

Updated 05/03/2024

Sign up

## Kafka training

3 days (21 hours)

### Presentation

Our [Kafka](#) training course will enable you to master a highly scalable and durable distributed system. Its structure enables it to offer a constant level of performance despite the use of several terabits of stored messages. It's the ideal solution for high availability requirements.

In this hands-on course, you'll learn how to build an application that can publish and subscribe to data on an Apache Kafka cluster. You'll also learn how to configure a distributed microservices architecture.

You'll learn how to use Kafka to manipulate data pipelines, master its concepts, use its main components, its basic architecture and call its APIs.

In addition to Core Kafka, Kafka Connect and Kafka Streams, the course also covers the Confluent platform, such as the Schema Registry and REST Proxy.

As with all our training courses, our Kafka course will introduce you to the latest version and its new features (at the time of writing: [Kafka 3.7](#)).

### Objectives

- Understand Kafka's architecture and how it works in a multi-client context
- Understand the principles of inter-application communications
- Good message distribution practices
- Configure Kafka to integrate data from different sources and formats
- Using KSQL
- Working safely
- Publish and retrieve data to a Kafka cluster
- Use tools and APIs to interact with a Kafka cluster
- Securing data to and from a Kafka cluster

- Create your own Consumers and Producers
- Implementing a multi-thread consumer
- Using a REST proxy
- Storing Avro data in Kafka with the Schema Registry
- Data ingestion with Kafka Connect

## Target audience

- Developers
- Data Scientists
- Architects
- System administrators
- DevOps

## Prerequisites

- Basic knowledge of a Unix system
- Knowledge of a modern development language (Java, Python, Scala)

## Software prerequisites (in the case of intra-company training)

- The latest OS updates
- The latest version of [JDK](#)
- [ZooKeeper Framework](#) updated

## RECOMMENDED READING BEFORE THE COURSE

- [Kafka in a nutshell](#) is an excellent way to understand the basics of the software.
- Official apache Kafka [documentation](#)
- An article to help you understand [Kafka jargon](#)

## Kafka training program

### Introduction

- System history and complexity
- Real-time processing: a business priority
- Kafka: A flow data platform

### The fundamentals

- An overview of Kafka and efficiency
- Producers
- Brokers
- Consumers
- Using ZooKeeper

## Installation & Administration

- Installation and advanced options
- Hardware considerations
- Managing Kafka

## Architecture

- Kafka's diary files
- Replication and reliability
- Ways of writing and reading Kafka
- Partitions, Consumer Groups and Scalability

## Developing with Kafka

- Using Maven for project management
- Programming access to Kafka
- Writing a Producer in Java
- Using the REST API to write a Producer
- Writing a Consumer in Java
- Use the REST API to write a Consumer

## Advanced development with Kafka

- Creating a Consumer Multi-Thread
- Offset management
- Consumer Rebalancing
- Manual management of commits
- Data partitioning
- Message Durability: managing message durability

## New features in Kafka 3.0

- New Kafka features
- API changeover
- KRaft enhancement
- Metadata enhancement
- Kafka Connect
- MirrorMaker 2 configuration option
- Depreciation of Java 8 and Scala 2.12 in Kafka

## Schema Management

- Introduction to Avro
- Avro Schemas
- Using the Schema Registry

## Kafka Connect: Data Movement

- Kafka Connect use cases
- The basics of Kafka Connect
- Modes: Standalone and Distributed
- Distributed mode configuration
- Tracking Offsets
- Connector Configuration
- Comparison of use with advanced options

## Kafka Streams

- Kafka Streams use cases
- Kafka Streams: the fundamentals Fundamentals
- Application study and use of Kafka Streams

## Advanced Tuning & Supervision (Optional Module) - 2 days

- Log Compaction
- Configuration Settings
  - Hardware and Capacity Planning
  - Processing Oversized Messages
- Monitoring & Alerting
  - ZooKeeper and OS-Level Monitoring
  - Key Kafka Metrics
- Cluster Administration
  - Key Cluster Administration Tasks
  - Replicating clusters with MirrorMaker
- Kafka Security
  - SSL for Encryption and Authentication
  - SASL for Authentication
  - Authorization
  - Migration to Secure Cluster

## ADDITIONAL MODULE IN ENGLISH ON REQUEST (+2 DAYS) - KAFKA DEVELOPER

- Training language : English
- Course level : Beginner to intermediate

Theory: 60% Hands-on: 40% Practical: 40

Audience:? Software Developers? Data Engineers? Architects

## Getting Started

- Kafka Overview
- Kafka Basic Concepts
- Zookeeper Overview

## Developing with Kafka

- First Kafka Producer (Java/Python)
- Inside Kafka Producer
- Advanced Kafka Producer (Java/Python)
- First Kafka Consumer (Java/Python)
- Inside Kafka Consumer
- Advanced Kafka Consumer (Java/Python)

## Replications and Reliability

- Basic Replication Concepts
- Recovery and Failures
- Exactly Once Semantics (EOS)
- Controller Broker
- Inside Kafka Broker

## Kafka Topic Management

- Topic Basic Operations
- Log Management
- Designing Topics
- Kafka Security

## Kafka Connect

- Kafka Connect Concepts
- Types of Connectors
- Kafka Connect Implementation
- Standalone and Distributed Modes
- Configuring the Connectors

## Kafka Streams

- Kafka Streams Concepts
- First Kafka Stream (Java/Python)
- Stateless KStreams
- Stateful KStreams

## Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

## Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

## Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

## Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

## Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

## Sanction

A certificate will be issued to each trainee who completes the course.