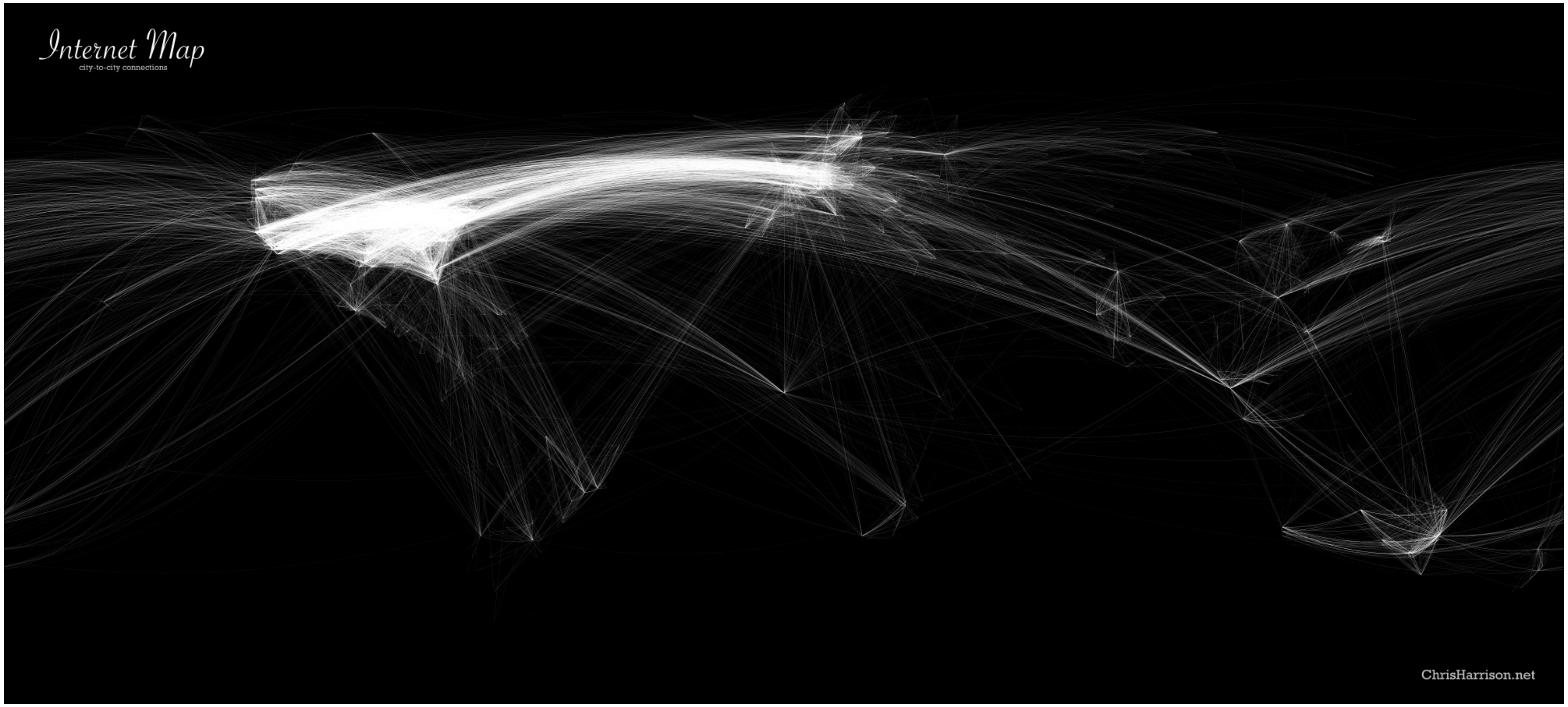
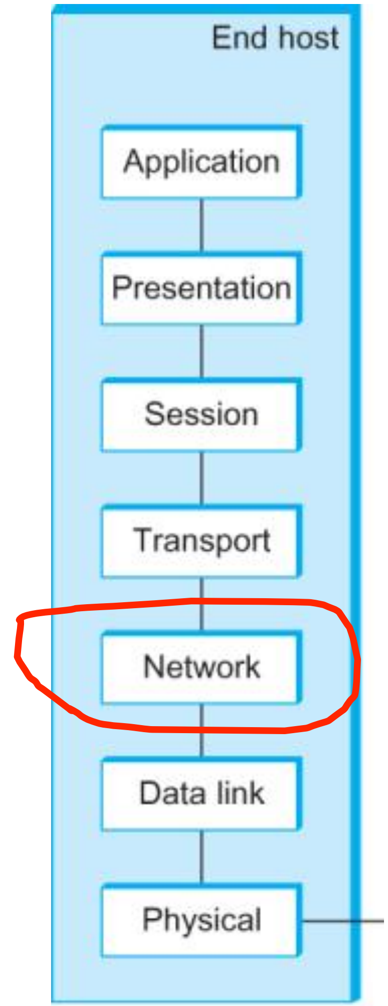
Basic Internetworking (IP)

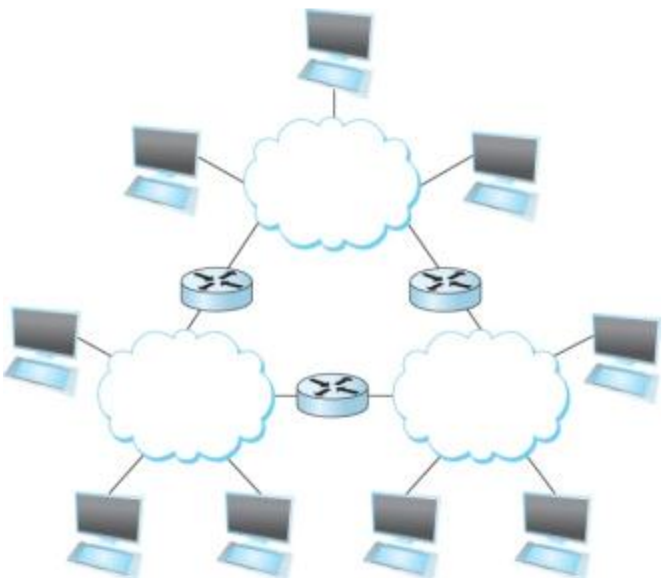


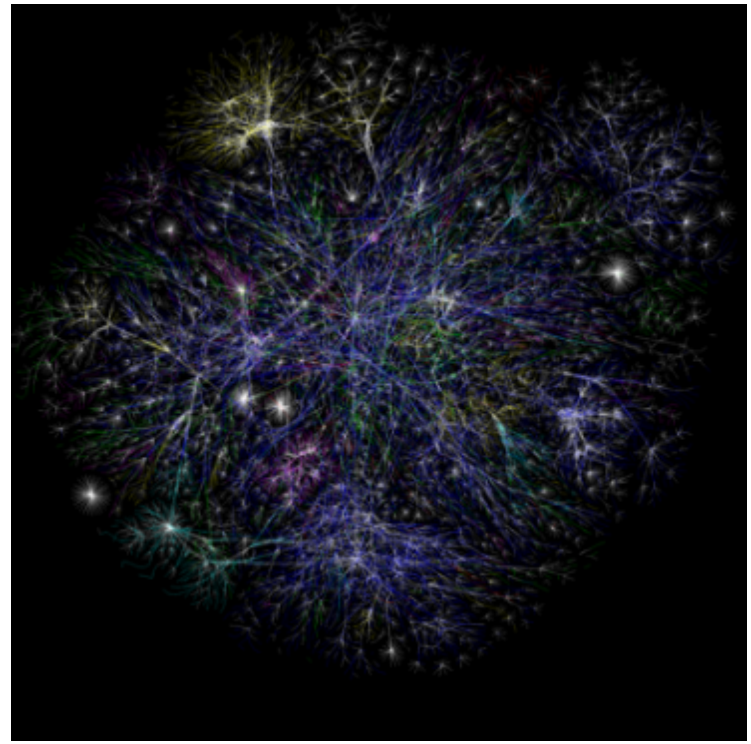
Overview

- Internetworking
  - Service model
  - Internet protocol (IP)
    - History
    - Packet format
    - Fragmentation
  - Global addressing
  - Discovering link-layer addresses
  - Assigning IP addresses



internet vs. Internet



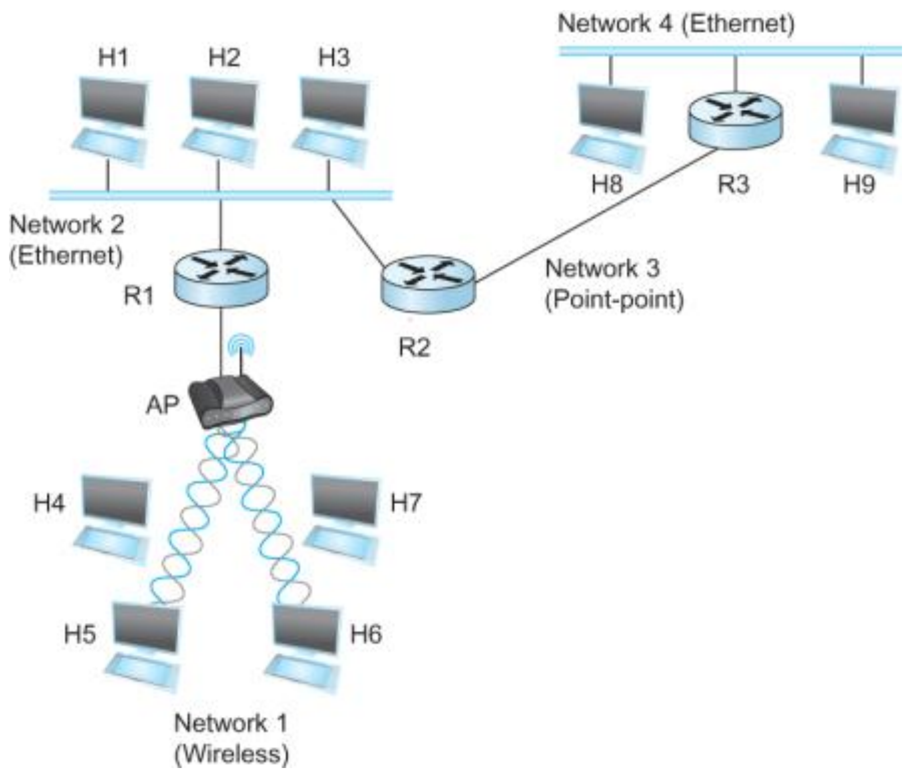


internetwork / internet:
A set of independent networks interconnected, could be completely walled off from world.

Internet
Global public network consisting of interconnected networks running TCP/IP.

Network of networks

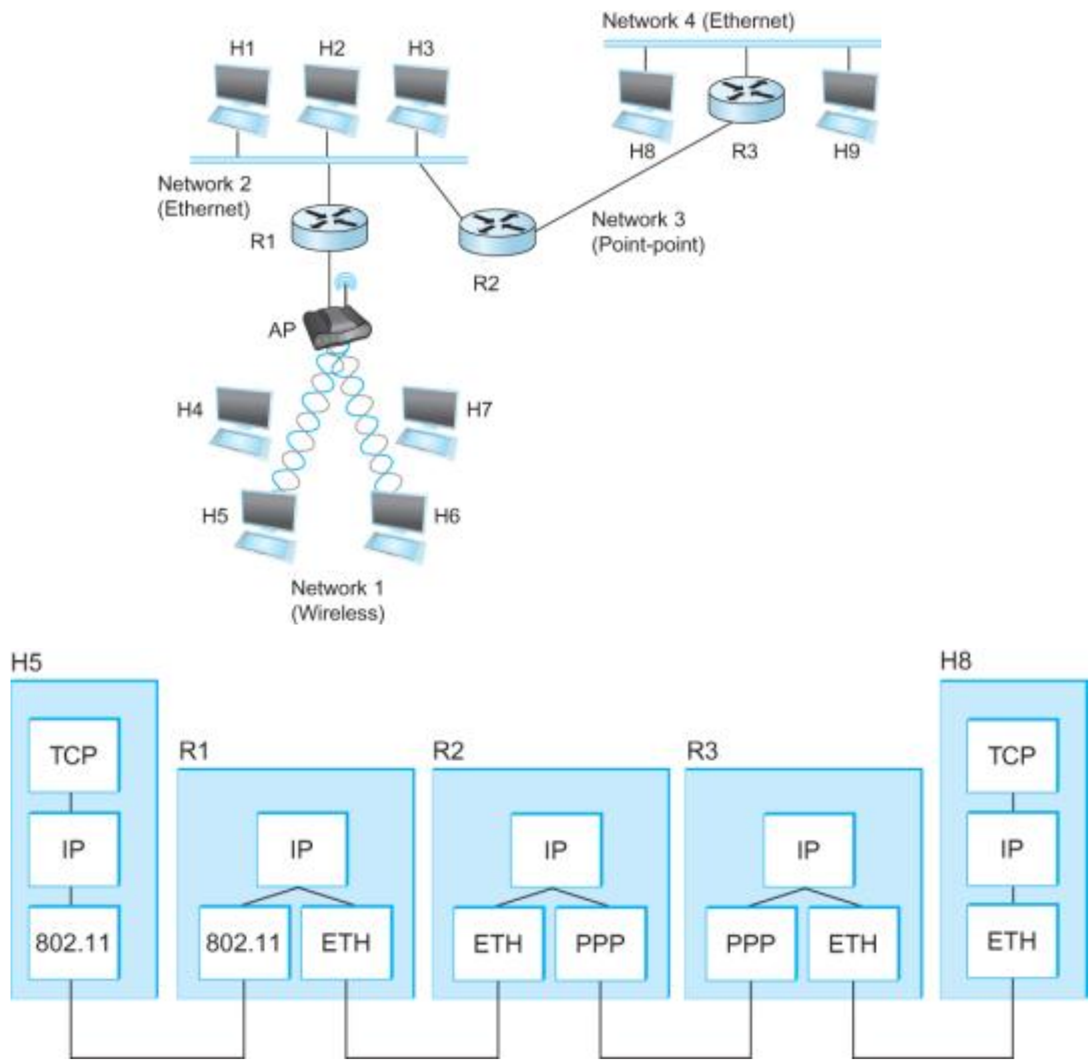
- Connect heterogeneous networks
  - Provide host-to-host packet delivery service



Application layer	Application gateway
Transport layer	Transport gateway
Network layer	Router
Data link layer	Bridge, switch
Physical layer	Repeater, hub

A simple internetwork, H1-9 are hosts, R1-R3 are routers

Communication from H5 to H8



Internet Protocol (IP)

- Packet delivery model
  - Connectionless
  - Best-effort (unreliable)
    - Packets may be lost
    - Packets may arrive out of order
    - Duplicate packets may occur
    - Packet may get delayed
- Global addressing scheme
  - How do we identify hosts on the network?

IP history and goals

- Internet Protocol (IP)
  - 1974 Cerf and Kahn propose common layer hiding network differences
  - Eventually split into TCP and IP
  - IP foundation of the modern Internet
  - Awarded 2004 Turing Award

A Protocol for Packet Network Intercommunication

VINTON G. CERF AND ROBERT E. KAHN,
MEMBER, IEEE

Abstract — A protocol that supports the sharing of resources that exist in different packet switching networks is presented. The protocol provides for variation in individual network packet sizes, transmission failures, sequencing, flow control, end-to-end error checking, and the creation and destruction of logical process-to-process connections. Some implementation issues are considered, and problems such as internetwork routing, accounting, and timeouts are exposed.

of one or more *packet switches*, and a collection of communication media that interconnect the packet switches. Within each HOST, we assume that there exist *processes* which must communicate with processes in their own or other HOSTS. Any current definition of a process will be adequate for our

IP history and goals

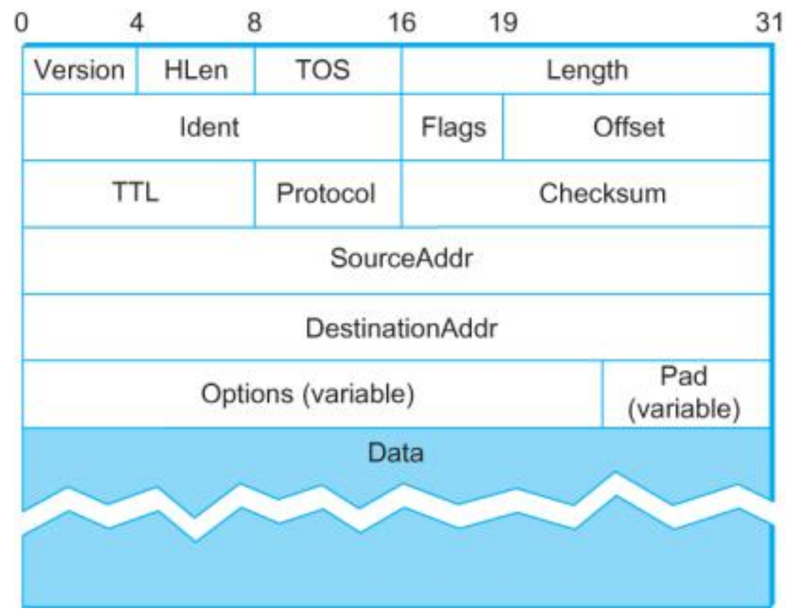
- Connect existing networks
  - Multiplex existing links such as radio networks
- Motivating application
  - Remote login to servers
    - Inherently bursty traffic, long silence periods
- Robust to failures
  - Survive equipment failure or attack
  - Traffic routes around trouble

IP history and goals

- Support multiple types of services
  - Differing requirements for speed, latency, reliability
- Heterogeneous networks
  - Minimal assumptions about underlying network
- Distributed management of resources
  - Node managed by different institutions
- Cost effective
  - Packet switched, share links via multiplexing

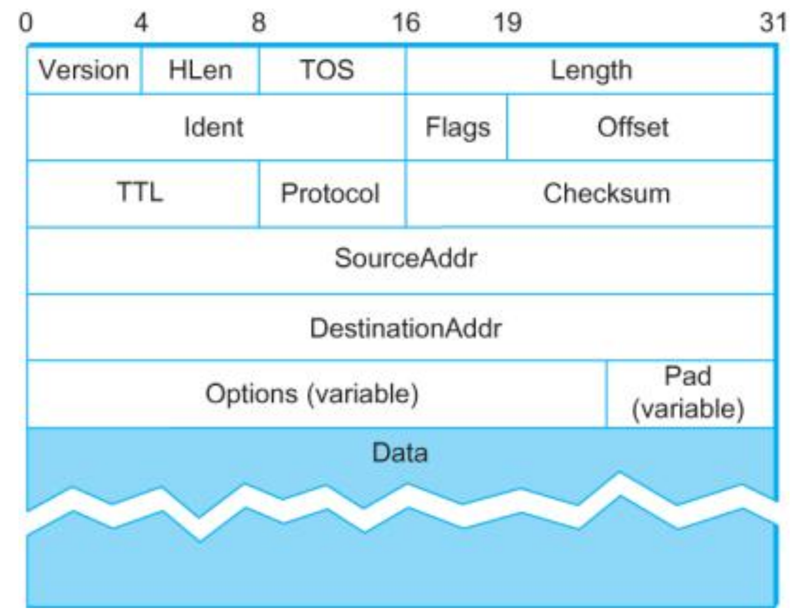
IP packet format

- Version
  - "4" IPv4, "6" IPv6
- HLen
  - # of 32-bit words in header
  - "5" for typical 20-byte IPv4 header
- TOS
  - Type of service
  - Allows for quality of service
- Length
  - Total length, max size 65535 bytes
  - Links may have small limits
- Ident / Flags / Offset
  - Used when packets are split up



IP packet format

- TTL
  - Time=to-live
  - Identify packets stuck in loop
- Protocol
  - Used to demux higher-level protocol
  - e.g. "6" Transmission Control Protocol (TCP), "17" User Datagram Protocol (UDP)
- Checksum
  - One's complement IP checksum algorithm
  - Not strong protection, but cheap to calculate



Time-to-live

- TTL
  - Original plan: TTL was time-to-live in seconds
  - Routers today just decrement by 1
  - Sender sets default value (e.g. 64)
  - If TTL reaches 0:
    - Host sends "time exceeded" message to source

Traceroute

- Traceroute tool exploits TTL
  - Send packets with TTL=1, 2, 3, ...
  - Record source of "time exceeded" message
- Windows
  - "tracert www.bbc.co.uk"
- Unix / MacOS
  - "traceroute www.bbc.co.uk"

Traceroute, London to tech

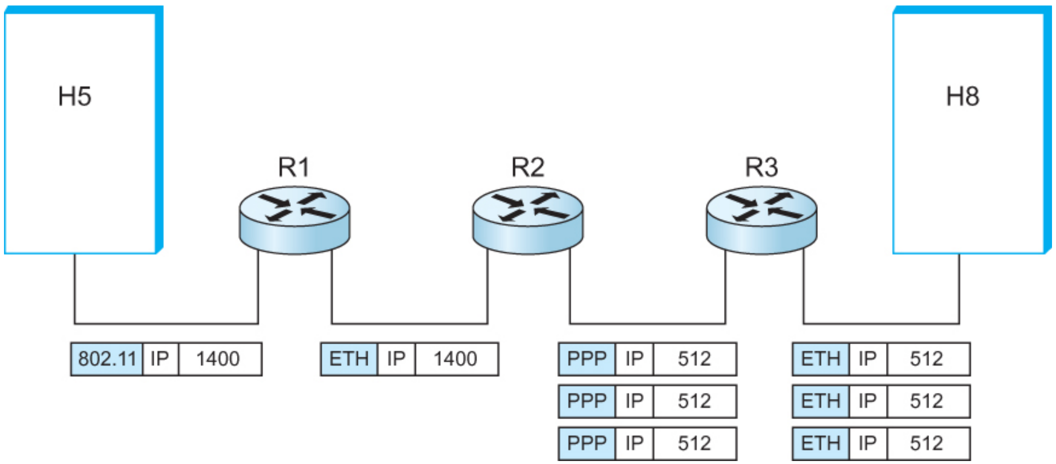
```
kvertanen@li264-110:~$ traceroute cs.mtech.edu
traceroute to cs.mtech.edu (150.131.202.136), 30 hops max, 60 byte packets
 1  212.111.33.229 (212.111.33.229)  0.750 ms  0.793 ms  0.818 ms
 2  212.111.33.233 (212.111.33.233)  0.449 ms  0.486 ms  0.533 ms
 3  te3-1-border76-01.lon2.telecity.net (217.20.44.217)  0.705 ms * *
 4  * * *
 5  IPP-Tiscali.lon1.telecity.net (213.200.79.133)  0.736 ms  0.723 ms  0.709 ms
 6  xe-8-1-0.lon11.ip4.tinet.net (89.149.186.21)  0.742 ms xe-3-2-0.lon11.ip4.tinet.net (89.149.187.126)  0.721 ms
xe-7-1-0.lon11.ip4.tinet.net (89.149.185.70)  0.747 ms
 7  te7-6.mpd02.lon01.atlas.cogentco.com (130.117.15.49)  1.511 ms  0.949 ms  1.617 ms
 8  te3-8.ccr01.lon01.atlas.cogentco.com (130.117.1.133)  1.335 ms  1.352 ms  1.390 ms
 9  te0-1-0-7.mpd21.ord01.atlas.cogentco.com (154.54.43.177)  92.391 ms
te0-3-0-7.mpd21.ord01.atlas.cogentco.com (154.54.24.54)  92.401 ms
te0-0-0-2.mpd21.ord01.atlas.cogentco.com (154.54.45.245)  92.413 ms
10  te3-1.ccr01.ord04.atlas.cogentco.com (154.54.24.178)  92.689 ms  92.803 ms  92.926 ms
11  te4-3.ccr01.msp01.atlas.cogentco.com (154.54.3.62)  106.209 ms  106.227 ms  106.212 ms
12  38.104.196.154 (38.104.196.154)  144.114 ms  144.199 ms  144.052 ms
13  192.73.48.129 (192.73.48.129)  144.961 ms  144.922 ms  144.994 ms
14  * * *
```

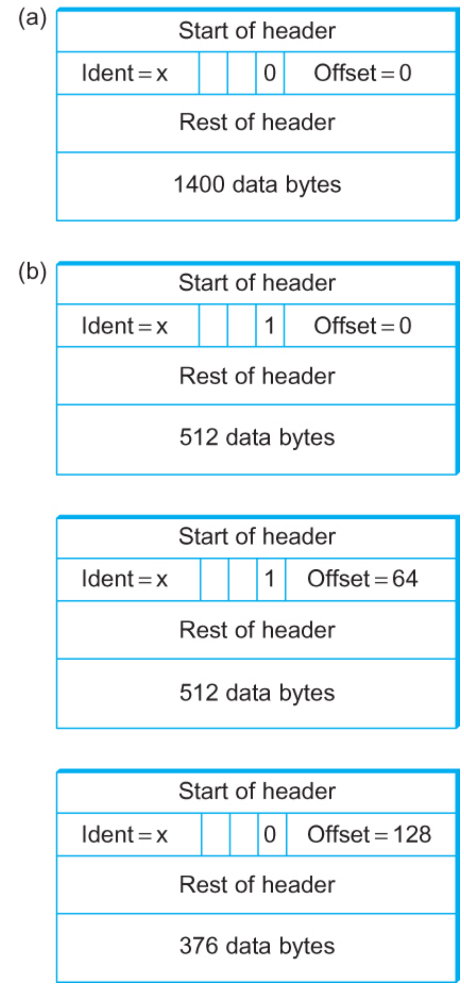
http://en.dnstools.ch/visual-traceroute.html

Fragmentation

- Different networks support different packet sizes
  - Maximum Transmission Unit (MTU)
  - e.g. Ethernet 1500 bytes, 802.11 2272 bytes
- Occurs at router
  - If inbound datagram > MTU of outbound network
  - Split into fragments
    - All fragments have same Ident field
    - Each is self-contained datagram

Fragmentation and reassembly



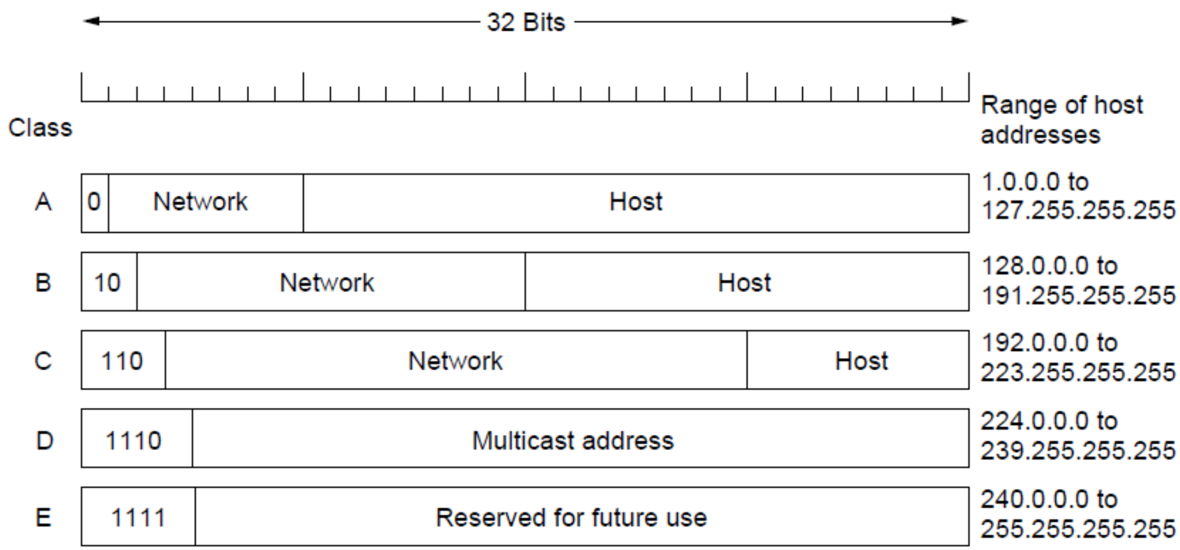


- Reassembly can be done independent of order of arrival
- Fragments may also be fragmented
- No attempt to recover if fragment missing
- Hosts can do MTU discovery
  - Probe message to determine max packet size

Global addressing

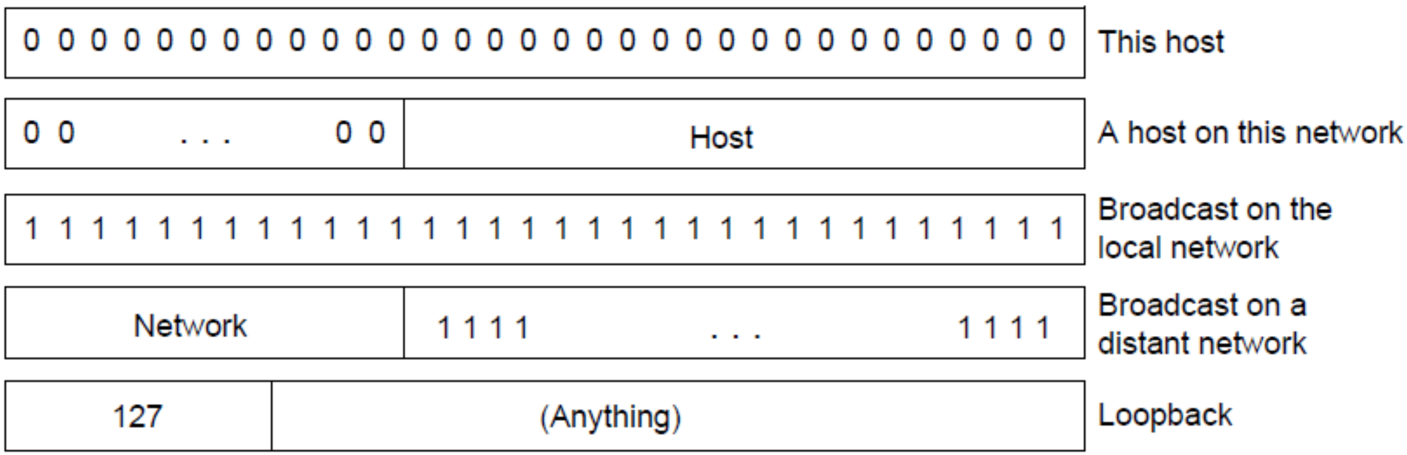
- IP service model
  - Assumes global addresses
- Why not 48-bit MAC address?
  - flat structure, no hierarchy
  - e.g. 01:23:45:67:89:ab
- IP addresses
  - IPv4 32 bits
  - network part
  - host part
  - e.g. 10.33.73.165

IPv4 address format



- Classful addressing (before 1993):
  - Class A: 128 networks with 16 million hosts
  - Class B: 16,384 networks with 65,536 hosts
  - Class C: 2 million networks, 256 hosts

Special IP addresses

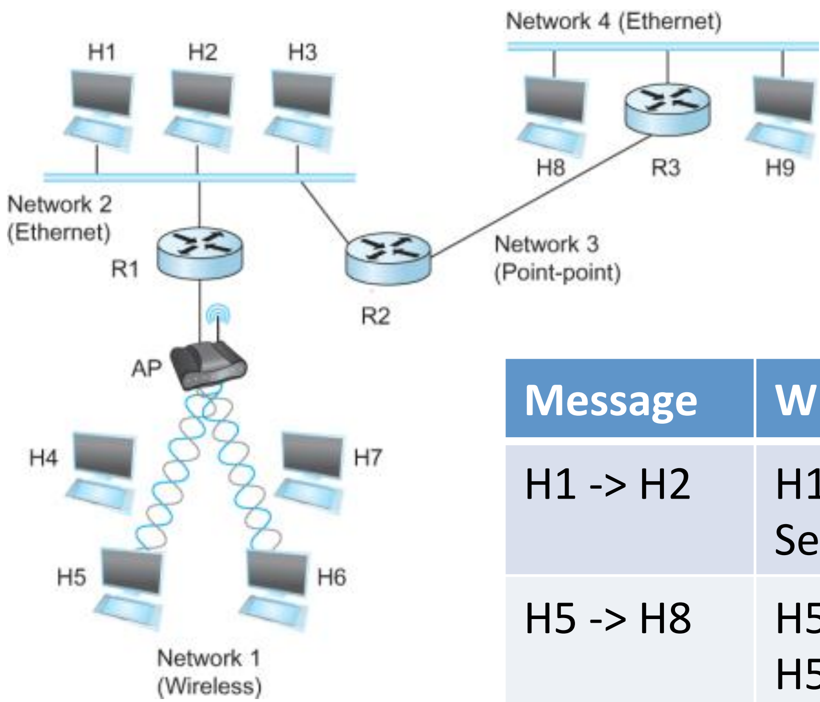


- Dot notation, each byte converted to decimal
  - 1000 0000 1101 0000 0000 0010 1001 0111
  - 128.208.2.151

IP datagram forwarding

- Why (network + host) address help?
  - Routers have a forwarding table
    - Network number -> next hop
  - Without hierarchy:
    - Tables in routers would be huge
    - Machines on same network wouldn't know it
  - Default router
    - Where to send things if not in your table

Routing example



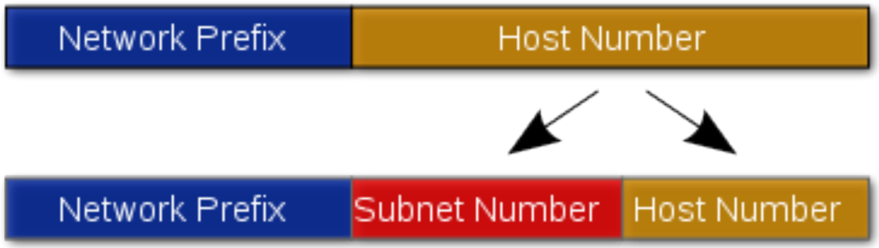
Message	What happens
H1 -> H2	H1 deduces on same network as H2 Sends Ethernet packet directly
H5 -> H8	H5 deduces H8 not on same network H5 sends message to default router R1 R1 can't delivery directly, send to its default router R2 R2 has a forwarding table showing H8 available from R3, sends to R3 R3 delivery to network 4.

The three bears problem

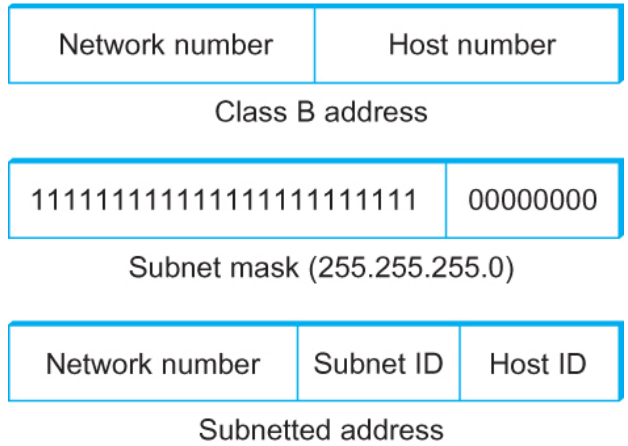
- For most organizations:
  - Class A network too big
  - Class C network too small
  - Class B network... just right
- Actually class B too large for most
  - Half of all class B holders had 50 or fewer hosts
  - 16,384 class B not enough for widespread popularity of the interpipes

Subnetting

- Original intent:
  - Network part of address uniquely identifies one physical network
  - Fewer hosts -> small, faster forwarding tables
- Problem: waste of address space to hand out IP block for each physical network
- Subnetting:
  - Split single classful network number into separate subnets
  - Subnet must be physically near each other



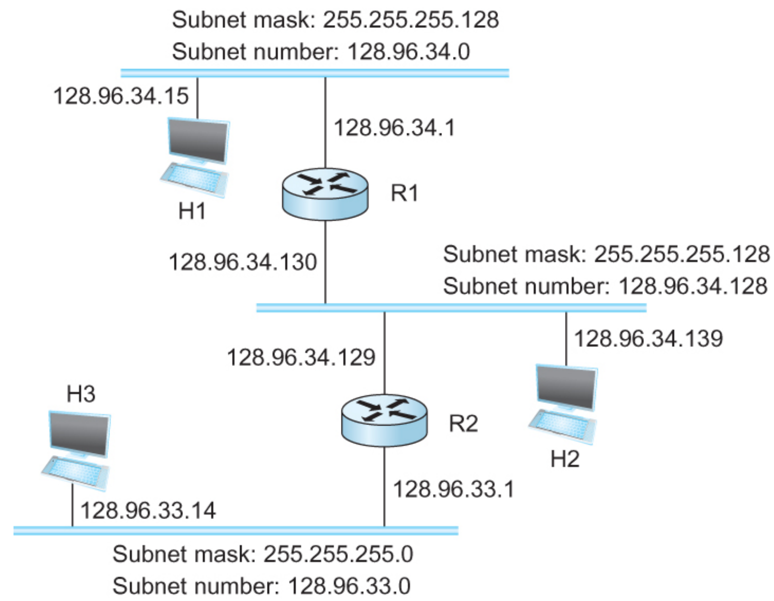
Subnetting examples



	Binary form	Dot-decimal notation
IP address	11000000.10101000.00000101.10000010	192.168.5.130
Subnet mask	11111111.11111111.11111111.**11**000000	255.255.255.192
Network prefix	11000000.10101000.00000101.10000000	192.168.5.128
Host part	00000000.00000000.00000000.00000010	0.0.0.2

	Binary form	Dot-decimal notation
IP address	11000000.10101000.00000101.10000010	192.168.5.130
Subnet mask	11111111.11111111.11111111.00000000	255.255.255.0
Network prefix	11000000.10101000.00000101.00000000	192.168.5.0
Host part	00000000.00000000.00000000.10000010	0.0.0.130

Subnetting




```
D = destination IP address
for each entry <SubnetNum,
                SubnetMask,
                NextHop>
     D1 = SubnetMask & D
     if D1 = SubnetNum
          if NextHop is an interface
             deliver datagram directly
             to destination
          else
             deliver datagram to NextHop
             (a router)
```

Datagram forwarding algorithm.

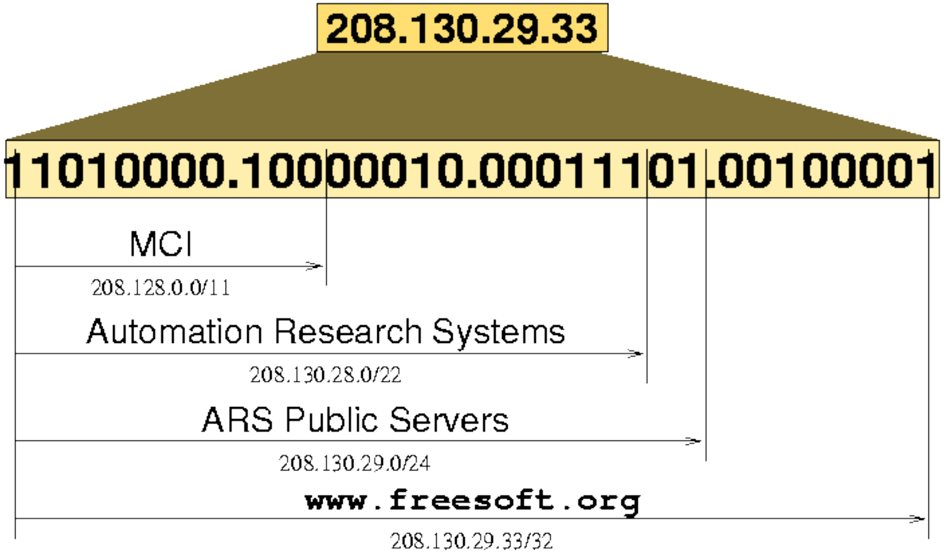
SubnetNumber	SubnetMask	NextHop
128.96.34.0	255.255.255.128	Interface 0
128.96.34.128	255.255.255.128	Interface 1
128.96.33.0	255.255.255.0	R2

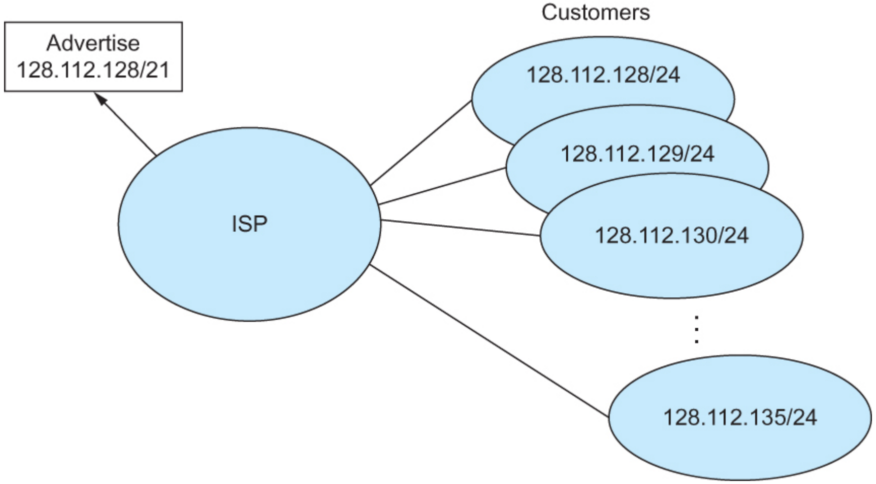
Forwarding table at R1.

Classless addressing

- Classless Interdomain Routing (CIDER)
  - We want:
    - Efficient address allocation
    - Small and fast forwarding tables
- Compromise:
  - Aggregate contiguous blocks of IP addresses
  - New /X notation
    - Specify how many prefix bits are network number
    - Like subnet mask, with X 1's and front then 0's
    - In CIDR 1's must be contiguous

CIDR examples





IP forwarding with CIDR

- CIDR prefixes 2-32 bits
  - May have overlapping prefixes in forwarding table
  - Example:
    - Forwarding table: 171.69 (16-bit prefix)
    - Forwarding table: 171.69.10 (24-bit)
    - Destination: 171.69.10.5, matches both
  - Router uses longest match

Private IP addresses

- Private networks (home networks, etc.)
  - Use specified part of IP address space
  - Not globally routable

IP address range	number of addresses	*classful* description	largest CIDR block (subnet mask)	host id size
10.0.0.0 – 10.255.255.255	16,777,216	single class A	10.0.0.0/8 (255.0.0.0)	24 bits
172.16.0.0 – 172.31.255.255	1,048,576	16 contiguous class Bs	172.16.0.0/12 (255.240.0.0)	20 bits
192.168.0.0 – 192.168.255.255	65,536	256 contiguous class Cs	192.168.0.0/16 (255.255.0.0)	16 bits

http://xkcd.com/742/

Address translation

- Problem:
  - How does a host send a message to someone on their own network?
  - Or to their default router?
  - IP address is not the link-level address (e.g. MAC)
- Solution:
  - Host maintains table: IP address -> link address
  - Uses the Address Resolution Protocol (ARP)

ARP procedure

- If destination IP in sender's ARP table:
  - fire off link-layer packet
  - otherwise send ARP query using broadcast address
- ARP query:
  - IP address you're looking for
  - Your own IP and hardware address
  - Destination responds with hardware address
  - Other hosts can ignore or refresh their ARP tables

```
vertanen@katie:/usr/sbin$ arp
Address                  HWtype  HWaddress           Flags Mask            Iface
10.33.73.39              ether   c8:2a:14:53:da:e5   C                     eth0
lugnut                   ether   00:16:3e:c2:af:f0   C                     eth0
10.33.73.254             ether   00:12:f2:81:ca:74   C                     eth0
10.33.73.120             ether   78:2b:cb:ac:9c:0a   C                     eth0
mtcs32.mtech.edu         ether   b8:ac:6f:45:56:ef   C                     eth0
10.33.73.121             ether   78:2b:cb:ac:9c:1b   C                     eth0
vertanen@katie:/usr/sbin$
```

Assignment of host IP

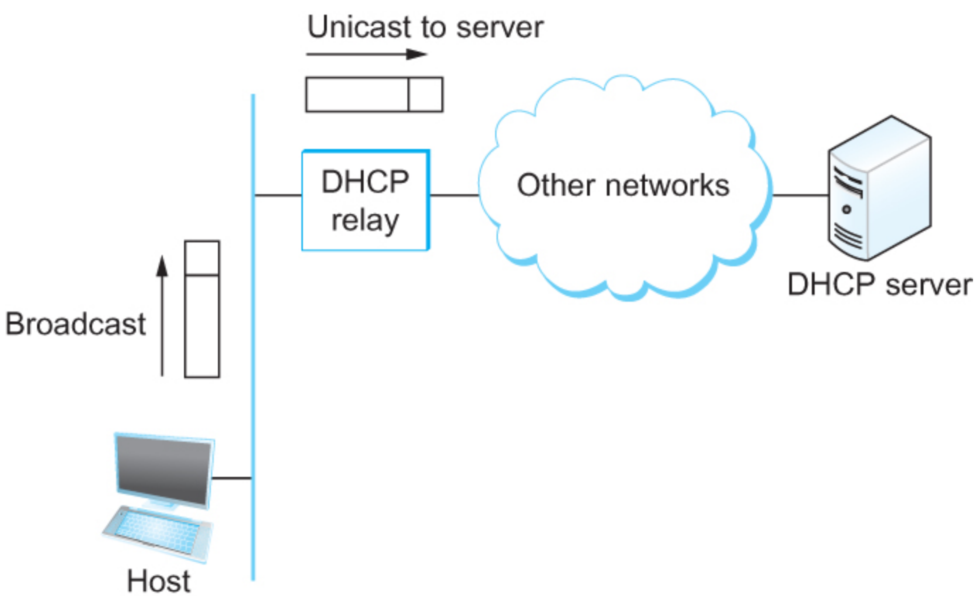
- Each hosts needs:
  - IP address (private or otherwise)
  - Default router IP address
- Configure each host manually
  - Tedious and error-prone
  - Hosts aren't on network, so can't remotely config

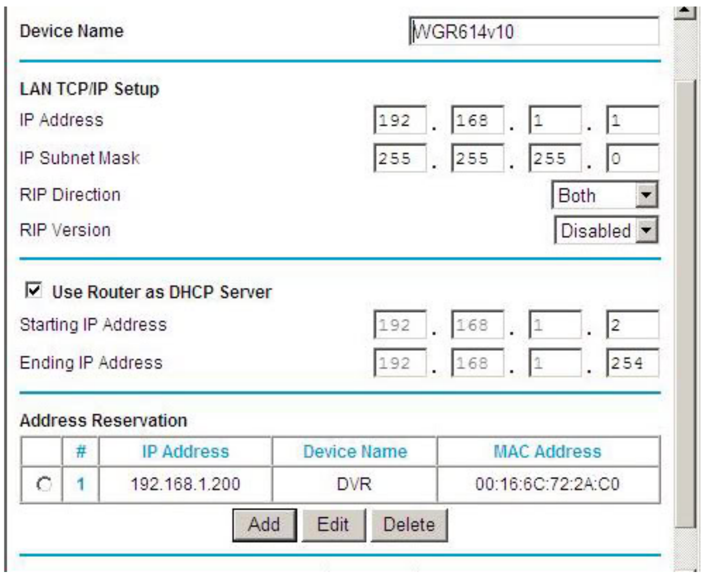
DHCP

- Dynamic Host Configuration Protocol (DHCP)
  - DHCP server provides config info
  - Gives out IP addresses and default router
    - DHCP server has pool of available IP addresses
    - Admin can set DHCP server to give out same IP to given hardware address
    - Addresses leased for a given time period
  - How do hosts discover DHCP server?
    - New hosts sends out broadcast DHCPDISCOVER message

DHCP relays

- If no DHCP server on a network segment
  - Use a relay to communication with DHCP server





Summary

- Internetworking
  - Internet protocol (IP)
    - Foundation of the Internet
    - Runs on virtually any type of network
  - Global IP addresses
  - IP -> hardware address via ARP
  - Hosts can get IP address and router via DHCP