



Модульная платформа
для высоконагруженных
инфраструктур, частного облака и
управления данными

Машина больших данных Скала^р МБД.Х

скала^р

Скала^р — модульная платформа

продукт Группы Rubytech

Для построения инфраструктуры
высоконагруженных
корпоративных и государственных
информационных систем

скала^р

10 лет
серийного
выпуска

700+

комплексов в промышленной
эксплуатации

10 тыс. +
вычислительных
узлов

Продуктовые направления Скала^р

решения для высоконагруженных корпоративных и государственных систем



Динамическая инфраструктура



Машины динамической инфраструктуры Скала^р МДИ

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

Инфраструктура ИИ



Машина искусственного интеллекта Скала^р

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

Управление данными



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

Машины больших данных Скала^р МБД

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

Машины хранения данных Скала^р МХД

- на основе технологии объектного хранения S3 для геораспределенных катастрофоустойчивых систем с миллиардами объектов различного типа и обеспечения быстрого доступа к ним
- решения на основе платформы S3 и российского ПО для комплексных задач резервного копирования и восстановления крупных массивов данных со встроенной иерархией хранения и обеспечением высокой доступности копий

Специализированные решения



Машина управления технологическими процессами Скала^р МСП.ТП

- высоконадежная инфраструктура для различных АСУ ТП промышленных предприятий с высокими требованиями к отказоустойчивости и информационной безопасности. Соответствует требованиям ЗОКИИ, в том числе критериям к доверенным ПАК

Машина специализированных банковских систем Скала^р МСП.БС

- на платформе Машин Скала^р для задач класса АБС и процессинговых решений с поддержкой высокой транзакционной и аналитической нагрузки, сегментирования баз данных и обеспечения ИБ

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



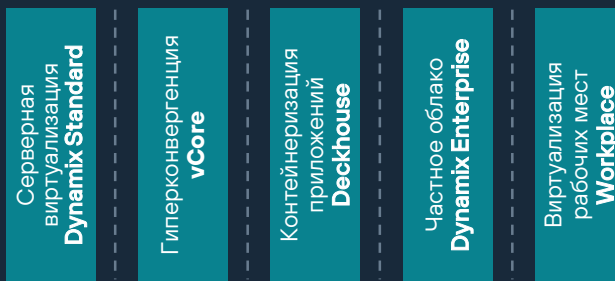
Мультитенантность



Автоматизация

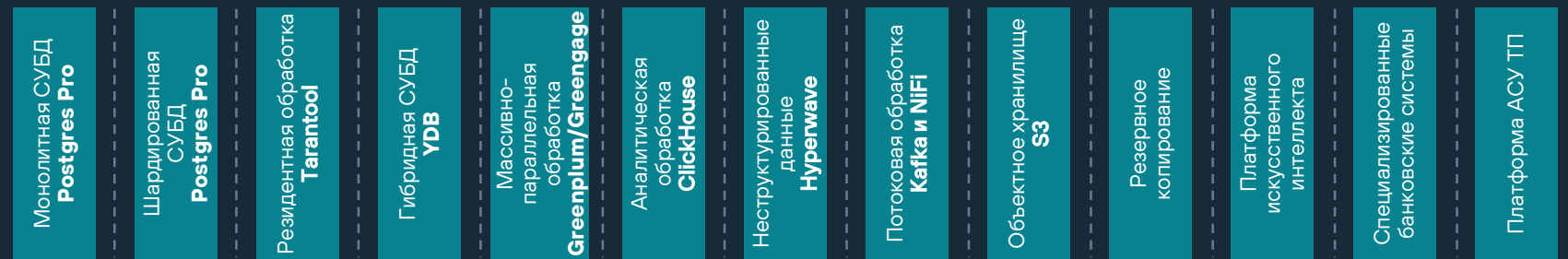
Модульная платформа

Динамическая инфраструктура



Динамическая инфраструктура

Инфраструктура управления данными



Транзакционная обработка

Управление данными

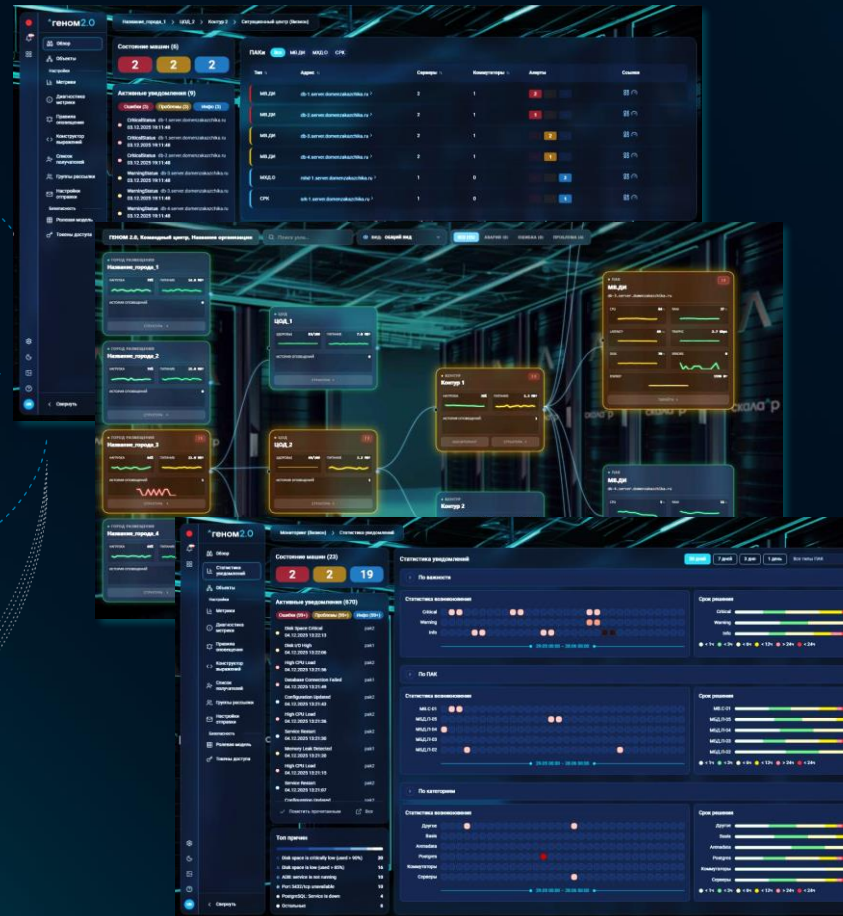
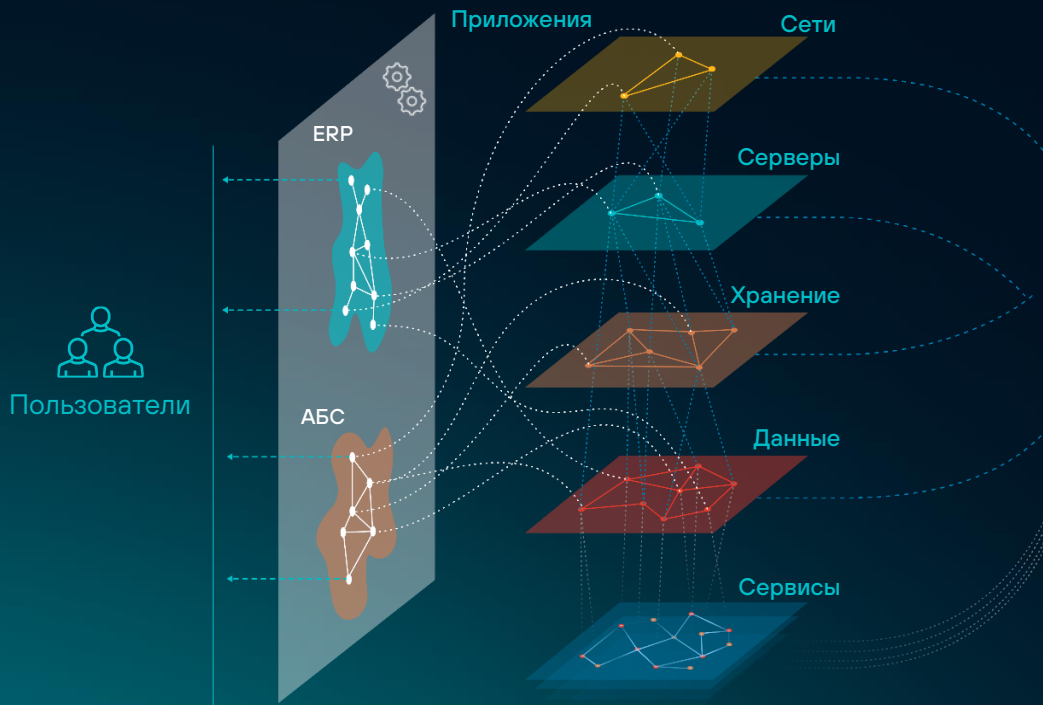
ИИ

Отраслевые решения

Развитие: Программная платформа Скала^р



объединение различных доменов управления в единую объектно-сервисную графовую модель - комплексное решение для эксплуатации инфраструктуры уровня ЦОД



- Единая точка обзора состояния контура
- Обозримость и удобство управления ЦОД
- Цифровой двойник инфраструктуры
- Контроль изменений оборудования и сервисов
- Моделирование изменений в инфраструктуре
- Высокая степень автоматизации

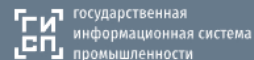
ПАК Скала^р в Реестрах РФ



Машины Скала^р (ПАК)

Модули Скала^р (ПАК)

Компоненты



государственная
информационная система
промышленности



Все сервисы ГИСП

Реестр промышленной продукции, произведенной на территории Российской Федерации

Программное обеспечение



РЕЕСТР
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Российский

Евразийский

ПАК Скала^р

Соответствуют критериям доверенного ПАК (ПП 1912)

- Технологическая независимость
- Информационная безопасность
- Функциональная устойчивость

Импортозамещение: сложность выбора

Отсутствие технологического лидерства



Глобальный ИТ-рынок

<p>Сетевая инфраструктура</p>	<p>Хранение данных</p>	<p>Виртуализация</p>
<p>Вычислительная инфраструктура</p>	<p>СУБД</p>	<p>Операционные системы</p>



Российский ИТ-рынок

<p>Сетевая инфраструктура</p>	<p>Хранение данных</p>	<p>Виртуализация</p>
<p>Вычислительная инфраструктура</p>	<p>СУБД</p>	<p>Операционные системы</p>

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 масштабирования
- Сокращение сроков внедрения и снижение рисков



Конкурентные преимущества оптимизированных решений Скала^р



Производительность

x2[↑]

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

x6[↑]

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 МБД.8 – это управляемая бизнес-инфраструктура



Гарантированная доступность

Высокая доступность и отказоустойчивость инфраструктуры из коробки

Предсказуемое масштабирование

Рост нагрузки без изменения архитектуры и с понятным бюджетом

Единая зона ответственности

Сопровождение оборудования и программного обеспечения всех компонентов Машин

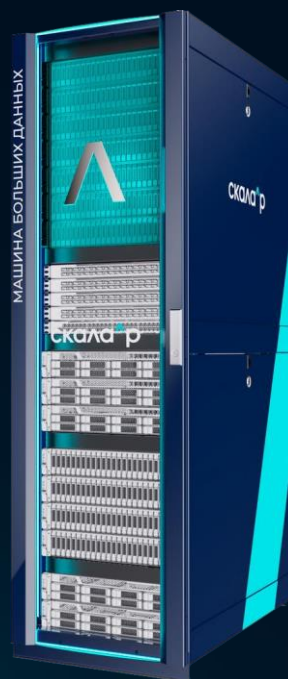
Экономия до 90%
на проектировании и внедрении

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-аппаратные комплексы и Модули платформы – включены
в Реестр российской промышленной и радиоэлектронной продукции, ПО Скала^p – в реестр Минцифры



Машины больших данных Скала[^]р МБД.8

Управление большими данными

высокопроизводительные хранилища и витрины данных
семейства Машин Скала^р МБД.8



Машина Скала ^р МБД.Г	Машина Скала ^р МБД.К	Машина Скала ^р МБД.С	Машина Скала ^р МБД.Х	Машина Скала ^р МБД.Т
Аналитическое хранилище данных	Витрины данных с реляционным доступом	Потоковая передача и обработка данных	Data Lake и неструктури- рованные данные	Обработка информации для задач реального времени
AD.B Arenadata DB	AD.QM Arenadata QuickMarts	AD.S Arenadata Streaming	AD.H Arenadata Hyperwave	PICODATA СУБД Picodata
Массивно- параллельная реляционная аналитическая СУБД на основе Greengage	Аналитическая обработка структурированных данных на основе колоночной БД ClickHouse	Система для потоковой обработки данных в режиме реального времени, на базе Apache Kafka и Apache NiFi	Экосистема сервисов Hyperwave для хранения, обработки и анализа больших объемов данных любого типа	Распределенные вычисления в оперативной памяти на основе Picodata



Управление Машинами больших данных Скала^р

Объединение лучшего оборудования и ПО для построения КХД



Производительность платформы

от 1 000 000

запросов в секунду

Максимальная ёмкость

> 100 Пбайт



Машина

Набор аппаратного и программного обеспечения в виде Модулей **Скала**^{^р}, соединенных вместе для обеспечения определенного метода обработки данных или предоставления ИТ сервиса с заданными характеристиками

Подсистемы

Логическое объединение компонентов Машины, собранных по функциональному признаку с целью пояснения состава и принципов действия машин

Модуль

Единица поставки Машин **Скала**^{^р}, функционально завершённый комплект аппаратных и-программных компонентов, выполняющих определенные функции в соответствии с их назначением.
Он является единым и неделимым элементом спецификации

Узел

Элемент Модуля, выполняющий определенную задачу в составе Модуля. Узлы размещаются в стойках

Состав Машин Скала^р линейки МБД.8



Используются максимально унифицированные Модули

Подсистема обеспечения базовых сервисов

Управление эксплуатацией:

- Автоматизация процедур обслуживания
- Мониторинг компонентов Машины

до **50%** Экономия на эксплуатации

Сетевая подсистема

- Объединение всех компонентов
- Обеспечение высокоскоростного взаимодействия
- Создание отказоустойчивой сети

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 кластерами

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.



Карта подсистем в МБД.8



Управление данными

Подсистема управления данными DG

Подсистема управления мастер данными (MDM)

Подсистема контроля качества данных DQ

Подсистема обезличивания данных Masking

Интеграция данных

МБД.С

Подсистема потоковой обработки больших данных

Подсистема обработки данных

Подсистема преобразования данных

Хранение и работа с данными

МБД.Т

Подсистема резидентной обработки

Оперативный кэш данных

МБД.КХ

Подсистема статистической обработки

Аналитика в реальном времени
Витрины данных

МБД.Г

Подсистема аналитической обработки

Data Warehouse:

- Детальный слой
- Витрины данных
- ELT-процессы

МБД.Х

Подсистема неструктурированной обработки

Data Lake (Legacy) / Lake House:
HDFS, HBase, Impala, Spark, Hive, Trino, Ozone (S3), Iceberg

Сохранность данных

Подсистема резервного копирования

Подсистема репликации данных

Базовые функции

Сетевая подсистема

Подсистема обеспечения базовых сервисов

Подсистема управления

Подсистема управления безопасностью

Подсистема авторизации

Легенда

Интеграция данных – набор подсистем, использующихся для загрузки и преобразования данных из внешних источников в МБД.8

Хранение и работа с данными – набор подсистем, выполняющих функции хранения и работы с разными типами данных

Базовые функции – набор подсистем, выполняющих функции обеспечения работоспособности и мониторинга подсистем, входящих в состав МБД.8

Управление данными – набор подсистем, которые позволяют упростить доступ к данным и их поиск, а также улучшить качество хранимых данных

Сохранность данных – набор подсистем, позволяющих организовать резервное хранение данных и репликацию данных в резервные ЦОД



Машина больших данных Скала[^]р МБД.Х

Машина Скала^р МБД.Х

Единая платформа для хранения и аналитики



Iceberg + Ozone (S3)/HDFS + SQL-движки

Сценарии применения

- Интерактивная аналитика в хранилище
- Построение витрин и BI без копирования данных
- Объединение batch и streaming обработки данных
- Единый доступ к данным для аналитики, BI и Data Science

Архитектура

- Ozone (S3) / HDFS — распределённое хранение данных
- Iceberg — таблицы с ACID, версионированием и управлением схемой данных
- Trino — распределённый SQL-движок (ad-hoc, BI)
- Spark — batch-обработка и подготовка данных

МБД.Х в сценарии Lakehouse объединяет хранение и аналитику данных, обеспечивая быстрый доступ к данным без дублирования и упрощая архитектуру аналитических систем

Особенности и ограничения

- Интерактивная аналитика без загрузки в DWH
- Снижение количества ETL и копий данных
- Независимое масштабирование хранения и вычислений без ограничений
- Поддержка batch, streaming и ad-hoc-сценариев
- Единая среда для хранения, обработки и анализа данных

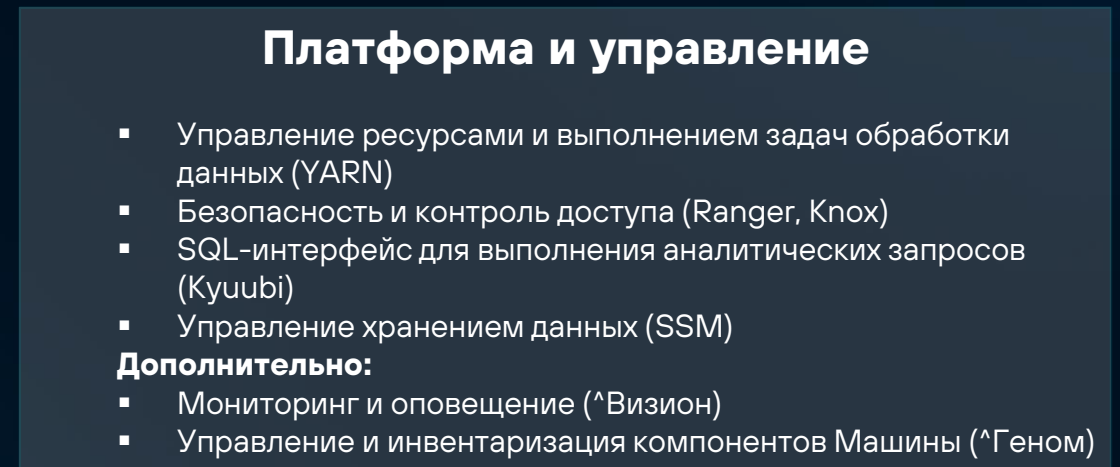
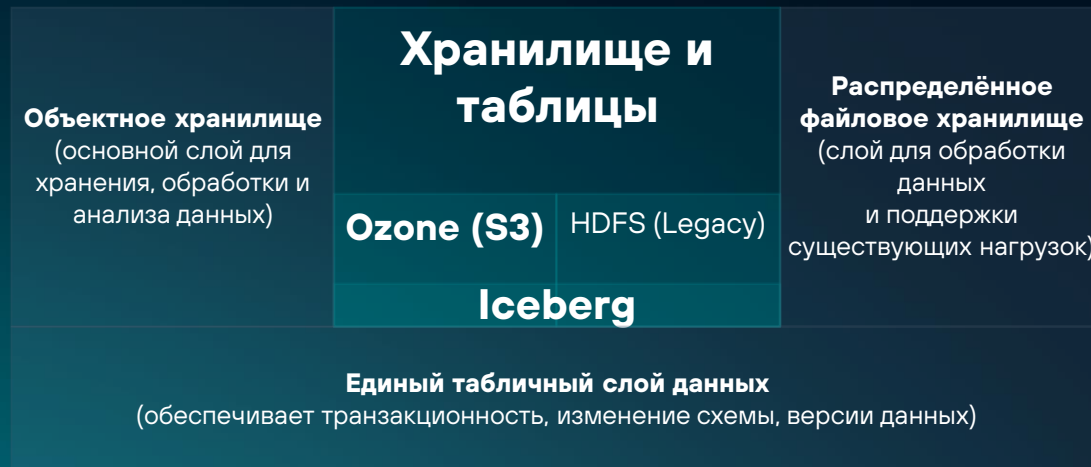
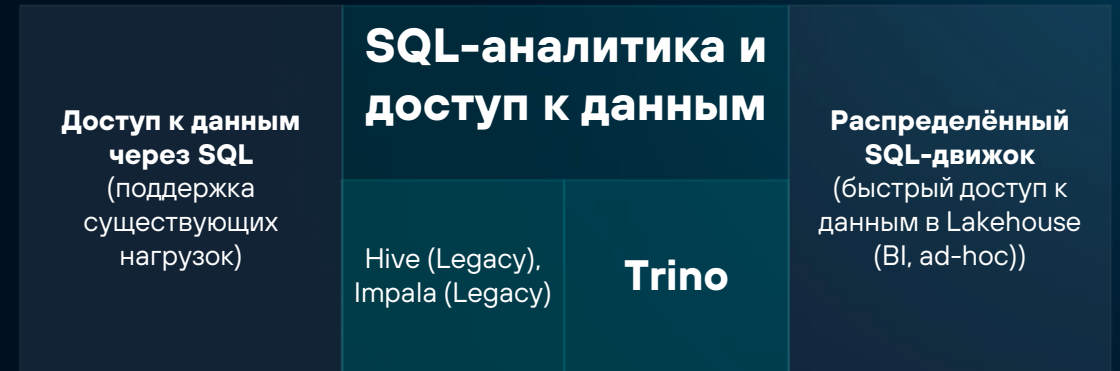
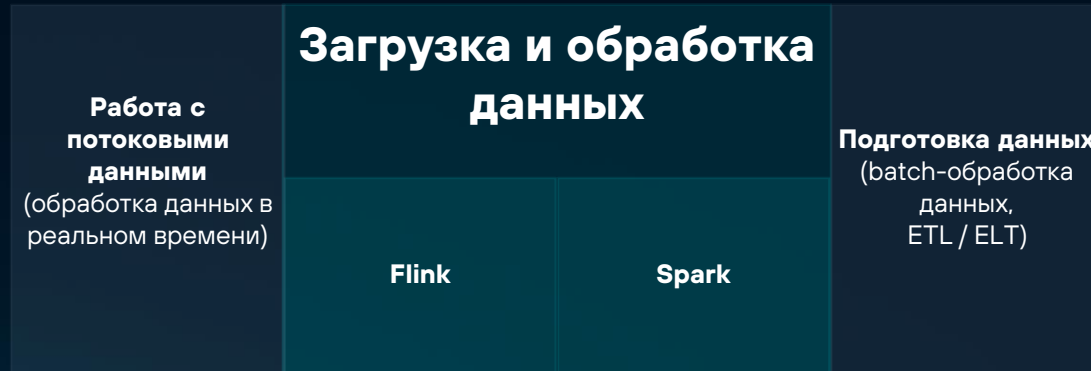
Объем хранения данных

от **38 Тбайт**
на Модуль,
без ограничений
на Машину

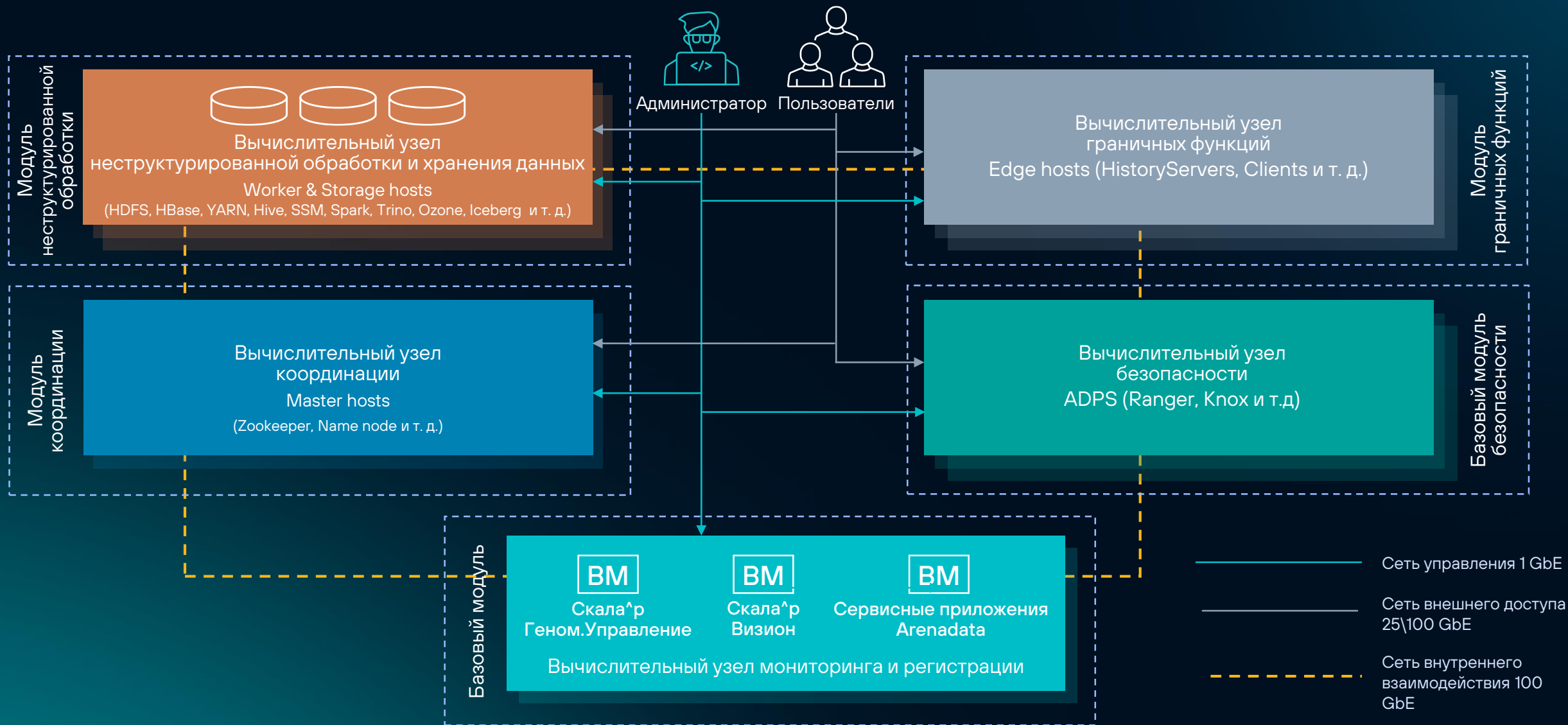
Избыточность
данных (при
отказоустойчивости)

60%

Основные функциональные компоненты МБД.Х



Общая архитектура Скала^р МБД.Х



Аппаратная составляющая Машин Скала^р

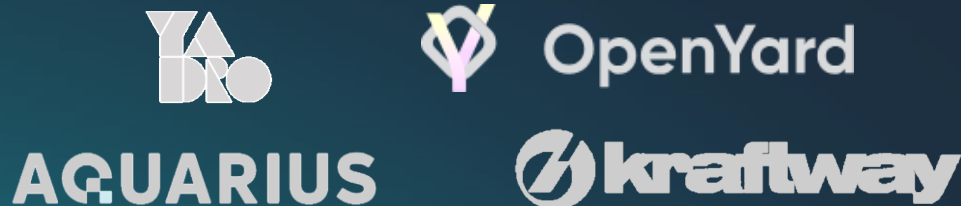


Машина Скала^р из реестра МПТ РФ должна:

- Быть собрана на единообразных 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



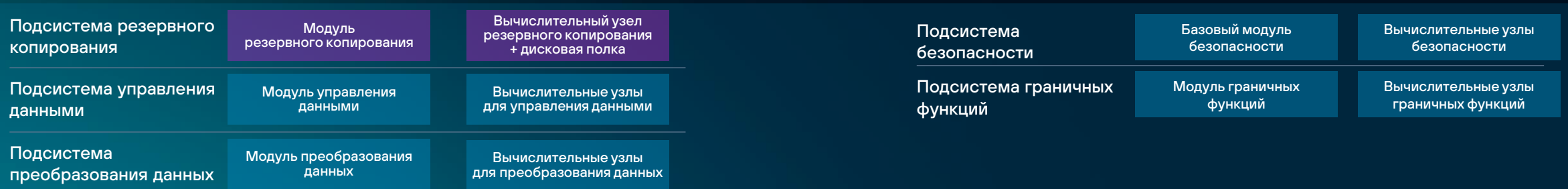
Выбранная системная плата (сервер) определяется по согласованию с Заказчиком Машины

* OEM — Original Equipment Manufacturer, производитель оригинального оборудования

Модульность Машины Скала[^]р МБД.Х



Дополнительный функционал



ПО в составе Машины Скала^p 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	ПО для управления ETL-процессами
Подсистема управления доступом	Базовый модуль безопасности	ADPS
Подсистема обработки данных	Модуль граничных функций	Дополнительное ПО для работы с ADH



Модули и компоненты Скала^р МБД.Х



Аппаратные составляющие

Серверы архитектуры x86-64

Платформы

- Yadro (R220 G3)
- OpenYard (RS2021)
- Аквариус
- Крафтвей

Сети и платы расширения

- Сетевые карты для интерконнекта: 2-х портовые (100 GbE) B4COM
- Сетевые карты для доступа к данным: 2-х портовые (25 GbE или 100GbE) B4COM
- Порты IPMI + порт 1 GbE (встроены в платформу)
- Бифуркатор Скала^р x16 для служебных дисков SSD NVMe M.2 и/или замка «Соболь»
- Сетевые узлы (коммутаторы) B4COM, ELTEX

Ключевые характеристики серверных платформ

- Имеют строго 2 CPU, формфактор 2U
- Новые поставки используют процессоры Intel Xeon Scalable IV/V поколения
- Оснащены накопителями 24 слота на передней панели
- Используется аппаратный рейд-контроллер Broadcom

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



МАШИНА БОЛЬШИХ ДАННЫХ



skala^p

skala^p



www.skala-r.ru

E-mail: info@skala-r.ru

skala^p