# 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

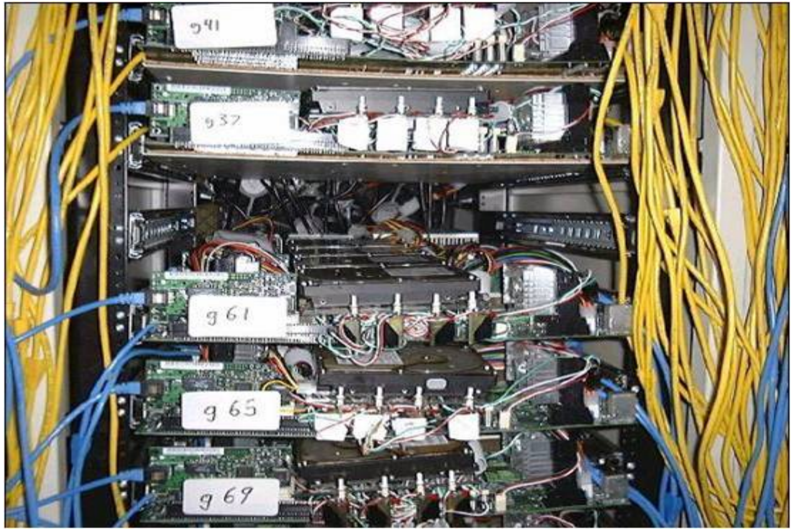
# DATA, DATA, DATA

"…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

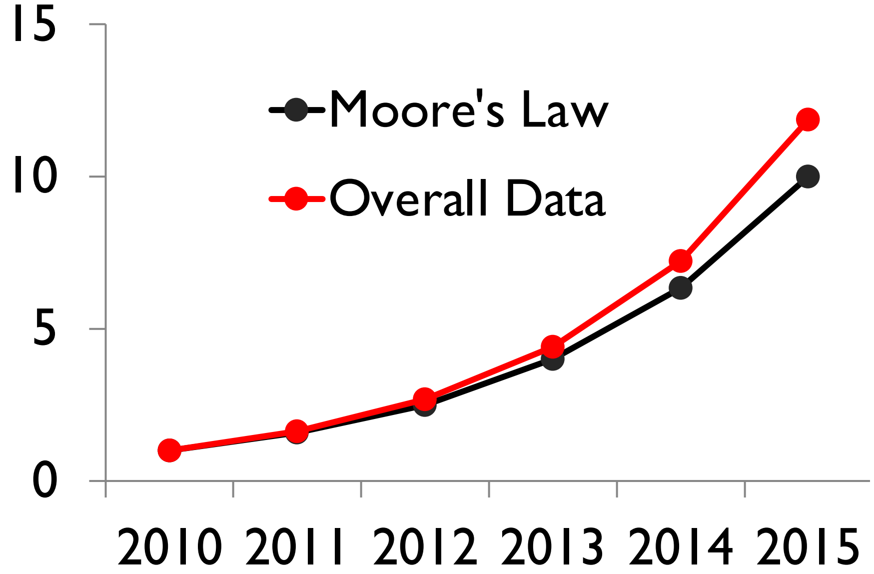
Low bandwidth network

Cheap !

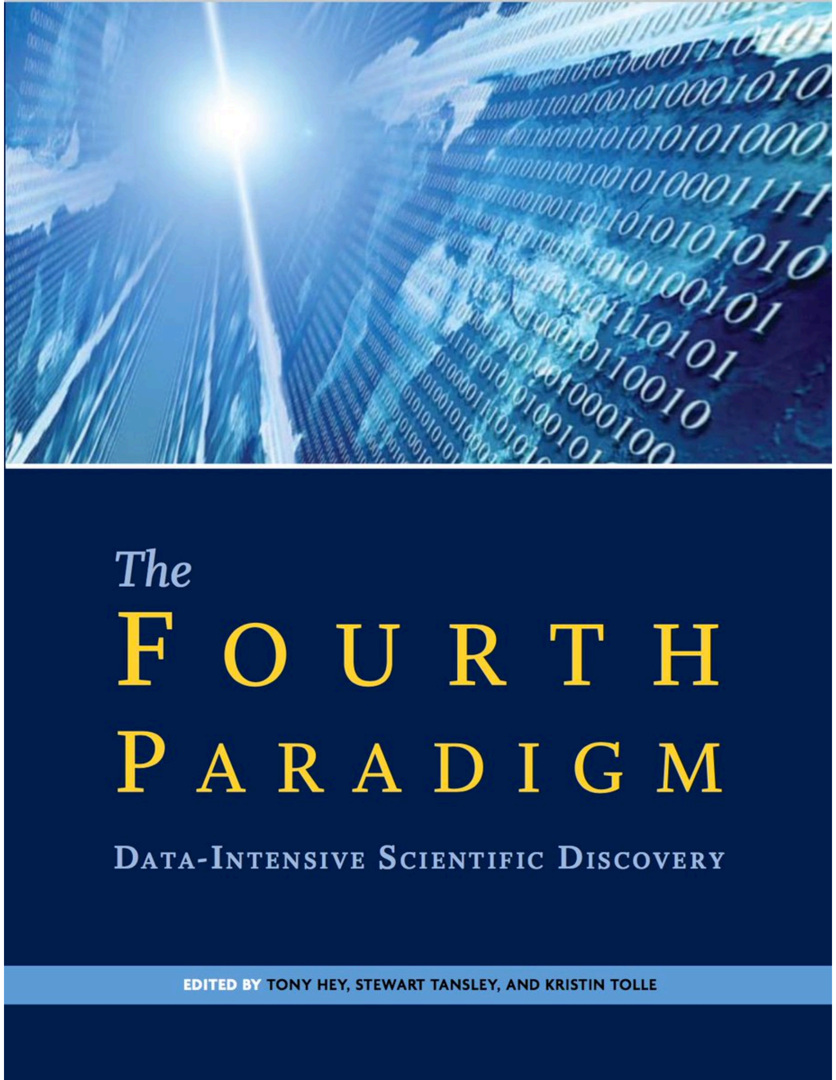
# DATACENTER EVOLUTION

Facebook's daily logs: 60 TB

Google web index: 10+ PB




(IDC report*)

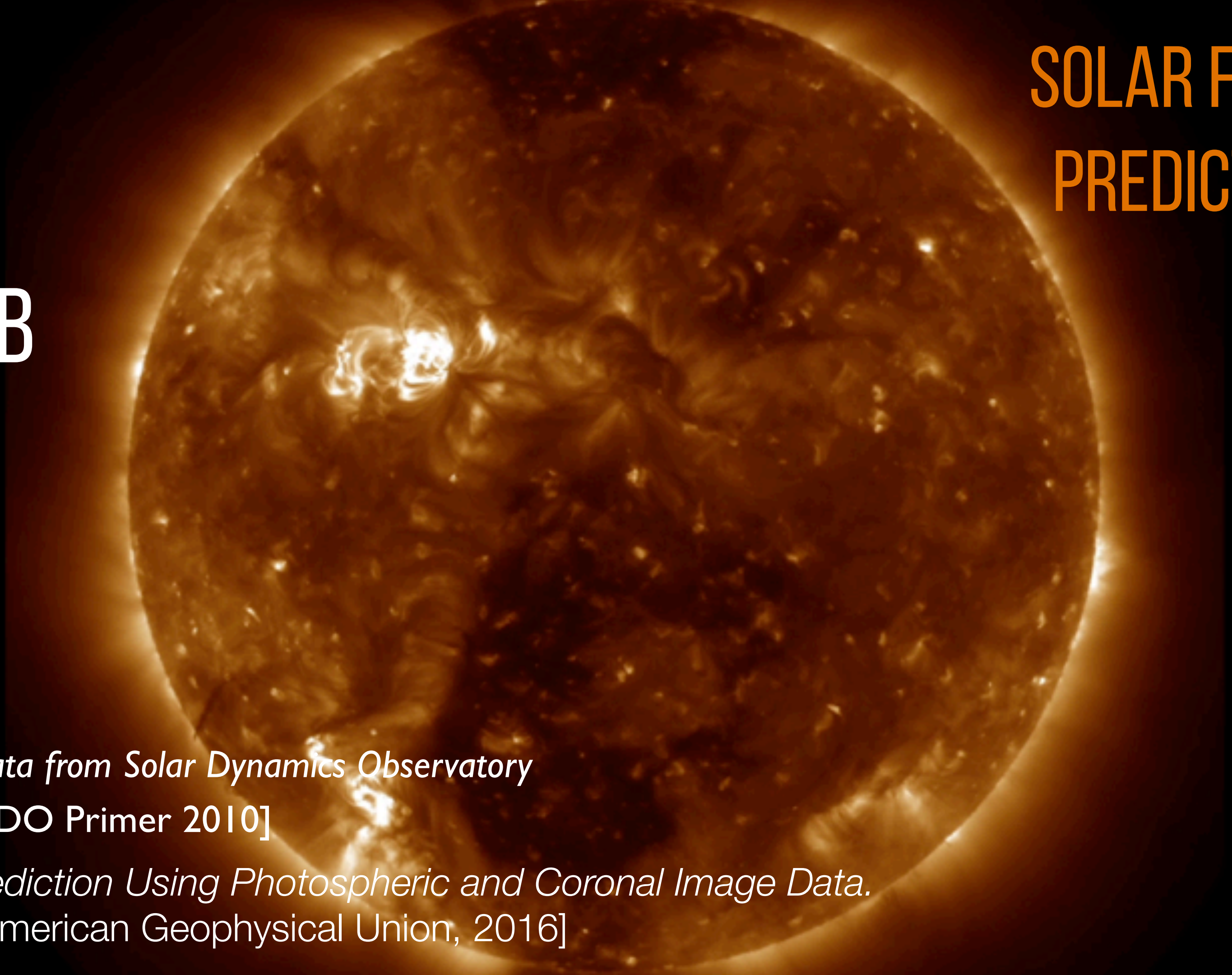



"scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets"

-- Jim Gray

# GRAVITY WAVE DETECTION

# SOLAR FLARE PREDICTION



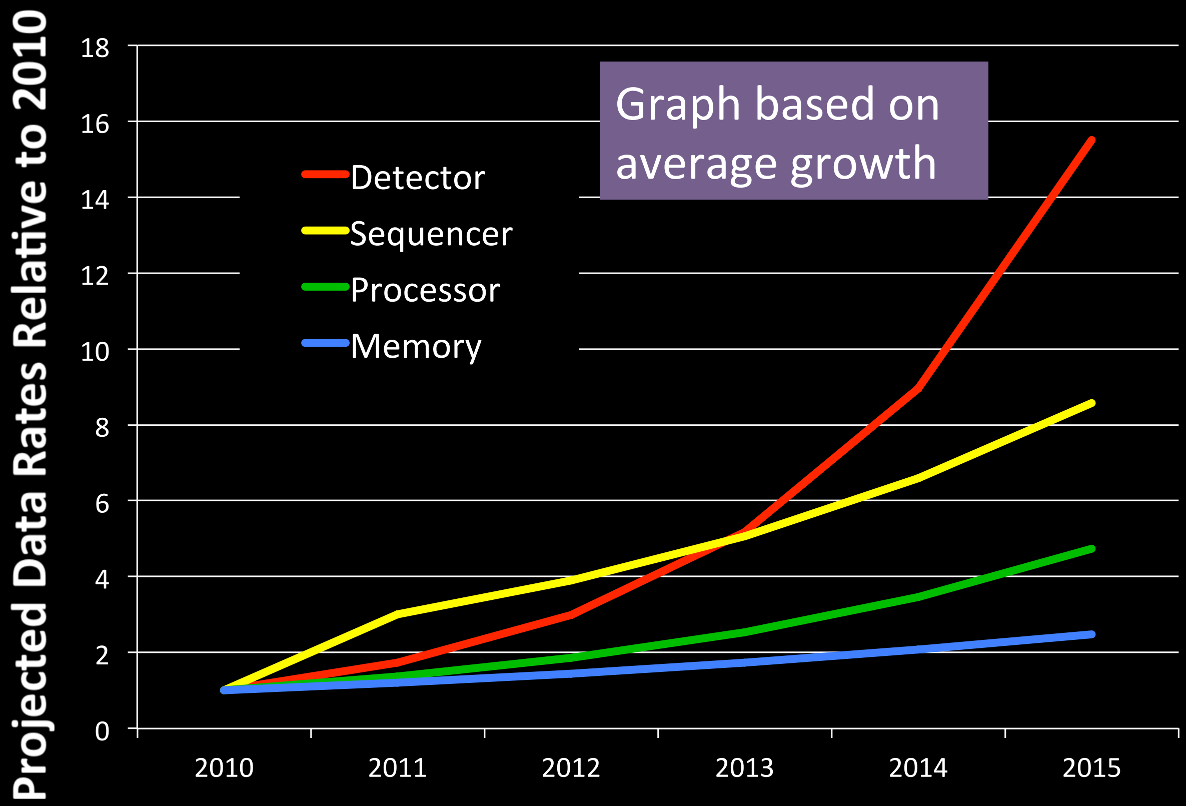
~ 2 PB

*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 ]

# 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*

557 Like

520 Tweet

A lightning strike has cau
for **Amazon** and **Microso**
many sites using Amazo
Microsoft's BPOS (Busin

# Official Gmail Blog

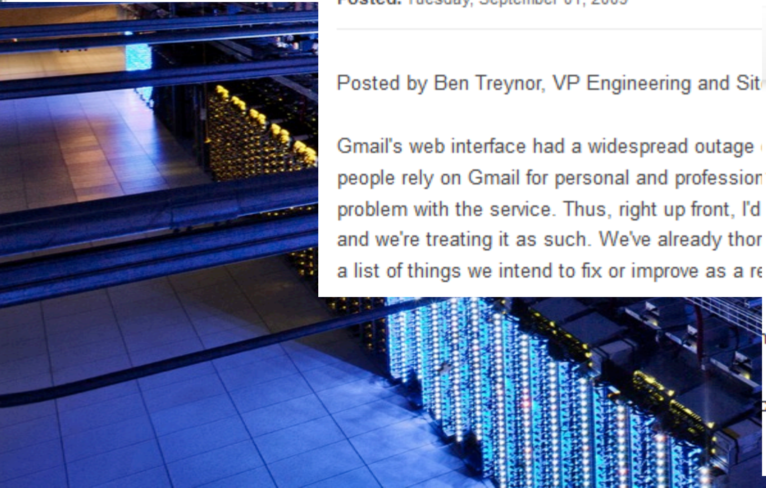
News, tips and tricks from Google's Gmail team and friends.

## More on today's Gmail issu

**Posted:** Tuesday, September 01, 2009

Posted by Ben Treynor, VP Engineering and Sit

Gmail's web interface had a widespread outage
people rely on Gmail for personal and professior
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



Sign Up

Entire Site

# Amazon EC2 and Amazon RDS Service Disruption

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.
cted 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

### Hadoop On-Premise
Cloudera, Hortonworks, MapR, Pivotal, IBM InfoSphere, bluedata, jethro

### Hadoop in the Cloud
amazon web services, Microsoft Azure, Google Cloud Platform, IBM InfoSphere, Cazena, Treasure Data, altiscale, Qubole

### Spark
databricks, GridGain, Tachyon Nexus

### Cluster Services
amazon web services, HPCC Systems, kubernetes, docker, Mesosphere, CoreOS, pepperdata, StackIQ

### NoSQL Databases
amazon DynamoDB, Google Cloud Platform, Oracle, Microsoft Azure, MarkLogic, mongoDB, DataStax, Aerospike, Couchbase, SequoiaDB, redislabs, influxdata

### NewSQL Databases
SAP HANA, Clustrix, Pivotal, paradigm4, NuoDB, memsql, splice machine, MariaDB, VoltDB, citusdata, deepdb, Trafodion, Cockroach Labs

### Graph Databases
neo4j, Giraph, OrientDB, InfiniteGraph

### MPP Databases
Teradata, Vertica, Netezza, Actian, kognitio, EXASOL, dremio

### Cloud EDW
amazon web services, Google Cloud Platform, Microsoft Azure, Pivotal, snowflake, Waterline Data, Infoworks

### Data Transformation
alteryx, talend, Trifacta, tamr, Paxata, StreamSets, Alation

### Data Integration
informatica, MuleSoft, snapLogic, Bedrock Data, xplenty

### Management / Monitoring
New Relic, AppDynamics, amazon web services, actifio, Numerify, splunk, Datadog, Rocana, Driven, Anodot

### Security
Tanium, illumio, Code42, DataGravity, CipherCloud, Vectra, sqrrl, BlueTalon

### Storage
amazon web services, Google Cloud Platform, Microsoft Azure, panasas, nimblestorage, Cohodata, Qumulo

### App Dev
apigee, Cask, Keen IO, Typesafe, Driven

### Crowdsourcing
amazon mechanical turk, CrowdFlower, WorkFusion

## Analytics

### Analyst Platforms
Palantir, Ayasdi, Quid, enigma, Digital Reasoning, Orbital Insight

### Analytics Platforms
Microsoft, guavus, Datameer, Bottlenose, interana

### Data Science Platforms
context relevant, Continuum Analytics, DataRobot, Alpine, plotly, Arimo, Mode, dataiku, Nutonian, Domino, Sense, yhat, Algorithmia

### Visualization
tableau, Google Cloud Platform, Qlik, looker, Roambi, Sisense, Zoomdata, datorama, Chartio

### BI Platforms
Power BI, amazon web services, Domo, salesforce Wave Analytics, GoodData, birst, kyvos insights, platfora, atscale, Arcadia, Sisense

### Statistical Computing
sas, SPSS, Matlab

### Log Analytics
splunk, sumologic, kibana, Cloud Physics, loggly

### Social Analytics
Hootsuite, Netbase, DataSift, Tracx, bitly, synthesio, simplereach

### Real-Time
amazon web services, Metamarkets, striim, confluent, DataTorrent, dataArtisans

### Machine Learning
Azure Machine Learning, amazon web services, H2O.ai, Dato, Skytree, rapidminer, DataRPM, deepsense.io, Vicenze, PredictionIO, glowfish

### Speech & NLP
NarrativeScience, Nuance, Wolfram Alpha, semantic machines, Gridspace, ARRIA, api.ai, maluuba, cortical.io, MindMeld, iDibon, Yseop

### Horizontal AI
IBM Watson, Cortana, sentient technologies, Viv, nervana, vicarious, nara, Numenta, HyperScience, Scaled Inference, Descartes Labs, clarifai, Dextro, MetaMind, Geometric Intelligence

### Search
hp Autonomy, Oracle Endeca, Exalead, Lucidworks, elastic, ThoughtSpot, Maana, swiftype, Algolia, Sinequa

### Data Services
Opera Solutions, Mu Sigma, EXL, DataScience, Silicon Valley Data Science, kaggle, datascope, DataKind

### For Business Analysts
OrigamiLogic, ClearStory Data, Cirro, import.io

### Web / Mobile / Commerce
Google Analytics, mixpanel, RJMetrics, BlueCore, Amplitude, granify, sumall, Airtable, retention science, custora

## Applications

### Sales & Marketing
Radius, Gainsight, bloomreach, Zeta, Everstring, livefyre, blueyonder, Lattice, kahuna, infer, Sailthru, persado, Aviso, 6sense, Quantifind, ActionIQ, fuse machines, Ngagio

### Customer Service
Medallia, Attensity, Clarabridge, ClickFox, StellaService, NGDATA, Preact, DigitalGenius, appuri, Wise.io

### Human Capital
gild, Connectifier, textio, entelo, hiQ

### Legal
Ravel, Judicata, Everlaw, Brevia, Premonition

### Ad Optimization
AppNexus, MediaMath, criteo, OpenX, rocketfuel, Integral Ad Science, theTradeDesk, Algorithms, dstillery, Liveintent, Tapad, DataXu, Appier, Moat

### Security
Cylance, CounterTack, cybereason, Area 1 Security, ThreatMetrix, SentinelOne, Recorded Future, Guardian Analytics, Fortscale, siftscience, Keybase, feedzai, Signifyd

### Vertical AI Applications
X.ai, facebook, Clara, Kasisto, lumiata

### Publisher Tools
outbrain, Taboola, quantcast, Chartbeat, yieldbot, Yieldmo

### Govt / Regulation
Socrata, OpenGov, FiscalNote, enigma, Predpol, mark43, OpenDataSoft

### Finance
affirm, LendingClub, OnDeck, Kreditech, zestfinance, LendUp, Kabbage, tidemark, Payoff, INSIKT, Zuora, Dataminr, Lenddo, Kensho, Aidyia, iSentium, Quantopian, sentient technologies

### Education / Learning
Knewton, Clever, Declara, Panorama Education, knowre

### Life Sciences
23andMe, Pathway Genomics, Counsyl, deep genomics, Recombine, Kyruus, Flatiron, zymergen, HealthTap, Metabiota, Zephyr Health, ovia, Ginger.io, transcriptic, Glow, Enlitic, AiCure, Atomwise

### Industries
Opower, eHarmony, RetailNext, duetto, Stitch Fix, WorkFusion, Blue River, Tachyus, SwiftKey, Seeq, FarmLogs, HowGood, celect, Sight Machine, statmuse, Boxever

## Cross-Infrastructure/Analytics
amazon web services, Google, Microsoft, IBM, SAP, sas, 1010 data, hp, Autonomy, Vertica, vmware, TIBCO, Teradata, Oracle, NetApp

## Open Source

### Framework
hadoop HDFS, hadoop MapReduce, YARN, Mesos, Spark, Tez, Apache Kylin, Flink, CDAP

### Query / Data Flow
Pig, Hive, Slamdata, Apache Drill, Google Cloud Dataflow

### Data Access
cassandra, HBase, Accumulo, mongoDB, kafka, CouchDB, SciDB, riak, nifi, OpenTSDB

### Coordination
talend, Apache Zookeeper, Apache Ambari

### Real-Time
Storm, Spark, Apex, Flink, Tachyon, druid

### Stat Tools
R, ScalaLab, NumPy, SciPy

### Machine Learning
mllib, Mahout, Aerosolve, Apache Singa, MADlib, Caffe, torch, CNTK, TensorFlow, scikit-learn, Weka, FeatureFu, jupyter, Veles, DIMSUM, DL4J

### Search
elasticsearch, Solr, Lucene

### Security
Apache Ranger

### Visualization
Zeppelin

## Data Sources & APIs

### Health
Apple, Jawbone, Garmin, practice fusion, fitbit, Withings, Validic, netatmo, kinsa, Human API

### IOT
Uptake, ThingWorx, helium, samsara, Augury, estimote

### Financial & Economic Data
Bloomberg, Dow Jones, Thomson Reuters, S&P Capital IQ, Yodlee, Premise, quandl, xignite, CB Insights, mattermark, StockTwits, estimize, Plaid

### Air / Space / Sea
Planet Labs, spire, Windward, Cruise, Skycatch, Airware, DroneDeploy

### Location / People / Entities
acxiom, Experian, Epsilon, InsideView, Garmin, foursquare, Streetline, esri, Crimson Hexagon, CartoDB, factual, PlaceIQ, Circulate, placemeter, Basis Technology, Sense360

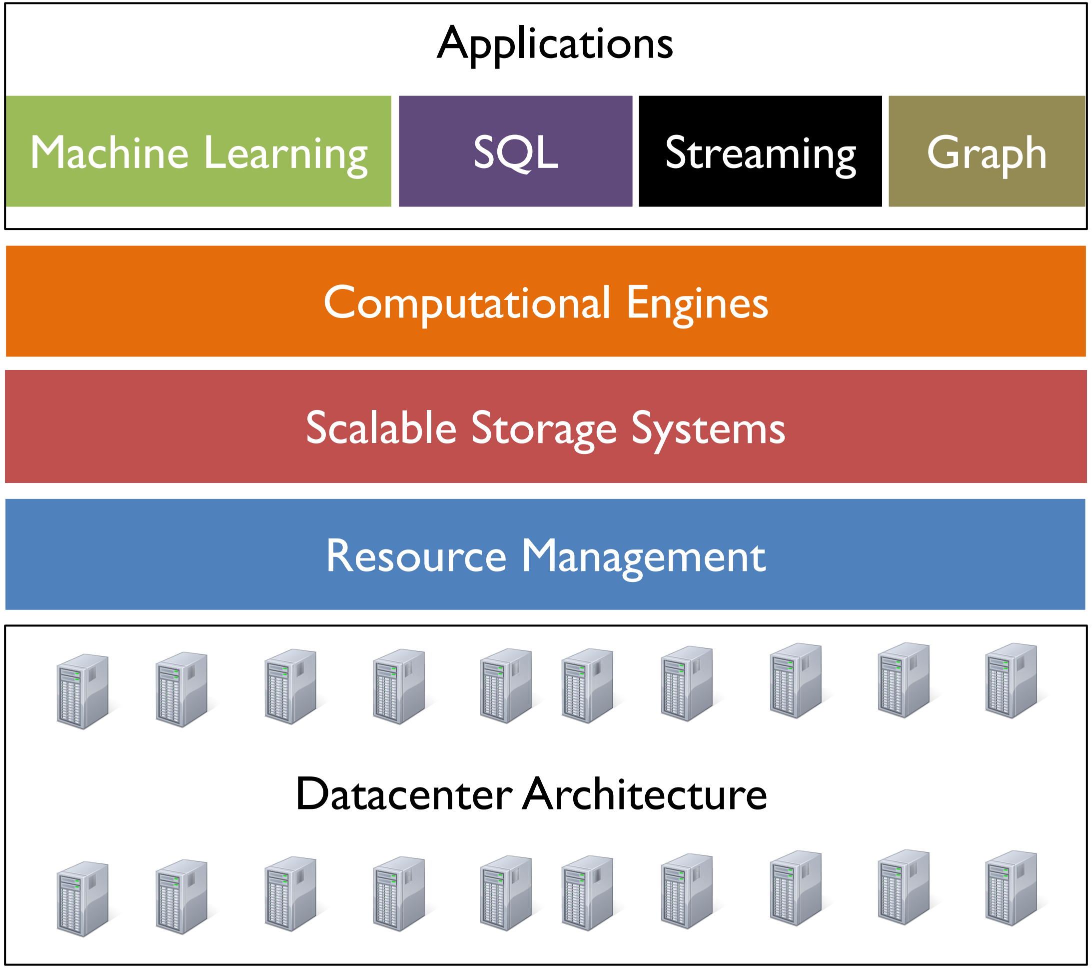
### Other
qualtrics, panjiva, Data.gov

## Incubators & Schools
GA, Pluralsight, DataCamp, Insight, DataElite, The Data Incubator, Metis

*Last Updated 3/23/2016*



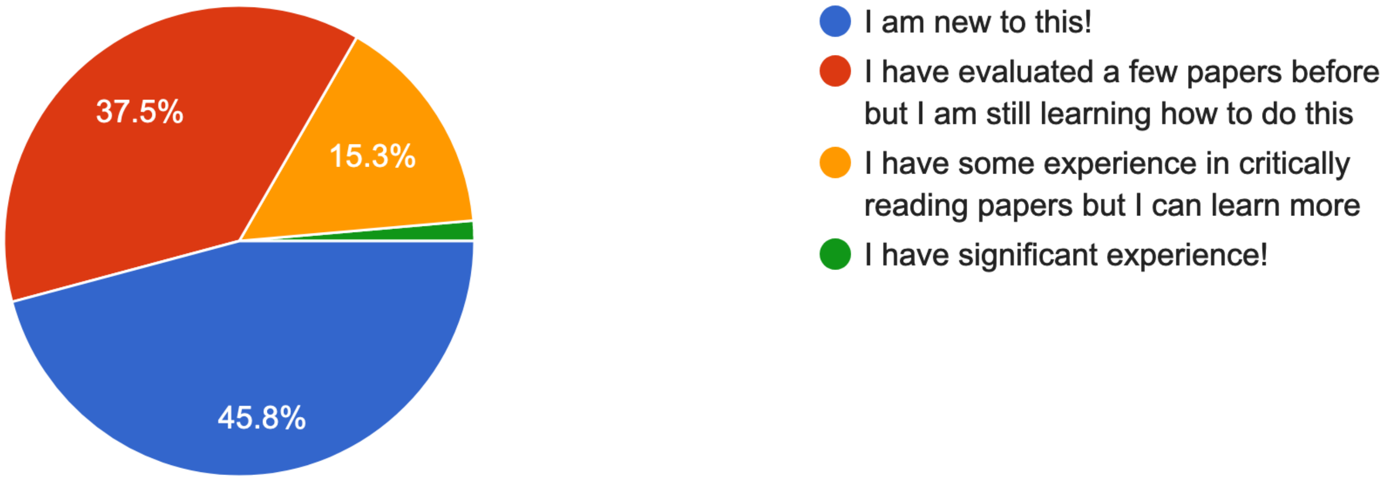
FIRSTMARK


Applications
Machine Learning
SQL
Streaming
Graph
Computational Engines
Scalable Storage Systems
Resource Management
Datacenter Architecture

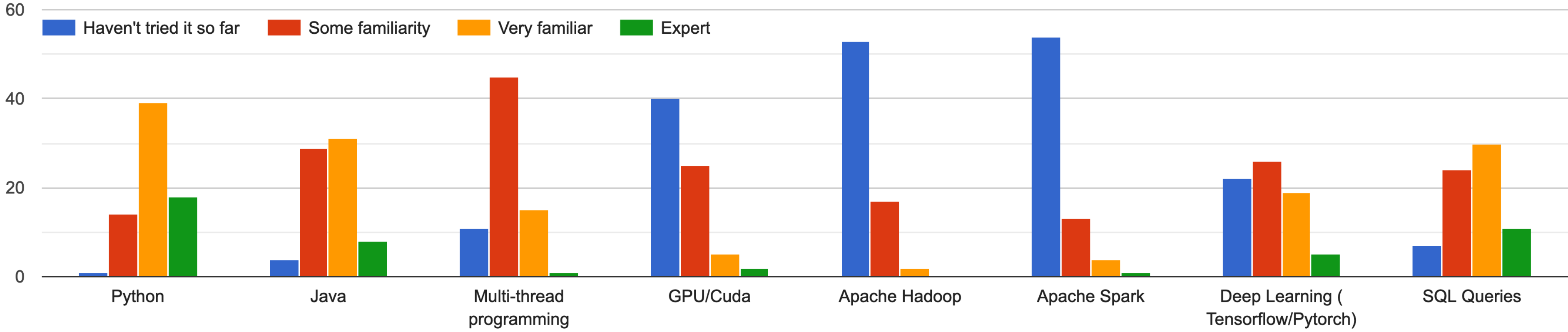
# COURSE SYLLABUS

How much experience do you have in critically reading and evaluating systems research papers?

72 responses

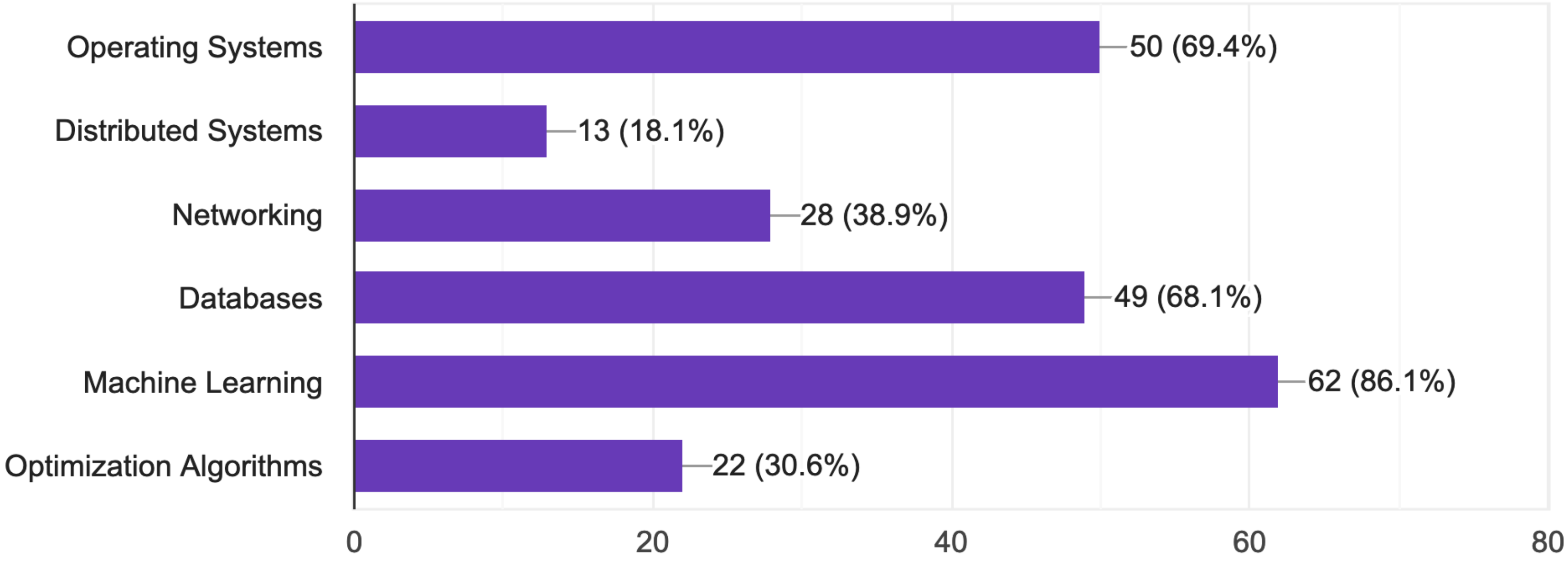


Select how familiar you with each of the topics below



## Select any related prior courses you have taken and completed satisfactorily (either at UW-Madison or any other school)

72 responses



# 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

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 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/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

# 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:

- 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