

“Горячая” миграция проекта с **20M** пользователями с **MySQL** на **PG**

Олег Правдин
CTO, LinguaLeo.com

Зачем ?

Снизить сроки разработки в 5+ раз

Снизить бюджет отдела разработки в 3+ раз

Снизить стоимость владения в 5+ раз

Закрыть доступ к базе через SQL-команды

Наконец-то задокументировать работу бэкенда

“Монолит”
на 1 млн. строк

400+ таблиц
100К шардов
Без описания

8 разработчиков
2 админа

Статус бэкенда

Доступ к базе
из сотен функций
“монолита”

1 - 2 месяца
на простую “фичу”

\$1000 / год
на 10,000
активных юзеров

Миграция как смена философии бэкенда

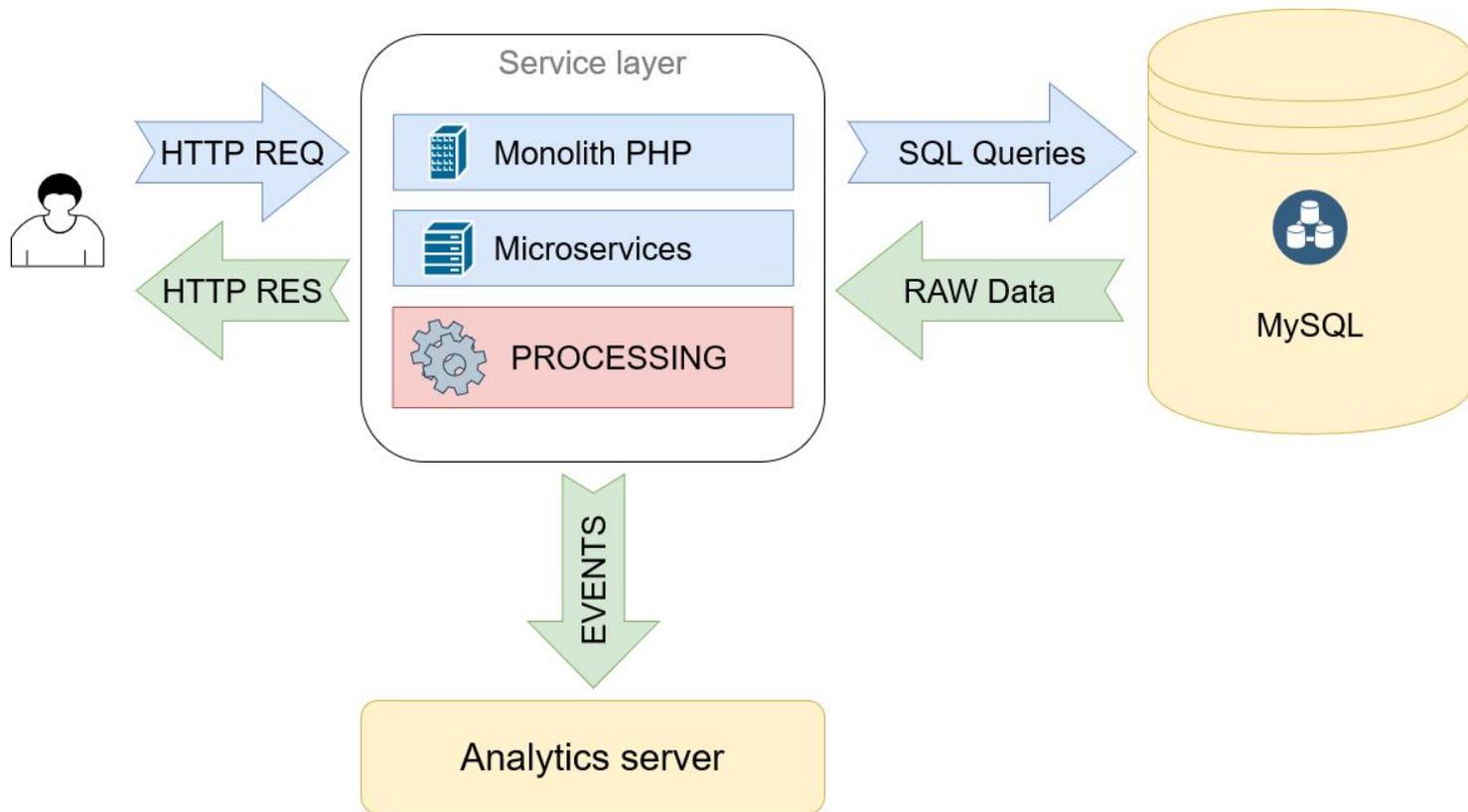
Перенос бизнес-логики
из “монолита”
в **базу данных**

SQL & **NoSQL**

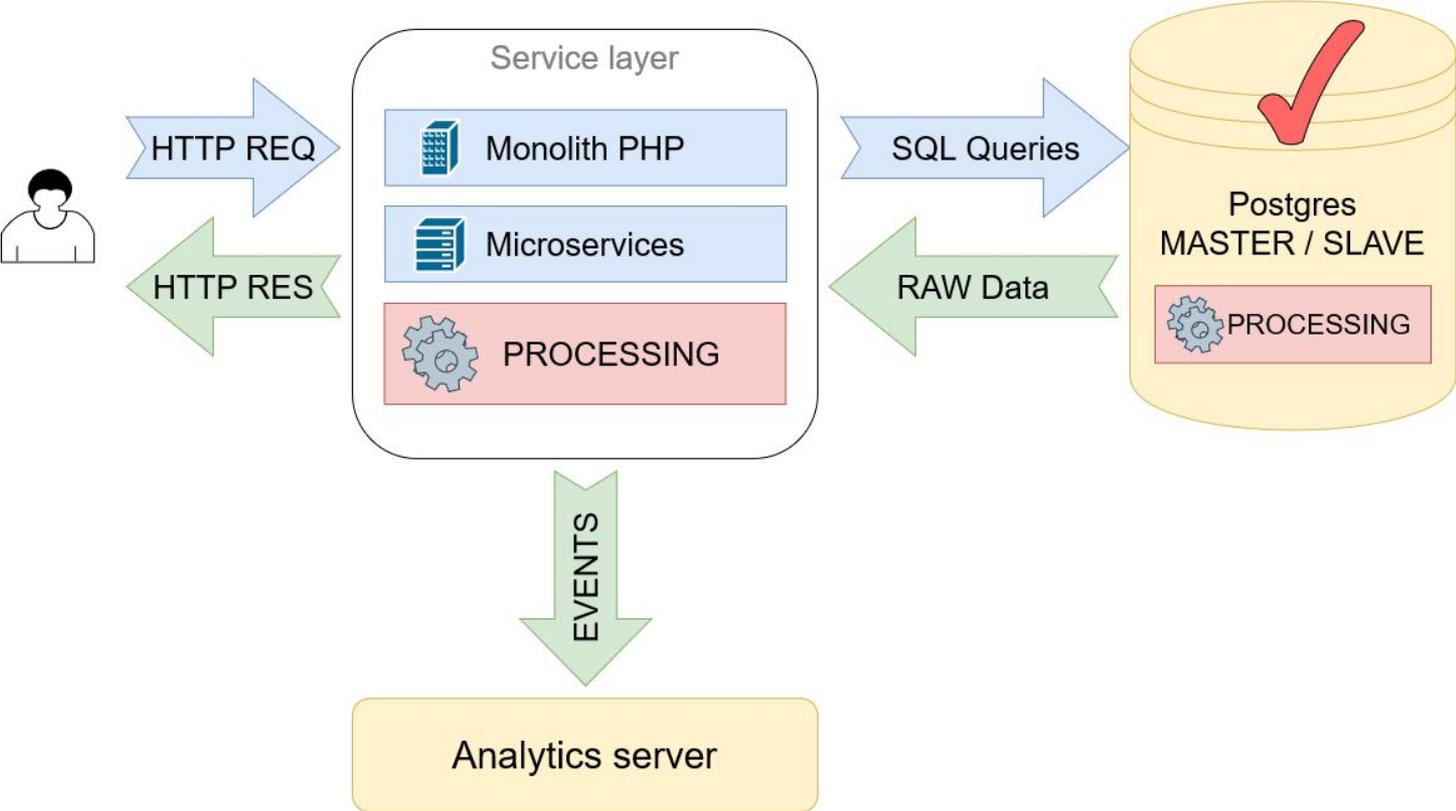
структура данных

Единая точка доступа
к базе данных
(“**proxy**” service)

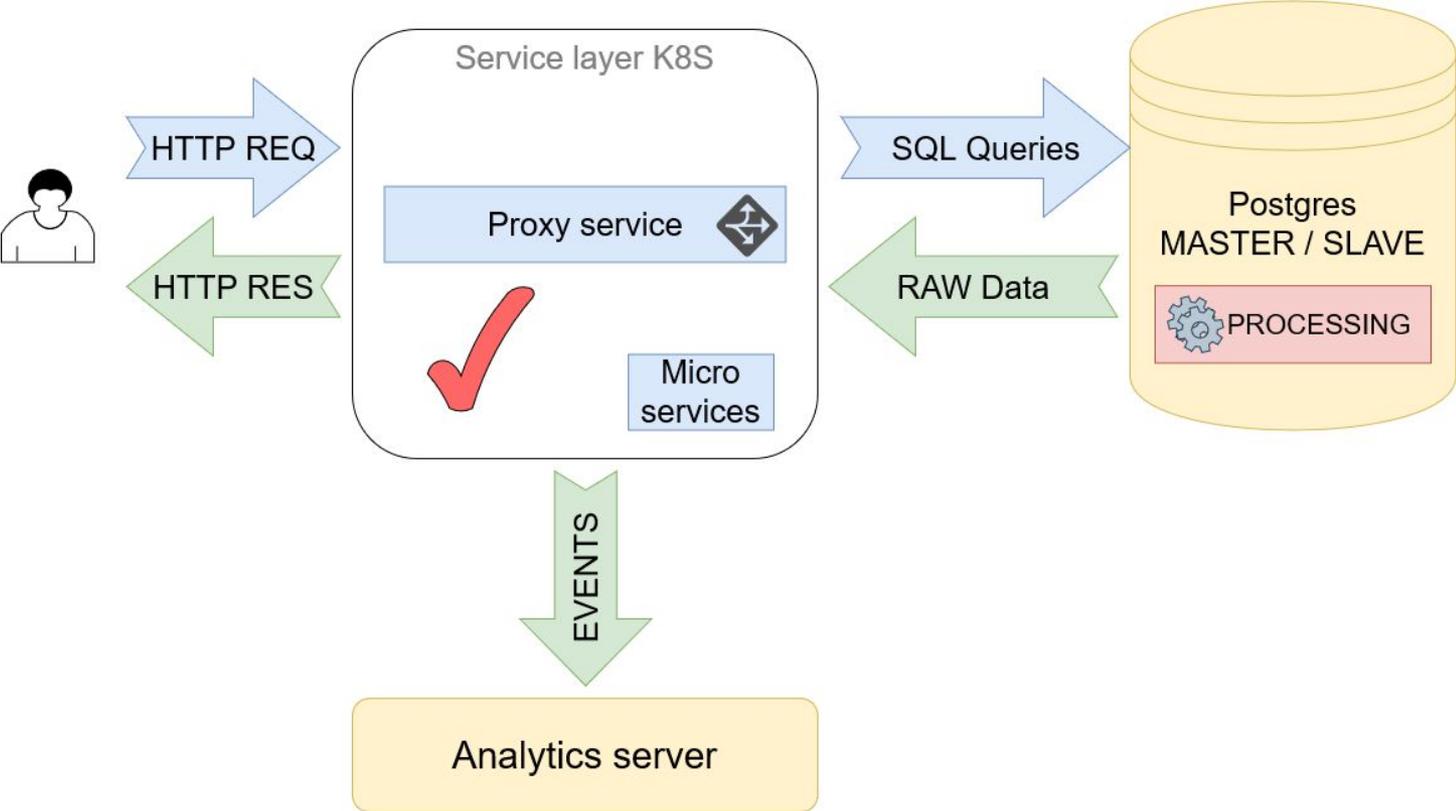
Архитектура до миграции



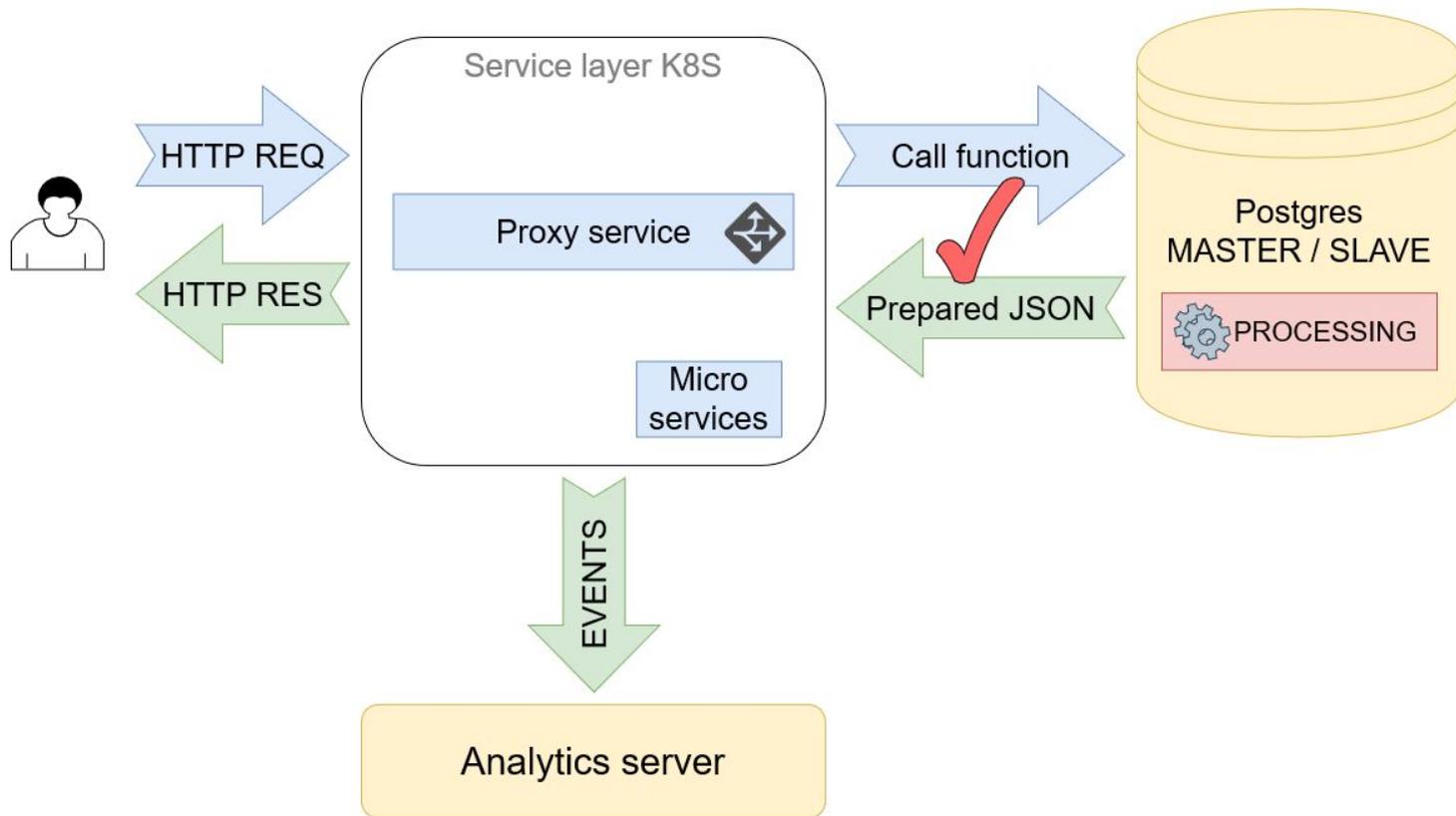
Перенос базы данных и бизнес-логики в PG



Монолит → микросервис



Архитектура после миграции



Данные и обработка данных в одном месте

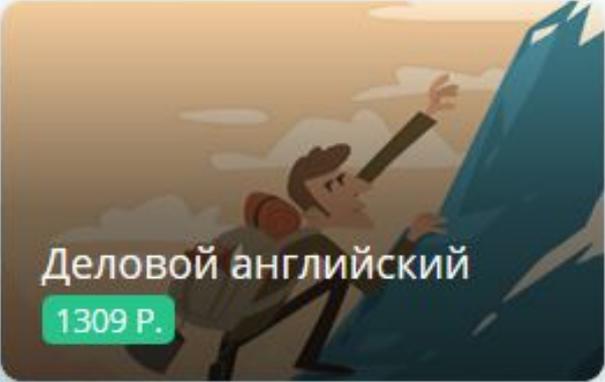
Основная бизнес-логика
в PG-хранимках

90+ % pgSQL-кода - это
SQL-запрос с CTEs

Любая хранимка:
json - вход => json-выход

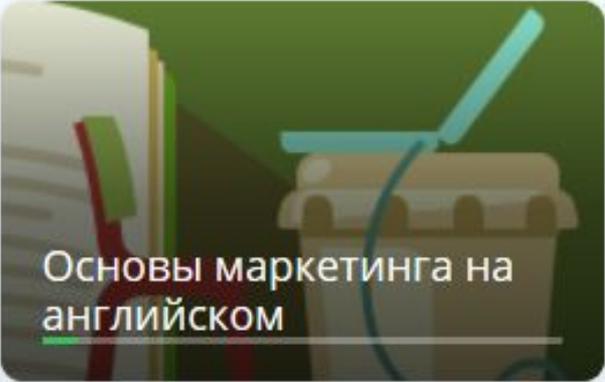
json-ответ включает
данные для аналитики

Веб-страница курсов английского языка



Деловой английский

1309 Р.

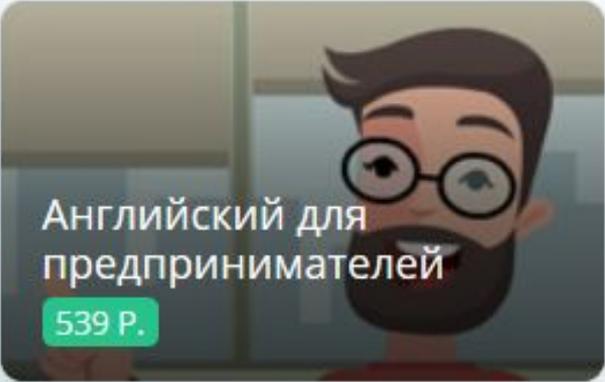


Основы маркетинга на английском



British Council: You are hired

БЕСПЛАТНО



Английский для предпринимателей

539 Р.

10+ критериев
по контенту и
пользователю

Бэкенд-код страницы курсов до миграции

Языки	PHP, SQL
Строк кода	10,000
SQL-запросов	12
Разработка (дней)	15
Выполнение (мсек)	600

Бэкенд-код страницы курсов после миграции

Языки	PHP, SQL	pgSQL
Строк кода	10,000	300
SQL-запросов	12	1
Разработка (дней)	15	2
Выполнение (мсек)	600	20

400+ SQL-таблиц ⇒ 8 таблиц с jsonb

100+ сущностей в одной таблице:

1. User ID
2. ID сущности
3. Ключ сущности
4. Date/Time
5. BigInt с битовыми масками
6. jsonb с любыми атрибутами

Primary key:

1, 2, 3

Index 1:

1, 2, 4, 5

Index 2:

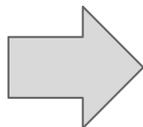
2, 4, 5

Типичный **SELECT** в хранимке

2-5 CTEs с

“**Index Only Scan**”

фильтрами



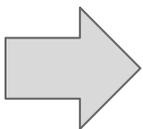
1+В строк => **10 .. 100 строк**

< 1 мсек

Обрабатываем

jsonb-атрибуты в

полученных строках



Быстрая выборка с
упаковкой в json

Стабильная
структура данных

Гибкость
разработки

Низкие сроки
разработки

Компактная
документация

Выше
качество

TOP-10
benefits

Простое
администрирование

Отличная
производительность

Без внешних
SQL-команд

Готовые данные
для аналитики

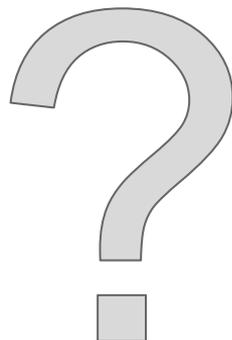
Балансировка
нагрузки

Что получилось

Код бэкенда:
новый

Код фронта:
без изменений

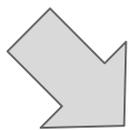
Данные:
400 => 8 таблиц



Как перенести данные
по 20М пользователям
без остановки бизнеса

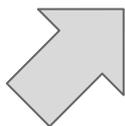
“Горячий” перенос данных в новую структуру

1. Данные авторизации

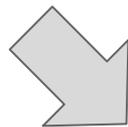


2. Пользователь авторизуется

3. Краткое приветствие



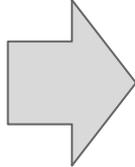
4. Перенос данных из MySQL в PG



Данные в новом бэкенде



Вариант быстрого переноса данных

1. Копия структуры MySQL-таблиц в PG
 2. Go сервис с подключением к MySQL и PG
 3. Хранимка для подготовки запросов к MySQL и PG
- 
1. SELECT к MySQL
INSERT к PG
 2. Конвертация данных в новый формат
 3. Для ускорения импорта запускали 8 потоков

Итоги миграции

Дней на разработку	30-60	5-10
Разработчики и админы	8 2	3 0
Стоимость владения	\$1000 / 10,000 и.	\$1000 / 100,000 и.
Безопасность	внешние SQL	только хранимки
Документация	так и не сделали(структура & API

Выводы

PostgreSQL ≈ Oracle

> **99%** бэкендов можно реализовать через PG-хранимки

Структура данных = не более **10 простых таблиц**

Для быстрой и качественной разработки достаточно отличного знания pgSQL

Возможности для **PG community**:

1. Миграция на PG окупается за 3 - 6 месяцев
2. Один PG expert может вести 2 - 5 проектов
3. Работа компании не зависит от прихотей разработчиков
4. Со стартапами: опцион за разработку бэкенда

Давайте проверим это вместе))

Спасибо!

Контактная информация:

Олег Правдин
СТО, LinguaLeo

o.pravdin@lingualeo.com

www.lingualeo.com

