### CS 744: BIG DATA SYSTEMS

Shivaram Venkataraman Spring 2024

### WHO AM 1?

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 3-4pm, CS 7367

TA: Tzu-Tao (Tommy) Chang

Office hours: Monday 5:30pm - 6:30pm, CS 1330

Wednesday 5:30pm - 6:30pm, CS 1330

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**



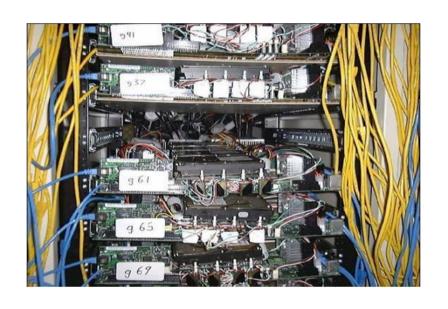
## DATA, DATA, DATA

"...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

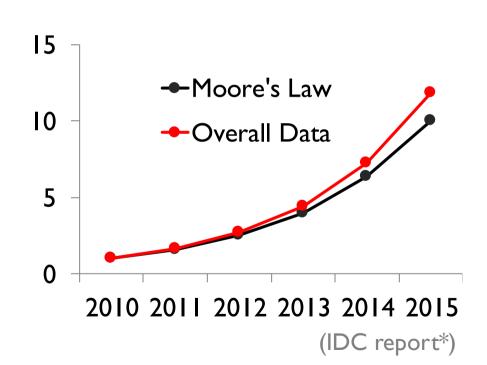
Low bandwidth network

Cheap!

#### DATACENTER EVOLUTION

Facebook's daily logs: 60 TB

Google web index: 10+ PB





The
FOURTH
PARADIGM

DATA-INTENSIVE SCIENTIFIC DISCOVERY

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

-- Jim Gray

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLL



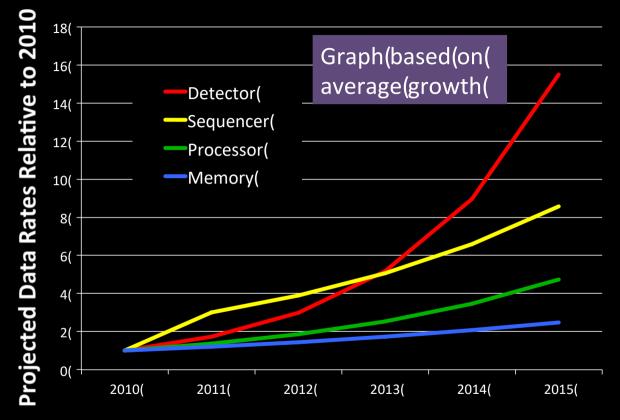


~ 2 PB

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

50 TB high resolution images



45 TB CommonCrawl data (2016-2019)

# DATACENTER EVOLUTION



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





Official Gmail Blog

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

A lightning strike has cau for Amazon and Microso Microsoft's BPOS (Busin

#### many sites using Amazo More on today's Gmail issu

Posted: Tuesday, September 01, 2009

Posted by Ben Treynor, VP Engineering and Sit

people rely on Gmail for personal and profession

problem with the service. Thus, right up front, I'd and we're treating it as such. We've already thor a list of things we intend to fix or improve as a re

> nctionality to all affected services, we would like to share more details with our customers about the events th our efforts to restore the services, and what we are doing to prevent this sort of issue from happening again. pted by this event, and as with any significant service issue, our intention is to share the details of what happe



#### 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** 

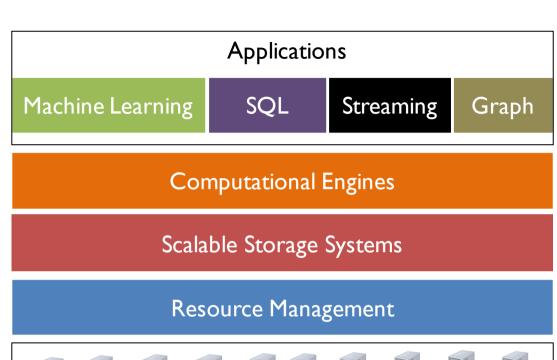


### How do we program this?



# **BIG DATA SYSTEMS**

#### Big Data Landscape 2016 (Version 3.0) Infrastructure **Analytics Applications** Data Science Human Hadoop Hadoop in Analyst Analytics Sales & Marketing Customer Service Visualization Spark Cluster Services Legal On-Premise the Cloud Platforms Platforms Capital Platforms RADIUS' Gainsight MEDALLIA +++ + a b l e a context relevant \*databrick Q Palantir Microsoft Abloomreach Zeta RAVEL CONTINUUM To DataRobo & kubernet ATTENCITY ... A EVERSTRING livefyre JUDICATA **QUAVUS** AYASDI CLARABRIDG MAPR Pivotal Qlik @ looker blue vonder Lattice GridGain in docker **CLICKFOX** IBM InfoSphere Connectifi Quid enigm Datameer MODE plothy ■Roamb Everlay @kahuna \_infer SAILTHRU MESOSPHERE STELLAService textic SISSINGS ZOOMDATA IBM InfoSphere TACHYON Digital Reasoning Core OS pepperdata Bottlenose. persado AVISO Ósense NG⇔DATA Preac ODOMINO Sense @Brevia entelo bluedata jethro 🐧 altiscale 🔃 bale ORBITAL INSIGHT inter ana S QUANTIFIND ACTIONIC Stack IQ That ALGORITHMIA CHARTIO **DigitalGenius** PREM®NITION hill appuri Wise.io NewSOL Databases BI Platforms Statistical Log Analytics Social **NoSQL Databases** Ad Optimization Security Vertical Al Analytics Computing amazon Ooogle Cloud Platforn Power BI amazon Clustrix Pivotal splunk> AppNexus MediaMat BCYLANCE Applications Hootsuite **S**sas ORACLE # paradigm4 sumologic sumologic CounterTack cybereason Microsoft Azure MarkLogic NETBASE criteo Threat Metrix. kıbana **DATASIFT** splice ≉birst DATASTAX SPSS OpenX \*rocketfuel mongoDB GoodData SentinelOne Recorded Future tracx bitly Integral () the Trade Desk O Clara **∢EROSPIKE** Couchbase citusdata platfora MATLAB **synthesio** deapdb Trafodion Cockroach LABS dstillery SequoiaDB redislabs @ influxdata loggly KASIST (1) atscale \*FORTSCALE \*siftscience SKeybase feedzai WSICNIFY DataXU Appier Graph MPP Data Data Speech & NLP Cloud EDW Real-Time Machine Learning Horizontal Al Databases Databases Transformation Integration amazon Watson Publisher Govt / Regulation Finance amazon informatica sentient alteryx **⊿ffirm IIILending**Club meo4j Tools METAMARKETS Socrata NUANCE VERTICA √ vicariou Microsoft Azure •talend Outbrain OnDeck> "Kreditech **Sstriim** MuleSoft Dato Pivotal (7) OPENGOV N NETEZZA TRIFACTA ngro 🕞 🕰 Numenta Tab<sub>2</sub>0la finance Lendlin Kabbage snapLogic Oction tamr mp quantcast FN FiscalNote Descartes clarifai tidemark (NSIKT Bedrock Data OrientDB DATATORRENT kognitio StreamSets MindMeld **7**Chartbeat Z UOra Dataminr Me Lenddo EXASOL Odremio dataArtisans PREDPOL" Infoworks Alation xplenty DIBON (2) △ yieldbot mark43 AIDYIA ISENTIUM Search Data Services For Business Web / Mobile Management Security Storage App Dev Crowd-\*\*\*\*\* OpenDataSoft Quantopian sentient Yieldmo / Monitoring TANIUM Analysts / Commerce sourcing apigee Google Analytics New Relic. 35 EXALERD Origamil.ogic Industries Life Sciences Education/ mixpanel C CODE 42 amazon OCTIFIO Lucidworks EXL PATHWAY GENOMICS OP@WER @Harmony CASK Keen IO Learning ClearStoru DataGravity panasas RJMetrics BLUECOR Numerify elastic ThoughtSpot RetailNext KNEWTON splunk> AMPLITUDE ( granify nimblestors Typesafe CIRRO STITCH FIX \*VECTRA M∧∧N∧ Ø swiftype Clever KYRUUS FLATIRON COHO sumal Airtable oeoezymergen HealthTap⊖ retention custora TACHYUS Seeo FarmLogs Qumulo import (b) @eclara METABIOTA ZEPHYR Cross-Infrastructure/Analytics PANORAMA HowGood celect @ MACHINE Ginger.io \* transcriptic Glow Cenlitic AiCure A. At statmuse B@XEVER amazon Google # Microsoft IBM SAX SSAS III (M) Amon VERTICA VMWATE TIBC TERADATA ORACLE Netapp **Open Source** Framework Query / Data Flow Data Access Coordination Real-Time Stat Tools Machine Learning Security accumulo HBASE mongoDP Apache SINGA MADlib. STORM Spork Anache Ranger talend Aerosolve cassandra ScalaLab YARN A MESOS Apache Zookeeper Caffe CNTK TensorFlo Visualization Solr SLAMDATA APACHE Spark TEZ VELES DIMSUM JUDYTER DL4J \*\* TACHYON - druid Flink QCDAP Apache Ambari **Data Sources & APIs Incubators & Schools** Location / People / Entities Health IOT Financial & Economic Data Air / Space / Sea Bloomberg D | DOW JONES PLURALSIGHT TAWBONE GARMIN **△** spire acxiem Experian Epsilon InsideView **Q** qualtrics DataCamp INSIGHT YEDLEE PREMISE CAPITAL IQ GARMIN GUISCHOLD STREETLINE @esri WINDWARD practice fusion # fitbit in helium in samsara 24 DataElite Crimson Hexagon CARTODB Factual. Place panjiva Withings VALIDIC netatmo quandl xignite CBINSIGHTS DATA.GOV The Data Incubator V Human API Airware ToneDeploy SteckTwits Gestimize RPLAID CIRCULATE | placemeter © Matt Turck (@mattturck), Jim Hao (@jimrhao), & FirstMark Capital (@firstmarkcap) FIRSTMARK Last Updated 3/23/2016





# **COURSE SYLLABUS**

#### **BACKGROUND SURVEY**

Take the survey! http://tinyurl.com/cs744-sp24-bgs

# **FAMILIARITY WITH TOOLS**

# **PRIOR COURSES**



### **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-sp24

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

Review: Fill out review form (link posted on Piazza) by I2pm

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

What if you cannot attend?

Best 15 responses (out of ~22)

# **COURSE FORMAT**

Recordings?

Important: In-class participation!

### **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/vcCokb4df5xb5hSBA

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 @ I 5% each)
- Final Project (in groups): 30%

### **ASSIGNMENTS**

Two homework assignments in Python using NSF CloudLab

- Assignment 0: Setup CloudLab account
- Assignment I: 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 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:

- Some 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 ~80 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: <a href="https://piazza.com/wisc/spring2024/cs744">https://piazza.com/wisc/spring2024/cs744</a>

Complete Assignment 0 (see website, Piazza)

Paper Reading: The Datacenter as a Computer