



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

8 лет

серийного
выпуска

400+

комплексов
в промышленной
эксплуатации

6500+

вычислительных
узлов

Линейка продуктов Скала^р



для высоконагруженных корпоративных и государственных систем
кластеры высокой доступности, катастрофоустойчивые и метрокластеры



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



Big Data & Data Science

Машины больших данных Скала^р МБД.8

на основе решений **Arenadata, Picodata** и **Datamart** для создания инфраструктуры хранения, преобразования и одновременной аналитической и статистической обработки больших объемов информации



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

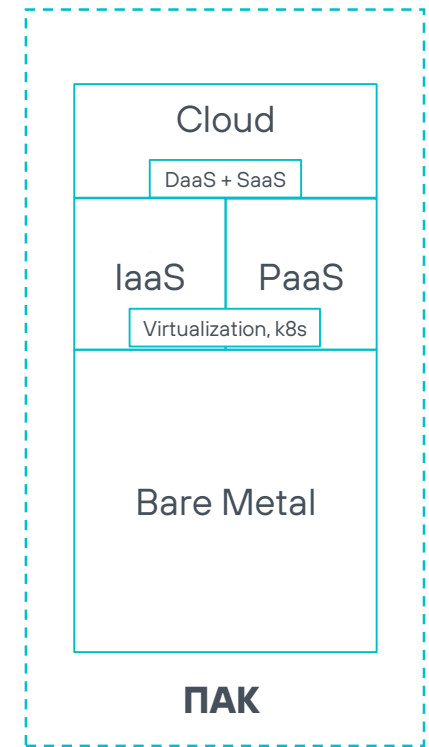
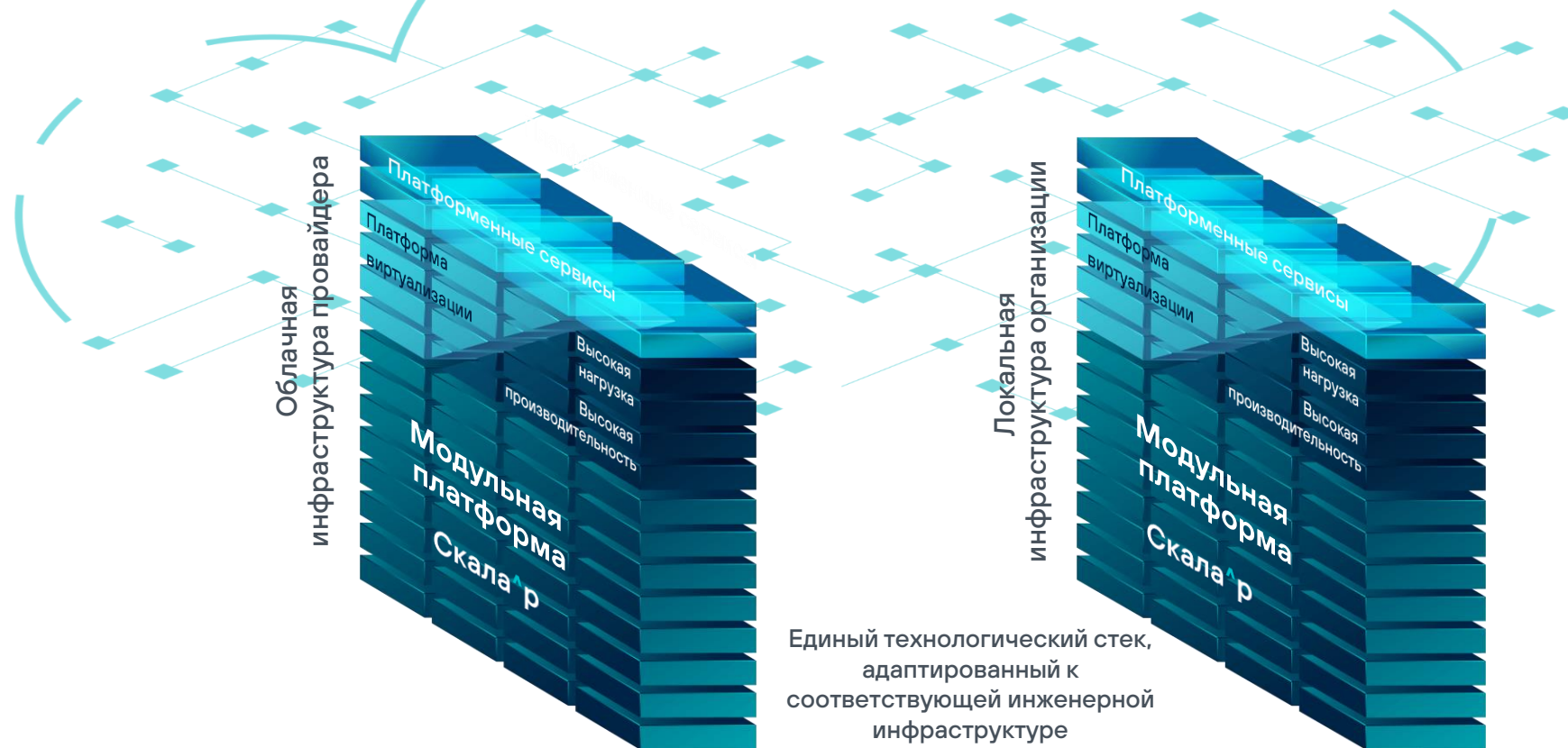
Использование опыта технологических лидеров (гиперскейлеров)

Использование самых зрелых и перспективных технологий в кооперации с технологическими лидерами российского рынка

в каждом из сегментов

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

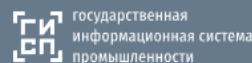
ПАК Скала^р в Едином реестре российской радиоэлектронной продукции



Машины

Модули

Компоненты



☰ Все сервисы ГИСП

Реестр промышленной продукции, произведенной на территории Российской Федерации

Машины

Программное обеспечение



Российский

Евразийский

- Продукция Скала^р включена в Единый реестр российской радиоэлектронной продукции
- Технические средства Машин и Модулей Скала^р включены в Единый реестр российской радиоэлектронной продукции
- Программные компоненты Машин Скала^р включены в Единый реестр российских программ для ЭВМ и баз данных


Единая система управления ресурсами и эксплуатацией Машин



 Цифровой двойник

 Развертывание

 Мониторинг

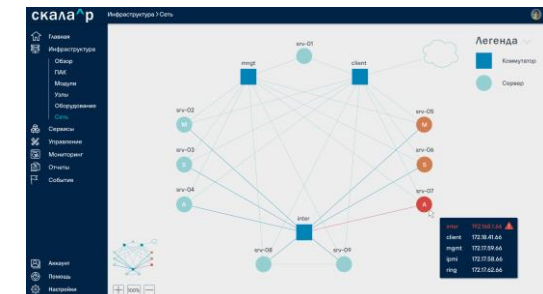
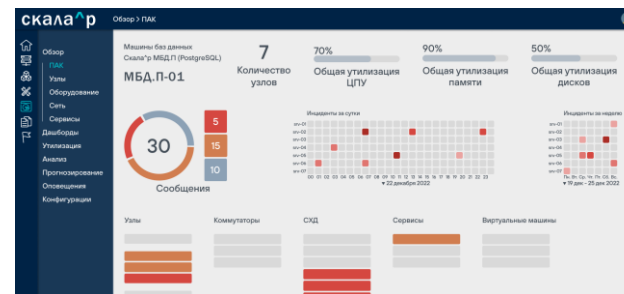
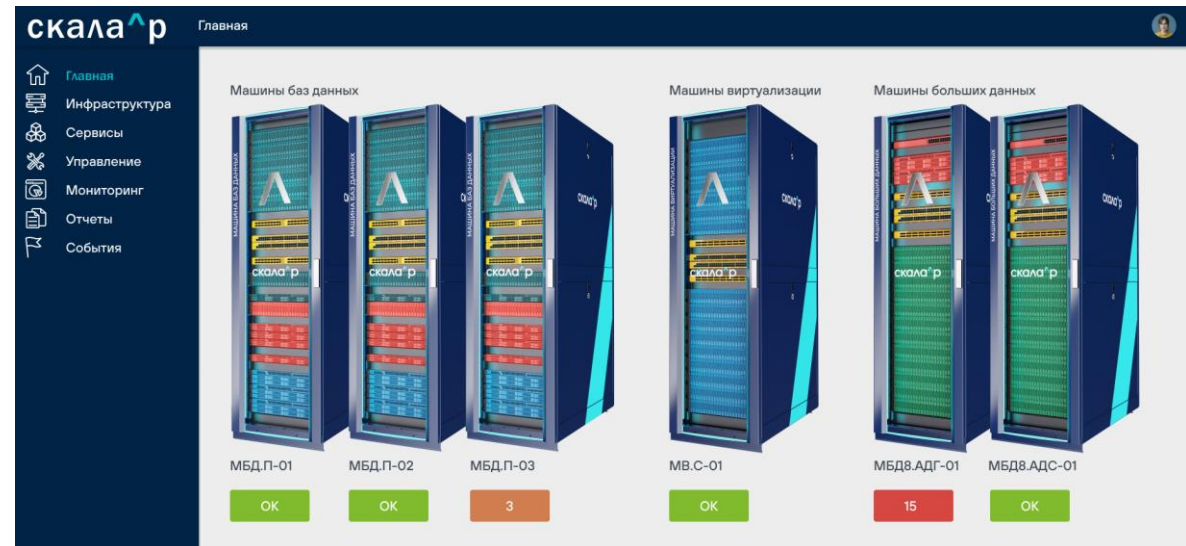
 Диагностика

 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:

- Высоконагруженные системы
- Катастрофоустойчивые инфраструктуры
- СУБД для частных и гибридных облаков

Преимущества:

- Высокая надежность и отказоустойчивость
- Высокопроизводительная сеть внутреннего взаимодействия 100 Гбит/с низкими задержками
- Высокопроизводительная обработка и хранение большого объема данных
- Объем баз до 150 ТБ
- Встроенная система резервного копирования до 4 ТБ/час

Замещаемые технологии:

- Oracle Exadata

Рекомендовано
от 12 000
транзакций в секунду

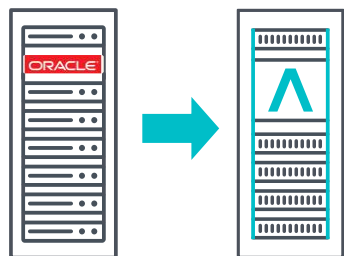
Пиковая
производительность
65 000+
транзакций/сек



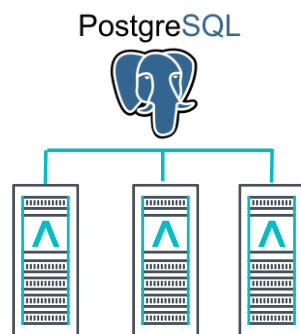
Машина баз данных Скала^р МБД.П



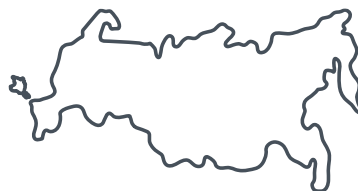
Сценарии применения



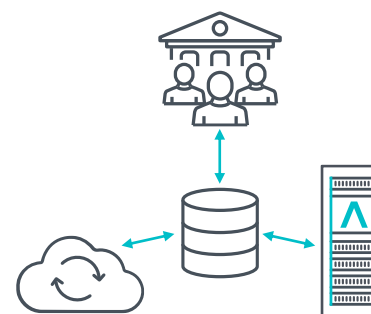
Замена Oracle
для транзакционных
нагрузок



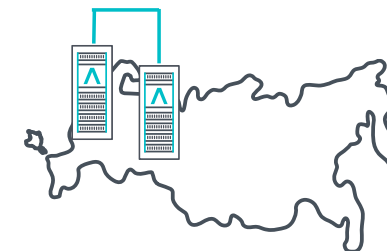
Высоконагруженные
системы
для скоростной
обработки транзакций
на основе Postgres



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

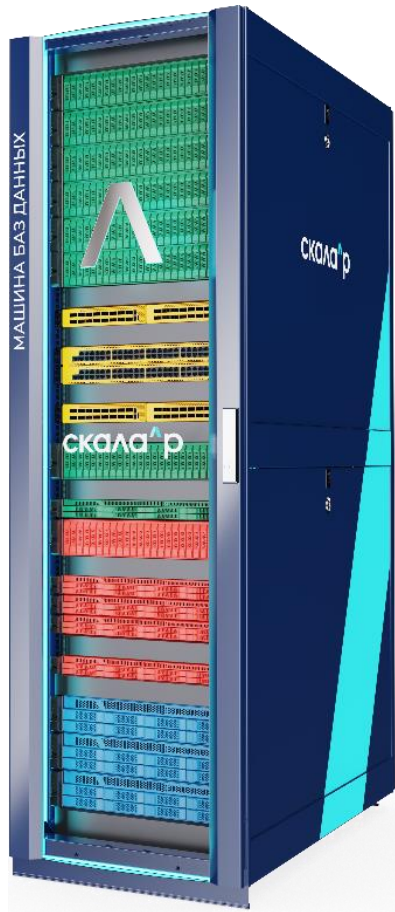


Создание
сервисов СУБД
для частных и
гибридных облаков
предприятий



Катастрофо-
устойчивые
инфраструктуры

Состав Машины Скала^р МБД.П



Блок вычисления и хранения

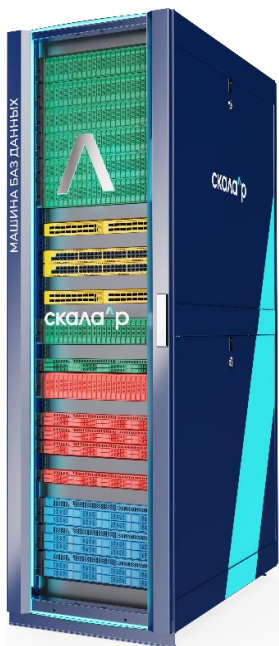
Блок коммутации и агрегации

Блок мониторинга и регистрации

Блок резервного копирования

- Кластер из трёх узлов (мастер, синхронная реплика, асинхронная реплика)
- Повышение производительности за счет записи на мастер, а чтения с синхронной реплики.
- Сети внешнего доступа
- Сеть внутреннего взаимодействия до 100 ГБ/с
- Выделенная сеть для управления и мониторинга
- Управление эксплуатацией и автоматизация критически важных процедур, снижение влияния человеческого фактора
- Мониторинг состояния всех компонент Машины
- 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 ТБ оперативной памяти
- От 4 до 150 ТБ 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 ГБ/с)
- Узел сетевого взаимодействия (100 ГБ/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
- Узлы хранения 700 ТБ резервная копия БД и 150 ТБ архивных файлов предзаписи

- 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:

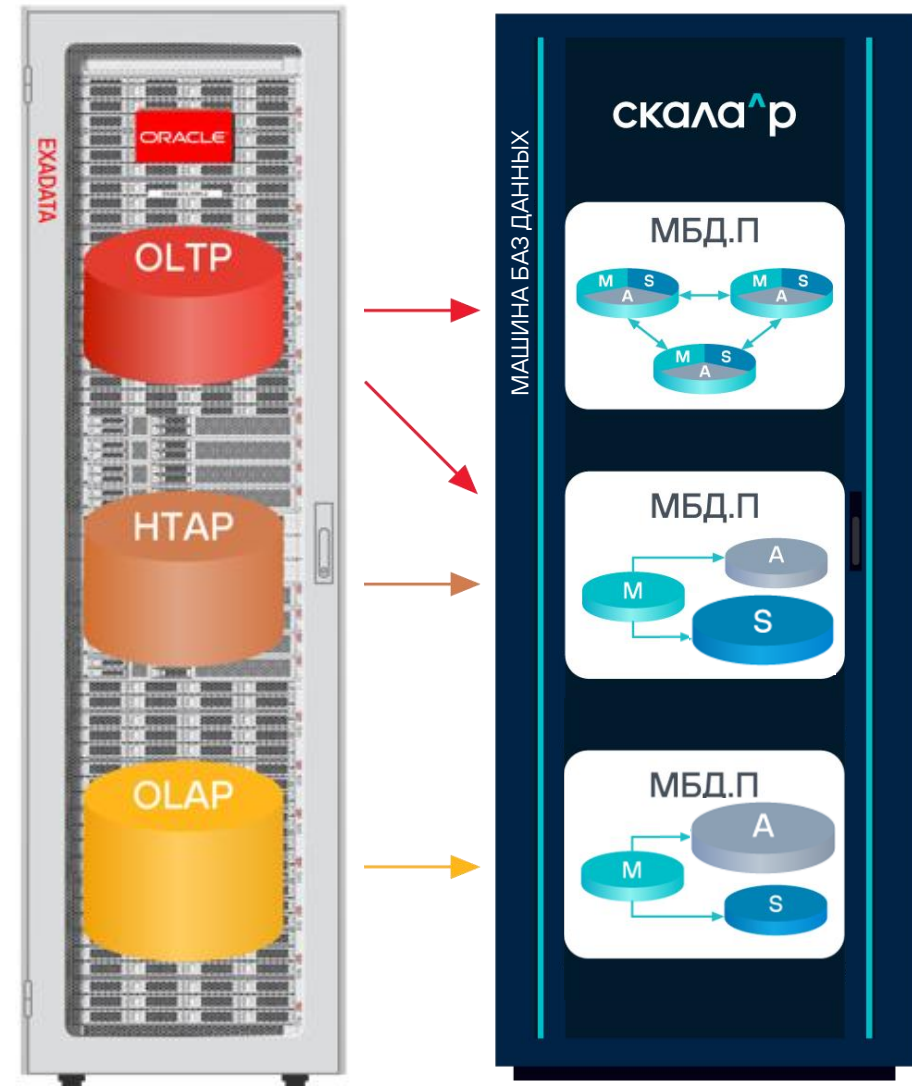
- Транзакционная (OLTP)
- Гибридная (HTAP)
- Аналитическая (OLAP)

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

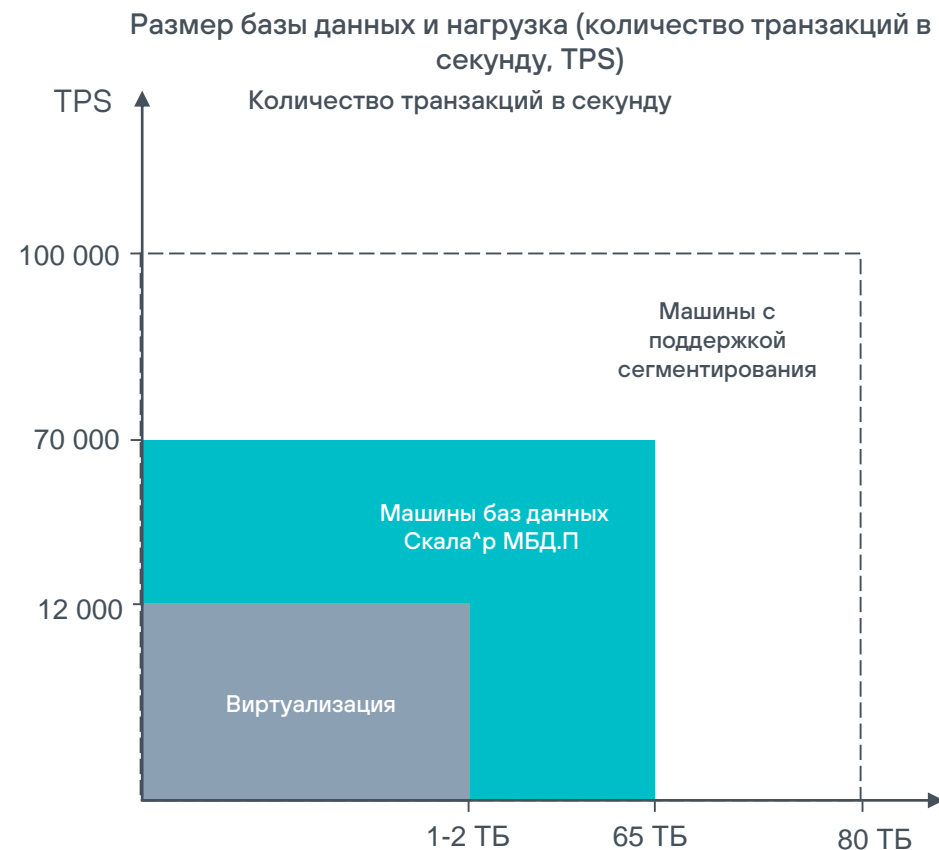


Машина баз данных Скала^р МБД.П



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



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

Машина баз данных Скала[^]р МБД.П



Технические показатели

Типовые размеры
томов данных

Тип 1 до **80 ТБ**

Тип 2 до **40 ТБ**

Производительность*

до **50 000 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 ГБ/с

на каждые 10 ТБ
объема хранения

* На объеме горячих данных 1ТБ (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

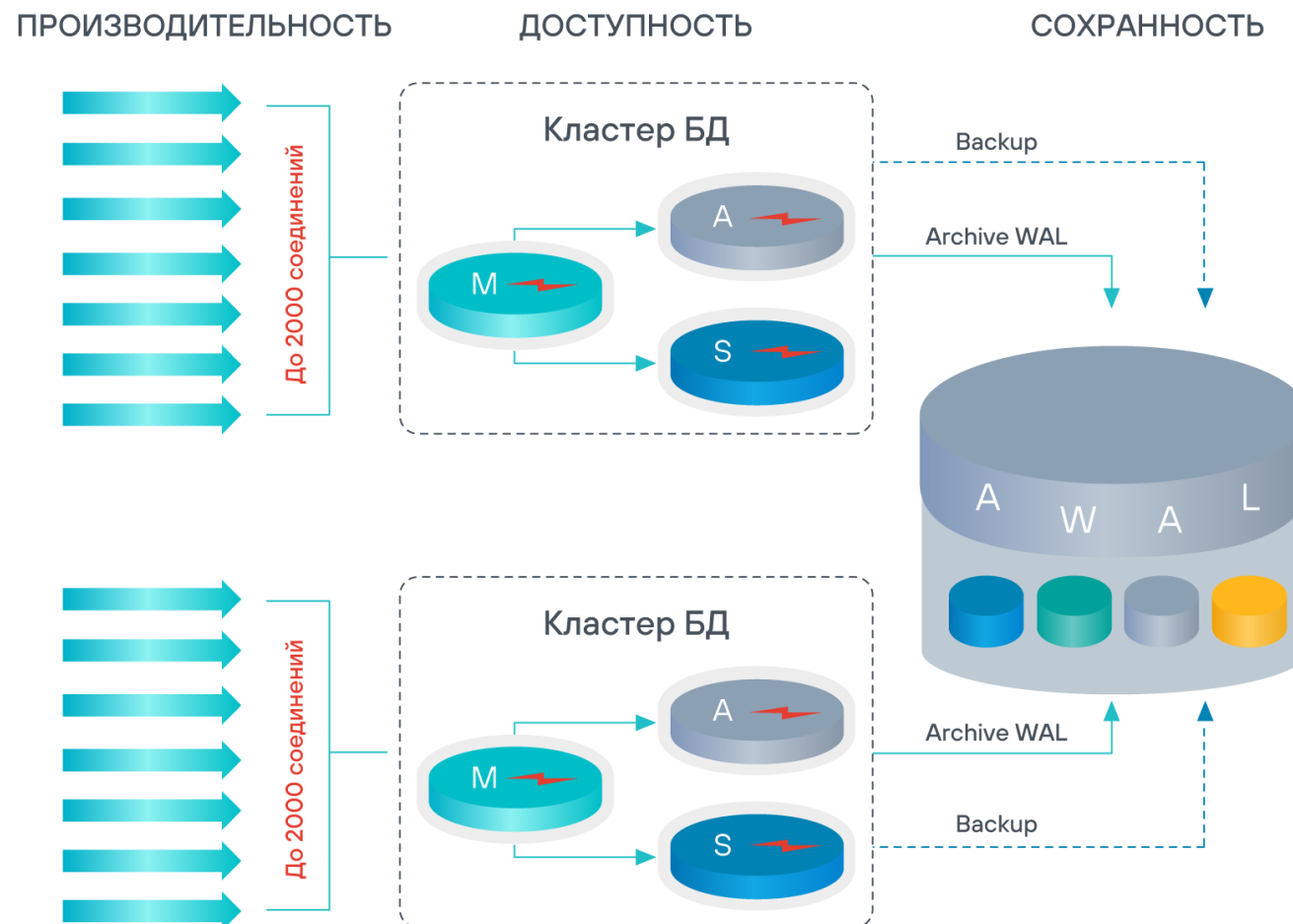
- Оптимизированная локальная система хранения
- Высокоскоростная сеть внутреннего взаимодействия
- Чтение с синхронной реплики

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

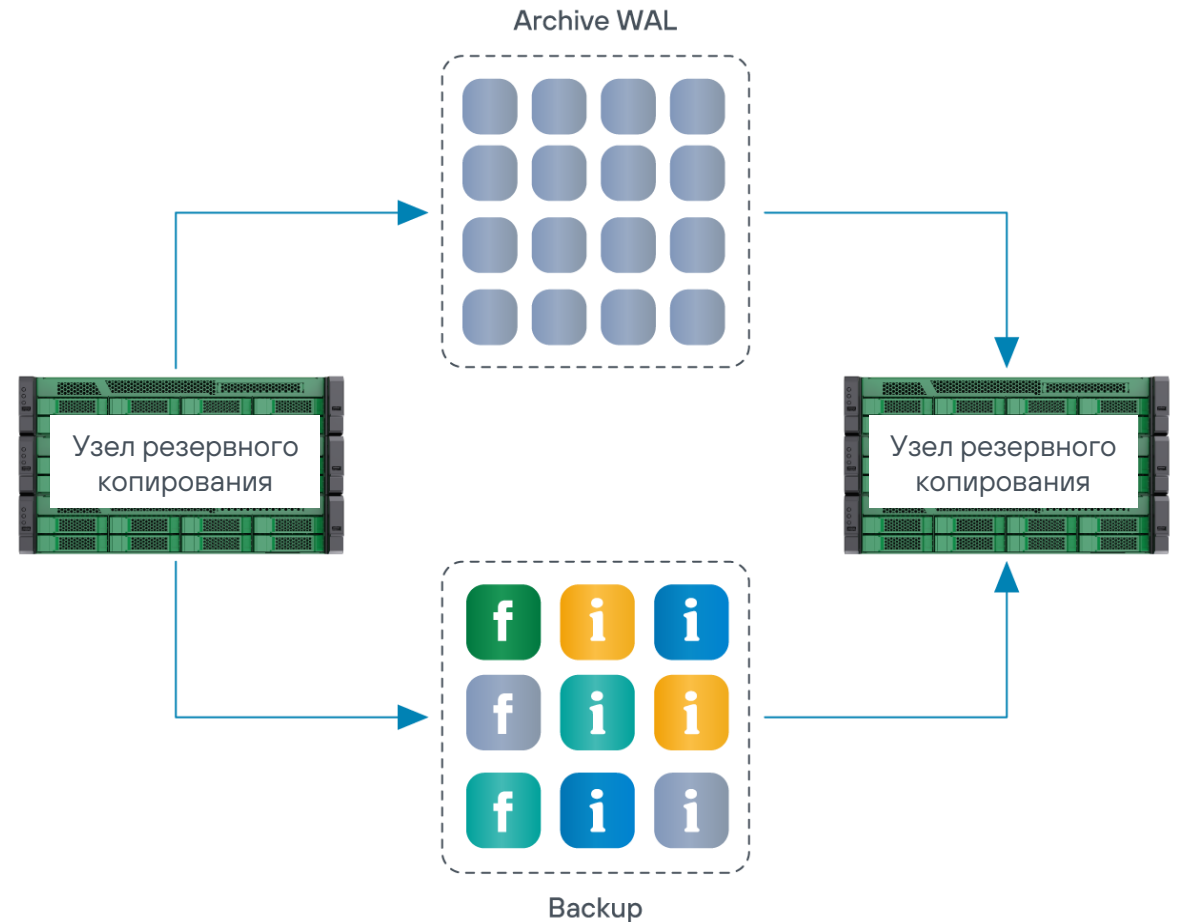
- Сохранность БД на внешних носителях
- Полное и инкрементальное копирование
- Кратное увеличение скорости копирования и восстановления за счет специализированного решения

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 rgbench для оценки конкурентных преимуществ

TPC-B — это эталонный тест для оценки производительности платформ и справедливого сравнения их между собой. Тест был разработан независимым Советом по оценке производительности обработки транзакций (TPC — Transaction Processing Performance Council, Сан-Франциско), и позволяет заказчикам выбрать платформу с оптимальным соотношением цена/производительность для решения своих задач.

Тест TPC-B

Тест TPC-B — простой тест для демонстрации совместной производительности СУБД и аппаратной платформы, при котором моделируется сеть контрагентов и транзакций между ними (простая, повторяемая единица работы в приложении):

- Сценарий тестирования включает пять команд SELECT, UPDATE и INSERT в одной транзакции
- Производительность определяется как число зафиксированных транзакций в секунду (TPS-B)
- Соотношение цена/производительность вычисляется как цена, отнесенная к максимальному числу зафиксированных транзакций в секунду (TPS-A) и выражается в тысячах долларов на TPS-A

Утилита rgbench

Rgbench — это простая программа для запуска тестов производительности PostgreSQL. Она многократно выполняет одну последовательность команд, возможно выполнение в параллельных сеансах базы данных, а затем вычисляет среднюю скорость транзакций (число транзакций в секунду). По умолчанию запускается тест типа TPC-B. Однако можно тестировать и другие сценарии, написав свои скрипты транзакций.

x2 производительность

по сравнению с аналогичным оборудованием



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

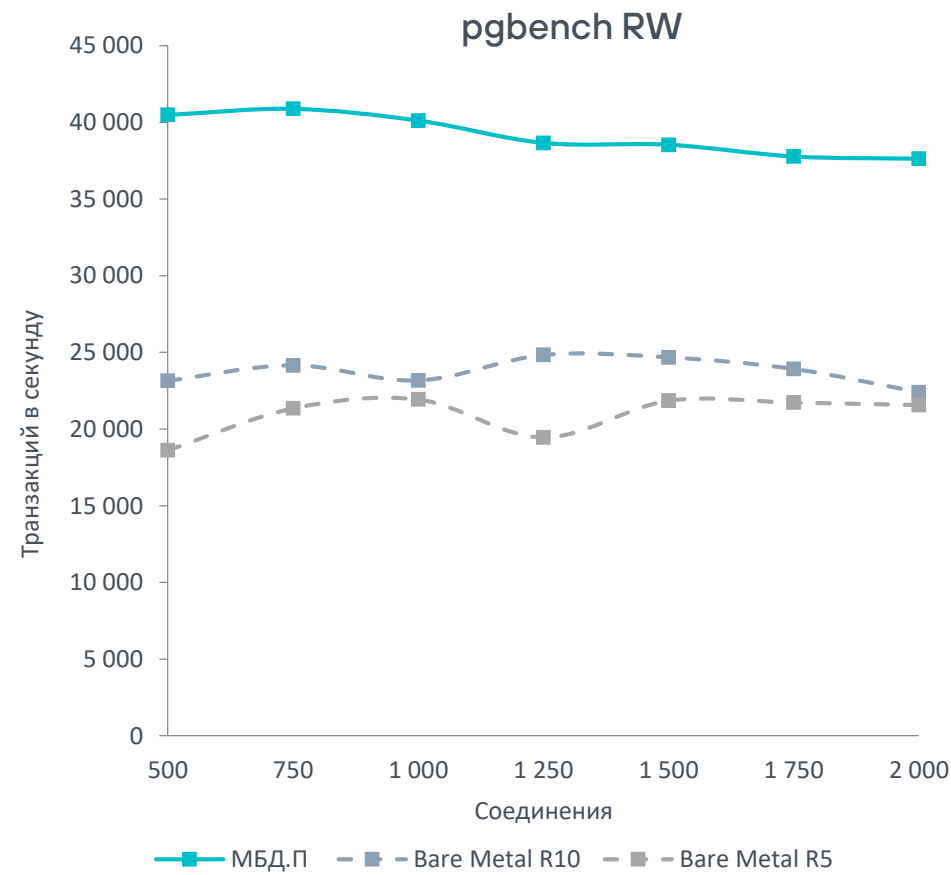
Аппаратные характеристики:

- 2x процессор Intel Xeon 24 ядра
- 512 ГБ оперативной памяти
- Контроллер LSI 9361-16i
- Диски 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 ТБ (ОЗУ в 3 раза меньше размера базы)

x4 скорость

обработки запросов, чем в виртуальных средах



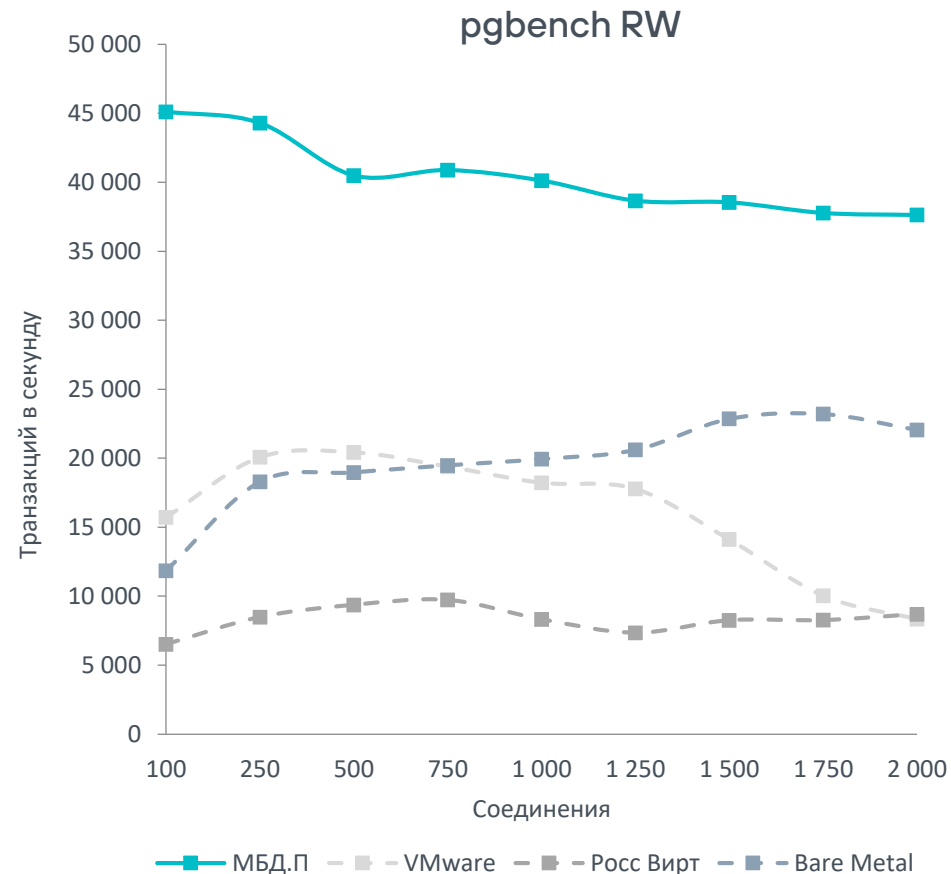
Сравнительные тесты Машин баз данных Скала^р МБД.П, виртуальных сред и Bare Metal

- Машина баз данных Скала^р МБД.П (диски SSD Intel)
- VMware (СХД Huawei OceanStor 5500)
- Российская виртуализация (СХД Ядро Татлин, NVMe)
- Bare Metal (диски SSD Intel)

x4↑

Машина баз данных Скала^р МБД.П использует сочетание высоко-производительного оборудования, возможностей аппаратной архитектуры и функционала PostgreSQL

При высокой нагрузке и большом числе соединений показывает в 4 раза большую производительность по сравнению с решениями в виртуальных средах и в 2 раза, чем Bare Metal с более быстрыми дисками



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

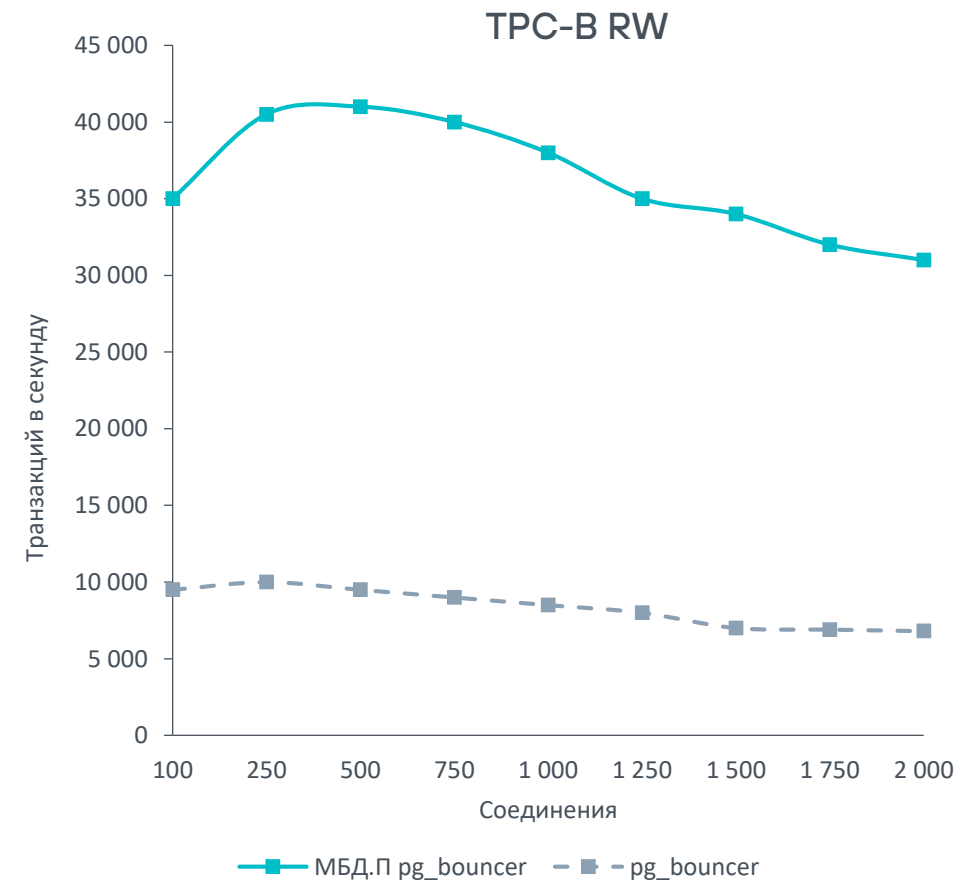
x4 производительность

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 ТБ


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

Параметры стенда:

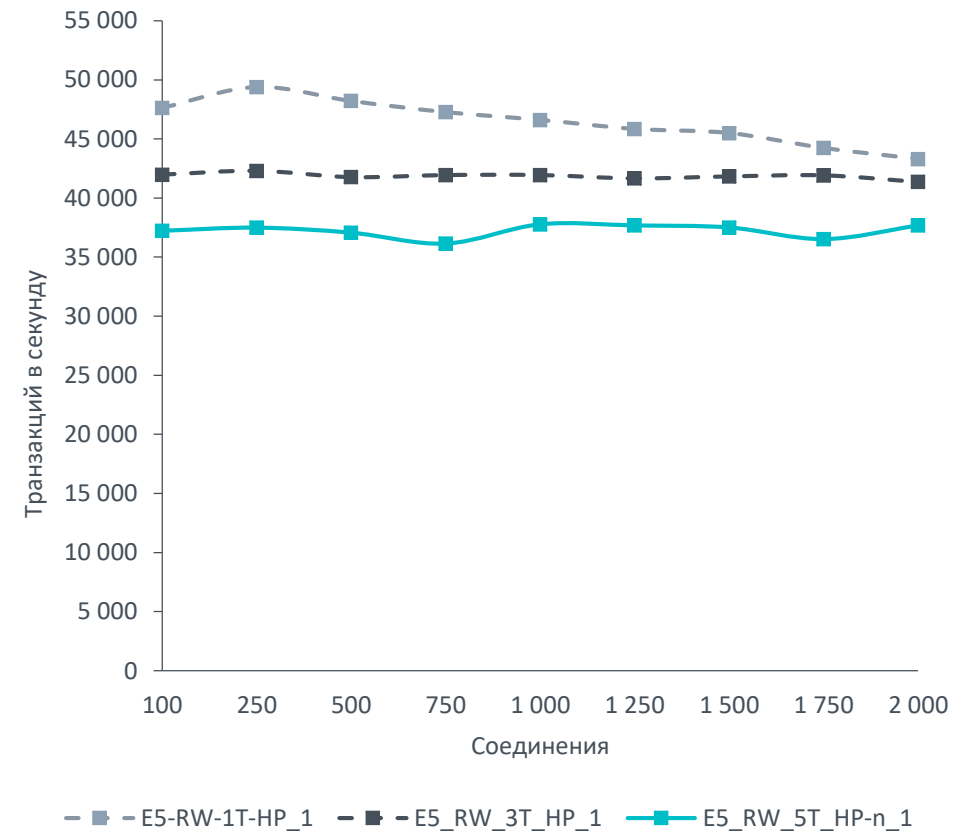
2x процессор, ОЗУ 768 ГБ, 8x SATA SSD 1920 ГБ data, 2x NVMe WAL, LSI 9560-8i

Настройки БД:

WAL segsize 64MB, Huge Page=on, max_wal_size 512 ГБ, 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 ТБ без существенного изменения производительности

pgbench RW

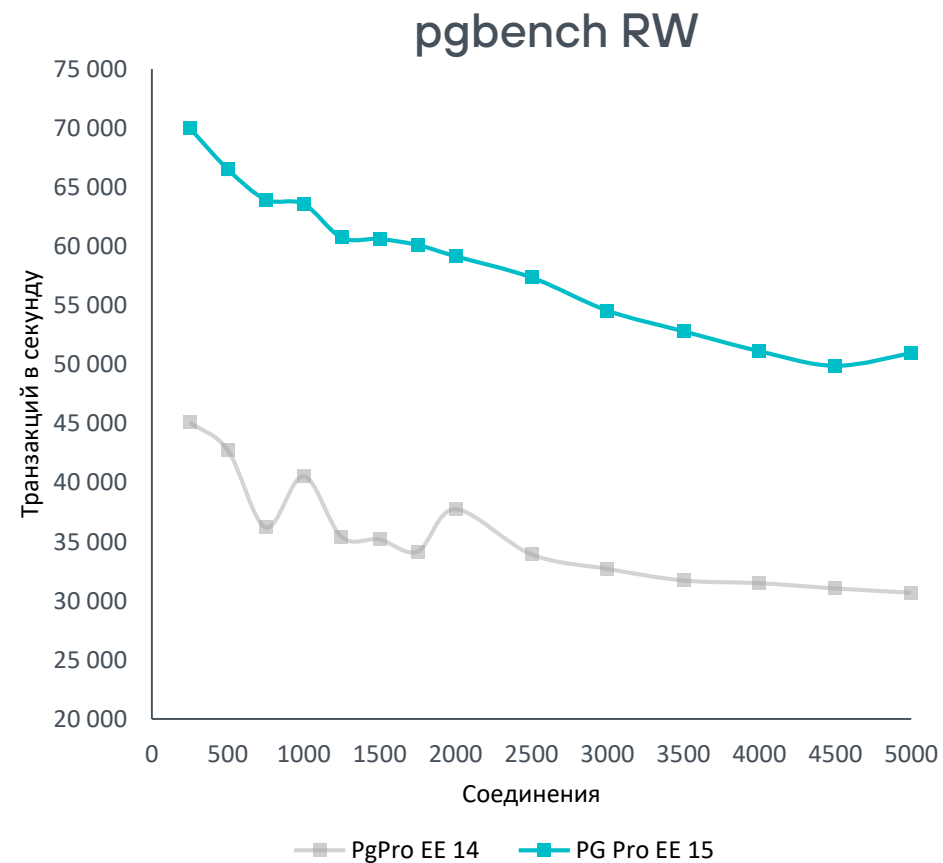


Результаты тестирования Postgres Pro 14 и 15



Сравнение производительности Postgres Pro на одинаковых аппаратных конфигурациях дисковой подсистемы

В тестовой лаборатории компании Скала^р были проведены сравнительные тесты оптимизированного решения на базе двух версий Postgres 14 и 15 версии и результаты показали достижимые показатели производительности до 70000 транзакций в секунду.



x1.6↑

Машина баз данных МБД.П на Postgres Pro версии 15 позволяет добиться повышения производительности до 60% в отличие от Машины на Postgres Pro версии 14

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 одного 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

[^]спектр

Кластер-менеджер

Конфигурация и управление:

- Кластеры PostgreSQL
- Настройка отказоустойчивости
- Резервное копирование и восстановление

[^]ВИЗИОН

Система мониторинга

Контроль параметров, сбор и хранение:

- Объектов управления
- Метрик
- Визуализаций параметров

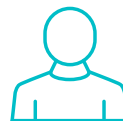
Скала^р Геном – система эксплуатации



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»

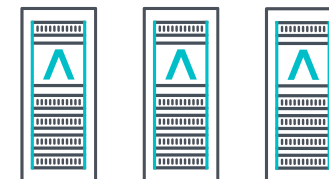
Геном



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

Управляющие воздействия

Машины Скала^р



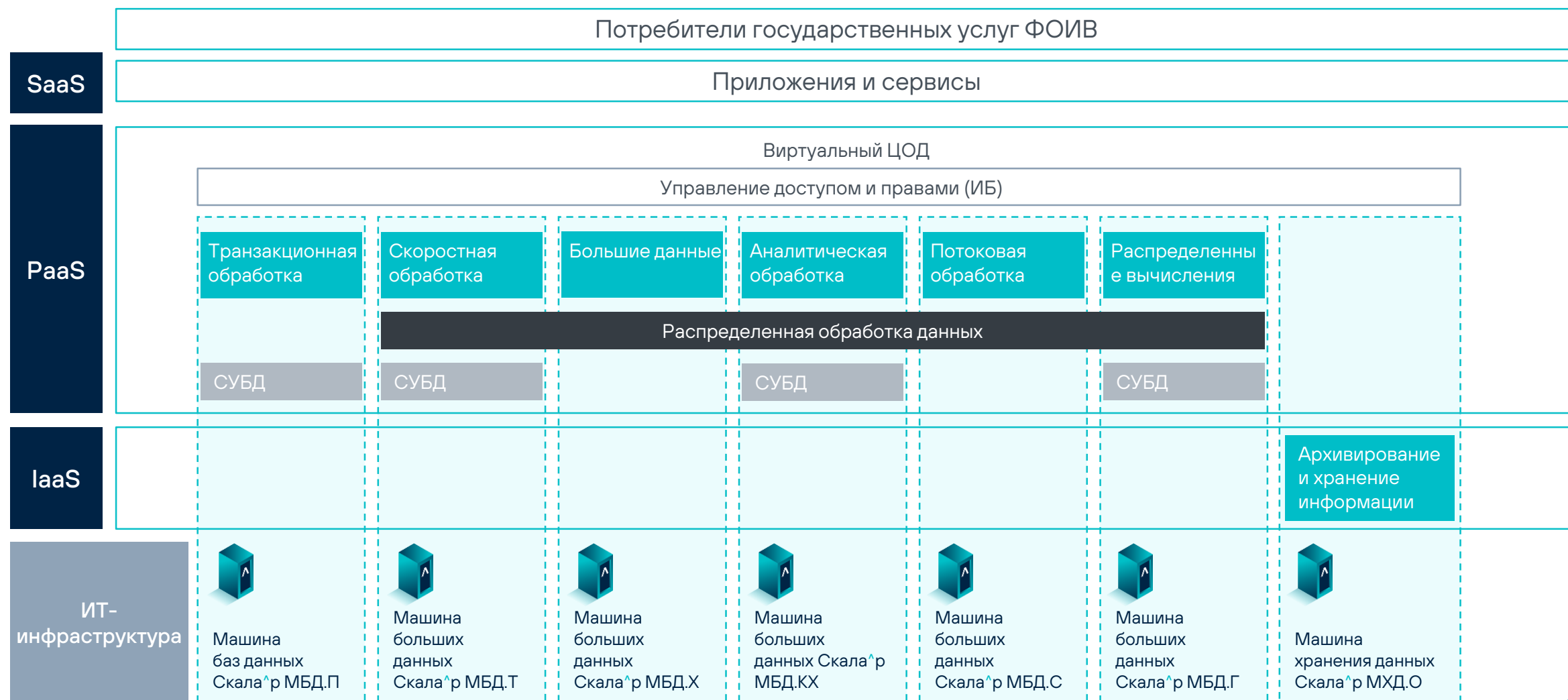
x6↑

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



Примеры реализованных проектов Скала[^]р

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



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