??????????? ????????? ??????????

> Бесплатный курс обучения по ClickHouse от Яндекс Практикум - https://yandex.cloud/ru/training/clickhouse

Кластер - логическая группа машин, обладающих всеми накопленными нормализованными событиями KUMA. Подразумевает наличие одного или нескольких логических шардов.

Shard (шард) - логическая группа машин, обладающих некоторой частью всех накопленных в кластере нормализованных событий. Подразумевает наличие одной или нескольких реплик. Если обьяснить на пальцах, механика работы кластера с несколькими шардами - это работа дисков в RAID0.

Увеличение количества шардов позволяет:

- накапливать больше событий за счет увеличения общего количества серверов и дискового пространства;
- поглощать больший поток событий за счет распределения нагрузки, связанной со вставкой новых событий;
- уменьшить время поиска событий за счет распределения поисковых зон между несколькими машинами.

> В случае выхода из строя машины с определенным шардом (при отсутствии репликации) данные продолжат накапливаться другими шардами, но в интерефейсе KUMA нельзя будет искать по событиям до тех пор, пока не ввести в строй обратно "упавшую" машину

Replica (реплика) - машина, являющаяся членом логического шарда и обладающая одной копией данных этого шарда. Если реплик несколько - копий тоже несколько (репликация данных). Если обьяснить на пальцах, механика работы кластера с несколькими репликами - это работа дисков в RAID1.

Увеличение количества реплик позволяет:

- улучшить отказоустойчиовость;
- распределить общую нагрузку, связанную с поиском данных, между несколькими машинами (однако для этой цели лучше увеличить количество шардов).

Keeper - машина участвующая в репликации и координации (распределенные Data Definition Language (DDL) запросы) данных на уровне всего кластера хранилищ, позволяющее иметь линеаризуемую запись и нелинейное чтение данных. На весь кластер минимально требуется хотя бы 1 реплика с данной ролью. Рекомендуемое и отказоустойчивое количество таких реплик - 3. Число реплик, участвующих в координации репликации, должно быть нечетным. Если Keeper перестанет работать, то реплики переходят в Read only, при этом поиск будет работать, но новые события записываться не будут.

> При разворачивании распределенного кластера, машины с ролью Keeper должны быть запущены **ДО** запуска / установки машин без роли Keeper

При работе keeper машины испольют RAFT алгоритм для определения лидера среди нескольких keeper машин. Лидер выполняет все операции записи и запускает автоматическое восстановление при отказе любого из подключенных серверов. Остальные узлы — подписчики или последователи, реплицируют данные с лидера и используются клиентскими приложениями для чтения. Подробнее можно почитать тут. Кластер может оставаться работоспособным при отказе определенного количества нод в зависимости от размера кластера. Например, для кластера из 3 нод (Keeper), алгоритм кворума продолжает работать при отказе не более чем одной ноды.

> Доп. информация: https://clickhouse.com/docs/en/architecture/horizontal-scaling#architecture-diagram

???????? ????????????? ?????????? ? ????????? ansible

```
storage:
  hosts:
    kuma-maybe-collector.example.com:
      ip: 1.1.1.1
      keeper: 1
    kuma-storage-1.example.com:
      ip: 0.0.0.0
      shard: 1
      replica: 1
      keeper: 2
    kuma-storage-2.example.com:
      ip: 0.0.0.0
      shard: 1
      replica: 2
      keeper: 3
```

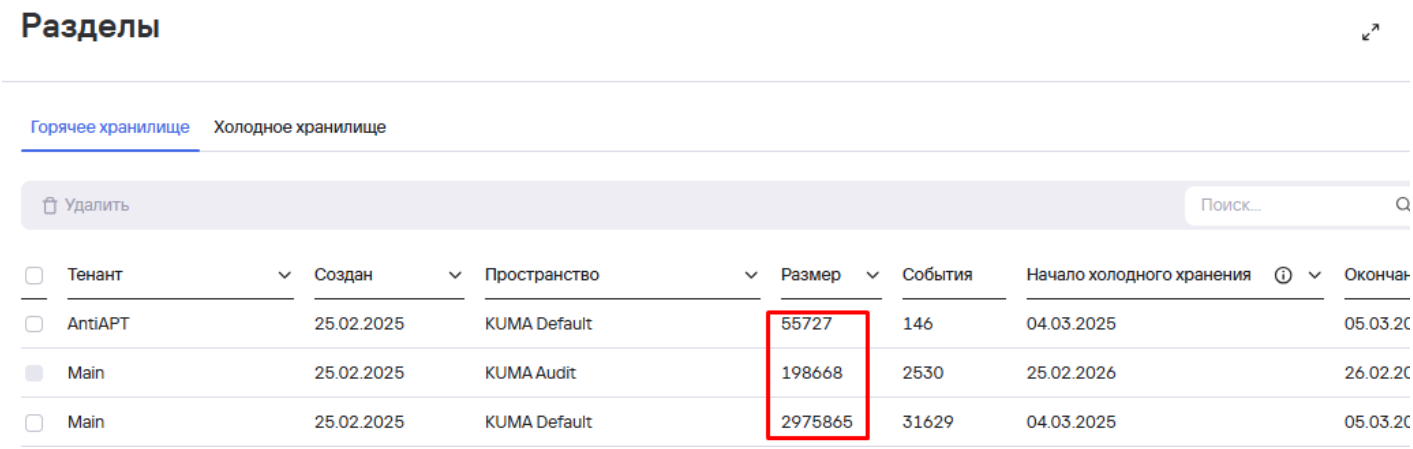
Если указаны параметры shard и replica, то машина является частью кластера и принимает участие в накоплении и поиске нормализованных событий KUMA. Если дополнительно указан параметр keeper, то машина также принимает участие в координации репликации данных на уровне всего кластера. Если keeper на хранилище не нужен, то указывается `keeper: 0`

Если указан только параметр keeper (отдельная машина с этой ролью), то машина не будет накапливать нормализованные события, но будет участвовать в координации репликации данных на уровне всего кластера. Значения параметра keeper должны быть уникальными, а значения shard и replica равны 0.

"Номера" реплик, шардов и киперов роли не играют. Например, конфиг с номерами реплик 1-2-3 работает так же хорошо как и с номерами 1-23-777, по смыслу это скорее айдишник, а не порядковый номер.

Если в рамках одного шарда определено несколько реплик, то значение параметра replica должно быть уникальным в рамках этого шарда. Пример для четырех серверов хранилища и трех отдельных киперов на рисунке ниже:

При просмотре разделов из веб интерфейса KUMA (Активные сервисы - ПКМ - Смотреть разделы) показывается сколько места занято партицией во всем кластере с учетом всех реплик.



> В точке назначения в KUMA прописываются URL всех хранилищ, отдельно установленные машины с ролью keeper указывать НЕ нужно

>

Буфер хранилища по умолчанию 128 Мб, при большом количестве шардов и заметной медленной работе поиска - увеличьте размер буфера, например, до 512 Мб и увеличте таймаут (интервал очистки буфера), например, до 60 сек.

Редкие большие вставки для БД ClickHouse лучше, чем частые и небольшие.

Установка службы хранилища - https://kb.kuma-community.ru/books/ustanovka-i-obnovlenie/page/ustanovka-sluzby-xranilishha-esli-etogo-ne-proizoslo-pri-ustanovke

Смотри также статью про схему взаимодействия KUMA - https://kb.kuma-community.ru/books/kuma-how-to/page/sxema-setevogo-vzaimodeistviia-kuma

---

Revision #21
Created 18 April 2024 09:58:24 by Boris Rzr
Updated 1 September 2025 07:29:12 by Boris Rzr