



**Крупнейшее
импортозамещение
Озера данных в России**



ОПИСАНИЕ ПРОЕКТА



Миграция аналитической платформы
Озеро данных на отечественную
платформу Arenadata Hadoop



РАЗМЕР ИМЕЕТ ЗНАЧЕНИЕ



ОБЪЕМ ДАННЫХ

> 6 Петабайт

> 1 000 пользователей

ПРИНЯЛИ УЧАСТИЕ

130 профильных команд

> 1 700 человек

Суммарно в проект инвестировано 338 400 чел. часов

ДОСТИГНУТЫЕ ЦЕЛИ



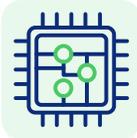
Обеспечено технологическое развитие аналитической платформы Банка без ограничений, связанных с legacy платформой Oracle BDA (Cloudera Hadoop) и другими иностранными компонентами



Реализована геораспределенная архитектура DR-решения высокой надежности



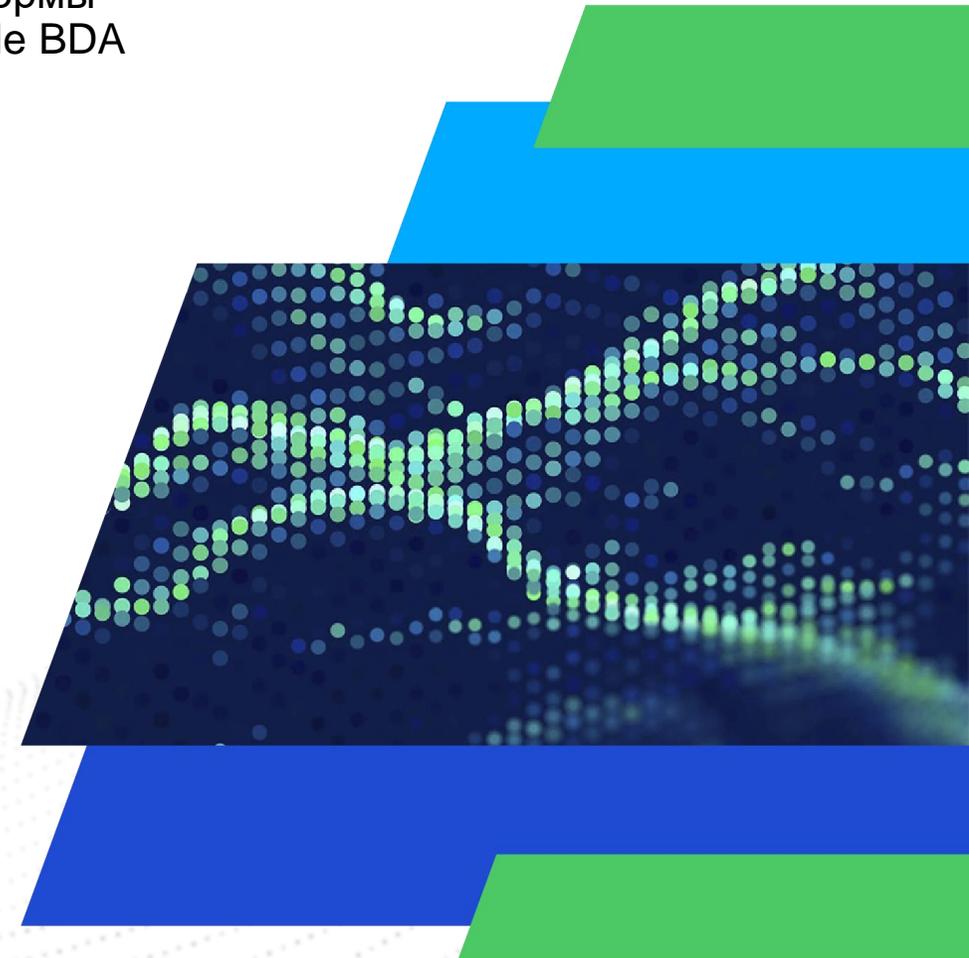
Сокращен T2M ИТ-сервисов для бизнеса за счет унификации и стандартизации процессов разработки и внедрения единых DevOps практик



Увеличены вычислительные мощности пользовательского кластера более чем в два раза



Достигнут технологический суверенитет



ИСПОЛЬЗОВАН СТЕК ТЕХНОЛОГИЙ



В рамках импортозамещения Банком разработаны ключевые компоненты решения:
ETL Framework и Hadoop Sync

Arenadata Hadoop
Spark
Hive
Zeppelin
Airflow
RabbitMQ
PostgreSQL



FluentD
Grafana
Kibana
Elastic search
ETL framework DDF
(собственная разработка)
Платформа репликации
Hadoop Sync
(собственная разработка)

OpenShift
DBeaver
RStudio
Jupyter Hub
ПАК Скала-ср/АДХ
Частное облако ВТБ

ДОСТИГНУТЫЕ РЕЗУЛЬТАТЫ



01

Бизнес-пользователям доступны данные из более 100 информационных систем-источников данных

02

Мигрированы данные, необходимые для моделей искусственного интеллекта и продвинутой аналитики

03

Обеспечена поддержка ключевых бизнес-процессов управленческой и регуляторной отчетности, а также разработки моделей продвинутой аналитики

04

Мигрирован аналитический функционал, включая более 300 витрин данных и 100+ пользовательских БД (песочниц) для различных направлений бизнеса

05

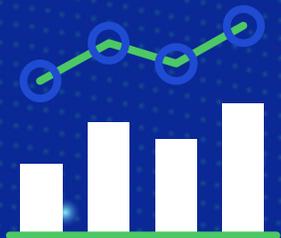
Увеличены вычислительные мощности пользовательского кластера более чем в два раза



СДЕЛАНО В РОССИИ!



Миграция Озера данных на отечественную платформу позволила ВТБ получить ряд **технологических и бизнес-преимуществ:**



Бизнес-пользователям доступны данные из более 100 информационных систем-источников, необходимые для моделей искусственного интеллекта, анализа, отчетности и продвинутой аналитики



Реализована геораспределенная архитектура DR-решения высокой надежности

100%

ТЕХНОЛОГИЧЕСКИЙ СУВЕРЕНИТЕТ