Course Outline

Data Warehousing on AWS

Course Summary

Description

Data Warehousing on AWS introduces you to concepts, strategies, and best practices for designing a cloud-based data warehousing solution using Amazon Redshift, the petabyte-scale data warehouse in AWS. This course demonstrates how to collect, store, and prepare data for the data warehouse by using other AWS services such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose, and Amazon S3. Additionally, this course demonstrates how to use business intelligence tools to perform analysis on your data.

Objectives

After taking this course, students will learn to:

- Discuss the core concepts of data warehousing.
- Evaluate the relationship between Amazon Redshift and other big data systems.
- Evaluate use cases for data warehousing workloads and review case studies that demonstrate implementation of AWS data and analytic services as part of a data warehousing solution.
- Choose an appropriate Amazon Redshift node type and size for your data needs.
- Discuss security features as they pertain to Amazon Redshift, such as encryption, IAM permissions, and database permissions.
- Launch an Amazon Redshift cluster and use the components, features, and functionality to implement a data warehouse in the cloud.
- Use other AWS data and analytic services, such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose, and
- Amazon S3, to contribute to the data warehousing solution.
- Evaluate approaches and methodologies for designing data warehouses.

- Identify data sources and assess requirements that affect the data warehouse design.
- Design the data warehouse to make effective use of compression, data distribution, and sort methods.
- Load and unload data and perform data maintenance tasks.
- Write queries and evaluate query plans to optimize query performance.
- Configure the database to allocate resources such as memory to query queues and define criteria to route certain types of queries to your configured query queues for improved processing.
- Use features and services, such as Amazon Redshift database audit logging, Amazon CloudTrail, Amazon CloudWatch, and Amazon Simple Notification Service (Amazon SNS), to audit, monitor, and receive event notifications about activities in the data warehouse.
- Prepare for operational tasks, such as resizing Amazon Redshift clusters and using snapshots to back up and restore clusters

Course Outline

Data Warehousing on AWS

Course Summary (cont.)

Topics

- Introduction to Data Warehousing
- Introduction to Amazon Redshift
- Launching Clusters
- Designing the Database Schema
- Identifying Data Sources
- Loading Data

- Writing Queries and Tuning Performance
- Amazon Redshift Spectrum
- Maintaining Clusters
- Analyzing and Visualizing Data

Audience

- Database architects
- Database administrators
-

- Database developers
- Data analysts and scientists

Prerequisite

- Courses taken: AWS Technical Essentials (or equivalent experience with AWS)

- Familiarity with relational databases and database design concepts

Duration

Three Days