CS 91: Cloud Systems & Datacenter Networks

Course Introduction

Kevin Webb
Swarthmore College
September 1, 2014
Instructor: Kevin Webb


- Hobbies: Making stuff (woodworking, ceramics, electronics), plants, PC games

- Research: Control platforms for networks

- I’m (relatively) new here!

- Please call me Kevin (or Professor/Dr. Webb)
Office Hours

• Monday: 11:30 – 12:30 (after class)
• Thursday: noon – 1:30
• By appointment

• 255 Science Center
Student Backgrounds

- CS 31
- Operating systems
- Networks
- Databases

- Some combination of the above
Course Format

• Primarily discussion of research papers

• Some days I’ll provide background info

• Expectations:
  – Read before class
  – Everyone will participate in discussion
Upcoming Schedule

• Wednesday:
  – Reading: Overview of cloud issues and concepts
  – Background for virtualization

• Friday:
  – MapReduce paper (for lab)

• Next week:
  – Virtualization papers
Full Schedule / Course Website


• Paper PDFs posted

• Link to online “text book”
  – Should be downloadable from on campus

• Schedule is tentative, shouldn’t change a ton
Other Course Resources

- Piazza
- Gradesource
- AWS (details later today)
Assignments

• Four labs (mostly 2 weeks)

• Labs meet on Fridays
  – Work in pairs, or solo
  – If possible, come to same lab as partner

• Labs go out on Friday, due midnight on Thursday before next lab

• Final project of your choice (2nd half of course)
Cloud Computing

• What is it?

What's your definition of cloud computing? Even the industry doesn't agree!

Linda posted on June 07, 2010 08:03  ★★★★★

Last week, I attended an industry meeting held by The ASCII Group. This meeting brought together over 100 technology business owners to share best practices and discuss trends in our industry.

There were several panel discussions moderated by ASCII members, with one of the most intriguing ones focusing on cloud computing. What made this panel so interesting was that the four panelists couldn't agree on a definition of cloud computing.
Cloud Computing

51% Of People Think Stormy Weather Affects 'Cloud Computing'

Most Americans Confused By Cloud Computing According to National Survey

Ninety-five percent of those claiming they never use the cloud actually do so via online banking and shopping, social networking, and storing photos and music.
Cloud Computing: Buzzword?

• Absolutely (also, “big data”)

• This course: uncover the technology behind the buzz.

• Combination of:
  – distributed systems, networks, OS, databases, web services, security, and many more!
Cloud Computing

• “The practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer.”
Cisco Official Icons (decades old)

• If you want to diagram a network, they have the icons for you.
Icons: Cisco Products

Router - Color and subdued
Router w/Silicon Switch
Wavelength

Workgroup Director
Network Management Appliance

Cisco MeetingPlace Express
Software-Based Router on File/Application Server

100BaseT Hub
uBR910 Cable DSU
CDDI/FDDI Concentrator

(12 pages of Cisco Product icons)
Example Diagram
Cloud abstractly represents:
- The Internet
- Someone else’s infrastructure
- Details we don’t need to worry about
Cloud Computing

• “The practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer.”

• Same properties:
  – (Usually) accessed over the Internet
  – Someone else manages the infrastructure
  – They provision it, keep it running, etc.
Why this material is exciting!

• Recent popularity / feasibility
  – Consumer Internet connections are fast
  – We’re still in the early stages: what’s next?
Why this material is exciting!

- Recent popularity / feasibility
- Opportunity for rapid innovation
  - The Internet is hard to change, despite challenges
  - Cloud / Data center systems evolve more rapidly
Why this material is exciting!

• Recent popularity / feasibility

• Opportunity for rapid innovation

• Enables unprecedented scale ("Big Data")
  – Provide resources on demand, only as needed
  – Solve old problems faster
  – Solve new problems that were intractable
Personalized Medicine

Cost per Genome

Moore's Law

NIH National Human Genome Research Institute

genome.gov/sequencingcosts
Personalized Medicine

• Perform data analysis on you

• Correlate with others to determine who is or isn’t likely to develop a disease

• Provide custom treatments
Business & Analytics

• Keep track of what customers are buying, looking at, ads clicked, etc.
  – Handle millions of purchases
  – Target sales/ads to users
  – Provide better products and services

• Analysis of markets
  – Stock / commodity trading
Large-scale Science

Also: brain simulations, climate forecasting, seismic prediction, etc...
Storage and Collaboration
___ as a Service

• Software: just run it in your browser
  – Google drive
  – Salesforce.com

• Infrastructure:
  – Amazon AWS
  – Microsoft Windows Azure
Misc. Data Collection / Processing
Misc. Data Collection / Processing

FiveThirtyEight
CS 91

• Peel away abstractions, see how it’s built

• Systems that are:
  – Large scale / distributed
  – High performance
  – Reliable
  – Secure
  – …
Amazon Web Services

• Graciously given us an education grant
  – $100 per person worth of time

• Gives us access to:
  – Elastic Compute Cloud (EC2)
  – Simple Storage Service (S3)
  – Elastic MapReduce (EMR)
  – Many more
Amazon Web Services

• These services charge by the hour.

• When you finish your work, **TURN IT OFF!**
• Systems that...
  – Google uses to process search data
  – Facebook uses to serve photos quickly
  – Amazon uses to store shopping cart information

• Many more that describe principles behind fast, reliable, parallel systems and networks...
  – Classics from the 70’s and 80’s
  – State of the art, recent research
Preview: Labs

• Process large corpus of movie data to make recommendations

• Control a (simulated) data center network

• Orchestrate message ordering in a distributed chat system
Preview: Final Project

• You get to choose what to work on

• Built on AWS or similar platform

• More details forthcoming...