Building Streaming Data Analytics Solutions on AWS

Duration: *1 Day(s)*

Course Overview

In this course, you will learn to build streaming data analytics solutions using AWS services, including Amazon Kinesis and Amazon Managed Streaming for Apache Kafka (Amazon MSK). Amazon Kinesis is a massively scalable and durable real-time data streaming service. Amazon MSK offers a secure, fully managed, and highly available Apache Kafka service. You will learn how Amazon Kinesis and Amazon MSK integrate with AWS services such as AWS Glue and AWS Lambda. The course addresses the streaming data ingestion, stream storage, and stream processing components of the data analytics pipeline. You will also learn to apply security, performance, and cost management best practices to the operation of Kinesis and Amazon MSK.

Review this course online at https://www.alta3.com/courses/AWS-DA-STREAM

Objectives

- Understand modern data architecture and AWS streaming services integration.
- Design and implement effective streaming data analytics solutions.
- Apply best practices for security, performance, and cost management.
- Select and deploy appropriate data ingestion and processing options.

Who Should Attend

- Data engineers and architects
- Developers seeking real-time streaming solutions
- IT professionals managing AWS environments
- Business analysts interested in data analytics

Prerequisites

- At least one year of data analytics experience or direct experience building real-time applications or streaming analytics solutions. We suggest the Streaming Data Solutions on AWS whitepaper for those that need a refresher on streaming concepts.
- Completed either Architecting on AWS or Data Analytics Fundamentals
- Completed Building Data Lakes on AWS

Course Outline

Module A: Overview of Data Analytics and the Data Pipeline

1. Data analytics use cases
2. Using the data pipeline for analytics

Module 1: Using Streaming Services in the Data Analytics Pipeline

3. The importance of streaming data analytics

4. The streaming data analytics pipeline
5. Streaming concepts

Module 2: Introduction to AWS Streaming Services

6. Streaming data services in AWS
7. Amazon Kinesis in analytics solutions
8. Demonstration: Explore Amazon Kinesis Data Streams
9. Practice Lab: Setting up a streaming delivery pipeline with Amazon Kinesis
10. Using Amazon Kinesis Data Analytics
11. Introduction to Amazon MSK
12. Overview of Spark Streaming

Module 3: Using Amazon Kinesis for Real-time Data Analytics

13. Exploring Amazon Kinesis using a clickstream workload
14. Creating Kinesis data and delivery streams
15. Demonstration: Understanding producers and consumers
16. Building stream producers
17. Building stream consumers
18. Building and deploying Flink applications in Kinesis Data Analytics
19. Demonstration: Explore Zeppelin notebooks for Kinesis Data Analytics
20. Practice Lab: Streaming analytics with Amazon Kinesis Data Analytics and Apache Flink

Module 4: Securing, Monitoring, and Optimizing Amazon Kinesis

21. Optimize Amazon Kinesis to gain actionable business insights
22. Security and monitoring best practices

Module 5: Using Amazon MSK in Streaming Data Analytics Solutions

23. Use cases for Amazon MSK
24. Creating MSK clusters
25. Demonstration: Provisioning an MSK Cluster
26. Ingesting data into Amazon MSK
27. Practice Lab: Introduction to access control with Amazon MSK
28. Transforming and processing in Amazon MSK

Module 6: Securing, Monitoring, and Optimizing Amazon MSK

29. Optimizing Amazon MSK
30. Demonstration: Scaling up Amazon MSK storage
31. Practice Lab: Amazon MSK streaming pipeline and application deployment
32. Security and monitoring

33. Demonstration: Monitoring an MSK cluster

Module 7: Designing Streaming Data Analytics Solutions

34. Use case review

35. Class Exercise: Designing a streaming data analytics workflow

Module B: Developing Modern Data Architectures on AWS

36. Modern data architectures