CPR E 563X: Advanced Data Storage Systems (Spring 2020)

How do **Google/Amazon/Facebook/IBM** store our data?
What's cool about **Intel 3D XPoint**?
What **failures** occurred in **cloud datacenters** & **national labs**?
What **bugs** exist in **Linux kernel**?
… … … … … …

CPR E 563X will answer the questions above, and let you learn more!

- Instructor: Dr. Mai Zheng
  - https://www.ece.iastate.edu/~mai/
- Two Sections Available
  - Regular on-campus section
  - Online section
- Cover Both **Background Knowledge** & **Latest Advances**
  - Hard Disk Drives & Linux I/O Stack
  - Redundant Array of Independent Disks (RAID)
  - Flash-based Solid State Drives & Persistent Memories
  - Ext4 – Default file system on all Linux distributions
  - F2FS – A flash friendly file system by Samsung Electronics
  - GoogleFS/HDFS – Foundation of large-scale machine learning
  - Lustre – High-performance parallel file system empowering >50% top 100 supercomputers
  - Dynamo – Amazon's highly available key-value store
  - RocksDB – Log structured persistent key-value store by Facebook
  - uDepot – IBM's in-memory key-value store for fast NVMe devices
  - Fault Injection, Model Checking, Fuzzing & Symbolic Execution
  - …
- Major Assignments
  - **Intensive reading** of selected papers published at top systems venues (e.g., SOSP/OSDI/FAST/ASPLOS/USENIX ATC)
  - **Critical thinking**: Write critiques & Discuss design tradeoffs
  - **Get hands dirty**: Design, implement, & present one project based on **High-Impact**, **Real-World**, **Open-Source** systems
- Prerequisites (recommended but not required)
  - Knowledge of operating systems and C programming language
  - Please contact the instructor if not sure
- Questions?
  - Send email to mai@iastate.edu, or stop by Durham 349