

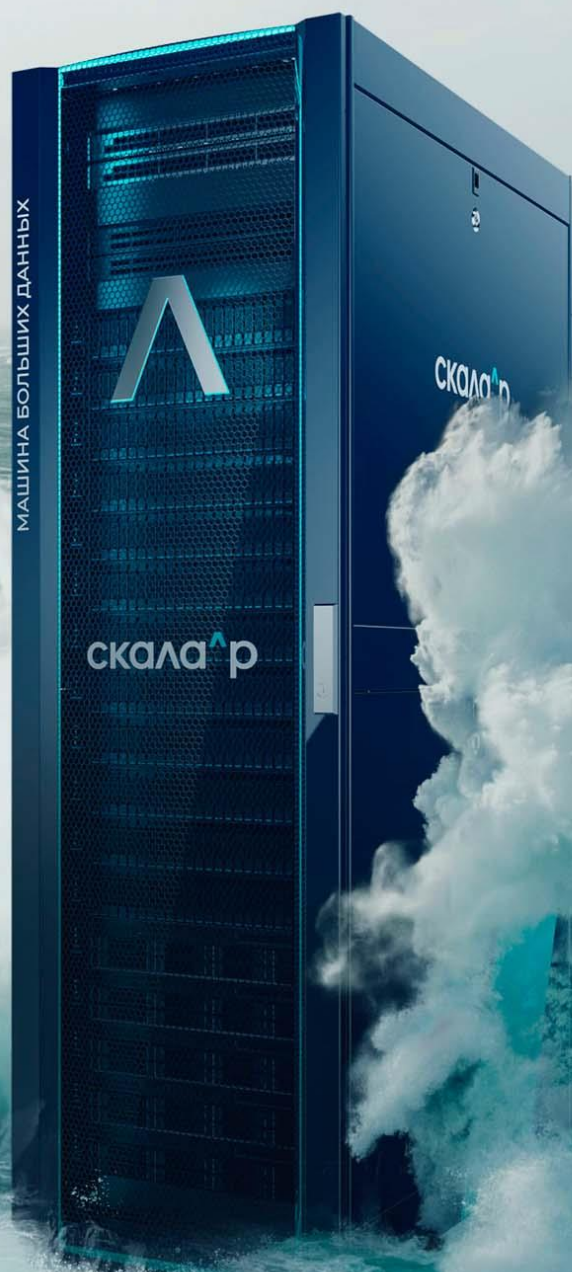


Машина больших данных Скала^р МБД.Т

Программно-аппаратный комплекс для развертывания высокопроизводительных программных систем на основе резидентной СУБД Picodata

Технический обзор

версия 3.3 от 30.12.2025



ОГЛАВЛЕНИЕ

Уведомление.....	4
Перечень терминов и сокращений.....	5
1. Предисловие	8
1.1 Описание документа	8
1.2 Аудитория.....	8
1.3 Обратная связь	8
2. Введение	9
3. Отличительные черты	10
4. Подтвержденная безопасность.....	12
5. Принципы создания Машины МБД.Т	16
5.1 Основные принципы.....	16
5.2 Основы резидентных СУБД.....	16
6. Состав Машины	18
6.1 Подсистемы.....	18
6.1.1 Подсистема обеспечения базовых сервисов и Сетевая подсистема	20
6.1.2 Подсистема резидентной обработки больших данных.....	20
6.1.3 Подсистема резервного копирования	20
6.1.4 Подсистема управления данными	21
6.1.5 Подсистема преобразования данных	21
6.2 Модули.....	21
6.2.1 Базовый модуль	21
6.2.2 Модуль резидентной обработки.....	22
6.2.3 Модуль резервного копирования	23
6.2.4 Модуль управления данными.....	23
6.2.5 Модуль преобразования данных.....	24
7. Специфичные черты	25
8. Гарантированное качество и полная готовность к промышленной эксплуатации	26
9. Реакция Машины на возможные отказы.....	28
10. Вариативность решения	29
11. Требования к размещению Машины.....	30

12. Техническая поддержка	31
13. Лицензирование ПО в составе модулей	33
13.1 Политика обновления ПО	33
О Компании	34

УВЕДОМЛЕНИЕ

Документ носит исключительно информационный характер и является актуальным на дату размещения.

Технические характеристики, приведенные в документе — справочные и не могут служить основанием для претензий.

Технические характеристики могут отличаться от приведенных вследствие модификации изделий.

Технические характеристики и комплектация изделий могут быть изменены производителем без уведомления.

Документ не является публичной офертой и не содержит каких-либо обязательств ООО «СКАЛА-Р».

ПЕРЕЧЕНЬ ТЕРМИНОВ И СОКРАЩЕНИЙ

Термин, сокращение	Определение
ACID	(англ. Atomicity, Consistency, Isolation, Durability) — набор требований к транзакционной системе, обеспечивающий наиболее надёжную и предсказуемую её работу — атомарность, согласованность, изоляцию, устойчивость
ACL	(англ. Access Control List) — список управления доступом, который определяет, кто или что может получать доступ к объекту (программе, процессу или файлу), и какие именно операции разрешено или запрещено выполнять субъекту (пользователю, группе пользователей)
Data Warehouse	«Хранилище данных» — это предметно-ориентированная, интегрированная, неизменяемая и поддерживающая хронологию совокупность данных, предназначенная для поддержки принятия управленческих решений, выполнения сложных аналитических запросов, агрегирующих большие объёмы исторических данных
ETL	(англ. Extract, Transform, Load) — процесс транспортировки данных, при котором информацию из разных мест преобразуют и кладут в новое место
MLAG	(англ. Multi-Switch Link Aggregation) — технология агрегации каналов, позволяющая одному или нескольким линкам с двух разных сетевых узлов быть объединёнными вместе таким образом, что для конечного устройства это выглядит как одиночное соединение
NFS	(англ. Network File System) — протокол сетевого доступа к файловым системам
OLTP	(англ. Online Transaction Processing) технология для оперативной обработки транзакций
RAID	(англ. Redundant Array of Independent Disks) — избыточный массив независимых накопителей, технология виртуализации данных для объединения нескольких физических дисковых устройств в логический модуль для повышения отказоустойчивости и производительности
RPC	(англ. Remote Procedure Call) — вызов удалённых процедур, класс технологий, позволяющих программам вызывать функции или процедуры в другом адресном пространстве (на удалённых узлах либо в независимой сторонней системе на том же узле)
RPO	(англ. Recovery Point Objective) — за какой период будут утеряны данные в случае сбоя

Термин, сокращение	Определение
RTO	(англ. Recovery Time Objective) — время на восстановление системы
snapshot	Полный снимок текущего состояния виртуальной машины, который позволяет в любой момент вернуть систему к сохранённой конфигурации
WAL	(англ. Write Ahead Log) — журнал предварительной записи транзакций
ГИС	Государственные информационные системы — системы, которые создаются для реализации полномочий государственных органов и обеспечения обмена информацией между ними, а также в иных установленных федеральными законами целях
ЗОКИИ	Значимый объект критической информационной инфраструктуры
ИСПДн	Информационные системы персональных данных. Совокупность информации, содержащейся в базах данных, и обеспечивающих её обработку с использованием информационных технологий и технических средств
Кластер	Отказоустойчивая архитектура функционала Машины
Машина	Набор аппаратного и программного обеспечения в виде Модулей Скала [^] р, соединенных вместе для обеспечения определённого метода обработки данных или предоставления ИТ-сервиса с заданными характеристиками. Зарегистрирована в ЕРРРП
Модуль	Функционально завершённый комплект сконфигурированного для выполнения заданных функций аппаратных и/или программных компонентов, аппаратных узлов и программного обеспечения (ПО), оформленный как самостоятельная единица продаж со своим кодом (part number) и стоимостью. Является единым и неделимым элементом спецификации. Зарегистрирован в ЕРРРП
ОС	Операционная система
ПАК	Программно-аппаратный комплекс
ПО	Программное обеспечение
Подсистема	Логическое объединение компонентов по функциональному признаку, с целью пояснения состава и принципов действия ПАК
СУБД	Система управления базами данных

Термин, сокращение	Определение
Узел	Вычислительный узел (сервер) или сетевой узел (коммутатор) в составе Модуля, в зависимости от контекста

1. ПРЕДИСЛОВИЕ

1.1 Описание документа

Этот технический обзор дает концептуальный и архитектурный обзоры **Машины больших данных Скала^р МБД.Т.**

Документ описывает то, как оптимизированные программно-аппаратные комплексы отвечают современным вызовам, и фокусируется на **Машине больших данных Скала^р МБД.Т.**

1.2 Аудитория

Эта брошюра предназначена для сотрудников компании **Скала^р**, партнёров и заказчиков, перед которыми ставятся задачи разработки, закупки, управления или эксплуатации **Машины больших данных Скала^р МБД.Т.**

1.3 Обратная связь

Скала^р и авторы этого документа будут рады обратной связи по нему.

Свяжитесь с командой **Скала^р** по электронной почте MBD8@skala-r.ru.

2. ВВЕДЕНИЕ

Машина больших данных Скала[^]р МБД.Т — это программно-аппаратный комплекс для обработки и хранения данных и построения высокопроизводительных решений на базе резидентной (In-Memory) СУБД Picodata. ПАК обеспечивает высочайший отклик при OLTP-нагрузках и объединяет аппаратную платформу и оптимизированное ПО, что позволяет работать с большими объёмами данных быстрее, надёжнее и с меньшими затратами.

Преимущества

- Проработанная интеграция аппаратного и программного обеспечения снижает стоимость обслуживания и сокращает расходы на сопровождение
- Возможность масштабирования с увеличением ёмкости хранения данных от 1 до 30 Тбайт без перебоев в работе
- Резервирование критичных компонентов, устойчивые сетевые протоколы и защита от простоев дают высокий уровень надёжности
- Использование современных NVMe-накопителей и стогигабитных сетей обеспечивают мгновенный отклик даже при пиковых нагрузках
- ПАК включён в Единый реестр российской радиоэлектронной продукции и использует отечественное программное обеспечение из реестра Минцифры РФ, что подтверждает его соответствие требованиям импортозамещения

Особенности архитектуры

- Программная основа — СУБД Picodata с поддержкой резидентных вычислений, строгой согласованности данных и безопасной среды исполнения на языке Rust
- Производительность — современные NVMe-накопители, стогигабитные сети, оптимизация структуры кластера для снижения времени отклика
- Отказоустойчивость — надёжные комплектующие, устойчивые сетевые протоколы, резервирование в каждом модуле
- Сетевые интерфейсы — высокоскоростной Ethernet (100 Гбит/с) для подключения к внешним системам
- Мониторинг и управление — встроенные средства контроля состояния оборудования и ПО, система резервного копирования

Области применения ПАК

- Обработка и аналитика больших массивов данных в реальном времени
- Сервисы с высокими SLA
- Системы, требующие масштабирования при росте нагрузки
- Проекты с требованиями по импортозамещению и информационной безопасности

3. ОТЛИЧИТЕЛЬНЫЕ ЧЕРТЫ

Резидентная обработка критичных данных

- Резидентное хранение и обработка данных в оперативной памяти для систем с требованиями к задержке в **единицы миллисекунд**
- Надёжный источник данных на платформе Picodata: транзакционная модель, журналирование изменений (WAL), репликация и сегментирование
- Производительность до **100 000 операций в секунду на одно CPU-ядро** при простых транзакционных сценариях (внутренние тесты)
- Стабильная работа при утилизации оперативной памяти **до 80% и выше** при смешанных нагрузках чтения и записи
- Практическая ёмкость «горячего» слоя данных — **от 1 до 30 Тбайт** за счёт In-Memory хранения и сжатия данных

Производительность и масштабирование

- Архитектура, оптимизированная под современные x86-серверы и **NVMe-накопители** для WAL и snapshot
- Горизонтальное масштабирование кластера с ростом нагрузки **без остановки приложений**
- Линейный рост производительности при добавлении вычислительных узлов в пределах «горячего» контура данных
- Минимальные задержки межузловое взаимодействия за счёт оптимизированного сетевого и In-Memory дизайна

Надёжность и отказоустойчивость

- ACID-транзакции для всех операций записи и обновления данных
- Синхронная и асинхронная репликация внутри площадки и между ЦОД
- Восстановление после сбоя за счёт snapshot и WAL с минимальным временем простоя (минуты, а не часы)
- Резервирование критичных компонентов на уровне платформы и инфраструктуры

Сохранность и контроль доступа

- Единый журнал изменений (WAL) для обеспечения целостности данных
- Гибкая ACL-модель разграничения доступа
- Регулярные snapshot и журналы изменений для восстановления данных и аудита
- Возможность реализации RPO ≈ 0 в конфигурациях с синхронной репликацией

Интеграция и удобство эксплуатации

- SQL-интерфейс, близкий к PostgreSQL, без необходимости переобучения команд
- Использование как высокопроизводительного оперативного слоя данных между OLTP-системами и аналитическими платформами
- Снижение нагрузки на классические транзакционные СУБД и Data Warehouse

Быстрый и безопасный запуск

- Типовые конфигурации, подтверждённые нагрузочными тестами
- Автоматизированное развёртывание, исключающее ошибки ручной настройки
- Гарантируемый уровень производительности при соблюдении рекомендованной архитектуры

Контроль, управление и эксплуатация

- Встроенный мониторинг платформы и аппаратных ресурсов
- Преднастроенные пороговые значения и алерты по ключевым метрикам
- Централизованное управление и администрирование кластера
- Единый сервисный контур поддержки и сопровождения

Эксплуатация и сервис

- Единый сервисный центр и единая ответственность
- Оперативные исправления и рекомендации
- Паспорт **Машины** в комплекте
- Обучение команды заказчика

Экономика

- Лицензии Picodata уже в составе ПАК
- Сжатые сроки ввода в эксплуатацию
- Только необходимые компоненты
- Минимальные затраты на комплексную поддержку

4. ПОДТВЕРЖДЕННАЯ БЕЗОПАСНОСТЬ

Машина больших данных Скала^{Ар} МБД.Т поставляется с одной из сертифицированных операционных систем:

- **ОС Альт 8 СП** (сертификат ФСТЭК 3866 от 10.08.2018, действует до 10.08.2028)
- **РЕД ОС** (сертификат ФСТЭК №4060 от 12.01.2019, действует до 12.01.2029)
- **Astra Linux Special Edition** (сертификат ФСТЭК №2557 от 27.01.2012, действует до 27.01.2026)

ОС Альт 8 СП

Может применяться для защиты информации в:

- значимых объектах критической информационной инфраструктуры 1 категории
- государственных информационных системах 1 класса защищённости
- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости
- информационных системах персональных данных при необходимости обеспечения 1 уровня защищённости персональных данных
- информационных системах общего пользования 2 класса

ОС соответствует требованиям следующих нормативных документов:

- «Требования безопасности информации к операционным системам» (ФСТЭК России, 2016) и «Профиль защиты операционных систем типа А четвертого класса защиты. ИТ.ОС.А4.ПЗ» (ФСТЭК России, 2017) по 4 классу защиты
- «Требования по безопасности информации к средствам контейнеризации» (ФСТЭК России, 2022) по 4 классу защиты
- «Требования по безопасности информации к средствам виртуализации» (ФСТЭК России, 2022) по 4 классу защиты
- «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) по 4 уровню доверия

РЕД ОС

Может применяться для защиты информации в:

- значимых объектах критической информационной инфраструктуры 1 категории
- государственных информационных системах 1 класса защищённости
- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости

- информационных системах персональных данных при необходимости обеспечения 1 уровня защищённости персональных данных
- информационных системах общего пользования 2 класса

ОС соответствует требованиям следующих нормативных документов:

- «Требования безопасности информации к операционным системам» (ФСТЭК России, 2016) и «Профиль защиты операционных систем типа А четвертого класса защиты. ИТ.ОС.А4.ПЗ» (ФСТЭК России, 2017) по 4 классу защиты
- «Требования по безопасности информации к средствам контейнеризации» (ФСТЭК России, 2022) по 4 классу защиты
- «Требования по безопасности информации к средствам виртуализации» (ФСТЭК России, 2022) по 4 классу защиты
- «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) по 4 уровню доверия

Astra Linux Special Edition

Может применяться для защиты информации в:

- информационных системах, не предназначенных для обработки сведений государственной тайны
- государственных информационных системах 1 класса защищённости
- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости
- информационных системах персональных данных при необходимости обеспечения 1 уровня защищённости персональных данных
- информационных системах общего пользования 2 класса

ОС соответствует требованиям следующих нормативных документов:

- «Требования безопасности информации к операционным системам» (ФСТЭК России, 2016), «Профиль защиты операционных систем типа А первого класса защиты. ИТ.ОС.А1.ПЗ» (ФСТЭК России, 2017) по 1 классу защиты и «Профиль защиты ОС (А второго класса защиты. ИТ.ОС.А2.ПЗ» по 2 классу защиты
- «Требования по безопасности информации к средствам контейнеризации» (ФСТЭК России, 2022) по 1 классу защиты
- «Требования по безопасности информации к средствам виртуализации» (ФСТЭК России, 2022) по 1 классу защиты

«Требования по безопасности информации к системам управления базами данных» (ФСТЭК России, 2023) по 1 классу защиты «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и

средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) по 1 уровню доверия

Протестирована совместимость с наложенными средствами защиты

Сертифицированное антивирусное средство защиты **Kaspersky Endpoint Security для Linux** (сертификат ФСТЭК №2534 от 27.12.2011, действует до 27.12.2030).

ПО может применяться для защиты информации в:

- государственных информационных системах 1 класса защищённости
- информационных системах персональных данных при необходимости обеспечения 1 уровня защищённости персональных данных
- значимых объектах критической информационной инфраструктуры 1 категории
- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости
- информационных системах общего пользования 2 класса

ПО соответствует требованиям следующих нормативных документов:

- «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) — по 2 уровню доверия
- «Требования к средствам антивирусной защиты» (ФСТЭК России, 2012)
- «Профиль защиты средств антивирусной защиты типа Б 2 класса защиты. ИТ.САВЗ.Б2.13» (ФСТЭК России, 2012)
- «Профиль защиты средств антивирусной защиты типа В второго класса защиты. ИТ.САВЗ.В2.ПЗ» (ФСТЭК России, 2012)
- «Профиль защиты средств антивирусной защиты типа Г второго класса защиты»
- «Требования к средствам контроля сменных машинных носителей информации» (ФСТЭК России, 2014)
- «Профиль защиты средств контроля подключения сменных машинных носителей информации второго класса защиты. ИТ.СКН.П2.ПЗ» (ФСТЭК России, 2014)

Сертифицированное средство доверенной загрузки ПАК «Соболь» версия 4 (сертификат ФСТЭК №4043 от 05.12.2018, действует до 05.12.2028).

ПО может применяться для защиты информации в:

- государственных информационных системах 1 класса защищённости
- информационных системах персональных данных при необходимости обеспечения 1 уровня защищённости персональных данных
- значимых объектах критической информационной инфраструктуры 1 категории

- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости
- информационных системах общего пользования 2 класса

ПО соответствует требованиям следующих нормативных документов:

- «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) — по 2 уровню доверия
- «Требования к средствам доверенной загрузки» (ФСТЭК России, 2013)
- «Профиль защиты средства доверенной загрузки уровня базовой системы ввода-вывода второго класса защиты. ИТ.СДЗ.УБ2.ПЗ» (ФСТЭК России, 2013)

Сертифицированная система единой аутентификации **Avanpost FAM** (сертификат ФСТЭК №4492 от 13.12.2021, действует до 13.12.2026).

ПО может применяться для защиты информации в:

- государственных информационных системах 1 класса защищённости
- информационных системах персональных данных при необходимости обеспечения 1 и 2 уровня защищённости персональных данных
- значимых объектах критической информационной инфраструктуры 1 категории
- автоматизированных системах управления производственными и технологическими процессами 1 класса защищённости

ПО соответствует требованиям следующих нормативных документов:

- «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий» (ФСТЭК России, 2020) — по 4 уровню доверия

5. ПРИНЦИПЫ СОЗДАНИЯ МАШИНЫ МБД.Т

Машина больших данных Скала[^]р МБД.Т — программно-аппаратный комплекс для размещения и работы с данными в оперативной памяти. В основе ПАК лежит Picodata. Архитектура и состав компонентов изначально ориентированы на конкретную реализацию резидентной (In-Memory) СУБД, что обеспечивает предсказуемые характеристики производительности, масштабирования и отказоустойчивости.

Традиционный подход к крупномасштабным реляционным СУБД для задач OLTP опирается на вертикальное масштабирование и один активный вычислительный узел (остальные — для резервирования). При росте объёма хранения это ведет к нелинейному усложнению транзакций и росту задержек. Переход к ПАК, спроектированному под конкретную транзакционную СУБД, частично снимает ограничения, но не решает фундаментальную проблему вертикального потолка.

Комплексный подход к крупномасштабным реляционным СУБД на базе **Машины больших данных Скала[^]р МБД.Т**, что используется для задач OLTP, реализует иной принцип: горизонтальное масштабирование за счёт распределения нагрузки между активными экземплярами кластера и хранения рабочей части данных в оперативной памяти или на локальных накопителях, в зависимости от требования к производительности.

5.1 Основные принципы

Репликация и отказоустойчивость

Инстансы объединяются в репликasetы с одной активной и одной/несколькими резервными репликами. При отказе активной реплики резервная автоматически принимает роль активной. Коэффициент репликации задается глобально, обеспечивая требуемый баланс между производительностью и отказоустойчивостью.

Координация запросов

При выполнении распределённого SQL любой узел может брать на себя роль координатора запроса: узел разбивает план, рассылает части на соответствующие бакеты/репликasetы по RPC и агрегирует результат. Это динамическая роль запроса, а не отдельный тип узла.

Интерконнект и масштабирование

Внутрикластерная сеть (интерконнект) — 100 Гбит/с. Поддерживаются конфигурации от 3 до 125 вычислительных узлов, что позволяет хранить в оперативной памяти более 100 Тбайт данных в зависимости от количества реплик; масштабирование выполняется без остановки сервисов.

Персистентность In-Memory

Данные в памяти персистентны за счёт WAL и периодических снимков; для пользователя журналирование прозрачно. Это обеспечивает строгую консистентность и быстрое восстановление (RPO/RTO задаются политиками).

5.2 Основы резидентных СУБД

Традиционные СУБД оптимизированы под хранение данных на энергонезависимых блочных носителях — жёстких дисках и твердотельных накопителях. И хотя они разработаны с учётом требований обработки транзакций, операции вставки, обновления,

частичной выборки в этих СУБД остаются весьма медленными из-за необходимости частой и дублирующей синхронизации с устройством постоянного хранения.

Заметного роста производительности можно достичь, разместив традиционную реляционную базу данных непосредственно в оперативной памяти. И ряд коммерческих СУБД, в том числе Oracle Database, IBM DB2, Microsoft SQL Server, допускают такие варианты применения.

Ещё больший эффект достигается за счёт отказа от использования реляционных структур и снижения требований к согласованности данных: резидентные NoSQL-СУБД классов «ключ — значение» и «семейство столбцов» позволяют достичь высокой производительности благодаря горизонтальному масштабированию в таких условиях. Нереляционные агрегатные модели данных дают возможность разделять базы данных на сегменты по формальному признаку, например, значению хэш-функции от ключа, и сегменты, будучи в нереляционных условиях фактически независимыми, могут управляться отдельными экземплярами СУБД, располагаемыми на различных аппаратных узлах. Снижение требований к согласованности позволяет сохранять уровень производительности на уровне близком к тому, как если бы база данных не была распределена по сети и не содержала реплик данных, полагаясь на «согласованность в конечном счёте», когда рано или поздно для отдельно взятой записи система придёт в согласованное состояние.

Использование распределённой схемы позволяет достичь не только горизонтальной масштабируемости и заметного роста производительности при добавлении новых узлов, но и увеличить устойчивость системы путем наличия нескольких копий данных. Как следствие — потеря отдельного узла и даже группы узлов — не приводит к потере данных, хотя общая производительность при этом может уменьшиться.

Резидентные технологии ускоряют доступ к данным и их обработку. Так, традиционная СУБД выдерживает нагрузку до пары десятков тысяч запросов в секунду, резидентная СУБД на аналогичном оборудовании способна **обрабатывать сотни тысяч запросов** за то же время.

СУБД Picodata, изначально разработанная как резидентная NoSQL-СУБД класса «ключ — значение», успешно применяется в качестве распределённых кэшей, для сервисов метаданных, организации брокеров сообщений. Применить её в качестве основной СУБД для систем, требующих строгой согласованности, позволяют доработки, выполненные в Picodata — благодаря им система прочно занимает место в классе, обозначаемом как «NewSQL», соединяя свойства горизонтальной масштабируемости, отказоустойчивости и сверхвысокой производительности из мира NoSQL и обеспечивая максимально строгие транзакционные требования, которые ранее удовлетворялись только в классических реляционных СУБД.

Обеспечение строгой согласованности, реализация Raft-консенсуса, поддержка языков запросов SQL и GraphQL, а также сервер приложений на языках Rust, встроенный в Picodata и обеспечивающий прямой доступ ко резидентной базе, позволяют работать с ней как с полноценным транзакционным решением корпоративного уровня, обладающим при этом свойствами горизонтальной масштабируемости и сверхвысокой производительностью, унаследованным от распределённых резидентных NoSQL-систем.

6. СОСТАВ МАШИНЫ

Ниже приведены термины, используемые для комплектации **Машины больших данных Скала^р МБД.Т.**

Машина — это набор аппаратного и программного обеспечения в виде **Модулей Скала^р**, соединенных вместе для обеспечения определенного метода обработки данных или предоставления ИТ-сервисов с заданными характеристиками.

Подсистема — логическое объединение компонентов по функциональному признаку, с целью пояснения состава и принципов действия ПАК.

Модуль — это единица поставки **Машин**, выполняющая определенные функции в соответствии с её назначением. Модуль является единым и неделимым элементом спецификации и содержит набор аппаратных узлов и ПО.

Узел — это элемент, выполняющий определенную задачу в составе Модуля.

Комплекты поставки

Машины больших данных Скала^р МБД.Т поставляются в виде функционально полного набора **Модулей Скала^р** и комплектуются в соответствии с показателями назначения, полученными от заказчика. **Машина** включает в себя базовый комплект и в случае необходимости дополняется комплектом модулей расширения и/или специальными модулями.

Базовый комплект — это набор **Модулей Скала^р**, минимально-необходимый для функционирования всех подсистем, обеспечивающих выполнение основного функционала **Машины**.

Комплект модулей расширения — это набор **Модулей Скала^р**, позволяющий, увеличить производительность, объём хранения и портовую ёмкость. Кроме того, можно добавить специальные **Модули Скала^р**, позволяющие расширить функциональность ПАК.

6.1 Подсистемы

Функции **Машины больших данных Скала^р МБД.Т** логически объединены в подсистемы. Часть подсистем обеспечивают основной функционал и всегда включены в **Машину**, а часть — дополнительный функционал и могут быть добавлены по требованию заказчика.

Основной функционал — это минимальный набор подсистем, необходимых **Машине больших данных Скала^р МБД.Т** для выполнения задач прямого назначения.

Дополнительный функционал — набор подсистем из Модулей, обеспечивающих расширение функций **Машины больших данных Скала^р МБД.Т.**

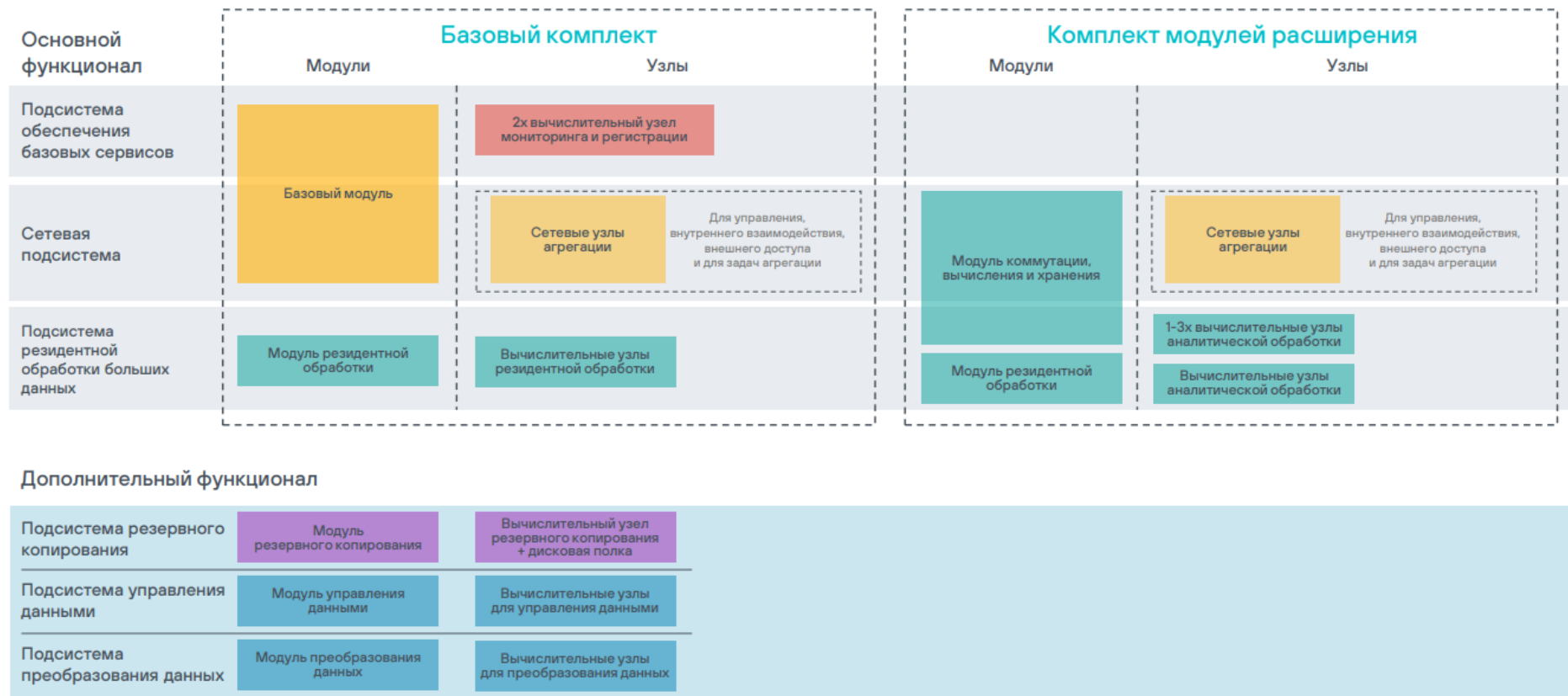


Рисунок 1. Комплектация Машины больших данных Скала^р МБД.Т

6.1.1 Подсистема обеспечения базовых сервисов и Сетевая подсистема

Подсистема обеспечения базовых сервисов отвечает за мониторинг и управление аппаратными и программными компонентами **Машины больших данных Скала^р МБД.Т.** В неё включены вычислительные узлы **Базового модуля** (см. п. 6.2.1), на которых предустановлено сервисное ПО **Скала^р Генум** и **Скала^р Визион**, выполняющее следующие функции:

- сбор, хранение и отображение данных на панелях мониторинга
- управление аппаратными компонентами
- управление пользователями и аутентификация (опционально)
- настройка программных компонентов
- настройка интеграции со сторонним ПО
- контроль и управление очередями подсистемы потоковой обработки данных

Архитектура подсистемы обеспечения базовых сервисов обеспечивает отказоустойчивый режим работы.

Сетевая подсистема выполняет функций организации сетевой связанности между всеми вычислительными узлами, входящими в состав **Машины больших данных Скала^р МБД.Т.**, и представляет собой набор сетевых узлов, которые организуют изолированные высокоскоростные сети:

- внутреннего взаимодействия (25 или 100 Гбит/с) — для организации быстрого функционирования между всеми компонентами ПАК
- внешнего доступа (25 или 100 Гбит/с) — для организации доступа к данным, что хранятся на узлах, входящих в состав подсистемы аналитической обработки данных
- управления (1 Гбит/с) — для организации передачи сервисной информации с вычислительных узлов, входящих в состав подсистемы аналитической обработки данных, на вычислительные узлы, входящие в состав подсистемы обеспечения базовых сервисов

Стартовый комплект сетевых узлов всегда размещается в **Базовом модуле** (см. п. 6.2.1).

6.1.2 Подсистема резидентной обработки больших данных

Основная подсистема, обеспечивающая работу основного функционального ПО. В этой подсистеме хранятся и обрабатываются данные в оперативной памяти или на локальных накопителях, в зависимости от конфигурации ПАК. Скорость обработки данных хранимых в модулях, входящих в подсистему, может отличаться в зависимости от SLA.

Подсистема реализуется **Модулем резидентной обработки** (см. п. 6.2.2).

6.1.3 Подсистема резервного копирования

Включает один или несколько **Модулей резервного копирования** (см. п. 6.2.3), в зависимости от объёма данных, подлежащих хранению в виде копий. В основе подсистемы лежит ПО, работающее по протоколу NFS.

6.1.4 Подсистема управления данными

Отвечает за организацию совместной работы с данными, интеграцию метаданных из различных систем обработки и анализа данных, а также предоставляет возможности поиска данных и совместной работы с метаданными, ведения корпоративного бизнес-гlossария и его тесной интеграции с каталогом данных. В основе подсистемы лежит ПО Arenadata Catalog.

Подсистема реализуется **Модулем управления данными** (см. п. 6.2.4).

6.1.5 Подсистема преобразования данных

Состоит из **Модулей преобразования данных** (см. п. 6.2.5), которые необходимы для создания аналитических платформ, а также для интеграции, выгрузки и обработки данных из любых источников. Является основой для выстраивания и оркестрации ETL/ELT-процессов.

6.2 Модули

6.2.1 Базовый модуль

Название в Едином реестре российской радиоэлектронной продукции — СКАЛА-Р Базовый модуль. Обеспечивает функционирование подсистемы обеспечения базовых сервисов и сетевой подсистемы.

Назначение

- Обеспечение сетевой связанности между компонентами
- Организация выделенной сети управления **Машиной**
- Организация подключения к сети заказчика
- Исполнение функций мониторинга и управления компонентами **Машины**

Узлы

- Два вычислительных узла мониторинга и регистрации, которые объединены в зеркальный кластер и используются для служебных функций
- Два сетевых узла 25/100 Гбит/с для организации внутреннего сетевого взаимодействия
- Два сетевых узла 25/100 Гбит/с для организации сети внешнего доступа (опционально)
- Сетевой узел 1 Гбит/с для организации работы сети управления, также может быть выполнен в отказоустойчивом исполнении
- Два сетевых узла 100 Гбит/с для организации агрегации, в случае добавления внутренних портов в крупных конфигурациях ПАК (опционально)

Отказоустойчивость обеспечена

- Резервированием вычислительных узлов, отвечающих за мониторинг и управление компонентами **Машины**
- Технологией RAID для накопителей вычислительных узлов
- Резервированием сетевых коммутаторов (объединение сетевых узлов в MLAG-пару)

Применяемое программное обеспечение

- **Скала^р Визион**
- **Скала^р Геном**
- ОС Альт 8 СП
- Сервисное ПО, входящее в состав Arenadata Cluster Manager (ADCM)
- ПО для управления пользователями и аутентификацией (опционально)

6.2.2 Модуль резидентной обработки

Название в Едином реестре российской радиоэлектронной продукции — СКАЛА-Р Модуль резидентной обработки. Обеспечивает функционирование подсистемы резидентной обработки больших данных.

Назначение

- Обработка и хранение данных в оперативной памяти с возможностью горизонтального масштабирования
- Обработка и хранения данных на локальных накопителях с возможностью горизонтального масштабирования
- Выполнение запросов и распределённые вычисления

Узлы

В состав Модуля входят вычислительные узлы, распределённые по 3 типам нагрузки:

- тип 1 – высокопроизводительный, необходимый для работы на высоких нагрузках
- тип 2 – наиболее сбалансированный, позволяющий хранить больше данных
- тип 3 – позволяющий хранить большой объем информации на локальных накопителях

Отказоустойчивость обеспечена

- Хранением данных минимум на 2 вычислительных узлах
- Технологией RAID для накопителей вычислительных узлов
- Резервированием сетевых коммутаторов (объединение сетевых узлов в MLAG-пару)

Применяемое программное обеспечение

- Picodata
- ОС Альт 8 СП

6.2.3 Модуль резервного копирования

Название в Едином реестре российской радиоэлектронной продукции — СКАЛА-Р Модуль резервного копирования. Обеспечивает функционирование подсистемы резервного копирования.

Назначение

- Резервирования и восстановления данных
- Хранение резервных копий

Узлы

В состав Модуля входит один вычислительный узел, обеспечивающий хранение до 94 Тбайт данных. Хранение осуществляется на накопителях NL-SAS

Отказоустойчивость обеспечена

- Технологией RAID для накопителей вычислительных узлов
- Резервированием сетевых коммутаторов (объединение сетевых узлов в MLAG-пару)

Применяемое программное обеспечение

Сетевая файловая система (NFS) — распределенная файловая система, которая обеспечивает пользователям доступ к файлам, расположенным на вычислительных узлах.

6.2.4 Модуль управления данными

Обеспечивает функционирование подсистемы управления данными (см. п. 6.1.4). Интерфейс основного функционального ПО Модуля предоставляет доступ к каталогу метаданных, бизнес-гlossарию, поиску, профилированию и проверке качества корпоративных данных.

Назначение

- Интеграции метаданных из различных систем обработки

- Анализа данных, поиска данных, совместной работы с метаданными
- Ведения корпоративного бизнес-гlossария и его интеграции с каталогом данных

Узлы

В состав Модуля входит не менее двух вычислительных узлов

Отказоустойчивость обеспечена

- Резервированием вычислительных узлов
- Технологией RAID для накопителей вычислительных узлов
- Резервированием сетевых коммутаторов (объединение сетевых узлов в MLAG-пару)

Применяемое программное обеспечение

- Arenadata Catalog (ADC)
- ОС Альт 8 СП

6.2.5 Модуль преобразования данных

Обеспечивает функционирование подсистемы преобразования данных (см. п. 6.1.5).

Назначение

Используется для решения задач, связанных с интеграцией данных, построения и наполнения хранилищ и витрин данных.

Узлы

В зависимости от модификации, Модуль может состоять из 2× или 3× вычислительных узлов.

Отказоустойчивость обеспечена

- Резервированием вычислительных узлов
- Технологией RAID для накопителей вычислительных узлов
- Резервированием сетевых коммутаторов (объединение сетевых узлов в MLAG-пару)

Применяемое программное обеспечение

- ПО для управления ETL-процессами

7. СПЕЦИФИЧНЫЕ ЧЕРТЫ

Проектирование и реализация **Машины больших данных Скала^р МБД.Т** осуществлялись с учётом ряда выбранных приоритетов, оказывающих непосредственное влияние на функциональные и эксплуатационные показатели.

Размещение данных в оперативной памяти и минимизация ввода-вывода

- Обеспечивает максимально возможную производительность подсистемы хранения
- Повышает производительность, так как данные хранятся в той же памяти, что и исполняемые программные процедуры
- Повышает надёжность работы **Машины**, потому что исключается влияние отказов накопителей на базу данных

Программно-определяемый приоритет повышенной устойчивости решения или повышенной производительности

Приоритет повышенной устойчивости:

- Гарантия сохранности данных при любых отказах
- В случае сбоя система продолжает работать без снижения производительности, простоя или потери данных (возможна кратковременная задержка на период назначения платформой другой реплики в качестве мастера вместо вышедшей из строя — несколько микросекунд)

Приоритет производительности:

- Снижение времени отклика за счёт большего распараллеливания

Выбор аппаратного решения для реализации процессов вычисления и хранения

- Обеспечивает максимум производительности, так как нет потерь на среду виртуализации и прочие потери сведены к минимуму
- Повышает надёжность **Машины**, потому что нет дополнительного программного уровня

Применение высоконадёжного и производительного оборудования в качестве платформы для размещения компонентов решения

- Обеспечивает надёжность **Машины** и стабильно высокий уровень производительности за счёт использования компонентов, проверенных временем
- Снижает стоимость сопровождения за счёт доступности элементов, требующих замены при выходе из строя

Выбор ПО с открытым кодом и отечественных разработок

- Повышает производительность за счёт доработки ПО силами **Скала^р** и партнёров
- Повышает надёжность **Машины**, так как снижены риски недоступности поддержки
- Снижает зависимость от импортных производителей ПО

8. ГАРАНТИРОВАННОЕ КАЧЕСТВО И ПОЛНАЯ ГОТОВНОСТЬ К ПРОМЫШЛЕННОЙ ЭКСПЛУАТАЦИИ

Качественные показатели **Машины больших данных Скала^р МБД.Т** обеспечиваются её соответствием проверенному стандартному варианту, соблюдением установленных норм и требований по формированию, реализацией работ высококвалифицированными специалистами на всех этапах жизненного цикла.

Производство и сборка: без компромиссов

- Используются только **проверенные комплектующие**, отобранные под реальные нагрузки
- Все компоненты собираются строго по регламенту, в соответствии с утверждённой схемой размещения
- Развёртывание программного обеспечения и первичная настройка выполняются **автоматизировано**, чтобы исключить «ручные» ошибки
- Все машины перед отгрузкой проходят функциональное тестирование и проверку на наличие известных уязвимостей
- Отклонения от типового решения **Машины больших данных Скала^р МБД.Т** исключены

Передача в эксплуатацию: все готово к работе

- **Машина** поставляется в полной готовности к работе — готова к эксплуатации сразу после подключения к сети заказчика
- В комплект входят: паспорт изделия и сертификат поддержки, полный пакет документации для прохождения аттестаций и согласований, обучение специалистов заказчика (по запросу)
- По запросу проводится обучение работе с **Машиной больших данных Скала^р МБД.Т** для специалистов заказчика

Техническая поддержка: от производителя, без посредников

- Поддержка входит в поставку (по умолчанию — 1 год, оптимально — 3 или 5 лет)
- Доступны пакеты технической поддержки 9×5 или 24×7, в зависимости от критичности системы заказчика
- Первая и вторая линии поддержки непосредственно от производителя или сертифицированного партнёра
- В сложных ситуациях в работу подключаются архитекторы и разработчики самой **Машины** — 3-я линия поддержки в России, без эскалации за рубеж

Сопровождение и развитие под задачи бизнеса

По запросу возможно:

- аппаратное расширение и модернизация
- горизонтальное и вертикальное масштабирование
- адаптация **Машины** под профильные нагрузки

Все доработки выполняются с участием тех, кто проектировал и создавал **Машину больших данных Скала^р МБД.Т**, что снижает риски и обеспечивает стабильность в эксплуатации.

9. РЕАКЦИЯ МАШИНЫ НА ВОЗМОЖНЫЕ ОТКАЗЫ

Отказы, связанные со стандартными элементами Скала^р МБД.Т

В **Машине больших данных Скала^р МБД.Т** обеспечена отказоустойчивость её основных элементов и процессов, в том числе:

- узлов (дублирование процессоров, источников питания и др.)
- подсистемы ввода-вывода (RAID)
- сети внутреннего взаимодействия (дублирование сетевых интерфейсов)
- системы резервного копирования

Отказы перечисленных элементов отрабатываются стандартными алгоритмами в соответствии с произведёнными настройками. Любой единичный отказ не повлияет на доступность системы в целом, хотя по конкретному сервису возможно небольшое снижение производительности. После устранения неисправности исходная производительность **Машины** также восстанавливается.

Отказы, связанные с узлами кластера

Для обеспечения бесперебойности доступа и сохранности данных в решении реализован многоузловой кластер, состоящий из набора сегментов (наборов реплик). В рамках каждого сегмента один из узлов выполняет функцию мастер-сервера БД с возможностью чтения и записи данных, остальные — функцию реплики с возможностью чтения.

В случае отказа любого узла кластера, исполняющего роль мастера, платформа в автоматическом режиме назначает на эту роль один из узлов с репликой, изменяет общую структуру кластера (уменьшив на одну реплику соответствующий сегмент) и продолжает обработку запросов в обновленной конфигурации. В случае отказа узла с репликой — переключение не требуется, выполняется только актуализация конфигурации кластера. Для пользователей такие отказы остаются незаметными, за исключением потенциального повышения времени отклика в периоды пиковых нагрузок.

После завершения обслуживания или устранения причины отказа и восстановления узла размещенные в сегменте данные будут скопированы на него средствами платформы и произведено переконфигурирование кластера (добавление узла в сегмент).

В случае полного отказа всех узлов сегмента платформа произведет переконфигурирование кластера путем выделения под данные этого сегмента узлов из других сегментов или путем включения данных сегмента в другой сегмент. При этом последующее восстановление данных сегмента осуществляется из актуальной резервной копии.

Конкретный алгоритм реагирования кластера определяется заказчиком в ходе формирования решения на платформе и может быть скорректирован в произвольный момент времени.

Детальный порядок обеспечения отказоустойчивости кластера и рекомендации по действиям при его администрировании в той или иной конкретной ситуации с конкретным экземпляром **Машины больших данных Скала^р МБД.Т** могут быть предоставлены по запросу.

10. ВАРИАТИВНОСТЬ РЕШЕНИЯ



Скала^р МБД.Т К-1



Скала^р МБД.Т К-2



Скала^р МБД.Т К-3

Рисунок 2. Варианты поставки МБД.Т (примеры)

Таблица 1. Типовые комплекты поставки Машины больших данных Скала^р МБД.Т

Параметры модели	К-1	К-2	К-3
Количество секций (стоек)	2	16	1
Общее энергопотребление	до 15 кВт	до 120 кВт	до 15 кВт
Энергопотребление на стойку	до 7,5 кВт	до 7,5 кВт	до 15 кВт
Количество серверов - узлов вычисления и хранения	16	128	16
Количество In-Memory платформ	1	1	1
Общий объем БД, Тбайт (при минимум двух репликах на один шард)	до 6	до 30	до 6
Объем хранения системы резервного копирования (СРК), Тбайт	до 180	до 1440	до 180

11. ТРЕБОВАНИЯ К РАЗМЕЩЕНИЮ МАШИНЫ

Машина больших данных Скала^Ар МБД.Т представляет собой комплект узлов для размещения в серверный монтажный шкаф 19", высота 42U и больше, с дальнейшей возможностью модульной расширяемости до 14 стоек.

Монтажный шкаф (стойка) может быть поставлена как опция.

Для подключения шкафа к системе электроснабжения должны быть предусмотрены два независимых входа электропитания.

Расчетная потребляемая мощность шкафа (задается параметрами ЦОД заказчика) определяет топологию размещения модулей и узлов в стойках ЦОД и учитывается при расчете **Машины**. От этого зависит количество дополнительного коммутационного оборудования в составе **Машины**.

В месте установки должны быть предусмотрены соответствующие мощности по отводу тепла.

Для подключения к локальной сети заказчика необходим резервированный канал до 4×100 Gigabit Ethernet или до 8×10/25 Gigabit Ethernet. Требуемые трансиверы определяются на этапе формирования спецификации **Машины**.

При развёртывании будут выполнены настройки сетевых адресов в соответствии со структурой сети заказчика. Заказчик должен предоставить необходимые данные в соответствии с номенклатурой компонентов **Машины больших данных Скала^Ар МБД.Т**.

В сети заказчика должны быть настроены соответствующие маршруты и права доступа.

Дальнейшие мероприятия по вводу в эксплуатацию осуществляются заказчиком путём настройки прикладных программных систем.

12. ТЕХНИЧЕСКАЯ ПОДДЕРЖКА

Поставка **Машин больших данных Скала^р МБД.Т** осуществляется с предварительными сборкой, тестированием и настройкой оборудования согласно требованиям заказчика. Качественная поддержка обеспечивается едиными стандартами гарантийного и постгарантийного технического обслуживания:

- пакет услуг по технической поддержке на первый год включен в поставку
- заказчик может выбирать пакет 9×5 или 24×7 (вариант для комплексов критической функциональности)
- срок начально приобретаемой технической поддержки может быть увеличен до 3-х и 5-и лет, также доступна пролонгация поддержки

Состав типовых пакетов услуг по технической поддержке **Машин больших данных Скала^р МБД.Т** представлен в таблице 2.

Таблица 2. Пакеты услуг по технической поддержке

Услуги	Пакет «9×5»	Пакет «24×7»
«Режим предоставления услуг 9×5» (в рабочее время по рабочим дням)	+	—
«Режим предоставления услуг 24×7» (круглосуточно)	—	+
Предоставление доступа к системе регистрации запросов/инцидентов Service Desk	+	+
Предоставление доступа к базе знаний по продуктам Скала^р	+	+
Предоставление обновлений лицензионного ПО Скала^р	+	+
Диагностика, анализ и устранение проблем в работе комплекса Скала^р, включая: ■ устранение аппаратных неисправностей ■ техническое сопровождение ПО	+	+
Консультации по работе комплекса Скала^р	+	+
«Защита конфиденциальной информации» (неисправные носители информации не возвращаются заказчиком)	Опция	Опция
Замена и ремонт оборудования по месту установки	+	+

Услуги	Пакет «9×5»	Пакет «24×7»
Доставка оборудования на замену за счет производителя	+	+
Расширенные параметры обслуживания	—	+
Времена реагирования и отклика, не более:		
Время регистрации обращений	30 минут, рабочие часы (9×5)	30 минут, круглосуточно (24×7)
Подключение специалиста к решению инцидентов критичного и высокого уровней	В течение 1 рабочего часа (9×5)	В течение 1 часа (24×7)

Примечание к срокам ремонта оборудования

Комплекс **Машина больших данных Скала[^]р МБД.Т** архитектурно является устойчивым к выходу из строя отдельных компонентов и даже узлов, поэтому нет необходимости в обеспечении дорогостоящего сервиса срочного восстановления оборудования в течение суток и менее. В комплексе предусмотрено как минимум двойное резервирование основных компонентов, позволяющее сохранять данные и работоспособность даже при выходе из строя нескольких дисков и/или серверов (узлов).

13. ЛИЦЕНЗИРОВАНИЕ ПО В СОСТАВЕ МОДУЛЕЙ

Команда **Скала^р** активно занимается развитием программных продуктов **Машин больших данных Скала^р МБД.8**. Направления развития формируются на основе анализа мирового опыта использования систем подобного класса и пожеланий заказчиков и партнеров. Новые функции реализуются в форме релизов, которые могут выходить несколько раз в год.

Программное обеспечение Arenadata Catalog лицензируется по количеству пользователей с правами администратора или модератора.

Программное обеспечение **Скала^р Геном**, **Скала^р Визион** поставляется исключительно в составе **Машин Скала^р** и лицензируется по метрикам комплекса в соответствии с количеством серверных узлов.

13.1 Политика обновления ПО

Команда **Скала^р** активно занимается развитием собственных программных продуктов. Направления развития формируются на основе анализа мирового опыта использования систем подобного класса и пожеланий заказчиков и партнеров. Новые функции реализуются в форме релизов. Обновления для **Машин**, находящихся в эксплуатации, производятся по согласованию с заказчиком.

О КОМПАНИИ

Скала^р — модульная платформа для построения высоконагруженной ИТ-инфраструктуры, продукт Группы Rubytch.

Программно-аппаратные комплексы (**Машины**) **Скала^р** выпускаются с 2015 года и представляют широкий технологический стек для построения динамических инфраструктур и инфраструктур управления данными высоконагруженных информационных систем.

Продукты **Скала^р** включены в Реестр промышленной продукции, произведенной на территории Российской Федерации, и в Единый реестр российских программ для ЭВМ и БД. Соответствует критериям доверенности и использованию для объектов критической информационной инфраструктуры (КИИ).

Машины Скала^р являются серийно выпускаемыми преднастроенными комплексами, которые быстро разворачиваются и вводятся в эксплуатацию. Глубокая интеграция технических средств и программного обеспечения в ПАК **Скала^р** позволяет получить расширенные возможности и функциональность, которые недоступны при использовании отдельных компонентов.

Модульный принцип обеспечивает интеграцию разнородных компонентов ИТ-инфраструктуры в единую платформу предприятий, корпораций и ведомств. Единые поддержка и сервисное обслуживание для всех продуктов линейки **Скала^р** от производителя обеспечивают оперативное разрешение инцидентов на стыке технологий.

Дополнительная информация — на сайте www.skala-r.ru.