



Корпоративная платформа
хранения и обработки данных

КТО МЫ

ARENADATA — ведущий российский разработчик платформ сбора и хранения данных для компаний, которые хотят изменить бизнес с помощью цифровизации процессов и сложной аналитики, основанной на использовании Big Data.

Наши решения строятся
на технологиях
с открытым кодом
и помогают работать с



Большие
данные



Интернет
вещей



Машинное
обучение



Связанные
технологии

Факты об ARENADATA

Специалисты из топовых data-driven компаний:



20+

Лидеров РФ выбрали нас
в качестве центральной
платформы данных

25 ПБ

Суммарный объём
данных, который наши
клиенты хранят в
платформе Arenadata



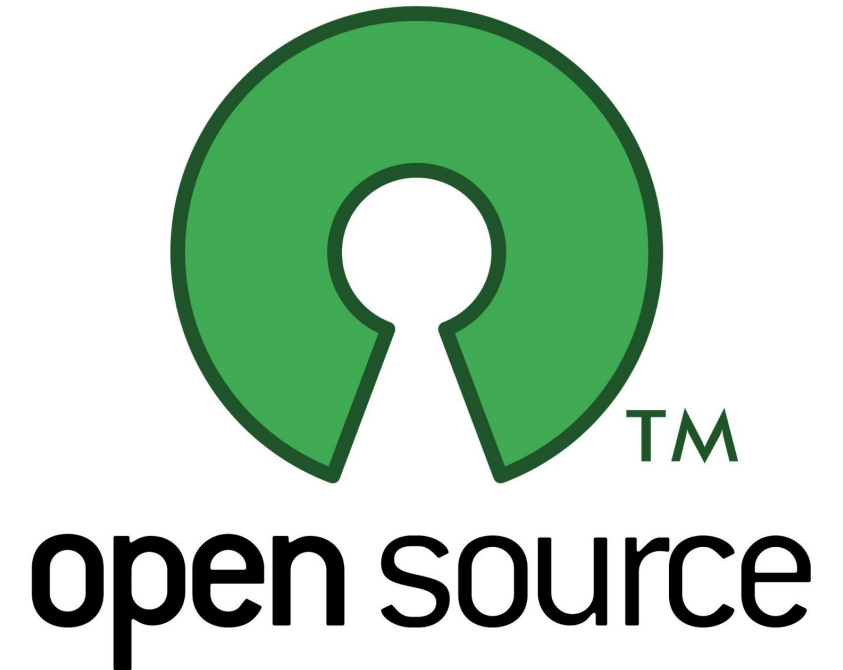
Развитая сеть партнёров
как в России, так и за
рубежом



Партнёрство
с ведущими вендорами
data-решений

Открытый ИСХОДНЫЙ КОД

- Вносим вклад в проекты Apache Software Foundation:
 - Apache Ambari
 - Apache Bigtop
 - Apache PXF
- 2^{ой} в мире контрибьютор Greenplum Database
- Активный контрибьютор ClickHouse и Tarantool
- Выпускаем **собственное ПО** в Community-редакциях



Российское программное обеспечение

Регистрация в
Реестре
Министерства
Связи

Отсутствие
валютных рисков -
цены в рублях

Центры разработки
в РФ - Москва,
Санкт-Петербург,
Хабаровск

Поддержка,
документация и
учебные курсы на
русском языке



Платформа для цифровой компании

Экспертиза зарекомендовавших себя партнёров-интеграторов, бизнес-кейсы и прикладное ПО, доказавшее свою эффективность

Оцифрованные бизнес-процессы

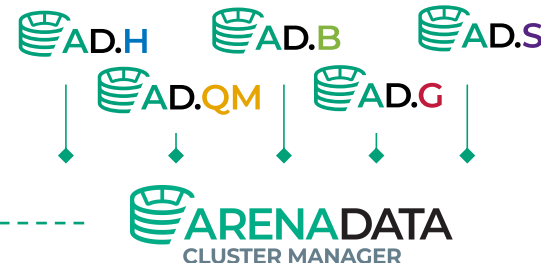


Множество инструментов для решения прикладных задач



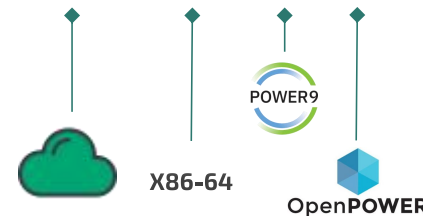
Arenadata Enterprise Data Platform с нативной интеграцией сервисов и компонент, централизованным мониторингом и управлением кластерами

Платформа данных



Любое ПО на любой инфраструктуре: эластично, отказоустойчиво, под любые реалии и задачи

Эластичная инфраструктура



Решаемые задачи



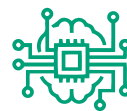
Корпоративные хранилища больших данных

- Классические КХД и хранилища больших данных с обновлением в реальном и околореальном времени.
- Виртуализация и федерация данных, доступ ко всем данным через единое окно с помощью SQL-запросов.



Разгрузка дорогостоящих инженерных систем (Teradata, Exadata)

- Импортзамещающие ПАК на базе массово-параллельных СУБД с открытым исходным кодом Arenadata DB и Arenadata QuickMarts.
- Прозрачная методика перехода, минимум рисков и сохранение всех преимуществ.



Машинное обучение и искусственный интеллект

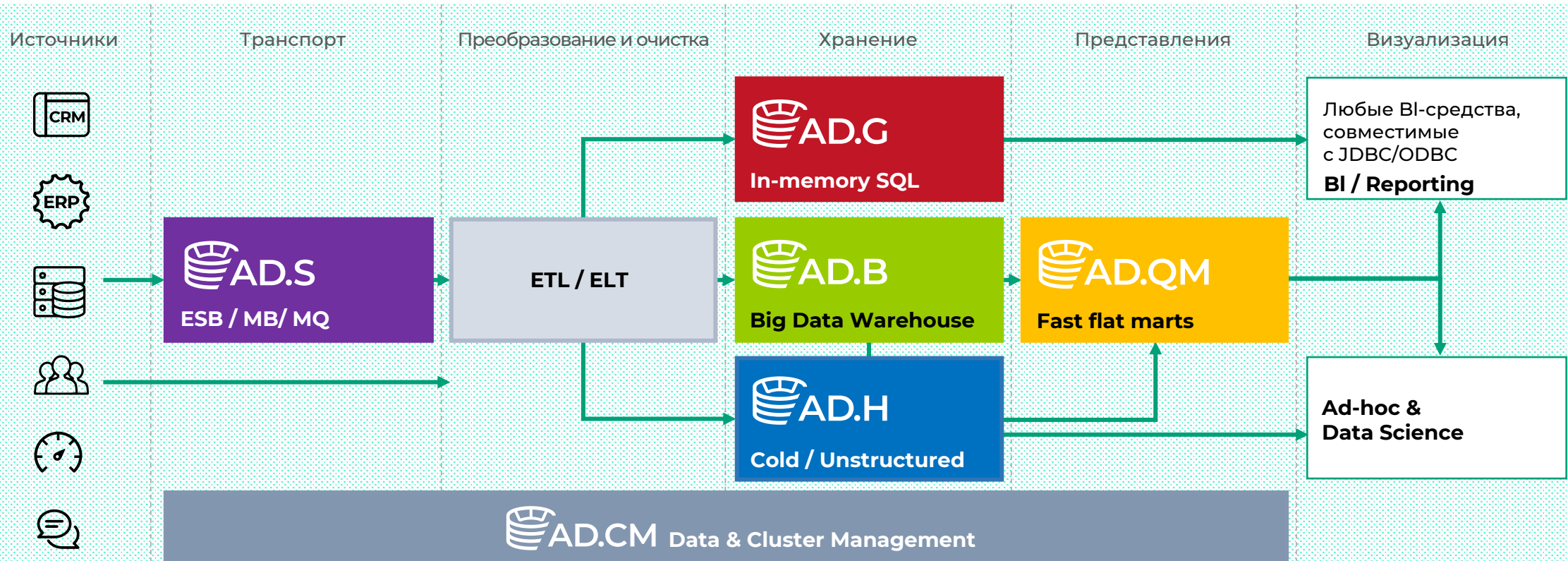
- Поддержка R, Python и Madlib.
- Distributed ML на базе Spark.
- Возможность эффективной эксплуатации моделей в SQL-среде с помощью встроенных функций Madlib.



Озёра и фабрики данных

- Единый хаб всех данных компании.
- Возможность быстрого развёртывания и сворачивания зон «песочниц» для пилотных проектов и проверки статистических гипотез.
- Работа со всеми аналитическими инструментами в единой среде.

Arenadata Enterprise Data Platform



On-premise

PAAS

Public Cloud

Private Cloud

Что внутри?

 AD.B Big Data Warehouse	 Greenplum
 AD.H Cold / Unstructured	 hadoop
 AD.QM Fast flat marts	 ClickHouse
 AD.G In-memory SQL	TARANTOO
 AD.S ESB / MB/ MQ	 kafka  nifi
 AD.CM Data & Cluster Management	 ANSIBLE

Каждый добавляемый компонент платформы доработан, чтобы стать её частью.

Производится:

- Анализ кода
- Интеграция с модулем мониторинга
- Проверки отказоустойчивости
- Подбор компетентного персонала поддержки
- Доработка возможностей параллельной интеграции
- Множественные прочие проверки и доработки

Компоненты платформы EDP

 AD.B	 AD.H	 AD.QM	 AD.G	 AD.S
Хранилище данных (MPP-СУБД)	Хранение холодных и слабо структурированных данных	Витрины данных	In-Memory SQL	Streaming (ESB / MB / MQ)
Arenadata DB (ADB) — аналитическая, распределённая СУБД с открытым исходным кодом, использующая концепцию MPP (massively parallel processing), построенная на базе СУБД Greenplum. Она реализована на кластере из множества серверов и предназначена для хранения и обработки больших объёмов данных — до десятков петабайт.	Arenadata Hadoop (ADH) — это корпоративный дистрибутив на базе Apache Hadoop, предназначенный для хранения и обработки слабоструктурированных и неструктурированных данных.	Arenadata QuickMarts (ADQM) — кластерная колоночная система управления базами данных, созданная на основе Yandex ClickHouse. С её помощью вы можете в режиме реального времени генерировать аналитические отчёты разного плана, используя большие объёмы информации, хранящейся в плоских витринах. ADQM многократно быстрее традиционных СУБД.	Arenadata Grid (ADG) — система распределённых вычислений в оперативной памяти, которая помогает реализовать real-time сценарии обработки данных в составе единой корпоративной платформы данных. ADG создана совместно с компанией Mail.Ru на базе технологии Tarantool.	Arenadata Streaming (ADS) — масштабируемая отказоустойчивая система для потоковой обработки данных в режиме реального времени, адаптированная для корпоративного использования и построенная на базе Apache Kafka и Apache Nifi.

Задачи для каждого компонента

 AD.B	 AD.H	 AD.QM	 AD.G	 AD.S
Хранилище данных (MPP-СУБД)	Хранение холодных и слабо структурированных данных	Витрины данных	In-Memory SQL	Streaming (ESB / MB / MQ)
Arenadata DB лучше всего справляется с: - Сложными запросами, обрабатывающими большие объёмы, в том числе содержащими сложные аналитические функции - ETL/ELT - Эффективным соединением больших таблиц - Работой с индексами - Data Science - Аналитическими функциями на процедурных языках - Ad-hoc аналитикой	Построение распределённой платформы хранения на базе Apache Hadoop, адаптированный для корпоративного использования для холодных и слабо структурированных данных, которые хранить на MPP-СУБД или проприетарных СУБД было бы слишком дорого.	Как использовать: - В связке с Arenadata DB / Greenplum - Аналитика над структурированными логами и событиями - Быстрые витрины - Распределённые вычисления для бизнеса и науки - Анализ серверных логов	Система класса решений In-Memory Data Grid (IMDG). Она позволяет построить распределённый кэш в памяти и добавить туда функционал базы данных. При этом пользователю предоставляется возможность крайне быстро получать ответы на запросы по точечным фильтрам, буквально в течение миллисекунды.	Используется как: - Корпоративная шина обмена данными для всех приложений - Среда консолидации больших потоков данных и управления ими в режиме реального времени - Инструмент разграничения прав доступа к потокам данных - Фреймворк для разработки потоковых аналитических приложений за считанные минуты в режиме реального времени и без единой строчки кода.

Преимущества Arenadata EDP



Хранение любых типов данных:

- структурированных
- неструктурированных



Обработка любого типа нагрузки на чтение и запись:

- OLAP
- OLTP
- Streaming



Поддержка гибридной ИТ-инфраструктуры или Multi-clouds



Единая точка входа в платформу для большей части пользователей



Модульность



Горизонтальная масштабируемость всех компонент



Open Source



ВОЗМОЖНОСТИ ПО ИНТЕГРАЦИИ

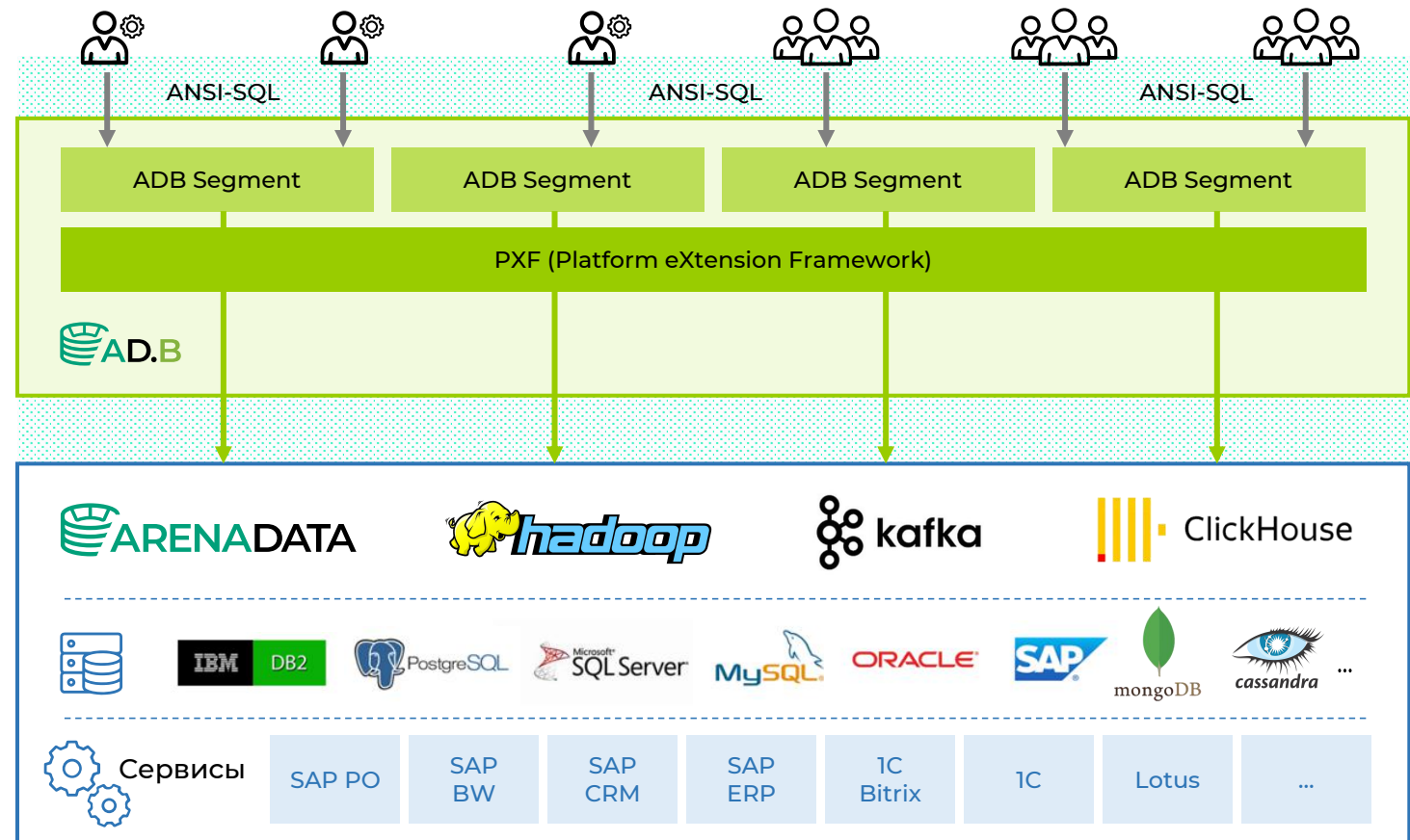


Platform
Extension
Framework
(PXF)

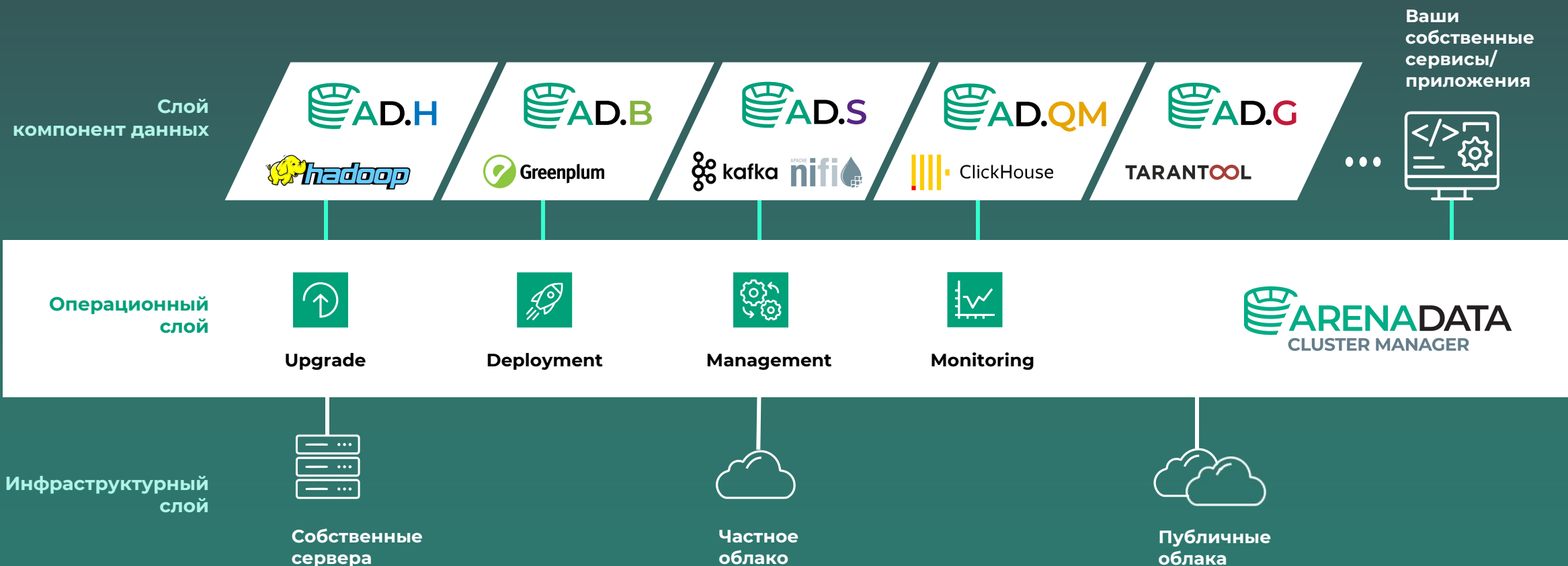
Виртуальное хранилище

Возможности Виртуального хранилища на базе Arenadata Database (ADB):

- Поддержка работы с любыми JDBC-совместимыми СУБД
- Параллельная запись во внешние СУБД (Postgres, Oracle, MS SQL, SAP HANA и пр.)
- Передача предикатов (Predicate Pushdown)
- Батчинг данных при передаче
- Соединение (JOIN) данных из источников в real-time ресурсами ADB
- Ролевая модель, разграничение прав доступа
- Запуск Python-кода поверх всех данных всех источников
- Запуск предварительного кода во внешнем источнике (pre-SQL) *скоро*



Гибридное корпоративное облако



Arenadata Cluster Manager (ADCM) - универсальный оркестратор гибридного ландшафта. Позволяет быстро устанавливать, настраивать все data-сервисы компании и управлять ими независимо от инфраструктуры.

Возможности работы с сервисами:

- Установка и настройка
- Обновление
- Управление
- Мониторинг
- Настройка прав доступа
- Интеграция с другими сервисами
- Контейнеризация

Возможности работы с инфраструктурой:

- Создание и удаление виртуальных машин
- Конфигурирование ОС
- Мониторинг
- Управление пользователями
- Настройка прав доступа

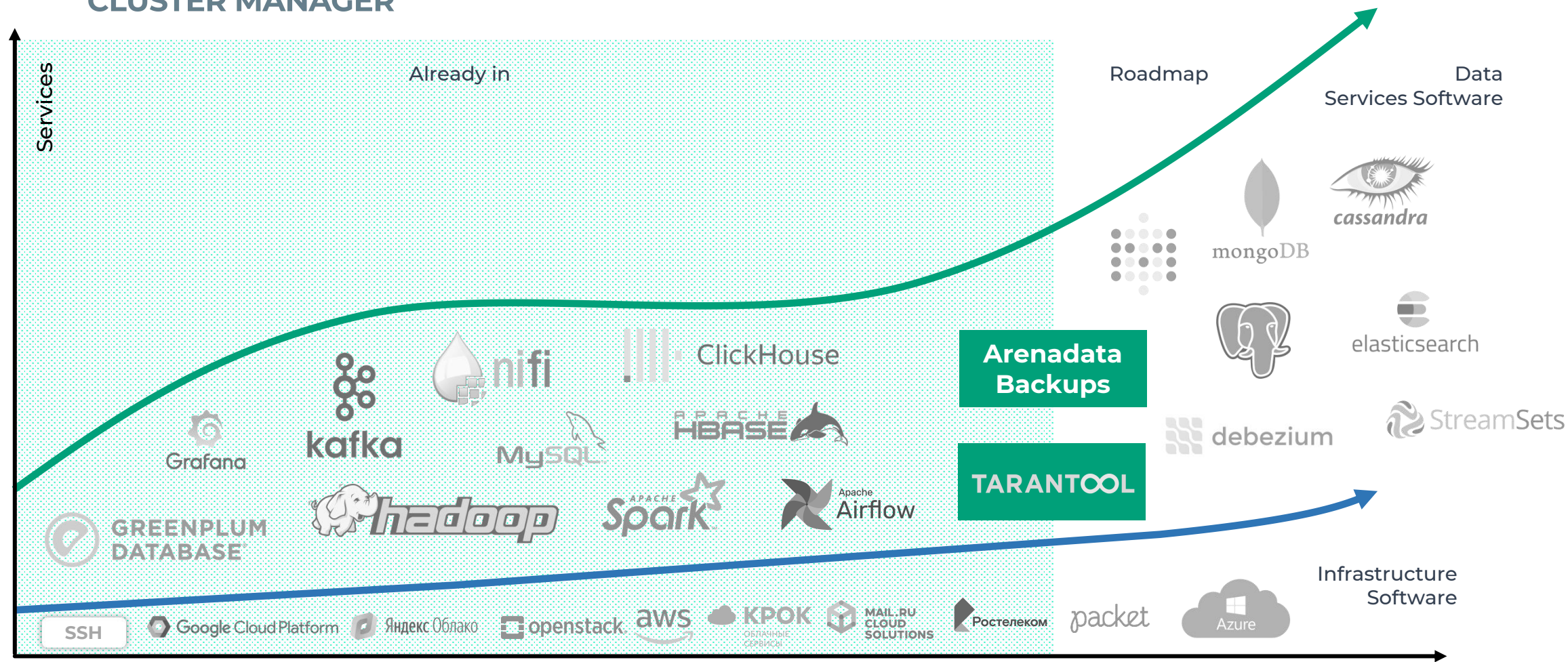
Дополнительные возможности:

- Настройка прав безопасности в ADCM
- Открытый ADCM API
- Открытый формат для создания новых бандлов - добавьте ваши собственные сервисы/инфраструктуру

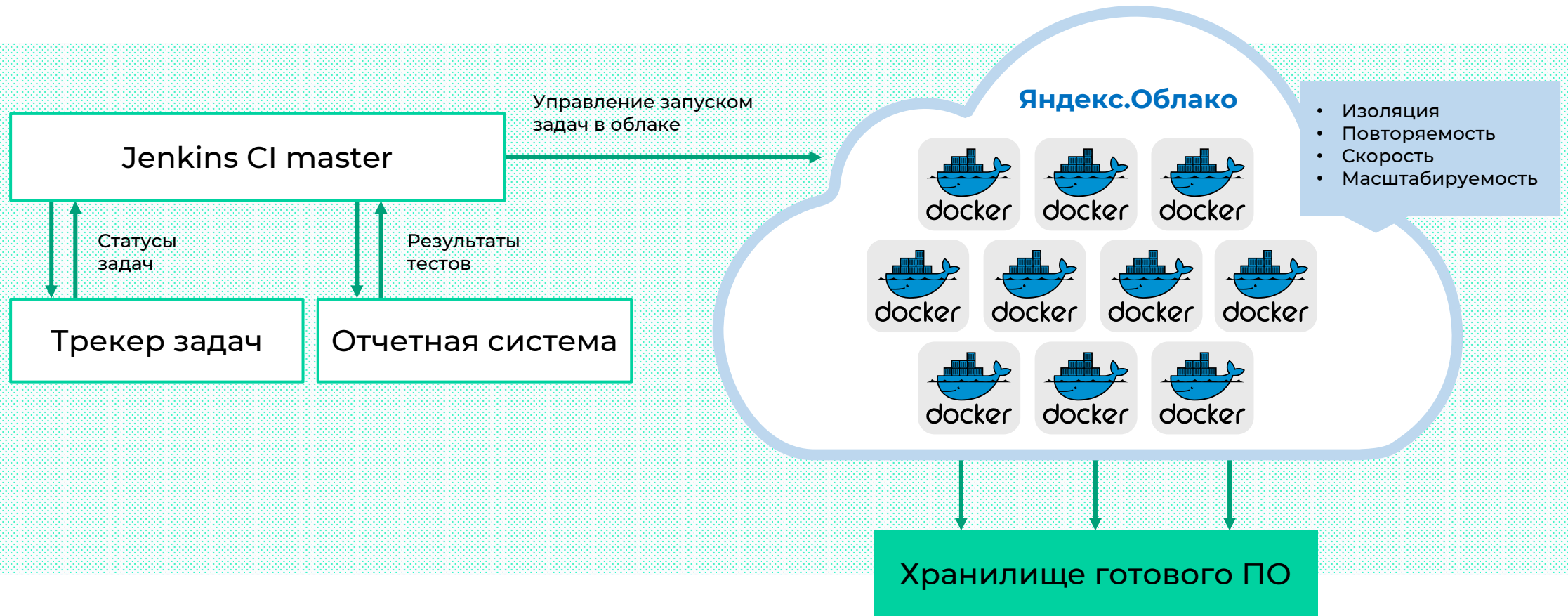


ARENADATA Roadmap

CLUSTER MANAGER



ARENADATA — среда сборки



Поддержка 24x7

Санкт-Петербург

Москва

Нур-Султан

Хабаровск

Сотрудники
распределены
по часовым
поясам

Услуги по поддержке клиентов

- Vendor support
- DBA as a service

24

часа
в сутки

7

дней
в неделю

365

дней
в году

VIP-пакет

- Выделенный сервис-менеджер
- Регулярный аудит
- Обновление минорных версий



Корпоративная платформа хранения
и обработки больших данных

Нам доверяют



Мнение клиентов



**Дмитрий
Алексеев**
заместитель
директора
департамента
технологического
развития
общественных
систем ВТБ



Все текущие хранилища данных банка будут поэтапно мигрировать на новое целевое единое хранилище, построенное на базе ядра розничного хранилища. При этом текущую технологическую платформу заменит отечественное решение от Arenadata.



**Михаил
Петров**
директор
департамента
Цифровой
трансформации
Счётной палаты РФ



С Arenadata мы построили продуктивное сотрудничество. Нам крайне радует, что рядом всегда есть специалисты, обладающие обширной экспертизой и высокой квалификацией.



**Сергей
Носов**
директор по
управлению
данными
ПАО Ростелеком



Посредством перевода централизованного хранилища на стек Open Source удалось сэкономить средства, сравнимые с годовым бюджетом построения хранилищ данных в крупных компаниях. С учетом требований, предписывающих компаниям с государственным участием переходить на отечественное ПО, это достижение особенно ценно.



**Иван
Черницын**
руководитель
Центра
аналитических
решений
Дирекции
региональных
продаж ПАО
«Газпром нефть»



Взаимодействие с командой Arenadata было весьма эффективным. Специалисты компании помогали с вопросами внедрения и безопасности, предлагали собственные решения и оперативно выпускали релизы с учётом специфических требований заказчика.



Наши партнёры

ВНЕДРЕНИЕ

accenture

КРОК

IBS

Jet

DIS
GROUP

АПЛАНА
ЦЕНТР РАЗРАБОТКИ

IBS

Platformix

ИТ ФИЛОСОФИЯ

POLYMEDIA



отр



ЛАНИТ

sapiens
solutions

РQTEX

Rubbles
SBDA Group

LUXMS
группа компаний

ARENADATA

ОБЛАЧНЫЕ ПРОВАЙДЕРЫ

MAIL.RU
CLOUD
SOLUTIONS

Yandex Cloud

Ростелеком

КРОК
ОБЛАКО

IBS DataFort

ТЕХНОЛОГИЧЕСКИЕ ПАРТНЁРЫ

accenture digital

@ mail.ru
group

Яндекс

YA
DR

IBM

Informatica

POLY//ATICA

VISIOLOGY

LUXMS
группа компаний

Форсайт
Аналитическая платформа



Корпоративная платформа хранения
и обработки больших данных

Оцените потенциал Big Data и Open Source вместе с Arenadata Enterprise Data Platform

- Узнайте больше на arenadata.tech
- Скачайте бесплатно на store.arenadata.io
- Получите консультацию по почте info@arenadata.io

