

УТВЕРЖДЕН

643.РЕНМ.02.05-01 32 01-ЛУ

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ  
«СИСТЕМА ХРАНЕНИЯ ДАННЫХ «ЕНИСЕЙ»

Руководство системного программиста

643.РЕНМ.02.05-01 32 01

Листов 34

|               |              |              |               |              |
|---------------|--------------|--------------|---------------|--------------|
| Инев. № подл. | Подп. и дата | Взам. инв. № | Инев. № дубл. | Подп. и дата |
|               |              |              |               |              |

## АННОТАЦИЯ

Настоящий документ «Программное обеспечение «Система хранения данных «Енисей» Руководство системного программиста» 643.РЕНМ.02.05-01 32 01 является руководством системного программиста по работе с программным обеспечением «Система хранения данных «Енисей» 643.РЕНМ.02.05-01 (далее – программа, ПО «СХД «Енисей»)), которое является составной частью ПАК СХД «Енисей».

Документ содержит общие сведения о программе, ее функциональном назначении, используемых технических и программных средствах, описание логической структуры, сведения о порядке установки, настройки и проверки программы, а также о сообщениях, выдаваемых системному программисту в ходе взаимодействия с программой.

Оформление документа выполнено согласно требованиям ГОСТ 19.106–78, ГОСТ 19.503–79.

## СОДЕРЖАНИЕ

|  |    |
|--|----|
| 1. Общие сведения о программе .....                              | 5  |
| 1.1. Наименование и обозначение программы.....                   | 5  |
| 1.2. Назначение программы .....                                  | 5  |
| 1.3. Основные функциональные возможности программы.....          | 5  |
| 1.4. Основные характеристики программы .....                     | 6  |
| 1.5. Требования к техническим средствам.....                     | 6  |
| 1.6. Требования к программному обеспечению.....                  | 7  |
| 1.7. Требования к ПО в части обеспечения интеграции .....        | 7  |
| 1.8. Общие характеристики входной и выходной информации.....     | 8  |
| 1.9. Языки программирования, на которых написана программа ..... | 8  |
| 1.10. Функциональные ограничения на применение .....             | 8  |
| 1.11. Требования к персоналу.....                                | 9  |
| 2. Структура программы .....                                     | 10 |
| 2.1. Структура программы с описанием функций ее компонентов..... | 10 |
| 2.1.1. Внешнее взаимодействие .....                              | 14 |
| 2.1.2. Внутреннее взаимодействие.....                            | 15 |
| 2.1.3. Хранение и восстановление данных .....                    | 16 |
| 2.1.3.1. Структура пула .....                                    | 19 |
| 2.1.3.2. Размещение данных .....                                 | 20 |
| 2.1.3.3. Восстановление данных.....                              | 21 |
| 2.1.3.4. Рекомендации по схемам защиты данных.....               | 22 |
| 2.1.4. Блочный доступ к данным .....                             | 22 |
| 2.1.5. Файловый доступ к данным.....                             | 23 |
| 2.2. Связи программы с другими программами .....                 | 23 |
| 2.3. Пользователи и группы.....                                  | 24 |
| 2.3.1. Учетные записи пользователей .....                        | 24 |
| 2.3.2. Группы пользователей.....                                 | 24 |
| 3. Настройка программы.....                                      | 25 |
| 3.1. Состав и содержание дистрибутива .....                      | 25 |
| 3.2. Настройка подключения к БД.....                             | 25 |
| 3.3. Установка программы и ее компонентов.....                   | 25 |
| 3.4. Настройка программы на состав ТС и ПС.....                  | 26 |
| 3.5. Настройка программы на условия конкретного применения ..... | 26 |

|   |    |
|---|----|
| 4. Проверка программы .....                             | 27 |
| 4.1. Проверка работоспособности .....                   | 27 |
| 4.2. Методы проверки .....                              | 27 |
| 4.2.1. Общие положения.....                             | 27 |
| 4.2.2. Получение информации о контроллерах СХД.....     | 28 |
| 4.2.3. Проверка доступности накопителей данных.....     | 29 |
| 4.2.4. Получение информации о сетевых интерфейсах ..... | 30 |
| 4.2.5. Получение информации о статусе кластера.....     | 31 |
| 5. Сообщения системному программисту .....              | 32 |
| Перечень сокращений .....                               | 33 |

## 1. ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ

### 1.1. Наименование и обозначение программы

Полное наименование: Программное обеспечение «Система хранения данных «Енисей».

Краткое наименование: ПО «СХД «Енисей».

Обозначение: 643.РЕНМ.02.05-01.

### 1.2. Назначение программы

ПО «СХД «Енисей» предназначена для развертывания высокопроизводительных систем хранения данных.

Область применения: ПО «СХД «Енисей» оптимально подходит для крупных организаций, которые работают с большим количеством данных в режиме 24/7. Гибкий функционал и возможности кластеризации системы позволяют демонстрировать отличные показатели производительности и отказоустойчивости в сегменте суперкомпьютеров.

Системы на базе ПО «СХД «Енисей» являются эффективным решением для организации хранилищ высокопроизводительного кластера, анализа большого объема данных, а также для корпоративных инфраструктур, в частности, для задач промежуточного хранилища с быстрым резервным копированием перед передачей данных на долгосрочное архивное хранение.

ПО «СХД «Енисей» выполняет основные задачи:

- хранение данных;
- резервирование данных;
- восстановление данных;
- управление и распределение большого объема данных;
- обслуживание запросов на чтение и запись данных.

Серверы пользователей, отправляющие запросы ПО «СХД «Енисей», называются хост-серверами. ПО «СХД «Енисей» предоставляет хост-серверам доступ к данным через каналы внешних сетей хранения данных.

### 1.3. Основные функциональные возможности программы

ПО «СХД «Енисей» обеспечивает функции:

- хранения данных (баз данных, "сырых" данных, резервных копий и так далее);
- обслуживания запросов на чтение и запись данных;
- управления и распределения большого объема данных;

- предоставления доступа к хранимым данным по различным протоколам связи для различных типов клиентов (потребителей этих данных);
- мониторинга ПО и аппаратной платформы;
- поддержки отказоустойчивых и само-восстанавливающихся дисковых массивов.

#### 1.4. Основные характеристики программы

В основе логической архитектуры ПО «СХД «Енисей» лежит виртуализация. Физические резервы программы распределяются между логическими (виртуальными) объектами, что позволяет абстрагироваться от аппаратной реализации и упрощает администрирование.

Для администрирования программы используются следующие логические объекты:

- кластер – логический объект ПО «СХД «Енисей», состоящий из одного или нескольких контроллеров СХД в единое пространство хранения данных. Объединение контроллеров СХД в кластер гарантирует высокую надёжность и бесперебойную работу ПАК СХД «Енисей».

- пул – логический объект ПО «СХД «Енисей», объединяющий пространства нескольких физических накопителей в единое пространство хранения данных. Структура пула обеспечивает надежное хранение данных с помощью технологии защиты данных RAID.

- том – логический объект ПО «СХД «Енисей», представляющий собой заданный объем пространства пула и доступный по заданным портам ПО «СХД «Енисей». В зависимости от типа доступа к данным тома делятся на блочные, файловые NFS и файловые SMB/CIFS. Тип тома задается при его создании и не может быть изменен в дальнейшем.

В программе реализован механизм создания снимков<sup>1</sup> для ресурсов с тонким и толстым типами резервирования.

Поддержка снимков осуществляется с использованием перенаправления записи (Redirect on Write) для блочных ресурсов.

#### 1.5. Требования к техническим средствам

Программа работает на технических средствах (далее – ТС) ПАК СХД «Енисей», сертифицированных и принятых в РФ.

---

<sup>1</sup> Снимок – «снимок» текущего состояния ресурса или группы ресурсов СХД. Снимки могут использоваться для восстановления данных при сбоях: путем возврата к состоянию предыдущего снимка или восстановления ресурсов из снимка. Также снимки могут использоваться в тестовых сценариях, которые могут оказать влияние на ресурсы СХД.

ПО «СХД «Енисей» сохраняет свою работоспособность при масштабировании и замене существующих ТС на средства, имеющие более высокие характеристики.

#### 1.6. Требования к программному обеспечению

Используемое при работе с программой ПО должно соответствовать требованиям, приведенным в таблице 1.

Таблица 1 – Требования к программному обеспечению

| ПО                        | Требования к ПО                |
|---------------------------|--------------------------------|
| Операционная система      | Debian 11                      |
| Веб-браузер               | Google Chrome                  |
| Дополнительные библиотеки | Интерфейс командной строки CLI |

#### 1.7. Требования к ПО в части обеспечения интеграции

ПО «СХД «Енисей» поддерживает интеграцию со сторонними системами по протоколам:

– iSCSI (Internet Small Computer System Interface) или интерфейсу малых компьютерных систем интернета. Это транспортный протокол, обеспечивающий передачу протокола хранения данных SCSI по TCP/IP через сетевое соединение, которым является Ethernet. iSCSI работает как метод организации распределенных хранилищ. iSCSI также используется как расширение для удаленного прямого доступа к памяти (RDMA). Его задачей является управление передачей данных непосредственно в буферы памяти компьютера SCSI, соединяющие компьютеры с устройствами хранения данных, и обратно без промежуточных копий данных и значительного вмешательства процессора.

– FCP (Fibre Channel Protocol) - транспортный протокол, обеспечивающий передачу протокола хранения данных SCSI по оптоволоконным сетям. Обеспечивает блочный доступ, высокую надёжность и производительность.

– SMB/CFIS – протокол общего доступа к сетевым файлам, который, как он реализован в Microsoft Windows, называется протоколом Microsoft SMB. Набор пакетов сообщений, определяющих определенную версию протокола, называется диалектом. Протокол CIFS является диалектом SMB. Как SMB, так и CIFS доступны в нескольких версиях Unix и других операционных системах.

– NFS – протокол сетевого доступа к файловым системам. Позволяет монтировать (подключать) удалённые файловые системы через сеть. NFS абстрагирован от типов файловых систем как сервера, так и клиента. NFS предоставляет клиентам прозрачный доступ к файлам и файловой системе сервера. Протокол NFS осуществляет доступ только к тем частям файла, к которым обратился процесс, и основное достоинство его в том, что он делает этот доступ прозрачным. Это означает, что любое приложение клиента, которое может работать с локальным файлом, с таким же успехом может работать и с NFS-файлом, без каких-либо модификаций самой программы.

#### 1.8. Общие характеристики входной и выходной информации

Входными и выходными данными для работы программы является неструктурированная информация, такая как файлы, архивы, резервные копии и т.д.

#### 1.9. Языки программирования, на которых написана программа

Программа реализована с применением следующих средств:

– языки программирования:

а) C;

б) Golang;

в) TypeScript (JavaScript) с набором фреймворков и библиотек:

1) Vue 3

2) Nuxt 3

3) PrimeVue 3

4) XTerm.js;

– язык написания запросов SQL;

– язык гипертекстовой разметки документа (веб-страницы) HTML;

– язык описания внешнего вида документа (веб-страницы) CSS;

– язык описания данных JSON;

– интерпретатор командной строки Bash.

#### 1.10. Функциональные ограничения на применение

ПО «СХД «Енисей» устанавливается, функционирует и поддерживается в актуальном состоянии на аппаратных комплексах «Енисей» под управлением ОС Debian 11.

Дополнительные функциональные ограничения на применение программы не предъявляются.



### 1.11. Требования к персоналу

Программа рассчитана на персонал, владеющий следующими основными знаниями и навыками:

- опыт работы на персональном компьютере;
- опыт работы с ОС;
- опыт работы с веб-приложениями (настройка типовых конфигураций браузера, навигация, работа с электронными и веб-формами, и другими типовыми интерактивными элементами);
- опыт работы с офисными приложениями;
- знание нормативной базы, регулирующей вопросы информационной и технической безопасности.

Системный программист дополнительно должен иметь следующие знания и навыки:

- наличие сертификата ОС уровня AL-1704 «Сетевое администрирование»;
- опыт использования ОС Debian 11 для автоматизации настройки и развертывания

ПО.

## 2. СТРУКТУРА ПРОГРАММЫ

ПО «СХД «Енисей» может работать как в одиночном режиме, так и в режиме кластера. Одиночный режим представляет собой базовый вариант реализации СХД, который имеет ряд ограничений по емкости, отказоустойчивости и т.д. Все алгоритмы ввода вывода и администрирования работают как в одиночном режиме, так и в режиме кластера.

Кластер представляет собой горизонтальное соединение (scale out) контроллеров СХД между собой по топологии, позволяющей всем видеть всех. В таком режиме ресурсы СХД – диски, тома, доступны на всех контроллерах СХД кластера. Кластер позволяет решать вопросы расширения емкости, балансировки нагрузки, устранения единой точки отказа, и т.д.

### 2.1. Структура программы с описанием функций ее компонентов

ПО "СХД Енисей" обеспечивает работу двухконтроллерного аппаратного комплекса в режиме Active-Active для достижения повышенной производительности, а также повышенной отказоустойчивости работы системы хранения данных. При аппаратном сбое одного из контроллеров, второй узел должен «подхватывать» его работу без остановки всей системы.

Физическая архитектура ПАК «СХД «Енисей» предусматривает дублирование внутренних аппаратных компонентов и каналов передачи данных для обеспечения отказоустойчивости.

Накопитель подключается к каждому из контроллеров СХД по двум линиям PCIe gen4.

Физическая архитектура ПАК «СХД «Енисей» показана на рисунке 1.

ПО «СХД «Енисей» нацелено на оптимальную и безотказную работу всех элементов СХД. Структура взаимодействия всех компонентов ПО «СХД «Енисей» представлена на рисунке 2.

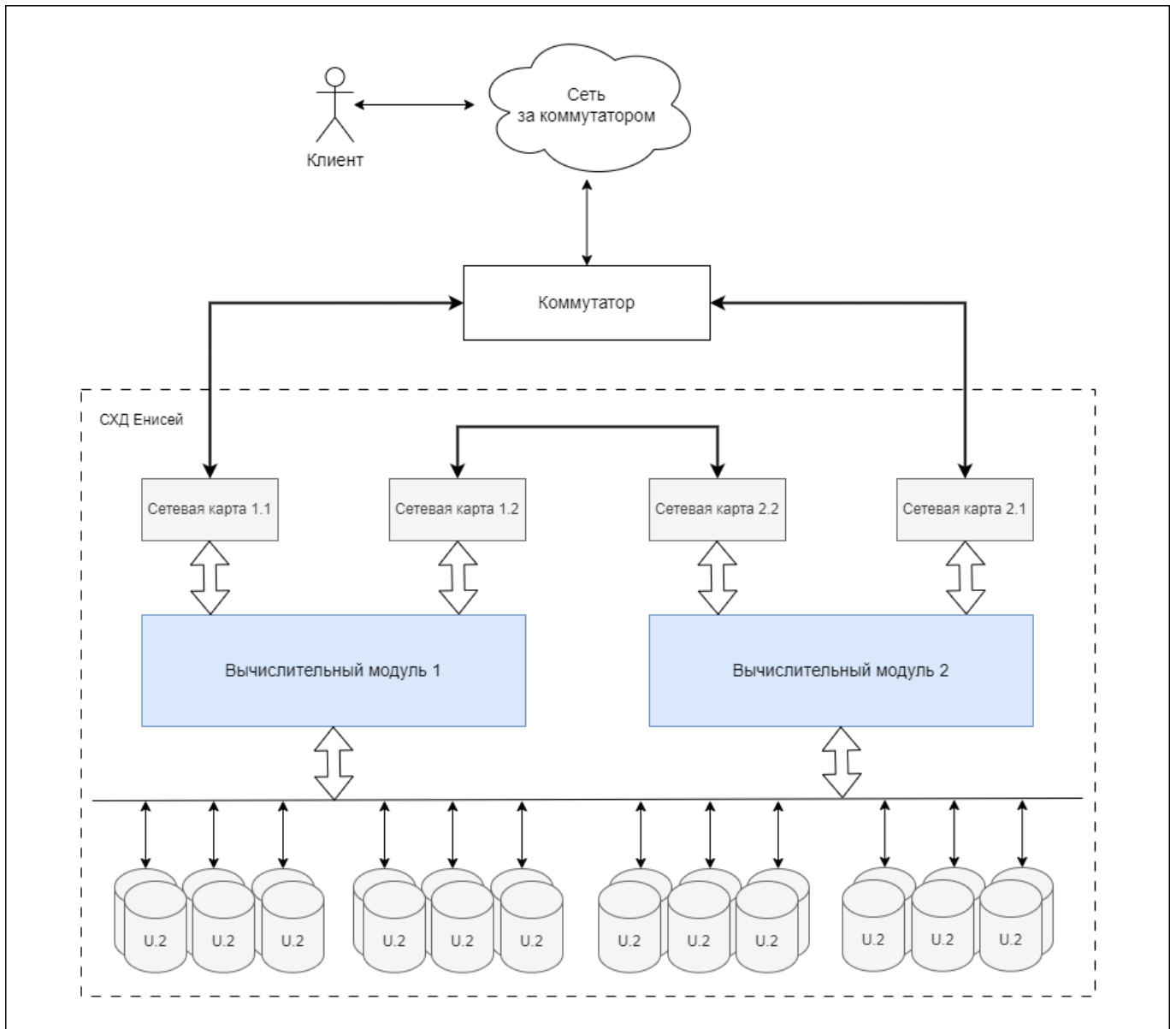


Рисунок 1 – Физическая архитектура ПАК «СХД «Енисей»

Аппаратная часть ПАК «СХД «Енисей» состоит из двух вычислительных модулей (далее ВМ), каждый из которых имеет как минимум по две сетевых карты.

Одна пара из сетевых карт, по одной на каждом из ВМ, соединяются друг с другом чтобы произвести объединение ВМ в вычислительный кластер.

Другие сетевые карты, как минимум по одной на каждом из ВМ, подключаются к одному или двум единицам коммуникационного оборудования, тем самым, обеспечивая резервирование канала связи с вычислительным кластером (через один или через другой ВМ).

Также такое соединение обеспечивает возможность распараллеливания операций чтения и записи данных или любых других, предусмотренных данным ПАК.

Каждый VM подключен к одним и тем же твердотельным двухпортовым дисковым накопителям, каждый накопитель связан через порт 1 с VM1 и через порт 2 с VM2.

Общее количество накопителей может быть от 6 до 24.

Такой тип накопителей и их подключения позволяет продолжить функционирование ПАК в случае выхода из строя одного из VM.

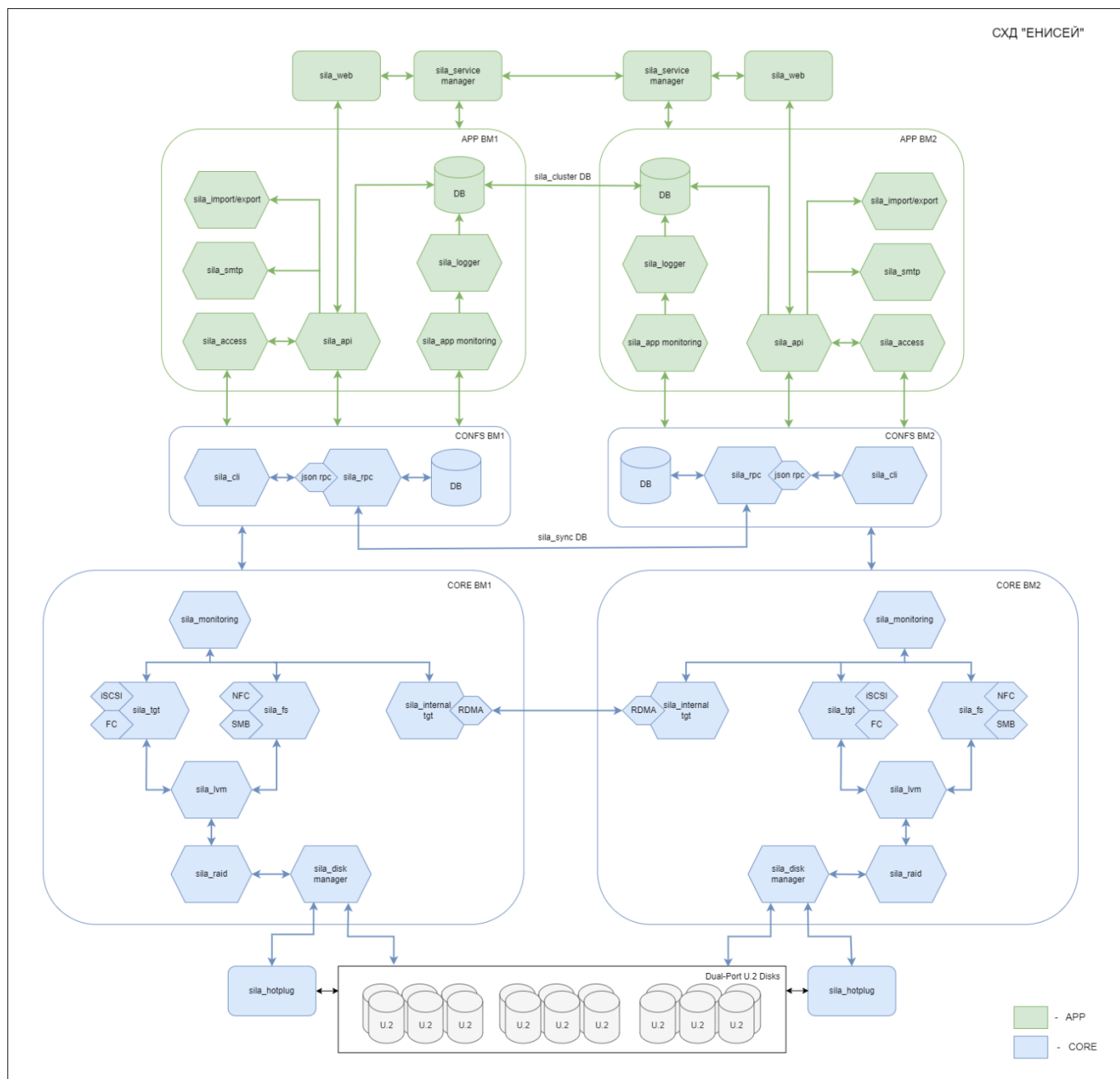


Рисунок 2 – Структурная схема ПО «СХД «Енисей»

В качестве основной Операционной Системы (далее ОС) на каждом из VM установлена операционная система на базе ядра Linux.

На каждом из VM установлено ПО «СХД «Енисей».

ПО «СХД «Енисей» состоит из двух основных компонентов:

- Программного обеспечения управления и администрирования СХД (APP);
- Ядро системы СХД (CORE) - системное программное обеспечение, основной функцией которого является арбитраж функций ввода вывода (чтения и записи) между клиентом и накопителями.

Данные компоненты взаимодействуют друг с другом посредством предопределенного программного интерфейса (далее API).

Основу программного обеспечения составляют модули, базирующиеся на микро сервисной архитектуре.

Ключевые модули ядра системы:

1. CORE – базовый модуль программы, обеспечивающий основу работы СХД в части функций ввода вывода (чтения и записи) между клиентом и накопителями. Функционирование обеспечивается следующими микро сервисами:
  - **sil\_a\_disk manager** – отвечает за работу с накопителями данных;
  - **sil\_a\_raid** – обеспечивает создание и управление массивов данных RAID;
  - **sil\_a\_lvm** – обеспечивает создание и управление логическими томами;
  - **sil\_a\_tgt** – отвечает за предоставление блочных устройств по протоколам iSCSI и FC;
  - **sil\_a\_fs** – отвечает за предоставление файлового доступа по протоколам NFS и SMB;
  - **sil\_a\_internal tgt** – обеспечивает взаимодействие между вычислительными модулями по протоколу RDMA;
  - **sil\_a\_monitoring** – обеспечивает функцию непрерывного мониторинга ядра системы СХД.
2. CONFES – модуль программы, отвечающий за хранение и синхронизацию настроек конфигурации, а также предоставляет API для взаимодействия с внешними модулями и системами. Функционирование обеспечивается следующими микро сервисами:
  - **DB** – база данных для хранения настроек конфигурации.

- **sil\_a\_rpc** – предоставляет API для взаимодействия с ядром системы по протоколу JSON-RPC;
  - **sil\_a\_cli** – предоставляет интерфейс командной строки для управления СХД.
3. **sil\_a\_hotplug** – модуль программы, обеспечивающий поддержку функции горячей замены накопителей данных.

Основные модули программного обеспечения управления и администрирования СХД:

1. APP – комплекс микро сервисов, обеспечивающий взаимодействие с ядром системы и управления СХД. Состоит из следующих элементов:

- **sil\_a\_api** – основной модуль, связывающий компоненты сервиса контроля и управления конфигурацией CORE и APP;
  - **sil\_a\_access** – сервис управления и контроля файловыми системами и их дальнейшего предоставления по протоколам NFS и SMB;
  - **sil\_a\_app monitoring** – сервис мониторинга и контроля параметров ядра системы СХД;
  - **sil\_a\_logger** – сервис внутреннего журналирования, обеспечивает сбор журналов с сохранением в базу данных;
  - **DB** – кластерная база данных, обеспечивающая хранение пользователей, сессий пользователей, журналов событий и метрик производительности;
  - **sil\_a\_smtp** – сервис отправки уведомлений на электронную почту;
  - **sil\_a\_import/export** – сервис подключения к внешним блочным ресурсам для импорта и экспорта данных с функцией посекторного копирования;
2. **sil\_a\_web** – веб приложение-клиент контроля управления и мониторинга.
3. **sil\_a\_service manager** - сервис контроля активного вычислительного модуля, управляет запуском и остановкой внутренних сервисов, предоставляет Virtual IP.

#### 2.1.1. Внешнее взаимодействие

Внешнее взаимодействие с ПО «СХД «Енисей» осуществляется по протоколам:

– iSCSI – Internet Small Computer System Interface или интерфейсу малых компьютерных систем интернета. Это транспортный протокол, обеспечивающий передачу протокола хранения данных SCSI по TCP/IP через сетевое соединение, которым является Ethernet. iSCSI работает как метод организации распределенных хранилищ.

– FCP (Fibre Channel Protocol) - транспортный протокол, обеспечивающий передачу протокола хранения данных SCSI по оптоволоконным сетям. Обеспечивает блочный доступ, высокую надёжность и производительность

– SMB/CFIS – протокол общего доступа к сетевым файлам, который, как он реализован в Microsoft Windows, называется протоколом Microsoft SMB. Набор пакетов сообщений, определяющих определенную версию протокола, называется диалектом. Протокол CIFS является диалектом SMB. Как SMB, так и CIFS доступны в нескольких версиях Unix и других операционных системах.

– NFS – протокол сетевого доступа к файловым системам. Позволяет монтировать (подключать) удалённые файловые системы через сеть. NFS абстрагирован от типов файловых систем как сервера, так и клиента. NFS предоставляет клиентам прозрачный доступ к файлам и файловой системе сервера. Протокол NFS осуществляет доступ только к тем частям файла, к которым обратился процесс, и основное достоинство его в том, что он делает этот доступ прозрачным. Это означает, что любое приложение клиента, которое может работать с локальным файлом, с таким же успехом может работать и с NFS-файлом, без каких-либо модификаций самой программы.

### 2.1.2. Внутреннее взаимодействие

Внутреннее взаимодействие компонентов ПО «СХД «Енисей» осуществляется по протоколам:

– RDMA – сетевая технология, обеспечивающая высокую пропускную способность и обмен данными с низкой задержкой, которая сокращает использование ЦП. RDMA поддерживает сети нулевого копирования, позволяя сетевому адаптеру передавать данные непосредственно в память приложения или из нее. RDMA означает, что сетевая карта (физическая или виртуальная) может предоставлять RDMA клиенту RDMA. На базе данной технологии обеспечивается высокоскоростное взаимодействие контроллеров СХД в кластерном решении.

– JSON-RPC – протокол удалённого вызова процедур, использующий JSON для кодирования сообщений. Это очень простой протокол (очень похожий на XML-RPC), определяющий только несколько типов данных и команд. JSON-RPC поддерживает уведомления (информация, отправляемая на сервер, не требует ответа) и множественные вызовы. На базе данного протокола обеспечена работа с БД, администрирование кластера и контроллеров СХД.

### 2.1.3. Хранение и восстановление данных

Одним из основных подходов для обеспечения надежности ПО «СХД «Енисей» является использование избыточного (корректирующего) кодирования данных. Принцип работы такого кодирования – добавление при записи к полезным данным избыточных данных. Избыточные данные используются для обнаружения и исправления ошибок, возникающих при чтении полезных данных.

ПО «СХД «Енисей» поддерживает классическую технологию защиты данных RAID (Redundant Array of Independent Disks – избыточный массив независимых (самостоятельных) дисков) – технология виртуализации данных для объединения нескольких физических дисковых устройств в логический модуль для повышения отказоустойчивости и (или) производительности.

Эта технология позволяет полностью использовать объем всех входящих в пул физических накопителей и выбрать уровень надежности хранения данных путем выделения необходимого избыточного пространства.

ПО «СХД «Енисей» поддерживает следующие базовые уровни защиты данных RAID:

– «RAID 0» (чередование дисков) (рисунок 3) – данный уровень задействует процесс разделения массива данных на единичные блоки с последующим их распределением по нескольким устройствам хранения, таким как жесткие диски («HDD») или твердотельные накопители («SSD»), объединенных в совокупную группу. Блоки распределяются последовательно и содержат каждый свою часть данных на отдельном диске. Поскольку чередование распределяет данные по большему количеству физических дисков, несколько дисков могут получить доступ к содержимому файла, что позволяет быстрее выполнять операции записи и чтения соответственно. Однако, в отличие от других уровней, «RAID 0» не имеет четности. Чередование дисков без данных четности в свою очередь не имеет избыточности или отказоустойчивости. Системы хранения выполняют чередование дисков по-разному. Система может чередовать данные на уровне байтов, блоков или разделов, или чередовать данные на всех или некоторых дисках кластера. Например, система хранения с десятью жесткими дисками может разделить блок размером «64 килобайта»



(«КБ») на первом, втором, третьем, четвертом и пятом дисках, а затем начать снова с первого диска. Другая система может удалить «1 мегабайт» («МБ») данных на каждом из десяти дисков, прежде чем вернуться на первый диск, чтобы повторить процесс.

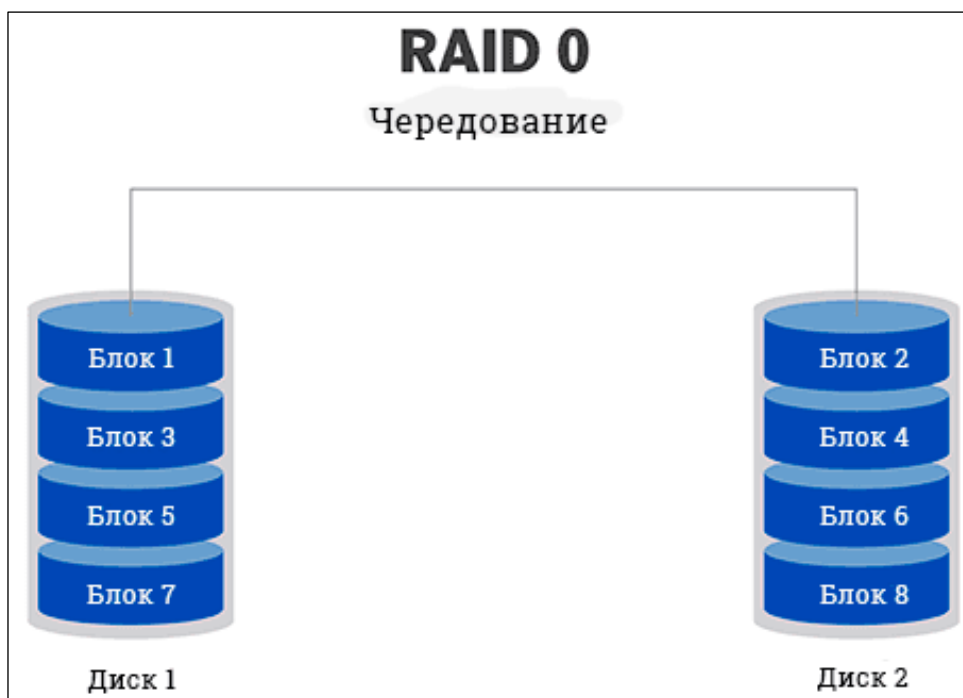


Рисунок 3 – Уровень «RAID 0»

– «RAID 1» (рисунок 4) – зеркальное отображение дисков, представляет собой обязательное размещение данных на два или более дисков, составляющих единый массив, с последующей синхронизацией содержимого копий. Однако функция чередования в данном варианте построения массива запоминающих устройств отсутствует. Зеркальное отображение диска – хороший выбор для приложений, требующих высоких значений производительности и доступности, например, ответственных за финансовые операции, отправку и получение электронной почты, функционирование операционной системы. Зеркальное отображение успешно работает как с жесткими дисками, так и с твердотельными накопителями.

Поскольку каждый диск, из входящих в массив, находится в рабочем состоянии, данные могут считываться с них одновременно, что значительно увеличивает скорость чтения. Массив «RAID 1» будет полноценно функционировать даже если останется работоспособным только один диск. Операции записи, однако, медленнее, потому что каждая операция записи выполняется отдельно для каждого диска (в случае наличия массива из двух дисков запись будет выполнена дважды).

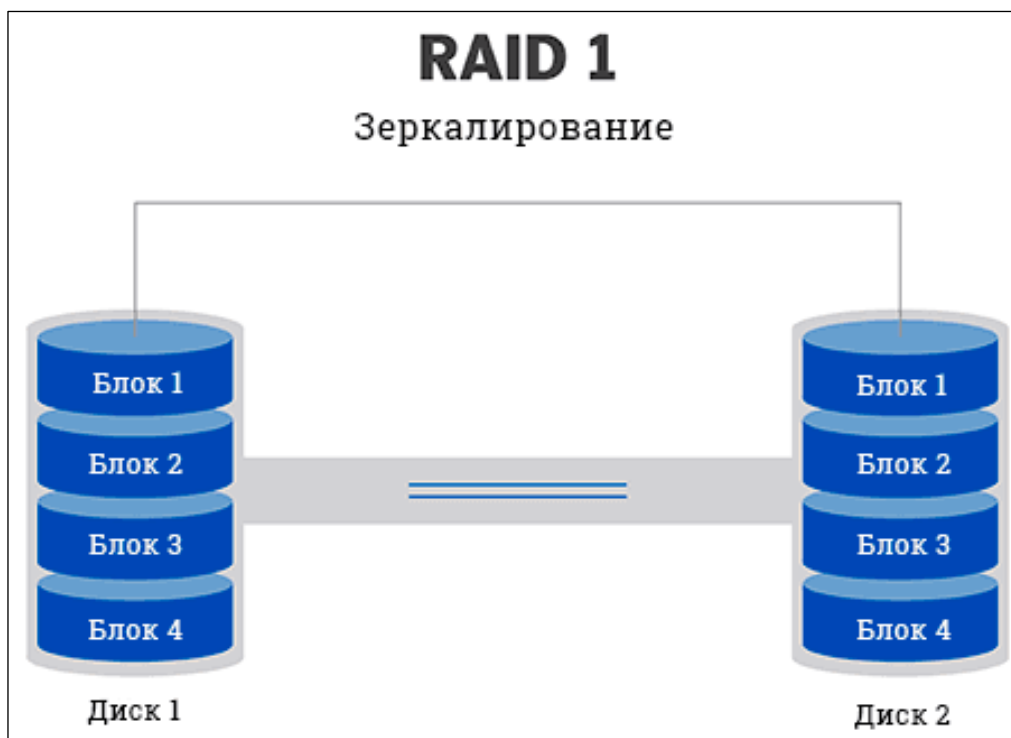


Рисунок 4 – Уровень «RAID 1»

– «RAID 5» (рисунок 5) – избыточный массив конфигурации независимых дисков подобного формата основан в соответствии с принципом чередования на уровне блоков четности. Данные и информация о четности распределяются равномерно по всем дискам, что влечет за собой снижение критичной уязвимости отдельного запоминающего накопителя из организованного объединения, и позволяет массиву «RAID 5» функционировать даже в случае сбоя какого-нибудь одного диска, а также даже восстановить данные за счет чередования, утраченные с проблемного устройства. В «RAID 5» информация о четности хранится по диагонали на всех дисках массива. В случае отказа одного из дисков исходные данные рассчитываются на основе информации о четности, оставшейся на остальных действующих запоминающих устройствах набора. Преимущества «RAID 5» в первую очередь заключаются в его совместном использовании чередования дисков и четности, что позволяет добиться существенного повышения пропускной способности и продуктивности совместно с высоким уровнем надежности.

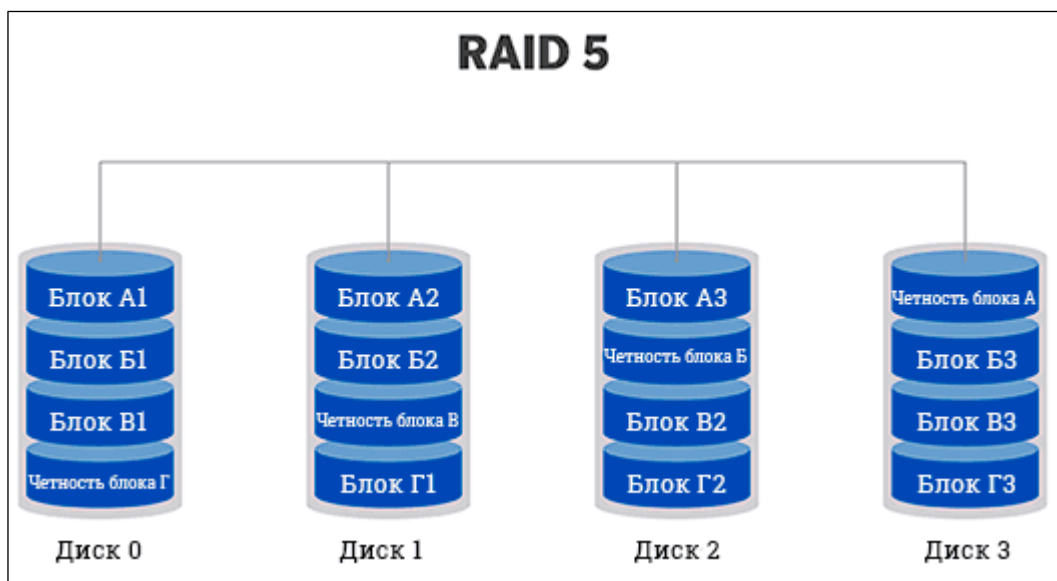


Рисунок 5 – Уровень «RAID 5»

### 2.1.3.1. Структура пула

Уровень надежности хранения данных в пуле определяется схемой резервирования данных. Пул содержит набор накопителей только одного вида и объема. Такой подход гарантирует возможность восстановления данных пула в случае отказа накопителей.

При создании нового пула доступны следующие схемы резервирования данных:

- «RAID 0» – Используется для хранилища, наличие защиты от потери данных в котором не критично, но требуется поддержка высокоскоростного режима чтения и записи. Кэширование потокового видео в реальном времени и редактирование видео композиций – основные области применения уровня построения массива «RAID 0» по причине востребованности в повышенных показателях скорости и производительности. Чередувание дисков без избыточности данных также может быть использовано для временных данных или в ситуациях, когда основная копия данных легко восстанавливается с другого устройства хранения.

- «RAID 1» – Используется для сценариев аварийного восстановления, поскольку обеспечивает мгновенное восстановление требуемых, для критически важных приложений, данных сразу после отказа диска. Если первичные диски в массиве повреждены или не могут быть использованы, произойдет переключение на вторичные или зеркальные резервные диски, обладающие работоспособными копиями, и функционирование приложений, обращающихся к востребованным данным, не будет нарушено.

- «RAID 5» – Используется для равномерного и сбалансированного выполнения операций чтения и записи на нескольких дисках, и, в настоящее время, данный вид массива

является одним из наиболее часто используемых форматов организации «RAID». Он имеет более полезное хранилище, чем конфигурации «RAID 1», и обеспечивает уровень работоспособности, эквивалентный параметрам массива «RAID 0». Группы «RAID 5» имеют минимум три жестких диска, но часто, из соображений эффективности, рекомендуется использовать как минимум пять дисков. Поскольку данные о четности распределены по всем дискам, «RAID 5» считается одной из самых безопасных конфигураций «RAID». ПО «СХД «Енисей» предусматривает следующие схемы резервирования на базе «RAID 5»:

- Объемный (5+1), 5 дисков под полезную нагрузку, 1 диск под четность;
- Высоконадежный (2+1), 2 диска под полезную нагрузку, 1 диск под четность;
- Сбалансированный (3+1), 3 диска под полезную нагрузку, 1 диск под четность.

Пулы с одинаковой схемой резервирования данных могут быть объединены в группу – Storage Pool. Данное объединение позволяет создавать единые расширяемые массивы данных.

#### 2.1.3.2. Размещение данных

Пул резервирует пространство хранения данных (логический том) для использующих его ресурсов. Поддерживаются следующие типы томов – «толстый» (Thick Provision), «тонкий» (Thin Provision).

При использовании «толстого» способа ресурс резервирует в пуле пространство хранения фиксированного размера и полностью предоставляет его хост-серверу.

При использовании «тонкого» способа ресурс резервирует в пуле пространство хранения, определяя его размер автоматически, с учетом заполнения свободного пространства пула и требований хост-сервера, использующего ресурс. Такой подход обеспечивает более эффективное и экономичное распределение пространства хранения, тем самым снижая затраты на администрирование ПО «СХД «Енисей».

Пространство каждого накопителя в пуле разделено на следующие части:

- пространство хранения данных;
- пространство хранения метаданных;
- резервное пространство.

Принцип хранения данных отображен на рисунке 6.



Рисунок 6 – Хранение данных

Пространство хранения данных используется для хранения пользовательских и избыточных данных. Это пространство разделено на одинаковые секции. ПО «СХД «Енисей» автоматически подбирает размер секций при создании пула.

Пространство метаданных содержит информацию о принадлежности секций накопителя определенному пулу.

Резервное пространство используется для восстановления данных при отказе части накопителей пула. В это пространство записываются данные отказавших накопителей, формируемые RAID на основе данных исправных накопителей. Для обеспечения гарантированной сохранности данных пула объем этого пространства выбирается равным совокупному объему максимального числа накопителей, которые могут отказать одновременно.

### 2.1.3.3. Восстановление данных

Восстановление данных происходит в фоновом режиме, когда пользовательские запросы к соответствующим блокам данных отсутствуют.

Технология RAID восстанавливает данные следующим образом:

- в резервном пространстве свободных накопителей пула выделяются новые секции;
- в новых секциях производится формирование данных отказавших накопителей на основе данных исправных накопителей;
- если блоки из одной цепочки записаны на накопители разных групп, запускается процесс перегруппировки данных;
- после замены отказавшего накопителя запускается процесс перемещения данных для равномерного использования накопителей.

В процессе восстановления данных (recovery), зарезервированное под восстановление данных пространство будет уменьшаться. После замены вышедшего из строя накопителя и прохождения процедуры восстановления данных, объем резервного пространства вернется к начальному значению на момент создания пула.

#### 2.1.3.4. Рекомендации по схемам защиты данных

При создании пула данных в ПО «СХД «Енисей» следует учитывать, что в массив дисков «RAID 0», при использовании, блоки данных записываются по очереди на каждый диск, что существенно увеличивает производительность операций чтения и записи. Избыточность данных отсутствует, поэтому выход из строя любого диска ведет к потере данных всей RAID-группы.

При создании пула данных в ПО «СХД «Енисей» следует учитывать, что в массив дисков «RAID 1», при использовании, блоки данных записываются на все диски одновременно, поэтому допустим выход из строя всех дисков, кроме одного. Данная конфигурация обладает наименьшей эффективностью использования дискового пространства, и, как следствие, наибольшей стоимостью при использовании трех и более дисков.

При создании пула данных в ПО «СХД «Енисей» следует учитывать, что в массив дисков «RAID 5», при использовании, блоки данных записываются с поблочным чередованием с одной контрольной суммой. При построении RAID 5 для контрольных сумм не выделяется отдельный диск, вместо этого данные циклически записываются на все диски. «RAID 5» позволяет производить параллельную запись, что существенно увеличивает производительность. Нагрузка на диск четности равномерно распределяется по всем дискам в массиве. «RAID 5» обеспечивает базовый уровень надежности, допустим выход из строя одного диска.

#### 2.1.4. Блочный доступ к данным

Доступ к данным в ПО «СХД «Енисей» предоставляется на уровне логических томов, а не пулов.

Приложение воспринимает блочный том как блочное хранилище данных заданного размера. Каждый логический том подключен к портам ввода-вывода ПО «СХД «Енисей», относящихся к контроллерам хранения. Чтение и запись данных с логического тома влияет только на загрузку портов ПО «СХД «Енисей», связанных с этим ресурсом.

Каждый ресурс характеризуется следующими идентификаторами:

- WWID,
- LUN ID.

WWID – глобальный идентификатор логического тома. Данный идентификатор формируется системой автоматически.

LUN ID – уникальный индекс логического тома. Этот идентификатор формируется системой автоматически (индекс ресурса по умолчанию) либо может быть назначен пользователем вручную.

Созданный блочный том становится немедленно доступным как SCSI LUN после предоставления доступа к нему инициатору.

#### 2.1.5. Файловый доступ к данным

Наряду с предоставлением доступа к данным на блочном уровне, ПО «СХД «Енисей» выполняет функции файлового обслуживания и хранения файловых данных.

ПО «СХД «Енисей» поддерживает доступ к файловым данным по протоколам прикладного уровня NFS или SMB/CIFS через каналы внешних сетей Ethernet.

Поддерживаемые версии протоколов доступа:

- SMB v2.4;
- NFS v4.

Файловый ресурс выполняет роль логической файловой системы и обеспечивает сетевой доступ к файлам. Приложение хост-сервера воспринимает файловый ресурс как файловое хранилище заданного размера. В основе каждого файлового ресурса лежит блочный ресурс.

Файловые ресурсы доступны локальным пользователям и группам, а также пользователям и группам LDAP. Администратор ПО «СХД «Енисей» разграничивает видимость ресурсов и определяет тип доступных операций для каждого пользователя или группы пользователей (чтение или чтение-запись).

ПО «СХД «Енисей» поддерживает разграничение доступа на уровне подсетей. Подсеть определяет диапазон IP-адресов хост-серверов, которые могут подключаться к заданным файловым ресурсам. Данный способ используется для разграничения доступа хост-серверов под управлением ОС Linux.

#### 2.2. Связи программы с другими программами

ПО «СХД «Енисей» поддерживает интеграцию со следующими внешними службами:

- служба каталогов LDAP – для добавления учетных записей пользователей из внешних баз данных;
- служба точного времени NTP – для синхронизации времени между ПО «СХД «Енисей» и остальными службами;

## 2.3. Пользователи и группы

### 2.3.1. Учетные записи пользователей

ПО «СХД «Енисей» поддерживает два типа учетных записей:

- локальные учетные записи, информация о которых хранится в системе;
- внешние учетные записи, информация о которых импортируется из службы каталогов LDAP.

По умолчанию в ПО «СХД «Енисей» создана одна локальная учетная запись «Администратор». Пользователь данной учетной записи может создать новые локальные учетные записи, а также настроить импорт учетных записей из внешней службы каталогов LDAP.

### 2.3.2. Группы пользователей

По аналогии с учетными записями в ПО «СХД «Енисей» поддерживается два типа групп пользователей: локальные и внешние.

По умолчанию в ПО «СХД «Енисей» созданы локальные группы – «Администратор» и «Мониторинг». Члены группы «Администратор» обладают возможностью использования всех функций в интерфейсах управления: управления учетными записями и группами пользователей (создание, удаление и редактирование), доступом к файловым ресурсам (в рамках заданных разграничений), возможность просмотра информации о конфигурации и состоянии ПО «СХД «Енисей», а также ее компонентов в интерфейсах управления.

Пользователи из группы «Администратор» могут создавать, редактировать и удалять локальные группы.

Пользователи из группы «Администратор» могут настраивать импорт групп из внешней службы каталогов LDAP.

Пользователи из группы «Мониторинг» обладают ограниченными правами и имеют лишь возможность просмотра данных СХД.



### 3. НАСТРОЙКА ПРОГРАММЫ

#### 3.1. Состав и содержание дистрибутива

Дистрибутив для установки программы находится на ОД «Программное обеспечение «Система хранения данных «Енисей». Дистрибутив» 643.РЕНМ.02.05-01, содержит архив: YNSY\_x.x.x.img.tar

#### 3.2. Настройка подключения к БД

В качестве БД в ПО «СХД «Енисей» используется нереляционная БД, также известная как база данных NoSQL, в которой для хранения данных используется простой метод «ключ-значение». Это дает возможность хранить данные как совокупность пар «ключ-значение», в которых ключ служит уникальным идентификатором. Как ключи, так и значения могут представлять собой что угодно: от простых до сложных составных объектов. Базы данных «ключ-значение» (или хранилища «ключ-значение») поддерживают высокую разделяемость и обеспечивают беспрецедентное горизонтальное масштабирование, недостижимое при использовании других типов БД. БД не имеет единого узла отказа, все данные реплицированы на всех контроллерах СХД кластера и автоматически синхронизируются.

Поскольку БД не ориентирована на хранение большого количества данных, а скорее на метаданные – настройки СХД, основным фокусом разработки является:

- контроль когерентности и актуальности данных на контроллерах СХД – синхронизация и версионирование;
- администрирование контроллеров СХД кластера – восстановление соединения, ограничение доступа к контроллерам СХД вне кластера;
- уведомление контроллеров СХД кластера об изменении данных, топологии соединения.

БД входит в состав ПО «СХД «Енисей», подключение осуществляется автоматически при старте системы.

#### 3.3. Установка программы и ее компонентов

ПО «СХД «Енисей» входит в состав ПАК «СХД «Енисей» и поставляется предустановленной в составе ПАК.

В текущей версии ПО «СХД «Енисей» не предусматривается активация, выпуск, распространение, управление лицензионными ключами.

### 3.4. Настройка программы на состав ТС и ПС

Произвести проверку доступности информационных портов, обеспечивающих работоспособность ПО «СХД «Енисей» (таблица 2).

Таблица 2 – Информационные порты, обеспечивающие работоспособность ПО «СХД «Енисей»

| Порт  | Протокол | Сервис   |
|-------|----------|--|
| 80    | HTTP     | Веб – интерфейс администратора ПО «СХД «Енисей»      |
| 443   | HTTPS    | Веб – интерфейс администратора ПО «СХД «Енисей»      |
| 12088 | HTTP     | API администрирования кластера ПО «СХД «Енисей»      |
| 2348  | TCP      | API администрирования кластера ПО «СХД «Енисей»      |
| 4420  | RDMA     | Порт кластерного соединения ПО «СХД «Енисей»         |
| 3260  | iSCSI    | Взаимодействие по протоколу iSCSI с ПО «СХД «Енисей» |
| 445   | SMB      | Взаимодействие по протоколу SMB с ПО «СХД «Енисей»   |
| 2049  | NFS      | Взаимодействие по протоколу NFS с ПО «СХД «Енисей»   |

### 3.5. Настройка программы на условия конкретного применения

Для подключения к ПО «СХД «Енисей» на АРМ оператора должен быть установлен веб-браузер Google Chrome.

Для установки соединения с веб-интерфейсом ПО «СХД «Енисей» необходимо в адресной строке ввести IP-адрес подключенного к локальной сети контроллера СХД, ввести имя и пароль учетной записи и нажмите Войти.

Подробное описание работы по настройке ПО «СХД «Енисей» изложены в Руководстве оператора.

#### 4. ПРОВЕРКА ПРОГРАММЫ

##### 4.1. Проверка работоспособности

Проверка работоспособности программы выполняется путем запуска ПО «СХД «Енисей».

Для запуска ПО «СХД «Енисей» на необходимо выполнить следующие действия:

- включить ПО «СХД «Енисей»
- выполнить подключение ПО «СХД «Енисей» к локальной сети;
- установить соединения с веб-интерфейсом ПО «СХД «Енисей».

При успешном выполнении вышеуказанных действий будет отображена страница авторизации ПО «СХД «Енисей» (рисунок 7).

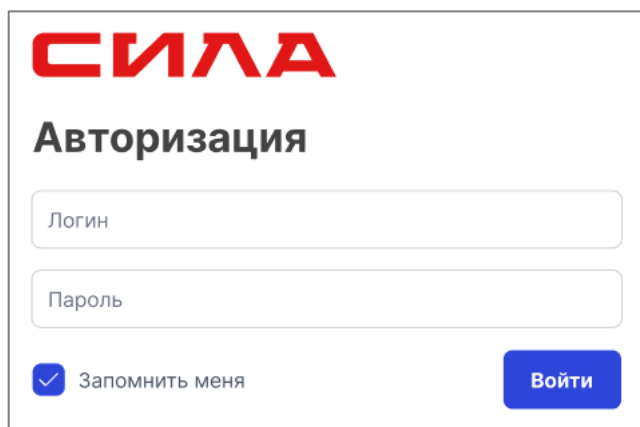


Рисунок 7 – Выполнение получения информации о контроллерах СХД

Программа работоспособна, если выполнены следующие условия:

- успешно пройдена авторизация;
- рабочее окно отображает главную страницу веб-интерфейса ПО «СХД «Енисей» и стартовые страницы его компонентов;
- отсутствуют сообщения об авариях и сбоях в работе.

##### 4.2. Методы проверки

###### 4.2.1. Общие положения

Проверка программы производится методом выполнения базовых сценариев ПО «СХД «Енисей»:

- «Получение информации о контроллерах СХД»;
- «Проверка доступности накопителей данных»;
- «Получение информации о сетевых интерфейсах»;
- «Получение информации о статусе кластера».

ПО «СХД «Енисей» должно быть подключено к локальной сети предприятия и запущено.

#### 4.2.2. Получение информации о контроллерах СХД

Сценарий описывает получение информации о контроллерах СХД ПО «СХД «Енисей» подключенных к локальной сети предприятия.

Шаги сценария выполняются с помощью CLI WEB-интерфейса в меню «Терминал» (рисунок 8).

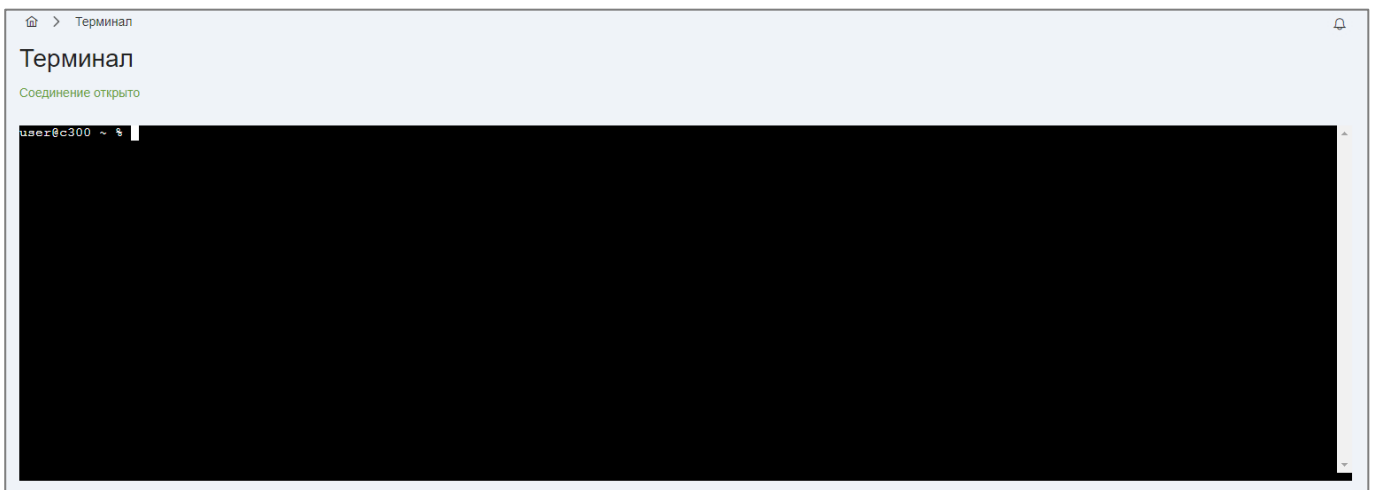


Рисунок 8 – Меню «Терминал» WEB-интерфейса

Сценарий предполагает получение информации о контроллерах СХД в CLI:

- войти в WEB-интерфейс ПО «СХД «Енисей» с правами администратора.
- в CLI из меню «Терминал» для получения списка контроллеров ввести команду:

```
node list
```

Результат представлен на рисунке 9.

```
user@c300 ~ % node list
name      nodeid                address      state
-----
c300     355FB998-BFD2-3F59-86F1-F5869C730D15  192.168.0.1  2
c301     A4C8272A-3D4F-37D1-9924-D1FE4225EE12  192.168.0.2  2
user@c300 ~ %
```

Рисунок 9 – Список контроллеров

- получить информацию о выбранном контроллере СХД командой:

```
node info -n c301
```

Результат представлен на рисунке 10.

```

user@c300 ~ % node info -n c301
name          c301
nodeid        A4C8272A-3D4F-37D1-9924-D1FE4225EE12
model
serial
address       192.168.0.2
state         2
version       0.1.1.555
start_date    Thu Jun  6 12:55:37 2024
user@c300 ~ % █

```

Рисунок 10 – Информация о контроллере

Получение информации свидетельствует о штатной работе контроллера СХД.

#### 4.2.3. Проверка доступности накопителей данных

Сценарий описывает получение информации о накопителях ПО «СХД «Енисей».

Шаги сценария выполняются с помощью CLI WEB-интерфейса в меню «Терминал» (см. рисунок 8).

Сценарий предполагает получение информации о накопителях ПО «СХД «Енисей» в CLI:

- войти в WEB-интерфейс ПО «СХД «Енисей» с правами администратора.
- в CLI из меню «Терминал» для получения списка накопителей ввести команду:

```
disk list
```

Результат представлен на рисунке 11.

```

user@c300 ~ % disk list
name          nguid                                size  used  state  node
-----
VirtioBlk0    4e7ba8a6-7156-5a5c-9808-e20893ef2c89 2GiB  pool_2 online c300
VirtioBlk0    11d9e45b-0a37-5390-9903-abf4f7fcb59f 2GiB  pool_2 online c301
VirtioBlk1    82480e03-b701-535c-9cdb-9dea57376177 2GiB  pool_2 online c300
VirtioBlk1    5b67085c-ebea-5756-bc8a-c0ce994a21f8 2GiB  pool_2 online c301
VirtioBlk2    050e2adf-cb1c-5e65-b296-1fa53d128c1e 2GiB  pool_2 online c300
VirtioBlk2    efee342d-7821-5470-90d7-2b462f1f1666 2GiB  pool_2 online c301
VirtioBlk3    f79b55fe-d922-57e2-bf4c-0ecfa20e7e4d 2GiB  pool_2 online c300
VirtioBlk3    efc94b3d-a634-5023-bc5f-13a1f6a7c6a4 2GiB  pool_2 online c301
VirtioBlk4    05afd55c-ebfd-5050-ac72-0b84d8deb66d 2GiB  pool_2 online c300
VirtioBlk4    3265012e-da55-5617-ab5c-8c5b01748d3c 2GiB  pool_2 online c301
user@c300 ~ % █

```

Рисунок 11 – Список дисков

- получить информацию о выбранном накопителе СХД командой:

```
disk info VirtioBlk1
```

Результат представлен на рисунке 12.

```

user@c300 ~ % disk info VirtioBlk1
name                VirtioBlk1
nguid               82480e03-b701-535c-9cdb-9dea57376177
node                c300
size                2GiB
used                pool_2
model_number
serial_number
firmware_revision
data_units_read
data_units_written
percentage_used
temperature_celsius
media_errors
num_err_log_entries
warning_temperature_time_minutes
critical_composite_temperature_time_minutes
user@c300 ~ % █

```

Рисунок 12 – Информация о диске

Получение информации свидетельствует о том, что накопители подключены к контроллерам СХД. Статус «online» означает, что накопители данных доступны.

#### 4.2.4. Получение информации о сетевых интерфейсах

Сценарий описывает получение информации о сетевых интерфейсах ПО «СХД «Енисей».

Шаги сценария выполняются с помощью CLI WEB-интерфейса в меню «Терминал» (см. рисунок 8).

Сценарий предполагает получение информации о сетевых интерфейсах ПО «СХД «Енисей» в CLI:

- войти в WEB-интерфейс ПО «СХД «Енисей» с правами администратора.
- в CLI из меню «Терминал» для получения списка сетевых интерфейсов ввести команду:

```
iface list
```

Результат представлен на рисунке 13.

```

user@c300 ~ % iface list
name  speed  address      netmask      type          state         node
-----
eth0  192.168.0.1  255.255.255.0  cluster     connected    c300
eth0  192.168.0.2  255.255.255.0  cluster     connected    c301
eth1  192.168.1.1  255.255.255.0  extern,control  connected    c300
eth1  192.168.1.2  255.255.255.0  extern,control  connected    c301
eth2  10.128.1.226 255.255.255.240 extern,control  connected    c300
eth2  10.128.1.227 255.255.255.240 extern,control  connected    c301
user@c300 ~ % █

```

Рисунок 13 – Список сетевых интерфейсов

- получить информацию о выбранном сетевом интерфейсе СХД командой:

```
iface info eth2
```

Результат представлен на рисунке 14.

```
user@c300 ~ % iface info eth2
mac          52:54:0:b7:a1:2e
speed
address      10.128.1.226
netmask      255.255.255.240
type         12
state        connected
node         c300

user@c300 ~ %
```

Рисунок 14 – Информация о сетевом интерфейсе

Статус «connected» свидетельствует о штатной работе сетевого интерфейса.

#### 4.2.5. Получение информации о статусе кластера

Сценарий описывает получение информации о кластере ПО «СХД «Енисей».

Шаги сценария выполняются с помощью CLI WEB-интерфейса в меню «Терминал» (см. рисунок 8).

Сценарий предполагает получение информации о кластере ПО «СХД «Енисей» в CLI:

- войти в WEB-интерфейс ПО «СХД «Енисей» с правами администратора.
- в CLI из меню «Терминал» для получения информации о кластере ввести команду:

```
cluster info
```

Результат представлен на рисунке 15.

```
user@c300 ~ % cluster info
clid          A6EF8874-F068-4191-A7EB-0F5AD03D3137
name          default
description
creation_date Mon Apr 15 15:42:10 2024
modification_date Mon Apr 15 15:48:29 2024
type          tcp
port          4420

user@c300 ~ %
```

Рисунок 15 – Информация о кластере

Получение информации свидетельствует о наличии созданного кластера.

## 5. СООБЩЕНИЯ СИСТЕМНОМУ ПРОГРАММИСТУ

Сообщения системному программисту представлены в таблице 3.

Таблица 3 – Сообщения оператору

| Сообщение         | Описание                  | Действие                  |
|-------------------|---------------------------|---------------------------|
| Incorrect command | Ошибка синтаксиса команды | Ввести корректную команду |
|                   |                           |                           |



## ПЕРЕЧЕНЬ СОКРАЩЕНИЙ

- API - Application Programming Interface, программный интерфейс приложения
- DB - Data Base, база данных
- FC - Транспортный протокол SCSI в сетях Fibre Channel
- JSON-RPC - Протокол удалённого вызова процедур, использующий JSON для кодирования сообщений.
- iSCSI - Транспортный протокол SCSI в сетях Ethernet
- NFS - Network File System, протокол сетевого доступа к файловым системам
- RAID - Redundant Array of Independent Disks, технология виртуализации данных для объединения нескольких физических дисковых устройств в логический модуль для повышения отказоустойчивости и (или) производительности
- RDMA - Remote Direct Memory Access, аппаратное решение для обеспечения прямого доступа к оперативной памяти другого компьютера при помощи высокоскоростной сети
- SCSI - Small Computer System Interface, представляет собой набор стандартов для физического подключения и передачи данных между компьютерами и периферийными устройствами
- SMB - Server Message Block, сетевой протокол прикладного уровня для удалённого доступа к файлам, принтерам и другим сетевым ресурсам
- BM - Вычислительный модуль
- ОС - Операционная система
- ПАК - Программно-аппаратный комплекс
- ПО - Программное обеспечение
- СХД - Система хранения данных
- ТС - Технические средства

