Overview

● Each student is required to work on one project based on the paper topics.

● Course schedule has suggested projects for each topic.

● All projects will be based on H-Store
  ○ Java / C++ / SQL / Python
  ○ Built-in benchmark framework.
  ○ Automatic deployment on Amazon EC2.
Overview (2)

● List of project descriptions are available on Github:

● Projects are tagged based on difficulty:
  - **EASY** - Single person group only.
  - **MEDIUM** - One or two person group.
  - **HARD** - Two person group only.

● Project difficulty does not mean more or less work.
H-Store

• Parallel, Main Memory Database System
  ○ [http://hstore.cs.brown.edu](http://hstore.cs.brown.edu)

• Java-based Front End
  ○ Query Planner
  ○ Transaction Coordinator
  ○ Stored Procedures
  ○ Client Interface

• C++ Execution Engine
  ○ Storage Manager
  ○ SQL Operator Executor
Easy Projects

● Eventually Consistent Queries
  ○ *Delayed execution for select queries.*

● Java Query Cache
  ○ *Fast caching of results for select queries.*

● Ad-Hoc Queries
  ○ *Execution of non-stored procedure queries*

● New Client Interfaces
  ○ *JSON, Memcache*
Medium Projects

- Checkpoint + Recovery
  - Write snapshots of data out to disk for recovery.

- Write-Ahead Logging
  - Write log of transaction requests for recovery.

- Distributed Query Optimizations
  - Make parallel queries run faster on multiple nodes.

- Port TPC-E to H-Store
  - Create a Java-only implementation.
Hard Projects

● Live Migration
  ○ *Dynamically add a new node to cluster.*

● MapReduce Transactions
  ○ *Execute analytical queries using MR-like txns.*

● Enhanced SQL Features
  ○ *Implement modern SQL functionality in H-Store.*

● Replicated Nodes
  ○ *Manage replicated nodes using Zookeeper.*
Road to Success

- Semester checkpoints:
  - Project Proposal
  - Milestone #1
  - Milestone #2
  - Final Project Deliverable

- Modern coding practices:
  - Functional Specifications.
  - Test Cases.
  - Code Reviews.
  - Experimental Analysis.
  - Documentation.
Road to Success (2)

• All students are expected to contribute equally to projects.

• Projects will be available as open-source on Github.
  ○ Github = Your Portfolio
  ○ Put pride in your work.

• Ability to extend project as MS thesis.
Due Next Class

- Pick a week to present papers.
- Pick a project.
- Pick a NewSQL system to present.
- Create a free Github account.
  - Upload a profile picture so we know who you are.
  - Read a tutorial on Git if you're not familiar with it.
Final Remarks

● Plagiarism will not be tolerated during any part of the course.
  ○ *It is not ok to copy entire blocks of text.*
  ○ *It is ok to paraphrase and cite.*

● When in doubt, ask somebody.

● See University Academic Code:
Final Remarks (Again)

- Plagiarism will not be tolerated during any part of the course.
  - *It is not ok* to copy entire blocks of text.
  - *It is ok to paraphrase and cite.*

- When in doubt, ask somebody.

- See University Academic Code: