

Очистка топика

Итак, мы успешно записали и прочитали сообщения, давайте посмотрим на механизм ретеншена.

```
./bin/kafka-topics.sh --describe --topic registrations --bootstrap-server localhost:9092
```

По умолчанию Кафка проверяет, нужно ли удалить данные по ретеншену каждые 5 минут. Давайте сделаем этот интервал меньше — каждую секунду.

Остановим брокер Кафки и открываем файл на редактирование:

```
vi config/server.properties
```

И выставляем там нужное значение:

```
log.retention.check.interval.ms=1000
```

После этого запускаем снова брокер.

Давайте теперь скажем Кафке, что мы хотим удалять сообщения после одной минуты:

```
./bin/kafka-configs.sh --bootstrap-server localhost:9092 --entity-type topics --entity-name registrations --alter --add-config retention.ms=60000
```

Запустив заново консьюмера мы увидим, что сообщения действительно удалились. В логе видим *“Found deletable segments with base offsets”*.

Давайте поподробнее разберем этот момент. Повторим эксперимент, слегка изменив настройки.

Скажем Кафке удалять данные после 10 секунд:

```
./bin/kafka-configs.sh --bootstrap-server localhost:9092 --entity-type topics --entity-name registrations --alter --add-config retention.ms=10000
```

А теперь запустим продюсера писать сообщения в цикле.

В одном терминале запустим вот такую конструкцию:

```
touch /tmp/data && tail -f -n0 /tmp/data | ./bin/kafka-console-producer.sh  
--topic registrations --bootstrap-server=localhost:9092 --sync
```

А во втором терминале - вот такую:

```
for i in $(seq 1 3600); do echo "test${i}" >> /tmp/data; sleep 1; done
```

Читая сообщения спустя минуту, мы по-прежнему видим старые сообщения!

```
./bin/kafka-console-consumer.sh --topic registrations --bootstrap-server  
localhost:9092 --consumer-property auto.offset.reset=earliest
```

Для того чтобы понять, что происходит, мы должны разобраться во внутренней структуре данных партиции.