Kafka Streams

Prerequisites: Basic knowledge of java and Apache Kafka

Duration: 4 Days (8 Hrs./Day)

Course Objective: This course is designed to provide a comprehensive understanding of Apache Kafka Streams, guiding you through its foundational concepts to advanced operations. You will explore Kafka Streams architecture, core concepts, and practical applications. By the end of the course, you'll be able to develop, deploy, and manage Kafka Streams applications, ensuring efficient stream processing and real-time data handling.

Apache Kafka version: Latest.

Lab Requirement: Koenig DC/Linux (CentOS 9) (customizable).

Module 1 - Introduction to Kafka Streams

What is Kafka Streams?

Use Case Examples

Lab: Running your first Kafka Streams Application: WordCount

Kafka Streams vs other stream processing libraries (Spark Streaming, NiFI, and Flink)

Module 2 - Kafka Streams core concepts

Kafka Streams Basic Operations

KTable

Serialization

Joins

Stateful Operations

Windowing

Time Concepts

Processor API

Testing

Error Handling

Internals

Stateful Fault Tolerance

Interactive Queries

Module 3 - Build your first Apache Kafka Streams application (WordCount) from scratch

Lab: Project setup

Lab: Kafka Streams Application Properties

Lab: Java 8 Lambda Functions - quick overview

Lab: Word Count Application Topology

Lab: Printing the Kafka Streams Topology

Lab: Kafka Streams Graceful Shutdown

Lab: Running Application from IntelliJ IDEA

Lab: Debugging Application from IntelliJ IDEA

Lab: Internal Topics for our Kafka Streams Application

Lab: Packaging the application as Fat Jar & Running the Fat Jar

Lab: Scaling our application

Module 4 - KStreams and KTables Simple Operations (Stateless)

KStream & KTables

Stateless vs Stateful Operations

MapValues / Map

Filter / FilterNot

FlatMapValues / FlatMap

Branch

SelectKey

Reading from Kafka

Writing to Kafka

Streams Marked for Re-Partition

Refresher on Log Compaction

KStream and KTables Duality

Transforming a KTable to a KStream

Transforming a KStream to a KTable

Module 5 - KStreams and KTables Advanced Operations (Stateful)

K Table groupBy

KGroupedStream / KGroupedTable Count

KGroupedStream / KGroupedTable

KGroupedStream / KGroupedTable Reduce

KStream peek

KStream Transform / TransformValues

What if I want to write to an external System?

Module 6 - Exactly Once Semantics

What's Exactly Once?

Exactly Once in Kafka 0.11

What's the problem with at least once anyway?

How to do exactly once in Kafka Streams

Lab: Achieving Exactly once semantics using kafka producer

Lab: Achieving Exactly once semantics using kafka streams

Module 7 - Joins - KStream to GlobalKTable

What are joins in Kafka Streams?

Join Constraints and GlobalKTables

The different types of joins: Inner Join, Left Join & Outer Join

Lab: Creating a join with UserEnrich Kafka Streams App