

Confluent Developer Training: Building Kafka Solutions

Course Objectives

In this three-day hands-on course you will learn how to build an application that can publish data to, and subscribe to data from, an Apache Kafka™ cluster. You will learn the role of Kafka in the modern data distribution pipeline, discuss core Kafka architectural concepts and components, and review the Kafka developer APIs. As well as core Kafka, Kafka Connect, and Kafka Streams, the course also covers other components in the broader Confluent Platform, such as the Schema Registry and the REST Proxy.

Hands-On Training

Throughout the course, hands-on exercises reinforce the topics being discussed. Exercises include:

- » Using Kafka's command-line tools
- » Writing Consumers and Producers
- » Writing a multi-threaded Consumer
- » Using the REST Proxy
- » Storing Avro data in Kafka with the Schema Registry
- » Ingesting data with Kafka Connect

Who Should Attend?

This course is designed for application developers, ETL (extract, transform, and load) developers, and data scientists who need to interact with Kafka clusters as a source of, or destination for, data.

Course Duration

This is a three-day training course.

Course Prerequisites

Attendees should be familiar with developing in Java (preferred) or Python. No prior knowledge of Kafka is required.

Course Contents

The Motivation for Apache Kafka

- Systems Complexity
- Real-Time Processing is Becoming Prevalent
- Kafka: A Stream Data Platform

Kafka Fundamentals

- An Overview of Kafka
- Kafka Producers
- Kafka Brokers
- Kafka Consumers
- Kafka's Use of ZooKeeper
- Kafka Efficiency

Kafka's Architecture

- Kafka's Log Files
- Replicas for Reliability
- Kafka's Write Path
- Kafka's Read Path
- Partitions and Consumer Groups for Scalability

Developing With Kafka

- Using Maven for Project Management
- Programmatically Accessing Kafka
- Writing a Producer in Java
- Using the REST API to Write a Producer
- Writing a Consumer in Java
- Using the REST API to Write a Consumer

More Advanced Kafka Development

- Creating a Multi-Threaded Consumer
- Specifying Offsets
- Consumer Rebalancing
- Manually Committing Offsets
- Partitioning Data
- Message Durability

Schema Management in Kafka

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

Kafka Connect for Data Movement

- The Motivation for Kafka Connect
- Kafka Connect Basics
- Modes of Working: Standalone and Distributed
- Configuring Distributed Mode
- Tracking Offsets
- Connector Configuration
- Comparing Kafka Connect with Other Options

Basic Kafka Installation and Administration

- Kafka Installation
- Hardware Considerations
- Administering Kafka

Kafka Streams

- The Motivation for Kafka Streams
- Kafka Streams Fundamentals
- Investigating a Kafka Streams Application

Confluent offers public training courses and private, onsite events. Please visit <http://confluent.io/training> for the public course schedule. For inquiries about onsite events, email training-admin@confluent.io