

Возможности колоночной аналитической СУБД ClickHouse

Научный руководитель

Панфилов Петр Борисович
Профессор, Высшая школа бизнеса



Докладчик

Волков Андрей Андреевич
Студент 4 курса, «Бизнес-информатика»

Мотивация



Мотивация



OLAP (интерактивная аналитическая обработка, англ. Online Analytical Processing) – технология обработки данных, заключающаяся в выборе агрегированной информации из больших массивов данных.

Мотивация



OLAP (интерактивная аналитическая обработка, англ. Online Analytical Processing) – технология обработки данных, заключающаяся в выборе агрегированной информации из больших массивов данных.

Сферы применения OLAP:

- системы поддержки принятия решения
- статистика веб-сайтов
- предиктивная аналитика
- финансовая отчетность

Мотивация



OLAP (интерактивная аналитическая обработка, англ. Online Analytical Processing) – технология обработки данных, заключающаяся в выборе агрегированной информации из больших массивов данных.

ClickHouse - колоночная аналитическая система управления базами данных (СУБД) для онлайн обработки аналитических запросов (OLAP).

Цель



Исследование возможностей колоночной аналитической СУБД ClickHouse в применении к задачам обработки данных и аналитике около реального времени, а также сбор и анализ метрик производительности при основных сценариях использования СУБД.

Цель



Исследование возможностей колоночной аналитической СУБД ClickHouse **в применении к задачам обработки данных и аналитике около реального времени**, а также сбор и анализ метрик производительности при основных сценариях использования СУБД.

Задачи



- исследовать возможности колоночной СУБД ClickHouse
- подготовить и развернуть инфраструктуру для тестирования
- провести тестирование основных возможностей ClickHouse
- проанализировать полученные результаты

Тестовая инфраструктура



- сервер ClickHouse
- экспортер метрик для ClickHouse
- сервис для сбора и хранения метрик Prometheus
- сервис для отображения метрик Grafana

Тестовые данные



В качестве данных были использованы случайно сгенерированные данные о погодных условиях, на сервере ClickHouse была создана таблица при помощи следующего запроса.

Данный запрос создает таблицу `weather`, в колонках которой содержится информация о погодных условиях. Также определено партиционирование по часам, первичный ключ и ключ сортировки задан по колонкам `country`, `timestamp`, `latitude`, `longitude`, `temperature`.

```
CREATE OR REPLACE TABLE weather (  
    country LowCardinality(String),  
    timestamp DateTime,  
    latitude Float64,  
    longitude Float64,  
    temperature Float64,  
    min_temperature Float64,  
    max_temperature Float64,  
    station_pressure Float64,  
    sea_level_pressure Float64,  
    dew_point Float64,  
    relative_humidity Float64,  
    actual_vapour_pressure Float64,  
    wind_speed Float64,  
    precipitation Float64,  
    fog UInt8  
)  
ENGINE = MergeTree()  
PARTITION BY toStartOfHour(timestamp)  
PRIMARY KEY (country, timestamp, latitude, longitude, temperature)  
ORDER BY (country, timestamp, latitude, longitude, temperature);
```

Обновление данных в реальном времени



В ClickHouse используется концепция таблиц с первичным ключом. Это значит, что данные инкрементально сортируются при помощи merge дерева в диапазоне первичного ключа, что способствует высокой скорости обработки данных. Благодаря данному подходу, поддерживается постоянная вставка данных в таблицу без каких-либо блокировок.

Кол-во записей	Средний размер батча (мб)	Среднее кол-во запросов в минуту	Среднее время обработки запроса (сек)	Пропускная способность (мб / сек)
1 000	0.11	18.8	0.127	0.866
10 000	0.99	18.2	0.198	5.000
50 000	4.86	14	0.851	5.710
100 000	9.91	10	1.85	5.357
200 000	18.98	6.27	4.63	4.099

Сжатие данных



ClickHouse, в отличие от других колоночных СУБД (MonetDB, InfiniDB CE), использует специальные алгоритмы для сжатия данных. Сжатие данных позволяет выделять меньше места на дисках и играет большую роль в демонстрации высокой производительности.

Сжатие данных



ClickHouse, в отличие от других колоночных СУБД (MonetDB, InfiniDB CE), использует специальные алгоритмы для сжатия данных. Сжатие данных позволяет выделять меньше места на дисках и играет большую роль в демонстрации высокой производительности.

100 МЛН

Записей было
сгенерировано

Сжатие данных



ClickHouse, в отличие от других колоночных СУБД (MonetDB, InfiniDB CE), использует специальные алгоритмы для сжатия данных. Сжатие данных позволяет выделять меньше места на дисках и играет большую роль в демонстрации высокой производительности.

100 МЛН

Записей было
сгенерировано

37,5 ГБ

Объем «сырых»
данных

Сжатие данных



ClickHouse, в отличие от других колоночных СУБД (MonetDB, InfiniDB SE), использует специальные алгоритмы для сжатия данных. Сжатие данных позволяет выделять меньше места на дисках и играет большую роль в демонстрации высокой производительности.

100 МЛН

Записей было
сгенерировано

7,8 ГБ

Сжатых данных

37,5 ГБ

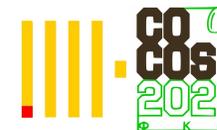
Объем «сырых»
данных

Наличие индекса



Поскольку данные в ClickHouse физически отсортированы по первичному ключу, то это позволяет получать данные для конкретных значений или их диапазонов с относительно небольшими задержками.

Наличие индекса



Запрос	Время выполнения запроса (сек)	Кол-во обработанных строк
Температура с разбивкой по перцентилям в Австралии	0.560	~ 585 000
Скорость ветра с разбивкой по перцентилям в Австралии	0.240	
Кол-во записей, в которых туман = 1 с разбивкой по часам в Австралии	0.278	
Температура с разбивкой по перцентилям во всех странах	6.700	~ 100 000 000
Скорость ветра с разбивкой по перцентилям во всех странах	10.40	
Кол-во записей, в которых туман = 1 с разбивкой по часам во всех странах	7.310	

Наличие индекса



Запрос	Время выполнения запроса (сек)	Кол-во обработанных строк
Температура с разбивкой по перцентилям в Австралии	0.560	~ 585 000
Скорость ветра с разбивкой по перцентилям в Австралии	0.240	
Кол-во записей, в которых туман = 1 с разбивкой по часам в Австралии	0.278	
Температура с разбивкой по перцентилям во всех странах	6.700	~ 100 000 000
Скорость ветра с разбивкой по перцентилям во всех странах	10.40	
Кол-во записей, в которых туман = 1 с разбивкой по часам во всех странах	7.310	

Поддержка приближённых вычислений



Для работы с большими объемами данных ClickHouse предоставляет возможность работать с частью данных (англ. *sample data*), что позволяет за меньшее время выполнять необходимые запросы с приближительной точностью.

Поддержка приближённых вычислений



Запрос (10% всех данных)	Время выполнения запроса - 100% (сек)	Время выполнения запроса (сек)	Кол-во обработанных строк
Температура с разбивкой по перцентилям в Австралии	0.560	0.200	~ 58 500
Скорость ветра с разбивкой по перцентилям в Австралии	0.240	0.171	
Кол-во записей, в которых туман = 1 с разбивкой по часам в Австралии	0.278	0.200	
Температура с разбивкой по перцентилям во всех странах	6.700	1.251	~ 10 000 000
Скорость ветра с разбивкой по перцентилям во всех странах	10.40	1.525	
Кол-во записей, в которых туман = 1 с разбивкой по часам во всех странах	7.310	0.769	

Поддержка приближённых вычислений



Запрос (10% всех данных)	Время выполнения запроса - 100% (сек)	Время выполнения запроса (сек)	Кол-во обработанных строк
Температура с разбивкой по перцентилям в Австралии	0.560	0.200	~ 58 500
Скорость ветра с разбивкой по перцентилям в Австралии	0.240	0.171	
Кол-во записей, в которых туман = 1 с разбивкой по часам в Австралии	0.278	0.200	
Температура с разбивкой по перцентилям во всех странах	6.700	1.251	~ 10 000 000
Скорость ветра с разбивкой по перцентилям во всех странах	10.40	1.525	
Кол-во записей, в которых туман = 1 с разбивкой по часам во всех странах	7.310	0.769	

Дальнейшие исследования



1. Собрать и проанализировать метрики утилизации ресурсов при разных сценариях использования
2. Составить бенчмарки для анализа скорости выполнения запросов и сравнить возможности ClickHouse с другими колоночными аналитическими СУБД (InfiniDB, SAP HANA, Vertica)

Заключение



Высокая пропускная способность и отличные **алгоритмы сжатия** позволяют данной СУБД обрабатывать большие объемы данных на вставку и хранить их без чрезмерных накладных ресурсов. **Наличие индекса** и **поддержка сэмплирования** данных позволяет быстро выполнять аналитические запросы к СУБД, избегая длительных задержек при выполнении.

Исходя из результатов исследования, СУБД ClickHouse отлично подходит для сбора и анализа аналитических данных, а также для задач аналитики около реального времени.

Список источников



1. Документация ClickHouse [Электронный ресурс] / Компания «Яндекс» – Режим доступа: <https://clickhouse.tech/docs/ru/>, свободный. (дата обращения: 05.04.21)
2. ClickHouse Research [Электронный ресурс] / Github/VolkovTech – Режим доступа: <https://github.com/VolkovTech/clickhouse-research>, свободный. (дата обращения: 05.04.21)
3. Отличительные возможности ClickHouse [Электронный ресурс] / Компания «Яндекс» – Режим доступа: <https://clickhouse.tech/docs/ru/introduction/distinctive-features/>, свободный. (дата обращения: 05.04.21)
4. ClickHouse for Time Series [Электронный ресурс] / Altinity Team – Режим доступа: <https://altinity.com/blog/clickhouse-for-time-series>, свободный. (дата обращения: 05.04.21)
5. Comparison of the Open Source OLAP Systems for Big Data: ClickHouse, Druid, and Pinot [Электронный ресурс] / Roman Leventov – Режим доступа: <https://leventov.medium.com/comparison-of-the-open-source-olap-systems-for-big-data-clickhouse-druid-and-pinot-8e042a5ed1c7>, свободный. (дата обращения: 05.04.21)



Спасибо за внимание!