
Lecture XXX
Carnegie Mellon University
Special Topics:
Self-Driving Database Management Systems
Project Discussion
@Andy_Pavlo // 15-799 // Spring 2022

TODAY'S AGENDA

System Overview
Implementation
Integration

PROJECT #2

Project deliverables:

~~→ Proposal~~

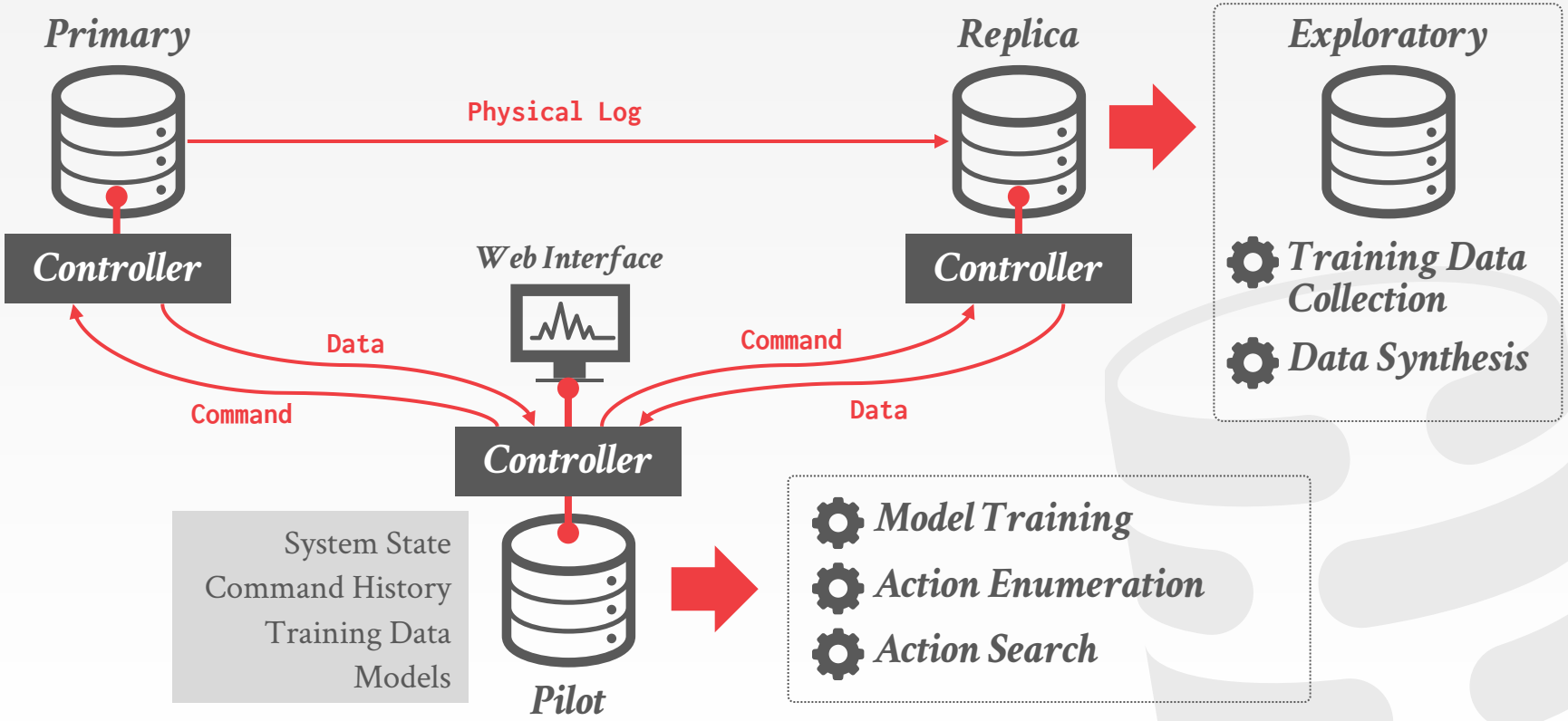
→ Final Presentation (April 20th)

→ Design Document

→ Code Review

→ Code Drop

SYSTEM OVERVIEW



GETTING STARTED

Assume that the ***Pilot*** will be able to get the data you need from the ***Primary*** node.
→ Ignore the distributed architecture.

Assume the ***Pilot*** will instruct your component on when it is time to do certain things (e.g., collect data, retrain models).

THINGS TO THINK ABOUT

What commands (SQL, commandline) do you need to interact with the DBMS.

→ Assume the ***Controller*** has necessary privileges to the DBMS and host OS.

Avoid operations that incur significant overhead on the ***Primary*** DBMS.

The controllers on ***Primary*** and ***Replica*** nodes should be stateless.

DISCUSSION

How are we going to integrate the different components together?

→ Should we assume a microservices architecture?

→ Is there a standard API that each component exposes?

How should we test projects?

→ Individual Tests vs. Integration Tests

→ Github Actions?

DISCUSSION

Does anyone need additional AWS credits?
→ I have 2018 Intel NUCs in my office.

Does anyone need workloads beyond BenchBase?

NEXT CLASS

The next four classes will discuss different approaches to building end-to-end systems.

Guest Speakers:

→ April 25th – Bailu Ding (MSR)

→ April 27th – Weiwei Gong (Oracle)