**Базовые технологии в системах хранения данных**

Информация становится одним из наиболее важных активов предприятия и определяет его конкурентоспособность. Одна из наиболее сложных проблем, которые встают перед менеджерами по информационным технологиям, — огромный объем распределенных данных, циркулирующих в информационной системе. Невозможность эффективно управлять ими может негативно влиять на прибыльность предприятия и ограничивать его способность к росту. Поэтому передача, хранение, защита и управление данными в условиях постоянного роста их объема и усложнения структуры стали сегодня основными проблемами для бизнеса.

Для многих система хранения данных ассоциируется с устройствами хранения и в первую очередь с дисковыми массивами. Действительно, дисковые массивы сейчас — основные устройства хранения данных, однако не стоит забывать, что обработка информации, формирование логической структуры ее хранения (дисковых томов и файловых систем) выполняется на сервере. В процедуры доступа к данным (помимо процессоров и памяти сервера) вовлечены установленные в нем адаптеры, работающие по определенному протоколу, драйверы, обеспечивающие взаимодействие» этих адаптеров с операционной системой, менеджер дисковых томов, файловая система и менеджер памяти ОС и т. д.

Как правило, система хранения данных содержит следующие подсистемы и компоненты: непосредственно устройства хранения (дисковые массивы, ленточные библиотеки), инфраструктуру доступа к устройствам хранения, подсистему резервного копирования и архивирования данных, ПО управления хранением, систему управления и мониторинга.

Традиционный подход к хранилищам данных состоит в непосредственном подключении серверов к системе хранения DAS (Direct Attached Storage). Помимо DAS, существуют устройства хранения данных, подключаемые к сети, — NAS (Network Attached Storage), a также компоненты сетей хранения данных — SAN (Storage Area Networks). И NAS-, и SAN-системы появились в качестве альтернативы архитектуре DAS. Причем каждое решение разрабатывалось как ответ на растущие требования к системам хранения данных и основывалось на использовании доступных в то время технологиях. Чтобы лучше понять преимущества сетевых моделей, начнем с традиционной.

**Системы хранения прямого подключения**

Устройства хранения DAS, также известные как SAS (Server Attached Storage), т. е. системы, подключаемые непосредственно к серверу, были разработаны много лет назад для расширения емкости хранения существующих серверов. В те времена при необходимости увеличения числа томов, связанных с приложениями, к серверу добавлялись новые диски либо приобретался новый сервер. Принимая во внимание технологические ограничения того времени (узкая полоса пропускания, медленные сети, дорогостоящие микропроцессоры) и относительно низкие требования к емкости и времени доступа, системы DAS были вполне адекватным решением.

Один из недостатков технологии DAS в сегодняшней ИТ-инфраструктуре — доступность дисковой системы, ассоциированной с данным сервером, только через это устройство, что приводит к созданию отдельных «островов» данных. Кроме того, файлы не могут совместно использоваться несколькими серверами, а для добавления дисковых накопителей к сети серверы должны быть переведены в автономный режим. Чтобы разрешить эти проблемы и обеспечить избыточность и высокую готовность данных, компаниям приходилось копировать данные на несколько серверов и использовать различные RAID-конфигурации. В настоящее время, невзирая на значительные изменения в объемах хранения и характере информационных систем, решения DAS остаются в основе своей неизменными. Поскольку с течением времени росли требования к системам хранения, компании просто добавляли диски к своим серверам и серверы к своим сетям. То, что когда-то было практичным решением, теперь привело к тому, что у компаний образовалась сложная инфраструктура систем хранения данных и огромные серверные фермы, которыми все труднее управлять.

**Сетевое хранение данных**

Архитектуры сетевых систем хранения были разработаны в 1990-х гг., и их задачей было устранение основных недостатков систем DAS. В общем случае сетевые решения в области систем хранения должны были реализовать три задачи: снизить затраты и сложность управления данными, уменьшить трафик локальных сетей, повысить степень готовности данных и общую производительность. При этом архитектуры NAS и SAN решают различные аспекты общей проблемы. Результатом стало одновременное сосуществование двух сетевых архитектур, каждая из которых имеет свои преимущества и функциональные возможности.

**Устройства хранения данных, подключаемые к сети NAS**

Основная задача систем NAS — упростить совместное использование файлов. На базовом уровне устройства NAS — это оборудование, которое подключается непосредственно к локальной сети. В этом состоит их основное отличие от систем с индивидуальными серверами с прямым подключением изолированных накопителей. Устройства NAS, часто называемые файлерами, состоят из единого головного устройства, выполняющего обработку данных и осуществляющего сетевое соединение цепочки дисков. Устройства NAS позволяют использовать системы хранения в сетях Ethernet, в них для организации совместного доступа к файлам применяется протокол TCP/IP. Эти устройства позволяют клиентам совместно использовать файлы, даже если клиентские системы работают под управлением различных ОС. В отличие от архитектуры DAS, в системах NAS не требуется переводить серверы в автономный режим для увеличения общей емкости; диски можно добавлять в структуру NAS простым подключением устройства в сеть.

NAS-устройства не могут совместно использовать носители других NAS-устройств в сети, при этом каждое добавляемое в сеть устройство хранения должно управляться как отдельный том. То, что NAS-устройства оперируют только в файловом формате и не способны разделять свою емкость хранения между несколькими узлами, — еще один ограничивающий фактор их применения.

**Сети хранения данных (SAN)**

Эти решения стали ответом не только на недостатки систем DAS и NAS, но, что более важно, на проблемы перегрузки каналов связи и задержки в локальных IP-сетях (10/100-Мбит/с). Впервые концепция SAN была предложена в 1998 г. Как и многие другие современные компьютерные технологии, она была заимствована из мира мэйнфреймов, где применялась, например, в центрах обработки данных для подключения компьютеров к системам хранения и распределенным сетям. Основное различие между SAN и NAS состоит в способе организации обмена данными между устройствами хранения и серверами. Вообще говоря, архитектура SAN нацелена на разрешение проблем, вызываемых интенсивными процедурами резервного копирования и обмена данными путем перенесения всей системы в выделенную подсеть. Основанные на протоколе Fibre Channel системы SAN позволяют в широких пределах изменять емкость системы хранения данных и гарантировать более высокую пропускную способность в пределах выделенной подсети (дисковые массивы и ленточные библиотеки, не оборудованные интерфейсами Fibre Channel, можно подключить к SAN, используя маршрутизаторы Fibre Channel-SCSI).

Оптимизация архитектуры SAN для протокола передачи данных на уровне блоков сделала естественным ее применение для работы с большими базами данных. Она стала рассматриваться как надежное решение, особенно для динамичного структурированного контента. В отличие от архитектуры NAS, сеть хранения не использует существующую локальную сеть для передачи данных между серверами и устройствами хранения; данные перемещаются по самой сети хранения, не вызывая в локальной сети избыточного трафика. Однако использование блочных команд ввода-вывода означает, что, несмотря на возможность совместного использования накопителей различными серверами, совместное использование файлов различными пользователями в общем случае невозможно, что вместе с высокой стоимостью и необходимостью привлечения значительных ресурсов для управления этими сложными решениями может ограничивать сферы применения данной технологии.

Таким образом, пользователи традиционно обращаются к устройствам DAS и SAN для передачи данных на уровне блоков. При этом можно более точно определить, где и как эти данные будут храниться. Однако в типичных клиент-серверных архитектурах предпочтительнее использовать серверы NAS для передачи данных на уровне файлов. Следует отметить, что в настоящее время созданы устройства, в которых обеспечивается совместимость как с блочными, так и с файловыми протоколами, т. е. работающие и как SAN, и как NAS.

**Управление жизненным циклом информации**

Вместе с увеличивающимся объемом данных растут и системы хранения. Однако, используя старые методы, управлять такими системами становится все труднее, а стоимость их резко возрастает. И при этом трудно реализовывать высокую доступность, достаточную производительность, адекватную темпам развития бизнеса, надежность хранения информации, которая соответствует требованиям бизнеса и государственных регулирующих органов по долгосрочному хранению информации.

Одно из решений этой проблемы — разделение данных по критерию их ценности для бизнеса и управление информацией с учетом изменения ее ценности во времени. Такой подход стали называть управлением жизненным циклом информации ILM (Information Life-cycle Management). Согласно этой концепции, наиболее важная на текущий момент информация должна автоматически перемещаться на самую быструю, надежную и защищенную систему хранения данных. В свою очередь менее важная перемещается на более дешевую и менее скоростную систему. Архивные данные, например, должны записываться на магнитные ленты и удаляться с рабочей системы, чтобы не снижать скорость доступа к самой критичной в данный момент информации. Уже ненужная информация будет автоматически удаляться из системы. Процесс этот должен быть цикличным.

Один из методов для разрешения противоречия между объемами хранимых данных и временем доступа к ним - это так называемое управление иерархическим хранением HSM (Hierarchical Storage Management). В HSM реализуется политика архивирования редко используемых данных; с дисков они переписываются на ленты с компрессией. Концепция ILM — шаг вперед по отношению к идеологии HSM. Большинство экспертов полагают, что ILM как бы «интеллектуализирует» традиционное понятие HSM, так как новый подход не просто оперирует блоками данных, а направлен на управление реальным содержанием данных на всем протяжении их жизненного цикла. Дело в том, что ILM — это стратегия проактивного управления информацией, она не сводится к какому-то определенному ПО или аппаратуре.

Управление жизненным циклом информации начинается с оценки и классификации имеющейся информации и ресурсов по ее хранению с точки зрения их бизнес-ценности. Первые два шага должна сделать организация, которая решила управлять своей информацией на основе ее жизненного цикла, а именно: определить цели управления информацией и классифицировать свою информацию на основе ее ценности для бизнеса.

Главное, что должно обеспечить ILM, — это расширенный набор политик, которые сможет задавать пользователь. После определения политик размещения данных могут применяться следующие технологические решения по их реализации: различные способы автоматической интеллектуальной миграции данных с систем хранения одного класса на другой, средства их географически удаленного и локального зеркалирования, инструменты резервного копирования и архивирования и т. д.

Внедрение ILM позволяет снизить совокупную стоимость владения за счет снижения затрат на оборудование и ПО. Автоматизация управления жизненным циклом информации означает такой уровень интеграции приложений, информационных ресурсов и сервисов управления хранением, при котором достаточно задать степень важности приложения или его информационных групп для того, чтобы соответствующий уровень обслуживания был применен автоматически.

**Резервное копирование**

Как известно, система резервного копирования — неотъемлемая часть системы хранения данных. Этот обязательный элемент обеспечения высокой доступности любых систем предназначен для создания резервных копий и восстановления данных. Кроме того, система резервного копирования — это один из необходимых методов обеспечения непрерывности бизнеса. Построение централизованной системы резервного копирования позволяет сократить совокупную стоимость владения ИТ-инфраструктурой благодаря оптимальному использованию устройств резервного копирования и сокращению расходов на администрирование (по сравнению с децентрализованной системой).

Централизованная система резервного копирования имеет многоуровневую архитектуру, в которую входят:

сервер управления резервным копированием, способный также совмещать функции сервера копирования данных;

один или несколько серверов копирования данных, к которым подключены устройства резервного копирования;

компьютеры-клиенты с установленными на них программами-агентами резервного копирования;

консоль администратора системы резервного копирования.

Администратор системы ведет список компьютеров-клиентов резервного копирования, устройств записи и носителей хранения резервных данных, а также составляет расписание резервного копирования. Вся эта информация содержится в специальной базе, которая хранится на сервере управления резервным копированием.

В соответствии с расписанием или по команде оператора сервер управления дает команду программе-агенту, установленной на компьютере-клиенте, начать резервное копирование данных в соответствии с выбранной политикой. Программа-агент собирает и передает данные, подлежащие резервированию, на сервер копирования, указанный ей сервером управления.

Сервер копирования сохраняет полученные данные на подключенное к нему устройство хранения данных. Информация о процессе (какие файлы копировались, на какие носители осуществлялось копирование и т. п.) сохраняется в базе сервера управления. Эта информация позволяет найти местоположение сохраненных данных при необходимости их восстановления на компьютере-клиенте.

Чтобы система резервного копирования сохраняла непротиворечивые данные компьютера-клиента, они не должны подвергаться изменениям в процессе их сбора и копирования программой-агентом. Для этого приложения компьютера-клиента должны завершить все транзакции, сохранить содержимое кэш-памяти на диск и приостановить свою работу. Этот процесс инициируется по команде программы-агента, которая передается приложениям компьютера-клиента.

Поскольку система резервного копирования предназначена для восстановления данных после сбоя или аварии, созданные резервные копии необходимо проверять на предмет целостности и работоспособности. Кроме того, при построении системы резервного копирования необходимо уложиться в сокращенное «окно» резервного копирования. Вообще говоря, требование круглосуточной работы информационных систем сокращает практически до нуля доступный временной интервал остановки приложений, необходимый для осуществления операции резервного копирования («окно» резервного копирования).

Необходимо уменьшить трафик данных резервного копирования в общей локальной сети. Структура системы резервного копирования предполагает передачу данных с компьютеров-клиентов на серверы копирования через эту локальную сеть. Поэтому очевидно, что при наличии большого объема данных, подлежащих резервному копированию, использование общей локальной сети для их передачи резко повысит трафик в сети и сделает ее недоступной для других приложений.

До появления сетей хранения данных SAN для сокращения трафика резервного копирования в основной сети применялась выделенная сеть резервного копирования, а также многоуровневая структура, содержащая несколько серверов копирования. Выделение сервера копирования и локализация трафика резервного копирования между этим сервером и «тяжелыми» серверами, несущими основную информационную нагрузку, позволяют сократить нагрузку на общую локальную сеть.

**Резервное копирование с использованием SAN**

Применение SAN позволяет полностью перенести трафик резервного копирования с локальной сети на сеть хранения. Существует два варианта реализации: без загрузки локальной сети, или внесетевое копирование (LAN-free backup), и без участия сервера, или внесерверное копирование (Server-free backup).

**Внесетевое копирование**

При внесетевом копировании данные с диска на ленту и обратно передаются внутри SAN. Исключение сетевого сегмента из пути резервного копирования данных позволяет избежать излишних задержек на передачу трафика через сеть IP и платы ввода-вывода. Нагрузка локальной сети падает, и резервное копирование можно проводить практически в любое время суток. Однако пересылку данных выполняет сервер, подключенный к SAN, что увеличивает нагрузку на него. Благодаря протоколу Fibre Channel с помощью одного оптического кабеля может быть организовано несколько каналов передачи данных. При этом весь объем резервируемых данных с backup-серверов хранения направляется на ленточное устройство, минуя локальную сеть. В этом случае локальная сеть необходима лишь для контроля работы самих backup-серверов со стороны главных серверов. Таким образом, только небольшой объем метаданных, которые содержат информацию о резервируемых данных, передается по локальной сети. Главные серверы отвечают в целом за политику резервного копирования данных в своем сегменте или зоне ответственности. Все backup-серверы по отношению к главному серверу являются клиентами. Считается, что рассматриваемый метод резервного копирования может максимально задействовать пиковую полосу пропускания Fibre Channel.

В качестве протокола, применяемого для передачи данных между серверами и библиотеками, могут использоваться как SCSI поверх Fibre Channel, так и IP поверх Fibre Channel, тем более что большинство FC-адаптеров и FC-концентраторов работают одновременно с обоими протоколами (IP и SCSI) на одном Fibre Channel-канале.

**Внесерверное копирование**

Вообще говоря, данный тип резервного копирования представляет собой дальнейшее развитие метода внесетевого копирования (LAN-free), поскольку уменьшает количество процессоров, памяти, устройств ввода-вывода, задействованных в этом процессе. Данный процесс архивирует разделы целиком, в отличие от пофайлового архивирования, но при этом позволяет восстанавливать отдельные файлы. По определению, при вне-серверном копировании данные копируются с диска на ленту и обратно без прямого участия сервера. Поскольку для резервного копирования требуется наличие некоторого дополнительного третьего узла, полностью отвечающего за процесс копирования, то отсюда происходит и другое название этого подхода — копирование с участием третьей стороны (Third\_-Party Copy, 3PC). Так, в качестве подобного оборудования может использоваться маршрутизатор хранилищ данных, который берет на себя функции, ранее выполнявшиеся сервером.

Одно из преимуществ архитектуры SAN — отсутствие жесткой привязки составляющих ее систем к каким-либо устройствам хранения данных. Это свойство и заложено в основу технологии резервного копирования без участия сервера. В данном случае к дисковому массиву может иметь прямой доступ как сервер данных, так и устройства, принимающие участие в копировании с дисковых массивов. Резервному копированию блоков данных, относящихся к какому-либо файлу, предшествует создание некоего индекса или списка номеров принадлежащих ему блоков. Это и позволяет в дальнейшем привлечь внешние устройства для резервного копирования.

Таким образом, внесерверное копирование позволяет напрямую перемещать данные между подключенными к сети SAN дисковыми массивами и библиотеками. При этом данные перемещаются по сети SAN и не загружают ни локальную сеть, ни серверы. Такое копирование считается идеальным для корпоративных сетей, которые должны функционировать в непрерывном режиме 24 часа в сутки, 7 дней в неделю. Особенно для тех, для которых временной период, в течение которого можно выполнять резервное копирование без существенного влияния на работу пользователей и приложений, становится недопустимо малым.

**Репликация данных**

Современные дисковые массивы обладают средствами создания копий данных внутри самого массива. Данные, созданные этими средствами, носят название Point-In-Time (PIT)-копий, т. е. фиксированных на определенный момент времени. Существует две разновидности средств создания PIT-копий: клонирование и «моментальный снимок» (snapshot). Под клонированием обычно понимают полное копирование данных. Для него требуется столько же дискового пространства, как и для исходных данных, и некоторое время. При использовании такой копии нет нагрузки на дисковые тома, содержащие исходные данные. Иными словами, нет дополнительной нагрузки на дисковую подсистему продуктивного сервера.

Механизм работы «моментальных снимков» иной и может быть реализован как программно на продуктивном сервере, так и аппаратно внутри массива. В момент, когда необходимо начать резервное копирование, программа-агент дает команду приложению завершить все транзакции и сохранить кэш-память на диск. Затем создается виртуальная структура — snapshot, представляющая собой карту расположения блоков данных, которую ОС и другое ПО воспринимает как логический том. Приложение прерывает стандартный режим работы на короткое время, необходимое для сохранения данных. После этого приложение продолжает работать в стандартном режиме и изменять блоки данных, при этом перед изменением старые данные блока с помощью драйвера snapshot копируются в область кэш-памяти snapshot и в карте расположения блоков данных указывается ссылка на новое местоположение блока. Таким образом, карта snapshot всегда указывает на блоки данных, полученные на момент завершения транзакций приложением. Блоки данных, которые не были изменены, хранятся на прежнем месте, а старые данные измененных блоков — в области кэш-памяти snapshot. Программа-агент копирует непротиворечивые данные, полученные на момент завершения транзакций приложением, осуществляя доступ к ним через драйвер snapshot, т. е. используя карту расположения блоков. Создание копий с помощью «моментальных снимков» экономит дисковое пространство, но создает дополнительную нагрузку на дисковую подсистему продуктивного сервера. Какой из методов создания PIT-копий выбрать, решается на этапе проектирования системы резервного копирования, исходя из бизнес-требований, предъявляемых к системе.

**Виртуализация ресурсов хранения**

Виртуализацию трудно отнести к совершенно новым технологиям — идеи виртуализации различных вычислительных ресурсов тем или иным образом реализовывались и ранее. А вот необходимость в виртуализации ресурсов хранения объясняется рядом причин. Прежде всего это, конечно, резкий рост объемов данных. Обострились проблемы хранения и управления большими объемами информации. Это связано и с широким распространением «островов» данных (напомним, что под ними понимаются данные, находящиеся на различных носителях в гетерогенных системах хранения, нередко территориально удаленных друг от друга и работающих под управлением разных ОС). Для обслуживания таких «островов», отличающихся сложностью конфигураций аппаратных и программных средств, а также разнообразием используемых технологий, необходимы дополнительные материальные и человеческие ресурсы. Ухудшилась оперативность доступа к данным, находящимся в гетерогенных системах хранения, что приводит к значительным финансовым потерям для компаний, бизнес которых связан с оперативной обработкой актуальной и особо важной информации. В той или иной степени перечисленные выше проблемы решаются с помощью технологий виртуализации ресурсов хранения.

Под виртуализацией ресурсов хранения обычно понимается отображение любого количества разнородных носителей, устройств и систем хранения (JBOD, RAID, RAIT и т. д.) в виде единого хранилища данных (так называемого виртуального пула), управление которым осуществляется централизованно. Можно сказать, что при применении технологий виртуализации «разрываются» физические связи между серверами и устройствами хранения разных типов, а физическая память преобразуется в единый логический пул, состоящий из отдельных гетерогенных устройств хранения, прозрачный доступ к которым обеспечивается независимо от их технических особенностей и территориального расположения. Технологии виртуализации ресурсов хранения позволяют:

оптимизировать использование имеющихся гетерогенных ресурсов хранения и управлять огромными пулами внешней памяти;

упростить управление гетерогенными системами хранения, разгрузив тем самым системных администраторов;

сократить срок резервного копирования и восстановления данных, повысив отказоустойчивость всей корпоративной информационной системы;

снизить совокупную стоимость владения корпоративной системой хранения (путем устранения «островов» данных, более эффективного использования уже имеющихся устройств и систем хранения, обеспечения работы с ресурсами хранения в гетерогенной среде, возможности построения корпоративной системы хранения из компонентов разных производителей, оптимальных по цене, и т.д.).

**Уровни виртуализации**

Итак, суть виртуализации — это отделение логического представления устройств от физического размещения данных, что позволяет преодолеть физические ограничения традиционных массивов. В пределах сетевой среды хранения данных, в которой может использоваться виртуализация, существуют три основных уровня: серверов, сети SAN и системный. В каждом конкретном случае для достижения максимальной эффективности эти уровни могут использоваться как совместно, так и независимо друг от друга.

На уровне серверов виртуализация может быть осуществлена путем использования ПО, постоянно находящегося на сервере и независимого от накопителей. При помощи этого ПО ОС заставляет сервер вести себя так, будто он связан с конкретным типом устройства хранения, хотя фактически он поддерживает связь с виртуальным диском. Виртуализация на уровне сервера может быть применена как в гомогенной SAN-среде, так и вне ее. Этот механизм характеризуется ограниченной способностью к взаимодействию с аппаратными или программными компонентами. Обычно такое решение рекомендуется для систем начального уровня, поскольку его довольно просто осуществить и использовать.

Более широко известна виртуализация на уровне сетевой структуры. Обычно она используется для открытой SAN-среды как с традиционными, так и с виртуализованными системами хранения. Она охватывает всю архитектуру SAN и здесь может применяться как основная реализация асимметричного объединения накопителей в общий пул в рамках среды SAN.

На системном уровне виртуализация осуществляется на основе контроллеров дисковых массивов, работа которых независима от активности узла. Контроллеры дисковых массивов создают виртуальные диски, выполняют мгновенные копии состояний системы и клонирование данных при взаимодействии с управляющим ПО. Виртуализация на уровне индивидуального контроллера системы хранения — это значительный шаг за пределы классической технологии RAID. Виртуализация на этом уровне особенно хорошо подходит для среды, требующей высокой производительности, готовности данных, отказоустойчивости, эффективности управления системой хранения, репликации данных и поддержки кластеров.

Упрощенное управление накопителями — еще одно значительное преимущество виртуализации на системном уровне, позволяющее администраторам оперировать атрибутами системы храпения, а не составляющими ее физическими объектами. Виртуализация позволяет пользователям создавать для своих систем хранения данных единую модель управления, не принимая во внимание тип носителей, тем самым позволяя избежать работы по физическому размещению данных. Такой подход снижает сложность развертывания системы хранения, позволяет администраторам управлять всеми накопителями как единым консолидированным пулом и переносит задачи управления с уровня индивидуальных накопителей на уровень всего пула.

Способность динамического расширения емкости виртуальных дисков без нарушения работы выполняющихся приложений значительно улучшает эффективность на системном уровне. Виртуализация позволяет следить за эффективностью использования емкости томов или всего пула накопителей и по мере необходимости выделять дополнительные ресурсы в динамическом режиме.

Возможность по желанию пользователя увеличивать (но не уменьшать) размер виртуального диска (в зависимости от наличия свободного дискового пространства в пуле) позволяет достичь высоких результатов в эффективности использования емкости носителей. Эта особенность виртуализации повышает эффективность использования накопителей как благодаря устранению ограничений на доступный размер емкости хранения, так и благодаря перераспределению данных по мере роста пула.

**Виртуальные библиотеки**

В последнее время все большую популярность приобретает идея ускорения резервного копирования с помощью специальных массивов из относительно дешевых жестких дисков с интерфейсом Serial ATA (SATA), которые служат для промежуточной записи копируемых данных. Реализуется эта идея, как правило, с помощью так называемых виртуальных библиотек VTL (Virtual Таре Library), которые «подставляют» серверу резервного копирования недорогой дисковый массив как обычную ленточную библиотеку. При этом, разумеется, нет необходимости в новом ПО для резервного копирования или модернизации старого.

Предпосылками этой тенденции, с одной стороны, было появление быстрых и недорогих жестких дисков SATA очень большой емкости, а с другой — потребность в проведении процедуры сохранения необходимого объема данных за ограниченное время, называемое «окном» резервного копирования.

«Окно» копирования определяется несколькими критериями. Поскольку создание резервной копии требует абсолютного доступа к данным, обычно этот процесс проводится в нерабочие часы, когда он оказывает минимальное влияние на работу персонала, загруженность серверов и локальной сети. Если объем переносимых во время одной сессии данных превышает размер картриджа магнитной ленты, возникает необходимость ручной замены носителя по требованию программы резервного копирования. Итак, помимо отличной скорости, система хранения на жестких дисках может помочь и в таком случае, предоставляя для записи резервной копии свободное пространство, превышающее возможности одного картриджа.

Достоинства сохранения резервных копий данных на жестких дисках (в первую очередь — высокая скорость сохранения и восстановления) и на магнитных лентах (низкая стоимость хранения и неограниченный объем) объединяются в решениях Disk-to-Disk-to-Tapc (D2D2T). Такой подход предполагает использование дисковой кэш-памяти в качестве промежуточного этапа в процедуре резервного копирования, конечная цель которого по-прежнему — магнитные ленты. Как правило, VTL обеспечивает 100%-ную аппаратную эмуляцию ленточных накопителей определенного типа. Таким образом, о совместимости нового оборудования с используемыми в организации операционными системами и программами резервного копирования беспокоиться обычно не приходится. Типовой сценарий подключения VTL может выглядеть примерно так. К адаптеру SCSI вместо ленточного накопителя подключается устройство VTL, к нему, в свою очередь, — требуемый ленточный накопитель. Устройство динамически настраивается на имитацию подключенного к нему стримера. Программа резервного копирования, не заметив подмены, продолжает общаться с ним так же, как с обычным ленточным накопителем.

**Список литературы**

Журнал Upgrade4\_08\_05