

**Building Streaming Data Analytics Solutions on AWS**

**Course Number:** AWS-176  
**Duration:** 1 day

**Overview**

This Building Streaming Data Analytics Solutions on AWS training course teaches attendees how to use Amazon Kinesis and Amazon MSK to build streaming data analytics solutions on AWS. Participants learn how to select and deploy appropriate options to ingest, transform, store, and analyze data.

Accelebrate is an AWS Training Partner (ATP) and this hands-on official AWS Classroom Training course is taught by an accredited Amazon Authorized Instructor (AAI).

**Prerequisites**

* One year of experience managing data analytics solutions or streaming data is helpful but not required
* Completed either [Architecting on AWS](file:////training/aws-architecture) or the self-paced course, [Data Analytics Fundamentals](https://aws.amazon.com/training/digital/data-analytics-fundamentals/)
* Completed Building Data Lakes on AWS

We suggest reviewing the [Streaming Data Solutions on AWS whitepaper](https://d0.awsstatic.com/whitepapers/whitepaper-streaming-data-solutions-on-aws-with-amazon-kinesis.pdf) for those that need a refresher on streaming concepts.

**Materials**

All AWS training students receive comprehensive courseware.

**Software Needed on Each Student PC**

A modern web browser and an Internet connection that allows connections by SSH or Remote Desktop (RDP) into AWS virtual machines.

**Objectives**

* Understand the features and benefits of 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
* Secure, monitor, and optimize Amazon Kinesis and Amazon MSK
* Apply cost management best practices

**Outline**

* Introduction
* 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
  + 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
  + Understanding producers and consumers
  + Building stream producers
  + Building stream consumers
  + Building and deploying Flink applications in Kinesis Data Analytics
  + Explore Zeppelin notebooks for Kinesis Data Analytics
  + 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
  + Provisioning an MSK Cluster
  + Ingesting data into Amazon MSK
  + 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
  + Monitoring an MSK cluster
* Designing Streaming Data Analytics Solutions
  + Use case review
  + Designing a streaming data analytics workflow
* Developing Modern Data Architectures on AWS
* Conclusion