CSE 461: Introduction

Ben Greenstein [Intel Research]

Jeremy Elson [Microsoft Research]

Fall 2008

Outline

- Administrative trivia
- Goals of the course
- How to study networks?

- Instructors: Ben Greenstein, Jeremy Elson
  - Email: ben@cs.washington.edu, jelson@gmail.com

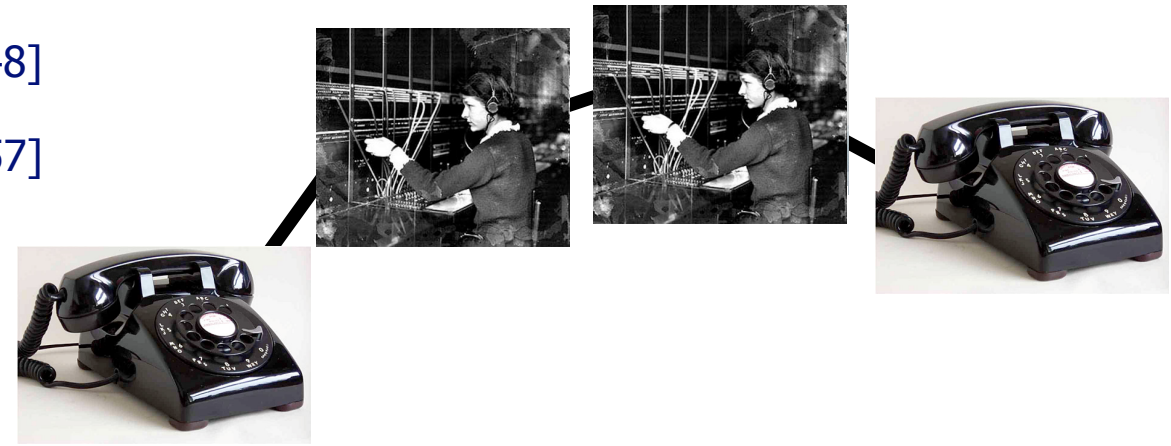
- Teaching assistants:
  - Ivan Beschastnikh: runs the sessions, manages projects
  - Alper Sarikaya: in charge of the homeworks
  - Email: ivan@cs.washington.edu, alpers@cs.washington.edu
  - Office hours: 2:30-3:30 MW CSE 214 (tentative)

Administrative Details

- Everything you need is on the course web page
  - http://www.cs.washington.edu/cse461
- Your TODO list:
  - Join the mailing list if not already on it: cse461@cs.washington.edu
  - Gain access to the CSE Labs (form for non-majors)
  - Get Computer Networks by Peterson and Davie, 4th Edition http://www.amazon.com/Computer-Networks-Approach-Kaufmann-Networking/dp/0123705487
  - Read chapters 1 and 2
  - Go to section
  - Obtain a Linksys WRT54GL wireless router from the TA (Ivan in CSE 391) tomorrow

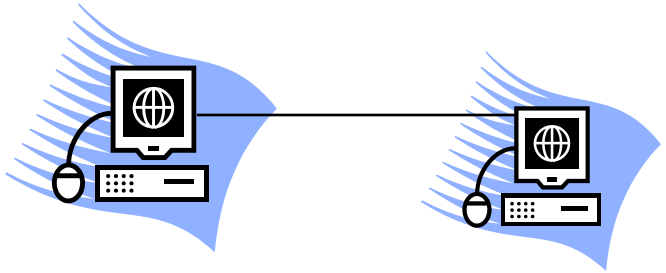
A Network in 461

- "Network" is clearly an overloaded word:
  - Economic networks, regulatory networks, social networks…
  - Important networks that we care less about…
    - Nollet [1746]
    - Morse/Cooke [1836]
    - Telephone (Bell/Gray) [1876]
    - Cable TV [1948]
    - Satellites [1957]

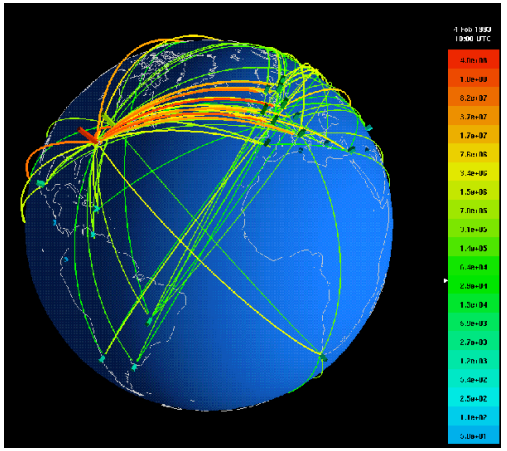


A Network in 461

- For 461, a network is what you get anytime you connect two or more computers together by some kind of a link.
  - I.e., Kleinrock, Kahn, Cerf, etc.



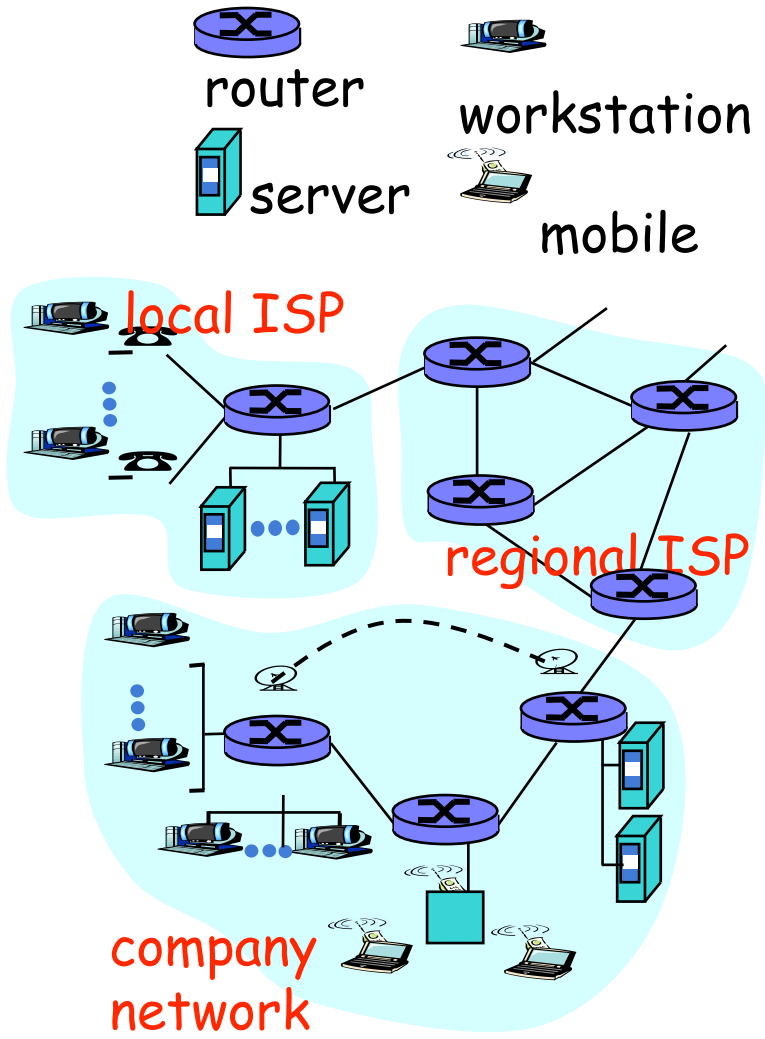
OR



Of particular interest: the Internet

- A network of networks
- Hierarchical structure
- millions of connected computing devices: *hosts, end-systems*
  - pc's workstations, servers
  - PDA's phones, toasters

  running *network apps*
- *communication links*
  - fiber, copper, radio, satellite
- *routers:* forward packets (chunks) of data thru network



“Cool” Internet Appliances

IP picture frame
http://www.ceiva.com/

Web-enabled toaster+weather forecaster
http://dancing-man.com/robin/toasty/

The networks we study

- We are interested in networks that are:
  - Large scale
  - Intrinsically unreliable
  - Distributed
  - Heterogeneous

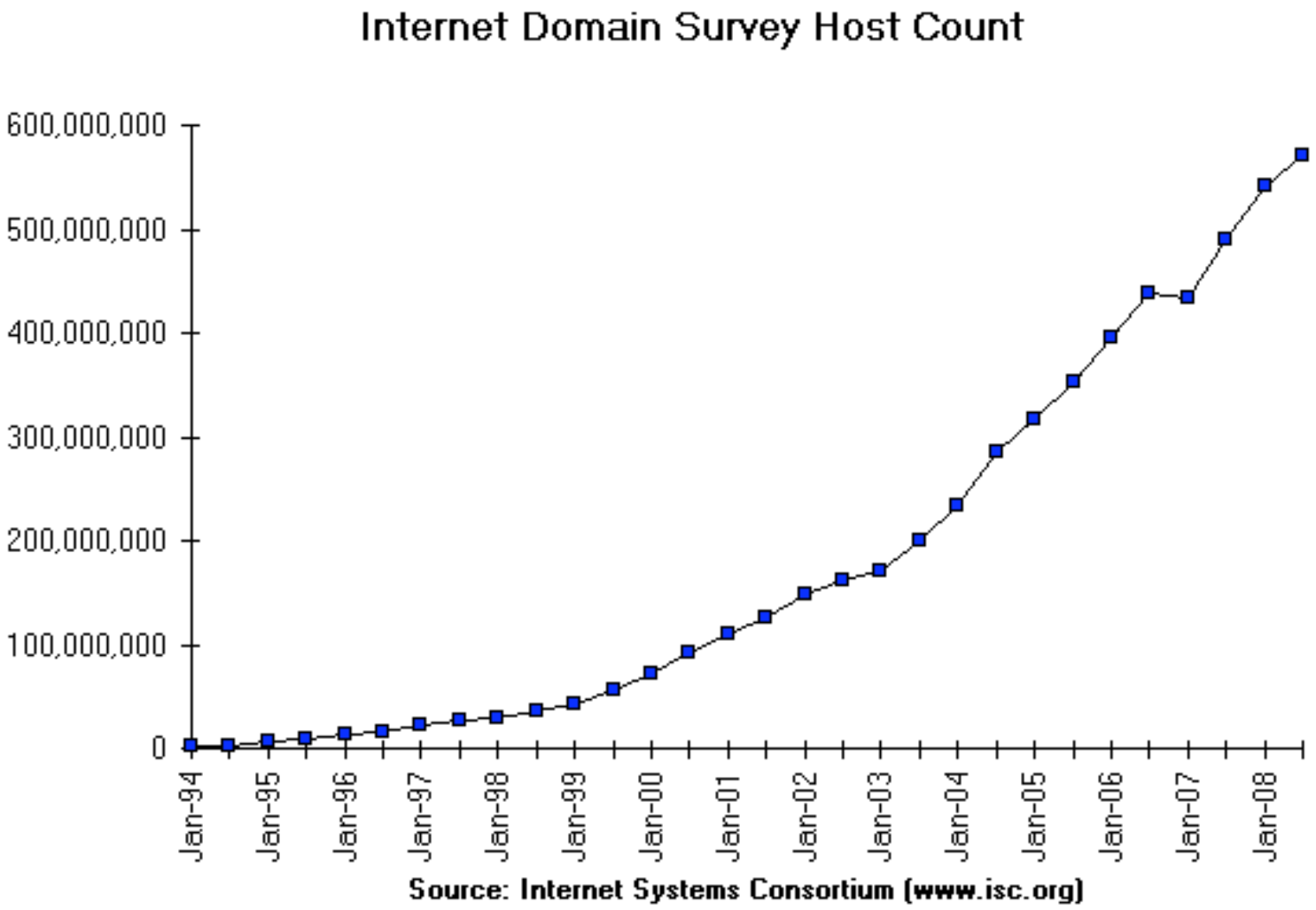
The meaning of "Large-scale"

TOP 20 COUNTRIES WITH HIGHEST NUMBER OF INTERNET USERS

#	Country or Region	Internet Users, Latest Data	Penetration (% Population)	% of World Users	Population (2008 Est.)	User Growth (2000 - 2008)
1	China	253,000,000	19.0 %	17.3 %	1,330,044,605	1,024.4 %
2	United States	220,141,969	72.5 %	15.0 %	303,824,646	130.9 %
3	Japan	94,000,000	73.8 %	6.4 %	127,288,419	99.7 %
4	India	60,000,000	5.2 %	4.1 %	1,147,995,898	1,100.0 %
5	Germany	52,533,914	63.8 %	3.6 %	82,369,548	118.9 %
6	Brazil	50,000,000	26.1 %	3.4 %	191,908,598	900.0 %
7	United Kingdom	41,817,847	68.6 %	2.9 %	60,943,912	171.5 %
8	France	36,153,327	58.1 %	2.5 %	62,177,676	325.3 %
9	Korea, South	34,820,000	70.7 %	2.4 %	49,232,844	82.9 %
10	Italy	34,708,144	59.7 %	2.4 %	58,145,321	162.9 %
11	Russia	32,700,000	23.2 %	2.2 %	140,702,094	954.8 %
12	Canada	28,000,000	84.3 %	1.9 %	33,212,696	120.5 %
13	Turkey	26,500,000	36.9 %	1.8 %	71,892,807	1,225.0 %
14	Spain	25,623,329	63.3 %	1.8 %	40,491,051	375.6 %
15	Indonesia	25,000,000	10.5 %	1.7 %	237,512,355	1,150.0 %
16	Mexico	23,700,000	21.6 %	1.6 %	109,955,400	773.8 %
17	Iran	23,000,000	34.9 %	1.6 %	65,875,223	9,100.0 %
18	Vietnam	20,159,615	23.4 %	1.4 %	86,116,559	9,979.8 %
19	Pakistan	17,500,000	10.4 %	1.2 %	167,762,040	12,969.5 %
20	Australia	16,355,388	79.4 %	1.1 %	20,600,856	147.8 %
TOP 20 Countries		1,115,713,572	25.4 %	76.2 %	4,388,052,548	284.5 %
Rest of the World		347,918,789	15.2 %	23.8 %	2,288,067,740	391.2 %
Total World - Users		1,463,632,361	21.9 %	100.0 %	6,676,120,288	305.5 %

http://www.internetworldstats.com/top20.htm

The meaning of “Large-scale”



Intrinsic Unreliability

- Information sent from a first place to a second
  - May not arrive
  - May arrive more than once
  - May arrive in garbled fashion
  - May arrive out of order
  - May be read by others
  - May be modified by others
- Why build intrinsically unreliable networks?

Distributed

- (Hopefully) independent failure modes
- Exposed and hidden dependencies

> *"A distributed system is a system in which I can't do my work because some computer has failed that I've never even heard of."* – Lamport

- Independent administrative controls
- Leads to…

Heterogeneous Networks

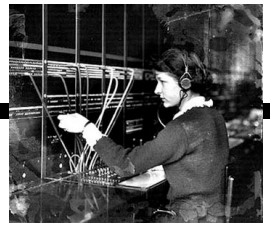
- Heterogeneous: Made up of different kinds of stuff
- Homogeneous: Made up of the same kind of stuff
- Principles
  - Homogeneous networks are easier to deal with
  - Heterogeneous networks have their own strengths

Fundamental question of this course:

- Given its enormous size, questionable reliability, wide-scale distribution and incredible heterogeneity, why does the Internet work?
  - (and why, at times, does it not?)
  - In the process of learning why, we'll cover 'how' and 'what' in some detail

Something to think about: Hasn't this been done before?

- Telephone network has worked well for over a century
  - Large scale
  - Fairly reliable
    - Each phone call gets a dedicated circuit (reserved path through the network)
  - Fairly centralized
    - Ma Bell used to operate almost all networks
  - Fairly homogenous
    - We all have telephones
  - Why is the Internet architected so differently?

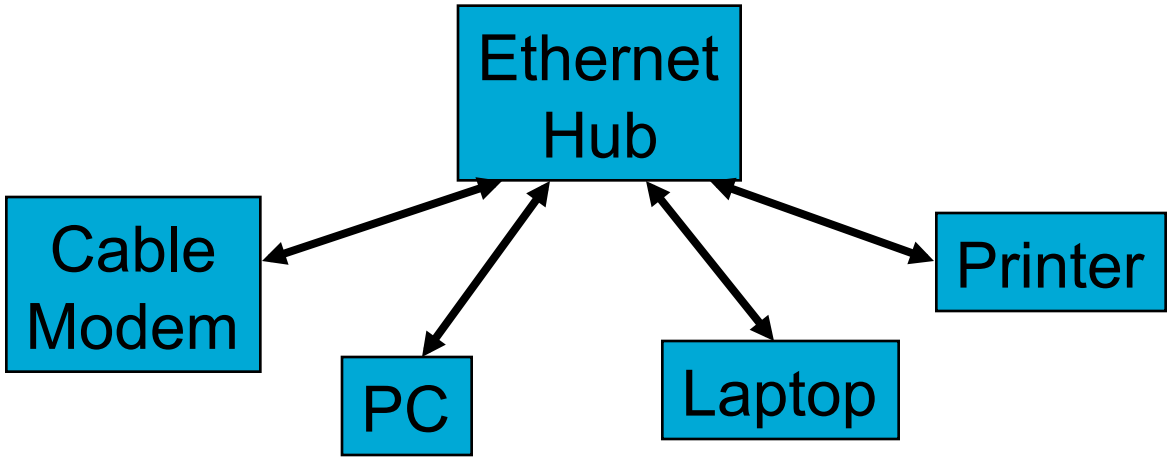




Model of a Network

- Links carry information (bits)
  - Wire, wireless, fiber optic, smoke signals …
  - May be point-to-point or broadcast
- Switches move bits between links
  - Routers, gateways,bridges, CATV headend, PABXs, …
- Hosts are the communication endpoints
  - PC, PDA, cell phone, tank, toaster, …
  - Hosts have names

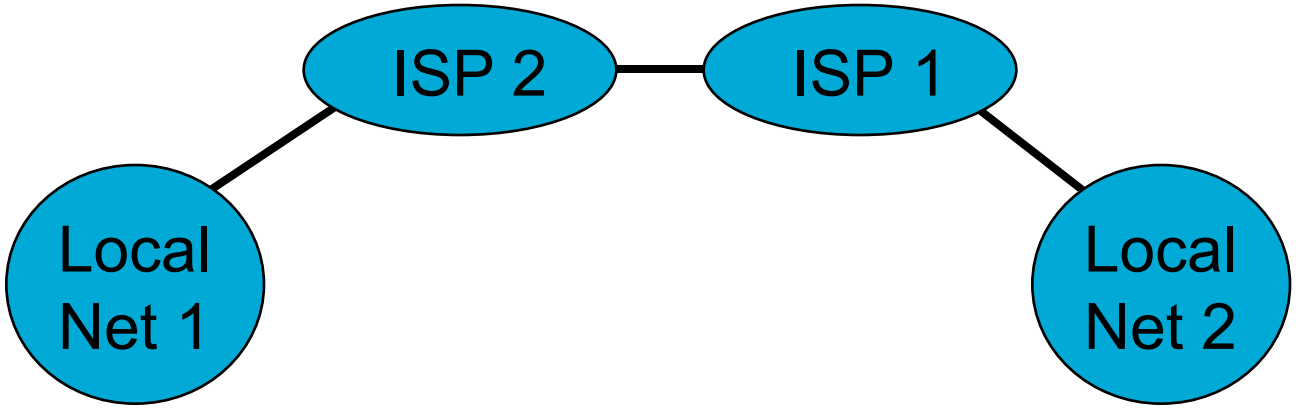
- Much other terminology: channels, nodes, intermediate systems, end systems, and much more
  - Beware of TLA (three letter acronym) overload

Example – Local Area Network



- Your home network
  - Ethernet is a broadcast-capable multi-access LAN

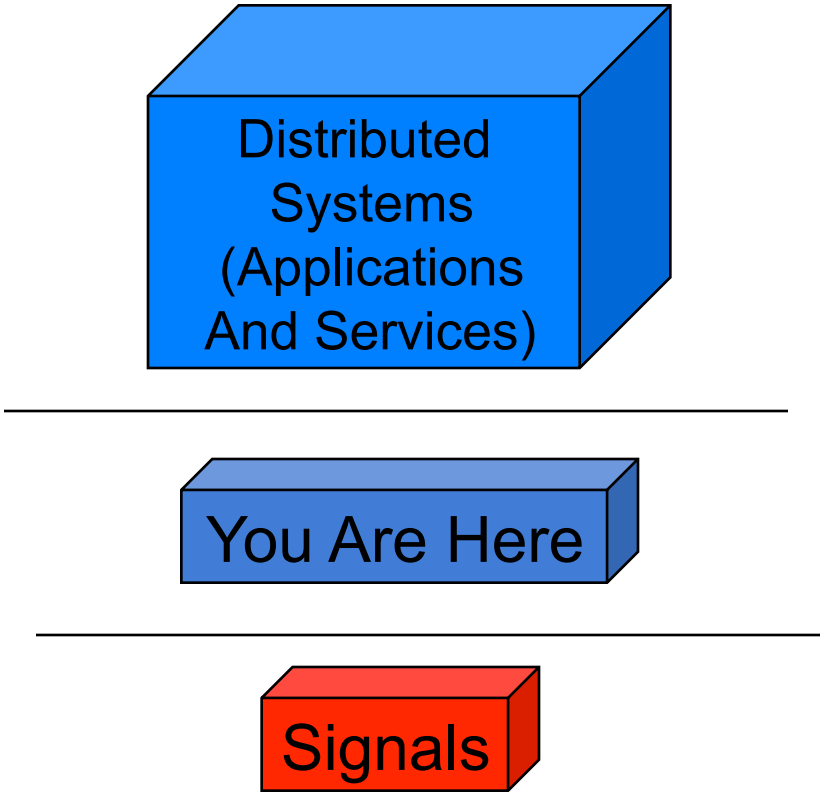
Example – An Internetwork



- Internetwork is a network of networks
- The Internet is a global internetwork in which all participants speak a common language
  - IP, the Internet Protocol

Goal of this Course

- You will understand how to design and build *large, distributed computer* networks.
  - Fundamental problems in building networks
  - Design principles of proven value
  - Common implementation technologies
- This is a systems course, not queuing theory, signals, or hardware design.
- We focus on networks, and a bit on applications or services that run on top of them (distributed systems).



How to study networks?

- Networks in general, and Internet in particular, are complex beasts
- Question: how do we begin to understand Internet's workings?
  - Hands-on programming
  - Tinkering, reverse-engineering the network

Programming Projects

- Develop services for a Linksys access point, including a broadcast tunnel and measurement tool
- WRT54GL can run Linux
- Develop on desktop, port to Linksys (need to cross-compile to MIPS)
- Programming in C
- Work alone, subject to the Gilligan's Island rule
  Each person will have a router for the quarter

Random Survey

- Do you have Internet access at home?
- Do you use a wireless router at home?
  - That you can unplug at will?
- Have you taken OS (CSE 451)?
- Have you written:
  - a program in C?
  - A program that runs on Linux?
  - Software for an embedded system?

Project Outline

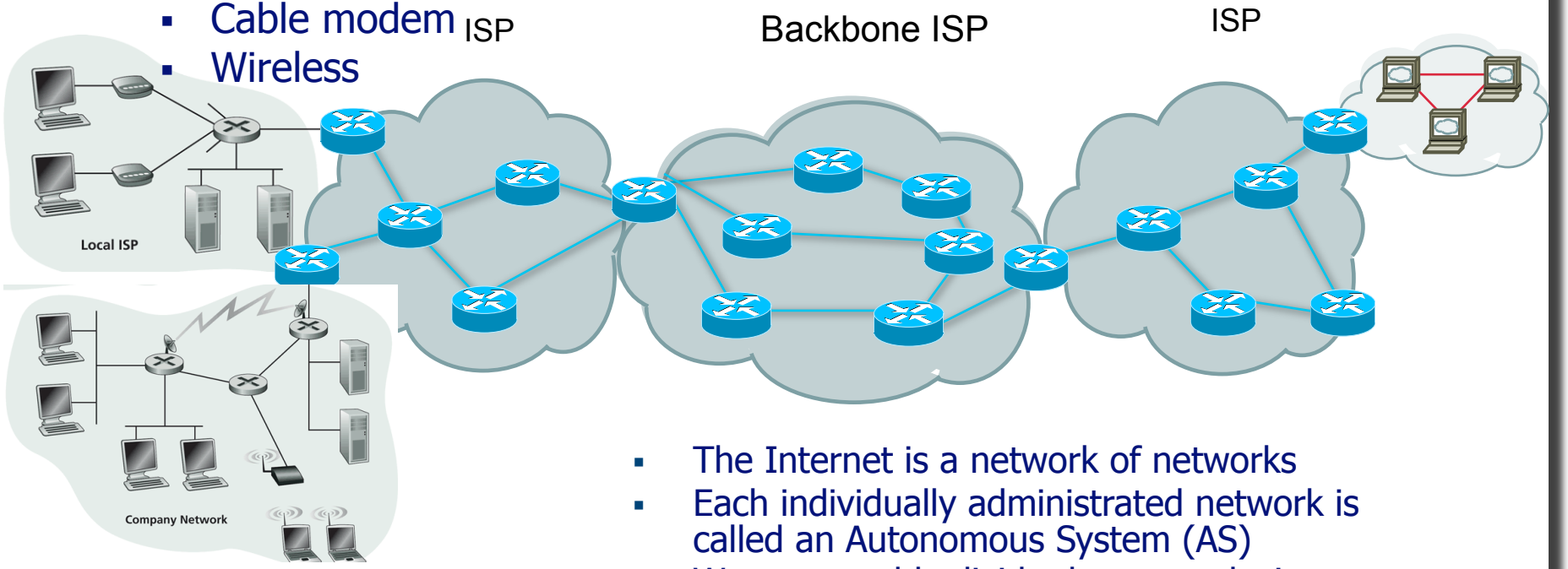
- Phase 1:
  - Chat with the Oracle
  - Write the client side of a protocol to communicate with an Oracle (a server we've written)
- Phase 2:
  - Measure traffic on the wireless LAN
  - Modify the router to keep track of the broadcast messages that are sent on the network
- Phase 3:
  - Tunnel broadcasts via the Oracle
  - Direct broadcast traffic from your network to all other students' networks by way of the oracle

- Goal: Service sharing across networks (e.g., for iTunes playlists)

Explore and reverse-engineer networks

Residential access

- Modem
- DSL
- Cable modem
- Wireless



Campus access

- Ethernet
- FDDI
- Wireless

- The Internet is a network of networks
- Each individually administrated network is called an Autonomous System (AS)
- We can roughly divide the networks into access networks and transit networks

A Connectivity Exploration Tool

- Traceroute:
  - Run traceroute host-name on unix machines
  - tracert host-name on windows
- Sends three probes to each intermediate node on the path to the final destination (more details later)
- Reports the IP address, a more readable name, and the round-trip latencies for the probes

Traceroute to an East Coast College

```
-bash-3.1$ traceroute planetx.scs.cs.nyu.edu
traceroute to planetx.scs.cs.nyu.edu (216.165.109.79), 30 hops max, 40 byte packets
 1  acar-hsh-01-vlan75.cac.washington.edu (128.208.2.100)  0.362 ms  0.353 ms  0.396 ms
 2  uwcr-hsh-01-vlan3904.cac.washington.edu (205.175.110.17)  0.407 ms  0.444 ms  0.478
    ms
 3  uwcr-hsh-01-vlan1901.cac.washington.edu (205.175.103.5)  0.592 ms  0.665 ms  0.687 ms
 4  uwbr-ads-01-vlan1902.cac.washington.edu (205.175.103.10)  50.060 ms  50.120 ms
    50.130 ms
 5  hnsp2-wes-ge-0-0-0-0.pnw-gigapop.net (209.124.176.12)  0.703 ms  0.729 ms  0.760 ms
 6  abilene-pnw.pnw-gigapop.net (209.124.179.2)  0.544 ms  0.561 ms  0.588 ms
 7  dnvrng-sttlng.abilene.ucaid.edu (198.32.8.50)  46.984 ms  46.969 ms  47.009 ms
 8  kscyng-dnvrng.abilene.ucaid.edu (198.32.8.14)  63.746 ms  62.699 ms  62.709 ms
 9  iplsng-kscyng.abilene.ucaid.edu (198.32.8.80)  57.320 ms  57.305 ms  57.344 ms
10  chinng-iplsng.abilene.ucaid.edu (198.32.8.76)  70.506 ms  71.011 ms  70.985 ms
11  buf-7600-abilene-chin.nysernet.net (199.109.2.1)  73.003 ms  72.942 ms  72.946 ms
12  nyc-gsr-buf-7600.nysernet.net (199.109.7.14)  81.995 ms  81.966 ms  81.936 ms
13  nyu-nyc-gsr.nysernet.net (199.109.4.22)  82.179 ms  82.249 ms  82.314 ms
14  WWLABGW.NYU.NET (192.76.177.75)  82.350 ms  82.188 ms  82.200 ms
15  delancy.scs.cs.nyu.edu (216.165.108.191)  82.307 ms  82.662 ms  82.558 ms
16  planetx.scs.cs.nyu.edu (216.165.109.79)  82.629 ms  82.493 ms  82.592 ms
```

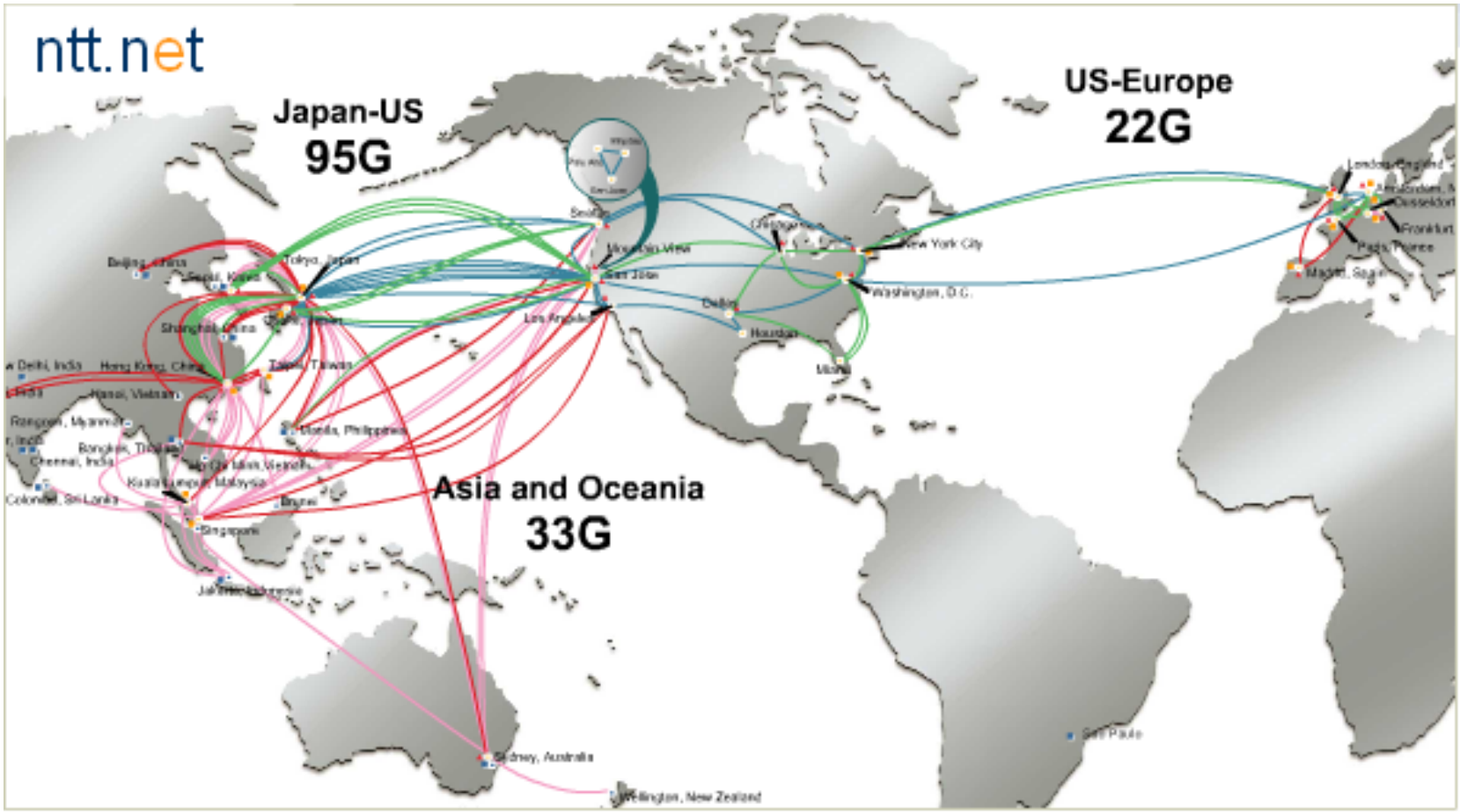
Abilene I2 Backbone

http://abilene.internet2.edu/maps-lists/

Traceroute to a commercial webserver

```
-bash-3.1$ traceroute www.nytimes.com
traceroute to www.nytimes.com (199.239.136.200), 30 hops max, 40 byte packets
 1  acar-hsh-01-vlan75.cac.washington.edu (128.208.2.100)  0.358 ms  0.357 ms  0.400 ms
 2  uwcr-hsh-01-vlan3904.cac.washington.edu (205.175.110.17)  0.426 ms  0.467 ms  0.502 ms
 3  uwcr-hsh-01-vlan1901.cac.washington.edu (205.175.103.5)  0.609 ms  0.639 ms  0.687 ms
 4  uwbr-ads-01-vlan1902.cac.washington.edu (205.175.103.10)  0.386 ms  0.428 ms  0.445 ms
 5  cnsp1-wes-ge-0-0-0-0.pnw-gigapop.net (209.124.176.8)  0.579 ms  0.643 ms  0.730 ms
 6  129.250.10.194 (129.250.10.194)  70.290 ms  66.878 ms  66.907 ms
 7  xe-1-2-0.r20.sttlwa01.us.bb.gin.ntt.net (129.250.2.206)  1.060 ms  1.063 ms  1.045 ms
 8  ae-0.r21.sttlwa01.us.bb.gin.ntt.net (129.250.2.54)  0.901 ms  0.901 ms  0.883 ms
 9  p64-2-0-0.r20.nycmny01.us.bb.gin.ntt.net (129.250.5.17)  74.106 ms  74.095 ms  74.103 ms
10  xe-4-1.r02.nycmny01.us.bb.gin.ntt.net (129.250.2.187)  141.125 ms  141.209 ms  141.305 ms
11  ge-1-1.a00.nycmny01.us.da.verio.net (129.250.30.113)  73.897 ms  73.997 ms  73.968 ms
12  * * *
13  * * *
14  * * *
```

A Commercial backbone: NTT



Traceroute to another commercial webserver

```
-bash-3.1$ traceroute www.nyse.com
traceroute to www.nyse.com (209.124.184.150), 30 hops max, 40 byte packets
 1  acar-hsh-01-vlan75.cac.washington.edu (128.208.2.100)  0.327 ms  0.353 ms  0.392 ms
 2  uwcr-hsh-01-vlan3904.cac.washington.edu (205.175.110.17)  0.374 ms  0.412 ms  0.443 ms
 3  uwcr-hsh-01-vlan1901.cac.washington.edu (205.175.103.5)  0.595 ms  0.628 ms  0.659 ms
 4  uwbr-ads-01-vlan1902.cac.washington.edu (205.175.103.10)  0.445 ms  0.472 ms  0.501 ms
 5  ccar1-ads-ge-0-0-0-0.pnw-gigapop.net (209.124.176.32)  0.679 ms  0.747 ms  0.775 ms
 6  a209.124.184.150.deploy.akamaitechnologies.com.184.124.209.in-addr.arpa (209.124.184.150)
      0.621 ms  0.456 ms  0.419 ms
```

What is going on?

```
-bash-3.1$ nslookup www.nyse.com
Name:    a789.g.akamai.net
Address: 209.124.184.137
Name:    a789.g.akamai.net
Address: 209.124.184.150
```

Points to note

- Multi-homed
- Certain routers don't respond
- Variability in response times
- Geography not apparent
  - Geography does not dictate paths
  - Sometimes paths are horribly inflated. Why?
- Content distribution networks operate by returning a nearby cache site

- Reverse engineering is fun!