
 ШКОЛА БОЛЬШИХ ДАННЫХ	ООО «Учебный центр «Коммерсант» © «Школа Больших Данных» www.bigdataschool.ru 2026	
---	--	--

Программа курса «ВОКА: Основы Apache Kafka»

О курсе: Архитектура и принципы работы самой популярной платформы потоковой передачи событий. Узнайте, как использовать Apache Kafka для асинхронной интеграции приложений в реальном времени за 16 контактных часов с преподавателем и 1 дополнительный день самостоятельной работы на кластере.

Аудитория:

- Администраторы кластера;
- Дата-инженеры;
- Архитекторы ИС;
- Разработчики и аналитики данных;
- Любые ИТ-специалисты, которые хотят освоить Apache Kafka.

Уровень подготовки:

- Уверенное знание базовых команд Linux: опыт работы с командной строкой, файловой системой, POSIX
- Владение текстовыми редакторами vi, nano

Продолжительность курса: 16 академических часов, 2 дня по 8 ак. часов дистанционно

Содержание программы

1. Введение в Kafka

- Что такое Apache Kafka?
- Сценарии использования Kafka
- Компоненты Kafka и Kafka экосистемы
- Архитектура Kafka

2. Основные концепции Kafka

- 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): Очереди Consumers.

3. Установка кластера Apache Kafka

- Требования по установке кластера Kafka
- Установка и настройка кластера Zookeeper: Настройка кворума. Конфигурация сервисов и файловой системы. Оптимизация и отказоустойчивость Zookeeper.
- Установка кластера Kafka без использования Zookeeper KIP-500
- Установка и настройка кластера Kafka: Планирование установки и sizing кластера Kafka.
- Установка кластера Kafka на Yandex Cloud
- Использование и настройка Yandex Managed service для Apache Kafka

4. Базовые операции в Apache Kafka

- Операции с Топиками (Topic): Изменение конфигурации топиков, масштабирование топиков
- Сегменты и индексы. Политика очистки логов (Log Cleanup Policy). Сжатие и консолидация данных в логах
- Публикация данных с консольного Producer
- Получение данных с консольного Consumer

5. Администрирование кластера Apache Kafka

- Утилиты командной строки для администрирования Kafka
- Управление группами потребителей (Consumer Group).
- Отказоустойчивость кластера Kafka
- Балансировка разделов(partition) и использование Partition Reassignment Tools
- Графические пользовательские интерфейсы GUI для работы с Kafka

6. Мониторинг производительности кластера Apache Kafka

- Оптимизация производительности компонент кластера
- Масштабирование кластера
- Best Practices Kafka/Confluent
- Мониторинг кластера Kafka
- Настройка мониторинга кластера Kafka с использованием Prometheus + Grafana

7. Интеграция с Apache Kafka

- Интеграция Kafka с внешними системами
- Утилита Kafka Connect: Архитектура Kafka Connect. Примеры настройки коннекторов
- Сценарии интеграции с Apache Kafka

8. Расширенная конфигурация кластера Kafka

- Kafka REST Proxy: отправка и получение сообщений в форматах JSON, бинарные и AVRO-файлы
- Организация доступа к Kafka через REST Proxy
- Schema Registry
- Что такое Kafka Schema registry
- Как работает Schema Registry

9. Основы Mirror Maker 2.0

- Зеркалирование кластера Apache Kafka и сценарии использования
- Утилита Mirror Maker 2.0

10. Основы безопасности кластера Apache Kafka

- Архитектура безопасности Apache Kafka
- Аутентификация и авторизация для кластера Apache Kafka