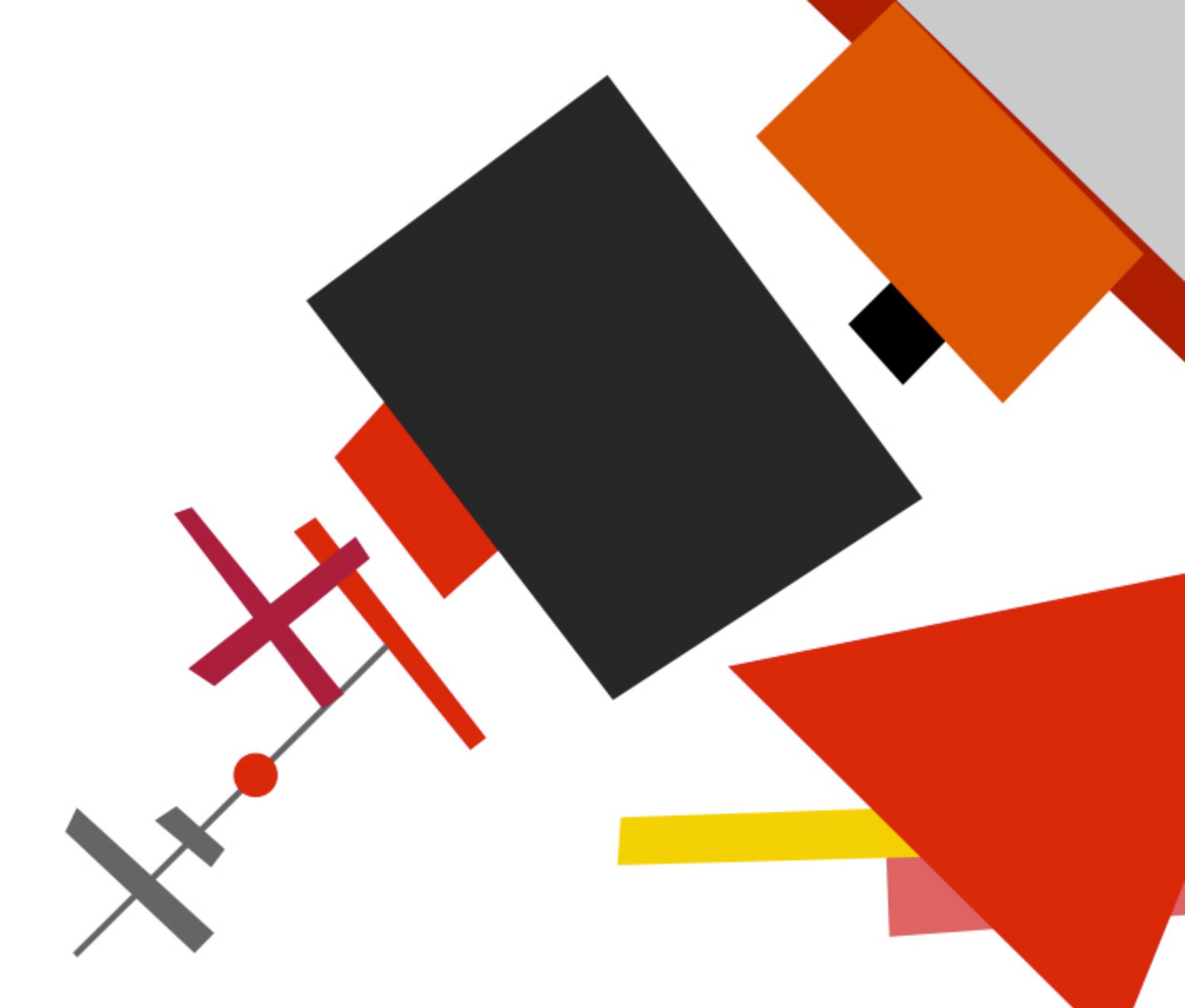


Яндекс



ClickHouse: очень быстро и очень удобно

Виктор Тарнавский



HighLoad⁺⁺

Профессиональная конференция
разработчиков высоконагруженных
систем

Пару слов обо мне

- › Разработчик, в прошлом разработка железа
- › Руководитель отдела аналитических продуктов
- › jkee@yandex-team.ru



ClickHouse

Немного истории



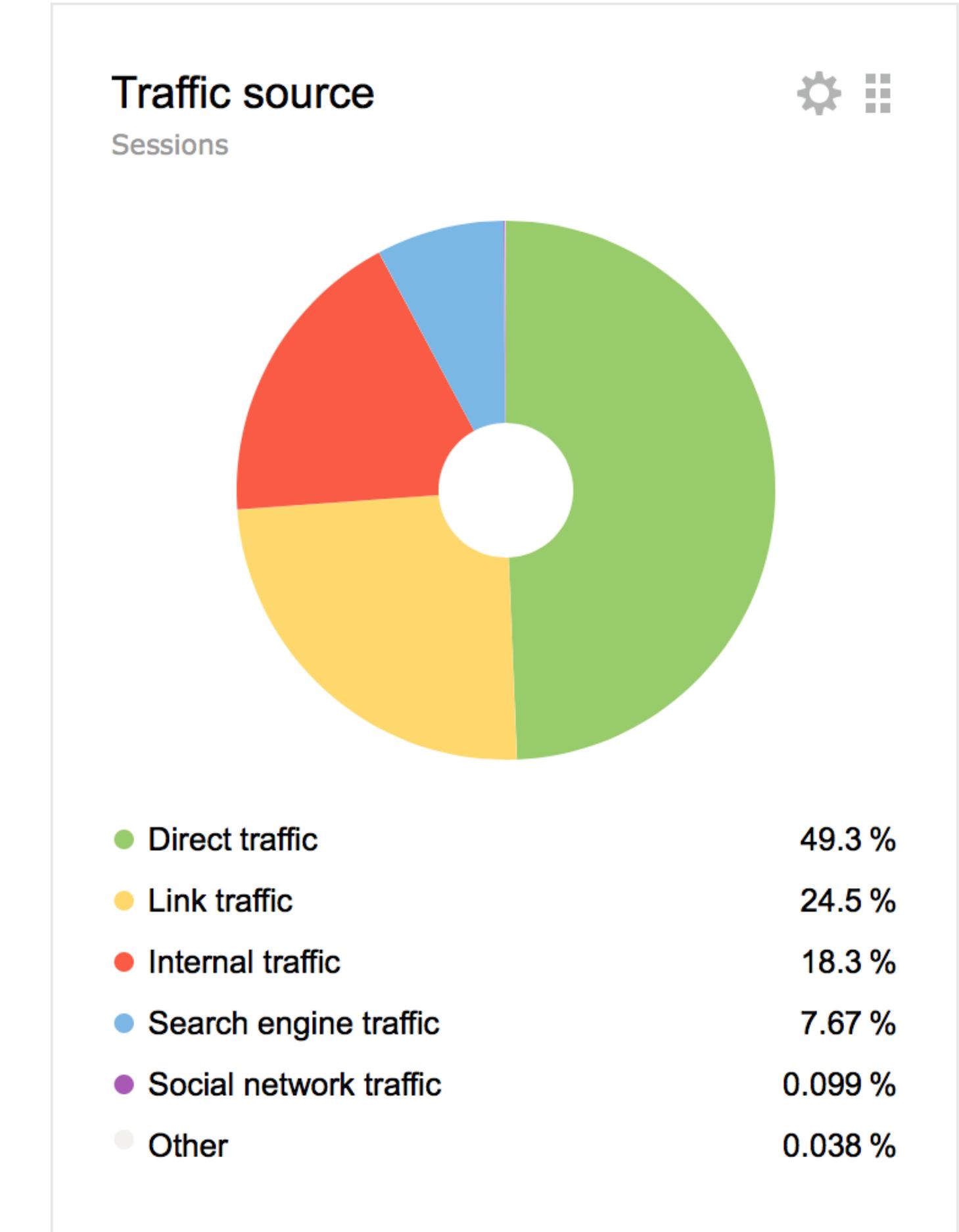
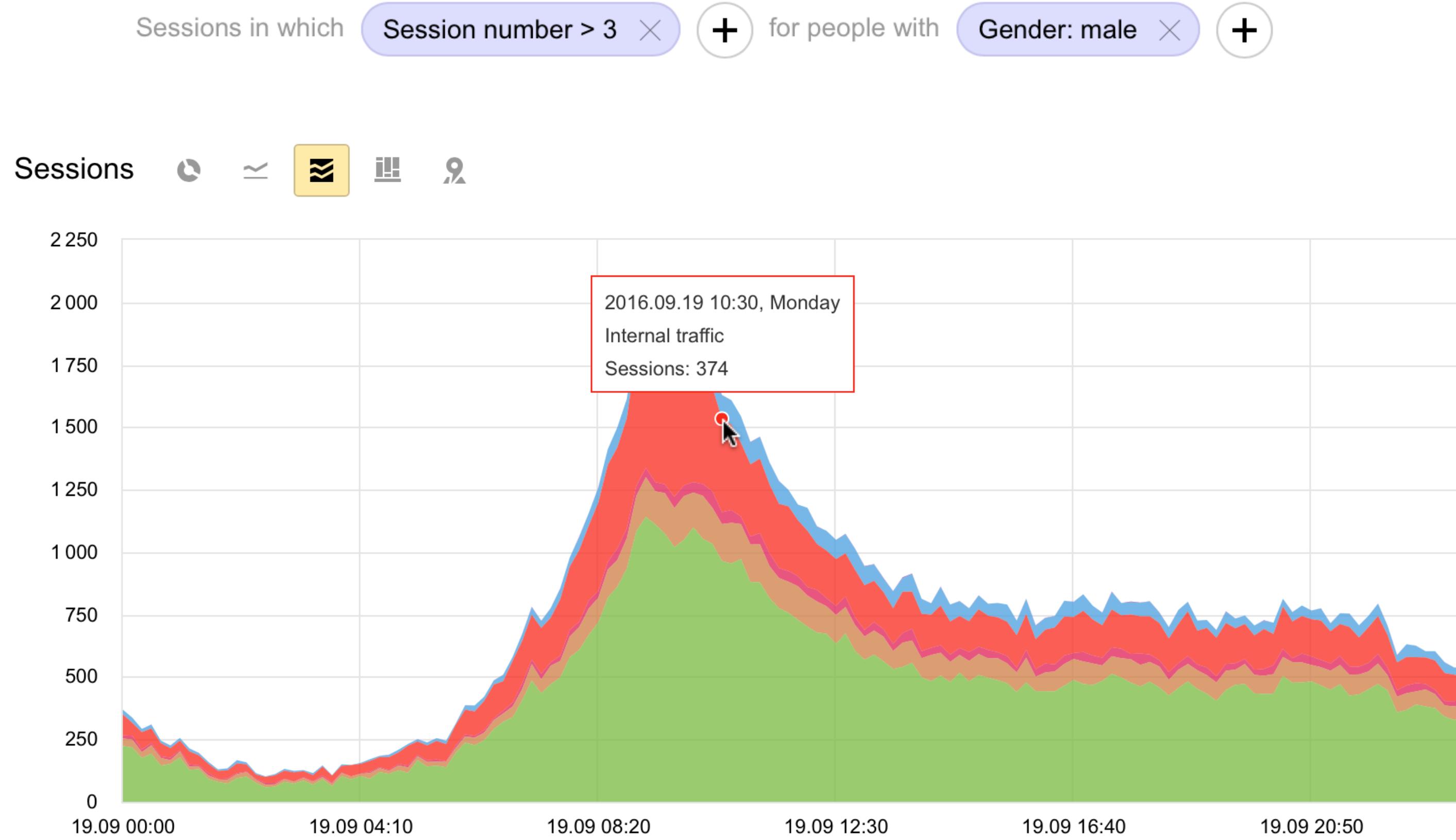
Яндекс.Метрика

Яндекс.Метрика - в тройке крупнейших веб-аналитических систем по количеству сайтов.

- › Более 20 млрд. событий в день
- › Более 1 млн. сайтов
- › Более 100 000 аналитиков каждый день

Яндекс Метрика

Яндекс.Метрика





Как хранить данные?

Классический подход

- › Агрегированные данные
- › Много разных таблиц

| Key → Metrics

Современный подход

- › Неагрегированные данные
- › Небольшое количество широких таблиц
- › Высокоэффективные фильтры и группировки

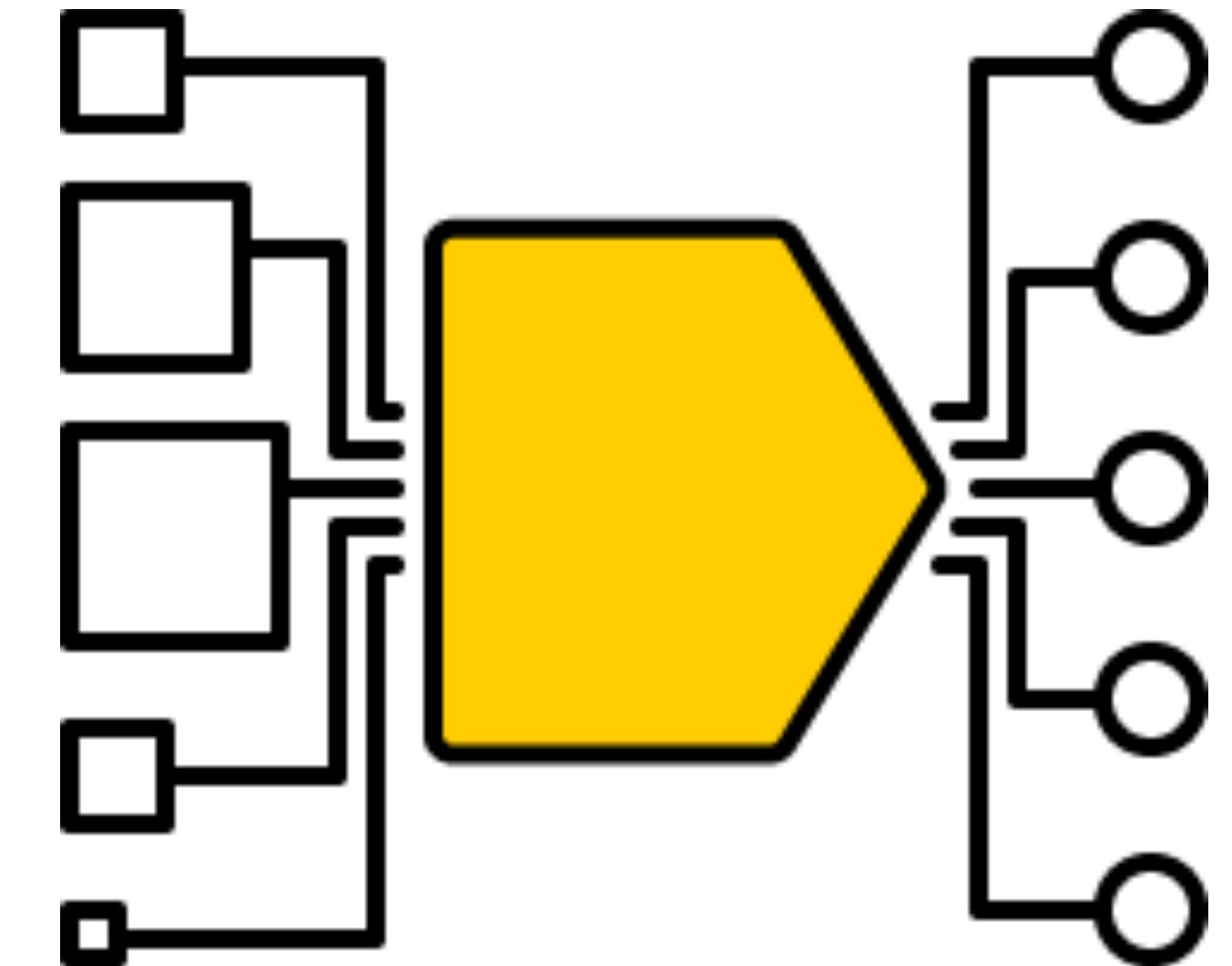
| Event → Properties



Какую базу данных
взять?

Что мы хотим от базы данных?

- › Максимально быстро выполнять запросы
- › Обработка данных в реальном времени
- › Возможность хранить петабайты данных
- › Отказоустойчивость в терминах датацентров
- › Гибкий язык запросов

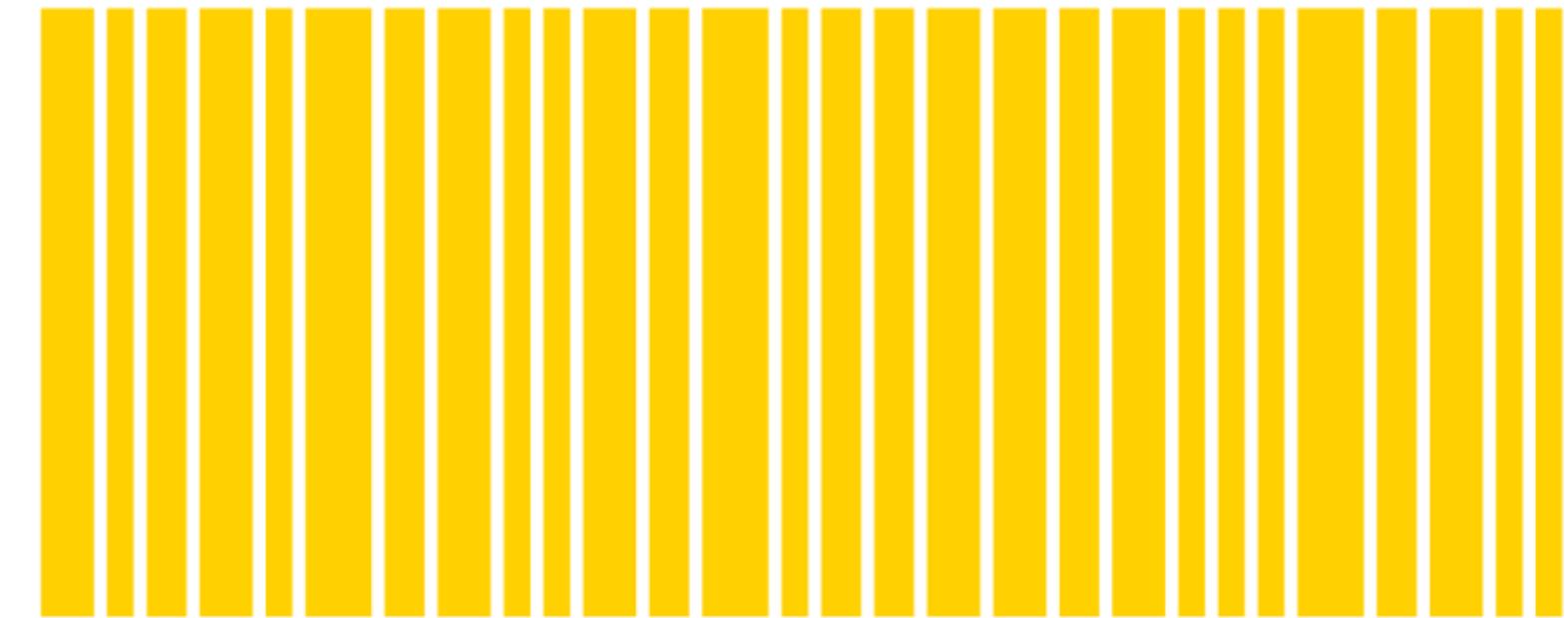




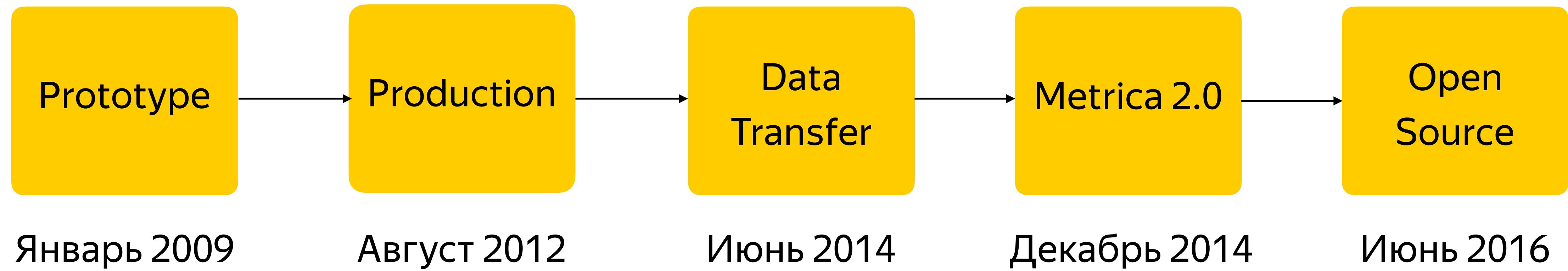
На рынке пустота!

Основные идеи при создании ClickHouse

- › SQL
- › Линейная масштабируемость
- › Фокус на быстром выполнении запросов
- › Column-oriented

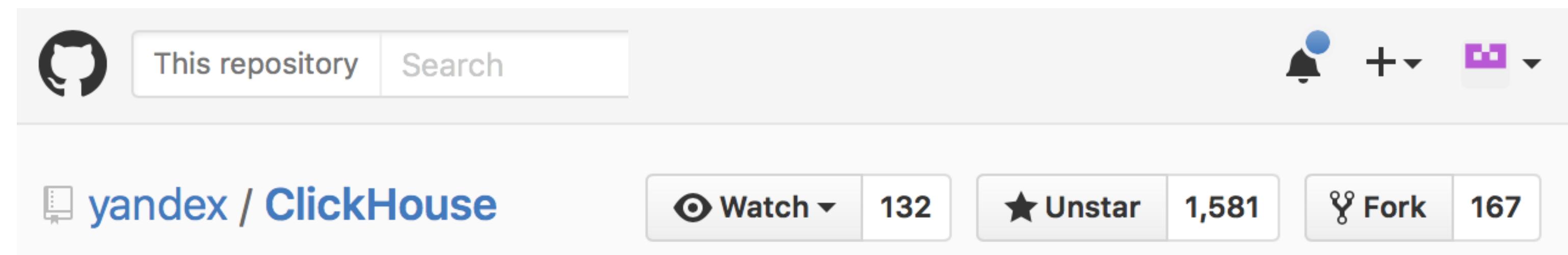
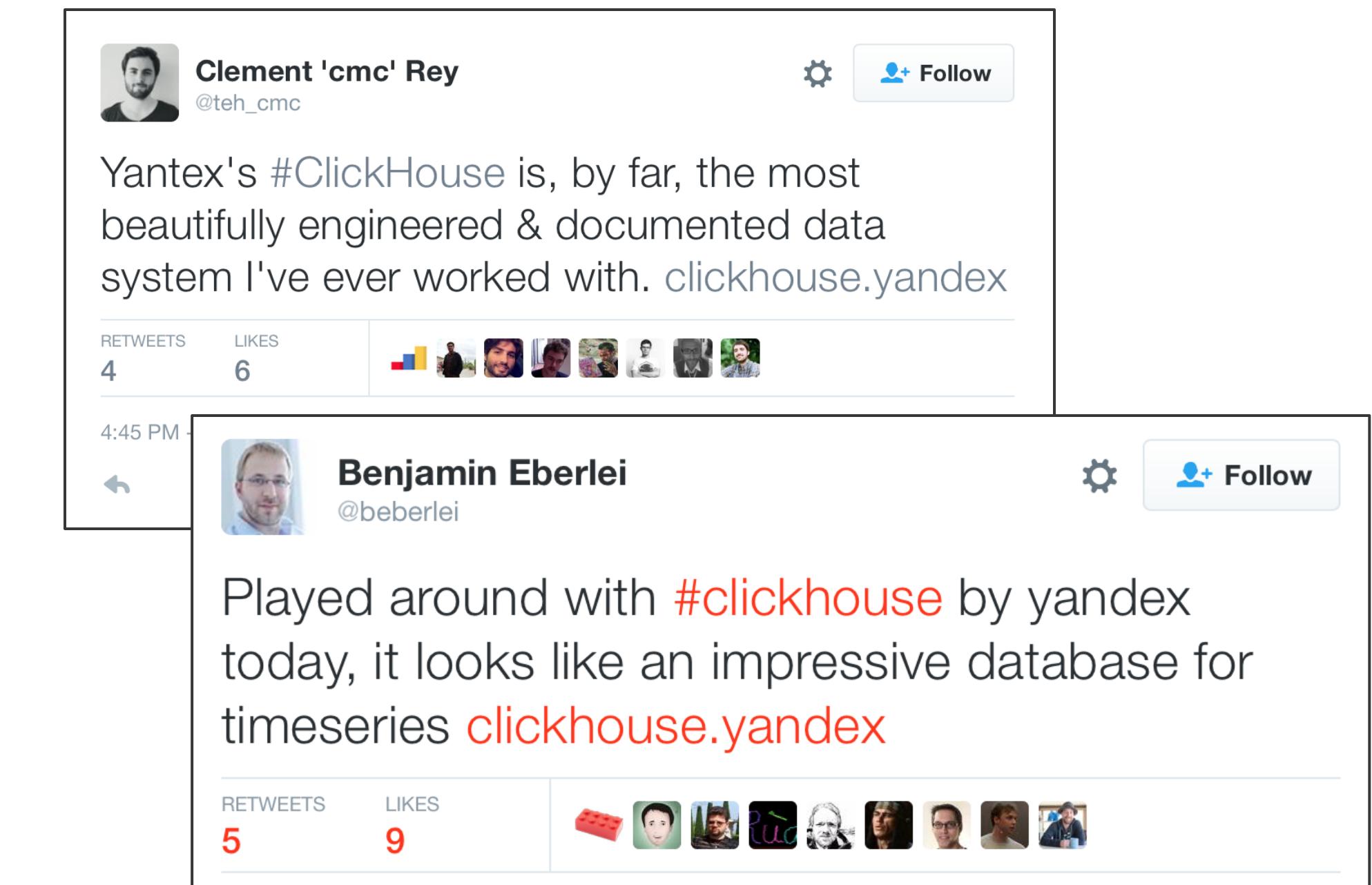


ClickHouse timeline



ClickHouse сегодня

- › Более 20 проектов в Яндексе
- › Open-source с Июля 2016
- › В продакшне вне Яндекса



ClickHouse

Возможности



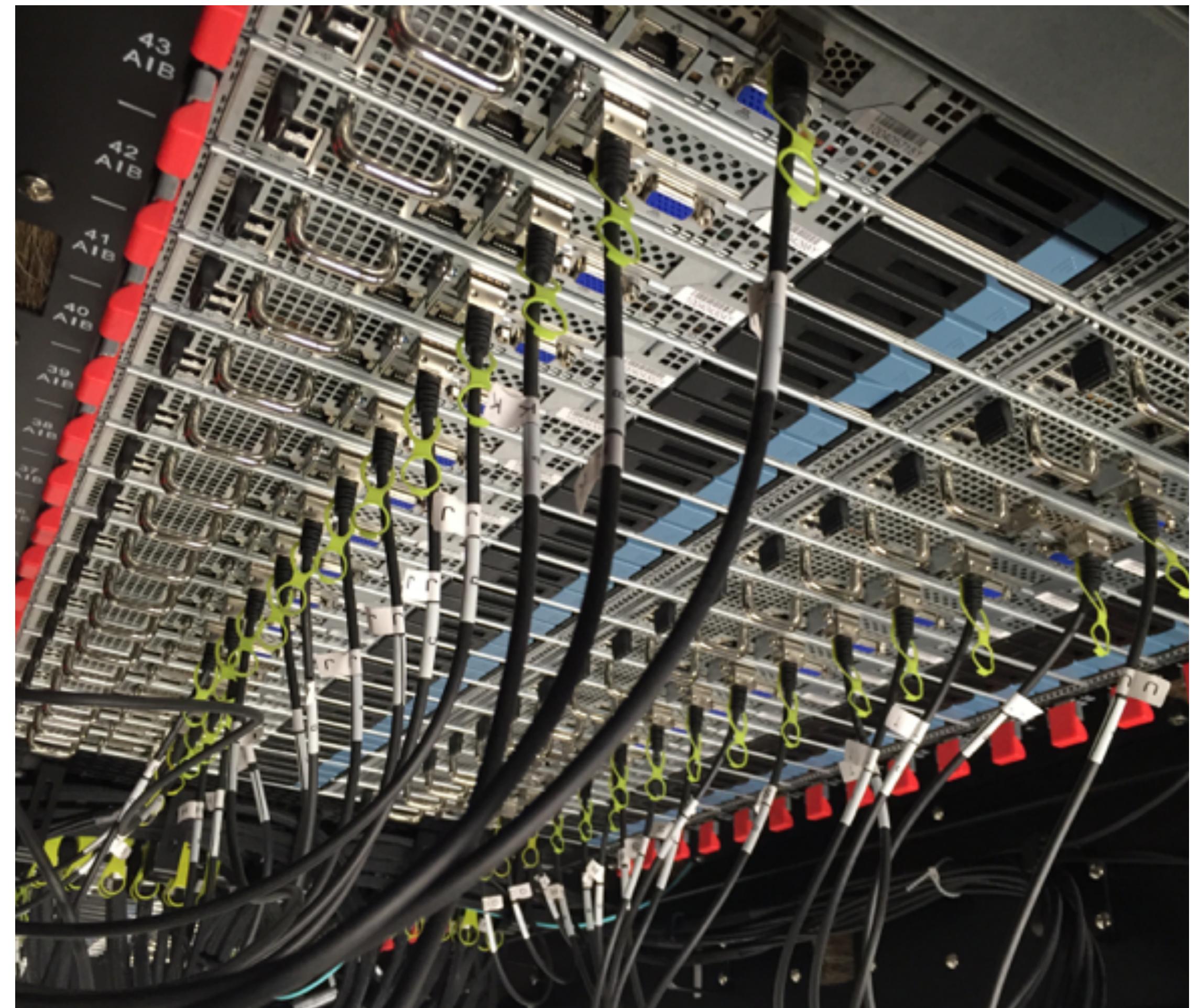
Линейная масштабируемость

- › Масштабируется до петабайт данных
- › Cross-datacenter
- › High availability
- › Сжатие данных



Кластер Яндекс.Метрики

- › 3 Pb
- › 412 серверов
- › 6 dataцентров
- › Часы даунтайма за 4 года



Запросы

- › SQL диалект
- › Приблизительные вычисления,
URI функции
- › Массивы, кортежи
- › Распределённые запросы из коробки
- › Внешние словари



Weekly traffic and audience

```
SELECT
    count() as visits,
    sum(PageViews) as hits,
    uniq(UserID) as users
FROM visits_all
WHERE StartDate > today() - 7
```

```
# Using external dictionary for regions
```

```
SELECT
    count() as visits,
    regionToName(regionToCountry(RegionID), 'en')
as country
FROM visits_all
WHERE StartDate > today() - 7
GROUP BY country
ORDER BY visits DESC
LIMIT 10
```

Скорость на грани фантастики

- › Типичные запросы быстрей чем за секунду
- › >100 раз быстрей чем Hadoop
>100 раз быстрей чем типичная DBMS
- › 100 млн - 1 миллиард строк в секунду на одной ноде
- › До 2 терабайт в секунду для кластера на 400 нод



У тебя нет ClickHouse



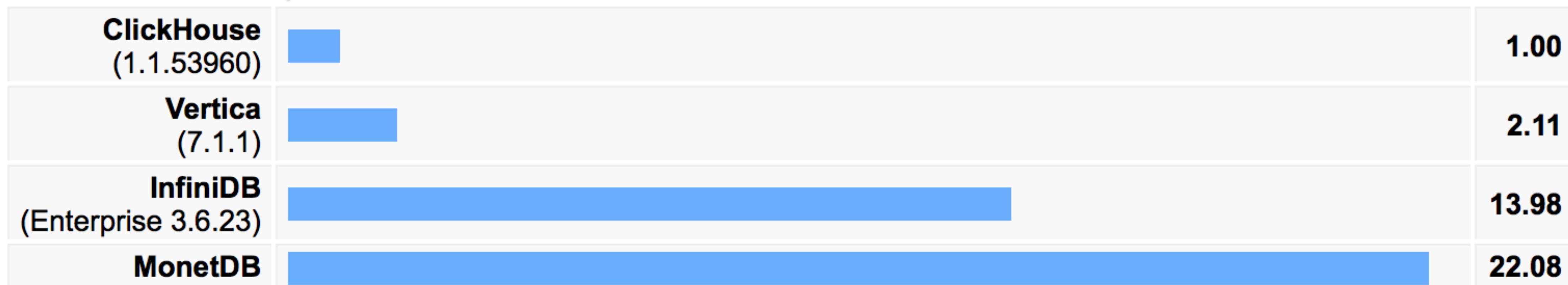
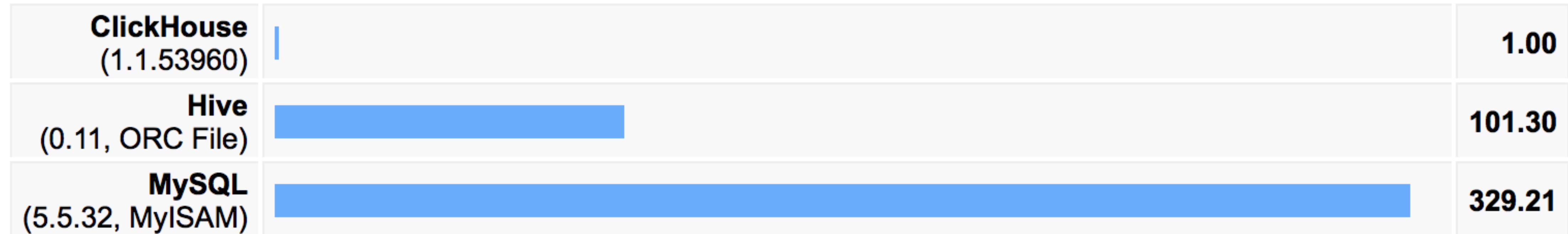
1. Делаем запрос
2. Идём пить кофе

У тебя есть ClickHouse



Проверь 37 идей за 5 минут!
ВНИМАНИЕ: развидеть нельзя

Relative query processing time (lower is better):



More info: <https://clickhouse.yandex/benchmark.html>

Интерфейсы

- › Консольный клиент
- › HTTP
- › JDBC
- › Python, PHP, NodeJS,
Go и Perl коннекторы



ClickHouse

Заглянем внутрь



Почему ClickHouse такой быстрый?

Код

- › Векторная обработка данных
- › Ориентированный на производительность C++14 код
- › Каждая фича оптимизируется по производительности



Почему ClickHouse такой быстрый?

Данные

- › Колоночное хранение
- › Merge Tree
минимум seek-ов
- › Обработка максимально близко
к данным



Почему ClickHouse такой быстрый?

Возможности

- › Семплирование из коробки
- › Функции с вероятностными алгоритмами
- › Оптимизация производительности на уровне запроса

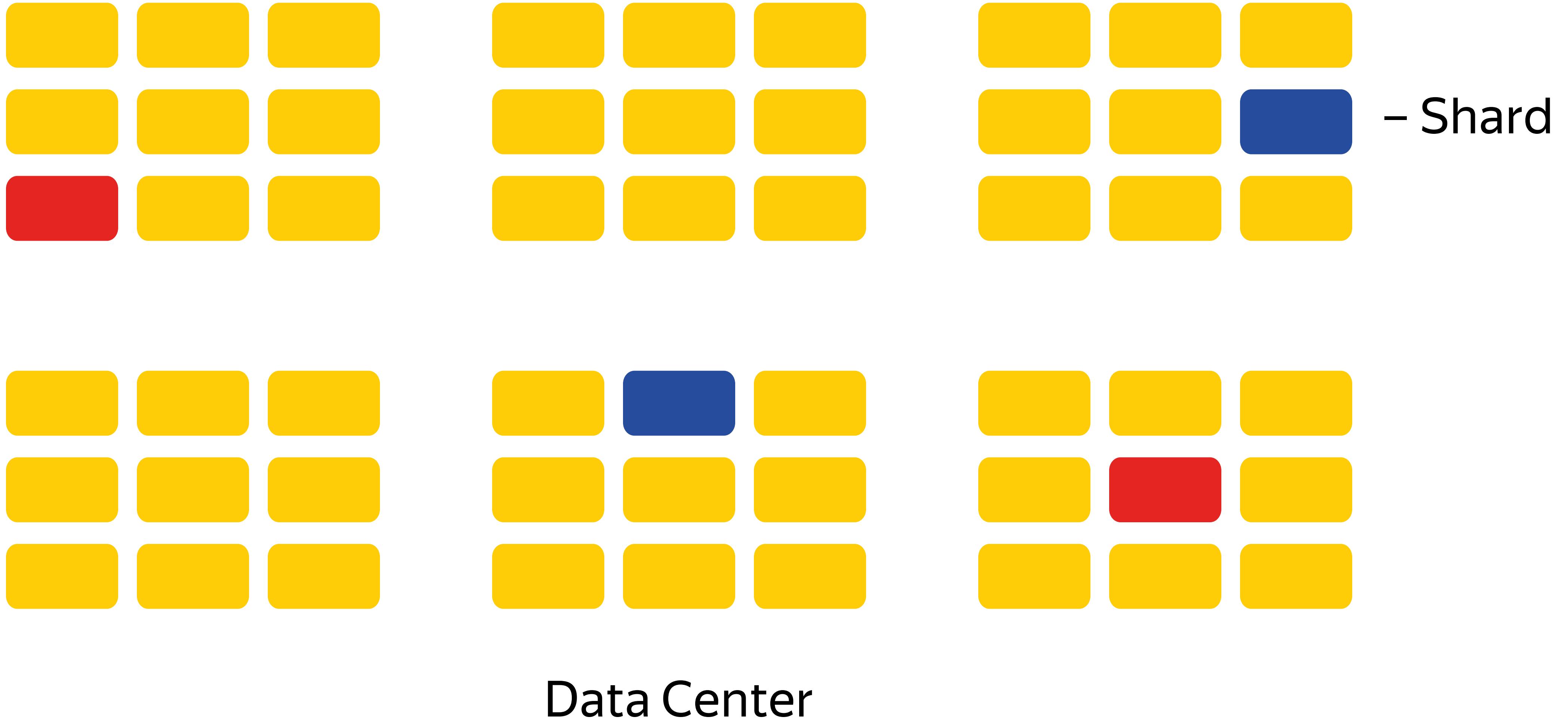


Масштабируемость и отказоустойчивость

- › Нет единой точки координации запросов
- › Асинхронная репликация
- › ZooKeeper для координации реплик (не используется в SELECT)



Кластер Яндекс.Метрики



ClickHouse

Кейсы



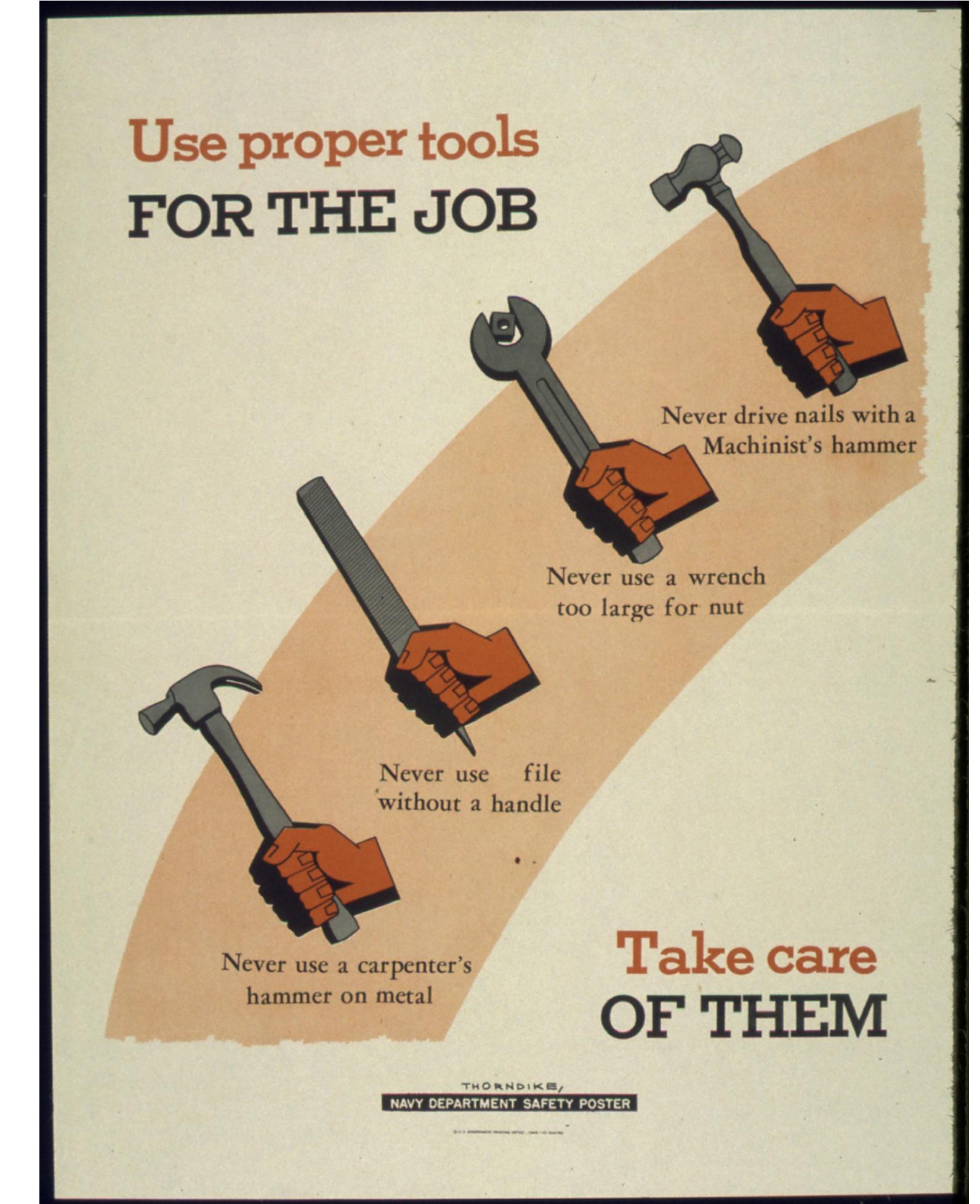
Когда НЕ надо использовать

- › OLTP
- › Key-value
- › Document-oriented
- › Модификация данных



ClickHouse: best practises

- › Широкие таблицы
МНОГО КОЛОНOK
- › QPS относительно небольшой
НО МНОГО данных используется
в каждом запросе
- › Большой поток входящих данных
- › Аналитика по петабайтам данных



Кейс: анализ серверных логов

С этого обычно пробуют ClickHouse

Что делать:

- › Вставить access и серверные логи в ClickHouse
примерное время: несколько часов

Что получаем:

- › Анализируем любые инциденты за секунды
- › Мониторинг и дашборды по ошибкам, таймингам и прочему

Кейс: внутренняя база для аналитики

Можно построить свой Data Warehouse и копать данные с невероятной скоростью

Что делать:

- › Берём Hadoop или другую "не такую быструю" базу
- › Копируем все данные в ClickHouse

Что получаем:

- › Realtime аналитика бизнес-процессов
- › Довольные аналитики и менеджеры
- › Легко построить внутренние дашборды/метрики

ClickHouse

С чего начать?



С чего начать?

- › Tutorial: <https://clickhouse.yandex/tutorial.html>
- › Пишите вопросы: clickhouse-feedback@yandex-team.com
- › GitHub: <https://github.com/yandex/ClickHouse>
- › Больше информации: <https://clickhouse.yandex>

ClickHouse

Итого



ClickHouse за один слайд

- › Open-source колоночная DBMS
- › Линейная масштабируемость
- › Невероятная скорость
- › SQL диалект с расширенными возможностями



Контакты

Виктор Тарнавский

Руководитель отдела аналитических продуктов



jkee@yandex-team.ru



clickhouse-feedback@yandex-team.ru