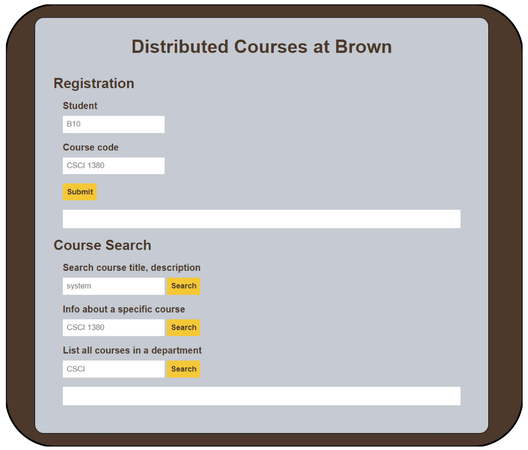
CourseCluster: A Distributed Approach to University Course Catalogs and Registration

Ben Bachmann, Alex Lin, Ethan Williams

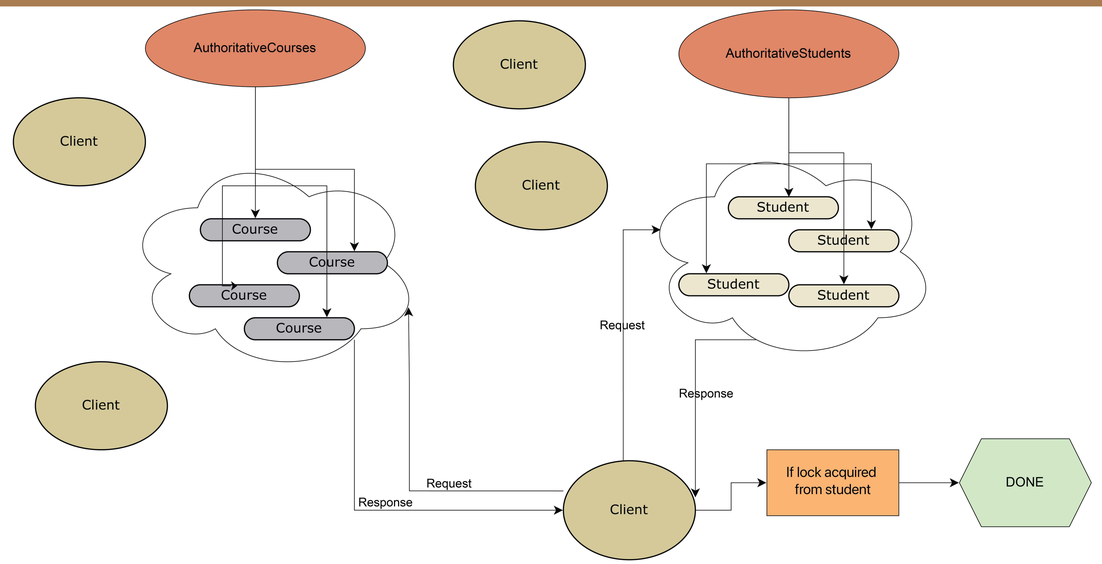
Demo

Introduction

We present CourseCluster, a distributed system designed to handle the complexities of course registration at a university level. CourseCluster is designed to manage large volumes of student and course data, ensuring scalability and data consistency across nodes. We deployed our system to an Elastic Compute Cloud (EC2) instance on Amazon Web Services (AWS) to enhance performance and stability across multiple nodes. We connected the system to a simple interface. CourseCluster enables students to search for courses by keyword, course ID, or academic department, and to register for a desired course with a single click. The entire functionality is hosted on a single web page.



System Architecture



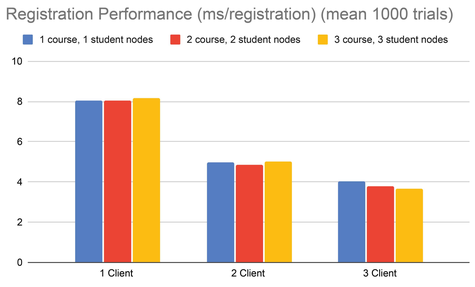
Performance & Metrics

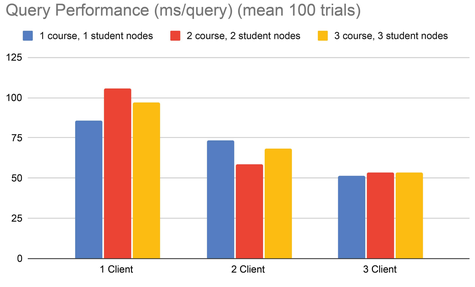
Table 1. Course Cluster Search, Query, and Registration Benchmarks On Different Node Configurations

Requests	1 client, 1 search, 1 course	1 client, 2 search, 2 course	1 client, 3 search, 3 course	2 client, 1 search, 1 course	2 client, 2 search, 2 course	2 client, 3 search, 3 course	3 client, 1 search, 1 course	3 client, 2 search, 2 course	3 client, 3 search, 3 course
	Search: "CSCI 1380"								
1		64.62	47.88	55.50	51.48	58.78	45.26		48.32
100	977.30	1,130.62	1,284.51	665.90	780.64	828.58	508.87	527.89	552.87
	Query: 'distributed systems'								
10	1,480.04	2,392.18	1,397.67	1,267.94	1,405.39	1,261.93	1,504.02	1,136.05	1,121.03
100	8,562.71	10,568.53	9,703.96	7,361.90	5,884.58	6,853.01	5,164.52	5,322.76	5,323.10
	Registration								
1	100.07	62.16	59.27	101.67	65.25	64.75	50.94	64.12	54.32
100	1,297.12	1,143.26	1,247.43	619.03	687.46	4,805.94	506.85	444.43	401.58
1000	8,032.95	8,039.47	8,162.19	4,973.94	4,850.22	5,015.11	4,051.19	3,798.15	3,646.23
	Time to complete (ms)								

Table 2. Course Cluster Cold Start Benchmarks On Different Node Configurations

1 client, 1 search, 1 course	1 client, 2 search, 2 course	1 client, 3 search, 3 course	2 client, 1 search, 1 course	2 client, 2 search, 2 course	2 client, 3 search, 3 course	3 client, 1 search, 1 course	3 client, 2 search, 2 course	3 client, 3 search, 3 course
8,516.54	7,621.84	7,720.08	8,349.01	7,250.73	7,583.58	8,453.30	7,193.80	8,031.38
Time to complete (ms)								





Key Details

Nodes Services

- authoritativeCourses
  - simulates db w/ data of all courses
- authoritativeStudents
  - simulates db w/ data of all students
- students
  - prepare, abort, commit services for course registration
  - listCourses to view all courses student is registered for
- courses
  - prepare, abort, commit services for course registration
  - search, TF-IDF indexing of course descriptions, names
  - search by text query, specific course, or by department
- clients
  - middleman between for frontend and distributed internal nodes

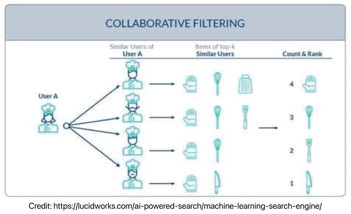
Future Work

ML Integration

- Course recommendations, intelligent search, or dynamically repartitioning courses

Education

- Given improved search and scalability, one could imagine our system scaling to millions of nodes. This enables the distribution of our course registration system across multiple universities.



Credit: https://lucidworks.com/ai-powered-search/machine-learning-search-engine/