INFM 603: Information Technology and Organizational Context

Session 1: Physical and Web Infrastructure

Jimmy Lin
The iSchool
University of Maryland

Thursday, September 5, 2013

Introduction

(How I got here)

Source: Wikipedia (All Souls College, Oxford)

Source: Wikipedia (Factory)

Source: Wikipedia (All Souls College, Oxford)

Introduction

(How you got here)

This course is about programming

(but the goal is not to make you into a programmer)

The key to surviving technology?

Agility

Source: Wikipedia (Cheetah

A brief history…

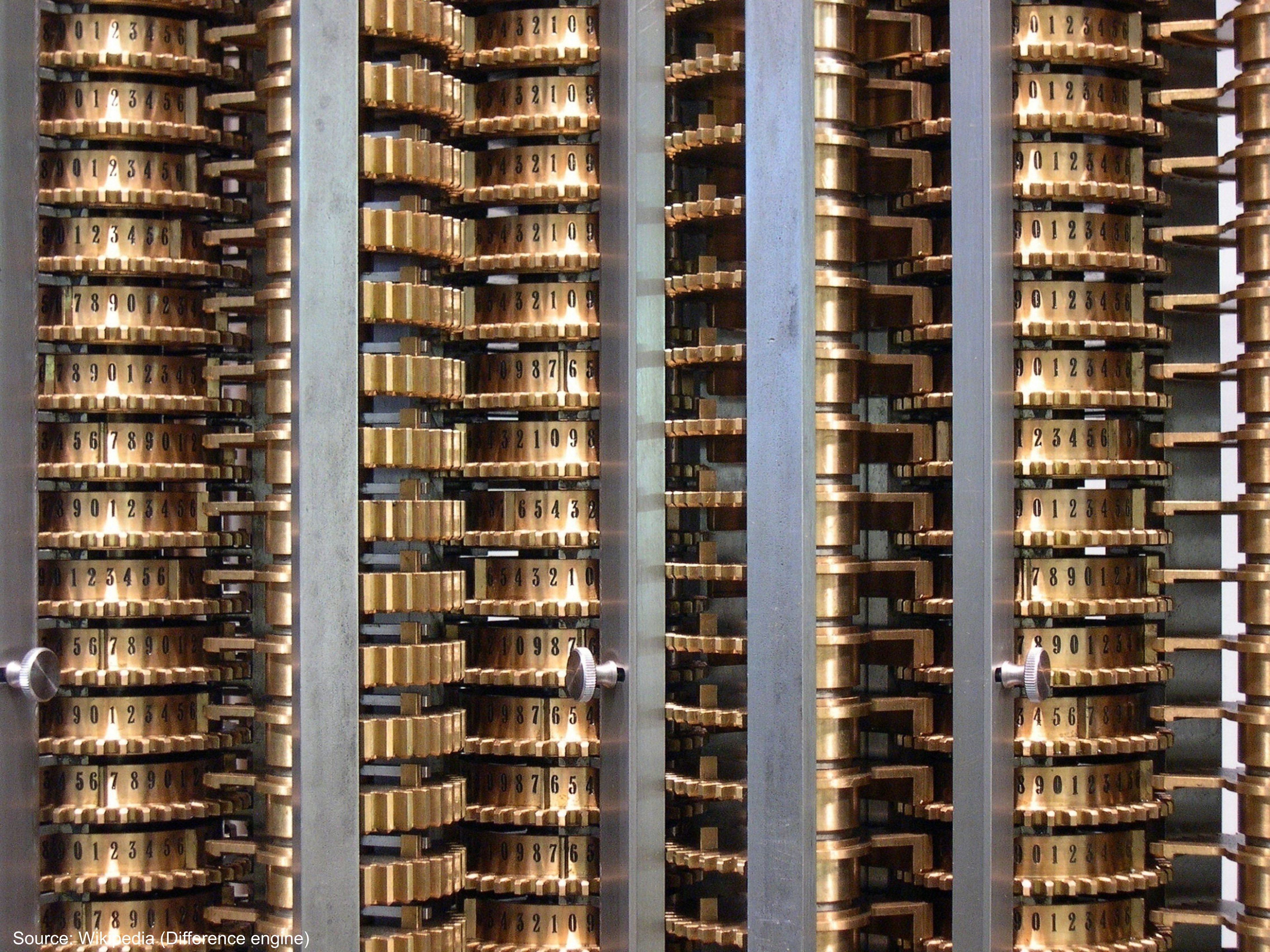
(How computing got here)

A COMPUTER WANTED.

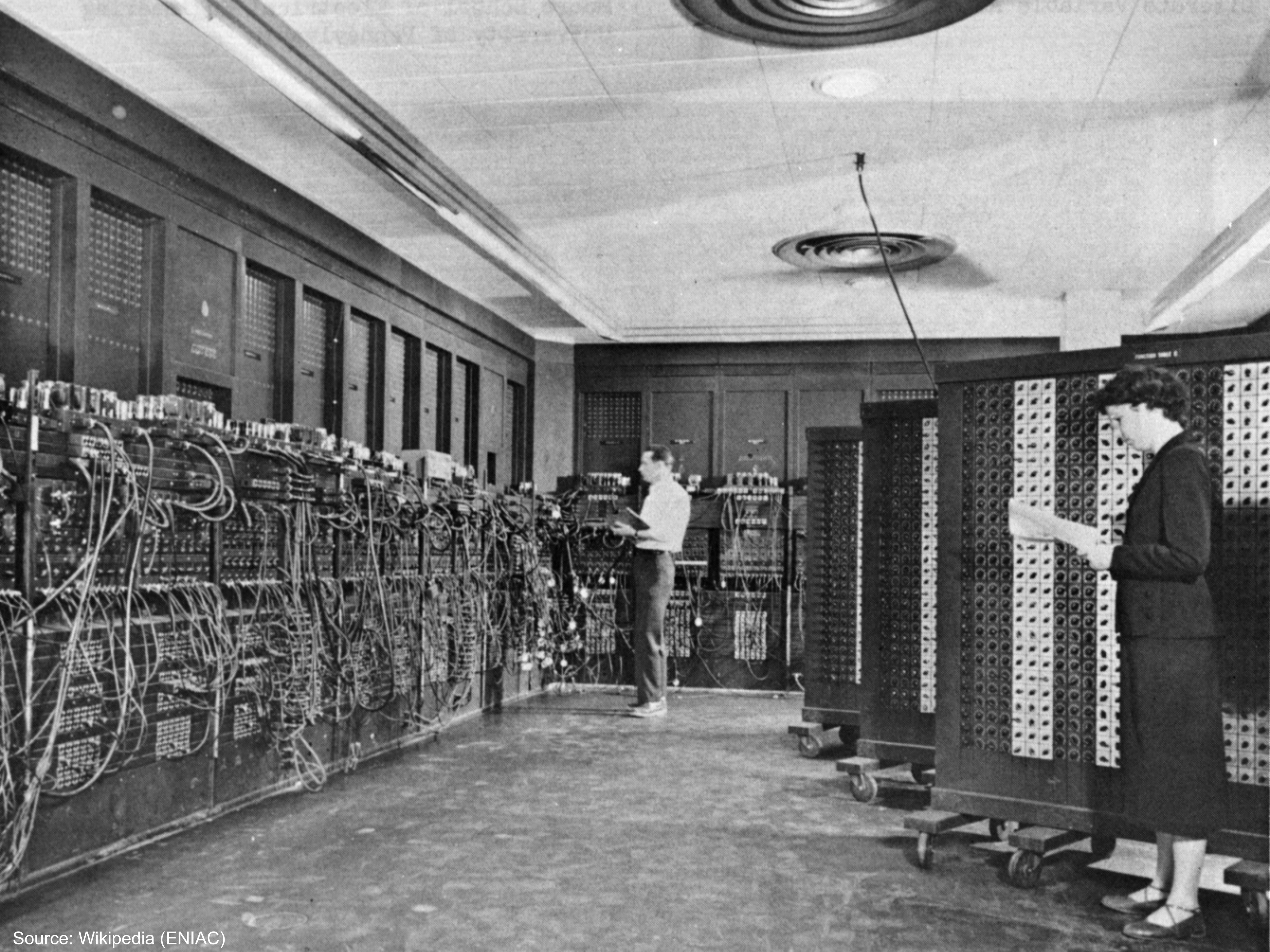
WASHINGTON, May 1.—A civil service examination will be held May 18 in Washington, and, if necessary, in other cities, to secure eligibles for the position of computer in the Nautical Almanac Office, where two vacancies exist—one at $1,000, the other at $1,400.

The examination will include the subjects of algebra, geometry, trigonometry, and astronomy. Application blanks may be obtained of the United States Civil Service Commission.

The New York Times
Published: May 2, 1892



Source: Wikipedia (Difference engine)



Source: Wikipedia (ENIAC)

Source: Wikipedia (IBM 704)

Source: Wikipedia (PDP-11)

Source: Wikipedia (Personal computer)

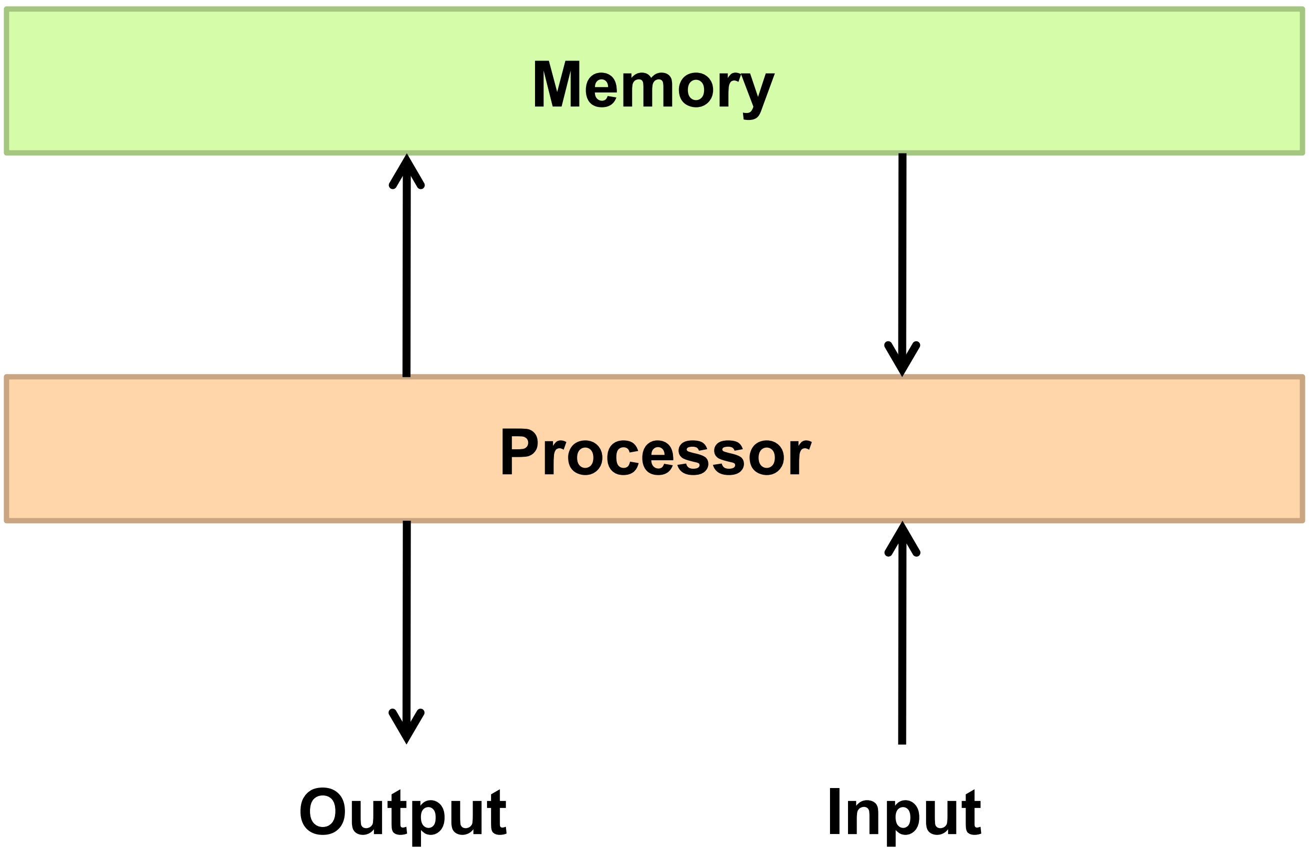
Source: Wikipedia

Source: Wikipedia


Source: Google

Source: Google

What is a computer?



The Processing Cycle

- Input comes from somewhere
  - Keyboard, mouse, touchpad, touch screen, microphone, camera, …
  - Fetch data from memory
- The computer does something with it
  - Add, subtract, multiply, etc.
- Output goes somewhere
  - Monitor, speaker, printer, robot controls, …
  - Store data back into memory

Source: Intel

Source: Wikipedia (Random-access memory)


Source: Wikipedia

Source: Solid-state drive

Networking

Why Networking?

- Sharing data
- Sharing hardware
- Sharing software
- Increasing robustness
- Facilitating communications
- Facilitating commerce

How did it all start?
How did it evolve?
How did we get here?

Packet vs. Circuit Networks

- Telephone system ("circuit-switched")
  - Fixed connection between caller and called
  - High network load results in busy signals
- Internet ("packet-switched")
  - Each transmission is broken up into pieces and routed separately
  - High network load results in long delays

Packet Switching

- Break long messages into short “packets”
  - Keeps one user from hogging a line
  - Each packet is tagged with where it’s going
- Route each packet separately
  - Each packet often takes a different route
  - Packets often arrive out of order
  - Receiver must reconstruct original message
  - How do packet-switched networks deal with continuous data?
  - What happens when packets are lost?

Different Networks Types

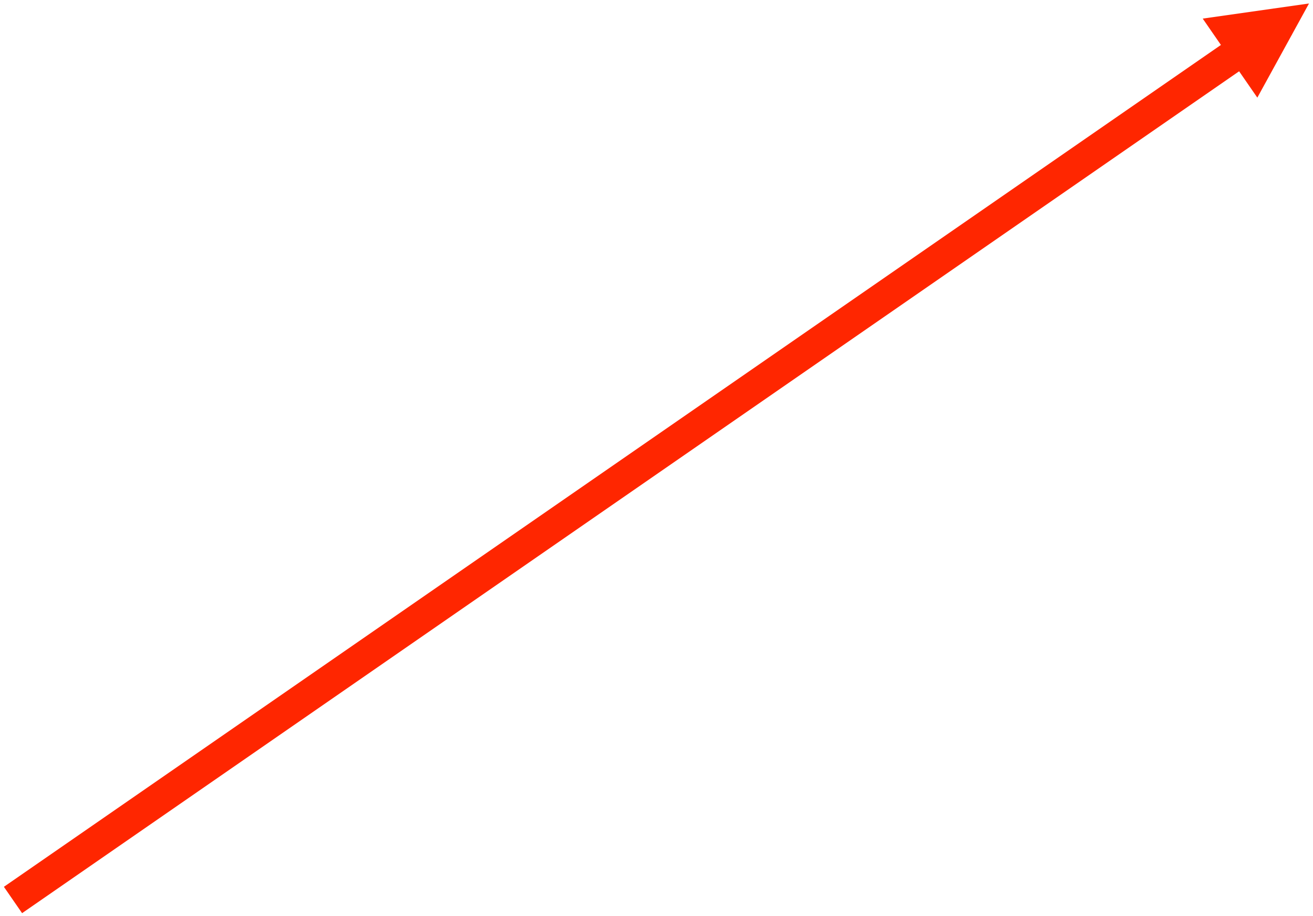
- Local Area Networks (LANs)
  - Connections within a building or a small area
  - Wireless or wired
- Wide Area Networks (WANs)
  - Connections between multiple LANs
  - May cover thousands of square miles
- The Internet
  - Collection of WANs across multiple organizations

The Internet

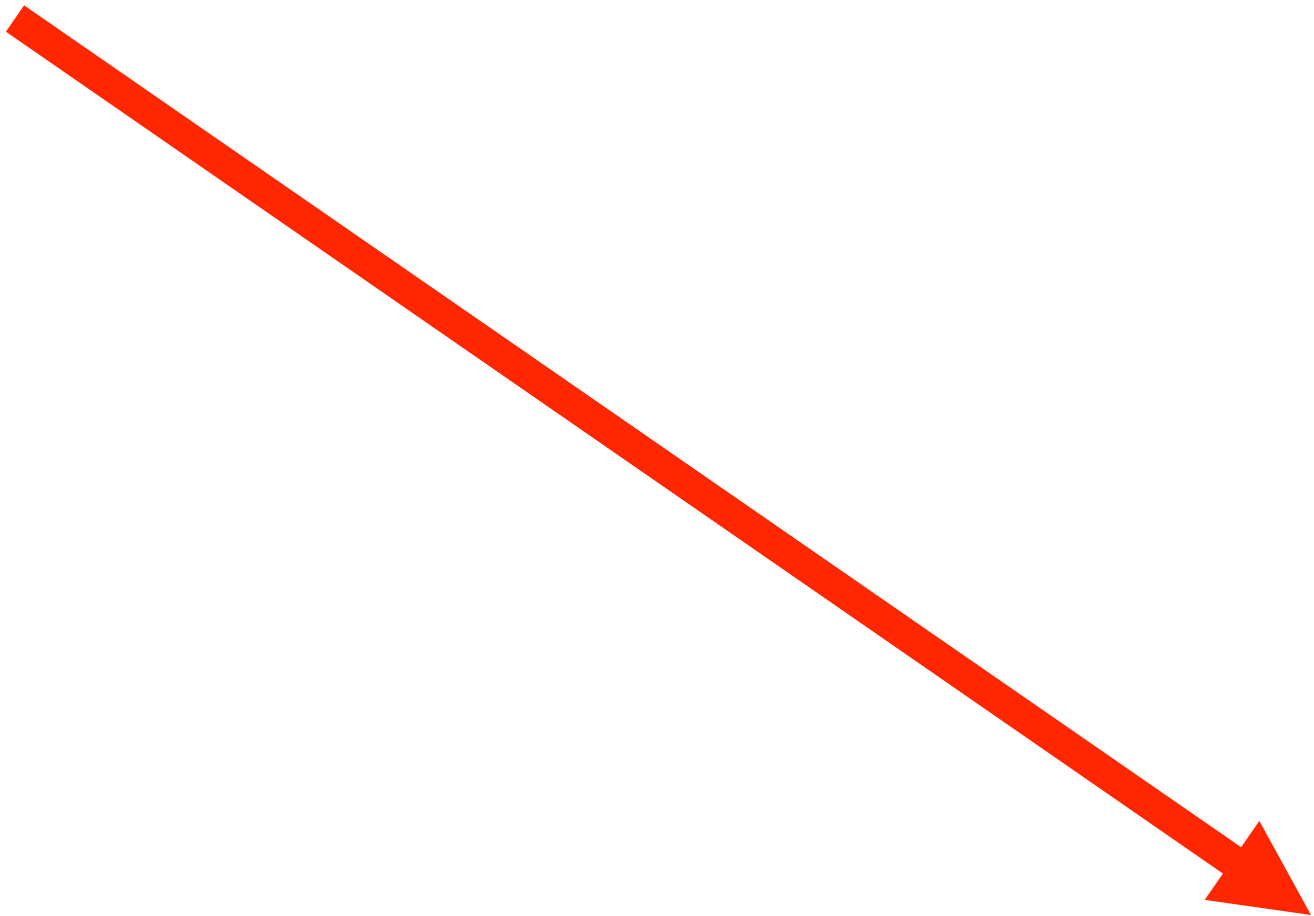
- Global collection of public networks
  - Private networks are often called “intranets”
- Use of shared protocols
  - TCP/IP (Transmission Control Protocol/Internet Protocol): basis for communication
  - DNS (Domain Name Service): basis for naming computers on the network
  - HTTP (HyperText Transfer Protocol): World Wide Web
- Next week: how does all of this work?

Characterizing Computing

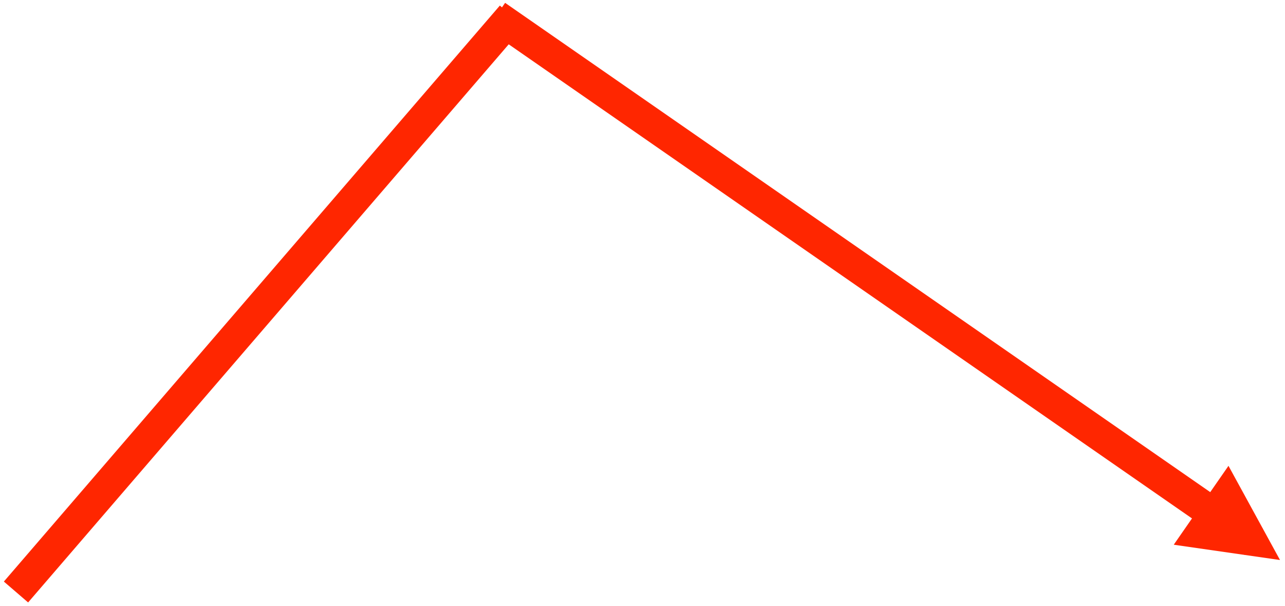
Trends in Computing: #1



Trends in Computing: #2



Trends in Computing: #3

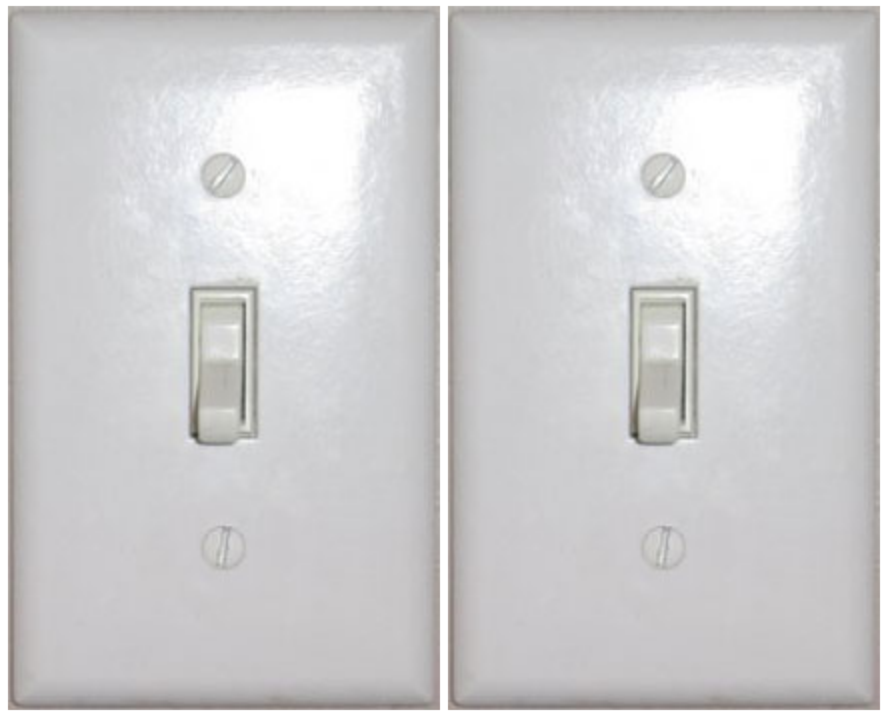


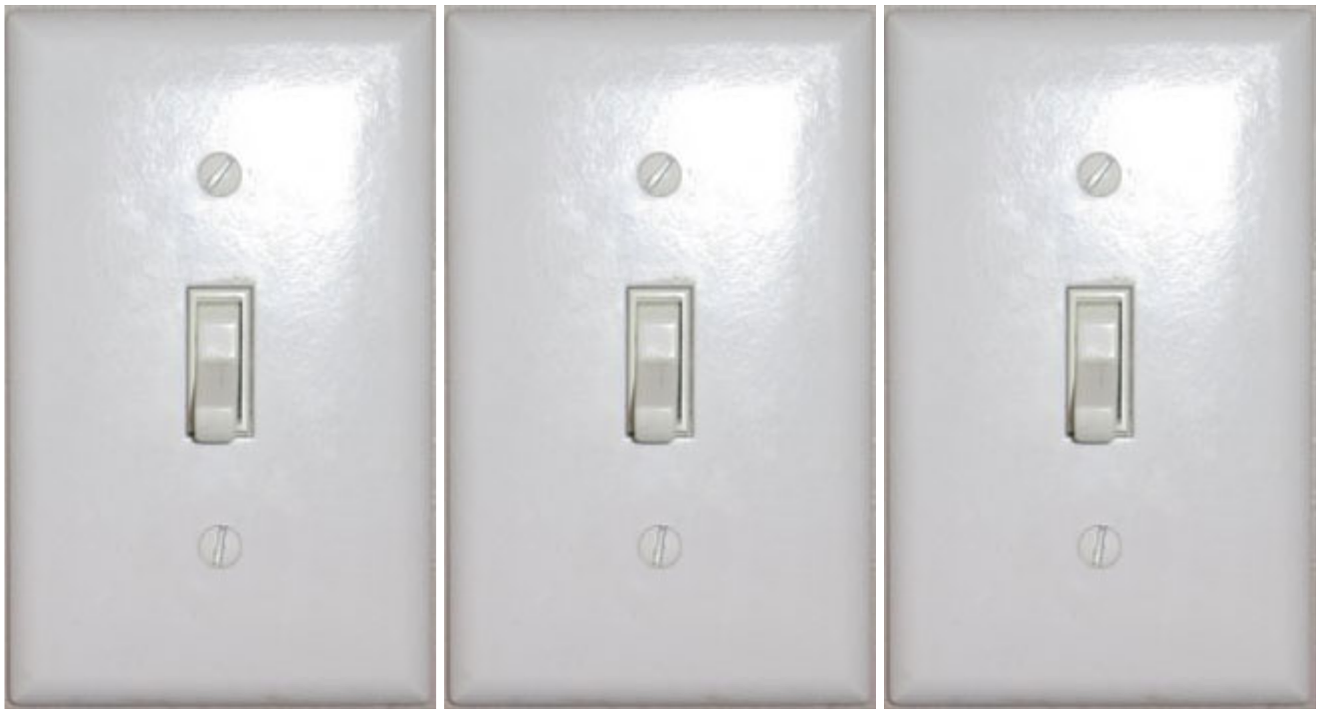
Ways to characterize computing

- How big?
- How fast?
- How reliable?

Computing is fundamentally about tradeoffs!

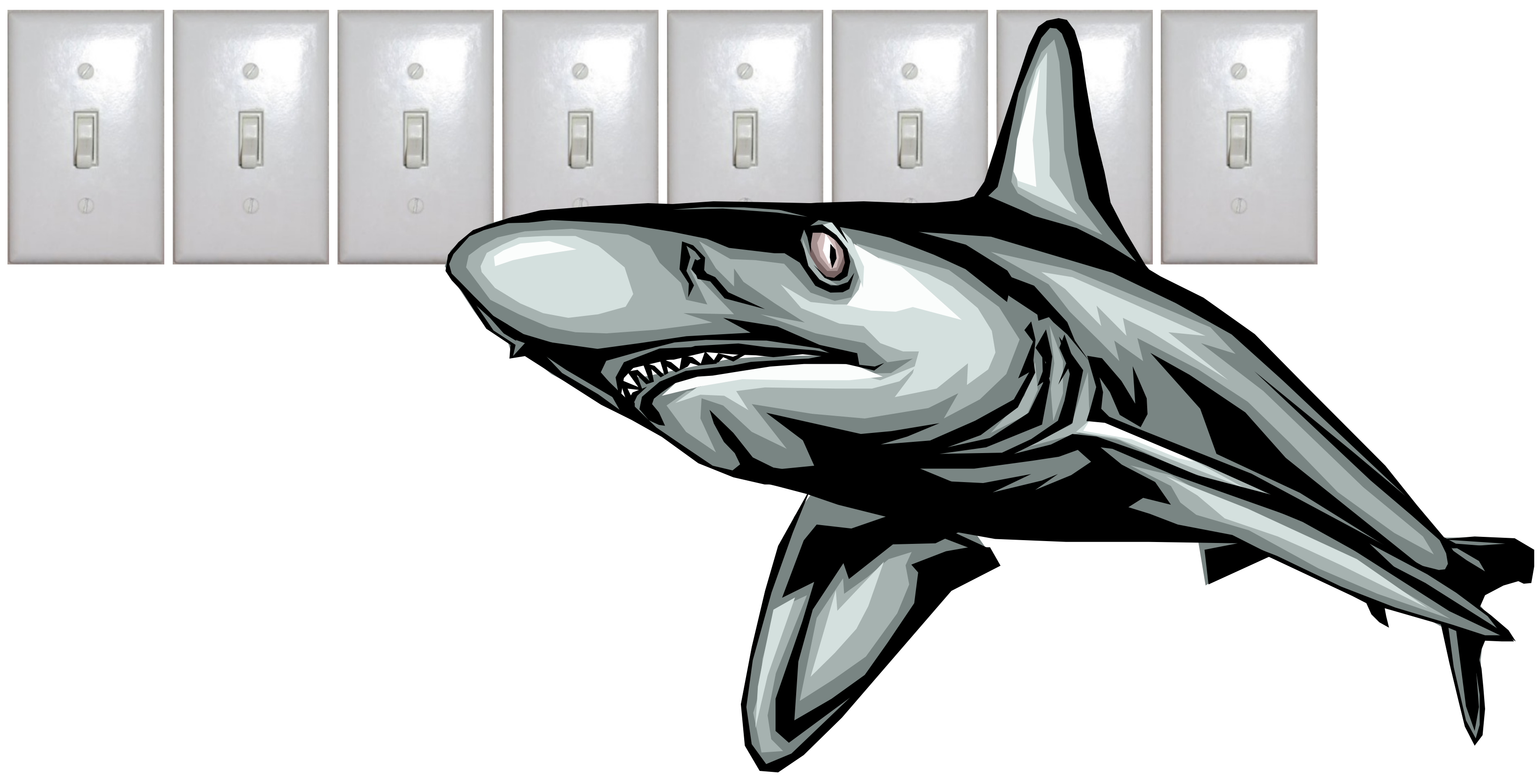
How big?





How many states can *n* bits represent?

(or the story of 18,446,744,073,709,551,615 grains of rice)



Data is represented via an encoding

American Standard Code for Information Interchange (ASCII)
= standard byte encoding used in PC's

01000001	= A	01100001	= a
01000010	= B	01100010	= b
01000011	= C	01100011	= c
01000100	= D	01100100	= d
01000101	= E	01100101	= e
01000110	= F	01100110	= f
01000111	= G	01100111	= g
01001000	= H	01101000	= h
01001001	= I	01101001	= i
01001010	= J	01101010	= j
01001011	= K	01101011	= k
01001100	= L	01101100	= l
01001101	= M	01101101	= m
01001110	= N	01101110	= n
01001111	= O	01101111	= o
01010000	= P	01110000	= p
01010001	= Q	01110001	= q
…		…	

Units of Size

Unit	Abbreviation	Size (bytes)
bit	b	1/8
byte	B	1
kilobyte	KB	2^{10} = 1,024
megabyte	MB	2^{20} = 1,048,576
gigabyte	GB	2^{30} = 1,073,741,824
terabyte	TB	2^{40} = 1,099,511,627,776
petabyte	PB	2^{50} = 1,125,899,906,842,624

In most cases, it's okay to approximate!

How small?

Units of Distance

Unit	Abbreviation	Fraction of a meter
meter	m	1
centimeter	cm	10^{-2} = 1/100
millimeter	mm	10^{-3} = 1/1,000
micrometer	µm	10^{-6} = 1/1,000,000
nanometer	nm	10^{-9} = 1/1,000,000,000
picometer	pm	10^{-12} = 1/1,000,000,000,000

Progression of Technology

Year	Feature Size
1971	10 µm
1975	3 µm
1982	1.5 µm
1985	1 µm
1989	800 nm
1994	600 nm
1995	350 nm
1997	250 nm
1999	180 nm
2002	130 nm
2004	90 nm
2006	65 nm
2008	45 nm
2010	32 nm
2012	22 nm

How large is a silicon atom? **~0.25 nm**

How fast?

Thinking About Speed

- Speed can be expressed in two ways:
  - How many things can you do in one second?
  - How long to do something once?
- Convenient units are typically used
  - 1 GHz instead of 1,000,000,000 Hz
  - 10 microseconds rather than 0.00001 seconds
  - When comparing measurements, convert units first!

Units of Frequency

Unit	Abbreviation	Cycles per second
hertz	Hz	1
kilohertz	KHz	$10^3 = 1{,}000$
megahertz	MHz	$10^6 = 1{,}000{,}000$
gigahertz	GHz	$10^9 = 1{,}000{,}000{,}000$

Units of Time

Unit	Abbreviation	Duration (seconds)
second	sec/s	1
millisecond	ms	10^{-3} = 1/1,000
microsecond	μs	10^{-6} = 1/1,000,000
nanosecond	ns	10^{-9} = 1/1,000,000,000
picosecond	ps	10^{-12} = 1/1,000,000,000,000
femtosecond	fs	10^{-15} = 1/1,000,000,000,000,000

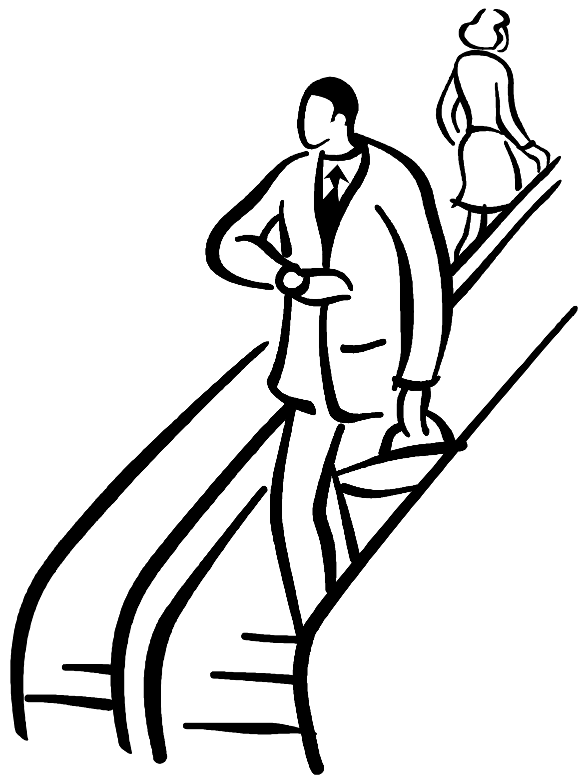
How far does light travel in one nanosecond? **0.3048 m**

How fast can we compute?

- Computation speed is limited by two factors:
  - Getting data to the CPU
  - Operating on the data in the CPU
- Two parts of moving data from here to there:
  - The delay between two locations
  - Amount of data you can move within a given amount of time
- Fundamentally, there's no difference:
  - Moving data from the processor to RAM
  - Saving a file to disk
  - Watching Netflix

Latency

Units in terms of time



Bandwidth

Units in terms of size per time

Discussion Point

- What's more important: latency or bandwidth?
  - Streaming audio (e.g., NPR broadcast over Web)
  - Streaming video (e.g., CNN broadcast over Web)
  - Audio chat
  - Video conferencing

How reliable?

Characterizing Reliability

“Nines”	Availability	Downtime (per year)
One nine	90%	36.5 d
Two nines	99%	3.65 d
Three nines	99.9%	8.76 h
Four nines	99.99%	52.56 m
Five nines	99.999%	5.256 m
Six nines	99.9999%	31.536 s

Time to roll up your sleeves…

Server

(Web? File?)

Source: http://www.studyzone.org/

Clients

Source: http://www.netofinancial.com/

Why Code HTML by Hand?

- The only way to learn is by doing
- WSIWYG editors…
  - Often generate unreadable code
  - Ties you down to that particular editor
  - Cannot help you manipulate backend databases
  - Little help when it comes to Javascript
- Hand coding HTML allows you to have finer-grained control
- HTML is demonstrative of other important concepts:
  - Structured documents
  - Markup
  - Metadata
  - …

Tips

- Edit files on your own machine, upload when you're happy
- Save early, save often, just save!
- Reload browser
- File naming
  - Don't use spaces!
  - Punctuation matters!
  - Case matters!