

C S 55B: DATABASE ESSENTIALS IN AMAZON WEB SERVICES

Foothill College Course Outline of Record

Heading	Value
Effective Term:	Summer 2024
Units:	4.5
Hours:	4 lecture, 2 laboratory per week (72 total per quarter)
Advisory:	C S 30A, C S 50A, C S 55A and C S 55C.
Degree & Credit Status:	Degree-Applicable Credit Course
Foothill GE:	Non-GE
Transferable:	CSU
Grade Type:	Letter Grade (Request for Pass/No Pass)
Repeatability:	Not Repeatable

Student Learning Outcomes

- A successful student will demonstrate an understanding of and be able to describe the basics of database technologies including their need, transactions, indexing, keys and components.
- A successful student will be able to demonstrate how to create a relational database and interface to an Amazon cloud based Relational Database System.
- A successful student will be able to describe that differences between uses of traditional relational and NoSQL database systems.

Description

This course addresses database fundamentals and cloud database design patterns and management. A wide variety of database needs are presented, such as structured, semi-structured, and unstructured datasets, and how those are supported in the cloud. Students learn to deploy a SQL database on infrastructure components and perform basic data operations on that infrastructure. Students then take those basic concepts and learn managed platform as a service solutions, such as Amazon RDS, Amazon DynamoDB, Amazon Kinesis stream processing/ analytics, in memory database accelerators, and ML big data tools. Basic database administration skills, such as migration, backups, restoration, retention, service configuration, high availability, and service scaling, are presented.

Course Objectives

The student will be able to:

1. Apply for an Amazon (AWS) account and Amazon Educate account
2. Understand and describe the basics of database technology, including their need, transactions, indexing, keys, and components
3. Demonstrate how to create and interface to a SQL based database on Amazon IaaS platform
4. Demonstrate how to create and interface to an Amazon cloud based Relational Database System (RDS) and compare to IaaS

5. Understand how to store and access relational data from the cloud RDS through application and programmatic methods
6. Explain the need for and principles of non-relational, unstructured database technology via Amazon DynamoDB
7. Demonstrate how to store and access non-relational data using the DynamoDB through application and programmatic methods
8. Understand and describe how to migrate exiting RDBMS instances into the Amazon RDS system using the AWS Database Migration service
9. Demonstrate the basics of monitoring, managing, backing up, and restoring RDS instances using Amazon tools
10. Understand the basics of backing up, restoring, and securing using encryption with Amazon DynamoDB instances
11. Understand how to accelerate NoSQL database workloads using DynamoDB Accelerator

Course Content

1. AWS access
 - a. AWS account acquisition
 - b. AWS Educate account acquisition
2. Database technology
 - a. Commercial database examples
 - b. Requirements for database technology vs. flat file storage
 - c. Transaction processing, including atomicity, consistency, isolation, and durability
 - d. Indexing principles
 - e. Modern Relational Database Systems (RDBMS) components
3. IaaS database services
 - a. MySQL service deployment to AWS EC2 IaaS
 - b. Remote console access to MySQL and basic SQL commands
 - c. Review of Database Administration duties
4. AWS Managed Database services
 - a. Amazon RDS service review including features and pricing
 - b. Amazon Aurora service review including features and pricing
 - c. Creation of RDS and Aurora databases including remote access via MySQL Workbench tool
5. Working with SQL data in the cloud
 - a. Review of data, metadata, and datatypes
 - b. Basic SQL commands
 - i. CREATE TABLE
 - ii. INSERT INTO
 - iii. DROP TABLE
 - iv. TRUNCATE TABLE
 - v. Primary Key
 - vi. Foreign key basics
 - vii. Complex queries using INNER JOIN and FOREIGN KEY
6. Course project: deploying a complex, real time, serverless data processing application using the following Amazon Services:
 - a. Amazon Cognito authorization/authentication
 - b. Amazon Kinesis data stream processing
 - c. Amazon Lambda serverless event processing
 - d. Amazon SQS queue service
 - e. Amazon DynamoDB unstructured NoSQL database
 - f. Amazon S3 object storage
 - g. Amazon Athena analytics

7. Semi-structured and unstructured databases
 - a. Introduction and need
 - i. Concepts of volume, velocity, variety, and veracity
 - ii. Introduction to Big Data and example of Amazon Rekognition image analysis service
 - iii. Introduction to Amazon DynamoDB NoSQL database service
 1. Tables and table creation
 2. Items
 3. Primary Key
 4. Secondary Index
8. Amazon DynamoDB usage
 - a. Creating and launching an Amazon DynamoDB table
 - b. Managing and securing the Amazon DynamoDB instance
 - c. Basics of JavaScript Object Notation (JSON)
9. Monitoring and managing RDS instances
 - a. Monitoring fundamentals
 - i. Scaling and high availability overview
 - ii. Performance metrics, including CPU, memory, disk space, and input/output operations (IOPS), throughput, and connections
 - iii. Monitoring via RDS console
 - iv. Monitoring via Trusted Advisor dashboard
 - v. Monitoring via Amazon CloudWatch
 - b. Backup and restore concepts
 - i. Strategies
 - ii. RDS backup
 - iii. Snapshot creation and restore
10. Managing Amazon DynamoDB instances
 - a. Backup and restore concepts
 - b. Creation and restore
 - c. Encryption
 - d. Acceleration using DynamoDB DAX connecting to Amazon DynamoDB via NoSQL Workbench
11. Database migration concepts
 - a. Understanding and using the Amazon Database Migration Service (AWS DMS)
 - b. Migrating data from source DB to target DB
 - i. Pricing model
 1. Source DB DBMS types

Lab Content

1. Create AWS working environment including AWS account, AWS Educate Account, and AWS Command line interface on local computer
2. Install node, npm, and the MySQL Workbench on local computer
3. Create an IaaS platform using Amazon EC2, remotely connect via SSH, and install/configure MySQL database
4. Execute SQL commands on an IaaS platform
5. Create a standard AWS Relational Database (RDS) and establish MySQL Workbench connectivity
6. Create a standard AWS Aurora Database and establish MySQL Workbench connectivity
7. Perform basic SQL operations with the AWS RDS via MySQL Workbench
8. Working with real world relational data in the cloud

- a. Install and verify the dataset of recent bank closures required by the FDIC, the Federal Deposit Insurance Corporation (dataset will be provided)
- b. Install and verify the dataset of about 1000 cameras with 13 properties, including weight, length, and price (dataset will be provided)
9. Perform advanced SQL operations with AWS RDS via MySQL Workbench
10. Analyze big data images with Amazon Rekognition
11. Install and connect to a Amazon DynamoDB NoSQL database using NoSQL Workbench
12. Perform a variety of experiments with the Amazon DynamoDB

Special Facilities and/or Equipment

1. Access to a computer with a web browser compatible with the Foothill learning management system and AWS Console.
2. A payment method for accessing AWS services (credit/debit/stored value card). AWS as a commercial service requires all accounts to be paid for accounts. AWS will provide credits and no actual spending with normal class use would be incurred.
3. A learning management system with an assignment posting component (through which all lab assignments are to be submitted) and a forum component (where students can discuss course material and receive help from the instructor). This applies to all sections, including on campus (i.e., face-to-face) offerings.
4. The college will provide a fully functional and maintained course management system through which the instructor and students can interact.
5. Students must have email accounts and ongoing access to computers with internet capabilities.

Method(s) of Evaluation

Methods of Evaluation may include but are not limited to the following:

Tests and quizzes

Written laboratory assignments, which include detailed instructions, sample runs, and documentation

Final examination

Completion of class project implementing and describing a stream based database capture and analysis and archival application

Method(s) of Instruction

Methods of Instruction may include but are not limited to the following:

Lectures which include motivation for the architecture of the specific topics being discussed

In-person or online labs (for all sections, including those meeting face-to-face/on campus), consisting of:

1. An assignment webpage located on a college-hosted course management system or other department-approved internet environment. Here, the students will review the specification of each assignment and submit their completed lab work
 2. A discussion webpage located on a college-hosted course management system or other department-approved internet environment. Here, students can request assistance from the instructor and interact publicly with other class members
- Detailed review of laboratory assignments which includes model solutions and specific comments on the student submissions

In person or online discussion which engages students and instructor in an ongoing dialog pertaining to all aspects of designing, implementing, and analyzing programs

When course is taught fully online:

1. Instructor-authored lecture materials, handouts, syllabus, assignments, tests, and other relevant course material will be delivered through a college-hosted course management system or other department-approved internet environment
2. Additional instructional guidelines for this course are listed in the addendum of C S department online practices

Representative Text(s) and Other Materials

All course materials provided by instructor through the online course management system.

Types and/or Examples of Required Reading, Writing, and Outside of Class Assignments

1. Reading:
 - a. Reading the supplied handouts and modules averaging 30 pages per week
 - b. Reading online resources as directed by instructor through links pertinent to the course
 - c. Watching video presentations by AWS and other cloud providers as contained in the course
 - d. Reading library and reference material directed by instructor through course handouts
2. Writing:
 - a. Writing technical prose documentation that supports and describes the programs that are submitted for grades

Discipline(s)

Computer Science