

Как сделать кластер отказоустойчивым:

- Отказоустойчивый кластер содержит минимум 3 зукипера;
- Отказоустойчивый кластер содержит минимум 2 брокера;
- Кластер следует размещать на разных серверах;
- Сервера должны размещаться также на разных стойках.

Controller:

Kafka является распределенной системой, и контроллер решает задачу по координации этой системы в случае различных изменений в ее состоянии.

Контроллером может стать абсолютно любой брокер в кластере. Контроллер отвечает за:

- За создание и удаление топиков;
- За добавление партиций и назначение им лидеров;
- За разрешение ситуаций с падениями брокеров или выходом их из кластера.

В обязанности контроллера также входит:

- Создание новых партиций и топиков;
- Удаление топиков.

Конфигурации брокеров, помогающие сделать кластер более отказоустойчивым:

- `auto.leader.rebalance.enable`
- `leader.imbalance.check.interval.seconds`
- `leader.imbalance.per.broker.percentage`

Данная группа настроек помогает держать нагрузку на брокеры кластера равномерной (даже после рестартов и падений). Эти настройки включены по умолчанию.

- `flush.messages`
- `flush.ms`

Данная группа настроек касается работы Kafka с подключениями.

- max.connections
- max.connections.per.ip
- max.connection.creation.rate

Данная настройка относится, к количеству партиций лога по умолчанию (для каждого топика).

- Num.partitions

Настройка числа партиций - не больше 4к партиций на брокер и не больше 200к партиций на кластер. Если партиций становится слишком много, то после жесткого падения ваш кластер может восстанавливаться продолжительное время (например, десятки минут). Это делает администрирование кластера более трудоемкой задачей.

Дополнительные материалы:

<https://www.confluent.io/blog/apache-kafka-supports-200k-partitions-per-cluster/>

Бэкапы:

- Бэкапы Zookeeper являются обязательными!
- Важно валидировать бэкапы в Zookeeper для проверки их работы (это легко и просто, поскольку бэкапы в Zookeeper легковесные).
- В случае Кафки в большинстве кейсов достаточно иметь репликацию данных (а данные уже не являются легковесными).
- Помните, что все кейсы разные, и вам нужно самостоятельно продумывать является ли оправданным или нет использование бэкапов в Kafka.

Практики, используемые для еще большего повышения отказоустойчивости:

- Disaster recovery plan
- Disaster тесты
- Runbooks и on-call дежурства
- Knowledge sharing и bus-factor 2+
- Разбор инцидентов

Ограничения при использовании одного Дата-центра:

- Потеря дата-центра = stop the world;
- Физически ограниченное место для горизонтального роста;
- Не всегда можно соблюсти требования юристов.

Важно помнить, что отвечать на вопрос нужен ли вашей компании второй/третий Дата-центр должен бизнес.