Hello!

CS 744: BIG DATA SYSTEMS

Shivaram Venkataraman
Fall 2021
WHO AM I?

Assistant Professor in Computer Science

PhD at UC Berkeley: System Design for Large Scale Machine Learning

Industry: Google, Microsoft Research
Open source: Apache Spark committer

Call Me: Shivaram or Prof. Shivaram
COURSE LOGISTICS

Shivaram Venkataraman
Office hours: Thursday 11-noon, CS 7367 (or Zoom?)

TA: Yien Xu
Office hours: Mon 5-6pm, Zoom?

Discussion, Questions: Use Piazza!
TODAYS AGENDA

What is this course about?

Why are we studying Big Data systems?

What will you do in this course?
BRIEF HISTORY OF BIG DATA
“…Storage space must be used efficiently to store indices and, optionally, the documents themselves. The indexing system must process hundreds of gigabytes of data efficiently…”

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page
GOOGLE 2001

Commodity CPUs
Lots of disks
Low bandwidth network
Cheap!
Facebook’s daily logs: 60 TB

Google web index: 10+ PB
“scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets”

-- Jim Gray
GRAVITY WAVE DETECTION
Solar Flare Prediction Using Photospheric and Coronal Image Data.

~ 2 PB

Working with data from Solar Dynamics Observatory
[Brown et. al SDO Primer 2010]

Solar Flare Prediction Using Photospheric and Coronal Image Data.
[Jonas et. al American Geophysical Union, 2016]
Data Growth is Outpacing Computing Growth

Source: More Data, More Science and... Moore’s Law  [Kathy Yellick]
Google data centers in The Dulles, Oregon
Datacenter Evolution

Capacity:
~10000 machines

Bandwidth:
12-24 disks per node

Latency:
256GB RAM cache
Outage in Dublin Knocks Amazon, Microsoft Data Centers Offline

By: Rich Miller
August 7th, 2011

A lightning strike has caused issues for Amazon and Microsoft data centers in Dublin, affecting many sites using Amazon's BPOS (Business Product and Services) and Microsoft's Azure cloud services.

More on today's Gmail issue

Posted: Tuesday, September 01, 2009

Posted by Ben Treynor, VP Engineering and Site Reliability

Gmail's web interface had a widespread outage where people rely on Gmail for personal and professional communication. We're investigating a problem with the service. Thus, right up front, I'd like to say that we're aware of this and we're treating it as such. We've already taken action by implementing a fix, and as a result, the majority of users are now seeing improved performance.

Amazon EC2 and Amazon RDS Service Disruption

Our efforts to restore the services, and what we are doing to prevent this sort of issue from happening again in the future, is our priority. We will continue to provide updates as we work to minimize the impact on our customers.
The Joys of Real Hardware

Typical first year for a new cluster:

- ~0.5 overheating (power down most machines in <5 mins, ~1-2 days to recover)
- ~1 PDU failure (~500-1000 machines suddenly disappear, ~6 hours to come back)
- ~1 rack-move (plenty of warning, ~500-1000 machines powered down, ~6 hours)
- ~1 network rewiring (rolling ~5% of machines down over 2-day span)
- ~20 rack failures (40-80 machines instantly disappear, 1-6 hours to get back)
- ~5 racks go wonky (40-80 machines see 50% packetloss)
- ~8 network maintenances (4 might cause ~30-minute random connectivity losses)
- ~12 router reloads (takes out DNS and external vips for a couple minutes)
- ~3 router failures (have to immediately pull traffic for an hour)
- ~dozens of minor 30-second blips for dns

- ~1000 individual machine failures
- ~thousands of hard drive failures

slow disks, bad memory, misconfigured machines, flaky machines, etc.

Long distance links: wild dogs, sharks, dead horses, drunken hunters, etc.
How do we program this?
Cloud computing storage

- Hierarchy
  - NVM
  - In-memory storage (Memcached/Redis)
  - SSD storage
  - Hard disk
  - Tape

Optimizing data storage

- Compression
- Data coding

BIG DATA SYSTEMS

Software systems that help process large datasets!
Scalable Storage Systems
Datacenter Architecture
Resource Management
Computational Engines
Applications

Machine Learning
SQL
Streaming
Graph

Frameworks used
Reliability, performance, capacity
Manage compute resources
Hardware
COURSE SYLLABUS
BACKGROUND SURVEY: PAPER READING

- I am new to this! (38.6%)
- I have evaluated a few papers before but I am still learning how to do this (12.3%)
- I have some experience in critically reading papers but I can learn more (45.6%)
- I have significant experience!
FAMILIARITY WITH TOOLS

Focus in this course

Python: Popular!
PRIOR COURSES

- Operating Systems: 44 (77.2%)
- Distributed Systems: 8 (14%)
- Networking: 31 (54.4%)
- Databases: 39 (68.4%)
- Machine Learning: 39 (68.4%)
- Optimization Algorithms: 13 (22.8%)
WHAT DO YOU HOPE TO LEARN FROM THE COURSE?

Understanding of Big Data system architectures and approach towards their design

I hope to be able to design and deploy end to end machine learning pipelines that are designed keeping the tenets of low latency and high throughput in mind.

Tools and techniques to handle big data! My current research project has me handling large computations that I'm woefully unprepared for.

I want to learn about different systems and see if I can get a Data Science job.

I am interested in Systems research and would like to read and understand advanced papers related to it.

…
LEARNING OBJECTIVES

At the end of the course you will be able to

- Explain the design and architecture of big data systems
- Compare, contrast and evaluate research papers
- Develop and deploy applications on existing frameworks
- Design, articulate and report new research ideas
LEARNING OBJECTIVES

At the end of the course you will be able to

• Explain the design and architecture of big data systems
• Compare, contrast and evaluate research papers
• Develop and deploy applications on existing frameworks
• Design, articulate and report new research ideas
CLASS FORMAT

Schedule: http://cs.wisc.edu/~shivaram/cs744-fa21
Reading: ~1 paper per class

Review: Fill out review form (link posted on Piazza) by 9am
Discussion: In-class group discussion, submit responses within 24 hours

What if you cannot attend?
  Best 15 responses (out of ~22)
  Discussion: Student/TA/Prof (write their name!) and submit
HOW TO READ A PAPER: EXAMPLE

The Google File System

Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung
Google

ABSTRACT
We have designed and implemented the Google File System, a scalable distributed file system for large distributed data-intensive applications. It provides fault tolerance while running on inexpensive commodity hardware, and it delivers high aggregate performance to a large number of clients.

While sharing many of the same goals as previous distributed file systems, our design has been driven by observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system assumptions. This has led us to reexamine traditional choices and explore radical

1. INTRODUCTION
We have designed and implemented the Google File System (GFS) to meet the rapidly growing demands of Google’s data processing needs. GFS shares many of the same goals as previous distributed file systems such as performance, scalability, reliability, and availability. However, its design has been driven by key observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system design assumptions. We have reexamined traditional choices and explored radically different points in the design space.
What are your goals in taking the Big Data Systems course?
How similar / different are goals among students in the group?

What were your main takeaways from "How to Read a Paper"?
PRACTICE DISCUSSION SUMMARY

Takeaways
- Pass based approach is helpful filter
- Passes become progressively longer
- Literature survey as well!
- Google Scholar, Citeseer, Conference website

When do I stop writing? / Time

References
- Which of these to read?
ASSESSMENT

- Paper reviews: 10%
- Class Participation, Discussion: 10%
- Assignments (in groups): 20% (2 @ 10% each)
- Midterm exams: 30% (2 @ 15% each)
- Final Project (in groups): 30%
ASSIGNMENTS

Two homework assignments in Python using NSF CloudLab
- Assignment 0: Setup CloudLab account
- Assignment 1: Data Processing
- Assignment 2: Machine Learning

Short coding based assignments. Preparation for course project
Work in **groups of three**
EXAMS

- Two midterm exams
- Open book, open notes
- Synchronous, in-class
- Focus on design, trade-offs

More details soon
COURSE PROJECT

Main grading component in the course!

Explore new research ideas or significant implementation of Big Data systems

Research: Work towards workshop/conference paper
Implementation: Work towards open source contribution
Example: Research

*How do we scheduling distributed machine learning jobs while accounting for performance, efficiency, convergence?*

Example: Implementation

*Implement a new module in Apache YARN that allows GPUs to be allocated to machine learning jobs.*
COURSE PROJECT

Project Selection:
- List of course project ideas posted
- Form groups of three
- Bid for one or more ideas or propose your own!
- Instructor feedback/finalize idea

Assessment:
- Project introduction write up
- Mid-semester check-in
- Poster presentation
- Final project report
WAITLIST

- Class size is limited to ~75 for this semester
- Focus on research projects, discussion
- Limited undergraduate seats

If you are enrolled but don’t want to take, please drop ASAP!
If you are on the waitlist, we will admit students as spots open up

If you want to audit the class:

→ online videos, Piazza → review
→ forms
→ assignment
BEFORE NEXT CLASS

Join Piazza: https://piazza.com/wisc/fall2021/cs744

Complete Assignment 0 (see website)

Paper Reading: The Datacenter as a Computer