



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+

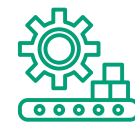
Лидеров РФ выбрали нас
в качестве центральной
платформы данных

25 ПБ

Суммарный объём
данных, который наши
клиенты хранят в
платформе Arenadata



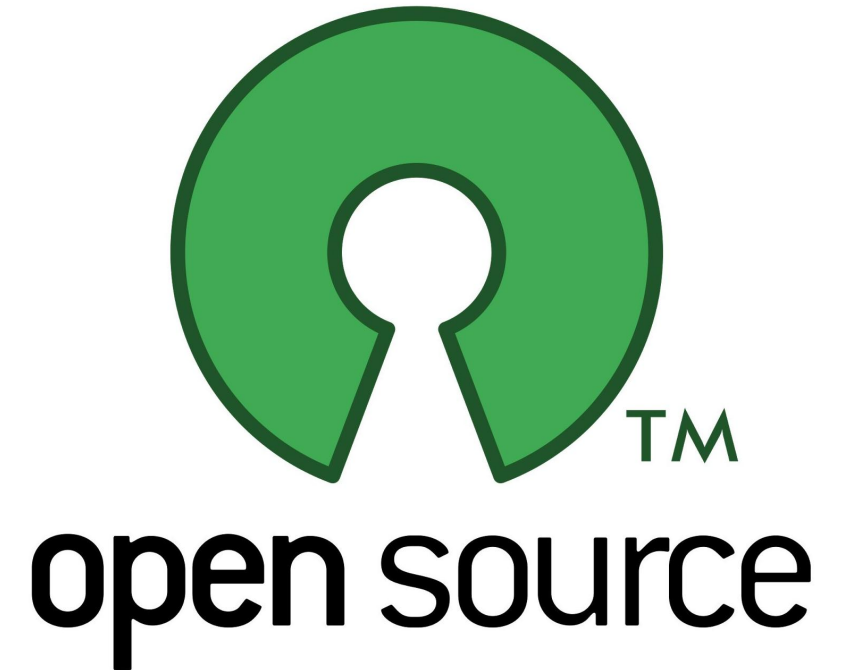
Развитая сеть партнёров
как в России, так и за
рубежом



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

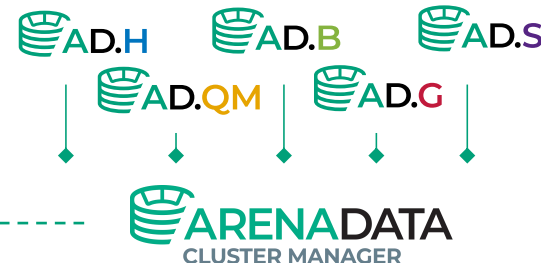


Множество инструментов для решения прикладных задач



Arenadata Enterprise Data Platform (EDP) 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

Enterprise Data Platform — решаемые задачи



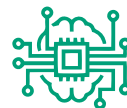
Корпоративные хранилища больших данных

- 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

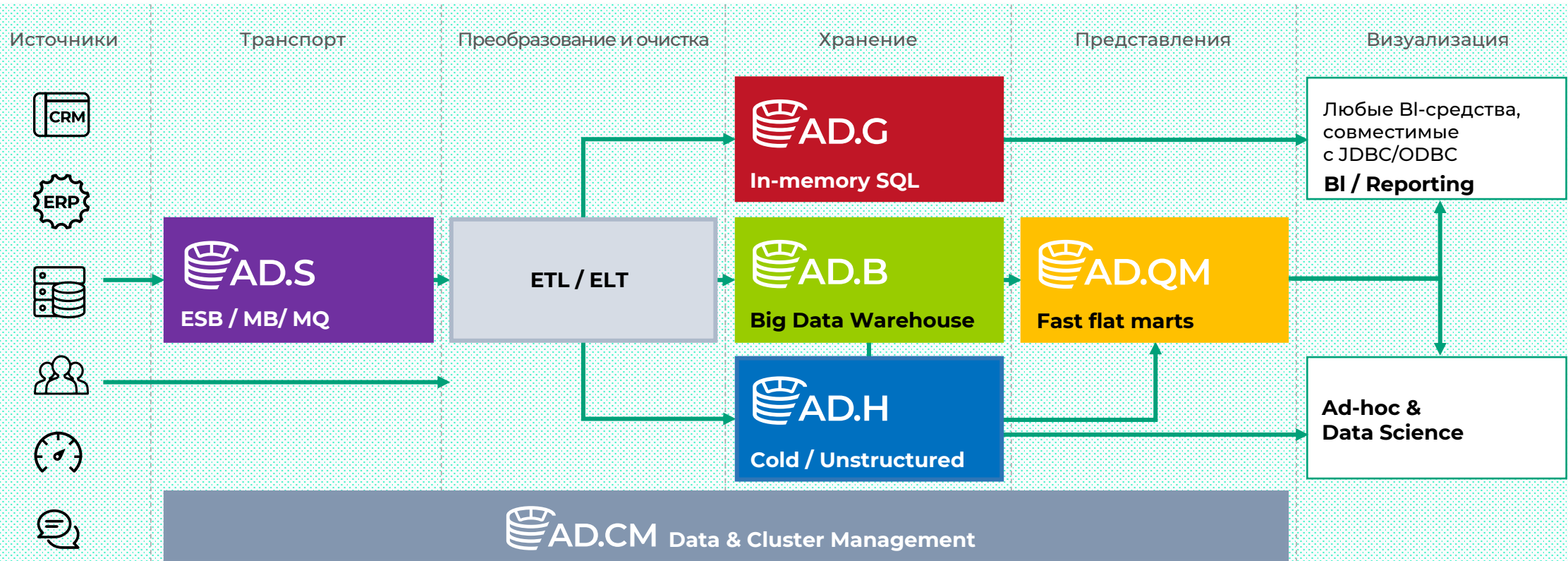
- Поддержка R, Python и Madlib.
- Distributed ML на базе Spark.
- Возможность эффективной эксплуатации моделей в SQL-среде с помощью встроенных функций Madlib.



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

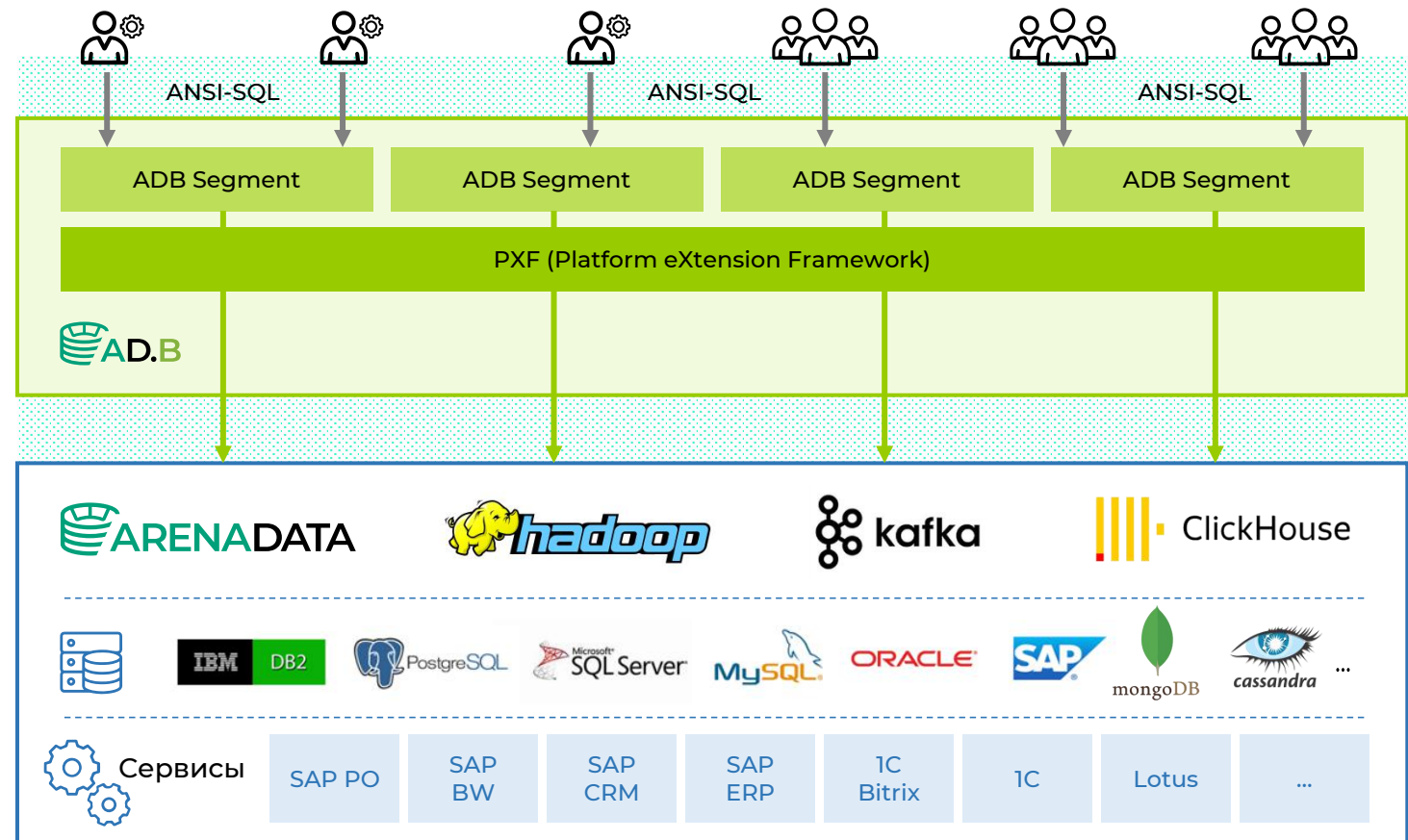
PAAS

Public Cloud

Private Cloud

Управление федерацией данных с PXF

- Единая точка входа для всех SQL-запросов
- Специализированный движок виртуализации данных
- Межкомпонентный массово-параллельный проброс предикатов запроса
- В том числе во внешние системы (SAP HANA, Oracle DB, MS SQL, PostgreSQL, etc)
- Кластеризация, шардирование и локальность данных в каждой компоненте хранения.



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

А что у мировых лидеров?



Крупнейшие мировые компании уже используют тот же самый стек технологий, только в виде разрозненных компонент



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

Скажите «ДА» современной платформе!



CLOUDERA
ORACLE®
teradata.
Exasol



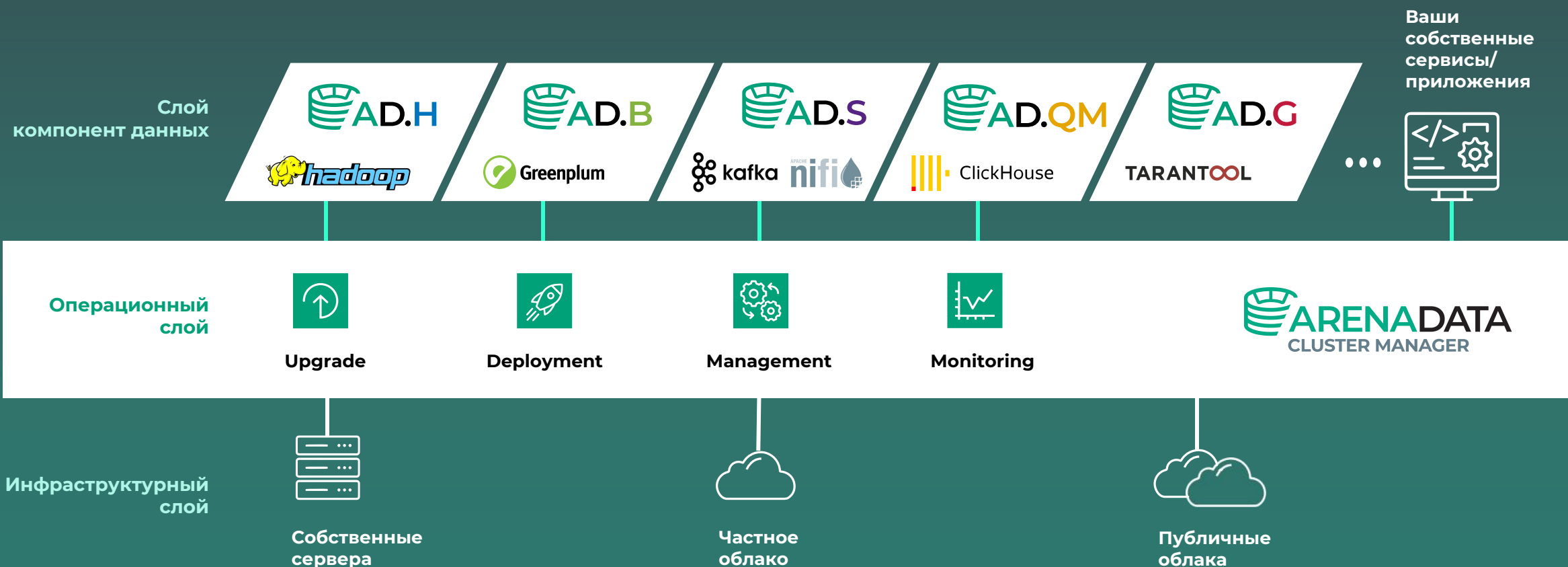
Скажите «ДО СВИДАНИЯ»

- закрытой архитектуре
- монолитной архитектуре
- vendor lock-in
- высокой TCO

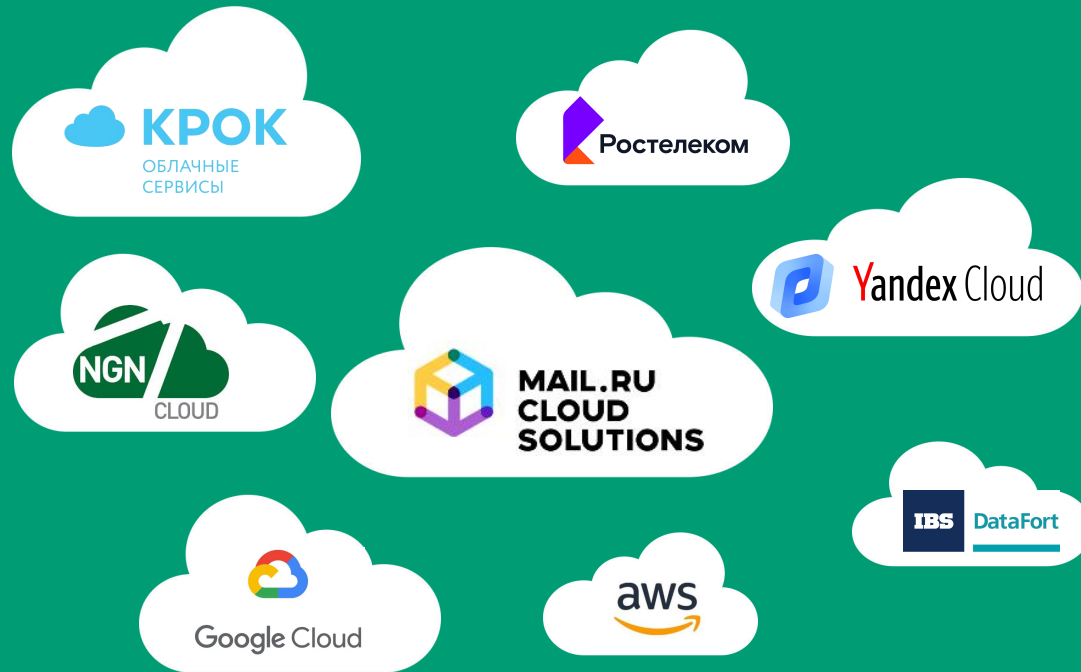


Мы обладаем опытом
миграции клиентов с
других систем

Гибридное корпоративное облако



Enterprise Multi-cloud платформа данных



- Вы можете развернуть любой компонент платформы Arendata EDP как на bare-metal оборудовании, так и в облаке, или же воспользоваться востребованным сегодня сценарием — гибридной ИТ-инфраструктурой или Multi-clouds.
- Установите EDP распределённо на разные инфраструктуры, обеспечив между ними сетевую доступность.

Компоненты платформы Arenadata EDP

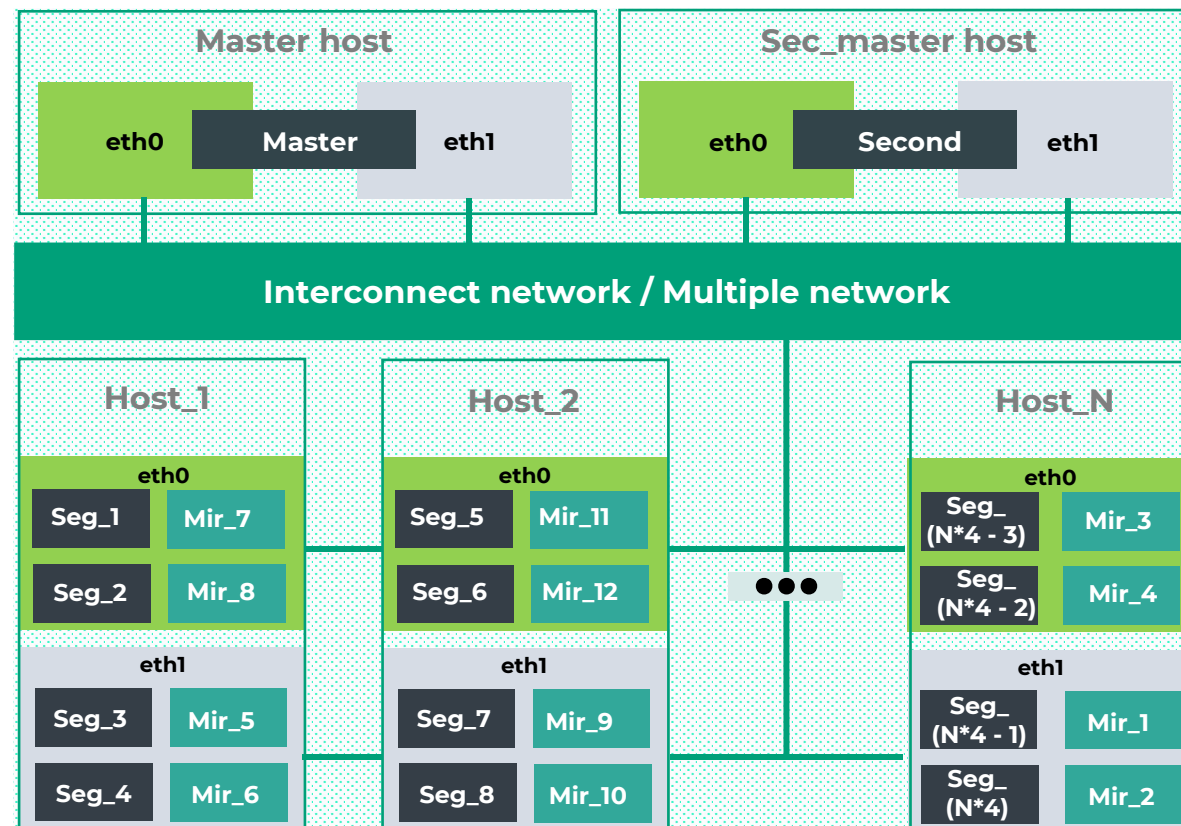


Arenadata DB (ADB) — аналитическая, распределённая MPP-СУБД с открытым исходным кодом.

ADB реализована на кластере из множества серверов и предназначена для хранения и обработки больших объёмов данных — до десятков петабайт.

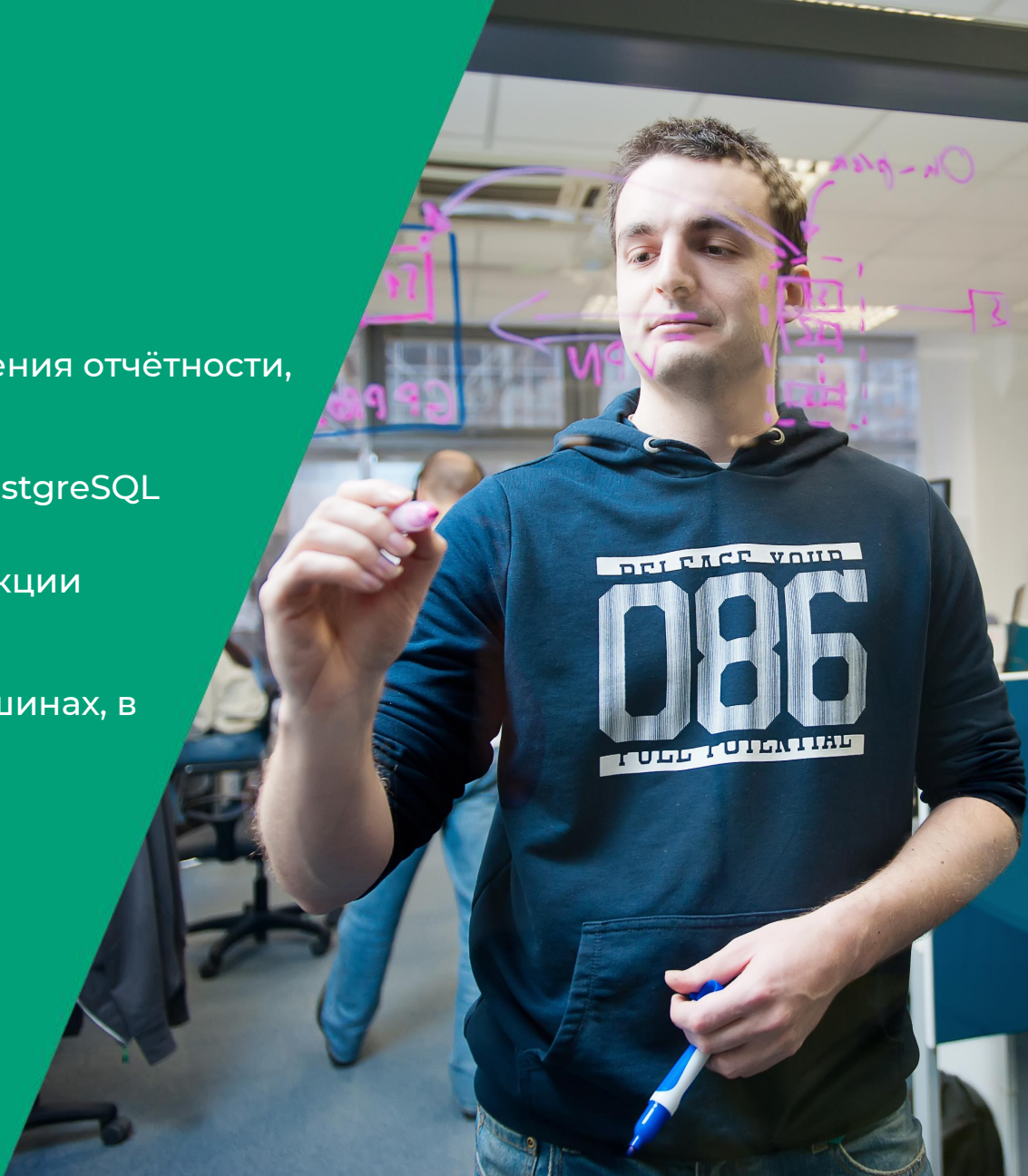
Arenadata DB лучше всего справляется с:

- Сложными запросами, обрабатывающие большие объёмы, в том числе сложными аналитическими функциями
- ETL/ELT
- Эффективным соединением больших таблиц
- Работой с индексами
- Data Science
- Аналитическими функциями на процедурных языках
- Ad-hoc аналитикой



Arenadata DB

- Arenadata DB подходит как для классических задач построения отчётности, так и для задач Data Science
- ADB на 100% совместима со всем ПО, которое работает с PostgreSQL
- Полное соответствие принципам строгой изоляции транзакции (принципы ACID)
- Может быть размещена в вашем ЦОДе, на виртуальных машинах, в частных и публичных облаках
- Эксплуатация ADB для администраторов не сложнее, чем работа с обычной PostgreSQL
- Набор коннекторов для параллельной загрузки и выгрузки данных из сторонних систем (Kafka, ClickHouse, Hadoop) существенно упростит процесс интеграции с источниками данных и аналитическими системами



Arenadata Hadoop (ADH) — это корпоративный дистрибутив на базе Apache Hadoop, предназначенный для хранения и обработки слабоструктурированных и неструктурированных данных.

- В состав дистрибутива ADH входят последние актуальные версии всех наиболее популярных инструментов. Некоторые из них существенно доработаны, что гарантирует минимальное количество ошибок ПО, самый полный существующий функционал и корректную интеграцию компонентов между собой.
- Arenadata Hadoop включает собственную систему управления, развёртывания и мониторинга Arenadata Cluster Manager, позволяющую устанавливать сервисы Hadoop и управлять ими как on-premise, так и в облаке.

Operations



Security &
Data Governance



Data Orchestration



Data Access & Processing



Persistence & Compute



Arenadata Hadoop

- Полный набор возможностей и инструментов для автоматической установки и настройки компонентов как на «голом железе», так и на виртуальных машинах (в облаке)
- Open-source дистрибутив Hadoop без vendor lock-in
- Проверенные версии компонентов стека Hadoop
- Собственная система управления
- Набор типовых пакетных сервисов по планированию, установке и аудиту системы
- Возможность доработки и кастомизации продукта под конкретные потребности заказчика
- Оригинальная документация на русском языке

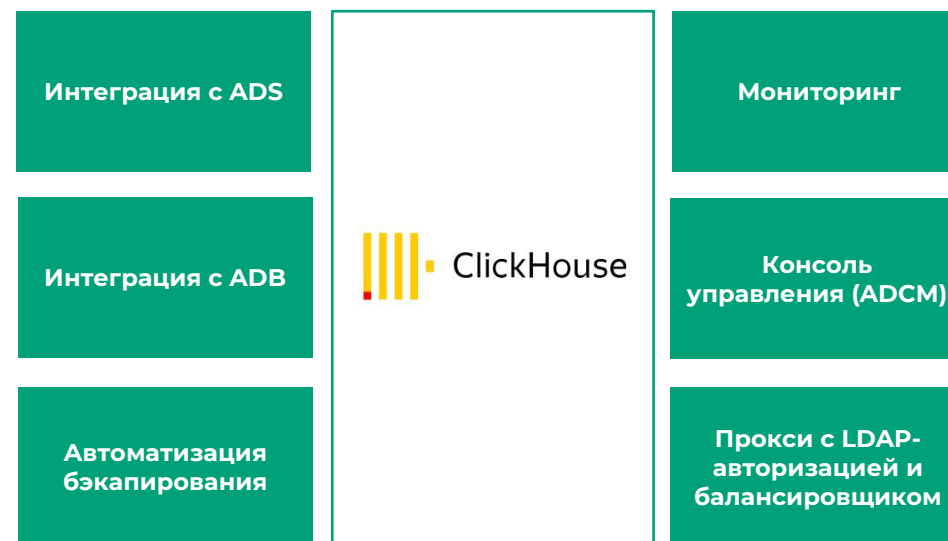


Arenadata QuickMarts (ADQM) — кластерная колоночная СУБД, созданная на основе Yandex ClickHouse. ADQM многократно быстрее традиционных СУБД.

С ADQM вы можете в режиме реального времени генерировать аналитические отчёты разного плана, используя большие объёмы информации, хранящейся в плоских витринах.

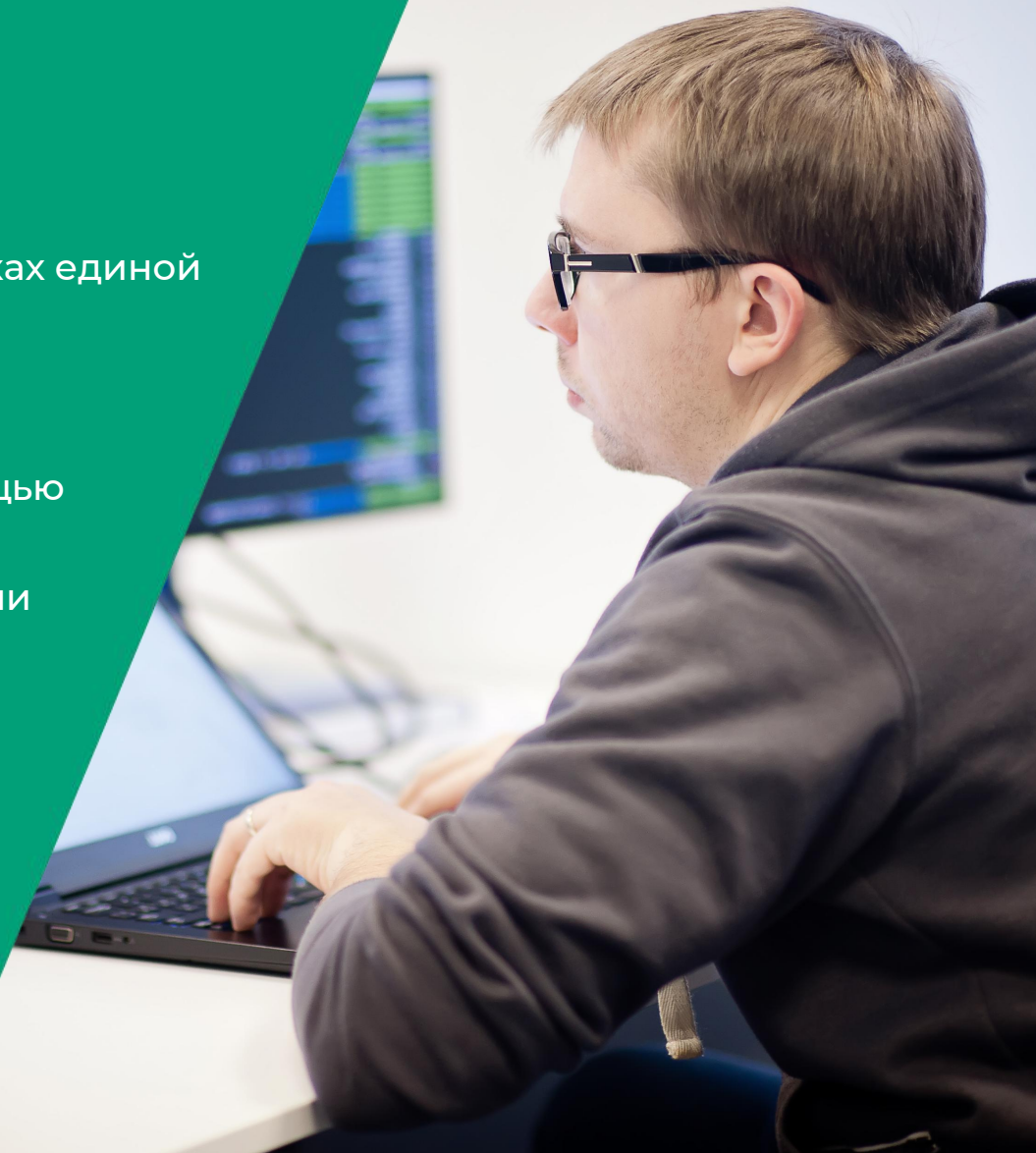
Скорость на грани фантастики:

- Типичные запросы быстрее чем за секунду
- > 100 раз быстрее чем Hadoop и обычные СУБД
- 100 млн - 1 миллиард строк в секунду на одной ноде
- До 2 терабайт в секунду для кластера на 400 нод



Arenadata QuickMarts

- Нативная интеграция с Arenadata DB, Hadoop, Streaming в рамках единой платформы Arenadata EDP
- Простота установки и настройки из Arenadata Cluster Manager
- Балансировка нагрузки и распределение доступа к БД с помощью Arenadata Baje
- Авторизация пользователей, в том числе поддержка интеграции с Kerberos Hadoop
- Мониторинг на базе Graphite и Grafana
- Дополнительное тестирование и контроль качества релизов
- Наличие российской поддержки
- Оперативная доработка и развитие функционала под требование заказчика

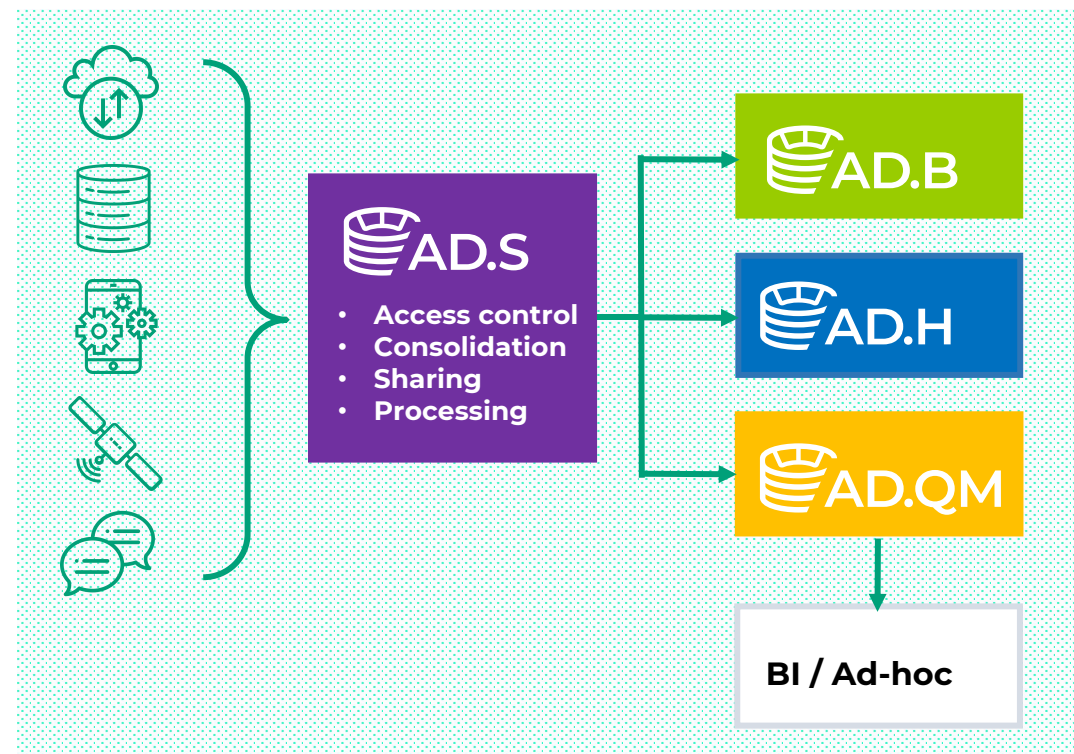


Arenadata Streaming (ADS) — масштабируемая отказоустойчивая система для потоковой обработки данных в режиме реального времени, адаптированная для корпоративного использования и построенная на базе Apache Kafka и Apache Nifi.

ADS является универсальным средством для решения задач, связанных с потоковой обработкой данных в режиме реального времени.

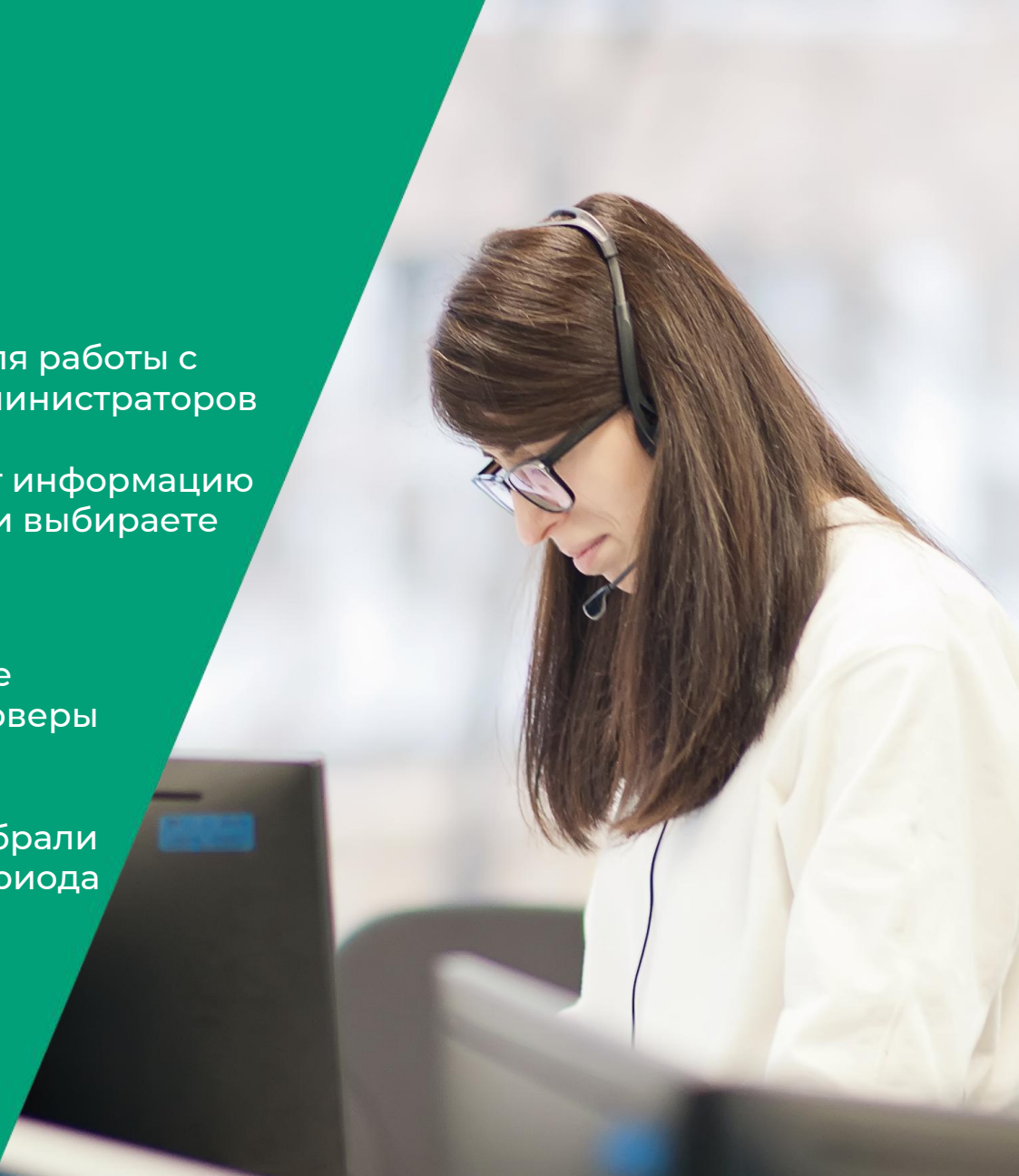
Используется как:

- 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.



Arenadata Streaming

- ADS позволяет снизить порог входа в технологию, так как для работы с продуктом не требуется высокий уровень компетенции администраторов
- В отличие от других брокеров сообщений, которые удаляют информацию сразу же после того, как она была доставлена, с ADS вы сами выбираете нужный период хранения данных
- Гибкость и масштабируемость системы позволяет ей расти вместе с вашими потребностями, подстраиваясь под текущие задачи. Так, вы всегда можете добавить в кластер новые серверы без остановки работы системы или же убрать ненужные
- ADS получает и хранит информацию независимо от того, забрали её или нет: она доступна в течение всего обозначенного периода хранения
- Широкий набор API для интеграции с другими внешними системами.



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



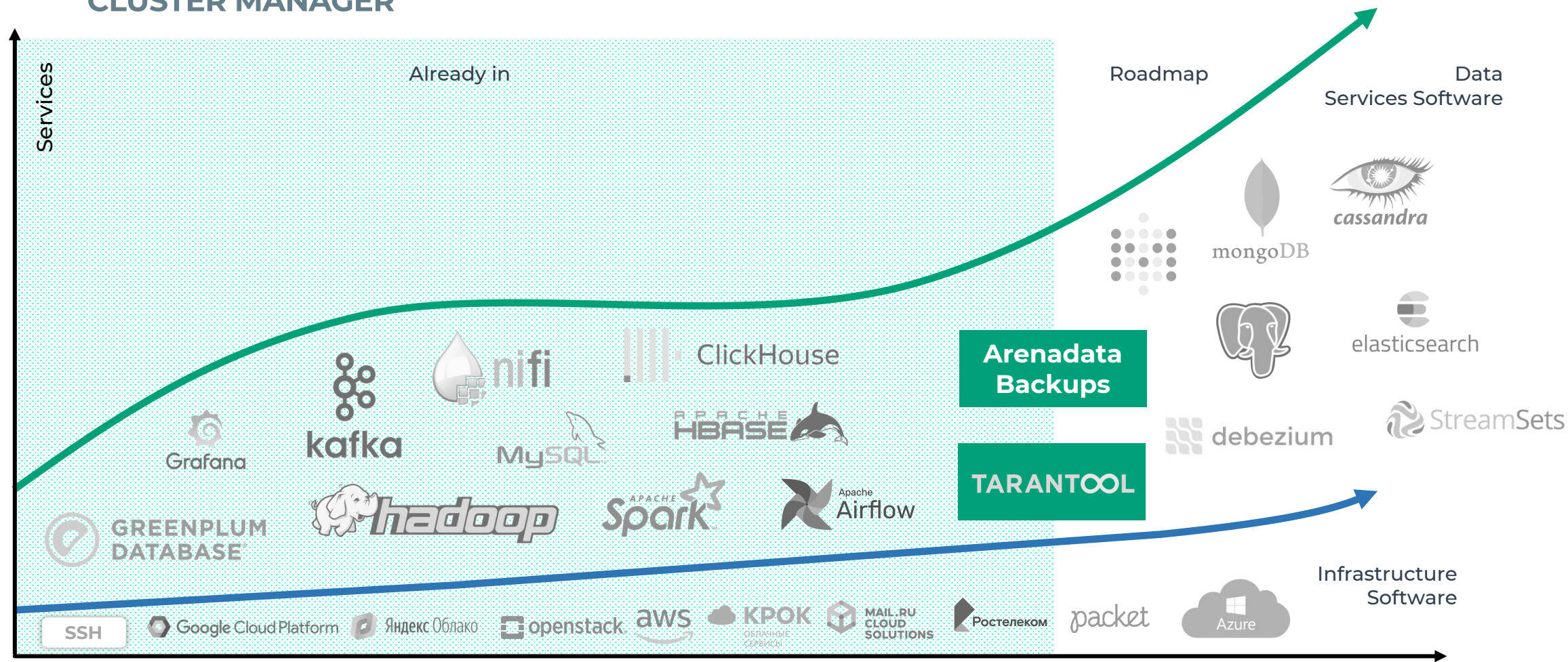
Arenadata Cluster Manager

- Arenadata Cluster Manager — удобный инструмент для установки, настройки, мониторинга и обновления как ПО Arenadata, так и сторонних сервисов
- Функционал приложений ADCM можно использовать для установки, настройки, обновления, управления, мониторинга и интеграции компонентов
- Модульная структура платформы предоставляет необходимый функционал для работы с приложениями и оборудованием по мере необходимости
- Вы значительно сокращаете время на установку и настройку элементов data-ландшафта, тестирование бизнес-сценариев и проверку гипотез
- Вы можете развернуть нужный продукт с помощью ADCM на удобной вам инфраструктуре - на собственных серверах, полностью в одном или нескольких облаках или распределённо на разные инфраструктуры



ARENADATA Roadmap

CLUSTER MANAGER



Услуги вендора

Вендорский консалтинг и профессиональные сервисы



Архитектурный контроль

Мы предоставляем проверенную вендорскую точку зрения на проект внедрения, позволяющую принять правильное архитектурное решение в непростых ситуациях.



Аудит цифрового ландшафта

Позволяет ИТ и менеджменту научиться отвечать на вопрос «Как должен измениться ландшафт данных и цифровых сервисов, чтобы обеспечить мои бизнес-потребности сейчас и в будущем?».



DBA as a service

Каждой СУБД необходим сведущий DBA. Достаточно обратиться к нам как к вендору СУБД, готовому предоставить услугу DBAaaS.



Smart Start

Это комплекс услуг по установке и настройке платформы Arenadata и интеграции наших продуктов в инфраструктуру данных заказчика. Мы также проводим инструктаж для администраторов, и выполняем ряд не столь значительных, но полезных вспомогательных действий.



Разгрузка дорогостоящих СУБД

Как вендор, мы поможем вам снизить ТСО владения проприетарными СУБД за счёт внедрения и интеграции аналитических MPP-систем под ключ.

DBA as a service

Каждой СУБД необходим администратор с нужным уровнем квалификации. Для этого достаточно обратиться к вендору СУБД, готовому предоставить услугу DBAaaS. И у Arenadata такая услуга есть.

DBA as a service позволит вам полностью абстрагироваться от человеческого фактора: вы получаете службу, поддерживающую ваши кластеры в режиме non-stop.

24

часа в сутки

7

дней в неделю

365

дней в году

Состав услуги DBAaaS:

- Поддержка работоспособности сервисов
- Администрирование сервисов
- Обеспечение резервирования и восстановления сервисов
- Мониторинг работы



Поддержка 24x7

Санкт-Петербург

Москва

Алма-Аты

Хабаровск

Сотрудники
распределены
по часовым
поясам

Услуги по поддержке клиентов

- Vendor support
- DBA as a service

VIP-пакет

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



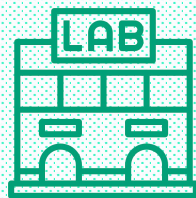
Корпоративная платформа хранения
и обработки больших данных

Курсы по продуктам Arenadata



Знания из первых рук

Наши преподаватели — практикующие специалисты, имеющие за плечами богатую экспертизу и значительное количество реализованных бизнес-кейсов.



Только 40% теории

Все курсы построены на реальных кейсах и практических примерах, а лабораторные работы проходят на стендах, симулирующих рабочую среду.



Сертификация

Пройдя обучение на курсах Arenadata и успешно сдав экзамен, слушатель станет обладателем именного сертификата, подтверждающего полученные знания.

Нам доверяют



Истории успеха



Организация «Умного озера данных» для комплексной системы анализа и управления данными Дирекции региональных продаж на базе Arenadata Hadoop.



Создание единой технологической платформы, способной объединить хранилища данных, создать систему отчётности и аналитики по миллионам клиентов, сокращение TCO на базе массово-параллельной СУБД Arenadata DB и Arenadata Streaming.



Создание платформы данных по всем данным компании с привязкой к системе Data Governance, способной выдерживать высокую нагрузку от пользователей (2000 аналитиков и более 800 конкурентных запросов) с помощью Arenadata DB, Arenadata QuickMarts и Arenadata Streaming.



Построение «Озера данных» для осуществления внешнего государственного аудита на новом качественном уровне. Используются данные всех возможных типов из совершенно разных источников, перечень которых заранее не определён. «Озеро данных» построено с помощью Arenadata Hadoop и Arenadata DB.



Разработка цифровых сервисов и моделей машинного обучения, создание удобного инструмента для вычислений и анализа, а также оптимизация производственных процессов и снижение себестоимости готовой продукции на базе Arenadata Hadoop.



Создание аналитического хранилища, отдельного от системы поддержки управленческой отчётности, объединение трёх «песочниц» и их миграция с SAP HANA на Arenadata DB.

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

РТЕХ

Rubbles
SBDA Group

LUXMS
группа компаний

ARENADATA

ОБЛАЧНЫЕ ПРОВАЙДЕРЫ

MAIL.RU
CLOUD
SOLUTIONS

Yandex Cloud

Ростелеком

КРОК
ОБЛАКО

IBS DataFort

ТЕХНОЛОГИЧЕСКИЕ ПАРТНЁРЫ

accenturedigital

@ mail.ru
group

Яндекс

YADRO

IBM

Informatica

POLY//ATICA

VISIOLOGY

LUXMS
группа компаний

Форсайт
Аналитическая платформа

3 шага к платформе Arenadata

1

Расскажите нам о задачах, ИТ-инфраструктуре и получите рекомендации по компонентам платформы и выбору партнёра.

2

Загрузите бесплатно необходимые пакеты для платформы данных с нашего сайта.

3

Разверните нашу платформу в своей инфраструктуре и отправьте нам технические и функциональные требования, чтобы получить индивидуальное предложение.





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

