

Deploy Distributed Database Course Project on Vocareum

Arizona State University

Jia Yu



THIS TALK

Course background

Deployment on Vocareum

Outcome

THIS TALK

Course background

Deployment on Vocareum

Outcome

COURSE BACKGROUND

- The principles of distributed data systems
- Data partitioning, query, sorting, join
- More than 150 students (more on Coursera)
- Hands-on experience is crucial



COURSE PROJECT

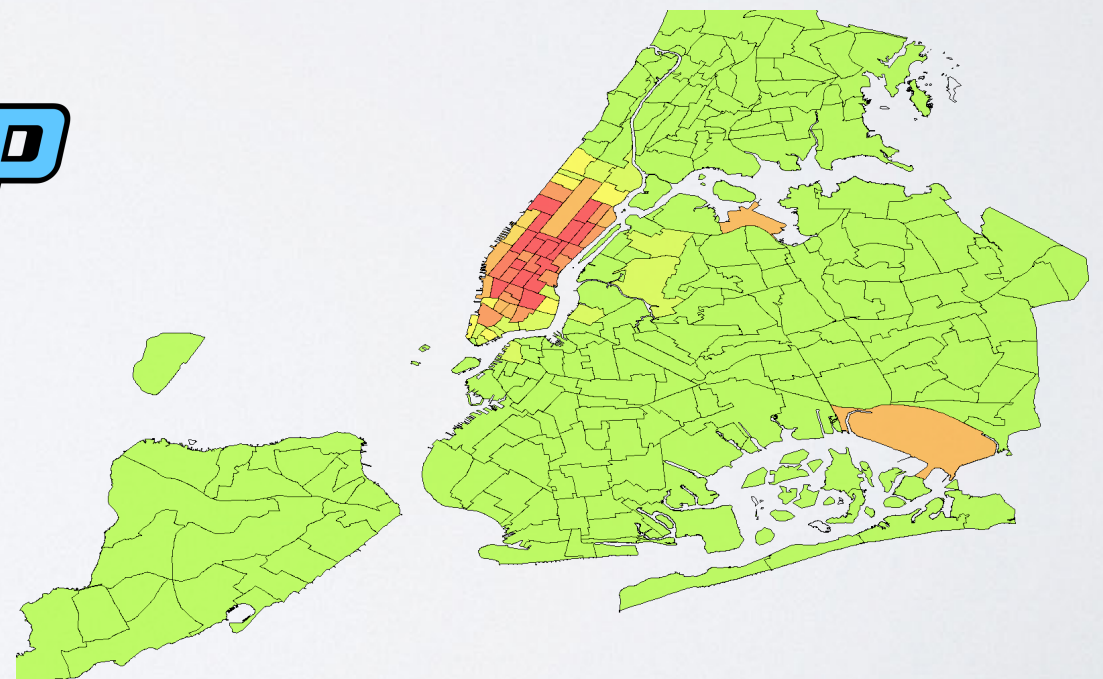
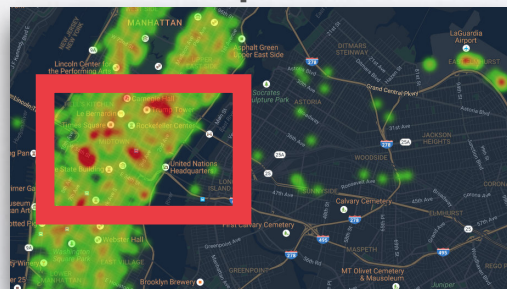
- Real life scenario: spatial data management



- Cutting-edge techniques: Apache Spark + HDFS



- Phase 1: spatial queries



- Phase 2: spatial analytics using spatial queries

THIS TALK

Course background

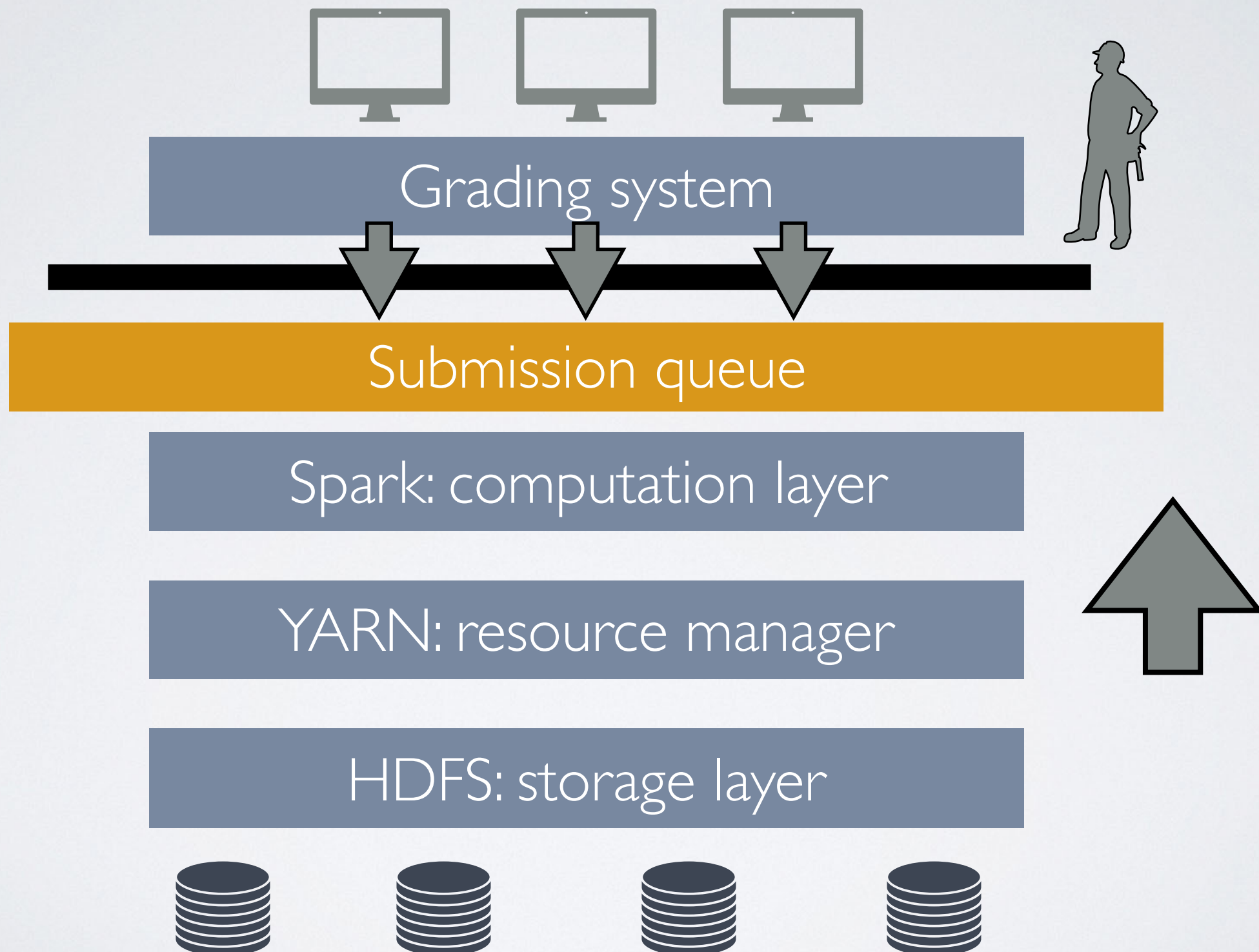
Deployment on Vocareum

Outcome

CHALLENGES

- Auto-grading system
- Centralized instructor interface to check submission status and grades
- A scalable cluster to handle heavy big spatial data analytics workload (up to 100 GB data, take more than half hour)
- Collect statistical information: runtime, memory utilization, CPU utilization
- Concurrent submissions

GRADING SYSTEM



INSTRUCTOR INTERFACE

Vocareum

(Ends - Dec 31 2017)

AssignmentsEnrollmentDashboardGradebookSettings

CSE512-Project-Phase2

SummarySubmissionsAuto-GradeGradersPlagiarismControlsExport

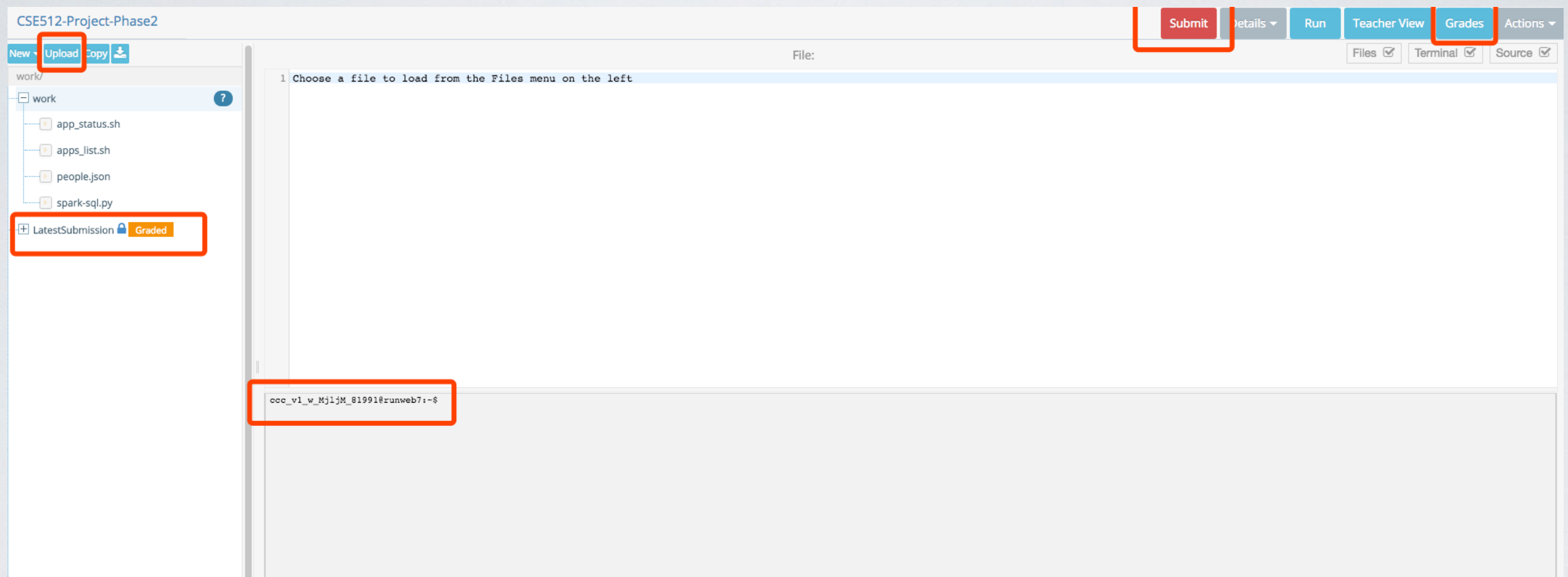
Refresh

User	Status	Scheduled	Start	End	Run-time	Queue-time
AlphaDB	Completed	2017/10/29 16:40:55	2017/10/29 16:40:56	2017/10/29 16:50:49	00:09:53	00:00:01
ASU DataDevils	Completed	2017/10/31 07:12:02	2017/10/31 07:12:07	2017/10/31 07:22:09	00:10:02	00:00:05
Aviated	Completed	2017/11/01 10:23:48	2017/11/01 10:23:55	2017/11/01 10:34:59	00:11:04	00:00:07
BigData & Spark	Completed	2017/11/01 16:19:32	2017/11/01 16:19:33	2017/11/01 16:41:21	00:21:48	00:00:01
Cloud9	Completed	2017/10/31 02:20:00	2017/10/31 02:20:03	2017/10/31 02:30:34	00:10:31	00:00:03
Dark Knights	Completed	2017/11/01 18:42:55	2017/11/01 18:42:57	2017/11/01 18:54:09	00:11:12	00:00:02
Data Invaders	Completed	2017/10/30 12:24:43	2017/10/30 12:24:48	2017/10/30 12:45:06	00:20:18	00:00:05
DataBassHunterz	Completed	2017/10/30 15:25:03	2017/10/30 15:25:09	2017/10/30 15:44:32	00:19:23	00:00:06
DataFalcons	Completed	2017/10/30 16:16:13	2017/10/30 16:16:14	2017/10/30 16:27:27	00:11:13	00:00:01
DBT	Completed	2017/10/30 22:41:01	2017/10/30 22:41:06	2017/10/30 23:02:19	00:21:13	00:00:05
DDS Project	Completed	2017/10/30 10:32:35	2017/10/30 10:32:40	2017/10/30 10:43:53	00:11:13	00:00:05
Deadpool	Completed	2017/10/29 18:08:57	2017/10/29 18:08:59	2017/10/29 18:19:44	00:10:45	00:00:02
Death Star	Completed	2017/11/01 09:26:20	2017/11/01 09:26:24	2017/11/01 09:36:56	00:10:32	00:00:04
Distributers	Completed	2017/11/01 13:54:25	2017/11/01 13:54:30	2017/11/01 14:04:23	00:09:53	00:00:05
Divide and Conquer	Completed	2017/11/01 14:06:56	2017/11/01 14:07:01	2017/11/01 14:17:42	00:10:41	00:00:05
Fantastic Four	Completed	2017/10/31 09:22:23	2017/10/31 09:22:26	2017/10/31 09:32:28	00:10:02	00:00:03

INSTRUCTOR INTERFACE

▼ Vocareum (Ends - Dec 31 2017)		Assignments			
▼ CSE512-Project-Phase2		Summary	Submissions	Auto-Grade	
User	Last submission	Graded	Grader	Grade	
AlphaDB	On-time	Y			
ASU DataDevils	On-time	Y			
Aviated	On-time	Y			
BigData & Spark	On-time	Y			
Cloud9	On-time	Y			
Dark Knights	On-time	Y			
Data Invaders	On-time	Y			
DataBassHunterz	On-time	Y			
DataFalcons	On-time	Y			
DBT	On-time	Y			
DDS Project	On-time	Y			
Deadpool	On-time	Y			
Death Star	On-time	Y			
Distributers	On-time	Y			
Divide and Conquer	On-time	Y			
Fantastic Four	On-time	Y			
First Aid Kit	On-time	Y			
...			

STUDENT INTERFACE



THIS TALK

Course background

Deployment on Vocareum

Outcome

OUTCOME

- Instructor perspective:
 - Unified grading environment
 - Lots of efforts on designing the grading system
 - Much less efforts on grading the submissions
- Student perspective:
 - More opportunities to submit
 - Powerful cluster for testing

SOME DATA

- ASU CSE 512 Distribute Data System
- 200 students, 48 groups
- Execution time: MIN=9min, AVG=13.5min, MAX=29min
- Boosted the student grade 4% as compared to the past offerings of the same class