

# CS 744: BIG DATA SYSTEMS

Shivaram Venkataraman

Spring 2025

# 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 3-4pm, CS 7367

TA: Tareq Mohammed

Office hours: TBD

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



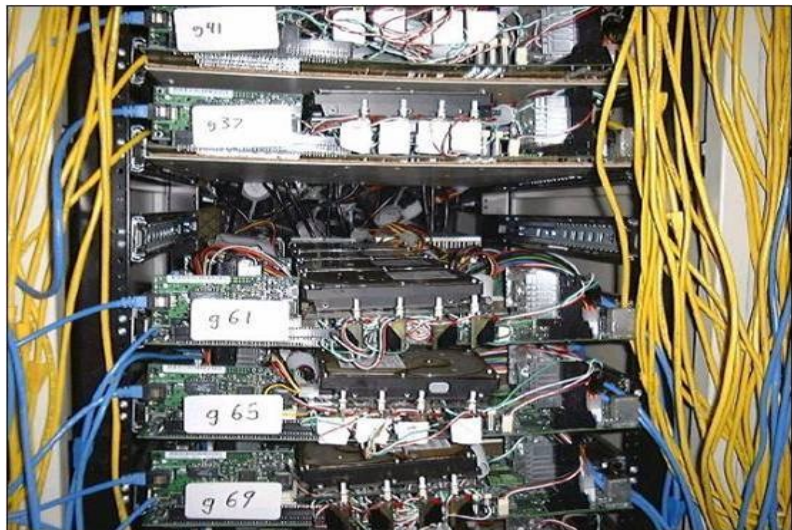
# DATA, DATA, DATA

“...**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

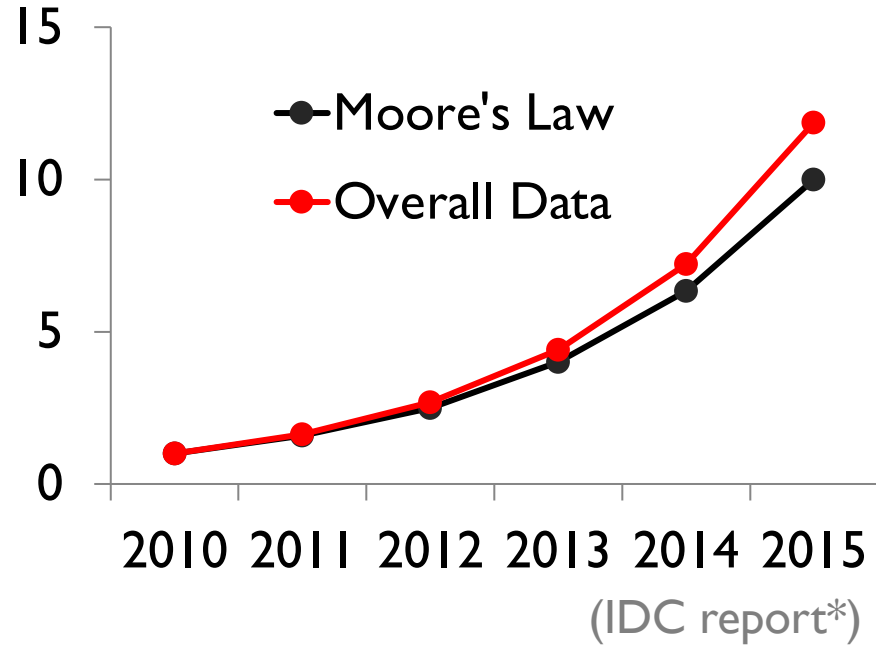
Low bandwidth network

Cheap !

# DATACENTER EVOLUTION

Facebook's daily logs: 60 TB

Google web index: 10+ PB





*The*  
**F O U R T H**  
**P A R A D I G M**

DATA-INTENSIVE SCIENTIFIC DISCOVERY

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLLE

“**scientific breakthroughs** will be powered by advanced computing capabilities that help researchers manipulate and explore **massive datasets**”

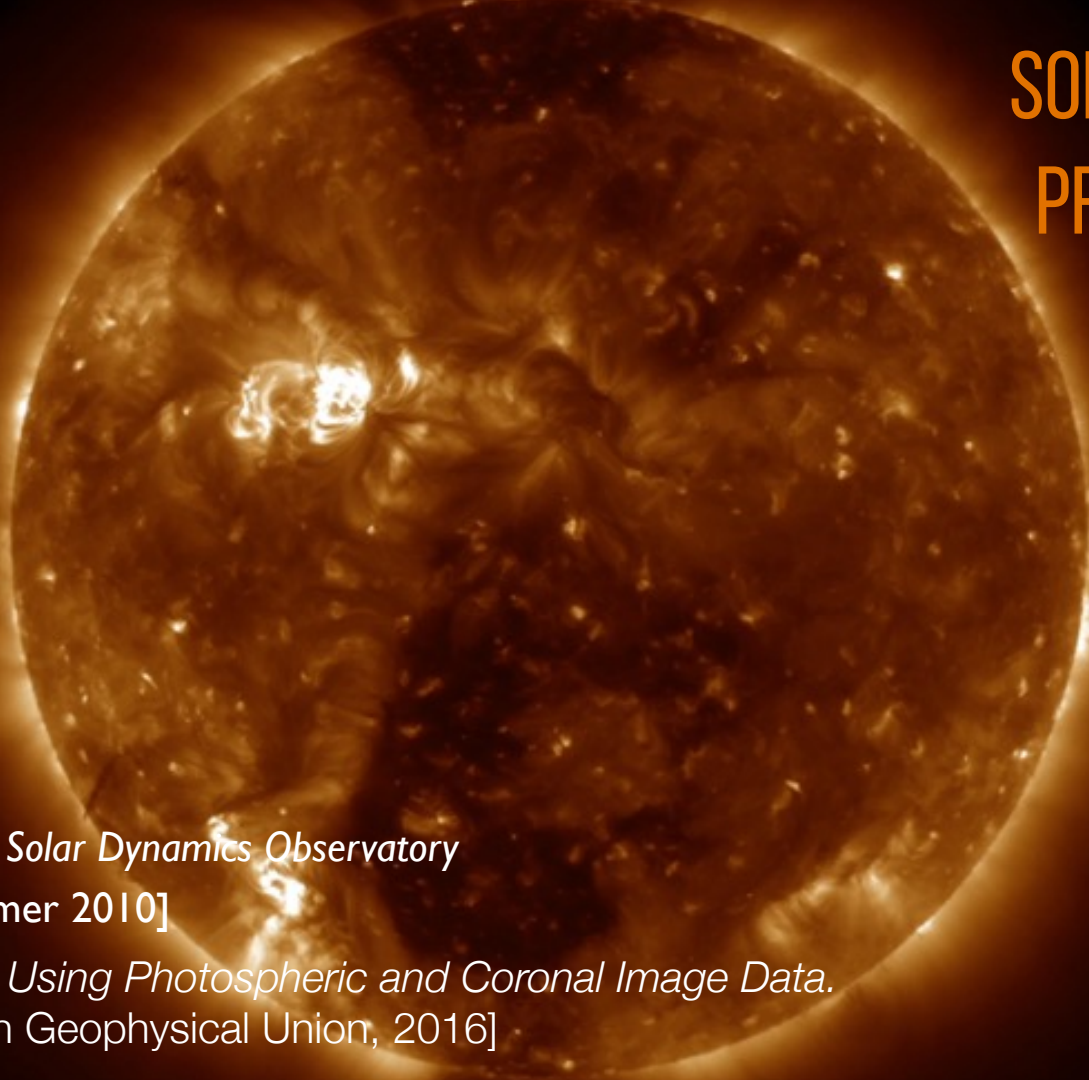
-- Jim Gray

# GRAVITY WAVE DETECTION



# SOLAR FLARE PREDICTION

~ 2 PB

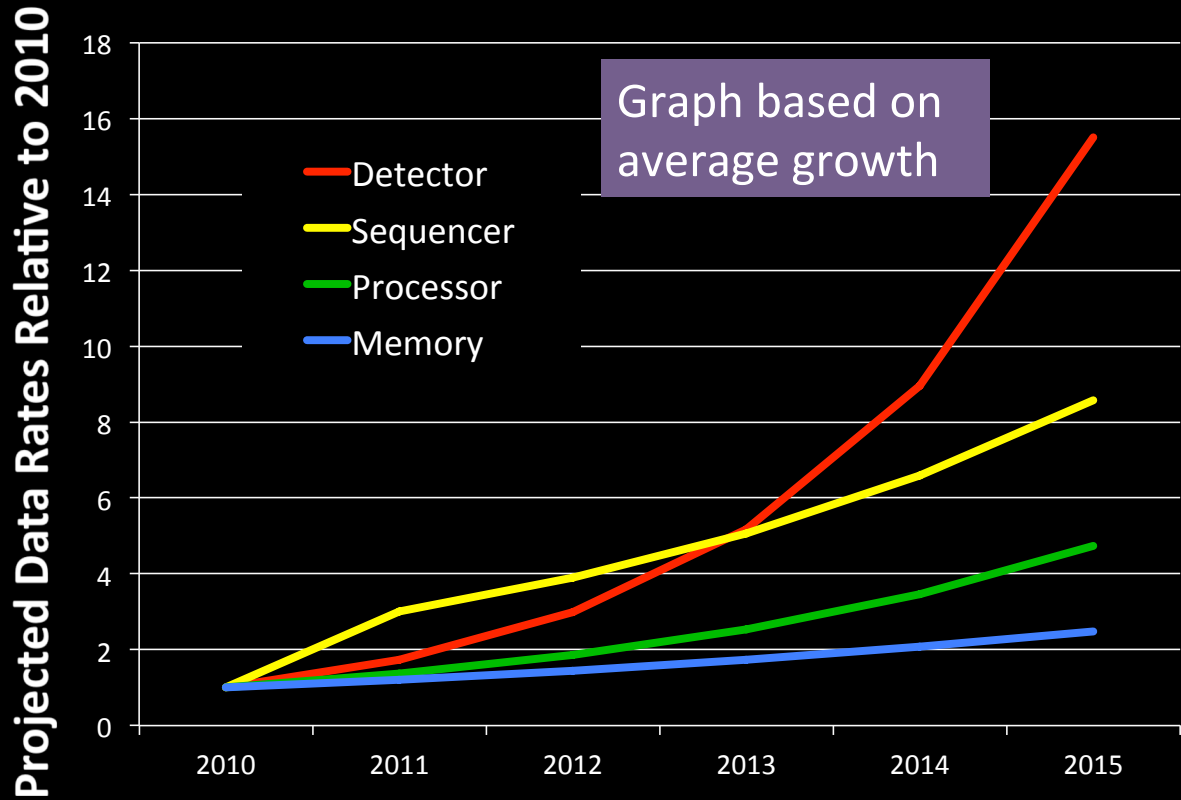


*Working with data from Solar Dynamics Observatory*

[Brown et. al SDO Primer 2010]

*Solar Flare Prediction Using Photospheric and Coronal Image Data.*

[Jonas et. al American Geophysical Union, 2016]



Source: More Data, More Science and... Moore's Law [Kathy Yellick ]

# LARGE ML MODELS

**LAION-5B: A NEW ERA OF  
OPEN LARGE-SCALE MULTI-  
MODAL DATASETS**



**GPT-3**

50 TB high resolution  
images

45 TB CommonCrawl  
data (2016-2019)

# DATACENTER EVOLUTION



Google data centers in The Dalles, 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



## Official Gmail Blog

News, tips and tricks from Google's Gmail team and friends.

557

520



A lightning strike has caused an outage for Amazon and Microsoft's BPOS (Business Process Outsourcing) sites using Amazon and Microsoft's BPOS (Business Process Outsourcing) 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 today. As many people rely on Gmail for personal and professional communication, this is a serious problem with the service. Thus, right up front, I'd like to say we're treating it as such. We've already thought through a list of things we intend to fix or improve as a result of this event.

## Amazon EC2 and Amazon RDS Service Disruption

In order to restore functionality to all affected services, we would like to share more details with our customers about the events that caused this outage, our efforts to restore the services, and what we are doing to prevent this sort of issue from happening again. We are also sorry for the inconvenience caused by this event, and as with any significant service issue, our intention is to share the details of what happened.



Sign Up

Entire Site

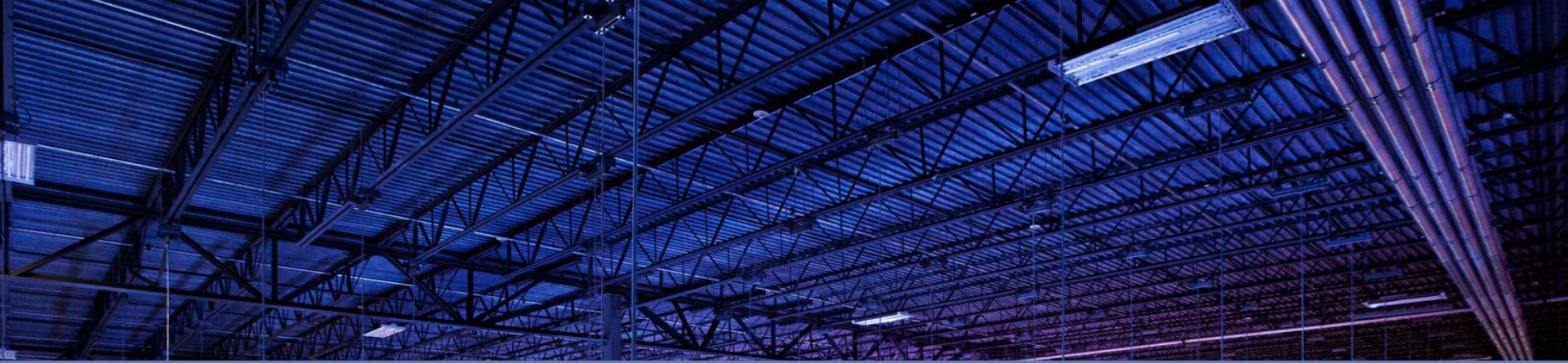
# The Joys of Real Hardware

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.**

**JEFF DEAN @ GOOGLE**

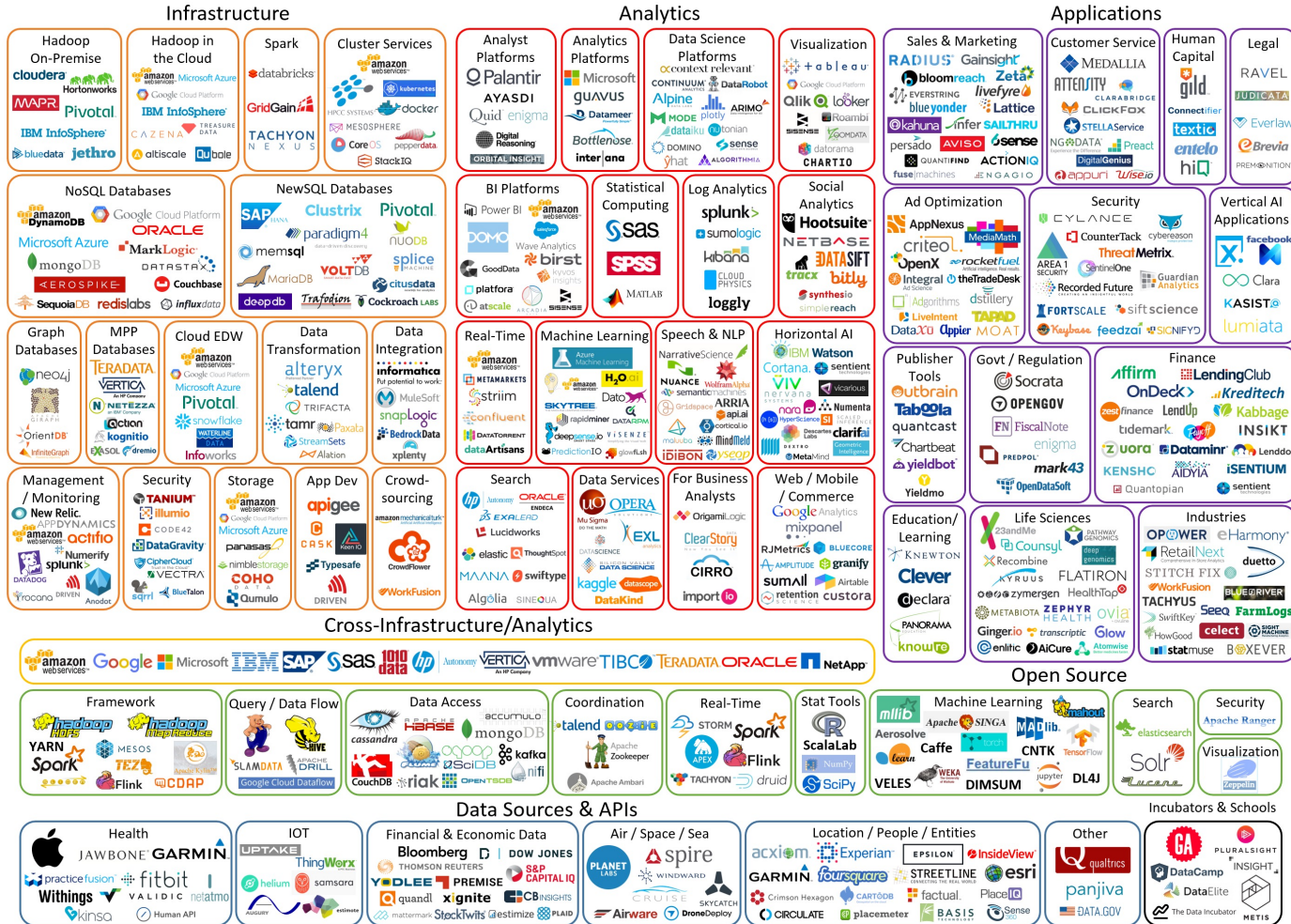


How do we program this ?



# BIG DATA SYSTEMS

# Big Data Landscape 2016 (Version 3.0)



## Applications

Machine Learning

SQL

Streaming

Graph

Computational Engines

Scalable Storage Systems

Resource Management



Datacenter Architecture



# COURSE SYLLABUS

# BACKGROUND SURVEY

Take the survey!

<http://tinyurl.com/cs744-sp25-bgs>



# FAMILIARITY WITH TOOLS

# PRIOR COURSES

**WHAT DO YOU HOPE TO LEARN FROM THE COURSE?**

# 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

# 744 VS 544?

## CS 544: Introduction to Big Data Systems

### **Learning Objectives**

- *Deploy distributed systems for data storage and analytics*
- *Demonstrate competencies with tools and processes...*
- *Write programs that use distributed platforms to efficiently analyze large datasets*
- *Produce meaning from large datasets by training machine learning models...*
- *Measure resource usage and overall cost of running distributed programs*
- *Optimize distributed analytics programs to reduce resource consumption...*
- *Demonstrate competencies with cloud services designed to store datasets ..*

# CLASS FORMAT

Schedule: <http://cs.wisc.edu/~shivaram/cs744-sp25>

Reading: ~1 **paper** per class. We will create reading groups (Canvas)!

Review: Fill out **review form** (link posted on Piazza) by 12pm

Discussion: **In-class group discussion**, submit responses within 24 hours

What if you cannot attend?

Best 15 responses (out of ~22)

# 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 rad-

### 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/ak3XxVkbSoHKDrc8A>

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

# 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 assignments. Preparation for course project

Work in **groups of three or four**

# EXAMS

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

More details including sample papers soon

**WHAT ABOUT CHATGPT?**

# 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

# COURSE PROJECT EXAMPLES

## Example: Research

*How do we schedule 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:

- Some course project areas posted
- Form groups
- 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 ~72 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

Meet me in office hours or on Thursday after class if reqd.

If you want to audit the class:

# BEFORE NEXT CLASS

Join Piazza: <https://piazza.com/wisc/spring2025/cs744>

Complete Assignment 0 (see website, Piazza)

Paper Reading: The Datacenter as a Computer