

## Программа курса

### «KAFKA: Администрирование кластера Kafka»

Версия 06.08.2025

**О курсе:** 3х дневный практический курс администрирования Apache Kafka создан для администраторов, дата-инженеров и специалистов, которые стремятся научиться строить высоконагруженные, отказоустойчивые конвейеры потоковых данных (data pipelines), обрабатывать миллионы сообщений в секунду и управлять сложными распределенными системами.

#### Аудитория:

- Админы и DevOps
- Дата инженеры
- Дата архитекторы
- Минимальные требования

#### Уровень подготовки:

- Уверенное знание команд Linux (опыт работы с командной строкой, файловой системой, POSIX, systemd, текстовыми редакторами vi, nano)
- Базовые знания настройки компонент безопасности (Kerberos, PKI, LDAP)

**Продолжительность курса:** 24 академических часа **+ 12 часов дополнительного обучения (видео материалы и лабораторные)**, 3 дня по 8 ак. часов дистанционно

## Содержание программы

### 1. Введение в Apache Kafka

- Что такое Apache Kafka
- Сценарии использования Kafka
- Компоненты Kafka и Kafka экосистемы
- **Архитектура Kafka 4**
- Эволюция Kafka: от Zookeeper к KRaft
- Новая архитектура без Zookeeper (KIP-500, KIP-866)
- Что нового в Apache Kafka 4.0: ключевые функции, изменения и совместимость
- Кластер **Zookeeper (опционально)**
  - **Лабораторная работа 1:** Установка single-node кластера Apache Kafka.

### 2. Основные концепции Kafka и компоненты

- Как работает Kafka: Топики (Topics) и партиций (Partitions)
- Репликация данных. Брокеры (Brokers): Развертывание. Конфигурация Multibroker
- Поставщики данных (Producers): Типы партиций. Синхронные и асинхронные Producers.
- Потребители данных (Consumers): Очередь Consumers.
- Архитектура с KRaft Controller'ами
- Новый механизм управления метаданными
- Логическая структура брокеров и quorum.

### 3. Установка кластера Apache Kafka

- Установка распределенного кластера Apache Kafka 4.0 в KRaft-режиме
- Минимальные требования
- Установка Kafka без Zookeeper
- Настройка KRaft controllers и quorum voters
- Примеры конфигурации kraft-controller mode

- Sizing и планирование развертывания кластера
- Установка Kafka с Zookeeper (обзорно)
- Процедуры миграции с Apache Kafka 3.x на 4.0
  - **Лабораторная работа 2:** Установка распределенного 3х узлового кластера Apache Kafka 4.0 в Kraft режиме. (дополнительно с Zookeeper).

#### 4. Базовые операции в Apache Kafka

- Операции с Топиками (Topic).
- Публикация данных с консольного Producer
- Получение данных с консольного Consumer
- Пользовательские интерфейсы для работы с Kafka: **kafka-manager, Conductor, RedPanda GUI**, etc
  - **Лабораторная работа 3:** Графические консоли GUI для управления и работы с Apache Kafka (**follow режим**).

#### 5. Оптимизация кластера Apache Kafka

- Оптимизация производительности компонент кластера
- Масштабирование кластера
- Балансировка разделов(partition) и использование Partition Reassignment Tools
- **Best Practices Kafka** — Мониторинг кластера Kafka
- Настройка мониторинга кластера Kafka с использованием **Prometheus + Grafana**
- Настройка **алертинга кластера Kafka**
  - **Лабораторная работа 4:** Балансировка партиций и тестирование отказоустойчивости кластера Apache Kafka.
  - **Лабораторная работа 5:** Мониторинг производительности кластера **Apache Kafka** с использованием **Prometheus, Grafana, JMX exporter, KMinion**.
  - **Лабораторная работа 6:** Настройка алертинга для кластера **Apache Kafka** с использованием **Prometheus, Grafana** и **Telegram(email)**.

#### 6. Интеграция с Apache Kafka (обзорно)

- Использование Apache NiFi
- Интеграция **Apache Hadoop** с Kafka
- Сценарии интеграции с Apache Kafka (ClickHouse, Greenplum)



**\*Дополнительные модули по интеграции для самостоятельного изучения (видеоматериалы)**

#### 7. Расширенная конфигурация кластера Kafka

- Расширенная конфигурация топиков (Topic): Изменение конфигурации топиков.
- Sizing (количество партиций и фактор репликации). Сегменты и индексы. Политика очистки логов (Log Cleanup Policy). Сжатие и консолидация данных в логах.
- **REST Proxy. Schema Registry (обзорно)**
- Репликация кластера с использованием **Mirror Maker 2**
- Новые подходы к rolling-обновлениям в кластере без Zookeeper
  - **Лабораторная работа 7:** Настройка **MirrorMaker 2.0** для катастрофоустойчивости (DR) для кластера Apache Kafka. (**follow режим**).

#### 8. Безопасность в Apache Kafka 3.0/4.0

- Шифрование SSL: Установка центра сертификатов (CA), настройка SSL/TLS для Kafka кластера и клиентов
- Аутентификация **SSL/SASL**
- Аутентификация **OAuth**
- Аутентификация **Kerberos**: Настройка Kerberos. Keytabs и principals. Настройка кластера Kafka и клиентов для поддержки Kerberos
- Авторизация в Kafka
- Управление списками управления доступом (ACL)
- Настройка ACL в KRaft-режиме и KIP-866

 ШКОЛА БОЛЬШИХ ДАННЫХ	ООО «Учебный центр «Коммерсант» ©	
	<p align="center"> <b>«Школа Больших Данных»</b>  <a href="http://www.bigdataschool.ru">www.bigdataschool.ru</a>  <b>2026</b> </p>	

- **Лабораторная работа 8:** Настройка шифрования TLS для кластера Apache Kafka.
- **Лабораторная работа 9:** Настройка двусторонней аутентификации с сертификатами TLS для кластера Apache Kafka.
- **Лабораторная работа 10:** Настройка аутентификации Kerberos (MIT) для кластера Apache Kafka.
- **Лабораторная работа 11:** Настройка авторизации и управления доступом на базе SASL/SSL для кластера Apache Kafka. (\*\* дополнительные материалы).
- **Лабораторная работа 12:** Настройка авторизации с использованием SASL/OAUTHBEARER на базе Keycloak для кластера Apache Kafka. (\*\* дополнительные материалы).

### **Бонус: +12 часов материалов для глубокого погружения**

Для подготовленных администраторов кластера Apache Kafka и дата инженеров которые выполняют задачи обслуживания и администрирования кластеров Kafka мы записали дополнительные модули, которые могут быть полезны для освоения сверх плановых модулей.

После просмотра видеоуроков, вы сможете задать вопросы преподавателю и выполнить дополнительные (\*\* дополнительные) лабораторные работы для закрепления материала на практике.

- **5 часов дополнительных лекций** по продвинутым темам администрирования, обслуживания и интеграции.
- **7 часов дополнительных** и для оттачивания навыков настройки и сопровождения распределенных кластеров по управлению Apache Kafka.