



Модульная платформа
для высоконагруженных
корпоративных и государственных
информационных систем

Машина больших данных Скала^р МБД.Х

скала^р

Скала^р — модульная платформа для построения инфраструктуры высоконагруженных корпоративных и государственных информационных систем



9 лет

серийного
выпуска

600+

комплексов
в промышленной
эксплуатации

8 500+

вычислительных
узлов

Продуктовые направления Скала^р

решения для высоконагруженных корпоративных и государственных систем



Динамическая инфраструктура

Машины виртуализации Скала^р МВ

на основе решений BASIS для создания динамической конвергентной и гиперконвергентной инфраструктуры ЦОД и виртуальных рабочих мест пользователей



Высокопроизводительные базы данных

Машины баз данных Скала^р МБД

на основе решений Postgres Pro для замены Oracle Exadata в высоконагруженных системах с обеспечением высокой доступности и сохранности критически важных данных



Инфраструктура для ИИ

Машина Скала^р для искусственного интеллекта

на основе оптимизированного программно-аппаратного стека для максимальной производительности при работе с моделями ИИ



Управление большими данными

Машины больших данных Скала^р МБД.8

на основе решений ARENADATA и PICODATA для создания инфраструктуры хранения, преобразования, аналитической, статистической обработки данных, а также распределенных вычислений



Интеллектуальное хранение данных

Машины хранения данных Скала^р МХД

на основе технологии объектного хранения S3 для геораспределенных катастрофоустойчивых систем с сотнями миллионов объектов различного типа и обеспечения быстрого доступа к ним

- Использование опыта технологических лидеров (гиперскейлеров)
- Использование самых зрелых и перспективных технологий в кооперации с технологическими лидерами российского рынка в каждом из сегментов

Модульная платформа Скала[^]р



Единый принцип модульной компоновки и платформенный подход

Единая облачная система управления сервисами



IaaS



PaaS



DBaaS

Единая система управления ресурсами и эксплуатацией



Разделение ресурсов



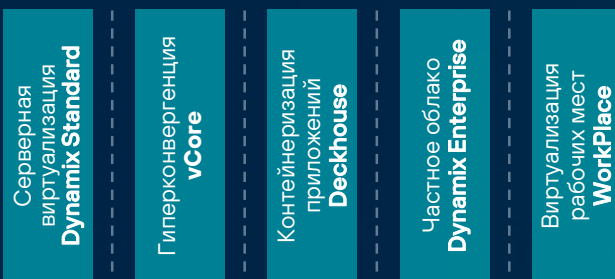
Мультитенантность



Автоматизация

Модульная платформа

Динамическая инфраструктура



Динамическая инфраструктура

Инфраструктура управления данными



Транзакционная обработка

Большие данные

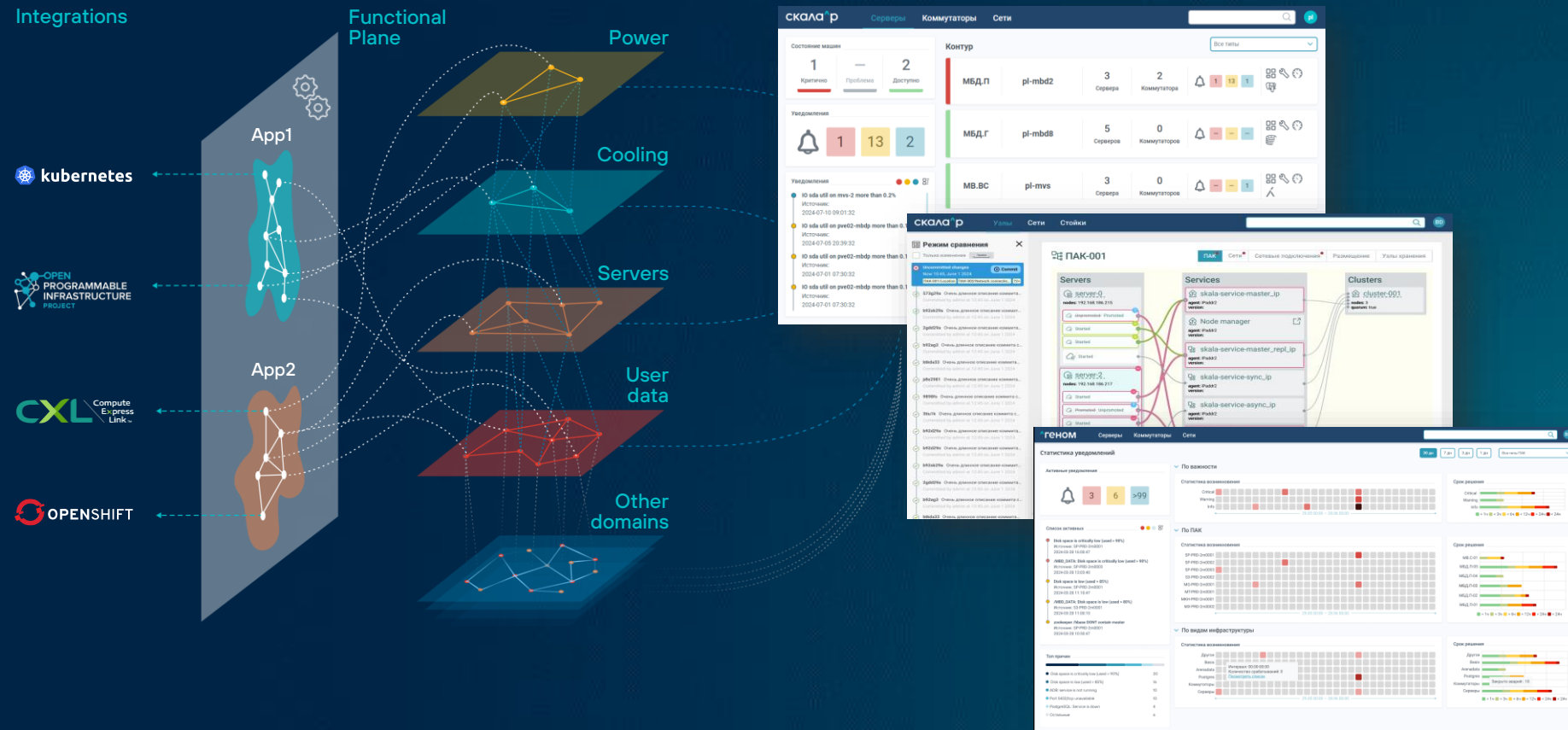
Специализир. решения

Отраслевые решения

Перспективная платформа Скала[^]р

Объединения различных доменов управления в единую функциональную графовую CMDb

Комплексное решение для эксплуатации инфраструктуры уровня ЦОД



- Единая точка обзора состояния контура
- Обозримость и удобство управления ЦОД
- Цифровой двойник инфраструктуры
- Контроль изменений быстроменяющихся топологий
- Моделирование изменений в инфраструктуре
- Высокая степень автоматизации
- Построение AI-Copilot для управления ЦОД

ПАК Скала[^]р в Реестрах РФ



Машины

Модули

Компоненты

ГТИ
государственная
информационная система
промышленности

МИНПРОМТОРГ
РОССИИ



Все сервисы ГИСП

Реестр промышленной продукции, произведенной на территории
Российской Федерации

Машины

Модули

Программное обеспечение



РЕЕСТР
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Российский

Евразийский

Машины

Модули

Программное обеспечение

Соответствуют критериям доверенного ПАК

- Технологическая независимость
- Информационная безопасность
- Функциональная устойчивость

Импортозамещение: Сложность выбора

Отсутствие технологического лидерства



Глобальный ИТ-рынок

Сетевая инфраструктура



Хранение данных



Виртуализация



Вычислительная инфраструктура



СУБД



Операционные системы



Российский ИТ-рынок

Сетевая инфраструктура



Хранение данных



Виртуализация



Вычислительная инфраструктура



СУБД



Операционные системы



Проблемы отсутствия ИТ-лидеров на российском рынке

- Отсутствие информации и практического подтверждения совместимости продуктов
- Время и ресурсы для подтверждения соответствия заявленной функциональности

- Проблема совместимости с продуктами из разных классов
- Размывание понятия «лидер»: в каждом сегменте существуют десятки на первый взгляд равноценных продуктов

Импортозамещение: Варианты перехода



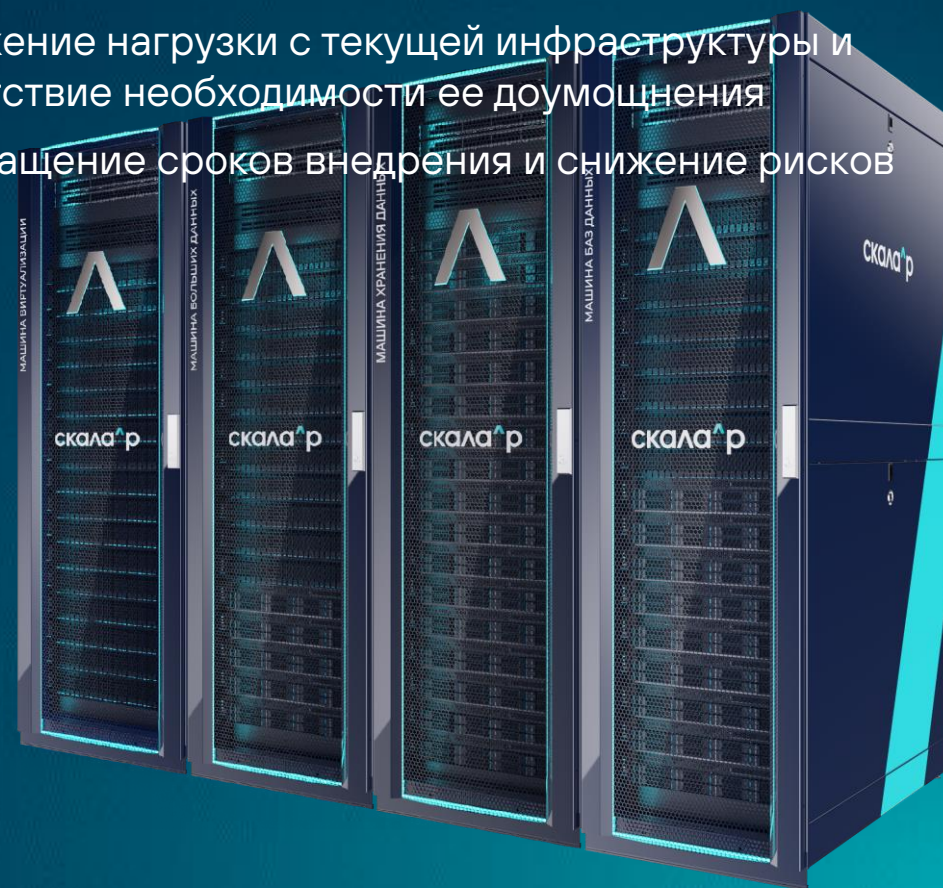
Покомпонентное замещение:

- Время на изучение вариантов, тестирование и выбор
- Лавина взаимосвязанных проектов по внедрению
- Сложность синхронизации дорожных карт развития
- Рост сроков внедрения и рисков на стыках



Создание целевой доверенной ИТ-инфраструктуры:

- Последовательный перевод систем на целевую доверенную ИТ-инфраструктуру
- Снижение нагрузки с текущей инфраструктуры и отсутствие необходимости ее доумощнения
- Сокращение сроков внедрения и снижение рисков



Почему ПАК Скала^р ?



- Гарантированно совместимые компоненты
- Отказоустойчивость на уровне архитектуры
- Оптимизация производительности
- Ответственность одного производителя за функционал и показатели назначения
- Решенные вопросы интеграции, эксплуатации, мониторинга, обеспечения ИБ, резервного копирования
- Поддержка и сервис из одного окна
- Серийность и преемственность
- Управляемая дорожная карта развития



Конкурентные преимущества оптимизированных решений



Производительность

x2[↑]

чем решения, использующие сопоставимые аппаратные средства за счет оптимизации ввода-вывода и интерконнекта и за счет разгрузки ЦПУ

x4[↑]

чем решения в виртуальной среде, использующие сопоставимые аппаратные средства за счет снижения латентности

x4[↑]

для систем с большим количеством сессий за счет использования специализированных пулеров и балансировщиков

RPO/RTO

x4[↓]

время выполнения резервного копирования и восстановления за счет специализированного встроенного модуля резервного копирования

x6[↓]

время полного восстановления узла в случае отказа за счет использования встроенной системы развертывания и цифрового двойника системы

Доступность

Кратное сокращение инцидентов связанных с ошибками эксплуатации и существенное увеличение доступности за счет использования специализированной системы управления ресурсами

ПАК — Машины Скала^р — преимущества перед самостоятельными проектами



Высокая отказоустойчивость

За счет специализированной модульной и кластерной архитектуры решений

Высокая производительность

Встречная оптимизация и устранение узких мест по всему стеку применимых технологий

Единая техническая поддержка

Сопровождение оборудования и программного обеспечения всех компонентов Машин

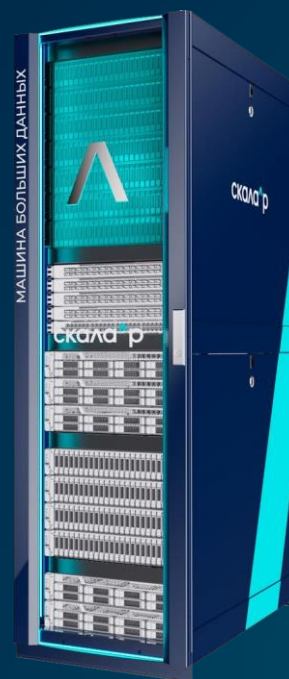
Экономия до 90%
на проектировании и внедрении

Продукты развиваются
с учетом пожеланий Заказчиков

Высокая доступность
и катастрофоустойчивость из коробки

Соответствие требованиям ИБ

Российское оборудование и ПО



Ускорение до 30%
проектов импортозамещения

Кратное сокращение инцидентов,
связанных с ошибками эксплуатации

Удобство закупочных процедур для ПАК и Модулей —
это номенклатурные позиции Реестра РЭП
Минпромторга РФ

Соответствие актуальному законодательству
по закупкам — преференции изделиям

Применение для КИС и ГИС,
включая доверенные ПАК для КИИ

Прямое взаимодействие с технологическими партнерами по развитию необходимого заказчиком функционала

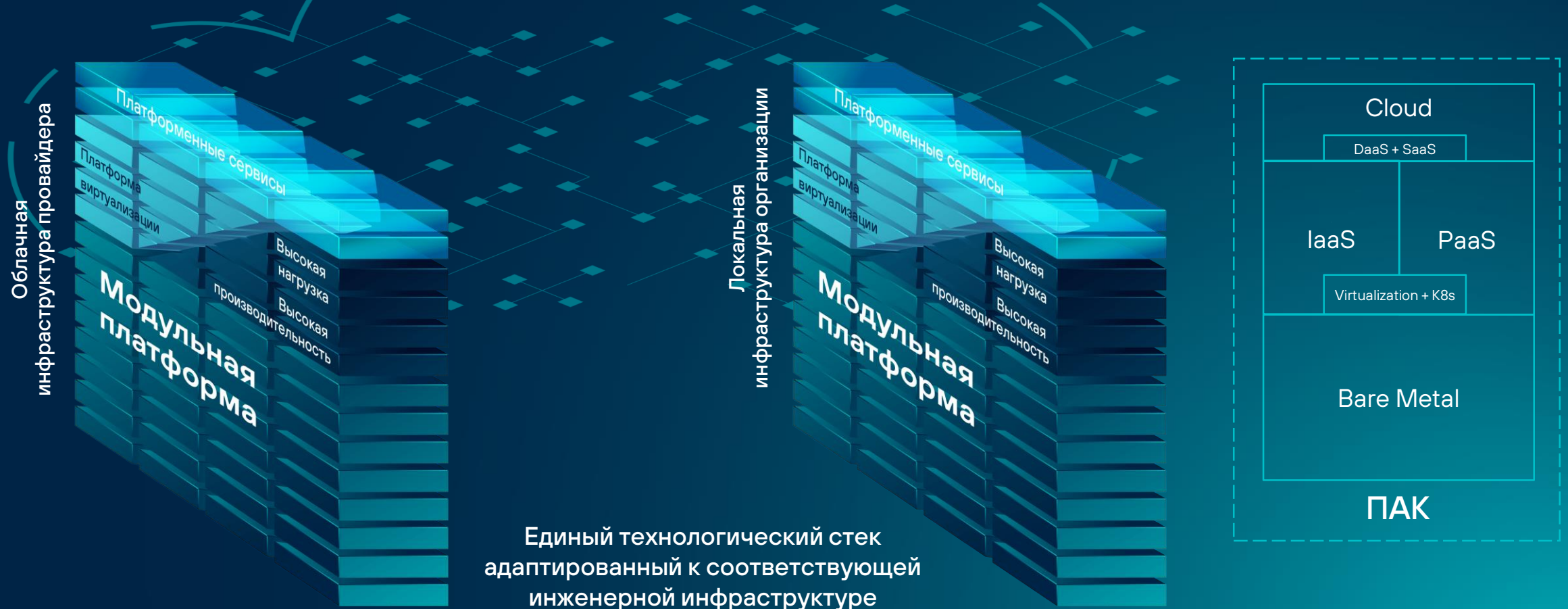
ПАК — Программно-аппаратный комплекс и модули платформы — включены в Единый реестр российской радиоэлектронной продукции и реестр Минцифры

Продуктово-технологическая концепция Скала^р

Миграция крупнейших организаций в распределенное облако

Построение локальных модульных инфраструктур с облачной системой управления от провайдера

Совместное использование локальных ресурсов и ресурсов провайдера из единой консоли управления



Если крупные корпоративные заказчики не идут в облако провайдера, то облако провайдера должно прийти к ним

Машины больших данных Скала^р МБД.8



высокопроизводительные хранилища и витрины данных
на базе продуктов Arenadata и Picodata

Скала^р МБД.Г ← Arenadata DB (ADB)

СУБД массивно-параллельной обработки (на основе Greenplum)

Скала^р МБД.Т ← Picodata

Распределенные вычисления в оперативной памяти (на основе Picodata)

Скала^р МБД.С ← Arenadata Streaming (ADS)

Потоковая обработка данных в реальном времени (на основе Kafka и NiFi)

Скала^р МБД.Х ← Arenadata Hadoop (ADH)

Машина для обработки больших данных средствами экосистемы Hadoop

Скала^р МБД.КХ ← Arenadata QuickMarts (ADQM)

Машина для быстрых аналитических витрин с реляционным доступом
(на основе ClickHouse)



Ожидания клиентов



Государственные организации

- Пополнение информации об объектах госорганов
- Подготовка данных для государственной отчетности
- Сбор статистической информации для прогнозирования
- Выявление скрытых зависимостей и противодействие мошенничеству
- Сбор информации с региональных и местных уровней



Банковский сектор

- Анализ доходов и расходов клиентов
- Сегментирование клиентской базы
- Анализ рисков и предотвращение мошенничества
- Анализ отзывов клиентов для повышения лояльности
- Предективная аналитика



Ритейл и e-commerce

- Исследование потребительской корзины
- Анализ скорости покупки товаров и пополнение склада
- Исследование причин простоя торгового оборудования и касс
- Исследование товарных предпочтений различных групп клиентов
- Точность предзаказа и автозаказа



Телеком

- Выявление аудиторий клиентов для маркетинговых компаний
- Оптимизация ценового предложения
- Предотвращение риска мошенничества
- Превентивный мониторинг трафика и выявление проблем
- Получение данных о поведении сетевого оборудования

Отвечая потребностям бизнеса



Возможная интеграция с любыми источниками информации

- OLTP, ERP, CRM
- Документы и почтовые сообщения
- Журналы веб-серверов, потоки посещений
- Данные социальных сетей
- Журналы промышленных систем
- Данные сенсоров и датчиков



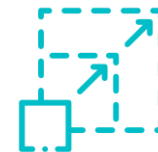
Выполнение бизнес-задач на неограниченном объеме данных

- Анализ больших данных которые постоянно пребывают измеряемые в ТБ
- Работа с данными типа «ключ – значение»
- Преобразование неструктурированных и частично структурированных данных в структурированную форму



Гарантия скорости обработки данных

- Использование параллельной обработки
- Распределенная обработка
- Инвариантная топология кластера для разных классов рабочих нагрузок



Горизонтальное масштабирование

- Рост системы в зависимости от потребностей бизнеса
- Расширение возможностей существующих аналитических систем
- Предельная загрузка оборудования
- Простое наращивание вычислительной мощности и полезной ёмкости системы

Управление Машинами больших данных Скала^р

Объединение лучшего оборудования
и программного обеспечения для построения КХД



Производительность
платформы:

от 1 000 000

запросов в секунду

Максимальная ёмкость:

более 100 Пбайт

Состав Машин больших данных Скала[^]р МБД.8



Основной функционал

Подсистема обеспечения базовых сервисов	Управление эксплуатацией: - Автоматизация процедур обслуживания - Мониторинг компонентов Машины	Экономия на эксплуатации
Сетевая подсистема	- Объединение всех компонентов - Обеспечение высокоскоростного взаимодействия - Создание отказоустойчивой сети	до 50% до 100 Гбит/с на порт
Основная, уникальная для каждой Машины функциональная подсистема	- Высокопроизводительные кластеры - Параллельные вычисления - Отказоустойчивая архитектура	от 2 вычислительных узлов
Подсистема управления	- Оркестрация запросов - Сервисные функции - Контроль и управление кластером	Интеллектуальное управление

Дополнительный функционал

Дополнительные подсистемы с уникальными возможностями, которые можно совмещать с основными	- Хранение резервных копий - Преобразование данных - Управление данными - Управление доступами к данным	Набор специализированного функционала: резервное копирование, ETL, Data Governance, безопасность и т.п.
--	---	---



Машина больших данных Скала^р МБД.Х гипермасштабируемая система

с применением продукта Arenadata Hadoop (ADH)



Сценарии применения

- Хранилище больших данных (озера данных)
- Унификация доступа к слабоструктурированным данным
- Применение в системах для машинного обучения и ИИ

Особенности

- Анализ больших данных
- Петабайты обрабатываемых данных
- Высокая доступность и отказоустойчивость

Замещаемые технологии

Oracle BigData Appliance, Teradata Appliance for Hadoop,
EMC DCA, Amazon Elastic MapReduce,
Google Cloud DataProc

Объем хранения данных
от 38 Тбайт
на Модуль,
до 30 Пбайт
на Машину

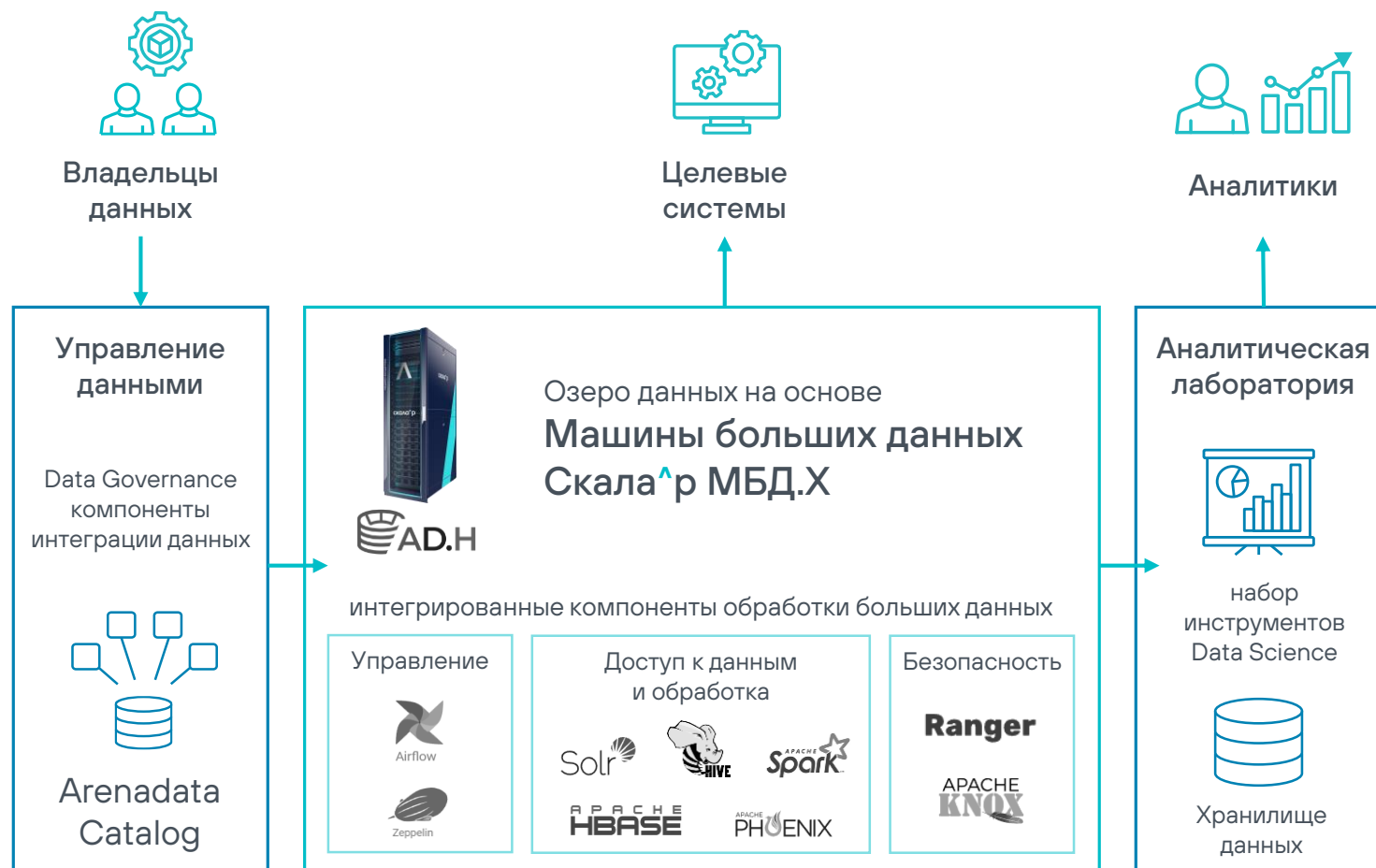
Сжатие данных
до 80%



Сценарий: Интеллектуальное озеро данных — полный контроль качества и чистоты данных



- Масштабируемое хранение больших данных
- Управление жизненным циклом данных
- Контроль изменений и обновлений
- Простое подключение к различным источникам
- Инструментарий управления данными
- Обеспечение безопасности данных
- Передача данных в связанные системы
- Наполнение корпоративного хранилища
- Скорость проработки гипотез и применение инструментов ИИ за счет встроенного ПО



Ожидания клиентов



Государственные организации

- Пополнение информации об объектах госорганов
- Подготовка данных для государственной отчетности
- Сбор статистической информации для прогнозирования
- Выявление скрытых зависимостей и противодействие мошенничеству
- Сбор информации с региональных и местных уровней



Банковский сектор

- Анализ доходов и расходов клиентов
- Сегментирование клиентской базы
- Анализ рисков и предотвращение мошенничества
- Анализ отзывов клиентов для повышения лояльности
- Предсказательная аналитика



Ритейл и e-commerce

- Исследование потребительской корзины
- Анализ скорости покупки товаров и пополнение склада
- Исследование причин простоя торгового оборудования и касс
- Исследование товарных предпочтений различных групп клиентов
- Точность предзаказа и автозаказа



Телеком

- Выявление аудиторий клиентов для маркетинговых компаний
- Оптимизация ценового предложения
- Предотвращение риска мошенничества
- Превентивный мониторинг трафика и выявление проблем
- Получение данных о поведении сетевого оборудования

Отвечая потребностям бизнеса



Возможная интеграция с любыми источниками информации

- OLTP, ERP, CRM
- Документы и почтовые сообщения
- Журналы веб-серверов, потоки посещений
- Данные социальных сетей
- Журналы промышленных систем
- Данные сенсоров и датчиков



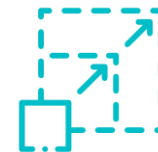
Выполнение бизнес-задач на неограниченном объеме данных

- Анализ больших данных которые постоянно пребывают измеряемые в Тбайт
- Работа с данными типа «ключ – значение»
- Преобразование неструктурированных и частично структурированных данных в структурированную форму



Гарантия скорости обработки данных

- Использование параллельной обработки
- Распределенная обработка
- Инвариантная топология кластера для разных классов рабочих нагрузок



Горизонтальное масштабирование

- Рост системы в зависимости от потребностей бизнеса
- Расширение возможностей существующих аналитических систем
- Предельная загрузка оборудования
- Простое наращивание вычислительной мощности и полезной ёмкости системы

Возможности Машины больших данных Скала^Λр МБД.Х



Полный комплект средств экосистемы Hadoop с корпоративным уровнем поддержки

Корпоративная платформа хранения и обработки данных

- Arenadata создала богатую экосистему data-сервисов, которая позволит заказчикам использовать не только Hadoop, но и полнофункциональные решения для обработки структурированных и потоковых данных

Собственная система управления

- Мы предоставляем открытую систему автоматического развёртывания и управления

Набор типовых пакетных сервисов по планированию, установке и аудиту системы

- Вам не придётся самостоятельно проводить оценку оборудования, которое потребуется для решения поставленной задачи. Наши специалисты настроят Arenadata Hadoop (удалённо или на месте), а в последствии проведут аудит системы и помогут разработать шаги для решения ваших проблем



Ключевые особенности и характеристики



Машина Скала^р МБД.Х для интеграции и обработки широкого спектра данных

Решаемые задачи:

- Объединение разрозненных данных из различных систем
- Интеграция данных из слабоструктурированных источников с другими системами
- Снижение времени внедрения для достижения новых возможностей
- Анализ быстро прибывающих больших данных
- Бизнес-запросы к сверхбольшим массивам, которые могут ожидать ответ около 1 минуты
- Работа с данными типа «ключ – значение»
- Извлечение данных и наборов текстовых журналов и преобразование слабоструктурированных данных в структурированную форму

Характеристики:

- Объём хранения данных от 30 Тбайт
- Встроенный мониторинг
- Практически неограниченное горизонтальное масштабирование

Может служить заменой импортных систем обработки больших данных



Преимущества Машины больших данных Скала[^]р МБД.Х



Масштабируемость

Способность хранить и обрабатывать терабайты больших данных, которые постоянно пополняются



Распределенные вычисления

Система позволяет распределять обработку запросов к данным одновременно на все узлы системы для быстрой подготовки результатов запросов



Управляемость

Дополнительные программные продукты и сервисы, позволяющие управлять и чувствовать каждый такт работы всей системы

Производительность



Максимальная скорость взаимодействия компонентов

Выделенный интерконнект

- Высокоскоростная высокопропускная сеть межсоединения ускоряет распределение заданий, ETL и ELT
- Параллельная обработка запросов на узлах обеспечивает суммирование мощностей всех узлов
- Создание параллельной синхронной копии не влияет на выполнение задания
- Все узлы кластера взаимодействуют между собой с одинаковой скоростью

Спрогнозированная нагрузка

- Производительность можно выбирать, подбирая нужное количество Модулей вычисления и хранения

RAID-защита для ведущих узлов

- Минимизация переключений на резервные узлы
- Управление процессорными потоками
- Минимальная просадка производительности в режиме восстановления

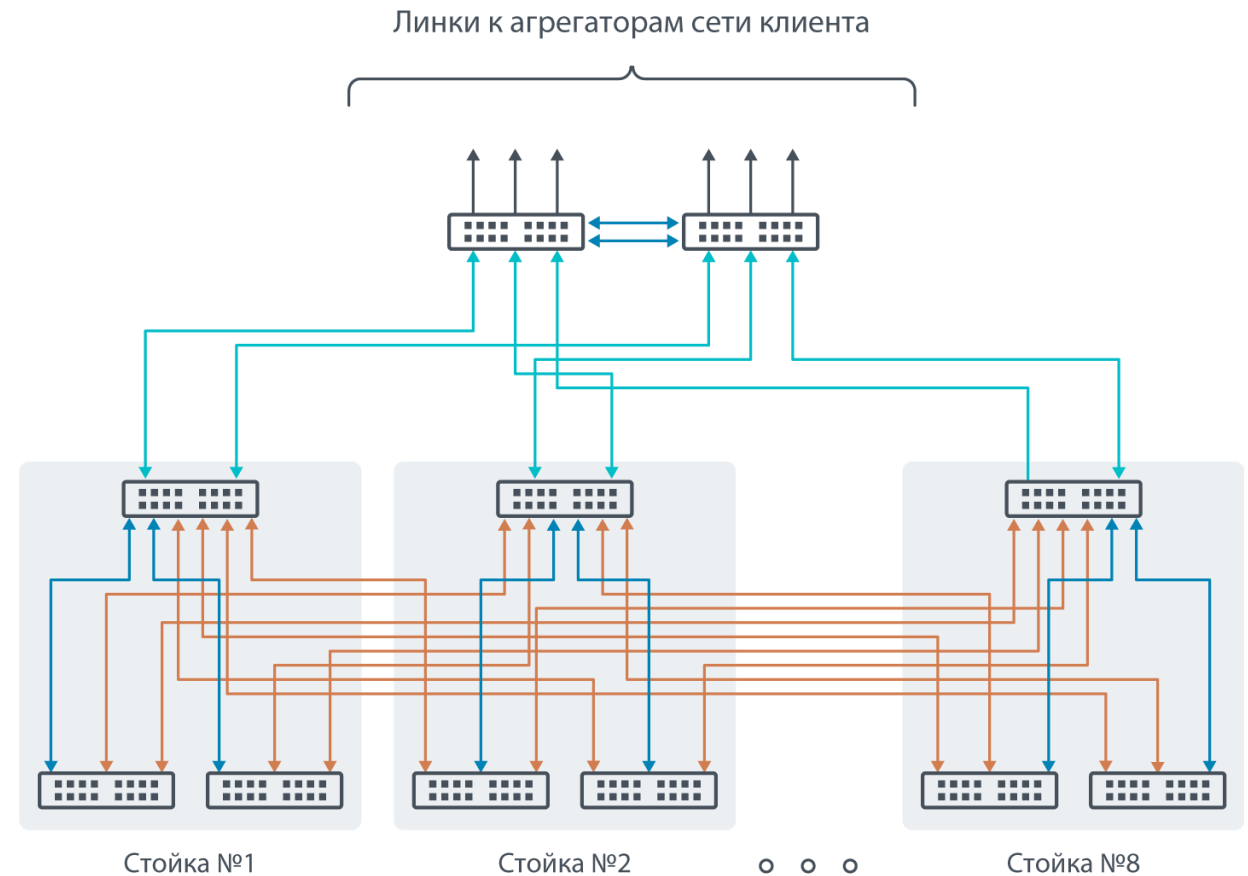


Схема внутренних соединений Leaf-Spine с увеличением скорости при горизонтальном масштабировании

Функциональная архитектура



Модульный принцип позволяет адаптироваться под любую нагрузку

Базовый модуль

- Коммутаторы (внешние соединения, интерконнекта)
- Узел мониторинга

Модуль управления и распределения

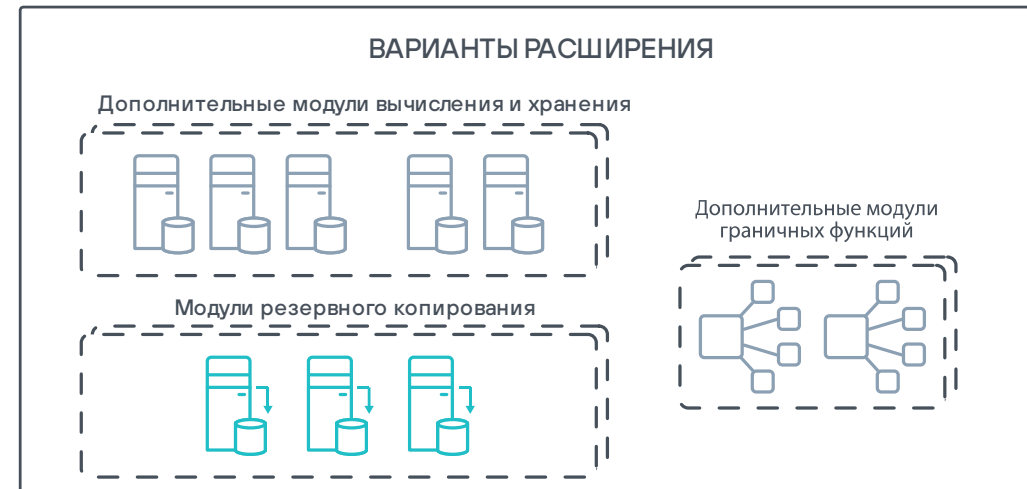
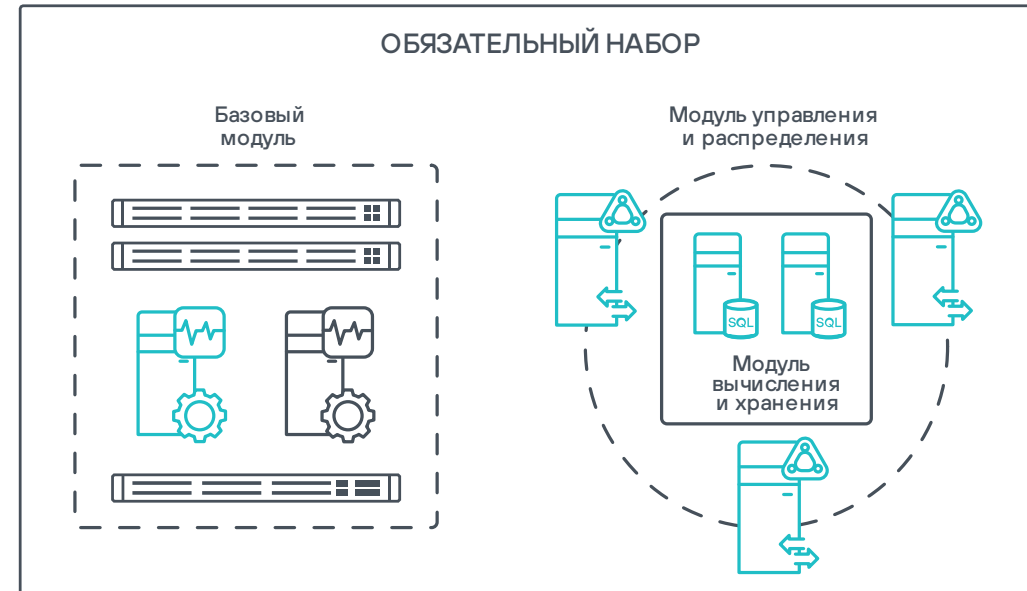
- 3 ведущих узла

Модуль вычисления и хранения

- 3 узла данных

Расширения

- Модули вычисления и хранения
- Модули резервного копирования
- Модули граничных функций



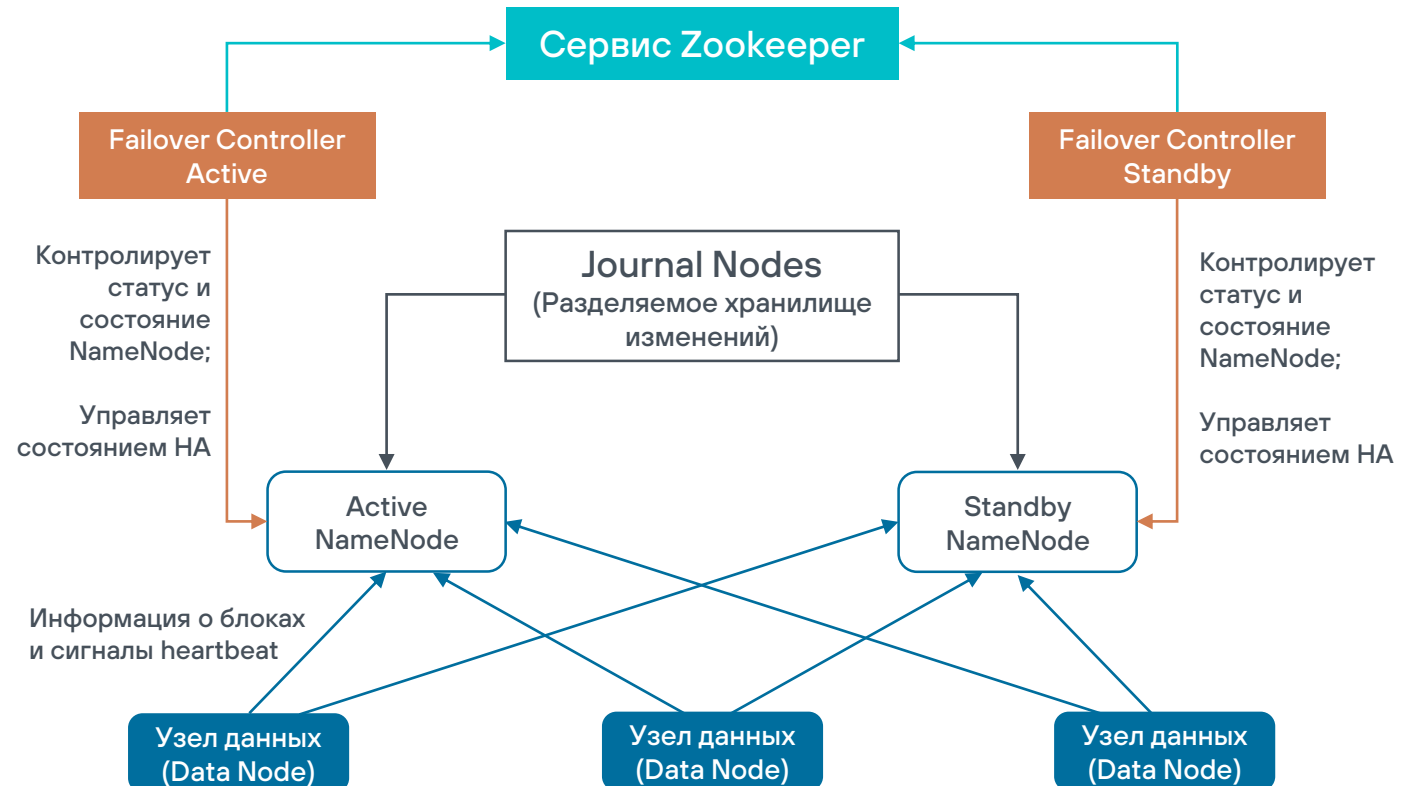
Отказоустойчивая архитектура



Позволяет гарантировать доступность сервисов

- Узлы Zookeeper обеспечивают координацию кластера, отслеживая:
 - состояние узлов имен
 - операционных систем
 - оборудования
- Узлы имен (NameNodes) работают в режиме «активный — резервный» (Active NameNode и Standby NameNode)
- Узлы данных (DataNodes) отвечают на запросы для хранения информации и объектов
- Распределённый журнал изменений ведётся на журнальных узлах (JournalNodes) — необходим для синхронизации активного и резервного узлов имен

Архитектура высокой доступности HDFS

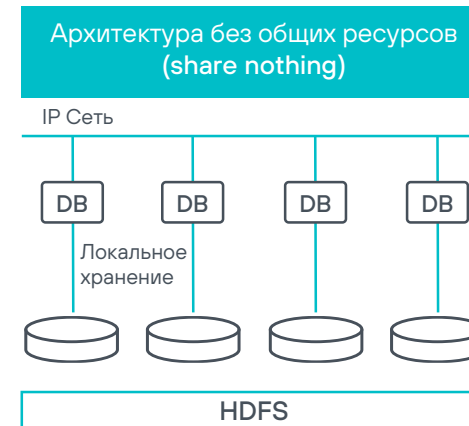


Доступность данных



Обеспечивается за счет отказоустойчивой архитектуры

- Архитектура без общих ресурсов (shared nothing)
- Каждый узел данных взаимодействует только со своими внутренними накопителями
- HDFS-блоки реплицируются 3 раза
- С целью повышения надежности для хранения 2-й и 3-й реплики выбираются те узлы данных, которые расположены в разных областях отказа (например, разных серверных стойках)
- Репликация выполняется в асинхронном режиме
- Проверка целостности данных находится в зоне ответственности приложения клиента. При создании файла рекомендуется рассчитывать контрольные суммы каждые 512 байт и сохранять их на узле имен
- HDFS (Hadoop Distributed File System) – файловая система, предназначенная для хранения файлов больших размеров, поблочно распределенных между узлами вычислительного кластера
- Позволяет производить вычисления параллельно для разных блоков на разных узлах

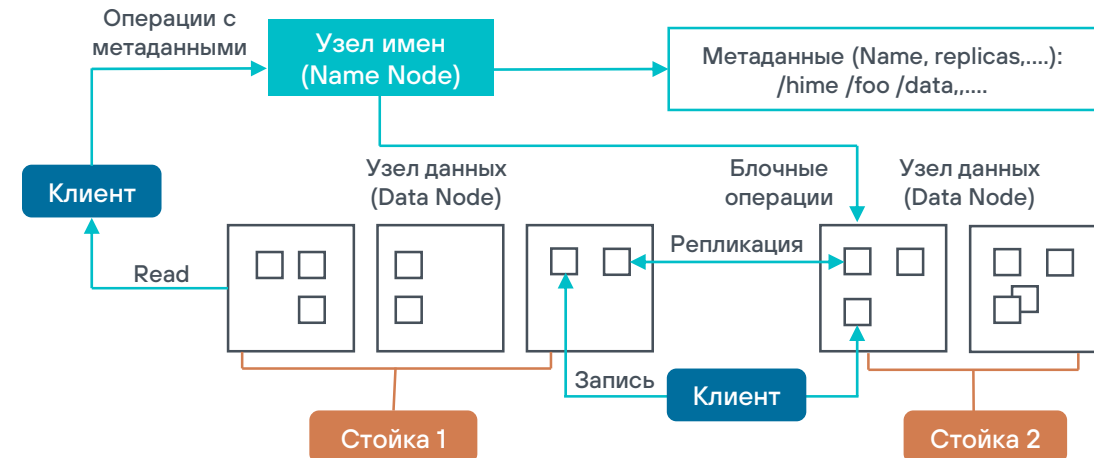


Узел имен – хранит метаданные файловой системы и метаинформацию о распределении Блоков, а так же список узлов данных.

Узел данных – непосредственно хранит блоки файлов. Узлы данных обрабатывают операции по записи и чтению данных.

Область отказа (в примере – стойка) – логическое объединение узлов данных согласно сетевой топологии кластера.

Архитектура HDFS



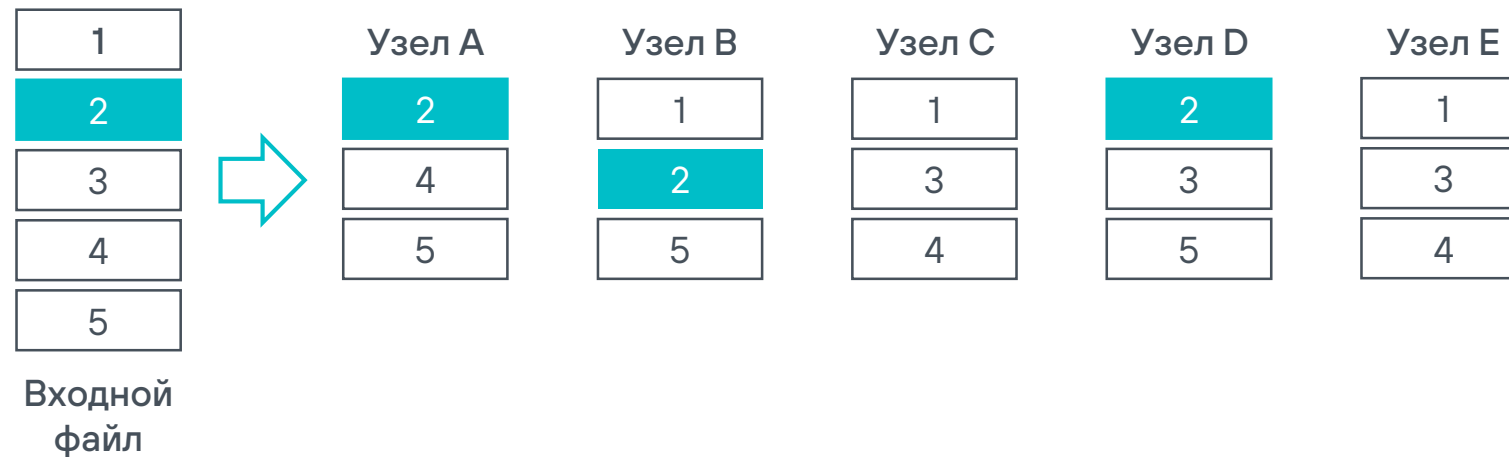
Репликация данных



Позволяет гарантировать защиту от одиночных сбоев

- Создание нового файла (операция записи)
- Обнаружение узлом имен отказа одного из узлов данных — если NameNode не получает от DataNode отклик на пульс-запросы, он запускает механизм репликации
- Повреждение существующих реплик
- Увеличение количества реплик, присущих каждому блоку

Распределение данных на HDFS



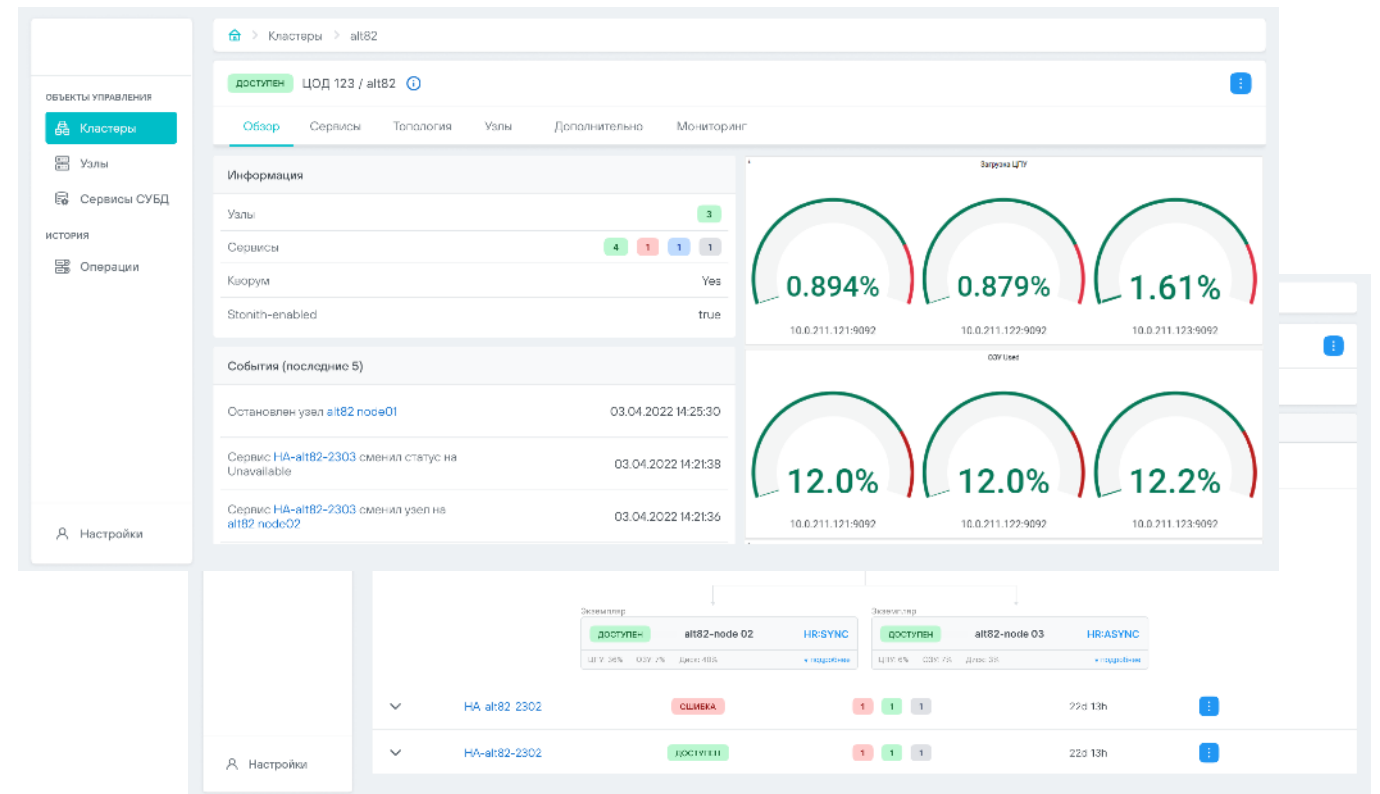
Управляемость

Система управления жизненным циклом Скала[^]р Геном



Данный программный продукт обеспечивает:

- Контроль развертывания компонентов Машины
- Ведение электронного паспорта Машины
- Отслеживание состояния узлов
- Отслеживание конфигурации программно-аппаратного состава Машины
- Снижение влияния человеческого фактора — сокращение рисков, связанных с ошибками эксплуатирующего персонала



Блок вычисления и хранения

Общий вид

Назначение:

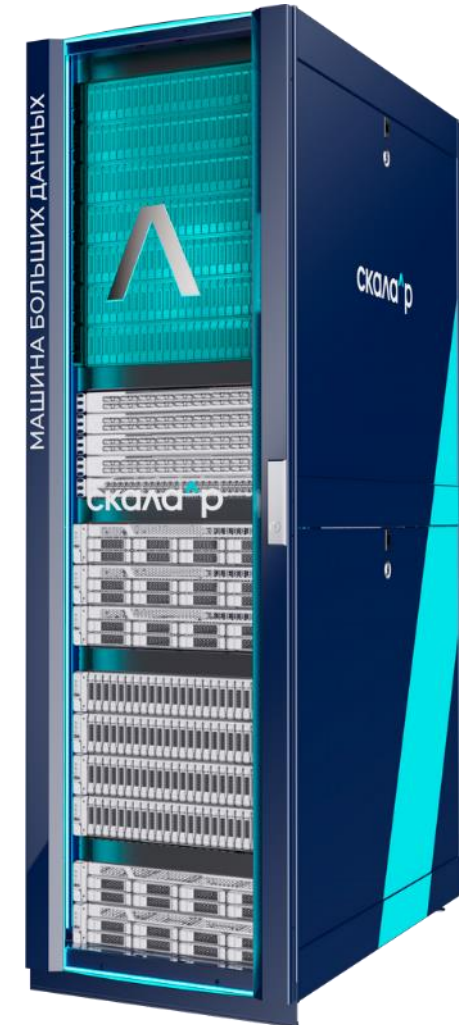
- Хранение данных и реплик
- Быстрое вычисление запросов

Модификации составляющих Модулей:

- По объему хранения и вычислений
- По производительности
- По назначению: для продуктивной среды или для разработки

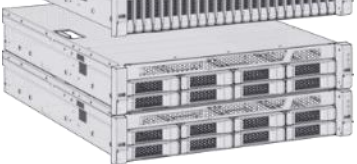
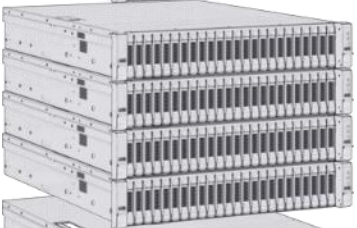
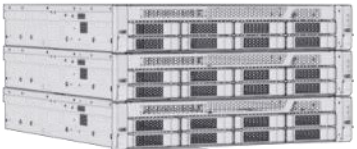
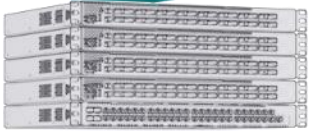
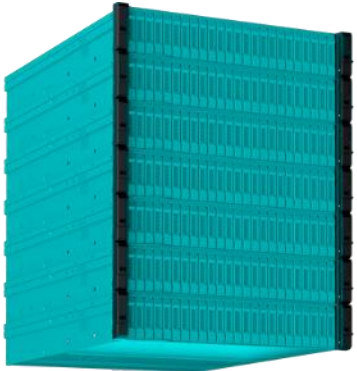
Расположение:

- В Базовом блоке
- В стойках расширения
- В Модулях расширения коммутации



Блок вычисления и хранения

Применимость и особенности



Применимость:

- По параметрам модулей данного Блока определяется производительность и объемы хранения МБД.Х
- Расширение производительного объема и повышение производительности всей системы в 50% случаев происходит за счет дополнения Модулей вычисления и хранения

Особенности:

- Самый высоконагруженный Блок в Машинах МБД.Х
- Хранение строится на SAS SSD 12G/24G или NVMe SSD
- Количество процессорных ядер — от 80 на модуль
- Оперативная память от 384 Гбайт до 1536 Гбайт на модуль в зависимости от исполнения

Блок коммутации и агрегации

Общий вид

Назначение:

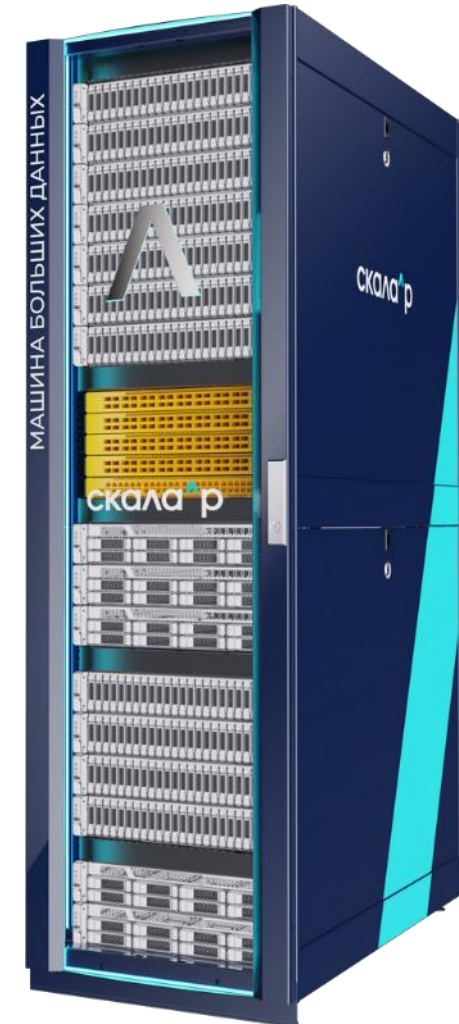
- Внутренний интерконнект на высокой скорости
- Агрегация по схеме Leaf-Spine или «звезда»
- Выделенная сеть для управления и мониторинга

Модификации составляющих Модулей:

- Модуль агрегации в Базовом блоке служит для соединения в одну сеть Модулей коммутации и связи с инфраструктурой
- Модуль коммутации в каждой активной стойке

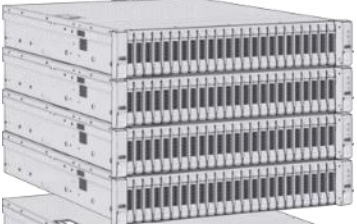
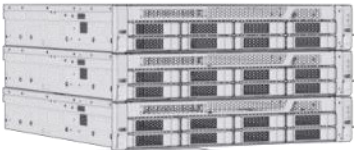
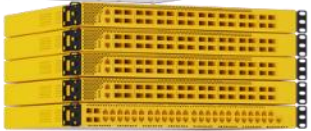
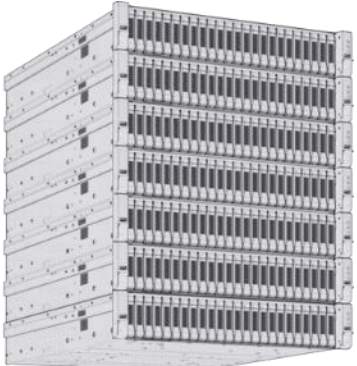
Расположение:

- В Базовом блоке
- В Модулях расширения коммутации



Блок коммутации и агрегации

Применимость и особенности



Применимость:

- Соединение с инфраструктурой заказчика
- Обеспечение скоростной внутренней коммутации
- Обеспечение отдельной сети для резервного копирования
- Обеспечение сетей для мониторинга и управления

Особенности:

- От трех до семи коммутаторов на стойку
- До трех параллельно действующих сетей для обеспечения отказоустойчивости

Блок управления и распределения

Общий вид

Назначение:

- Управление синхронизацией реплик
- Поддержание отказоустойчивого кластера

Модификации Модулей:

- Трехузловой кластер — стандартное решение
- Семиузловой кластер — решение для поддержания копий в резервном ЦОД с удаленным арбитром

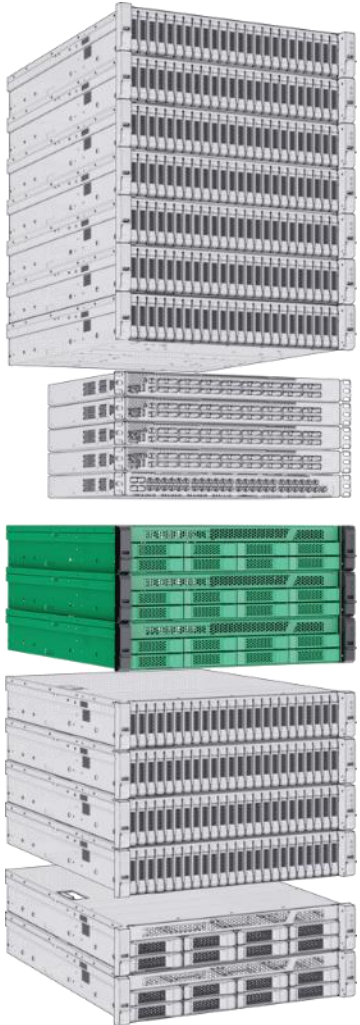
Расположение:

- В 99% случаев в Базовом модуле



Блок управления и распределения

Применимость и особенности



Применимость:

- Является основой для поддержания репликации данных
- Может быть расширен резервными узлами

Особенности:

- Зафиксированы оптимальные конфигурации
- В отдельных случаях может использовать внешние относительно Модуля базы данных для хранения метаданных

Блок мониторинга и регистрации

Общий вид

Назначение:

- Управление Машиной от бизнес-модели до конкретных аппаратных компонентов
- Управление развертыванием, обновлением, жизненным циклом Машины

Модификации составляющих Модулей:

- Один узел — абсолютный минимум без резервирования
- Два узла — взаимное резервирование с ручным переключением
- Четыре узла — стандартная отказоустойчивость высокой доступности с распределенным хранилищем

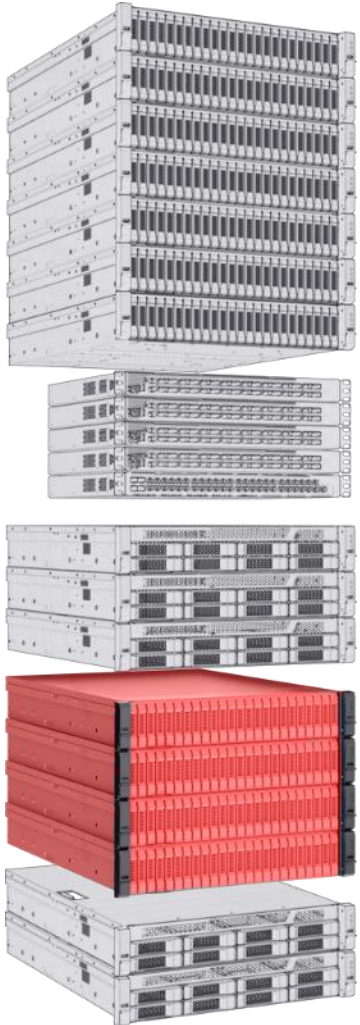
Расположение:

- В 99% случаев в Базовом модуле



Блок мониторинга и регистрации

Применимость и особенности



Применимость:

- Присутствует в любой Машине Больших Данных Скала^р
- Обязательно содержит ПО:
 - Скала^р Геном
 - Скала^р Визион
 - Arenadata Cluster Manager
 - Arenadata Enterprise Tools
 - Arenadata Monitoring

Особенности:

- Всегда в виртуальной среде
- Система управления виртуализацией входит в комплект

Блок резервного копирования

Общий вид

Назначение:

- Хранение резервных копий
- Хранение настроек и метаданных
- Пространство для ETL

Модификации составляющих Модулей:

- 2 размера базы
 - Неделя + инкременты
- 3 размера базы
 - Неделя + инкременты + текущий
- 4 размера базы
 - 2 недели + неделя + инкременты + текущий

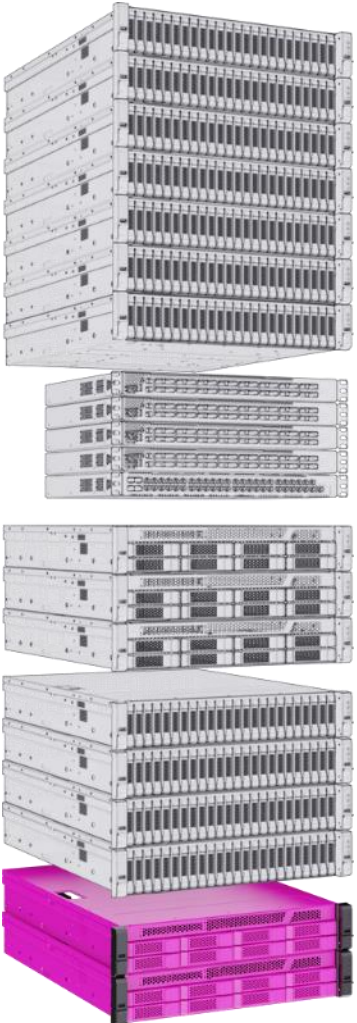
Расположение:

- В стойках Машины равномерно



Блок резервного копирования

Применимость и особенности



Применимость:

- Элемент, от которого можно отказаться с понижением надежности
- Возможно совмещение платформ для формирования теплого резерва (асинхронное копирование)
- Возможно использование для очень холодных данных

Особенности:

- Дисковое хранение
- RAID50
- Возможно подключение в выделенной параллельной сети
- Возможно иерархическое хранение (в разработке)

Модульность Скала[^]р МБД.Х



Основной функционал

	Базовый комплект		Комплект модулей расширения	
	Модули	Узлы	Модули	Узлы
Подсистема обеспечения базовых сервисов		2x Вычислительный узел мониторинга и регистрации		
Сетевая подсистема	Базовый модуль	Сетевые узлы управления Сетевые узлы внутреннего взаимодействия Сетевые узлы внешнего доступа (опция) Сетевые узлы агрегации Применяется при наличии комплекта модулей расширения при первой поставке	Модуль коммутации, вычисления и хранения	Сетевые узлы управления Сетевые узлы внутреннего взаимодействия Сетевые узлы внешнего доступа (опция) Сетевые узлы агрегации Применяется при первом расширении машины и в случае отсутствия узлов в базовом модуле
Подсистема неструктурированной обработки больших данных	Модуль неструктурированной обработки	Вычислительные узлы неструктурированной обработки данных	Модуль неструктурированной обработки	1-3 Вычислительные узлы неструктурированной обработки данных Вычислительные узлы неструктурированной обработки данных
Подсистема управления	Модуль координации	3x Вычислительные узлы координации (Zookeeper, Name Node и т.п.)		

Дополнительные функции

Подсистема резервного копирования	Модуль резервного копирования	Вычислительный узел резервного копирования + дисковая полка	Подсистема безопасности	Базовый модуль безопасности	Вычислительные узлы безопасности
Подсистема управления данными	Модуль управления данными	Вычислительные узлы для управления данными	Подсистема граничных функций	Модуль граничных функций	Вычислительные узлы граничных функций
Подсистема преобразования данных	Модуль преобразования данных	Вычислительные узлы для преобразования данных			

ПО в составе Скала^р МБД.Х

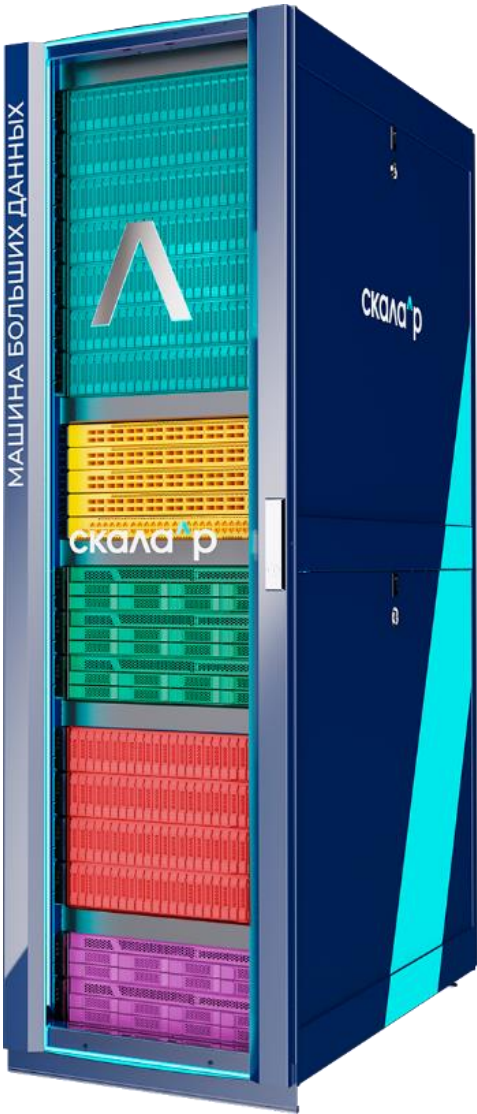


Основной функционал



Дополнительные функции

Подсистема резервного копирования	Модуль резервного копирования	NFS хранилище
Подсистема управления данными	Модуль управления данными	Arenadata Catalog
Подсистема преобразования данных	Модуль преобразования данных	ФормИТ, Cedrusdata
Подсистема управления доступом	Базовый модуль безопасности	ADPS
Подсистема обработки данных	Модуль граничных функций	Дополнительное ПО для работы с ADH



Модульность Скала[^]р МБД.Х



Основной функционал

	Базовый комплект		Комплект модулей расширения	
	Модули	Узлы	Модули	Узлы
Подсистема обеспечения базовых сервисов		2x Вычислительный узел мониторинга и регистрации		
Сетевая подсистема	Базовый модуль	Сетевые узлы управления Сетевые узлы внутреннего взаимодействия Сетевые узлы внешнего доступа (опция) Применимо при наличии комплекта модулей расширения при первой поставке Сетевые узлы агрегации	Модуль коммутации, вычисления и хранения	Сетевые узлы управления Сетевые узлы внутреннего взаимодействия Сетевые узлы внешнего доступа (опция) Применимо при первом расширении машины и в случае отсутствия узлов в базовом модуле Сетевые узлы агрегации
Подсистема неструктурированной обработки больших данных	Модуль неструктурированной обработки	Вычислительные узлы неструктурированной обработки данных	Модуль неструктурированной обработки	1-3 Вычислительные узлы неструктурированной обработки данных Вычислительные узлы неструктурированной обработки данных
Подсистема управления	Модуль координации	3x Вычислительные узлы координации (Zookeeper, Name Node и т.п.)		

Дополнительные функции

Подсистема резервного копирования	Модуль резервного копирования	Вычислительный узел резервного копирования + дисковая полка	Подсистема безопасности	Базовый модуль безопасности	Вычислительные узлы безопасности
Подсистема управления данными	Модуль управления данными	Вычислительные узлы для управления данными	Подсистема граничных функций	Модуль граничных функций	Вычислительные узлы граничных функций
Подсистема преобразования данных	Модуль преобразования данных	Вычислительные узлы для преобразования данных			

ПО в составе Скала^р МБД.Х

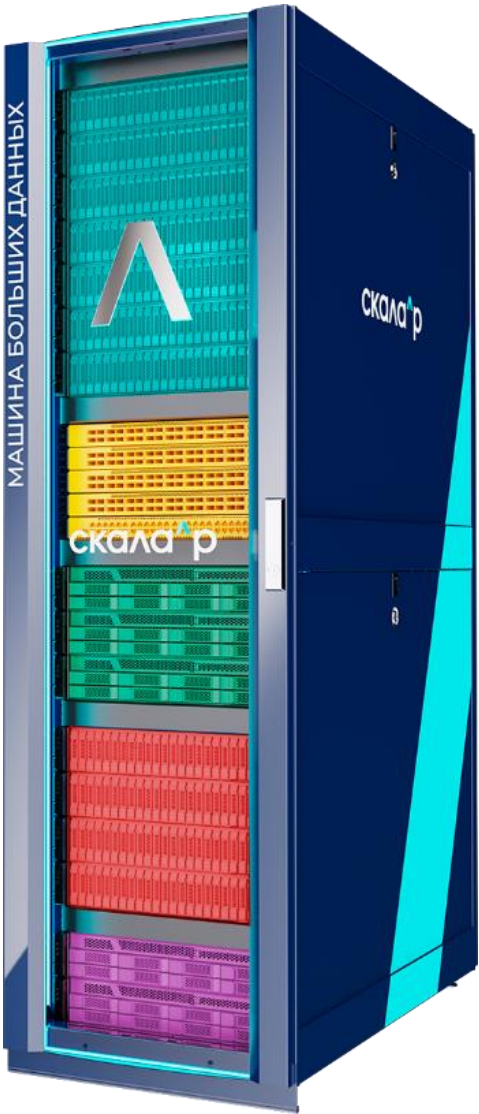


Основной функционал

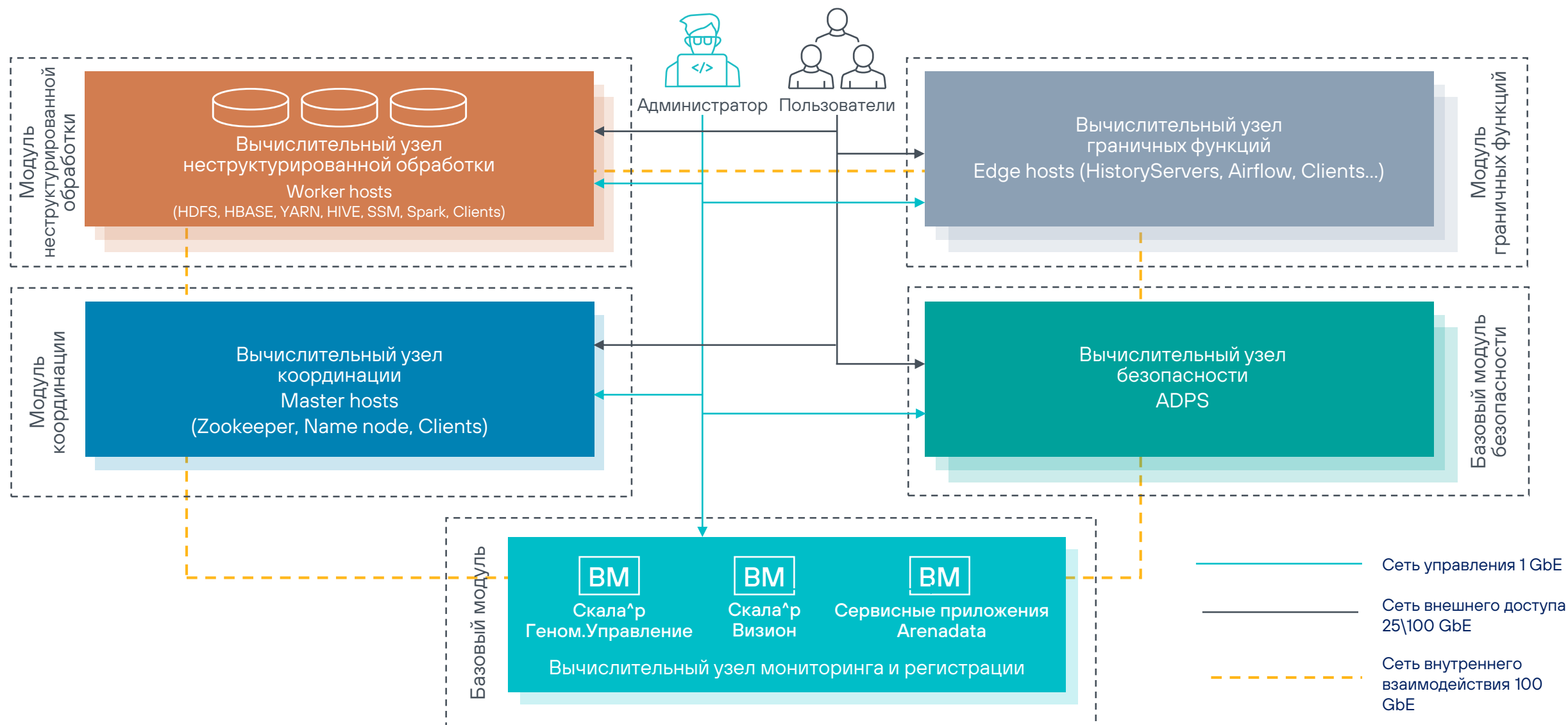
Подсистема обеспечения базовых сервисов	Базовый модуль Модуль коммутации, вычисления и хранения (при расширении)* Модуль неструктурированной обработки Модуль координации	Сервисное ПО Arenadata (ADCM и т.д.) ^геном ^ВИЗИОН
Сетевая подсистема		B4com и Eltex (микрокод)
Подсистема неструктурированной обработки данных		 Хранение данных HDFS, UI для работы с данными Работа с данными (Spark, Airflow и т.п.)
Подсистема управления		 Координаторы для контроля репликации данных (Zookeeper, Name Node и т.д.)

Дополнительные функции

Подсистема резервного копирования	Модуль резервного копирования	NFS хранилище
Подсистема управления данными	Модуль управления данными	Arenadata Catalog
Подсистема преобразования данных	Модуль преобразования данных	ФормИТ, Cedrusdata
Подсистема управления доступом	Базовый модуль безопасности	ADPS
Подсистема обработки данных	Модуль граничных функций	Дополнительное ПО для работы с ADH



Общая архитектура Скала^р МБД.Х



Аппаратная составляющая Машин Скала^р



Машина Скала^р из реестра МПТ РФ должна

- Быть собрана на едиобразных системных платах производства РФ
- Содержать российские компоненты, если только у них нет аналогов производства РФ
- Использовать устанавливаемое ПО из реестра Минцифры
- Компоненты машины имеют отказоустойчивую архитектуру

В Машинах Скала^р применяются:

OEM* серверы производства:



Сетевое оборудование и сетевые карты производства РФ



Выбранная системная плата (сервер) определяется по согласованию с Заказчиком Машины

* OEM — Original Equipment Manufacturer, производитель оригинального оборудования

Примеры внедрений Машин Скала^Ар



Цель

Сделано



Создание Единого корпоративного хранилища для повышения скорости и качества аналитической информации

Корпоративное хранилище данных



Создание Новой платформы данных

Новая платформа данных работает в продакшене



Создание Единого корпоративного хранилища и миграция с Oracle

Единое аналитическое хранилище



Объединить ключевые социальные сервисы в рамках Единой Централизованной Цифровой платформы, которая объединяет ключевые социальные сервисы для граждан РФ

Создание инфраструктуры и модернизация действующих систем и встраивание их в ГИС ЕЦП.

Техническая поддержка и услуги



Машины Скала[^]р поставляются с пакетами услуг технической поддержки:



техническая
поддержка из
«одного окна»

24×7

с поддержкой
служб эксплуатации
в круглосуточном режиме



возможность авансовой замены и ремонта
оборудования по месту установки;
опция невозврата накопителей с данными

1–5

лет
с возможностью
продления



Круглосуточно

- 8-800-234-23-25
- tac@skala-r.ru
- личный кабинет Service Desk
- <https://tac.skala-r.ru>



В программу поддержки входит:

- решение инцидентов
- консультации по эксплуатации Машин
- предоставление обновлений ПО



Дополнительные
профессиональные услуги



Программы дополнительных консультаций
администрирования и эксплуатации Машин

Почему Заказчики выбирают Скала^р



Глубокая интеграция и встречная оптимизация **компонентов** от платформенного ПО до микроконтроллеров:

- Высочайшая устойчивость
 - Экстремальная производительность
 - Стабильные показатели на предельных нагрузках
-
- Серийный выпуск, поддержка и сервисное обслуживание 24×7
 - Быстрое развертывание и ввод в эксплуатацию
 - Соответствие требованиям к критичным, высоконагруженным информационным системам
 - Снижение совокупной **стоимости** владения (TCO)





www.skala-r.ru

E-mail: info@skala-r.ru

skala^p