



Модульная платформа
для высоконагруженных
корпоративных и государственных
информационных систем

Машина баз данных Скала^р МБД.П

СКАЛА^р

Скала^р сегодня:



разработка и производство модульной платформы для высоконагруженных государственных и корпоративных информационных систем

9 лет

серийного
выпуска

600+

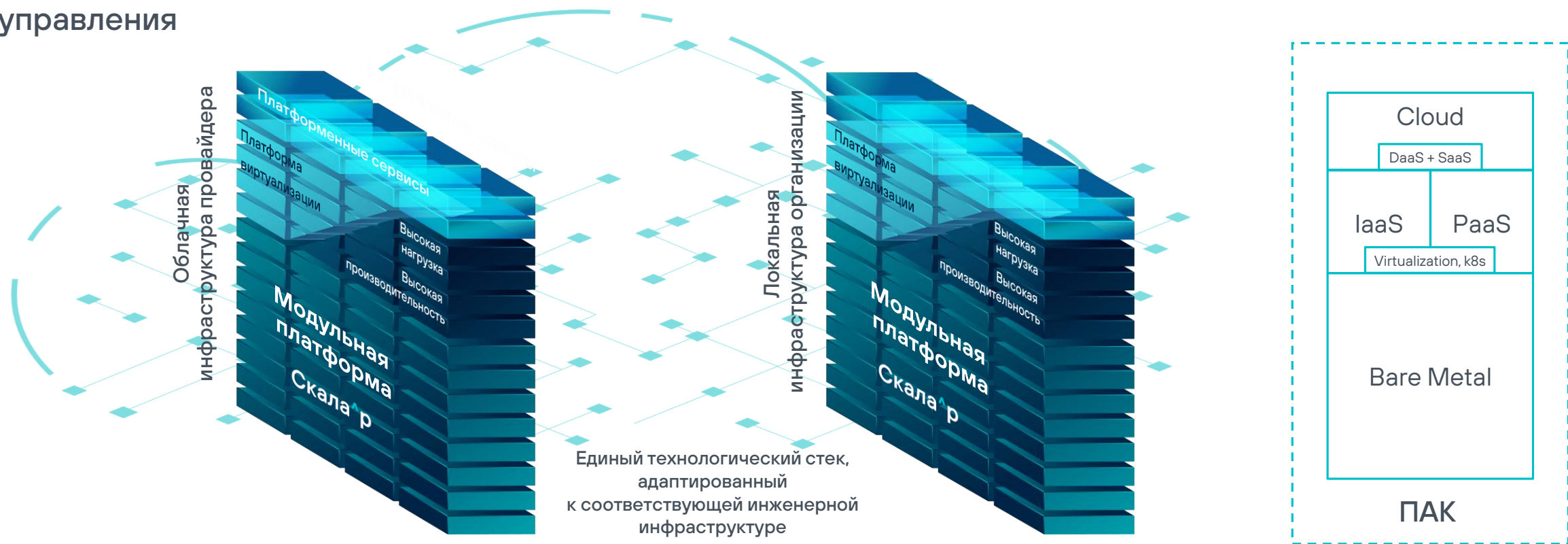
комплексов
в промышленной
эксплуатации

8 500+

вычислительных
узлов

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

Оптимизация под различные нагрузки и различный программный стек



Нагрузки

HPC AI/ML BPM Applications Containers Serverless IaC OLTP In memory BIG Data OLAP Streaming ETL Storage

Единая облачная система управления сервисами и интеграция с облачными платформами провайдеров

Модульная платформа



Единая система управления ресурсами и эксплуатацией

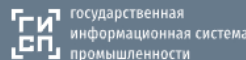
ПАК Скала[^]р в Реестрах РФ



Машины

Модули

Компоненты



государственная
информационная система
промышленности



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

Единая система управления ресурсами и эксплуатацией Машин



Цифровой двойник



Развертывание



Мониторинг



Диагностика



Оптимизация



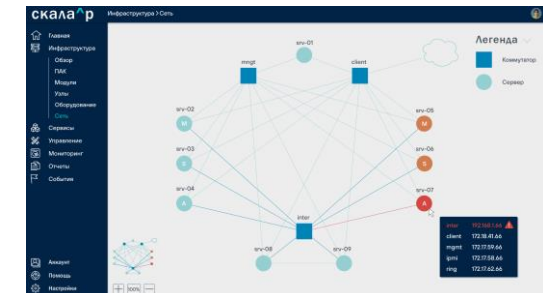
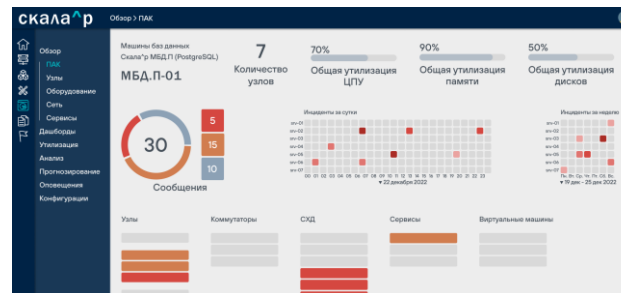
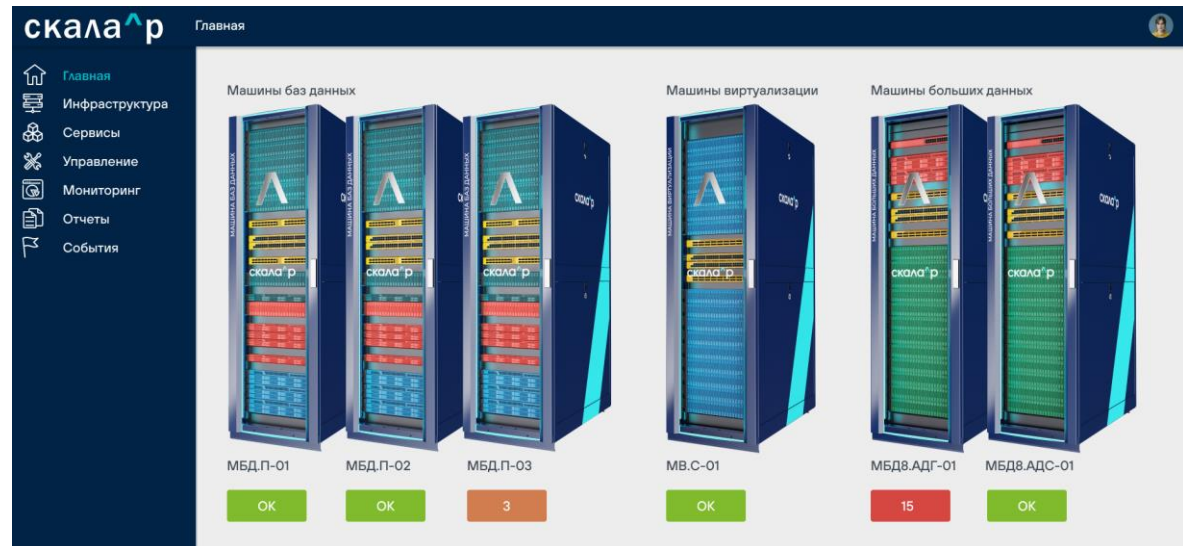
Инвентаризация



Обслуживание



Восстановление





Машина баз данных Скала[^]р МБД.П

Машина баз данных Скала^р МБД.П

Программно-аппаратный комплекс на основе СУБД Postgres Pro



Сценарии применения

- СУБД для частных и гибридных облаков
- СУБД для нагруженных приложений (1С ERP, АБС)
- Высоконагруженные отказоустойчивые системы
- Катастрофоустойчивые инфраструктуры

Особенности

- Высокая надежность и отказоустойчивость
- Высокая производительность
- Сеть внутреннего взаимодействия 100 Гбит/с с низкими задержками
- Встроенная система хранения данных со скоростью резервного копирования и восстановления от 13 до 44 Тбайт/час
- NUMA оптимизация
- Средства управления кластером
- Соответствие требованиям ФСТЭК

Замещаемые технологии

- Oracle Exadata

Рекомендовано
от 12 000
транзакций в секунду

от 1 до 160 Тбайт
размер баз данных

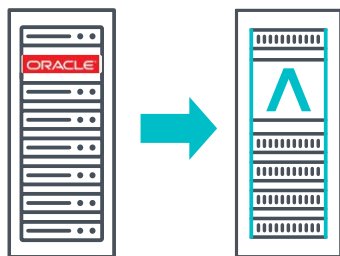
Пиковая
производительность
75 000+
транзакций/сек



Машина баз данных Скала[^]р МБД.П



Сценарии применения

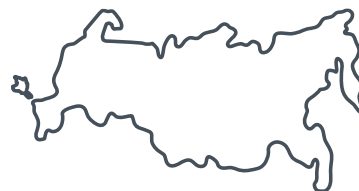


Замена Oracle
для транзакционных
нагрузок

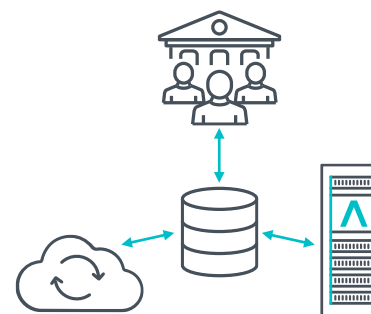
PostgreSQL



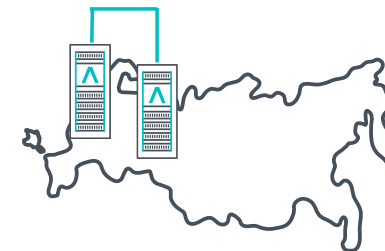
Высоконагруженные
системы
для скоростной
обработки транзакций
на основе Postgres Pro



Платформа
баз данных
для ответственных
приложений
федерального уровня
с высокими
требованиями
к надежности

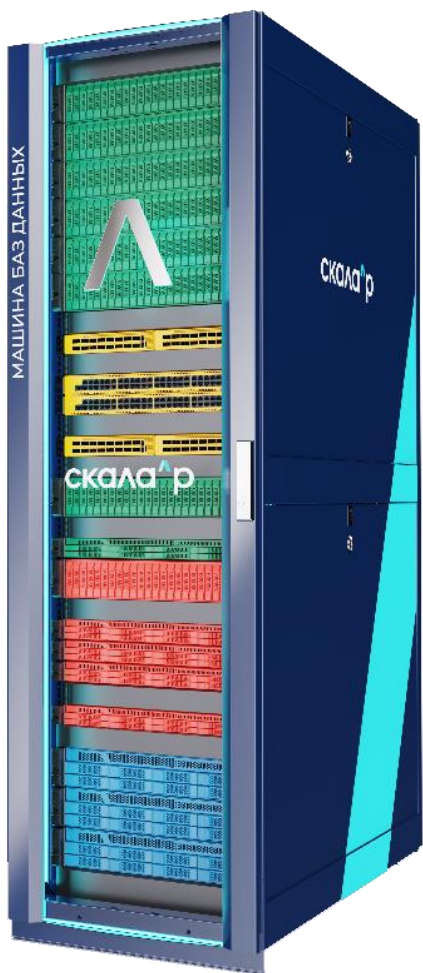


Создание
сервисов СУБД
для частных
и гибридных облаков
предприятий



Катастрофо-
устойчивые
инфраструктуры

Состав Машины Скала^р МБД.П



Блок вычисления и хранения

- Кластер из трёх узлов (мастер, синхронная реплика, асинхронная реплика)
- Повышение производительности прикладного ПО за счет распределения нагрузки между мастером и репликами

Блок коммутации и агрегации

- Сети внешнего доступа
- Сеть внутреннего взаимодействия до 100 Гбит/с
- Выделенная сеть для управления и мониторинга

Блок мониторинга и регистрации

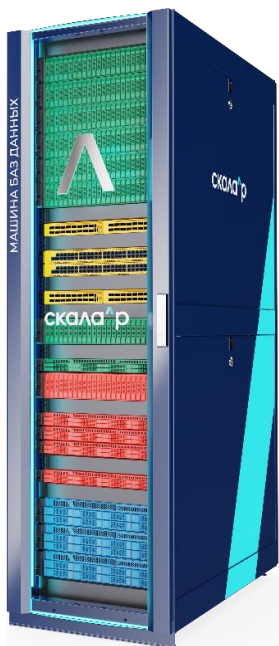
- Управление эксплуатацией и автоматизация критически важных процедур, снижение влияния человеческого фактора
- Мониторинг состояния всех компонент Машины

Блок резервного копирования

- Хранение резервных копий БД и WAL
- Отказоустойчивый кластер

Машина Скала^р МБД.П

Блоки и модульная архитектура



Блок вычисления и хранения

- Кластер из трёх узлов (мастер, синхронная реплика, асинхронная реплика)
- Повышение производительности за счет записи на мастер, а чтения с синхронной реплики

Модуль баз данных

3x Вычислительный узел баз данных

3 вычислительных узла, в каждом:

- 48 ядер 2.8 ГГц
- 1 Тбайт оперативной памяти
- От 4 до 160 Тбайт дискового пространства для БД

Блок коммутации и агрегации

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

Базовый модуль

2x Узлы внешнего доступа

2x Узлы сетевого взаимодействия

2x Сетевые узлы управления

2x Служебные узлы

Сетевой компонент:

- Узел внешнего доступа (до 100 Гбит/с)
- Узел сетевого взаимодействия (100 Гбит/с)

Компонент управления:

- Сетевой узел управления
- Служебный узел

Блок мониторинга и регистрации

- Управление эксплуатацией и автоматизация критических процедур
- Мониторинг состояния всех компонент Машины

Блок резервного копирования

- Хранение резервных копий БД и WAL
- Отказоустойчивый кластер

Модуль хранения

2x Отказоустойчивый узел хранения

1-2 Узла расширения хранения

Компонент хранения:

- Отказоустойчивый узел хранения
- Узел хранения на каждый Модуль БД - до 1 Пбайта

- Машина опционально может поставляться в телекоммуникационных шкафах
- Машина масштабируется Модулями баз данных и Модулями хранения

- до 4x Модулей баз данных
- до 2x Модулей хранения

Пути миграции с Oracle Exadata на Скала^р



Oracle Exadata консолидирует 3 типа нагрузки:

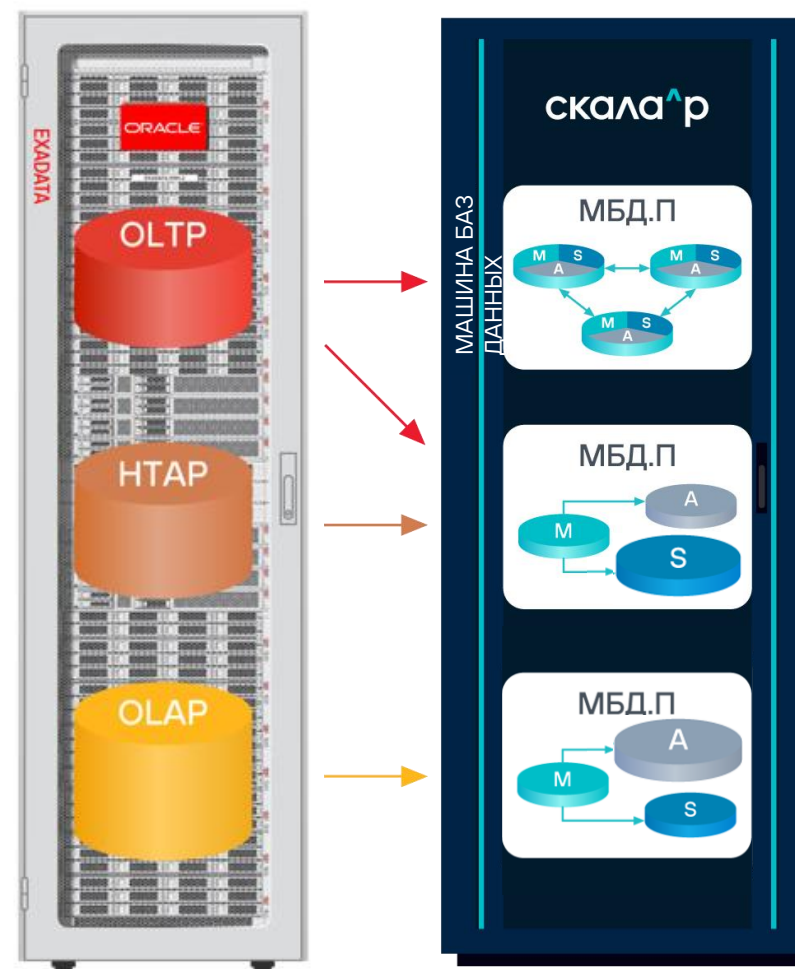
- Транзакционная (OLTP)
- Гибридная (HTAP)
- Аналитическая (OLAP)

Транзакционная и гибридная типы нагрузок мигрируют на Скала^р МБД.П с максимально возможным уровнем производительности, доступности и сохранности данных:

- OLTP – до 3 сервисов СУБД на кластер
- HTAP – 1 сервис СУБД, использование синхронной реплики

Аналитические нагрузки мигрируют на:

- Асинхронную реплику МБД.П



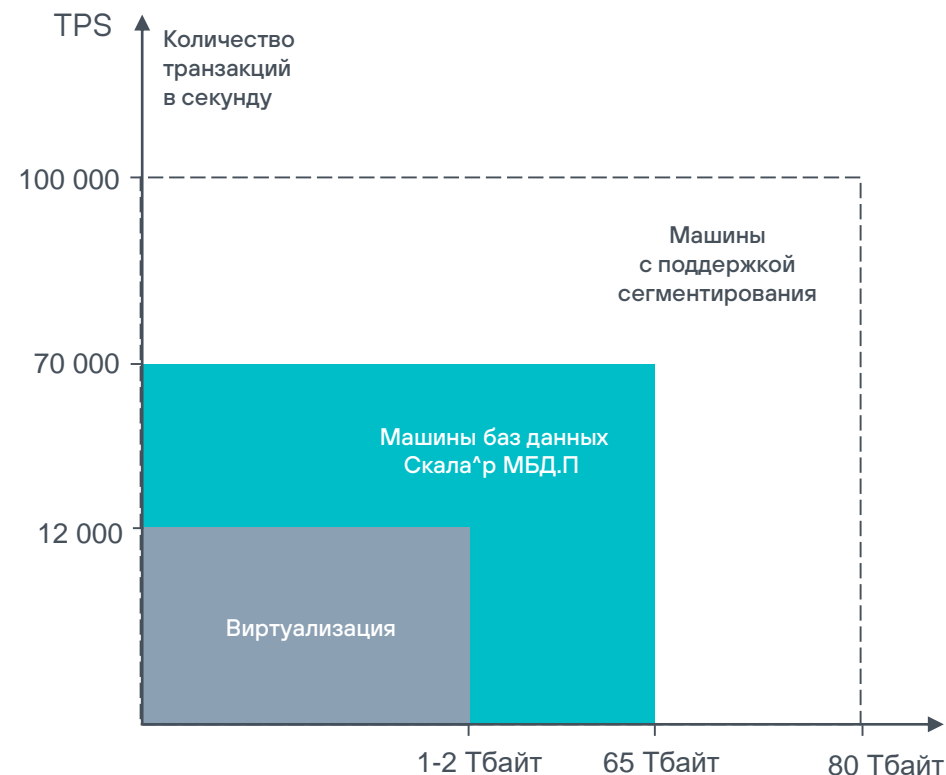
Машина баз данных Скала^р МБД.П



Границы применимости

- Машина баз данных Скала^р МБД.П обладает более высокой производительностью, чем в инфраструктуре с виртуализацией
- Виртуализация вносит дополнительные накладные расходы на операции ввода-вывода
- Использование виртуализации для СУБД оправдано в случаях незначительного объема данных или малых нагрузках на БД

Размер базы данных и нагрузка
(количество транзакций в секунду, TPS)



Ограничения для использования виртуальных сред для СУБД PostgreSQL: менее 12000 транзакций в секунду, и до 1-2 Тбайт хранимых данных

Машина баз данных Скала[^]р МБД.П



Технические показатели

Типовые размеры
ТОМОВ данных

Тип 1 до **80** Тбайт

Тип 2 до **40** Тбайт

Производительность*

до **75 000** TPS

Время обработки
сетевых отказов
уровня коммутации
не более

500 мс

Время переключения
сервиса СУБД между
узлами в случае отказа
не более

1-2 минуты

Реальная
пропускная способность
каждого узла СУБД
составляет
не менее

2 Гбит/с

на каждые 10 Тбайт
объема хранения

* На объеме горячих данных 1 Тбайт (объем с которым постоянно работает приложение) и до 5000 пользователей

Машина баз данных Скала[^]р МБД.П



Конкурентные преимущества

Производительность

x2[↑]

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

x4[↑]

чем решения в виртуальной среде, использующие сопоставимые аппаратные средства

x4[↑]

при использовании пулера соединений за счет использования специализированного решения

RTO, RPO

x4[↓]

время выполнения резервного копирования и восстановления баз 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

Число инцидентов

Кратное сокращение инцидентов, связанных с ошибками эксплуатации, за счет использования специализированных средств управления и превентивного мониторинга Машины баз данных

Архитектурные принципы обеспечения высоких значений целевых показателей



Производительность

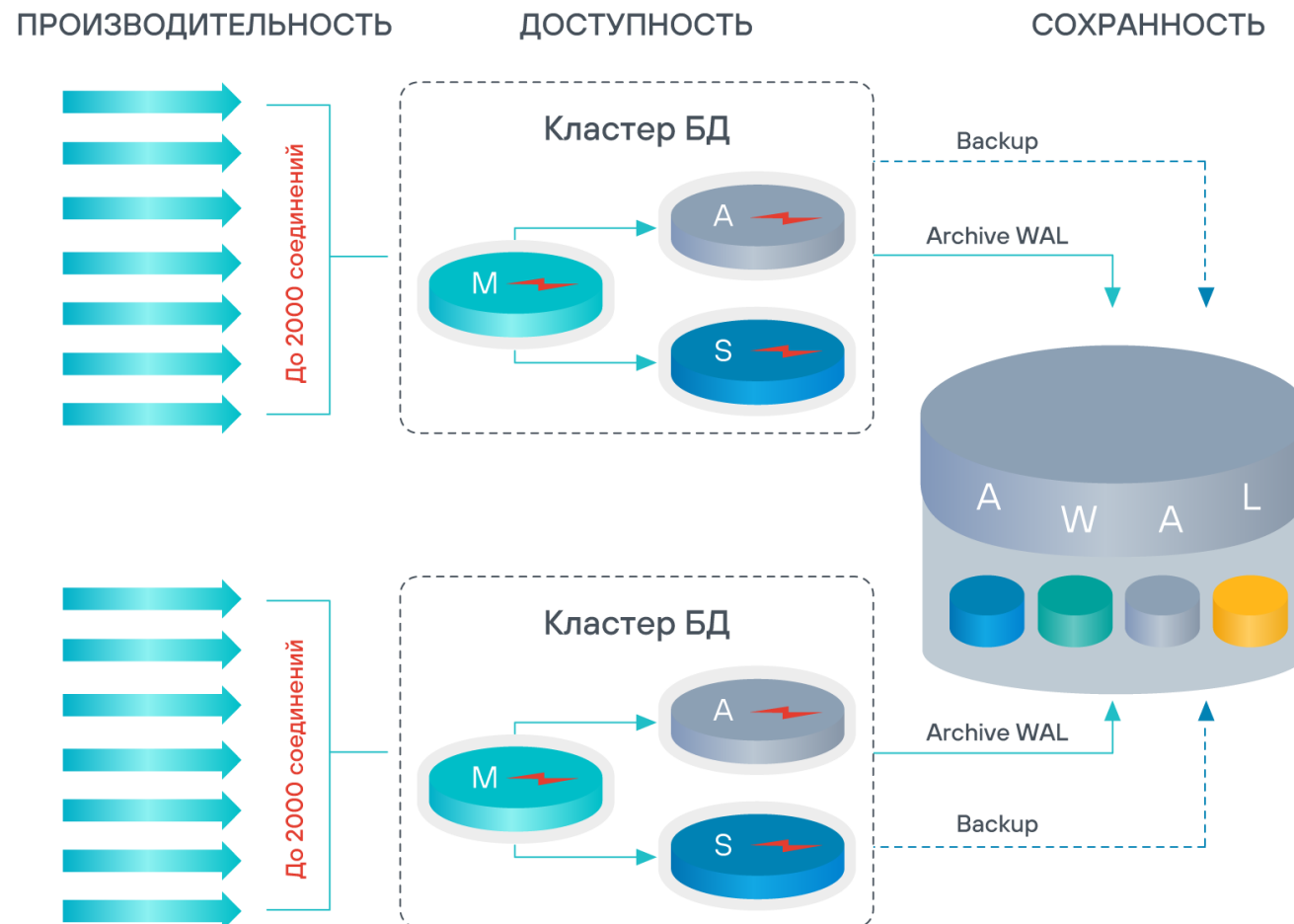
- Оптимизированная локальная система хранения
- Высокоскоростная сеть внутреннего взаимодействия
- Чтение с синхронной реплики

Доступность

- Кластерное ПО
- Программный RAID

Сохранность

- Полные и инкрементальные копии БД
- Хранение архивных журналов
- Защита данных программным RAID



Сохранность данных

Встроенный функционал резервного копирования



Резервное копирование БД

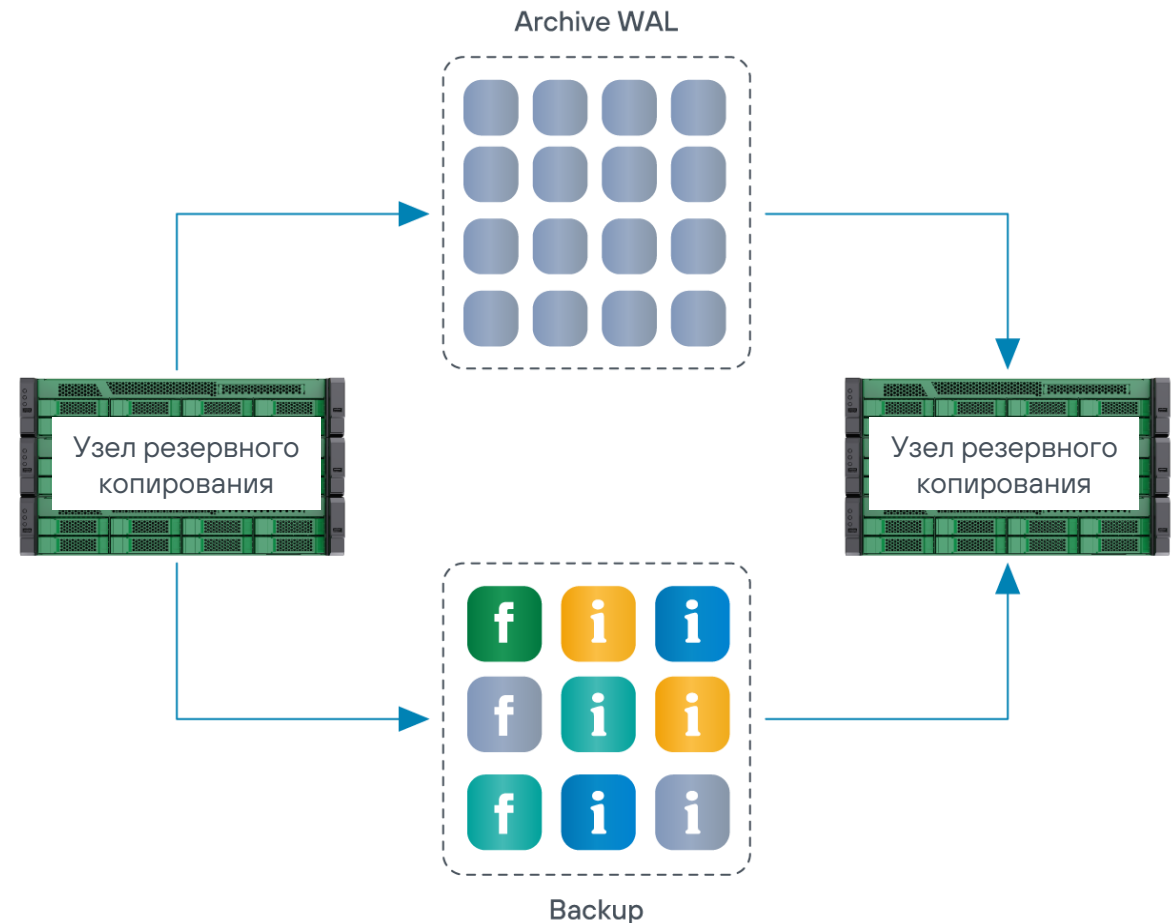
- Сохранность БД на внешних носителях
- Полное и инкрементальное копирование
- Кратное увеличение скорости копирования и восстановления за счет специализированного решения

Архивирование журналов предзаписи

- Доступность архива журнала для восстановления на момент времени (point-in-time recovery)
- Возможность восстановления ведомого сервера БД при большом отставании от ведущего сервера

Использование программного RAID

- Защита от ошибок записи (write hole)
- Уменьшение физического износа диска



Машина баз данных Скала^р МБД.П



TPC-B и pgbench для оценки конкурентных преимуществ

Тест TPC-B

Тест TPC-B — эталонный тест для оценки производительности СУБД и справедливого сравнения платформ между собой. Его результаты демонстрируют совместную производительность СУБД и аппаратной платформы и позволяют заказчикам выбрать платформу с оптимальным соотношением цена / производительность для решения своих задач. Тест был разработан независимым Советом по оценке производительности обработки транзакций и благодаря своей простой логике считается в сообществе пользователей Postgres стандартным.

Утилита pgbench

Pgbench — это программа для запуска тестов производительности PostgreSQL. Она многократно выполняет одну последовательность команд, а затем вычисляет среднюю скорость транзакций (число транзакций в секунду). Помимо TPC-B, pgbench включает тесты типов TPC-C и TPC-E с более сложной логикой.

Тест типа **TPC-B** хорошо характеризует «отзывчивость» производительности СУБД на более производительное оборудование - он минимально задействует механизмы блокировок, не использует сложные транзакции, откаты и т.п. В ходе теста TPC-B моделируется сеть контрагентов и транзакций между ними. Производительность определяется как число зафиксированных транзакций в секунду по тесту B (tpsB).

x2 производительность

по сравнению с аналогичным оборудованием



Сравнительные тесты Машины баз данных
Скала[^]р МБД.П и сопоставимых решений

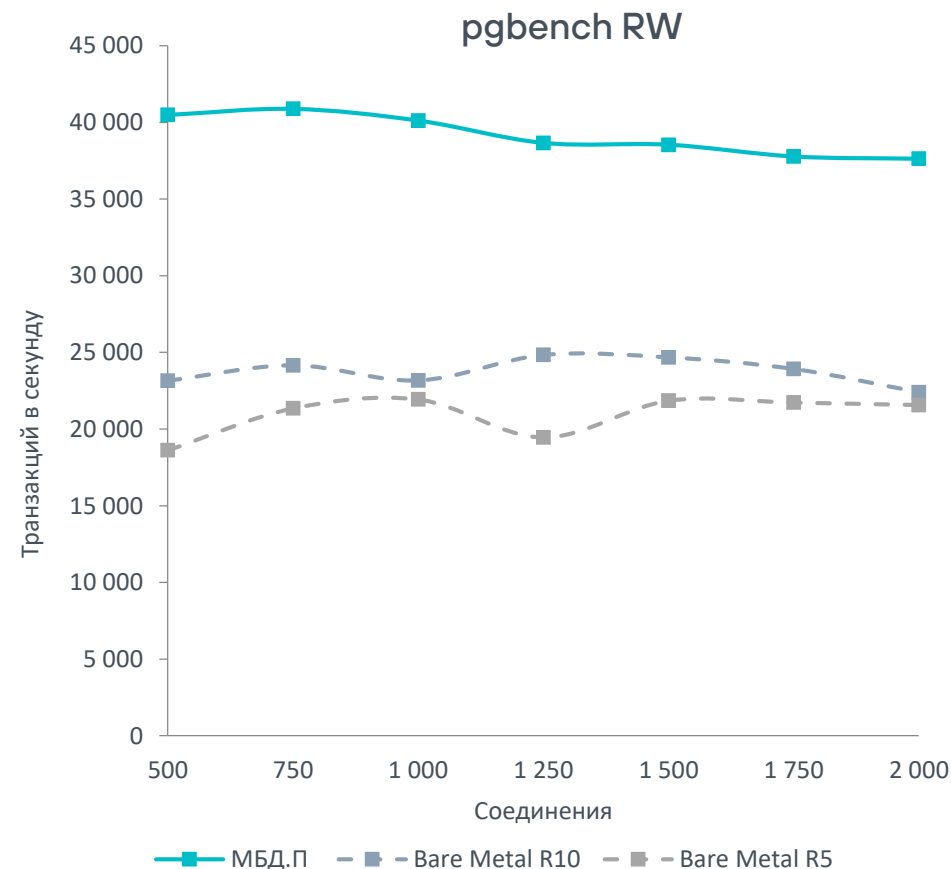
Аппаратные характеристики:

- 2 × ЦПУ Intel Xeon 24 ядра
- 512 Гбайт оперативной памяти
- Контроллер LSI 9361-16i
- Тип накопителей - SSD

x2[↑]

Машина баз данных Скала[^]р МБД.П использует специализированный программный RAID, оптимизированный для PostgreSQL, а также специальное операционное окружение

При росте количества соединений выигрыш в производительности до двух раз



* Тесты pgbench на БД размером 1,5 Тбайт (ОЗУ в 3 раза меньше размера базы)

x4 производительность

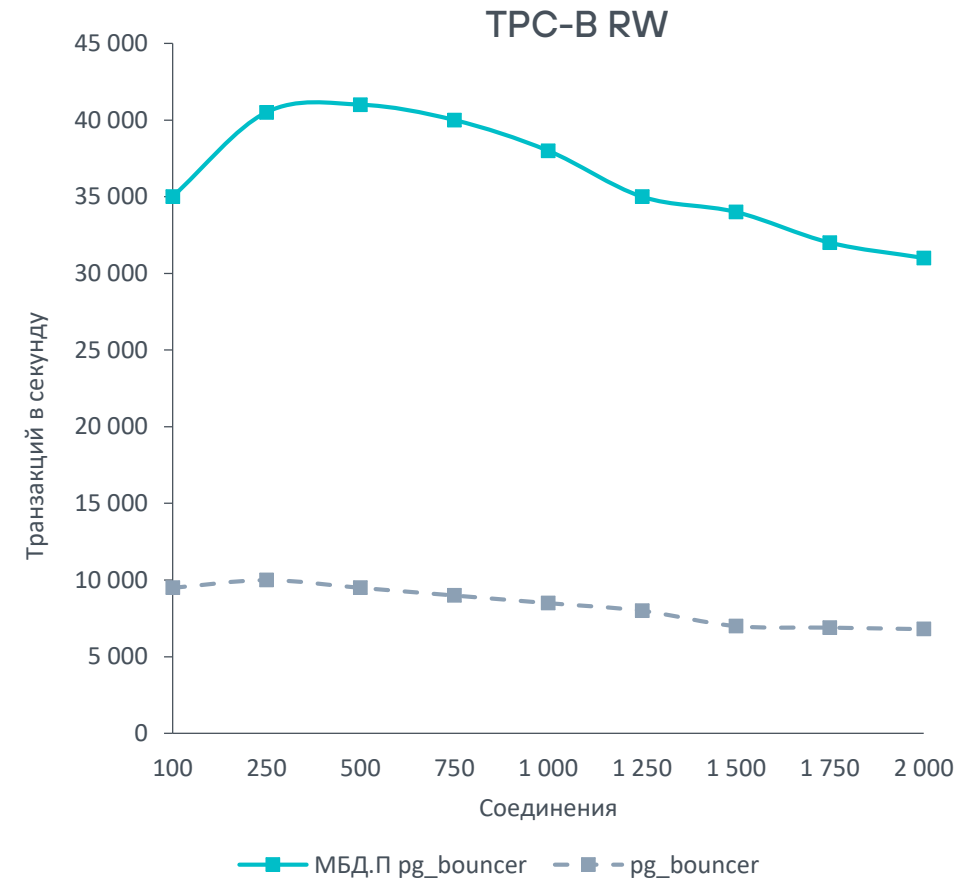
при использовании пулера соединений

- Пулер соединений позволяет осуществлять подключения не напрямую к PostgreSQL, а к сервису соединений, что уменьшает использование ресурсов и сокращает время выполнения сложных запросов
- Типовой пулер pg_bouncer ограничивает производительность PostgreSQL при OLTP-нагрузке

x4↑

Специализированное решение, используемое в Машине баз данных Скала[^] р МБД.П, позволяет увеличить производительность СУБД в 4 раза в режиме использования пулера соединений*

* Доступ к базе данных возможен через пулер и напрямую



Производительность Машин баз данных Скала^р

Для объемов горячих данных 1-3-5 Тбайт

- В большинстве случаев даже при больших БД объем «горячих» данных, с которыми идет активная работа, не превышает 1 Тбайт
- Даже если объем горячих данных выше (3 Тбайт, 5 Тбайт), снижение производительности Машины баз данных Скала^р МБД.П составляет около 10-20%

Параметры стенда:

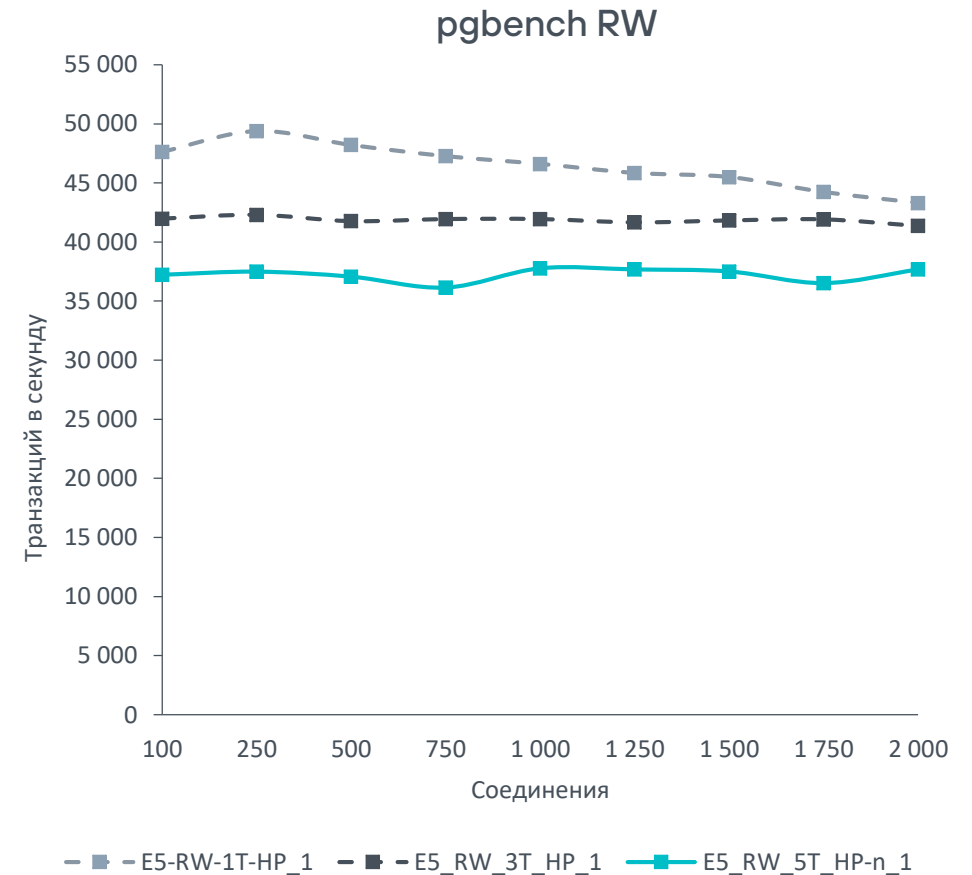
2 × ЦПУ, ОЗУ 768 Гбайт, 8x SATA SSD 1920 Гбайт data, 2x NVMe WAL, LSI 9560-8i

Настройки БД:

WAL segsize 64MB, Huge Page=on, max_wal_size 512 Гбайт, Autovacuum=on
Checkpoint = 5 min



Машина баз данных Скала^р МБД.П готова к работе под экстремальными нагрузками с большим объемом горячих данных 3-5 Тбайт без существенного изменения производительности



x4 скорость резервного копирования



Скорость выполнения резервного копирования	
Специализированный модуль СРК Машины баз данных Скала [^] р МБД.П	~ 4 ТБ/час
Классическая система СРК Huawei OceanStor, NetApp FAS	~ 1 ТБ/час

x4 [↑]	Резервное копирование
------------------------	-----------------------

Средства развертывания и восстановления баз данных	+	Встроенный модуль СРК	➔	Ускоряют время полного восстановления узла в случае отказа в 6 раз
----------------------------------------------------	---	-----------------------	---	--------------------------------------------------------------------

x6 [↑]	Скорость восстановления
------------------------	-------------------------

Экстремальная устойчивость



Машина баз данных Скала^р МБД.П показала сопоставимые результаты с системами лучших мировых производителей

По статистике возникает
не более 1 инцидента
без нарушения оказания сервиса
за 8 месяцев

за 5 лет эксплуатации
Машин баз данных Скала^р МБД.П
не было ни 1 инцидента
с потерей данных



Продукты Скала[^]р – программная платформа

Программная платформа Скала[^]р



Единая система управления ресурсами и эксплуатацией Машин

[^]геном

Система эксплуатации

Автоматизация операций:

- Обновление
- Восстановление
- Инвентаризация

[^]визион

Система мониторинга

Контроль параметров,
сбор и хранение:

- Объектов управления
- Метрик
- Визуализаций параметров

[^]спектр

Кластер-менеджер

Конфигурация и управление:

- Кластеры PostgreSQL
- Настройка отказоустойчивости
- Резервное копирование и восстановление

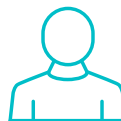
Скала^р Геном – система эксплуатации



Единая система эксплуатации для всех Машин:

- Оптимизация эксплуатации и техподдержки
- Типизация конфигураций
- Исключение случайных ошибок
- Увеличение скорости развертывания
- Восстановление Машины в случае сбоя
- Хранение информации о конфигурации и шагах подготовки – «Цифровой двойник»

Геном



- Параметры конфигурации Машины
- Ход развертывания
- Результаты развертывания
- (Паспорт ПАК)

Управляющие воздействия

Машины Скала^р

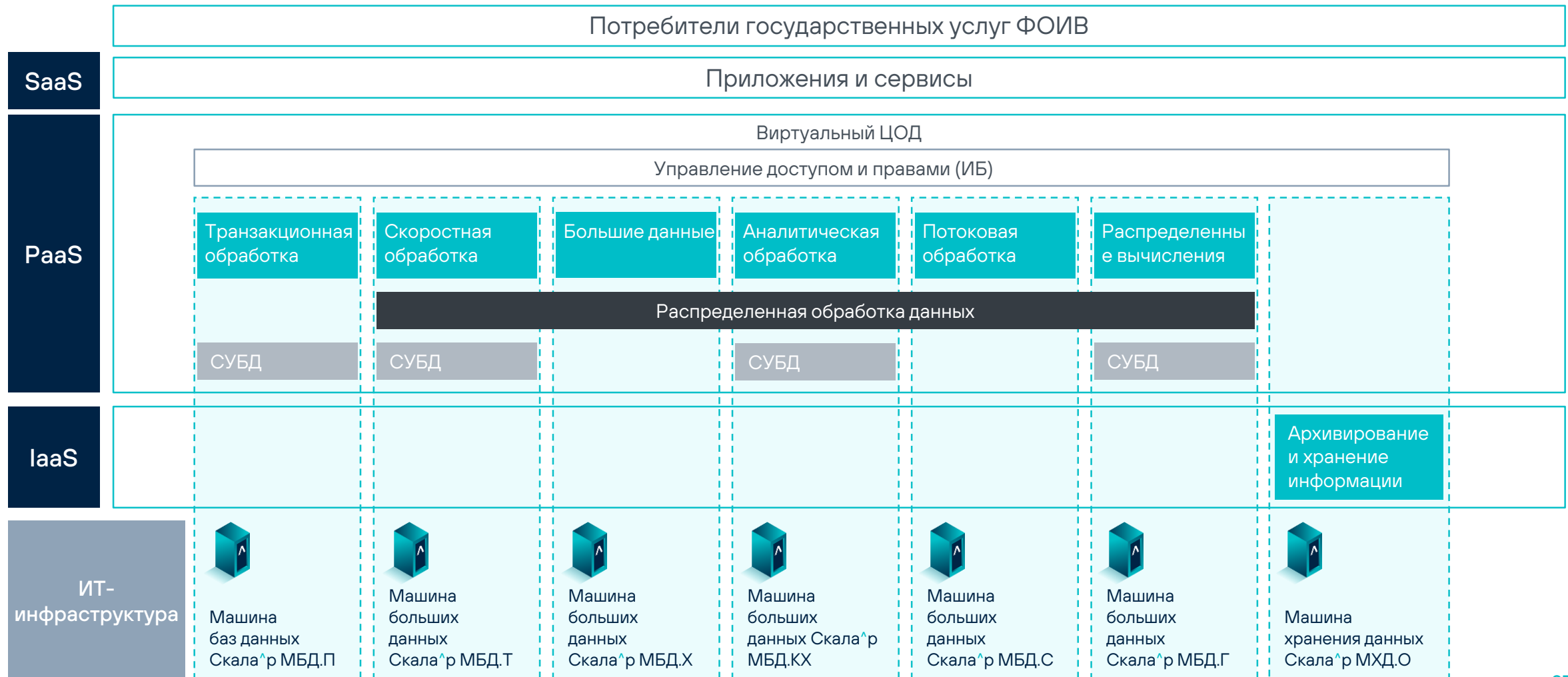


x6↑

Ускорение восстановления узлов Машины в 6 раз за счет применения автоматизации встроенной в Скала^р Геном

Пример реализованных проектов Скала^р

Создание единой облачной платформы сервисов ГИС в рамках Гособлака. Сервисная модель оказания услуг



Техническая поддержка и услуги



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



техническая
поддержка из
«одного окна»

24x7

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



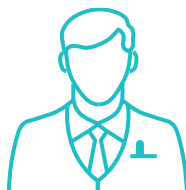
Круглосуточно

- 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)





Модульная платформа
для высоконагруженной
инфраструктуры по модели
распределенного облака