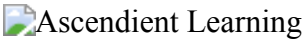

IBM - Apache Kafka Administration by Confluent

Course Number: ADMINKAFKAG Duration: 3 days

Course Description

Gain hands-on expertise in Apache Kafka and the Confluent Platform. Learn to deploy, secure, and optimize clusters using proven industry standards, and build seamless data pipelines with Kafka Connect through practical, scenario-based exercises.

Skills Gained

Course Objectives

During this hands-on course you will learn how:

- Kafka and the Confluent platform work, and how their main subsystems interact
- To set up, manage, monitor and tune your cluster
- To use industry best practices developed by the world?s foremost Apache Kafka experts

Hands-on Training

Throughout the course, hands-on exercises reinforce the topics being discussed. Exercises include:

- Using Kafka?s command-line tools
- Exploring configuration
- Using Kafka?s administrative tools
- Tuning Producer and Consumer performance
- Securing the cluster
- Building data pipelines with Kafka Connect

Who Can Benefit

This course is designed for System Administrators and Operations staff responsible for building, managing, monitoring, and tuning Kafka clusters and components of Confluent Platform. Architects and DevOps staff who want a tour on how the Central Services (Kafka, KRaft, Schema Connect, Registry, etc.) work are also welcome.

Prerequisites

Attendees should have a working knowledge of the Kafka architecture. It is also important to have strong knowledge of Linux/Unix and understand basic TCP/IP networking concepts. Familiarity with Java Virtual Machine (JVM) is helpful. Participants are required to provide a laptop computer with unobstructed internet access to fully participate in the class.

Course Details

Outline

Bridging From Fundamentals

- Fundamentals Review
- Replication Review

Replicating Data: A Deeper Dive

- Which messages can be consumed?
- Managing Replica Placement
- Partition Leadership Tracking
- Partition Follower Responsiveness

Producing Messages Reliably

- Message Production Acknowledgement
- Dealing with same-session duplicates in Producers
- Transactional Messages Production

Providing Durability in Other Ways

- How Kafka Organizes Partition Data Files in Local Storage
- Beyond Local Storage (Tiered Storage)
- Persistence of Consumer Offsets

Configuring a Kafka Cluster

- Static Kafka Server Configuration
- Dynamic Server Configuration
- Topic Dynamic Configuration Overrides

Managing a Kafka Cluster

- Notes on Installation and Upgrading
- Kafka Server Roles: Controllers and Brokers (KRaft)
- Basics of Kafka Monitoring
- Message Retention Policies
- Replica Workload Management (Self-Balancing Cluster)
- Elastic Cluster: adding and removing Brokers

Balancing Load with Consumer Groups

- How do Partitions and Consumer Scale
- How does Kafka Manage Consumer Groups

Introduction to Performance Optimization

- Sending Multiple Messages in one Request (Batching)
- Broker Anatomy for Produce and Fetch Requests
- Metrics and Configuration related to Broker Anatomy
- Client Quotas
- Performance Tests

Securing a Kafka Cluster

- Security Concepts applicable to Kafka
- Options to Secure Kafka/Confluent Deployments
- Authorization (Access Control Lists - ACL)

Kafka Connect

- Concepts related to Kafka Connect
- Connect Workers and Connector Configuration
- Finding and Deploying Connectors

Deploying Kafka in Production

- Reference Architecture for Apache Kafka and the Confluent Platform
- Capacity Planning