

Exercise 1: Kafka – Getting Started

Deadline: 05.03.2024; 23:59 CET

Exercise Abstract

In the first week's lecture on Event-driven and Process-oriented Architectures you will learn about *Apache Kafka*¹ as one of the main technologies that we will be using throughout the course.

① Labs: Getting Started with Kafka

In the practical part of this week's session we ask you to work on 3 **Labs** to get started with Kafka:

- Lab01 – Kafka Getting Started²: This lab helps you to get set up with Kafka in Docker on your machine and run included scripts for testing Kafka's messaging functionality.
- Lab02Part1 – Simple Kafka Producer and Consumer³: This lab provides you with plain Java code for implementing a simple Kafka *Producer* and *Consumer* of messages that connect to the Kafka Docker instance from Lab01.
- Lab02Part2 – Kafka Producers and Consumers in an Eye-tracking System⁴: This lab provides you with Java code implementing different Kafka *Producers* and *Consumers* in the context of an eye-tracking system. The implementations of the producers and the consumers demonstrate many of the concepts presented in the lecture.

② Suggestions for Software Project: Experiments with Kafka

Concerning the **Software Application** we suggest you to experiment with Kafka and report on the results. Here you are free to choose which aspects to investigate. Some suggestions are given in the lecture, e. g., to investigate:

- the impact of load and batch size on processing latency,
- the risk of data loss due to dropped messages,
- the outage of Zookeeper,
- the risk of data loss due to consumer lag,
- the risk of data loss due to offset misconfigurations.

Additionally, you can for example investigate the behavior of Kafka, producers and consumers with multiple brokers, consumers and consumer groups, topics and partitions, as well as acknowledgment configurations. Feel free to use the provided data sets for your experiments.

Software Project Expectations

The labs in Part 1 of this sheet are meant to be done during class and are **not graded**. In the **Project** part we expect you to prepare a report on your experiments with Kafka with your findings for the software application. Ensure that you describe your experiments at a level of detail that the experiments are reproducible. Please work in groups of 2 to 3 students.

¹<https://kafka.apache.org/>

²<https://github.com/scs-edpo/lab01-kafka-getting-started>

³<https://github.com/scs-edpo/lab02Part1-kafka-producer-consumer>

⁴<https://github.com/scs-edpo/lab02Part2-kafka-EyeTracking>

Hand-in Instructions

The report for the project-related Part 2 will **be graded** as part of the software application for your group to be handed in for the first part of the lecture until after the semester break. An intermediate hand-in of the results from Exercise 1 is expected together with the results from Exercise 2 via Canvas by the deadline indicated on the top of this sheet. Each group member **must explicitly indicate** which part she/he/they has/have been working on. Please approach the tutors for individual feedback regarding your submission for this assignment.