???? ???????? ?????? ? KUMA

В KUMA существует три типа пространства для хранения событий:

- Горячее
- Холодное
- Архивное

Для оптимизации использования дискового пространства и ускорения выполнения запросов в KUMA введено несколько уровней устройств хранения:

- **Горячее (hot)** - оперативное хранение, обычно состоит из быстродействующих устройств с ограниченным объемом пространства [Диски, например: NVMe или SSD]. Поиск по событиям доступен из веб-интерфейса KUMA.
- **Холодное (cold)** - медленные устройства, но большого объема [Диски, например: HDD SAS или HDD SATA]. Поиск по событиям доступен из веб-интерфейса KUMA.

Основная идея разделения хранилищ на "горячие" и "холодные" состоит в том, что доступ к данным сохраняется, но при этом увеличиваются задержки. Используется сочетание настроек политики хранения ClickHouse и механизма переноса разделов таблиц между дисками. Плюсом подхода является возможность использовать в качестве хранилища любое примонтированное в качестве каталога Linux устройство, хранилища **HDFS** (используется функционал хранения, поиск делается с ClickHouse), а также **S3**.

> Планируется прекращение поддержки **HDFS** и рекомендуется запланировать перенос данных в **S3**.

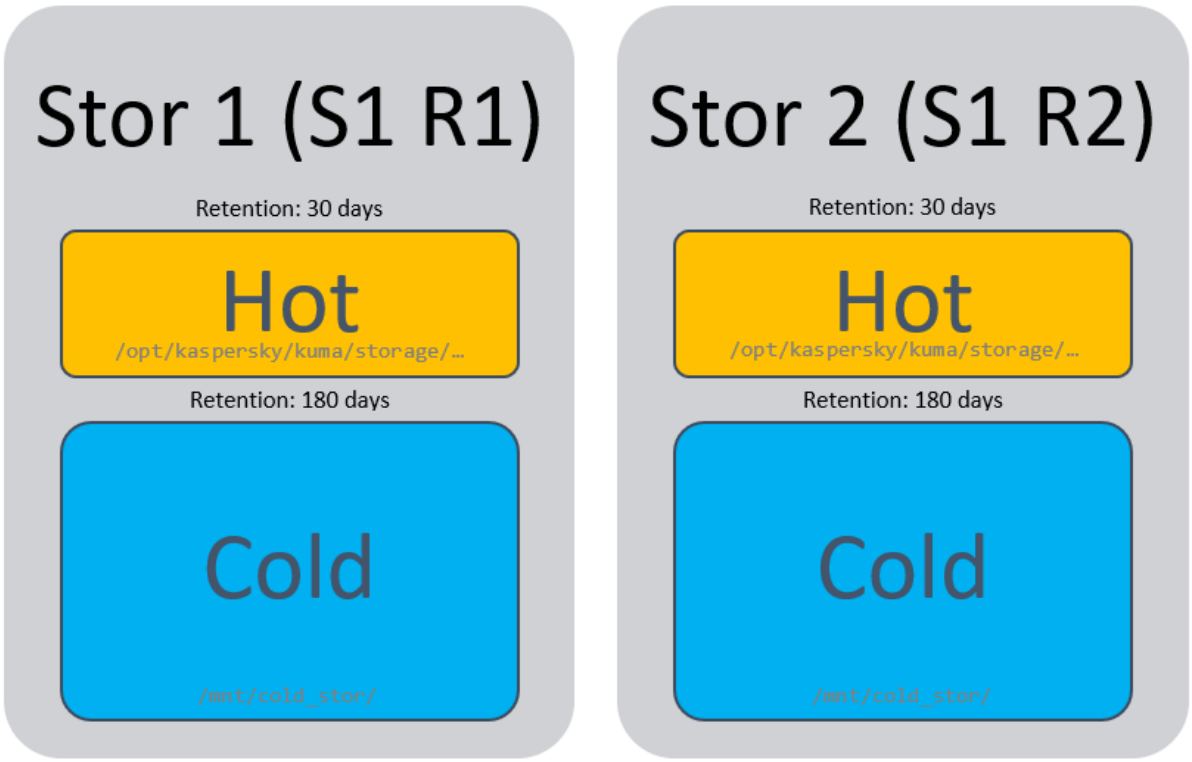
> Подробнее про холодное хранение - https://support.kaspersky.ru/kuma/4.2/221257?page=help

Для холодного по объему пространства нужно столько же, сколько и для горячего (нет дополнительной компресии), сами диски можно использовать менее производительней и подешвле.

С версии 3.4 в связи с расширением функционала, изменился подход к определению срока хранения событий при использовании холодного хранения: Общий срок хранения событий определяется параметром TTL события - отсчет начинается от момента попадания события в хранилище. Значение TTL указывает, сколько времени событие будет храниться в KUMA. Время нахождения события в горячем хранилище определяется Вариантами условий хранения и может быть определено в днях, гигабайтах или процентах дискового пространства. Для горячего хранения можно применить до двух политик. Данные будет

находиться в горячем хранилище до срабатывания одной из политик, после чего раздел с самыми старыми данными будет перемещен в холодное хранилище и будет находиться в холодном хранилище до истечения TTL. Общее время = горячее + холодное.

Холодное пространство монтируется на сами хранилища (в случае локального типа холодного хранения) в нужном объеме, на рисунке ниже схематично показано, как это выглядит для двух реплик (R) и одного шарда (S):



Владельцем папки по точке монтирования холодного хранения должна быть УЗ kuma, команда: `chown -R kuma:kuma /mnt/cold_stor/`

Если добавить второй локальный диск, то данные между ними будут распределяться по алгоритму round-robin до тех пор, пока кусок данных, который мы хотим вставить не окажется больше свободного зарезервированного места на диске. Когда это случится все записи будут падать по round-robin в следующие диски. Если дисков всего 2, то соответственно в тот, где есть место

Монтирование диска в fstab с правами для kuma

Узнайте UID и GID пользователя kuma

```
id -u kuma
id -g kuma
```

Добавить запись в конец файла /ets/fstab, пример:

```
/dev/sdb1 /mnt/cold_stor/ auto defaults,uid=988,gid=984,umask=770 0 0
```

- **Архивное хранение** — (отщелкивание индексов ClickHouse) по архивным данным поиск не возможен, только если вручную, либо автоматизировано разархивировать и аттачить партиции. [Диски, например: Лента, HDD SATA, USB FLASH, ленты, в сейфе]. Операция архивирования выполняется не автоматически [функционал не из коробки], есть скрипт не официальный, который может выполнять эту задачу (доступен из комьюнити в Community-Pack - см. **тут**), либо использовать автоматизацию. Объем занимаемого пространства примерно на **40% меньше**, чем при горячем/холодном хранении (например, при потоке событий в 1000 EPS, для его хранения в течении 1 месяца требуется 1 ТБ места в хранилище (без учета реплик), в архиве будет занимать 600-700 GB).
- С версии 4.2 из веб-интерфейса стала возможным настройка автоматического архивирования + ручной режим, соответственно и импорт также возможен через веб

> KUMA позволяет гибко настроить политику хранения событий (retention-период) по разным условиям: По дате, По объему, По проценту от занятого места на диске

Revision #19
Created 2023-09-05 12:25:21 UTC by Boris Rzr
Updated 2026-05-18 10:40:02 UTC by Koala