скала^р

Динамическая инфраструктура

Машина серверной виртуализации Скала^р МДИ.В

Масштабируемый и отказоустойчивый ПАК
для серверной виртуализации

Технический обзор

версия 1.95 от 23.03.2026

skala-r.ru info@skala-r.ru

Оглавление

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 изготовленных ПАК могут отличаться от приведенных вследствие модификации изделий.

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

В документе используются обозначения программных продуктов компании Базис в нотации лицензионных именований и записей в реестре ПО Минцифры РФ. Они могут отличаться от маркетинговых именований, например, продукт Basis Virtual Security (как указано на сайте производителя) имеет лицензионное написание Базис.Virtual Security, что является синонимом.

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

Перечень терминов и сокращений

Термин, сокращение	Определение
MLAG	(англ. Multi-Switch Link Aggregation) Технология агрегации каналов, позволяющая одному или нескольким линкам с двух разных сетевых узлов быть объединенными вместе таким образом, что для конечного устройства это выглядит как одиночное соединение
NFS	(англ. Network File System) Протокол сетевого доступа к файловым системам
OCFS2	(англ. Oracle Cluster File System 2) Разделяемая журналируемая кластерная дисковая файловая система, позволяющая нескольким узлам одновременно читать и писать на общее хранилище
RAID	(англ. Redundant Array of Independent Disks) Избыточный массив независимых дисков, технология виртуализации данных для объединения нескольких физических дисковых устройств в логический модуль для повышения отказоустойчивости и/или производительности
SDS	(англ. Software Defined Storage) Программно-определяемое хранилище, типично собранное из локальных дисков нескольких унифицированных серверных узлов, объединенных производительной сетью
SSD	(англ. Solid-State Drive) Запоминающее устройство на основе микросхем памяти
ЗОКИИ	Значимый объект критической информационной инфраструктуры
ИСПДн	Информационные системы персональных данных. Совокупность информации, содержащейся в базах данных, и обеспечивающих ее обработку с использованием информационных технологий и технических средств
Кластер	Отказоустойчивая архитектура функционала Машины
Машина	Автономный масштабируемый модульный программно-аппаратный комплекс (ПАК, изделие с кодом ОКПД 26.20.14.160 из реестра радиоэлектронной продукции Минпромторга РФ), решающий функциональную задачу хранения, обработки и передачи данных согласно предустановленному системно-прикладному ПО и предоставляющий необходимые для задачи ресурсы вычислений и хранения

Термин, сокращение	Определение
Модуль	Функционально завершенный комплект сконфигурированных аппаратных узлов и программного обеспечения (ПО) для выполнения заданных функций в составе Машины или автономно, оформленный как самостоятельная единица продаж со своим артикулом и стоимостью. Является единым и неделимым элементом спецификации. Зарегистрирован как ПАК (с кодом ОКПД 26.20.14.160) в реестрах Минпромторга РФ
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
Узел	Вычислительный узел (сервер вычисления и\или хранения) или сетевой узел (например, коммутатор) или узел хранения (СХД) в составе Модуля, в зависимости от контекста

1. Предисловие

1.1 Описание документа

Настоящий документ дает концептуальный и архитектурный обзоры ПАК **Машина серверной виртуализации Скала^р МДИ.В** из семейства **Машин динамической инфраструктуры Скала^р**.

До сентября 2025 года ПАК поставлялся под названием **Машина виртуализации Скала^р МВ.ДИ**, исполнение «Стандарт».

Далее в документе применяются как полное, так и сокращенные названия ПАК: **Машина Скала^р МДИ.В, Машина**.

1.2 Аудитория

Документ предназначен для партнеров **Скала^р** и заказчиков, перед которыми ставятся задачи разработки, закупки, управления или эксплуатации **Машины серверной виртуализации Скала^р МДИ.В**.

1.3 Обратная связь

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

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

2. Введение

Машина серверной виртуализации Скала^р МДИ.В является программно-аппаратным комплексом, предназначенным для создания горизонтально масштабируемой и отказоустойчивой среды серверной виртуализации, а также инфраструктуры для виртуальных рабочих мест пользователей.

В состав **Машины серверной виртуализации Скала^р МДИ.В** входят унифицированные сетевые и вычислительные узлы. В качестве подсистемы хранения данных в **Машине** используется гиперконвергентное программно-определяемое распределенное хранилище, т.е. вычислительные узлы в **Машине** выполняют функцию в том числе и хранения данных. В качестве опции **Машина серверной виртуализации Скала^р МДИ.В** может быть также укомплектована конвергентным программно-определяемым распределенным хранилищем с блочным доступом по протоколу iSCSI и выделенными (отделёнными от роли вычислительных узлов) серверными узлами хранения, либо классическими аппаратными СХД с блочным доступом (также по протоколу iSCSI) с последующим развертыванием на их базе отказоустойчивых вычислительных кластеров.

Конструктивно **Скала^р МДИ.В** собирается из служебных и функциональных **Модулей Скала^р**, каждый их который, как и **Машина**, является самостоятельным изделием (ПАК) с записью в реестре Минпромторга.

Масштабирование **Машины** предусмотрено на уровне функциональных модулей, содержащих один или несколько узлов вычислений и/или хранения.

Машина серверной виртуализации Скала^р МДИ.В обеспечивает следующие преимущества:

- гарантия совместимости — **Скала^р** проводит R&D работы по выбору и валидации оборудования вместе с используемым программным обеспечением. Вычислительные узлы, контроллеры, накопители, сетевые узлы, программное обеспечение виртуализации, управления тестируются в соответствии с разработанными архитектурой и технологическими картами ПАК именно в той комбинации, которая будет поставляться заказчикам;
- **Машина** поставляется с предварительно установленным системным, служебным и функциональным ПО и готова к эксплуатации немедленно после ее монтажа и подключения у заказчика;
- серийное производство позволяет создать легко тиражируемое решение с регламентными сроками доступности к поставке. Оборудование из состава Машины является не только технически валидированным, но и подтвержденным OEM-соглашениями с ведущими поставщиками узлов и компонентов;
- собственная сервисная служба позволяет обеспечить техническую поддержку «из одного окна» по всем составляющим комплекса. Инциденты, даже на стыке аппаратных и программных компонентов, будут разбираться службой поддержки **Скала^р**, специалисты которой, при необходимости, уже сами обратятся в поддержку производителя конкретного компонента;
- получение новых функций с обновлениями системы. Функциональность реализована не в аппаратном, а в программном обеспечении, поэтому добавление новых функций не будет ограничено тем, что «приобретен комплекс предыдущего поколения»;
- быстрая адаптация к требованиям бизнеса. **Машина серверной виртуализации Скала^р МДИ.В** базируется на типовом серверном и сетевом оборудовании с широкой доступностью на российском рынке, с возможностью гибкого подбора его конфигураций;

- надежная производительная работа комплекса при больших нагрузках. Конвергентная программно-определяемая система хранения данных поддерживает масштабирование производительности как вертикально (выбором более быстрых накопителей и интерфейсов их подключения), так и горизонтально — выбором большего числа накопителей/узлов хранения. Кроме того, такая система состоит из большого числа равнозначных узлов хранения и архитектурно не имеет единственного компонента, который мог бы стать узким местом или единой точкой отказа всего хранилища;
- возможность вывести из эксплуатации любой элемент (или даже несколько) системы: вычислительный узел или узел хранения без существенного влияния на общую работоспособность и производительность системы;
- нечувствительность к единичным отказам;
- более низкие требования к ширине экспертизы специалистов, эксплуатирующих решение, за счет меньшей номенклатуры компонентов;
- возможность гибкого и почти мгновенного перераспределения ресурсов между задачами;
- удобное централизованное управление вычислительными ресурсами из единого интерфейса с назначением приоритетов и правил общего доступа;
- возможность переноса работающих виртуальных машин между вычислительными узлами без остановки;
- все функции сервера управления средствами виртуализации доступны через программный интерфейс (API), что дает возможность реализации полной автоматизации операций над инфраструктурой, и, как следствие, возможность встраивания в частные и гибридные облачные решения;
- механизмы сетевой изоляции, непрерывности и высокой доступности для запущенных виртуальных машин, автоматической балансировки вычислительной нагрузки в рамках одного вычислительного кластера;
- мониторинг текущего состояния и утилизации ресурсов как аппаратных, так и программных компонентов **Машины**, с возможностью настройки длительности хранения собираемых метрик и их пороговых значений и параметров оповещения;
- управление полным жизненным циклом **Машины**: установка и обновление всех программных компонентов в составе **Машины**, автоматическое формирование отчетов о версиях установленного ПО.

3. Отличительные черты

Главными отличительными чертами **Машины серверной виртуализации Скала^р МДИ.В** являются поставка программно-аппаратного комплекса «под ключ» и ориентация на высокую надежность и высокую производительность на уровне архитектуры:

- как гипервизор, так и программно-определяемая система хранения данных создавались и развиваются с упором на надежность и безотказную работу;
- резервирование данных на хранилище гибко настраивается под задачу: для особо критичных виртуальных машин можно повысить количество копий при резервировании данных с рекомендуемых «трех копий» до требуемого значения. При деградации хранилища ниже определенного заданного числа копий, хранилище автоматически перейдет в режим «только чтение», чтобы не допустить возможные полные потери данных или переход в неконсистентное состояние. В случае защиты данных помехоустойчивым кодированием (Erasure Coding) можно также выбрать оптимальные параметры кода Рида-Соломона, отдавая предпочтение или более высокой производительности или уменьшению накладных расходов полезного объема хранилища;
- для всех виртуальных машин доступен кластер высокой доступности, минимизирующий время простоя при отказах вычислительных узлов.

Также, все комплектующие и составные части оборудования **Машины** проходят всестороннее тестирование и валидацию еще на этапе R&D, затем на этапе входного тестирования на производстве, и на этапе выходного тестирования готового ПАК. При этом, конфигурация **Машины** предусматривает резервирование всех потенциальных единых точек отказа. Это не только вычислительные узлы и накопители, но и сетевые контроллеры и сетевые узлы в составе **Машины**.

Последнее по порядку, но не по значению — служба технической поддержки «из одного окна».

Подробный перечень отличительных свойств по категориям приводится ниже.

Высокая доступность и производительность

- защита от отказа единичных узлов и аппаратных компонентов, установленных в вычислительные узлы (жесткие диски, блоки питания и т.п.), обеспечивается их дублированием и схемами программной защиты данных;
- высокоскоростные сети внутренней и внешней связности;
- архитектурная оптимизация производительности;
- специальные настройки программного обеспечения;
- встроенная система мониторинга и анализа состояния **Машины** собственной разработки **Скала^р**.

Отказоустойчивость на всех уровнях

- надежные комплектующие;
- входное тестирование всех компонентов на совместимость и корректность функционирования;
- резервирование значимых компонентов на аппаратном уровне;
- отказоустойчивая архитектура;
- оперативное восстановление при сбоях.

Гибкая система управления

- централизованное управление из единого web-интерфейса;
- Расширения функций опциональным интегрированным ПО.

Линейная масштабируемость

- компоненты **Машины** подобраны и сбалансированы для раскрытия всего потенциала масштабируемости;
- наращивание количества вычислительных узлов и/или узлов хранения обеспечивает максимальную производительность с сохранением экономической эффективности и надлежащего уровня эксплуатационного качества.

Безопасность на всех уровнях

- комплекс сертифицированных средств защиты от различных производителей;
- предварительный анализ защищенности и анализ уязвимостей релизов **Машин**, для своевременно обнаружения слабых мест в защите и предотвращения возможных атак.

Обеспечение качества при развертывании

- оптимальность настроек подтверждена значительным количеством установок;
- автоматизированное развертывание ПАК на производстве **Скала^р** снижает риск человеческой ошибки;
- стандартизация развертывания гарантирует соответствие продукта заявленным характеристикам.

Непрерывный контроль состояния Машины

- мониторинг работоспособности ПО и оборудования;
- установленные пороговые значения критичных параметров;
- различные каналы информирования системой мониторинга об отклонениях.

Гибкие возможности администрирования

- проработанные рекомендации по выполнению процедур обслуживания;
- предустановлено дополнительное ПО для управления.

Поддержка в эксплуатации

- централизованная техническая поддержка ПАК;
- выпуск предварительно проверяемых патчей;
- обучение персонала заказчика.

Экономическая эффективность

- специальные условия лицензирования;
- сокращенные сроки ввода в эксплуатацию;
- только обоснованно необходимые компоненты.

Альтернатива VMware vSphere, Microsoft Hyper-V, Citrix Hypervisor

- полностью российское решение;
- отлаженные инструменты и процессы миграции;
- высокие надежность и производительность;
- качество, подтвержденное опытом практического применения.

4. Подтвержденная безопасность

Машина серверной виртуализации Скала^р МДИ.В поставляется с сертифицированной **ОС Альт 8 СП** (сертификат ФСТЭК 3866 от 10.08.2018, действует до 10.08.2028). ОС используется в качестве гостевой в служебных виртуальных машинах комплекса или основной для выделенных служебных вычислительных узлов.

ОС может применяться для защиты информации:

- в значимых объектах критической информационной инфраструктуры 1 категории, в государственных информационных системах 1 класса защищенности;
- в автоматизированных системах управления производственными и технологическими процессами 1 класса защищенности;
- в информационных системах персональных данных при необходимости обеспечения 1 уровня защищенности персональных данных;
- в 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, приказ № 118) по 4 классу защиты;
- «Требования по безопасности информации к средствам виртуализации» (ФСТЭК России, 2022, приказ № 187) по 4 классу защиты;
- «Требования по безопасности информации к системам управления базами данных» (ФСТЭК России, 2023) 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, приказ № 76) по 4 уровню доверия.

Машина серверной виртуализации Скала^р МДИ.В опционально поставляется с сертифицированным средством защиты виртуальной инфраструктуры Базис.Virtual Security (сертификат ФСТЭК 4348 от 24.12.2020, действует до 24.12.2030). Применяется Исполнение 3 продукта BVS согласно Формуляру изделия.

Базис.Virtual Security применяется для защиты информации:

- в государственных информационных системах до 1 класса защищенности включительно;
- в информационных системах персональных данных при необходимости обеспечения 1 уровня защищенности персональных данных;
- значимых объектах критической информационной инфраструктуры 1-ой категории значимости;
- в информационных системах общего пользования 2 класса.

Базис.Virtual Security соответствует требованиям следующих нормативных документов:

- «Требования о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах», утвержденным приказом ФСТЭК России от 11.02.2013 г. №17 с изменениями, внесенными приказом ФСТЭК России от 15.02.2017 г. № 27 и приказом ФСТЭК России от 28.05.2019 г. №106;
- «Состав и содержание организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных» введенным в действие приказом ФСТЭК России №21 от 18.02.2013 г. с изменениями, внесенными приказом ФСТЭК России от 23.03.2017 г. №49 и приказом ФСТЭК России от 14.05.2020 г. №68;
- «Требования по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации», введенным в действие приказом ФСТЭК России № 239 от 25.12.2017 г. с изменениями, внесенными приказом ФСТЭК России от 09.08.2018 г. №138, приказом ФСТЭК России от 26.03.2019 г. №60 и приказом ФСТЭК России от 20.02.2020 г. №35;
- «Требования к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды», утвержденным приказом ФСТЭК России от 14.03.2014 г. №31 с изменениями, внесенными приказом ФСТЭК России от 23.03.2017 г. №49, приказом ФСТЭК России от 09.08.2018 г. №138 и приказом ФСТЭК России от 15.03.2021 г. №46;
- «Требования по защите информации, содержащейся в информационных системах общего пользования», введенным в действие приказом ФСТЭК России №489 от 31.08.2010 г.

Протестирована совместимость с наложенными средствами защиты

Машина серверной виртуализации Скала^р МДИ.В совместима и может включать в себя при поставке сертифицированное средство единой аутентификации **«**Аванпост FAM**»** (сертификат ФСТЭК 4492 от 13.12.2021 до 13.12.2026), которое может быть применено в следующих информационных системах, не предназначенных для обработки сведений государственной тайны:

- для применения в государственных информационных системах 1 класса защищенности;
- для применения в информационных системах персональных данных при необходимости обеспечения 1 уровня защищенности персональных данных;
- для применения в автоматизированных системах управления производственными и технологическими процессами 1 класса защищенности.

«Аванпост FAM» 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, приказ № 76) по 1 уровню доверия.

Машина серверной виртуализации Скала^р МДИ.В совместима и может использовать антивирусное средство защиты **«Kaspersky Security для виртуальных сред 5.2 Легкий агент»** (сертификат ФСТЭК 3883 от 14.02.2018, действует до 14.02.2026), которое соответствует следующим документам:

- «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;
- «Требования к средствам антивирусной защиты» (ФСТЭК России, 2012);
- «Профиль защиты средств антивирусной защиты типа Б четвертого класса защиты. ИТ.САВ3.Б4.ПЗ» (ФСТЭК России, 2012);
- «Профиль защиты средств антивирусной защиты типа В четвертого класса защиты. ИТ.САВ3.В4.ПЗ» (ФСТЭК России, 2012).

Указанное ПО не входит в поставку **Машины**.

Также совместимо с **Машиной** и может быть установлено (но не входит в поставку **Машины**) Сертифицированное антивирусное средство защиты **Kaspersky Endpoint Security для Linux** (сертификат ФСТЭК 2534 от 27.12.2011, действует до 27.12.2030) соответствует документам:

- «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.САВ3.Б2.13» (ФСТЭК России, 2012);
- «Профиль защиты средств антивирусной защиты типа В второго класса защиты. ИТ.САВ3.В2.ПЗ» (ФСТЭК России, 2012);
- «Профиль защиты средств антивирусной защиты типа Г второго класса защиты».

Машина серверной виртуализации Скала^р МДИ.В совместима и может включать в себя при поставке сертифицированное средство доверенной загрузки **ПАК «Соболь»** версия 4 (сертификат ФСТЭК №4043 от 05.12.2018, действует до 05.12.2028), которое может быть применено для защиты информации:

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

ПАК «Соболь» 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 платы расширения второго класса защиты. ИТ.СДЗ.ПР2.ПЗ» (ФСТЭК России, 2013).

5. Сценарии использования

Машина серверной виртуализации Скала^р МДИ.В является универсальным модульным компонентом для построения инфраструктурной платформы и может применяться в разнообразных сценариях и их комбинациях. Несколько примеров приведены ниже.

Во всех сценариях важное значение имеют такие преимущества **Машины серверной виртуализации Скала^р МДИ.В**, как подтвержденная совместимость программных и аппаратных компонентов, быстрый и простой ввод в эксплуатацию за счет поставки полностью предустановленного решения, единая техническая поддержка и доступ к новым функциям путем обновления программной части **Машины**.

Универсальная серверная инфраструктура

Машина серверной виртуализации Скала^р МДИ.В служит надежным фундаментом для инфраструктуры серверной виртуализации общего назначения.

От четырех до нескольких десятков вычислительных узлов, объединенных в один или несколько отказоустойчивых вычислительных кластеров, с управлением из одного окна единой системы управления. Оптимальное количество вычислительных узлов в одном кластере составляет 20-28 узлов.

В зависимости от требований это может быть как единая **Машина**, так и несколько несвязанных ПАК (например, обслуживающих организационно разные задачи под управлением различных отделов).

Доверенные ПАК

Ранее компании использовали технологии и ПО преимущественно иностранного происхождения ввиду их широкой доступности и технологической зрелости, но в сложившейся ситуации они уже не отвечают требованиям государственной безопасности и надежности ввиду рисков несанкционированного доступа к данным, отсутствия гарантий бесперебойности в работе и из-за использования в качестве инструмента для ослабления критически важных отраслей государства.

В ноябре 2023 года вышло постановление правительства РФ №1912, обязывающее перевести все значимые объекты критической информационной инфраструктуры (ЗОКИИ) на доверенные программно-аппаратные комплексы до 31 декабря 2029 года. С 1 сентября 2024 года субъекты КИИ не могут использовать недоверенные ПАК, приобретенные после этой даты.

Для признания ПАК доверенным необходимо выполнить ряд требований:

- ПАК находится в реестре российской промышленной и радиоэлектронной продукции Минпромторга РФ;
- программное обеспечение ПАК зарегистрировано в реестре российского ПО;
- если программно-аппаратный комплекс реализует функцию защиты информации, то ПО, реализующее такую функцию, должно иметь сертификат ФСТЭК;
- ПАК должен предоставлять функциональную устойчивость.

Таким образом, доверенный ПАК должен состоять исключительно из российских оборудования и программного обеспечения.

Скала^р использует подход, позволяющий организациям любого масштаба и направленности, в том числе субъектам критической информационной инфраструктуры (КИИ), оперативно выстраивать высоконагруженный, обозримый и легко масштабируемый ИТ-ландшафт полностью на российских технологиях, обеспечивая при этом максимальную

совместимость всех компонентов системы. В результате, сроки проектов по импортозамещению значительно сокращаются.

Государственные информационные системы (ГИС)

ГИС могут отличаться рядом формальных требований, таких как наличие **Машины** в Едином реестре российской промышленной и радиоэлектронной продукции, или необходимостью аттестации построенной инфраструктуры, что влечет за собой требование сертификации ответственных компонентов (ФСТЭК) и соответствия требованиям уровней информационной безопасности.

Эти требования могут быть выполнены при помощи **Машины серверной виртуализации Скала^р МДИ.В**.

6. Принципы создания Машины

Машина серверной виртуализации Скала^р МДИ.В создана для работы с масштабируемой виртуальной инфраструктурой

Целью разработки **Машины** было создание программно-аппаратного комплекса (ПАК), специально адаптированного для ПО Базис.vCore и vControl в целях создания высокопроизводительной платформы серверной виртуализации. В **Машине** использованы преимущества тонкой настройки всех компонентов под функции и потребности конкретного ПО и тем самым обеспечен максимум ее производительности.

Комплексное размещение компонентов, применение высокопроизводительных протоколов и устройств хранения также способствуют достижению этой цели.

Использование **Машины серверной виртуализации Скала^р МДИ.В** обеспечивает:

- горизонтальное масштабирование — наращивание вычислительных ресурсов и ресурсов хранения **Модулями Скала^р** без простоев;
- высокую отказоустойчивость — кластер высокой доступности (HA) для ВМ, репликацию блоков данных и алгоритмы Erasure Coding на уровне программно-определяемого хранилища, технологии RAID при применении СХД;
- гибкую архитектуру — поддержка подсистемы хранения как на основе высокопроизводительного гиперконвергентного программно-определяемого хранилища, так и на выделенных узлах хранения программно-определяемого хранилища или классической СХД с универсальным блочным доступом по протоколу iSCSI по сети Ethernet;
- единое управление — web-интерфейс ПО Базис.vControl или посредством REST API, также ряд функций может быть реализован в служебном ПО **Скала^р Геном**.

Технологические принципы

- дублирование критичных компонентов оборудования;
- применение высокопроизводительных компонентов узлов;
- горизонтальное масштабирование вычислительных ресурсов и ресурсов хранения;
- сохранение работоспособности при отказе отдельных элементов системы;
- многоуровневое тестирование ПАК, его узлов и компонентов при производстве;
- комплексный подход к безопасности и защите данных.

Технические решения

- архитектура основана на модулях и подсистемах;
- специальное ПО для управления и мониторинга ПАК;
- глубокая адаптация компонентов для совместной работы в составе **Машины**.

Надежность на уровне архитектуры

- система управления виртуализацией может быть развернута на собственном выделенном кластере высокой доступности и, таким образом, отделена от вычислительной нагрузки в рамках ПАК, что повышает ее безопасность и доступность;
- система мониторинга и управления жизненным циклом ПАК также может быть развернута на выделенных служебных узлах, чтобы исключить воздействие

вычислительной нагрузки на показания регистрируемых метрик и ее зависимость в целом от работы продуктивных вычислительных кластеров.

Высокая производительность на уровне архитектуры

- благодаря выделенной сети внутреннего взаимодействия, все узлы взаимодействуют между собой с одинаково высокой скоростью.

Проработанность всех программных компонентов

Основные программные элементы **Машины серверной виртуализации Скала^р МДИ.В:**

- ПО гипервизора 1-го типа Базис.vCore и системы управления виртуализацией Базис.vControl;
- Программная платформа обслуживания и мониторинга ПАК **Скала^р Геном**.

В **Машине** обеспечены оптимизация, тонкая настройка и доработка перечисленных компонентов для обеспечения их наибольшей производительности и функционального соответствия потребностям заказчика.

В ходе создания **Машины** была разработана методика оптимизации настройки ядра ОС каждого из вычислительных узлов **Машины** под конкретный вариант ее применения. Заказчик получает **Машину серверной виртуализации Скала^р МДИ.В**, настроенную под функциональные требования проекта.

Инсталляция узлов **Машин** на производстве осуществляется при помощи разработанного в **Скала^р** специального ПО, что исключает риск человеческой ошибки.

Практическое применение тиражируемых экземпляров **Машин серверной виртуализации Скала^р МДИ.В** продемонстрировало высокую производительность при эксплуатации.

Команда инженеров и архитекторов **Скала^р** изучают области для дальнейшего развития **Машин** и продолжают работу по оптимизации ПАК, кроме того, постоянно ведется деятельность по развитию систем мониторинга и управления собственной разработки.

Все перечисленное формирует целостную архитектуру и динамику развития **Машины серверной виртуализации Скала^р МДИ.В**.

Сопровождение и поддержка

Важным дополнением ко всему перечисленному является полная ответственность производителя за **Машину** в целом, включая все программные и аппаратные компоненты. Это означает не только уверенность в работоспособности изделия в целом, но и последующую поддержку от единого поставщика в режиме «одного окна», а не от нескольких разных поставщиков, как бывает при самостоятельном подборе, развертывании и настройке компонентов в случае традиционного подхода.

7. Состав Машины серверной виртуализации Скала^р МДИ.В

Ниже приведены термины, используемые для комплектации **Машины серверной виртуализации Скала^р МДИ.В**.

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.

Модуль — это единица комплектации **Машин**, выполняющая определенные функции в соответствии с ее назначением. Модуль является единым и неделимым элементом спецификации и содержит набор аппаратных узлов и ПО. **Модуль**, как и **Машина** целиком, является ПАК согласно классификации ОКПД2.

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

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

Машина серверной виртуализации Скала^р МДИ.В (базовый комплект) состоит из Модулей, представленных на рисунке 1.

Рисунок 1. Состав Машины серверной виртуализации Скала^р МДИ.В

Комплекты поставки

Машины серверной виртуализации Скала^р МДИ.В поставляются в виде функционально необходимого набора **Модулей Скала^р** и комплектуются в соответствии с

показателями назначения, полученными от заказчика. **Машина** включает в себя Базовый модуль и дополняется комплектом Модулей расширения и/или специальными Модулями.

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

Комплект Модулей расширения — это набор **Модулей Скала^р**, позволяющий масштабировать ПАК, например, когда не хватает портовой емкости, или есть необходимость увеличить производительность или объем хранения данных. Кроме того, можно добавить специальные **Модули Скала^р**, позволяющие расширить функциональность ПАК.

На диаграмме ниже (Рисунок 2) представлена комплектация **Машины серверной виртуализации Скала^р МДИ.В** (Базовый комплект и комплект Модулей расширения), а также соответствие Модулей **Машины** функциональным подсистемам.

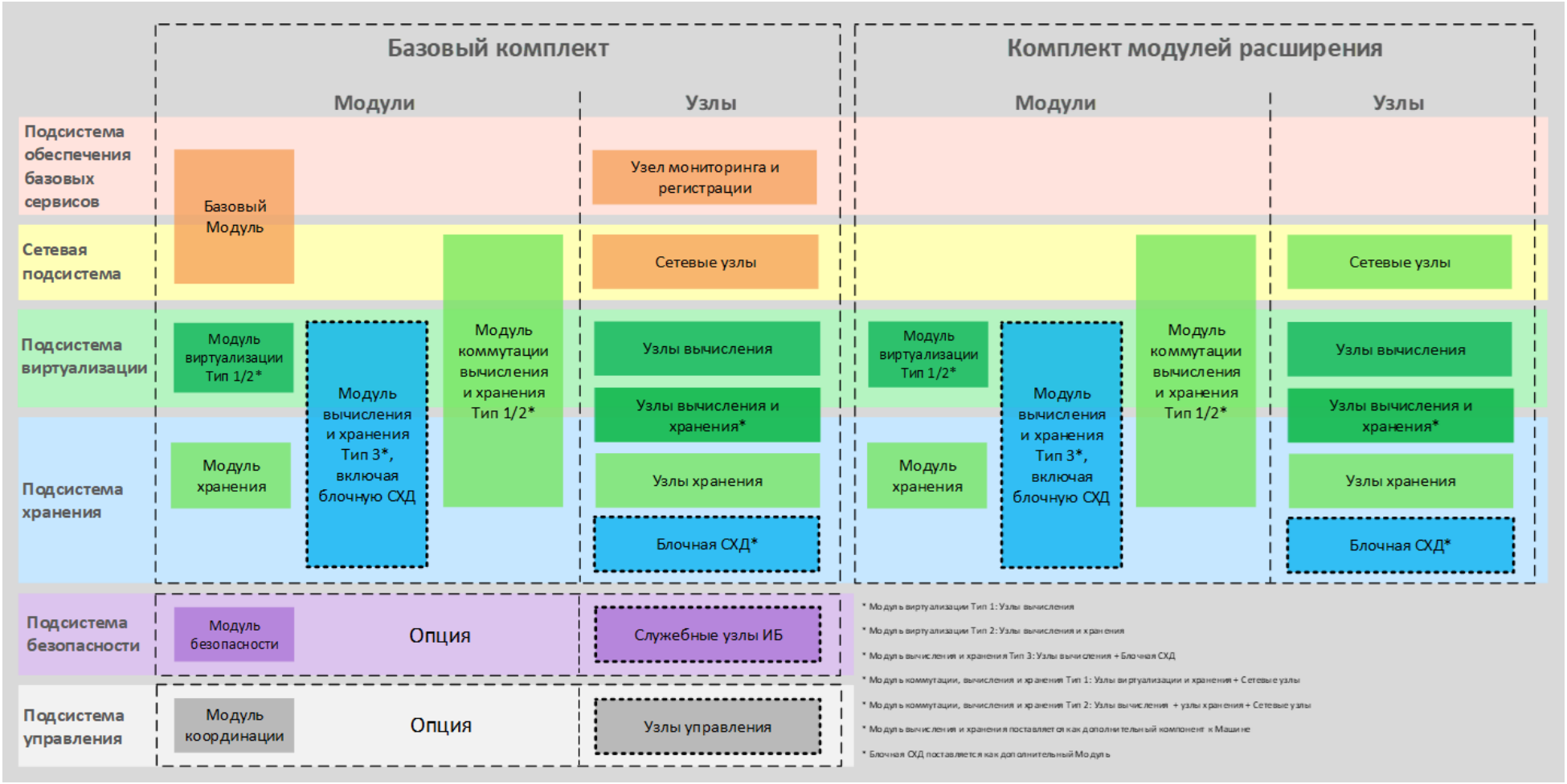

Рисунок 2. Комплектация и подсистемы Машины серверной виртуализации Скала^р МДИ.В

7.1 Подсистемы Машины

Функции **Машины серверной виртуализации Скала^р МДИ.В** логически объединены в подсистемы (Рисунок 2). Часть подсистем обеспечивают основной функционал и всегда включены в **Машину**, а часть — предоставляют дополнительный функционал и могут быть добавлены по требованию заказчика.

Основной функционал обеспечивается набором подсистем, необходимых **Машине серверной виртуализации Скала^р МДИ.В** для выполнения задач прямого назначения.

Дополнительный функционал предоставляется набором подсистем из Модулей, обеспечивающих расширение функций **Машины серверной виртуализации Скала^р МДИ.В**.

7.1.1 Подсистема обеспечения базовых сервисов

Подсистема обеспечения базовых сервисов отвечает за мониторинг и управление аппаратными и программными компонентами **Машины Скала^р МДИ.В**. В нее включены два узла мониторинга и регистрации Базового модуля (см. 7.2.1), на которых предустановлена программная платформа **Скала^р Геном**, выполняющая следующие функции:

- сбор, хранение и отображение данных мониторинга;
- настройка правил оповещения;
- отправка оповещений о состоянии ПАК;
- управление аппаратными компонентами;
- настройка программных компонентов;
- настройка интеграции мониторинга и управления компонентами со сторонним ПО.

Архитектура подсистемы обеспечения базовых сервисов обеспечивает отказоустойчивый режим работы с двойным резервированием.

Подсистема обеспечения базовых сервисов реализуется в **Базовом модуле** (см. 7.2.1).

7.1.2 Сетевая подсистема

Сетевая подсистема выполняет функций организации сетевой связанности между всеми узлами, входящими в состав **Машины**, и представляет собой набор сетевых узлов (коммутаторов), которые организуют изолированные высокоскоростные сети:

- внутреннего взаимодействия (в зависимости от требований заказчика 25 Гбит/с или 100 Гбит/с) — для организации быстрой связности с минимальными задержками между всеми компонентами ПАК, в том числе, для работы гиперконвергентного программно-определяемого хранилища;
- доступа к хранилищу (25 Гбит/с), при использовании классической СХД в составе **Машины** (может быть отдельно от интерконнекта или совмещена с ним);
- внешнего доступа (в зависимости от требований заказчика 25 Гбит/с или 100 Гбит/с) — для организации доступа к виртуальным машинам и элементам управления виртуализацией и администрирования ПАК;
- управления (1 Гбит/с) — для организации выделенной сети низкоуровневого управления всеми узлами ПАК.

Стартовый комплект сетевых узлов всегда относится к **Базовому модулю** (см. 7.2.1).

7.1.3 Подсистема виртуализации

Подсистема виртуализации обеспечивает вычислительными ресурсами среду виртуализации (центральный процессор, оперативная память, дисковые ресурсы) и состоит из узлов вычисления и хранения (в случае использования гиперконвергентного программного

хранилища), либо только вычислительных узлов (в случае использования вычислительных кластеров на основе конвергентного программного хранилища или классических аппаратных СХД) с гипервизором 1-го типа Базис.vCore, который управляет разделяемым доступом виртуальных машин к ресурсам вычислительного узла.

Основные функции подсистемы виртуализации:

- предоставление вычислительных ресурсов вычислительных узлов виртуальным машинам;
- управление разделяемым доступом к вычислительным ресурсам, динамическое распределение и оптимизациях их использования;
- реализация жизненного цикла виртуальных машин: создание, запуск, остановка, удаление.

Подсистема реализуется **Модулями виртуализации** (см. 7.2.2) и также может расширяться ими или узлами **Модуля коммутации, вычисления и хранения** (см. 7.2.5).

7.1.4 Подсистема хранения

Подсистема хранения может иметь вариант исполнения как гиперконвергентного (реализуемого на вычислительных узлах) программно-определяемого распределенного хранилища данных (SDS) под управлением ПО **Базис.uStor**, состоящего как минимум из 4-х унифицированных узлов вычисления и хранения. Гиперконвергентное решение - наиболее производительный вариант в части подсистемы хранения, т.к. доступ осуществляется на уровне QEMU-драйвера модуля ядра гипервизора.

Возможны и два иных исполнения хранения – с физически выделенной ролью хранения в составе Машины – вариант конвергентного программно-определяемого хранилища под управлением ПО **Базис.uStor** (в специальной расширенной версии), и вариант классической аппаратной СХД, оба с блочным доступом по протоколу iSCSI и дальнейшем развертыванием кластера OCFS2 в подсистеме управления виртуализацией. Эти варианты дают меньшую производительность, чем гиперконвергентный вариант, но предоставляют более универсальный доступ и возможность подключения к сторонним системам и сервисам.

Допускается также совместное использование разных вариантов подсистемы хранения в рамках одной Машины, но в разных вычислительных кластерах.

Подсистема хранения обеспечивает дисковыми ресурсами подсистемы виртуализации и подсистему управления, здесь размещаются образы дисков всех виртуальных машин, разворачиваемых в виртуальной среде.

Подсистема реализуется узлами **Модулей виртуализации** (см. 7.2.2) и/или узлами **Модулей хранения** (см. 7.2.3) и/или узлами **Модулей коммутации, вычисления и хранения** (см. 7.2.5).

В случае применения аппаратной СХД, подсистема реализуется **Модулем вычисления и хранения**, дополнительным к комплектации **Машины**.

7.1.5 Подсистема безопасности

В подсистеме безопасности (для сертифицированной ФСТЭК версии ПО виртуализации) функционируют три выделенных физических служебных узла безопасности с развернутым на них в кластерном отказоустойчивом исполнении ПО Базис.Virtual Security – средством защиты виртуальной инфраструктуры от несанкционированного доступа.

Подсистемой безопасности обеспечивается выполнение следующих функций:

- доверенная загрузка виртуальных машин;
- контроль целостности виртуальных машин;
- регистрация и логирование событий безопасности;

- управление доступом к виртуальной инфраструктуре;
- ограничение программной среды в средстве виртуализации;
- управление потоками информации в среде виртуализации;
- идентификация и аутентификация пользователей.

Подсистема реализуется **Базовым модулем безопасности** (см. 7.2.6).

7.1.6 Подсистема управления

Задача подсистемы управления – обеспечить независимую от продуктивной среды собственную отказоустойчивую среду виртуализации для размещения на своих вычислительных мощностях всех служебных инфраструктурных виртуальных машин и управление их жизненным циклом.

Этот логический компонент ПАК в полной максимальной конфигурации состоит из четырех вычислительных узлов управления – контроллеров, собранных в отказоустойчивый гиперконвергентный кластер, на котором функционирует все программное обеспечение, отвечающее за работоспособность платформы виртуализации, включая внутренние служебные базы данных, интерфейсы управления GUI и REST API, со своим собственным программно-определяемым хранилищем под единым управлением программного продукта Базис.vControl. В случае использования в качестве общей разделяемой системой хранения данных классической аппаратной СХД количество узлов кластера сокращается до трех.

Подсистема реализуется **Модулем координации** (см. 7.2.7).

Примечание: в минимальной конфигурации все служебные и управляющие программные компоненты могут размещаться на ресурсных узлах, вместе с продуктивными виртуальными машинами заказчика. В этом случае Подсистема управления размещается на вычислительных узлах **Базового модуля** и узлах **Модулей виртуализации**.

7.2 Модули Машины

В разрезе модульности, **Машина серверной виртуализации Скала^р МДИ.В** состоит из определенного набора следующих функциональных модулей **Скала^р** (Рисунок 2):

- **Базового модуля**;
- **Модуля виртуализации**;
- **Модуля хранения**;
- **Модуля вычисления и хранения** (только при применении аппаратной СХД, как дополнительный к Машине Модуль);
- **Модуля коммутации, вычисления и хранения**;
- **Базового модуля безопасности**;
- **Модуля координации**.

Часть модулей обеспечивают основной функционал и всегда включены в **Машину**, а часть – дополнительный функционал и могут быть добавлены по требованию заказчика.

7.2.1 Базовый модуль

Название в Едином реестре Минпромторга – **СКАЛА-Р Базовый модуль**.

Базовый модуль используется для обеспечения сетевой связанности между компонентами, служебных функций, а также для организации выделенной сети управления **Машиной**.

В составе **Базового модуля** есть необходимые компоненты для подключения **Машины** к вышестоящим сетевым узлам.

Типовой состав для продуктивной инсталляции ПАК:

- в максимальной комплектации это два вычислительных узла, объединенных в зеркальный кластер для служебных функций и исполнения Программной платформы **Скала^р Геном**, включая мониторинг ПАК. В минимальной конфигурации эти функции могут размещаться на ресурсных узлах **Модулей виртуализации** вместе с виртуальными машинами заказчика, тогда в **Базовый модуль** включаются один или более дополнительных ресурсных вычислительных узла);
- два сетевых узла 100 Гбит/с для внутреннего сетевого взаимодействия в случае использования программно-определяющего хранилища;
- два сетевых узла 25 Гбит/с для внутреннего сетевого взаимодействия в случае использования аппаратной СХД с блочным доступом;
- два сетевых узла 25 Гбит/с для сети внешнего доступа;
- один сетевой узел 1 Гбит/с для выделенной сети управления.

Типовые характеристики вычислительных узлов **Модуля** представлены в таблице ниже:

Категория	Характеристика
Процессор	2x CPU (2.1 ГГц / 24 ядра)
Оперативная память	512 Гбайт
Диски (накопители)	диски ОС; 12x 1.92 Тбайт SSD
Сетевые карты	100 Гбит/с, dual port 10/25 Гбит/с, dual port 1 Гбит/с RJ45 (на материнской плате) 1 Гбит/с RJ45 IPMI (на материнской плате)

Характеристики сетевых узлов **Модуля**:

- сетевые узлы 100 Гбит/с – маршрутизирующие коммутаторы с 32 портами 100 GbE.
- сетевые узлы 25 Гбит/с – маршрутизирующие коммутаторы с 48 портами 25 GbE и 8 портами 100 GbE.
- сетевой узел 1 Гбит/с снабжен 48 портами 1 GbE и 4 портами 10 GbE.

Применяемое программное обеспечение:

- Операционная система Альт 8 СП или Астра SE;
- ПО **Скала^р Геном**;
- ПО Аванпост FAM (опционально).

7.2.2 Модуль виртуализации

Название в Едином реестре Минпромторга – **СКАЛА-Р Модуль виртуализации**.

В состав **Модуля** входит один или более вычислительный узел виртуализации в исполнении для конвергентного варианта **Машины** (без локальных дисков для создания хранилища) или узел вычисления и хранения для гиперконвергентного варианта **Машины** (с дисками для SDS).

Вычислительные узлы **Модуля виртуализации** образуют единый кластер вычисления. Это физические узлы, которые формируют вычислительные мощности для системы виртуализации (ЦПУ/ОЗУ). В состав **Машины** может входить до 100 узлов вычисления, предоставляющих вычислительные ресурсы для ВМ в рамках разворачиваемой виртуальной

инфраструктуры, при этом в один вычислительный кластер, определяющий максимально возможный домен обеспечения высокой доступности для ВМ, рекомендуется включать не более 20-28 вычислительных узлов.

Вычислительные узлы виртуализации в **Модулях виртуализации**, а также в **Модулях коммутации, вычисления и хранения** должны быть однотипные по вычислительным ресурсам в рамках сборки каждого кластера.

Типовые характеристики вычислительных узлов Модуля представлены в таблице ниже:

Категория	Характеристика
Процессор	2x CPU (2.6 ГГц / 28 ядер)
Оперативная память	1024 Гбайт
Диски (накопители)	диски ОС; диски (только для SDS) – от требований заказчика
Сетевые карты	2x 100 Гбит/с, dual port 10/25 Гбит/с, dual port 1 Гбит/с RJ45 (на материнской плате) 1 Гбит/с RJ45 IPMI (на материнской плате)

Вычислительный узел типично предоставляет под полезную нагрузку до 50 ядер pCPU без переподписки и учета гипертрединга, и 1024 Гбайт ОЗУ.

Применяемое в **Модуле** программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore;
- Основное функциональное ПО виртуализации Базис;
- ПО управления хранением Базис (при применении SDS).

7.2.3 Модуль хранения

Название в Едином реестре Минпромторга – **СКАЛА-Р Модуль хранения**.

Модуль хранения предназначен для организации конвергентного (отдельного от вычислений) исполнения программно-определяемого распределенного хранилища **Базис.uStor** с выделенными узлами хранения. На них не размещаются виртуальные машины.

Модуль может использоваться для организации хранилища как для узлов **Модулей координации** (опция), так и для «бездисковых» узлов **Модуля виртуализации**. Программное хранилище подключается через выделенную сеть хранения данных, поддерживаемый протокол работы с СХД – iSCSI с последующим развертыванием кластера OCFS2 для обеспечения файлового доступа подсистеме виртуализации.

Типовые характеристики вычислительных узлов с ролью только хранения (SDS) представлены в таблице ниже:

Категория	Характеристика
Процессор	2x CPU (3 ГГц / 12 ядер)
Оперативная память	256 Гбайт
Диски (накопители)	диски ОС; диски для SDS – от требований заказчика
Сетевые карты	4x 100 Гбит/с, dual port 1 Гбит/с RJ45 (на материнской плате) 1 Гбит/с RJ45 IPMI (на материнской плате)

Применяемое в **Модуле** программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore;
- ПО управления хранением Базис.uStor.

7.2.4 Модуль вычисления и хранения[1]

Название в Едином реестре Минпромторга – **СКАЛА-Р Модуль вычисления и хранения**.

Состав (минимально): один или более вычислительных узлов, масштабирующих ресурсы иных Модулей Машины, и узел хранения на основе блочной аппаратной СХД.

Система хранения на основе блочной СХД используется в качестве опции для организации постоянного хранилища как для узлов **Модулей координации** (опция), так и для узлов **Модуля виртуализации** (или иных модулей с ролью ресурса виртуализации). Подключается через выделенную сеть хранения данных, поддерживаемый протокол работы с СХД – iSCSI с последующим развертыванием кластера OCFS2 для обеспечения в подсистеме виртуализации файлового доступа.

Типовые характеристики СХД в составе **Модуля** представлены в таблице ниже:

Категория	Характеристика
Контроллеры	Два независимых
Кэш	512 Гбайт
Диски (накопители)	56x 1.92 Тбайт SSD (количество и тип дисков могут быть уточнены)
Сетевые карты	8x 10/25 Гбит/с, dual port 2x 1 Гбит/с RJ45

Применяемое в **Модуле** программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore на вычислительном узле (узлах);
- Основное функциональное ПО виртуализации Базис на вычислительном узле (узлах);
- ПО управления СХД от вендора СХД.

7.2.5 Модуль коммутации, вычисления и хранения

Название в Едином реестре Минпромторга – СКАЛА-Р Модуль коммутации, вычисления и хранения.

Модуль используется для обеспечения дополнительной сетевой связности между компонентами **Машины** при масштабировании, а также обязательно содержит набор вычислительных узлов (от одного и более), масштабирующих ресурсы иных **Модулей** ПАК.

В состав **Модуля** входят необходимые сетевые компоненты для расширения портов коммутации и их агрегации (при необходимости):

- два сетевых узла (пара) 100 Гбит/с для внутреннего сетевого взаимодействия в случае использования программно-определяющего хранилища;
- или два сетевых узла (пара) 25 Гбит/с для сети доступа в случае использования аппаратной СХД с блочным доступом;

[1] Примечание. На дату публикации данного документа аппаратная СХД может быть добавлена в ПАК **Машина МДИ.В** только как дополнительная позиция вне Машины, отдельным ПАК СКАЛА-Р Модуль вычисления и хранения.

- два сетевых узла (пара) 25 Гбит/с для сети внешнего доступа к ресурсам виртуализации;
- один сетевой узел 1 Гбит/с для расширения сети управления;
- в зависимости от шага масштабирования и при необходимости: по паре сетевых узлов 100 Гбит/с для агрегации сети внутреннего взаимодействия, для сети внешнего доступа, или для сети доступа к СХД соответственно.

Характеристики сетевых узлов **Модуля**:

- сетевые узлы 100 Гбит/с снабжены 32 портами 100 GbE.
- сетевые узлы 25 Гбит/с снабжены 48 портами 10/25 GbE и 8 портами 100 GbE.
- сетевой узел 1 Гбит/с снабжен 48 портами 1 GbE и 4 портами 10 GbE.

Применяемое в **Модуле** программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore **на** вычислительном узле (узлах);
- Основное функциональное ПО виртуализации Базис и/или ПО управления хранением Базис (при применении SDS) на вычислительном узле (узлах).

7.2.6 Базовый модуль безопасности

Название в Едином реестре Минпромторга – **СКАЛА-Р Базовый модуль безопасности**.

Базовый модуль безопасности всегда состоит из трех вычислительных узлов.

Базовый модуль безопасности является компонентом безопасности на основе программной платформы Базис.Virtual Security, применяемой в составе сертифицированной ФСТЭК версии ПО виртуализации в составе **Машины**.

Базис.Virtual Security является программным средством со встроенными средствами защиты от несанкционированного доступа в виртуальной инфраструктуре

Типовые характеристики вычислительных узлов **Модуля** представлены в таблице ниже:

Категория	Характеристика
Процессор	2x CPU (2.1 ГГц / 24 ядра)
Оперативная память	512 Гбайт
Диски (накопители)	диски ОС; 12x 1.92 Тбайт SSD
Сетевые карты	100 Гбит/с, dual port 10/25 Гбит/с, dual port 1 Гбит/с RJ45 (на материнской плате) 1 Гбит/с RJ45 IPMI (на материнской плате)

Применяемое программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore;
- ПО Базис.Virtual Security;
- ПО Axiom JDK или аналог.

7.2.7 Модуль координации

Название в Едином реестре Минпромторга – **СКАЛА-Р Модуль координации**.

Модуль координации представляет собой выделенный кластер управления высокой доступности, построенный из четырех (в случае SDS) или трех (в случае аппаратной СХД)

физических узлов управления на базе гипервизора Базис.vCore с разделяемой системой хранения (программной или аппаратной) под управлением программного продукта Базис.vControl.

Задача **Модуля координации** – обеспечить независимую от продуктивной среды собственную отказоустойчивую среду виртуализации для размещения на своих вычислительных мощностях служебных инфраструктурных ВМ.

Типовые характеристики вычислительных узлов **Модуля** представлены в таблице ниже:

Категория	Характеристика
Процессор	2x CPU (2.1 ГГц / 24 ядра)
Оперативная память	512 Гбайт
Диски (накопители)	диски ОС; 12x 1.92 Тбайт SSD (при применении SDS)
Сетевые карты	100 Гбит/с, dual port 10/25 Гбит/с, dual port 1 Гбит/с RJ45 (на материнской плате) 1 Гбит/с RJ45 IPMI (на материнской плате)

Применяемое программное обеспечение:

- ПО гипервизора 1-го типа Базис.vCore;
- Основное функциональное ПО виртуализации Базис;
- ПО управления хранением Базис (при применении SDS).

7.3 Производительность хранилищ различной архитектуры в Машине

В зависимости от технического решения при комплектации **Машины**, будут достигаться различные уровни и показатели производительности хранилища. Различия указаны ниже.

Ожидаемый уровень производительности гиперконвергентного хранилища (прямой файловый доступ к хранилищу через QEMU-драйвер): на один вычислительный узел при использовании NVMe SSD дисков – от 700 000 до 1 млн IOPS для последовательных операций чтения блоками по 4K со средней задержкой 1-2 миллисекунд (в зависимости от конкретных моделей SSD дисков и их количества в узле). Наращивание количества вычислительных узлов приводит к практически линейному росту общей производительности хранилища.

Ожидаемый уровень производительности выделенного конвергентного программного хранилища (блочный доступ по iSCSI с кластерной файловой системой OCFS2): на один узел хранения при использовании NVMe SSD дисков – от 250 000 до 500 000 IOPS для последовательных операций чтения блоками по 4K со средней задержкой 1-2 миллисекунд (в зависимости от конкретных моделей SSD дисков и их количества в узле). Наращивание количества вычислительных узлов приводит к практически линейному росту общей производительности хранилища.

Ожидаемый уровень производительности подсистемы хранения в случае использования аппаратной СХД (блочный доступ по iSCSI с кластерной файловой системой OCFS2): определяется характеристиками СХД конкретного производителя.

8. Архитектура Машины серверной виртуализации Скала^р МДИ.В

8.1 Общая (максимальная) конфигурация

Конфигурация **Машины** в общем случае (максимальном варианте конфигурации) представлена на рисунке 3. Данная конфигурация подходит для решения большинства стандартных задач, связанных с обеспечением работы информационных систем. Различные варианты исполнения общей конфигурации представлены в разделе 8.2.

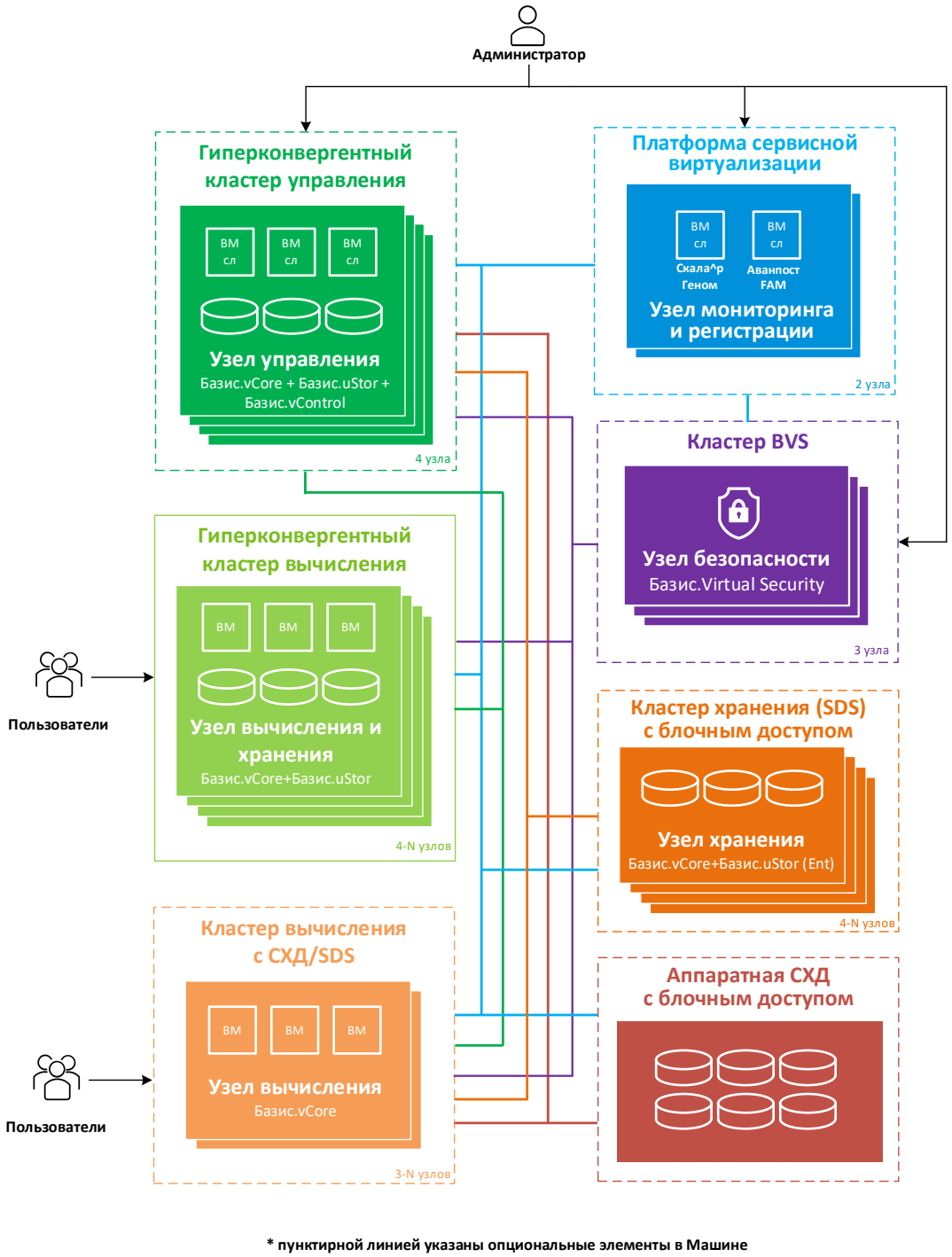

Рисунок 3. Схема Машины серверной виртуализации Скала^р МДИ.В в общей (максимальной) конфигурации

Максимальное количество узлов вычисления в одном комплексе **Машины** определяется только способностью обеспечить их сетевую связность с общими разделяемыми узлами хранения (обеспечивающими классическую аппаратную или распределенную программно-определяемую систему хранения данных), а также фактической утилизацией этих узлов хранения вычислительными узлами.

Однако, при прочих равных условиях, рекомендуется включать в состав одного вычислительного кластера до 28 вычислительных узлов для удобства обслуживания, уменьшения домена отказа и снижения нагрузки на узлы хранения. Большее количество вычислительных узлов объединяются в несколько вычислительных кластеров со своими собственными узлами хранения.

Тем не менее, технически верхнего предела для количества вычислительных узлов нет, как и не ограничено количество вычислительных кластеров. На настоящий момент в промышленной эксплуатации находятся конфигурации **Машины** с централизованным управлением, состоящие из более чем 10 отказоустойчивых вычислительных кластеров.

При обоснованной целесообразности в состав **Машины** могут быть добавлены как бездисковые вычислительные узлы, так и узлы, из ресурсов которых в кластере используются только диски. Программные компоненты в этом смысле не накладывают никаких ограничений на вычислительные узлы в составе кластеров.

8.2 Варианты исполнения Машины

Варианты исполнения **Машины** отличаются наличием или отсутствием гиперконвергентного либо выделенного конвергентного хранилища **Базис.uStor** (см. раздел 9.5), наличием или отсутствием аппаратной СХД, реализацией функционала базовых сервисов и кластера управления на выделенных вычислительных узлах либо в виртуальных машинах на общих ресурсных узлах, и наличием или отсутствием кластера BVS. Варианты исполнения представлены в таблице 1 и подробно описаны в подразделах ниже.

Таблица 1. Различные варианты исполнения Машины МДИ.В

№	Варианты исполнения	Служебные узлы	Кластер управления	Кластер BVS	Узлы вычисления и хранения	Узлы вычисления	Узлы хранения	Аппаратная СХД
	Исполнения **Машины Скала^р МДИ.В** с гиперконвергентным хранилищем Базис.uStor (SDS)							
1	Минимальное исполнение с гиперконвергентным SDS	□	□	□	■	□	□	□
2	Минимальное исполнение ФСТЭК с гиперконвергентным SDS	□	□	■	■	□	□	□
3	Максимальное исполнение ФСТЭК с гиперконвергентным SDS	■	■	■	■	□	□	□
	Исполнения **Машины Скала^р МДИ.В** с выделенным конвергентным хранилищем Базис.uStor (SDS)							
4	Минимальное исполнение с выделенным SDS	□	□	□	□	■	■	□
5	Минимальное исполнение ФСТЭК с выделенным SDS	□	□	■	□	■	■	□
6	Максимальное исполнение ФСТЭК с выделенным SDS	■	■	■	□	■	■	□

№	Варианты исполнения	Служебные узлы	Кластер управления	Кластер BVS	Узлы вычисления и хранения	Узлы вычисления	Узлы хранения	Аппаратная СХД
	Исполнения **Машины Скала^р МДИ.В** с аппаратной СХД							
7	Минимальное исполнение с аппаратной СХД	□	□	□	□	■	□	■
8	Минимальное исполнение ФСТЭК с аппаратной СХД	□	□	■	□	■	□	■
9	Максимальное исполнение ФСТЭК с аппаратной СХД	■	■	■	□	■	□	■

В минимальных исполнениях все программные компоненты обеспечения базовых сервисов (обслуживание, мониторинг, ПО Аванпост) и управления виртуализацией vControl размещаются в виртуальных машинах на продуктивных вычислительных узлах **Машины**. Это уменьшает общее количество требуемых вычислительных узлов, но одновременно сокращает объем доступных заказчику ресурсов на продуктивных вычислительных узлах. При расчете **Машин** для заказчика эти факторы, безусловно, учитываются.

Минимальная конфигурация **Машины Скала^р МДИ.В** с гиперконвергентным программно-определяемым хранилищем (п. 8.2.1) – четыре узла вычисления и хранения и сетевая подсистема с не менее чем 2+1 сетевыми узлами.

Минимальная конфигурация **Машины Скала^р МДИ.В** с выделенным программно-определяемым хранилищем (п. 8.2.4) – четыре узла хранения с тремя узлами вычисления и сетевая подсистема с не менее чем 2+2+1 сетевыми узлами.

Машина с программно-определяемым хранилищем в минимальном исполнении обеспечивает непрерывность доступа к данным и их сохранность при потере любого одного узла вычисления и хранения для гиперконвергентного варианта, и одновременной потере любых одного узла вычисления и одного узла хранения для варианта выделенного SDS.

В случае использования вычислительным кластером в качестве разделяемой системы хранения аппаратной СХД, число вычислительных узлов в минимальной конфигурации, как и в случае с выделенным SDS, может быть сокращено до трех с не менее чем 2+1 сетевыми узлами (п. 8.2.7).

8.2.1. Исполнение 1 (минимальное исполнение с гиперконвергентным хранилищем SDS)

В данном исполнении представлены только вычислительные узлы с гиперконвергентным хранилищем Базис.uStor. Все программные компоненты обеспечения базовых сервисов (обслуживание, мониторинг, ПО Аванпост) и управления виртуализацией vControl размещаются в виртуальных машинах на вычислительных узлах комплекса, сокращая объем доступных заказчику ресурсов. Схема данного исполнения представлена на рисунке 4. Для ПАК с малым количеством узлов роли коммутаторов внешнего и внутреннего взаимодействия могут быть объединены в одной паре сетевых узлов.

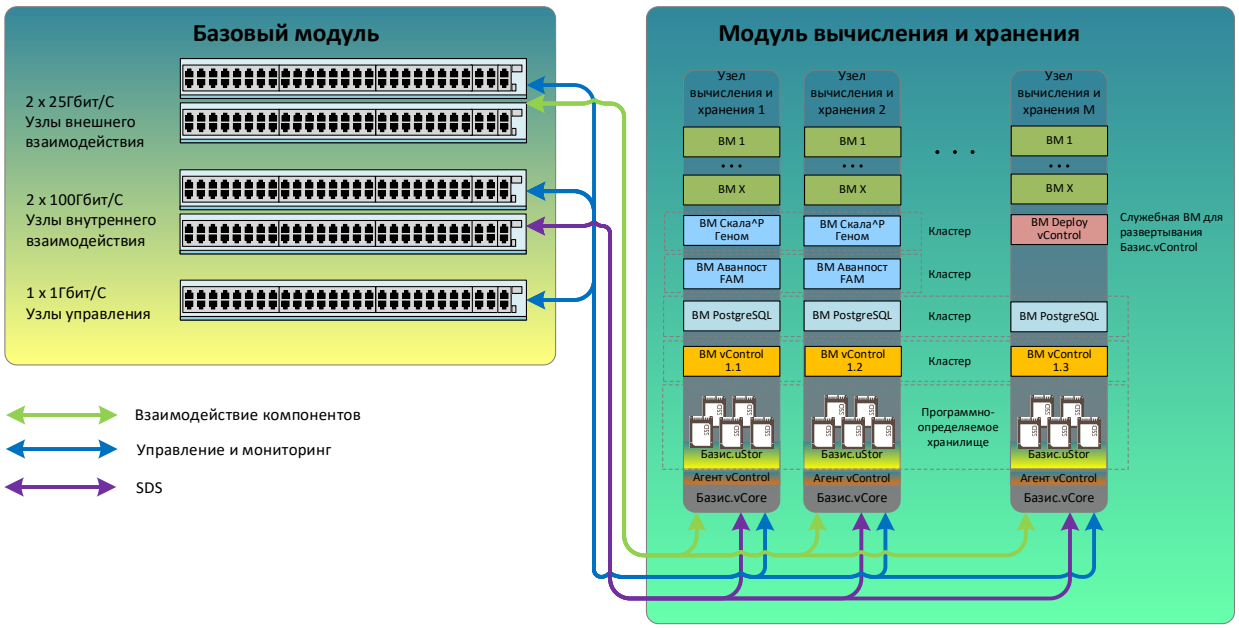

Рисунок 4. Схема исполнения 1
(минимальное исполнение с гиперконвергентным хранилищем SDS)

8.2.2. Исполнение 2 (минимальное исполнение ФСТЭК с гиперконвергентным хранилищем SDS)

Так же, как и в случае с Исполнением 1, используется минимальная конфигурация **Машины** с гиперконвергентным хранилищем, однако, для случаев, когда требуется аттестация ФСТЭК, в состав **Машины** необходимо добавить **Базовый модуль безопасности** с ПО Базис.Virtual Security – сертифицированного средства защиты виртуальной инфраструктуры. Схема такого исполнения представлена на рисунке 5.

Для ПАК с малым количеством узлов роли коммутаторов внешнего и внутреннего взаимодействия могут быть объединены в одной паре сетевых узлов.

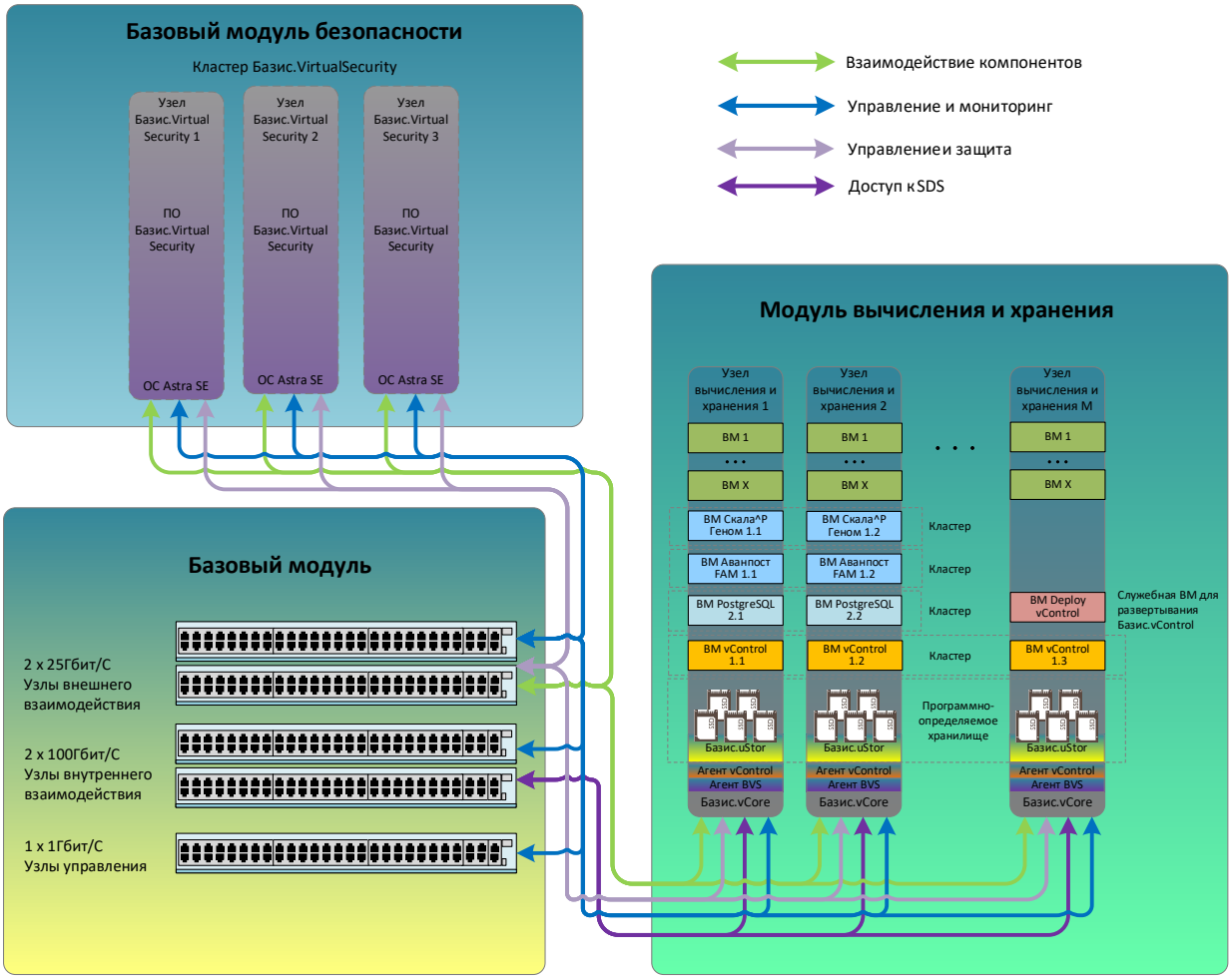

Рисунок 5. Схема исполнения 2
(минимальное исполнение ФСТЭК с гиперконвергентным хранилищем SDS)

8.2.3. Исполнение 3 (максимальное исполнение ФСТЭК с гиперконвергентным хранилищем SDS)

Максимальное исполнение **Машины Скала^р МДИ.В** с гиперконвергентным программно-определяемым хранилищем, с выделенным кластером управления, с выделенными служебными узлами, с кластером Базис.Virtual Security (сертификат ФСТЭК) для возможности аттестации решения по требованиям ИБ.

Схема такого исполнения представлена на рисунке 6.

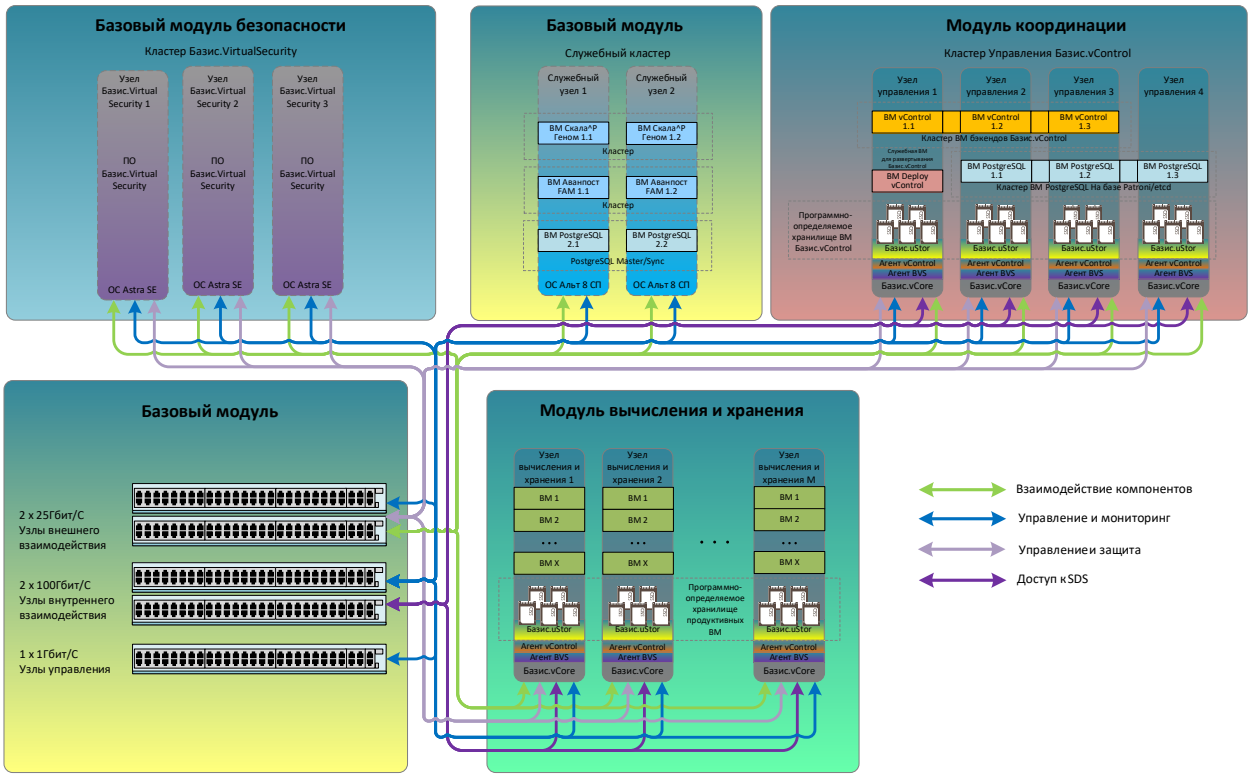

Рисунок 6. Схема исполнения 3
(максимальное исполнение ФСТЭК с гиперконвергентным хранилищем SDS)

8.2.4. Исполнение 4 (минимальное исполнение с выделенным конвергентным хранилищем SDS)

В данном исполнении представлены только вычислительные узлы и узлы хранения с конвергентным хранилищем Базис.uStor. Все программные компоненты обеспечения базовых сервисов (обслуживание, мониторинг, ПО Аванпост) и управления виртуализацией vControl размещаются в виртуальных машинах на вычислительных узлах комплекса, сокращая объем доступных заказчику ресурсов (компенсируется добавлением ресурсов на этапе сайзинга).

Схема данного исполнения представлена на рисунке 7.

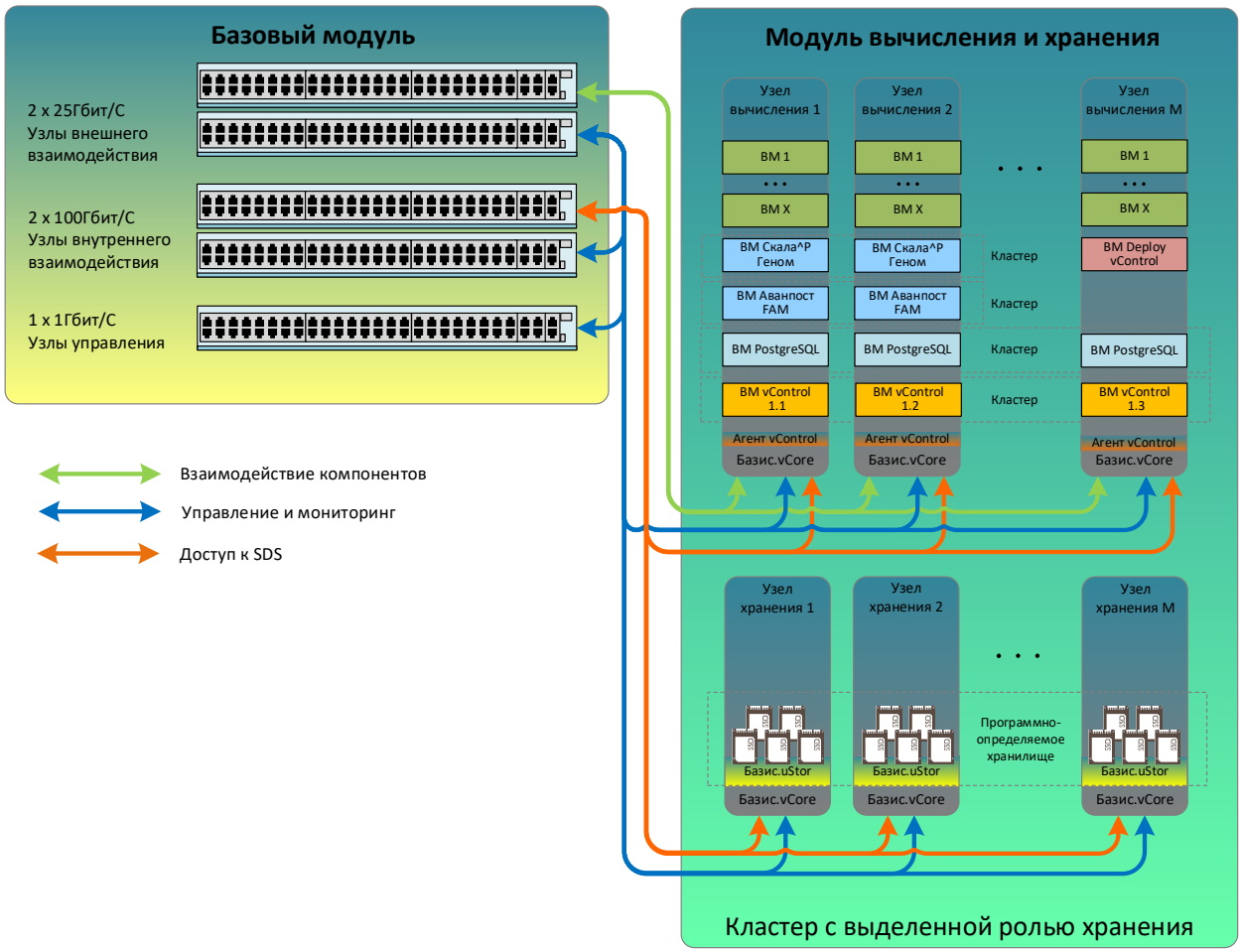

Рисунок 7. Схема исполнения 4
(минимальное исполнение с выделенным конвергентным хранилищем SDS)

8.2.5. Исполнение 5 (минимальное исполнение ФСТЭК с выделенным конвергентным хранилищем SDS)

Так же, как и в случае с Исполнением 4, используется минимальная конфигурация **Машины** с выделенным конвергентным хранилищем Базис.uStor, однако, для случаев, когда требуется аттестация ФСТЭК, в состав **Машины** необходимо добавить **Базовый модуль безопасности** с ПО Базис.Virtual Security – сертифицированного средства защиты виртуальной инфраструктуры.

Схема такого исполнения представлена на рисунке 8.

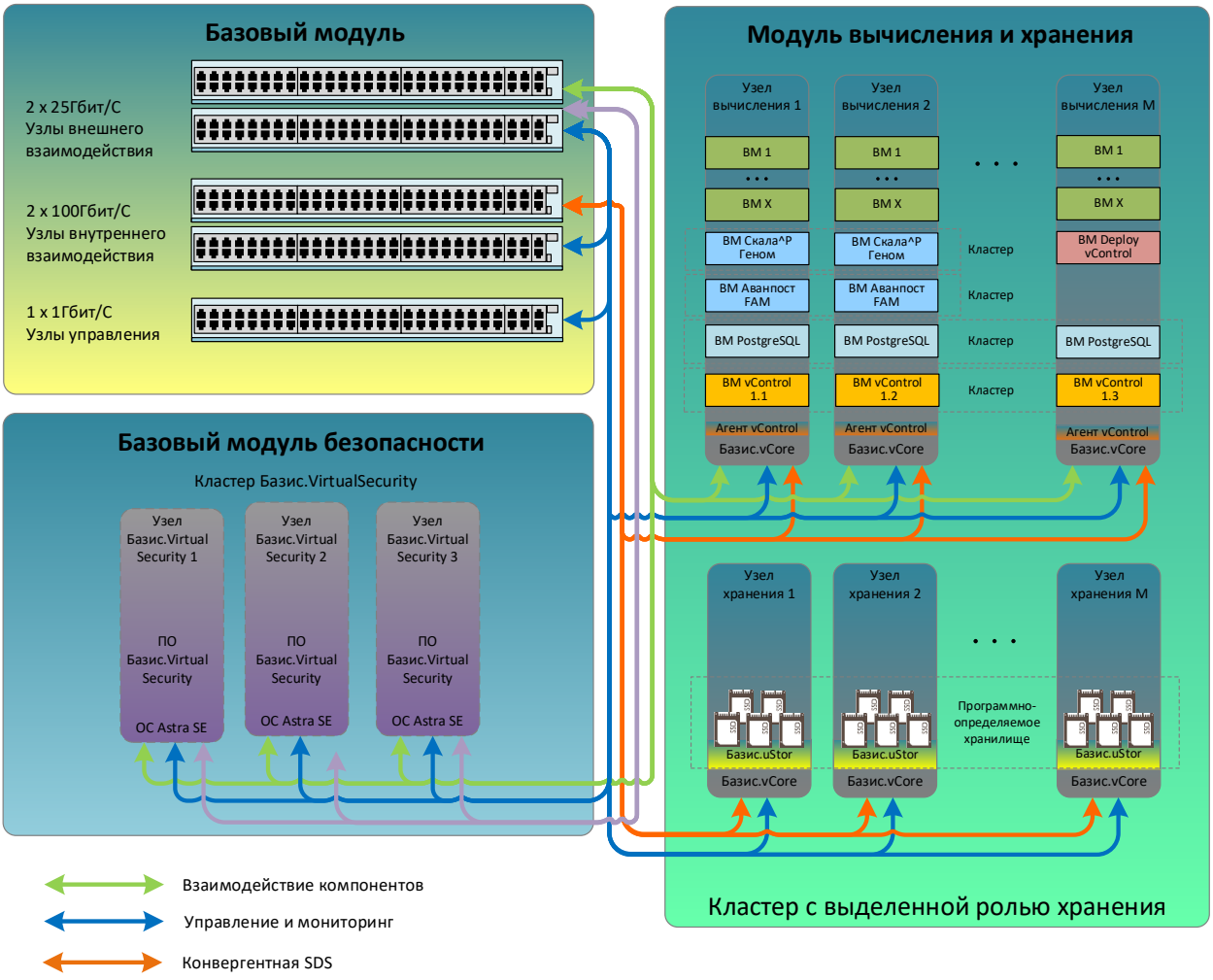

Рисунок 8. Схема исполнения 5
(минимальное исполнение ФСТЭК с выделенным конвергентным хранилищем SDS)

8.2.6. Исполнение 6 (максимальное исполнение ФСТЭК с выделенным конвергентным хранилищем SDS)

Максимальное исполнение **Машины Скала^р МДИ.В** с выделенным программно-определяемым хранилищем, с выделенным кластером управления, с выделенными служебными узлами, с кластером Базис.Virtual Security (сертификат ФСТЭК) для возможности аттестации решения по требованиям ИБ.

Схема такого исполнения представлена на рисунке 9.

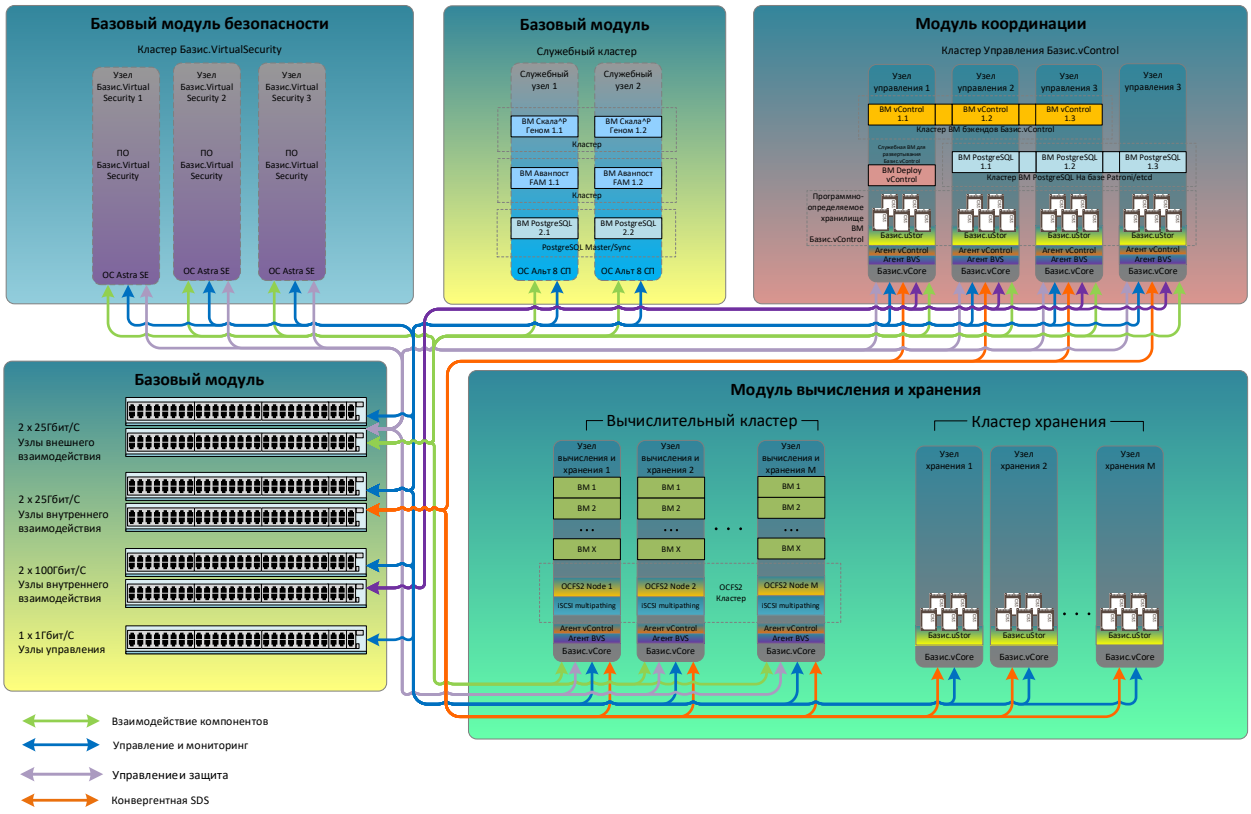

Рисунок 9. Схема исполнения 6
(максимальное исполнение ФСТЭК с выделенным конвергентным хранилищем SDS)

8.2.7. Исполнение 7 (минимальное исполнение с аппаратной СХД)

Данное исполнение аналогично Исполнению 1, за исключением использования в качестве разделяемого общего хранилища классической аппаратной СХД с блочным доступом и построения кластера OCFS2 для обеспечения файлового доступа со стороны гипервизоров. Все программные компоненты обеспечения базовых сервисов (обслуживание, мониторинг, опция ПО Аванпост) и управления виртуализацией vControl размещаются в виртуальных машинах на продуктивных вычислительных узлах комплекса, сокращая объем доступных заказчику ресурсов. Схема данного исполнения представлена на рисунке 10.

Для ПАК с малым количеством узлов роли коммутаторов внешнего и внутреннего взаимодействия могут быть объединены в одной паре сетевых узлов.

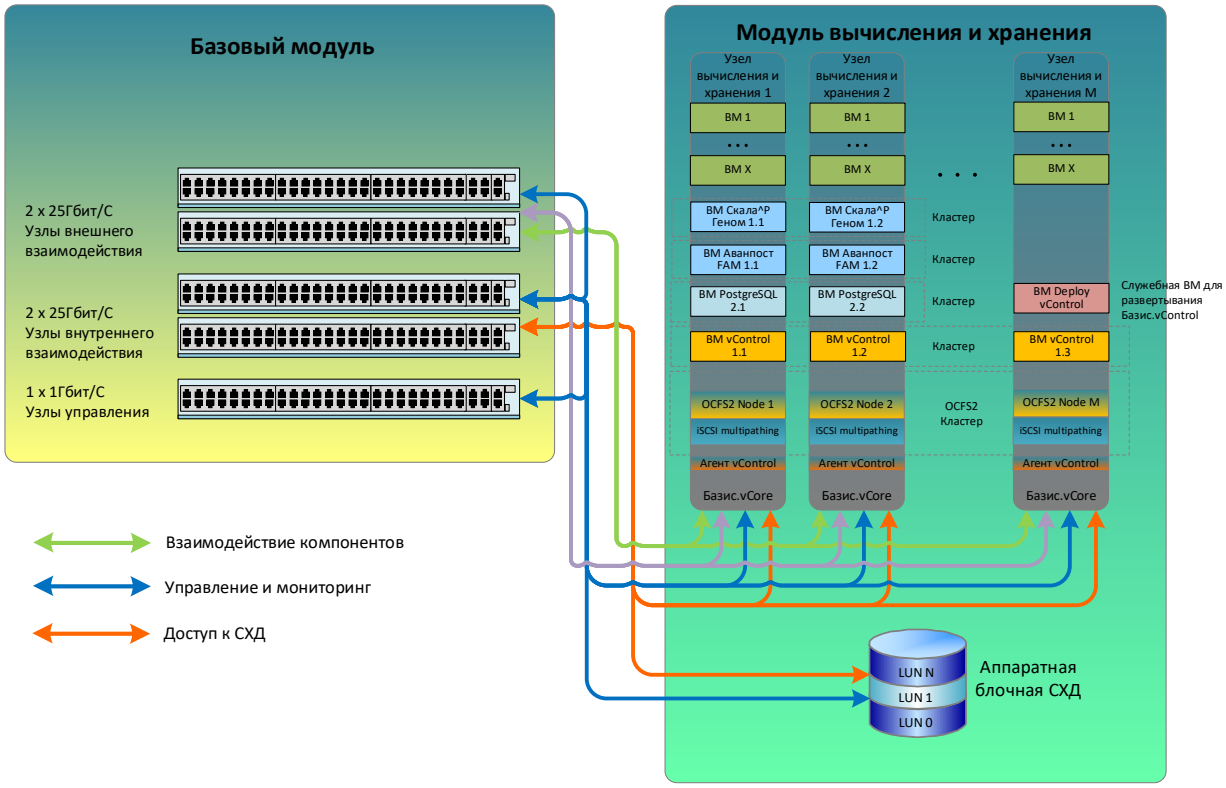

Рисунок 10. Схема исполнения 7 (минимальное исполнение с аппаратной СХД)

8.2.8. Исполнение 8 (минимальное исполнение ФСТЭК с аппаратной СХД)

Минимальное исполнение **Машины** с классической аппаратной СХД и с кластером Базис.Virtual Security для возможности аттестации ПАК со стороны ФСТЭК. Схема исполнения представлена на рисунке 11.

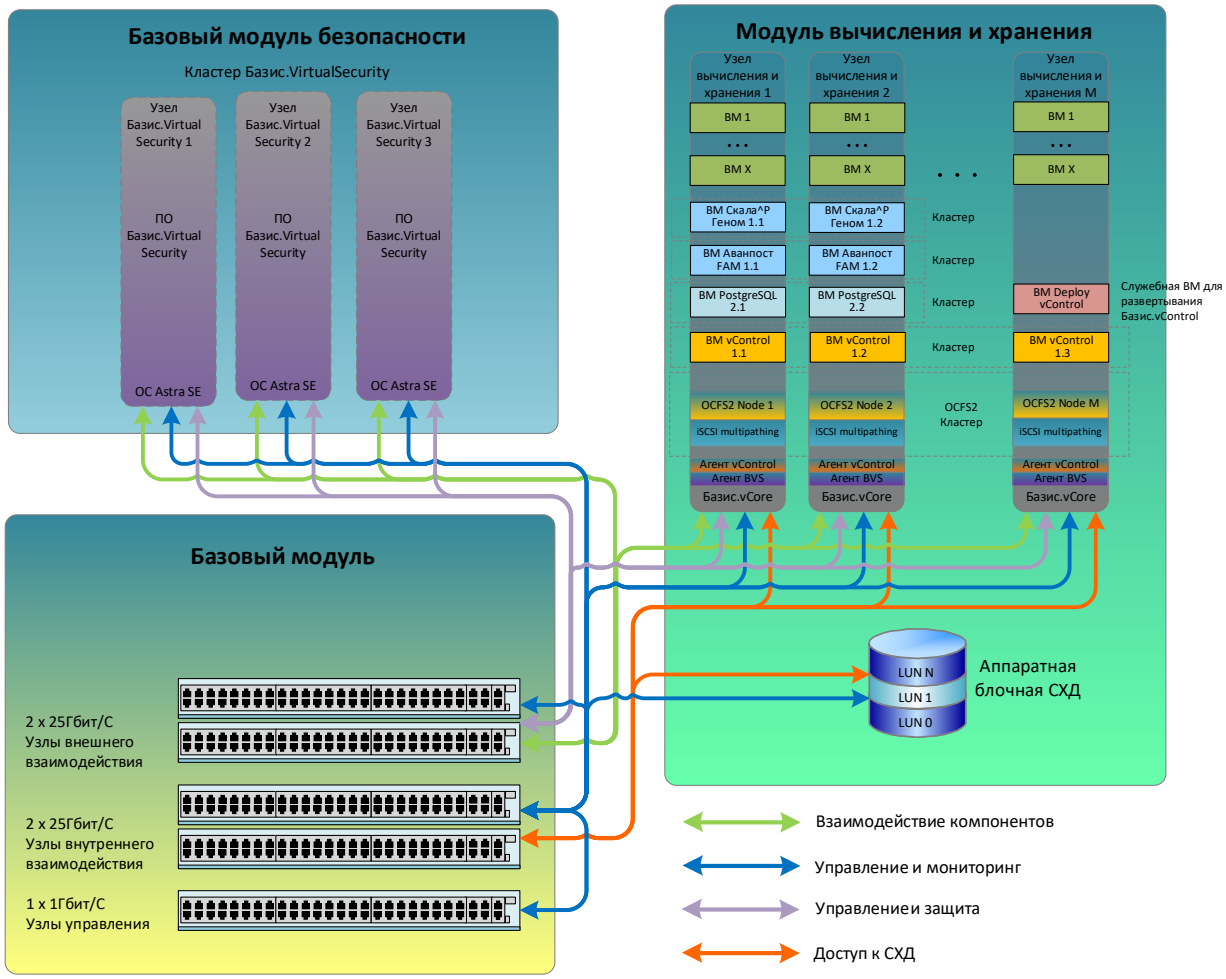

Рисунок 11. Схема исполнения 8 (минимальное исполнение ФСТЭК с аппаратной СХД)

8.2.9. Исполнение 9 (максимальное исполнение ФСТЭК с аппаратной СХД)

Максимальное исполнение **Машины Скала^р МДИ.В** с классической аппаратной СХД, с выделенным кластером управления из трех вычислительных узлов, с выделенными служебными узлами, с кластером Базис.Virtual Security для возможности аттестации ПАК со стороны ФСТЭК. Схема исполнения представлена на рисунке 12.

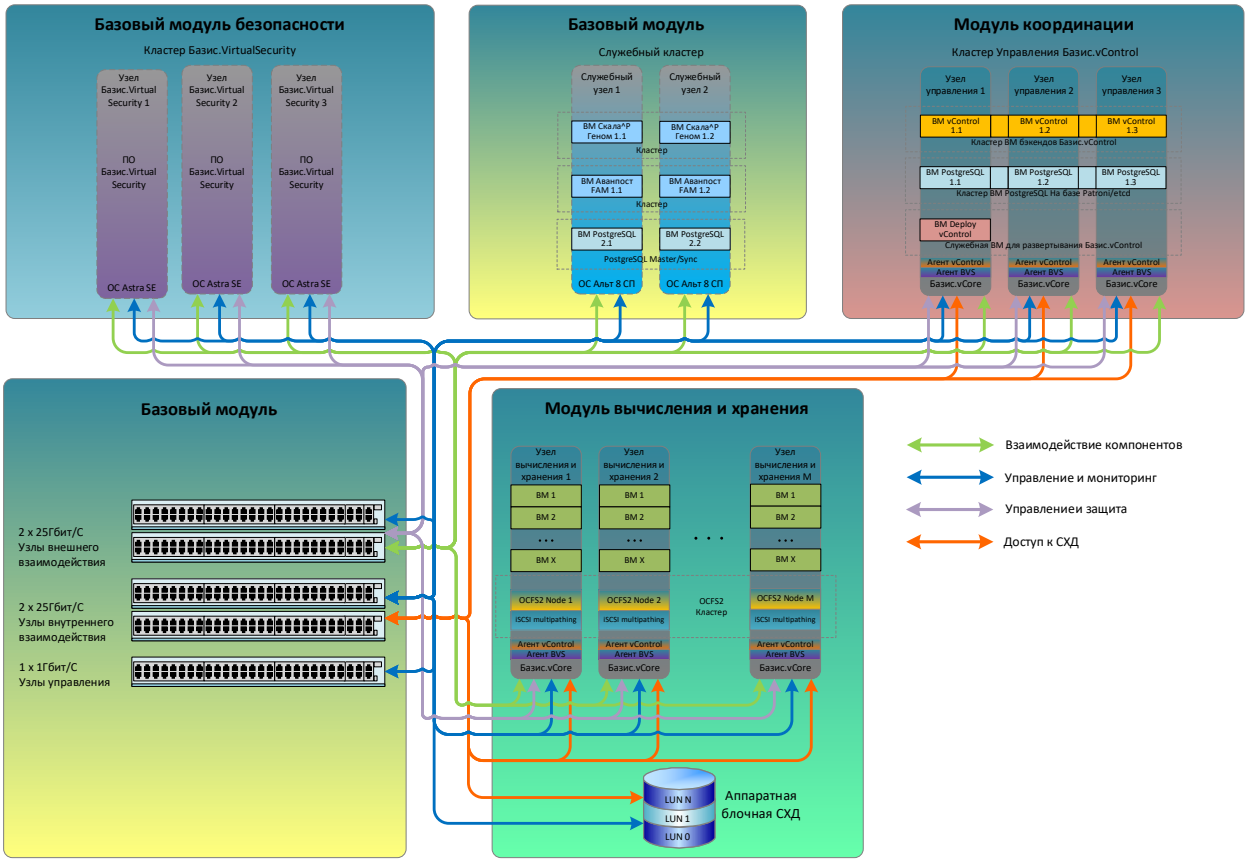

Рисунок 12. Схема исполнения 9 (максимальное исполнение ФСТЭК с аппаратной СХД)

8.3 Кластер управления

Кластер управления – главный логический компонент виртуальной инфраструктуры, развернутый на четырех (в случае гиперконвергентного кластера) или трех (в случае кластера с аппаратной СХД) узлах управления – контроллерах. В кластере управления запускается все программное обеспечение, отвечающее за работоспособность платформы, внутренние служебные базы данных, интерфейсы управления, включая GUI и REST API, которое управляют всеми вычислительными узлами кластера.

Кластер управления подсистемы управления представляет собой выделенный кластер высокой доступности, построенный на вычислительных узлах (узлах управления) с гипервизором Базис.vCore и общей разделяемой системой хранения данных (аппаратной или гиперконвергентной программной) под единым управлением программного продукта Базис.vControl.

Задача кластера управления – обеспечить независимую от продуктивной среды собственную отказоустойчивую среду виртуализации для размещения на своих вычислительных мощностях служебных инфраструктурных виртуальных машин в составе:

- кластер из 3-х виртуальных машин с ПО Базис.vControl для управления всей виртуальной инфраструктурой ПАК;
- кластер из 3-х виртуальных машин с ПО СУБД PostgreSQL (кластер собран на базе стека кластеризации Patroni/etcd) – служебная база данных, необходимая для работы ПО Базис.vControl.

8.4 Кластер вычисления

Кластер вычисления в гиперконвергентном исполнении обеспечивает вычислительными ресурсами и ресурсами хранения среду виртуализации (центральный процессор, оперативная память, дисковые ресурсы). На вычислительных узлах кластера работает непосредственно гипервизор 1-го типа Базис.vCore, который управляет разделяемым доступом виртуальных машин к вычислительным ресурсам вычислительных узлов.

При этом вычислительные узлы могут образовывать единый кластер вычисления на базе одного или нескольких разделяемых систем хранения данных (аппаратных или программных), либо набор независимых кластеров вычисления. Общее число узлов вычисления, входящих в состав **Машины**, может достигать 200 узлов, предоставляющих вычислительные ресурсы для виртуальных машин в рамках разворачиваемой виртуальной инфраструктуры, однако, в один вычислительный кластер, определяющий максимально возможный домен обеспечения высокой доступности для виртуальных машин, рекомендуется включать не более 20-28 вычислительных узлов.

8.5 Кластер BVS

Кластер BVS – опциональный кластер, выполняет функции авторизации, обеспечения безопасности системы управления и системы виртуализации. Предназначен для использования в государственных информационных системах до 1 класса защищенности включительно, в информационных системах персональных данных до 1 уровня защищенности включительно, в автоматизированных системах до класса 1Г включительно.

Кластер состоит из трех вычислительных узлов с ОС Astra Linux 1.8 на базе программного продукта Базис.Virtual Security (BVS).

Функции Базис.Virtual Security подробно рассмотрены в разделе 9.4.

8.6 Платформа служебной виртуализации

Платформа служебной виртуализации обеспечивает среду для исполнения следующих виртуальных машин:

- виртуальная машина с установленной программной платформой **Скала^р Геном** – ПО, предназначенное для управления жизненным циклом **Машины**, мониторинга основных аппаратных и программных компонентов ПАК, оповещение при сбоях, построение отчетов по собираемым данным. Программная платформа **Скала^р Геном** детально рассматривается в разделе 9.6;
- виртуальная машина с установленным ПО Аванпост FAM – используется как единая платформа аутентификации и авторизации к сервисам управления, предоставляемым ПО **Скала^р Геном**. ПО Аванпост FAM может отсутствовать, если есть его установка или аналог в инфраструктуре заказчика.

Данные виртуальные машины могут быть размещены на своих выделенных **служебных узлах – узлах мониторинга и регистрации** (см. раздел 10.2).

Для экономичных или малых конфигураций ПАК данные ВМ могут быть размещены на общих ресурсных узлах, без выделенных служебных узлов для них.

8.7 Сетевое взаимодействие

Схема сетей **Машины** представлена на рисунке 13. Сетевое взаимодействие обеспечивается двумя или тремя наборами сетевых узлов для организации сетей под следующие задачи:

- **сети внутреннего взаимодействия** – интерконнекта – служебные немаршрутизируемые сети для трафика репликации и доступа к программно-определяемому хранилищу, трафика доступа к аппаратной СХД (роль сети хранения), трафика репликации служебных узлов (назначение данных узлов будет раскрыто ниже);
- **сети внешнего взаимодействия** (сети доступа) – множество сетей продуктивных виртуальных машин. Подключения к внешней сети (uplink) осуществляются с помощью сетевых узлов, входящий в комплект **Машины**;
- **сети управления** — выделенная сеть доступа к IPMI-контроллерам вычислительных узлов из состава **Машины** и сеть доступа к консоли управления гипервизора по протоколу SSH.

Типовые конфигурации сетевых узлов — 2+2+2+1, 2+2+1 или 2+1 сетевых узла.

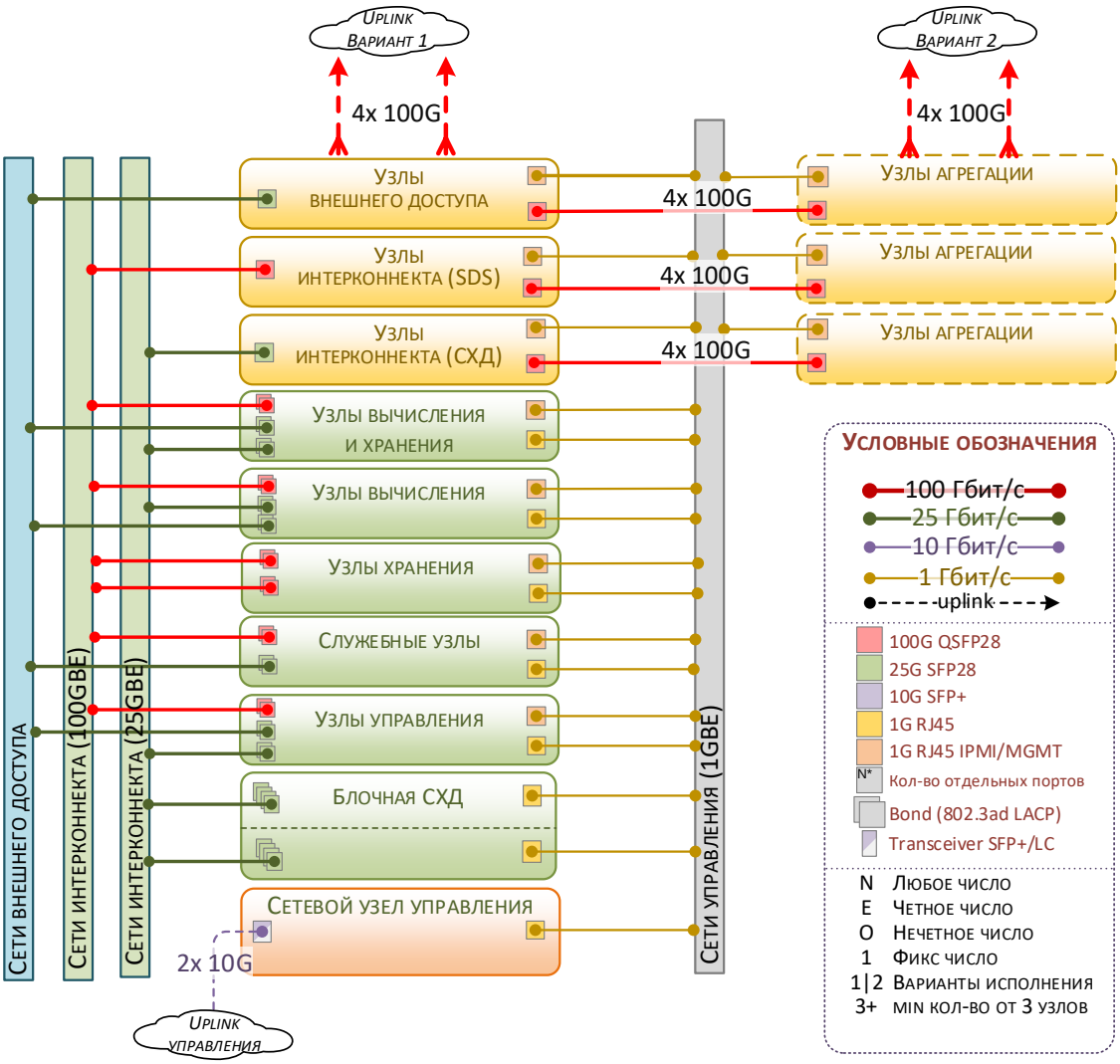

Рисунок 13. Физическая организация сетей в Машине Скала^р МДИ.В

Конфигурация 2+2+1

В этой конфигурации первая пара сетевых узлов обеспечивает сети внешнего доступа, вторая пара — сети интерконнекта для узлов хранения (доступ к СХД и/или организация программно-определяемого хранилища), и один сетевой узел обеспечивает сеть для контролеров управления IPMI и сеть SSH гипервизоров.

В последнем случае, сетевой узел не дублирован в силу невозможности подключения контролеров IPMI более чем одним сетевым интерфейсом.

Конфигурация 2+2+2+1

В этой конфигурации первая пара сетевых узлов обеспечивает сети внешнего доступа, вторая пара — сети интерконнекта для программно-определяемого хранилища, третья пара сетевых узлов – сеть хранения СХД. И один сетевой узел обеспечивает сеть для контролеров управления IPMI и сеть SSH гипервизоров. Используется в многокластерных смешанных конфигурациях.

Конфигурация 2+1

В этой конфигурации одна пара сетевых узлов обеспечивает и сети интерконнекта программно-определяемого хранилища, и сети внешнего доступа, и сети хранения СХД. Разделение сетей достигается за счет их логической изоляции через VLAN. Сеть управления IPMI реализуется так же, как в первом варианте, на одном выделенном сетевом узле.

Выбор того, какую конфигурацию использовать, зависит от сайзинга, в частности, ожидаемой нагрузки на сети внешнего доступа и интерконнекта.

Кроме того, в зависимости от ряда факторов могут быть выбраны сетевые узлы с портами 25 или 100 Гбит/с. Сетевой узел для сети управления IPMI всегда с портами 1 Гбит/с.

Агрегация

В случае нехватки портовой емкости коммутаторов внешнего доступа или коммутаторов интерконнекта в поставке **Машины** на каждую высокоскоростную сеть используется пара агрегирующих коммутаторов 100 Гбит/с с использованием технологий агрегации VXLAN или MLAG. Сеть управления не агрегируется, но каскадируется.

9. Программные компоненты

Программная часть **Машины Скала^р МДИ.В** реализована с помощью компонентов, перечисленных ниже (Таблица 2).

Таблица 2. Программные компоненты Машины серверной виртуализации Скала^р МДИ.В

Наименование ПО	Назначение
ПО Базис.vControl	Сервер управления средой виртуализации. Обеспечивает единый графический интерфейс управления виртуальной инфраструктурой **Машины**. Управляет гипервизорами, кластерами высокой доступности, виртуальными машинами, выполняет балансировку нагрузки, и т.д.
ПО Базис.vCore	Гипервизор 1-го типа. Обеспечивает разделение вычислительных ресурсов вычислительного узла между виртуальными машинами.
ПО Базис.uStor	Гиперконвергентное или выделенное в отдельные узлы хранения программно-определяемое хранилище (SDS). Обеспечивает распределенный дисковый массив с гибкой конфигурацией производительности и резервирования данных.
ПО Базис.Virtual Security – для версии с сертификатом ФСТЭК	ПО защиты виртуальной среды со встроенными средствами защиты от несанкционированного доступа и контроля целостности виртуальной среды.
ПО **Скала^р Геном**	ПО, используемое для управления жизненным циклом **Машины**, мониторинга аппаратных и программных компонентов, оповещения при сбоях, построение отчетов по собираемым данным.
ПО Аванпост FAM (Federated Access Manager) – опция	ПО используется как единая платформа аутентификации и авторизации к сервисам управления, предоставляемым ПО **Скала^р Геном**
ОС Альт СП (Альт 8 СП)	Операционная система со встроенными программными средствами защиты информации, которая сертифицирована ФСТЭК России, используется как базовая ОС на всех сервисных и служебных виртуальных машинах в ПАК.
ОС Астра SE	Операционная система со встроенными программными средствами защиты информации, которая сертифицирована ФСТЭК России, используется как базовая ОС на узлах с Базис.Virtual Security.

ПО PostgreSQL / ПО Postgres Pro Certified / ПО Jatoba Commercial / ПО Jatoba Certified	Объектно-реляционная система управления базами данных, работающая на базе языка SQL, используется в качестве служебной СУБД в подсистеме виртуализации. СУБД PostgreSQL с открытым кодом не может применяться с защищенными версиями ПО.

9.1 Размещение программных компонентов

На рисунке 14 показана максимальная архитектура **Машины серверной виртуализации Скала^р МДИ.В** с размещением программных компонентов на узлах в составе **Машины**. В максимальной архитектуре используется дезагрегированная конфигурация физических вычислительных узлов, когда каждый из программных компонентов размещается на своих собственных узлах. Такая конфигурация является оптимальной с точки зрения вычислительной производительности, минимизации взаимного влияния компонентов **Машины** и возможности дальнейшего будущего масштабирования. Однако, для небольших инсталляций (обычно, до 5-10 вычислительных узлов) возможно размещение программных компонентов на общих вычислительных узлах. Так, виртуальные машины с ПО **Скала^р Геном**, виртуальная машиной с ПО Аванпост FAM и виртуальные машины с программными компонентами ПО Базис.vControl могут быть размещены на общих вычислительных узлах в едином отказоустойчивом кластере, где будут размещаться продуктивные виртуальные машины. Таким образом, из типовой архитектуры остаются только два типа выделенных узлов: узлы кластера для ПО Базис.Virtual Security и непосредственно вычислительные узлы с гипервизором под продуктивную нагрузку с развернутым на них же распределенным программно-определяемым хранилищем Базис.uStor.

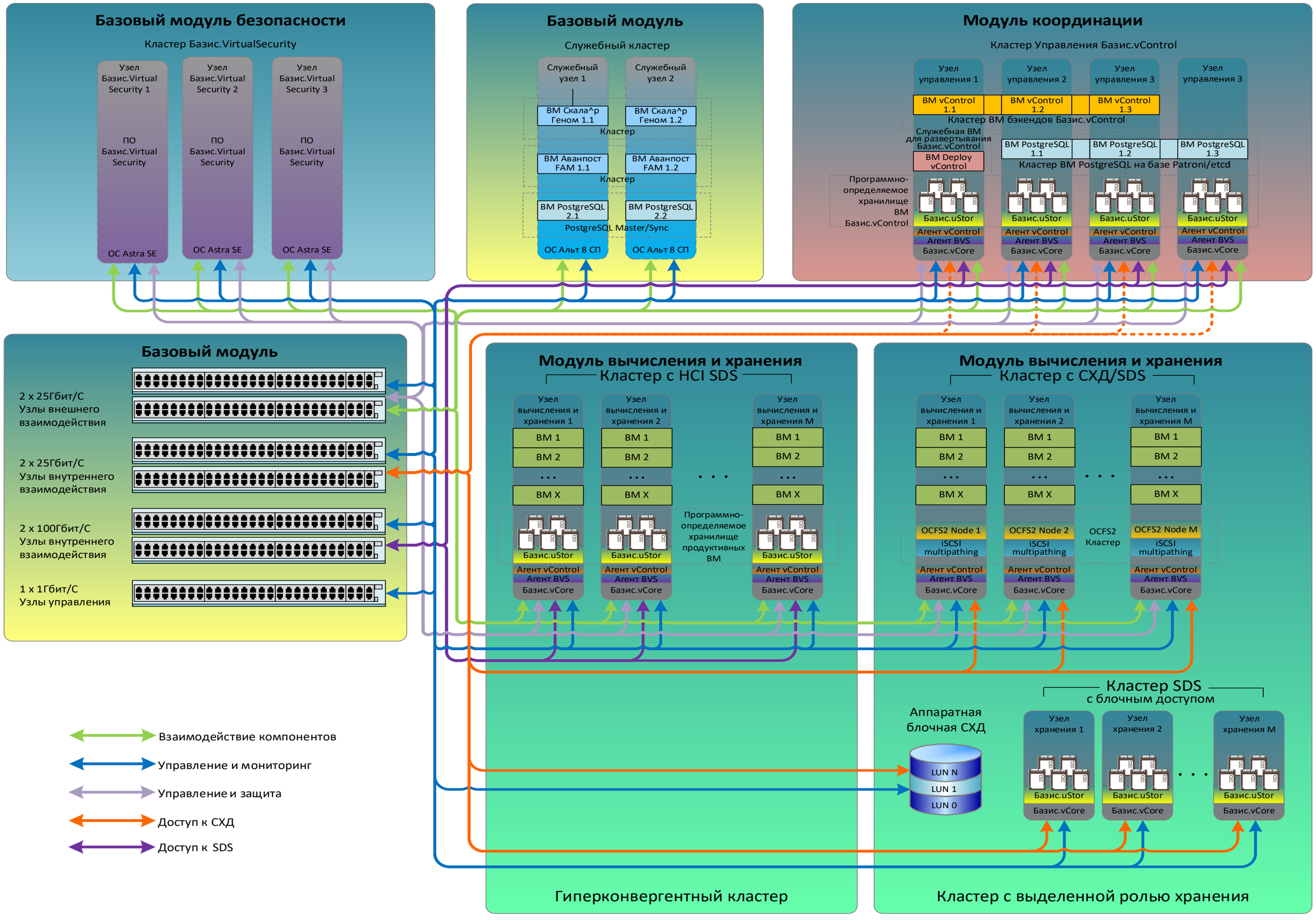

Рисунок 14. Узлы и программные компоненты в максимальной архитектуре Машины серверной виртуализации Скала^р МДИ.В

9.2 Базис.vControl

ПО Базис.vControl — это гибкая система управления и мониторинга среды виртуализации, в первую очередь, адаптированная и оптимизированная для взаимодействия с гипервизорами собственной разработки Базис.vCore.

ПО Базис.vControl позволяет выполнять следующие (административные) функции:

- Заводить учетные записи пользователей, управлять доступом пользователей к системе и ее объектам;
- Ограничивать доступные пользователям ресурсы с помощью пулов ресурсов;
- Управлять жизненным циклом виртуальных сред:
  - создавать новые виртуальные среды;
  - настраивать конфигурацию виртуальной среды и устанавливать дополнительное ПО в гостевую систему виртуальной среды;
  - управлять состоянием работы виртуальной среды: запускать, перезагружать, приостанавливать и выключать;
  - управлять правами доступа к объектам виртуальной среды для пользователей или групп пользователей;
  - создавать и управлять резервными копиями и снимками виртуальных сред;
- Осуществлять перемещение виртуальных сред между узлами виртуализации (миграцию);
- Группировать виртуальные среды в логическую структуру для упорядочивания работы и последующего делегирования управления;
- Создавать кластеры и управлять ими:
  - регулировать объем доступных ресурсов в кластере путем изменения его конфигурации;
  - отслеживать суммарные значения потребления ресурсов и состояние виртуальных сред, созданных в рамках кластера;
- Подключать хосты виртуализации и управлять ими:
  - предоставлять физические ресурсы (дисковое пространство, RAM, CPU) для развертывания и работы виртуальных сред;
  - просматривать информацию об использовании ресурсов;
  - изменять сетевую конфигурацию;
- Использовать общее хранилище шаблонов виртуальных сред и ISO-образов на всех кластерах, подключенных к системе управления.

ПО Базис.vControl состоит из следующих программных компонентов (см. Рисунок 15):

- **Служебная база данных** на основе СУБД PostgreSQL – для хранения информации о виртуальных ресурсах (виртуальные машины, сети), пользователях, группах, ролях, а также информации о вычислительных узлах, входящих в кластер, событиях, алертах и т.д. Программные пакеты Patroni и etcd, реализующие механизм высокой доступности кластерной СУБД;
- **Бэкенд vControl** – основное приложение, реализующее управление платформой Базис.vControl. Служит для обеспечения REST API для Фронтенда Базис.vControl, взаимодействует с агентами, выполняет периодические задачи;

- **Агент vControl** – запускается на управляемых узлах вычисления (физических вычислительных узлах). Управляет гипервизором, запускает стандартные Linux-команды, а также осуществляет мониторинг состояния вычислительного узла и гипервизора;
- **Менеджер агентов** – осуществляет взаимодействие с агентами vControl, запущенными на физических серверах с установленным гипервизором vCore. Двухсторонний протокол взаимодействия между Менеджером агентов и агентами управления, установленными на хостах, построен на базе библиотеки ZeroMQ;
- **Хранилище метрик** на основе БД ClickHouse – выполняет хранение значений метрик для хостов и ВМ. Метрики напрямую отправляют агенты, установленные на хостах. ClickHouse Keeper реализует механизм высокой доступности кластерной СУБД;
- **Кэш-хранилище** на основе БД Redis – реализует КЭШ для хранения очереди команд в системе управления, а также выполняет хранение пользовательских сессий и скриншоты ВМ. Служба Redis Sentinel реализует механизм высокой доступности кластерной СУБД;
- **Websocket Server** – обеспечивает двухстороннюю связь между Бэкендом и Фронтендом Базис.vControl. После авторизации пользователя Фронтенд Базис.vControl устанавливает соединение с WebSocket Server, чтобы получать сообщения о всех изменениях на Бэкенде Базис.vControl;
- **Фронтенд vControl** – реализует WebUI Графический интерфейс, который запускается в браузере пользователя.

Все программные компоненты ПО Базис.vControl разворачиваются в отказоустойчивой кластерной 4-х узловой конфигурации (3-х узловой для варианта с аппаратной СХД) на виртуальных машинах, размещенных в кластере Управления на узлах управления, либо на вычислительных узлах продуктивной нагрузки (в зависимости от сайзинга и «размера» **Машины**). Агенты ПО Базис.vControl устанавливаются на всех вычислительных узлах с гипервизорами.

Вычислительные узлы кластера управления обеспечивают вычислительные ресурсы и ресурсы хранения для служебных виртуальных машин. Кластер управления продолжит работу даже после выхода из строя до двух вычислительных узлов. Для обеспечения высокой доступности в решении реализована соответствующая функциональность — кластер высокой доступности на уровне гипервизора. В качестве основы он использует собственную разделяемую гиперконвергентную программную систему хранения данных или общую для ПАК аппаратную СХД.

Для всех служебных виртуальных машин в кластере управления включается функция «высокой доступности». В этом случае соответствующие компоненты в каждом гипервизоре начинают отслеживать доступность каждого вычислительного узла в кластере и вести учет того, какие виртуальные машины работают на каждом вычислительном узле. В случае отказа вычислительного узла, работавшие на нем виртуальные машины также «упадут», но будут перезапущены на прочих вычислительных узлах **Машины**, минимизировав тем самым время простоя.

В гипервизоре Базис.vCore работает отдельный сервис, обеспечивающий высокую доступность для виртуальных машин – Fenix HA. Эти сервисы, запущенные на каждом из вычислительных узлов **Машины**, проверяют доступность друг друга по SNMP, тайм-аут недоступности вычислительного узла — минута. При выборе вычислительного узла для перезапуска «упавшей» виртуальной машины реализована логика выбора наименее загруженного. Для проверки соединения Агент vControl с равным интервалом посылает heartbeat-сообщения в сторону Менеджера агентов. При обрыве соединения Агент пытается пересоздать соединение. При пропуске нескольких сообщений подряд Менеджер агентов

определяет, что Агент недоступен, устанавливает статус недоступности сервера, а также создает соответствующее уведомление.

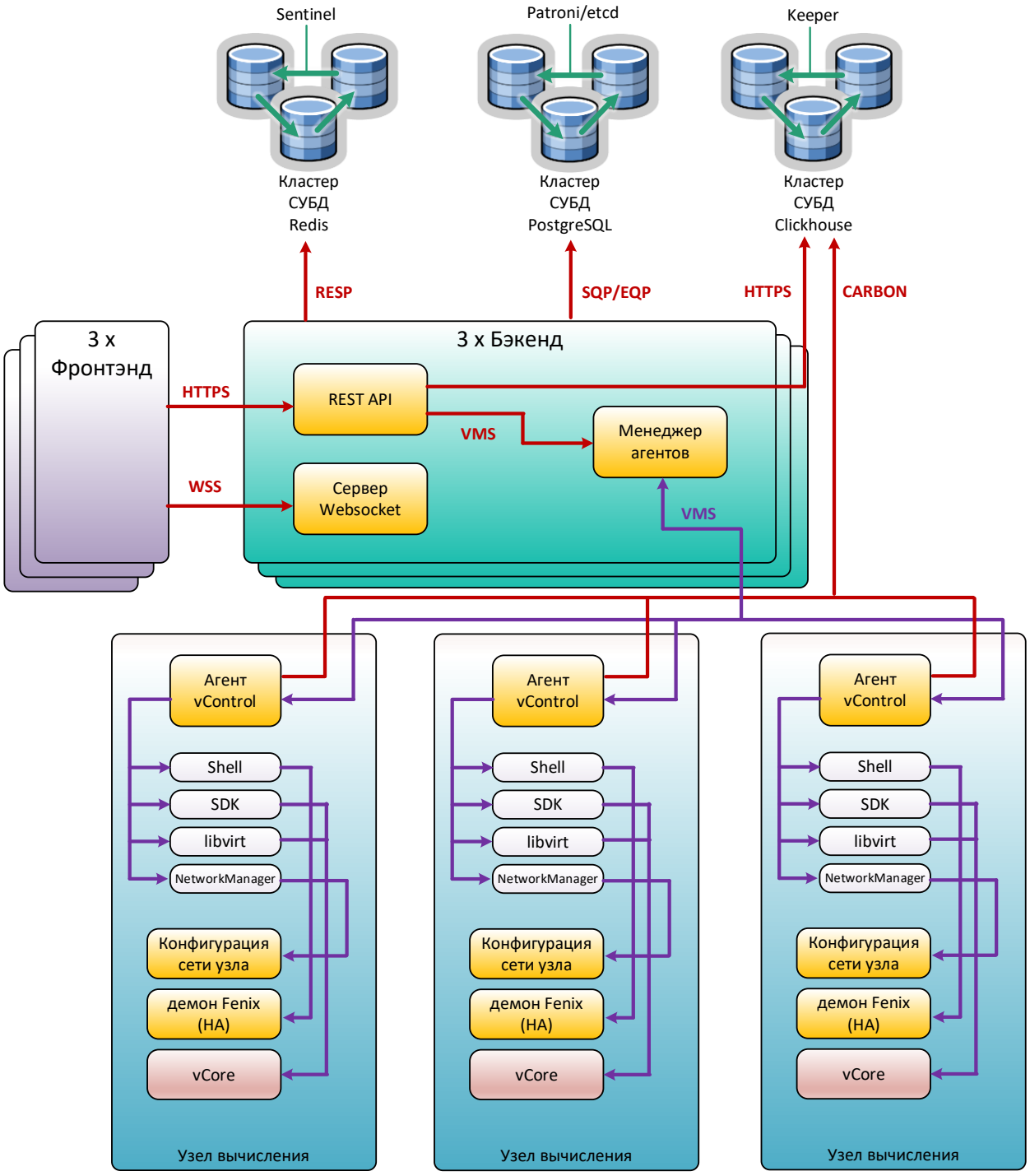

Рисунок 15. Схема взаимодействия программных компонентов ПО Базис.vControl

Агент vControl использует для управления виртуальной инфраструктурой следующие интерфейсы:

- **SDK виртуализации** — основной интерфейс взаимодействия с гипервизором vCore;
- **libvirt** — интерфейс библиотеки используется для работы с пробросом физических устройств в ВС и для взаимодействия с QEMU-агентом;

- **shell** — командный интерфейс управления (команды исполняются в консоли сервера);
- **NetworkManager** — управление сетевой конфигурацией хоста.

Варианты установки системы управления виртуализацией

ПО Базис.vControl для управления виртуальной инфраструктурой комплекса **Скала^р МДИ.В** можно развернуть в двух разных конфигурациях:

1. Все служебные роли объединяются на одной виртуальной машине. Запускаются несколько экземпляров таких виртуальных машин (обычно три) в кластерном исполнении;
2. Разные служебные роли разносятся по разным виртуальным машинам с дублированием программных компонентов (дезагрегацией и кластеризацией).

Для компактных сборок ПАК используется конфигурация первого варианта, с объединением ролей Базис.vControl в единой виртуальной машине и созданием кластера таких виртуальных машин высокой доступности (HA) на узлах вычислений и хранения, совместно с полезной нагрузкой заказчика.

Второй вариант (с дезагрегацией компонентов Базис.vControl) обеспечивает и более высокую доступность управления виртуализацией, и выдерживает более высокие нагрузки, что целесообразно для инфраструктур с повышенными требованиями к доступности, и/или с большим числом вычислительных узлов и виртуальных машин. В этих случаях, под системы управления выделяется свой отдельный выделенный кластер с собственным программно-определяемым хранилищем. Обычно, это становится целесообразно, когда система управления начинает потреблять ресурсы, сопоставимые с ресурсами 2-3 вычислительных узлов виртуализации. При этом, при дезагрегации ролей компонентов Базис.vControl каждая роль, при необходимости, может быть кластеризована на двух или более виртуальных машинах. Такой подход дает больше возможностей для точечного масштабирования конкретных программных компонентов, испытывающих потребность в вычислительных ресурсах.

Защита работы сервера управления (виртуальной машины с ролью Базис.vControl) от аппаратного отказа обеспечивается функцией высокой доступности (high availability, HA), работа которой не зависит от работоспособности самого ПО Базис.vControl. Даже если ВМ сервера управления окажется на вычислительном узле, который откажет из-за какого-либо аппаратного сбоя, функция HA отработает и восстановит работоспособность виртуальной машины на одном из доступных узлов вычислительного кластера. Обычно период недоступности в таком сценарии составляет 1–5 минут (в зависимости от конфигурации комплекса **Скала^р МДИ.В** и нагрузки на него).

Защитой от программных отказов самого сервера управления также является его периодическое резервное копирование.

В обоих конфигурация ПО управления Базис.vControl может управлять несколькими вычислительными кластерами **Скала^р МДИ.В** из единой консоли.

Такая реализация позволяет настраивать единые для всех вычислительных кластеров **Скала^р МДИ.В** хранилища шаблонов виртуальных машин, а также обеспечивать клонирование/перемещение виртуальных машин между отдельными вычислительными кластерами.

9.3 Базис.vCore

ПО Базис.vCore – это разработанный компанией Базис гипервизор 1-го типа, в основе которого лежит широко используемое решение с открытым исходным кодом KVM (Kernel-based Virtual Machine), используемое совместно с эмулятором аппаратного обеспечения QEMU (Quick Emulator) и библиотекой программной интеграции функций виртуализации Libvirt (Virtualization Library), он обеспечивает все необходимые для виртуализации функции для платформ x86-x64 и устанавливается на каждом вычислительном узле.

Основные свойства гипервизора:

- поддержка всех основных гостевых операционных систем, включая российские;
- полноценное управление жизненным циклом виртуальных машин;
- добавление устройств к виртуальной среде (машине) в процессе ее работы: ЦПУ, память, диски, сетевые интерфейсы и др.;
- динамическое перераспределение памяти между виртуальными средами для увеличения физически доступной памяти (за счет освобождения неиспользуемой);
- резервное копирование виртуальных сред;
- имеет встроенный конфигуратор гипервизора, предоставляющий графический интерфейс администратору для управления настройками подсистем гипервизора;
- реализует сетевую подсистему, обеспечивающую функции сетевого обмена в среде виртуализации, сегментацию сетевого трафика как на уровнях вычислительных узлов, так и на уровнях виртуальных машин.

Гипервизор Базис.vCore позволяет создавать виртуальные машины со следующими характеристиками на виртуальную машину (приведенные значения протестированы и поддерживаются производителем, технически максимальные конфигурации могут быть больше):

- количество виртуальных ядер — до 64 шт.;
- объем оперативной памяти — до 1 Тбайт;
- объем диска — до 16 Тбайт.

Гипервизор Базис.vCore состоит из следующих программных компонентов:

- **ядро гипервизора (KVM)**, реализующее основные функции виртуализации и обеспечивающее функционирование других подсистем гипервизора 1 типа;
- **подсистема виртуализации**, в которую входят библиотека для программной интеграции функций виртуализации Libvirt и эмулятор аппаратных платформ и виртуализированного оборудования QEMU;
- **конфигуратор гипервизора**, предоставляющий графический интерфейс администратору для управления настройками подсистем гипервизора;
- **сетевая подсистема**, обеспечивающая реализацию функций сетевого обмена в среде виртуализации, а также сегментации сетевого трафика на уровнях вычислительных узлов и виртуальных машин;
- **дисковая подсистема**, которая должна отвечать за взаимодействие средства виртуализации с локальными и внешними системами хранения данных;
- **virtual-security-agent** – агент сервера Базис.Virtual Security, который функционирует на всех узлах с гипервизором (вычислительных узлах), выполняет подсчет контрольных сумм, управление правилами фильтрации сетевого трафика и т.д.;
- **vcore-agent (vcontrol-agent)** – компонент управления, обеспечивающий выполнение транспортных функций в программном модуле и гарантирующий доставку

различных управляющих команд между программными компонентами серверной части и сервисом-агентом vCore, который функционирует на вычислительных узлах, а также предоставляет другие интерфейсы централизованного управления гипервизорами 1 типа «Базис.vCore»;

- **сервис-агент (JSagent)**, обеспечивающий информационный обмен между внутренними подсистемами гипервизора и программным модулем, который отвечает за централизованное управление виртуальной инфраструктурой.

9.4 Базис.Virtual Security

ПО Базис.Virtual Security является опциональным продуктом в составе ПАК, и необходимо, когда требуется аттестация ПАК как доверенного и применение защищенной версии управления виртуализацией.

Базис.Virtual Security – программный компонент ПАК со встроенными средствами защиты от несанкционированного доступа в виртуальной инфраструктуре и предназначено для использования в государственных информационных системах до 1 класса защищенности включительно, в информационных системах персональных данных до 1 уровня защищенности включительно, в автоматизированных системах до класса 1Г включительно.

В виртуальной инфраструктуре ПО Базис.Virtual Security реализует следующие функции безопасности:

- доверенная загрузка виртуальных машин;
- контроль целостности виртуальных машин;
- регистрация и логирование событий безопасности;
- управление доступом к виртуальной инфраструктуре;
- ограничение программной среды в средстве виртуализации;
- управление потоками информации в среде виртуализации;
- идентификация и аутентификация пользователей.

Ниже рассмотрим каждую из реализуемых функций безопасности более подробно.

Доверенная загрузка виртуальных машин

При создании ВМ и каждой ее загрузке Базис.Virtual Security осуществляет подсчет контрольных сумм файлов конфигурации виртуального оборудования ВМ. В случае несоответствия контрольных сумм эталонным значениям, Базис.Virtual Security блокирует запуск виртуальной машины и вносит соответствующую запись в журнал событий безопасности.

Контроль целостности

Базис.Virtual Security контролирует целостность в процессе загрузки и динамически в процессе функционирования средства виртуализации объектов контроля самостоятельно. При каждом запуске объекта контроля и динамически в процессе функционирования средства виртуализации осуществляется подсчет контрольных сумм файлов и сравнивается с эталонными значениями. Эталонные значения контрольных сумм сохраняются в базе данных Базис.Virtual Security. При несоответствии контрольных сумм эталонным значениям, Базис.Virtual Security блокирует запуск объекта контроля и вносит соответствующую запись в журнал событий безопасности.

Регистрация событий безопасности

В Базис.Virtual Security реализована подсистема регистрации событий, связанных с функционированием средств виртуализации, и оповещение администратора безопасности средства виртуализации о наступивших событиях безопасности.

Сбор, запись и хранение информации о событиях безопасности осуществляется с помощью встроенной подсистемы регистрации событий безопасности, согласно заданным администратором правилам. Состав регистрируемых событий соответствует ГОСТ Р 59548-2022 «Защита информации. Регистрация событий безопасности. Требования к регистрируемой информации».

Регистрация событий безопасности, связанных с функционированием средства виртуализации, фиксируется и сохраняется в журнале регистрации событий безопасности Базис.Virtual Security.

Журнал событий безопасности средства виртуализации доступен только для чтения. При исчерпании области памяти, отведенной под журнал событий безопасности средства виртуализации, осуществляется архивирование журнала с последующей очисткой высвобождаемой области памяти.

Управление доступом

В Базис.Virtual Security реализован ролевой метод управления доступом с четырьмя ролями пользователей: разработчик виртуальной машины, администратор безопасности средства виртуализации, администратор средства виртуализации, администратор виртуальной машины. Средствами Базис.Virtual Security обеспечивается:

- возможность формирования идентификатора, который однозначно идентифицирует пользователя;
- возможность удаления идентификатора пользователя;
- возможность блокировки идентификатора пользователя;
- создание пользователей;
- создание ролей.

Ограничение программной среды в средстве виртуализации

В Базис.Virtual Security при исполнении Изделия с базовой сертифицированной операционной системой, соответствующей требованиям по безопасности информации к средствам виртуализации, средствами сертифицированной ОС обеспечивается, в соответствии с требованиями по безопасности информации к средствам виртуализации 4-го класса защиты, следующие функции безопасности:

- контроль за запуском компонентов программного обеспечения, обеспечивающий выявление и блокировку запуска компонентов, не включенных в перечень (список) компонентов, разрешенных для запуска;
- выявление и блокировку запуска компонентов, целостность которого нарушена;
- блокировку запуска компонентов, не прошедших аутентификацию с использованием свидетельств подлинности модулей (в том числе цифровых сигнатур производителя или иных свидетельств подлинности модулей).

Управление потоками информации в среде виртуализации

В Базис.Virtual Security при исполнении Изделия с базовой сертифицированной операционной системой реализовано, в соответствии с требованиями по безопасности информации к средствам виртуализации 4-го класса защиты, следующие функции безопасности:

- Управление потоками информации между виртуальными машинами и информационными (автоматизированными) системами на канальном и сетевом уровнях самостоятельно или с применением сертифицированных средств управления потоками информации (коммутаторов, маршрутизаторов) и (или) межсетевых экранов, а также контроль взаимодействия виртуальных машин между собой;
- Централизованное управление правилами осуществляется администратором Базис.Virtual Security с использованием собственного сервис-агента, который взаимодействует с программной библиотекой управления средством виртуализации libvirt.

Идентификация и аутентификация пользователей

В Базис.Virtual Security реализован механизм идентификации и аутентификации пользователей в среде виртуализации с учетом требований разделов 4 – 7 ГОСТ Р 58833–2020 «Защита информации. Идентификация и аутентификация. Общие положения». Основной функционал в этой части:

- в случае неуспешной идентификации и аутентификации пользователей в средстве виртуализации их доступ блокируется;
- аутентификация пользователей осуществляется при предъявлении идентификатора и пароля пользователя;
- пароль пользователя для первичной аутентификации устанавливается администратором средства виртуализации или администратором безопасности;
- обеспечена возможность смены установленного администратором средства виртуализации пароля пользователя средства виртуализации после его первичной аутентификации;
- невозможность установления одинаковых идентификаторов и паролей для разных пользователей;
- при попытке ввода неправильного значения идентификатора или пароля пользователя выводиться сообщение с приглашением ввести правильный идентификатор и пароль еще раз;
- при исчерпании установленного максимального количества неуспешных попыток ввода неправильного пароля учетная запись пользователя средства виртуализации блокируется с возможностью разблокировки администратором средства виртуализации или с возможностью автоматической разблокировки по истечении временного интервала, устанавливаемого администратором средства виртуализации;
- защита пароля пользователя средства виртуализации обеспечивается при его вводе за счет отображения вводимых символов условными знаками;
- обеспечивается хранение аутентификационной информации пользователя средства виртуализации в защищенном формате или в защищенном хранилище;
- обеспечивается взаимная идентификация и аутентификация пользователей и средства виртуализации при удаленном доступе с использованием сетей связи общего пользования;

- пароль пользователя средства виртуализации содержит не менее 8 символов при алфавите пароля не менее 70 символов. Максимальное количество неуспешных попыток аутентификации (ввода неправильного пароля) до блокировки – 4.

ПО Базис.Virtual Security устанавливается **всегда** на три выделенных вычислительных узла (Узла BVS) – в режиме BareMetal, образуя собственный отказоустойчивый кластер на базе стека кластеризации Pacemaker/Corosync и состоит из следующих программных компонентов (Рисунок 16):

- **Nginx** – прокси-сервер, с помощью которого реализуется проксирование запросов до сервисов «Базис.Virtual Security» и реализуется поддержка TLS;
- **virtual-security-admin** – основной компонент, который реализовывает бизнес-логику, обеспечивает интеграцию между подсистемами программного модуля и предоставляет API;
- **virtual-security-web** – программный компонент, который предоставляет графический интерфейс администратору для централизованного управления программным модулем и различными функциями безопасности, которые данный программный модуль реализует;
- **virtual-security-auth** – сервис, который отвечает за реализацию функций безопасности, связанных с идентификацией, аутентификацией пользователей в средстве виртуализации и в других внешних системах с использованием протоколов OpenID Connect (OIDC), SAML 2.0, Kerberos, LDAP/Active Directory;
- **ldap-server** – LDAP-сервис, который обеспечивает реализацию протокола легковесного доступа к каталогам (LDAP, Lightweight Directory Access Protocol) и позволяет выполнять операции создания, хранения, изменения и удаления различных записей в указанных каталогах, в том числе управлять жизненным циклом учетных записей пользователей;
- **virtual-security-agent** – программный компонент, который функционирует в составе гипервизора 1типа «Базис.vCore» или в составе сертифицированной хостовой операционной системы, и выполняет функции безопасности, связанные с контролем жизненного цикла виртуальных машин, вычислительных сетей и узлов виртуализации, а также с контролем целостности и доверенной загрузкой;
- **virtual-security-proxy** – сервис проксирования, который отвечает за управление доступом к API-интерфейсам защищаемых информационных и автоматизированных систем;
- **vcore-broker** – это шина обмена сообщениями, которая выполняет транспортные функции в программном модуле и гарантирует доставку различных управляющих сигналов между программными компонентами серверной части и сервисом-агентом virtual-security-agent, который функционирует на вычислительных узлах;
- **basis-vbalancer** – сервис балансировки нагрузки, который реализует распределение запросов между нескольким узлами кластера управления программного модуля «Базис.Virtual Security»;
- **PostgreSQL** – объектно-реляционная система управления базами данных;
- **virtual-security-pg-pam** – сервис, который обеспечивает аутентификацию по токену сессии, выданному в соответствии со стандартом OpenID Connect, при удаленном или локальном доступе систем, обращающихся к СУБД Postgres от имени ранее аутентифицированных пользователей.

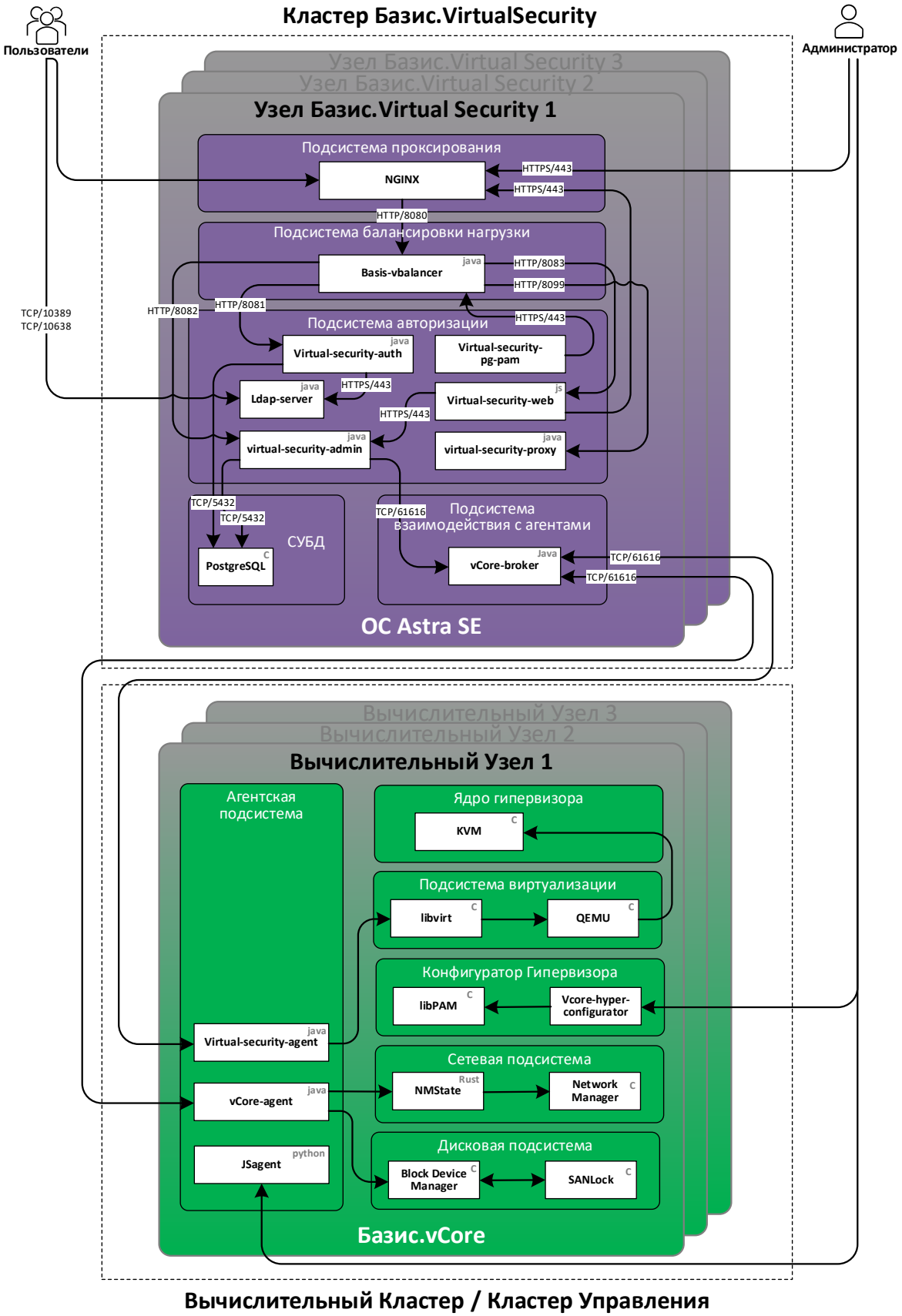

Рисунок 16. Схема взаимодействия программных компонентов ПО Базис.Virtual Security

9.5 Базис.uStor

Базис.uStor – это распределенная программно-определяемая система хранения данных с горизонтальным масштабированием и отказоустойчивой архитектурой без единой точки отказа, с ориентацией на высокую производительность. Возможны два варианты реализации Базис.uStor: гиперконвергентный с развертыванием системы хранения на узлах вычисления и конвергентный (в специальной расширенной версии ПО Базис.uStor) с развертыванием системы хранения на своих собственных выделенных узлах хранения. При этом в первом случае доступ к программному хранилищу осуществляется на уровне QEMU-драйвера модуля ядра гипервизора, а во втором случае доступ осуществляется по универсальному протоколу iSCSI как к блочному устройству. Все узлы кластера Базис.uStor (узлы хранения) абсолютно равнозначны в процессах обработки ввода-вывода данных, при этом на части из них дополнительно запускаются сервисы управления кластером (узлы управления и хранения).

Программные компоненты распределенного хранилища Базис.uStor устанавливаются на каждом из узлов хранения в ПАК совместно с гипервизором Базис.vCore в качестве хостовой операционной системы. Базис.uStor позволяет объединить все дисковые накопители, установленные в вычислительных узлах, в единое дисковое пространство. Оно используется любой из виртуальных машин **Скала^р МДИ.В** в среде виртуализации для хранения файлов образов своих дисковых устройств (в формате RAW - неформатированного логического раздела), а также может быть использовано внешними потребителями по протоколу блочного доступа iSCSI (для активации этого функционала требуется редакция ПО Базис.uStor Enterprise).

Для данных на хранилище доступны различные схемы резервирования, за счет чего обеспечивается высокая доступность данных и устойчивость к сбоям хранилища при единичных отказах вычислительных узлов и накопителей.

Основные функции и свойства программно-определяемого хранилища Базис.uStor:

- алгоритмы обеспечения избыточности данных: хранение двух и более реплик данных на дисковых накопителях разных вычислительных хостов ПАК **Скала^р МДИ.В** или хранение блоков четности/избыточности (Erasure Coding с использованием кодов Рида-Соломона). Немного упрощая, при использовании EC поддерживаются алгоритмы, логически похожие на RAID 1, RAID 6;
- настраиваемый домен отказа (место хранения одной из копий реплики или блока четности) хранилища: дисковый накопитель, вычислительный узел, группа вычислительных узлов;
- гибкие возможности модернизации и обслуживания узлов хранения без прерывания работы хранилища – легкость горизонтального и вертикального масштабирования хранилища без прерываний в работе;
- создание и презентация пулов хранения с различными схемами резервирования на одном наборе дисковых накопителей.

В продуктивных решениях хранилище Базис.uStor может быть развернуто минимум на 4 вычислительных узлах (в тестовых инсталляциях допустимо использование 3 вычислительных узлов), минимум три из которых также наделяются ролью управляющих узлов. Хранилище легко наращивается без каких-либо простоев в работе добавлением новых вычислительных узлов с дисковыми накопителями. При этом производительность хранилища растет линейно, с каждым новым узлом за счет параллелизма операций ввода-вывода.

Узлы хранилища могут выполнять одновременно роль непосредственно хранилищ данных и роль управляющих контроллеров. Например, 3–5 узлов в инсталляции могут быть выделены как управляющие (они же могут хранить данные), а дополнительные узлы (до нескольких десятков) могут добавляться уже только с ролью хранения для масштабирования

емкости хранилища и производительности. Все узлы с ролью хранения объединены по сети (в том числе, поддерживается высокоскоростной Fabric over RDMA (RoCEv2)), и вместе образуют единое отказоустойчивое хранилище.

Далее рассмотрим подробно каждую подсистему хранилища Базис.uStor и его программные компоненты.

Подсистема управления

Подсистема управления отвечает за конфигурирование и административное управление хранилищем Базис.uStor. Она предоставляет программные интерфейсы для администратора (графический Web UI и командную строку), а также внутренние сервисы для обработки команд и координации изменений конфигурации хранилища. В состав подсистемы управления входят следующие программные компоненты:

- **Nginx** - web-сервер и обратный (реверс) прокси, обслуживающий пользовательский графический интерфейс и API. Nginx принимает входящие HTTP(s) запросы от web-клиента или внешних систем и перенаправляет их на внутренние сервисы (Web UI и uStor API). Он обеспечивает единую точку доступа к UI/API, балансировку и, при необходимости, терминирование SSL-соединений;
- **uStor Web UI** – графический web-интерфейс управления хранилищем, реализованный в виде браузерного приложения, через которое администратор может выполнять все основные операции: создание и настройка пулов хранения, управление томами, просмотр состояния системы и т.д., упрощает эксплуатацию, предоставляя визуальные дашборды и формы вместо прямого обращения к командам. Web UI взаимодействует с API для выполнения действий и получения данных о состоянии хранилища;
- **uStor API** - программный интерфейс управления (REST API) для автоматизации и интеграции, предоставляет собой набор HTTP API-методов для всех функций управления (создание/удаление томов, настройка защиты данных, добавление узлов и пр.). Возможна интеграция с внешними системами для оркестрации (например, системы оркестрации облаков или резервного копирования могут создавать тома через API). API-сервис валидирует запросы, применяет бизнес-логику (например, проверка прав или корректности параметров) и передает задания на выполнение во внутренние сервисы uStor (Worker/Utils). При этом, все изменения конфигурации через API фиксируются в системе (в т.ч. в аудит-логе);
- **uStor CLI** - утилита командной строки для управления системой, которая позволяет администраторам выполнять команды управления кластером из терминала или скриптов. CLI вызывает соответствующие методы uStor Utils, тем самым достигается тот же эффект, что и через Web UI. CLI удобна для автоматизации (скрипты администрирования) и в ситуациях, когда нет доступа к web-интерфейсу;
- **uStor Utils** - вспомогательный служебный модуль (набор утилит и библиотек) для внутренних служебных нужд системы. Он содержит общий функционал, используемый различными сервисами (парсинг конфигураций, утилиты для работы с хранилищем, вспомогательные скрипты и т.д.). uStor Utils обеспечивает унифицированный подход к выполнению задач внутри кластера, чтобы компоненты Worker и Consumer могли повторно использовать готовые функции (например, подготовка дисков, сбор диагностической информации);
- **uStor Worker** - сервис-планировщик, отвечающий за обработку входящих управленческих задач. Когда администратор инициирует любую операцию (например, создание нового тома или изменение настроек пула хранения), API передает эту задачу на выполнение Worker'у. Worker размещает задачу во внутренней очереди (взаимодействуя с etcd) и координирует ее выполнение. Он разбивает сложные

операции на шаги, распределяет их между узлами (при необходимости) и контролирует ход выполнения. По сути, uStor Worker выступает как диспетчер задач, подготавливая задания для выполнения;

- **uStor Consumer** - сервис-фоновый, обеспечивающий обработку результатов выполнения команд на узлах.

Авторизация и аутентификация в Базис.uStor

Доступ к REST API в Базис.uStor осуществляется с обязательной проверкой пользователя и его прав. По умолчанию система включает встроенный механизм аутентификации, поддерживающий локальные учетные записи и базовые роли (например, администратор, оператор, наблюдатель). Однако в продуктивных средах и особенно в режиме безопасной эксплуатации рекомендуется интеграция с внешними системами централизованной аутентификации, такими как Active Directory (AD) или LDAP. Подключение к внешнему каталогу позволяет централизованно управлять доступом, применять существующие политики безопасности организации и использовать единый вход (SSO) для всех администраторов.

При использовании AD/LDAP REST API uStor проверяет учетные данные пользователя через подключенный каталог, а затем сопоставляет его группы или атрибуты с внутренними ролями системы (RBAC). Это обеспечивает гибкое разграничение прав. Все запросы к API проходят авторизацию: проверяется, имеет ли пользователь соответствующие разрешения на выполнение конкретного действия (создание тома, удаление узла, изменение настроек и т.д.). При этом каждое действие фиксируется в журнале аудита, что обеспечивает полную трассировку активности.

Такой подход позволяет Базис.uStor вписываться в требования корпоративной безопасности, обеспечивать единые политики доступа на уровне всей ИТ-инфраструктуры и защищать управление кластером от несанкционированных действий даже в условиях распределенной среды.

Подсистема управления тесно связана с подсистемой контроллера – она отправляет туда запросы на изменение состояния и получает оттуда информацию о текущем состоянии кластера для отображения. Также она взаимодействует с подсистемой мониторинга (например, Web UI может отображать метрики здоровья, собираемые uStor Health).

Подсистема контроллера

Подсистема контроллера отвечает за глобальное состояние кластера хранилища, консенсус между узлами, хранение метаданных и контроль над распределением данных. Она обеспечивает согласованность конфигурации во всем кластере и отслеживает работоспособность компонентов. В состав контроллерной подсистемы входят:

- **uStor etcd** - централизованное распределенное хранилище конфигурации и служебных данных кластера. uStor использует etcd как надежное key-value хранилище, в котором хранятся сведения о всех объектах кластера: зарегистрированные узлы и их роли, списки томов и их характеристик, привязка томов к носителям (карта размещения данных), параметры пулов (схемы резервирования) и др. Etcd обеспечивает согласованность этих данных между всеми контроллерами в кластере хранения за счет распределенного консенсуса (Raft) – это фундамент отсутствия единой точки отказа, так как конфигурационные данные реплицируются на нескольких узлах и доступны даже при выходе из строя одного из них. Компоненты Worker/Consumer через etcd ставят задачи и фиксируют изменения, Службы uStor Disk (OSD) регистрируют свое состояние и получают актуальные конфигурации, а сервисы мониторинга могут отслеживать изменения через механизмы watch и heartbeat;

- **uStor Monitor** - сервис мониторинга и управления кластером (не путать с подсистемой мониторинга состояния, о которой ниже). Этот компонент выполняет роль контроллера, следящего за топологией и целостностью данных. uStor Monitor отслеживает состояние узлов и сервисов: получает информацию об активных uStor Disk (OSD) на каждом узле, их здоровье и заполненность, фиксирует подключение/отключение узлов. При сбоях uStor Monitor инициирует процедуры самовосстановления – например, при отказе одного из дисков или целого узла, он определяет, какие данные утрачены или находятся в риске, и запускает процессы восстановления (rebuild) на других узлах. Он хранит критическую информацию в etcd (например, метаданные о том, какие копии данных, где находятся, кто в кластере «лидер» для определенной группы данных и т.п.).

Отказоустойчивость uStor Monitor обеспечивается за счет распределенной архитектуры: сервис разворачивается сразу на нескольких контроллерных узлах, при этом каждый экземпляр может полноценно выполнять функции мониторинга и координации. Между экземплярами происходит синхронизация состояния через etcd, а логика работы ориентирована на консенсусную модель, где отсутствует единая точка отказа. Даже при выходе из строя одного или нескольких управляющих узлов другие экземпляры uStor Monitor продолжают функционировать, сохраняя управляемость кластером и способность выполнять восстановительные процедуры. Такая архитектура гарантирует высокую доступность сервиса и непрерывный контроль над распределением и целостностью данных в любых сценариях деградации кластера. Таким образом, uStor Monitor по функциональности аналогичен классическому метадата-серверу/монитору в распределенных СХД, но превосходит его по устойчивости благодаря горизонтальному дублированию и тесной интеграции с отказоустойчивым хранилищем метаданных etcd.

- **База данных mind-audit**: подсистема журналирования действий и событий реализована как распределенная отказоустойчивая база данных, обеспечивающая сохранность и консистентность логов даже при сбоях отдельных узлов. Каждый критически важный запрос к системе (например, операции создания, удаления, изменения параметров, а также системные события) фиксируется в журнале аудита, который доступен для последующего анализа и аудита. Архитектура хранилища журнала построена таким образом, что кластер продолжает работу и сохраняет новые события даже при частичной недоступности управляющих узлов — механизм консенсуса и репликации обеспечивает устойчивость к сбоям и защиту от потери данных без потери целостности записи. Это гарантирует, что информация о действиях пользователей и внутреннем состоянии кластера всегда остается доступной и достоверной.

Подсистема контроллера работает в тесной связке с подсистемой хранения данных: в uStor etcd хранятся указания для uStor Disk какие данные, где держать, а uStor Monitor опирается на информацию от OSD для принятия решений. За счет децентрализованной природы uStor etcd и дублирования сервисов мониторинга на нескольких узлах, достигается отсутствие единой точки отказа, контроллерная логика продолжает работать даже при выходе из строя отдельного управляющего узла.

Подсистема хранения данных

Подсистема хранения данных отвечает за непосредственное хранение данных (образов дисков виртуальных машин) на дисках узлов хранения с обеспечением необходимого уровня резервирования и целостности. Основной (и единственный) компонент этой подсистемы – это uStor Disk (OSD), объектный сервер хранения данных.

- **uStor Disk (OSD)** - данный компонент запускается на каждом узле хранения в виде OSD-сервиса и управляет одним или несколькими локальными дисковыми устройствами. Он принимает от подсистемы блочного доступа команды на чтение/запись блоков данных и осуществляет эти операции на локальном носителе. uStor Disk

отвечает за хранение фрагментов данных (чанков) и их репликацию или кодирование. При записи нового блока данных OSD может сохранить его у себя и параллельно передать копии другим узлам (для схемы с репликацией) либо вычислить кодовые блоки четности для хранения на нескольких узлах (при использовании помехоустойчивого кодирования EC). Конкретная схема резервирования настраивается на уровне группы хранения (Пула) – **Базис.uStor** поддерживает как репликацию с произвольным числом копий, так и Erasure Coding по схеме K+N. Например, можно задать три копии данных либо кодирование 3+2 и т.п.

Каждый OSD следит за состоянием своих носителей. В случае отказа диска или узла uStor Monitor инициирует реконструкцию: другие OSD читают уцелевшие копии данных и создают недостающие копии на новых дисках. Базис.uStor умеет выполнять частичную реконструкцию – восстанавливать только поврежденные части групп данных, а не весь объем целиком, что в значительной степени ускоряет ребилд. После замены неисправного накопителя или добавления нового узла система автоматически ребалансирует данные – перераспределяет существующие фрагменты по кластеру хранения, чтобы задействовать новую емкость и выровнять нагрузку. Также реализована проактивная проверка целостности (scrubbing) – регулярная проверка данных на дисках, позволяющая обнаружить скрытые повреждения и восстановить их из резервных копий.

OSD взаимодействует с uStor etcd/Monitor для получения актуальной карты размещения фрагментов данных и информации о том, какие копии данных должны храниться на данном узле. Он периодически отправляет heartbeat сигналы или обновляет в uStor etcd свой статус (например, занятость, наличие ошибок) для контроллера. Таким образом, подсистема хранения представляет собой распределенный массив дисков, управляемых OSD-процессами, которые совместно хранят данные с заданной избыточностью и высокой надежностью.

Подсистема блочного доступа к данным

Подсистема блочного доступа реализует программный слой, предоставляющий клиентам доступ к распределенному хранилищу в виде блочных устройств. Она связывает приложения/сервера с физическими данными на службах фрагментов uStor Disk (OSD), выполняя роль «транспорта» для команд чтения/записи. В эту подсистему входят следующие программные компоненты:

- **uStor NBD** – компонент интеграции с Network Block Device (NBD), позволяющий на стороне клиента (вычислительном узле) подключить удаленный том uStor как локальное блочное устройство /dev/nbdX. Для этого он устанавливает сетевое соединение с кластером хранения uStor и передает блоковые команды ввода-вывода. По сути, uStor NBD – это клиентский драйвер/сервис, облегчающий использование нативного протокола хранилища uStor через стандартный интерфейс NBD. При этом, NBD-клиент может быть реализован как модуль ядра ОС или как пользовательская программа, устанавливаемая на вычислительные узлы, которые будут подключаться к хранилищу;
- **uStor Storage Daemon (ustord)** - представляет собой блочный проксирующий сервис, который с клиентской стороны реализует поддержку стандартных высокопроизводительных протоколов, таких как **vhost-user-blk** и **vhost-vduse**, предоставляя виртуальное блочное устройство в привычной для гипервизоров и виртуальных машин форме. Внутри кластера хранения uStor он преобразует поступающие команды в собственный распределенный протокол ustor, с помощью которого происходит взаимодействие с системой хранения. Такой подход позволяет обеспечить эффективный доступ к данным без необходимости вносить изменения в архитектуру или поведение пользовательского ПО — гипервизоры, драйверы и другие клиенты продолжают работать с виртуальными дисками через стандартные интерфейсы, в то время как ustord маршрутизирует операции чтения и записи к нужным

узлам и сервисам фрагментов OSD внутри кластера хранения. Это позволяет абстрагировать клиентские приложения от внутренней структуры программно определяемого хранилища и гарантировать масштабируемый, отказоустойчивый и производительный доступ к распределенному хранилищу uStor;

- **uStor Client Library** - библиотека клиентского доступа к хранилищу, реализующая собственный протокол ustor и логику взаимодействия с кластером хранения. Она инкапсулирует детали работы с uStor etcd и OSD, предоставляя высокоуровневый API для операций с томами (read/write/flush, и т.д.). uStor Client Library используется внутри uStor NBD, драйверов для гипервизоров (например, QEMU) и сервисов типа uStor iSCSI, чтобы упростить их реализацию. В библиотеке заложены алгоритмы выбора оптимального пути к данным: например, выбор ближайшего или наиболее свободного узла для чтения, параллельное чтение с нескольких реплик, сборка данных из фрагментов при избыточном кодировании EC, повтор запроса к другой копии при недоступности узла. При записи библиотека обеспечивает отправку данных на все необходимые OSD.

Подсистема блочного доступа обеспечивает максимальную производительность и минимальные задержки при работе с распределенным хранилищем. Нативный протокол ustor специально оптимизирован для работы в сети и учитывает особенности горизонтально масштабируемого хранения (минимизирует накладные расходы, позволяет эффективно чередовать обращения к разным узлам). Благодаря ему Базис.uStor может предоставлять высокую скорость доступа к данным с минимальной загрузкой CPU на узлах хранения. Также, эта подсистема позволяет легко интегрироваться с различными платформами: для KVM-гипервизоров есть специальный драйвер QEMU, напрямую работающий с uStor Client Library, для Linux-узлов без виртуализации – NBD, а для общих случаев – стандартный протокол (iSCSI) через отдельную подсистему.

Подсистема доступа по протоколу iSCSI

Для совместимости с промышленными стандартами и высокопроизводительного доступа Базис.uStor включает подсистему предоставления данных по популярному протоколу iSCSI. Она представлена программным компонентом:

- **uStor iSCSI** - специализированный сервис на узлах кластера хранения, реализующий сервер (target) для протокола iSCSI. Этот компонент позволяет клиентам, не использующим нативный ustor-протокол, подключаться к хранилищу как к стандартной блочной СХД.

В режиме iSCSI компонент создает targets и LUN, соответствующие томам, настроенным в Базис.uStor, и принимает подключения от iSCSI-инициаторов по сети (TCP/IP). При подключении и последующих SCSI-командах (чтение/запись) он преобразует эти команды в операции с распределенным хранилищем: через uStor Client Library обращается к OSD для выполнения запросов. iSCSI-таргет может поддерживать многопутевой доступ (MPIO) для повышенной отказоустойчивости: например, несколько узлов кластера могут выступать iSCSI порталами, и в случае отказа одного, инициатор переключится на другой без прерывания доступа.

Подсистема мониторинга в uStor

Подсистема мониторинга предназначена для наблюдения за работоспособностью и производительностью кластера хранения Базис.uStor, а также для передачи этой информации внешним системам мониторинга. Она обеспечивает сбор метрик, их агрегирование и экспорт, а также оповещение об аварийных ситуациях. В подсистему мониторинга входят следующие компоненты:

- **uStor Health** - компонент контроля «здоровья» всей системы. Он собирает телеметрию с каждого узла хранения: нагрузка на CPU/RAM, использование дискового

пространства, количество операций ввода-вывода, задержки, состояние сетевых соединений, аппаратные показатели (например, SMART-статусы дисков) и т.д. uStor Health реализован как агент на узлах, отправляющий данные в uStor etcd. Его задача – предоставить актуальную картину состояния всех компонентов. Например, Health будет знать, какой OSD помечен как «Degraded» из-за проблем с диском, или что у какого-то узла высокая загрузка по CPU, – и сигнализировать об этом. Health не только собирает метрики, но и может сравнивать их с порогами, генерируя события (оповещения) если параметры выходят за допустимые пределы;

- **uStor SNMP** - интеграционный модуль мониторинга, представляющий интерфейс SNMP (Simple Network Management Protocol). Он позволяет внешним системам сетевого мониторинга получать сведения о состоянии **Базис.uStor** стандартными средствами. uStor SNMP реализует SNMP-агент, поддерживающий набор OID, соответствующих ключевым метрикам СХД (емкость, число активных/неисправных дисков, IOPS, задержки, статус узлов и т.д.). Администратор может подключить **Базис.uStor** к любой SNMP-совместимой существующей системе мониторинга и получать от нее данные – либо активными запросами (GET), либо путем получения SNMP Trap. uStor SNMP отправляет trap-сообщения при важных событиях (например, отказ диска, потеря узла, исчерпание места) на указанные адреса. Этот компонент действует как мост между внутренней системой здоровья **Базис.uStor** и традиционной экосистемой мониторинга по SNMP;
- **uStor Health Exporter** - компонент экспорта метрик, ориентированный на современные системы мониторинга, такие как Prometheus. Health Exporter собирает детальные метрики производительности и состояния от uStor Health и предоставляет их через HTTP-интерфейс в формате, пригодном для опроса Prometheus-сервером. С помощью этого экспортера администраторы могут интегрировать Базис.uStor в стек мониторинга Grafana/Prometheus: данные будут собираться и визуализироваться наряду с метриками других частей инфраструктуры. Exporter обеспечивает низкоуровневый доступ к метрикам с минимальной задержкой и нагрузкой на сам кластер.

Подсистема мониторинга тесно взаимодействует с контроллером и хранилищем: uStor etcd предоставляет ей информацию о состоянии узлов/дисков, uStor Health получает данные с OSD и ОС узлов хранения. На основе этих данных могут строиться дашборды администрирования (в Web UI) и генерироваться оповещения. Благодаря наличию Prometheus- и SNMP-интеграции, uStor может легко встраиваться в существующие процессы мониторинга дата-центра.

Общая схема взаимодействия программных компонентов в составе программно-определяющего хранилища Базис.uStor совместно с узлами вычисления показана на рисунке 17 (для гиперконвергентного варианта хранилища) и на рисунке 18 (для выделенного конвергентного варианта хранилища).

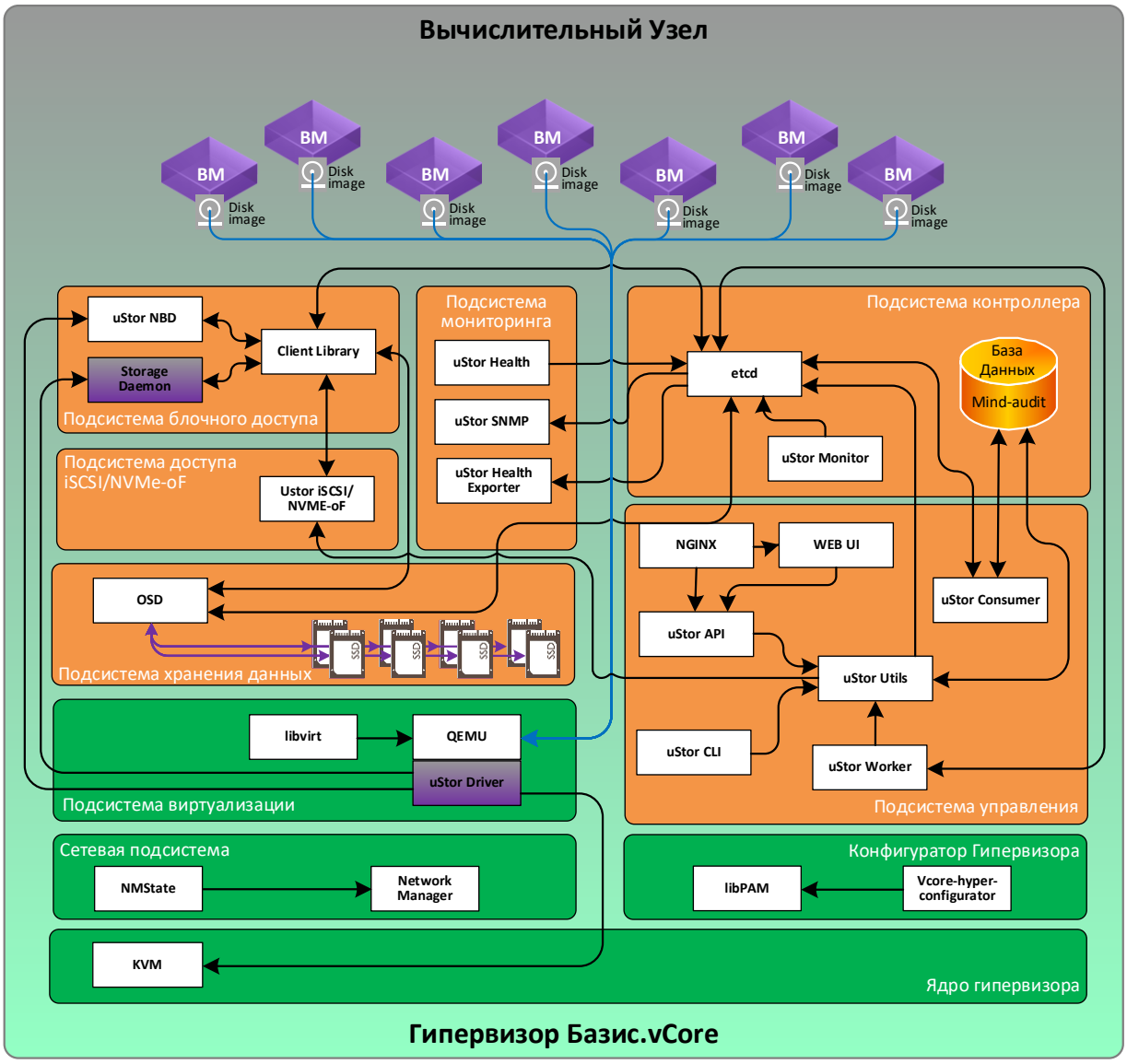

Рисунок 17. Схема взаимодействия программных компонентов программно-определяемого хранилища Базис.uStor и гипервизора Базис.vCore для гиперконвергентного варианта хранилища

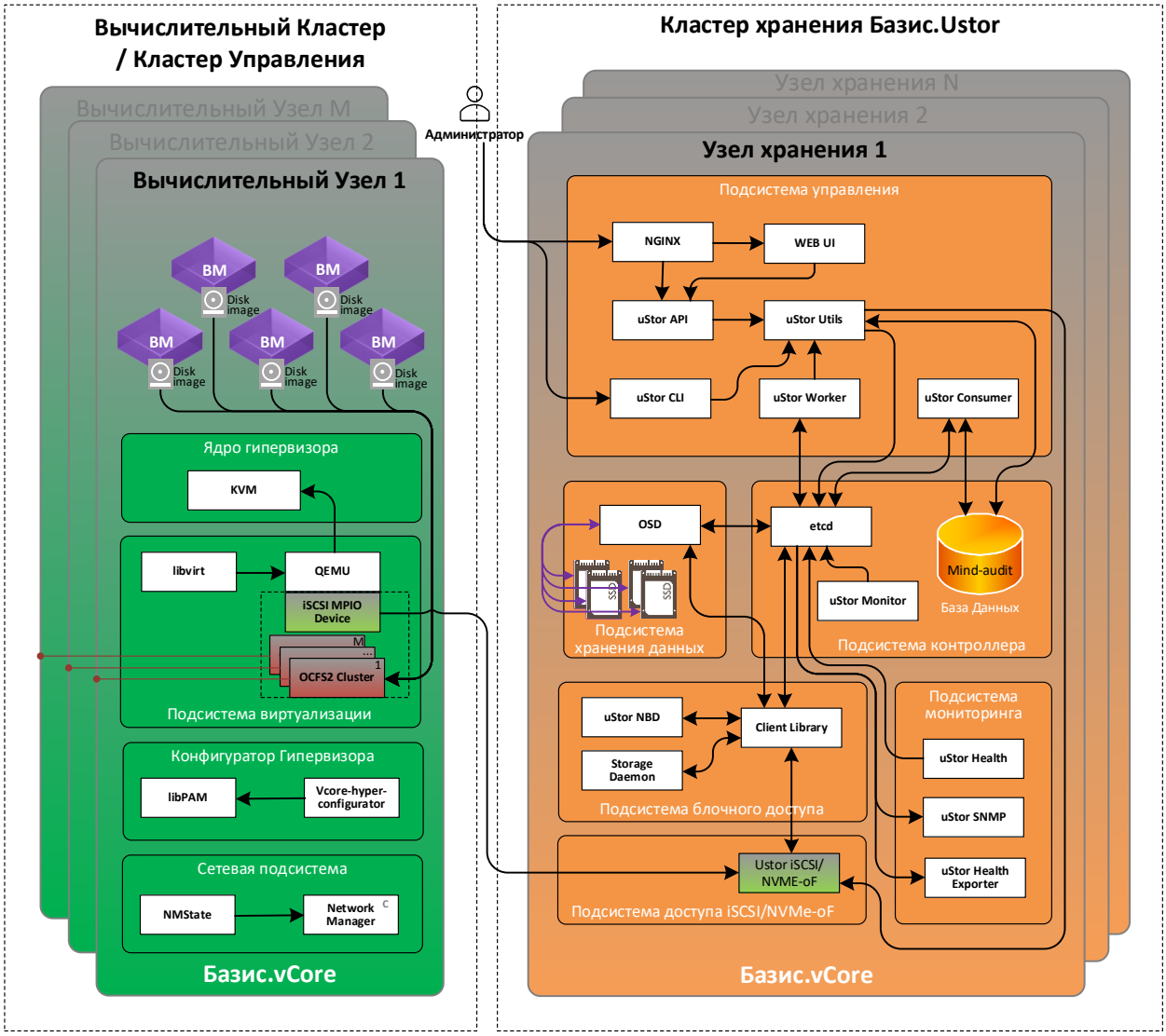

Рисунок 18. Схема взаимодействия программных компонентов программно-определяемого хранилища Базис.uStor и гипервизора Базис.vCore для выделенного конвергентного варианта хранилища

9.6 Программная платформа Скала^р Геном

Программная платформа **Скала^р Геном** (Рисунок 19) предназначена для управления жизненным циклом ПАК, диагностики, мониторинга, обслуживания и управления как отдельным ПАК, так и инфраструктурой, состоящей из множества различных ПАК **Скала^р**. ПО **Скала^р Геном** обладает, в частности, обладает следующим функционалом:

- ведение электронного паспорта **Машины**;
- отслеживание состояния узлов;
- предоставление доступа к IPMI всех узлов **Машины**;
- вывод узла **Машины** в режим обслуживания;
- загрузка и запуск обновления ПО;
- сбор данных о конфигурации элементов **Машины**;
- сбор данных, отображение, мониторинг элементов ПО, активных компонентов Модулей **Машины**, служебных сервисов и сервисов БД;
- конфигурирование метрик мониторинга, настройка уведомлений;
- конфигурирование графического отображения на информационных панелях в виде графиков, отдельных значений, диаграмм, таблиц;
- хранение метрик с возможностью настройки глубины хранения и управления жизненным циклом хранимых данных;
- отображение в пользовательском графическом интерфейсе данных о состоянии объектов мониторинга;
- контроль изменений объектов мониторинга в режиме, близком к реальному времени;
- сбор и мониторинг логов;
- мониторинг сервисов, специфичных для различных типов **Машин**.

Объектом мониторинга **Скала^р Геном** может быть любой физический или логический объект, например, память, процессор, файловая система, количество пользователей, очередь файлов на обработку, объем обработанного трафика, значение температуры, и другие.

Отличительной особенностью ПО **Скала^р Геном** являются возможности мониторинга специфичных параметров ПАК, обеспечивающих его надежность и производительность, что позволяет выполнять быстрый и качественный анализ причин возникновения внештатных ситуаций, строить прогнозы развития ситуации в будущем.

Сбор данных с узлов ПАК осуществляется с помощью установленных агентов ОС и СУБД.

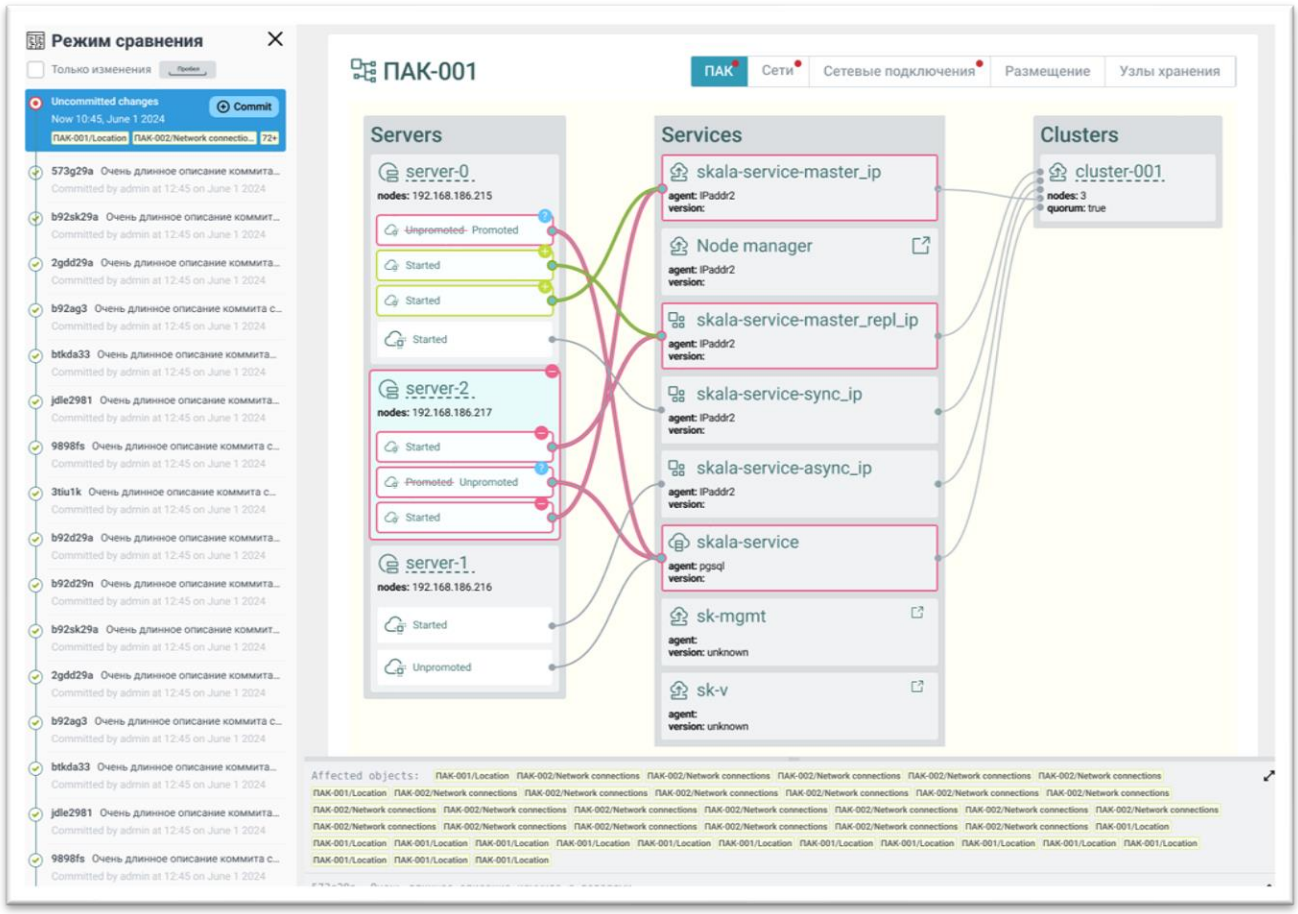

Рисунок 19. Пример интерфейса ПО Скала^р Геном (отображение схемы подключений узлов ПАК)

ПО **Скала^р Геном** позволяет проводить настройку появлении новых критических уведомлений, условия их получения гибко настраиваются в соответствии с текущими потребностями (Рисунок 20). Можно управлять формированием почтовых уведомлений: регулировать их группировку, частоту отправки и т.д.

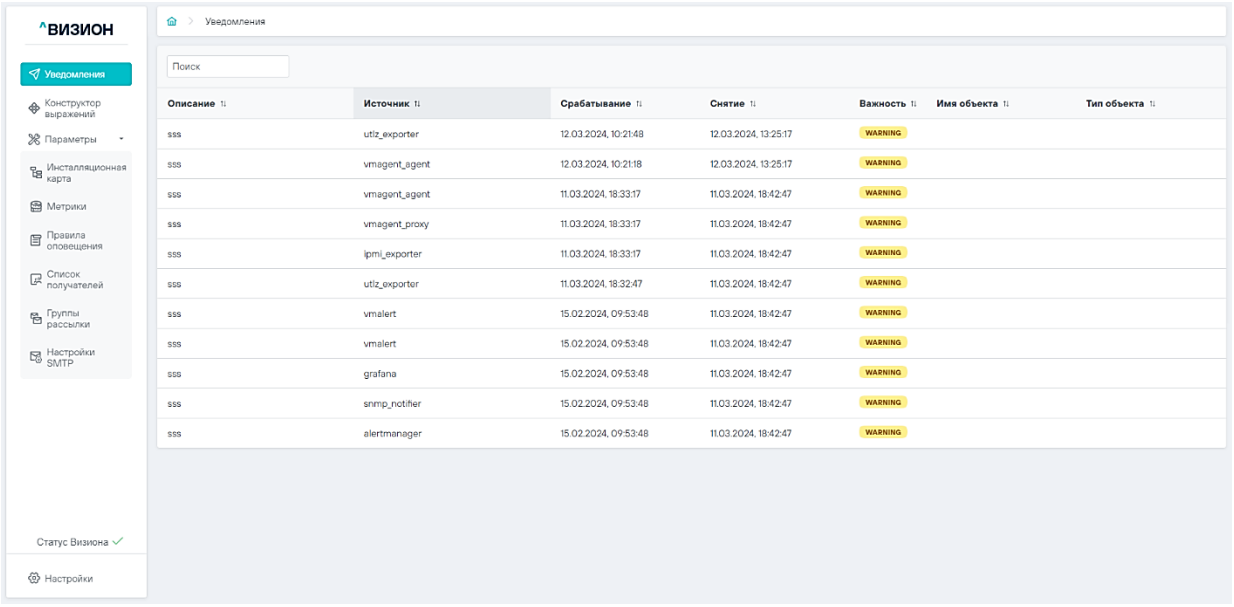

Рисунок 20. Пример интерфейса ПО Скала^р Геном (пример окна настройки уведомлений мониторинга)

Если при эксплуатации ПАК используются дополнительные каналы доставки уведомлений, возможна настройка их трансляции во внешнюю систему управления оповещениями об инцидентах. Эта настройка производится в пользовательском интерфейсе раздела мониторинга ПО **Скала^р Геном** (Рисунок 21).

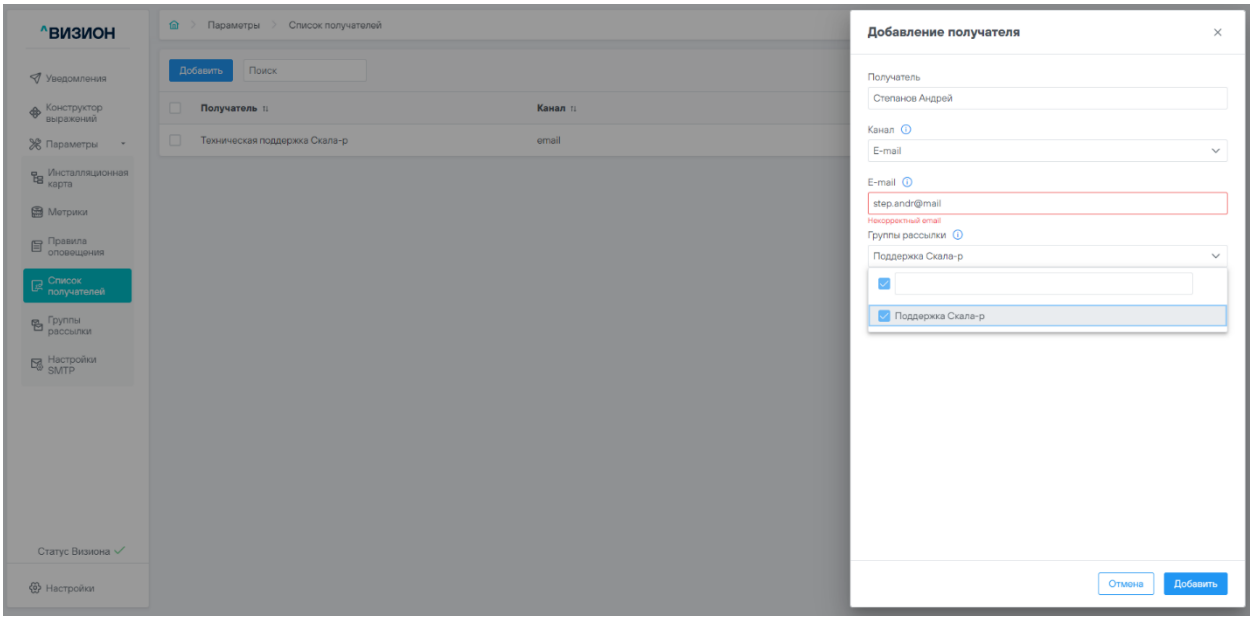

Рисунок 21. Пример интерфейса ПО Скала^р Геном (настройка пересылки сообщений)

Архитектурно, программная платформа **Скала^р Геном** состоит из следующих сервисных виртуальных машин, вместе образующих программную платформу обслуживания, управления и мониторинга **Скала^р Геном** (Рисунок 22): ВМ Платформы (сервер управления), ВМ Топологии (сервер топологии, CMDB) и ВМ Мониторинга (Визион).

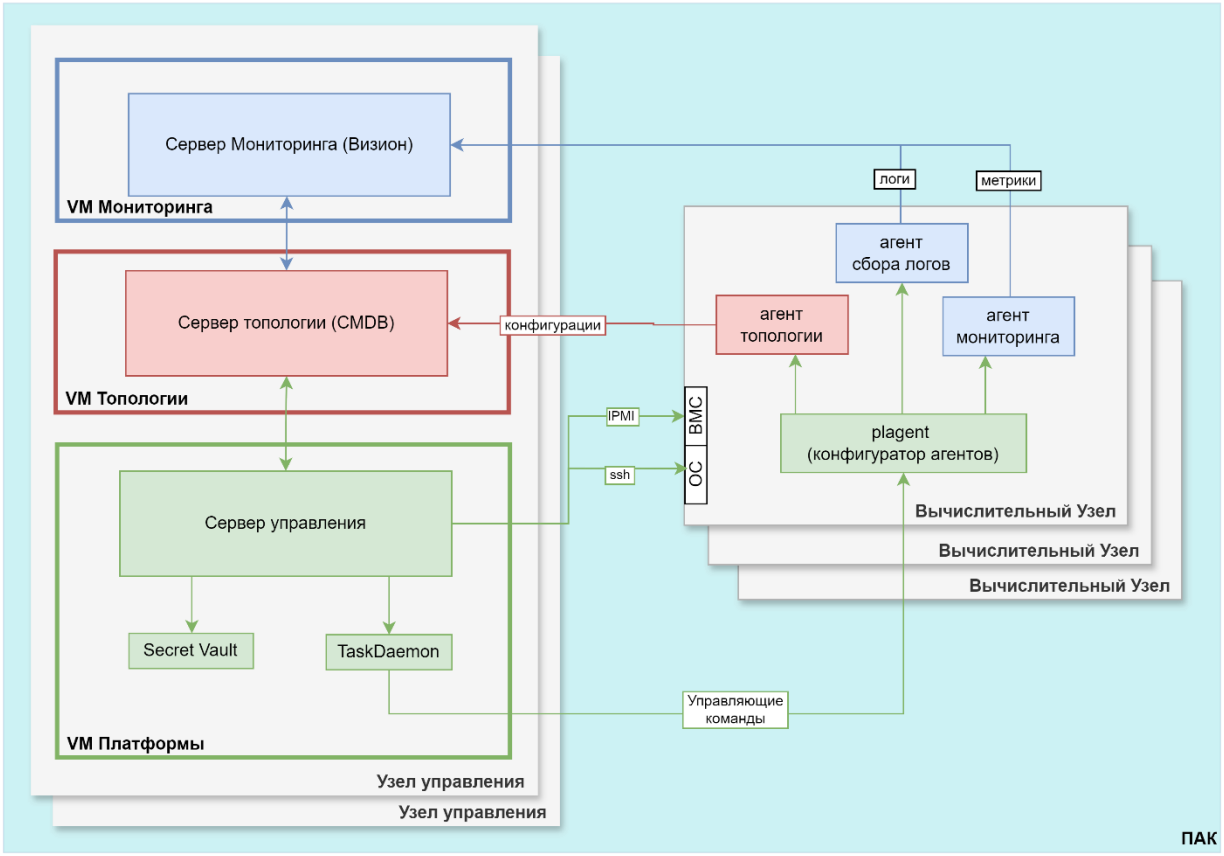

Рисунок 22. Основные компоненты программной платформы Скала^р Геном

Программная платформа **Скала^р Геном** может быть развернута как в режиме, обеспечивающем обслуживание одного ПАК (Рисунок 22), так и в режиме централизованного управления несколькими ПАК, обеспечивающим управление инфраструктурой, состоящей из множества ПАК **Скала^р** (Рисунок 23). В последнем случае, на служебных узлах подчиненных ПАК разворачиваются ВМ прокси-сервера топологии и ВМ прокси-сервера мониторинга, передающие соответствующие данные на служебные узлы основного ПАК. При этом, сервер управления, развернутый на ВМ Платформы основного ПАК, обеспечивает непосредственное управление агентами, установленными на вычислительных узлах подчиненных ПАК.

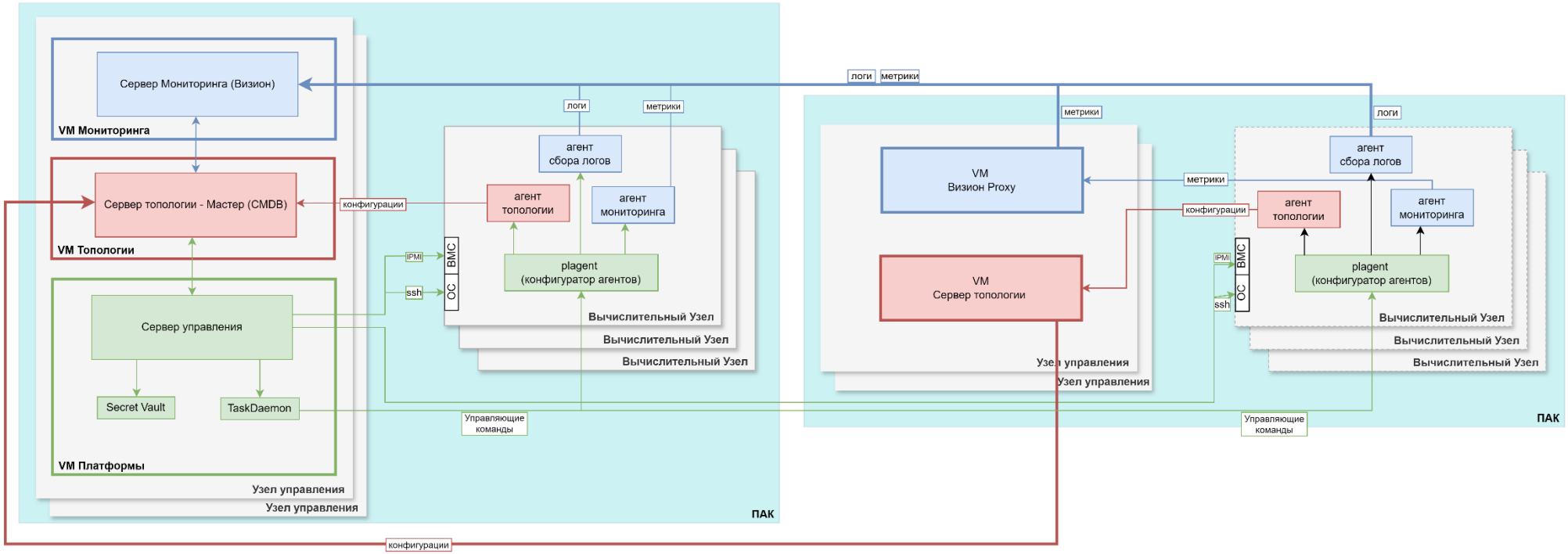

Рисунок 23. Программная платформа Скала^р Геном в режиме централизованного управления несколькими ПАК

10. Высокая доступность и защита данных

Для обеспечения высокой доступности в решении реализована соответствующая функциональность — кластер высокой доступности на уровне гипервизора Базис.vCore. В качестве основы кластер использует гиперконвергентное программно-определяемое хранилище, реализующее высокую доступность данных виртуальных машин.

Для организации резервного копирования реализован сервер резервного копирования, интегрированный в систему управления **Машиной**.

10.1 Кластер высокой доступности

Предполагается, что для вычислительных узлов из состава **Машины** включается функция «высокой доступности». В этом случае соответствующие компоненты в каждом гипервизоре начинают отслеживать доступность каждого прочего вычислительного узла **Машины** и вести учет того, какие виртуальные машины работают на каждом узле. В случае отказа вычислительного узла, работавшие на нем виртуальные машины так же «упадут», но будут перезапущены на прочих вычислительных узлах **Машины**, минимизировав тем самым время простоя.

Следует учитывать, что вопрос доступности ВМ находится на стыке собственно функции высокой доступности в гипервизоре и программно-определяемой системы хранения.

Для функции высокой доступности практически несущественно, сколько вычислительных узлов откажут одновременно: виртуальные машины будут перезапущены на оставшихся. Но с точки зрения данных виртуальных машин, некоторые из них могут стать недоступны уже при втором-третьем одновременном отказе. Это нормальное поведение системы, и, если будет поставлена задача обеспечить высокую доступность и в случае множественного отказа вычислительных узлов, она сможет быть решена путем указания соответствующих политик резервирования данных.

В гипервизоре работает отдельный сервис, обеспечивающий высокую доступность для виртуальных машин.

Эти сервисы, запущенные на каждом из вычислительном узле **Машины**, проверяют доступность друг друга по snmp, тайм-аут недоступности узла — минута.

При выборе вычислительного узла для перезапуска «упавшей» виртуальной машины реализована логика выбора наименее загруженного.

10.2 Служебные узлы с двойным резервированием

В **Машине** содержатся два служебных вычислительных узла, объединенные в зеркальный кластер для выполнения служебных функций. **Служебные узлы** выполняют функции обеспечения базовых сервисов, отвечает за мониторинг и управление аппаратными и программными компонентами **Скала^р МДИ.В.**

Установлены программная платформа **Скала^р Геном** и ПО Аванпост FAM (опция), выполняющие следующие функции:

- сбор, хранение и отображение данных мониторинга;
- настройка правил оповещения;
- отправка оповещений о состоянии ПАК;
- управление аппаратными компонентами;
- настройка программных компонентов;

- настройка интеграции со сторонним ПО;
- авторизацию пользователей (опция FAM).

В роли базовой операционной системы на каждом служебном узле используется ОС Альт Сервер 8 СП Release 10 – сертифицированная ФСТЭК России и включенная в Единый реестр российских программ для электронных вычислительных машин и баз данных. В качестве гипервизора используется связка из программного модуля ядра KVM и эмулятора ввода-вывода QEMU из дистрибутива базовой ОС.

Отказоустойчивость работы служебных ВМ с ПО **Скала^р Геном** и ПО Аванпост FAM на служебных узлах достигается размещением двух копий каждой сервисной виртуальной машины на разных служебных узлах, при этом одна из них активна, другая работает в пассивном режиме.

В каждой из этой пары сервисной виртуальной машине в качестве дискового устройства используется прямое подключение неформатированного логического раздела (RAW), расположенного на программном массиве RAID1 (зеркало) на двух SSD дисках. Логический раздел активной сервисной виртуальной машины реплицируется, в свою очередь в логический раздел пассивной копии (выключенной) сервисной виртуальной машины. Синхронизация данных между локальным и удаленными разделами идет через протокол TCP (без шифрования и аутентификации), по умолчанию используется порт TCP/3260 (может быть изменен).

Основной логический диск и его реплика на удаленном сервисном узле работают в режиме соответственно первичного (primary) узла и вторичного (secondary). Вторичный (резервный) узел хранит реплику, но не позволяет осуществить к ней локальный доступ, первичный же позволяет осуществить локальный доступ. Как только происходит повышение роли логического диска до первичного – доступ открывается, а сам диск больше не будет принимать реплику, а наоборот, будет отдавать свою реплику на удаленный узел.

Логическая схема работы служебных узлов **Скала^р МДИ.В** представлена на рисунке 24.

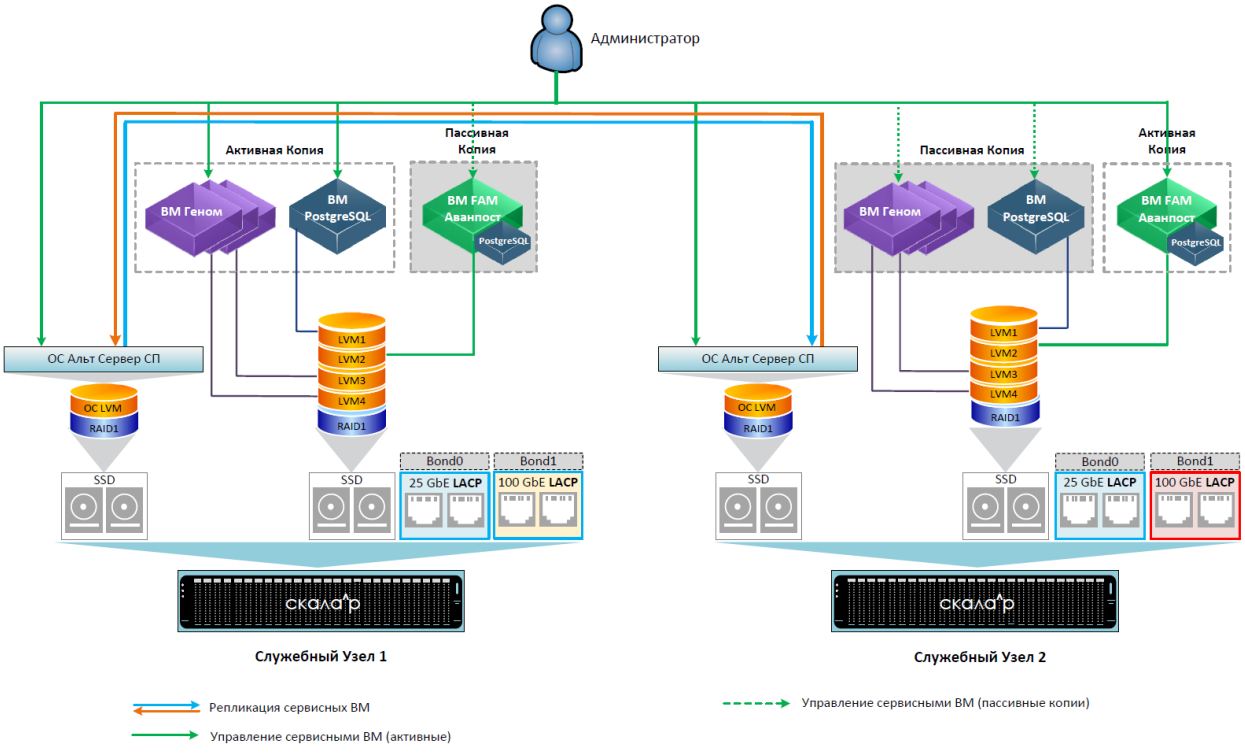

Рисунок 24. Логическая схема работы служебных узлов Скала^р МДИ.В

10.3 Высокая доступность данных программно-определяемого хранилища

Программно-определяемая подсистема хранения **Скала^р МДИ.В** в виде гиперконвергентной **Базис.uStor** может реализовать избыточность данных двумя алгоритмами:

- хранение заданного количества реплик - полных копий данных;
- хранение блоков избыточного кодирования (Erasure Coding) на основе кода Рида-Соломона.

В первом случае для каждого «блока» данных создается указанное число копий, притом никакие копии этого блока не будут расположены на одном и том же узле (Рисунок 25).

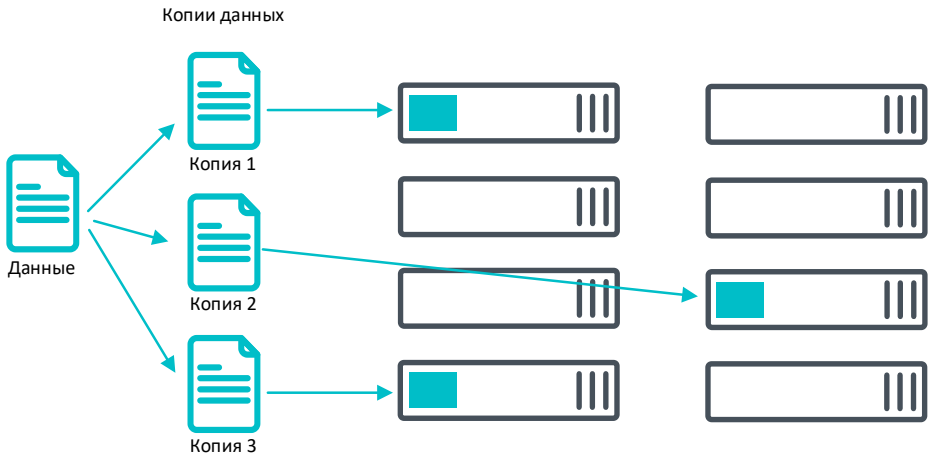

Рисунок 25. Создание реплик, полных копий данных

За счет этого в случае отказа узлов или накопителей в них данные виртуальной машины продолжат оставаться доступными.

Обычно используется политика хранения в 2 или 3 копиях данных.

Во втором случае резервирование обеспечивается добавлением блоков четности/избыточности. Могут использоваться следующие варианты (данные + четность), в зависимости от желаемого уровня доступности и количества хостов в системе: 3 + 2, 5 + 2, 7 + 2, 17 + 3. Схема реализации технологии избыточного кодирования для случая 5 + 2 приведена ниже (Рисунок 26).

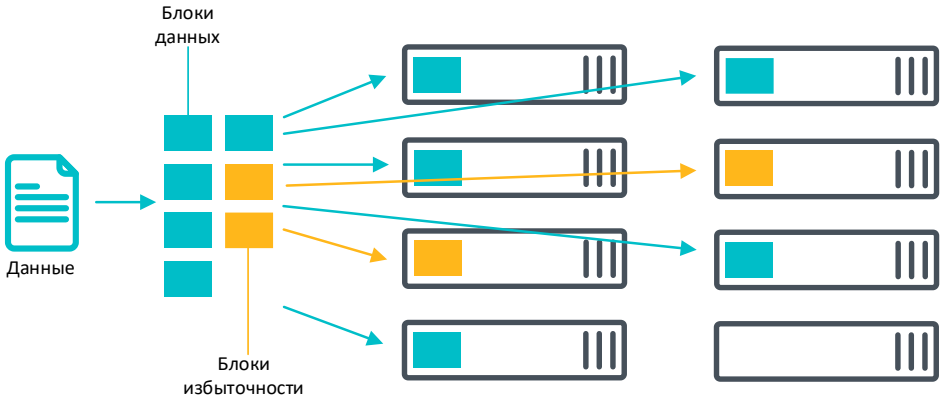

Рисунок 26. Избыточное кодирование 5 + 2

Преимущества и ограничения этих вариантов

Преимущества реализации избыточности данных репликами:

- высокая производительность дисковой подсистемы;
- быстрое восстановление недостающих копий данных при отказе одного или двух узлов;
- процесс восстановления недостающих реплик практически не влияет на общую производительность дискового хранилища;
- чаще целесообразен для небольших комплексов.

К недостаткам режима создания реплик можно отнести высокие накладные расходы на хранение данных.

Режим избыточного кодирования позволяет более эффективно использовать дисковое пространство, но у него имеются следующие ограничения:

- для достижения целевой производительности может потребоваться больше накопителей / выбор более быстрых накопителей в некоторых сценариях;
- более низкая скорость восстановления недостающих копий данных (rebuild) при потере одного или двух узлов;
- дополнительная загрузка процессоров;
- для его использования требуется пять узлов и больше.

Встроенный инструмент резервного копирования

В сервере управления виртуальной инфраструктурой с ПО Базис.vControl реализован встроенный механизм осуществления резервного копирования виртуальных машин.

Доступны выбор ВМ для резервного копирования, управление расписанием автоматизированного резервного копирования, выбор места хранения резервных копий. Поддерживается инкрементальное и полное резервное копирование.

По умолчанию резервные копии хранятся на программно-определяемой системе хранения **Машины**, этого достаточно для защиты от логических ошибок (например, случайного удаления ВМ).

Если модель угроз предусматривает защиту от отказа программно-определяемой системы хранения, то поддерживается и рекомендуется использовать внешнее хранилище резервных копий, с подключением по протоколу NFS.

Использование третьих систем резервного копирования также возможно, совместимость конкретных продуктов и версий следует уточнять в документации.

11. Поставка и лицензирование ПО в составе Машины

Необходимое программное обеспечение **Машины Скала^р МДИ.В** поставляется исключительно в составе данного ПАК и лицензируется по метрикам комплекса в соответствии с его размером.

Основное функциональное ПО (Таблица 3) лицензируется в составе Лицензионного пакета Модулей, составляющих **Машину** (и может облагаться нулевым НДС).

Специализированное и системное ПО входит в структуру Модулей и **Машины**, и не выделяется в обособленный пакет. Таблица 4 содержит информацию о номенклатуре данного ПО.

На состав лицензируемого ПО, эксплуатируемого в составе **Машины Скала^р МДИ.В**, также влияет выбор в пользу защищенного исполнения функционального ПО виртуализации (сертифицированная ФСТЭК версия). Кроме того, выбор типа применяемой системы хранения, а именно: гиперконвергентное программно-определяемого распределенного хранилище данных uStor либо аппаратная СХД, используемой для организации постоянного хранилища, также учитывается при формировании необходимого лицензионного пакета. Необходимая для заданного уровня расширенной гарантии на ПАК техническая поддержка входит в лицензии.

Ниже приведен перечень ключевых компонентов Лицензионного пакета **Машины Скала^р МДИ.В**.

Таблица 3. Компоненты Лицензионного пакета

Краткое наименование компонента	Назначение	Единица лицензирования
ПО **Скала^р Геном**. Лицензия на систему эксплуатации инфраструктур на базе **Машин Скала^р**	ПО обеспечения эксплуатации инфраструктур на базе **Машин Скала^р**, а также развертывания, управления, мониторинга и конфигурации.	Вычислительный/сетевой узел в составе ПАК
Защищенное исполнение системы виртуализации		
Лицензия на ПО Базис.Virtual Security (Исполнение 3), включая Базис.vCore и Базис.vControl, с функционалом СЗИ	ПО серверной виртуализации (защищенное исполнение)	2 физических процессора одного физического сервера с ролью вычислений (гипервизора)
Медиа-комплект ПО Базис.Virtual Security, включая Базис.vCore и Базис.vControl.	Носитель ПО и формуляр программного изделия	На 1 ПАК

Краткое наименование компонента	Назначение	Единица лицензирования
Стандартное исполнение системы виртуализации		
Лицензия на ПО Базис.vControl, включая Базис.vCore	ПО серверной виртуализации	2 физических процессора одного физического сервера с ролью вычислений (гипервизора)
Лицензия на ПО Базис.vControl, стартовый пакет	ПО серверной виртуализации	На 1 ПАК
Вариант: программно-определяемая система хранения данных (SDS)		
Лицензия на ПО Базис.vControl, модуль SDS, опция uStor Standard	ПО гиперконвергентной распределенной программно-определяемой системы хранения данных	*При объеме дисковых накопителей сервера SDS свыше 50 Тбайт:* 1 Тбайт сырого объема дисковых накопителей ---------- *При объеме дисковых накопителей сервера SDS до 50 Тбайт:* На 1 физический сервер SDS
Лицензия на ПО Базис.vControl, модуль SDS, опция uStor Enterprise	ПО с опцией выделенной (конвергентной) распределенной программно-определяемой системы хранения данных с поддержкой дополнительных функций	1 Тбайт сырого объема дисковых накопителей
Вариант: Аппаратная СХД		
Система хранения данных (СХД)	Хранение данных	Дополнительного лицензирования не требуется

Таблица 4. Специализированное и системное ПО в составе ПАК

Краткое наименование компонента	Назначение	Единица лицензирования
Защищенное исполнение (для лицензионного компонента Базис.Virtual Security (Исполнение 3))		
Средство защиты информации "Среда разработки и исполнения Java Axiom JDK"	Защищенная среда разработки и исполнения приложений на Java	Вычислительный узел с ролью BVS и вычислительный узел с ролью вычислений (гипервизора)
СУБД Postgres Pro Certified СКАЛА-Р	Служебная защищенная СУБД для работы компонентов BVS и vControl (без возможности хранения пользовательских данных)	Вычислительный узел с СУБД, специальная облегченная лицензия только для ПАК **Скала^р**
ОС Астра или ОС Альт Сервер, сертифицированные версии	ОС для исполнения на узлах с ролью BVS	Вычислительный узел с ролью BVS
Защищенное и стандартное исполнение, в случае выделенных служебных узлов для ПО Геном		
ОС Астра или ОС Альт Сервер, с подсистемой виртуализации, сертифицированные версии	ОС узлов для запуска служебных виртуальных машин (компонентов Геном, включая мониторинг, обслуживание и др.)	Служебный вычислительный узел
ПО Системы единой аутентификации **Машин Скала^р** Аванпост FAM	*Опционально*: Организация универсальной сертифицированной платформы авторизации и аутентификации (IAM) в составе ПАК при отсутствии ее аналогов в инфраструктуре заказчика	4 процессорных ядра, специальная облегченная лицензия только для ПАК **Скала^р**, не более 100 учетных записей

Собственное ПО **Скала^р** и специальные лицензии только для ПАК **Скала^р** поставляются исключительно и только в составе ПАК **Скала^р**.

12. Гарантированное качество

Качественные показатели **Машины серверной виртуализации Скала^р МДИ.В** обеспечиваются ее соответствием проверенному стандартному варианту, соблюдением установленных норм и требований по формированию, реализацией работ высококвалифицированными специалистами на всех этапах жизненного цикла.

Производство (комплектование и развертывание ПО)

- При производстве используются высококачественные комплектующие;
- Сборка продукции осуществляется строго в соответствии с утвержденным планом размещения компонентов;
- Первичное развертывание ПО осуществляется в автоматическом режиме;
- Дополнительные настройки ПО осуществляются в соответствии с утвержденной методикой и пошаговой инструкцией;
- Осуществляется функциональное тестирование сформированной **Машины**;
- Отклонения от типового решения **Скала^р МДИ.В** исключены.

Передача в эксплуатацию

- **Скала^р МДИ.В** полностью сформирована, протестирована, готова к размещению в сети заказчика и размещению прикладного ПО;
- В комплекте со **Скала^р МДИ.В** передаются паспорт, сертификат на поддержку;
- Проводится обучение специалистов заказчика работе со **Скала^р МДИ.В** (по запросу).

Поддержка

- **Скала^р МДИ.В** поставляется с годовой поддержкой (более выгодный вариант — на 3 или 5 лет), которая включает в себя решение вопросов, связанных с нарушениями работоспособности как комплекса в целом, так и его отдельных аппаратных компонентов и программного обеспечения;
- Первая и вторая линия поддержки предоставляются непосредственно производителем **Скала^р** или сертифицированным партнером **Скала^р**;
- У заказчика есть возможность выбора варианта поддержки (9x5 или 24x7);
- В сложных случаях в решении проблем на третьей линии поддержки участвуют архитекторы и инженеры, разработчики ПО и **Машины серверной виртуализации Скала^р МДИ.В**.

Дополнительные требования

Возможна реализация дополнительных требований по модернизации или развитию **Скала^р МДИ.В** (по запросу), в том числе:

- аппаратная модернизация решения;
- горизонтальное или вертикальное масштабирование нового или имеющегося решения;
- изменение функциональности компонентов дистрибутивов ПО, их доработка;
- тестирование приложений, производительности приложений или иное другое запрошенное тестирование.

Работы выполняются с участием архитекторов и инженеров, разработчиков **Машины** и ПО **Скала^р МДИ.В**.

13. Требования к размещению

Изделие представляет собой серверный монтажный шкаф 19”, высота 42U, с дальнейшей возможностью модульной расширяемости до 14 стоек.

Наполнение шкафа оборудованием и совокупный вес зависят от выбранного варианта решения и могут составлять от 400 до 800 кг.

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

Расчетная потребляемая мощность шкафа составляет от 6 до 11 кВт.

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

Требования для подключения **Машины Скала^р** к локальной сети заказчика определяются на этапе формирования спецификации **Машины**.

При развертывании решения на нем будут выполнены настройки сетевых адресов в соответствии со структурой сети заказчика. Заказчик должен предоставить необходимые данные в соответствии с номенклатурой компонентов решения.

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

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

14. Техническая поддержка

Поставка **Скала^р МДИ.В** осуществляется с предварительными сборкой, тестированием и настройкой оборудования согласно требованиям заказчика. Качественная поддержка **Скала^р МДИ.В** 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 9x5 или в расширенном режиме 24x7 (опция для критической функциональности);
- Срок начально приобретаемой годовой технической поддержки может быть увеличен до 3 и 5 лет, также доступна пролонгация поддержки по окончании сроков;
- Возможно включение в состав стандартных пакетов дополнительных опций и услуг.

Состав типовых пакетов услуг по технической поддержке представлен в таблице ниже (Таблица 5).

Таблица 5. Пакеты услуг по технической поддержке Скала^р МДИ.В

Услуга	Пакет «9x5»	Пакет «24x7»
Режим «Обслуживание комплекса **Скала^р МДИ.В** в режиме 9x5» (в рабочее время по рабочим дням)	+	–
Режим «Обслуживание комплекса **Скала^р МДИ.В** в режиме 24x7» (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 МДИ.В** включая: ▪ устранение аппаратных неисправностей; ▪ техническое сопровождение ПО.	+	+
Консультации по работе комплекса **Скала^р МДИ.В**	+	+

Услуга	Пакет «9x5»	Пакет «24x7»
«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	+	+
Расширенные опции обслуживания	–	+
Времена реагирования и отклика, не более:		
Время регистрации обращений	30 минут, рабочие часы (9x5)	30 минут, круглосуточно (24x7)
Подключение специалиста к решению инцидентов критичного и высокого уровней	В течение 1 рабочего часа (9x5)	В течение 1 часа (24x7)

Примечание к срокам ремонта оборудования

Комплекс **МДИ.В** 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 вычислительных узлов (серверов).

Полное описание услуг поддержки доступно на сайте skala-r.ru.

15. О компании

Скала^р — модульная платформа для построения высоконагруженной ИТ-инфраструктуры (продукт Группы Rubytech). Лидер российского рынка ПАК (по версии CNews Analytics, 2024).

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 развертываются 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.