Apache Kafka

Course Description

Apache Kafka is a distributed streaming platform that is widely used for building real-time data pipelines and streaming applications.

This course is designed to provide a deep understanding of Apache Kafka, a powerful distributed streaming platform used by thousands of companies for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications. Whether you are a beginner looking to get started with Kafka or an experienced user aiming to sharpen your skills, this course will guide you through the fundamentals to advanced concepts and best practices.

Key Learning Objectives

- Understand the core concepts and architecture of Apache Kafka.
- Learn how to install and configure Kafka clusters.
- Master Kafka Producers and Consumers for building efficient data pipelines.
- Gain hands-on experience with Kafka Streams and KSQL for real-time data processing
- Gain hands-on experience with Kafka Producer, Consumer and Kafka Streams.

Course Prerequisite

- Java Programming (Basic)
- Python Programming (Basics)

Course Outline

Week	Module	Planned Date
Day 01	**Module 01 - Getting Started with Kafka and Core API's** 4 Hrs	Jun 24, 2024

Week	Module	Planned Date
Day 02	**Advance Kafka Configuration** 1 Hrs **Exploring Kafka Core API** 3 Hrs	Jun 25, 2024
	Deep Dive Kafka Core API using Spring Boot 4 Hrs	Jun 26, 2024
	Kafka Streams 4 Hrs	Jun 27, 2024

Course Content

Getting Started with Kafka and Core API's

- [ ] Integration between components
- [ ] What is Kafka
- [ ] Components of Messaging System
- [ ] Understanding Kafka components in detail
    - [ ] Producer
    - [ ] Consumer
    - [ ] Broker
    - [ ] Cluster
    - [ ] Topic
    - [ ] Partitions
    - [ ] Offset
    - [ ] Consumer groups

- [ ] Message Retention in kafka
- [ ] Kafka Commit Log
- [ ] Kafka
    - [ ] Starting Zookeeper
    - [ ] Starting Kafka Server
    - [ ] Topic operations: create, list, delete, describe
    - [ ] Publishing data to a topic using console producer
    - [ ] Publishing data to a topic using console consumer
    - [ ] Sending and receiving messages
- [ ] **Hands on** – Kafka Cluster with Multiple Brokers
    - [ ] Creating separate configuration files for brokers
    - [ ] Launching multiple brokers
    - [ ] Getting cluster information and broker details from Zookeeper
- [ ] **Hands on** – Topic with multiple partitions
    - [ ] Creating topic with multiple partitions
    - [ ] How messages are spread across partitions
    - [ ] Reading messages from specific partitions
    - [ ] Reading messages from specific offset in specific partition

Advanced Kafka Configuration

- [ ] Broker Configuration
- [ ] Producer Configuration
- [ ] Consumer Configuration
- [ ] Topic Configuration
- [ ] 3.6. Performance Tuning

Deep Dive Kafka Core API using Spring Boot

- [ ] Understanding Producer Partitioning Mechanism using Java
- [ ] Different ways to implement partitioning mechanism
    - [ ] Providing partition number
    - [ ] Using Round Robin
    - [ ] Key Hashing
- [ ] Messaging Sending
- [ ] Producer API

- [ ] Synchronous Send
- [ ] ASynchronous Send

Module 05 - Kafka Streams

- [ ] Introduction to Stream Processing
- [ ] Introduction to Kafka Streams
- [ ] Other Stream Processing Framework
- [ ] Adding Kafka Stream Dependency
- [ ] Implementing Kafka Streams