

Гипервизоры и СХД: часть 2

Обзор особенностей интеграции гипервизора Citrix XenServer с компонентами уровня хранения данных. Продолжение темы, начатой в предыдущем номере SN (SN № 4/48 (2011), www.storagenews.ru/48/AVRORAID_VMware_Store_48.pdf).



Сергей Платонов — менеджер по продуктам, компания AVRORAID.

Введение

Очередной мажорный релиз Citrix XenServer был представлен в конце сентября 2011 г., а публичная бета была доступна уже в конце первого месяца лета. Citrix позиционирует гипервизор как платформу виртуализации для поставщиков облачных ресурсов. Для управления ресурсами виртуализации вендор ориентируется на использование *Microsoft System Center 2012 Virtual Machine Manager*.

С точки зрения взаимодействия с системами хранения данных, многие пользователи этой платформы виртуализации ожидали, в первую очередь, аналога Storage vMotion. Но эта функция не была реализована в новой версии гипервизора от Citrix и, скорее всего, не стоит ожидать ее появления и в будущем.

Подходы к использованию СХД

Citrix использует совершенно иной подход, чем VMware. Вместо интеграции различных функций СХД внутри гипервизора компания позволяет тесно интегрироваться с системами хранения посредством технологии StorageLinks и использовать их возможности.

Неофициально XenServer позволяет использовать mdraid, что открывает дополнительные возможности для пользователей. Например, можно выполнять зеркалирование двух неотказоустойчивых систем хранения данных или же объединять локальные диски в RAID-массив при отсутствии бюджета на хороший аппаратный контроллер.

При создании SR на блочных носителях используется Logical Volume Manager (на котором размещаются образы VHD), к производительности которого многие пользователи Linux имеют претензии.

Кроме того, при использовании локальных носителей возможно использование файловой системы ext v3, далеко не новой версии файловой системы.

Мы выполнили тестирование производительности виртуальных дисков с использованием приложения IOMeter 2006. За 100% принята производительность, снятая с жесткого диска без виртуализации. Как показано на рис. 1 и 2, производительность виртуальных дисков, особенно на потоковых операциях, требующих высокой пропускной способности, сильно отстает от производительности СХД. Для обеспечения высокой производительности дисковой подсистемы необходимо использовать Raw LUN Mapping.

При использовании файловых хранилищ можно рассчитывать только на NFSv3.

Из нововведений в XenServer 6.0 в части работы с файловыми хранилищами добавлена только возможность поддержки NFS для обеспечения отказоустойчивости. HA (high availability) Heartbeat диск теперь может быть размещен на NFS-разделе — решение предсказуемое в связи с тем, что в ряде случаев Citrix рекомен-

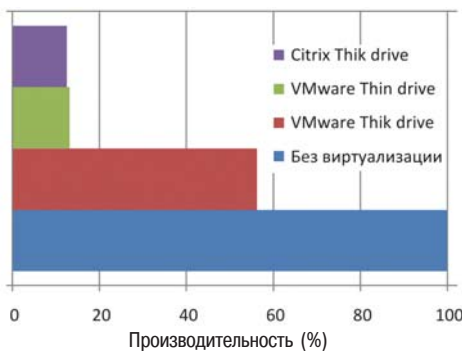


Рис. 1. Производительность (в процентах) для последовательной записи блоками по 256 Кбайт.

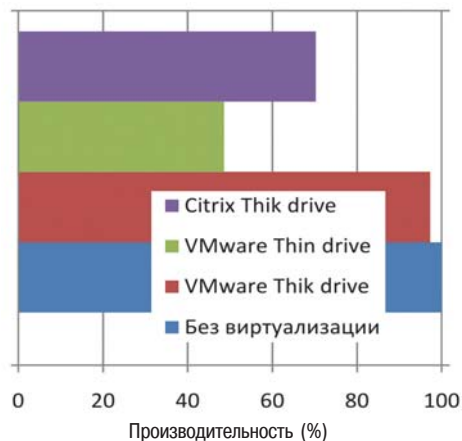


Рис. 2. Производительность (в процентах) для случайных обращений блоками 8Кбайт (67% — чтение, 33% — запись).

дует использовать файловые хранилища в качестве разделяемого ресурса.

Сама архитектура Storage подсистемы не слишком сложна, и администраторам нужно разобраться только в типах репозитория и различиях между ними. Для контраста можно вспомнить более десятка аббревиатур, используемых VMware. Но, несмотря на всю сложность архитектуры, управление VMFS Datastores в ESX значительно проще. Такие операции, как добавление LUN в DS, увеличение размера LUN, обновление VMFS проходят без дополнительных операций, прерывающих работу виртуализированной инфраструктуры. Вкупе с отсутствием аналога Storage vMotion трудности с наращиванием дискового пространства для существующих SR не позволяют XenServer обеспечить требуемый уровень эластичности.

Поддержка протоколов

Репозитории (SR) в XenServer могут использовать локальные диски, файловые хранилища по протоколу NFS, а также блочные СХД, подключенные по протоколам FC, iSCSI, SAS. Мы не видим программного инициатора FCoE, также ограничена поддержка Infiniband. В списке известных трудностей упоминается о проблемах с производительностью файловых хранилищ в средах 10G. При ориентации на облака, в дальнейшем, компания Citrix возможно более внимательно посмотрит на данные протоколы.

Thin Provisioning

Выделение пространства по требованию — Thin Provisioning (TP) — является одной из самых востребованных функций для построения облаков.

Основанные на LVM Storage Repositories не поддерживают выделение пространства по требованию для активных виртуальных дисков. Но теперь Thin Provisioning доступен для мгновенных снимков (снэпшотов).

При использовании файловых хранилищ обеспечивается полная поддержка выделения пространства. Мы могли бы порекомендовать использование файловых хранилищ, если бы не производительность NFSv3.

RAW LUN MAPPING

Citrix XenServer не поддерживает SR размером более 2 Тбайт (за исключением некоторых массивов, использующих технологию Storage Links).

Все рассматриваемые гипервизоры поддерживают прямое назначение LUN вир-

туальным машинам для обеспечения возможности подключения логических дисков LUN большого размера и/или высокой производительности. Использование данной возможности несет за собой как некоторые преимущества, так и недостатки. Но в XenServer подключение логических дисков LUN большого размера также связано и с некоторыми сложностями по настройке подключения. Проброс блочных устройств в XenServer потребует от администратора умений обращаться с консолью, а также поиска необходимых инструкций в неофициальных источниках.

Кроме того, Citrix опять не оправдал ожидания пользователей – не добавил поддержку NPIV.

В настоящее время практически всех пользователей систем хранения интересует уровень Quality of Service (QoS). Например, СХД на платформе AVRORAID предоставляют максимальную гибкость в сравнении с гипервизорами, поддерживая приоритизацию инициаторов и/или приложений для гарантированного уровня производительности (QoS настраиваются на конкретные группы WWN). Также СХД на платформе AVRORAID предоставляет выполнение реконструкции в 4 раза быстрее других СХД и гарантию сохранения заявленной скорости чтения и записи при выходе 1 или 2 дисков в RAID6.

Интеграция с СХД

Для интеграции с системами хранения данных Citrix использует технологию StorageLink. В XenServer 6.0 больше не требуется специальная виртуальная машина с ОС Windows для организации работы StorageLinks. Эта технология позволяет с максимальной эффективностью использовать возможности систем хранения данных, такие как: выделение пространства по требованию (Thin Provisioning), дедупликация, мгновенные снимки и другие.

StorageLinks предоставляет единый интерфейс взаимодействия со всеми поддерживаемыми СХД и “общается” с этими СХД на понятном им языке. Кроме того, SL предоставляет набор API, позволяющих интегрировать среды XenServer и Nureg-V со сторонними средствами резервного копирования и восстановления, а также со средствами управления инфраструктурами.

Использование нативных функций СХД является более эффективным решением, чем копирование данного функционала в файловую систему гипервизора. Но интеграцией с SL обладает довольно незначительный список систем хранения данных, и Citrix придется развивать сообщество для доказательства состоятельности данного решения. К сожалению, отказ от StorageLink Gateway привел к потере ряда поддерживаемых СХД.

Катастрофоустойчивость

Развивая, с одной стороны, партнерские отношения с Microsoft в части использования средств управления виртуальными

ресурсами, Citrix, с другой стороны, пытается отойти от зависимости некоторых своих функций от использования иных продуктов Microsoft. Так, для использования встроенных функций защиты от катастроф больше не требуется установка специальной виртуальной машины с ОС Windows.

Новая функция Site Recovery заменяет собой StorageLink Gateway DR, т.е. перестает быть частью SL. Это приводит к тому, что Site Recovery поддерживает практически все массивы, обладающие функцией репликации данных, подключенные к репозиториям iSCSI или Hardware HBA.

Табл. 1. Сравнение функциональных особенностей VMware vSphere 5 и Citrix XenServer 6.0

| Критерий | VMware vSphere 5 | Citrix XenServer 6.0 |
|---|--|---|
| Поддерживаемые интерфейсы подключения СХД | FC, iSCSI (аппаратная и программная поддержка), FCoE (аппаратная и программная поддержка), SAS, Infiniband (SRP) | FC, iSCSI (аппаратная и программная поддержка), FCoE (только аппаратная поддержка), SAS |
| Использование файловых хранилищ | NFSv3 | NFSv3 |
| Возможность использования локального хранилища | Да | Да |
| Возможность использования программного RAID, в том числе для зеркалирования СХД | Нет | Да, неофициально |
| Поддержка MPIO | Да | Да |
| Использование SSD | Как swap | Нет |
| Загрузка из сети | Да | Да |
| Формат виртуальных дисков | vmdisk поверх VMFS-5 и NFS | vhd поверх ext3 и NFS, vhd поверх LVM |
| Максимальный размер виртуального диска | 2ТБ | 2ТБ, 15ТБ для некоторых СХД |
| Расширение виртуальных дисков “на лету” | Да | Нет |
| Возможность расширения хранилища | Добавление LUN в имеющееся хранилище без перебора в работе работы Расширение размера LUN без перебора в работе | Расширение размера LUN требует отключения от СХД или перезагрузку хоста |
| Возможность классифицировать хранилища и объединять их в кластера | Да | Нет |
| Возможность “прокинуть” LUN | Да, 2 режима | Да |
| Возможность “прокинуть” HBA | Да | Ограниченно |
| Поддержка NPIV | Да, для RDM | Нет |
| Поддержка “тонких” дисков | Да, в том числе thinreclamation | Только для файловых хранилищ |
| “Горячая” миграция виртуальных машин между хранилищами | Да, в том числе между хранилищами разных типов | Только “холодная” миграция |
| QoS для СХД | SIOC обеспечивает QoS, в том числе для NFS DRS обеспечивает перемещение и правильное размещение гостей на хранилищах | Только для блочных хранилищ Только для гостей, доступных с одного хоста |
| Обеспечение отказоустойчивости без внешней СХД | VSA, прототип CloudFS | Неофициальная возможность использования DRBD |
| Дополнительные средства обеспечения безопасности SAN | Да, изоляция LUN | Нет |
| Интеграция с СХД | VAAI – API для взаимодействия с СХД и перемещения ресурсоемких операций на сторону СХД, частично стандартизован VASA – API, позволяющее определять параметры СХД VAMP – позволяет разрабатывать собственные плагины для MPIO | StorageLink – позволяет использовать функции СХД, такие как ThinProvisioning, мгновенные снимки и копии |
| Связанные образы | Нет | Да, при клонировании из шаблонов, PVS |
| Обеспечение катастрофоустойчивости | SRM | Site Recovery |

Балансировка нагрузки

Теперь для функций балансировки нагрузки используется Virtual appliance с установленной ОС GNU/Linux вместо Microsoft Windows, что позволило значительно сократить используемый им размер.

Возможность балансировки нагрузки по операциям ввода-вывода и предоставление QoS для СХД в XenServer появились раньше, чем у VMware. Но для платформы виртуализации от Citrix данные возможности так и остались на базовом уровне.

QoS в XenServer доступен только для блочных СХД и применяется лишь для виртуальных дисков, доступных с одного хоста, и не распространяется на пул.

Для построения облачных инфраструктур (как частных, так и публичных) QoS является необходимой функцией, и Citrix возможно ее будет развивать.

Связанные образы

Использование связанных образов и сервиса Provisioning Services является уникальным для Citrix. Эта технология позволяет значительно сократить используемое дисковое пространство и упростить управление большими инфраструктурами.

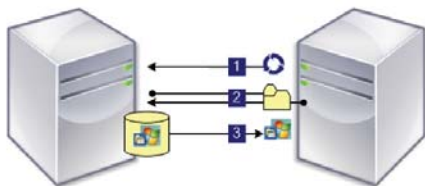


Рис. 3. Процесс загрузки VM с помощью Provisioning Server (PVS):

- (1) — целевое устройство начинает взаимодействие с PVS и получает необходимую информацию;
- (2) — целевое устройство отправляет запрос на получение файла нагрузки и получает этот файл;
- (3) — целевое устройство загружает назначенный образ виртуального диска.

При создании виртуальных машин с использованием шаблонов для каждой из них выполняется запись только измененных блоков, неизменные блоки при этом читаются прямо с мастер-образа VHD.

Инфраструктура Provisioning Services основана на технологии потоковой передачи программного обеспечения, что позволяет выполнять инициализацию и повторную инициализацию компьютеров в реальном времени с одного образа на общем диске. Такой подход полностью устраняет необходимость администрирования отдельных систем и установки в них исправлений. Вместо этого все операции выполняются с главного образа. Локальный жесткий диск каждой системы можно использовать для кэширования данных во время выполнения, а в некоторых случаях удалять из системы, что позволяет экономить электроэнергию, снижать частоту отказов и уменьшать риски безопасности.

С помощью Provisioning Services администраторы готовят специальное устройство (главное целевое устройство) для работы с образами и устанавливают на него все необходимое программное обеспечение (рис. 3). Далее с жесткого диска главного устройства создается образ виртуального диска и сохраняется в сети (на сервере Provisioning Server или запоминающем устройстве).

Когда виртуальный диск становится доступным в сети, целевое устройство может работать без жесткого диска и загружаться непосредственно по сети. Provisioning Server выполняет потоковую передачу содержимого образа виртуального диска в целевое устройство по запросу, в режиме реального времени. При этом целевое устройство работает так же, как в случае использования локального диска. В отличие от технологии тонкого клиента, обработка выполняется на целевом устройстве.

Provisioning Services 6.0 вышел вслед за XenServer 6.0 и предоставляет следующие дополнительные возможности:

- **Integrated vDisk Version Management:** позволяет создавать и управлять версиями vDisk прямо из консоли Provisioning Services Console. Версии могут быть созданы для различных устройств и групп для обеспечения жизненного цикла виртуальных дисков разных версий с сохранением дискового пространства;
- **Automated vDisk Updates:** позволяет создавать запланированные задачи, автоматизирующие процесс обновления vDisk через Microsoft System Center Configuration Manager (SCCM) и Windows Server Update Services (WSUS);
- **Support for vDisk Replication:** позволяет использовать решения по репликации хранилищ, такие как Microsoft DFS Replication для распределения дисков по географически разнесенным сайтам. Provisioning Servers обслуживает набор доступных vDisks и их версий и балансирует нагрузку, а также обеспечивает разграничение доступа серверов к хранилищам.

Выводы

Подход Citrix XenServer является верным: пусть функциями СХД занимаются сами СХД. Но Storage Links поддерживается весьма ограниченным числом производителей систем хранения данных, поскольку, во-первых, есть проблемы с поддержкой набирающих популярность интерфейсов, да и, во-вторых, просто отсутствуют некоторые важные функции, значительно осложняющие жизнь интеграторам и администраторам облачных сервисов. При всем этом архитектура гипервизора позволяет Citrix расширить в будущем функционал до желаемого уровня, а такие функции, как PVS, не имеют аналогов у конкурентов.

Сравнительный анализ рассмотренных платформ виртуализации — VMware vSphere 5 и Citrix XenServer 6.0 — представлен в табл. 1.

Сергей Платонов,
компания AVRORAID

RAIDIX TECHNOLOGY
AVRORAID

AVRORA 2.0*

СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ

*новый продукт компании AvroRAID

- ✓ полная отказоустойчивость
- ✓ высокая производительность
- ✓ широкий модельный ряд