

# Cloudera Training for Apache Kafka

## Course Content –

### Kafka Overview

- High-Level Architecture
- Common Use Cases
- Cloudera's Distribution of Apache Kafka

### Deploying Apache Kafka

- System Requirements and Dependencies
- Service Roles
- Planning Your Deployment Deploying Kafka Services
- Exercise: Preparing the Exercise Environment
- Exercise: Installing the Kafka Service with Cloudera Manager
- Exercise (optional): Create Metrics Dashboards
- Exercise (optional): Using the CM API

### Kafka Command Line Basics

- Create and Manage Topics
- Running Producers and Consumers

### Using Streams Messaging Manager (SMM)

- Streams Messaging Manager Overview
- Producers, Topics, and Consumers
- Data Explorer
- Brokers
- Topic Management
- Exercise: Managing Topics using the CLI
- Exercise: Connecting Producers and Consumers from the Command Line

## Kafka Java API Basics

- Overview of Kafka's APIs
- Topic Management from the Java API
- Exercise (optional): Managing Kafka Topics Using the Java API
- Using Producers and Consumers from the Java API
- Exercise: Developing Producers and Consumers with the Java API

## Improving Availability through Replication

- Replication
- Exercise: Observing Downtime Due to Broker Failure
- Considerations for the Replication Factor
- Exercise: Adding Replicas to Improve Availability

## Improving Application Scalability

- Partitioning
- How Messages are Partitioned
- Exercise: Observing How Partitioning Affects Performance
- Consumer Groups
- Exercise: Implementing Consumer Groups
- Consumer Rebalancing
- Exercise: Using a Key to Control Partition Assignment

## Improving Application Reliability

- Delivery Semantics
- Demonstration (optional): ISRs vs. ACKs
- Producer Delivery
- Exercise: Idempotent Producer
- Transactions
- Exercise: Transactional Producers and Consumers
- Handling Consumer Failure
- Offset Management
- Exercise: Detecting and Suppressing Duplicate Messages

- Exercise: Handling Invalid Records
- Handling Producer Failure

## Analyzing Kafka Clusters with SMM

- End-to-End Latency
- Notifiers
- Alert Policies
- Use Cases

## Monitoring Kafka

- Monitoring Overview
- Monitoring using Cloudera Manager
- Charts and Reports in CM
- Monitoring Recommendations
- Metrics for Troubleshooting
- Diagnosing Service Failure
- Exercise: Monitoring Kafka

## Managing Kafka

- Managing Kafka Topic Storage
- Demonstration (optional): Message Retention Period
- Log Cleanup and Collection
- Rebalancing Partitions
- Cruise Control
- Exercise: Installing Cruise Control
- Exercise: Troubleshooting Kafka Topics
- Unclean Leader Election
- Exercise: Unclean Leader Election
- Adding and Removing Brokers
- Exercise: Adding and Removing Brokers
- Best Practices

## Message Structure, Format, and Versioning

- Message Structure
- Schema Registry
- Defining Schemas
- Schema Evolution and Versioning
- Schema Registry Client
- Exercise: Using an Avro Schema

## Improving Application Performance

- Message Size
- Batching
- Compression
- Exercise: Observing How Compression Affects Performance

## Improving Kafka Service Performance

- Performance Tuning Strategies for the Administrator
- Cluster Sizing
- Exercise: Planning Capacity Needed for a Use Case

## Securing the Kafka Cluster

- Encryption
- Authentication
- Authorization
- Auditing