

Introduction to Apache Kafka

A 2-day hands-on workshop on Apache Kafka.

Apache Kafka is an open-source stream-processing software platform developed by the Apache Software Foundation, written in Scala and Java. The project aims to provide a unified, high-throughput, low-latency platform for handling real-time data feeds.

The course explains Kafka Producer and Consumer API and Kafka clustering in details. Explores the recommended development patterns of Kafka usage from creating a simple Kafka application to building a cluster to support High Availability with zero data loss and FIFO ordering with process only once semantics with the help of a Redis cache. The course introduces administration and monitoring of Kafka. The course includes lots of hands on lab and demos

Students participating in the workshop will be introduced and experienced the best practices of using Apache Kafka.

AUDIENCE

Developers, Team Leaders, Architects

KNOWLEDGE REQUIREMENTS

Java

LENGTH

2 Days

BONUS

Hands-on lab sessions

SYLLABUS

Course Introduction
Apache Kafka Introduction
Simple Producer and Consumer API
Serializers API
Producer Partitioner and Additional Properties
Consumer Offset Management
Consumer Groups Rebalance Listener
Introduction to Redis as Cache
Storing Offsets outside of Kafka
Transactions
Kafka Basic Cluster
Kafka Cluster Advanced Topics
Utilities and ZooKeeper
Additional Kafka API
Summary

HARDWARE AND SOFTWARE REQUIREMENTS

Computer Requirements

- RAM: minimum 6 GB of RAM required for exercises and platform to operate, 8 GB recommended.
- Disk Space: At least 40 GB of free disk space
- Internet connection
- All machines connected to same Network

Supported Operating Systems

- Linux Ubuntu VM is provided.
- Machine should be able to run VMware Player and host the Ubuntu VM

Additional Software Requirements

- PDF Reader
- Zip software

Class HW required (If delivered on customer site)

- Projector 1024*768 minimum resolution (HD preferred)
- White Board
- Erasable Markers
- Desktops or Laptops (see HW Requirements)
- 12-24 ports Switch
- Internet connectivity
- Electricity outlets for all computers/monitors and other equipment.
- At least 3 electricity outlets next to instructor location.

DAY 1 - AGENDA	
Day 1: Lesson 1: Course Introduction	Duration: 1 hour
<ul style="list-style-type: none"> • Course Introduction • lab 	
Day 1: Lesson 2: Apache Kafka Introduction	Duration: 1.5 hours
<ul style="list-style-type: none"> • Basic Concepts - MOM • What is Apache Kafka? • Getting Started • Lab 	
Day 1: Lesson 3: Simple Producer and Consumer API	Duration: 1 hour
<ul style="list-style-type: none"> • Kafka Topic Basics • Producer API • Consumer API • Lab 	
Day 1: Lesson 4: Serializers API	Duration: 1 hour
<ul style="list-style-type: none"> • Built-in Serializers • Custom Object Serializes • Common 3rd party Serializers • Lab 	
Day 1: Lesson 5: Producer Partitioner and Additional Properties	Duration: 1 hour
<ul style="list-style-type: none"> • The Default Partitioner • Custom Partitioner • Additional Producer Properties • Maintaining producer FIFO ordering and Zero Data Loss • Lab 	
Day 1: Lesson 6: Consumer Offset Management	Duration: 1.5 hour
<ul style="list-style-type: none"> • Consumer API revisited • Kafka Offsets • Lab 	
Day 1: Lesson 7: Consumer Groups Rebalance Listener	Duration: 1 hour
<ul style="list-style-type: none"> • Consumer Groups Rebalance Listener • Lab 	

DAY 2 - AGENDA	
Day 2: Lesson 8: Introduction to Redis as Cache	Duration: 2 hours Including Lab
<ul style="list-style-type: none"> • Introduction to Redis • Redis Data model and API introduction • Redis and Java – Lettuce • Lab 	
Day 2: Lesson 9: Storing Offsets outside of Kafka	Duration: 1 hour
<ul style="list-style-type: none"> • Storing offsets outside of Kafka • Storing offsets outside of example • Additional consumer considerations • Lab 	
Day 2: Lesson 10: Transactions	Duration: 1 hour
<ul style="list-style-type: none"> • Transactions and Kafka • Producer Transaction • Consumer read_committed • Read-process-write • Lab 	
Day 2: Lesson 11: Kafka Basic Cluster	Duration: 1 hour
<ul style="list-style-type: none"> • Kafka Cluster Basics • Configuring a Multi-Broker Cluster • Lab 	
Day 2: Lesson 12: Kafka Cluster Advanced Topics	Duration: 1 hour
<ul style="list-style-type: none"> • Kafka Architecture • Partition Replication • Zero Data Loss • Lab 	
Day 2: Lesson 13: Utilities and ZooKeeper	Duration: 1.0 hour
<ul style="list-style-type: none"> • ZooKeeper Configuration • Additional Useful Commands • Monitoring Kafka 	
Day 2: Lesson 14: Additional Kafka API	Duration: 0.5 hour
<ul style="list-style-type: none"> • Additional Kafka API 	
Day 2: Lesson 15: Summary	Duration: 0.5 hours Including Lab
<ul style="list-style-type: none"> • Summary 	