

Confluent Kafka

Duration: 3 Days

Description

Confluent Kafka is a distributed data streaming platform built on top of and extended by Apache Kafka. Apache Kafka is a distributed data streaming platform used to ingest, store, process and transmit large amounts of data in real time. Confluent offers an enhanced and managed version of the Kafka ecosystem, making this platform easier to use and more powerful.

Confluent Kafka is used in big data processing, real-time analytics, event-driven application development and many more scenarios. It helps businesses manage and evaluate their data more effectively.

Outline

Module 1: Introduction to Basic Kafka Concepts and Confluent Kafka

- What are Apache Kafka and Confluent Kafka?
- Distributed data flow concepts and their importance
- Use cases of Confluent Kafka

Module 2: Apache Kafka Fundamentals

- Kafka architecture: brokers, subjects, scores, producers, consumers, etc.
- Data distribution and replication
- Kafka message format and structure

Module 3: Confluent Platform Overview

- Introduction to Confluent Platform components
- Confluent Control Center: Monitoring and management
- Confluent Schema Registry: Data serialization and compliance
- Kafka Connect: Integrating external systems with Kafka
- Confluent KSQL: Data processing with SQL

Module 4: Designing and Creating Kafka Topics

- Planning and designing Kafka topics
- Creating and configuring Kafka topics

- Selection of the number of partitions and replication factors

Module 5: Data Generation for Kafka Topics

- Writing Kafka generators using Confluent APIs
- Generating serialized data to subjects with Json
- Configuring Kafka producer settings

Module 6: Data Consumption from Kafka Topics

- Kafka consumer development using Confluent APIs
- Consumer offsets and department assignments management
- Processing data serialized with json from subjects

Module 7: Data Integration with Kafka Connect

- Introduction to Kafka Connect architecture
- Install and configure source and destination connectors
- Handling data format conversions

Module 8: Data Serialization and Compliance with Schema Registration

- Understanding and advantages of the euro
- Save and refine schemas in the Schema Record
- Ensuring backward and forward compatibility

Module 9: Data Stream Processing with KSQL

- Introduction to data stream processing concepts
- Write KSQL queries for real-time data processing
- Create and manage KSQL flows and tables

Module 10: Scaling and Managing Confluent Kafka

- Cluster sizing and hardware requirements
- Consider Confluent Kafka deployment
- Manage and monitor Confluent Kafka clusters

Module 11: Real-World Use Cases and Best Practices

- Implement application scenarios with Confluent Kafka
- Design patterns and best practices for building resilient applications
- Troubleshooting common problems and challenges

Prerequisites

A basic understanding of computer programming and databases.