NETAPP TECHNICAL REPORT

Windows Multipathing и Data ONTAP: Fibre Channel и iSCSI

Ryan Hardin, NetApp

Июнь 2013 | TR-3441

Версия 3.0

Коротко о главном:

Этот документ содержит описание различных методов организации *multipathing*, то есть многопутевого подключения к системами хранения данных по протоколам iSCSI и Fibre Channel для хост-OS семейства Microsoft® Windows® применительно к NetApp® Data ONTAP®. «За» и «против» каждого метода подробно рассмотрены с тем, чтобы дать читателю возможность выбрать среди них наилучший, применительно к его системе.

Оглавление

1 Введение

Для построения высокодоступной сети хранения данных (*storage area network*, SAN), необходимо предпринять ряд шагов, чтобы отдельный локальный отказ в ней не вел к общему отказу доступности хранимых данных. В этом документе мы рассмотрим методы организации избыточности в каналах передачи данных между хост-системами и системами хранения данных для построения надежной инфраструктуры SAN.

2 Цели документа и его аудитория

Этот документ предназначен для системных архитекторов и администраторов систем хранения данных, разрабатывающих решение с использованием протоколов iSCSI и FC (Fibre Channel) и использующее системы хранения NetApp. Мы рассчитываем, что:

- Читатель имеет общее представление об устройстве и работе контроллеров систем хранения NetApp, и их программной «начинке», по крайней мере в части, связанной с блочным доступом к данным.
- Читатель знаком с блочными протоколами, такими как *Fibre Channel* и *iSCSI*.

В этом документе мы особенно сосредотачиваемся на функциональности *NetApp Data ONTAP 8.2* и *Microsoft Windows Server® 2012*, но также включаем и информацию о предыдущих версиях *Data ONTAP* и *Microsoft Windows Server* в тех случаях, когда это необходимо.

Полный список справочных материалов приведен в конце документа.

3 Multipathing

Multipathing, или *многопутевое подключение* – это возможность использовать несколько путей доступа с сервера к хранимым на системе хранения данным. *Multipathing* защищает от аппаратных сбоев и отказов (обрыв кабелей, отказ коммутаторов и их портов, выход из строя HBA, и так далее), и обеспечивает более высокую производительность за счет использования *агрегированной*, то есть объединенной полосы пропускания нескольких используемых каналов. Когда один путь к данным или соединение становится неработоспособным, то ПО обеспечения *multipathing* автоматически перемещает операции на другие доступные пути. *Multipathing* часто делится на две категории: так называемый режим *active-active* и *active-passive*. Чаще всего *multipathing* рассматривается как решение *active-active* , при котором ввод-вывод к одному LUN идет по нескольким путям одновременно.

3.1 Устранение Единой Точки Отказа (Single Points of Failure)

Любые электрические и механические компоненты в любой момент могут перестать работать. Способом достичь высокой надежности является метод «*устранения единой точки отказа*» в системе. При этом при отказе каждого индивидуального компонента системы система в целом остается доступной пользователям. Любое решение *multipathing* должно использовать раздельные адаптеры, кабеля, и коммутаторы, чтобы устранить наличие одной уязвимой «*точки отказа*».

3.2 Стек хранения в Windows

Когда приложение пишет данные на диск, эти данные проходят по последовательности уровней стека хранения хост-OS, и затем по каналу передачи данных к системе хранения (например это может быть *parallel SCSI*, *serial SCSI (SAS), Fibre Channel*, *iSCSI*, и так далее).

Рисунок 1 схематично показывает устройство стека хранения в *Microsoft Windows Server 2012*.

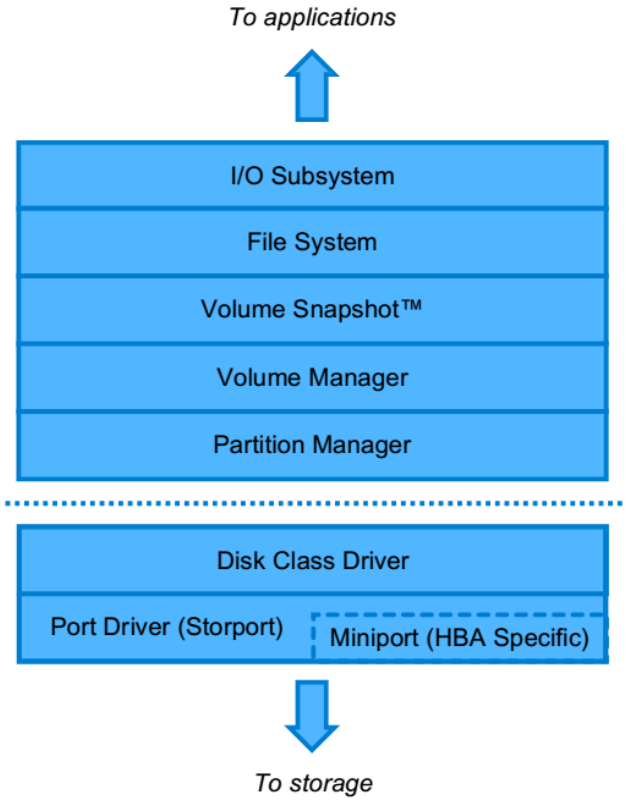

Рис. 1) Стек хранения в Microsoft Windows.

Multipathing обеспечивается некоторым усложнением отдельных уровней этого стека. Приложения осуществляют записи на одну файловую систему на raw-устройстве. Слой обеспечения *multipathing* принимает запросы и перенаправляет их в один из нижележащих путей передачи данных. Это перенаправление происходит прозрачно для других уровней стека, как для лежащих выше, так и ниже уровня *multipathing*. Процесс разделения одного на несколько путей может осуществляться разными способами, и каждый из них имеет свои достоинства и недостатки, которые будут перечислены ниже.

3.3 Выбор правильного решения Multipathing

Три поддерживаемых решения *multipathing*, о которых пойдет речь в документе, это: *link aggregation*, *iSCSI с множественными соединениями в сессии* (*multiple connections per session*, *MCS*), и *MPIO*.

Link aggregation может использоваться совместно с двумя другими методами, MCS и MPIO, и имеет ряд преимуществ и недостатков, описанных детальнее в главе 4.

Между *MCS* и *MPIO*, NetApp рекомендует выбирать *MPIO* по следующим причинам:

- MCS не поддерживается для *clustered Data ONTAP*. Так как MPIO может использоваться как с *clustered Data ONTAP* , так и с *Data ONTAP 7-Mode*, то NetApp рекомендует MPIO как решение организации многопутевого подключения для обоих этих архитектур.
- MCS это только для iSCSI. Так как MPIO может использоваться как с FC, так и с iSCSI, NetApp рекомендует MPIO как решение организации многопутевого подключения для обоих этих протоколов.

- MCS это только для Windows. Так как MPIO есть для большинства поддерживаемых OS хоста, NetApp рекомендует MPIO как решение организации многопутевого подключения для разных хост-OS.

Подробнее о поддерживаемых возможностях смотрите **NetApp Interoperability Matrix Tool**.

4 Link Aggregation

Одна из возможных точек разделения трафика это уровень драйвера сетевой карты, в котором можно использовать метод *TCP/IP link aggregation*. *Link aggregation* это техника, которая позволяет взять несколько отдельных физических линков Ethernet, и организовать их в единый линк. Этот механизм описан в спецификации IEEE 802.3ad. Трафик отправляется в один из линков группы, в соответствии с алгоритмом. Эта технология имеет множество названий, например «*channel bonding*» и «*NIC teaming*». Механизм *link aggregation* не относится только к системам хранения, он может быть использован для любого трафика.

С выходом *Windows Server 2012*, *link aggregation* на хосте с *iSCSI software initiator* теперь поддерживается. Напомним, что в *Windows Server 2008 R2* и более ранних версиях, *Microsoft iSCSI software initiator* НЕ ПОДДЕРЖИВАЛ *link aggregation* на стороне хоста. *Link aggregation* на стороне системы хранения (так называемый «*VIF*» в *Data ONTAP 7G* и «*ifgrp*» в *Data ONTAP 8.0* и новее) поддерживается как Microsoft, так и NetApp для всех версий хост-OS.

Достоинства:

- Прозрачен для всех сетевых протоколов: преимущества *link aggregation* доступны не только для iSCSI, но также и для других типов сетевого трафика, например NFS, SMB (CIFS).
- Хорошо известная и отработанная технология.

Недостатки:

- Не поддерживается на хосте с *Microsoft iSCSI software initiator* в *Windows Server 2008 R2* и ранее.
- Агрегируемые интерфейсы должны подключаться в одну общую сеть, часто в один коммутатор или карту в коммутаторе, что ограничивает возможности физической изоляции множественных путей.
- Зависит от поддержки агрегации на коммутатор и в драйверах.
- Механизм балансировка нагрузки определяется соответствующим алгоритмом на агрегированном интерфейсе, использующий один физический линк на конкретный IP или MAC получателя. Вследствие этого, вы сможете получить более эффективное использование линков с помощью MPIO, чем с помощью *link aggregation*.
- В Windows, включение MPIO активирует дополнительный механизм перепосылки (retry), который недоступен при использовании только лишь *NIC teaming*.

5 Multiple Connections per Session (MCS)

Сессии iSCSI с множественными соединениями в одной сессии, также называемые *MCS*, или *MC/S*, это часть спецификации iSCSI. Они создают множественные пути в рамках одной сессии iSCSI, используя отдельные соединения TCP. На обеих сторонах процесса требуется поддержка MCS, как на стороне *iSCSI initiator* (хост) так и у *iSCSI target* (система хранения). Текущие версии *NetApp Data*

ONTAP 7-Mode и *Microsoft Windows* обе поддерживают MCS. Начиная с *Data ONTAP 8.2 7-Mode*, значение по умолчанию для числа соединений в сессии равно 32. Смотрите **NetApp Interoperability Matrix** для наиболее актуальной информации о поддерживаемых версиях Data ONTAP и инициаторов.

Хотя iSCSI с MCS поддерживается в некоторых средах с использованием *7-Mode*, он пока не поддерживается в *clustered Data ONTAP*.

Рисунок 2 показывает место *iSCSI multiconnection sessions* в стеке хранения.

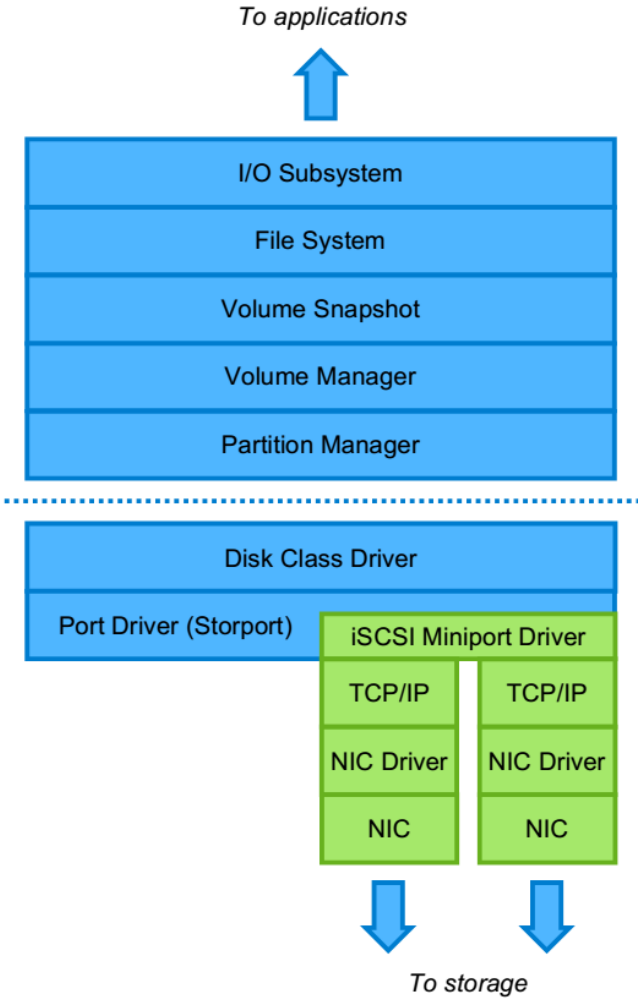

Рис. 2) Стек хранения Microsoft Windows с MCS.

Microsoft iSCSI initiator не поддерживает *multiconnection sessions* по одному или нескольким путям. *SnapDrive® for Windows* будет работать с уже существующими соединениями iSCSI, на которых включен *multiconnection sessions*, но не создает новые соединения с *multiconnection session*, и не сможет узнать о их существовании, если они будут созданы вручную. Смотрите документ **Microsoft iSCSI Initiator User's Guide** для указаний о настройке *multiconnection session* для соединений iSCSI.

iSCSI multiconnection sessions могут использоваться с одним портом таргета или портом-инициатором, или же использовать несколько портов на любом из концов. Если используются несколько таргет-портов, все таргет-интерфейсы для соединения должны быть в одной портал-группе таргета. По умолчанию, каждый интерфейс находится в своей собственной портал-группе

таргета. Подробнее про *iSCSI target portal groups* вы можете прочитать в документе **Block Access Management Guide**.

Хотя технически и возможно совместно и одновременно использовать *iSCSI multiconnection sessions* и *MPIO multipathing* для одного и того же LUN, это не поддерживается Microsoft и NetApp.

Преимущества:

- Является частью спецификации iSCSI.
- Не требуется дополнительных программных средств multipathing.
- Не зависит от средств агрегирования Ethernet в сетевой инфраструктуре.

Недостатки:

- Не поддерживается в *clustered Data ONTAP*.
- Не управляется из *SnapDrive iSCSI connection wizard*.
- Не поддерживается для *Microsoft software initiator boot* (см. **Windows Host Utilities Release Notes**).
  Не поддерживается совместно с MPIO.
- Балансировка осуществляется на сессию; все LUN-ы в сессии iSCSI используют одну и ту же политику балансировки.

6 Multipathed I/O (MPIO)

Классический способ организовать многпутевое подключение это добавить в стек хранения дополнительный уровень, обеспечивающий *multipathing*. Этот метод не определяет какие-то нижележащие транспорты и является стандартным методом для организации многопутевого доступа по протоколам iSCSI, Fibre Channel, и даже к обычным *parallel* и *serial SCSI* таргетам. Существует несколько различных реализаций *multipathing* на различных OS. В Microsoft Windows, каждый производитель систем хранения предоставляет так называемый *device-specific module* (DSM) для своих систем. Кроме этого Microsoft также разработала и предлагает универсальный DSM (в *Windows 2000 Server* и *Windows 2003 Server* только для iSCSI, и универсальный для Fibre Channel и iSCSI в *Windows Server 2008* и новее).

Рисунок 3 показывает, где находится уровень MPIO в стеке хранения.

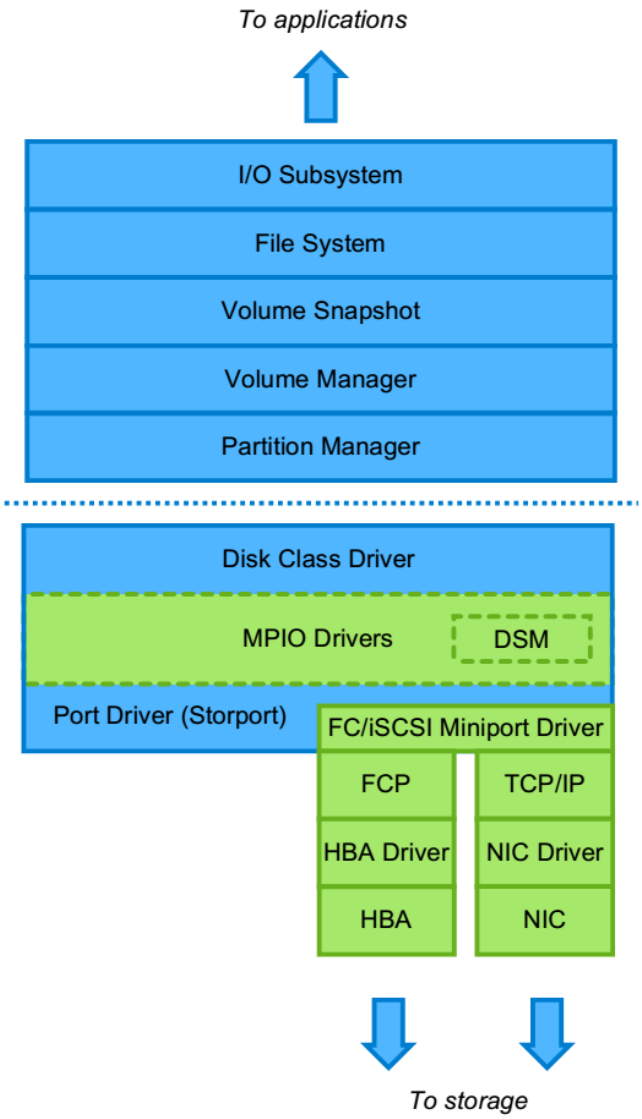

Рис. 3) Стек хранения Microsoft Windows с MPIO.

Так как MPIO находится выше уровня минипорт-драйвера, драйвер MPIO видит только абстрактное SCSI устройство, и не знает об используемом для работы с ним транспортном протоколе. Это позволяет использовать пути Fibre Channel и iSCSI к одному и тому же LUN. Так как протоколы имеют различные характеристики доступа и производительность, NetApp рекомендует, если вы используете их оба, реализовать для этих путей схему *active-passive*, при которой определенные пути указываются резервными, и используются только в случае, когда перестают работать назначенные основными.

Преимущества:

- Не зависит от инфраструктуры Ethernet и его методов агрегирования каналов.
- Очень давно известная и хорошо изученная реализация.
- Можно смешивать пути с разными протоколами (например, iSCSI и Fibre Channel).
- Каждому LUN можно назначить свою политику балансировки.

Недостатки:

- Необходимо использование дополнительного слоя *multipathing*.

6.1 Asymmetrical Logical Unit Access (ALUA)

Не все пути доступа к LUN имеют одинаковые свойства доступа. В *clustered Data ONTAP*, один узел кластера владеет LUN, однако остальные узлы могут также иметь доступ к его данным. В HA-паре контроллеров, работающих в *7-Mode*, один узел владеет LUN, но, в случае использования Fibre Channel, порты обоих узлов имеют доступ к его данным. Путь доступа входит в контроллер, после чего запрос к не-собственному LUN передается по кластерному интерконнекту на узел-владелец, и лишь затем с него достигает LUN. В *clustered Data ONTAP*, этот трафик передается по каналу кластерного интерконнекта с шириной полосы 10 Gigabit Ethernet. В *Data ONTAP 7-Mode*, этот трафик передается по кабелю так называемого *HA cluster interconnect* между парой контроллеров. Путь доступа, идущий к портам контроллера-владельца LUN, называются «*primary*» или «*optimized*». Пути к портам других узлов, по которым также можно получить доступ к данным этого LUN-а, называются в разных руководствах «*unoptimized*», «*partner*», «*proxy*», или «*secondary*»-порты. Рисунок 4 иллюстрирует путь прямого (direct) и непрямого (indirect) пути к ресурсам двухузловолго кластера *clustered Data ONTAP*, а на рисунке 5 показана аналогичная схема для системы *7-Mode*. Выделенные цветом линии показывают пути передачи данных.

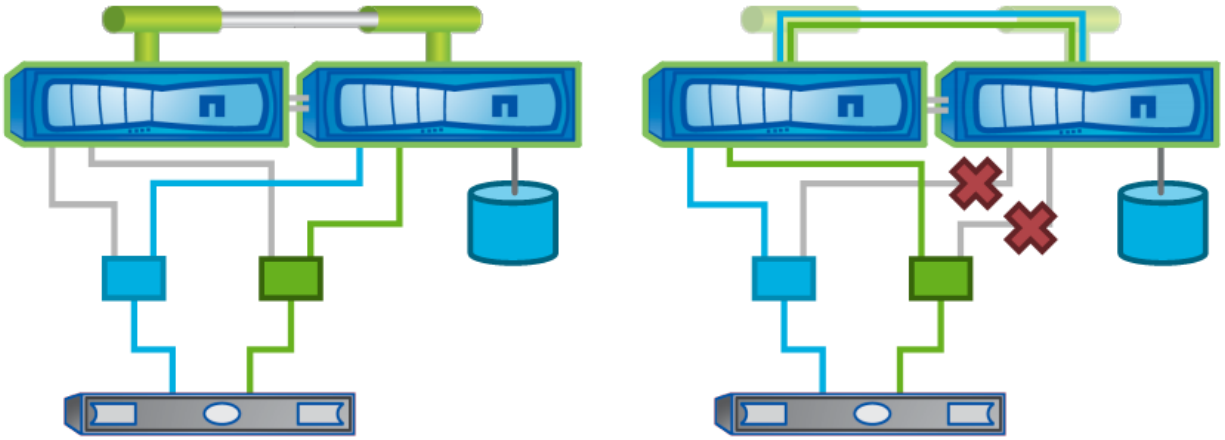

Рис. 4) Отказ пути в *clustered Data ONTAP*.

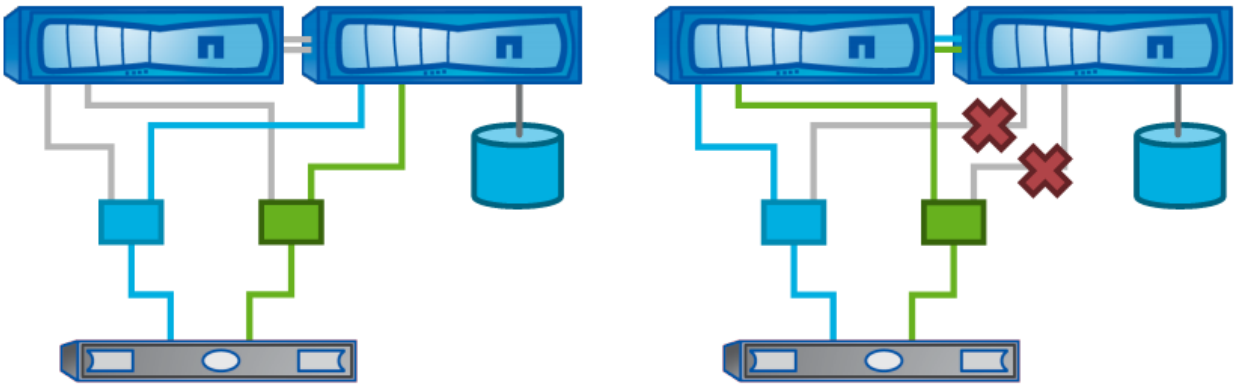

Рис. 5) Отказ пути по протоколу Fibre Channel в *Data ONTAP 7-Mode*.

Чтобы быть уверенным в том, что данные передаются только по primary-пути, хост должен связаться с контроллером системы хранения, чтобы определить, какой из путей – прямой и предпочтительный, а какой – нет. Обычно это делалось ранее с использованием ПО многопутевого подключения, разработанного самим вендором системы хранения. Однако стандартизированный метод, добавленный в спецификацию SCSI, позволил делать это

«вендоронезависимо». Он называется *Asymmetrical Logical Unit Access* (ALUA) и был реализован впервые в Data ONTAP 7.2 и Windows 2008. ALUA позволяет инициатору запросить таргет о атрибутах пути, например о том, какой из путей будет предпочтительным для доступа к этому таргету. В результате, программное решение поддержки многопутевого подключения может поддерживать любую систему хранения, использующую ALUA.

В *clustered Data ONTAP*, ALUA должен быть разрешен на стороне хоста для того, чтобы быть уверенным в прямом подключении к LUN. ALUA используется для всех протоколов SAN, то есть Fibre Channel, FCoE, и iSCSI.

В отличие от *clustered Data ONTAP*, в *Data ONTAP 7-Mode* ALUA не поддерживается для соединений iSCSI. Это так потому, что для iSCSI нет *proxy path* и механизм переключения линка при отказе работает иным способом, чем в случае Fibre Channel.

6.2 Опции DSM

Три основных опции *multipathing*, которые доступны с версии Windows Server 2008 и новее, это: встроенный *Microsoft MPIO* с использованием собственного *Microsoft DSM*, поставляемый NetApp *Data ONTAP DSM*, а также *Symantec™ Veritas™ Dynamic Multipathing* (DMP). Кроме того, *Windows Server 2003* может также работать с *Data ONTAP DSM* или *Microsoft iSCSI DSM*.

В *Windows Server 2008* была представлена новая функция «*нативного*» MPIO, умеющего использовать ALUA для выбора пути. Эта возможность появилась в Windows Server 2008, и поддерживает как Fibre Channel, так и iSCSI. В стандартный набор политик балансировки входят такие политики, как: *failover only*, *round robin*, *round robin with subset*, *least queue depth*, и *weighted paths*. В *Windows Server 2008 R2* добавилась также политика *least blocks*. Политикой по умолчанию для соединений *Fibre Channel* является *round robin with subset*, а для iSCSI – *failover*.

Если для подключения к LUN по Fibre Channel используется *Microsoft MPIO* , то параметр ALUA в igroup данных инициаторов должен быть включен Команда для включения ALUA из командной строки: `igroup set <igroup_name> alua yes`. Это включение должно быть проделано до того, как LUN будет увиден хостом.

NetApp Data ONTAP DSM предлагает стандартные политики балансировки нагрузки и простой в использовании интерфейс управления (GUI и CLI). Поскольку это продукт NetApp, он поддерживается в составе системы хранения NetApp, что обеспечивает простую техническую поддержку и поиск решения при возникновении проблем. Хосты под управление *Windows Server 2003*, подключаемые к *clustered Data ONTAP* с помощью Fibre Channel или iSCSI должны использовать *Data ONTAP DSM*, так как *Windows Server 2003* не имеет собственных компонентов *multipathing*.

Symantec разработал собственное решение обеспечения *multipathing*, под названием *Veritas DMP*. Оно предлагает богатый набор возможностей и политик балансировки нагрузки, хорошо подходящий для систем, со строгими требованиями к производительности.

Veritas DMP доступен для VMware®, Linux®, и UNIX®, поэтому использование *Veritas DMP* может помочь построить гомогенный интерфейс управления платформами с различными OS на них.

Внимание: На момент публикации данного документа *Symantec Veritas DMP* не поддерживался с системами, работающими под управлением *clustered Data ONTAP*. Для уточнения вопросов поддержки, смотрите **NetApp Interoperability Matrix Tool**.

6.3 Политики балансировки загрузки (Load Balance Policies)

Когда к LUN есть несколько путей доступа, необходимо определить метод использования всех доступных путей. Этот метод принято называть *load balance policy*. Существует шесть стандартных политик в Windows Server 2012, применяемых к MCS и MPIO:

- **Failover only**: Только один из путей активен в данный момент времени, альтернативный путь зарезервирован на случай отказа основного пути.
- **Round robin**: Операции ввода-вывода посылаются в каждый из доступных путей по очереди.
- **Round robin with subset**: Часть доступных путей используется в режиме *round robin*, остальные становятся активными только в случае отказа первых.
- **Least queue depth**: Операции ввода-вывода посылаются по пути, имеющему наименьшую на данный момент загрузку очереди ввода-вывода.
- **Least blocks**: Операции ввода-вывода посылаются по пути, с наименьшим числом одновременно отправленных в них блоков.
- **Weighted paths**: Каждому пути назначается приоритет, или «вес», для ввода-вывода используется путь с низшим доступным весом, то есть наивысшим приоритетом.

При использовании Data ONTAP DSM добавляется также политика *Auto Assigned*:

- **Auto assigned**: Политика *Auto Assigned* это политика типа «*active-passive*». Для каждого LUN, может быть использован только один путь в данный момент времени. Если активный путь переходит в пассивный, то политика выбирает следующий активный путь. Политика *Auto Assigned* не передает загрузку по всем доступным путям одновременно.

При использовании *clustered Data ONTAP*, *Data ONTAP DSM 3.5* и новее добавляет возможности приоритезировать пути FC выше путей iSCSI, используя для этого параметр реестра `iSCSILeastPreferred`. Параметр `iSCSILeastPreferred` указывает, что *Data ONTAP DSM* будет предпочитать пути FC путям iSCSI для того же LUN. Вы можете включить этот параметр в случае, если планируете использовать пути iSCSI как резерв для путей FC. По умолчанию, *Data ONTAP DSM* использует *ALUA access states* для приоритезации путей. Это не приоритезируется по протоколам, все блочные протоколы при этом равны. Если вы включите эту настройку, то DSM будет учитывать и *ALUA state,* и тип протокола, при этом пути FC будут предпочтительнее путей iSCSI. *DSM* будет использовать оптимальные по ALUA пути к LUN через iSCSI только в случае, если не будет доступно оптимальных по ALUA путей FC.

Эта настройка применима к LUN, использующим политику балансировки вида *least queue depth* или *round robin with subset*.

Внимание: Соединения iSCSI в 7-Mode не поддерживают ALUA, и, следовательно, не могут быть использованы в igroup со смешанными протоколами, то есть iSCSI и FC одновременно. По этой причине настройка `iSCSILeastPreferred` не применима в системах с использованием 7-Mode.

Подробнее смотрите **Data ONTAP DSM for Windows MPIO Installation and Administration Guide**.

7 Рекомендации по построению сети iSCSI

7.1 Сетевые топологии iSCSI

Протокол iSCSI был определен и описан рабочей группой *Internet Engineering Task Force* в опубликованном документе, под номером *RFC 3270*. Копия текущего стандарта доступна по ссылке http://www.ietf.org/rfc/rfc3720.txt.

Первое решение, которое должен принять пользователь при построении сети хранения с использованием iSCSI, это запуск ее трафика по физически выделенной сети. Выделенная инфраструктура iSCSI Ethernet может включать в себя выделенные коммутаторы или VLAN-ы. Для конфигураций меньшего размера, хосты могут подключаться непосредственно к отдельным узлам системы хранения, используя кабели прямого подключения (crossover).

Внимание: NetApp рекомендует, чтобы, если вы используете множественные пути или сессии iSCSI, каждый путь был изолирован в свою собственную подсеть.

На системе хранения, на порту, который не используется для обслуживания сессий iSCSI, этот сервис должен быть в явном виде запрещен. После того, как сервис iSCSI на данном порту будет запрещен, будут отклоняться все попытки установить сессии iSCSI через этот интерфейс. Такое решение увеличивает безопасность, позволяя работу iSCSI только по заранее выбранным портам.

Начиная с версии Data ONTAP 7.3, появился так называемый *iSCSI access lists* , что дает более высокую гранулярность управления и доступа. Доступ можно предоставлять конкретным инициаторам, входящим через конкретные интерфейсные порты системы хранения.

NetApp рекомендует использовать сетевую топологию, которая минимизирует риск неавторизованного доступа или модификации данных, передаваемых по сети. Вы можете свести к минимуму сторонний доступ к данным с помощью использования выделенных для передачи данных физических кабелей, коммутируемой среди, виртуальных сетей (VLAN), и выделения под работу протоколов сети хранения определенных портов на системе хранения.

Для построения топологии сети хранения с использованием протокола iSCSI, может быть использовано три схемы. Каждая из них имеет свои достоинства и недостатки.

7.2 Совместно используемая коммутируемая сеть iSCSI

Совместно используемая коммутируемая сеть предполагает передачу как трафика iSCSI, так и другого трафика Ethernet по той же физической сети. Так как эта сеть совместно используема разными трафиками разных хостов, этот вариант менее безопасен, чем выделенная сеть; вам следует продумать использование дополнительных средств обеспечения безопасности для снижения рисков.

NetApp рекомендует использовать VLAN-ы для разделения трафика iSCSI от трафиков любых других сетей в совместно используемой сетевой среде. Использование VLAN предоставляет возможности дополнительной защиты безопасности и упрощает поиск причин сетевых проблем, когда они возникают. Система хранения NetApp может быть сконфигурирована на использование VLAN, организованных с помощью тегов VLAN, или же с помощью выбора фиксированных портов коммутатора, и прозрачно для контроллеров системы хранения.

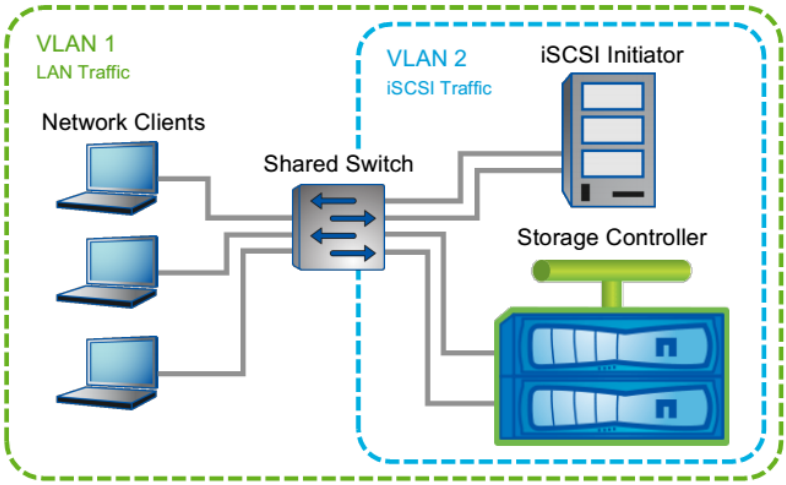

Рис. 6) Топология сети iSCSI в совместно используемой сетевой среде.

Преимущества:

- Возможно использование Link aggregation, если это поддерживает коммутатор.
- Для отказоустойчивости можно использовать несколько коммутаторов.
- Число хостов и систем хранения ограничены только доступным числом портов на коммутаторе.
- Можно использовать имеющуюся инфраструктуру коммутации Ethernet, экономя деньги.
- Каждому LUN можно назначить свою политику балансировки нагрузки.

Недостатки:

- Полоса сети разделяется между всеми участниками сети, то есть публичной LAN и сетью iSCSI, так как LAN, а также инициаторы и таргеты iSCSI-сети подключены в один и тот же коммутатор.
- Коммутатор должен поддерживать VLAN-ы.

7.3 Выделенная коммутируемая сеть iSCSI

В этой конфигурации коммутаторы Ethernet и кабели выделены исключительно для транспортировки трафика iSCSI между хостами iSCSI и системами хранения. Эта конфигурация очень похожа на *фабрику* Fibre Channel, в которой в ее сети ходит исключительно один протокол сети хранения данных. Такая схема требует дополнительных затрат на выделенное оборудование коммутации Ethernet, но зато имеет преимущества с точки зрения безопасности и производительности.

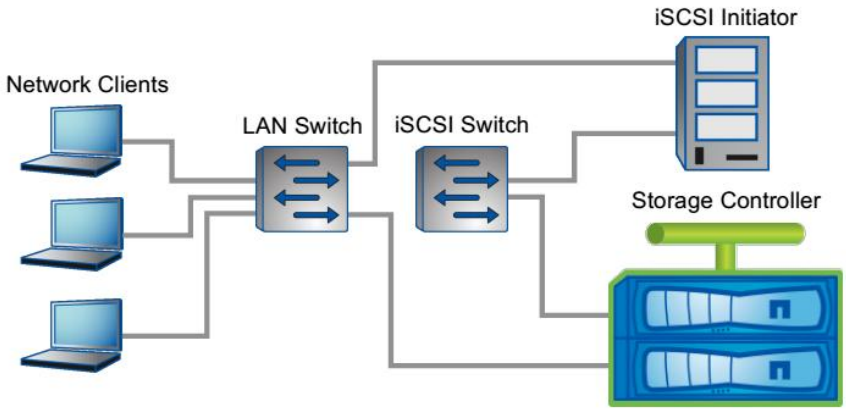

Рис. 7) Топология сети iSCSI в выделенной сетевой среде.

Преимущества:

- Максимально безопасно: трафик iSCSI физически изолирован от трафика сети public LAN.
- Для использования доступна вся полоса пропускания.
- Возможно использование Link aggregation, если это поддерживает коммутатор.
- Для отказоустойчивости можно использовать несколько коммутаторов.
- Число хостов и систем хранения ограничены только доступным числом портов на коммутаторе.
- Вы можете использовать недорогой неуправляемый коммутатор, так как поддержка VLAN не требуется.

Недостатки:

- Один и более коммутаторов должны быть выделены под использование с трафиком iSCSI.
- Конфигурирование и администрирование более сложное, чем при прямом подключении.

7.4 Сеть iSCSI с прямым подключением

Хосты включаются непосредственно в порты системы хранения, без коммутатора.

Это наиболее безопасная с точки зрения физической безопасности схема, она также позволяет использовать всю доступную полосу интерфейса между инициатором и таргетом.

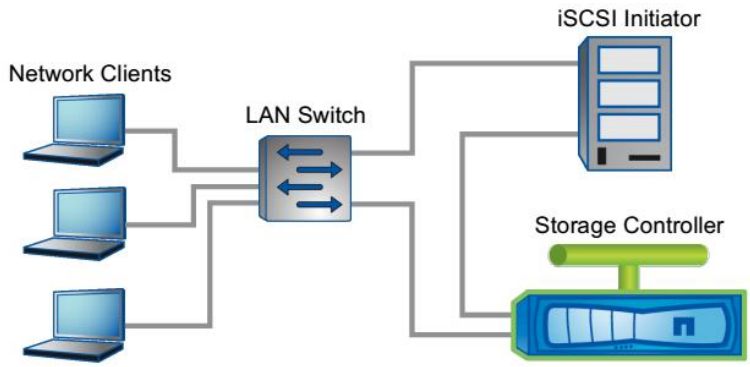

Рис. 8) Прямое включение iSCSI.

Преимущества:

- Низкая стоимость: не требуется коммутатор Ethernet.
- Очень безопасно: нет возможности включения в действующий канал передачи.
- Легко установить и настроить.
- Доступна вся полоса пропускания интерфейса.

Недостатки:

- Число инициаторов и/или путей доступа ограничено числом сетевых портов контроллера.
- Расстояние между инициатором и таргетом ограничено.
- Не поддерживается *HA failover* на системе хранения.

7.5 Использование Jumbo Frames

По умолчанию, Ethernet посылает в одном своем кадре 1500 байт данных. Это хорошо работает для приложений, отсылающих небольшие объемы данных, например для клиентских приложений. Однако, для передачи больших блоков данных, например как это работает обычно в iSCSI, кадры Ethernet большего размера дадут некоторую выгоду и будут более эффективны. Так как размер заголовка кадра фиксирован, увеличение размера кадра увеличивает его полезную нагрузку за счет снижения сетевого оверхеда, и позволяет передать больше данных на один запрос. Это увеличение повышает общую пропускную способность, снижает накладные расходы на передачу, а также нагрузку в процессе передачи на процессор системы и схемы NIC, занятых обработкой служебной информации кадра.

Термин «*jumbo frame*» обычно означает кадр Ethernet с 9000 байт данных в нем, однако, с технической точки зрения, он применим к любому кадру с размером более 1500 байт данных. В отличие от стандартного размера кадра, размер *jumbo frame* не стандартизирован. Каждое сетевое устройство должно быть сконфигурировано на определенный максимум передаваемого кадра. Следовательно, каждое сетевое устройство в системе должно поддерживать одинаковый для системы в целом размер для *jumbo frames*.

Системы хранения NetApp поддерживают *jumbo frames* размером 9000 байт данных на всех интерфейсах 1 и 10 Gigabit Ethernet.

7.6 Использование Flow Control

Для соединений *1 Gigabit Ethernet* NetApp рекомендует выставлять значение *flow control* на сетевом коммутаторе в значение «*full*», а на всех *iSCSI targets* и *initiators* в значение «*send*». Для соединений *10 Gigabit Ethernet*, NetApp советует отключить (disable) *flow control* для всех *iSCSI targets*, *initiators*, и портов коммутатора. Для подробностей об использовании *flow control*, смотрите **Data ONTAP Network Management Guide**.

7.7 NetApp Host Utilities

NetApp предлагает специальный набор утилит под названием *SAN Host Utilities kit* для каждой поддерживаемой OS. Этот набор содержит набор приложений по сбору диагностических данных и конфигурационные скрипты, которые устанавливают правильные значения таймаутов SCSI и путей подключения, для работы с системами хранения NetApp.

Набор утилит также включает инструменты для улучшения качества поддержки хоста в среде NetApp SAN, например путем сбора конфигураций и логов, а также просмотра деталей всех подключенных с NetApp LUN-ов.

Внимание: При использовании *Data ONTAP DSM 3.5* и новее, использовать *Host Utilities kit* для установки таймаутов не требуется. *Data ONTAP DSM* делает эти настройки в хост-OS самостоятельно.

8 Рекомендации по построению Fibre Channel Fabric

В рамках обсуждаемой темы *multipathing* мы говорим только о MPIO в применении к Fibre Channel. Три основных топологии, используемых для среды Fibre Channel, это:

- **Прямое подключение**: Инициаторы и таргеты в 7-Mode соединяются непосредственно кабелем.

Обратите внимание: по причине использования NPIV и архитектуры кластерной Data ONTAP, прямое подключение соединений Fibre Channel в clustered Data ONTAP не поддерживается.

- **Single fabric**: Все порты инициаторов и таргетов подключаются к одному коммутатору, или фабрике коммутаторов.
- **Multifabric**: Некоторые порты инициаторов и таргетов подключаются к раздельных фабрикам, для обеспечения избыточности.

Эти конфигурации детально рассмотрены в **Data ONTAP SAN Configuration Guide** и включают в себя диаграммы и поддерживаемые топологии для различных платформ NetApp.

Как мы уже упоминали ранее, NetApp рекомендует для любого решения SAN использовать избыточные компоненты, чтобы снизить риски и устранить единую точку отказа. В случае использования Fibre Channel это означает использование нескольких HBA, коммутаторов/фабрик и кластеризация хранилищ.

Справочные документы

При составлении этого TR использовались материалы следующих источников:

• Clustered Data ONTAP SAN Configuration Guide
http://support.netapp.com/documentation/productlibrary/index.html?productID=30092

• Data ONTAP SAN Configuration Guide for 7-Mode
http://support.netapp.com/documentation/productlibrary/index.html?productID=30092

• NetApp Interoperability Matrix Tool
http://support.netapp.com/matrix/mtx/login.do

История изменений