

Система хранения данных "Resilient Cloud Storage"

ОГЛАВЛЕНИЕ

1. Описание системы	3
1.1 Введение	3
1.2 Архитектура «Resilient Cloud Storage»	4
1.2.1 Модули СХД «Resilient Cloud Storage»	4
1.2.3 Логическая структура	8
1.3 Построение СХД "Resilient Cloud Storage"	9
2. Использование.....	11
2.1 Подключение СХД "Resilient Cloud Storage" к клиентам.....	11
2.2 Виртуализация существующих СХД	11
2.3 Примеры использования	11
3. Преимущества	12
3.1 Эффективное программное обеспечение.....	12
3.2 Горизонтальное масштабирование	12
3.3 Разделение групп потребителей.....	12
3.4 Кэширование данных.....	12
3.5 Быстрое клонирование и тонкие диски	12

Глава 1

Описание системы

1.1 Введение

"Resilient Cloud Storage" — это эластичная и распределённая система хранения данных (СХД).

Благодаря её горизонтально масштабируемой архитектуре, небольшую систему из нескольких модулей можно без лишних усилий превратить в систему с несколькими тысячами модулей хранения общим объёмом в несколько сотен петабайт.

СХД "Resilient Cloud Storage" не имеет в своём составе традиционных централизованных контроллеров, которые передают через себя данные — всё взаимодействие клиенты СХД осуществляют напрямую с модулями хранения.

Модульная архитектура и равномерное распределение данных по модулям позволяет создать систему с массовым параллелизмом, в которой операции ввода-вывода выполняются в распределённой среде. В результате увеличение количества модулей хранения приводит как к увеличению ёмкости, так и к линейному повышению производительности системы в целом.

СХД "Resilient Cloud Storage" содержит в своём составе SSD-носители для приложений, требующих высокой производительности, и классические диски для архивного хранения больших объёмов информации.

СХД "Resilient Cloud Storage" обладает функциональностью репликации данных, тонких томов, их клонирования и создания моментальных снимков (snapshots), что позволяет эффективно управлять пространством и обеспечивать защиту от сбоев и катастроф.

1.2 Архитектура «Resilient Cloud Storage»

СХД "Resilient Cloud Storage" является программно-аппаратным комплексом, имеющим в своём составе модули хранения, укомплектованные быстрыми и медленными носителями, объединённые в единую сеть под управлением централизованной системы управления и мониторинга.

1.2.1 Модули СХД «Resilient Cloud Storage»

RCSC - Координатор системы хранения



Управляет модулями хранения данных и координирует их работу в единой архитектуре управления данными. Он также содержит информацию о топологии и состоянии всей системы, и о распределении данных внутри хранилища.

RCSDM - Модуль хранения



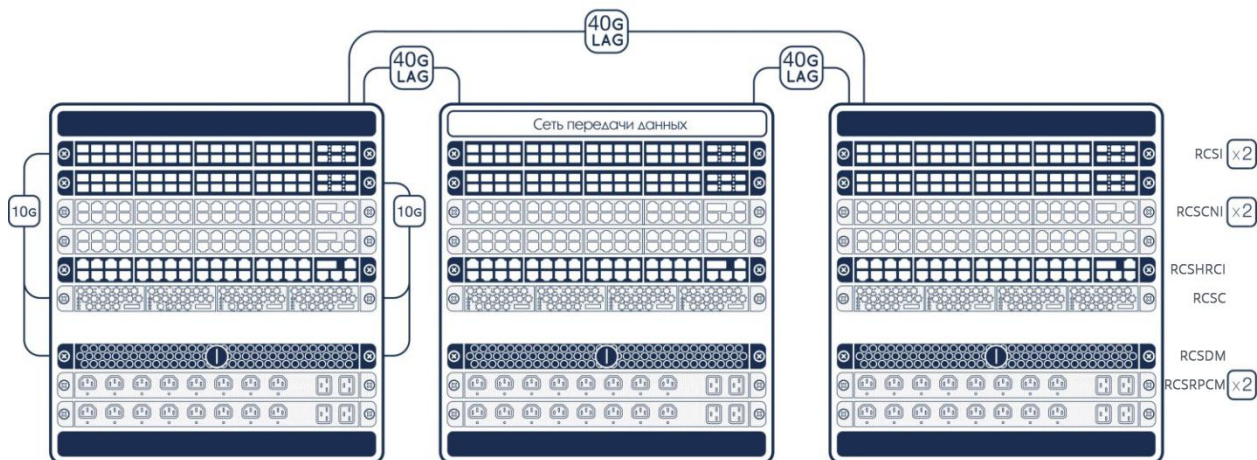
Обеспечивает эффективное хранение и управление данными с высокой производительностью, а также предоставление доступа клиентам СХД к информации. Кроме того, предоставляет такие возможности, как мгновенное восстановление данных, клонирование, репликация данных, создание мгновенных резервных копий.

Содержит в своём составе SSD и HDD носители. В наличии имеются различные конфигурации модулей.

RCSI - Интерконнект системы хранения



Шина для объединения модулей в единую систему хранения данных. Обладает высокой пропускной способностью и низкими задержками для взаимодействия модулей системы хранения между собой и доступа клиентов к информации, хранимой в СХД.

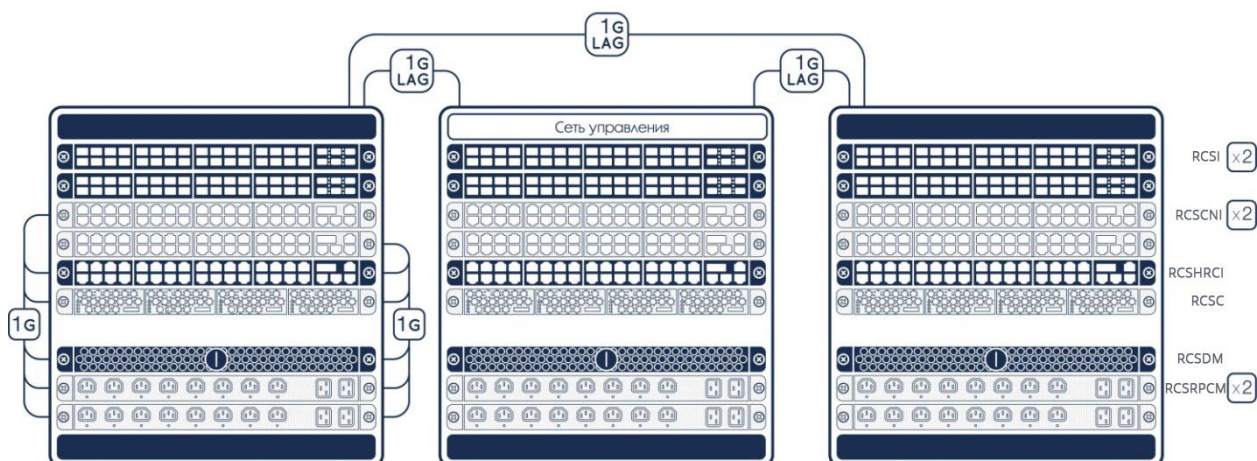


Пример организации сети передачи данных в случае расположения оборудования в трех стойках

RCSCNI - Интерконнект сети управления



Шина управления модулями хранения системы. Служит для мониторинга и управления компонентами СХД с помощью как встроенных средств, так и внешних по протоколам SNMP/SSH.

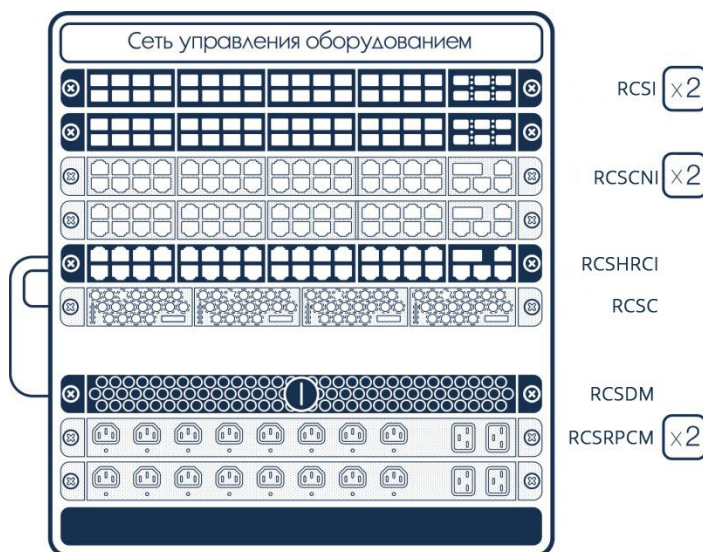


Пример организации сети управления в случае расположения оборудования в трех стойках

RCSHRCI - Интерконнект управления оборудованием



Шина удалённого управления оборудованием. Обеспечивает возможность удалённого управления электропитанием и консольным доступом, независимо от основной сети управления координаторами и модулями хранения.

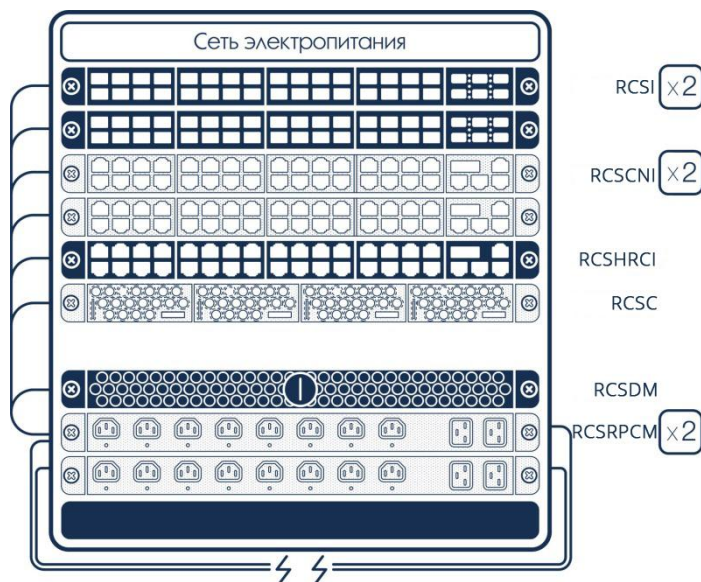


Пример организации сети управления оборудованием внутри одной стойки

RCSRPCM - Удалённое управление питанием



Позволяет производить удаленное включение, отключение и перезагрузку оборудования СХД. Обладает функцией автоматического переключения между источниками питания (ABP), максимальный ток - 16 А в непрерывном режиме. Переключаемые розетки и интерфейс локальной сети (Ethernet) обеспечивают возможность управления отдельными потребителями электропитания в режиме реального времени или в соответствии с задаваемой программой. Дает возможность администратору устанавливать индивидуальные пороги срабатывания аварийной сигнализации в соответствии с заданными параметрами электропитания.



Пример организации электропитания внутри стойки

RCSIM – Интерфейсный модуль



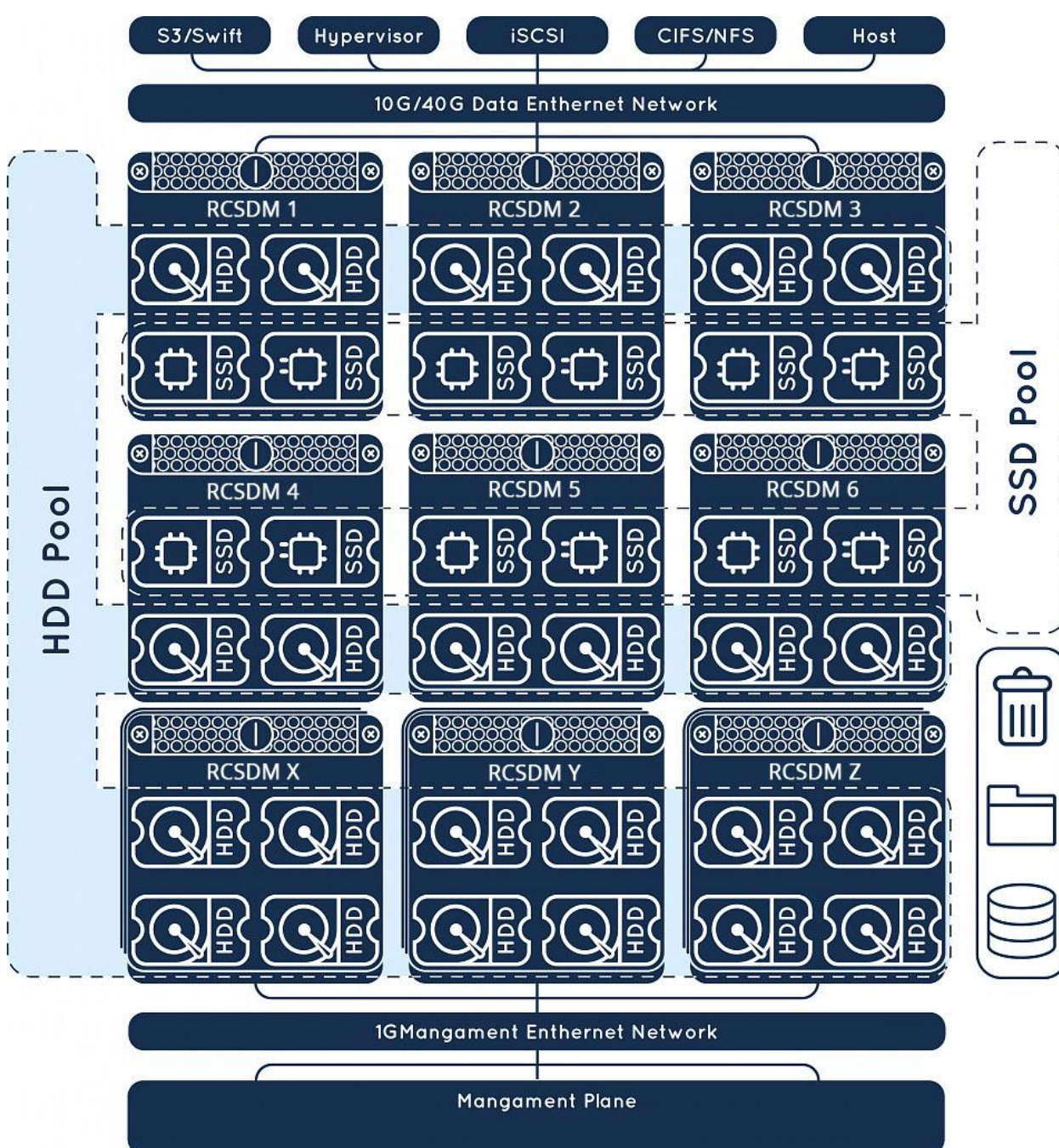
Обеспечивает подключение к системе хранения по протоколам iSCSI, Fibre Channel, InfiniBand.

RCSVCA - Модуль подключения к виртуальной инфраструктуре

Обеспечивает подключение системы хранения к платформе виртуализации. Является программным обеспечением (виртуальной машиной), устанавливаемое внутрь среды виртуализации.

1.2.3 Логическая структура

1. **Инсталляция** - программно-аппаратный комплекс "Resilient Cloud Storage"
2. **Пул** - объединение носителей в группу в соответствии с их характеристиками либо нуждами потребителей
3. **Носитель** - устройство для хранения данных
4. **Образ** - виртуальный диск (LUN)
5. **Клиент** - платформа виртуализации, виртуальный/физический хост и т.д.
6. **Snapshot** - мгновенный снимок данных виртуального диска



Пример логической структуры СХД Resilient Cloud Storage

1.3 Построение СХД 'Resilient Cloud Storage'

СХД 'Resilient Cloud Storage' использует следующий механизм хранения информации: данные разбиваются на объекты фиксированного размера, которые равномерно распределяются по всем носителям в пуле пропорционально их ёмкости. Каждая часть информации (объект) хранится минимум в трёх копиях. Возможна также индивидуальная настройка количества копий на уровне каждого пула.

Правила выбора места размещения копий определяют область отказоустойчивости, при выходе из строя которой СХД сохраняет полную функциональность. В качестве такой области могут выступать носитель/сервер/стойка/ряд стоек/ЦОД. Также есть возможность произвольного определения области отказоустойчивости в соответствии с требованиями клиента. По умолчанию областью отказоустойчивости является стойка, и дальнейшее описание принципов построения СХД будет исходить из того, что область отказоустойчивости определена как одна стойка.

В случае использования трех копий базовой конфигурацией СХД является следующая:

1. Три координатора
2. Три модуля хранения
3. Два интерконнекта сети хранения
4. Два интерконнекта сети управления
5. Один интерконнект управления оборудованием
6. Три модуля управления питанием

Координаторы не участвуют в обмене данными между модулями хранения и клиентами (они взаимодействуют между собой напрямую), поэтому трёх координаторов достаточно для работы системы больших масштабов (сотни модулей хранения). Для очень больших инсталляций их количество может быть увеличено до 5-7 шт. для обеспечения большей отказоустойчивости.

Модули хранения хранят и обрабатывают информацию. В случае использования трех копий для хранения данных модули хранения наращиваются комплектами по три штуки и устанавливаются по одному в каждую из областей отказоустойчивости (как правило – стойка).

Интерконнекты объединяют модули хранения в высокопроизводительную сеть хранения. В этой сети идёт обмен с потребителями (клиентами) и репликация данных между модулями хранения. Каждый интерконнект имеет как минимум одно соединение в сторону клиентов. На начальном этапе (до 30 модулей хранения) достаточно двух интерконнектов для соединения всех модулей хранения. Дальнейшее наращивание (до 60 модулей хранения) требует установки еще двух интерконнектов в соседней стойке. Наращивание СХД до 90 модулей хранения требует установки еще двух интерконнектов в третью стойку. Следующее наращивание объёма идёт по такому же сценарию в следующих трёх стойках.

Для управления модулями хранения используются выделенные сети управления и отдельные интерконнекты сети управления. Для управления оборудованием используются соответствующие интерконнекты. Их наращивание происходит аналогичным с интерконнектами сети управления образом.

При добавлении новых модулей хранения они включаются в карту инсталляции, определяющую области отказоустойчивости (правила размещения по модулям, носителям и стойкам, рядам и ЦОДам). Эти правила в свою очередь назначаются пулам данных. Сами данные в виде образов хранятся в этих пулах и распределяются по носителям в соответствии с правилами, присвоенными пулу.

При изменении количества носителей данные перераспределяются по доступным (определяется правилом размещения) носителям. Это справедливо как для увеличения количества носителей при добавлении новых модулей, так и для уменьшения в случае их выхода из строя — данные перераспределяются самостоятельно по доступным носителям, автоматически восстанавливая заданный уровень избыточности. Таким образом, при выходе из строя носителя нет необходимости срочно проводить его замену, если в наличии есть свободное пространство.

Благодаря эластичности архитектуры, масштабирование системы хранения проводится без особых усилий: добавляются новые модули с новыми носителями, и данные автоматически распределяются по всем доступным носителям. Вместе с объёмом также растёт и производительность системы хранения.

Глава 2

Использование

2.1 Подключение СХД "Resilient Cloud Storage" к клиентам

СХД "Resilient Cloud Storage" подключается к потребителям интерфейсами:

- Ethernet 10 или 40 Гбит/с.
- FibreChannel 8 или 16 Гбит/с.

СХД "Resilient Cloud Storage" является универсальной СХД и поддерживает блочное, объектное и файловое хранение:

- Блочные: RBD, iSCSI, FC
- Файловые: S3, CIFS, NFS.

2.2 Виртуализация существующих СХД

Технологии постоянно развиваются, и то, что несколько лет назад было high end, со временем не дотягивает до уровня средней системы.

СХД "Resilient Cloud Storage" позволяет повторно использовать существующие системы хранения, пока их производительность и затраты на обслуживание делают такое использование целесообразным. При этом дисковое пространство существующих систем хранения представляется в качестве виртуальных модулей хранения. Администрирование при этом не отличается от работы с оригинальными модулями СХД «Resilient Cloud Storage».

2.3 Примеры использования

Благодаря эластичной структуре, СХД "Resilient Cloud Storage" оптимально подходит для решения широкого спектра задач, где требуется плавное расширение ёмкости и увеличение мощности в связи с возрастающими требованиями. При этом одни данные, к которым требуется интенсивный доступ, могут размещаться на SSD-носителях, а другие, например, предназначенные для архивного хранения могут размещаться на классических медленных дисках.

Сочетание этих свойств позволяет использовать СХД "Resilient Cloud Storage" для широкого спектра высоконагруженных систем, таких как платформы виртуализации и виртуальных десктопов, обычных серверов, файловых серверов общего доступа.

Для стыковки с платформой виртуализации VMware используется следующее полностью программное решение: на каждом сервер ESXi ставится несколько виртуальных модулей, взаимодействующие с СХД Resilient Cloud Storage по протоколу RBD, а с сервером ESXi по iSCSI с использованием multipathing через виртуальный сетевой коммутатор. Такая архитектура позволяет эффективно распараллеливать обращения к СХД и избежать единственной точки отказа.

Для платформы виртуализации на базе гипервизора KVM организуется подключение по поддерживаемому данным гипервизором протоколу RBD.

Использование технологии моментальных снимков позволяет эффективно организовать резервное копирование данных без необходимости их регулярного копирования.

СХД "Resilient Cloud Storage" идеально подходит для длительного хранения видеозаписей с камер видеонаблюдения, медиа-архивов, резервных копий информации и других неструктурированных данных больших объёмов.

Глава 3

Преимущества

3.1 Эффективное программное обеспечение

СХД "Resilient Cloud Storage" является программно-аппаратным комплексом и включает в себя, помимо ПО, разработанного компанией RCNTEC, ряд Open Source технологий.

3.2 Горизонтальное масштабирование

В процессе увеличения количества носителей в "классической" системе хранения неизбежно возникает предел в виде производительности контроллеров. СХД "Resilient Cloud Storage" не имеет в своем составе централизованных контроллеров - клиенты общаются напрямую с модулями хранения данных, что позволяет создавать практически неограниченно масштабируемые системы хранения. С увеличением количества модулей наряду с объемом растет производительность и пропускная способность интерфейсов.

3.3 Разделение групп потребителей

Так как в СХД "Resilient Cloud Storage" данные организованы в пулы, то при необходимости можно изолировать данные разных клиентов (ДМЗ, ЦОД, VDI и др.) на этом уровне. При наличии особых требований можно выделить отдельные диски некоторым пулам, однако рекомендованным является использование дисков совместно несколькими пулами, что позволяет максимально использовать производительность большого количества дисков, преимущества отказоустойчивости и самовосстановления наряду с разделением данных клиентов.

3.4 Кэширование данных

Медленные диски большой емкости хорошо подходят для долговременного хранения данных, с которыми не предполагается интенсивная работа, но скорость этих дисков, как правило, невелика. СХД "Resilient Cloud Storage" позволяет использовать часть SSD носителей для кэширования данных медленного пула, что позволяет разместить большие объемы данных на дешевом хранилище и обеспечить высокую производительность чтения.

3.5 Быстрое клонирование и тонкие диски

Часто возникают регулярные задачи создания копии системы. Как правило, в этом случае из резервной копии создается новый сервер с копией системы. Это требует значительного времени и занимает дисковое пространство, равное объему оригинала. Для оптимизации данной задачи можно использовать функцию клонирования данных СХД "Resilient Cloud Storage" для быстрого создания копии системы, которая не будет занимать дополнительное пространство. Так как после клонирования копия и оригинал начинают "жить" независимо, то это особенно эффективно при создании копий больших систем или большого количества копий одной системы.

СХД "Resilient Cloud Storage" также поддерживает функцию "тонких дисков", что позволяет эффективно использовать дисковое пространство. Однако в этом случае необходимо следить за объемом свободного пространства в силу возможного превышения выделенного клиентам логического объема над физическим.