

Машина хранения данных Скала[^]р МХД.О

Программно-аппаратный комплекс
на основе технологии объектного хранения S3
для хранения миллиардов объектов и быстрого
доступа к ним

Технический обзор

версия 2.5 от 21.04.2025



ОГЛАВЛЕНИЕ

| | |
|---|-----------|
| Перечень терминов и сокращений | 5 |
| 1. Предисловие | 7 |
| 1.1 Описание документа | 7 |
| 1.2 Аудитория | 7 |
| 1.3 Обратная связь | 7 |
| 2. Введение | 8 |
| 3. Отличительные черты | 9 |
| 4. Производство в Российской Федерации | 11 |
| 5. Архитектура | 13 |
| 6. Структура | 14 |
| 6.1 Модули Машины хранения данных Скала^р МХД.О | 14 |
| 6.1.1 Базовый модуль | 15 |
| 6.1.2 Модули хранения | 16 |
| 6.1.3 Модули коммутации и хранения..... | 18 |
| 6.1.4 Базовые модули агрегации..... | 19 |
| 6.2 Программное обеспечение | 20 |
| 6.2.1 ПО Скала^р Спектр S3..... | 20 |
| 6.2.2 ПО Скала^р Визион..... | 21 |
| 6.2.3 ПО Скала^р Геном | 21 |
| 6.3 Распределенный дисковый массив..... | 21 |
| 6.4 Отказоустойчивость | 22 |
| 7. Модернизация и обслуживание | 23 |
| 8. Сценарии использования | 24 |
| 8.1 Локальное хранилище S3 | 24 |
| 8.2 Распределенное хранилище S3 с асинхронной георепликацией..... | 25 |
| 9. Планирование инфраструктуры | 26 |
| 10. Границы применимости | 27 |
| 11. Техническая поддержка | 28 |
| 12. Поставка и лицензирование ПО | 30 |
| О Компании | 31 |

Информация, представленная в документе, носит исключительно информационный характер, является актуальной 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 |
|--------------------|--|
| Active Directory | Сервис корпорации Microsoft для управления и хранения информации о пользователях, компьютерах и других ресурсах в компьютерных сетях. Он предоставляет централизованное хранилище для всех этих данных, позволяя организовывать и контролировать доступ к ним. |
| API | (англ. Application Programming Interface) Интерфейс программирования приложений |
| BLOB | (от англ. Binary Large Object — «двоичный большой объект») Тип данных, предназначенный для хранения больших объёмов двоичных данных: изображений, видео, мультимедийных файлов и т. д. |
| Erasurе Coding | Метод защиты данных, используемый в проектировании систем для обеспечения надёжности и доступности данных. |
| GW | Сервисный компонент обработки запросов S3, реализующий методы S3 API |
| HDD | (англ. Hard disk drive) Твердотельный накопитель без подвижных частей |
| IAM | (англ. Identity and Access Management) Сервис, набор инструментов для аутентификации и авторизации пользователей в прикладных приложениях |
| IP адрес | (от Internet Protocol) Адрес сетевого устройства в сети Internet |
| LDAP | (англ. Lightweight Directory Access Protocol) Протокол для работы с данными, чаще всего учетными записями пользователей, организованными в виде дерева (каталог) |
| NS и OS | Сервисы S3, обеспечивающие хранение (мета)данных объектов |
| NTP | (англ. Network Time Protocol) Протокол для синхронизации времени между устройствами (коммутаторы, серверы, рабочие станции и др.) с источником точного времени, рассчитанный на работу в сетях низкой надежности |
| Postgres | Название СУБД (системы управления базами данных) |
| SSD | (англ. Solid-State Drive) Запоминающее устройство на основе микросхем памяти |

| Термин, сокращение | Определение |
|--------------------|---|
| S3 (Amazon S3) | (англ. Simple Storage Service, «три S») Облачная система хранения в составе Amazon Web Services, организованная по объектному принципу |
| S3 хранилище | Сервис хранения файлов с данными в форме объектов. От обычного хранения файлов хранение объектов в S3 хранилище отличается форматом хранения, наличием метаданных и уникальных идентификаторов объектов, которые дают возможность организации хранения миллиардов файлов, что не доступно для большинства файловых хранилищ |
| БД | База данных, сокр. |
| ЖАМ | Аббревиатура от «журналирование, аудит, мониторинг»; совокупность операций для управления доступом к данным, особенно с точки зрения обеспечения безопасности |
| ИТ-индустрия | (сокр. от «информационные технологии») — сегмент науки и технологии, занимающийся созданием, внедрением и поддержкой цифровых инноваций и сложных технических систем. |
| ОС | Операционная система |
| ПО | Программное обеспечение |
| СУБД | Система управления базами данных, сокр. |
| Тенант | Логический объект, включающий в себя совокупность вычислительных ресурсов и пользователей, выделенных в пользование Заказчику |
| ЦОД | Центр обработки данных |
| ЭВМ | Электронная вычислительная машина, сокр. — комплекс технических, аппаратных и программных средств, предназначенных для автоматической обработки информации, вычислений, автоматического управления. |

1. ПРЕДИСЛОВИЕ

1.1 Описание документа

Этот технический обзор дает концептуальный и архитектурный обзоры **Машины хранения данных Скала^р МХД.О**

Брошюра описывает то, как оптимизированные программно-аппаратные комплексы отвечают современным вызовам, и фокусируется на **Машине хранения данных Скала^р МХД.О** как одном из лидирующих решений в этом сегменте.

1.2 Аудитория

Эта брошюра предназначена для сотрудников компании **Скала^р**, партнеров и Заказчиков, перед которыми ставятся задачи разработки решения, закупки, управления или эксплуатации **Машины хранения данных Скала^р МХД.О**.

1.3 Обратная связь

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

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

2. ВВЕДЕНИЕ

Машина хранения данных Скала^р МХД.О предназначена для создания горизонтально масштабируемого объектного хранилища, совместимого с Amazon S3.

S3 хранилище — это сервис хранения файлов с данными в форме объектов. От обычного хранения файлов хранение объектов в S3 хранилище отличается форматом хранения, наличием метаданных и уникальных идентификаторов объектов, которые дают возможность организации хранения миллиардов файлов, что не доступно для большинства файловых хранилищ. Одной из актуальных задач S3 хранилища является перенос в него массивов не табличных данных (файлов, бинарных объектов и т.п.) из реляционных СУБД для повышения их производительности и упрощения эксплуатации.

Основными преимуществами S3 хранилища **Машины хранения данных Скала^р МХД.О** являются:

- Высокая надежность и отказоустойчивость
- Масштабируемость — возможность создания хранилищ практически неограниченных размеров
- Хранение неограниченного количества объектов — одно из основных преимуществ, достигаемое благодаря тому, что адреса доступа к объектам хранятся в виде ссылок, а не как имена
- Наличие версионности объектов
- Поддержка мультитенантности
- Обеспечение катастрофоустойчивости за счет использования георепликации;
- Классы обслуживания – возможность при сохранении объекта определять, на каком типе хранилища он будет помещен
- Высокая эффективность хранения за счет сжатия данных
- Интеграция с IAM
- Расширенный аудит, мониторинг и журналирование работы системы
- Ролевая модель доступа
- Обеспечение целостности данных на уровне хранилищ

В настоящий момент **Машина хранения данных Скала^р МХД.О** является единственным продуктом российского производства, обеспечивающим реализацию объектного хранилища практически любого размера и совместимого с Amazon S3.

Начиная с 2014 года продукты **Скала^р** планомерно развивались, превращаясь из специальных решений, предназначенных для провайдеров интернет, в расширенные системы корпоративного класса. На сегодняшний день решения **Скала^р** являются безусловными лидерами на российском рынке корпоративных средств виртуализации и хранения данных за счет своих функциональных возможностей и довольно простых процессов внедрения и эксплуатации.

3. ОТЛИЧИТЕЛЬНЫЕ ЧЕРТЫ

1. Надежное хранение больших объемов данных

- Объем хранения данных — до 8 Пбайт на одну **Машину**
- Хранение нескольких копий или использование технологии Erasure Coding
- Проработанные варианты для типовых применений с разными требованиями к объему хранения и производительности
- Настраиваемая отказоустойчивость **Машины хранения данных Скала^р МХД.О**
- Защита репликацией на географически удаленные кластеры **Машины хранения данных Скала^р МХД.О**
- Формирование катастрофоустойчивых решений

2. Высокая производительность

- Производительность до 5000 операций/с на один узел (для объектов размером 40 Кбайт) на каждый узел хранения **Машины хранения данных Скала^р МХД.О**
- Архитектурная оптимизация производительности
- Специальные настройки программного обеспечения
- Сбалансированный комплект оборудования
- Поддержка многопоточной загрузки (Multiload)

3. Отказоустойчивость на всех уровнях

- Резервирование значимых компонентов на аппаратном уровне
- Обеспечение отказоустойчивости на программном уровне
- Автоматическое восстановление работоспособности при сбоях
- Настройка доменов отказоустойчивости с учетом физического размещения накопителей

4. Обеспечение качества при развертывании

- Оптимальность настроек, проверенная тестами
- Автоматизированное развертывание исключает человеческие ошибки
- Стандартизация развертывания гарантирует соответствие решения заявленным характеристикам

5. Непрерывный контроль состояния

- Мониторинг работоспособности ПО хранилища и оборудования
- Преднастроенные пороговые значения критичных параметров
- Различные каналы информирования об отклонениях

6. Гибкие возможности администрирования

- Проработанные рекомендации по выполнению отдельных операций
- Использование подходов, уменьшающих вероятность критических ошибок
- Выполнение любых операций администрирования без остановки **Машины хранения данных Скала^р МХД.О** и без прерывания сервиса конечным пользователям
- Интеграция с IAM и ЖАМ, возможна интеграция с Active Directory/LDAP

7. Поддержка эксплуатации

- Централизованная поддержка решения (одно окно)
- Единая ответственность за весь комплекс
- Выпуск патчей и рекомендаций
- Паспорт **Машины хранения данных Скала^р МХД.О** в комплекте
- Обучение персонала Заказчика

8. Экономическая эффективность

- Высокая эффективность хранения за счет сжатия объектов
- Организация на базе одной **Машины хранения данных Скала^р МХД.О** нескольких независимых сервисов (мультиотенантность)
- Возможность полного лицензирования емкости хранения или использования подписки на определенную емкость хранения
- Сокращенные сроки ввода в эксплуатацию
- Только обоснованно необходимые для корпоративных решений компоненты

9. Альтернатива Amazon S3 и Ceph Object Gateway

- Совместимость с протоколом Amazon S3
- Сбалансированный набор управляющих команд и простота администрирования
- Локальная реализация S3 в вашем ЦОД или частном облаке
- Качество, подтвержденное опытом практического применения

4. ПРОИЗВОДСТВО В РОССИЙСКОЙ ФЕДЕРАЦИИ

Специалистами компании **Скала^р** была проведена существенная работа по созданию схем и конструктивного исполнения **Машины хранения данных Скала^р МХД.О**, основанного на принципе модульности. Результаты проведенной работы на сегодняшний день не имеют аналогов на рынке РФ.

Машина хранения данных Скала^р МХД.О присутствует в Едином реестре российской радиоэлектронной продукции Минпромторга РФ (РЭП МПТ) согласно Постановлению Правительства РФ № 878.

Машина хранения данных Скала^р МХД.О может поставляется единым комплексом, одной номенклатурной позицией как Программно-аппаратный комплекс (ПАК). При этом **Машина** состоит из набора отдельных модулей (каждый из которых также является изделием в реестре РЭП МПТ), что обеспечивает гибкость комплектации и модернизации товарными позициями из реестра.

Машина хранения данных Скала^р МХД.О признана произведенным в РФ товаром, в соответствии с Правилами выдачи заключения о подтверждении производства промышленной продукции на территории Российской Федерации, утвержденными постановлением Правительства от 17 июля 2015 г. № 719.

Машина хранения данных Скала^р МХД.О соответствует постановлению Правительства РФ № 616 от 30 апреля 2020 г. о запрете на закупку импортной радиоэлектронной продукции и постановлению Правительства РФ № 925 от 16 сентября 2016 г. о приоритете российской радиоэлектронной продукции в 30%.

Машина хранения данных Скала^р МХД.О соответствует постановлению Правительства РФ № 2013 и № 2014 от 03 декабря 2020 г. о минимальной доле закупок товаров российского происхождения.

ВНИМАНИЕ! Реестровое написание наименования Машины СКАЛА-Р МХД.О отличается от маркетингового написания с применением товарного знака **Скала^р**.

Товарные позиции Машин и Модулей представлены ниже (Таблица 1).

Таблица 1. Товарные позиции Машин и Модулей

| Код по ОКПД 2 | Примечание |
|---|-------------|
| 26.20.14.160. Программно-аппаратные комплексы, созданные на серверах или устройствах, содержащие в своем составе один или более вычислительных узлов | Для Машин |
| 26.20.14.160. Программно-аппаратные комплексы, созданные на серверах или устройствах, содержащие в своем составе один или более вычислительных узлов и 26.20.14. Машины вычислительные электронные цифровые, поставляемые в виде систем для автоматической обработки данных | Для Модулей |

Наличие **Машины хранения данных Скала^р МХД.О** на сайте государственной информационной системы промышленности показано ниже (Рисунок 1, фрагмент страницы [Каталог продукции ГИСП](#)).

Машина хранения данных СКАЛА-Р МХД.О (РМБГ.466535.002-444)

ООО "СКАЛА-Р"



Рисунок 1. Машина хранения данных Скала^р МХД.О на сайте государственной информационной системы промышленности (ГИСП)

Подробная информация о **Машинах** и Модулях **Машины хранения данных Скала^р МХД.О**, включенных в Единый реестр российской радиоэлектронной продукции, представлена в таблицах ниже (Таблица 2 и Таблица 3).

Машины и Модули различаются исполнением (код .0X в номере конструкторской документации). Предлагается на момент написания данного документа 6 исполнений, различающихся серверной платформой (материнской платой производства РФ).

Таблица 2. Подробная информация о Машине хранения данных Скала^р МХД.О, включенных в РЭП МПТ

| Наименование Машины (разработан согласно Техническим условиям РМБГ.466535.002ТУ) | Код изделия по ОКПД2 |
|---|----------------------|
| Машина хранения данных СКАЛА-Р МХД.О (РМБГ.466535.002-444) | 26.20.14.160 |

Таблица 3. Подробная информация об основных Модулях Машины хранения данных Скала^р МХД.О, включенных в РЭП МПТ

| Наименование Модуля (разработан согласно Техническим условиям РМБГ.466535.003ТУ) | Код изделия по ОКПД2 |
|---|----------------------|
| СКАЛА-Р Базовый модуль (РМБГ.466535.003-210) | 26.20.14.160 |
| СКАЛА-Р Модуль хранения (РМБГ.466535.003-262) | 26.20.14.160 |
| СКАЛА-Р Модуль коммутации и хранения (РМБГ.466535.003-265) | 26.20.14.160 |
| СКАЛА-Р Базовый модуль агрегации ¹ (РМБГ.466535.003-11.0x) | 26.20.14 |

¹ Устанавливается для подключения к Базовому модулю дополнительных Модулей коммутации и хранения

5. АРХИТЕКТУРА

Машина хранения данных Скала^Ар МХД.О реализуется на узлах с архитектурой x86-64 с установленными накопителями, предназначенными для хранения данных. Архитектура Машины представлена на рисунке (Рисунок 2).

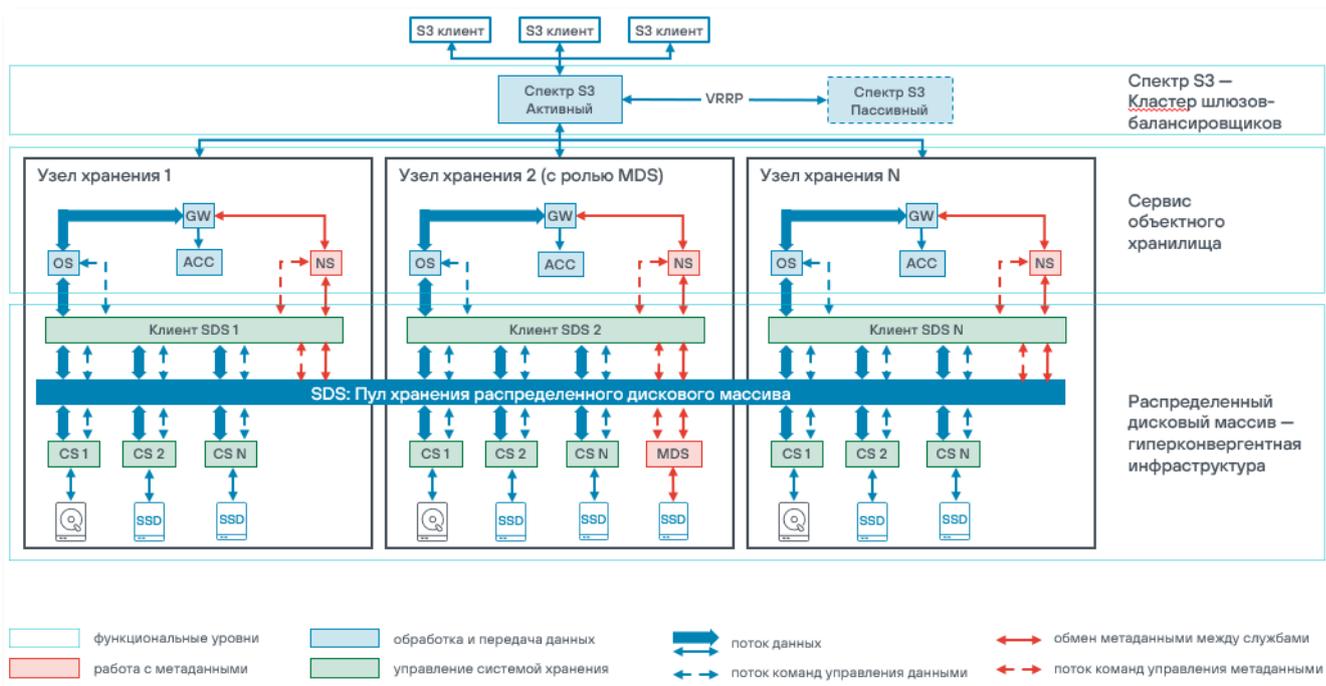


Рисунок 2. Архитектура Машины хранения данных (Схема внутренней коммутации)

Машина хранения данных Скала^Ар МХД.О включает следующие основные подсистемы:

- **Спектр S3** обеспечивает балансировку запросов между узлами хранения для обеспечения равномерной нагрузки на **Машину**, выполняет сжатие объектов, интеграцию с провайдерами аутентификации (IAM), возможность интеграции с LDAP или Active Directory и расширенное журналирование
- **GW** - сервисный компонент обработки запросов S3, реализующий методы S3 API. GW получает и обрабатывает запросы, сформированные с использованием протокола Amazon S3, выполняет аутентификацию пользователей S3 и проверку списков контроля доступа (ACL)
- Сервис **OS** обеспечивает хранение данных объектов на распределенном дисковом хранилище
- Сервис **NS** обеспечивают хранение метаданных объектов, включающих имя объекта, его размер, список контроля доступа (ACL), расположение, владельца, и др.
- Распределенное дисковое хранилище представляет собой кластер с функционалом обеспечения высокой доступности служб и данных. Функционал высокой доступности задействуется и для обеспечения гарантированной доступности сервисов S3 (NS и OS) за счет их переноса на работающие узлы в случае сбоя узла, на котором они работали.

Программные компоненты **Машины хранения данных Скала^Ар МХД.О** запускаются на узлах как сервисы, поэтому для работы сервиса S3 не нужны виртуальные среды, что упрощает эксплуатацию комплекса. **Спектр S3** реализуется на двух выделенных узлах и работает в режиме отказоустойчивого кластера.

6. СТРУКТУРА

Машина хранения данных Скала^р МХД.О имеет модульную структуру, включающую Модули хранения данных и Модули, обеспечивающие служебные функции. В зависимости от требований к производительности и емкости хранения **Машины хранения данных Скала^р МХД.О** состав Модулей конечного изделия меняется — **Машина хранения данных Скала^р МХД.О** может поставляться в разных исполнениях.

Машина хранения данных Скала^р МХД.О поставляется как готовый преднастроенный комплекс, однако в процессе эксплуатации состав Модулей может расширяться для повышения емкости хранимых данных или увеличения производительности.

Машина хранения данных Скала^р МХД.О является отказоустойчивым комплексом, допускающим единичные отказы оборудования без потери работоспособности. Для случаев, требующих реализации катастрофоустойчивости, **Машина хранения данных Скала^р МХД.О** предоставляет возможность создания копии данных на удаленной площадке с использованием асинхронной репликации (георепликация). Катастрофоустойчивое решение реализуется на базе двух (и более) **Машин хранения данных Скала^р МХД.О**.

6.1 Модули Машины хранения данных Скала^р МХД.О

В состав **Машины хранения данных Скала^р МХД.О** (Рисунок 3), в зависимости от конфигурации, могут входить следующие Модули:

- Базовый модуль — шасси с установленными узлами, реализующими базовый набор сервисов, необходимых для работы **Машины хранения данных Скала^р МХД.О**, включая обеспечение сетевой связности и балансировку запросов пользователей
- Модули хранения — состоят из узлов с установленными накопителями, предназначенными для хранения данных
- Модули коммутации и хранения — предназначены для расширения возможностей Базового модуля по подключению Модулей хранения
- Базовые модули агрегации — служат для подключения Модулей коммутации и хранения к Базовому модулю

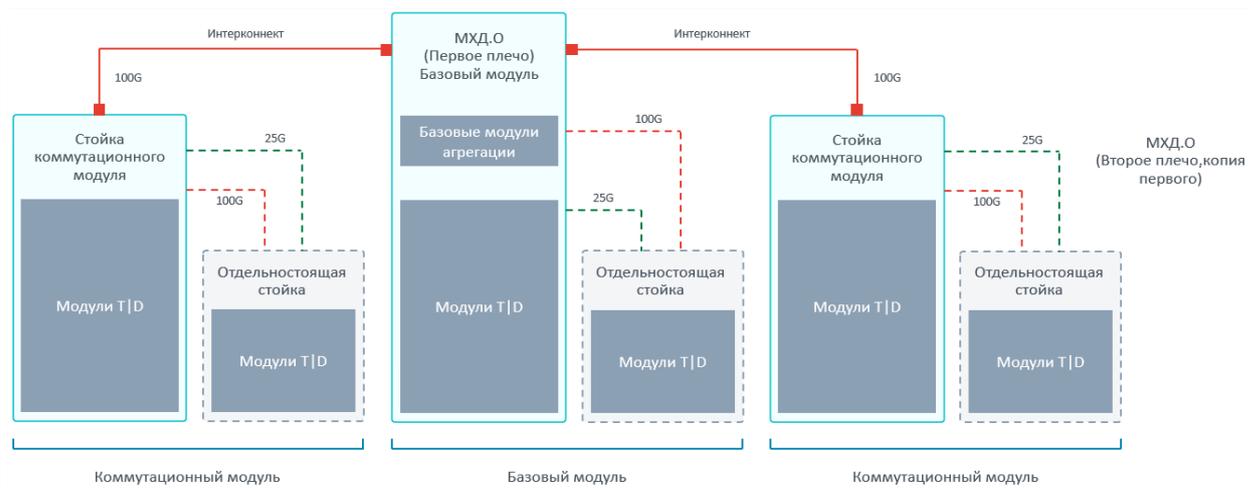


Рисунок 3. Состав Машины хранения данных Скала^р МХД.О

6.1.1 Базовый модуль

6.1.1.1 Общие сведения

Схема Базового модуля приведена на рисунке (Рисунок 4).

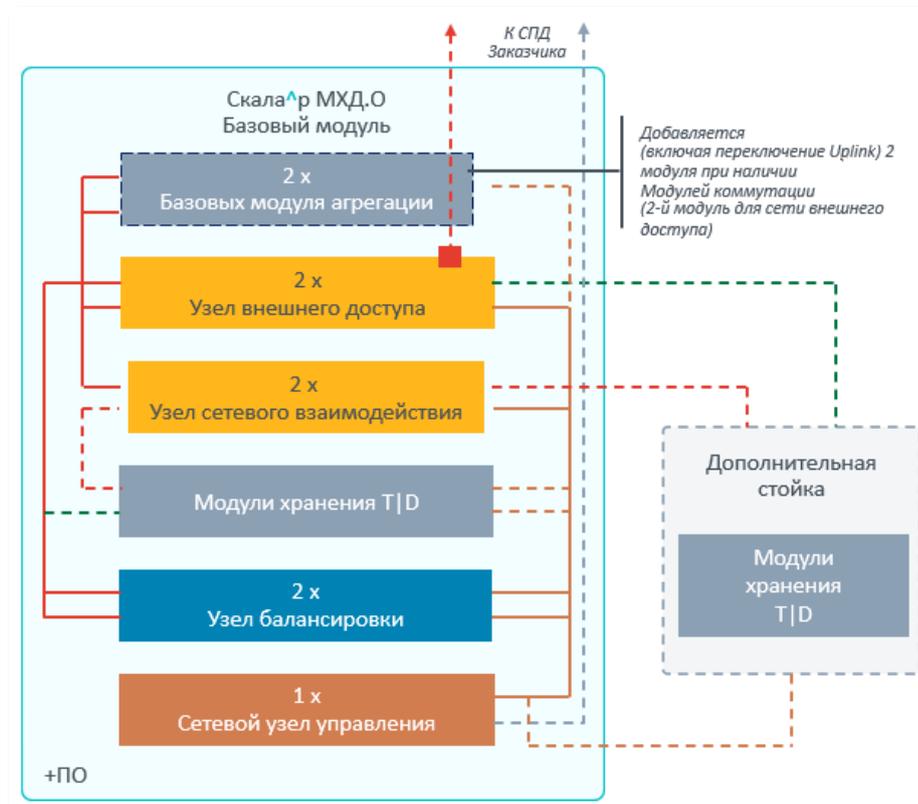


Рисунок 4. Схема Базового модуля

В состав Базового модуля входят следующие узлы:

- Узлы внешнего доступа — для сети пользователей **Машины хранения данных Скала^р МХД.О**
- Узлы сетевого взаимодействия — внутренний сегмент (интерконнект)
- Узлы балансировки — функции Спектр S3
- Сетевой узел управления — служебный коммутатор менеджмента

Базовый модуль может быть интегрирован со следующими Модулями:

- Модули хранения — комплекты узлов хранения **Машины хранения данных Скала^р МХД.О** для расширения объемов хранения
- Базовые модули агрегации — опция для расширения состава **Машины хранения данных Скала^р МХД.О**, средствами сетевого взаимодействия

Узлы внешнего доступа предназначены для обеспечения отказоустойчивого доступа потребителей сервиса хранения объектов к Модулям хранения. Подключение пользователей осуществляется через узлы балансировки, на которых установлено ПО **Спектр S3**.

Узлы сетевого взаимодействия реализуют отказоустойчивое подключение узлов хранения в единую сеть (интерконнект) для организации распределенного дискового массива, обеспечивающего хранение объектов.

Сетевой узел управления обеспечивает подключение всех аппаратных компонентов **Машины хранения данных Скала^р МХД.О** в сеть низкоуровневого мониторинга и управления.

Состав администраторов – для каждого тенанта реализуется роль администратора тенанта, которая может быть назначена одному или нескольким пользователям.

В состав Базового модуля включатся, как минимум, по одному модулю хранения Т и D (для типов модулей 1-3, 5, 6) или 3 модуля хранения Т (для типа модуля 4). Базовый модуль обеспечивает подключение от 5 до 26 узлов хранения без подключения Модулей коммутации и хранения.

В состав Базового модуля могут входить Модули агрегации, которые используются при расширении ресурсов **Машины хранения данных Скала^р МХД.О** за счет подключения дополнительных Модулей коммутации и хранения.

6.1.1.2 Характеристики

Базовый модуль поставляется в монтажной стойке (опция) высотой 42U и глубиной не менее 100 см. Характеристики Базового модуля приведены в таблице ниже (Таблица 4).

Таблица 4. Характеристики Базового модуля

| Наименование Модуля | Энергопотребление (кВт) | Количество свободных юнитов для монтажа Модулей хранения (U) |
|------------------------|-------------------------|--|
| Начальная конфигурация | 2,1 | 33 |

6.1.2 Модули хранения

6.1.2.1 Общие сведения

Модули хранения **Машины хранения данных Скала^р МХД.О** выполняют одновременно две функции:

- Организацию распределенного дискового массива для хранения объектов и метаданных
- Обеспечение доступа к объектам пользователей по протоколу, совместимому с Amazon S3

В составе **Машины хранения данных Скала^р МХД.О** могут поставляться:

- Модули хранения Т, включающие три узла хранения (Рисунок 5)
- Модули хранения D, включающие два узла хранения (Рисунок 6)

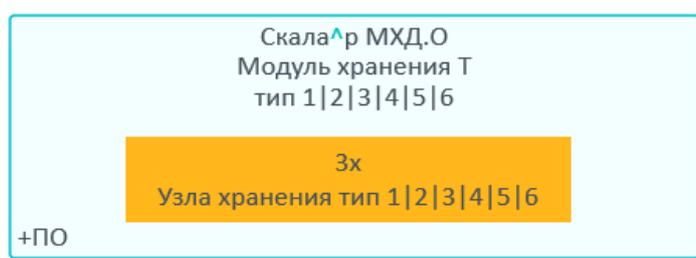


Рисунок 5. Модуль хранения Т

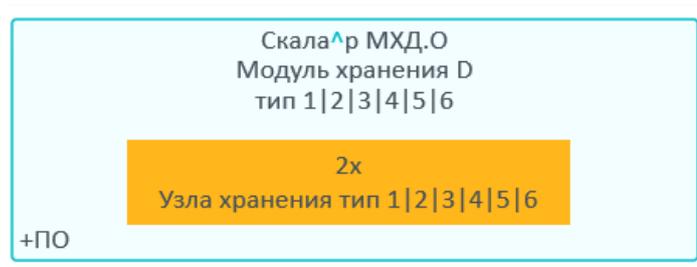


Рисунок 6. Модуль хранения D

Модули хранения представлены в 6-ти типах, предназначенных для реализации разных по требованиям к производительности и емкости хранения задач.

6.1.2.2 Характеристики

Модули хранения поставляются вместе с Базовым модулем или в составе Модуля коммутации и хранения, предназначенного для расширения емкости хранения Базового модуля. **Машина хранения данных Скала^р МХД.О** может иметь несколько типов Модулей хранения в своем составе, что позволяет организовать хранение объектов с разными классами обслуживания. Характеристики Модулей хранения приведены в таблице ниже (Таблица 5).

Таблица 5. Характеристики Модулей хранения

| Наименование Модуля | Тип накопителей | Полезная емкость (Тбайт) | Энергопотребление (кВт) | Требуемое место для монтажа (U) |
|-------------------------|-----------------|--------------------------|-------------------------|---------------------------------|
| Модуль хранения D тип 1 | SATA HDD | 77 | 1,2 | 4 |
| Модуль хранения T тип 1 | SATA HDD | 117 | 1,8 | 6 |
| Модуль хранения D тип 2 | SATA SSD | 17 | 1 | 4 |
| Модуль хранения T тип 2 | SATA SSD | 25 | 1,5 | 6 |
| Модуль хранения D тип 3 | SATA SSD | 34 | 1 | 4 |
| Модуль хранения T тип 3 | SATA SSD | 51 | 1,5 | 6 |
| Модуль хранения D тип 4 | SATA HDD | 1825 | 1,2 | 4 |
| Модуль хранения T тип 4 | SATA HDD | 274 | 1,8 | 6 |
| Модуль хранения D тип 5 | SATA SSD | 24 | 1 | 4 |
| Модуль хранения T тип 5 | SATA SSD | 37 | 1,5 | 6 |
| Модуль хранения D тип 6 | SATA SSD | 59 | 1 | 4 |
| Модуль хранения T тип 6 | SATA SSD | 74 | 1,5 | 6 |

Минимальное количество Модулей хранения в составе **Машины хранения данных Скала^р МХД.О**:

- Для типов 1-3, 5, 6: 1 модуль хранения D + 1 модуль хранения T
- Для типа 4: 3 модуля хранения T

Возможные режимы обеспечения сохранности данных в хранилище, определенные для Модулей хранения:

- Модули хранения D или T, тип 1-3: Метод: прямая репликация данных по формуле 3:2 (3 копии блоков данных хранятся строго на разных узлах)
- Модули хранения D или T, тип 4: защита избыточным кодированием Erasure Coding по формуле 5+2 (5 фрагментов данных и 2 блока четности, все на разных узлах)

6.1.3 Модули коммутации и хранения

6.1.3.1 Общие сведения

Модули коммутации и хранения используются для расширения емкости хранения **Машины Скала^р МХД.О** в случаях, когда портовая емкость Базового модуля не позволяет подключить дополнительные Модули хранения. Для использования добавляемых Модулей коммутации и хранения в Базовый модуль устанавливаются две пары коммутаторов Модулей агрегации, одна — для объединения узлов внешнего доступа, а вторая — для объединения узлов сетевого взаимодействия.

Схема Модуля коммутации и хранения представлена на рисунке (Рисунок 7).

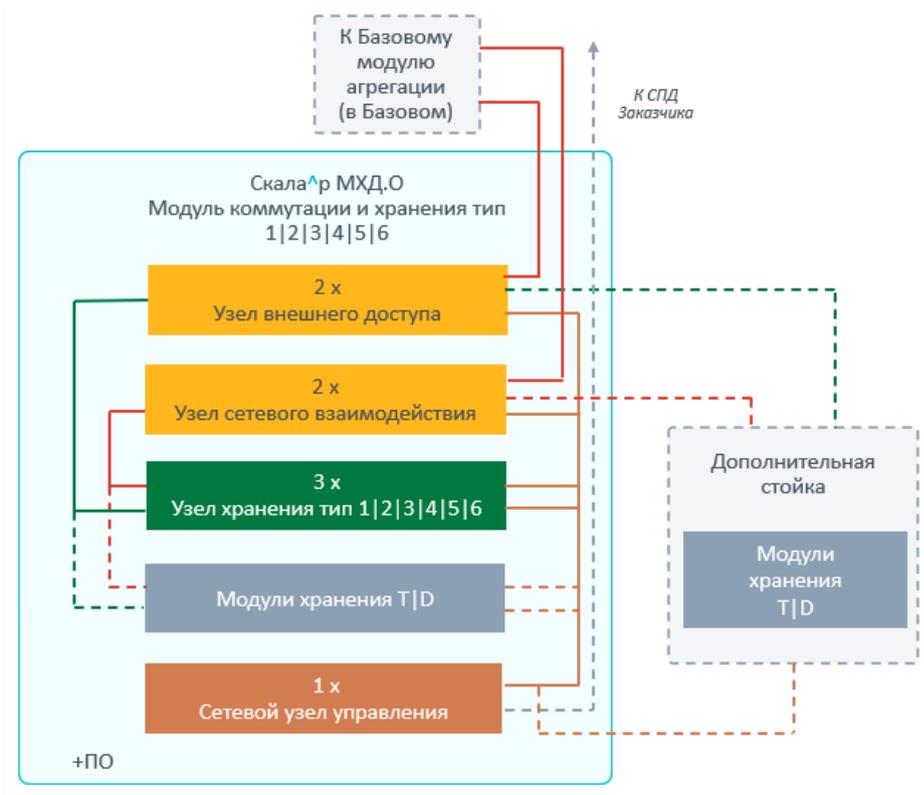


Рисунок 7. Схема Модуля коммутации и хранения

Модуль коммутации и хранения предназначен только для увеличения числа подключенных Модулей хранения. Модуль коммутации и хранения обеспечивает подключение до 7 модулей хранения T или до 11 модулей хранения D.

Модуль коммутации и хранения, по аналогии с Базовым модулем, включает сетевое оборудование для организации внешнего доступа, сети распределенного дискового массива, и сети управления, которые подключаются к аналогичным сетям Базового модуля через Модули агрегации.

В состав Модуля коммутации и хранения обязательно включено три узла хранения. Тип Модуля коммутации и хранения определяется типом установленных в него узлов хранения.

6.1.3.2 Характеристики

Модуль коммутации и хранения поставляется в монтажной стойке (опция) высотой 42U и глубиной не менее 100 см. Характеристики Модулей коммутации хранения приведены в таблице ниже (Таблица 6).

Таблица 6. Характеристики Модулей коммутации и хранения

| Наименование Модуля | Тип накопителей | Полезная емкость (Тбайт) | Энергопотребление (кВт) | Количество свободных юнитов для монтажа Модулей хранения (U) |
|------------------------------------|-----------------|--------------------------|-------------------------|--|
| Модуль коммутации и хранения тип 1 | SATA HDD | 117 | 2,3 | 31 |
| Модуль коммутации и хранения тип 2 | SATA SSD | 25 | 1,9 | |
| Модуль коммутации и хранения тип 3 | SATA SSD | 51 | 1,9 | |
| Модуль коммутации и хранения тип 4 | SATA HDD | 274 | 2,3 | |
| Модуль коммутации и хранения тип 5 | SATA SSD | 37 | 1,9 | |
| Модуль коммутации и хранения тип 6 | SATA SSD | 74 | 1,9 | |

Режимы обеспечения сохранности данных, определенные для Модулей коммутации и хранения повторяют принципы избыточности, описанные ранее для Модулей хранения:

- Модули коммутации хранения тип 1-3, 5, 6: 3 копии данных на разных узлах
- Модуль коммутации хранения тип 4: Защита избыточным кодированием Erasure Coding 7+2 (7 фрагментов данных 2 блока четности на разных)

6.1.4 Базовые модули агрегации

6.1.4.1 Общие сведения

Базовые модули агрегации состоят из отказоустойчивых пар коммутаторов и предназначены для подключения внешних и внутренних сетей Модулей коммутации и хранения к Базовому модулю. Пара Базовых модулей агрегации обеспечивает подключение до 6 Модулей коммутации и хранения, независимым образом объединяя узлы внешнего доступа, а также узлы сетевого взаимодействия (интерконнект) **Машины хранения данных Скала^р МХД.О.**

Схема Базового модуля агрегации представлена на рисунке (Рисунок 8).

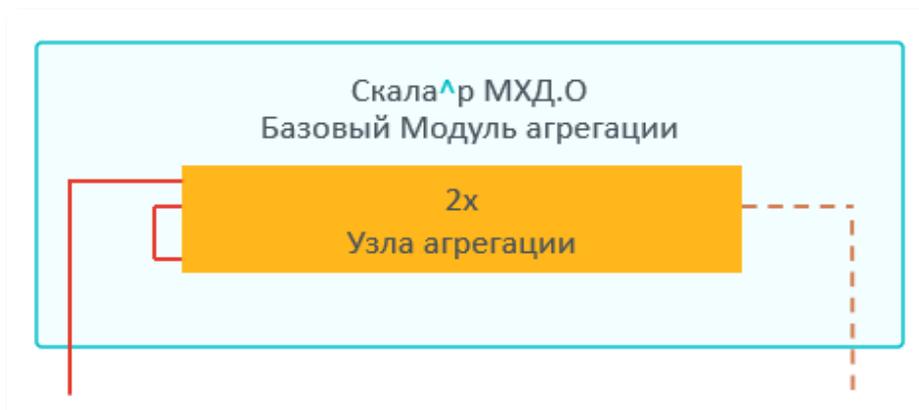


Рисунок 8. Схема Базового Модуля агрегации

6.1.4.2 Характеристики

Модули агрегации предназначены для установки в Базовый модуль при расширении **Машины** и устанавливаются парой. Характеристики одного Модуля агрегации приведены в таблице ниже (Таблица 7).

Таблица 7. Характеристики Модуля агрегации

| Наименование Модуля | Энергопотребление (кВт) | Количество занимаемых юнитов для монтажа (U) |
|---------------------|-------------------------|--|
| Модуль агрегации | 1,1 | 2 |

6.2 Программное обеспечение

В составе **Машины хранения данных Скала^р МХД.О** используется следующее программное обеспечение:

- Скала^р Спектр S3
- Скала^р Визион
- Скала^р Геном
- Распределенный дисковый массив

ПО **Скала^р Спектр S3** реализует протокол S3, доступ к данным, сжатие данных и балансировку запросов пользователей.

ПО **Скала^р Визион** предназначено для контроля состояния компонентов **Машины хранения данных Скала^р МХД.О**.

ПО **Скала^р Геном** обеспечивает обслуживание аппаратных и программных компонентов **Машины хранения данных Скала^р МХД.О**.

6.2.1 ПО Скала^р Спектр S3

Хранилище объектов

ПО **Скала^р Спектр S3** предоставляет API-интерфейс, подобный интерфейсу Amazon S3, который является одним из самых распространенных API-интерфейсов объектного хранилища. Конечные пользователи могут работать с **Машиной хранения данных Скала^р МХД.О** так же, как они работают с Amazon S3. Можно использовать привычные приложения для S3 и продолжать работу с ними после миграции данных из Amazon S3 на **Машину хранения данных Скала^р МХД.О**.

Хранилище объектов реализует интерфейс, который позволяет управлять данными в виде объектов (как в хранилище данных типа «ключ-значение»), в противоположность файлам в файловых системах или блокам в блочном хранилище. Каждый объект в хранилище содержит данные и метаданные, которые его описывают, а также уникальный идентификатор, позволяющий найти объект в хранилище. Хранилище объектных данных оптимизировано для хранения миллиардов объектов любого размера благодаря сочетанию очень высокой масштабируемости с высокой доступностью и согласованностью данных.

По сравнению с другими типами хранилищ, ключевое отличие хранилища объектов состоит в том, что части объекта нельзя изменить. При изменении объекта формируется его новая версия, а старая остается доступной, если иное не предусмотрено политиками хранения. Изменение объекта как единого целого устраняет проблему конфликтов, а объект с самой недавней меткой времени считается текущей версией.

6.2.2 ПО Скала^р Визион

Машины хранения данных Скала^р МХД.О поставляются со средствами мониторинга, настроенными на контроль характерных для **Машин Скала^р** параметров с заданными пороговыми значениями, позволяя из единой консоли проводить полноценный мониторинг всех компонентов **Машины хранения данных Скала^р МХД.О**.

Основные функциональные возможности:

- Мониторинг серверного оборудования
- Мониторинг сетевого оборудования
- Мониторинг программно-определяемого хранилища и сервисов объектного хранения
- Централизованное хранение и анализ лог-файлов
- Передача данных во внешние системы мониторинга

6.2.3 ПО Скала^р Геном

Скала^р Геном — решение, предназначенное для управления жизненным циклом **Машины хранения данных Скала^р МХД.О**.

Основные преимущества **Скала^р Геном**:

- Упрощает поддержку **Машин хранения данных Скала^р МХД.О**, установленных у Заказчика, включая вопросы обновления компонентов
- Снижает требования к квалификации эксплуатирующего **Машину** персонала

6.3 Распределенный дисковый массив

Распределенный дисковый массив создается на базе накопителей, установленных в узлы Модулей хранения, объединенных с использованием узлов сетевого взаимодействия. От функционирования узлов сетевого взаимодействия зависит работоспособность всего распределенного дискового массива, поэтому очень важно исключить несанкционированное изменение настроек этих узлов.

Для хранения данных в распределенном дисковом массиве используется две разных технологии:

- Создание копий блоков данных (реплики)
- Формирование контрольных сумм (технология кодирования Erasure Coding)

Создание копий данных предполагает запись одних и тех же данных на накопители, размещенные на разных узлах Модулей хранения. Такая технология используется на Модулях хранения Тип 1-3, 5, 6. В промышленной эксплуатации создается две дополнительных копии записанных данных, т.е. всего хранится три экземпляра одних и тех

же данных, что позволяет без потери данных пережить потерю одновременно до двух узлов Модулей хранения.

Формирование контрольных сумм предполагает деление записываемых данных на части, которые записываются на физически разные узлы Модулей хранения. В дополнение к записанным частям рассчитываются две контрольные суммы, которые также записываются на разные узлы Модулей хранения, не занятые частями записанных данных. Принцип очень похож на принцип работы RAID 6. Такая технология используется на Модулях хранения Тип 4. За счет создания двух контрольных сумм при использовании этой технологии отказ одновременно двух узлов Модулей хранения не приведет к потере данных.

Процессы восстановления копий при авариях — реализованы как фоновые, стартуют автоматически, и в общем случае не требуют вмешательства администраторов (кроме случаев обслуживания и случаев нехватки узлов или места на них).

Технология создания копий данных имеет более высокую производительность, однако обладают высокой избыточностью. Технология формирования контрольных сумм обладает невысокой избыточностью, но имеет заметно более низкую производительность на операциях записи.

6.4 Отказоустойчивость

Отказоустойчивость **Машины хранения данных Скала^р МХД.О** достигается набором технологий, реализованных на разных уровнях программного и аппаратного обеспечения:

- Применение отказоустойчивого распределенного дискового хранилища, не имеющего единой точки отказа
- Технология высокой доступности для служб сервиса имен (NS) и сервиса хранения (OS). Стоит также отметить, что сервис шлюза (GW) не хранит собственных данных и перезапускается автоматически при сбое платформы
- Наличие узлов шлюзов S3, организованных в кластер высокой доступности

Отказоустойчивость хранения данных обеспечивается за счет хранения нескольких копий данных на разных узлах Модулей хранения или применения технологии Erasure Coding с распределением частей данных и контрольных сумм данных по разным Модулям хранения. Использование любой из этих технологий гарантирует сохранность данных при одновременном отказе до двух узлов Модулей хранения в типовых решениях **Машины хранения данных Скала^р МХД.О**.

На уровне служб хранилища S3 используется механизм автоматического перезапуска чувствительных к отказам служб NS и OS. В случае выхода из строя узла хранения, на котором они функционируют, службы автоматически перезапускаются на других узлах хранения. Служба GW не хранит собственных данных и автоматически перезапускается при сбоях узлов хранения.

Для исключения возможности перенаправления запроса пользователя к узлу хранения, вышедшему из строя, в состав **Машины хранения данных Скала^р МХД.О** включены узлы шлюзов S3, которые осуществляют постоянный контроль работоспособности узлов хранения и распределяют запросы пользователей только между работоспособными узлами хранения.

Узлы шлюзов S3 организованы в кластер высокой доступности.

7. МОДЕРНИЗАЦИЯ И ОБСЛУЖИВАНИЕ

Машину хранения данных Скала^р МХД.О легко обслуживать и модернизировать. Основные возможности:

- Замена вышедшего из строя накопителя без остановки узла и компонентов сервиса S3
- Удаление нормально функционирующего накопителя (например, для замены на другой) без остановки узла и компонентов сервиса S3
- Замена вышедшего из строя узла хранения или узла балансировки без остановки **Машины Скала^р МХД.О**
- Добавление дополнительного Модуля хранения или Модуля коммутации и хранения (для увеличения емкости хранения) без остановки **Машины Скала^р МХД.О**
- Временный вывод из состава **Машины хранения данных Скала^р МХД.О** любого узла Модулей хранения для его модернизации или профилактических и регламентных работ и возвращение его в активный пул без остановки **Машины**.

8. СЦЕНАРИИ ИСПОЛЬЗОВАНИЯ

Благодаря своей универсальности, S3 хранилище на базе **Машины хранения данных Скала^р МХД.О** может использоваться для хранения файлов (объектов) самого разного типа, включая:

- Объекты, которые ранее хранились в СУБД
- Резервное копирование (полностью совместимо с pg_probackup)
- Разнообразные документы и их образы
- Большие данные (big data)
- Архивы и резервные копии
- Статичные данные веб-сайтов
- Фотографии и видеозаписи

То есть S3 хранилище на базе **Машины хранения данных Скала^р МХД.О** может использоваться практически в любом технологическом процессе, приложении или сервисе, где требуется хранение большого количества файлов или бинарных объектов.

S3 хранилища дают возможность повысить производительность решений для управления данными за счет выноса функции хранения больших объектов во внешнюю систему. Например, при использовании СУБД Postgres, хранение большого числа объектов в самой базе увеличивает ее размер и существенно снижает производительность. Хранение объектов за пределами СУБД Postgres в S3 хранилище существенно уменьшает объем базы и повышает ее производительность, при этом доступ к объектам осуществляется по статичным ссылкам, хранимым в базе.

В связи с необходимостью перехода на технологии, свободные от санкционных рисков, **Машина хранения данных Скала^р МХД.О**, как российский продукт, является оптимальным решением в области импортозамещения, особенно в связке с **Машинами баз данных Скала^р МБД.П**.

Существует две типовые схемы использования **Машины хранения данных Скала^р МХД.О**:

- Локальное хранилище S3
- Распределенное хранилище S3 с асинхронной георепликацией.

8.1 Локальное хранилище S3

Локальное хранилище S3 является стандартным вариантом использования комплекса. К основным достоинствам относятся:

- Возможность создания хранилищ объемом до 8 Пбайт
- Возможность построения высокопроизводительного комплекса со скоростью чтения/записи в десятки Гбит/с
- Высокая устойчивость к единичным отказам оборудования, простота восстановления после сбоя

Машина хранения данных Скала^р МХД.О обеспечивает возможность использования различных Модулей хранения, которые могут быть использованы для организации классов хранения объектов S3 хранилища с использованием атрибута запроса:

```
x-amz-storage-class: [StorageClass]
```

Это позволяет создавать решения, которые обеспечивают оптимальное хранение данных с разными уровнями востребованности или реализуют иерархический принцип хранения объектов (Рисунок 9).

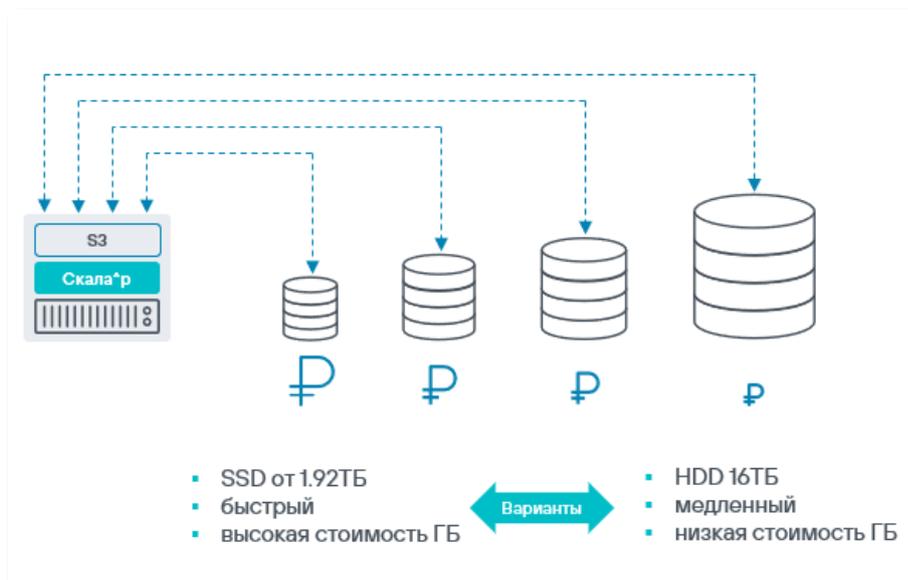


Рисунок 9. Различные уровни хранения объектов

Для создания разных уровней используются разные типы накопителей и разные технологии хранения данных, которые позволяют создать хранилища с практически плавной градацией соотношения производительность/емкость хранения.

8.2 Распределенное хранилище S3 с асинхронной георепликацией

Несколько комплексов **Скала^р МХД.О** могут быть объединены в единую систему, где есть головное и резервные хранилища S3. К основным достоинствам такого решения относятся:

- Катастрофоустойчивость
- Относительно невысокие требования к каналу связи между головным и резервными хранилищами

Схема предоставления сервиса S3 распределенной системой **Машины хранения данных Скала^р МХД.О** конечным пользователям может быть реализована в режимах active-passive или active-active. В последнем случае пользователи смогут одновременно изменять объекты на разных площадках и последней версией будут считаться объекты с более поздней меткой времени создания.

9. ПЛАНИРОВАНИЕ ИНФРАСТРУКТУРЫ

Для выбора необходимой конфигурации **Машины** можно использовать следующие характеристики по предельной емкости и производительности **Машины хранения данных Скала^р МХД.О**, задаваемые использованием Модулей хранения разного типа, представленные в таблице ниже (Таблица 8).

Таблица 8. Характеристики по предельной емкости и производительности **Машины хранения данных Скала^р МХД.О**

| Тип | Тип накопителей | Достижимая емкость на одну Машину хранения данных (Тбайт) | Производительность, на Тбайт хранения при чтении* объектов 40 Кбайт | Производительность, на Тбайт хранения при чтении* объектов 5 Мбайт |
|-----|-----------------|---|---|--|
| 1 | SATA HDD | 5850 | 14 оп/с | 4 оп/с |
| 2 | SATA SSD | 1250 | 61 оп/с | 16 оп/с |
| 3 | SATA SSD | 2550 | 30 оп/с | 8 оп/с |
| 4 | SATA HDD | 7998 | 5 оп/с | 2 оп/с |
| 5 | SATA SSD | 1850 | 59 оп/с | 14 оп/с |
| 6 | SATA SSD | 3700 | 29 оп/с | 7 оп/с |

* Производительность при записи ниже в 1,4 раза.

Типы 1 и 4 рекомендуется использовать для задач резервного копирования и «холодного» хранения данных.

Для интеграции **Машины хранения данных Скала^р МХД.О** в существующую ИТ-инфраструктуру необходимо выполнить следующие шаги:

- Выделить внутренние IP-адреса для сети распределенного дискового хранилища
- Выделить IP-адреса для сети доступа **Машины хранения данных Скала^р МХД.О**
- Выделить IP-адреса для сети управления
- Предоставить доступ узлам **Машины хранения данных Скала^р МХД.О** к сервису времени NTP
- Сформировать имя для сервиса S3
- Сформировать сертификат с использованием удостоверяющего центра или выдать самоподписанный сертификат для защиты трафика S3 (подходит только для локального варианта использования)

10. ГРАНИЦЫ ПРИМЕНИМОСТИ

Применение **Машины хранения данных Скала^р МХД.О** (Рисунок 10) рекомендовано при выполнении одного или нескольких следующих условий:

- Необходимость организации сервиса S3 в частном облаке
- Реализация высокопроизводительного хранилища объектов для разгрузки СУБД (вынос BLOB на внешнее хранилище)
- Замена файловых сервисов в информационных системах для повышения безопасности и надежности хранения (за счет использования технологии ключ-значения)
- Хранение резервных копий, вместо дисковых массивов и ленточных накопителей
- Реализация распределенных хранилищ с одновременным доступом на разных площадках (катастрофоустойчивое хранение документов/резервных копий), без ограничения расстояния.

Машину хранения данных Скала^р МХД.О целесообразно применять при объеме хранения данных от 40 Тбайт до 8 Пбайт.



Рисунок 10. Границы применимости Машины хранения данных Скала^р МХД.О

11. ТЕХНИЧЕСКАЯ ПОДДЕРЖКА

Поставка **Машины хранения данных Скала^р МХД.О** осуществляется с предварительными сборкой, тестированием и настройкой оборудования согласно требованиям Заказчика. Качественная поддержка **Машины хранения данных Скала^р МХД.О** обеспечивается едиными стандартами гарантийного и постгарантийного технического обслуживания:

- Пакет услуг по технической поддержке на первый год включен в поставку
- Заказчик может выбирать пакет в базовом режиме 9×5, или в расширенном режиме 24×7 (опция для 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 представлен в таблице ниже (Таблица 9).

Таблица 9. Пакеты услуг по технической поддержке Машины хранения данных Скала^р МХД.О

| Услуги | Пакет «9×5» | Пакет «24×7» |
|--|-------------|--------------|
| «Обслуживание комплекса Скала^р в режиме 9×5» (в рабочее время по рабочим дням) | + | — |
| «Обслуживание комплекса Скала^р в режиме 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: <ul style="list-style-type: none"> ▪ 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 | + | + |
| Доставка оборудования на замену за счет производителя | + | + |

| Услуги | Пакет «9×5» | Пакет «24×7» |
|--|---------------------------------|--------------------------------|
| Расширенные опции 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

Машина хранения данных Скала^р МХД.О 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 **Машине хранения данных Скала^р МХД.О** предусмотрено, как минимум, двойное резервирование основных компонентов, позволяющее сохранять данные и работоспособность даже при выходе из строя нескольких дисков и/или узлов.

12. ПОСТАВКА И ЛИЦЕНЗИРОВАНИЕ ПО

Команда **Скала^р** активно занимается развитием программных продуктов **Машины хранения данных Скала^р МХД.О**. Направления развития формируются на основе анализа мирового опыта использования систем подобного класса и пожеланий Заказчиков и партнеров. Новые функции реализуются в форме мажорных и минорных релизов: мажорные релизы выпускаются раз в квартал, минорные релизы выпускаются при необходимости более быстрого введения в эксплуатацию небольших улучшений в системе.

Программное обеспечение **Скала^р Спектр S3**, **Скала^р Визион**, **Скала^р Геном** поставляется исключительно в составе **Машины хранения данных Скала^р МХД.О** и лицензируется по метрикам комплекса в соответствии с его размером.

Лицензирование

В **Машине хранения данных Скала^р МХД.О** поставляются следующие лицензии:

- Лицензии ПО **Скала^р Спектр S3**
- Лицензии ПО **Скала^р Визион** и **Скала^р Геном**

Условия лицензирования можно уточнить по запросу через наших партнеров.

О КОМПАНИИ

Скала^р (продукт **Группы Rubytech**) с 2015 года выпускает программно-аппаратные комплексы (**Машины**) и сегодня предлагает 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. Соответствуют критериям доверенности и использованию для объектов критической информационной инфраструктуры (КИИ).

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