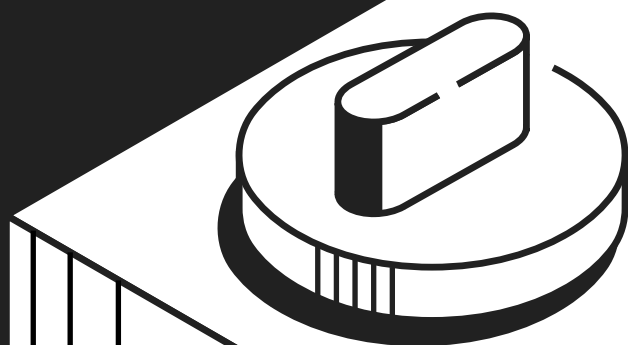




CLOUD.RU SOFTWARE-DEFINED STORAGE

Программно-определяемое хранилище
для частных и гибридных инсталляций

Версия: Февраль 2026

О продукте	3
Сценарии использования и развертывания	4
Основные характеристики	5
Архитектура	6
Безопасность, аутентификация и авторизация	7
Требования к оборудованию	8
Расчет полезной емкости	9
Лицензирование	11

О продукте

Cloud.ru Software-Defined Storage (SDS) – программное решение для построения отказоустойчивого и производительного блочного хранилища.

Собственная разработка Cloud.ru.

Основная идея SDS — легко и быстро трансформировать устройства хранения данных (HDD, SSD) и сетевое оборудование в защищенное кластерное SAN/NAS-хранилище корпоративного уровня. Такое хранилище оптимизировано для хранения больших объемов данных и обеспечивает их репликацию и высокую доступность. Оно может использоваться для безопасного размещения виртуальных машин и контейнеров, а также их миграции между физическими хостами без перерывов в предоставлении этих ресурсов.

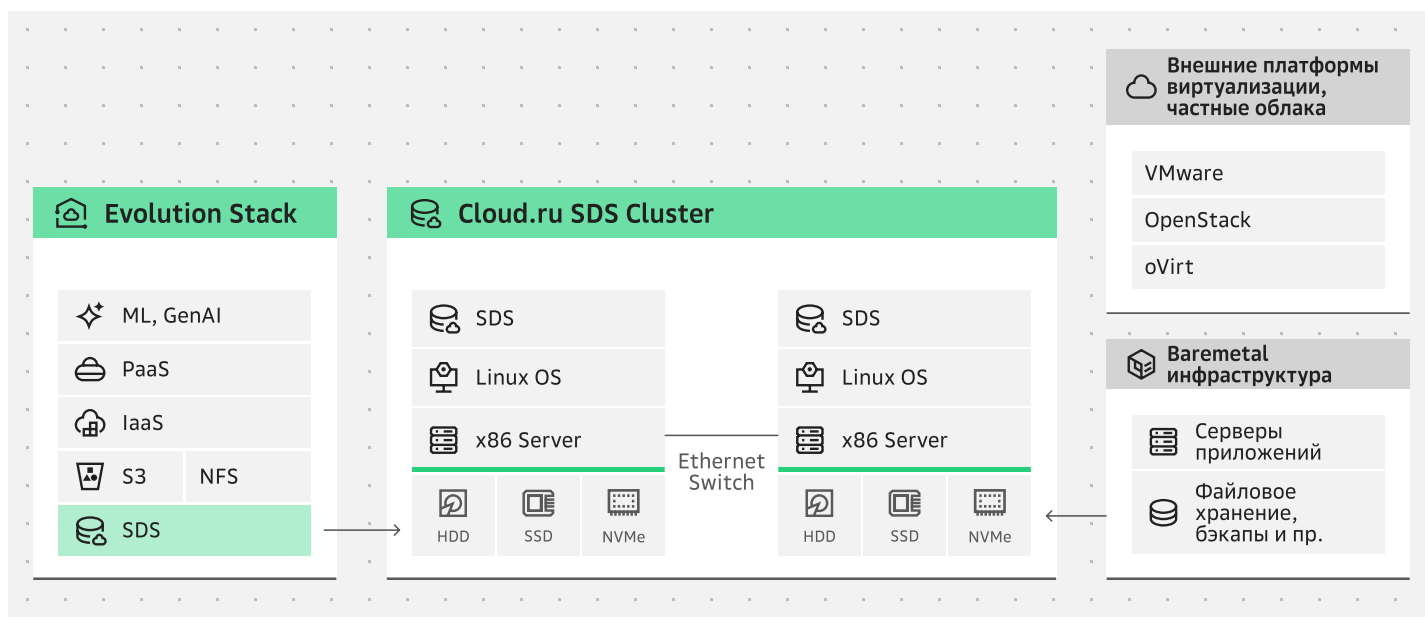
Ряд технологических ноу-хау позволяет экономить общий объем аппаратных ресурсов, не предъявляя повышенные требования к техническим характеристикам узлов SDS.

Ключевые преимущества

- **Горизонтальное масштабирование:** добавление дополнительных ресурсов (новые узлы и диски) без прерывания работы системы.
- **Снижение совокупной стоимости владения** в сравнении с традиционными решениями за счет использования схемы кодирования с исправлением ошибок (Erasure Coding — EC).
- **Высокая доступность данных** при сбоях отдельных компонентов (CS, MDS).
- **Обеспечение целостности данных:** реализация механизмов репликации данных и коррекции ошибок.
- **Высокая производительность операций чтения и записи**, минимальные задержки и высокая пропускная способность за счет оптимизации работы с метаданными и эффективного распределения нагрузки между узлами кластера.
- **Экономия вычислительных ресурсов гипервизора** и кратное уменьшение числа подключений к кластеру SDS за счет использования технологии доступа vhost-user.

SDS Cloud.ru входит в реестр российского программного обеспечения ([№19189 от 23.09.2023](#)).

Сценарии использования и развертывания



Блочное хранилище облачной платформы Cloud.ru Evolution Stack

SDS является опорным решением для блочного хранения служебных и клиентских данных в составе облачной платформы Cloud.ru Evolution Stack, а также обеспечивает технологический слой для сервиса объектного хранения S3 в составе платформы.

1

Блочное хранилище для платформ виртуализации

Использование SDS как самостоятельного продукта для блочного хранения в составе сторонних on-premise решений корпоративной виртуализации на базе VMware, OpenStack, oVirt, других вендорских продуктов и в качестве основного клиентского хранилища для провайдеров облачных услуг, заменяя зарубежные СХД.

2

Программно-аппаратные комплексы блочного и S3-хранения

SDS выступает программной основой для программно-аппаратных комплексов объектного хранения Cloud.ru Evolution Stack, а также может использоваться для создания клиентских ПАК блочного и объектного хранения.

3

Гиперконвергентные конфигурации

Для небольших инсталляций, пилотных сред и сред разработки возможна унификация вычислительных узлов (стандартных серверов) для размещения на одном узле как SDS-части, так и вычислительной части платформы Cloud.ru Evolution Stack или других платформ, поддерживающих подобную схему.

4

Основные характеристики

Общие параметры

Кодовая база: собственная разработка (C++, Go).

Поддержка гибридных (SSD+HDD) и all-flash (SAS SSD, SAS SSD + NVMe, all-NVMe) конфигураций SDS.

Масштабирование:

- рекомендация для одного кластера: 7-16 узлов,
- поддерживаемый объем: 24 узла.

Отказоустойчивость

Механизмы обеспечения избыточности данных:

- Erasure coding (EC): настраиваемые схемы. Схема по умолчанию: 4+2.
Максимальная схема: 21+5 (Max Stripe Depth=21, Max Redundancy=5).
Схема без redundancy: 1+0.
- Реплики: настраиваемое количество: 1+0, 1+1, 1+2 и далее.

End to End Data Protection (CRC): все данные хранятся с контрольными суммами (CRC32) и проверяются на каждом чтении. Защита от silent corruption.

Периодическая фоновая проверка целостности данных: скраббинг (данные и CRC) и глубокий скраббинг (с использованием EC).

Настраиваемые домены отказа: диск, сервер (по умолчанию), стойка, POD (группа стоек).

Поддержка протоколов

- vhostd (приложение) для работы с OpenStack (vhost-user-blk),
- NVMe-over-TCP (H1 2026),
- NVMe-over-RDMA (RoCE v2) (H1 2026).

Производительность

Ориентировочный предел в одном кластере: 2+ млн IOPS.

Базовый ориентир для одного узла: 150-200k IOPS 8k 65/35 r/w (60–100 ТБ сырого пространства).

Функциональность

Tiering: данные всегда пишутся в быстрый tier (Hot Storage). По мере «остывания» данные кодируются и «сбрасываются» на более медленный tier (Cold Storage).

QoS: настройка ограничений пропускной способности или числа IO-операций в секунду на томе (volumes).

Балансировка: данные ребалансируются для равномерной утилизации дисков.

Поддержка мгновенных снимков состояния тома, full и incremental.

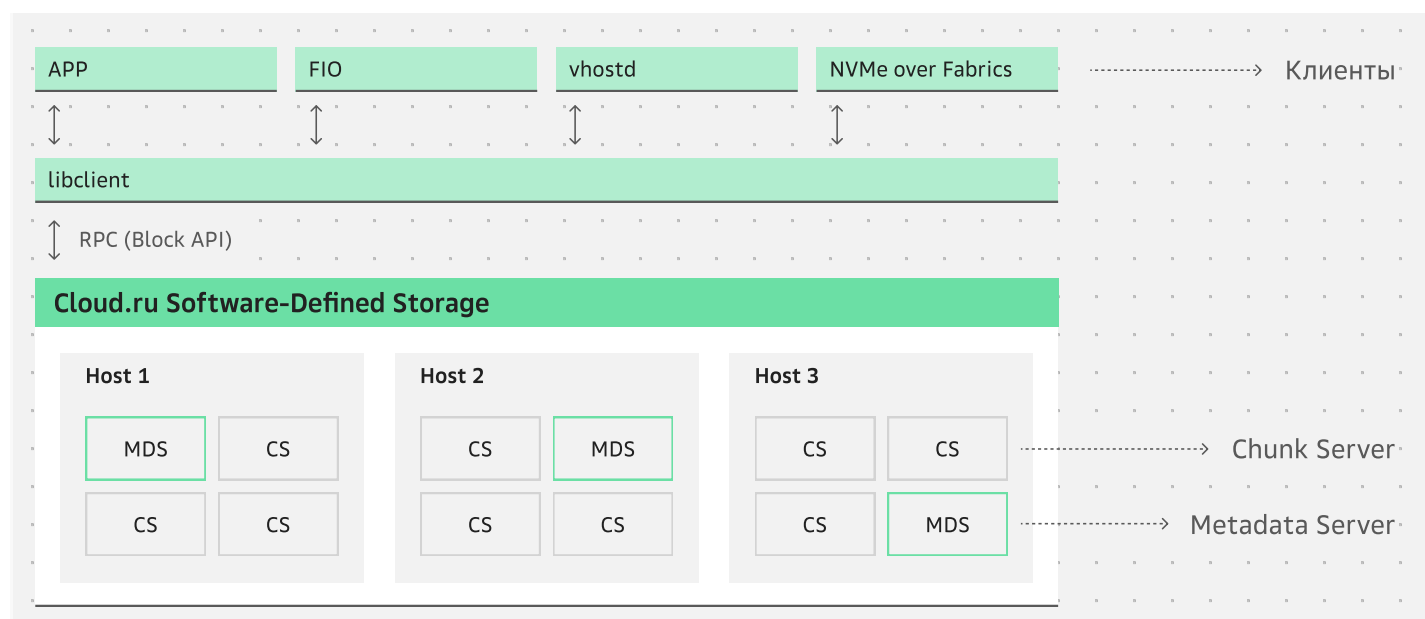
Поддержка клонирования томов.

Мониторинг и алертинг: широкий набор преднастроенных дашбордов мониторинга.

Шифрование на стороне клиента.

Ansible плейбуки для быстрого развертывания и масштабирования решения.

Архитектура SDS



Сервер метаданных (MDS)

Отвечает за поддержание всех метаданных блочных устройств в постоянном хранилище ключ-значение с репликацией RAFT. Это включает в себя имена блочных устройств, отображение блочного устройства на чанки, состояния чанков и реестр включенных в кластер серверов чанков (CS).

Контролирует действия на уровне всей системы, такие как управление арендой, восстановление чанков в случае сбоя CS, управление списком активных CS и миграцию чанков между CS.

Сервер чанков (CS)

Volumes делятся на фиксированные чанки. Каждый чанк идентифицируется неизменным и глобально уникальным 64-битным идентификатором чанка, присвоенным ведущим MDS в момент создания чанка.

CS хранит части чанков на локальных дисках в виде файлов Linux и осуществляет чтение или запись данных чанка, указанных идентификатором чанка и диапазоном байтов.

Для обеспечения надежности каждый чанк реплицируется на нескольких CS или кодируется с помощью кодов Рида-Соломона. CS взаимодействует с другими CS для репликации записанных данных или обновления блоков четности.

Клиент

Общается с ведущим MDS и CS для чтения или записи данных от имени приложения. Клиент получает запросы ввода-вывода от приложений и передает эти запросы в кластер хранения.

Клиент взаимодействует с ведущим MDS для операций с метаданными. Вся коммуникация, связанная с данными, идет напрямую к CS для обеспечения строгой консистентности данных даже при сбоях CS или MDS.

3

Инструмент управления

Используется для управления кластером и общается с ведущим MDS через gRPC API. Может добавлять новые MDS в кластер или исключать существующие, отвечает за создание томов, пользователей и CS.

4

Безопасность, аутентификация и авторизация

Шифрование

Поддерживается шифрование данных volume на клиенте (vhostd и пр.). Дальнейшая передача данных по сети происходит в зашифрованном виде. При чтении расшифровка данных происходит также на клиенте. Ключ шифрования клиент получает от MDS.

Ключи шифрования хранятся в MDS и реплицируются с помощью алгоритма RAFT. Ключи хранятся в зашифрованном виде с помощью мастер-ключа.

Помимо шифрования с помощью мастер-ключа поддерживается интеграция с KMS и для шифрования томов персональным ключом пользователя при использовании SDS в составе облачной платформы Cloud.ru Evolution Stack.

Ролевая модель

Для авторизации пользователей в кластере и разграничения прав доступа к различным методам MDS API предусмотрена базовая ролевая модель с возможностью изменения ролей для существующих пользователей.

Требования к оборудованию

Архитектура SDS позволяет использовать единые конфигурации серверов для построения и бесшовного масштабирования кластера.

Требования к универсальному узлу хранения:

	Минимальные характеристики	Рекомендованные характеристики
CPU	1 CPU, 24+ cores 2.0GHz	2 CPU, 24+ cores 2.6GHz
RAM	256 GB	256 GB
Диск (данные) *	2 диска по 1.92 TB NVMe	8–12 дисков по 7.68 TB NVMe
Диск (метаданные)	1 диск 512 GB NVMe/SSD	1 диск 1024 GB NVMe/SSD
Диск (ОС)	2 диска по 240 GB SSD RI	2 диска по 240 GB SSD RI
Сетевые интерфейсы	2 сетевых карты по 2x25 GB	2 сетевых карты по 2x25 GB 1 сетевая карта 2x10 GB



* Выбор объема и типа диска всегда обусловлен задачами проекта и целевыми сценариями. Поддержка гибридных конфигураций HDD+SSD позволяет подобрать под каждый проект оптимальный объем и тип дисков для получения нужного соотношения между полезным объемом и производительность SDS.

Минимальное количество узлов в кластере определяется целевым механизмом обеспечения отказоустойчивости. Например:

- Erasure Coding 4+2: минимум 6 узлов (7 и более – рекомендованная конфигурация);
- Erasure Coding 6+3: минимум 9 узлов (10 и более – рекомендованная конфигурация);
- Реплика 1+2: 3 узла и более.

Для обеспечения наиболее эффективного использования дисковой емкости в рамках одного кластера рекомендуется использовать серверы идентичной конфигурации.

Расчет полезной емкости

Фактор репликации для разных схем обеспечения избыточности данных

Схема кодирования	Минимальное количество узлов	Рекомендованное количество узлов	Количество недоступных узлов без потери данных	Фактор репликации
1+0 (1 реплика)	1	1	-	1
1+2 (3 реплики)	3	4	2	3
3+2 (erasure coding)	5	6	2	1.66
4+2 (erasure coding)	6	7	2	1.5
6+3 (erasure coding)	9	10	3	1.5
9+3 (erasure coding)	12	13	3	1.33

Фактор репликации определяет соотношение между сырым и полезным объемом данных для выбранного механизма обеспечения избыточности. Чем меньше значение фактора репликации, тем более эффективно используется общее сырое дисковое пространство. Помимо фактора репликации ключевым параметром для выбора того или иного механизма обеспечения избыточности являются требования к отказоустойчивости кластера и максимальному количеству узлов, выход которых из строя не должен приводить к потере данных.

Полезная емкость для кластера AllFlash

Параметр	Значение	Комментарий
Количество серверов	10	2 CPU, 24+ cores 2.6GHz
Диски для хранения данных (CS)	10 дисков по 7.68 TB	76.8 ТБ сырого пространства на 1 сервер
Схема кодирования	4+2 (erasure coding)	Фактор репликации: 1.5
Размер Hot Storage	5%	
Резерв	10%	По умолчанию резервируется 5% оставшегося места, т.е. места на Cold Storage. В текущем расчёте - 10%. Параметр изменяемый. При заполнении диска на 95% кластер перейдет в режим readonly. Не рекомендуется утилизация более 90%
Всего дисков для хранения данных	100 штук	10 серверов x 10 дисков
Суммарный сырой объём (лицензируемый)	768 TB	10 серверов x 10 дисков x 7.68 TB
Объём пространства для хранения без учета резерва	486 TB (442 TiB)	768 TB / 1.5 фактор репликации * (100%-5% Hot Storage)
Объём пространства для хранения с учётом резерва	438 TB (398 TiB)	768 TB / 1.5 фактор репликации * (100%-5% Hot Storage) * (100%-10% резерв)

Пример конфигурации кластера

Лицензирование

Лицензионная метрика: 1 ТВ сырой дисковой емкости. Учитывается только диски, используемые для хранения данных. Диски для хранения метаданных и диски операционной системы не требуют лицензирования.

Например, кластер из 10 узлов в конфигурации с 10 дисками 7.68 ТВ для хранения данных потребует лицензирования 768 ТВ сырой дисковой емкости (10 серверов + 10 дисков) x 7.68 ТВ.



Есть вопросы? Пишите на hybrid-sales@cloud.ru

Любое копирование и воспроизведение содержания
(в том числе частичное) без разрешения
правообладателя запрещено

2025 © Cloud.ru