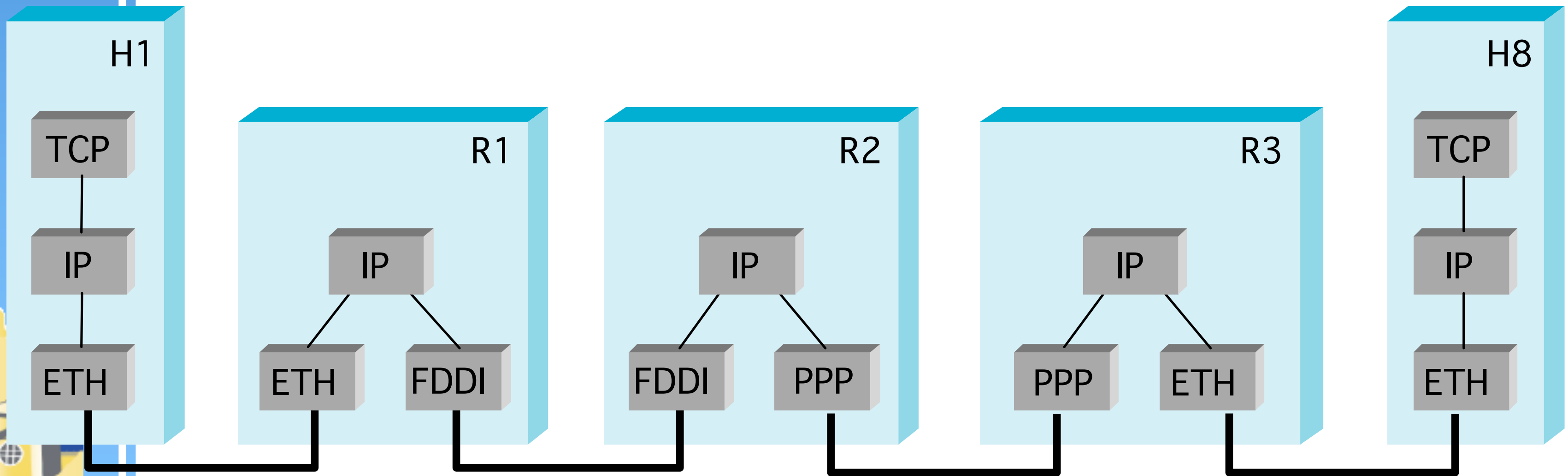
Internet

- collection of networks, each of which can be different types of networks



IP Internet

- Protocol Stack view
  - The packet goes through different network protocols as it traverses different types of networks
    - R's are routers and H is the hosts
  - Remember, bridges do not perform protocol translation and so wouldn't show up in this figure. E.g., there may be a bridge between R1 and R2



Service Model

- Connectionless (datagram-based)
  - Each datagram carries the destination address
- Best-effort delivery (unreliable service)
  - packets may be lost
  - packets can be delivered out of order
  - duplicate copies of a packet can be delivered
  - packets can be delayed for a long time

Issues

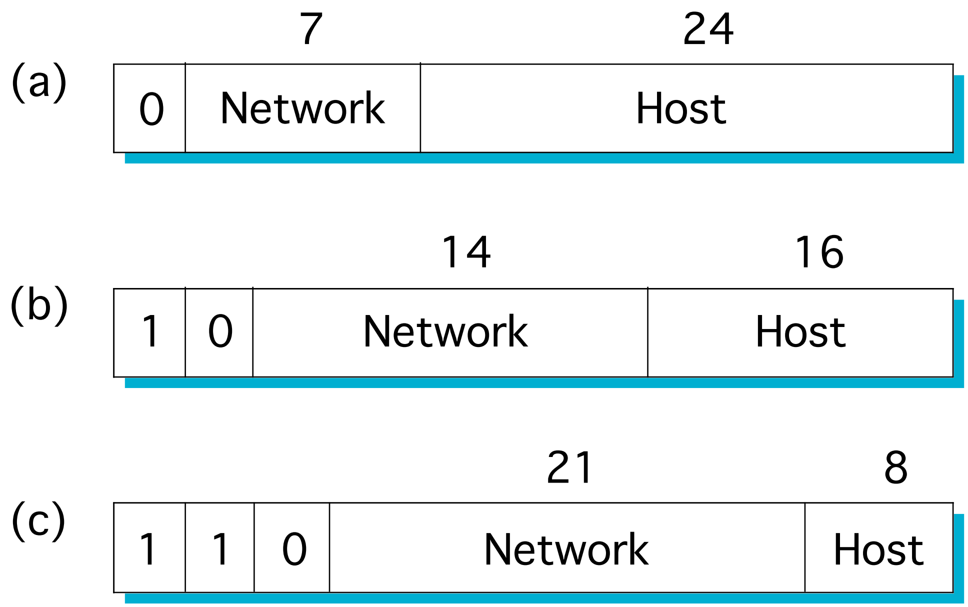
1. Global naming (IP addresses) and mechanisms to translate to human usable form (e.g. www.nd.edu) (DNS)
2. Fragmentation - not all networks can deal with a given packet size
3. Routing, mapping ip address to physical address etc..

Issue 1: Global Addresses

- As we connect different types of networks, need a way to address each host, independent of the network type (ethernet, FDDI etc.)
  - Darwin.cc.nd.edu's address is 08:00:20:7e:16:cc
- Properties (www.iana.org/assignments/ipv4-address-space)
  - globally unique
  - hierarchical: network + host, so that we can use it to route
  - http://www.caida.org/outreach/resources/learn/ipv4space/

Issue 1: ipv4 address classes

- Dot Notation to describe the 32 bit address:
  - 10.3.2.4
  - 128.96.33.81
  - 192.12.69.77



- Class D (224 - 239) Multicast
- Class E (240 – 247) Experimental

Darwin.cc.nd.edu

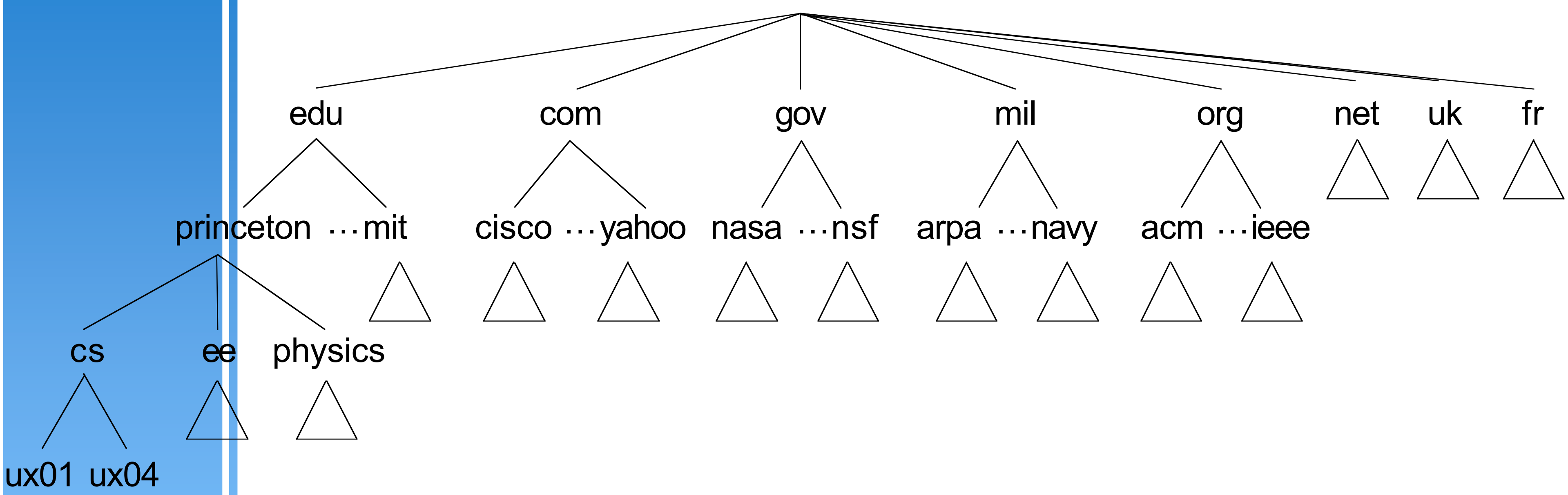
- Darwin.cc.nd.edu = 129.74.250.114
  - 10000001.01001010.11111010.01110010
- Class B address, ND owns 129.74.X.X

Domain Name Service (DNS)

- Provides Internet domain name to IP address translation
  - Domain name translation (nd.edu)
  - Hostname translation (wizard.cse.nd.edu)
  - Service location (MX records, mail service for ND)
- Use nslookup command to interact with DNS

Domain Naming System Hierarchy

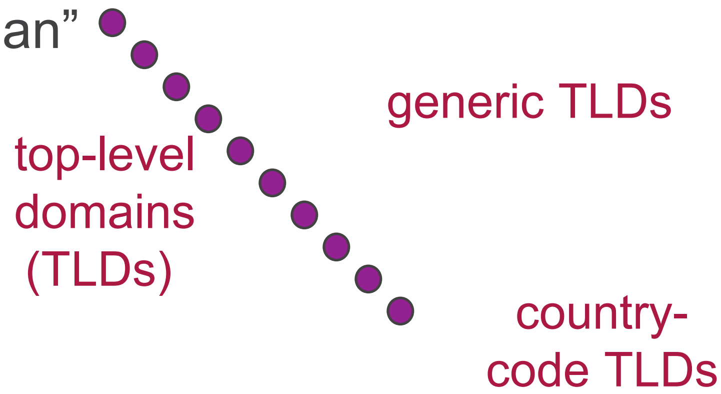
- DNS is hierarchical

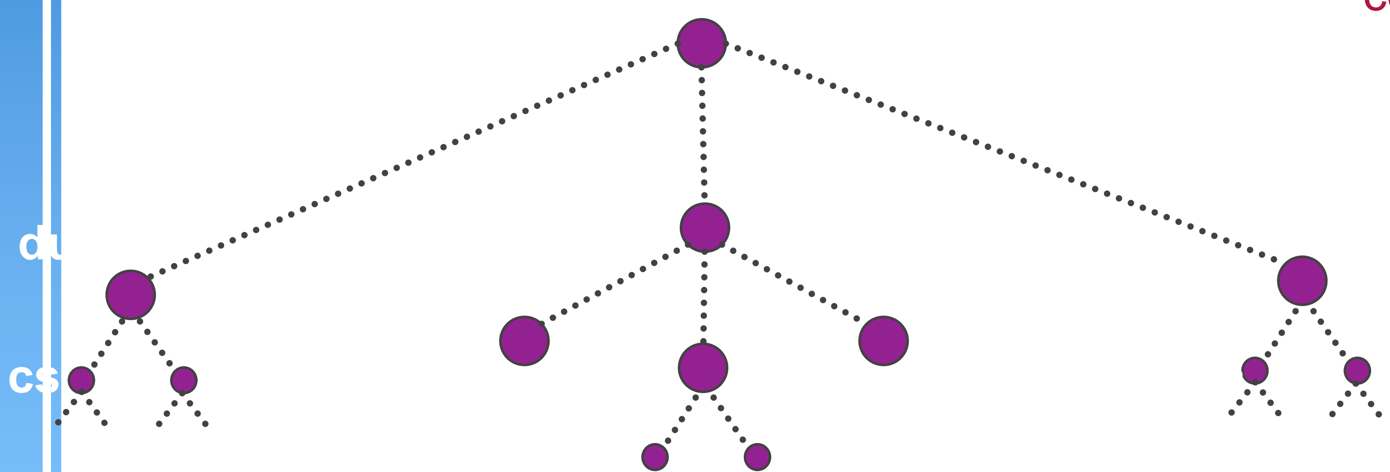


DNS hierarchy

DNS name space is *hierarchical*:

- fully qualified names are "little endian"
    - scalability
    - decentralized administration
    - domains are naming *contexts*





Source: Jeff Chase

DNS Name Server Hierarchy

DNS servers are organized into a hierarchy that mirrors the name space.

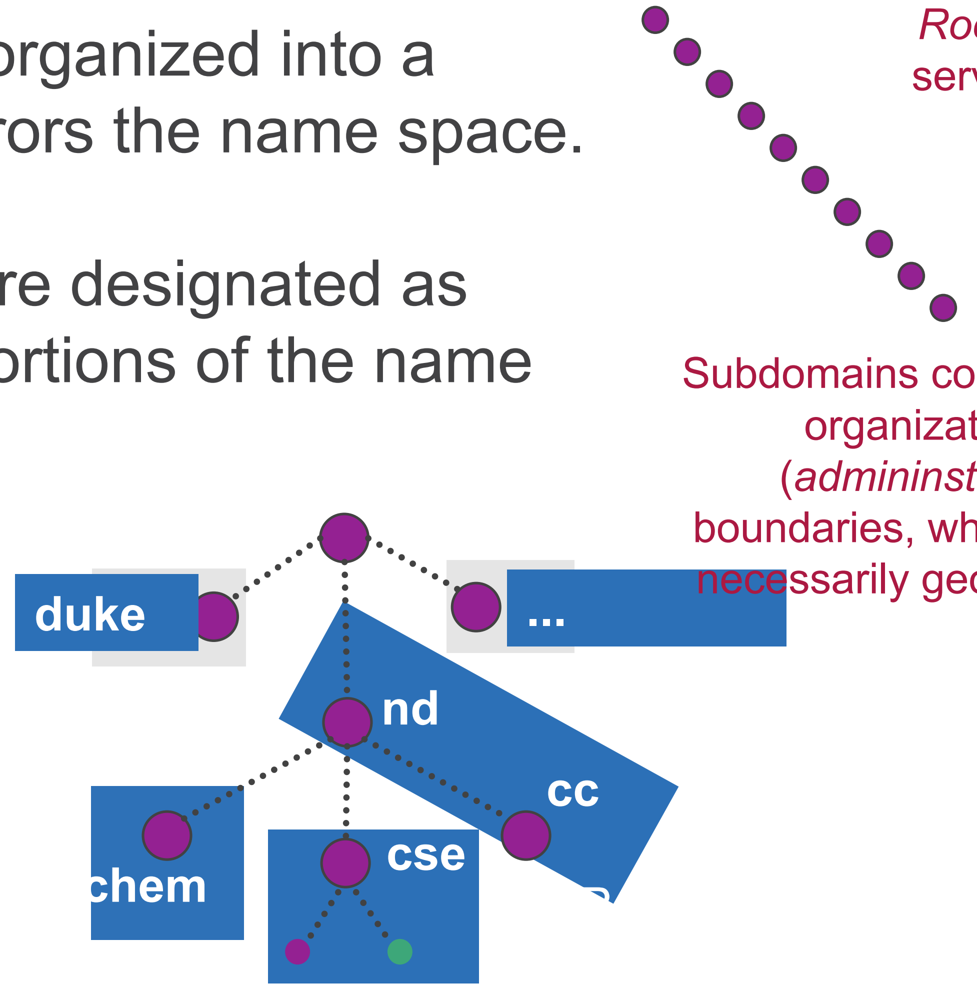
Specific servers are designated as *authoritative* for portions of the name space.

Root servers list servers for every TLD.

Subdomains correspond to organizational (*admininstrative*) boundaries, which are not necessarily geographical.

Servers may delegate management of *subdomains* to child name servers.

Parents refer subdomain queries to their children.

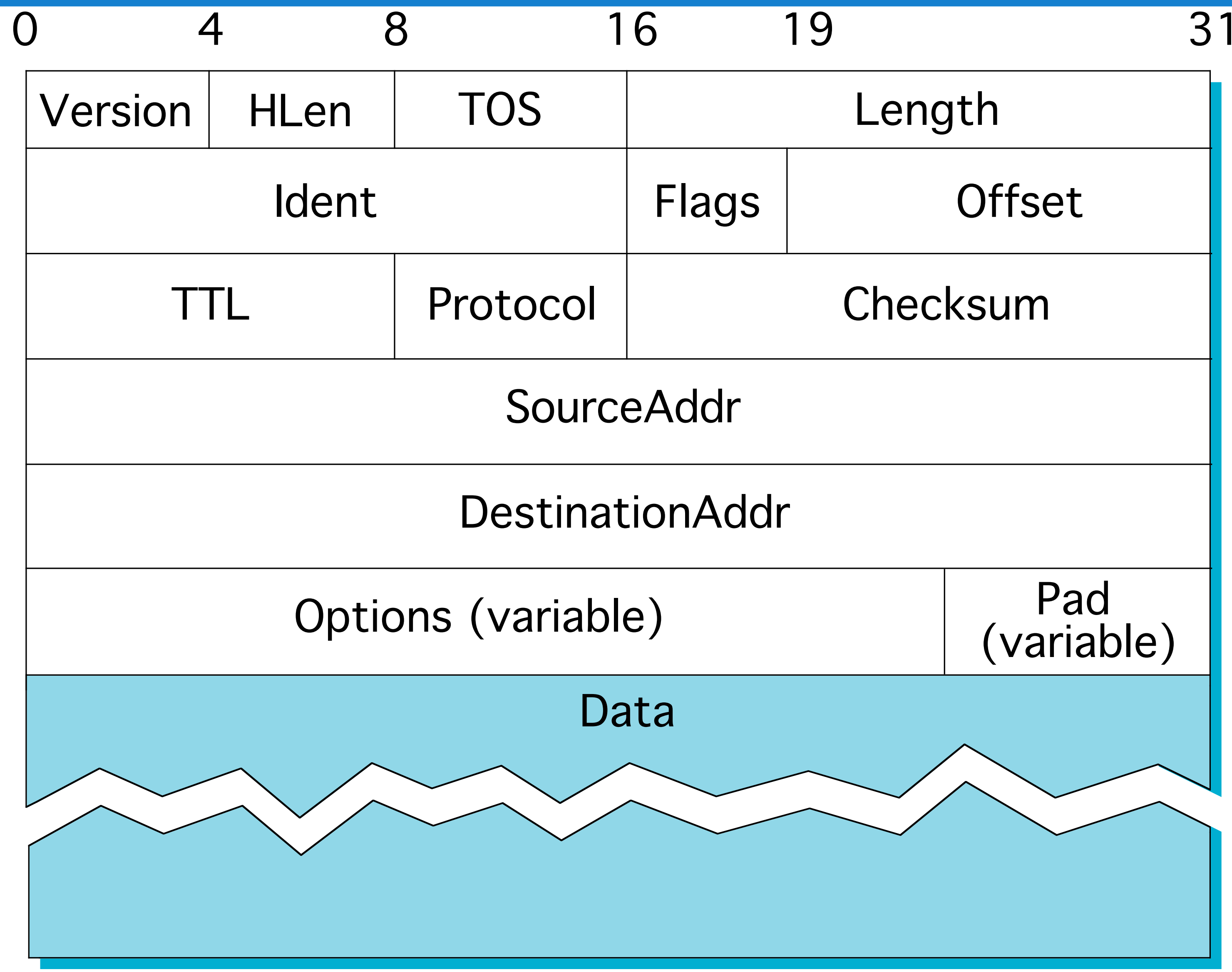


Source: Jeff Chase

Multicast DNS (Apple Rendezvous)

- Situations where there are no DNS servers (local intranet - for example home users who want to name the machines without dealing with the complexities of maintaining DNS servers) or where you may not know where the DNS servers are
- Create names in the .local domain. For example, my laptop can be called surendar.local.
- Zeroconf initiative

IP v4 Datagram format



IP v6 format

- Developed so that we can address more than 2^32 hosts (ipv4)
- http://ipv6.internet2.edu/boston/presentations/09-ipv6-under-the-hood.ppt
  - Version (4 bits) – only field to keep same position and name
  - Class (8 bits) – was Type of Service (TOS), renamed
  - Flow Label (20 bits) – new field
  - Payload Length (16 bits) – length of data, slightly different from total length
  - Next Header (8 bits) – type of the next header, new idea
  - Hop Limit (8 bits) – was time-to-live, renamed
  - Source address (128 bits)
  - Destination address (128 bits)

IPv4 and IPv6 headers

Version	Class	Flow Label	
Payload Length		Next Header	Hop Limit
Source Address			
Destination Address			

Version	IHL	Type of Service	Total Length	
Identification			Flags	Fragment Offset
Time-to-live		Protocol	Header Checksum	
Source Address				
Destination Address				
Options				Padding

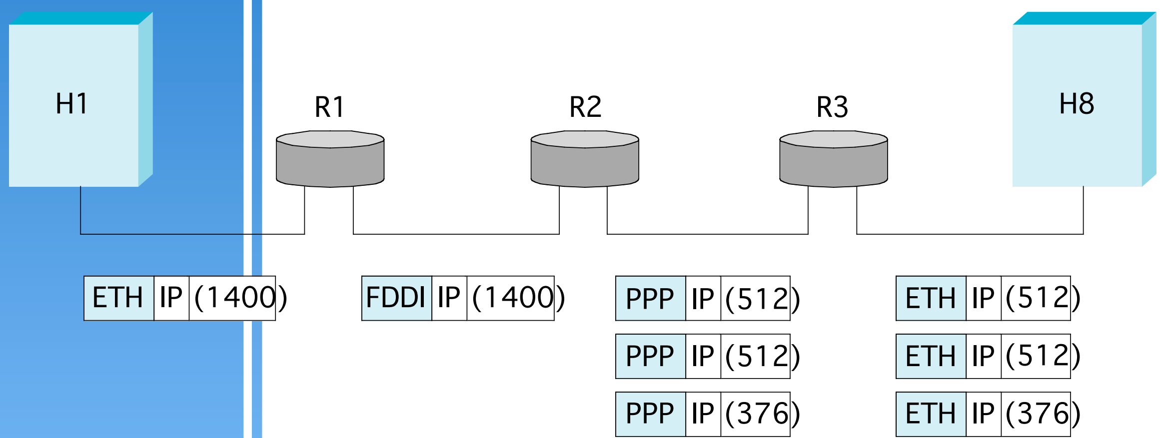
Basic Headers - IPV6

- Simplifications
  - Fixed length of all fields, not like old options field – IHL, or header length irrelevant
  - Remove Header Checksum – rely on checksums at other layers
  - No hop-by-hop fragmentation – fragment offset irrelevant
    - MTU discovery (IPv4 also support Path MTU discovery)
  - Add extension headers – next header type (sort of a protocol type, or replacement for options)
  - Basic Principle: Routers along the way should do minimal processing

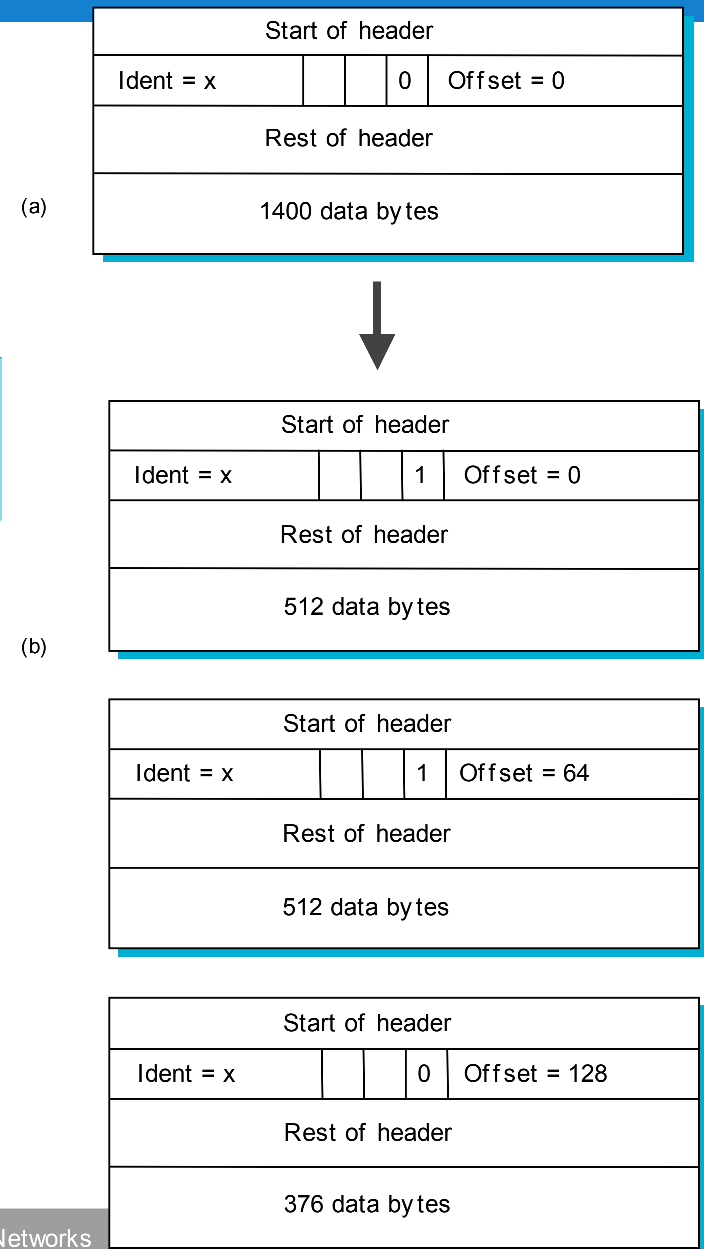
Issue 2: Fragmentation and Reassembly

- Cannot expect all networks to deal with the same packet size, can choose the absolute smallest - but that would mean poor performance for all networks
  - Each network has some MTU (maximum transmission unit)
- Design decisions
  - fragment when necessary (MTU < Datagram)
  - re-fragmentation is possible
  - fragments are self-contained datagrams
  - use CS-PDU (not cells) for ATM
  - delay reassembly until destination host
  - do not recover from lost fragments
  - try to avoid fragmentation at source host

Example



H1
R1
R2
R3
H8
ETH IP (1400)
FDDI IP (1400)
PPP IP (512)
PPP IP (512)
PPP IP (376)
ETH IP (512)
ETH IP (512)
ETH IP (376)



(a)
Start of header
Ident = x
0
Offset = 0
Rest of header
1400 data bytes
(b)
Start of header
Ident = x
1
Offset = 0
Rest of header
512 data bytes
Start of header
Ident = x
1
Offset = 64
Rest of header
512 data bytes
Start of header
Ident = x
0
Offset = 128
Rest of header
376 data bytes

Gateway13 example

- Ifconfig eth0 in gateway13.cse.nd.edu

Link encap:Ethernet HWaddr 00:07:E9:3C:8F:80

inet addr:129.74.154.198 Bcast:129.74.155.255
 Mask:255.255.252.0

inet6 addr: fe80::207:e9ff:fe3c:8f80/64 Scope:Link

UP BROADCAST RUNNING MULTICAST
 MTU:9000 Metric:1

Issue 3: Datagram Forwarding

- Strategy
  - every datagram contains destination's address
  - if connected to destination network, then forward to host
  - if not directly connected, then forward to some router
  - forwarding table maps network number into next hop
  - each host has a default router
  - each router maintains a forwarding table

Example (R2)

Network number	Next
1	R3
2	R1
3	Interface 1
4	Interface 0

Address Translation

- Map IP addresses into physical addresses
  - destination host
  - next hop router
- Techniques
  - encode physical address in host part of IP address
  - table-based
- Mechanism to map IP to physical address: ARP
  - table of IP to physical address bindings
  - broadcast request if IP address not in table
  - target machine responds with its physical address
  - table entries are discarded if not refreshed

ARP Details

- Request Format
  - HardwareType: type of physical network (e.g., Ethernet)
  - ProtocolType: type of higher layer protocol (e.g., IP)
  - HLEN & PLEN: length of physical and protocol addresses
  - Operation: request or response
  - Source/Target-Physical/Protocol addresses
- Notes
  - table entries timeout in about 10 minutes
  - update table with source when you are the target
  - update table if already have an entry
  - do not refresh table entries upon reference

ARP Packet Format

0	8	16	31
Hardware type = 1		ProtocolType = 0x0800	
HLen = 48	PLen = 32	Operation	
SourceHardwareAddr (bytes 0 — 3)			
SourceHardwareAddr (bytes 4 — 5)		SourceProtocolAddr (bytes 0 — 1)	
SourceProtocolAddr (bytes 2 — 3)		TargetHardwareAddr (bytes 0 — 1)	
TargetHardwareAddr (bytes 2 — 5)			
TargetProtocolAddr (bytes 0 — 3)			

Sample arp table in darwin.cc.nd.edu

- arp -a

Net to Media Table: IPv4

```
Device    IP Address                    Mask        Flags    Phys Addr
___ __________ ________- ___- ________-
hme0    eafs-e06.gw.nd.edu    255.255.255.255          00:d0:c0:d3:aa:40
hme0    bind.nd.edu           255.255.255.255          08:00:20:8a:5f:cf
hme0    honcho-jr.cc.nd.edu   255.255.255.255          00:b0:d0:82:83:7f
hme0    mail-vip.cc.nd.edu    255.255.255.255          02:e0:52:0c:56:c4
hme0    john.helios.nd.edu    255.255.255.255          08:00:20:85:db:c4
hme0    casper.helios.nd.edu  255.255.255.255          08:00:20:b1:f8:e1
hme0    pinky.helios.nd.edu   255.255.255.255          08:00:20:a9:88:30
```

ARP problems

- ARP trusts any response - no authentication method
  - Works great at home, how about Notre Dame
- Replies which do not correspond to requests are allowed to update cache in many instances
- New information must supercede old info

Internet Control Message Protocol (ICMP)

- Mechanisms to notify of errors (not mandatory)
  - Echo (ping)
  - Redirect (from router to source host)
  - Destination unreachable (protocol, port, or host)
  - TTL exceeded (so datagrams don't cycle forever)
  - Checksum failed
  - Reassembly failed
  - Cannot fragment