




CLOUD COMPUTING

Lecture 26 - CS2110 - Fall 2009

Cloud Computing

2

- Last time we talked about distributed computing
 - Basically, the technology of the Web
 - We use it all the time
- But what happens when these systems get very big?
 - The world has a *lot* of people in it
 - ... and plenty use Facebook



Cloud Computing concept

3

- What if we start to offload things from personal computers into the web?
 - Email becomes gmail, hotmail, ...
 - Files can be shared: Flickr, Picassa, ...
 - Online tools for creating documents: OpenOffice,
 - Potentially: put the whole world online

How much data?

4

- Telephone call: 56kbts/second or less
- Photo: 1 few megabytes
- DVD download: 600-700 Mbytes
- 11.5 Billion web pages in 2005, probably 30 Billion today
- Add to this “sensors” such as satellites, surveillance cameras, weather monitors, etc
- Adds up to a whole bunch of data, all of it online

Search

5

- In addition to “hosting” content, cloud systems need to be able to find what you want
 - “That adorable picture of Johnny when he tried to blow out the birthday candles and fell into the cake”
 - “Grateful Dead Live at Fillmore East on 1970-09-19”
 - “Index to coffee shops in Amsterdam”
 - “Best restaurant in Trumansburg New York”

Key elements to the cloud

6

- Huge amounts of data stored
- Indexed “offline” for fast retrieval
 - Basic idea is to associate vector of terms with object
 - For each set of terms, pre-compute the best objects
 - The more pages point to something, the more likely that something is to be what you want
 - But many, many refinements on this
- Many saw this opportunity but Google was first to do a really good job of answering queries

Cloud computing is big business

7

- Fastest sector for growth in the industry right now
 - Cloud computing systems are BIG
 - **One** Microsoft data center is 12x the size of a football field
 - Entirely packed with “containers” full of computers
 - Built near a dam: Power from next door
 - In a cool place: Not air conditioned; just uses outside air
 - A single system like this may have more horsepower than all the worlds supercomputers combined!



Clouds in Science

8

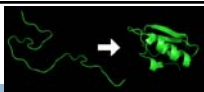
- Cloud computing is revolutionizing scientific research in many fields
 - Collect vast amounts of data
 - Pose questions about “reality” rather than needing to develop abstract models...



.... a new paradigm!

Examples: Folding

9

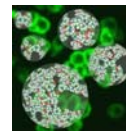


- Protein folding problem is computationally very hard and probably not really solvable
- But the physical world knows how to fold a protein and does it all the time.
- Could we somehow create software that uses a database of protein folding examples?
 - Software would take a protein as input
 - Then look for “familiar patterns” within it and see how those folded in the database

Examples: Smart drugs

10

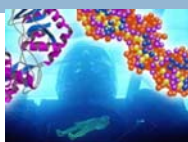
- Smart drug design
 - Given a target, like a virus, design a drug that can attack that target
 - Do it by looking for “matching shapes” in a massive collection of real-world data



Examples: Digital Human

11

- This project is at the Univ. of Michigan
 - Goal: create a digital simulator for a human being
 - Use it for virtual medical experiments, practicing surgery, understanding the human body



Examples: Bridging the Rift

12



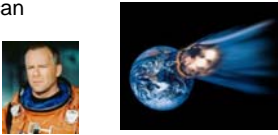
- A project to create a massive “library of life” documenting all the species in the Rift Valley on the border of Israel and Jordan
- Seek to understand life there... and relationships

The Bridging the Rift Center, straddling the border between Israel and Jordan, will someday house a database of information about every living thing on the planet.

Examples: Search for Asteroids

13


- Cornell is using Arecibo radio telescope to capture massive numbers of sky images
- Then writing software to search those images for asteroids (they occlude background stars)
 - ▣ Goal is to learn about our solar system
 - ▣ And maybe also learn about threats to the planet so that Bruce can save us...



Examples: Global Warming

14

- Researchers agree that something is definitely happening (this is *not* disputed at all)
- Less clear precisely why. Or what it implies
- Goal: Capture huge amount of data about climate that can be studied directly



Demo

15

- Live Objects
 - ▣ Created to combine cloud-hosted content with distributed computing concepts
 - ▣ Actually uses the same ideas we've learned about in cs2110!


Challenges of "big data"

16

- It has become much easier to collect data than to make sense of it
 - ▣ We're drowning in the stuff!
 - ▣ So the huge challenge is now to build tools that let us understand what we're seeing
- Much of Cloud Computing seems to focus on "silly" social applications like Twitter... but the bigger issues are universal and fascinating

Where's Waldo?

17

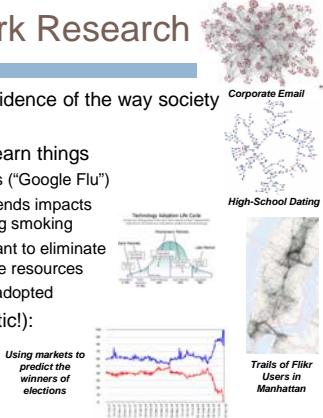


- When you access a file, like
 - ▣ http://en.wikipedia.org/wiki/World_Wide_Web
- ... how does your request get to a server?
- A cloud system has lots of computers sitting at the internet address "http://en.wikipedia.org"
 - ▣ In fact it may even have more than one data center!
- Requests are "load balanced" over the machines

Social Network Research


18

- The cloud captures evidence of the way society is organized
- We can study this to learn things
 - ▣ Like how the flu spreads ("Google Flu")
 - ▣ Or how your circle of friends impacts your likelihood of quitting smoking
 - ▣ Or who to treat if you want to eliminate TB but don't have infinite resources
 - ▣ How new products are adopted
- Spring Course (fantastic!): CS 2850 - Networks



Clouds: Pro and Con

19

- Cloud computing could give us anytime, anywhere access to all our stuff
 - ▣ No need to carry a heavy PC
 - ▣ Just take your phone and talk to it... they can convert voice to a query and fetch what you need!
- A personal butler with perfect memory 
- But search engines can't fetch what they haven't seen

Clouds: Pro and Con

20

- A personal butler with perfect memory...
 - ▣ Including stuff you might wish to forget
 - ▣ And stuff you didn't think it knew, like that your aunt Hilda died of a heart attack, and uncle Fred went insane
 - ▣ Employers and insurance companies are using this kind of data already and may do it more in the future!
 - ▣ Some employers don't like gay employees, or people who have had abortions, or who donate to PETA
- On the web, everything is public and permanent

Hey, I have a right to privacy!

21

- Really?
 - ▣ You do have a right to freedom from "search and seizure" and also to not "incriminate yourself"
 - ▣ But where does it say that you have a right to take down the Facebook pictures from Spring Break?
 - And even if you take them down... did Facebook keep copies?
- This is the core issue!
 - ▣ Technology is moving way faster than the law
 - ▣ Lessig: "East Code versus West Code"



Right now: A mess

22

- I don't have a right to spy on you with a parabolic microphone, but I do have a right to take pictures with a telephoto lens
- It isn't legal to tape a telephone conversation, but if I leave a message on your Android phone, it IS legal for Google to create a transcript. And index it.
- It isn't legal to download music from free services but people do it a lot



Googled

23

- One real worry is that free search could wipe out entire industries
 - ▣ Right now this seems farfetched, like believing that global warming could melt the ice at the poles
 - ▣ Yet newspapers are failing left and right, the music and film industries are making fewer films, small bookstores are going under
- The world changes. But does it necessarily get better?



Googled

24

- What about bad data?
 - ▣ E.g., if someone lies about you on the web
- And who's responsible?
 - ▣ For much of a week, Google images pulled up a racist image of Michelle Obama
 - They wouldn't take it down, pointing out that it wasn't on their web site... yet their web "images" search would find it and display it in thumbnail
 - Also asserted right of "Hot Girls.com" to free speech
 - ▣ Who was insulting the First Lady? Google? "Hot Girls.com"?
- What if an image search runs into child pornography:
 - ▣ Does this make you "guilty" of downloading child porno images?



Googled

25



- Google tools are free, right?
 - ▣ Actually, you are “selling them your data”
 - ▣ And they provide services in exchange
- They also place advertising, very selectively



Building a “mission critical cloud”

26

- The cloud works well for search, and saves money
 - ▣ By some metrics, cloud computing is 10x cheaper than computing with private computers in your office
 - ▣ So naturally there is a trend to shift medical records, banking, other “critical” things to the cloud
- Can we make the cloud secure and reliable enough?
 - ▣ A challenge that may require a mixture of laws and technology...

Building a “mission critical cloud”

27

- Examples of research issues:
 - ▣ Once data is in the cloud, how confident can we be that it won't “rot”?
 - ▣ Are cloud platforms at risk of viruses?
 - ▣ What if a big cloud company goes bankrupt?
 - ▣ If we “depend” on cloud systems, who is at fault if one of them has a bug
 - E.g. if a doctor makes a mistake because she saw erroneous data
 - Or if Wall Street stampedes and wipes out a company because its financial state was inaccurately reported?
- The list really goes on and on!

The Future is Coming...

28



- A future without secrets.... And someday, other people will be in charge at Google, Microsoft, Yahoo
- What if we have a new McCarthy era?
- Could Google figure out your deepest thoughts?
- Could organized criminals “data mine” Google?
- What about insider trading based on “snooping” digital data from people who are naïve?

So what can I do?

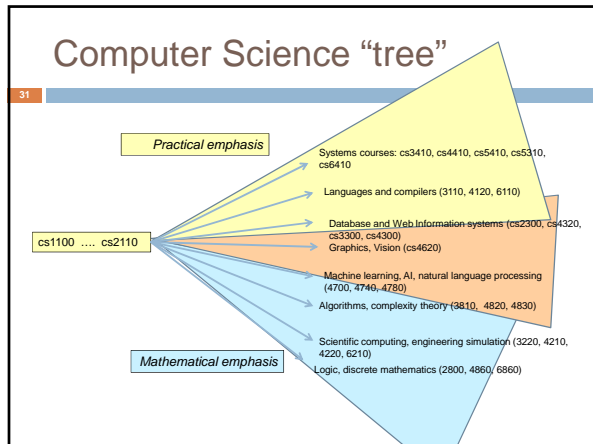
29

- Don't “opt out”... opt in but be an activist
 - ▣ Impossible to just walk away from a societal trend
- We need to find ways to push back
 - ▣ Change the laws. Cloud computing is bound by law.
 - ▣ Help others understand the issues
 - ▣ And build better computing systems!
 - CS and IS courses can help you learn how...

CS courses to consider

30

- CS 2850 - Networks
- CS 3110 – Functional Programming
- CS 3410 – Computer Architecture
- CS 3700 – Artificial Intelligence
- Helps to think in terms of “areas”
 - ▣ Systems (databases, operating systems, cloud)
 - ▣ Programming Languages
 - ▣ Machine learning, NLP, AI
 - ▣ Scientific computing
 - ▣ Graphics, computer vision



We hope you enjoyed CS2110 and learned some useful things too!

32

... Have a great winter vacation!

The slide features three photographs of winter scenes. The top-left photo shows a person in red gear climbing a steep, snow-covered mountain peak. The top-right photo shows a golden retriever standing in a snowy forest with tall trees. The bottom photo shows a person in winter gear skiing down a snowy slope.