CS 744: BIG DATA SYSTEMS

Shivaram Venkataraman

Fall 2019

WHO AM I ?

Assistant Professor in Computer Science

PhD Thesis 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

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

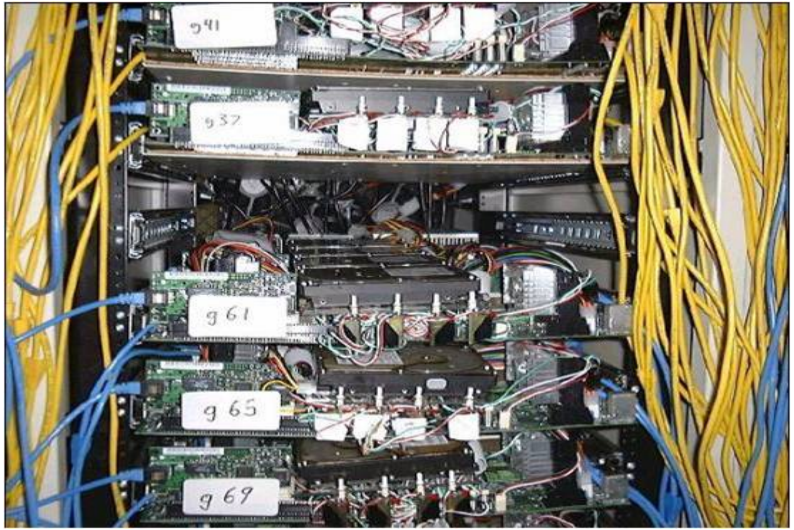
DATA, DATA, DATA

"…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

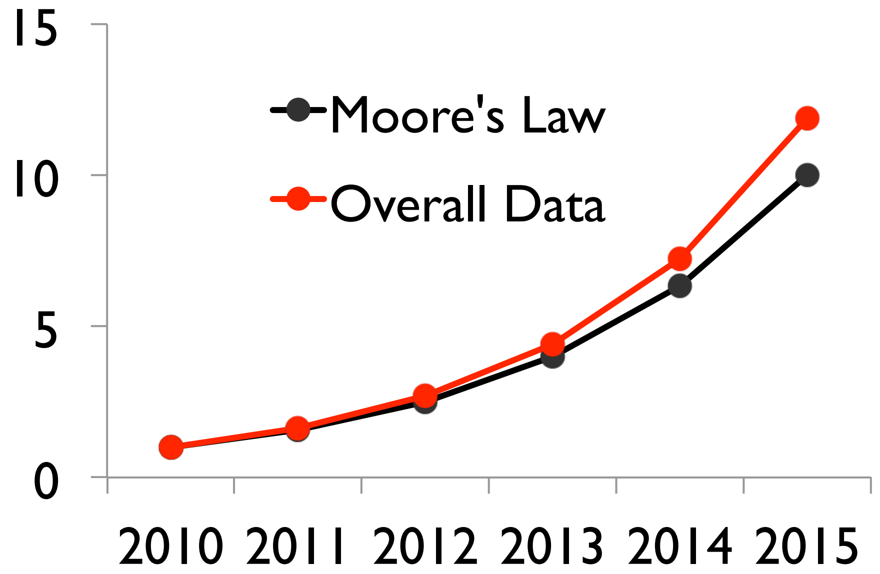
Low bandwidth network

Cheap !

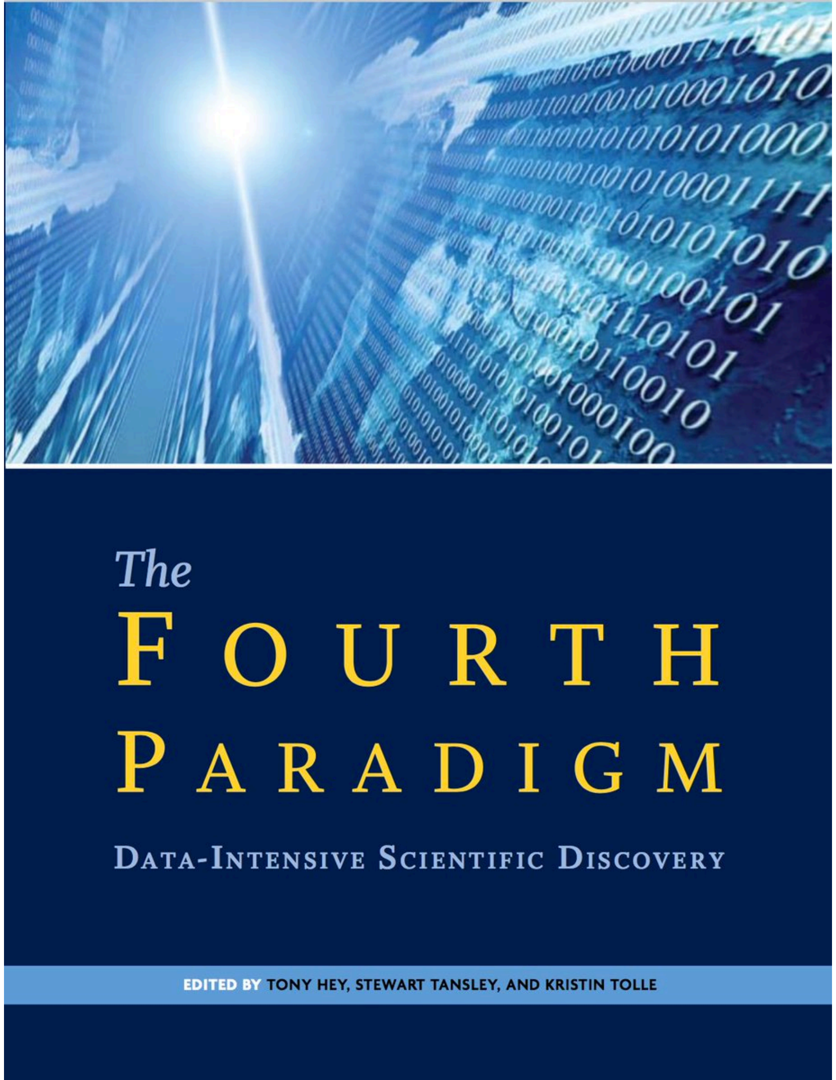
DATACENTER EVOLUTION

Facebook's daily logs: 60 TB

Google web index: 10+ PB



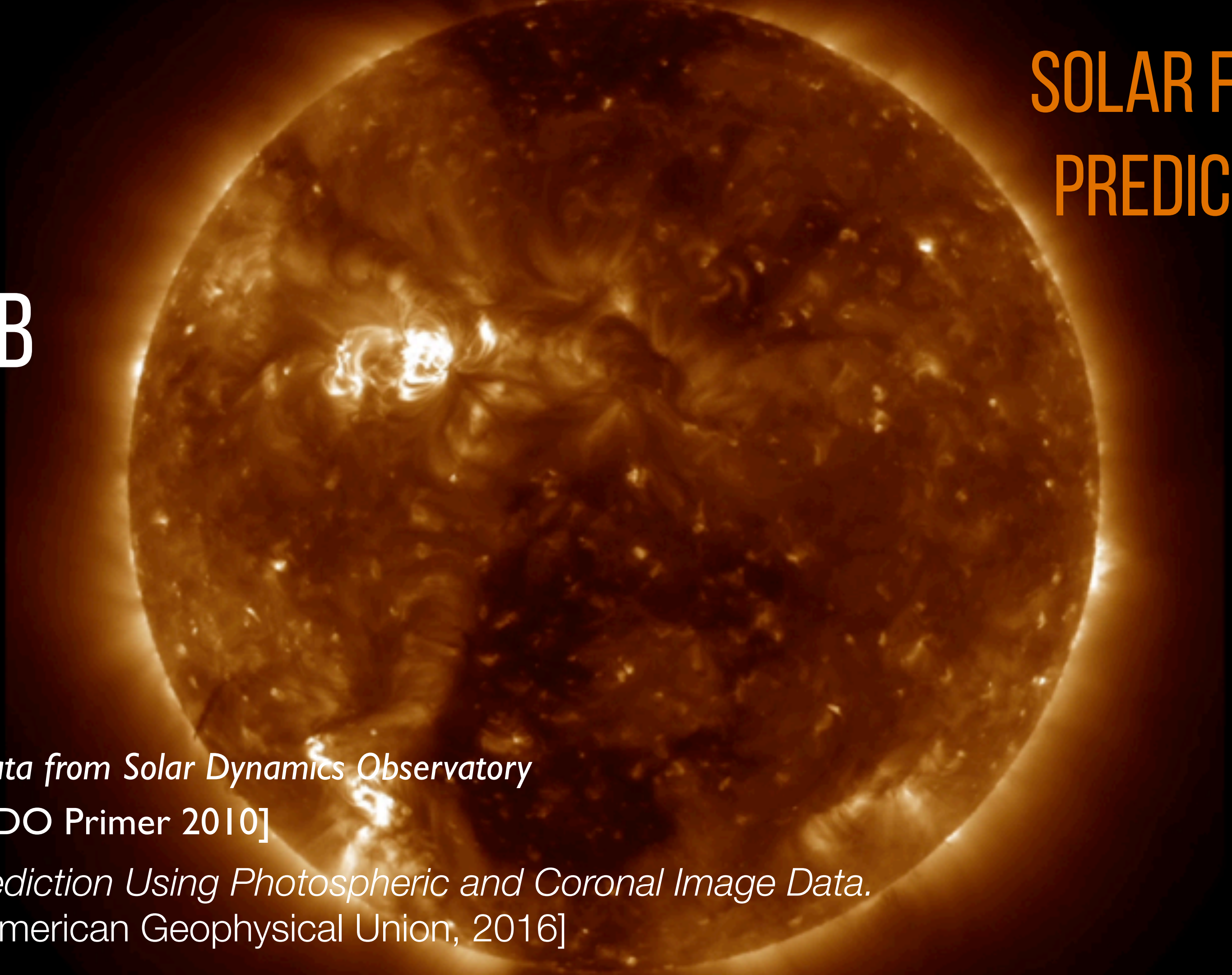
(IDC report*)



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

-- Jim Gray

SCIENTIFIC APPLICATIONS



SOLAR FLARE PREDICTION

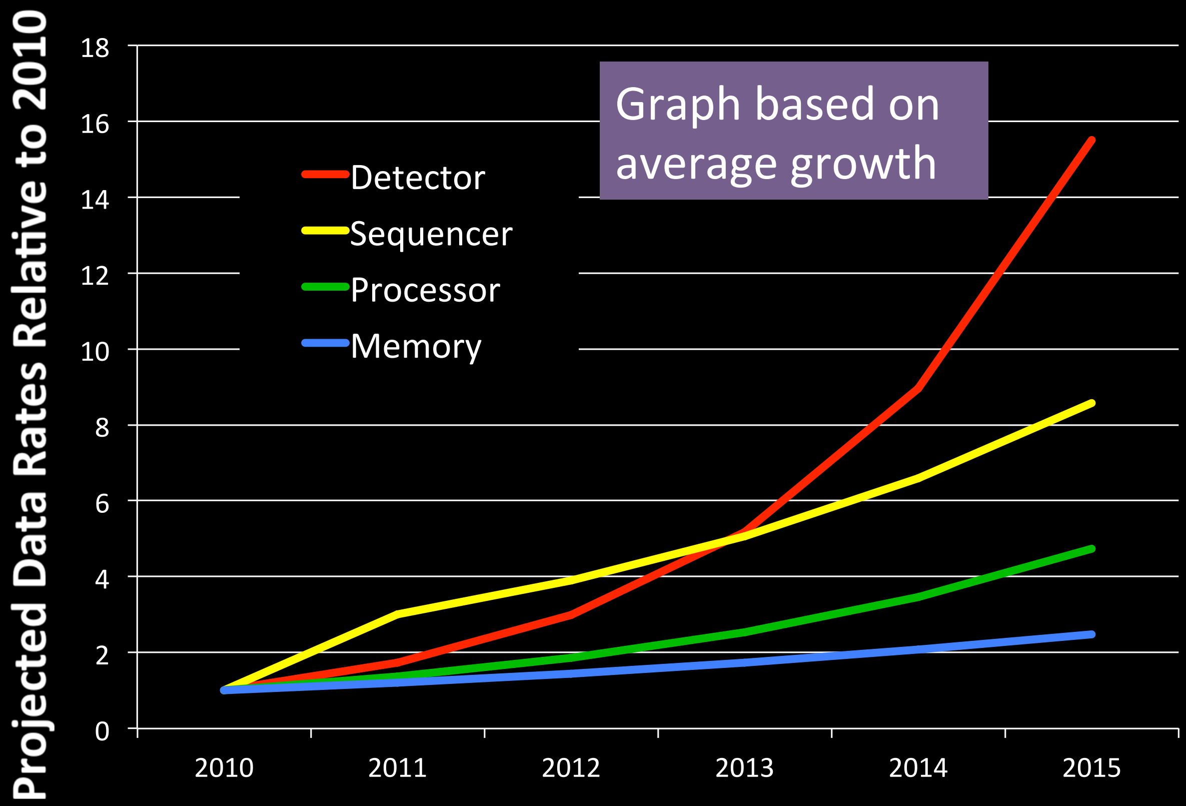
~ 2 PB

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

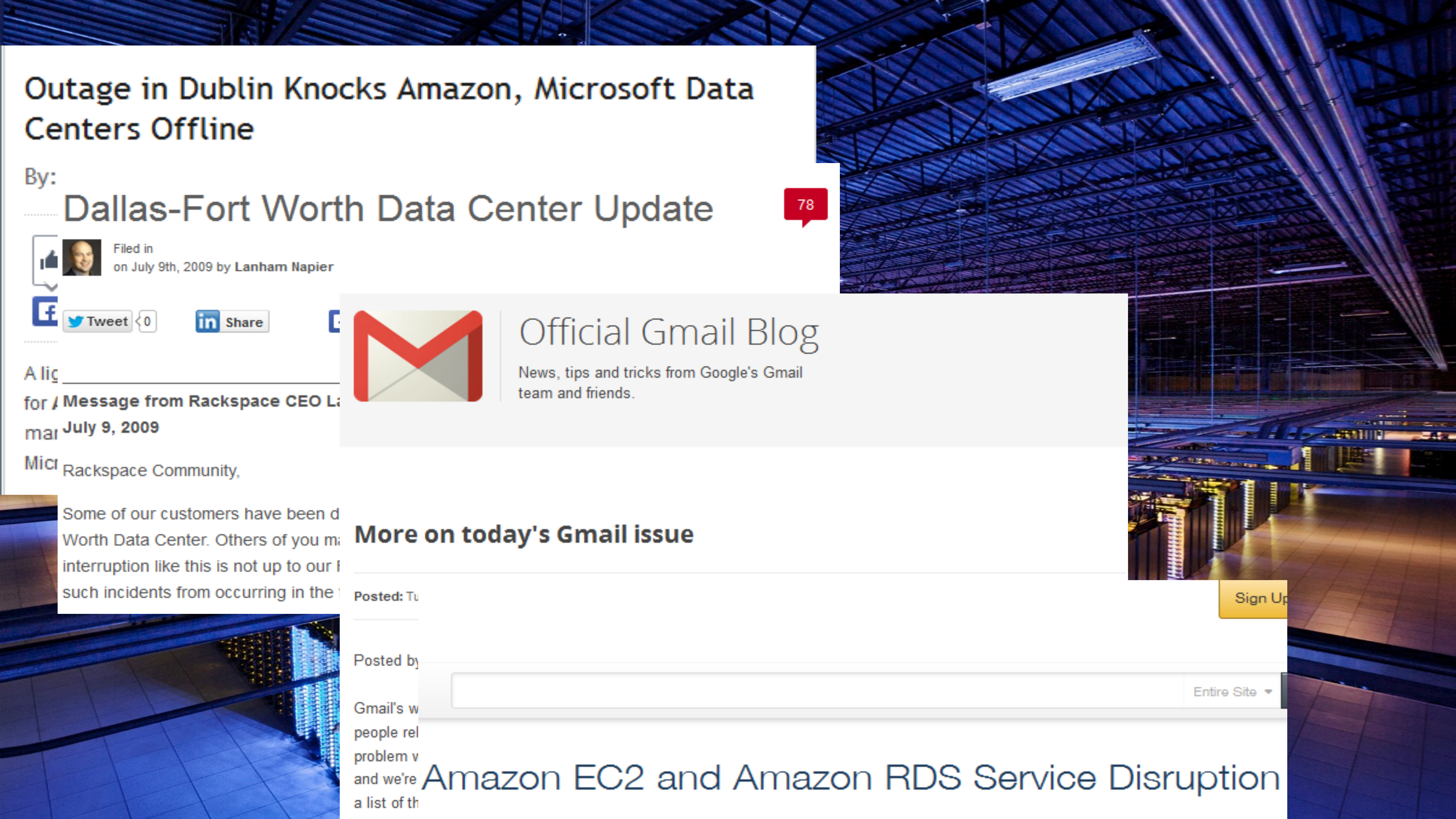
DATACENTER EVOLUTION

Capacity:

~10000 machines

Bandwidth:
12-24 disks per node

Latency:
256GB RAM cache


Outage in Dublin Knocks Amazon, Microsoft Data Centers Offline
By:
A lig
for A
mai
Micr
Dallas-Fort Worth Data Center Update
78
Filed in
on July 9th, 2009 by Lanham Napier
Tweet 0
Share
Message from Rackspace CEO La
July 9, 2009
Rackspace Community,
Some of our customers have been d
Worth Data Center. Others of you ma
interruption like this is not up to our
such incidents from occurring in the
Official Gmail Blog
News, tips and tricks from Google's Gmail team and friends.
More on today's Gmail issue
Posted: Tu
Posted by
Gmail's w
people rel
problem w
and we're
a list of th
Sign Up
Entire Site
Amazon EC2 and Amazon RDS Service Disruption

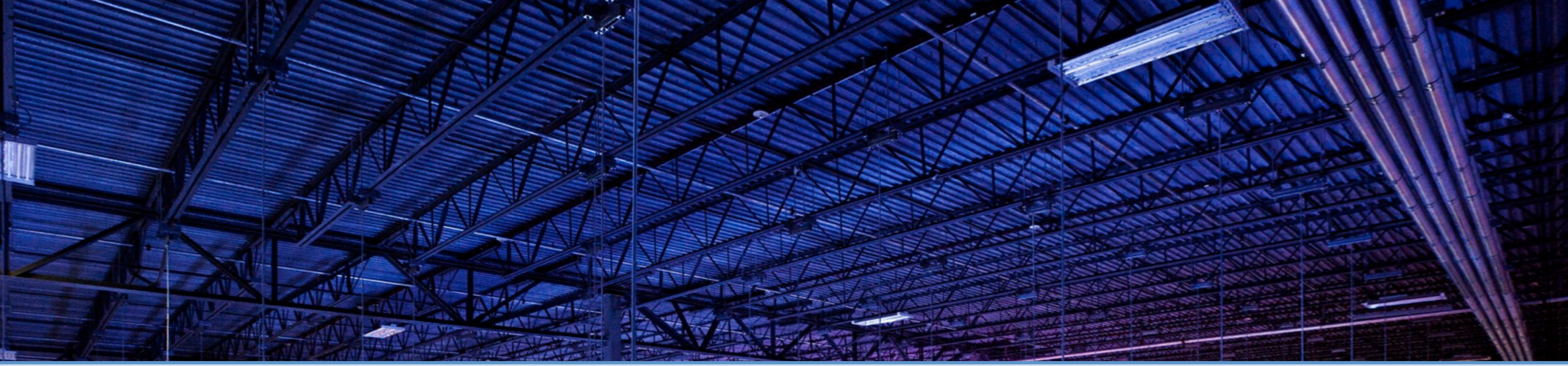
The Joys of Real Hardware

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.

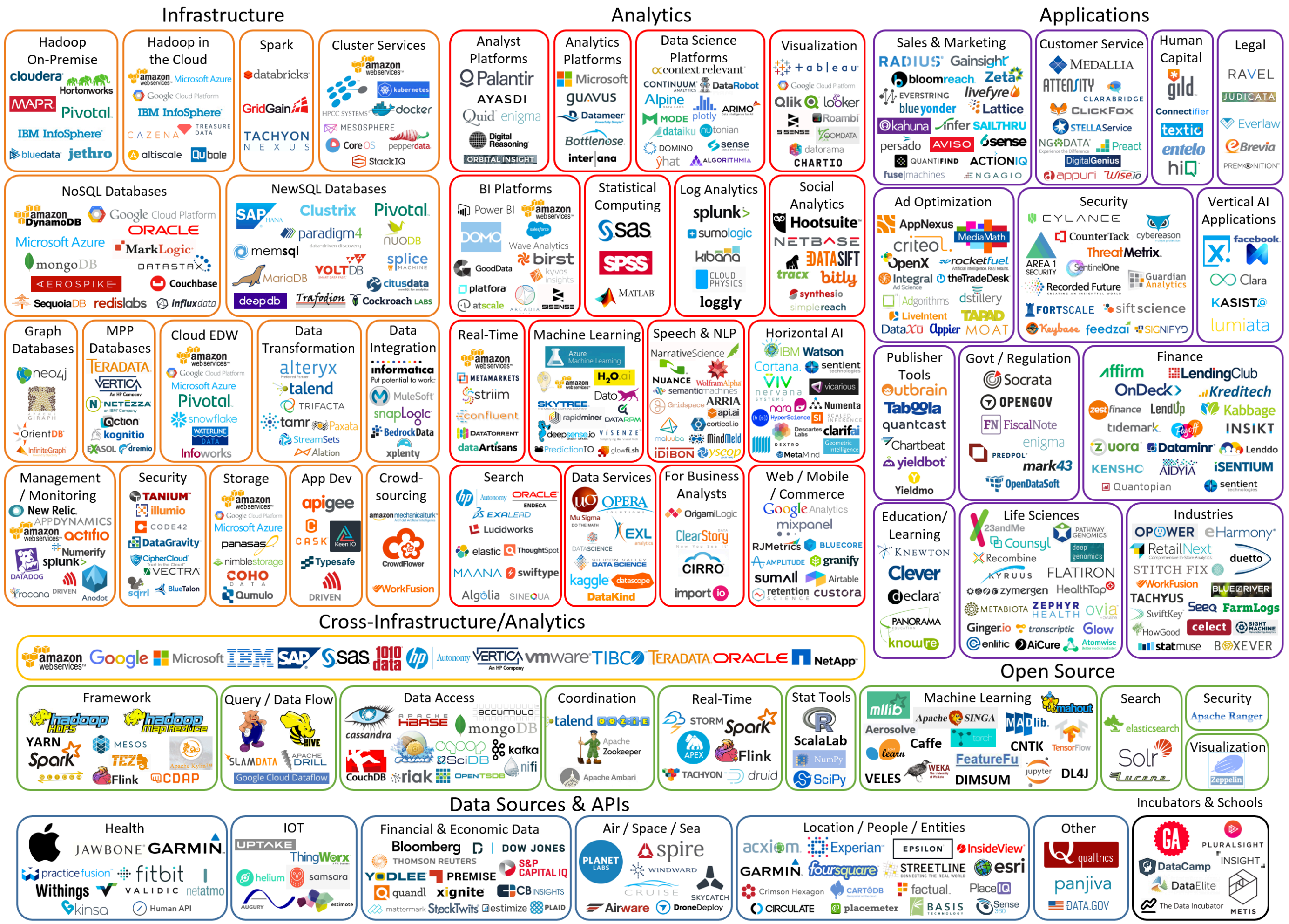
JEFF DEAN @ GOOGLE



How do we program this ?

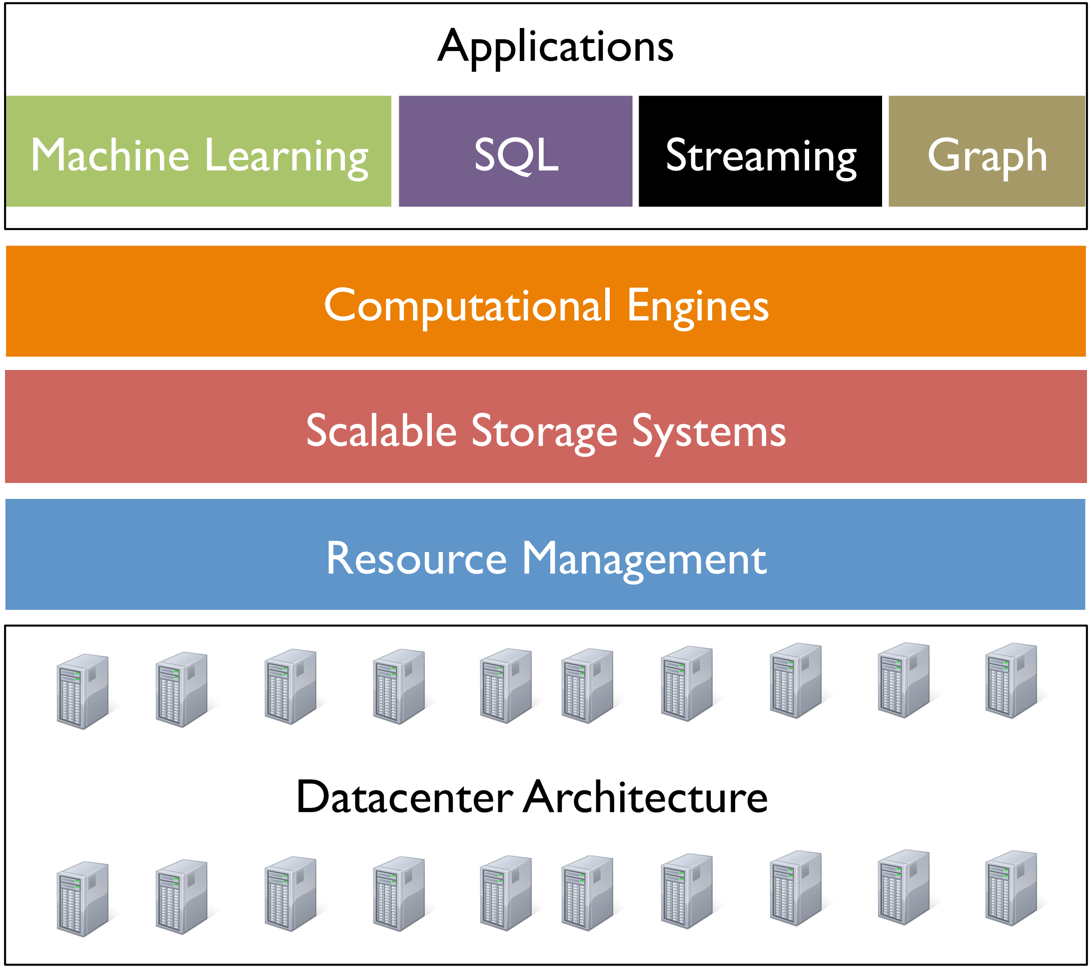
BIG DATA SYSTEMS

Big Data Landscape 2016 (Version 3.0)



Last Updated 3/23/2016

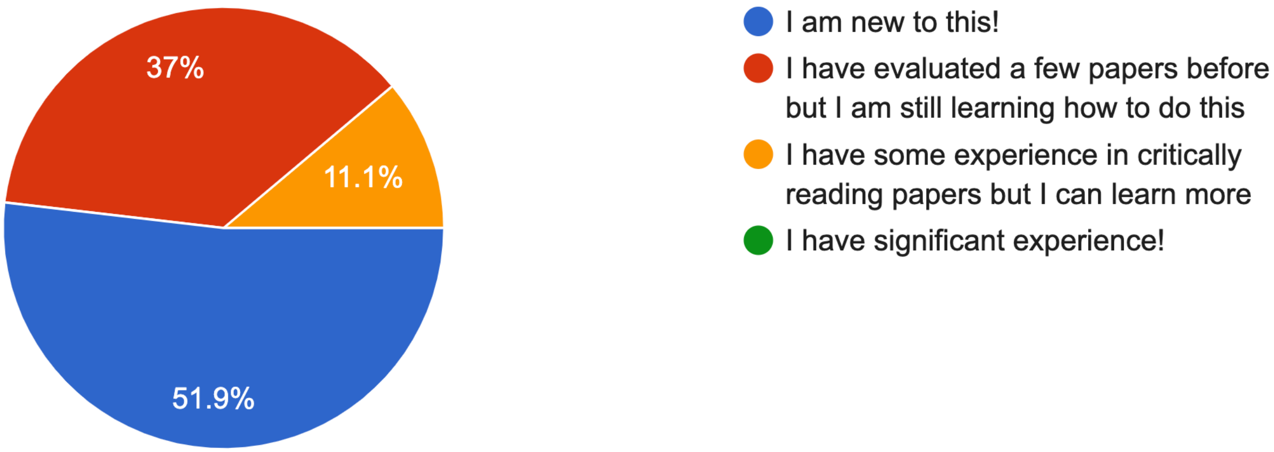
FIRSTMARK

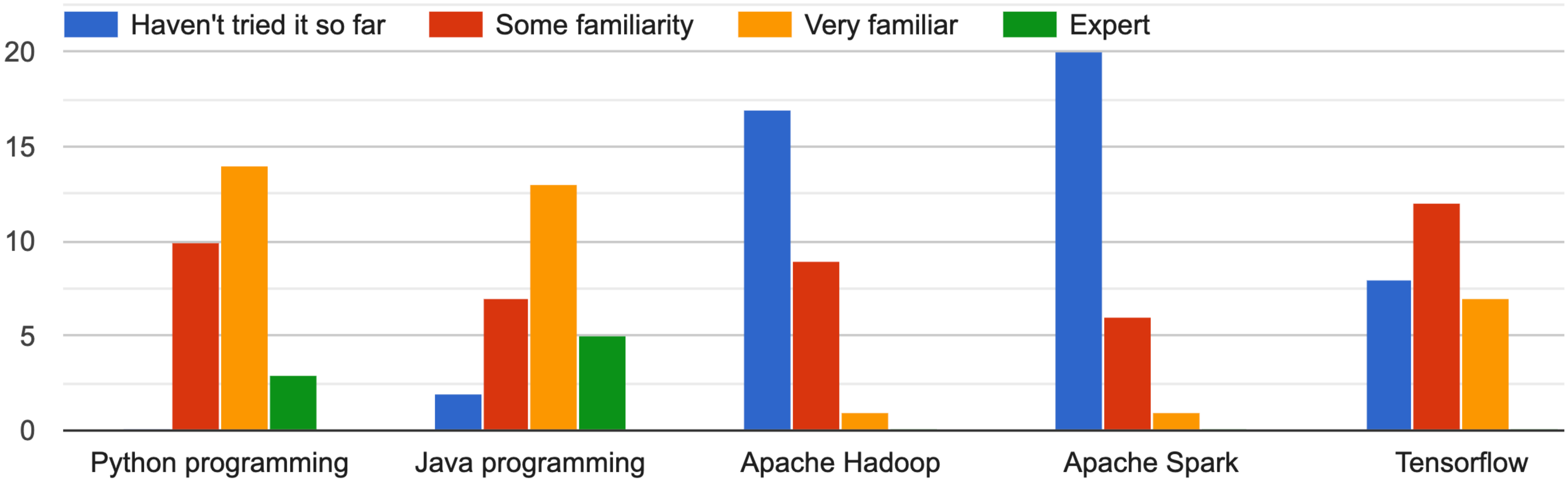

Applications
Machine Learning
SQL
Streaming
Graph
Computational Engines
Scalable Storage Systems
Resource Management
Datacenter Architecture

COURSE SYLLABUS

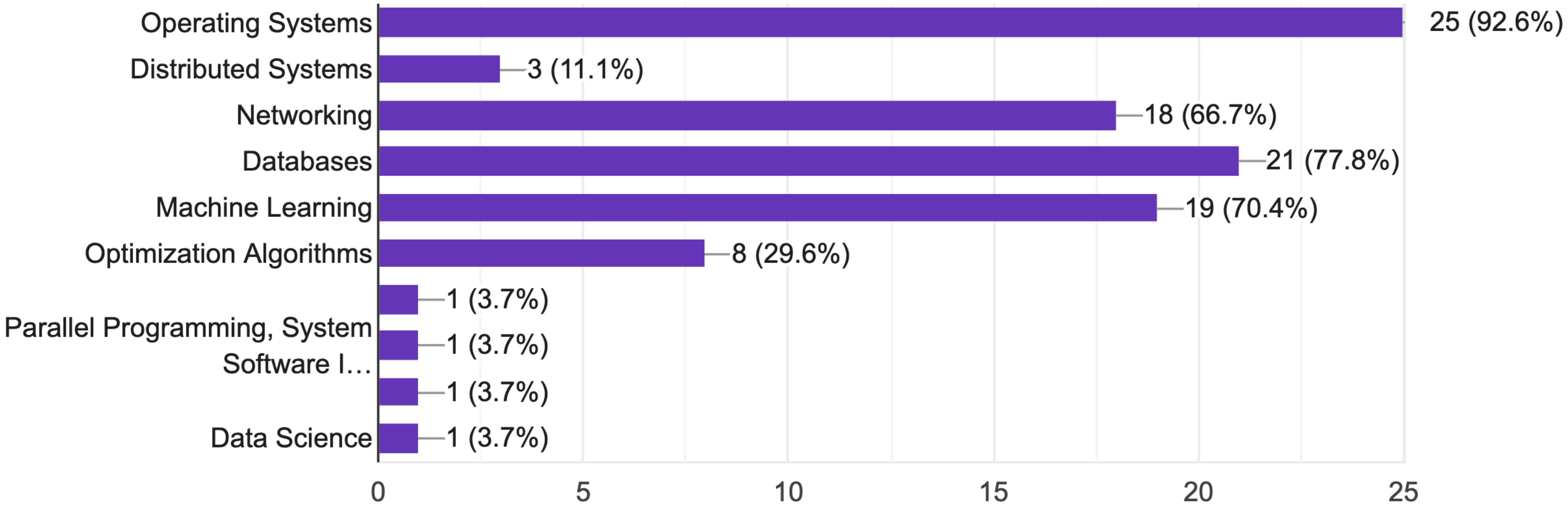
How much experience do you have in critically reading and evaluating systems research papers?

27 responses





Haven't tried it so far
Some familiarity
Very familiar
Expert
20
15
10
5
0
Python programming
Java programming
Apache Hadoop
Apache Spark
Tensorflow


Operating Systems
Distributed Systems
Networking
Databases
Machine Learning
Optimization Algorithms
Parallel Programming, System Software I…
Data Science
25 (92.6%)
3 (11.1%)
18 (66.7%)
21 (77.8%)
19 (70.4%)
8 (29.6%)
1 (3.7%)
1 (3.7%)
1 (3.7%)
1 (3.7%)
0
5
10
15
20
25

WHAT DO YOU HOPE TO LEARN FROM THE COURSE?

To be able to evaluate the research papers more effectively…

I hope learn to design systems used for big data processing…

Learn about current day technologies that are used to manage large amounts of data…

Learn how to implement a machine learning project on big data.

Both theory and applications of big data systems, i.e., how to design, how to implement and how to evaluate.

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-fa19

Reading: 1 paper per class

Review: Fill out review form (posted on Piazza) by 9am

Discussion: In-class group discussion, submit responses

(Best15 out of 20 responses)

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 assumptions. We have reexamined traditional choices and explored radically different points in the design space.

HOW TO READ A PAPER: SUMMARY

1st pass: Read abstract, introduction, section headings, conclusion
2nd pass: Read all sections, make notes

Some key points

- What is the problem being considered?
- What are the main contributions? How do they compare to prior work?
- What workloads, setups were considered in the evaluation?
- What parts of the claims are adequately backed up?

…

PAPER REVIEW, DISCUSSION

Examples

- One or two sentence summary of the paper
- Description of the problem or assumptions made
- Comparison to other papers discussed in class
- One flaw or thing that can be improved
- Experimental setup and what do the results mean

ASSESSMENT

- Paper reviews: 10%
- Class Participation: 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/Spark
- Assignment 2: Machine Learning/Tensorflow

Short coding based assignments. Preparation for course project

Work in groups of three

COURSE PROJECT

Main grading component in the course!

Goal: Explore new research ideas or significant implementation in the area 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 will be posted around (9/12)
- Form groups of three
- Pick one or more ideas or propose your own!
- Submit project ideas, instructor feedback/finalize idea (9/26),

Assessment:

- Project introduction write up
- Poster presentation
- Final project report

COURSE LOGISTICS

Instructor office hours:	Mon 11-12am at 7367 CS
Ainur's office hours:	Mon 2-3pm and Thu 2-3pm at 4291 CS
Discussion, Questions:	Use Piazza!

WAITLIST

- Class size is limited to 60 for this semester
- Focus on research projects, discussion
- Course is offered both semesters
- 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 email

BEFORE NEXT CLASS

Join Piazza: https://piazza.com/wisc/fall2019/cs744

Complete Assignment 0 (see website)