

[Презентация к уроку 9.3.1](#)

Текстовая расшифровка видео:

БРОКЕР СООБЩЕНИЙ АРАСНЕ КАФКА

План:

- Apache Kafka vs RabbitMQ;
- Apache Kafka vs ActiveMQ;
- Что такое Apache Kafka;
- Использование Kafka;
- Kafka в компаниях.

Apache Kafka vs RabbitMQ

Разница между Apache Kafka и RabbitMQ заключается в том, что RabbitMQ больше специализируется на маршрутизации сообщений, а Apache Kafka на том, чтобы их хранить, и хранить в том порядке, в каком они пришли.

RabbitMQ гарантирует это, а Apache Kafka не убеждает в том, что консьюмеры забрали сообщения. Сдвиг офсетов и работа над получением сообщений лежит полностью на «плечах» консьюмера. RabbitMQ же занимается маршрутизацией, и консьюмеры не нагружаются. В тоже время, например, в кейсе, когда очень быстро обрабатываются одни процедуры и медленно – другие, RabbitMQ не будет нигде хранить данные. Он будет отправлять эти сообщения. Таким образом, если они не будут успевать обрабатываться, сообщения будут теряться. В случае с Apache Kafka такого не происходит.

Apache Kafka vs ActiveMQ

Здесь отличие Apache Kafka от очередей в следующем:

- Если не настроить хранение сообщений, то Kafka, если сообщение не будет доставлено, не будет производить по умолчанию повторных попыток, и сообщения могут теряться.
- Очереди ActiveMQ, RabbitMQ гарантируют доставку один раз. Они необходимы в тех случаях, когда гарантия доставки критична.

Еще одно отличие:

- При считывании сообщений из очереди, брокер не удаляет его. Сообщение остается там храниться.
- Сообщения из очереди Kafka могут считываться разными консьюмерами сколько угодно раз.

Что такое Apache Kafka

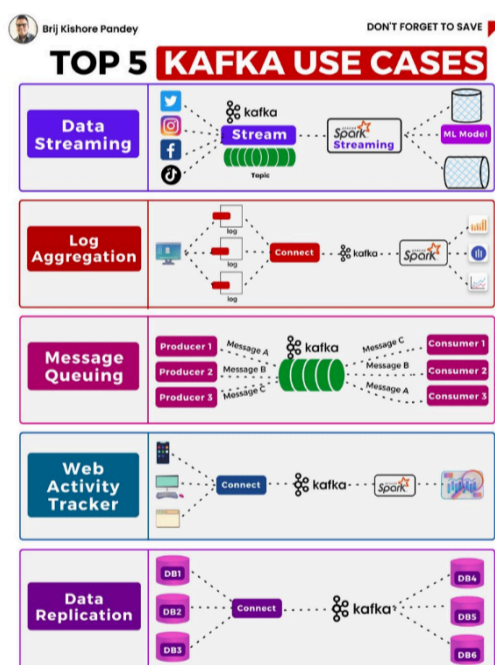
Apache Kafka – это платформа для потоковой передачи событий.

Apache Kafka спроектирована так, чтобы передавать достаточно большие объемы данных с минимальными задержками. Журнал записей Kafka позволяет записывать данные в историческом порядке, как они пришли, а также упорядочивать их по разделам, что увеличивает пропускную способность на чтение. Также Apache Kafka позволяет настраивать долговременное хранилище, что дает возможность читать исторические данные и real-time одновременно из одного места.

Apache Kafka очень удобен для тех систем, где нужно иметь высокую пропускную способность, обрабатывать очень большие данные и настраивать долговременное хранилище (в некоторых случаях).

Использование Kafka

Где можно использовать:



- В потоковой обработке данных;
- В сборе логов и их агрегации;
- В отправке сообщений;
- Для отслеживания активности в интернете;
- Для репликации данных;
- Для разгрузки микросервисов.

Kafka в компаниях

- В **LinkedIn** используется в рекомендательной системе, когда рекомендует новые контакты.
- В **Netflix** тоже в рекомендательной системе, когда предлагает посмотреть какие-либо сериалы.