CS 744: BIG DATA SYSTEMS

Shivaram Venkataraman
Fall 2020
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: Tuesday 11-noon, BBCollaborate

TA: Saurabh Agarwal
Office hours: Wed 3-4pm, BBCollaborate

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
GOOGLE 1997
“... 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

IDO report*
“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.

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

Graph based on average 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 an outage for Amazon and Microsoft's BPOS (Business Productivity Online Suite) data centers in Dublin, Ireland. Many sites using Amazon's EC2 and RDS services are affected.

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 that affected many people relying on Gmail for personal and professional work. We're aware of the problem with the service. Thus, right up front, I'd like to say that we're taking this very seriously, and we're treating it as such. We've already discussed a list of things we intend to fix or improve as a result.

Amazon EC2 and Amazon RDS Service Disruption

In an effort to quickly restore service to as many customers as possible, we are taking the following steps:

1. Hotfix: We have released a hotfix that should improve the performance of the service. We will continue to monitor its effectiveness.
2. Monitoring: We are increasing the monitoring of the service to ensure that any further issues are quickly identified and addressed.
3. Communication: We will continue to keep customers updated on the progress of the resolution.

As with any service disruption, we understand the impact it can have on our customers. We apologize for the inconvenience and appreciate your patience as we work to resolve this issue.
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% packet loss)
~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?
BIG DATA SYSTEMS
<table>
<thead>
<tr>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Learning</td>
</tr>
</tbody>
</table>

**Computational Engines**

**Scalable Storage Systems**

**Resource Management**

**Datacenter Architecture**
COURSE SYLLABUS
How much experience do you have in critically reading and evaluating systems research papers?
72 responses

- 45.8% I have significant experience!
- 37.5% I am new to this!
- 15.3% I have evaluated a few papers before but I am still learning how to do this
- 4.8% I have some experience in critically reading papers but I can learn more
Select how familiar you with each of the topics below

- Python
- Java
- Multi-thread programming
- GPU/Cuda
- Apache Hadoop
- Apache Spark
- Deep Learning (Tensorflow/Pytorch)
- SQL Queries
Select any related prior courses you have taken and completed satisfactorily (either at UW-Madison or any other school)

72 responses

- Operating Systems: 50 (69.4%)
- Distributed Systems: 13 (18.1%)
- Networking: 28 (38.9%)
- Databases: 49 (68.1%)
- Machine Learning: 62 (86.1%)
- Optimization Algorithms: 22 (30.6%)
WHICH TIMEZONE ARE YOU WORKING FROM?

>90% are in Central
~few in Pacific
~few other time zones
WHAT DO YOU HOPE TO LEARN FROM THE COURSE?

Learn about the design decisions and challenges involved in building big data systems…

How to efficiently read a paper, how to write a paper through the project, learn more about big data stacks…

To get a better sense of what it covers. It sounds like a totally new (but interesting) field to…

I am interested in ML and would like to gain experience in dealing with large datasets.

To get a practical sense of how big data systems work, understand theoretical concepts…
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

Paper Review
Discussion
Assignment
Project
CLASS FORMAT

Schedule: [http://cs.wisc.edu/~shivaram/cs744-fa20](http://cs.wisc.edu/~shivaram/cs744-fa20)

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

(Best 15 out of 20 responses for both)
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 new ones.

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.
PRACTICE DISCUSSION!

https://forms.gle/oiWGjujBJG8iEwDS6
PRACTICE DISCUSSION SUMMARY
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
• Mostly synchronous
• 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

Peer Review!
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 and have a pressing case, send me an email

If you want to audit the class:
BEFORE NEXT CLASS

Join Piazza: https://piazza.com/wisc/fall2020/cs744

Complete Assignment 0 (see website)

Paper Reading: The Datacenter as a Computer