Building Streaming Data Analytics Solutions on AWS

Amazon

- **Nível:**
- **Duração:** 7h

Sobre o curso

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.

Destinatários

This course is intended for:

- Data engineers and architects
- Developers who want to build and manage real-time applications and streaming data analytics solutions

---

Objetivos

In this course, you will learn to:

- Understand the features and benefits of a modern data architecture. Learn how AWS streaming services fit into a modern data architecture.
- Design and implement a streaming data analytics solution
- Identify and apply appropriate techniques, such as compression, sharding, and partitioning, to optimize data storage
- Select and deploy appropriate options to ingest, transform, and store real-time and near real-time data
- Choose the appropriate streams, clusters, topics, scaling approach, and network topology for a particular business use case
- Understand how data storage and processing affect the analysis and visualization mechanisms needed to gain actionable business insights
- Secure streaming data at rest and in transit

- Monitor analytics workloads to identify and remediate problems
- Apply cost management best practices

Pré-requisitos

We recommend that attendees of this course have:

- 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

Programa

- Overview of Data Analytics and the Data Pipeline
- Using Streaming Services in the Data Analytics Pipeline
- Introduction to AWS Streaming Services
- Using Amazon Kinesis for Real-time Data Analytics
- Securing, Monitoring, and Optimizing Amazon Kinesis
- Using Amazon MSK in Streaming Data Analytics Solutions
- Securing, Monitoring, and Optimizing Amazon MSK
- Designing Streaming Data Analytics Solutions
- Developing Modern Data Architectures on AWS

Overview of Data Analytics and the Data Pipeline

- Data analytics use cases
- Using the data pipeline for analytics

Using Streaming Services in the Data Analytics Pipeline

- The importance of streaming data analytics
- The streaming data analytics pipeline
- Streaming concepts

Introduction to AWS Streaming Services

- Streaming data services in AWS
- Amazon Kinesis in analytics solutions
- Demonstration: Explore Amazon Kinesis Data Streams
- Practice Lab: Setting up a streaming delivery pipeline with Amazon Kinesis
- Using Amazon Kinesis Data Analytics
- Introduction to Amazon MSK
- Overview of Spark Streaming

Using Amazon Kinesis for Real-time Data Analytics

- Exploring Amazon Kinesis using a clickstream workload
- Creating Kinesis data and delivery streams
- Demonstration: Understanding producers and consumers
- Building stream producers
- Building stream consumers
- Building and deploying Flink applications in Kinesis Data Analytics
- Demonstration: Explore Zeppelin notebooks for Kinesis Data Analytics
- Practice Lab: Streaming analytics with Amazon Kinesis Data Analytics and Apache Flink

Securing, Monitoring, and Optimizing Amazon Kinesis

- Optimize Amazon Kinesis to gain actionable business insights
- Security and monitoring best practices

Using Amazon MSK in Streaming Data Analytics Solutions

- Use cases for Amazon MSK
- Creating MSK clusters
- Demonstration: Provisioning an MSK Cluster
- Ingesting data into Amazon MSK
- Practice Lab: Introduction to access control with Amazon MSK
- Transforming and processing in Amazon MSK

Securing, Monitoring, and Optimizing Amazon MSK

- Optimizing Amazon MSK
- Demonstration: Scaling up Amazon MSK storage
- Practice Lab: Amazon MSK streaming pipeline and application deployment
- Security and monitoring
- Demonstration: Monitoring an MSK cluster

Designing Streaming Data Analytics Solutions

- Use case review
- Class Exercise: Designing a streaming data analytics workflow

Developing Modern Data Architectures on AWS

- Modern data architectures