



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Факультет компьютерных наук ООО Яндекс Программный проект

TTL (настраиваемое время хранения) данных в таблицах ClickHouse и в отдельных столбцах

Выполнил студент группы БПМИ-173

Попов Антон Дмитриевич

Научный руководитель:

Миловидов Алексей Николаевич,

руководитель группы разработки

ClickHouse, Яндекс

ClickHouse - распределенная аналитическая колоночная СУБД, поддерживающая SQL-подобный диалект для запросов. Исходный код находится в open-source.

- Таблица состоит из *кусков* данных, отсортированных по первичному ключу
- При вставке в таблицу создаются отдельные куски данных
- Данные, относящиеся к разным партициям, разбиваются на разные куски
- В фоновом режиме ClickHouse выполняет слияния кусков данных в рамках одной партиции для более эффективного хранения

ОСНОВНЫЕ ПОНЯТИЯ, ОПРЕДЕЛЕНИЯ, ТЕРМИНЫ

Block – минимальная единица обработки данных. Всегда помещается в оперативную память

DataStream – часть выполнения конвейера запроса. Выполняет какую-то операцию над блоком, после чего передает блок на чтение/запись следующему потоку.

TTL – время хранения данных.

Партиция – набор записей в таблице, объединенных по какому-либо критерию. Например, партиция может быть по месяцу, по дню или по типу события. Каждая партиция состоит из отдельных фрагментов или так называемых *кусков данных*

Постановка задачи

- Старые данные не актуальны для аналитики. Можно периодически их удалять для экономии места на диске.
- Возможное решение - выполнять запрос drop partition
- Хотелось бы автоматизировать данный процесс

Автоматизация контроля за хранением данных.

- Экономия места на диске
- Соблюдение законодательства, регулирующего хранение персональных данных

Есть запросы пользователей в виде issue на GitHub.



Сложность задачи

- Данные хранятся в неизменяемых кусках.
- Единственный способ изменить данные - сделать слияние кусков или удалить целиком.
- Необходимо учитывать TTL при слияниях, а также назначать их с учетом TTL.

Цель и задачи работы

Цель работы

Добавить возможность для таблиц семейства MergeTree указывать время хранения данных для столбцов и всей таблицы, по прошествии которого эти данные удаляются

Задачи работы

1. Реализовать при запросе создания таблицы семейства MergeTree возможность указания TTL в качестве параметра столбца таблицы и/или параметра таблицы.
2. Добавить логику при слиянии кусков, которая учитывает TTL.
3. Реализовать удаление данных с истекшим TTL.

Создание таблицы

Указание TTL для столбца:

```
CREATE TABLE t (d DateTime, a Int TTL d + INTERVAL 1 DAY)
```

Указание TTL для таблицы:

```
CREATE TABLE t (d DateTime, a Int) engine = MergeTree PARTITION  
BY toDayOfMonth(d) ORDER BY a TTL d + INTERVAL 1 DAY
```

Ограничения на выражение для TTL

- Нельзя устанавливать TTL на все ключевые столбцы
- Выражение должно зависеть хотя бы от одного Date/DateTime столбца
- Результатом выражения должен быть столбец с типом Date или DateTime

Информация о TTL для кусков

- Для каждого куска сохраняется файл `ttl.txt`, в котором хранится минимальное и максимальное время TTL для каждого столбца и для всей таблицы.
- Формат хранения - JSON.
- Данная информация обновляется при вставке и при слиянии кусков.

Слияния кусков

При выборе кусков для слияния с учетом TTL используется TTLMergeSelector. Он работает, используя жадный алгоритм: выбирается кусок с самым старым истекшим TTL, после чего выбирается наибольший отрезок кусков с истекшим TTL, который содержит нужный кусок.

При этом периодичность использования TTLMergeSelector ограничена настройкой `merge_with_ttl_timeout`.

Удаление устаревших данных

- Фильтрация данных происходит в `TTLBlockInputStream`
- Значения с истекшим TTL заменяются на значения по умолчанию
- Данные столбцов, в которых все значения были занулены, стираются с диска
- Строки с истекшим TTL удаляются

Основные результаты работы

Поставленные задачи были выполнены.

Результат - открытый Pull Request

<https://github.com/yandex/ClickHouse/pull/4212>

Направления дальнейшей работы

- Добавить тест производительности
- Реализовать запрос ALTER MODIFY TTL для изменения выражения TTL



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Спасибо за внимание!

Попов Антон Дмитриевич
adpopov_2@edu.hse.ru

Москва - 2019