

# МИГРАЦИЯ С ПЛАТФОРМЫ EXADATA С ОБЕСПЕЧЕНИЕМ ПОКАЗАТЕЛЕЙ SLA:

ЦЕЛЕВАЯ АРХИТЕКТУРА И РАСПРЕДЕЛЕНИЕ  
АНАЛИТИЧЕСКИХ НАГРУЗОК МЕЖДУ GREENPLUM,  
CLICKHOUSE И DATA LAKE

12 МАРТА 2026

БАЛЕВ АЛЕКСАНДР

# OTP GROUP

75 лет на рынке

41 тыс. сотрудников

17 млн. клиентов

12 стран присутствия

# OTP RUSSIA

57 тыс. точек потребительского кредитования

6,5 тыс. сотрудников

2,1 млн. клиентов

82 отделения

# ТЕКУЩАЯ ИНФРАСТРУКТУРА ДАННЫХ



## Функциональные области хранилища на базе Oracle:

- Подготовка ad-hoc аналитики и регуляторной отчетности
- Формирование витрин данных и управленческой отчетности
- Обеспечение интеграционных потоков с внешними информационными системами

*Описываемая платформа является критически важным компонентом ИТ-инфраструктуры в промышленной эксплуатации.*

## 100 ТБ

Объем данных

*Общий объем хранилища*

## 1000+

Пользователи

*Активных пользователей системы*

## T-1

Актуальность

*Частота обновления витрин*

## <92%

SLA

*Стабильность обновления данных в витринах*

# ОБОСНОВАНИЕ МИГРАЦИИ

*Реализация проекта по импортозамещению стала не только ответом на внешние ограничения, но и возможностью для комплексной архитектурной оптимизации информационных систем компании.*

## **Минимизация стратегических рисков**

*Снижение зависимости от проприетарного программного обеспечения и инфраструктуры зарубежных вендоров.*

## **Оптимизация совокупной стоимости владения (ТСО)**

*Снижение затрат на лицензирование и эксплуатацию, вызванных нелинейным ростом стоимости инфраструктуры при увеличении объёмов данных.*

## **Масштабируемость инфраструктуры**

*Переход на архитектуру, поддерживающую эффективное горизонтальное масштабирование ресурсов в условиях растущей нагрузки.*

## **Повышение скорости разработки (Time-to-Data)**

*Сокращение цикла вывода новых аналитических инструментов и витрин данных на промышленную эксплуатацию.*

# ОТКАЗ ОТ СТРАТЕГИИ ПРЯМОЙ МИГРАЦИИ (1:1)



1

## Неэффективность прямой миграции

*Перенос хранилища методом «lift-and-shift» приводит к сохранению текущих архитектурных ограничений в новой среде.*

2

## Эволюция профиля нагрузки

*Увеличение объемов данных и новые требования к обновлению данных (Intraday) требуют адаптации к изменившимся условиям эксплуатации.*

3

## Формирование современной платформы данных

*Переход к архитектуре многокомпонентных систем данных вместо использования монолитной СУБД является целевым решением для оптимизации бизнес-процессов.*

# ПРИНЦИПЫ ЦЕЛЕВОЙ АРХИТЕКТУРЫ



## Декомпозиция уровней хранения и вычислений

*Реализация принципа Separation of Storage and Compute для обеспечения независимого масштабирования ресурсов.*



## Технологическая дифференциация

*Подбор инструментов в зависимости от специфики бизнес-задач и профиля нагрузки, исключая привязку к конкретным вендорам.*



## Отказоустойчивость и масштабируемость

*Обеспечение горизонтального масштабирования системы без внесения архитектурных ограничений.*

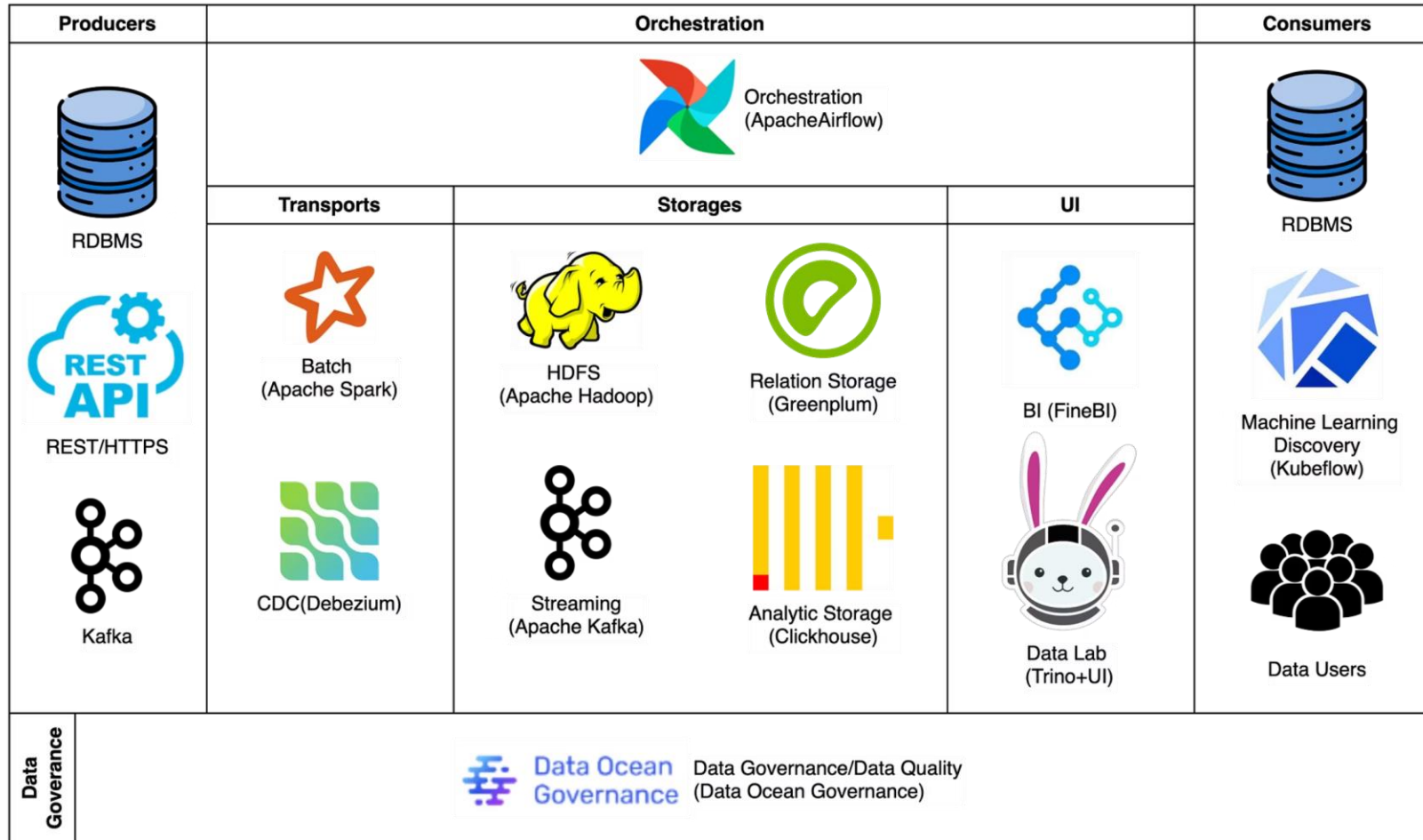


## Централизованное управление ETL

*Внедрение оркестрации с обеспечением полной прозрачности и прослеживаемости данных (Data Lineage).*

# ЦЕЛЕВОЙ ТЕХНОЛОГИЧЕСКИЙ СТЕК

Наше архитектурное решение построено на трехслойной модели обработки данных, обеспечивающей эффективный поток от источников до конечных потребителей.



# ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ КОМПОНЕНТОВ ЦЕЛЕВОЙ АРХИТЕКТУРЫ

Сценарий использования	Технологическое решение	Технические преимущества
Долгосрочное хранение первичных данных	Data Lake	Оптимизация затрат на хранение, горизонтальная масштабируемость
Трансформация данных и формирование корпоративного слоя	Greenplum	Применение MPP SQL для комплексной аналитики и построения витрин данных
Оперативная аналитическая отчетность	ClickHouse	Колоночное хранение, высокая производительность при параллельных запросах
Оркестрация и управление регламентами	Apache Airflow	Использование DAG-шаблонов, обеспечение прозрачности и мониторинга пайплайнов

- Архитектурное разделение задач между специализированными системами предотвращает избыточную нагрузку на отдельные компоненты и обеспечивает оптимальную производительность каждой подсистемы.

# МЕТОДОЛОГИЯ НЕПРЕРЫВНОГО ПЕРЕХОДА



Реализация проекта осуществлялась посредством поэтапного перехода, что позволило избежать рисков, связанных с единовременным внедрением (Big Bang).

Основными условиями достижения целевых показателей стали параллельная эксплуатация систем и непрерывный мониторинг качества данных на каждом этапе интеграции.

# КЛЮЧЕВЫЕ ТЕХНОЛОГИЧЕСКИЕ И ОРГАНИЗАЦИОННЫЕ ВЫЗОВЫ

## Миграция PL/SQL-логики

Перенос хранимых процедур потребовал значительных трудозатрат ввиду необходимости ручной адаптации кода.

## Оптимизация и производительность

Различия в архитектуре оптимизаторов и особенностях целевых систем потребовали дополнительной доработки процессов.

## Контроль качества данных

Обеспечение сопоставимости результатов между legacy-системой и целевой платформой стало критическим фактором успеха.

## Развитие компетенций команды

Процесс обучения сотрудников новым технологиям осуществлялся параллельно с выполнением миграционных задач.

## Управление ожиданиями стейкхолдеров

Выстраивание прозрачной коммуникации с заинтересованными сторонами потребовало системного подхода на всех этапах проекта.

# КЛЮЧЕВЫЕ МЕТРИКИ ВЛИЯНИЯ ПРОЕКТА

## 3-7 дн.

**Сокращение Time-to-Data**

*Снижение времени готовности данных для анализа с 4–6 недель до до 3–7 дней*

## 70%

**Сокращение времени разработки**

*Снижение затрат времени на разработку аналитических решений за счет использования шаблонизации*

## 46%

**Сокращение затрат на хранение**

*Снижение совокупных затрат на хранение данных благодаря оптимизации архитектуры*

## T-1 → H-1

**Сокращение времени обновления данных**

*Обновление данных в течение часа после их получения вместо одного дня*

## 92% → 95%

**Улучшение выполнения SLA**

*Повышение уровня соблюдения SLA по обновлению ключевых витрин*

# ИТОГИ ПРОЕКТА



## Реализация миграции данных

*До конца 2026 будет закончен полный переход с СУБД Oracle на open source решения без прерывания операционной деятельности.*



## Оптимизация расходов

*Снижение совокупной стоимости владения (TCO) на 46% вследствие реструктуризации лицензионной политики и инфраструктурных затрат.*



## Повышение производительности

*Сокращение времени обработки данных и формирования отчетных витрин в 2-3 раза.*



## Масштабируемость инфраструктуры

*Развернута отказоустойчивая архитектурная модель, адаптированная к перспективному масштабированию и текущим бизнес-требованиям.*



## Внедрение процессов Data Mesh & Data Governance

*Внедрена децентрализованная модель управления данными (Data Mesh) с четкими принципами Data Governance, обеспечивающей прозрачность, качество и безопасность данных на всех уровнях организации.*

# СПАСИБО!