Course Outline

# Introduction to Apache Kafka – 3 Days

## Course Summary

### Description

This course will teach Apache Kafka - a popular distributed messaging system. We will cover streaming

### Objective

Upon completion of this course, the student will learn:

- Streaming technologies and architecture
- Kafka concepts and architecture
- Programming using Kafka API
- Kafka Streams API
- Kafka Connect
- KSQL
- Monitoring Kafka
- Tuning / Troubleshooting Kafka
- Best practices

### Topics

- Introduction to Streaming Systems
- Introducing Kafka
- Using Kafka APIs
- Kafka Streams API
- Monitoring and Instrumenting Kafka
- Confluent Kafka Platform
- Kafka Connect
- Kafka Best Practices
- Kafka Case Studies

### Audience

This course is designed for Developers and Architects.

### Prerequisite

- Recommended: Comfortable with Java programming language and Java development tools (Eclipse, Maven) - programming exercises are in Java
- Nice to have: Comfortable in Linux environment (be able to navigate Linux command line, run commands)

### Duration

Three Days

Course Outline

# Introduction to Apache Kafka – 3 Days

## Course Outline

### I. Introduction to Streaming Systems

A. Understanding Fast data
B. Streaming terminologies
C. Understanding at-least-once / at-most-once / exactly once processing patterns
D. Popular streaming architectures
E. Lambda architecture
F. Streaming platforms overview

### II. Introducing Kafka

A. Comparing Kafka with other queue systems (JMS / MQ)
B. Kafka Architecture
C. Kaka concepts: Messages, Topics, Partitions, Brokers, Producers, commit logs
D. Kafka & Zookeeper
E. Producing messages
F. Consuming messages
G. Consumers, Consumer Groups
H. Message retention
I. Scaling Kafka
- Labs:
  - Getting Kafka up and running
  - Using Kafka utilities

### III. Using Kafka APIs

A. Configuration parameters
B. Producer API - sending messages to Kafka
C. Consumer API - consuming messages from Kafka
D. Producer send modes
E. Message compression
F. Commits, Offsets, Seeking
G. Managing offsets - auto commit / manual commit
- Labs:
  - Writing Producer / Consumer
  - Benchmarking Producer send modes
  - Comparing compression schemes
  - Managing offsets
  - Clickstream processing

### IV. Kafka Streams API

A. Introduction to Kafka Streams library
B. Features and design
C. Streams concepts: KStream / KTable / KStore
D. Streaming operations (transformations, filters, joins, aggregations)
E. Using Streams API: foreach / filter / map / groupby
- Labs:
  - Kafka Streaming APIs

### V. Monitoring and Instrumenting Kafka

A. Monitoring Kafka metrics
B. Introduction to Metrics library
C. Instrumenting Kafka applications with the Metrics library
D. Using Grafana to visualize metrics
- Labs
  - Monitor Kafka cluster
  - Instrument Kafka applications with the metrics library

### VI. Confluent Kafka Platform

A. Introduction to Confluent platform
B. KSQL
C. KSQLdb
D. Avro Schema Registry

Course Outline

# Introduction to Apache Kafka – 3 Days

## Course Outline (cont.)

### *VII. Kafka Connect*

A. Connect ecosystem
B. Popular connectors
C. Sample configurations

### *VIII. Kafka Best Practices*

A. Avoiding common mistakes
B. Hardware selection
C. Cluster sizing
D. Partition sizing
E. Zookeeper settings
F. Compression and batching
G. Message sizing
H. Monitoring and instrumenting
I. Troubleshooting

### *IX. Kafka Case Studies*

This section will feature case studies from various companies using Kafka solve real world problems