

Примеры использования БД ClickHouse в компаниях

Окко: Сбор и хранение событий трекинга событий пользователей (ODD-слой)

Специфика

Пользователи, при взаимодействии с сервисом, генерируют события при переходах на разделы сервиса и активности на экранах (клик, фокус, иные интерации на экранах).

Применение

Использование ClickHouse для сбора и хранения фактов трекинга событий пользователя. Формирование ODD-слоя данных, для последующей трансформации и агрегации данных. Данные экспортируются топиков Kafka и представляют собой MergeTree таблицы.

Источники

1. Проект текущего места работы

Окко: Интеграция с внутренними сервисами платформы для получения требуемых справочников в сыром слое данных

Специфика

Инфраструктура платформы включает ряд сервисов, которые имеют различные интерфейсы интеграции (RestAPI, GraphQL) для получения данных, требуемых в различных сценариях аналитики и задачах машинного обучения.

Применение

Используется S3 движок для предоставления результатов экспорта данных из сторонних систем в ClickHouse в различных форматах (csv, json, xml).

Источники

1. Проект текущего места работы

VK: Хранение логов микросервисов компании

Специфика

Требуется хранить отладочные логи высоконагруженных систем (конвертация видое, обработка файлов) для последующего мониторинга и алертинга.

Применение

Использование ClickHouse для хранения логов сервисов. Экспорт логов осуществляется через самописный сервис экспорта событий, который предназначен для обхода ограничений ClickHouse (частая вставка, гарантия доставки данных).

Источники

1. [Использование ClickHouse в VK, или Зачем мы написали KittenHouse](#)

Ответы на вопросы из лекции

К каким классам систем относится ClickHouse

ClickHouse является OLAP-системой, предназначенной для вставки больших объемов данных и быстрого формирования отчетов на накопленных исторических данных.

Источники:

1. [Что такое OLAP?](#)

Какую проблему вы бы решили используя ClickHouse, а какую бы не стали?

Какую проблему вы бы решили используя ClickHouse

Формирование RAW-слоя данных, который предполагает хранение данных различных форматов из внешних систем.

Какую проблему вы бы не стали решать используя ClickHouse

Формирование DDS слоя с высокой степенью нормализации (к примеру [Data Vault](#)).

Источники

1. [Data Lake – от теории к практике. Сказ про то, как мы строим ETL на Hadoop](#)

Где можно получить помощь по ClickHouse и куда сообщать о багах?

Где можно получить помощь:

1. [Русскоязычный канал сообщества ClickHouse](#)
2. [Англоязычный канал сообщения ClickHouse](#)
3. [Официальная документация ClickHouse](#)
4. [База знаний проекта Altinity](#)

Куда сообщать о багах:

1. [Официальный репозиторий ClickHouse на GitHub](#)