Report abuse

Learn more

## CourseSchedule-CC-24-fall

Updated automatically every 5 minutes

This list of papers and readings are under development, please check back here frequently.

Generally lectures will be monday and sprint demos will be wednesday

- Week 1: Sept 4: Course Overview
  - Wed Sept 4:
    - Introduction to cloud computing
    - course structure, schedule & project descriptions
  - o Fri Sept 6
    - **■** Fill in the skills survey
    - **■** Fill in the project preference from
    - We will assign project teams Monday and you will have some time to meet with your team
- Week 2: Sept 9 11: Overview of Cloud Computing
  - Monday Sept 9: Ata Turk Cloud DCs
    - Quiz 1
    - What is cloud computing & Review paper reviewing guidelines, project description guidelines...
    - The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines (Chapters 1 through 3)
  - Wed Sept 11: Michael Daitzman Agile
    - First agile methods lecture, before class read and watch the video in this document (Agile Lecture 1) (1-2 hours)
- Week 3: Sept 16 18 Michael Daitzman Agile
  - Mon Sep 16: Second agile methods lecture
    - Before the class read <u>Agile Section 2</u>
  - Wed Sep 18: Ata Turk Cloud File Systems, Cloud compute platforms
    - Quiz 2
    - The Google File System
    - MapReduce: simplified data processing on large clusters
  - Fri Sept 20:
    - Project descriptions: Due Friday Sept 20th, update your Readme.md file
- Week 4: Sept 23 Sept 25
  - o Mon Sept 23: Orran Krieger Virtualization
    - Quiz 3
    - Container-based Operating System
       Virtualization: A Scalable, High-performance
       Alternative to Hypervisors
    - Xen and the art of virtualization
      - Additional reading:



Report abuse

Learn more

# CourseSchedule-CC-24-fall

Updated automatically every 5 minutes

- Quiz 4
  - Remaining Sprint 1 demos
  - BigTable: A Distributed Storage System for Structured Data
  - Spark: Cluster Computing with Working Sets
- Wed Oct 2: Cloud compute platforms stream processing
  - Quiz 5
  - Storm@twitter [ video (Nathan Marz presenting Storm) ]
  - Spark: Discretized Streams: An Efficient and Fault-Tolerant Model for Stream Processing on Large Clusters [ video (Tathagata Das presenting the paper) 1
- Week 6: Oct 7 9
  - Mon Oct 7: Behind the Scenes of a Distributed Containerized Web Application - Evren Sen (Lead **AWS Principal Solutions Architect)**
  - Wed Oct 9:
    - Sprint demo 2
- Week 7: Oct 14 16
  - Mon Oct 14: Indigenous People's Day Holiday, Classes Suspended
  - Tue Oct 15: Substitute a Monday schedule of classes - Peter Desnoyers - Cloud DBs
    - - Quiz 6
      - Ceph: a scalable, high-performance distributed file system; also check: CRUSH, data placement algorithm of CEPH and RADOS, underlying object store of CEPH.
      - Flat Datacenter Storage [ video (Jeremy Elson presenting the paper) I
  - Wed Oct 16 Ata Turk Cloud Computing Platforms Serverless
    - Quiz 7
    - Bonus Workshop 1:
      - Go through the following workshop, AWS: Build a Serverless Web Application using Generative Al. take screenshots showcasing you completed the workshop and send them to the TA. You have 1 week to complete.
    - Serverless Computation with OpenLambda
    - A Berkeley View on Serverless Computing
      - Additional reading
        - Serverless in the Wild: Characterizing and Optimizing the Serverless Workload at a Large Cloud Provider

Report abuse

Learn more

## CourseSchedule-CC-24-fall

Updated automatically every 5 minutes

#### Systems

- Quiz 8
- <u>Large-scale cluster management at Google</u> with Borg
- Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center
- Bonus Workshop 2:
  - Go through the following workshop up to the Observability section (first 9 sections), <u>Building Web Applications</u> <u>based on Amazon EKS</u>, take screenshots showcasing you completed the workshop sections and send them to the TA. You have 1 week to complete.
- Wed Oct 23
  - Invited Talk: Dan Schatzberg, Research Scientist Meta
    - Presenter: Dan Schatzberg is a
      Research Scientist at Meta working on
      Resource Control. His work focuses
      on safely increasing resource
      utilization within Meta's vast fleet of
      servers. Prior to joining Meta, he was
      pursuing his PhD at Boston University
      with a focus on customizable operating
      systems for datacenter applications.
    - Abstract: Facebook is composed of thousands of software services incorporating functionality such as web serving, caching, machine learning, and many others. Over the last few years, we've overhauled our internal infrastructure for supporting these many services in order to reduce operational costs and better utilize our fleet of servers. In this talk I'll describe the Infrastructure as a Service project and how it differs from public laaS offerings.
- Week 9: Oct 28- Oct 30
  - o Mon Oct 28:
    - Sprint demo 3
  - Wed Oct 30 Ata Turk Cloud DBs
    - Quiz 9
    - Spanner: Google's Globally-Distributed
       Database
    - Bonus Workshop 3:
      - Go through the first 2 sections of the following workshop (Create a new RDS SQL Server, High Availability),



Report abuse

Learn more

## CourseSchedule-CC-24-fall

Updated automatically every 5 minutes

- vveek 10: Nov 4 6
  - Mon: Nov 4 Surya Jayanthi
    - Quiz 10
    - Kafka: a Distributed Messaging System for Log Processing
    - Bonus Workshop 4:
      - Go through the following workshop, File Streaming with AWS Media Services, take screenshots showcasing you completed the workshop and send them to the TA. You have 1 week to complete.
  - Wed Nov 6:
    - Sprint demo 4
- Week 11: Nov 11 13
  - o Mon: Nov 11
    - Sprint demo 4 continued.
    - Invited Talk: Stefan Mationg, AWS
      - AWS Lambda: Run code without provisioning or managing servers
  - Wed: Nov 13
    - Quiz 11
    - Invited Talk: Dan Lambright, Apple
      - Dan Lambright has worked the last 4
        years in iCloud for Apple. In prior years
        he worked at Meta, AWS, and Red
        Hat. He enjoys building file systems
        and databases. He has an MSCS from
        the University of Arizona and a BSCS
        from RPI.
      - This talk will be about one the databases underpinning iCloud (FoundationDB), and the SQL-like layer above it called the Record layer. An overview of the architecture and how applications talk to it will be discussed.
- Week 12: Nov 18 20
  - Monday Nov 18
    - Dynamo: amazon's highly available key-value store
  - Wed Nov 20
    - Sprint demo 5
- Week 13: Nov 25 27
  - Monday: Nov 25 LLMs Ravi Bandaru/Surya Jayanthi
    - Quiz 12
    - Lost at C: A User Study on the Security
       Implications of Large Language Model Code
       Assistants

Report abuse

Learn more

## CourseSchedule-CC-24-fall

Updated automatically every 5 minutes

Penetration Testing

- Characterization of Large Language Model Development in the Datacenter
- Wed: Nov 27
  - Thanksgiving break
- Week 14: Dec 2 4: Cloud Observability Surya Jayanthi
  - Monday Dec 2:
    - Quiz 13
    - Zero Overhead Monitoring for Cloud-native Infrastructure using RDMA
  - Wednesday Dec 4:
    - Quiz 14
    - The History of Logging @ Facebook (Abridged)
    - Invited Talk: TBD
- Week 15: Dec 9-11
  - Monday Dec 9:
    - Final presentations

#### Additional Reading:

- Cloud Backup: <u>Data Domain Cloud Tier: Backup here</u>, backup there, deduplicated everywhere!
- Cloud Tracing: X-trace: a pervasive network tracing framework
- Cloud Datacenter Networks: Data Center TCP (DCTCP)
- Empowering Azure Storage with RDMA
- CICD: AWS CICD blog posts: Automating safe, hands-off deployments, Going faster with continuous delivery, and Ensuring rollback safety during deployments
- Resource Management: <u>Twine: A Unified Cluster</u> Management System for Shared Infrastructure
- Serverless: On-demand Container Loading in AWS Lambda
- Storage for Serverless: <u>Pocket: Elastic Ephemeral Storage</u> for Serverless Analytics
- NoSQL: <u>Cassandra A Decentralized Structured Storage</u> System
- LLM Checkpointing: <u>Gemini: Fast Failure Recovery in</u>
  Distributed Training with In-Memory Checkpoints
- BigQuery insides: Overview of BigQuery storage
- Big Metadata: When Metadata is Big Data
- Cloud security: <u>Credit Karma: Understanding Security</u> <u>Implications of Exposed Cloud Services through Automated</u> Capability Inference
- Cloud hw security: <u>The Security Design of the AWS Nitro System</u>
- Cloud DBs: FoundationDB: A Distributed Key Value Store

#### Workshops to look at:

 AWS: Build a Serverless Web Application using Generative Al



Report abuse

Learn more

CourseSchedule-CC-24-fall

Updated automatically every 5 minutes