Szkolenie: Cloudera

Cloudera Training for Apache Kafka

DOSTĘPNE TERMINY

2026-11-30 | 4 dni | Virtual Classroom

Cel szkolenia:

This four-day instructor-led course begins by introducing Apache Kafka, explaining its key concepts and architecture, and discussing several common use cases. Building on this foundation, you will learn how to plan a Kafka deployment, and then gain hands-on experience by installing and configuring your own cloud-based, multi-node cluster running Kafka on the Cloudera Data Platform (CDP).

You will then use this cluster during more than 20 hands-on exercises that follow, covering a range of essential skills, starting with how to create Kafka topics, producers, and consumers, then continuing through progressively more challenging aspects of Kafka operations and development, such as those related to scalability, reliability, and performance problems. Throughout the course, you will learn and use Cloudera’s recommended tools for working with Kafka, including Cloudera Manager, Schema Registry, Streams Messaging Manager, and Cruise Control.

What Skills You Will Gain

During this course, you learn how to:

- Plan, deploy, and operate Kafka clusters
- Create and manage topics
- Develop producers and consumers
- Use replication to improve fault tolerance
- Use partitioning to improve scalability
- Troubleshoot common problems and performance issues

Who Should Take this Course?

This course is designed for system administrators, data engineers, and developers. All students are expected to have basic Linux experience, and basic proficiency with the Java programming language is recommended. No prior experience with Apache Kafka is necessary.

Plan szkolenia:

- Kafka Overview

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

- 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

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

- 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

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

- 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

Wymagania:

This course is designed for system administrators, data engineers, and developers. All students are expected to have basic Linux experience, and basic proficiency with the Java programming language is recommended. No prior experience with Apache Kafka is necessary.

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

Poziom trudności



Certyfikaty:

The participants will obtain certificates signed by Cloudera (course completion).

Upon completion of the course, attendees are encouraged to continue their study and register for the CDP Data Developer exam
https://www.cloudera.com/about/training/certification/cdp-datadev-exam-cdp-3001.html

Certification is a great differentiator. It helps establish you as a leader in the field, providing employers and customers with tangible evidence of your skills and expertise.

Prowadzący:

Certified Cloudera Instructor

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001