

АНГАРА

Система хранения данных



Документ содержит следующую информацию:

- Функциональная спецификация
- Техническая спецификация
- Конфигурация продукта и варианты интеграции с инфраструктурой SAN
- Монтаж и установка оборудования

Категория документа

Документ предназначен для публичного доступа

Система хранения данных Ангара® является системой корпоративного класса, которая позволяет обеспечить высокий уровень надёжности для хранения и обработки критичных для бизнеса данных

Unified Storage

- Поддерживает NAS и SAN технологии хранения. Поддерживаются основные mainstream протоколы передачи данных – iSCSI, SCSI, S3, NFS

Высокая скорость доступа к данным

- В аппаратной платформе используются 64-битные многоядерные микропроцессоры, позволяющие организовать быстрый обмен данными и обеспечить необходимые вычислительные мощности для работы внутренних сервисов системы хранения данных

Ангара предлагает целый набор инструментов для ускорения операций работы с данными, используя кэш-память, SSD (Flash) носители информации, а также интеллектуальные алгоритмы Tiering для перемещения данных между медленными SAS (nearline) и быстрыми SSD носителями информации

Настройка и конфигурация

- Ангара предлагает двухконтроллерную конфигурацию в режиме кластера Active-Active, которая позволяет обеспечить бесперебойную работу потребителей системы хранения данных в режиме 365 дней 24x7. Дополнительно предусмотрена возможность расширения количества контроллеров от 2 до 96 в рамках единой системы хранения данных. Ангара® поддерживает работу в режиме одного контроллера

Quality of Service

- Качество обслуживания клиентов системы хранения данных Ангара обеспечивается архитектурой коммутационной шины ETHERNET, которая позволяет гарантировать время отклика и пропускную способность системы и таким образом защитить важные информационные системы клиента в зависимости от роста нагрузки со стороны пользователей. Дополнительно существуют механизмы изолирования ресурсов, таким образом позволяя резервировать порты ввода-вывода, кэш-память для обслуживания конкретных клиентов системы хранения данных

Отказоустойчивость и надёжность системы

- Режим работы Active-Active подразумевает, что в рамках двухконтроллерной системы сбой одного контроллера не приводит к остановке в обслуживании клиентов, доступ к данным осуществляется через порты ввода-вывода двух контроллеров одновременно
- Используется Adaptive RAID технология, которая позволяет обеспечить защиту от сбоев носителей информации, за счёт архитектуры распределённого контроля чётности. RAID конфигурация подразумевает единовременную установку по 4 диска, в рамках единой отказоустойчивой RAID-группы, позволяя объединять несколько RAID-групп и увеличивать ёмкость хранилища
- Доступность дисков в рамках системы хранения данных обеспечивают резервные каналы передачи данных через специальные порты ввода-вывода System IO, которые объединяются в специальную шину ETHERNET на базе коммутаторов.

Отказоустойчивая шина передачи данных ETHERNET

- Передача данных внутри системы хранения Ангара осуществляется через коммутаторы на базе протоколов xTiger IO и Ethernet. Каждый контроллер имеет основной и резервный канал передачи данных с возможностью увеличения количества резервных подключений до 4 независимых каналов.
- Шина передачи данных ETHERNET используется только для связи контроллеров и JBOD шасси между собой, а также используется в работе RAID, репликации данных, создания мгновенных снимков, а также выполнения операций резервного копирования данных

Защита контроллеров от сбоев системы электропитания

- Используется энергонезависимая батарея, которая в момент полного отключения электропитания обеспечивает безопасное сохранение данных в кэш-памяти на специальном диске system cache.

Защита контроллеров от аппаратных сбоев

- В случае аппаратного сбоя контроллера, данные записанные в кэш-память дублируются в кэш-памяти соседнего контроллера и ввод-вывод продолжается без остановки через соседний контроллер. В качестве защиты от ситуации split-brain, контроллеры имеют резервные каналы связи между собой и непрерывно обмениваются информацией о своём состоянии.
- В случае гарантированной защиты от режима split-brain используются выделенные диски через SAS канал ввода-вывода, который выполняет роль кворума и обеспечивает 100% гарантию работоспособности системы хранения данных в случае аппаратных сбоев.
- Передача данных от диска до клиентской системы защищается с помощью технологии EEDP – End-to-End-Data-Protection, которая гарантирует проверку контроля целостности данных при выполнении передачи данных от диска через каналы ввода-вывода

Защита носителей информации от сбоев

- Система хранения данных использует диагностические команды протокола SCSI и интерфейс S.M.A.R.T для диагностики дисков и обнаружения возможных сбоев дисков. В целях контроля целостности данных используется информация DIF (Data Integrity Field) на самом носителе в рамках каждого сектора размером 512(520)/4096(4106) байт, из которых 8 байт отводится для CRC метки DIF
- В случае преждевременного выхода из строя диска, Ангара анализирует дополнительно параметры работы диска, выявляя возможные отклонения в его работе и инициирует копирование данных на выделенный spare диск, либо в его отсутствие инициирует восстановление данных, используя механизм

Инструменты управления системы хранения данных

- Имеется встроенный в контроллер HTML5 Web GUI инструмент управления всеми операциями системами хранения данных. Поддерживается также режим управления через интерфейс командной строки, установленной на клиентской системе и перенос системы Web GUI на выделенный модуль сервисной консоли

Высокая доступность системы хранения данных (high availability)

- Система хранения данных позволяет создавать мгновенные снимки томов данных (snapshots), реализуя полные копии данных (Full Copy) и копии данных по запросу (Copy on Write). В момент создания копии тома данных, создаётся новый том и указывается направление копирования данных
- Целостность данных в работе мгновенных снимков, обеспечивается за счёт журналирования операций записи данных в специальной области bitmap, расположенной непосредственно в рамках адресного пространства тома, которая гарантирует актуальную копию данных в любой момент времени.
- В случае сценария удалённой репликации данных между системами хранения данных Ангара, используется репликация томов данных, которая работает в синхронном и асинхронном режимах, обеспечивая в первом случае гарантированную копию данных на удалённом томе данных в случае недоступности основной системы хранения данных
- В целях гарантированной доступности данных при использовании удалённой репликации, применяется миграция LUN в прозрачном режиме для клиента, таким образом удалённая система хранения данных обслуживает нагрузку связанную с конкретным LUN, пока на основной системе хранения данных выполняются профилактические работы

Безопасность данных

- Доступ к данным осуществляется с помощью механизма selective storage presentation, в рамках которого назначаются права доступа для клиентских систем к заданным LUN в системе хранения данных
- Любые операции по предоставлению доступа к данным дублируются в журнале администратора безопасности и доступны для анализа в любой момент времени.
- Данные передаваемые по открытым каналам передачи данных могут кодироваться в соответствии с алгоритмом AES

Используемые термины и сокращения

- **Носитель информации** – физический объект способный хранить информацию на долговременной основе без использования дополнительных устройств (в данном случае понимается Hard Disk Drive или Solid State Drive в форм-факторе 2.5 или 3.5 дюйма)
- **Система хранения данных** – программно-аппаратный комплекс представляющий собой комплекс, состоящий из аппаратного обеспечения с установленными носителями информации и программного обеспечения, позволяющего получить интерфейс для доступа и хранения данных
- **ПО СХД** – Программное обеспечение системы хранения данных
- **Программно-определяемая система хранения данных Ангара®** – программное обеспечение устанавливаемое на аппаратную платформу с целью получить заданный функционал для безопасного и долговременного хранения данных
- **PCIe** – Peripheral Component Interconnect Express шина передачи данных
- **RAID** – Redundant Array of Independent Disks, методы объединения группы носителей информации с целью получения единого пространства для хранения информации с заданными параметрами отказоустойчивости
- **RAID группа** – группа носителей информации объединённая в единый логический элемент с заданным именем с целью получения единого адресного пространства с заданными характеристиками распределения данных контроля чётности
- **Логический том (Volume)** – виртуальное и неделимое адресное пространство, размещённое в рамках адресного пространства одной или нескольких RAID групп
- **Ангара** – зарегистрированная торговая марка компании ООО СП «М-Энергоналадка»

1. Архитектура и спецификация оборудования	8
1.1. Основные компоненты	8
1.2. Архитектура и компоненты Ангара	9
1.2.1. Архитектура и дизайн шасси контроллера Ангара	9
1.2.2. Архитектура и дизайн шасси JBOD Ангара	11
1.2.3. Модуль ввода-вывода IO Module	12
1.2.4. Индикаторы состояния системы хранения данных	13
1.2.5. Шасси для установки дисков 2.5	14
1.2.6. Сервисная консоль Ангара	15
1.2.7. Коммутатор Ангара LAN	16
1.2.8. Коммутаторы System IO	17
2. Решения на базе архитектуры Ангара	19
2.1. Катастрофоустойчивые решения	19
2.2. Высокопроизводительные решения	21
2.3. Серверная виртуализация	22
2.4. Миграция данных с помощью External Storage	23
3. Архитектура программного обеспечения	24
3.1. Архитектура системного программного обеспечения Ангара	24
3.2. Архитектура Volume Copy	26
3.3. Архитектура RAID Ангара	27
3.4. Архитектура Thin Provisioning	28
3.5. Обновление программного обеспечения Ангара	28
4. Конфигурирование систем хранения данных Ангара	29
4.1. Базовый блок Ангара	29
5. Программно-определяемая система хранения данных Ангара® StorageOS	30
5.1. Требования к аппаратной платформе для установки ПО СХД	30
5.2. Требования к операционной системе для установки ПО СХД	31
5.3. Функциональные возможности ПО СХД	32
5.3.1. Менеджер управления физическими носителями	32
5.3.2. Менеджер управления логическими томами	33
5.3.3. Менеджер управления доступом к логическим томам в ОС StorageOS®	34
6. Спецификация базового функционала программного обеспечения	35
7. Поддерживаемые операционные системы программно-определяемой ОС StorageOS®	40

Архитектура и дизайн шасси контроллера Ангара

Модель	Контроллер Ангара (шасси)	JBOD (шасси)
Ангара	Форм-фактор 2U шасси 24 диска NVMe/SAS/SATA	Форм-фактор 2U шасси 24 диска NVMe/SAS/SATA



Рисунок 1 – Шасси контроллера Ангара (front view)



Рисунок 2 – Шасси контроллера Ангара (rear view)

Архитектура и дизайн шасси контроллера Ангара

Архитектура шасси построена на активном резервировании основных компонент, включая блоки питания, порты ввода-вывода. В составе контроллера имеется элемент удалённого управления электропитанием и предусмотрена энергонезависимая батарея для защиты кэш-памяти контроллера.

В составе контроллера Ангара имеются 24 независимых диска 2.5'' с интерфейсами NVMe или SAS.

Поддерживаются диски с реализацией однопортовых и двухпортовых вариантов подключения по протоколам SAS/NVMe.



Рисунок 3 – Архитектура контроллера P200

1 – два блока питания Power supply 640W (1+1)

2,3- контроллер



Рисунок 3 – Архитектура контроллера P200

1 – панель индикаторов состояния шасси контроллера

2- носитель информации NVMe/SAS (24 слота) hot swap

3 – сервисный порт USB

Диск system cache (standby)

Диск system cache (PN:DZ-SSD-01) предназначен для хранения конфигурации системы хранения данных и выполняет роль хранилища данных из кэш-памяти в случае экстренных проблем с электропитанием.

Тип диска	Кол-во дисков	Объём диска, Гб	Объём кэш-памяти, Гб
System cache	1	512	8 - 256
System cache	2	512	256 - 512

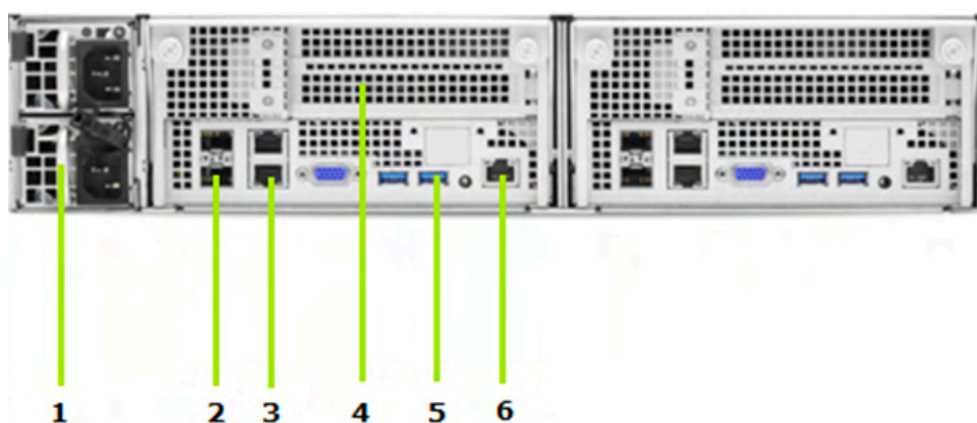


Рисунок 5 – Основные компоненты контроллера P200

- 1 – блок питания 650W (1+1)
- 2- Ethernet SFP+ 10G (2 разъёма) System IO
- 3 – Ethernet RJ45 1G (2 разъёма)
- 4 – слот для установки IO Module (2 слота)
- 5 – USB порт
- 6 – сервисный порт (RJ45)

Архитектура и дизайн шасси JBOD Ангара

Шасси JBOD имеет модульную архитектуру и поддерживает SAS интерфейс на базе модулей ввода-вывода PN:DZ-HBA-01 для объединения в единую систему хранения данных вместе с контроллерами Ангара и другими JBOD модулями. Архитектура Ангара позволяет соединять JBOD в режиме стекирования (daisy-chain) с целью расширения объема дискового пространства.

Ангара поддерживает до 12 последовательно включенных JBOD шасси, каждое из которых имеет как минимум одно подключение к одному из двух контроллеров Ангара.



Рисунок 6 – JBOD шасси Ангара (rear view) для 108 дисков SAS/SATA

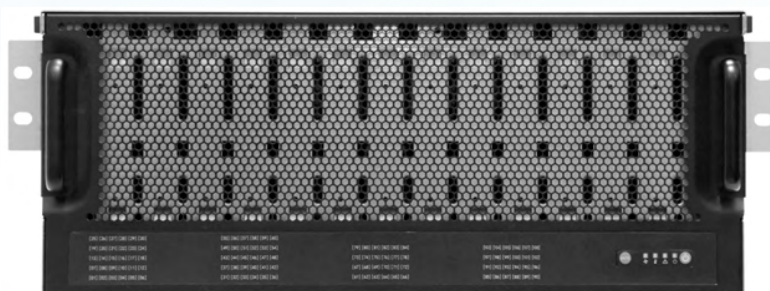


Рисунок 7 – JBOD шасси Ангара (front view) для 108 дисков SAS/SATA



Рисунок 8 – JBOD шасси Ангара (front view) для 24 дисков SAS/SATA



Рисунок 9 – JBOD шасси Ангара (rear view) для 24 дисков SAS/SATA

Модуль ввода-вывода IO Module

Модуль ввода-вывода обеспечивает передачу данных между системой хранения данных и клиентом, используя существующие возможности протоколов передачи данных.

В целях интеграции с сетями хранения данных, используются протоколы Ethernet (iSCSI, CIFS, NFS), SAS.

Основной протокол логического уровня, используемый для передачи данных между системой хранения данных и клиентом, является протокол SCSI, реализованный согласно стандарту США ANSI и спецификации SPC-3/SPC-4.

Протокол SCSI является промышленным стандартом передачи данных, который реализован в большинстве операционных систем и поддерживается для совместимости и интеграции систем хранения данных с IT-инфраструктурой клиентов.

Наименование модуля	Скорость подключения (G – Гигабит)
Ethernet	10G, 100G
SAS	12G, 24G, 36G, 48G



Рисунок 10 – Модуль ввода-вывода 4 port SAS 12Gb

Тип модуля	Поддерживаемые протоколы	Тип соединения	Part number
4 port SAS 12Gb PCIe x8	SAS, SCSI, System IO	SFF-8644 miniSAS HD (copper, optical cables)	#DZ-HBA-01
2 port Ethernet 10Gb (встроенный) PCIe x8	System IO, Ethernet, IP, TCP/IP, HTTP, NFS, CIFS, System IO	SFP+ (copper cable)	#DZ-HBA-04
2 port PCI express 4.0 x16	SystemIO / Ethernet	Mini SAS HD	#DZ-HBA-05

Индикаторы состояния системы хранения данных

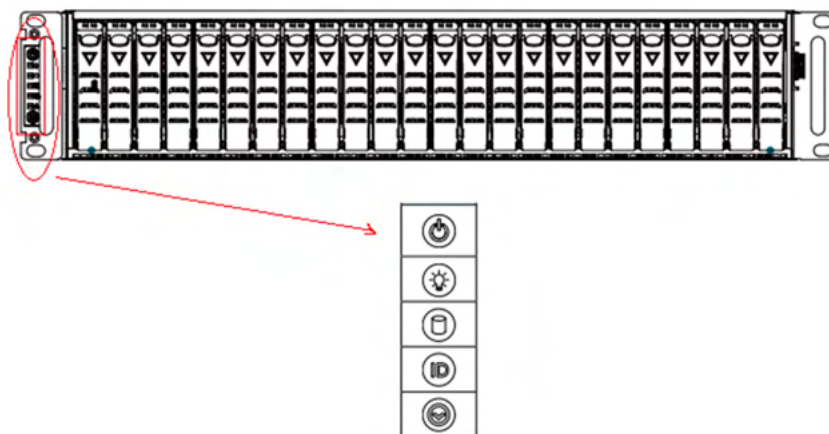


Рисунок 11 – Индикаторы контроллера Ангара

Индикатор LED	Обозначение	Цвет	Назначение
Идентификатор System Power		Зеленый	Включен: Система подключена к сети электропитания Выключен: Система не подключена к сети электропитания
Идентификатор System ID LED		Голубой	Активен: Мигающий сигнал Не активен: Сигнала нет
Кнопка Power button			
Кнопка System Reset button			
Идентификатор Drive Activity LED		Зеленый	Активен: Мигающий сигнал Не активен: Сигнала нет

Шасси для установки дисков 2.5

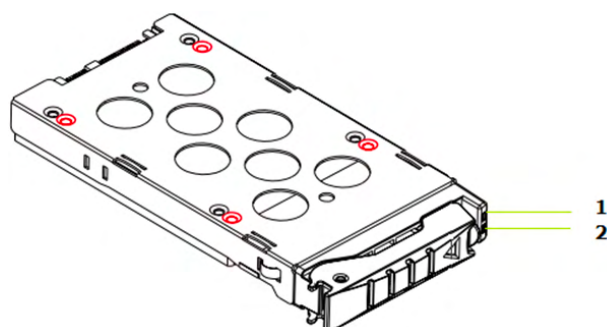


Рисунок 12 – Шасси для установки носителя информации

1 – индикатор активности
носителя информации (голубой)

2 – индикатор состояния носителя
информации (голубой – активен,
чёрный – не активен)

Интерфейс носителя информации	Тип разъёма носителя информации
NVMe	U.2
SATA/SAS	U.2

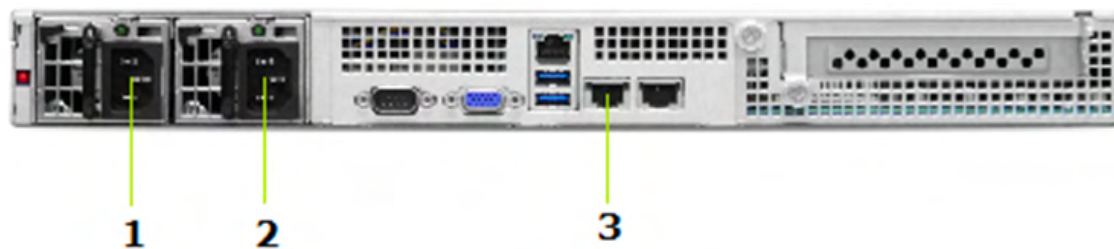
Сервисная консоль Ангара

Сервисная консоль предоставляет защиту от ситуации split-brain в архитектуре Ангара и обеспечивает 100% доступность системы хранения данных в случае аппаратных сбоев контроллера, а также в случае сценария репликации данных между двумя системами хранения Ангара



Рисунок 13 – Сервисная консоль Ангара

Сервисная консоль Ангара (PN:DZ-SP-01) устанавливается в рамках одной системы хранения данных и обслуживает до 96 контроллеров



1 – Блок питания 450W (1+1)

2 – Блок питания 450W (1+1)

3 – RJ45 Ethernet 10G порт сервисной консоли

Сервисная консоль подключается к коммутатору Ангара LAN (PN: DZ-LC-01), который обеспечивает сетевое IP соединение с контроллерами системы хранения данных

Коммутатор Ангара LAN

Коммутатор Ангара LAN (PN:DZ-LC-01) обеспечивает сетевое соединение сервисной консоли и контроллеров системы хранения данных, обеспечивая от 16 до 48 доступных сетевых порта Ethernet 10G

В условиях масштабирования системы хранения данных, необходимо установить дополнительный коммутатор, используя порты 10G Uplink

Количество контроллеров Ангара	Количество коммутаторов LAN
2 – 24	1
24 – 48	2



Рисунок 14 – Коммутатор Ангара LAN 24 порта

- 1 – блок питания 450W
- 2 – сервисный Ethernet RJ45 порт управления коммутатором
- 3 – порт Ethernet 1G
- 4 – порт Ethernet 10G SFP+ (Uplink)

Коммутаторы System IO

Коммутатор System IO используется для расширения дисковой ёмкости системы хранения данных Ангара и выполняет следующие функции:

- Расширение дисковой ёмкости СХД при добавлении новых контроллеров и модулей расширения JBOD
 - Обеспечивает функционирование СХД в режиме NSPOF (no single point of failure - без единой точки отказа



Рисунок 15 – Коммутатор System IO over Ethernet

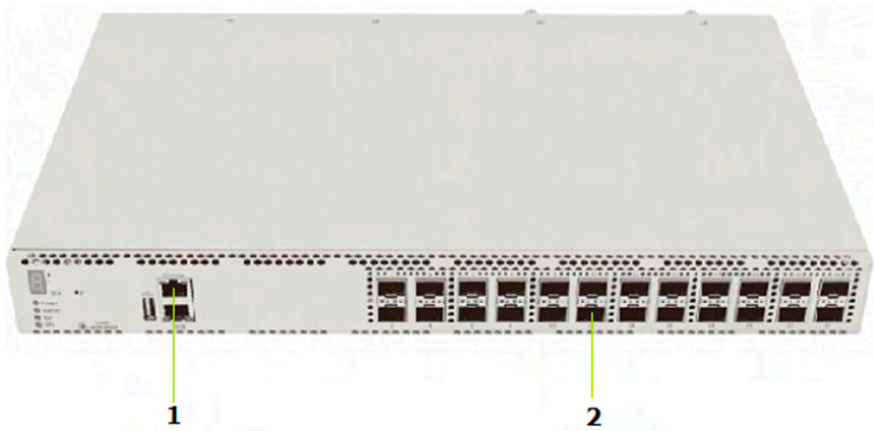
Коммутатор	Количество портов	Протокол	Скорость, Гигабит/с	Part Number
ETHERNET 24 10G+	24	System IO	10	PN: DZ-DSW-02

Коммутатор System IO используется для подключения портов System IO размещённых в контроллере Ангара.

В рамках одной системы хранения данных используется не менее 2 коммутаторов System IO.

Каждый контроллер имеет минимум два одновременных подключения к паре коммутаторов System IO.

Существует возможность расширения портов System IO от 2 до 8 в рамках одного контроллера Ангара, используя модуль ввода-вывода PN:DZ-HBA-04 или PN: DZ-HBA-01.



1 – Сервисный порт Ethernet 1GB

2 – порты System IO SFP+ 10G

Контроллер Ангара имеет встроенную пару портов System IO, таким образом позволяя сконфигурировать минимальную конфигурацию СХД из двух контроллеров.

Рекомендуется установка коммутаторов Ангара LAN в случае использования более 2 контроллеров в рамках системы хранения данных Ангара, а также в случае подключения расширения ёмкости в виде шасси JBOD.

Использование коммутаторов System IO целесообразно в случае конфигурации системы хранения данных в режиме без единой точки отказа (NSPOF – no single point of failure).

Доступные опции для коммутаторов System IO:

- трансиверы 10G SFP+ (9/125) Long Wave 1270/1330nm (LC connector)
- трансиверы 10G SFP+ (50/125) Short Ware 850nm (LC connector)

В системе Ангара можно использовать также коммутаторы на базе протокола Fibre Channel, поддерживаются любые модели 8G и 16G коммутаторов от вендоров: Brocade и Cisco.

Ниже в таблице указаны рекомендации по выбору коммутатора для получения отказоустойчивой архитектуры Ангара.

Тип коммутатора	Количество контроллеров P200	Тип модуля ввода-вывода	PN модуля ввода-вывода
System IO over Ethernet	От 2 до 30 JBOD	2 port Ethernet 10/100G	PN:DZ-HBA-04
System IO over FC	От 8 до 15 JBOD	2 port Ethernet 16/32G	PN:DZ-HBA-02/03

Катастрофоустойчивые решения

Катастрофоустойчивая конфигурация реализуется с помощью топологии распределённого хранения данных, построенной в рамках отдельных систем хранения данных в физически изолированных ЦОД.

Системы хранения данных Ангара обеспечивают доступность данных в случае недоступности инфраструктуры одного из двух ЦОД за счёт технологии репликации данных, позволяя обеспечить доступ приложений к данным в любой момент времени.

Конфигурация двух распределённых систем хранения данных Ангара обеспечивает непрерывную доступность данных с помощью механизма репликации данных Volume Copy. В конфигурации участвуют два ЦОД: основной (ЦОД 1) и резервный (ЦОД 2), таким образом подразумевается, что основная копия данных находится в ЦОД 1, а резервная копия данных перемещается с помощью Volume Copy в ЦОД 2.

Ангара поддерживает следующие режимы репликации данных с помощью механизма Volume Copy:

- Синхронный режим
- Асинхронный режим

Синхронный режим репликации Volume Copy обеспечивает целостную и актуальную копию данных, размещённую в удалённом ЦОД, таким образом гарантируя доступность актуальных данных в любой момент времени.

Асинхронный режим репликации Volume Copy позволяет значительно повысить скорость репликации данных и одновременно обеспечить гарантированную доставку данных в заданное временное окно.

Существует два основных параметра – целевое время репликации данных (RTO) и целевое временное окно репликации данных (RPO), которые необходимо рассчитать для выбора режима репликации данных.

Режим репликации	RTO (time objective)	RPO (point objective)
Синхронный	0	0
Асинхронный	Настраивается, зависит от типа и способа реализации сети	Настраивается, зависит от типа и способа реализации сети

Таблица 1 – Сравнительный анализ режимов репликации

В Таблице 1 указаны основные отличия в результате сравнения параметров RTO и RPO для актуального состояния данных резервного ЦОД (см. Рис.22). Таким образом, чтобы получить в любой момент времени актуальное состояние данных в резервном ЦОД, необходимо использовать режим синхронной репликации. Параметры RTO и RPO рассчитываются в зависимости от качества канала передачи данных между основным ЦОД и резервным ЦОД.

Параметр RPO (Recovery point objective) предлагается рассчитывать как интервал времени, который требуется для создания идентичной копии данных в резервном ЦОД с учётом репликации заданного объёма данных (delta copy) и характеристик канала передачи данных между основным и резервным ЦОД. Объём данных для репликации может существенно меняться в зависимости от задач: полная репликация данных, частичная репликация изменений (delta copy).

Частичная репликация изменённых данных (delta copy) возможна благодаря карте изменённых блоков в функционале Volume Copy, таким образом возможно передавать только изменённые данные, существенно снижая нагрузку на канал передачи данных и снижая время RTO и RPO.

Параметр RTO (Recovery time objective) предлагается рассчитывать как интервал времени, который требуется для передачи актуальной копии данных из основного ЦОД в резервный ЦОД.

Конфигурация репликации данных с помощью Volume Copy подразумевает наличие основной (primary) системы хранения данных, размещённой в основном ЦОД и резервной (remote) системы хранения данных, размещённой в резервном ЦОД.

Режим синхронной репликации гарантирует в любой момент времени наличие идентичной копии данных в основной (primary) и удалённой (remote) системах хранения данных.

Режим асинхронной репликации гарантирует доставку данных между основной (primary) и удалённой (remote) системами хранения данных в заданный временной интервал (Recovery Time Objective) и в рамках заданного временного окна репликации данных (Recovery Point Objective).

Топология катастрофоустойчивого решения подразумевает готовность распределённого хранилища к полному отключению одного из ЦОД, поэтому в качестве защиты используется специальный ресурс Quorum, который определяет состояние каждого из двух ЦОД и назначает роли (primary, remote) для каждой из систем хранения данных.

Высокопроизводительные решения

Ангара предлагает высокопроизводительные технологии хранения данных, используя механизм тонкой разметки (thin provisioning), которые объединяются в пулы (pool) с помощью функционала Dynamic Data Tiering (DDT) и позволяют размещать данные на разных типах носителей информации (HDD,SSD). Концепция DDT предполагает разделение всей ёмкости хранилища на высокопроизводительные носители (NVMe), твёрдотельные носители (SSD) и механические носители информации (SAS/SATA).

Ангара позволяет создавать и настраивать правила для перемещения данных между пулами, управлять производительностью приложений и таким образом влиять на следующие показатели:

- Время отклика (latency)
- Пропускная способность (IOPS, throughput)

Функционал DDT позволяет определять правила для динамического использования свободного пространства на основе пулов, исходя из пороговых настроек IOPS (io per second – количество операций IO в секунду) и throughput (Mb/s – количество Мегабайт в секунду).

Серверная виртуализация

В системе Ангара реализована поддержка систем виртуализации VMware, Hyper-V на основе тесной интеграции с протоколами SCSI, Fibre Channel, SAS

В случае катастрофоустойчивых решений, система Ангара позволяет выполнять миграцию виртуальных машин без остановки с помощью функционала Transparent Migration (TM).

Функционал TM реализован с помощью механизма виртуализации WWN портов системы хранения данных Ангара, который позволяет без downtime реализовать миграцию всех операций ввода-вывода в резервный ЦОД, используя также механизм Volume Copy.

Механизм TM также предназначен для миграции данных с систем хранения сторонних вендоров практически без downtime в режиме 24x7.

Миграция данных с помощью External Storage

Система Ангара позволяет предоставлять доступ к ресурсам, которые расположены физически в системах хранения данных сторонних вендоров (механизм External Storage).

Используя механизм Transparent Migration совместно с External Storage, система Ангара выполняет миграцию данных, позволяя одновременно предоставлять доступ к данным и выполнять операции ввода-вывода.

Механизм Transparent Migration позволяет выполнять миграцию данных в режиме Online 24x7, одновременно мигрируя данные с систем хранения данных сторонних пользователей.

Механизм External Storage позволяет выполнять миграцию данных с систем хранения данных сторонних производителей в режиме 24x7, обеспечивая минимальных downtime и выполняя миграцию данных с помощью Volume Copy.

Архитектура системного программного обеспечения Ангара

Программное обеспечение Ангара имеет модульную архитектуру и поставляется вместе с оригинальной аппаратной платформой.

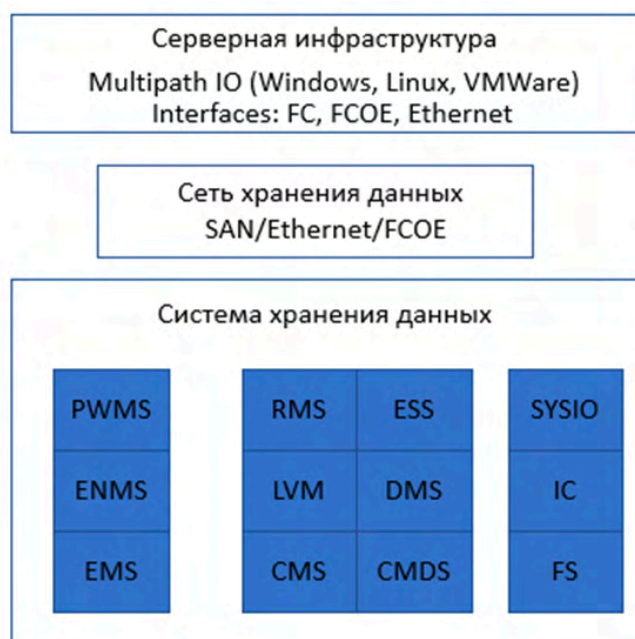


Рисунок 16 – Программная архитектура Ангара

В составе программного обеспечения Ангара имеются следующие модули:

Power Management Service (PWMS)

Модуль PWMS обеспечивает управление системой электропитания всех компонент контроллера Ангара системы хранения данных, а также отвечает за ремонт и замену всех компонент в горячем режиме, т.е. без выключения контроллера Ангара

Enclosure Management Service (ENMS)

Программный модуль ENMS отвечает за мониторинг состояния всех компонент системы хранения данных, включая мониторинг состояния носителей информации и исключает системные сбои контроллера Ангара.

Environment Management Service (EMS)

Программный модуль EMS отвечает за мониторинг состояния среды в ДатаЦентре, включая мониторинг температуры и влажности.

Environment Management Service (EMS)

Программный модуль EMS отвечает за мониторинг состояния среды в ДатаЦентре, включая мониторинг температуры и влажности.

Environment Management Service (EMS)

Программный модуль EMS отвечает за мониторинг состояния среды в ДатаЦентре, включая мониторинг температуры и влажности.

Raid Management Service (RMS)

Программный модуль RMS управляет установкой, заменой носителей информации и выполняет функцию RAID для защиты данных.

Logical Volume Manager (LVM)

Программный модуль LVM управляет ёмкостью системы хранения данных, выделяя требуемый объём хранилища по запросу системного администратора. LVM непосредственно работает с модулем RMS и обеспечивает размещение данных непосредственно через сервис RMS.

Cache Management Service (CMS)

Программный модуль CMS управляет кэш-памятью контроллера Ангара и обеспечивает высокую скорость доступа к данным и минимальное время отклика для внешних приложений.

Extended System Services (ESS)

Программный модуль ESS обеспечивает функции репликации данных, управляет мгновенными копиями томов данных.

Data Management Service (DMS)

Программный модуль DMS обеспечивает интеллектуальное размещение данных, используя логические тома данных.

Command Management Service (CMDs)

Программный модуль CMDs обеспечивает выполнение команд основных транспортных протоколов, таким образом обслуживая и выполняя команды чтения и записи данных.

System IO (SYSIO)

Программный модуль System IO выполняет функции внутренней системной шины передачи данных ETHERNET для обеспечения работы всех основных функций системы хранения данных, включая масштабирование и увеличение ёмкости хранилища.

Interconnect Service (IC)

Программный модуль IC обеспечивает работу модулей ввода-вывода в составе системы хранения данных, обеспечивая интеграцию контроллера Ангара в составе сетей хранения данных (SAN, Ethernet, FCOE).

Filesystem Service (FS)

Программный модуль FS обеспечивает работу файловых сервисов NFS, CIFS системы хранения данных, а также протокол резервного копирования данных NDMP.

Архитектура Volume Copy

Функционал Volume Copy предназначен для репликации данных (передачи данных) между двумя и более системами хранения данных серии Ангара на базе контроллера P200.

Существует два режима репликации данных – режим полного копирования и режим передачи изменённых данных.

В рамках системы Ангара единицей репликации является Volume Pair (пара), которая объединяет два тома: основной том (мастер) и его копию на локальной или внешней системе хранения данных.

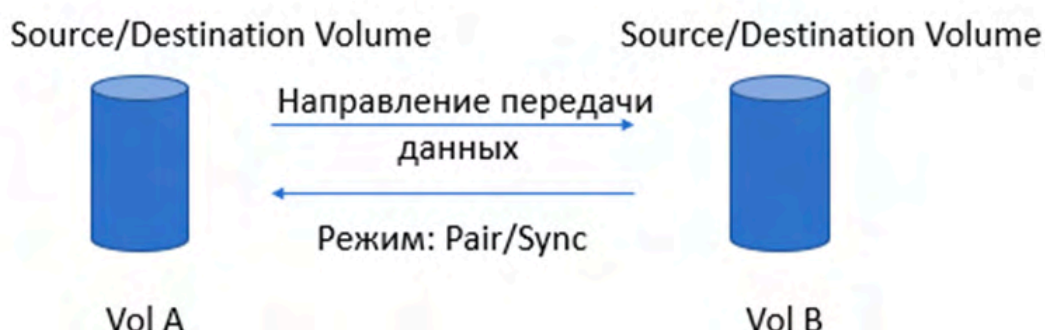


Рисунок 17. Схема репликации Volume Copy

В режиме репликации для заданного тома Vol A (см. Рис.17) создаётся пара на базе тома Vol B, таким образом определяются два тома, между которыми будет производиться репликация данных. Каждому тому назначается роль – master или copy. В момент создания пары система Ангара создаёт карту изменённых блоков главного (master) тома и позволяет временно остановить репликацию либо возобновить в любой момент времени. Направление репликации определяется заданными ролями – от главного (master) тома к его копии (copy).

Volume Copy позволяет организовать каскадный режим репликации (не более трёх томов) (см. Рис. 18).

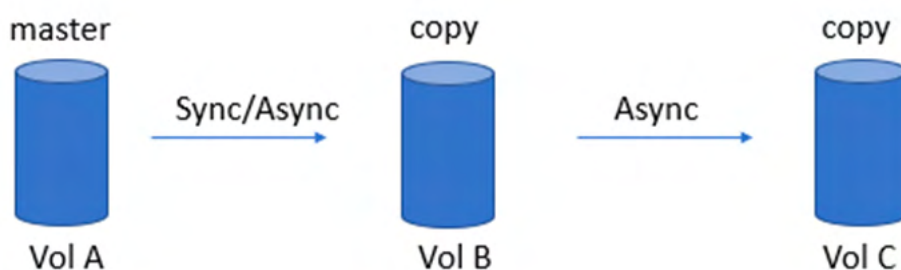


Рисунок 18. Каскадный режим репликации Volume Copy

Поддерживаются следующие режимы репликации данных: асинхронный (Async) и синхронный (Sync). В режиме синхронной репликации гарантируется одновременное изменение данных для основного (master) тома и его копии (copy том). В режиме асинхронной репликации гарантируется перенос данных с задержкой во времени в соответствии с изменениями записанными в карту изменённых блоков (bitmap) основного тома (master).

Архитектура RAID Ангара

Архитектура RAID системы Ангара реализована на базе отказоустойчивых алгоритмов, поддерживающих следующие сервисные операции:

- Замена носителей информации (HDD) без остановки функционирования операций RAID в режиме hot swap
- Восстановление информации в случае обнаружения дефектов

Поддерживаются следующие типы RAID:

- RAID50 (3D+1P)
- RAID60 (2D+2P)
- RAID10 (2D+2D)

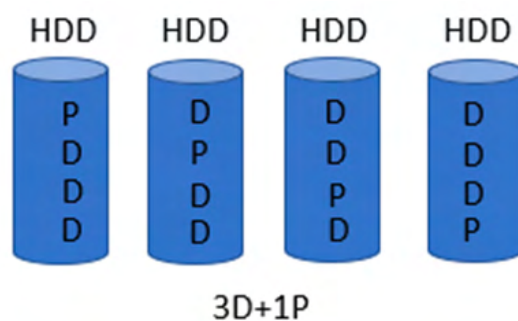


Рисунок 19. Конфигурация группы RAID50 3D+1P

Распределение данных (D) и данных контроля чётности (P) в конфигурации 3D+1P выполняется таким образом (см. Рис.19), чтобы потеря любого носителя (HDD) не влияла на доступность данных в RAID.

Ангара позволяет увеличивать ёмкость хранилища, используя операцию объединения групп дисков, получая следующие конфигурации RAID:

- RAID50 (6D+2P, 7D+1P, 12D+4P, 14D+2P, 28D+4P)
- RAID60 (2D+2P, 4D+4P, 6D+2P, 8D+4P, 12D+4P)
- Используя механизм Thin Provisioning, возможно создавать множество пулов, используя различную конфигурацию RAID групп. В рамках одного пула возможно объединять до 64 RAID групп

Архитектура Thin Provisioning

Архитектура Thin Provisioning разработана с учётом возможности over provisioning (выделение пространства сверх доступного объёма), таким образом система Ангара позволяет создавать тома для размещения данных без ограничения доступной физической ёмкости хранилища.

Обновление программного обеспечения Ангара

Обновление программного обеспечения выполняется с помощью механизма «NonStop Upgrade»® в системе Ангара, который обеспечивает функционирование всех сервисов и компонентов системы во время обновления программного обеспечения или во время выполнения операций сервисного обслуживания системы Ангара.

Базовый блок Ангара

Базовый блок системы Ангара является двухконтроллерной системой на базе двух шасси P200. По умолчанию в конфигурацию базового блока входит опция high availability для защиты кэш-памяти контроллера в режиме active-active.

Базовый блок P200 в режиме active-active позволяет обеспечить конфигурацию P200 без единой точки отказа.



Рисунок 29. Конфигурация active-active Ангара®

Режим active-active (см. Рис.29) обеспечивается с помощью функционала шины передачи данных ETHERNET и протокола System IO, который реализован в базовом блоке на основе адаптера #SQ-HBA-04.

5.1. Требования к аппаратной платформе для установки ПО СХД

Программно-определяемая система Ангара® StorageOS доступна для установки в виде программных модулей для установки на совместимом оборудовании.

Требования к установке программных модулей Ангара® доступны в руководстве пользователя и также доступны для скачивания в разделе <https://www.datazeh.ru/downloads>.

Ангара® StorageOS поддерживает различные архитектуры на базе процессоров указанных в Таблице 2.

Архитектура	Модель процессора
Intel x86	Xeon (Broadwell, Haswell)
ARM v64	Qualcomm ARMv64, Baikal S

Таблица 2 – Совместимые архитектуры процессоров

5.2. Требования к операционной системе для установки ПО СХД

Перечень поддерживаемых ОС Linux указан в Таблице 3.

Наименование ОС	Уровень ядра
Ubuntu Server 22.04	Официальный в поставке

Таблица 3 – Список поддерживаемых операционных систем

5.3. Функциональные возможности ПО СХД

5.3.1. Менеджер управления физическими носителями

Управление физическими носителями производится с помощью утилиты Ангара® raidtool. В Таблице 4 перечислены основные функции утилиты raidtool.

Текущие обновления raidtool и описание актуальных изменений можно найти в дистрибутиве ПО СХД в документе raidtool_release_notes<версия>.pdf, где <версия> определяет дату и номер релиза ПО СХД.

Функция raidtool	Описание функции
Сканирование шины и определение адреса и параметров носителя информации	Определение доступности носителя информации и параметров доступа (скорость, отказоустойчивость, способ доступа, адрес носителя и его уникальный номер wwn)
Считывание состояния носителя информации	Считывание параметров и характеристик носителя информации, учитывая его износ и прогнозирование возможных сбоев в будущем
Создание raid группы	Конфигурирование группы носителей информации в составе raid группы с заданными параметрами (RAID10,50,60)
Физическое добавление или удаление носителя информации	Безопасная остановка или запуск носителя информации в режиме горячей замены без остановки аппаратной платформы

Таблица 4 – Функции утилиты raidtool

Менеджер управления логическими томами RAIDLVM обеспечивает функционал указанный в Таблице 5.

RAID LVM обеспечивает функционирование RAID групп и обеспечивает доступ к данным непосредственно расположенным на носителях информации.

Функция RAID LVM	Описание функции
Выполнение операций ввода-вывода в рамках заданных RAID групп	Предоставление API и обеспечение доступа к носителям информации с целью перемещения данных между носителями информации и внешним приложением
Хранение конфигурации RAID групп	Обеспечение целостности и консистентности конфигурации в рамках всего комплекса СХД
Изменение конфигурации RAID групп	Предоставление API для изменения конфигурации RAID групп
Создание и размещение логических томов	Выделение единого адресного пространства с заданным именем в рамках одной или нескольких RAID групп
Хранение конфигурации логических томов	Обеспечение целостности и консистентности конфигурации в рамках всего комплекса СХД

Таблица 4 – Функции утилиты raidtool

5.3. Функциональные возможности ПО СХД

5.3.3. Менеджер управления доступом к логическим томам в ОС StorageOS®

Менеджер управления доступом к логическим томам IOPATH обеспечивает функционал указанный в Таблице 6.

IOPATH обеспечивает доступ к логическим томам с помощью доступных физических адаптеров установленных в СХД (см. 1.2.3) и типов используемых протоколов.

IOPATH поддерживает протоколы доступа Ethernet, SYSTEM IO, iSCSI.

Функция IOPATH	Описание функции
Выполнение операций ввода-вывода в рамках протокола SCSI и стандартов ANSI SPC-3/SPC-2	Выполнение операций ввода-вывода в соответствии с протоколом SCSI
Управление доступом для чтения и записи данных в рамках логических томов	Модуль аутентифицирует и предоставляет доступ клиентов для чтения и записи данных в рамках логических томов

Таблица 6 – Функции IOPATH

Спецификация СХД Ангара®	Описание
Блочное хранилище	
Максимальное количество подключенных серверов	• Ethernet - 1024 • iSCSI ports: 1024
Максимальное количество серверов в группе	8
RAID уровни	10,50,60
Максимальное количество LUN	4096
Максимальное количество LUN доступных одному хосту	255
Максимальное количество дисков в одной RAID группе	256
Минимальное количество дисков в одной RAID группе	4
Максимальное количество storage pools	16
Минимальный размер LUN	512Kb
Максимальный размер LUN	16TB

Спецификация СХД Ангара®	Описание
Эмуляция файлового хранилища	
Максимальное количество файловых систем	1024
Минимальный размер файловой системы	1GB
Максимальный размер файловой системы	8PB
Максимальное количество файлов на одной файловой системе	2 миллиарда
Максимальный размер файла	128TB
Максимальное количество SMB объектов общего доступа	128
Максимальное количество NFS объектов общего доступа	128
Максимальное количество HTTP соединений на один контроллер	64
Максимальное количество NDMP потоков на один контроллер	8

Максимальное количество локальных пользователей	1000
Максимальное количество локальных групп пользователей	1000
Максимальное количество пользователей в группе	1000
Максимальная длина пути к файлу	1024 bytes
Максимальная длина имени файла или директории	256 bytes
Максимальное количество файлов открытых одновременно на одном контроллере	2048

Спецификация Ангара® Volume Copy	
Максимальное количество пар томов для репликации (synchronous + asynchronous)	256
Максимальное количество вторичных LUN в паре	• Synchronous: 1:2 • Asynchronous: 1:10
Максимальное количество внешних подключаемых хранилищ данных	128
Максимальное количество пар томов на один контроллер	128

Спецификация Ангара® Thin Provisioning

Максимальное количество «тонких» LUNs (аналогично Максимальное количество LUN)	4096
Максимальный размер «тонкого» LUN	256TB
Гранулярность «тонкого» тома	128KB

Спецификация Ангара® Tiering

Максимальное количество поддерживаемых уровней хранения	3 (NVME, SAS SSD, SAS NL)
Типы политик для перемещения данных по уровням	• Based on IOPS and latency settings • Based on Quota settings in MBs
Поддерживаемый тип LUN	Работает только с Thin Provisioning

Спецификация Ангара® File Services

NFS	NFS v4/v3
CIFS	2.0/3.0

Спецификация Ангара® SCSI Services

SCSI

SPC-3/SPC-4

Спецификация Ангара® transport protocols

SAS

12 Gbit

Ethernet

10/100 Gbit

Операционная система	Версия ядра
Linux	3.10 и выше
Windows Server	2019 и выше

Авторские права

Авторские права на торговые марки, материал в рамках данного документа, а также информация о составе и структуре программного обеспечения Ангара® принадлежит компании ООО СП «М-Энергоналадка».

Пользователь может просматривать информацию в рамках данного документа исключительно с целью некоммерческого использования.

Разрешение использовать данный документ предоставляется со следующими ограничениями:

1. Необходимо сохранить указание на авторские права и другие защищенные права.
2. Никакие изменения в материалах не допускаются.
3. Никакие материалы не могут быть скопированы, воспроизведены, озвучены, переданы или пересланы каким-либо образом без письменного разрешения обладателя прав торговой марки ООО СП «М-Энергоналадка».