

Группа компаний

**Я**ндекс Маркет

# Как устроена платформа управления данными в Яндекс.Маркет

Денис Хуртин, руководитель разработки DMP Яндекс.Маркет

# Data management platform (DMP)

- › Data Lake (raw data)
- › Data Warehouse (structured data + summary)
- › Data Marts (views)
  
- › Reporting (dashboards)
- › Analyzing (data mining)
  
- › Data Catalog (metadata)
- › Security (permissions + encryption)

# Основная проблема – быстрый рост

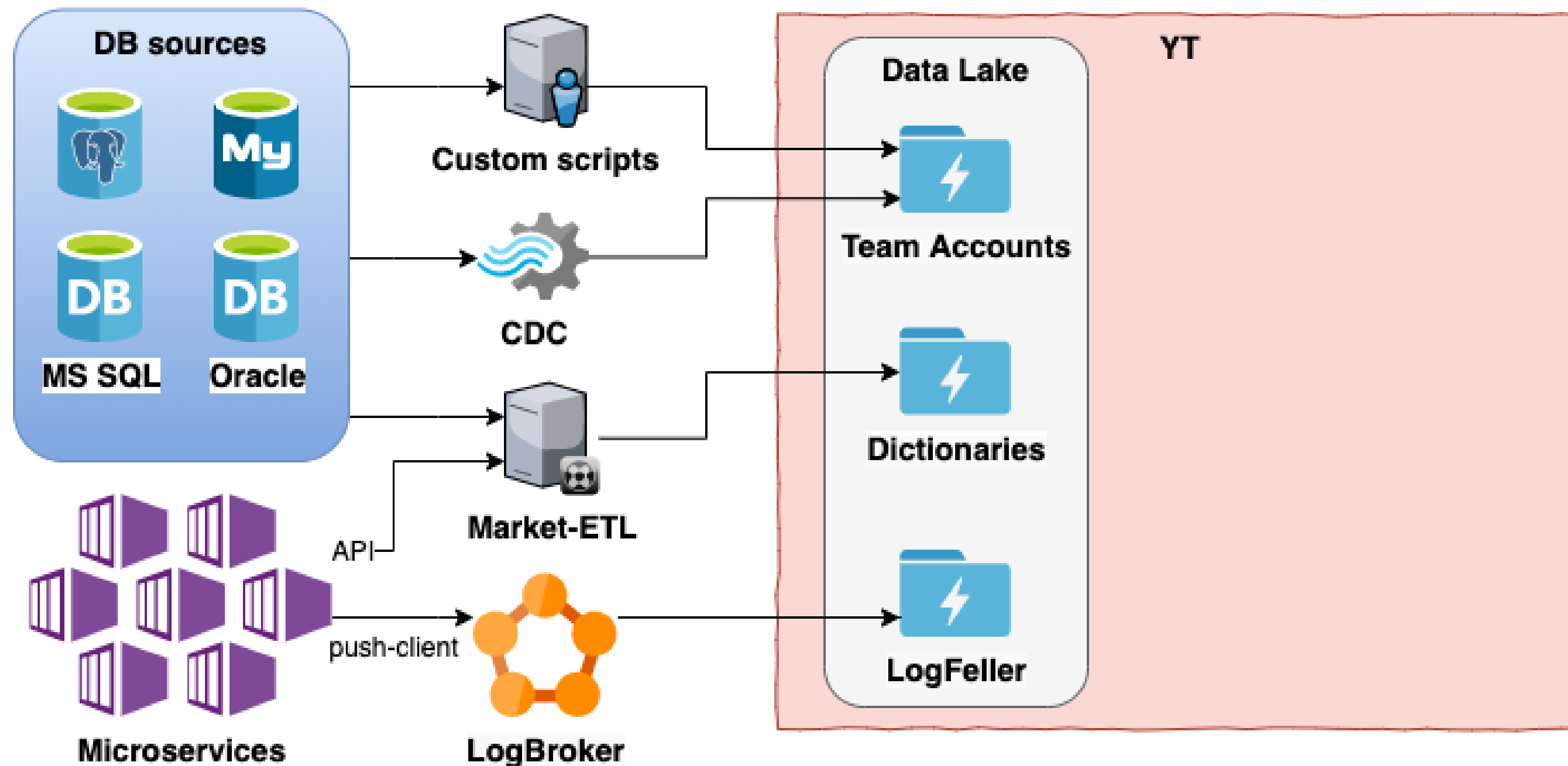
- › Наши данные растут в 2-3 раза каждые два года
  - новые сервисы и эксперименты
  - органический рост
- › Наша инфраструктура – своя + яндексовая
- › Наша инфраструктура – живой организм
  - модернизация и переезд на новые технологии

# Архитектура снизу вверх

- › Масштабируем там, где перестанет работать
- › Оптимизируем неэффективное
- › Ускоряем там, где перестает удовлетворять потребностям
- › Автоматизируем ручную работу

Data Lake

# Data Lake на Yandex Tables (YT)



# Нечестный Data Lake

1. Парсим и сохраняем в виде таблиц со схемой
2. Храним только последние N дней / месяцев / лет

# Data Lake в цифрах

65+

Уникальных  
систем-источников  
данных

450+

Уникальных  
таблиц ежедневно

50Tb+

Ежедневный объем  
данных

5Pb+

Общий объем на  
каждом из 2-х  
кластеров

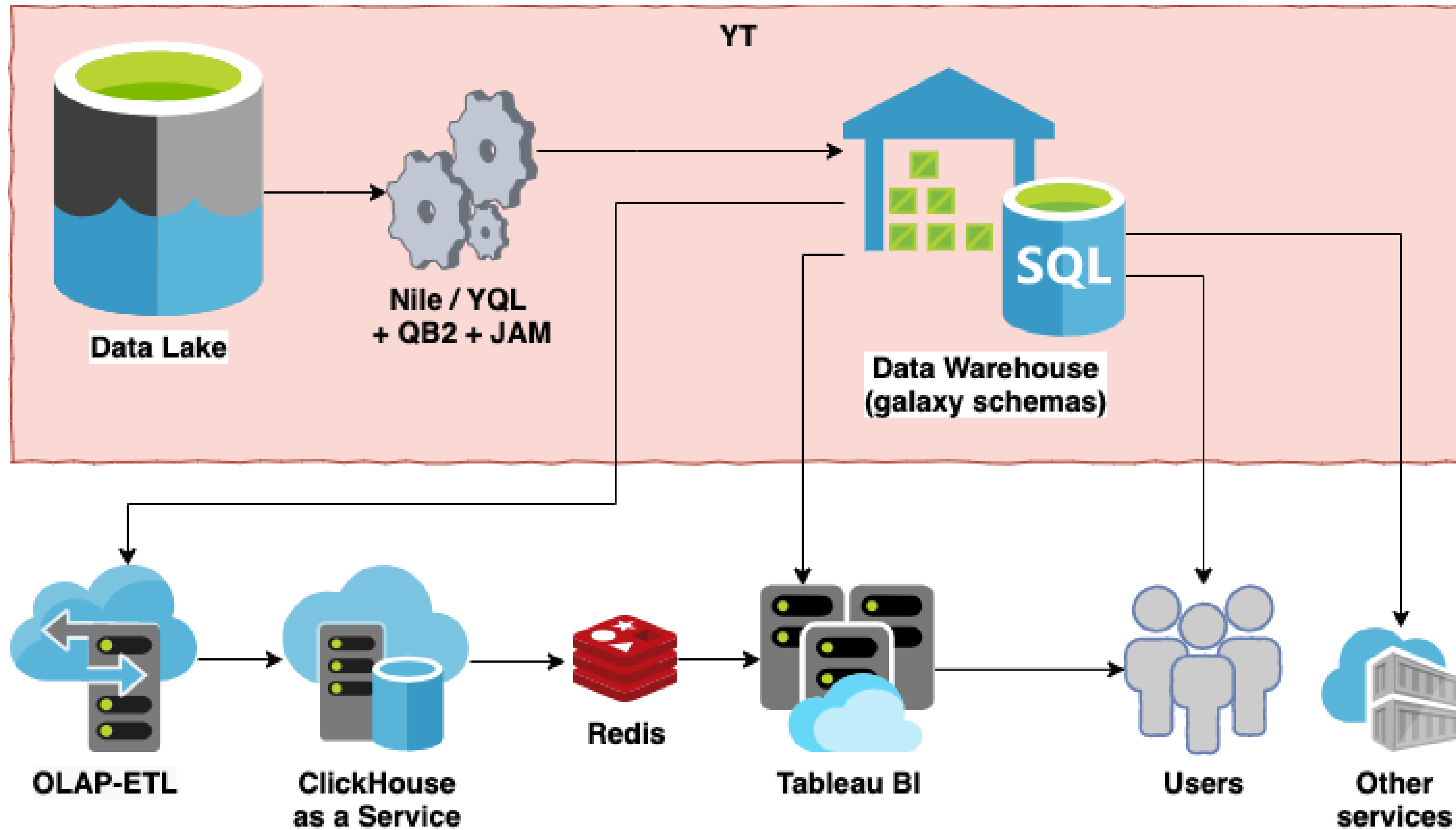


# Data Lake: текущие проблемы и планы

- › Добавление новой выгрузки – коммит в конфигури
  - прикрутить API и/или WebUI
- › Много дневных snapshot-выгрузок, пока не из всех баз данных поддерживается CDC
  - перейти на CDC и микробатчинг
- › Частенько что-то отваливается на стороне источников
  - отдать источникам задачу поставки своих данных в Data Lake, пользуясь нашей платформой

Data Warehouse (DWH) + Data Marts (DM)

# DWH + DM



# DWH на YT в цифрах

**50Tb+**

Объем исходных данных, обрабатываемых ежедневно

**7100**

Ядер в квоте на ежедневную подготовку данных и расчет отчетов

**70+**

Широких таблиц готовится для ClickHouse (~15Tб)

**5Pb+**

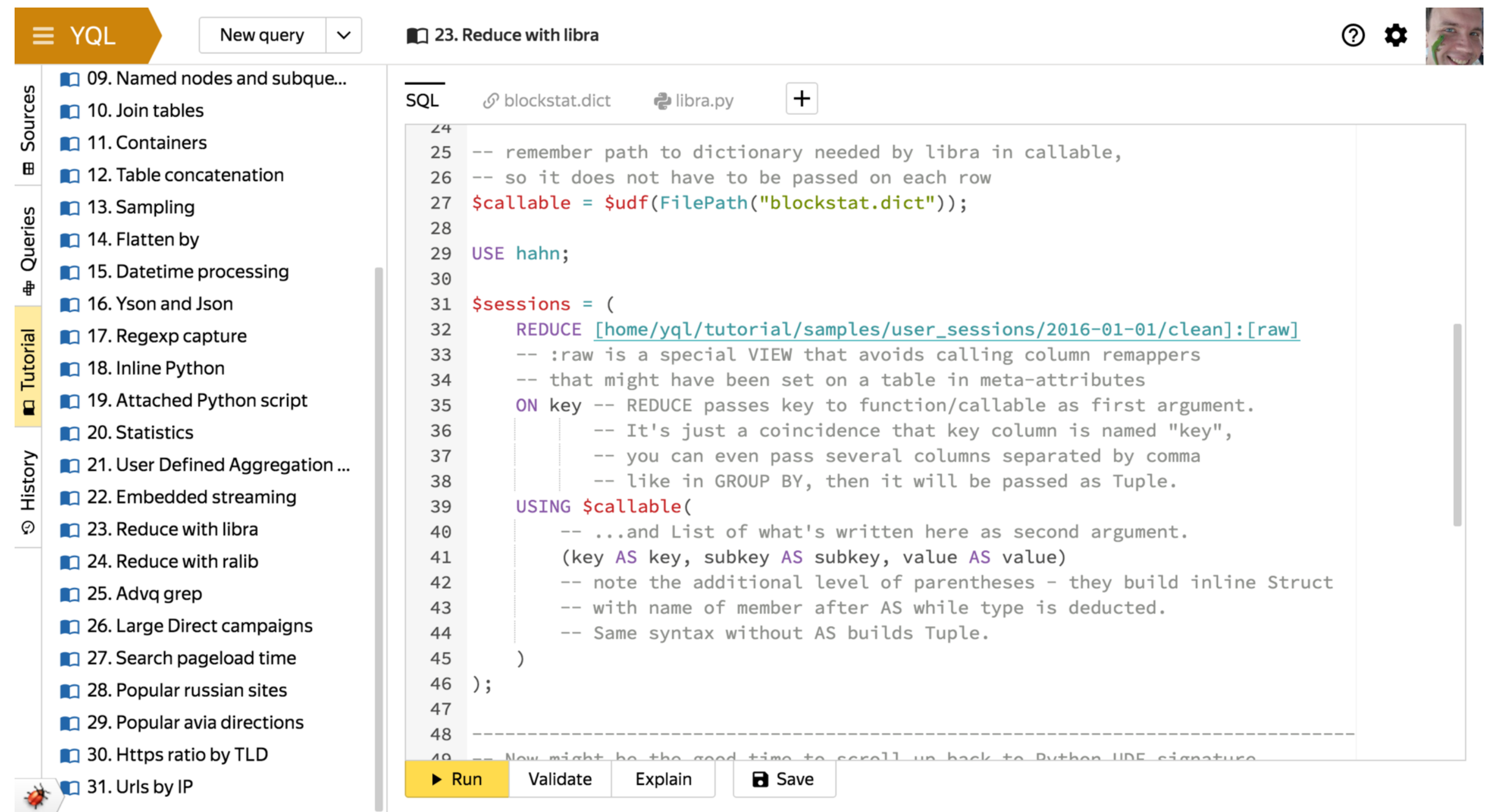
Общий объем на каждом из 2-х кластеров

# DWH: Почему YТ?!

- › Уже используется в маркете (индексация для поиска, ...)
- › Дорабатывает и поддерживает команда разработки YТ
- › Шаринг неутилизированных ресурсов
- › Хорошая масштабируемость на «обычном железе»
- › Можно не только MapReduce-ом обрабатывать (СН, Spark)

# YQL – Yandex Query Language

- Написан на C++
- Запускает MapReduce
- UDF на разных языках
- Провязан со многими внутренними инструментами (оркестрации, визуализации)
- Разрабатывается и поддерживается инфраструктурой Яндекса
- Подробнее в статье - <https://habr.com/ru/company/yandex/blog/312430/>



The screenshot displays the YQL web interface. On the left, a sidebar menu lists various tutorials and queries, including '09. Named nodes and subque...', '10. Join tables', '11. Containers', '12. Table concatenation', '13. Sampling', '14. Flatten by', '15. Datetime processing', '16. Yson and Json', '17. Regexp capture', '18. Inline Python', '19. Attached Python script', '20. Statistics', '21. User Defined Aggregation ...', '22. Embedded streaming', '23. Reduce with libra', '24. Reduce with ralib', '25. Advq grep', '26. Large Direct campaigns', '27. Search pageload time', '28. Popular russian sites', '29. Popular avia directions', '30. Https ratio by TLD', and '31. Urls by IP'. The main area shows a query editor for '23. Reduce with libra'. The query is written in SQL and includes a Python UDF named 'libra.py'. The query text is as follows:

```
24
25 -- remember path to dictionary needed by libra in callable,
26 -- so it does not have to be passed on each row
27 $callable = $udf(FilePath("blockstat.dict"));
28
29 USE hahn;
30
31 $sessions = (
32   REDUCE [home/yql/tutorial/samples/user_sessions/2016-01-01/clean]:[raw]
33   -- :raw is a special VIEW that avoids calling column remappers
34   -- that might have been set on a table in meta-attributes
35   ON key -- REDUCE passes key to function/callable as first argument.
36   -- It's just a coincidence that key column is named "key",
37   -- you can even pass several columns separated by comma
38   -- like in GROUP BY, then it will be passed as Tuple.
39   USING $callable(
40     -- ...and List of what's written here as second argument.
41     (key AS key, subkey AS subkey, value AS value)
42     -- note the additional level of parentheses - they build inline Struct
43     -- with name of member after AS while type is deducted.
44     -- Same syntax without AS builds Tuple.
45   )
46 );
47
48 -----
49 -- Now might be the good time to scroll up back to Python UDF signature
```

At the bottom of the editor, there are buttons for 'Run', 'Validate', 'Explain', and 'Save'.

# Nile+QB2

- Nile – аналог cascading, написанный на python для YT
- QB2 – правила извлечения информации (можно заменить на UDF). Запускается как mapper на YT.
- Могут выполняться на YQL-движке
- Разрабатывается и поддерживается инфраструктурой Яндекса
- Подробнее в статье - <https://habr.com/ru/company/yandex/blog/332688/>

```
stream = job.table('//yandex-app-log/2017-04-11')
users_stage = stream.qb2(
    log = 'yandex-app-log',
    fields = [
        'api_key', 'device_id', 'event_date',
        qbe.json_log_field('EventValue').rename('event_value'),
        qbe.dictitem('stage', from_='event_value'),
    ],
    filters = [
        qbf.defined('device_id'),
        qbf.equals('api_key', 10321),
    ]
)
result = users_stage \
    .groupby('event_date', 'stage') \
    .aggregate(users=na.count_distinct('device_id')) \
    .put('//tmp/result')

job.run()
```

# JAM – Jobs and Actions Manager

- Запуск произвольного бинаря по готовности данных на кластере (событийная обработка данных)
- Запуск произвольного бинаря по расписанию (cron)
- Высокоуровневый планировщик pipeline-а обработки данных



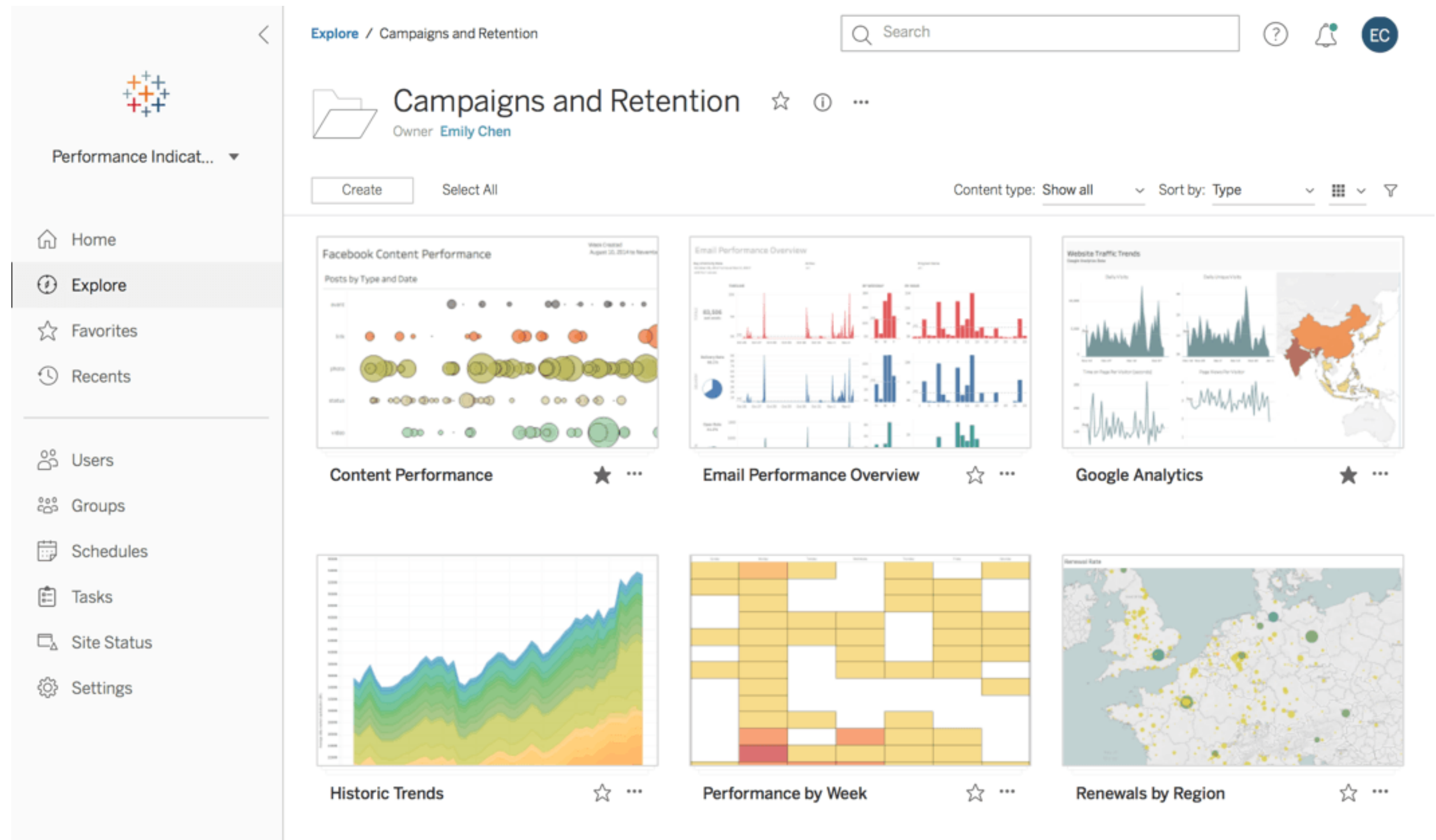
# DWH: текущие проблемы и планы

- › Много времени уходит на оптимизацию расчетов
  - высокоуровневые профилирующие инструменты
  - нагрузочное тестирование в релизных процессах
  - проектирование хранилища (Data Vault 2.0 ?)
- › Тестирование на больших данных
  - тестовая среда с семплированными данными
- › Много ручной работы в управлении метаданными
  - централизовать управление метаданными в единый сервис Data Catalog и провязать его со всей платформой

Reporting & Analyzing

# Tableau как основной BI инструмент

- Dashboards с бизнес-метриками
- Анализ данных аналитиками и руководителями



# BI-система в цифрах

**350+**

Уникальных  
пользователей

**220+**

Уникальных  
источника данных

**3000+**

Уникальных  
dashboards

**16000+**

Уникальных views

# BI: текущие проблемы и планы

- › Extract-ы перестали помещаться в памяти
  - перейти на широкие таблицы в ClickHouse
- › Легко сделать неоптимальный дашборд
  - вести обучение пользователей
  - ввести релизные процессы на критичные отчеты
- › Выгрузка больших данных из Tableau в Excel
  - сделать отдельные инструменты

Security

# Permissions

- › Аутентификация через TVM-токены
- › Авторизация через систему ролей в IDM  
(внутренний сервис в Яндексе для управления ролями)
- › ACL во всех используемых инструментах
  - в YT на уровне таблиц и колонок (row based не используется)
  - в BI на уровне Projects, Dashboards, Views

# Encryption

Сейчас прорабатываем следующую схему:

- › Чувствительные данные шифруются источником (на уровне конкретных колонок) и выгружаются в Data Lake
- › Все витрины содержат зашифрованные колонки (подготовка данных их не трогает)
- › Пользователь BI обладает ключом расшифровки на своей локальной машине (используя TabPy)



# Основные принципы

# Платформизация и девопсизация

- › Платформа – повторно используемые инструменты (не под конкретную задачу, а под класс задач)
- › Разработчики платформы сами ее поддерживают
- › Процветает НИН (not invented here)
  - свой продукт легче эксплуатировать, масштабировать и интегрировать в экосистему
  - цена инструмента на рынке часто завышена, дешевле разработать свой (конечно есть исключения)

# Эффективность

- › При большой нагрузке (HighLoad генерирует BigData) технические оптимизации приносят существенную экономию денег
- › Приборы (измерительные инструменты) дают возможность видеть, что сейчас хуже всего работает (сначала измерь, потом оптимизируй)
- › DMP позволяет создавать приборы для бизнеса

Вопросы

Группа компаний


**Я**ндекс Маркет

# Спасибо

**Денис Хуртин**

Руководитель разработки DMP

 [dkhurtin@yandex-team.ru](mailto:dkhurtin@yandex-team.ru)

 [lo4loo](https://t.me/lo4loo)