

Содержание

Введение	2	Директоры Brocade X6	14
Технологии		Служебные лезвия, обязательно присутствующие	
Основные понятия и определения	2	в шасси	15
SAN и Fibre Channel	2	Оптические ICL	15
ASIC	3	Лезвия с портами	16
Топологии, масштабирование и резервирование	3	Коммутатор Brocade G620	16
Планирование топологии	3	Оборудование Brocade для IP Extension	17
Топология каскада	3	Brocade 7800	17
Топология кольца	3	Brocade FX8-24	17
Топология полного графа (full mesh)	4	Brocade 7840 Extension switch	17
Двухуровневая топология «ядро-периферия»	4	Варианты использования	18
Резервированные фабрики	4	Результаты тестирования	19
Зонирование	5	Brocade SX6 Extension blade	19
Типы зонирования	5	Встроенные коммутаторы для шасси	
Разработка плана зонирования	6	блейд-серверов	19
Зонирование с одним инициатором		Brocade Network Advisor	19
(single initiator zoning)	6	Brocade Fabric Vision	20
Peer Zoning — новый подход к зонированию	6	Архитектура Fabric Vision	20
Target Driven Zoning (TDZ)	7	Monitoring and Alerting Policy Suite (MAPS)	20
Виртуальные фабрики	7	Flow Vision	21
Шлюз доступа	8	Brocade Network Advisor Dashboards	21
Масштабирование фабрик SAN до уровня Meta SAN	8	Slow Drain Device Quarantine (SDDQ)	21
Связь центров обработки данных на расстоянии	9	COMPASS	22
Темная оптика (FC over Dark Fiber)	10	Диагностический порт ClearLink	22
Оптическое мультиплексирование (FC over xWDM)	10	Forward Error Correction (FEC)	23
Буферные кредиты	10	IO Insight	23
Режимы передачи на большие расстояния	11	VM Insight	23
FC over IP	11	Поддержка возможностей Fabric Vision в	
Будущее — технологии NVMe и NVMe over Fabrics	12	различных поколениях коммутаторов Brocade	23
Продукты	12	Расширение возможностей Fabric Vision на все	
Коммутаторы Brocade Fibre Channel		устройства в сети	23
четвертого поколения	12	Analytics Monitoring Platform (AMP)	24
Brocade 300	12	Лицензии коммутаторов	25
Коммутаторы Brocade Fibre Channel		Лицензии на аппаратную часть	25
пятого поколения	12	Лицензии для повышения производительности	25
Директоры DCX 8510 Backbone	13	Лицензии на расширенные функции	25
Служебные лезвия, обязательно присутствующие		Лицензии на мониторинг и управление	25
в шасси	13	Enterprise Bundle	25
Лезвия с портами	13	Особенности лицензирования коммутаторов	
Специальные лезвия	14	FC шестого поколения	25
Коммутаторы	14	Дополнительная информация	26
Коммутатор Brocade 6520	14	Обучение и сертифицированные курсы Brocade	26
Коммутатор Brocade 6510	14	Программа авторизованного обучения и	
Коммутатор Brocade 6505	14	сертификации Brocade	26
Функционал коммутаторов пятого поколения	14	Вопросы и ответы	27
10 Гбит/с Fibre Channel	14	Источники информации	28
Сжатие и шифрование трафика ISL	14	Заключение	28
Коммутаторы Brocade Fibre Channel			
шестого поколения	14		

Введение

Данный документ описывает технологии, продукты и решения Brocade в области Fibre Channel SAN по состоянию на август 2016 года. Изложение ведется на языке, понятном специалистам в области систем хранения данных. Брошюра может использоваться в качестве пособия для первоначального ознакомления, а также в качестве повседневного справочного материала.

Новости Brocade в России и СНГ выкладываются в группе LinkedIn "Brocade Russia SAN", www.linkedin.com/grp/home?gid=4246353 – там идет живая дискуссия и обмен мнениями между более чем 750 экспертами SAN нашего региона – присоединяйтесь! Эта и другие русскоязычные брошюры Brocade выложены в формате pdf на сайте www.brocade.ru в разделе «Материалы по SAN».

Технологии

Основные понятия и определения

Сеть хранения данных (Storage area network, SAN) – специализированная сеть, предназначенная для соединения серверов и систем хранения. Служит для передачи данных, которые должны быть записаны на СХД или прочитаны с СХД. В отличие от локальных вычислительных сетей (LAN) не предназначена для передачи данных пользователей между серверами и рабочими станциями.

Коммутатор Fibre Channel (FC) – устройство, работающее в сети хранения данных (SAN) и предназначенное для коммутации пакетов, передаваемых по протоколу Fibre Channel. FC-коммутаторы Brocade работают под управлением операционной системы Brocade Fabric OS. Коммутаторы могут быть монолитные (с фиксированным числом портов), модульные (директоры), встроенные и специализированные. В монолитных коммутаторах количество портов фиксировано, и все они физически присутствуют в коммутаторе. При этом часть портов может быть заблокирована, и для их активации требуется отдельная лицензия. Модульные коммутаторы (директоры) представляют собой шасси, в которое можно устанавливать аппаратные модули расширения (лезвия) с портами и/или специализированные модули – набирая, тем самым, функционал коммутатора. Встроенные коммутаторы используются в шасси блейд-серверов и некоторых систем хранения. Функционал коммутаторов может активироваться программно – посредством установок лицензий на заданный функционал. Специализированные коммутаторы могут выполнять специализированные функции, например коммутатор AMP позволяет отслеживать в сети SAN задержки приложений.

Коммутатор расширения (extension switch) – это специализированный коммутатор Fibre Channel, обладающий дополнительным функционалом. Помимо коммутации пакетов FC, позволяет связывать разрозненные сегменты сети Fibre Channel SAN, используя стандартные iP-сети (LAN, WAN). Для этого используется протокол Fibre Channel over IP (FCIP), который позволяет передавать пакеты FC через защищенный IP туннель. Коммутатор расширения может также быть представлен в виде модуля для директора (Extension blade).

Оптический модуль (transceiver) – модуль, устанавливаемый в порты коммутатора. Предназначен для преобразования электрического сигнала в оптический и наоборот. Состоит из источника света – оптического диода (LED) либо лазера и фотодетектора. Коммутатор осуществляет коммутацию сигналов в виде электрических импульсов. По сети сигналы передаются в виде оптических импульсов. Существуют различные форм-факторы оптических модулей. Наиболее часто используемые: SFP, SFP+, QSFP. Оптические модули FC в современных коммутаторах Brocade могут работать со скоростью 32, 16, 8, 4, 2 Гбит/с. Традиционно, оптические SFP модули в коммутаторах Brocade автоматически поддерживают три частоты – основную и на две ниже. Например, модуль SFP+ 16Гбит/с поддерживает скорости 16, 8 и 4 Гбит/с.

Фабрика – один или несколько коммутаторов Fibre Channel, соединенных между собой. Вне зависимости от топологии соединения фабрика работает как единая логическая сущность. Фабрика имеет единую точку управления. Все коммутаторы, входящие в фабрику, поддерживают единый набор сервисов управления. Как правило, сеть SAN строится из двух независимых фабрик (dual fabric), когда сервера и системы хранения подключаются друг к другу по двум физически независимым путям.

Meta-SAN – структура из нескольких независимых сетей SAN, между которыми настроена маршрутизация. Благодаря маршрутизации устройства из разных сетей SAN могут взаимодействовать между собой, в то время как сети SAN остаются полностью независимыми, как с точки зрения параметров сетей, так и с точки зрения управления сетями. Meta-SAN применяется для локализации сегментов больших сетей SAN. Это позволяет локализовать возможные ошибки в сетях и облегчить управление сетями.

Основные типы портов Fibre Channel

В процессе инициализации порты Fibre Channel могут принимать различные логические роли (типы). В зависимости от типа порта его функционал меняется. В зависимости от текущего положения в топологии сети один и тот же физический порт может получать при инициализации различные логические типы. Основные типы портов Fibre Channel представлены ниже:

- U_Port (Universal port) – порт на коммутаторе в промежуточном состоянии инициализации. После инициализации станет портом другого возможного типа;
- G_Port (Generic port) – порт на коммутаторе в промежуточном состоянии инициализации. Ожидает переход в состояние F_Port или E_Port;
- F_Port (Fabric port) – порт на коммутаторе, к которому подключаются оконечные устройства (серверы или СХД);
- N_Port (Node port) – порт на сервере или СХД. Подключается к F_Port на коммутаторе;
- E_Port (Expansion port) – порт на коммутаторе, с помощью которого коммутатор подключается к другому коммутатору. Два E_Port на разных коммутаторах образуют Inter Switch Link (ISL);
- EX_Port – маршрутизирующий порт. Специальный тип E_Port, применяемый для созда-

ния Meta-SAN. С помощью EX_Port можно соединить между собой две независимые фабрики. Пара портов EX_Port и E_Port образуют Inter Fabric Link (IFL);

- D_Port (Diagnostic Port) – порт, предназначенный для передачи диагностической информации для проведения тестирования оборудования и канала связи, используя технологию Brocade ClearLink Diagnostic;
- VE_Port, VEX_Port – виртуальные порты E_Port и EX_Port. Применяются в технологии IP Extension при создании туннелей по протоколу FCIP;
- AE_Port, AF_Port – специальные типы портов, используемые при подключении коммутаторов к устройству Analytics Monitoring Platform (AMP);
- M_Port (Mirror port) – порт на коммутаторе, применяющийся для зеркалирования трафика;
- FL_Port, NL_Port – специализированные типы F_Port и N_Port, применявшиеся ранее для подключения устройств по протоколу Fibre Channel Arbitrated Loop (FCAL). В пятом и шестом поколениях коммутаторов Brocade Fibre Channel эти типы портов более не поддерживаются.

Традиционные этапы развития SAN и Meta-SAN организации

1. Построение SAN в одном центре обработки данных (ЦОД1). SAN начинает расти от двух коммутаторов и может вырасти до тысяч портов FC.
2. Построение ЦОД2 для синхронной репликации данных между ЦОД1 и ЦОД2. ЦОД1 и ЦОД2 расположены на расстоянии не более нескольких десятков километров друг от друга, SAN связаны между собой в режиме маршрутизации через FC-каналы. Режим работы ЦОДов: основной-основной или основной-резервный.
3. Построение ЦОД3 на удалении в несколько сотен километров или более от ЦОД1 и ЦОД2 для асинхронной репликации данных в ЦОД3 по технологии FCIP – ЦОД3 подключен по каналам IP и используется как резервный.

SAN и Fibre Channel

Storage Area Network (SAN) – Сеть Хранения Данных. Предназначена для подключения систем хранения данных (СХД) к серверам. В основе передачи данных в сетях SAN лежит протокол блочного доступа к данным SCSI. Исторически сети SAN предназначались для обеспечения блочного доступа к данным, хотя современные сети SAN могут работать и с другими протоколами доступа, например, файлового или объектного. Основное отличие сетей SAN от сетей передачи данных (локальных сетей, LAN) лежит именно в области применения. Сети LAN предназначены, в основном, для передачи пользовательских данных между серверами и рабочими станциями пользователей. Основное назначение сетей SAN – передача данных, предназначенных для хранения на СХД (операции чтения и записи). Сети SAN могут использовать те же протоколы передачи данных, что и сети LAN – например TCP/IP. Но исторически основным протоколом для построения сетей SAN является Fibre Channel.

Fibre Channel (FC) – стек специализированных протоколов, основанных на открытых отраслевых стандартах (IEEE T11 committee). Основное предназначение – построение сетей SAN. FC является транс-

портным протоколом и может использоваться для транспортировки данных любых типов (в том числе, IP-пакетов). В настоящее время основной способ его использования — передача команд протокола SCSI в среде открытых систем для построения сетей SAN с блочным доступом к данным. Кроме того, существует вариант протокола FC для подключения мэйнфреймов к CXD — FICON. Сети Fibre Channel и FICON могут сосуществовать вместе, используя единую коммутирующую и кабельную инфраструктуру.

Вне зависимости от используемых протоколов передачи данных рекомендуется физически разделять сети SAN и LAN для исключения взаимного влияния передаваемых данных. Кроме того, в силу архитектурных причин и различия в типах передаваемых данных требования к дизайну сетей SAN и LAN значительно отличаются. В некоторых случаях, для сокращения расходов и упрощения сетевой инфраструктуры, имеет смысл строить совместные сети, использующие единую инфраструктуру для передачи пользовательских данных и данных хранения. Часто подобные сети называются конвергентными. Они могут использоваться как стандартные сетевые протоколы для передачи хранимых данных (TCP/IP), так и специализированные, например, Fibre Channel over Ethernet (FCoE).

Fibre Channel является наиболее зрелым и востребованным протоколом для построения сетей хранения данных, в силу того, что в его основе лежат одни из самых надежных и высокопроизводительных технологий передачи информации, что особенно важно при работе с бизнес-критичными приложениями. Использование технологий Fibre Channel позволяет утилизировать на 100 процентов доступную полосу пропускания каналов связи, обеспечить минимальные задержки при передаче данных и гарантированно выполнять требуемые SLA (соглашения об уровне обслуживания).

Остальная часть брошюры будет посвящена продуктам и решениям Brocade в области классического протокола Fibre Channel, так же речь пойдет об использовании туннельного протокола Fibre Channel over IP, предназначенного для соединения сегментов сети Fibre Channel через туннель, построенный на основе протокола TCP/IP.

ASIC

В основе большинства продуктов Brocade лежат интегральные схемы ASIC (application specific integrated circuit) собственной разработки.

Табл. 1. Поколения ASIC Brocade.

Название	Год выпуска	Скорость	Продукты
Stitch	1997	1 Гбит/с	Silkworm 1000
LOOM	1999	1 Гбит/с	Silkworm 6400 Silkworm 2x00
BLOOM/BL OOM-2	2001	1/2 Гбит/с	Silkworm 12K, 24K, Silkworm 3x00, 3x50
Condor/Gol denEye	2004	1/2/4 Гбит/с	48K, 4100, 4900, 200, 7500, 4G Embedded
Egret	2008	10 Гбит/с	FC10-6 Blade
Condor2/Go ldenEye2	2008	1/2/4/8 Гбит/с	DCX, 5100, 5300, 300, 7800, 8G Embedded
Condor3	2011	2/4/8/10/16 Гбит/с	DCX 8510, 6510, 6505, 6520, 7840, 16G Embedded
Condor4	2016	4/8/10/16/32 Гбит/с	X6-8, X6-4, SX6 extension blade, G620

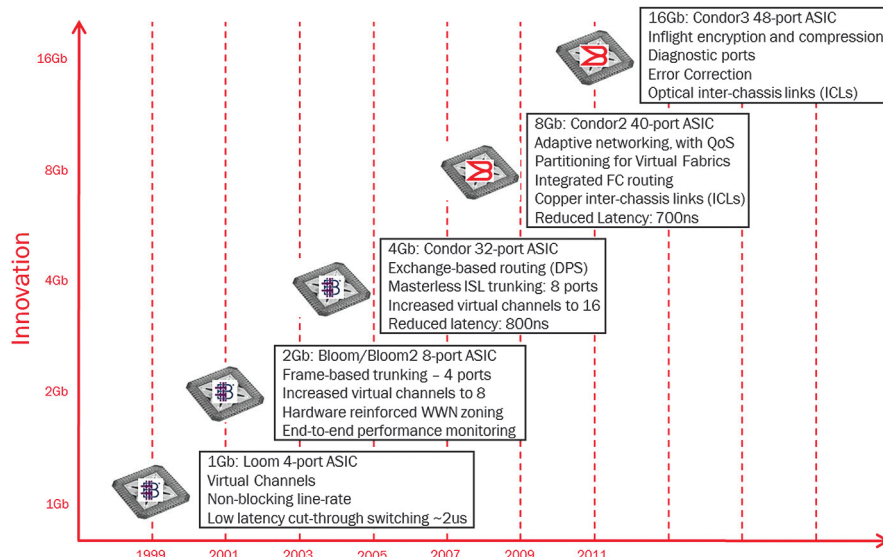


Рис. 1. Поколения ASIC Brocade.

Табл. 2. Сравнение ASIC Condor 3 и Condor 4.

Feature	Condor 3	Condor 4
Clock Speed	425 MHz	566 MHz
Data Path	64 bits	96 bits
Number of Ports	48	64
Speed	16/10/8/4 Gbps	32/16/10/8/4 Gbps
Data Switching	652 Gbps	1.74 Tbps
FC Encryption	32 Gbps encryption	64 Gbps encryption
FC Compression	64 Gbps compression	64 Gbps compression
L2 Frame Latency	700 ns	700 ns
Total Number of Buffers	8192	15360
IO Insight Latency Monitoring	No	Yes

1) Не поддерживается в FOS 8.0.1. Поддержка будет добавлена в последующих версиях Fabric OS.
2) Мониторинг скорости времени выполнения команд ввода/вывода SCSI для определения проблемных точек в сети SAN.

Использование таких чипов позволяет на аппаратном уровне реализовать большую часть функционала, что в итоге приводит к более высокой производительности, большей надежности и меньшему энергопотреблению (табл. 1, рис. 1). Condor4 — наиболее современный и функциональный ASIC для Fibre Channel на сегодняшний день. Он является основой для линейки коммутаторов Brocade шестого поколения.

Основные характеристики Condor4:

- скорость: 4/8/10/16/32/128 Гбит/с Fibre Channel;
- количество портов на одну микросхему: 64;
- типы портов: D_Port (ClearLink Diagnostic Port), E_Port, EX_Port, F_Port, AE_Port, NPIV-enabled N_Port;
- пропускная способность: 2 Тбит/с;
- сжатие и шифрование трафика на лету;
- буферные кредиты: 15360.

Сравнение ASIC Condor 3 и Condor 4 — табл. 2.

Топологии, масштабирование и резервирование

Планирование топологии

У любой проектируемой сети есть своя топология, то есть схема расположения и соединения коммутаторов, маршрутизаторов и другого се-

тевого оборудования. При составлении схемы сети эти элементы образуют различные геометрические фигуры, которые обычно легко идентифицируются и во многом определяют свойства сети. Такие параметры фабрики SAN, как производительность, доступность и масштабируемость могут зависеть, прежде всего, от топологии.

Чтобы не усложнять сеть, большинство архитекторов используют только несколько топологий. Обычно это топология, ориентированная на системы хранения, каскад, кольцо, полный граф либо «центр-периферия» (Core/Edge, CE). Практически любая задача построения SAN может быть решена с помощью одной из этих топологий или их комбинации.

Топология каскада

Каскадная сеть состоит из группы коммутаторов, соединенных последовательно с использованием одного или нескольких соединений (ISL) между соседними коммутаторами.

Такая архитектура не обладает масштабируемостью, не обеспечивает высокой производительности и не отличается высокой надежностью. Для масштабирования SAN необходимо добавить коммутаторы на концах цепочки, что увеличивает число переходов между концами сети и создает риск запаздывания пакетов и перегруженности сети. Другими словами, производительность сети падает по мере добавления в нее новых устройств и коммутаторов. Это полностью противоречит целям построения сети SAN, призванной улучшить производительность (рис. 2).

На практике каскадные сети применяются редко, поскольку по мере роста их производительность падает. Кроме того, если выйдет из строя коммутатор в середине сети, то сеть распадется на отдельные сегменты. Коммутатор в данной случае является потенциальной точкой отказа сети.

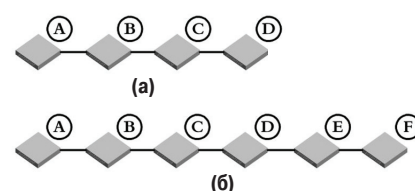


Рис. 2. Топология каскада из четырех (а) и из шести (б) коммутаторов.

Из-за этих недостатков архитекторам не следует использовать каскадное подключение, если важны производительность, доступность и масштабируемость. Тем не менее, каскады остаются базовой топологией, которую можно просто и с небольшими затратами внедрить, поэтому они подходят для самых маленьких SAN-сетей, прежде всего для сетей только из двух коммутаторов. В таких небольших сетях перечисленные выше проблемы не могут возникнуть, поскольку архитектура с двумя коммутаторами является самой упрощенной версией каскадной топологии. Ее можно также использовать, когда трафик носит, в основном, локальный характер (в пределах отдельных коммутаторов), а ISL используются для управления и/или второстепенных и редко запускаемых приложений.

Топология кольца

Кольцо похоже на каскад, но в нем конечные точки соединены между собой (рис. 3).



Рис. 3. Топология кольца из шести коммутаторов.

Преимущество колец по сравнению с каскадами состоит в наличии альтернативного маршрута, если выйдет из строя коммутатор или ISL в кольце, то данные будут передаваться по кольцу в другом направлении.

Помимо улучшения доступности, кольца увеличивают производительность. В сети Fibre Channel данные всегда передаются между коммутаторами по самому короткому из доступных маршрутов, выбор которого осуществляется протоколом FSPF (Fabric Shortest Path First). Например, устройство, подключенное к коммутатору А, будет передавать данные для коммутатора F, через соединение А-В, а не через маршрут, который связывает коммутаторы В и Е. В каскадной сети дополнительное соединение отсутствует, поэтому нельзя выбрать кратчайший маршрут. Хотя во многих отношениях кольца работают лучше каскадов, у обеих топологий есть общие проблемы. Например, при добавлении в кольцо коммутатора увеличивается число последовательных связей между коммутаторами («хопов»), что ухудшает производительность и надежность SAN.

В определенных случаях кольцо даже хуже каскада. Например, добавление коммутатора приводит к необходимости временно отключить часть кольца. Если потребуются установить седьмой коммутатор в сеть, показанную на рисунке, и сохранить топологию кольца, то придется отключить один ISL, что неизбежно повлияет на все операции ввода/вывода, которые идут через это соединение. Например, если коммутатор устанавливается на дальней стороне F, то нужно временно отсоединить ISL между А и F, что приведет к обрыву идущих через него потоков данных.

Как и каскадные сети, кольца подходят для небольших проектов, где трафик в основном носит локальный характер. Если же в SAN больше четырех коммутаторов, то следует использовать другую топологию. Кольца также можно использовать для некоторых распределенных решений, в которых топология существующей сети MAN/WAN определяет топологию внедряемой на ее основе SAN.

Топология полного графа (full mesh)

Топология полного графа предусматривает прямое соединение каждого коммутатора со всеми остальными коммутаторами в сети. Полный граф решает многие проблемы каскадов и колец — расстояние до любого коммутатора минимально, а добавление коммутаторов в полный граф не увеличивает число «хопов» между коммутаторами. Установка дополнительных коммутаторов не нарушает работу сети, поскольку они устанавливаются не между уже соединенными коммутаторами. Можно использовать альтернативные маршруты, причем, в отличие от каскада, их несколько, поэтому полный граф обладает высокой отказоустойчивостью (рис. 4).

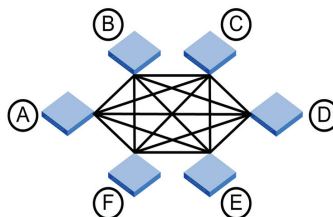


Рис. 4. Топология полного графа (mesh) с шестью коммутаторами.

У топологии полного графа есть несколько существенных недостатков.

Стоимость. Даже в небольшой сети слишком много ISL по отношению к числу полезных портов, и по мере расширения сети число доступных портов коммутаторов сокращается.

Из-за ограничений, указанных в предыдущем пункте, полный граф нельзя масштабировать — при добавлении коммутатора расходуется по одному порту на всех остальных и возникает проблема экспоненциального роста числа ISL. Для соединения в полный граф двух коммутаторов требуется в сумме два порта, для трех коммутаторов — шесть портов, для четырех коммутаторов — двенадцать и так далее. В результате топология полного графа быстро становится крайне невыгодной по стоимости и ее расширение уже не имеет смысла после того, когда половина портов коммутаторов выделена для ISL. Например, из 16-портовых коммутаторов можно создать сеть, состоящую максимум из восьми доменов. Если в такую сеть добавить еще один коммутатор, то число свободных портов коммутаторов сети уменьшится.

Из-за этих ограничений имеет смысл использовать полный граф только в SAN с четырьмя доменами. Хотя, если каждый домен — это 384-портовый директор, то получится достаточно большая сеть. Полный граф также часто используется при построении небольших MAN/WAN, в которых каждая площадка является вершиной графа. Часто используется гибридная топология: Core-Edge внутри каждой площадки и полный граф либо кольцо для соединения площадок. Такой подход применим, когда имеется небольшое количество площадок; большие сети MAN/WAN обычно используют неполный граф или варианты Core-Edge.

Важное развитие данной топологии стало возможно с выходом линейки DCX 8510. Оптические порты UltraScale ICL высокой производи-

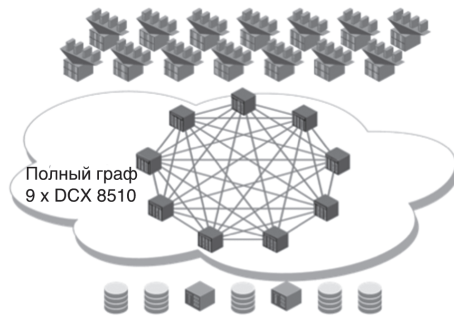


Рис. 5. Полный граф из 9 DCX 8510.

тельности позволяют соединять между собой директоры DCX 8510 без использования «полезной» портовой емкости. Это сразу снимает ограничения на масштабирование полного графа для сетей, построенных только на таких директорах. 32 порта ICL на DCX 8510-8 дают возможность построить полный граф из 9 шасси DCX. В результате получается максимально надежная и производительная сеть SAN размером до 4600 портов с пропускной способностью 256 Гбит/с между любой парой DCX.

Это решение — настоящая революция в построении SAN для самых крупных и требовательных заказчиков. При использовании ICL высвобождается большое количество портов, в 4-8 раз уменьшается количество кабелей между коммутаторами, и несмотря на огромное количество портов, количество коммутаторов в SAN остается небольшим, что значительно упрощает управление, регламентные работы и поиск неисправностей (рис. 5).

Двухуровневая топология «ядро-периферия»

Топология «ядро-периферия» (Core-Edge, CE) — это эволюция популярной в мире сетей передачи топологии «звезда». На рисунке показана отказоустойчивая сеть CE Fibre Channel. Коммутаторы в нижней части рисунка формируют «ядро» сети, а соединенные через них коммутаторы относятся к периферийной группе (рис. 6).

Топология Core-Edge стала основной для SAN по нескольким причинам:

- **хорошая протестированность**, поскольку большинство применяемых фабрик, а также тестовые лаборатории Brocade и другие лаборатории, занимающиеся тестированием SAN, используют топологию Core-Edge в той или иной форме;
- **хорошая сбалансированность** — ее симметричность обеспечивает балансировку нагрузки и резервирование, трафик между граничными коммутаторами можно распределять между всеми центральными коммутаторами;

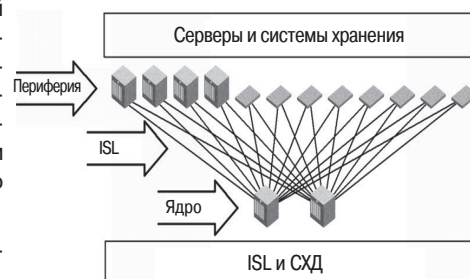


Рис. 6. Топология ядро-периферия.

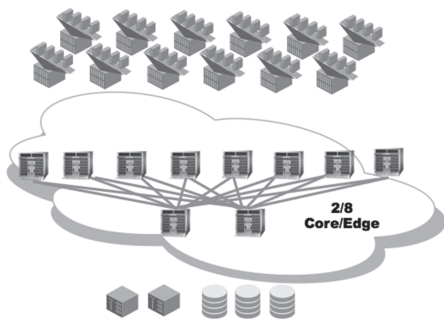


Рис. 7. Топология ядро-периферия из 10 DCX 8510.

- **детерминированность:** скорость передачи данных между двумя периферийными коммутаторами никак не влияет на скорость между любыми двумя другими коммутаторами;
- **экономичность:** имеются варианты для разных соотношений цены и производительности, если требуется обеспечить более высокую пропускную способность для конкретного коммутатора периферийной группы, это можно сделать, увеличив число соединений этого коммутатора с ядром сети, без дополнительных расходов на расширение SAN;
- **легкая адаптация и модификация сети,** поскольку каждый центральный коммутатор можно дублировать, а периферийные коммутаторы взаимозаменяемы;
- **топология Core-Edge имеет простую структуру, проста для понимания, документирования и устранения сбоев;**
- **простота масштабирования** без нарушения работы сети при подключении новых периферийных коммутаторов к свободным портам центральных или постепенной заменой центральных коммутаторов на модели с большим числом портов.

Хотя другие топологии используются в небольших SAN, практически все крупномасштабные внедрения основаны на вариантах CE.

С помощью ICL на DCX 8510 также можно строить фабрики очень больших размеров (до 10 DCX) по топологии Core-Edge (рис. 7).

Резервированные фабрики

Резервирование — это использование избыточности для предотвращения сбоя всей системы из-за сбоя одного компонента. В нашем случае «система» — это вся SAN. Резервированные подсети (фабрики) являются компонентами этой системы; они дублируют функции друг друга так, чтобы сбой одной фабрики не привел к отказу всей инфраструктуры SAN.

Отказоустойчивые фабрики пытаются сохранить работоспособность при сбоях в фабрике. Отказоустойчивые фабрики надежны, однако **никакая отдельная фабрика не может рассматриваться как настоящее решение высокой надежности**, поскольку фабрика целиком может выйти из строя из-за аварии, ошибки оператора или сбоя программного обеспечения.

Для защиты от таких ошибок необходим еще один уровень доступности: резервированные фабрики (Dual-Fabric или A/B SAN). Эта архитектура предусматривает использование двух независимых физически изолированных фабрик (как в

отказоустойчивом кластере используются два физически изолированных сервера). Дублирование компонентов и программное обеспечение для переключения при сбоях — это самые эффективные средства обеспечения высокой доступности серверных платформ. Аналогичным образом, архитектура с дублированием фабрик — лучший способ обеспечить высокую доступность SAN (рис. 8).



Рис. 8. Резервирование фабрик.

Без физической изоляции решение не является резервированным — оно только отказоустойчиво и в нем может возникнуть сбой, так же, как и в отказоустойчивой фабрике. Сегментация директора с помощью разделов, виртуальных фабрик (VSAN или VF) или зон не увеличит доступность SAN.

Сбой всего директора может произойти по разным причинам. Например, все VSAN работают с образом операционной системы, и дефект в коде ОС или ошибка оператора при ее обновлении приведет к сбою всего шасси, что одновременно нарушит оба пути (рис. 9).

Это не значит, что схемы разделов, такие, как VSAN, не имеют перспектив. Brocade разработала несколько таких механизмов и поддерживает их применение для решения многих проблем администрирования SAN, в том числе безопасности, совместимости и управляемости. Механизмы сегментации Brocade включают зонирование и виртуальные фабрики (Virtual Fabrics). Например, зонирование внутри фабрики предотвращает взаимодействие серверных адаптеров,

поэтому проникнувший в один сервер злоумышленник не сможет попасть на другой сервер через SAN. Один адаптер не сможет случайно вызвать отказ в работе другого. Также можно независимо управлять дисковыми подсистемами разных серверов. Все эти функции полезны, и применение зонирования приветствуется при проектировании SAN. Такие механизмы сегментирования, как зонирование и Virtual Fabrics, эффективны, но не для обеспечения доступности.

Применение полностью резервированной фабрики позволяет обеспечить непрерывную работу узлов даже при сбое всей фабрики или ее отключения для обслуживания. При описании характеристик доступности на первом месте стоит доступность пути. Если одно соединение оборвется, но система хранения данных будет доступна по другим путям, то не произойдет сбой приложений, использующих SAN. Хотя возможно некоторое ухудшение производительности серверов приложений, но, главное, они продолжат работать.

Чтобы резервированная архитектура работала правильно, две фабрики нужно использовать с несколькими HBA, несколькими RAID-контроллерами и средствами параллельного доступа (multipathing), обслуживающими устройства SAN, которым нужна максимальная доступность.

Отметим, что резервирование фабрики не исключает применение отказоустойчивости, и лучше комбинировать эти два подхода. В изображенной на рисунке схеме каждая фабрика отказоустойчива, и они резервированы между собой.

Итак, можно сформулировать еще один принцип: «Решения HA SAN должны использовать полностью резервированную архитектуру 'A/B' и предпочтительно, чтобы каждая фабрика была отказоустойчивой».

Зонирование

С помощью зонирования администратор может разбить фабрику на зоны, в которые объединены взаимодействующие между собой устройства, и применить лучшие в своем классе механизмы блокировки связи между устройствами из разных зон. Зоны могут перекрываться и поддерживать как очень простые, так и сложные схемы передачи трафика. Зонирование можно

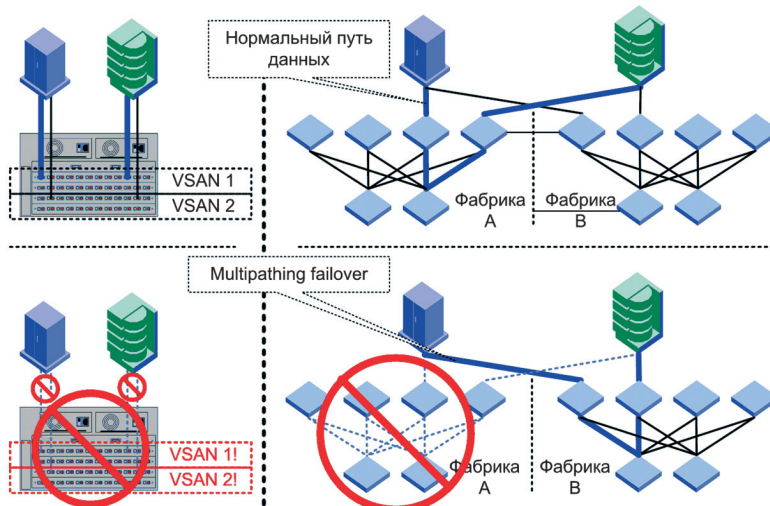


Рис. 9. Сравнение физического резервирования фабрик с логическим делением коммутатора на VSAN.

рассматривать как комбинирование лучших аспектов технологии VLAN с мощными функциями безопасности, обычно используемыми только в межсетевых экранах корпоративного класса.

Зоны удобны для создания барьеров между разными операционными средами, создания внутри фабрики нескольких функциональных групп или тестовой среды и/или областей, изолированных от остальной фабрики. Если в фабрике используются зоны, то никакой сервер или устройство хранения не смогут использовать фабрику до тех пор, пока оно явно не включено хотя бы в одну зону авторизованным администратором фабрики.

Зонирование — это инструмент уровня фабрики, и конфигурации зон автоматически передаются на все ее коммутаторы. Фабрика обеспечивает максимум резервирования и надежности, поскольку каждый коммутатор хранит информацию о зонах и может ее передать новому коммутатору, добавленному в фабрику.

Типы зонирования

Самое простое зонирование — это зонирование на базе портов, то есть физические порты коммутатора распределяются между разными зонами. Например, задается такое правило: только устройства, подключенные к порту 1 коммутатора 1 могут «разговаривать» с устройствами на порте 2 коммутатора 3.

В настоящее время предпочтительным методом стало зонирование на базе WWN, которое ограничивает доступ по WWN устройств вместо идентификатора порта. Это более гибкий метод, поскольку любые узлы сети всегда остаются в назначенной им зоне. Однако у него есть один недостаток — если заменить устройство, то у него изменится WWN (адрес порта останется прежним), поэтому потребуется вносить обновление в конфигурацию зон. Разумеется, переключение устройства от одного порта к другому при использовании зонирования по WWN выполняется проще, поэтому этот недостаток не считается принципиальным.

Зонирование бывает жесткое (аппаратное) и мягкое (программное). При мягком зонировании разбивка на зоны выполняется только с помощью операционной системы коммутатора — обычно сервер имен фабрики предоставляет конечным узлам только ограниченную информацию. Узлы в фабрике информируются о существовании других узлов только с помощью сервера имен, и если этот сервис ограничивает его ответы на запросы в соответствии с данными зон, то на практике никакое устройство не может обращаться к устройствам из других зон (однако пакеты могут передаваться между разными зонами). Этот подход работает эффективно до тех пор, пока зоны не надо будет изменить, пока оборудование не записывает в кэш таблицы сервера имен или пока пользователю не потребуется гарантировать, что никакие пакеты случайно или преднамеренно не могут выйти за пределы определенных зон.

Устройства Brocade в подавляющем большинстве случаев используют жесткое (аппаратное) зонирование, реализованное в микросхемах портов. После активации зонной конфигурации, ее содержимое компилируется и прописывается в CAM-таблицы в памяти коммутатора. Та-

кое зонирование более быстрое и надежное с точки зрения безопасности.

Мягкое зонирование включается в случае, если зоны определены в разных форматах: и по WWN, и по номерам портов одновременно. Такая схема зонирования не рекомендуется.

Разработка плана зонирования

Хотя текущим управлением зонированием обычно занимается администратор SAN, архитектор SAN должен продумать, как построить зоны, поскольку от этого зависит, какие потоки данных между устройствами будут разрешены. Он также должен еще на фазе проектирования проанализировать характер потоков с точки зрения производительности, что упростит разработку плана зон.

Здесь важен творческий подход, поскольку не может быть только одной «правильной» конфигурации зон для данной сети. Далее описывается один из распространенных подходов.

Сначала необходимо составить список всех серверов и систем хранения данных, которые нужно объединить в зоны (обычно к этому моменту уже имеется список подключенных к SAN устройств). В зависимости от привычных для архитектора SAN утилит этот список может быть представлен в виде электронной таблицы, текстового файла ASCII или другом формате. Список должен включать имя, физическое расположение, WWN, идентификатор порта и функцию каждого устройства.

Затем надо оценить потребности каждого сервера в хранении, и если ему выделено отдельное устройство, то указать это в списке, а если нет, то определить, сколько устройств хранения и какого типа требуется серверу. Далее надо найти ближайшее устройство хранения, которое соответствует требованиям сервера и выделить серверу емкость из этого дискового массива. В списке надо указать, что серверу требуется доступ к этому массиву. Необходимо предусмотреть сценарии переключения при сбоях, несколько портов под управление программно обеспечения multipathing и доступ к серверам резервного копирования и лентам.

После того, как у архитектора появится понимание, какому порту сервера нужен доступ к какому порту устройства хранения, надо записать эти сведения.

На этом этапе можно создать общий план зонирования, где отмечен необходимый доступ. Например, архитектор может сконфигурировать одну зону из одного порта устройства хранения и всех серверов, которые обращаются к нему. При этом один сервер может обращаться к другому серверу из этой же зоны, что в SAN обычно не требуется и даже нежелательно. Поэтому лучше иметь по одной зоне на пару устройств и использовать перекрывающиеся зоны. В перекрывающихся зонах адаптеры серверов совместно используют один или несколько портов систем хранения, но сами адаптеры изолированы друг от друга. В идеале фабрика состоит из большого числа зон с парами портов или WWN и определяет связь между парой «ини-

циатор-цель» (SCSI Initiator/Target). Такие зоны иногда называют point-to-point. Если инициатор сконфигурирован на доступ к нескольким портам устройств хранения данных, то следует включить его в несколько зон вместо того, чтобы объединять несколько портов в одну зону устройств хранения, тогда будет возможен только тот доступ, который необходим.

Зонирование с одним инициатором (single initiator zoning)

В очень больших сетях применение зон point-to-point не всегда удобно, поскольку оно означает создание большого числа зон и необходимость многочисленных изменений в конфигурации фабрики при добавлении и перемещении устройств. Более популярный метод — это зоны с одним инициатором, при котором достигается баланс между необходимостью ограничить число зон и упрощением управления. Для каждого инициатора в фабрике создается отдельная зона, которая также содержит все порты систем хранения для этого инициатора.

Кроме обеспечения повышенной безопасности, этот подход помогает оптимизировать сервисы фабрики. Зоны оптимизируют сервисы фабрики, такие как распространение уведомлений (RSCN) и ответ сервера имен, и ограничивают обнаружение «лишних» устройств. Правильная конфигурация зон необходима для больших сетей (даже если для них безопасность не критична), и методы point-to-point и «одного инициатора» являются оптимальными (рис. 10).

Архитектору также следует позаботиться об удобной системе имен для зон. Для улучшения масштабируемости и управляемости лучше всего использовать короткие и понятные имена. Трудно придумать короткое, но в то же время понятное имя, поэтому рекомендуется сокращать имена так, чтобы они не потеряли своего значения.

Зоны должны содержать только псевдонимы (alias) устройств, но не номера портов или WWN, для того, чтобы можно было определить устройство только один раз и использовать его в нескольких зонах без переопределений. Если устройство вышло из строя, было перемещено или модернизировано, то потребуется изменить только WWN, привязанный к его псевдониму, после чего это изменение отразится на всех зонах, в которых он был прописан. Как имена зон, псевдонимы должны быть по возможности короткими и понятными, чтобы их мог легко идентифицировать администратор SAN.

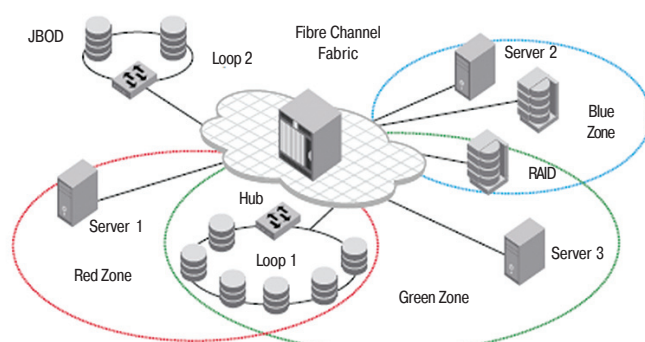


Рис. 10. Зонирование с одним инициатором.

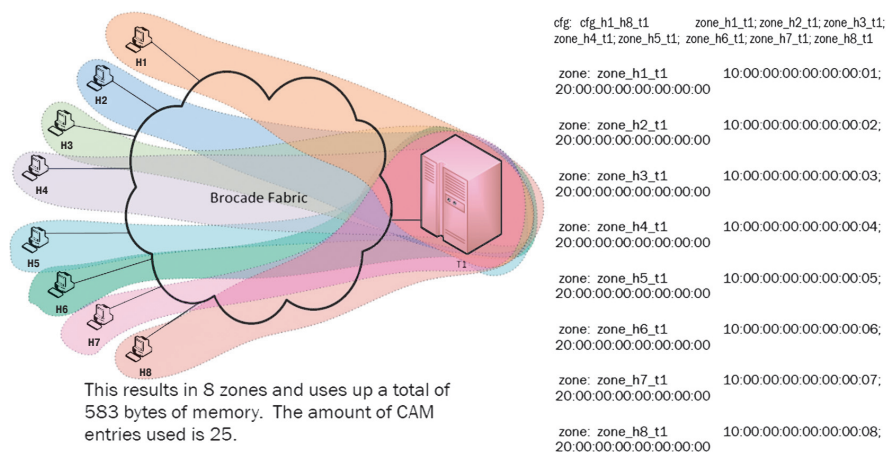


Рис. 11. Традиционный способ зонирования требует создания 8 различных зон.

Зонирование стоит использовать в комбинации с другими методами, такими как маскирование HBA- или LUN-контроллера хранения (LUN mapping, LUN masking).

Peer Zoning – новый подход к зонированию

В операционную систему Fabric OS начиная с версии 7.4 был внесен достаточно интересный функционал.

До версии Fabric OS 7.3 включительно рекомендуемым и наиболее часто используемым методом зонирования было так называемое зонирование с одним инициатором (single initiator zoning, рис. 11)

При подобном методе зонирования для каждого инициатора в фабрике создается отдельная зона, которая также содержит все порты систем хранения, с которыми работает данный инициатор. Этот метод считался наиболее популярным в течение долгого времени, так как позволял достичь баланса между числом сконфигурированных в фабрике зон и простотой управления.

В настоящее время, в связи с ростом размеров фабрик и числа подключенных к фабрикам устройств подобный метод перестал удовлетворять требованиям по масштабируемости и управляемости. Количество сконфигурированных зон быстро увеличивается, все сложнее становится ими управлять. В каждом коммутаторе сети SAN имеется специальная база данных (zone database), в которой хранится информация о зонной конфигурации в фабрике. Размер памяти, отведенной для данной базы данных, конечен. При росте числа зон возможна ситуация, когда размер базы данных достигнет своего верхнего предела и дальнейшая работа с зонами станет невозможна.

Для упрощения конфигурирования зон и сокращения количества записей в базе данных стандартами протоколов Fibre

Channel следующего поколения (FC-SW-6 и FC-GS-7) предусмотрен новый тип зонирования – Peer Zoning. Хотя этот стандарт обязателен для оборудования Fibre Channel следующего поколения (Gen 6), с выходом версии FOS 7.4 он стал доступен и на существующем оборудовании Gen 5.

При конфигурации зонирования методами Peer Zoning в базе данных создаются два типа участников – основной участник (principal member) и неосновной участник (peer member).

Зонная конфигурация позволяет взаимодействовать между собой основным и неосновным участникам. Но любые два неосновных участника не могут взаимодействовать между собой. Так же не могут взаимодействовать между собой любые два основных участника.

Для типичной конфигурации зонирования количество инициаторов (серверов) значительно превышает количество target устройств (систем хранения). Поэтому инициаторы (серверы) конфигурируются как неосновные участники (peer members), а системы хранения – как основные участники (principal members). В одну зону можно включить необходимое количество как principal members, так и peer members (рис. 12).

Преимущества подобной конфигурации следующие:

- значительно упрощается конфигурация зонирования за счет сокращения количества зон

по сравнению с традиционными методами зонирования;

- сокращается количество широковещательных сообщений в фабрике (RSCN distribution);
- эффективно используется база данных зонирования, сокращается нагрузка на внутренние процессоры коммутатора.

В случае использования Peer Zoning конфигурация сокращается до 1 зоны, в которую входят 8 peer members и один principal member.

Применяемый в Fabric OS 7.4 метод конфигурирования и использования Peer Zoning полностью соответствует новым стандартам FC-SW-6 и FC-GS-7. При этом зонная конфигурация фабрики может содержать как традиционные зоны, так и новые зоны с использованием Peer Zoning одновременно. Peer Zoning также может использоваться при маршрутизации между фабриками (Fibre Channel Routing).

Target Driven Zoning (TDZ)

Другой интересной возможностью, доступной начиная с версии Fabric OS 7.4 является Target Driven Zoning (TDZ). Фактически этот функционал предоставляет возможность устройствам, подключенным к сети SAN, создавать зонную конфигурацию самостоятельно.

Функционал также является частью новых стандартов Fibre Channel FC-SW-6 и FC-GS-7 и основан на использовании описанного выше метода зонирования Peer Zoning.

Устройства, подключенные к сети SAN, могут вносить изменения в зонную конфигурацию, используя свое подключение к коммутаторам по протоколу Fibre Channel. В основном данный функционал предназначен для систем хранения данных. Управляющее ПО системы хранения данных определяет, какие серверы (инициаторы) могут работать с СХД, и на основании этого создаются необходимые зоны в фабрике.

Преимущества подобного метода управления зонной конфигурацией:

- интеграция процесса создания и изменения зон с процессом управления системой хранения данных;
- упрощение и ускорение процесса подключения СХД к серверам за счет снижения требуемого количества шагов и количества различных интерфейсов управления.
- снижение количества возможных ошибок за счет автоматизации операций управления зонного конфигурирования.

Функционал Target Driven Zoning полностью зависит от поддержки его возможностей системами хранения данных и их управляющего ПО. В настоящее время возможность подобного автоматического конфигурирования ПО доступна для систем хранения данных различных производителей.

Виртуальные фабрики

Для более эффективного использования ресурсов, изоляции ошибок и повышения масштабируемости в операционную систему Brocade FOS был добавлен механизм виртуальных фабрик

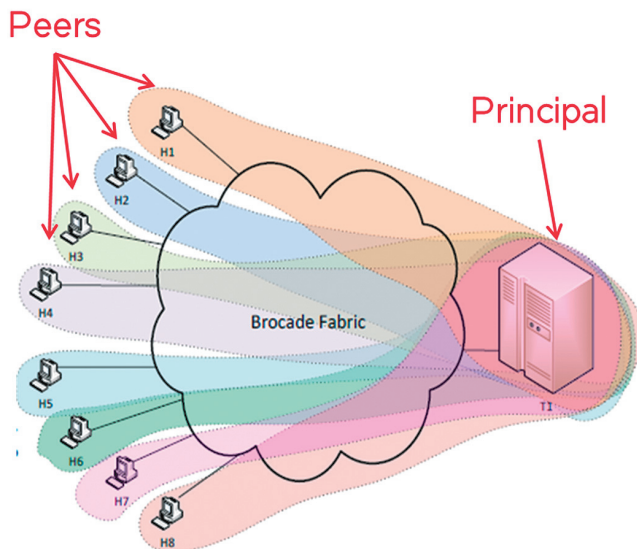


Рис. 12. В случае использования Peer Zoning конфигурация сокращается до одной зоны, в которую входят 8 peer members и один principal member.

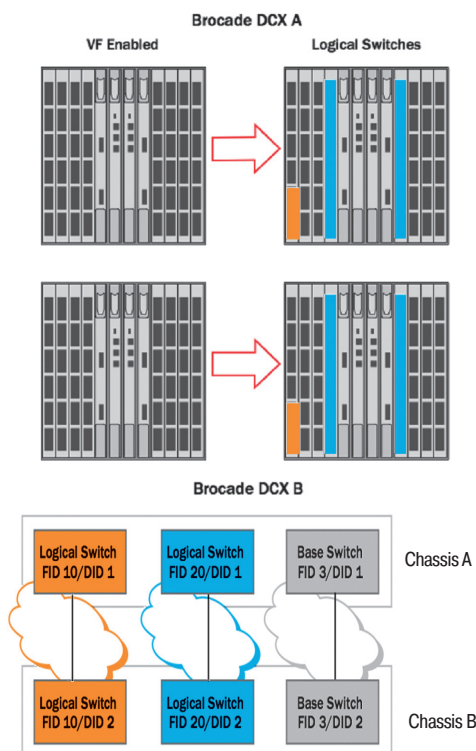


Рис. 13. Выделение логических коммутаторов в шасси DCX.

(virtual fabric, VF). Виртуальные фабрики могут состоять из одного или нескольких логических коммутаторов (logical switch, LS), а также физических коммутаторов (не разделенных на логические). У каждой виртуальной фабрики должен быть уникальный идентификатор FID (рис. 13).

Логический коммутатор — это набор портов на физическом коммутаторе, выделенный в отдельную единицу управления. Логические коммутаторы могут соединяться между собой с помощью различных видов ISL и создавать виртуальные фабрики. Кроме того, виртуальные фабрики могут участвовать в Meta SAN и выступать в роли бэкбонных или периферийных (рис. 14).

Функционал виртуальных фабрик отдельно не лицензируется, он присутствует по умолчанию

Табл. 3. Максимальное количество логических коммутаторов в зависимости от платформы

Platform	Maximum number of logical switches
Brocade X6 Directors	8
Brocade DCX 8510 Backbones	8
Brocade 6510	4
Brocade 6520	4
Brocade G620	4
Brocade 7840	4

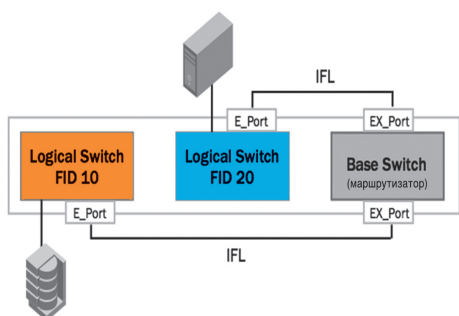


Рис. 14. Маршрутизация между виртуальными фабриками.

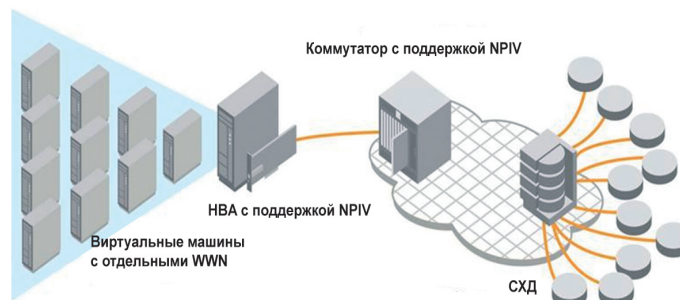


Рис. 15. Хост виртуальных машин, подключенный в SAN через NPIV.

на большинстве коммутаторов Brocade. Как правило, он используется в крупных сетях хранения данных для создания выделенных фабрик репликации и СРК, виртуальных бэкбонных фабрик, выделенных фабрик для разных подразделений, а также в случае одновременного использования FCP и FICON. Видео-пример настройки виртуальных фабрик и маршрутизации можно посмотреть на YouTube: <http://youtu.be/nEp4y0Da6Xc>.

Шлюз доступа

Две насущные проблемы в мире Fibre Channel — это совместимость между собой сетевого оборудования разных вендоров и масштабирование крупных SAN.

Несмотря на то, что существуют стандарты Fibre Channel, у разных производителей оборудования есть собственные наработки, которые улучшают работу SAN, но не являются стандартными (это касается в основном коммутаторов — серверные адаптеры и контроллеры систем хранения меньше страдают от проблем с совместимостью). Для того, чтобы совместить между собой коммутаторы разных производителей, необходимо привести их функционал к «общему знаменателю», отключив нестандартные функции. На практике, такое урезание функционала приводит к тому, что производительность, надежность и масштабируемость фабрики существенно снижается. Исходя из этих соображений, Brocade не поддерживает подключение коммутаторов других производителей через E-порт или EX-порт.

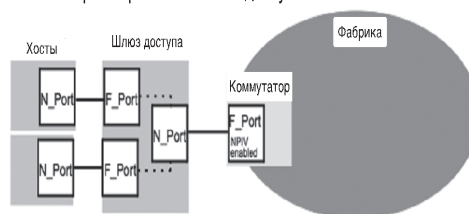
Масштабируемость SAN также становится все более насущным вопросом, особенно по мере того, как набирают популярность решения на блейд-серверах, когда в каждом шасси может быть установлено несколько FC коммутаторов. Максимальное количество коммутаторов в фабрике ограничено 56, однако, чем больше доменов в фабрике, тем больше времени и трафика нужно для работы внутренних процессов — построения таблиц маршрутизации, операций с зонированием и т.д. Поскольку весь служебный трафик имеет максимальный приоритет, то большое его количество может вызвать задержки при передаче полезного трафика приложений. Особенно неприятный эффект бывает при «дребезжании» оптических линков — в ситуации, когда сила сигнала находится на границе уверенного приема. В случае, если такое дребезжание возникает на ISL без транка, соответствующий маршрут в фабрике постоянно пропадает, а потом снова появляется. Такое поведение приводит к перестроению фабрики (build fabric, BF). Это штатная ситуация, которая не приводит к прерыванию пользовательского трафика, но когда пропадание линка происходит слишком часто,

объем служебной информации между коммутаторами растет и в какой-то момент может вытеснить трафик приложений.

N_Port ID Virtualization (NPIV) — протокол, разработанный в 2002 году для обеспечения возможности нескольким устройствам со своими адресами использовать один физический N-порт в фабрике. Например, один физический сервер, подключенный к порту коммутатора, виден в фабрике как один N-порт со своим 24-битным адресом, WWN и т.д. В случае, если на порту включен режим поддержки NPIV, а на сервере работают несколько виртуальных машин, то в фабрике появятся N-порт, 24-битные адреса и WWN для каждой из виртуальных машин (это полезно, например, для зонирования — с помощью зон можно будет ограничивать доступ к СХД каждой виртуальной машине, а не всему физическому хосту). Режим NPIV можно рассматривать как SAN аналог NAT (рис. 15).

Шлюз доступа (access gateway, AG) — это функционал в FOS, позволяющий перевести коммутатор в специальный режим, в котором он становится аналогом многопортового HBA с поддержкой NPIV. При включении AG, коммутатор перестает видиться в фабрике как отдельный домен, на нем перестает выполняться часть команд, касающихся сервисов фабрики (например, команды зонирования), а за коммутационную логику начинает отвечать вышестоящий коммутатор. AG подключается в фабрику уже не с помощью E-порта и ISL, а как обычный хост с помощью N-порта к F-порту на вышестоящем коммутаторе (рис. 16).

Порты в режиме шлюза доступа AG



Порты в обычном режиме фабрики (native fabric)

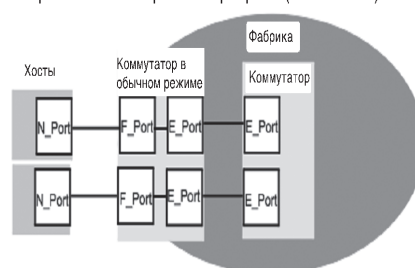


Рис. 16. Типы портов в режимах фабрики и шлюза доступа.

С функцией шлюза доступа масштабируемость фабрики растет в несколько раз за счет того, что небольшие коммутаторы работают в режиме AG, не являясь доменом в фабрике и, соответственно, не увеличивая объем служебного трафика внутри SAN. В случае неисправности линка между шлюзом доступа и коммутатором, страдают только хосты, подключенные к шлюзу доступа — для фабрики это не является аварийной ситуацией, требующей перестроения. Brocade рекомендует переводить небольшие коммутаторы, и особенно коммутаторы в шасси блейд-серверов, в режим Access Gateway.

Поскольку при работе в режиме NPIV большая часть функционала на шлюзе доступа отключается, это позволяет решить проблемы с совместимостью: Brocade поддерживает подключение коммутаторов в режиме Access Gateway к SAN других производителей, а также подключение коммутаторов других производителей, работающих в аналогичных NPIV режимах к коммутаторам Brocade.

Несмотря на то, что AG — это упрощенный режим работы коммутатора, в нем есть дополнительные функции, обеспечивающие повышенную надежность, масштабируемость и управляемость — режимы failover, N_Port trunking, каскадирование, группирование портов и балансировка нагрузки. Подробнее про этот функционал можно почитать в Fabric OS Access Gateway Administrator's Guide.

Режим шлюза доступа является базовым функционалом FOS и не требует дополнительного лицензирования. Режим шлюза доступа поддерживается на большом количестве аппаратных платформ — практически на всех коммутаторах 4,8,16 и 32 Гбит/с за исключением моделей с большим количеством портов — директоров DCX, DCX8510, X6, а также Brocade 5300 и 6520. Необходимо учитывать, что ранее для работы в режиме AG на коммутаторах Brocade с фиксированным количеством портов (300, 6505, 5100, 6510) требовалось наличие всех лицензий Ports on Demand. Начиная с версии FOS 7.3.0 это требование отменено, режим Access Gateway работает и на коммутаторах, где присутствуют неактивированные порты.

Масштабирование фабрик SAN до уровня Meta SAN

Технология Fibre Channel Routing (FCR) позволяет передавать данные между сетями (SAN) без слияния сетей. Эту технологию можно сравнить с IP-маршрутизацией из мира LAN, поэтому ее иногда называют Layer 3 routing (рис. 17).

Как видно из рисунка, можно соединить несколько «островов» SAN (изолированных фабрик), подключив их к FC маршрутизаторам с помощью особых соединений (Inter-Fabric Links, IFL). В данном случае полученная большая сеть называется Meta SAN, поскольку она реализует уровень иерархии выше традиционной SAN. В Meta SAN каждая фабрика идентифицируется с помощью уникального байта Fabric Identifier (FID). Для соединения устройств из разных фабрик Meta SAN создаются специальные зоны (Logical Storage Area Networks, LSAN zones), охватывающие несколько фабрик. FC маршрутизаторы обеспечивают подключение и могут использоваться для увеличения архитектурной и практической мас-

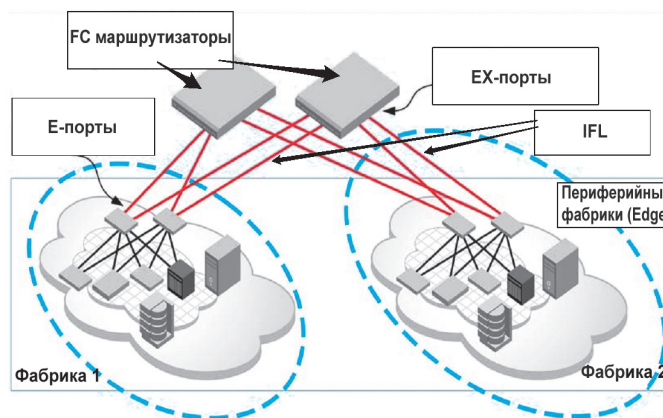


Рис. 17. MetaSAN и основные термины.

штабируемости на несколько порядков, поскольку они добавляют еще один байт к адресному пространству фабрики и решают проблему масштабирования, которая до сих пор ограничивала максимальный размер фабрики. Тем не менее, у каждой фабрики сохраняется собственная копия сервисов фабрики, она может управляться отдельно от остальных фабрик Meta SAN и их устройств.

Настройка маршрутизации проходит в несколько этапов. Сначала порты на маршрутизаторах конфигурируются для работы в режиме EX, чтобы подключенные к ним фабрики не сделали попытку объединиться. После этого, назначаются уникальные идентификаторы FID каждой из подключенных фабрик. После того, как маршру-

тизаторы увидели все периферийные фабрики, наступает время экспортировать устройства, которые должны общаться между собой. На рисунке видно, что хост-инициатор из фабрики 2 экспортируется в фабрику 1, а СХД, с которой он работает, экспортируется из фабрики 1 в фабрику 2. В обеих фабриках создаются «фантомные» виртуальные свитчи-домены, к которым логически подключены экспортированные устройства. Обо всей топологии Meta SAN знают только маршрутизаторы. Для периферийных фабрик процесс импортирования устройств выглядит как добавление нового коммутатора с подключенными серверами и СХД.

Механизм экспортирования использует проверенную технологию зонирования — в обоих пе-

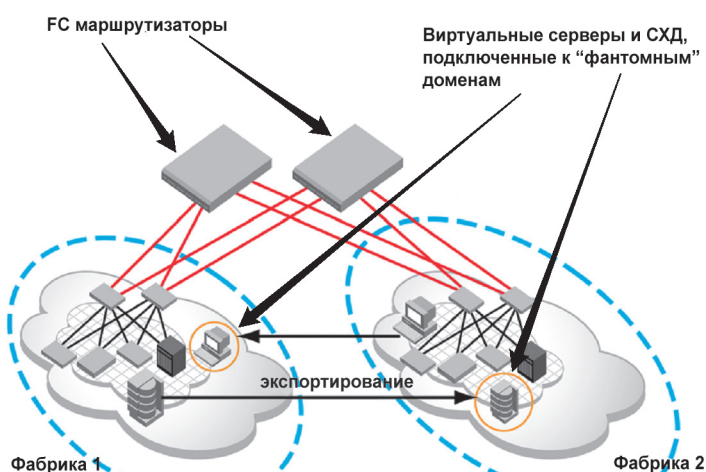


Рис. 18. Экспортирование устройств с помощью маршрутизации FC.

Периферийные фабрики расположены в разных помещениях

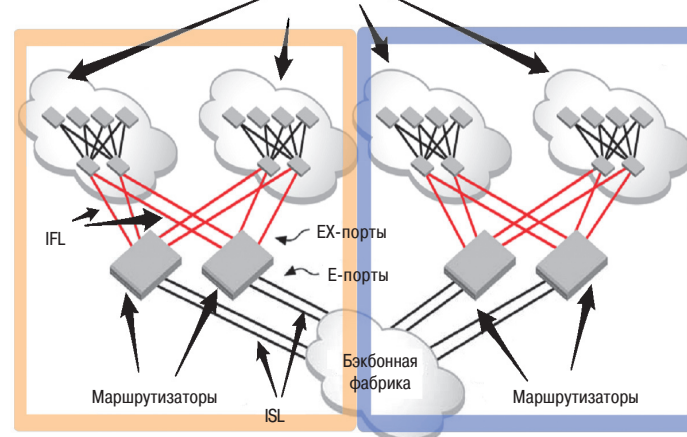


Рис. 19. Связь между площадками с помощью FC-маршрутизации и бэббонной фабрики.

риферийных фабриках создаются зоны, в названии которых должен быть префикс «LSAN_». Содержимое этих зон в обеих фабриках должно быть одинаковым — это WWN экспортируемых устройств. Маршрутизатор постоянно опрашивает Name Server в периферийных фабриках, и как только он видит эти специальные зоны, он экспортирует устройство в нужную фабрику (рис. 18).

В крупных SAN зачастую используется не пара маршрутизаторов, а целые фабрики, созданные специально для маршрутизации трафика. Такие фабрики называются бэкбонными (backbone). Как правило, такие фабрики организовываются в случае, когда необходимо маршрутизировать трафик между разными площадками и/или серверными помещениями. При такой организации SAN, как правило, пара маршрутизаторов ставится в каждое серверное помещение. Между собой маршрутизаторы объединяются посредством ISL и E-портов в бэкбонную фабрику, а периферийные фабрики подключаются через IFL и EX-порты. При этом, каналы связи между площадками и серверными помещениями, как правило, остаются внутри бэкбонной фабрики. Таким образом, неполадки на этих каналах передачи данных повлияют только на трафик, идущий по бэкбону (а это, как правило, трафик репликации и СРК), но не вызовут перестроения периферийных фабрик, в которых работают критичные для бизнеса приложения (рис. 19).

Технология FC маршрутизации поддерживается почти на всех современных коммутаторах Brocade (кроме моделей 300 и 6505). Для того, чтобы коммутатор получил возможность работать в режиме маршрутизатора и включать EX-порты, необходимо наличие лицензии Integrated Routing.

Связь центров обработки данных на расстоянии

За последние годы резко выросла потребность в решениях класса Disaster Recovery (DR) и Business Continuity (BC), что вызвало существенный спрос на высокоскоростную и высоконадежную передачу данных на уровне блоков между системами хранения данных, установленными на разных площадках. Кроме DR и BC имеются и экономические факторы, способствующие построению SAN, которые объединяют несколько площадок, в том числе требования к ИТ-подразделениям достигать большего, используя меньше ресурсов. Более эффективное использование систем хранения корпоративного класса также стало стимулом для роста SAN, а технологии обеспечения связи на больших расстояниях позволяют реализовать преимущества SAN для нескольких площадок.

Рассмотрим основные технологии соединения SAN.

Темная оптика (FC over Dark Fiber)

Трафик Fibre Channel передается по ISL или IFL по выделенной оптической линии, соединяющей площадки. Такой подход обеспечивает высокую надежность и производительность, поскольку не используется промежуточное оборудование, которое может отказать, не требуется преобразование протоколов, и полоса пропускания ис-

пользуется только приложениями SAN. Для передачи данных на большие расстояния используются длинноволновые модули (LWL/ELWL SFP) в сочетании с одномодовым оптоволоконным кабелем. Эта технология меньше всего влияет на задержки, и является предпочтительной для случаев, когда требуется синхронная репликация данных между системами хранения данных либо требуется обеспечить доступ серверов к данным на удаленной площадке в случае аварии. Поскольку в этих случаях производительность приложения напрямую зависит от задержек доступа к дискам, и соответственно от расстояния между площадками (при скорости сигнала в оптоволокне = 2/3 от скорости света в вакууме, каждый километр трассы добавляет 5 мкс задержек), имеет смысл использовать технологию FC в темной оптике на расстояниях прямой досягаемости для длинноволновых SFP (80-100 км). На больших расстояниях придется использовать регенераторы или усилители сигнала, которые также добавляют задержек и стоимости всему решению, а суммарное время прохождения сигнала между парой устройств станет слишком большим для синхронной репликации.

Оптическое мультиплексирование (FC over xWDM)

Соединения Fibre Channel осуществляются с подключением к аппаратуре оптического мультиплексирования (DWDM или CWDM). Трафик FC передается на выделенной длине волны, но по волокну между площадками могут работать другие сервисы, использующие иную длину волны. По производительности этот подход аналогичен темному волокну и не добавляет задержек кроме тех, что связаны со скоростью света или ограничением пропускной способности. Подход практически так же надежен, как темное волокно (если не считать риск отказа WDM-устройств). В то же время он поддерживает передачу на большие расстояния, чем темное волокно, если промежуточные узлы WDM используются как повторители. Следует отметить, что в некоторых случаях коммутаторы или маршрутизаторы FC на обоих концах должны поддерживать большое число буферных кредитов, если средства WDM не обеспечивают этого. Некоторые WDM-устройства позволяют использовать свои буферы для связи на дальние расстояния (технология buffer-to-buffer credit spoofing).

Технология пассивного мультиплексирования также очень популярна в решениях Fibre Channel. Аппаратно эта технология реализуется в SFP трансиверах с разными длинами волн в диапазоне CWDM или DWDM. Эти модули устанавливаются в коммутаторы вместо обычных SWL/LWL SFP. Сигналы разных «цветов» затем поступают на пассивную призму, которая мультиплексирует несколько длин волн и передает получившийся сигнал по одному волокну на аналогичную призму на другой стороне канала. Там, в свою очередь, сигнал демultipлексируется и попадает на трансиверы с соответствующей длиной волны. Brocade поддерживает пассивные CWDM и DWDM трансиверы производства компаний SmartOptics и Finisar. Список поддерживаемых модулей приведены в документе: «Brocade Fabric OS Compatibility Matrix».

Буферные кредиты

Важным для связи на расстоянии является понятие буферных кредитов. Fibre Channel был создан для передачи SCSI трафика по сети, и поэтому при нормальной работе должен обеспечивать передачу трафика без отбрасывания пакетов. Это отличает FC от, например, Ethernet, где в случае возникновения узких мест в сети, пакеты попросту отбрасываются, а протокол верхнего уровня TCP обеспечивает целостность передачи данных. В FC порт-отправитель не может передавать пакеты, если порт-получатель не готов их принять.

Возможно, что порт-получатель не может сразу же передать пакет дальше, поэтому ему нужен буфер для временного хранения кадров. Все устройства в SAN имеют ограниченное число буферов, поэтому необходим механизм, с помощью которого устройства могут сообщать другим устройствам, что у них есть свободные буферы, до того, как к ним будет послан кадр, то есть отправитель должен знать, способен ли получатель принять кадр. Контроль потока (Flow control) используется для того, чтобы никакой порт в цепочке между отправителем и получателем не отбрасывал кадры FC.

Существуют различные способы реализации контроля потока в сети. В Fibre Channel используется механизм буферных кредитов (buffer-to-buffer credits, BB_credits). Все устройства Fibre Channel (коммутаторы, маршрутизаторы, HBA, JBOD-диски и контроллеры дисковых массивов) используют этот механизм. Он гарантирует, что кадры не будут посылаться до тех пор, пока отправитель не получит от получателя подтверждение о наличии свободных буферов.

Когда порт устройства первоначально подключается к фабрике, то устройство-получатель сообщает отправителю число доступных буферов, а отправитель сообщает получателю, сколько буферов ему нужно. Устройства согласуют число доступных буферов и выделяют буферные кредиты до начала передачи данных.

Когда кадры передаются от узла к коммутатору, они записываются в память коммутатора, и передающее устройство делает отметку об использовании буфера. Передача прекращается, если больше не осталось свободных буферов (кредитов). Между тем, получатель попытается освободить буферы, пересылая пакеты дальше, и оповещает отправителя, если ему удастся освободить буферы (командой R_RDY или VC_RDY). Таким образом, кредиты возвращаются узлу максимально быстро и, если нет проблем в фабрике, устройство может постоянно передавать пакеты со скоростью физической линии. Если где-то (в сети или на принимающей стороне) возникнет переполнение линии, то, вскоре, все буферные кредиты будут исчерпаны; передающее устройство будет ждать освобождения буферов и возобновит пересылку данных, только когда ему будет предоставлен хотя бы один кредит.

Коммутаторы Brocade FC обычно согласуют буферные кредиты прозрачно для пользователей. Например, таким локальным портам, как F-порт (порт для подключения оконечных устройств — серверов или систем хранения) не нужно много буферов для поддержания скорости линии. По

умолчанию, продукты Brocade выделяют восемь буферов на F-порт, что более чем достаточно для поддержания максимальной скорости передачи.

При использовании Fibre Channel на больших расстояниях (например, свыше 500 метров) использование буферных кредитов очень важно, особенно при очень больших расстояниях (10–500 км) — при передаче по темному волокну, с помощью оборудования WDM или SONET/SDH, поскольку передача кадра FC по оптическому кабелю занимает определенное время (скорость света в оптическом кабеле приблизительно равна 2/3 от скорости света в вакууме).

Разумеется, при такой скорости это время не может быть очень большим. Но при достаточно большой длине кабеля и скорости соединения одновременно в кабеле может находиться сразу несколько кадров. С помощью несложного математического расчета можно точно посчитать это количество:

Для примера возьмем кабель длиной 20 км, по которому передаются данные со скоростью 8 Гбит/с.

Максимальный размер пакета — 2112 байт, при этом, учитывая перекодирование 8b/10b, для передачи каждого байта информации используется 10 бит. При скорости 8 Гбит/с этот пакет передается за время, равное $(2112 \times 10) / (8 \times 10^9) \approx 2.5 \times 10^{-6}$ с.

Учитывая скорость света в оптическом кабеле — $300000 \times 2/3 = 200000$ км/с, можем посчитать, что один пакет занимает в кабеле $200000 \times 2.5 \times 10^{-6} = 0.5$ км.

Соответственно, в 20 км кабеле одновременно в каждую сторону может передаваться количество фреймов равно $20/0.5 = 40$.

Для стопроцентного использования пропускной способности линка, необходимо, чтобы буферов с каждой стороны хватало, чтобы передать эти пакеты, а также дожидаться подтверждения освобождения буфера — R_RDY, то есть необходимо $40 \times 2 = 80$ буферных кредитов.

Для определения числа буферов (кредитов) для заданного расстояния используется следующая упрощенная формула:

«Число кредитов» = «число километров» * («скорость в Гбит/с» / 2)

Стоит отметить, что соединения смогут работать, если число свободных буферов недостаточно, но передача будет менее эффективна. Трасса 8 Гбит/с длиной 20 км сможет работать с 50 кредитами, но максимальная пропускная способность при этом составит примерно 5 Гбит/с.

Важно помнить, что число передаваемых кадров равно числу кредитов, то есть, например, если нужно обеспечить максимальную скорость передачи для трассы длиной 20 км в режиме 8 Гбит/с, то нужно передавать 80 кадров. В данном примере подразумевается, что длина кадра — 2 КБ; если будут использоваться кадры 1 КБ, то потребуется вдвое больше кредитов. Эта особенность помогает лучше понять механизм буферных кредитов Fibre Channel.

Учитывая то, что в реальных условиях длина пакета отличается от 2КБ, рекомендуется выделять буферные кредиты исходя из среднего

размера пакета или, если его сложно посчитать, с некоторым запасом прочности.

Также стоит отметить, что ISL нельзя «разогнать» выделением большого числа буферов.

Режимы передачи на большие расстояния

Устройства Brocade имеют несколько режимов передачи на большие расстояния

Существуют простые правила выбора нужного режима:

- все локальные порты E-port и EX-port работают в режиме **LO**, буферные кредиты выделяются по умолчанию;
- режим **LE** используется на дистанциях до 10 км, независимо от скорости передачи; этот режим не требует лицензии;
- динамический режим **LD** автоматически определяет длину трассы и с помощью сложного алгоритма вычисляет число необходимых кредитов, исходя из расстояния и скорости передачи (стоит учитывать, что LD считает количество буферов, исходя из размера пакета 2КБ);
- статический режим **LS** — это самый гибкий режим для более опытных пользователей, который позволяет обеспечить полное управление в сложных случаях. Например, при настройке трассы длиной 20 км в режиме 8 Гбит/с с размером кадров 1 КБ вместо обычных 2 КБ, в этой ситуации требуется вдвое больше буферных кредитов (160). Для выделения 160 кредитов необходимо выставить порт в режим LS и указать вдвое большую дистанцию — 40 км.

FC over IP

FCIP — это основной способ связи ЦОД в случаях, когда передача FC по темной оптике или через xWDM невозможна или экономически нецелесообразна: на расстоянии в сотни и тысячи километров или в случае отсутствия оптоволоконной между площадками даже на небольшом расстоянии. Основное преимущество FCIP — возможность относительно недорого подключаться к практически везде доступным сервисам MAN/WAN.

При использовании FCIP, пакеты Fibre Channel инкапсулируются в TCP/IP, а затем передаются через IP-туннель между парой FCIP устройств.

Поддерживается как прямое подключение FCIP устройств друг к другу, так и соединение через MAN/WAN инфраструктуру (необходимо, чтобы задержки в MAN/WAN между парой FCIP маршрутизаторов не превышали 200 мс, а количество потерянных пакетов было бы не больше 1%). Поддержка Jumbo Frames не требуется.

Технология FCIP в реализации Brocade расширена дополнительным функционалом, который позволяет сделать соединение дата-центров максимально надежным и производительным.

Протокол TCP был спроектирован для надежной работы в почти любых средах передачи данных, вне зависимости от скорости передачи, задержек, потерь и дубликации пакетов, изменения порядка сегментов. Разработчики Brocade оптимизировали TCP для передачи трафика SAN (storage-optimized TCP, SO-TCP) с помощью изменения некоторых количественных показателей:

- для улучшения производительности в сетях высокой пропускной способности, растянутых на большие расстояния (long fat networks, LFN) размер TCP cwnd увеличен до 20 МБ;
- влияние потерянных сегментов на производительность снижено за счет изменения процента уменьшения cwnd во время Fast Retransmit/Fast Recovery;
- для более быстрого набора максимальной скорости увеличены значения initial window, ssthresh;
- уменьшено количество ретрансмиссий за счет использования механизма выборочного подтверждения (selective acknowledgment, SACK, RFC 2188).

Одна из проблем при передаче SCSI трафика на большие расстояния — это время, которое тратится на инициализацию операции записи. По стандарту SCSI, необходимо сначала отправить запрос на разрешение записи (FCP_CMD_WRT), получить разрешение на запись (FCP_XFR_RDY), после чего уже можно отправлять данные. Время, которое тратится на подготовку к записи увеличивается по мере того, как растет расстояние между устройствами — сначала всю длину линка должна пройти команда FCP_CMD_WRT, а потом через эти же километры кабеля должно вернуться подтверждение FCP_XFR_RDY. Все это время данные не передаются, пропускная способность тратится впустую (рис. 20).

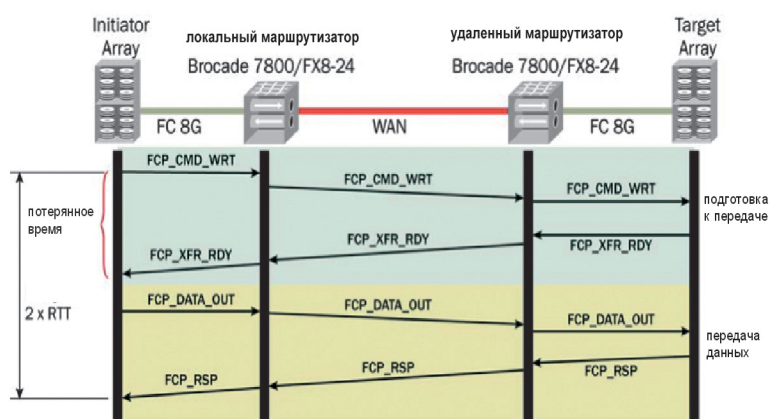


Рис. 20. Операция записи без FastWrite.

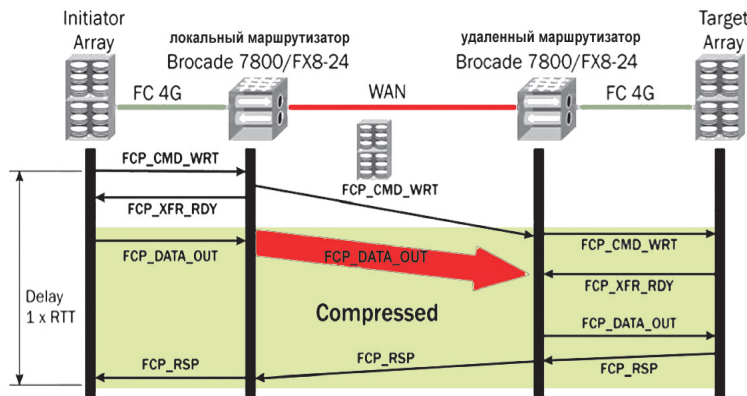


Рис. 21. Операция записи с использованием ускорения FastWrite.

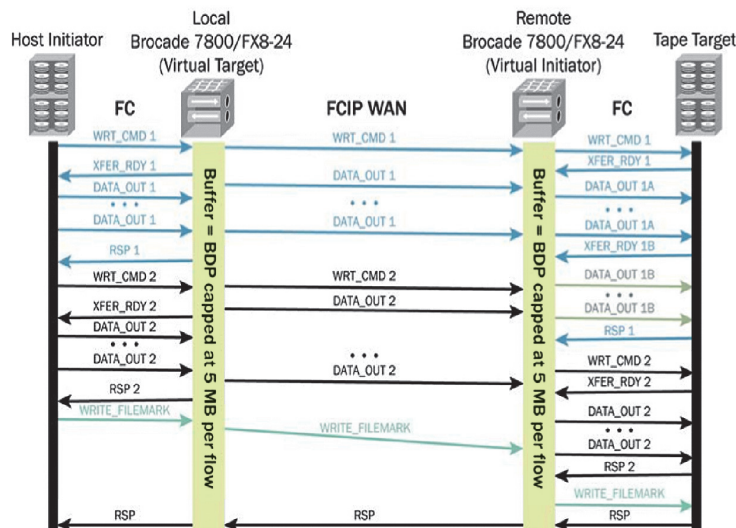


Рис. 22. Оптимизация записи на ленту с помощью Brocade OSTP.

Технология Brocade FastWrite предусматривает, что локальный маршрутизатор перехватывает команду на начало операции записи FCP_CMD_WRT и сразу же отправляет подтверждение инициатору операции записи в течение пары микросекунд. Таким образом на время RTT (время прохождения сигнала в обе стороны, return trip time) сокращается длительность каждой операции записи, а небольшие операции записи и вове укладываются в одно RTT, то есть задержки на канале передачи компенсируются на 50 процентов. Целостность данных при этом гарантируется тем, что подтверждение удачной записи FCP_RSP никогда не буферизуется и никак не изменяется (рис. 21).

Похожим образом оптимизируется и запись на ленту. Локальный маршрутизатор подменяет XFR_RDY от библиотеки и буферизует данные, предназначенные для записи на ленту, принимая в расчет доступную пропускную способность и длину канала между площадками. Переданные данные повторно буферизуются на удаленном маршрутизаторе, чтобы равномерно записывать их на ленту, без частых операций остановки и запуска ленты, истощающей ресурс привода. Как и в случае с FastWrite, окончательные записи WRITE_FILEMARK и подтверждение FCP_RSP никак не буферизуются и не модифицируются (рис. 22).

Подробнее про способы оптимизации FCIP можно почитать в документе «Brocade Open Systems Tape Pipelining, FastWrite, and Storage-Optimized TCP».

Для дополнительной оптимизации в FCIP маршрутизаторах также используется сжатие тра-

фика, передающегося через WAN. В зависимости от пропускной способности канала, можно выбрать более или менее агрессивные методы сжатия.

Для защиты данных при передаче через IP сети общего назначения можно использовать IPsec.

Будущее — технологии NVMe и NVMe over Fabrics

Традиционная технология Fibre Channel подразумевает передачу по сети SAN стандартных команд протокола SCSI. Тот протокол, который мы чаще всего называем Fibre Channel, с точки зрения стандартов правильнее называть SCSI over FC. Почему так произошло? Начиная с 80-х годов XX века протокол SCSI был и остается основным протоколом записи и чтения данных на магнитные диски. Варианты протокола с течением времени значительно изменились, увеличилась скорость работы, тип интерфейса сменился с параллельного на последовательный — Serial Attached SCSI (SAS). При этом сама структура протокола и принцип буферизации команд оставались в основном неизменными.

До тех пор, пока основным устройством хранения данных были магнитные диски, использование SCSI представлялось оправданным и достаточным для обеспечения необходимой скорости доступа. В последнее время широкое распространение получили устройства записи на твердотельной памяти — твердотельные диски, флэш-память. Подобные устройства (они часто называются NVM — Non-Volatile Memory) могут

обеспечивать значительно более высокую скорость чтения/записи и значительно меньшее время доступа по сравнению с традиционными магнитными дисками. Для подобных устройств использование протокола SCSI для обмена данными является неоптимальным вариантом с точки зрения производительности.

Для подобных устройств был разработан новый стандарт протоколов обмена данными. Новый стандартный протокол называется NVMe Express (NVMe). NVMe — программный протокол передачи данных для устройств, использующих физическую шину PCI Express. Протокол NVMe не использует команды SCSI для обмена данными. Это позволяет добиться очень высоких скоростей передачи при малых временных задержках.

В настоящее время протокол NVMe поддерживают практически все производители твердотельных устройств хранения данных. Большинство современных операционных систем имеют поддержку NVMe на уровне встроенных драйверов.

Вместе с тем у протокола NVMe есть ряд особенностей, ограничивающих его применение. В первую очередь это протокол прямого подключения дисков к серверу, используя шину PCIe. Он обладает низкой масштабируемостью, не поддерживает большие расстояния, не подходит для построения больших сетей хранения данных. Практически все крупные заказчики используют сети SAN и централизованные СХД, поэтому необходима адаптация протокола NVMe к задачам построения сетей SAN. NVMe over Fabrics — это расширение протокола NVMe для его использования в сетях SAN с опорой на существующую инфраструктуру сетей. Так как большинство сетей SAN построено на основе протокола Fibre Channel, логично использовать существующую инфраструктуру для работы с данными по протоколу NVMe. В настоящее время ведется разработка стандарта NVMe over Fibre Channel. Появление первых промышленных реализаций ожидается в 2016 году. Предполагается, что все коммутаторы Brocade Fibre Channel пятого и шестого поколения будут поддерживать NVMe over Fibre Channel. Для работы с NVMe over Fibre Channel потребуются сетевые адаптеры (HBA) шестого поколения. Таким образом, вся инфраструктура для подключения NVMe устройств в сетях SAN по протоколу Fibre Channel практически готова.

Продукты

Ниже описаны продукты Brocade Fibre Channel, предлагаемые компанией Brocade и нашими OEM-партнерами по состоянию на сентябрь 2016 года. Более подробные технические характеристики продуктов приведены в отдельном документе в форме таблицы.

Коммутаторы Brocade Fibre Channel четвертого поколения

Несмотря на выход коммутаторов шестого поколения Fibre Channel, в продуктовой линейке Brocade продолжают присутствовать коммутаторы четвертого поколения FC. Это базовые модели, хорошо зарекомендовавшие себя на рынке и обеспечивающие необходимый функционал.

Brocade 300 — это коммутатор начального уровня размером 1U, предназначенный для построения небольших SAN, либо для использования в качестве периферийного коммутатора для сетей типа «ядро-периферия» (рис. 23). Коммутатор поддерживает 24 порта FC 8 Гбит/с без переподписки, но может быть заказан в конфигурации

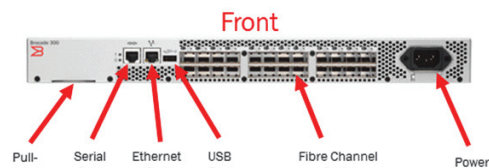


Рис. 23. Коммутатор Brocade 300.

8 и 16 портов, а также с SFP-модулями (трансиверами) 4 Гбит/с. Настройка коммутатора максимально упрощена благодаря программе EZSwitchSetup, которая позволяет сконфигурировать коммутатор из графического интерфейса. Для подключения к SAN других производителей или поддержания компактных размеров фабрики Brocade 300 может быть сконфигурирован в режиме Access Gateway, который позволяет подключать в SAN серверы по технологии NPIV и снимает ограничения по совместимости. Коммутатор Brocade 300 имеет только один блок питания, что ограничивает его применение в решениях повышенной отказоустойчивости.

Коммутаторы Brocade Fibre Channel пятого поколения

Директоры DCX 8510 Backbone

Директоры Brocade DCX 8510 Backbone аппаратно реализованы в виде модульных шасси размером 14U (DCX 8510-8) и 9U (DCX 8510-4). В шасси расположены блоки питания и вентиляторы с возможностью горячей замены, пассивный backplane, 2 управляющих лезвия (Control Processor Blade, CP) в отказоустойчивой конфигура-

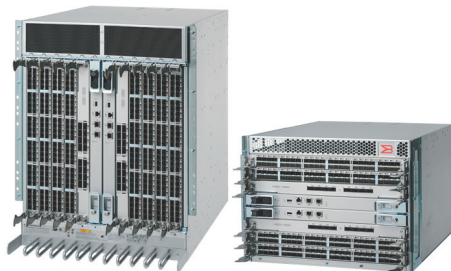


Рис. 24. Brocade DCX 8510-4 и DCX 8510-8.

ции и два коммутирующие лезвия (Core Blade, CR). Все модули DCX резервированы, включая управляющие и коммутирующие лезвия, блоки питания, вентиляторы охлаждения и платы, хранящие информацию о шасси (рис. 24).

Основной функционал коммутатора реализован в лезвиях, устанавливаемых в шасси.

Служебные лезвия, обязательно присутствующие в шасси

CP8 — управляющее лезвие, Control Processor.

На этом лезвии находится образ операционной системы Fabric OS, процессоры и память для ее работы. Там же расположены сетевые и консольные управляющие интерфейсы. Через backplane

два лезвия CP8, работающие в отказоустойчивом кластере, обмениваются информацией о конфигурации и heartbeat-сообщениями.

CR16-8 — коммутирующее лезвие, Core Blade.

Пара этих лезвий выполняет функцию матрицы коммутации для DCX 8510-8. На каждом CR16-8 установлено 4 ASIC'a Condor3, что в сумме на двух лезвиях дает 384 внутренних канала по 16 Гбит/с каждый. 256 каналов через backplane связаны с 8 слотами для установки лезвий с портами, обеспечивая 512 Гбит/с пропускной способности на каждый слот.

Оставшиеся 128 каналов используются для функционала Inter-Chassis Link (ICL): на каждом лезвии CR16-8 расположено 16 портов **UltraScale ICL**, аппаратно реализованных с помощью оптических QSFP трансиверов, пропускной способностью 64 Гбит/с каждый.

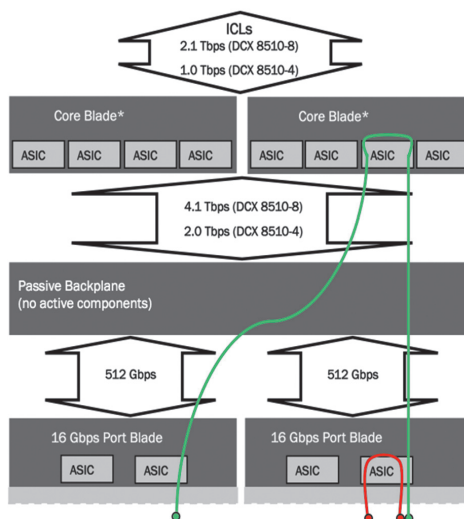
CR16-4 — модель Core Blade для 4-слотового коммутатора DCX 8510-4.

2 ASIC'a Condor3 обеспечивают 512 Гбит/с на слот, аналогично большому DCX 8510-8. Количество ICL портов — 8 на каждом лезвии.

Лезвия с портами

FC16-32 — лезвие с линейной производительностью; 32 порта 16 Гбит/с Fibre Channel. Отсутствие переподписки достигается за счет использования двух ASIC Condor3. По 16 каналов с каждого Condor3 выходят на SFP разъемы для подключения оптики (front-end порты), а еще по 16 каналов связаны через backplane с Core Blade (back-end порты). Отношение пропускной способности front-end и back-end портов 1:1 позволяет гарантировать отсутствие узких мест в архитектуре лезвия и линейную производительность.

FC16-48 — лезвие с переподпиской 1.5:1; 48 портов 16 Гбит/с Fibre Channel. Порты пары Condor3 распределены в пропорции 1.5:1 — 24 front-



*) CR16-8 blade (DCX 8510-8): четыре ASIC'a
CR16-4 blade (DCX 8510-4): два ASIC'a

Локальная коммутация (трафик между устройствами в одной группе портов не покидает ASIC)

Рис. 25. Движение трафика в коммутаторах DCX 8510.

end и 16 back-end. В отличие от 32-портового лезвия, архитектура FC16-48 предполагает переподписку, что в теории может привести к возникновению узких мест при высокой нагрузке, когда трафик на одном ASIC превысит пропускную способность back-end

портов. На практике с такими узкими местами помогает бороться механизм локальной коммутации (local switching). Его смысл заключается в том, что трафик между устройствами, подключенными к портам одного ASIC коммутируется локально, не задействуя back-end порты и не выходя за пределы лезвия. Схематичное изображение движения трафика внутри шасси DCX 8510 представлено на рис. 25. Таким образом, ситуация, в которой на лезвии FC16-48 могло бы возникнуть узкое место по пропускной способности является скорее теоретической и может сложиться разве что в лабораторной обстановке или при грубых ошибках в проектировании. Тем не менее, некоторые заказчики предпочитают свести к нулю риск и используют для подключения наиболее критичных к производительности систем лезвия без переподписки.

FC8-32E, FC8-48E — лезвия без переподписки на 32 и 48 портов 8 Гбит/с Fibre Channel.

Данные лезвия используют 16 Гбит/с ASIC Condor3, но его front-end порты программно ограничены на работу со скоростью 8 Гбит/с. При этом back-end порты работают на скорости 16 Гбит/с, что позволяет избежать переподписки даже на 48-портовом лезвии. На этих лезвиях недоступна часть функций Condor3 — D_port, сжатие, шифрование, 10 Гбит/с FC, FEC. Также не предусмотрено лицензий, с помощью которых можно было бы включить функционал 16 Гбит/с.

FC8-64 — лезвие с переподпиской 2:1 на 64 порта 8 Гбит/с Fibre Channel.

Это лезвие используется в том случае, когда необходимо получить максимальное количество портов на шасси DCX. Полностью заполненный такими лезвиями DCX 8510-8 позволяет получить в одном шасси 512 портов, а DCX 8510-4 — 256 портов. Отличительной особенностью данного лезвия является использование уменьшенных в габаритах SFP трансиверов — mSFP. Это необходимо учитывать при планировании кабельной системы, так как не все кабели с разъемом LC можно включить в mSFP, зачастую необходимо приобретение специальных кабелей или гидр-разветвителей. Кроме того, лезвие создано на основе ASIC Condor2, поэтому максимальная скорость back-end портов ограничена 8 Гбит/с, и это лезвие не может полностью использовать пропускную способность слота DCX 8510. Как следствие, FC8-64 работает с переподпиской 2:1.

FC16-64 — лезвие с переподпиской 2:1 на 64 порта 16 Гбит/с Fibre Channel (рис. 26).

Это лезвие можно использовать, когда необходимо получить максимальное количество портов на шасси DCX. Полностью заполненный такими лезвиями DCX8510-8 позволяет получить в одном шасси 512 портов 16 Гбит/с, а DCX8510-4 — 256 портов (рис. 27).

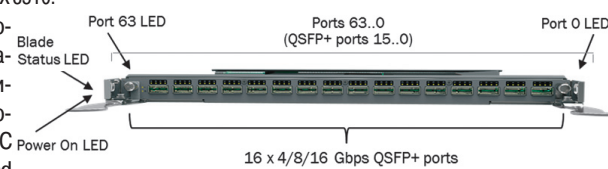


Рис. 26. Лезвие FC16-64.

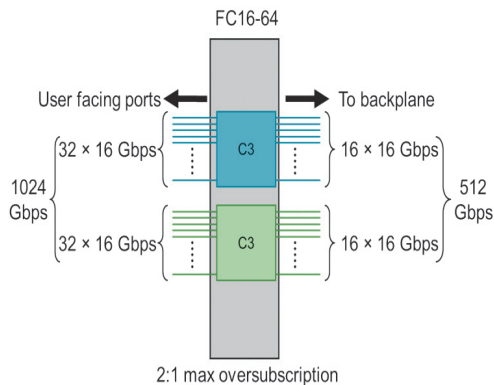


Рис. 27. Внутренняя коммутация в лезвии FX16-64.

В линейке продуктов Brocade есть подобное лезвие – с 64 портами, работающими на скорости 8 Гбит/с. Отличительной особенностью этого лезвия является то, что оно использует уменьшенные в габаритах SFP-трансиверы (mSFP). Использование подобных трансиверов требует также специальных оптических кабелей, так как mSFP-трансиверы имеют нестандартный LC-разъем. Причина подобного решения – невозможность разместить на лицевой панели 64 стандартных трансивера SFP. Все эти требования приводят к усложнению кабельной инфраструктуры и удорожанию общего решения.

Новое лезвие лишено подобных недостатков. Вместо использования большого числа mSFP портов на лезвии размещено 16 портов стандарта QSFP+ (аналогичных портам ICL на коммунитирующих лезвиях Core Blade). Каждый такой порт представляет собой фактически четыре отдельных порта 16 Гбит/с. Для соединения QSFP-портов со стандартными портами SFP есть различные варианты кабелей:

- QSFP – разветвитель на 4 разъема LC;
- QSFP – разъем MPO для патч-панели;
- QSFP – QSFP.

Порты QSFP могут работать на стандартных скоростях 16/8/4 Гбит/с и использоваться как для подключения конечных устройств, так и для соединения коммутаторов между собой

Специальные лезвия

FX8-24 – лезвие для FCIP маршрутизации. 12 портов 8 Гбит/с FC, 10 портов 1GbE, 2 порта 10GbE.

Лезвие предназначено для связи дата-центров с помощью технологии FCIP. Для этого могут использоваться 10 портов гигабитного Ethernet или с дополнительной лицензией 2 порта 10-гигабитного Ethernet. Аппаратно реализуется сжатие трафика, а также шифрование IPsec.

FCOE10-24 – лезвие с 24 портами 10 Гбит/с FCOE.

Используется для подключения конвергентных сетей и устройств по протоколу FCoE. Лезвие не является новым. Ранее оно было доступно для установки в директоры предыдущего поколения (Gen 4) DCX. Теперь стало доступным использование этого лезвия в современных директорах DCX8510-8.

Коммутаторы

Коммутатор Brocade 6520

Коммутатор Brocade 6520 имеет высокую плотность FC-портов: в форм-факторе 2U вмещает-

ся 96 портов FC 16 Гбит/с без переподписки (рис. 28). Схема лицензирования – 48 портов в базовой модели и возможность два раза использовать лицензию Ports on Demand по 24 порта. Основу коммутатора составляют шесть ASIC Condor3 – четыре на front-end и два – на back-end. За счет использования четырех ASIC на front-end, коммутатор поддерживает шифрование и сжатие трафика объемом до 128 Гбит/с. Также поддерживается FC-маршрутизация и связь на

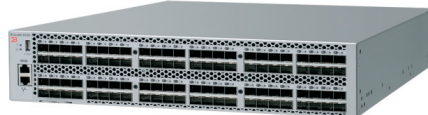


Рис. 28. Коммутатор Brocade 6520.

дальнем расстоянии (4256 буферных кредитов на каждые 48 портов). Поддерживается направление охлаждения с выдувом со стороны портов (PSE, port-side exhaust) либо со стороны блоков питания (NPSE, non port-side exhaust).

Коммутатор Brocade 6510

Коммутатор Brocade 6510 пришел на смену модели Brocade 5100 – в 1U теперь вмещается 48 портов 16 Гбит/с, суммарной производительностью 768 Гбит/с без переподписки (рис. 29). Гибкая схема лицензирования позволяет постепенно увеличивать количество активных портов с 24 до 48 с шагом в 12 портов. Brocade 6510 спроектирован с использованием одного ASIC Condor3,



Рис. 29. Коммутатор Brocade 6510.

и как следствие работает с минимальной задержкой (700 нс), поддерживает использование до 8000 буферных кредитов, режим Access Gateway, FC-маршрутизацию, FICON, и совместим с предыдущими поколениями FC коммутаторов Brocade. Энергопотребление – самое низкое на рынке – 0.14 Вт/Гбит/с. Поддерживается направление охлаждения с выдувом со стороны портов (PSE, port-side exhaust) либо со стороны блоков питания (NPSE, non port-side exhaust).

Коммутатор Brocade 6505

Коммутатор Brocade начального уровня – Brocade 6505 – это простая, и в то же время эффективная платформа для создания SAN. С помощью лицензий количество доступных портов можно увеличить с 12 до 24 (рис. 30). Поддерживаются скорости 2, 4, 8 и 16 Гбит/с. В базовой поставке один блок питания с интегриро-



Рис. 30. Коммутатор Brocade 6505.

ванными вентиляторами, но для повышения надежности можно заказать второй блок питания. Коммутатор обладает высокой производительностью, работает без переподписки и поддерживает использование до 8000 буферных кредитов.

Коммутатор Brocade 7840 Extension switch

Коммутатор Brocade 7840 Extension Switch предназначен для соединения инфраструктуры хра-

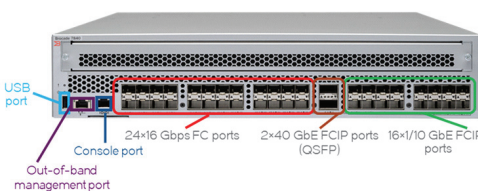


Рис. 31. Коммутатор Brocade 7840 Extension switch.

нения данных центров обработки данных на расстояниях до десятков тысяч километров по сетям WAN (рис. 31).

Функционал коммутаторов пятого поколения

Развитие облачных технологий предъявляет все больше требований к связи между удаленными центрами обработки данных. Для еще большего упрощения и удешевления передачи трафика СХД между площадками, в новых продуктах Brocade появился ряд функциональных нововведений.

10 Гбит/с Fibre Channel

Для максимального использования оптического волокна между площадками заказчики часто используют аппаратуру спектрального уплотнения каналов xWDM. Использование этих технологий позволяет по одному волокну передавать большое количество сигналов, сдвигая их частотные характеристики. При этом зачастую эти сигналы могут нести разную «смысловую» нагрузку – например, на одной длине волны может передаваться трафик СХД, на другой – Ethernet, на третьей – видео, и т.д. Несмотря на то, что поддержка 16 Гбит/с Fibre Channel заявлена многими производителями xWDM оборудования, большинство уже инсталлированных устройств работает на скоростях до 10 Гбит/с и умеет передавать трафик Fibre Channel на скоростях 4 и 8 Гбит/с. Имея в эксплуатации дорогую xWDM платформу, у заказчика возникает проблема – либо понижать скорости между FC коммутаторами до 8 и 4 Гбит/с, либо проводить дорогой апгрейд xWDM до 16 Гбит/с.

Brocade помогает защитить инвестиции в имеющееся оборудование, давая возможность включать порты на 16-гигабитных коммутаторах в режим 10 Гбит/с. Большинство xWDM поддерживают передачу на этой скорости, а линейные карты 10 Гбит/с обычно конструктивно проще и дешевле, чем 8- или 4-гигабитные.

Порты лезвий FC16-xx DCX 8510 или Brocade 65xx могут работать в режиме 10 Гбит/с FC – достаточно поставить лицензию и установить 10 Гбит/с SFP.

Сжатие и шифрование трафика ISL

Помимо этого, для более эффективного использования канала между площадками в новый ASIC добавлена функция сжатия трафика FC по алгоритму McLZO (алгоритм с максимальной скоростью распаковки, коэффициент сжатия до 2:1).

Для защиты данных, передаваемых между площадками, можно использовать встроенные в ASIC функции кодирования, которые шифруют данные, передаваемые между 16 Гбит/с коммутаторами с помощью 256-битного ключа.

Функции сжатия и шифрования канала входят в базовую поставку Fabric OS и являются бес-

платными. Производительность аппаратных комплексов шифрования и компрессии — по 32 Гбит/с на каждый ASIC. Шифрование и сжатие трафика недоступно на модели начального уровня 6505.

Видео-пример настройки сжатия и шифрования можно посмотреть на YouTube: <http://youtu.be/nEp4y0Da6Xc>.

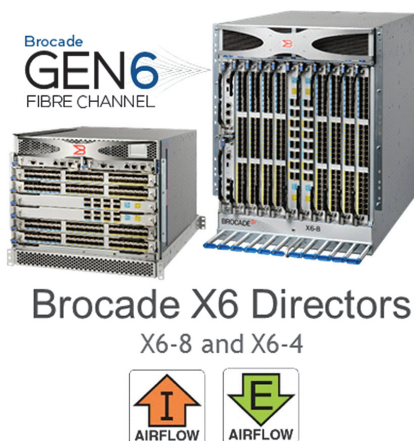
Коммутаторы Brocade Fibre Channel шестого поколения

В 2016 году компанией Brocade были выпущены первые продукты шестого поколения Fibre Channel. Отличительные особенности шестого поколения коммутаторов:

- **новые скорости работы.** Стандартная скорость порта FC составляет 32 Гбит/с. При этом обеспечивается обратная совместимость с оборудованием предыдущих поколений. Порты коммутаторов могут работать на скорости 4/8/10/16/32/128 Гбит/с. Возможно подключение любого имеющегося оборудования;
- **впервые появилась возможность использования параллельного протокола Fibre Channel.** Специализированные Q-Flex порты обеспечивают скорость работы 128 Гбит/с (на первом этапе — как 4x32 Гбит/с). Это дает возможность подключения современного быстродействующего оборудования;
- **поддержка протокола NVMe over Fabrics.** Возможность включения в сеть Fibre Channel современных быстродействующих СХД с интерфейсами NVMe;
- **IO Insight** — глубокий анализ производительности систем хранения данных с использованием технологий Fabric Vision;
- **VM Insight** — расширение возможностей мониторинга, включенных в пакет Fabric Vision, до уровня отдельных виртуальных машин.

Директоры Brocade X6

Первое в отрасли семейство директоров Fibre Channel 6-го поколения обеспечивает непревзойденную производительность приложений, уровень надежности по стандартам центров обработки данных и высокую динамичность, что позволяет ускорить доступ к данным, адаптироваться к меняющимся требованиям бизнеса и гарантировать бесперебойное функционирование.



Brocade X6 Directors
X6-8 and X6-4

Рис. 32. Brocade X6-8 и X6-4 Directors.

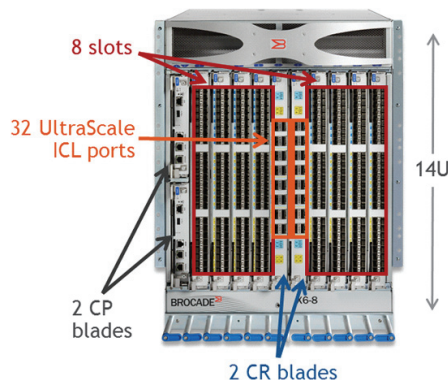


Рис. 33. Brocade X6-8 Director.

ние бизнес-процессов. Директоры семейства X6 поддерживают до 384 портов 32 Гбит/с, работающих одновременно на полной скорости (без переподписки) и до 32 портов UltraScale ICL (Inter-chassis links, коммутация между шасси директоров) 128 Гбит/с, благодаря чему суммарная пропускная способность системы достигает 16 Тбит/с. Благодаря пропускной способности 32 Гбит/с на 72% сокращается время отклика приложений, устраняются узкие места в подсистемах ввода-вывода данных и полностью раскрывается потенциал флэш-накопителей и накопителей следующего поколения на базе протокола NVMe.

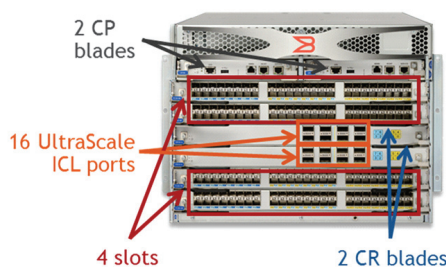


Рис. 34. Brocade X6-4 Director.

Аппаратно коммутаторы реализованы в виде модульных шасси размером 14U (X6-8) и 9U (X6-4). В шасси расположены блоки питания и вентиляторы с возможностью горячей замены, пассивный backplane, 2 управляющих лезвия (Control Processor Blade, CP) в отказоустойчивой конфигурации и два коммутающих лезвия (Core Blade, CR). Все модули директоров зарезервированы, включая управляющие и коммутающие лезвия, блоки питания, вентиляторы охлаждения и платы, хранящие информацию о шасси.

Коммутаторы поддерживают выбор направления обдува охлаждающих вентиляторов. Возможно направление выхода воздуха как в сторону коммутационных портов, так и в сторону блоков электропитания. Направление обдува выбирается на этапе заказа оборудования и не может быть изменено впоследствии (рис. 32, 33, 34).

Основной функционал коммутатора реализован в лезвиях, устанавливаемых в шасси.

Служебные лезвия, обязательно присутствующие в шасси

Brocade X6 Control Processor — управляющее лезвие, Control Processor.

На этом лезвии находится образ операционной системы Fabric OS, процессоры и память для ее работы. Там же расположены сетевые и кон-

ольные управляющие интерфейсы. Через backplane два лезвия Brocade X6 Control Processor, работающие в отказоустойчивом кластере, обмениваются информацией о конфигурации и heartbeat-сообщениями.

Brocade CRX-8 — коммутающее лезвие, Core Blade для 8-слотового коммутатора X6-8 Director.

Пара этих лезвий выполняет функцию матрицы коммутации для X6-8 Director. На каждом CRX-8 установлено 4 ASIC'a Condor4, что в сумме на двух лезвиях дает 512 внутренних каналов по 32 Гбит/с каждый. 384 канала через backplane связаны с 8 слотами для установки лезвий с портами, обеспечивая 1,5 Тбит/с пропускной способности на каждый слот.

Оставшиеся 128 каналов используются для функционала Inter-Chassis Link (ICL): на каждом лезвии CRX-8 расположено 16 портов **UltraScale ICL**, аппаратно реализованных с помощью оптических QSFP трансиверов, с пропускной способностью 128 Гбит/с каждый.

Brocade CRX-4 — коммутающее лезвие, Core Blade для 4-слотового коммутатора X6-4 Director.

Пара этих лезвий выполняет функцию матрицы коммутации для X6-4 Director. На каждом CRX-4 установлено 2 ASIC'a Condor4, что в сумме на двух лезвиях дает 256 внутренних каналов по 32 Гбит/с каждый. 192 канала через backplane связаны с 4 слотами для установки лезвий с портами, обеспечивая 1,5 Тбит/с пропускной способности на каждый слот.

Оставшиеся 64 канала используются для функционала Inter-Chassis Link (ICL): на каждом лезвии CRX-4 расположено 8 портов **UltraScale ICL**, аппаратно реализованных с помощью оптических QSFP трансиверов, с пропускной способностью 128 Гбит/с каждый.

Оптические ICL

Одной из важнейших особенностей коммутаторов Brocade X6 Director является использование специальных оптических соединений для связи между собой нескольких шасси. Эти соединения называются UltraScale Inter-Chassis Link (ICL), и они присутствовали на предыдущих поколениях DCX8510 и DCX в виде оптических кабелей (DCX8510) и многожильных медных кабелей длиной 2м (DCX).

С помощью этих кабелей можно было соединить до 3 шасси DCX между собой (рис. 35), сэкономив большое количество оптических портов для подключения серверов и дисковых массивов (пропускная способность ICL — 512 Гбит/с, что позволяло сэкономить до 64 портов на каждом DCX).



Рис. 35. Медные ICL на 8-гигабитных DCX.

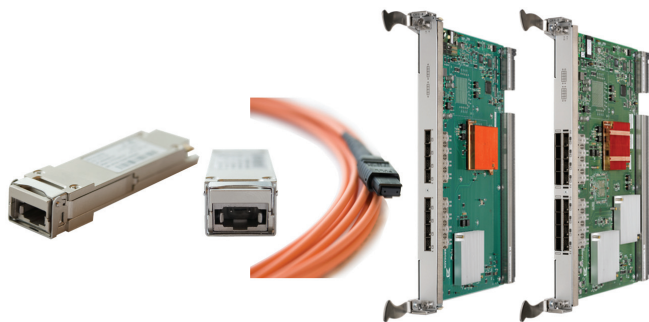


Рис. 36. Оптические модули QSFP, кабель с MPO разъемом и коммутационные лезвия CR16-4 и CR16-8.

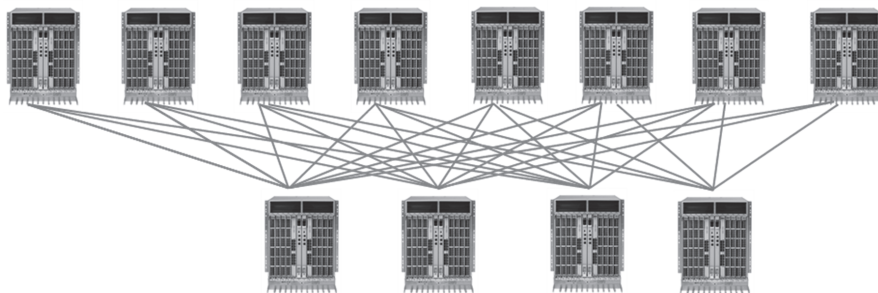


Рис. 37. Соединение до 12 коммутаторов с помощью оптических ICL Links.

На DCX8510 ICL сделаны еще более производительными, а также намного более удобными для использования. Вместо медных кабелей используются стандартные многожильные оптические кабели длиной до 100 м с MPO разъемами и трансиверы QSFP (quad-SFP) – рис. 36. Общая пропускная полоса ICL для 8-слотового шасси выросла до 2 Тбит/с (32 ICL порта по 64 Гбит/с), и 1 Тбит/с – для 4-слотового (16 ICL портов), и соединять теперь можно до 12 шасси DCX 8510 (рис. 37). Использование оптических ICL позволяет заказчику сэкономить до 128 портов на каждое шасси DCX8518.

Возможность соединять между собой до 12 шасси позволяет сделать сети хранения данных еще более плоскими, упрощая задачу проектировщиков и администраторов, делая при этом более предсказуемой производительность SAN инфраструктуры.

Кроме того, увеличилась возможная длина оптических соединений ICL. Если ранее длина кабеля ICL составляла не более 100 метров, что требовало размещения всех шасси в одном серверном помещении либо в смежных помещениях, то теперь длина соединения увеличилась до 2 км, что позволяет размещать коммутаторы в различных частях одного здания или в разных зданиях, находящихся на значительном расстоянии друг от друга.

Для такого соединения используются специальные активные одномодовые (Single mode) кабели с соединениями типа MTP – MTP или MTP – patch panel (рис. 38).



Рис. 38. Одномодовые (Single mode) кабели с соединениями типа MTP – MTP.

Использование таких кабелей позволяет задействовать имеющуюся кабельную инфраструктуру с помощью подключения через патч-панели.

Для коммутаторов шестого поколения возможности использования ICL были значительно расширены. Прежде всего, благодаря использованию новых скоростей 32 Гбит/с производительность соединений удвоилась. Коммутатор X6-8 Director имеет 32 специализированных порта ICL, каждый из них имеет пропускную способность 128 Гбит/с (Используются стандартные QSFP модули – 4 порта по 32 Гбит/с). Коммутатор X6-4 Director имеет 16 специализированных портов ICL, каждый из них имеет пропускную способность 128 Гбит/с. Общая пропускная полоса ICL для 8-слотового шасси выросла до 4 Тбит/с (32 ICL порта по 128 Гбит/с), и 2 Тбит/с – для 4-слотового (16 ICL портов), и соединять теперь можно до 12 шасси – так же, как и для DCX 8510. Коммутаторы X6 Director поддерживают подключение через ICL порты к коммутаторам предыдущего поколения DCX 8510, обеспечивая таким образом совместимость с установленными ранее системами.

Лезвия с портами

FC32-48 – лезвие с линейной производительностью; 48 портов 32 Гбит/с Fibre Channel. Отсутствие переподписки достигается за счет использования двух ASIC Condor4. По 24 каналов с каждого Condor4 выходят на SFP разъемы для подключения оптики (front-end порты), а еще по 24 канала связаны через backplane с Core Blade (back-end порты). Отношение пропускной способности front-end и back-end портов 1:1 позволяет гарантировать отсутствие узких мест



48x32 Gbps ports (SFP+)

Рис. 39. Коммутирующее лезвие FC32-48.

в архитектуре лезвия и линейную производительность (рис. 39).

SX6 Extension blade – лезвие для FCIP маршрутизации, а также для репликации систем хранения данных по протоколу IP. 16 портов 32 Гбит/с FC, 16 портов 1/10 GbE, 2 порта 40 GbE (рис. 40).



Рис. 40. Лезвие SX6 Extension blade.

Лезвие предназначено для связи дата-центров с помощью технологии FCIP. Для этого могут использоваться 16 портов 10-гигабитного Ethernet или 2 порта 40-гигабитного Ethernet. Аппаратно реализуется сжатие трафика, а также шифрование IPsec. Подробно о возможностях лезвия можно прочитать в разделе «Связь центров обработки данных на расстоянии».

Коммутатор Brocade G620

Коммутатор Brocade G620 имеет наиболее высокую плотность FC-портов на рынке: в форм-факторе 1U уместается 64 порта FC 32 Гбит/с, суммарной производительностью 2048 Гбит/с без переподписки (рис. 41). Впервые доступны также параллельные порты Fibre Channel со скоростью работы 128 Гбит/с. Общая структура доступных для подключения портов выглядит следующим образом: 48 портов FC 32 Гбит/с плюс четыре Q-Flex порта с максимальной скоростью работы 128 Гбит/с. Порты Q-Flex представляют из себя стандартные счетверенные разъемы, предназначенные для установки оптических модулей QSFP. Порты Q-Flex можно использовать различным образом. В настоящее время они могут использоваться только для соединения коммутаторов G620 между собой. В дальнейшем планируется возможность различных методов подключения: разделение каждого Q-Flex порта на четыре стандартных порта FC 32 Гбит/с, использование порта Q-Flex как единого порта FC на скорости 128 Гбит/с для подключения конечных устройств, а также соединение с QSFP-портами на лезвиях FC16-64 для директоров DCX8510 для формирования гибкой архитектуры подключения core-edge. Гибкая схема лицензирования позволяет постепенно увеличивать количество активных портов с 24 до 48 с шагом в 12 портов. Кроме того, отдельно доступны для лицензирования четыре Q-Flex порта. Brocade G620 спроектирован с использованием одного ASIC Condor4, и как следствие работает с минимальной задержкой 700 нс (900 нс при включенном функционале FEC), поддерживает использование до 15360 буферных кредитов, режим Access Gateway, FC-маршрутизацию, FICON, и совместим

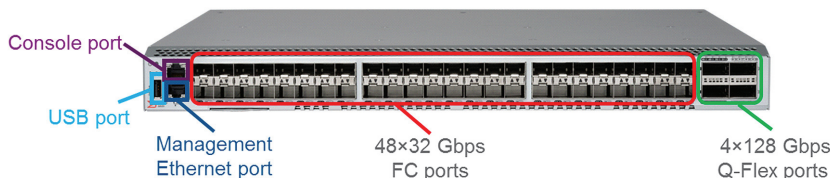


Рис. 41. Brocade G620.

с предыдущими поколениями FC коммутаторов Brocade. Энергопотребление — самое низкое на рынке — 0,10 Вт/Гбит/с.

Коммутатор поддерживает выбор направления обдува охлаждающих вентиляторов. Возможно направление выхода воздуха как в сторону коммутационных портов, так и в сторону блоков электропитания. Направление обдува выбирается на этапе заказа оборудования и не может быть изменено впоследствии.

Оборудование Brocade для IP Extension

Сегодня ИТ-отделы находятся под давлением растущих объемов трафика данных между центрами обработки данных и изменений внутри виртуализированных рабочих нагрузок в пределах окружения Fibre Channel и среды хранения IP. Также они сталкиваются с ростом требований к «соглашениям об уровне обслуживания» (SLA) и высоким ожиданиям от центров обработки данных в области аварийного восстановления инфраструктуры для обеспечения быстрого, непрерывного доступа и репликации критически важных данных в любой точке мира. Администраторам систем хранения необходимо обеспечить копирование больших объемов данных быстро, надежно и просто при минимизации капитальных затрат и операционные расходов.

WAN-соединениям присущи постоянные проблемы: большие задержки, потеря пакетов, ограниченная пропускная способность. Для WAN-соединений характерны частые сбои, число которых растет каждый год. Характерен большой круг потенциально возможных ошибок: полная потеря связи, деградация соединения, асимметричные условия передачи в линии. Такие сети состоят из большого количества соединений, образованных оборудованием различных поставщиков.

Репликация данных по каналам IP заставляет взглянуть на задачу копирования данных по-новому:

- использование выделенных виртуальных каналов позволяет выделить требуемую пропускную способность и изолировать пользовательский трафик от трафика репликации;
- консолидация данных репликации от нескольких СХД;
- максимально эффективное использование транспортной сети и резервирование путей репликации;
- обеспечение отказоустойчивости решений с помощью объединения независимых сетей в единый пул репликации;
- защита данных при передаче по сети между ЦОД;
- полный контроль над прохождением трафика, мониторинг возникающих проблем.

Для задач IP extension компания Brocade предлагает оборудование, относящееся к различным поколениям FC.

Brocade 7800

Brocade 7800 — это FC-маршрутизатор форм-фактора 1U, относящийся к четвертому поколению Fibre Channel, на борту которого установлено 16 портов FC 8 Гбит/с и 6 портов 1GbE, предназначенных для передачи данных на большие расстояния по технологии FCIP (рис. 42).



Рис. 42. Коммутатор Brocade 7800.

Коммутатор Brocade 7800 позволяет создать быструю, надежную и недорогую сетевую инфраструктуру для осуществления удаленной репликации, резервного копирования и миграции данных. Brocade 7800 предоставляет высоконадежное и безопасное удаленное соединение благодаря таким инновационным функциям, как маршрутизация Fibre Channel (FCR), аппаратное сжатие, шифрование IPsec, а также оптимизация SCSI-операций чтения и записи. Коммутатор может поставляться в трех конфигурациях с различным количеством активных портов, что делает его идеально гибким решением для расширения SAN любого размера.

Brocade FX8-24

Brocade FX8-24 — лезвие, предназначенное для установки в директоры DCX и DCX8510. Относится к четвертому поколению Fibre Channel. Функционально аналогичен Brocade 7800. Основное отличие от Brocade 7800 — наличие двух портов 10 Гбит Ethernet. Возможна организация туннеля FCIP между продуктами Brocade 7800 и Brocade FX8-24.

Brocade 7840 Extension switch

Коммутатор Brocade 7840 Extension Switch (рис. 20) предназначен для соединения инфраструктуры хранения данных центров обработки данных на расстояниях до десятков тысяч километров по сетям WAN (рис. 43). Brocade 7840 является идеальной платформой для расширения инфраструктуры центра обработки данных, высокопроизводительной репликации и резервного копирования данных. Он использует любой тип WAN-транспорта для соединения открытых систем и мейнфреймов на любом расстоянии. Без использования подобных решений репликация данных на большие расстояния часто невозможна или нецелесообразна.

Основные функциональные особенности:

- перемещение больших объемов данных между центрами обработки данных, для удовлетворения растущих потребностей задач аварийного восстановления с высочайшим уровнем производительности и масштабируемости;

- балансировка нагрузки и повышение отказоустойчивости сетевой инфраструктуры с использованием функционала Extension Trunking для повышения эффективности использования WAN и защиты от отказов каналов WAN;
- консолидация Fibre Channel, FICON и IP-трафика от различных устройств для обеспечения высокой скорости, высокой доступности и безопасности обмена данными между центрами обработки данных;
- непрерывная работа оборудования при обновлении встроенного программного обеспечения;
- расширенный проактивный мониторинг для автоматического обнаружения аномалий WAN и исключения незапланированных простоев;
- защита канала шифрованием при передаче по линиям связи с полной линейной скоростью и без потери производительности;
- включает масштабируемость оборудования «по требованию» с расширением функционала с помощью электронных лицензий.

Brocade 7840 соответствует самым высоким требованиям систем аварийного восстановления. Коммутатор имеет двадцать четыре 16 Гбит/с FC/FICON® порта, шестнадцать портов 1/10 Gigabit Ethernet (GbE), и два порта GbE 40. Это обеспечивает плотность портов и пропускную способность, необходимую для максимальной производительности приложений при использовании сетей WAN.

Конструкция Brocade 7840 обеспечивает простую возможность масштабирования «по требованию» с помощью активации лицензий на программное обеспечение. Полоса пропускания WAN может быть установлена в широком диапазоне — от 5 Гбит/с до 40 Гбит/с. При включении сжатия данных организации могут масштабироваться до 80 Гбит/с пропускной способности приложений, в зависимости от типа данных и характеристик соединения.

Базовая конфигурация Brocade 7840 включает в себя широкий набор базовых сервисов: FCIP, IP Extension, Brocade Fabric Vision technology, Extension Trunking, Adaptive Rate Limiting, IPsec, compression, Open Systems Tape Pipelining (OSTP), Fast Write, Adaptive Networking, и Extended Fabrics. Кроме того, доступен ряд дополнительных лицензий для расширения базового функционала репликации как для открытых систем, так и для мейнфреймов: Integrated Routing (FCR), FICON Management Server (CUP), и Advanced FICON Accelerator.

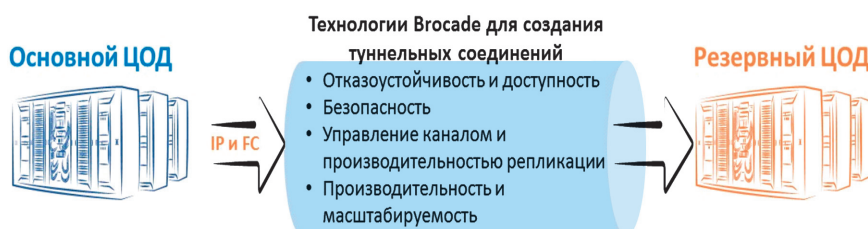


Рис. 43. IP Extension для репликации данных.

Поддерживая Время задержки (Round-TripTime, RTT) до 250 мс, Brocade 7840 позволяет строить решения расширения на расстояния до 37 500 км.

Brocade 7840 поддерживает набор технологий для ускорения репликации и резервного копирования и расширения пропускной способности канала на расстоянии, используя сжатие данных, ускорение дисковых операций и операций чтения/записи на ленты, WAN-оптимизированный протокол TCP, и другие. Расширенные функции и технологии включают в себя:

- **чрезвычайно эффективную аппаратную архитектуру** – позволяет производить обработку IP-датаграмм и Fibre Channel / FICON кадров с высокой скоростью и низкой латентностью, расширяя доступные дистанции для синхронной репликации;
- **WAN-optimized TCP** – агрессивный стек TCP, оптимизация размера окна TCP и управления потоком данных, ускорение транспорта TCP для приложений с высокими требованиями к пропускной способности;
- **Streams** – функционал WAN-optimized TCP, который используется для предотвращения блокировок (Head-of-LineBlocking, HoLB) при передаче данных по сетям WAN;
- **Per Priority TCP Quality of Service (PTQ)** – обеспечивает передачу данных с различными приоритетами (высокий, средний, низкий) для потоков данных Fibre Channel и IP Extension по одному туннелю в сети WAN. Использует индивидуальные сессии TCP для каждого уровня приоритета QoS;
- **Extension Trunking** – объединение нескольких соединений по сети WAN в единое логическое соединение (транк) с обеспечением активной балансировки нагрузки и защитой от аварийных ситуаций в канале;
- **Lossless Link Loss (LLL)** – часть функционала Extension Trunking, обеспечивает восстановление данных «на лету» в случае выхода из строя канала (линка). Со стороны приложения, работающего с данными, сбой даже не будет замечен, так как все данные будут доставлены в порядке очереди;
- **Adaptive Rate Limiting** – динамическая балансировка полосы пропускания между минимальным и максимальным значением для оптимизации использования доступных ресурсов и оптимального реагирования на сбой в сети WAN;
- **Advanced compression architecture** – возможность выбора режима сжатия данных для разных требований к пропускной способности;
- **FCIP Fast Write (FCIP-FW)** – ускорение обработки операций записи SCSI, увеличивает производительность синхронных и асинхронных задач репликации при работе по каналам WAN с высокими задержками и на больших расстояниях;
- **IPsec** – безопасная передача данных по сетям WAN путем шифрования канала «на лету» с аппаратной реализацией стандартного 256-разрядного алгоритма AES без потери производительности или добавления задержек в канале;

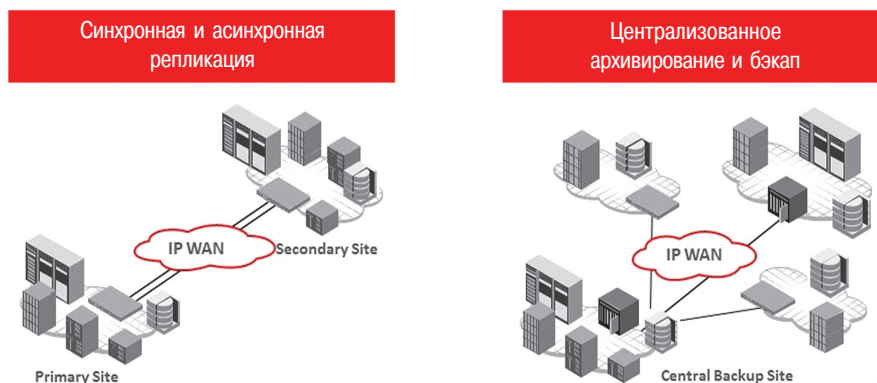


Рис. 44. Варианты использования IP Extension.

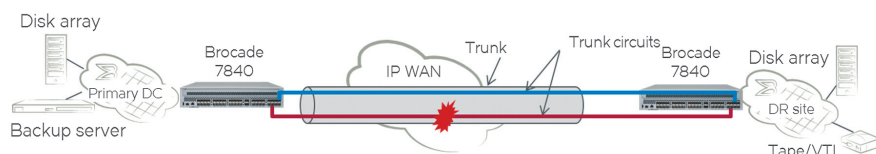


Рис. 45. Использование функционала Extension trunking.

- **Open Systems Tape Pipelining (OSTP)** – ускорение операций чтения и записи на ленточных накопителях при работе на больших расстояниях. Позволяет значительно снизить время резервного копирования и восстановления при использовании удаленных ленточных накопителей;
- **Advanced Accelerator for FICON** – использование передовых сетевых технологий, технологий управления данными и анализа существующих соединений для ускорения IBM zGM, операций чтения и записи на ленточных накопителях и подключения z/OS к аналитическим системам Teradata на расстоянии.

Коммутатор поддерживает выбор направления обдува охлаждающих вентиляторов. Возможно направление выхода воздуха как в сторону коммутационных портов, так и в сторону блоков электропитания. Направление обдува выбирается на этапе заказа оборудования и не может быть изменено впоследствии.

Варианты использования

Brocade® 7840 Extension switch позволяет строить эффективные решения для организаций, которым требуется инфраструктура для объединения сетей удаленных центров обработки данных и защиты данных, находящихся на СХД с различными типами доступа – блочным, файловым, объектным, а также на ленточных системах хранения, при использовании синхронной/асинхронной репликации, централизованного архивирования и бэкапа (рис. 44).

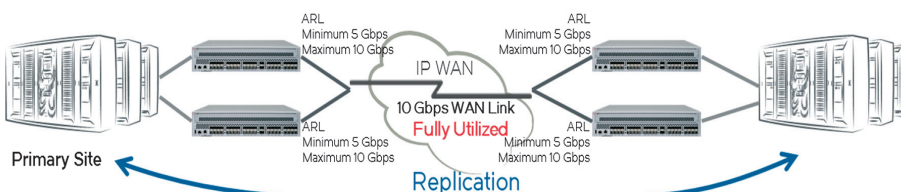


Рис. 46. Технология Adaptive Rate Limiting позволяет использовать всю возможную полосу пропускания канала на 100% при выходе из строя одного из пары коммутаторов.

1) Использование функционала Extension trunking Задача

Каждый раз, когда WAN-соединение рвется, пакеты данных, которые передавались в этой линии, теряются. Репликация данных останавливается, а в случае синхронной репликации – останавливается весь ввод-вывод приложения. Для многих приложений, управляющих репликацией данных, это означает возврат назад к моменту, предшествующему потере данных, и повторную передачу данных.

Решение

При использовании функционала Extended trunking потерянные пакеты данных автоматически передаются повторно по работоспособному соединению. Репликация данных продолжается без остановки (рис. 45).

2) Использование технологии Adaptive Rate Limiting Задача

Если один коммутатор из пары резервирования отказывает или не может работать (например, в результате сбоя по питанию), полоса пропускания WAN-соединения, зарезервированная за этим устройством, не будет использоваться. Таким образом, мы теряем 50% доступной полосы пропускания.

Решение

Используя технологию Adaptive Rate Limiting, парный коммутатор автоматически определяет, что второй коммутатор неработоспособен, и автоматически начинает задействовать всю возможную полосу пропускания канала на 100% от имеющихся ресурсов (рис. 46).



Рис. 47. Использование функционала MAPS и Fabric Vision позволяет автоматически определять узкие места и находить причину медленной работы в сети хранения данных, что ускоряет поиск ошибок и позволяет избежать внеплановых остановок.



Рис. 48. Использование технологии Flow Vision совместно со встроенным в Brocade7840 генератором трафика позволяет заранее протестировать необходимые характеристики канала и убедиться в его соответствии необходимым требованиям.

3) Использование технологии Fabric Vision

Задача

Организация использует решение, основанное на репликации данных, при этом фактическая скорость репликации меньше, чем требуется для нормальной работы приложения. При этом причина медленной работы непонятна. Если не удастся локализовать неисправность, приложение может утратить работоспособность.

Решение

Использование функционала MAPS и Fabric Vision позволяет автоматически определять узкие места и находить причину медленной работы в сети хранения данных. Это ускоряет поиск ошибок и позволяет избежать внеплановых остановок (рис. 47).

4) Использование технологии Flow Vision

Задача

Планируется внедрение нового решения по репликации данных. Необходимо проверить, что все настроено правильно и производительность канала соответствует требованиям по быстродействию и уровню ошибок в канале. В дальнейшем это поможет быстро запустить решение и избежать длительного процесса поиска неисправностей.

Решение

Использование технологии Flow Vision совместно со встроенным в Brocade7840 генератором трафика позволяет заранее протестировать необходимые характеристики канала и убедиться в его соответствии необходимым требованиям (рис. 48).

Результаты тестирования

Тестирование решения проводилось в ходе репликации данных по протоколу IP между двумя системами хранения. Расстояние между СХД – чуть больше 1000 км. Задержка в канале достигала 25 мс, частота потери пакетов в канале – 0,1%.

Полученные результаты тестирования:

- без использования нашего продукта Brocade 7840, используя передачу данных непосредственно между системами хранения: скорость передачи 6 Мбайт/с;
- при использовании Brocade 7840 для создания защищенного канала и ускорения передачи данных с помощью встроенного функционала сжатия данных и оптимизации характеристик канала: скорость передачи 320 Мбайт/с.

Табл. 4. Скорость репликации данных по протоколу IP (в Мбайт/с) в зависимости от задержек и объема потери пакетов с использованием (синим) Brocade 7840 и без него (красным).

Latency (ms) RTT	No packet drop	0.1% packet drop	0.5% packet drop
0	320 320	320 33	320 16
5	320 320	320 18	320 10
10	320 262	320 8	320 6
25	320 178	320 6	320 6
50	320 62	320 6	320 6
100	320 40	320 6	320 6
200	310 5	310 4	260 3
250	200 4	195 2	170 2

- – IP Storage + IP Extension
- – IP Storage native

Полные результаты тестирования приведены в табл. 4. Тестирование проводилось с использованием одного порта 10 Гбит/с Ethernet и одного набора реплицируемых данных со степенью сжатия 2:1.

Brocade SX6 Extension blade

Brocade SX6 Extension blade – лезвие для установки в коммутаторы X6 Director. Относится к шестому поколению Fibre Channel. Функционал, связанный с работой по протоколу IP, аналогичен коммутатору Brocade 7840 Extension switch. Возможно соединение по протоколу IP и организация туннеля между продуктами Brocade 7840 Extension switch и Brocade SX6 Extension blade.

Основное отличие от Brocade 7840 Extension switch – новое лезвие построено на аппаратной платформе Fibre Channel 6 поколения (Condor 4 ASIC). Это обеспечивает возможность работы портов Fibre Channel на скорости 32 Гбит/с и поддержке современных технологий мониторинга VM Insight и IO Insight.

Встроенные коммутаторы для шасси блейд-серверов

Для шасси блейд-серверов ведущих мировых производителей Brocade предлагает встраиваемые FC-коммутаторы, построенные на микросхемах Brocade (4, 8 и 16 Гбит/с). Эти комму-

таторы работают под управлением операционной системы Fabric OS и легко интегрируются в SAN, построенные на FC-коммутаторах Brocade любых поколений. Встраиваемые коммутаторы могут также работать в режиме Access Gateway, что позволяет подключать их к SAN других вендоров, и лучше масштабировать сеть, так как коммутаторы в режиме Access Gateway выполняют роль концентратора портов и не участвуют в работе служб SAN.

Кроме того, для некоторых моделей шасси блейд-серверов выпускаются конвергентные коммутаторы, поддерживающие Fibre Channel или DCB.

За дополнительной информацией по встраиваемым коммутаторам Brocade Вы можете обратиться к производителям блейд-решений (Dell, Fujitsu, HDS, HP, Huawei, IBM, Lenovo и др.).

Brocade Network Advisor

Brocade Network Advisor (BNA) – это основной программный продукт Brocade для управления сетевой инфраструктурой. BNA позволяет из одной точки управлять сетями Brocade FC SAN, Fibre Channel over Ethernet (FCoE), коммутаторами и маршрутизаторами IP Layer 2/3, сетями Layer 4/7 и Multiprotocol Label Switching (MPLS). BNA также интегрируется с ведущими партнерскими решениями, такими как VMware vCenter, Microsoft System Center Operations Manager, IBM Systems Director и IBM Tivoli.

Для SAN Brocade Network Advisor позволяет конфигурировать, управлять и отслеживать состояние всех коммутаторов и директоров Brocade, а также адаптеров HBA и CNA. Мониторинг производительности в реальном времени, хранение истории производительности SAN, автоматизированное обновление микрокода на коммутаторах, управление зонированием, настройка FCR, FCIP и FICON – это только часть функционала BNA для SAN (рис. 49).

Brocade Network Advisor поддерживает разделение пользовательских ролей (RBAC) как для SAN, так и для IP-части.

Brocade Network Advisor – это программное обеспечение, устанавливаемое на сервере или рабочей станции. Для его функционирования необходимо подключение к коммутаторам Brocade по протоколу IP.

ПО Brocade Network Advisor имеет несколько версий (табл. 5).

Бесплатная версия BNA Professional позволяет управлять двумя фабриками размером до 300 портов.

Версия BNA Professional Plus увеличивает этот лимит до 100 фабрик, 2560 портов и позволяет сохранять исторические данные о производительности. Эта версия платная.

Для самых крупных SAN нужно использовать BNA Enterprise – это платная версия Network Advisor, позволяющая одновременно управлять 100 фабриками суммарным количеством портов до 15000.

Brocade Network Advisor лицензируется по инсталляциям. Каждый экземпляр установленного ПО требует наличия лицензии (платной или бесплатной). При этом нет необходимости по-

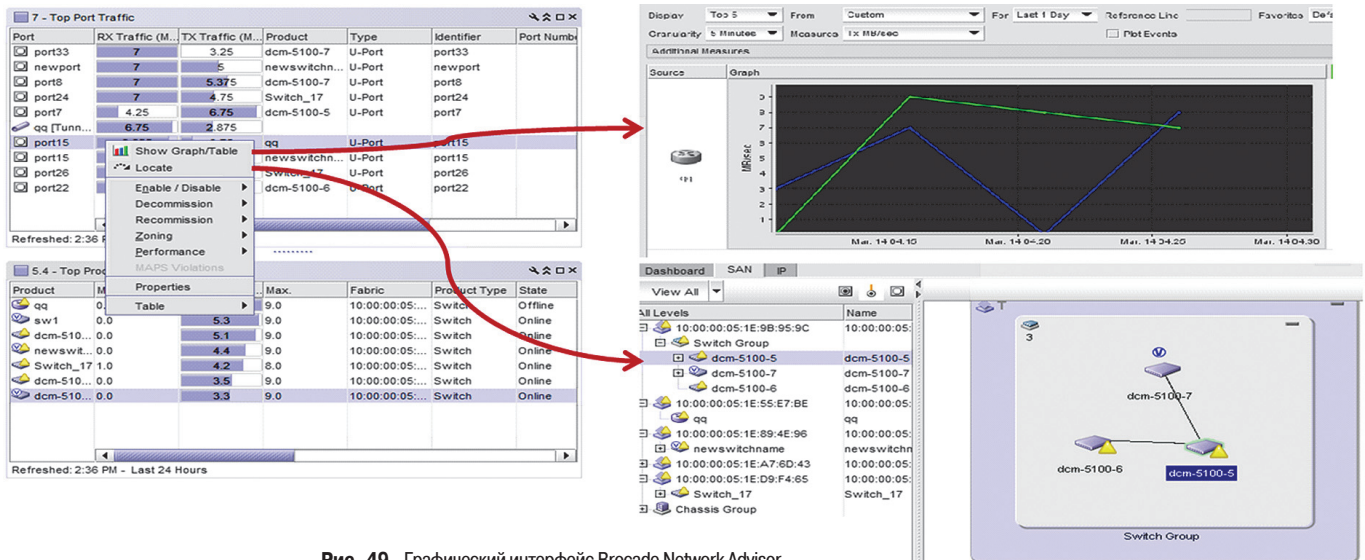


Рис. 49. Графический интерфейс Brocade Network Advisor.

Табл. 5. Версии Brocade Network Advisor.

SAN Professional (No License)	SAN Professional Plus	SAN Enterprise
Free download		
Up to 300 SAN switch ports and 2 SAN fabrics	Up to 2560 SAN switch ports (~40 devices) and 100 SAN fabrics	Up to 15,000 SAN switch ports (~400 devices) and 100 SAN fabrics
For full features list by product version and license type, refer to the Brocade Network Advisor Software Licensing Guide		

купать лицензии для устройств, которыми BNA может управлять.

Brocade Network Advisor можно загрузить с веб-сайта Brocade. Эта версия будет работать в полнофункциональном режиме Enterprise 120 дней, а затем потребует лицензию или перейдет в режим Professional. Заказчики, использующие DCFM, EFCM или Fabric Manager с действующей сервисной поддержкой, могут произвести миграцию на BNA.

Brocade Fabric Vision

Архитектура Fabric Vision

Brocade Fabric Vision – набор технологий и программных средств в сетях Fibre Channel, предназначенных для автоматизированного обнаружения и исправления ошибок, мониторинга текущего состояния фабрики, мониторинга производительности и анализа потоков данных в сетях SAN, а так же средств автоматизированного конфигурирования различных компонентов сетей (рис. 50).

Архитектура Fabric Vision базируется на трех составляющих: специализированных микросхемах (Condor 3 и Condor 4 ASIC), операционной системе Fabric OS и ПО управления и мониторинга Network Advisor. Главные задачи Fabric Vision – предотвращение аварий, ускорение восстановления, оптимизация производительности приложений и ускорение внедрения новых систем.

В состав Fabric Vision входят различные компоненты. Некоторые составляющие Fabric Vision появились уже давно и широко использовались пользователями коммутаторов Brocade: Bottleneck Detection, Forward Error Correction, Clearlink Diagnostics (D_port). Другая же часть – это новый функционал Monitoring and Alerting Policy

Suite (MAPS), Flow Vision, Fabric Performance Impact Monitoring (FPI), COMPASS. С выходом новых версий FOS функционал Fabric Vision постоянно дополняется новыми возможностями. Кроме того, возможности Fabric Vision зависят и от поколения коммутаторов. В шестом поколении Fibre Channel появился новый функционал – IO insight и VM Insight.

Monitoring and Alerting Policy Suite (MAPS)

MAPS – ключевой компонент архитектуры Fabric Vision. Эта технология представляет собой простой в управлении комплекс мониторинга и оповещения, который проактивно следит за производительностью и здоровьем инфраструктуры СХД.

Система мониторинга и оповещения, срабатывающая при превышении заданных пороговых значений, существовала в Fabric OS с 2000 года – это механизм Fabric Watch. Пользователь мог устанавливать пороговые значения на поведение различных элементов коммутатора или фаб-

рики, и пересечение этих пороговых значений приводило к отсылке уведомления или даже блокированию неисправного порта. Механизм Fabric Watch обладал достаточно богатым функционалом, но на практике его применение осложнялось трудной настройкой. Основная проблема заключалась в том, что пользователь зачастую не знал, какое пороговое значение для того или иного элемента было бы правильным для его фабрики. Если с компонентами коммутатора все довольно однозначно (выход модуля из строя приводит к отправке уведомления), то с другими элементами дело обстоит сложнее. Было непонятно с первого раза правильно угадать, например, какое значение счетчика “Invalid Transmission Words” или “Invalid CRC” должно приводить к блокированию порта или отсылке уведомления по SNMP. Как следствие, крайне полезный функционал Fabric Watch использовался очень редко, да и то в основном заказчиками, которые досконально знали поведение своих сетей хранения данных.

Технология MAPS призвана максимально упростить настройку мониторинга за счет использования шаблонов, созданных на основе накопленного опыта. Три ключевых понятия, которые используются при работе с MAPS, это – *группы, правила и политики*.

Группа – это набор похожих объектов, за которыми можно следить, используя одинаковые пороговые значения. Сгруппировав объекты таким образом и присвоив каждой группе свой набор правил, можно радикально упростить настройку мониторинга. Для наглядности приве-

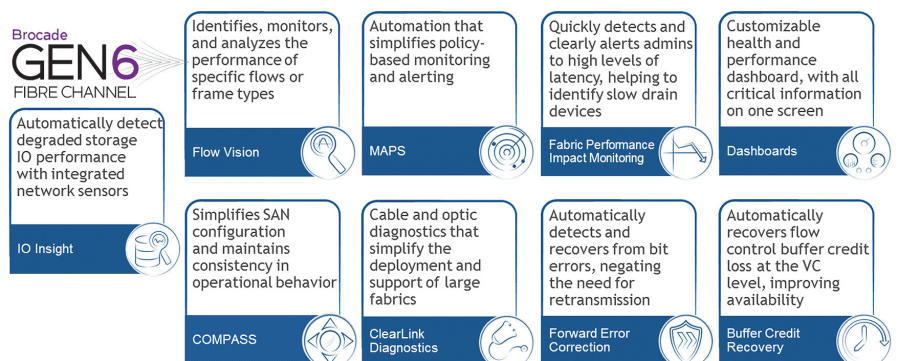


Рис. 50. Набор компонентов Fabric Vision.

дем пример некоторых групп, создаваемых по умолчанию:

- все E-порты (ALL_E_PORTS);
- все коротковолновые 16G SFP (ALL_16GSWL_SFP);
- все вентиляторы (ALL_FAN);
- все порты CXD (ALL_TARGET_PORTS).

Кроме групп по умолчанию, которых в FOS более 20, администратор имеет возможность создавать свои собственные группы, например: все порты UNIX-систем, SFP от конкретного производителя, наиболее критичные ISL-порты или линки FCIP и т.д.

Правила привязываются к группам. Правило – это *условие* и *действие*, которое совершается в случае, если условие выполнено. Условие содержит счетчик, пороговое значение и промежуток времени. Пример условия: “если счетчик CRC изменился более чем на 5 единиц за одну минуту”, “если загрузка порта (передатчика) выше 70%”, “если изменения в конфигурации зон происходят чаще, чем раз в две минуты”. Примеры действий, которые могут быть выполнены: запись в журнал (RASlog), уведомление по SNMP, уведомление по электронной почте, отключение порта.

Затем мы совмещаем группы и правила, например:

- “если на любом F-порту наблюдается больше 10 ошибок “C3 discard в минуту”, уведомить администратора, оставить запись в RASlog и заблокировать порт”;
- “если на любом длинноволновом 16G SFP уровень передаваемого сигнала ниже 2 dBm, перевести SFP в статус “Marginal”.

Совокупность всех правил и групп, которые могут одновременно действовать на коммутаторе, называется *политикой*.

Механизм *политик* обеспечивает простоту настроек и управления, которая является ключевым отличием MAPS в сравнении с Fabric Watch. При включении MAPS на коммутаторе появляются три политики по умолчанию: консервативная, средняя и агрессивная. Каждая из этих политик содержит более 190 правил, относящихся к группам по умолчанию. Все эти правила основаны на рекомендациях из лучших практик Brocade.

Консервативная политика включается по умолчанию и является наименее требовательной. Уведомления отправляются только в случае, когда накопилось достаточно большое количество ошибок.

Агрессивная политика – наоборот, очень чутко реагирует на малейшие неполадки в SAN, не ограничиваясь уведомлениями, а принимая действия вплоть до отключения портов с избыточным количеством ошибок.

В самом простом случае администратор может использовать на коммутаторе одну из политик по умолчанию, но можно пойти дальше. Механизм MAPS не позволяет изменять политики по умолчанию, но вместо этого дает возможность клонировать их, а получившиеся клоны уже дорабатывать по своему усмотрению, комбинируя правила из разных политик или создавая пользовательские группы и правила. Создан-

ные политики можно с помощью Network Advisor разослать на все коммутаторы в фабрике, либо экспортировать в XML, чтобы применить на другой площадке или поделиться с коллегами.

Если на коммутаторах уже имеется настроенная конфигурация Fabric Watch, ее можно нажатием одной кнопки преобразовать в новую политику, которую затем также можно клонировать и дорабатывать. А вот миграция в обратную сторону от MAPS к Fabric Watch не поддерживается – по мнению разработчиков, однажды включив MAPS, возвращаться обратно не захочет уже никто.

Flow Vision

Технология Flow Vision раскрывает потенциал микросхем Condor3 и Condor4, позволяя идентифицировать потоки данных, проходящих по фабрике, и собрать статистическую информацию по этим потокам. Собранная статистика позволяет лучше понимать поведение устройств, подключенных к SAN, быстрее реагировать на нештатные ситуации, проактивно избегать узких мест в фабрике и оптимизировать использование ресурсов.

Краеугольным камнем для понимания принципов работы Flow Vision является понятие *потока (flow)*. В терминологии Flow Vision потоком называется уникальный набор фреймов Fibre Channel, обладающих определенными свойствами. К параметрам, позволяющим выделить поток, относятся входящий или исходящий порт коммутатора, свойства или поля фрейма (например, SID, DID, WWN, номер LUN-а или тип фрейма) и направление потока. Примеры определенных потоков:

- все фреймы от хоста А к таргету В, выходящие из порта С на коммутаторе D;
- фреймы scsi3reserve от любых хостов к LUN0 таргета Е, входящие в порт F на коммутаторе G;
- все фреймы scsi3read, выходящие из порта H на коммутаторе I.

У Flow Vision есть три составляющие: *Flow Monitor*, *Flow Generator* и *Flow Mirror*. Задача Flow Monitor – обнаруживать и собирать статистику по потокам, проходящим через коммутатор. Собираются, в частности, данные по количеству отправленных и полученных фреймов, пропускной способности, количеству операций чтения и записи, IOPS и другие.

Flow Generator позволяет генерировать потоки данных с полной загрузкой в режиме 16 Гбит/с для тестирования фабрики перед введением в эксплуатацию. Для этого в фабрике создаются специальные SIM-порты, которые эмулируют настоящие устройства (серверы и дисковые массивы).

FlowMirror дает возможность создавать копию потока и перенаправлять его на управляющий процессор коммутатора для более глубокого анализа. Так, например, можно проанализировать поток данных от проблемного устройства (“slow-drain device”), чтобы понять, что именно вызывает замедление его работы.

Стоит отметить, что решения для анализа трафика от третьих производителей были доступны на рынке и раньше, но все они требовали вмешательства в работу сети для подключения к точкам мониторинга специальных оптических разветвителей (TAP, Traffic Access Point), что приводило к прерыванию трафика и ослаблению оптического сигнала. В случае, когда сеть хранения данных испытывает проблемы и необходимо найти их причину, вмешательство такого рода крайне нежелательно. Более того, отключение порта для установки разветвителя часто приводит к тому, что проблема на время пропадет из-за повторной инициализации соединения, что значительно усложняет диагностику, а в некоторых случаях делает ее невозможной.

Brocade Network Advisor Dashboards

Любые настройки MAPS и Flow Vision можно выполнять из командной строки, но намного удобнее – из Network Advisor: создавать политики мониторинга и рассылать их на все коммутаторы в фабрике, создавать определения потоков и SIM-порты, следить за обнаружением потоков данных на порту и за собранной статистикой – для всего этого намного лучше подходит графический интерфейс. Любые счетчики из набора правил MAPS или графики производительности по потокам можно вывести на главную вкладку BNA в виде виджета (рис. 51).

Бесплатно скачать полнофункциональную 120-дневную пробную версию Brocade Network Advisor можно на сайте brocade.com.

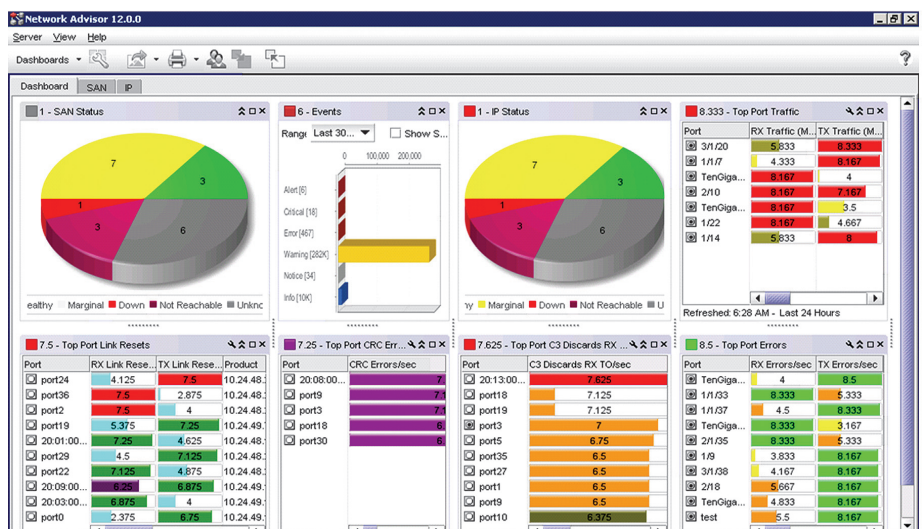


Рис. 51. Brocade Network Advisor Dashboards.

Slow Drain Device Quarantine (SDDQ)

Новая возможность Fabric Vision, появившаяся в версии 7.4 Fabric OS – функционал Slow Drain Device Quarantine (SDDQ, изоляция медленных устройств в фабрике). Функционал SDDQ является расширением возможностей другой функции – Fabric Performance Impact Monitoring (FPI), которая появилась ранее в версии 7.3 Fabric OS. SDDQ опирается в своей работе на FPI, упрощая конфигурирование и автоматизируя необходимые действия.

Почему возникла необходимость в подобном функционале? Как известно, для обеспечения гарантированной доставки пакетов в сети Fibre Channel используется механизм контроля потока (Flow control), основанный на использовании буферных кредитов (buffer-to-buffer credits, BB_credits). Все устройства Fibre Channel используют этот механизм. Он гарантирует, что пакеты не будут посылаться по сети до тех пор, пока отправитель не получит от получателя подтверждение о наличии свободных буферов для временного хранения пересылаемых пакетов. Механизм буферных кредитов существует с момента появления первых коммутаторов Fibre Channel, входит во все существующие стандарты и хорошо зарекомендовал себя с точки зрения эффективности контроля потока.

В то же время у механизма буферных кредитов есть узкое место. Если в сети SAN появляется медленное устройство (Slow Drain Device), которое не может принимать поток данных с той скоростью, с которой его пересылает отправитель, то потенциально оно может замкнуть на себя все имеющиеся буферные кредиты. Если

получатель и отправитель подключены к одному коммутатору SAN, подобное поведение никак не повлияет на другие устройства в сети. Если же получатель и отправитель подключены к разным коммутаторам сети SAN, связанным между собой межкоммутаторными соединениями (Inter switch links, ISL), потеря буферных кредитов может оказать влияние на весь поток данных, передающихся по этим соединениям, поскольку для него также не будет свободных кредитов (рис. 52).

Частично проблема была решена ранее с помощью разделения ISL на несколько виртуальных каналов (virtual channels, VC) с различным уровнем приоритета. При использовании VC медленное устройство будет оказывать влияние только на те данные, которые передаются в том же виртуальном канале. Но это решает проблему лишь частично, так как по умолчанию всем данным всех устройств в сети присваивается одинаковый приоритет. Ранее для изоляции медленных устройств необходимо было ручное вмешательство. Также необходимо было предварительно настроить систему мониторинга в сети таким образом, чтобы можно было оперативно обнаруживать медленные устройства и создаваемые ими проблемы.

Использование функционала Slow Drain Device Quarantine, реализованного в FOS, позволило автоматизировать процесс мониторинга и изоляции медленных устройств (рис. 53).

Теперь программное обеспечение коммутатора может автоматически обнаруживать медленные устройства и перемещать весь поток данных для этого устройства в виртуальный канал с низким приоритетом. Таким образом обеспечивается эффективная изоляция медленных устройств и их негативного влияния на передачу данных

в фабрике. Для работы SDDQ требуется наличие лицензии Fabric Vision на коммутаторах сети. Лицензия Fabric Vision необходима для активации функционала мониторинга на коммутаторах.

COMPASS

Новая возможность Fabric Vision, появившаяся в версии 7.4 Fabric OS – COMPASS (Configuration and Operational Monitoring Policy Automation Services Suite, рис. 54). Чем же вызвана необходимость этой утилиты конфигурирования и мониторинга сетевых устройств?

Для большой фабрики задача одинакового конфигурирования большого числа устройств достаточно сложна. Разные администраторы фабрики имеют, как правило, свои предпочтения по конфигурированию параметров. Со временем различные параметры на устройствах меняются в силу разных причин, и отследить эти изменения достаточно проблематично. Иногда необходимо менять определенные параметры коммутаторов для выполнения различных задач. Потом все забывают, что параметры менялись, и понять причину некорректной работы сети достаточно сложно.

COMPASS позволяет упростить выполнение двух важных задач:

- развертывание в сети большого числа устройств с автоматическим конфигурированием необходимых параметров коммутаторов и сервисов в фабрике;
- мониторинг состояния настроек и уведомление об отличии сконфигурированных параметров от их шаблонных значений. Уведомления о нарушении политик конфигурирования могут быть представлены в графическом виде в Brocade Network Advisor.

The problem

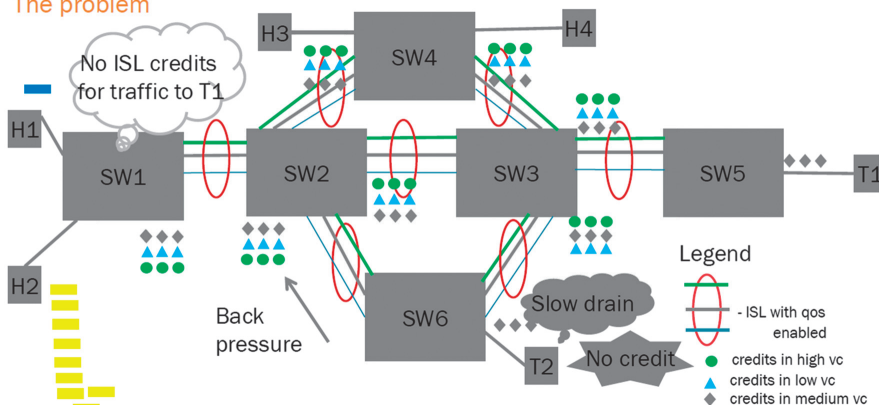


Рис. 52. Влияние Slow Drain Device на производительность сети SAN.

The solution – Auto VC Quarantine

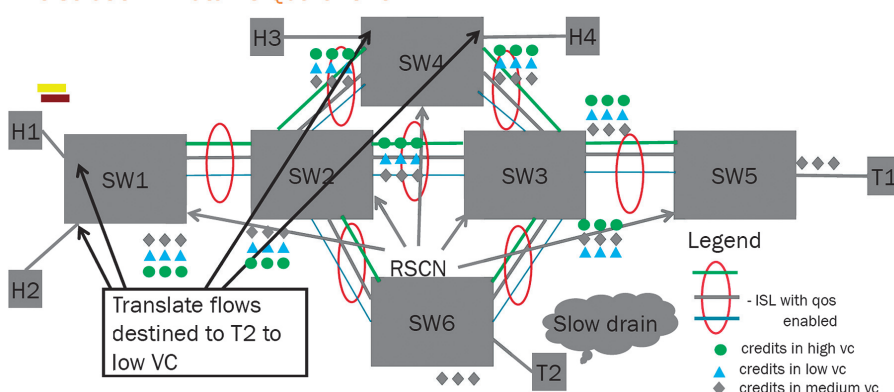


Рис. 53. Работа Slow Drain Device Quarantine.

Диагностический порт ClearLink

Для упрощения диагностики канала между коммутаторами, а также для тестирования канала перед введением в эксплуатацию, на новых устройствах появился новый тип порта – D_Port (диагностический). Включение порта в режим D_Port позволяет запустить интенсивное стресс-тестирование канала с насыщением его FC-пакетами, которое должно «отловить» возможные ошибки на кабеле, патч-панелях, в SFP или на самом ASIC. Кроме того, D_Port позволяет с точностью до нескольких метров померить общую длину канала и задержки при прохождении трафика. В режиме D_port можно перевести порты с установленными модулями SFP 16 Гбит/с или 10 Гбит/с. Также поддерживается диагностический порт между коммутаторами пятого поколения и некоторыми серверными адаптерами 16 Гбит/с производства Emulex и Qlogic. Сам по себе функционал ClearLink не требует дополнительного лицензирования, но для запуска диагностического порта между коммутаторами Brocade и HBA производства Emulex и Qlogic требуется лицензия Fabric Vision (рис. 55).

Forward Error Correction (FEC)

С ростом скорости передачи данных Fibre Channel становится все труднее поддерживать кабельную инфраструктуру в рабочем состоянии. Те оптические кабели, которые позволяли надежно передавать сигнал между устройствами

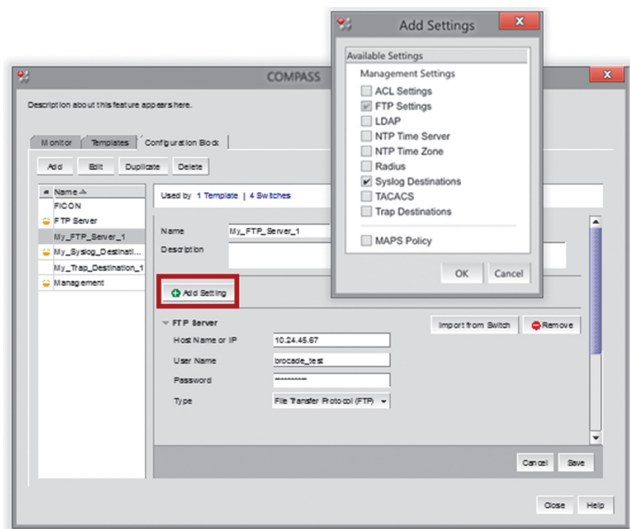


Рис. 54. Работа COMPASS.

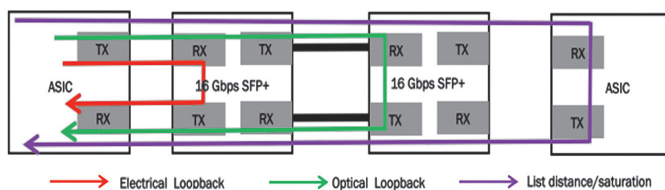


Рис. 55. Схема работы D_Port.

FC на скорости 8 Гбит/с и ниже, на более высоких скоростях часто неспособны обеспечить безошибочную передачу данных.

FEC позволяет упростить решение этой задачи, хотя поддержание кабельной инфраструктуры в рабочем состоянии — постоянный и непрерывный процесс в любой организации. FEC использует свободные биты внутри последовательностей сигнального протокола для генерации контрольного кода (Error Correcting Code, ECC), что позволяет обнаруживать и исправлять однократные ошибки. Использование FEC эквивалентно добавлению дополнительных 2.5 dB к бюджету канала (link loss budget) с точки зрения затухания передаваемого сигнала (рис. 56).

FEC работает на скорости передачи данных 10, 16, 32 Гбит/с и выше. Для скоростей 10 и 16 Гбит FEC является обязательной функцией, для скоростей 32 и 128 Гбит/с FEC включен постоянно. Все коммутаторы Brocade FC пятого и

шестого поколения поддерживают и используют FEC для передачи данных между собой. Кроме того, FEC может использоваться при обмене данными между коммутаторами Brocade и HBA производства Emulex и Qlogic.

IO Insight

Возможности технологий Fabric Vision превосходят традиционные ограничения средств сетевого мониторинга благодаря наличию IO Insight, первого в отрасли интегрированного сетевого датчика, который позволяет получить более подробное представление о производительности подсистем ввода-вывода инфраструктуры сетей хранения данных. Проблемы с производительностью приложений на уровне сервера и СХД могут оперативно выявляться и устраняться. IO Insight отслеживает в реальном времени производительность и поведение подсистем ввода-вывода, что позволяет получить исчерпывающие сведения о проблемах и

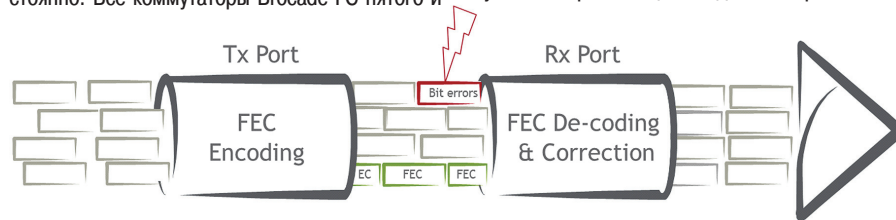


Рис. 56. Схема работы FEC.

Proactive IO monitoring:

- Total IOs
- First response max/average
- IO latency (ECT) max/average
- Outstanding IO max/average

Performance metrics provided based on data size

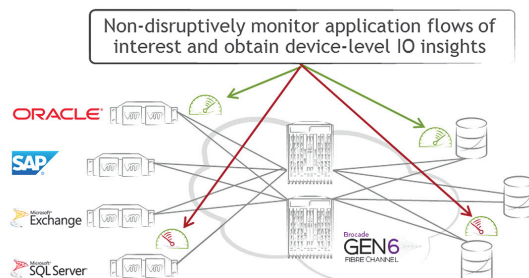


Рис. 57. IO Insight.

гарантировать соблюдение требуемых уровней обслуживания. Технология IO Insight впервые появилась в коммутаторах шестого поколения Fibre Channel, является их отличительной особенностью (рис. 57) и доступна уже сегодня.

Использование IO Insight расширяет возможности технологии Flow Vision. Flow Vision позволяет собирать статистику по потокам данных различных устройств и приложений, но не имеет средств углубленного анализа команд протокола SCSI, которые лежат в основе любого взаимодействия сервера и СХД. IO Insight собирает расширенную статистику по командам SCSI и является незаменимым инструментом производительности приложений и СХД.

VM Insight

Технология Fabric Vision поддерживает функцию VM Insight, которая позволяет получить комплексное представление о производительности подсистемы хранения данных отдельных виртуальных машин. Без этой информации трудно обеспечить оптимальную производительность и доступность виртуальных машин в виртуализованном центре обработки данных. В VM Insight используется маркировка виртуальных машин на основе отраслевых стандартов, что позволяет отслеживать проблемы с производительностью приложений на уровне отдельных виртуальных машин в сети хранения данных Brocade Fibre шестого поколения. Опираясь на эту информацию, администраторы СХД могут задать требуемый базовый уровень производительности приложений и оперативно выявлять аномалии, настраивая инфраструктуру в соответствии с требованиями конкретного уровня обслуживания. Эту информацию также можно оперативно сопоставить с другими показателями Fabric Vision, чтобы выявить первопричину проблем до того, как они окажут негативное влияние на бизнес-процессы. Функционал VM Insight является новым и требует поддержки со стороны производителей гипервизоров и СХД, поскольку именно на конечных устройствах происходит маркировка принадлежности пакетов данных различным виртуальным машинам (рис. 58). Ожидается, что технология VM Insight войдет в состав набора решений Brocade Fabric Vision в первой половине 2017 года.

Поддержка возможностей Fabric Vision в различных поколениях коммутаторов Brocade

Наиболее полно все возможности Fabric Vision представлены в шестом поколении коммутаторов

• Proactively monitor:

- Read/Write statistics
- IO per second
- Avg/max IO
- Total IO

• Latency metrics:

- First Response Time
- Exchange Completion Time
- Pending IO Time

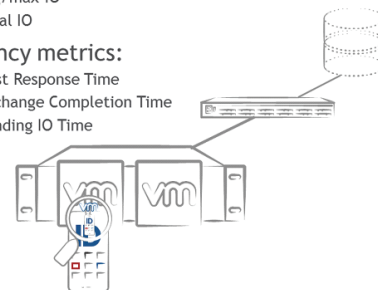


Рис. 58. VM Insight.

ров Fibre Channel. В предыдущих поколениях доступна только часть возможностей в связи с ограничениями программного и аппаратного обеспечения (табл. 6).

Расширение возможностей Fabric Vision на все устройства в сети

Ранее были рассмотрены многочисленные воз-

Табл. 6. Поддержка возможностей Fabric Vision в различных поколениях коммутаторов Brocade.

Features	8 Gbps Platforms	Gen 5 Platforms	Gen 6 Platforms
Forward Error Correction (FEC)	No	Yes	Yes
VC-level BB_Credit Recovery	No	Yes	Yes
Brocade ClearLink Diagnostics (D_Port)	No	Yes	Yes
MAPS	Yes	Yes	Yes
Fabric Performance Impact (FPI) Monitoring	Yes	Yes	Yes
Flow Learning	No	Yes	Yes
Flow Monitor	Yes, with some limitations	Yes	Yes
Flow Mirroring	No	Yes	Yes
Flow Generator	No	Yes	Yes
COMPASS	Yes	Yes	Yes
IO Insight	No	No	Yes

можности, которые открывают технологии Fabric Vision для мониторинга и управления сетью SAN. Но ранее все эти возможности работали только на коммутаторах Fibre Channel производства Brocade. Кроме коммутаторов, в сеть SAN входит большое число оконечных устройств, подключаемых к сети с помощью сетевых адаптеров (Host Bus Adapters, HBA). Совместная работа Brocade и основных производителей HBA, компаний Broadcom (Emulex) и Qlogic позволила расширить функционал Fabric Vision на всю фабрику, включая HBA (рис. 59).

Преимущества подобного подхода:

- возможность управления всеми устройствами в сети с помощью единого ПО управления – Brocade Network Advisor;
- быстрое развертывание новых устройств в фабрике и замена существующих с помощью автоматизации процедур конфигурирования устройств;

- управление производительностью в сети;
- повышение надёжности работы сети за счет возможности углубленного мониторинга сетевых устройств и применения автоматических методов исправления ошибочных ситуаций.

Для HBA доступен следующий функционал Fabric Vision:

- автоматизация развертывания новых устройств в сети с помощью функционала Fabric Assigned PWWN;
- автоматическое обнаружение загрузочных томов в сети SAN;
- диагностический порт (ClearLink Diagnostic);
- расширение возможностей качества обслуживания (QoS) до сетевых адаптеров и виртуальных машин;
- Forward Error Correction (FEC);
- Link Cable Beacon – возможность визуальной идентификации сетевых компонентов;
- Read Diagnostic Parameters – возможность сбора расширенной информации по конфигурированию и диагностике сетевых адаптеров (HBA) и трансиверов (SFP).

Analytics Monitoring Platform (AMP)

Дальнейшим расширением функционала Flow Vision и IO Insight является специализированное устройство мониторинга производительности сетей SAN, которое называется Analytics Monitoring Platform (AMP). Устройство AMP было представлено компанией Brocade в 2015 году. В августе 2016 года вышло новое поколение устройства AMP, включающее значительное обновление и расширение программного функционала. Внешне устройство выглядит как обычный коммутатор Fibre Channel, но оно предназначено не для коммутации, а для сбора и анализа информации о потоках данных в сети SAN (рис. 60).

С помощью AMP можно подключиться одновременно к 20 коммутаторам Fibre Channel, работающим в четырех различных фабриках. AMP позволяет собирать расширенную статистику производительности в сети без прерывания работы сети и без влияния на подключенные устройства. Графическое предоставление информации о производительности в сети с использованием AMP представлено на рис. 61.



Рис. 60. Analytics Monitoring Platform (AMP).

AMP расширяет и дополняет возможности Flow Vision и IO Insight. Если работа IO Insight привязана к конкретному коммутатору, то с помощью AMP возможно осуществлять сквозной мониторинг всей сети, отслеживать большее число потоков данных. Кроме того, основным преимуществом AMP является возможность глубокого и полного анализа собранных данных – как в реальном времени, так и в исторической перспективе – возможны хранение и анализ данных на срок до двух лет. Сравнение различных технологий мониторинга Brocade – рис. 62.

Лицензии коммутаторов

Коммутаторы Brocade обладают обширным функционалом, часть которого может лицензироваться дополнительно. Лицензии обычно продаются в виде ключа (Transaction Key), который необходимо активировать на сайте MyBrocade (my.brocade.com). Для активации необходимо знать WWN (уникальный адрес) коммутатора и ключ. Часть функций можно открыть временными лицензиями, действующими 45 или 60 дней. Эти лицензии можно получить, направив запрос в московский офис Brocade.

Ниже приведены основные дополнительные лицензии, которые не входят в базовую поставку на FC-коммутаторах:

Лицензии на аппаратную часть:

- *Ports on Demand license (POD)*: лицензии на включение (активацию) FC-портов. Коммутатор может поставляться с частью физических портов, которые не активированы и не могут использоваться до активации лицензии;
- *ICL License*: комплект из лицензии и QSFP модулей для активации функционала ICL на DCX 8510 и X6 Director. Одна лицензия активирует 16 портов ICL (1 Тбит/с). На DCX 8510-8 можно поставить две лицензии, на DCX 8510-4 – одну. Для X6 Director правила лицензирования немного изменились. Для X6-8 и X6-4 лицензии ICL различные. Каждая лицензия активирует половину портов директора. Таким образом, для активации всех портов необходимо приобрести две лицензии ICL. Кабели приобретаются отдельно (p/n кабелей можно посмотреть в DCX 8510 ICL FAQ на сайте Brocade).
- *Enterprise ICL License*: лицензия должна быть установлена на всех DCX 8510, использующих ICL, в случае, если этот DCX соединяется с 4 или более другими DCX. Для коммутаторов X6 Director лицензия Enterprise ICL License не нужна.
- *7800 Upgrade*: лицензия для коммутатора Brocade 7800. В базовой поставке – 4 порта FC 8 Гбит/с и 2 GbE. С помощью лицензии функционал коммутатора расширяется до максимального – 16 FC.

Временные лицензии на аппаратную часть не предоставляются.

Shared End-to-End Gen 5 Technology

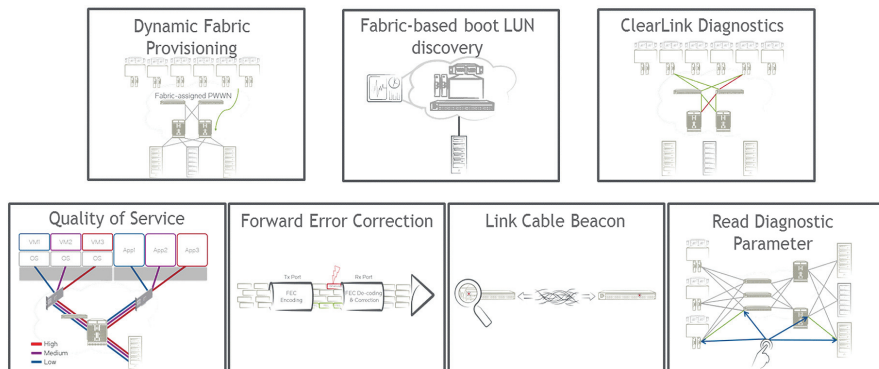


Рис. 59. Расширение возможностей Fabric Vision на все устройства в сети.

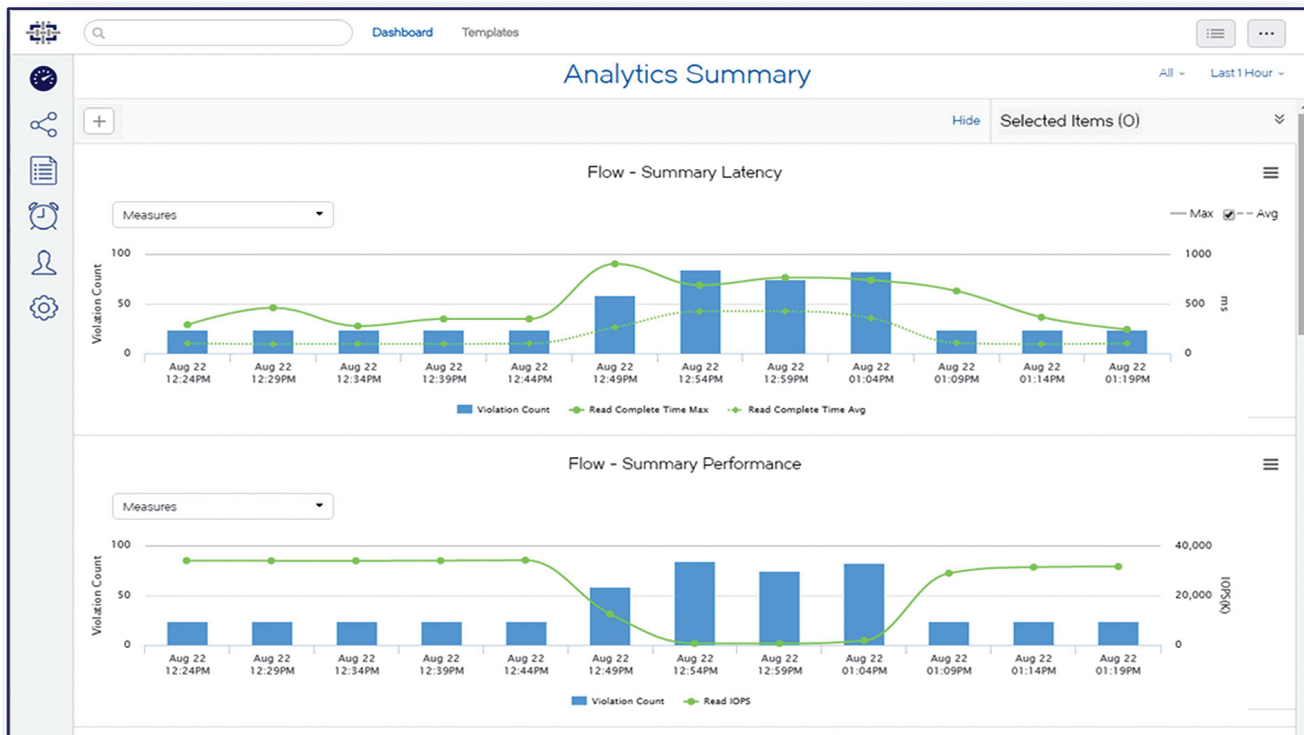


Рис. 61. AMP – графическое представление информации о производительности в сети.

Brocade Monitoring and Analytics Technologies

Brocade Analytics Monitoring Platform

Brocade Fabric Vision Technology with IO Insight

Brocade Fabric Vision Technology on Gen 5

Monitoring, management, and diagnostic tools

- Predefined, threshold-based rules/policies
- Dashboard views
- Live monitoring

Gen 6 storage IO-level monitoring

- SCSI R/W response/latency monitoring
- Alert for host and storage outliers
- Identify latency and congestion
- 100 percent non-disruptive

Use cases:

- Flow-level monitoring of a device port
- Visibility to understand storage performance
- Monitoring SLA compliance

Gen 5 and Gen 6 end-to-end IO monitoring

- Full visibility into all SCSI commands
- Alert for host, fabric, and storage outliers/individual IO flows
- Automatic LUN/flow learning
- Direct fabric latency monitoring
- 100 percent non-disruptive

Use cases:

- Monitoring mission-critical workloads
- Extensive diagnostics and troubleshooting
- Highly scalable from device level to fabric level

Рис. 62. Сравнение различных технологий мониторинга Brocade.

Лицензии для повышения производительности:

- **Extended Fabrics:** лицензия, позволяющая выделять дополнительные буферные кредиты для обеспечения максимальной производительности на расстояниях больше 10 км;
- **ISL Trunking:** лицензия на возможность агрегировать межкоммутаторные соединения, формируя на аппаратном уровне высокоэффективные «транки», в которых балансировка нагрузки ведется на уровне пакетов FC, а сбои отдельных соединений отрабатываются прозрачно и без перестроения фабрики;
- **Accelerator for FICON:** лицензия для оптимизации использования WAN-сети для протокола FICON. Используется для репликации типа XRC и работы с ленточными приводами FICON через FCIP-каналы.
- **Advanced Extension:** лицензия для расширенного функционала FCIP, в частности, для FCIP Trunking. Рекомендуется для большинства FCIP внедрений;
- **WAN Rate Upgrade 1 License и WAN Rate Upgrade 2 License** – лицензии для коммутатора

Brocade 7840 Extension switch. Управляют полосой пропускания сети WAN на коммутаторе. В базовой поставке полоса пропускания ограничена 5 Гбит/с. Первая лицензия расширяет полосу пропускания до 10 Гбит/с, вторая – до максимально возможной. Также вторая лицензия открывает порты 40 Гбит/с.

- **10GbE:** лицензия на использование 10G портов. Подходит как для FCIP (лезвие FX8-24), так и для 10G FC в продуктах пятого поколения Fibre Channel. При лицензировании DCX8510 нужно закупать одну лицензию на каждое лезвие, где планируется использовать 10G. Для коммутаторов шестого поколения Fibre Channel лицензия не нужна.

Лицензии на расширенные функции:

- **Integrated Routing:** лицензия для коммутаторов, позволяющая осуществлять передачу данных между независимыми фабриками, не объединяя их;
- **Adaptive Networking:** возможность использовать QoS на FC-соединениях, а также ограничивать потоки данных оконечных устройств (функция Ingress Rate Limiting) для более тон-

кой настройки производительности. Начиная с версии FOS 7.2 данная лицензия более не используется. Функционал включен в базовую поставку FOS.

- **Server Application Optimization (SAO):** лицензия для управления QoS со стороны сервера с помощью политик (SLA). При использовании NPIV каждая виртуальная машина может устанавливать свой QoS. Начиная с версии FOS 7.2 данная лицензия более не используется. Функционал включен в базовую поставку FOS.

Лицензии на мониторинг и управление:

- **Advanced Performance Monitoring:** лицензия для точного мониторинга производительности, отслеживания максимально нагружающих фабрик устройств (Top Talkers), объема трафика между любой парой устройств. Начиная с версии FOS 7.4 функционал Advanced Performance Monitoring более не поддерживается и не может использоваться. Вместо него необходимо использовать Flow Vision;
- **Fabric Watch:** проактивный мониторинг SAN, позволяет контролировать множество параметров работы коммутатора, от количества отказавших вентиляторов до количества ошибок определенного типа на порту. При достижении установленного порогового значения Fabric Watch отправляет сообщения в системный журнал, syslog и по протоколу SNMP. Начиная с версии FOS 7.4 функционал Fabric Watch более не поддерживается и не может использоваться. Вместо него необходимо использовать MAPS;
- **Fabric Vision:** новая лицензия, заменяющая одновременно Advanced Performance Monitoring и Fabric Watch. Позволяет пользоваться функционалом Flow Vision и MAPS. Если на коммутаторе были установлены лицензии Fabric Watch и Advanced Performance Monitoring,

toring, то лицензия Fabric Vision добавится автоматически при переходе на FOS 7.2. Если одной из двух лицензий не хватает, ее можно докупить;

- **FICON CUP**: лицензия для in-band мониторинга и управления коммутаторами с мейнфреймов, через интерфейс Control Unit Port (CUP).

Enterprise Bundle

Лицензии могут поставаться по отдельности или в виде наборов (Bundle). Наиболее популярный набор – Enterprise Bundle. Его состав менялся в течение времени. В настоящий момент Enterprise Bundle включает следующий набор функций:

- Trunking (Trk);
- Fabric Vision (FV);
- Extended Fabric (EF).

Enterprise Bundle может быть уже предустановлен на коммутаторе при его приобретении – в зависимости от модели оборудования.

Особенности лицензирования коммутаторов FC шестого поколения

Для коммутаторов шестого поколения Fibre Channel схема лицензирования значительно упрощена. Значительно сокращено количество лицензий. Ниже приведены доступные для шестого поколения лицензии:

- ICL POD License;
- Ports on Demand license (POD);
- Integrated Routing;
- Enterprise Bundle (Trunking, Fabric Vision, Extended Fabric);
- Mainframe Enterprise Bundle (Trunking, Fabric Vision, Extended Fabric, FICON CUP).

Отдельно лицензии Trunking, Fabric Vision, Extended Fabric и FICON CUP не предлагаются.

Дополнительная информация

Обучение и сертифицированные курсы Brocade

Существует несколько вариантов обучения технологиям Brocade SAN:

- очные курсы с преподавателем на английском или русском языке;
- веб-курсы с инструктором на английском языке;
- веб-курсы и материалы для самостоятельной подготовки на английском языке.

Все веб-курсы доступны на портале MyBrocade. Часть курсов – платно, часть – бесплатно. В частности, для партнеров Brocade Alliance Partner Network уровней Select, Premier или Elite, веб-обучение бесплатно. На уровень Select может авторизоваться практически любой системный интегратор, заинтересованный в партнерстве с Brocade. Общая структура курсов Brocade представлена рисунке. Руководства для самоподготовки к сертификационным экзаменам, пробные сертификационные тесты – бесплатно

доступны на сайте Brocade – www.brocade.com/education/. Базовыми курсами для обучения является бесплатный веб-курс FC101 и курсы CFA200 и CFP300 (доступны как очные с преподавателем, там и в форме веб-курсов).

Очное обучение в России и СНГ доступно в учебных центрах в Санкт-Петербурге, Москве и Киеве. Помимо чтения курсов силами своих преподавателей, все эти учебные центры периодически приглашают англоязычных преподавателей для чтения авторизованных курсов Brocade:

- **Unix Education Center** (www.unixedu.ru), тренер Юрий Яворский, Санкт-Петербург – авторизованные курсы Brocade (BASA200, BPSA300, CFD200, SAN-TS300, EFF200 и EFS300) на русском языке и авторские курсы по продуктам Brocade;
- **Учебный центр HPE в Москве** (www.hpe.ru/education/), раздел «Системы хранения данных – HPE Storage & SAN», тренер Антон Чернышов – собственные курсы HPE по SAN (Brocade) по-русски в Москве;
- **Учебный центр МУК, Киев** (training.muk.ua), тренер Евгений Мельничук, Brocade Distinguished Architect, сотрудник HPE Украина – собственные курсы по SAN (Brocade) по-русски или по-украински.

Рекомендуемые бесплатные веб-курсы (по-английски):

- Introduction to Fibre Channel Concepts (FC 101-WBT);
- Brocade 16 Gbps Fibre Channel Product Training (SAN 140-WBT);
- Introduction to Brocade Network Advisor (BNA 101-WBT);
- Brocade Network Advisor Implementation, Migration, and Configuration (BNA 150-WBT);
- Managing Storage Area Networks with Brocade Network Advisor (BNA 160-WBT);
- Brocade Network Advisor for SAN – PDF-Only (BNA 200-SAN-PDF);
- Introduction to Brocade Fabric Operating System (FOS 101-WBT);

- Fabric OS Update Training (FOS 110-WBT);
- Introduction to Brocade SAN Fabric Resiliency (FR 100-WBT);
- Introduction to FICON (FICON 101-WBT);
- Introduction to the Brocade 7840 Extension Switch (EXT 110-WBT).

Программа авторизованного обучения и сертификации Brocade

С 1 ноября 2016 года в программу сертификации Brocade вносятся изменения, связанные с появлением новых продуктов, в том числе коммутаторов Brocade 6 поколения. Будут доступны обновленные учебные курсы и сертификационные экзамены.

Данные курсы доступны, в частности, в учебном центре НОУ Unix Education Center, Санкт-Петербург.

BASA 200 BROCADE ASSOCIATE SAN ADMINISTRATOR

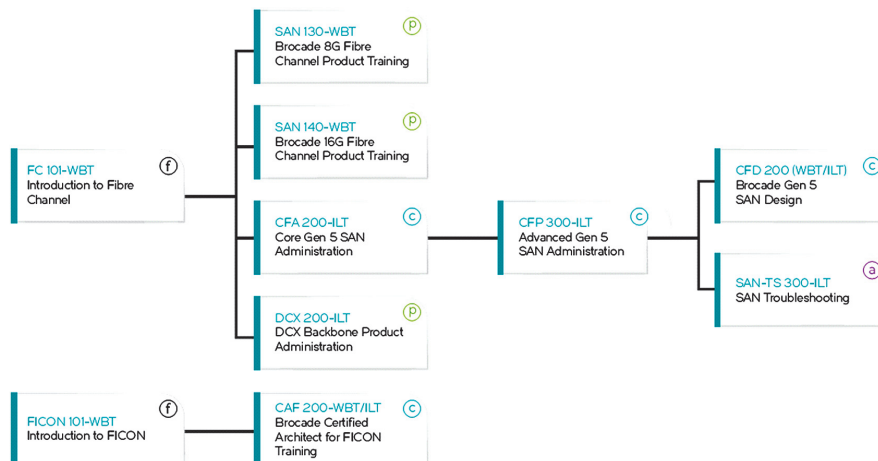
Курс дает слушателям знания и навыки работы с продуктами Brocade пятого и шестого поколений (Gen 5 и Gen 6), а также уделяет внимание основным возможностям Fabric OS v8.0.1. В ходе курса рассматривается введение в технологию Fibre Channel, настройка коммутаторов, зонирование, MAPS, ISL, увеличение расстояний, маршрутизация, ISL транкинг, ICL, установка нового ПО, резервное копирование конфигураций, основы поиска и исправления неисправностей. Практические занятия проводятся на основе удаленной лаборатории Brocade при помощи утилит управления Brocade – telnet и Network Advisor 14.0.1.

Курс предназначен для администраторов хранилищ и сетей хранения данных, проектировщиков СХД, системных инженеров и операторов СХД, которые осуществляют настройку, поддержку, мониторинг, оптимизацию и базовое устранение ошибок в работе сетей хранения Brocade.

Данный курс готовит к успешной сдаче сертификационного экзамена Brocade Associate SAN Administrator 2017 exam (Exam #143-430).

Очный курс, рассчитанный на 4 дня обучения. Обучение проводится на русском языке.

Структура учебных курсов Brocade



SUPPLEMENTAL



LEGEND

- Ⓡ fundamentals
- Ⓟ product training
- Ⓒ certification or accreditation
- Ⓜ advanced/troubleshooting

BPSA 300 BROCADE PROFESSIONAL SAN ADMINISTRATOR

Курс дает слушателям знания и навыки работы с расширенными возможностями Fabric OS v8.0.1. В ходе курса рассматриваются вопросы маршрутизации FibreChannel (FC-FC Routing), Fibre Channel over Internet Protocol (FCIP), AccessGateway, VirtualFabric (VF), Monitoring and Alerting Policy Suite (MAPS), Flow Vision и утилиты управления Brocade – telnet и Network Advisor 14.0.1.

Курс предназначен для администраторов хранилищ и сетей хранения данных, проектировщиков СХД, системных инженеров и операторов СХД, которые осуществляют настройку, поддержку, мониторинг, оптимизацию и устранение ошибок в работе СХД Brocade - FC-FC Routing, FCIP, NPIV, AccessGateway и VirtualFabric.

Данный курс подготавливает к успешной сдаче сертификационного экзамена Brocade Professional SAN Administrator 2017 exam (Exam #143-090).

Очный курс, рассчитанный на 3 дня обучения. Обучение проводится на русском языке.

CFA 200 Core SAN Administration – Brocade Certified Fabric Administrator

Курс Brocade Certified Fabric Administrator дает слушателям знания и навыки работы с продуктами Brocade, а также уделяет внимание основным возможностям Fabric OS v7.2.0. В ходе курса рассматривается введение в технологию Fibre Channel, настройка коммутаторов, зонирование, ISL, маршрутизация, ISL-транкинг, установка нового ПО, резервное копирование конфигураций, основы поиска и исправления неисправностей. Очный курс, рассчитанный на 3 дня обучения. Курс включает практические занятия на оборудовании Brocade. Читается по-русски.

Данный курс способствует получению и укреплению навыков, необходимых для подготовки к экзамену: 143-425 – Brocade Certified Fabric Administrator Gen 5.

CFP 300 Advanced SAN Administration – Brocade Certified Fabric Professional

Курс Brocade Certified Fabric Professional дает слушателям знания и навыки работы с расширенными возможностями Fabric OS v7.2.0. В ходе курса рассматриваются вопросы маршрутизации Fibre Channel (FC-FC Routing), Fibre Channel over IP (FCIP), Adaptive Networking, Access Gateway, Virtual Fabric (VF) и утилиты управления Brocade. Слушатели выполняют практические задания по управлению коммутаторами с использованием утилит telnet, Brocade Web Tools и Brocade Network Advisor v12.1.3. Очный курс, рассчитанный на 3 дня обучения. Курс включает практические занятия на оборудовании Brocade. Читается по-русски.

Данный курс способствует получению и укреплению навыков, необходимых для подготовки к экзамену: 143-085 – Brocade Certified Fabric Professional Gen 5.

CFD200 BROCADE GEN 5 SAN DESIGN

Курс охватывает вопросы проектирования SAN на основе аппаратных и программных компонент Brocade Gen5. Программа обучения содержит фундаментальные составляющие проектирования и внедрения архитектуры сетей хране-

ния Brocade с прицелом на максимально эффективное решение.

Курс ориентирован на инженеров и pre-sale специалистов Brocade, которые планируют повысить свой уровень до архитекторов SAN, а также тем, кто хочет подготовиться к сдаче экзамена Brocade Certified Fabric Designer Gen 5 (Exam #143-280).

Очный курс, рассчитанный на 3 дня обучения. Обучение проводится на русском языке.

SAN-TS 300 Brocade SAN Troubleshooting

Этот 3-дневный курс под руководством инструктора дает студентам технические навыки, необходимые для поиска и устранения неисправностей, а также задач сбора данных в коммутируемых и маршрутизируемых SAN на основе продуктов 16 Гбит/с и Brocade Fabric OS v7.1.0. Лекции охватывают основные концепции с прицелом на минимум теории и максимум практических занятий, в ходе которых слушатели на деле закрепляют все рассматриваемые в лекциях процессы. Обучение проводится на русском языке.

Данный курс способствует получению и укреплению навыков, необходимых для подготовки к экзамену: 160-110 – Brocade Accredited Technical Support Specialist.

EFF200 ETHERNET FABRIC FOUNDATIONS

Курс дает слушателям знания и навыки работы с продуктами Brocade VDX. В ходе обучения рассматриваются концептуальные составляющие технологии конвергентных сетей VCS Fabrics – преимущества и варианты использования, стандарты Data Center Bridging (DCB), TRILL и FSPF, передача трафика в фабрике и базовые процессы ее конфигурирования. Также в программу курса входит дополнительная информация по настройке VLANs, Spanning Tree и vLAGs средствами NOS v7.0. Вопросы обслуживания охватывают обновление программного обеспечения, базовый поиск и устранение неисправностей, а также управление фабрикой средствами Brocade Network Advisor.

Курс предназначен для администраторов и системных инженеров, которые планируют осуществлять установку и базовую настройку фабрик VCS, а также их администрирование и мониторинг.

Очный курс, рассчитанный на 3 дня обучения. Обучение проводится на русском языке.

EFSA300 ETHERNET FABRIC STORAGE ADMINISTRATION

Курс является введением в VCS Storage Fabrics. Темы курса охватывают вопросы администрирования VCS Fabrics, связанные с передачей трафика сетей хранения. В ходе обучения детально разбираются процессы стандартов DCB и FCoE, а также интеграция продуктов VDX с традиционными сетями хранения Fibre Channel посредством Access Gateway и маршрутизации FCR. Дополнительные темы охватывают передачу трафика iSCSI и NAS в VCS Fabrics, технологии AMP, QoS и Buffering, а также мониторинг фабрик на основе функциональности MAPS.

Курс предназначен для администраторов, системных инженеров и архитекторов, которые планируют либо уже осуществляют разработку ре-

шений, настройку, поддержку, мониторинг, оптимизацию и устранение ошибок в среде Brocade VCS Fabrics.

Очный курс, рассчитанный на 5 дней обучения. Обучение проводится на русском языке.

Вопросы и ответы

Как купить оборудование Brocade FC?

Все продажи и поддержка SAN-оборудования Brocade в регионе Россия и СНГ идут только через OEM-партнеров Brocade. В настоящее время у компании Brocade существует 11 OEM-партнеров. Это компании Dell, EMC, Fujitsu, Hitachi, Hewlett Packard Enterprise, Huawei, IBM, Lenovo, NetApp, Oracle, Sugon (<http://www.brocade.com/en/partners/oem-partners.html>). Компания Brocade не продает и не поддерживает оборудование Fibre Channel напрямую в регионе Россия и СНГ.

Всякий авторизованный OEM-партнером дистрибутор, системный интегратор или сервисный партнер может продавать или обслуживать оборудование Brocade SAN.

Нет и не планируется прямых дистрибуторов или системных интеграторов для продажи SAN-оборудования Brocade в регионе Россия и СНГ.

Продажа оборудования по OEM-соглашениям означает, что с юридической точки зрения оборудование является продуктом соответствующего OEM-партнера.

Гарантийная и постгарантийная поддержка Brocade FC?

Техническую поддержку на продукты Brocade FC осуществляют OEM-партнеры – поставщики соответствующего оборудования, а также, по их поручению – соответствующие сервисные партнеры. Длительность и условия технической поддержки могут различаться у различных OEM-партнеров. Компания Brocade не осуществляет прямую техническую поддержку продуктов Fibre Channel на территории России и СНГ. По всем вопросам обслуживания продуктов FC необходимо обращаться к соответствующему OEM-партнеру.

При невозможности самостоятельного решения технической проблемы OEM-партнер может привлечь к решению специалистов Brocade. Решение об эскалации технической проблемы принимает OEM-партнер.

Когда прекращается поддержка?

Стандартная политика компании Brocade – после прекращения поставки продуктов их техническая поддержка со стороны Brocade осуществляется еще пять лет. В течение этого срока возможны (при наличии соответствующего контракта на техподдержку) замена вышедших из строя компонентов и обновление программного обеспечения. После окончания пятилетнего периода техническая поддержка продуктов, как правило, не осуществляется. Точный срок окончания технической поддержки необходимо уточнять у OEM-партнеров – поставщиков соответствующего оборудования.

Как стать партнером Brocade SAN?

Партнерство Brocade Partner Network учитывает продажи SAN-оборудования через OEM партнеров.

Регистрация партнером начального уровня Select дает бесплатный доступ к учебным веб-курсам (<http://my.brocade.com/wps/portal/registration>).

Можно ли покупать апгрейды Brocade FC через другого OEM-партнера?

Так как все продукты Brocade Fibre Channel, продаваемые в России и СНГ, являются OEM-продуктами соответствующих компаний, то покупка апгрейдов (программных и аппаратных) через другого OEM-партнера, отличного от изначального продавца коммутатор, юридически является покупкой оборудования другого вендора. Технически, такое решение скорее всего будет работоспособным, но появляются риски по поддержке. Следует ясно определить, кто будет поддерживать это решение, и получить подтверждение данной сервисной организации, что решение будет действительно ими поддерживаться.

Как обновлять ПО Brocade FC?

Компания Brocade периодически выпускает новые версии программного обеспечения, например Brocade Fabric OS и Brocade Network Advisor. Поскольку данные продукты поставляются только через OEM-партнеров, обновление ПО следует получать только от соответствующих OEM-партнеров – поставщиков оборудования в соответствии с их лицензионным соглашением. Как правило, ПО можно загрузить с веб-сайта соответствующего OEM-партнера при наличии действующей технической поддержки. Подробности необходимо уточнять в службе технической поддержки поставщика оборудования или ПО.

Какая версия ПО Fabric OS рекомендована к установке?

Компания Brocade периодически выпускает новые версии Fabric OS. Это могут быть как глобальные обновления ПО, связанные с появлением нового оборудования или нового программного функционала, так и патч-релизы, предназначенные для устранения выявленных программных ошибок. Выбор требуемой версии Fabric OS зависит от следующих факторов:

- поддержка требуемой версии на имеющемся оборудовании. Коммутаторы предыдущих поколений могут не поддерживать последние версии FOS;
- поддержка требуемой версии соответствующим OEM-партнером. Поставщики оборудования могут поддерживать только определенные версии FOS;
- требования к функционалу. Новейший функционал может быть доступен только в последних версиях ПО;
- требования к стабильности работы. Последние версии ПО могут содержать новейший функционал, но также могут содержать различные программные ошибки. Brocade публикует документ под названием «Brocade Fabric OS Target Path Selection Guide», в котором указаны версии FOS, которые компания Brocade считает наиболее стабильными с точки зрения надежности работы. Мы рекомендуем использовать версии Fabric OS, указанные в данном документе.

Как определить совместимость с Brocade FC коммутаторами?

Компания Brocade периодически публикует документ под названием «Brocade Fabric OS Compatibility Matrix». В нем приведен список оборудования, совместимого с оборудованием Brocade. Мы так же рекомендуем уточнять совместимость оборудования у производителей серверов, СХД и ПО, которые будут применяться в данном решении и у тех компаний, которые будут осуществлять техническую поддержку решения.

Каковы пределы масштабируемости сетей Brocade FC?

Для построения сетей SAN, обеспечивающих максимальную стабильность работы, советуем при планировании сети учитывать рекомендации, приведенные в документе «Brocade SAN Scalability Guidelines». Документ описывает, каких максимальных значений могут достигать различные сетевые параметры. Компания Brocade проводила тестирование конфигураций вплоть до указанных максимумов и может гарантировать работоспособность и стабильность указанных конфигураций. При превышении максимальных значений поведение сетевых устройств может быть непредсказуемым.

Чем сотрудники Brocade могут практически помочь?

Пожалуйста, обращайтесь к нам за помощью по всем вопросам связанным с SAN-инфраструктурой. Мы можем:

- помочь найти правильный контакт у OEM-партнера;
- независимо оценить подготовленные решения по развитию SAN-инфраструктуры. На основе отчета SAN Health и беседы с инженерами заказчика дать рекомендации по текущему состоянию SAN;
- предоставить демо-технику (из демофонда Brocade в Москве), а также удаленное тестирование в лабораториях по миру;
- провести семинары/лекции по интересующей теме;
- выдать временные лицензии для FC-коммутаторов.

Контакты сотрудников Brocade по этому и другим направлениям можно найти на сайте www.brocade.ru.

Источники информации

Группа Brocade Russia на LinkedIn (<http://www.linkedin.com/groups?gid=4246353>)

В этой группе выкладываются технические материалы и ведутся дискуссии, относящиеся к технологиям Brocade SAN. Группу ведет российская команда Brocade SAN.

Brocade SAN Health – бесплатное ПО проверки состояния сети SAN (www.brocade.com/sanhealth):

- используется для обнаружения устройств, документирования и планирования;
- работает на всех продуктах McDATA и Brocade;

- начиная с версии 3.2.0 также собирает информацию с Cisco MDS;
- имеет утилиты миграции и переноса зон с фабрик Cisco на Brocade.

Литература для чтения по теме Brocade SAN:

- Раздел *Education* на сайте Brocade;
- Книга «Основы проектирования SAN» Джоша Джадда;
- *SAN Fabric Administration Best Practices Guide*;
- *Brocade Fabric OS Compatibility Matrix*;
- *Brocade SAN Scalability Guidelines*;
- *BROCADE MAINFRAME CONNECTIVITY SOLUTIONS*;
- *SECURING FIBRE CHANNEL FABRICS*;
- *SAN Migration Guide: Migrating from Cisco to Brocade*;
- *IBM RedBook: Introduction to Storage Area Networks*;
- *IBM RedBook: Gen 5 16 Gbps Switches and Network Advisor*;
- *IBM RedBook: Implementing SAN with 8 Gbps Directors and Switches*;
- *Руководство по SAN дизайну и лучшим практикам*;
- www.brocade.ru.

Сайт my.brocade.com – здесь находится официальная документация по коммутаторам Brocade, Fabric OS, Network Advisor (Installation, Administration, Licensing, Command Reference, Troubleshooting and Diagnostics guides, Release Notes etc.)

Обучение Brocade (Раздел Brocade Training на сайте my.brocade.com):

- руководства для самоподготовки к сертификационным экзаменам, пробные сертификационные тесты – бесплатно доступны на сайте Brocade;
- Web-курсы (платные и бесплатные), бесплатно – пробные сертификационные тесты, курсы по SAN и IP. Продвинутые платные курсы;

Заключение

Мир проходит через цифровую трансформацию SMAC (Social media, Mobility, Analytics, Cloud computing). Появляется все больше данных, растут требования к ИТ-инфраструктуре. Системы хранения на флэш-технологиях совершенствуются быстрее, чем ожидали. В работе – стандарт по сетевому доступу к энергонезависимой памяти NVMe over Fibre Channel. Коммутаторы Fibre Channel Gen6 готовы решать новые задачи. Технология Fibre Channel жива и развивается.

Brocade – основной игрок на рынке Fibre Channel SAN. Мы специализируемся на построении больших инфраструктур. 96% банков, страховых компаний и предприятий розничной торговли мира используют сети Fibre Channel для построения сетей хранения данных. Крупнейшие банки и мобильные операторы России доверяют свои данные Brocade. Brocade – это надежно.