

Сети хранения данных на базе интерфейса Fibre Channel

Эту главу можно рассматривать как введение в сети хранения данных (storage area network — SAN) в общем и в сети хранения данных на базе интерфейса Fibre¹ Channel в частности. Хотя сети хранения данных могут создаваться и на основе технологий, отличных от Fibre Channel, большинство из них будут использовать Fibre Channel еще достаточно долгое время. Именно поэтому интерфейсу Fibre Channel уделяется в этой главе основное внимание. Сети хранения данных, основанные на других технологиях, например iSCSI, рассматриваются в главе 8.

Интерфейс Fibre Channel — это технология межсистемного взаимодействия, которая объединяет в себе возможности высокоскоростного ввода-вывода и сетевого обмена данными. Когда эта книга готовилась к печати, сети хранения данных на основе Fibre Channel обеспечивали быстроедействие 1 Гбит/с; кроме того, растет количество сетей, поддерживающих скорость 2 Гбит/с.

В терминологии Fibre Channel устройства называются *узлами* (nodes). Это весьма напоминает узлы в терминологии сетей IP. Узел Fibre Channel может иметь несколько портов, как и узел IP, который зачастую получает несколько адресов IP. Разница между ними в том, что порт Fibre Channel представляет собой физический элемент, а порт IP — логический. Каждый узел Fibre Channel имеет уникальное 64-разрядное имя WWN (World Wide Name), которое назначается производителем. Это напоминает уникальные адреса MAC, которые назначаются производителями сетевым адаптерам Ethernet. Каждый порт сети хранения данных на базе кольца с разделением доступа Fibre Channel имеет 8-битовый адрес, а порт в коммутируемой связанной архитектуре — 24-битовый. При подключении *кольца с разделяемым доступом* (arbitrated loop) к *коммутатору связанной архитектуры* (fabric switch),

¹Написание Fiber было заменено на Fibre, чтобы указать, что технология Fibre Channel может использовать медные и оптические носители.

коммутатор представляет 8-битовый адрес в виде 24-битового. Оба идентификатора порта назначаются динамически. Концепция портов и различные их типы обсуждаются в разделе 4.5.

Сфера применения Fibre Channel рассматривается в следующем разделе, после чего технология SAN сравнивается с NAS (см. главу 3). В этой главе сначала описываются принципы технологии Fibre Channel, затем внимание акцентируется на протоколах нижнего уровня и наконец рассматриваются различные элементы (включая устройства), из которых и создаются системы SAN на базе интерфейса Fibre Channel.

4.1 Сферы применения технологии Fibre Channel

В технологии Fibre Channel предпринята попытка объединить лучшее из двух миров — каналов передачи данных и сетей. Термин *канал* впервые стал использоваться в мире мэйнфреймов и описывал структурированный механизм передачи данных. В большинстве случаев передача данных выполняется между компьютерной системой и периферийным устройством, например жестким диском или накопителем на магнитной ленте. К таковым каналам относятся интерфейсы SCSI (Small Computer System Interface) и HIPPI (High-Performance Parallel Interface). Работа каналов обычно реализуется средствами аппаратного обеспечения.

По сравнению с каналом *сеть* представляет собой более универсальный механизм для передачи данных, который, однако, менее структурирован. Кроме того, сеть может работать на значительно большем расстоянии и подключаться к большему количеству устройств, чем канал. В отличие от каналов, сети в основном реализуются средствами программного, а не аппаратного уровня.

Один из подходов в объединении систем хранения данных и сетей заключается в том, что сеть становится ключевым элементом, к которому добавляются новые возможности с одновременной компенсацией недостатков подобного подхода. Речь идет о технологии хранения данных на базе протокола IP (см. главу 8).

Другой подход состоит в использовании центрального хранилища данных (канальная система) и расширения существующих технологических функций. На базе этого метода создавалась технология Fibre Channel. Одно из основных преимуществ Fibre Channel по сравнению с IP Storage заключается в разработке продуктов Fibre Channel уже в течение 10 лет, в то время как решения для IP Storage появились сравнительно недавно.

Учитывая, что Fibre Channel основана на структуре каналов, стоит рассмотреть недостатки другой известной технологии — SCSI.

- Максимальная скорость передачи данных — 80 Мбайт/с (впоследствии скорость возросла до 320 Мбайт/с, но уже после появления технологии Fibre Channel), чего явно недостаточно для хранилищ данных большого объема.
- Адаптер поддерживает только 16 устройств.
- Преимущество, которое одновременно является и недостатком, — обратная совместимость. Стандарт SCSI развивался много лет, и производители устройств смогли обеспечить обратную совместимость для нескольких поколений устройств. Но администратор должен следить, чтобы к шине не подключались устройства предыдущих поколений, поскольку шина автоматически перейдет в режим, поддерживающий работу самого старого устройства.
- Поддерживается длина кабелей, составляющая несколько десятков метров. Этого недостаточно для создания обычных, а тем более географически распределенных кластеров.
- Существуют и другие альтернативы SCSI, например SSA (Serial Storage Architecture), которые еще находятся в рамках архитектуры Intel или вообще представляют собой открытый стандарт.

4.2 Сравнение SAN и NAS

В главе 3 рассматривается технология NAS. Прежде чем знакомиться с архитектурой сетей хранения данных на базе Fibre Channel, следует провести сравнение принципов создания хранилищ. В табл. 4.1 описываются различия и общие черты этих технологий.

Таблица 4.1. Сравнение технологий NAS и SAN

NAS	SAN
NAS предоставляет доступ к файлам по сети. Поскольку пользователи знают о существовании файлов, например файлов электронных таблиц Microsoft Excel или файлов документов Word, устройства NAS более интуитивны для восприятия	SAN предоставляют доступ к дисковым блокам по сети. Пользователи обычно знают о существовании файлов, а файловая система предоставляет уровень абстракции для сокрытия информации о том, в каком дисковом блоке находится определенный файл
NAS использует сетевые протоколы доступа к файлам, например CIFS и NFS. Эти протоколы обычно пользуются услугами протокола TCP/IP, который, в свою очередь, зависит от технологии Ethernet	SAN предоставляет доступ к дисковым блокам по протоколу Fibre Channel. Верхний уровень этого протокола обычно представляет собой протокол SCSI

Окончание табл. 4.1

NAS	SAN
Устройства NAS — это, по сути, серверы, к файлам которых получают доступ клиенты	Устройства SAN — это устройства хранения данных, к которым могут получать доступ программы-инициаторы
Устройства NAS несложно внедрять в локальную сеть. NAS не требует усложнять структуру сети, однако все данные передаются средствами локальной сети	SAN требует создания новой сетевой архитектуры, которая обычно основана на интерфейсе Fibre Channel. Новая сеть снижает вероятность “заторов” в локальной сети и позволяет выполнять такие операции, как резервное копирование вне локальной сети
Подключение устройства NAS к локальной сети не требует больших затрат	Подключение SAN к локальной сети обходится довольно дорого
Обычно устройства NAS подключаются с помощью интерфейса Ethernet со скоростью передачи 100 Мбит/с или 1 Гбит/с, если используется технология Gigabit Ethernet	Сети хранения данных на основе Fibre Channel обычно работают на скоростях 1 или 2 Гбит/с

4.3 Преимущества Fibre Channel

Рассмотрим преимущества использования SAN на базе Fibre Channel. Хотя возможности, описанные в разделах 4.3.1–4.3.7, все еще широко применяются, не забывайте, что такая ситуация может измениться в любой момент, так как технологический прогресс не стоит на месте.

4.3.1 Масштабируемость

Сети хранения данных обладают способностью к масштабированию в аспекте объема и скорости передачи данных. Можно добавить коммутатор и расширить кольцо отдела до магистрали, которая каскадом соединяет иерархии коммутаторов. Конечно, сети IP также поддерживают масштабирование, что является основной причиной такого быстрого развития технологии IP Storage (см. главу 8).

4.3.2 Сегрегация хранилищ

При корректной настройке и управлении SAN предоставляет наилучшие возможности обоих вариантов построения хранилищ данных: хранилище отделено от сервера приложений и в то же время обеспечивает защиту и целостность данных. Например, можно разделить данные отделов, которые физически будут находиться в одном пуле устройств хранения данных.

4.3.3 Централизация и управление хранилищем

Сети хранения данных позволяют консолидировать устройства хранения, оптимизировать их использование и управление. Речь идет о том, что дублировать данные, которые размещены вне инфраструктуры SAN, нет необходимости. Еще одним преимуществом является возможность распределения хранилищ. Это позволяет избежать ситуации, когда у одного сервера объем хранилища избыточен, а второму серверу объема не хватает. Эффективное распределение объема хранилища позволяет сократить накладные расходы, связанные с добавлением устройств хранения к определенному серверу. Например, если управление дисками проводится централизованно, хранилище одного сервера может быть передано в использование другому серверу, которому не хватает свободного дискового пространства. Еще одним примером может служить хранилище, необходимое для создания “моментального снимка” в процессе резервного копирования. В классической схеме каждый сервер имеет для этих целей собственный жесткий диск. В технологии SAN все серверы совместно используют несколько дисков для создания временных моментальных снимков.

4.3.4 Поддержка устаревших устройств

Сети хранения данных позволяют защитить инвестиции в устаревшие устройства, например в устройства хранения SCSI, или, с ультрасовременной точки зрения, в кольцевые концентраторы. Устройства-мосты, которые обеспечивают взаимодействие между старыми устройствами и интерфейсом Fibre Channel, часто являются основой для поддержки устаревших устройств. Конечно, такая схема позволяет использовать далеко не все устаревшие устройства.

4.3.5 Поддержка большего количества устройств

Сети хранения данных предоставляют доступ к большому количеству устройств. Кольцо с разделяемым доступом поддерживает теоретический максимум, равный 127 портам (обычно подключается до 15 устройств, а значительные задержки в работе возникают после подключения 50 устройств).

Сеть хранения данных на основе связанной архитектуры теоретически поддерживает около 15 млн. (2^{24}) портов.

4.3.6 Расстояние

Сети хранения данных позволяют размещать устройства хранения гораздо дальше от узлов (серверов и рабочих станций), чем другие интерфейсы. Шина SCSI поддерживает расстояния порядка десятков метров, а Fibre Channel — десятков километров.

4.3.7 Новые возможности

Интерфейс Fibre Channel позволяет использовать совершенно новые функции. Один из примеров — резервное копирование данных без использования локальной сети, когда резервная копия передается по сети хранения данных, а локальная сеть остается свободной для запросов клиентов и серверов к устройствам хранения. Еще одним примером является перенос данных между двумя подсистемами хранения без участия сервера. Кроме того, файловые системы SAN (они рассматриваются в главе 6) позволяют нескольким серверам получить одновременный доступ к одним и тем же томам.

4.4 Топологии Fibre Channel

В разделах 4.4.1–4.4.3 рассматриваются различные топологии подключения устройств, которые формируют сеть хранения данных на базе Fibre Channel. Топологии “точка–точка” (point to point), кольцо с разделяемым доступом (arbitrated loop) и коммутируемая связанная архитектура (switched fabric) перечислены в порядке возрастания их сложности.

4.4.1 Топология “точка–точка”

Технология Fibre Channel поддерживает подключение по топологии “точка–точка”. В этом случае сервер обычно подключается к выделенной подсистеме хранения, причем данные не используются совместно. На рис. 4.1 показана сеть, построенная на основе этой топологии.

Для реализации топологии “точка–точка”, как минимум, необходим сервер, адаптер Fibre Channel (*адаптер шины*) и устройство хранения (например, жесткий диск или накопитель на магнитной ленте), оснащенное интерфейсом Fibre Channel.

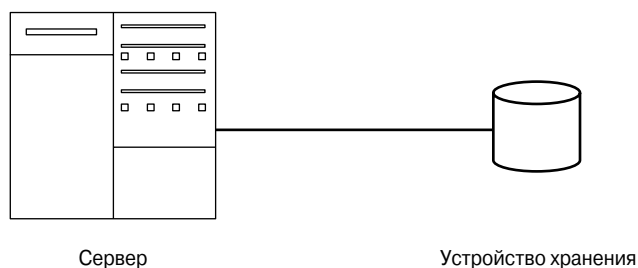


Рис. 4.1. Топология “точка–точка”

4.4.2 Кольцо с разделяемым доступом

Кольцо — это схема логического подключения устройств, при котором данные передаются по логически замкнутому контуру. В *кольце с разделением доступа* (arbitrated loop) протокол описывает порядок, в котором узел получает разрешение на передачу данных. Кольцо Fibre Channel с разделением доступа (Fibre Channel arbitrated loop — FC-AL) может быть реализовано на базе различных устройств хранения (жестких дисков, накопителей на магнитной ленте), серверов, адаптеров шины и устройств для их подключения. В качестве устройств подключения могут выступать концентраторы или коммутаторы Fibre Channel (см. раздел 4.7.4). На данный момент достаточно указать, что устройства подключения играют важную роль в организации инфраструктуры кольца, в его работе и управлении им.

На рис. 4.2 показан пример кольца Fibre Channel с разделением доступа. Конфигурация аналогична физической звезде и логическому кольцу, используемому в локальных сетях на базе технологии Token Ring. Кроме того, как и в сетях Token Ring, данные перемещаются по кольцу в одном направлении. Но, в отличие от Token Ring, устройство может запросить право на передачу данных, а не ждать получения пустого маркера от коммутатора.

Команды Fibre Channel поддерживают согласование и доступ к кольцу для передачи данных. Кроме того, предоставляются команды для назначения адресов портов кольца с разделением доступа (arbitrated-loop port addresses — AL-PA) различным узлам кольца. Каждый узел в управляемом кольце Fibre Channel имеет контур для собственного отключения от кольца и сохранения непрерывности кольца в случае ошибки.

Кольца Fibre Channel с разделением доступа могут адресовать до 127 портов, в частности типа NL (дополнительная информация приводится в разделе 4.5), что обусловлено методом указания адреса AL-PA. Один из этих портов зарезервирован для подключения к коммутатору связанной архитектуры (см. раздел 4.7.4.3). Остальные 126 портов доступны для предоставления узлам. Практика показывает, что типичные кольца с разделением доступа

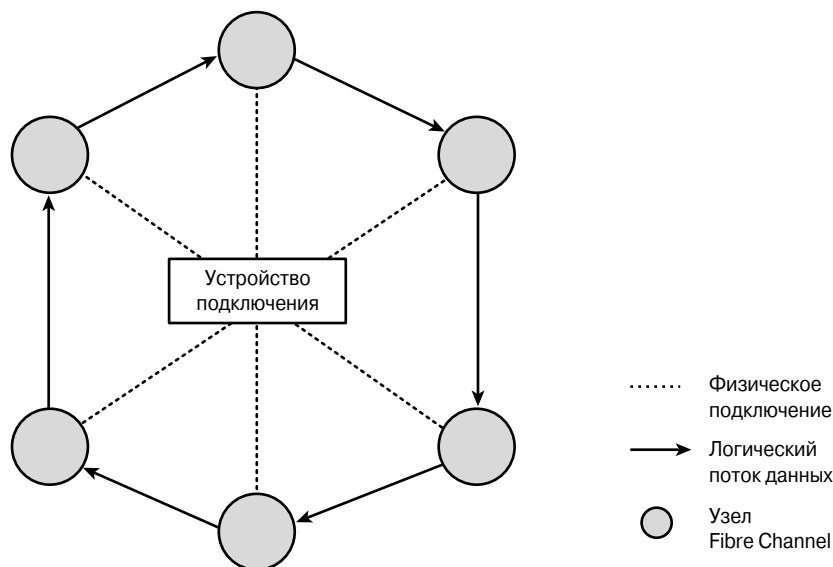


Рис. 4.2. Топология кольца Fibre Channel с разделением доступа

содержат до 12 узлов, а после подключения 50 узлов производительность падает до такой степени, что имеет смысл перейти на коммутатор связанной архитектуры. Кольца с разделяемым доступом позволяют использовать все преимущества Fibre Channel при значительно меньших денежных затратах. Однако стоимость коммутаторов Fibre Channel стремительно снижается, поэтому применение коммутируемых связанных архитектур становится все более предпочтительным.

В разделах 4.4.2.1–4.4.2.3 описываются важные концепции, связанные с кольцом Fibre Channel с разделяемым доступом: инициализация кольца, управление кольцом и различные типы колец (закрытые и открытые кольца).

4.4.2.1 Инициализация кольца Fibre Channel

Протокол инициализации кольца Fibre Channel обладает четко определенной структурой. Инициализация выполняется в таких случаях:

- при первой установке и запуске кольца;
- при подключении нового устройства;
- при отключении или перезапуске существующего устройства.

Кольца инициализируются с помощью специальных управляющих кадров. В процессе инициализации выполняются описанные ниже действия.

- На время до следующей инициализации выбирается хозяин кольца (loop master). Одной из наиболее важных задач хозяина кольца является назначение адресов различным портам в кольце.
- Выполняется назначение адресов в кольце. Значения адресов, присвоенных устройствам, используются для решения конфликтов, когда несколько устройств пытаются одновременно передавать данные. Коммутаторы связанных архитектуры имеют наибольший приоритет; они могут выбрать любой произвольный адрес. Портам разрешается получать адрес, который использовался ранее.

4.4.2.2 Разделение доступа в кольце Fibre Channel

Если через один порт организуется передача данных на другой порт, сначала следует выяснить владельца кольца. Для этого порт отправляет специальный кадр, который называется **ARB primitive**. В этом кадре содержится адрес порта, который намерен получить управление над кольцом. При получении портом кадр **ARB primitive** передается дальше, если порт не собирается передавать данные. Если порт-получатель намерен сам передавать данные, он сравнивает адрес из **ARB primitive** с собственным адресом. Если у получателя адрес имеет меньшее значение (большой приоритет), то кадр отправляется дальше, но содержит адрес данного порта. В противном случае кадр **ARB primitive** отправляется дальше с неизменным адресом. Через некоторое время порт получает кадр **ARB primitive** со своим адресом обратно. Именно на этом этапе порт получает управление кольцом.

Порт, который выиграл управление кольцом, отправляет кадр **OPN primitive** тому порту, которому нужно передать данные. Этот кадр передается дальше промежуточными портами, пока не достигнет целевого порта. Порт отвечает другим кадром **ARB primitive**, после получения которого портом-инициатором может начинаться сеанс передачи данных. После завершения отправки данных порт-инициатор отправляет кадр **CLS primitive**. Однако целевой порт может продолжать отправку кадров для завершения сеанса передачи данных, а порт-инициатор должен быть готов принять эти кадры даже после отправки **CLS primitive**. После завершения отправки кадров целевой порт может ответить на кадр **CLS primitive** собственным кадром **CLS primitive**. На этом этапе кольцо готово к началу следующего сеанса передачи данных.

В стандарте Fibre Channel описан необязательный алгоритм предотвращения ситуации, когда устройства с низким приоритетом никогда не получают доступ к кольцу из-за устройств с высоким приоритетом. Этот необязательный алгоритм запрещает устройствам с высоким приоритетом получать

кольцо в управление сразу после завершения сеанса связи, пока возможность передачи данных не появится у устройств с низким приоритетом.

4.4.2.3 Открытые и закрытые кольца

Закрытое кольцо представляет собой изолированное кольцо Fibre Channel с разделением доступа (кольцо не подключено к связанной архитектуре, которая рассматривается в разделе 4.4.3). *Открытое кольцо* — это кольцо с разделением доступа, подключенное к связанной архитектуре через коммутатор. В данном случае обеспечение совместимости зависит непосредственно от коммутатора, как описано ниже.

- Коммутатор связанной архитектуры должен обеспечивать уникальность адресов в пределах кольца и связанной архитектуры, а также заменять трехбайтовые адреса в связанной архитектуре на однобайтовые адреса в пределах кольца. Этот процесс обычно называется *подменой адресов*.
- От коммутатора связанной архитектуры требуется, чтобы подключенные к ней устройства в пределах кольца с разделением доступа выглядели как подключенные к кольцу.
- Обнаружение устройств необходимо для поддержки передачи данных в рамках общей и частной транспортной инфраструктуры. Устройства Fibre Channel, поддерживающие связанную архитектуру, самостоятельно регистрируются на сервере SNS — Simple Name Server (см. раздел 4.4.3.1), однако устройства Fibre Channel для кольца с разделением доступа не поддерживают такую возможность. На коммутатор возлагается задача по обнаружению каждого устройства в кольце Fibre Channel и добавлению характеристик устройства на SNS. Это необходимо для устройств Fibre Channel, поддерживающих работу в рамках связанной архитектуры.

4.4.3 Коммутируемая связанная архитектура

В топологии коммутируемой связанной архитектуры Fibre Channel (Fibre Channel switched-fabric) каждое устройство имеет логическое подключение к любому другому устройству. Обратите внимание на слово *логическое*. Обеспечение физического подключения устройств по топологии “каждый с каждым” потребовало бы огромных затрат, так как для N устройств необходимо N^2 портов и физических подключений. В реальности каждое устройство подключается к коммутатору, а коммутатор поддерживает логические подключения между всеми своими портами.

На рис. 4.3 показан простейший вариант топологии коммутируемой связанной архитектуры Fibre Channel. Несколько узлов (устройств хранения и компьютерных систем) подключены к коммутатору Fibre Channel. Коммутатор —

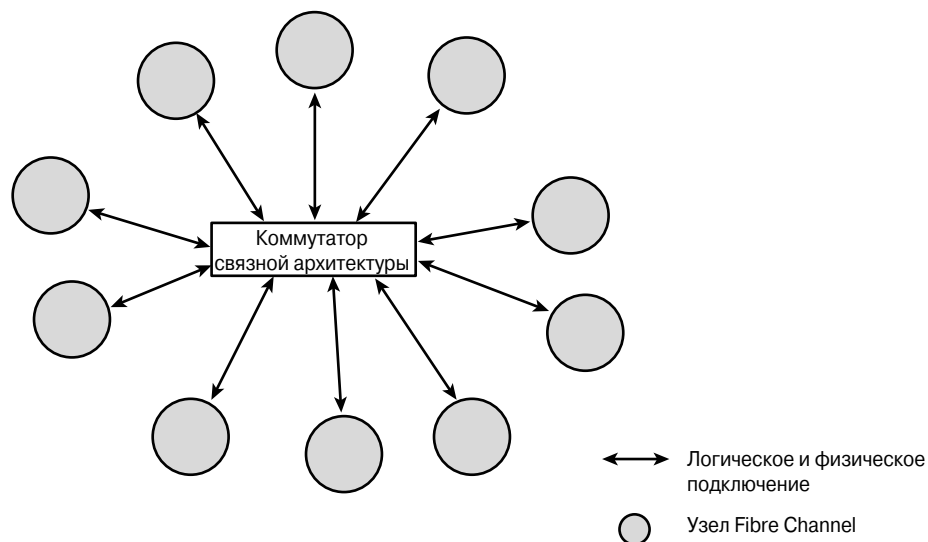


Рис. 4.3. Топология коммутируемой связанной архитектуры Fibre Channel

это высокоскоростное устройство, которое обеспечивает подключение по схеме “каждый с каждым” и обрабатывает несколько одновременных подключений. Кроме того, коммутатор поддерживает такие службы, как Fabric Login (см. раздел 4.4.3.6). Коммутаторы могут быть подключены каскадно (в виде иерархии) или в виде сети, что позволяет формировать более сложные конфигурации. Очень часто строится трехуровневая иерархия, на самом нижнем уровне которой размещены кольца с разделением доступа Fibre Channel (FC-AL), подключенные к малопроизводительным коммутаторам. Эти коммутаторы, в свою очередь, подключаются к высокоскоростным коммутаторам, обеспечивающим максимально возможную пропускную способность.

Несколько коммутаторов могут быть соединены друг с другом. В топологии связанной архитектуры используется 3-байтовый идентификатор (24 бит) для уникальной идентификации каждого устройства, что *теоретически* допускает подключение по схеме “точка–точка” для 2^{24} устройств (примерно 15 млн.). Конечно, на практике сети хранения данных включают в себя намного меньше устройств.

Поскольку количество потенциальных устройств может быть достаточно большим, от сетей хранения данных на базе Fibre Channel требуется предоставление служб по управлению адресами, а также дополнительных служб; в противном случае ручное управление сетями хранения данных было бы невозможным. Учитывая, что обеспечение взаимодействия Fibre Channel и протоколов верхнего уровня (например, SCSI) проводится самими узлами, а ад-

ресное пространство весьма велико, динамическое управление адресами на инфраструктурном уровне будет обязательным (не требуя ручной настройки).

Топология связанной архитектуры реализуется на основе коммутатора связанной архитектуры (см. раздел 4.4.3.5). Кроме предоставления подключений “точка–точка”, коммутатор предоставляет такие службы, как SNS (Simple Name Server), RSCN (Registered State Change Notification), FAN (Fabric Address Notification), Broadcast Server IP-FC, Principal Switch и Fabric Login. Эти службы кратко рассматриваются в разделах 4.4.3.1–4.4.3.6. Обратите внимание, что службы основаны на модели клиент/сервер. Серверный компонент реализован в виде коммутатора связанной архитектуры, а клиентский компонент представляет собой такие устройства, как адаптеры шины и контроллеры RAID.

4.4.3.1 Сервер SNS

Этот сервер представляет собой базу данных коммутатора связанной архитектуры. База данных отслеживает два типа информации.

1. Имя устройства и информация об адресе.
2. Поддерживаемые протоколы высокого уровня.

Сервер SNS позволяет инициатору сеанса связи, например адаптеру шины сервера, делать запрос к службе имен и получать список доступных устройств хранения данных. Это обеспечивает автоматическое управление адресами, что имеет особое значение, поскольку адресное пространство сетей хранения данных поддерживает до 15 млн. адресов.

При инициализации устройство связанной архитектуры отправляет базовый пакет регистрации (в терминологии Fibre Channel он называется *FLOGI*) на известный адрес 0xFFFFFE. Запрос на регистрацию вместо адреса заполняется нулевым значением, что обозначает запрос адреса.

Как только устройство подключается к коммутатору, иницируется другой базовый запрос (*PLOGI*), отправляемый серверу SNS, который имеет фиксированный адрес 0xFFFFFC. Устройство отправляет соответствующую информацию, например адрес в формате WWN (World Wide Name), тип порта, адрес коммутатора, протоколы верхнего уровня (SCSI), которые поддерживаются устройством, и другие служебные параметры.

Некоторые устройства не отправляют SNS всю информацию о себе. Отдельные модели коммутаторов отправляют устройствам запросы целевого определения и обновляют значения соответствующих полей в базе данных SNS.

4.4.3.2 Служба RSCN

Служба RSCN (Registered State Change Notification) позволяет устройству регистрироваться для получения уведомлений о подключении к коммутатору или отключении от него других устройств. Например, сервер может зарегистрироваться для получения уведомлений при подключении к коммутатору или отключении от него связанной архитектуры устройств хранения. Таким образом, RSCN дополняет сервер SNS: сервер позволяет компоненту Fibre Channel (например, узлу) обнаруживать другие компоненты (например, единицы хранения). В то же время служба RSCN позволяет узлу (компоненту Fibre Channel) регистрироваться для получения уведомлений об изменениях в составе компонентов Fibre Channel.

Пакеты RSCN могут прерывать передачу данных. Кадры RSCN используются получателем для обновления собственных таблиц маршрутизации. Некоторые поставщики предоставляют возможности по остановке широковещательного распространения пакетов RSCN на уровне коммутатора. Хотя это полезная функция, существует опасность, что из-за неправильной настройки кадр RSCN никогда не достигнет устройства, которое должно его получить.

4.4.3.3 Служба FAN

Служба FAN (Fabric Address Notification) обеспечивает ликвидность транзакций, выполняемых устройствами. Обычно это необходимо в том случае, когда инициализация кольца прерывает активную передачу данных между двумя компонентами Fibre Channel.

4.4.3.4 Служба Broadcast Server

Интерфейс Fibre Channel поддерживает протоколы более высокого уровня, требования для реализации которых должны быть выполнены. К одному из таких протоколов относится IP. Так как в рамках сетей IP поддерживается вторичный протокол ARP (Address Resolution Protocol), посредством которого устройства отправляют широковещательные запросы, этот же протокол должен поддерживаться и устройствами Fibre Channel.

4.4.3.5 Служба Principal Switch

При использовании в сети хранения данных на основе Fibre Channel нескольких коммутаторов связанной архитектуры необходим протокол выбора одного из коммутаторов в качестве главного коммутатора (principal switch). Выбранный главный коммутатор будет работать как коммутатор с адресом

0xFFFFFE и обеспечивать уникальность адресов устройств в сети хранения данных на основе Fibre Channel.

4.4.3.6 Служба Fabric Login

Устройства связанной архитектуры всегда стремятся воспользоваться службой Fabric Login. Для этого на фиксированный адрес 0xFFFFFE отправляется специальный кадр *FLOGI*. Кадр FLOGI содержит адрес отправляющего устройства, который устанавливается в нулевое значение, если пакет используется для запроса нового адреса.

Служба Fabric Login позволяет выполнять следующие задачи:

- получение информации о параметрах связанной архитектуры, например списков поддерживаемых служб;
- назначение адресов;
- инициализация буфера для контроля передачи данных.

4.5 Типы портов Fibre Channel

В стандарте Fibre Channel определено несколько типов портов, зависящих от топологии сети хранения данных и от устройства, к которому относится порт. Различные типы портов представлены в табл. 4.2.

Таблица 4.2. Типы портов Fibre Channel

Тип порта	Устройство, поддерживающее порт	Описание	Типы портов, которые можно подключать к этому порту
F	Коммутатор связанной архитектуры	Может работать только в рамках сети на базе связанной архитектуры. Никогда не является источником или точкой назначения данных. Работает как посредник для обеспечения связи между устройствами	N
FL	Коммутатор связанной архитектуры	Может работать только в кольце. Обеспечивает подключение кольца Fibre Channel к сети на базе связанной архитектуры. Никогда не является источником или точкой назначения данных	NL

Окончание табл. 4.2

Тип порта	Устройство, поддерживающее порт	Описание	Типы портов, которые можно подключать к этому порту
E	Коммутатор связанной архитектуры	Порт расширения, который используется для подключения коммутаторов связанной архитектуры друг к другу. Никогда не является источником или точкой назначения данных. Выполняет роль посредника	E
N	Адаптер шины или устройство хранения	Не поддерживает работу в кольце. <i>Всегда</i> представляет собой источник или точку назначения данных. Реализован в сетях “точка–точка” или в сетях на базе связанной архитектуры	F или N
NL	Адаптер шины или устройство хранения	Аналогичен типу N, однако поддерживает работу в кольце. <i>Всегда</i> представляет собой источник или точку назначения данных	NL или FL
U		Универсальный неиспользуемый порт. После активизации может выступать в роли портов E, F или FL	Нет
L		Универсальный термин для описания портов NL или FL	
G		Может выступать в роли портов E или F	E или N

Чтобы облегчить жизнь администраторам сетей хранения данных, поставщики устройств обычно предоставляют порты с автоматическим конфигурированием. Такое конфигурирование проводится следующим образом.

1. Порт пытается инициализироваться как порт кольца (FL). Если инициализация проходит успешно, порт настраивается как порт FL.

2. Если инициализация кольца завершается неудачно, делается попытка инициализации в виде порта E; таким образом, проверяется возможность подключения к сети на базе связанной архитектуры.
3. Если инициализация в виде порта E завершается неудачно, порт пытается активизировать службу Fabric Login. Если регистрация выполняется успешно, порт иницируется как порт F.

Не стоит рассчитывать, что все поставщики устройств устанавливают на них динамически конфигурируемые порты, которые могут работать вместо любого порта. Такая возможность встроена лишь в некоторые устройства. Это позволяет переконфигурировать порты без перезагрузки или отключения питания. Остальные устройства для получения определенного типа порта требуют замены прошивки или установки платы расширения, что может потребовать выключения питания.

4.6 Протокол Fibre Channel

В общем контексте *Fibre Channel* — это набор стандартов, разработанных в Национальном институте стандартизации США. Интерфейс Fibre Channel предоставляет высокопроизводительное последовательное подключение между хостом и единицами хранения, а также между самими единицами хранения. Стандарт позволяет обеспечить высокоскоростную передачу данных в сетях с топологией “точка–точка” и кольцо. Более того, Fibre Channel предоставляет все эти возможности вместе с проверкой ошибок.

В стандарте Fibre Channel определено пять функциональных уровней: от FC-0 до FC-4. Уровни рассматриваются в разделах 4.6.1–4.6.5. Обратите внимание, что по практическим соображениям уровни FC-0, FC-1 и FC-2 реализуются аппаратно.

4.6.1 Уровень FC-0

Определяет физические характеристики интерфейса и носителя. В частности, посредством FC-0 определяются спецификации уровней сигналов, носителя и получателей/отправителей. Уровень FC-0 позволяет использовать несколько интерфейсов, что дает возможность выбирать разные скорости передачи данных и различные передающие среды. В качестве примера физической передающей среды можно привести медный провод, однопроводный и многопроводный кабели. Скорость передачи варьируется от 12,5 до 106,25 Мбайт/с.

Те, кто знаком с семиуровневой сетевой моделью ISO OSI², могут заметить, что FC-0 соответствует седьмому уровню модели ISO OSI.

4.6.2 Уровень FC-1

Определяет схемы кодирования и декодирования данных, сигналов и специальных символов, а также управление ошибками. Кроме того, уровень FC-1 отвечает за обслуживание линий связи.

Уровень FC-1 использует схему кодирования, которая называется *8B/10B*. Схема проектировалась для того, чтобы обеспечить следующее:

- эффективную синхронизацию данных;
- расширенное обнаружение ошибок;
- эффективное обнаружение управляющих символов;
- упрощенное проектирование аппаратного обеспечения приемников/передатчиков.

Схема кодирования 8B/10B преобразует каждые 8 бит в два возможных значения, объемом 10 бит. Эти 10 бит используются в виде *Ann.m*, где *A* — значение *K* для индикации команды или *D* для индикации данных; *nn* — десятичное значения последних пяти битов байта; *M* — десятичное значение первых трех битов байта.

Два возможных значения появляются потому, что посредством спецификации выбирается одно из значений для кодирования данных при их передаче на базе недавней истории передачи. Это необходимо для того, чтобы обеспечить минимальное количество переходов состояния (между 0 и 1), что повысит эффективность передачи. Недавняя история передачи называется *динамическим рассогласованием*.

Как уже отмечалось, все данные кодируются с помощью 10 бит. Некоторые неиспользованные 10-битовые символы (в контексте данных) применяются для отделения фреймов и сигналов, включая сигналы о готовности порта для принятия данных, а также другие типы сигналов. Основное внимание уделяется обнаружению и исправлению ошибок на этапе передачи. Данные Fibre Channel всегда передаются группами по 4 байта, которые называются *словами передачи* (transmission words).

4.6.3 Уровень FC-2

Определяет передачу данных от одного узла к другому, т.е. непосредственно транспортный механизм. Уровень FC-2 формирует кадры, опреде-

²ISO — International Organization for Standardization (Международная организация по стандартизации);

OSI — Open System Interconnection (взаимодействие открытых систем).

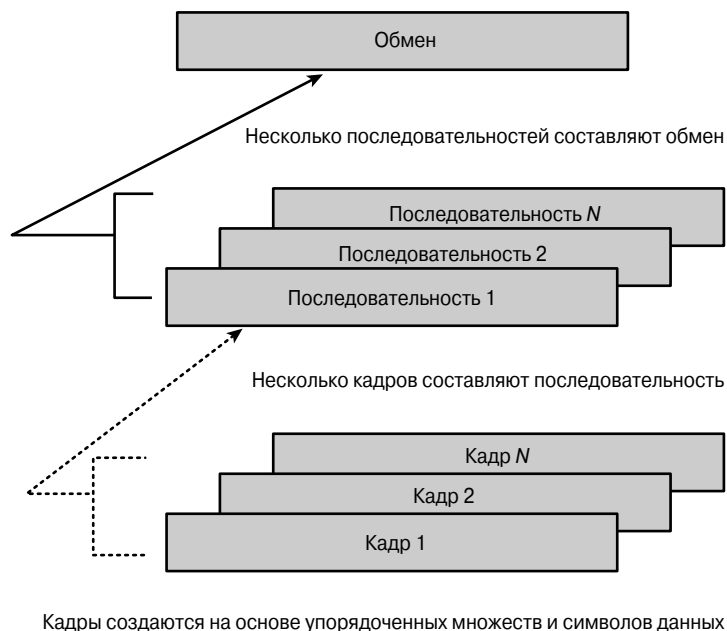


Рис. 4.4. Иерархия передачи данных Fibre Channel

ляет классы обслуживания и службы регистрации связной архитектуры или портов. Этот уровень можно представить в качестве аналога уровню MAC (Media Access Control) в модели ISO OSI.

Уровень FC-2 определяет:

- иерархию передачи данных Fibre Channel, которая включает в себя упорядоченные множества, кадры, последовательности и обмены;
- управление потоком Fibre Channel;
- протоколы FC-2;
- классы обслуживания FC-2.

Эти компоненты рассматриваются в разделах 4.6.3.1–4.6.3.7, но перед этим следует выяснить, как соединяются различные структурные элементы иерархии.

В Fibre Channel данные передаются с помощью *кадров* (frame). Кадр представляет собой эквивалент пакета TCP/IP. Кадры создаются из *упорядоченных множеств* и символов данных. Несколько кадров группируются вместе и формируют *последовательность*, а несколько последовательностей формируют *обмен* (exchange). Это демонстрируется на рис. 4.4.

В разделах 4.6.3.1–4.6.3.4 все эти компоненты рассматриваются более подробно.

4.6.3.1 Упорядоченные множества Fibre Channel

Упорядоченные множества — это структуры последовательной передачи данных объемом 4 байта, которые представляют собой специальные символы или сигналы линии связи. Далее приведены примеры подобных множеств.

- Разделители кадров SOF (Start Of Frame) и EOF (End Of Frame), которые являются аналогами пакетов SOF и EOF в сетях Ethernet. В отличие от Ethernet, в Fibre Channel определено несколько вариантов SOF и EOF, поскольку уровнем FC-1 используется схема кодирования, формирующая несколько представлений для каждого передаваемого символа.
- Два базовых сигнала для указания состояния порта.
 - *Idle* — указание, что порт готов для передачи или приема данных.
 - *Receiver Ready* — указание, что буфер интерфейса (устройства взаимодействия) готов для приема данных.
- Базовая последовательность. Простое упорядоченное множество, которое регулярно передается для указания особого статуса порта. К особым статусам относятся:
 - *Not Operational* (NOS) — используется только в сетях с топологией “точка-точка” или в связанной архитектуре (но не в кольце с разделением доступа) для указания на отказ в работе линии связи или появление определенной ошибки;
 - *Offline* (OLS) — передается во время инициализации порта или при получении базового статуса NOS; таким образом, в ответ на NOS порт отправляет ответ OLS;
 - *Link Reset* (LR) — используется для указания на необходимость повторной инициализации линии связи;
 - *Link Reset Response* (LS) — используется для указания, что данные LR получены и обработаны.

4.6.3.2 Кадр Fibre Channel

Как пакет IP является базовым элементом протокола Internet (IP), так и кадр представляет собой основной структурный элемент интерфейса Fibre Channel. Существует три типа кадров.

1. *Кадры управления линией связи* (link control frames), используемые для отправки команды управления линией связи.

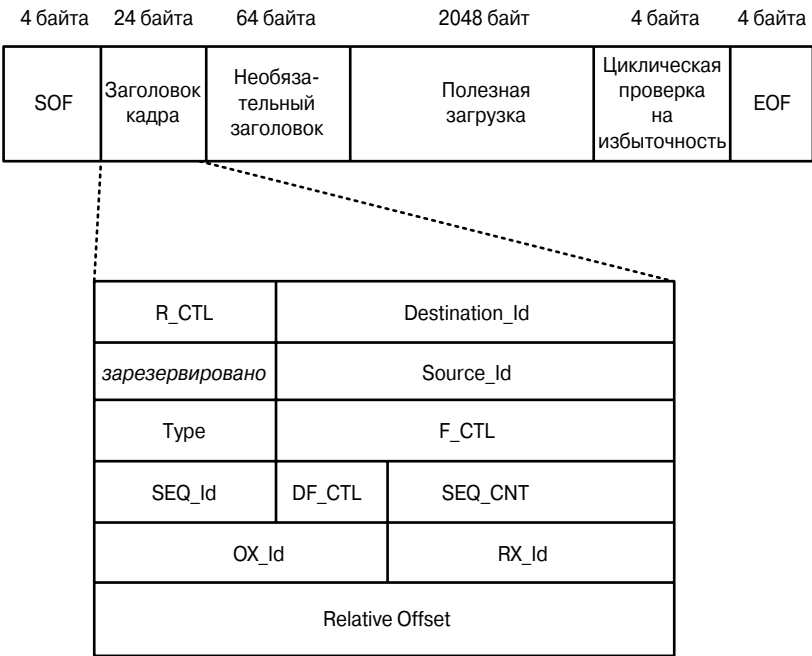


Рис. 4.5. Заголовок кадра Fibre Channel

- 2. *Кадры данных линии связи (link data frames)*, используемые для отправки данных, необходимых для управления линией связи.
- 3. *Кадры данных устройства (device data frames)*, которые содержат данные для протоколов более высокого уровня, например данные, считанные с жесткого диска.

На рис. 4.5 показан заголовок кадра Fibre Channel. Кадр проектировался для передачи 2048 байт данных и необязательного заголовка размером 64 байт. Подобный размер кадра позволяет передавать за один раз большой объем данных с минимальными накладными расходами (около 1,5%). Однако при этом другому узлу придется ждать, пока завершится передача большого кадра, что подразумевает увеличение задержек в передаче. Сравним это с протоколом АТМ (Asynchronous Transfer Mode), где кадр имеет размер 53 байт и накладные расходы протокола составляют около 10%. Это позволяет снизить задержки, но время передачи определенного объема данных возрастает.

Каждый кадр начинается и заканчивается специальным ограничителем, как и в других сетевых протоколах. Это SOF и EOF соответственно. Каждый кадр имеет заголовок, выполняющий несколько функций. Одна из них — это предоставление адресов назначения и источника для обеспечения коммута-

ции данных. Еще одна задача состоит в переносе информации для управления линией связи, включая управление самой передачей.

Другие поля заголовка кадра Fibre Channel рассматриваются ниже.

- Поле **Destination_Id** используется для маршрутизации кадра. В топологиях “точка–точка” и кольцо с разделением доступа маршрутизация может проходить обычным образом, что не относится к топологии коммутируемой связной архитектуры. Поле **Source_Id** предназначено для передачи сообщений об ошибках и для предотвращения циклов данных при маршрутизации.
- Поля **R_CTL** и **Type** используются для сортировки различных кадров уровня FC-4 по их прибытию в точку назначения. Таким образом, эти поля указывают, содержит ли прибывший кадр данные интерфейса SCSI, протокола IP или другие данные. Значения поля **Type** описываются в табл. 4.3.
- Поле **R_CTL** используется для указания содержимого кадра. Кадр может содержать данные или информацию для управления линией связи; в последнем случае кадры могут быть запрошенными и незапрошенными.

Таблица 4.3. Значения поля **Type** в кадрах Fibre Channel

Значение	Описание
0x00	Базовое обслуживание линии связи
0x01	Порт Fibre Channel — расширенное обслуживание линии связи
0x02	Порт памяти
0x03	Инициализация памяти
0x04	Интерфейс IEEE 802.2
0x05	Протокол Internet (IP)
0x06	Подчиненное устройство IPI (Intelligent Peripheral Interface)
0x07	Основное устройство IPI
0x08	Инициатор SCSI
0x09	Целевое устройство SCSI
0x0A	Интерфейс HIPPI
0x0B	Код SBCCS (Single-Byte Command Code Sets)
0x0C	Зарезервировано для новых типов шины
0x10	Зарезервировано
0xD0	Уникально для производителя
0xF0	Прямой канал

- Поле `F_CTL` используется для описания информации кадра, например первой или последней последовательности.
- Поле `DF_CTL` указывает на присутствие или отсутствие необязательных заголовков.
- Поля `SEQ_Id` и `SEQ_CNT` уникально идентифицируют счетчик последовательности обмена (см. раздел 4.6.3.3).
- Поле `OX_Id` (идентификатор обмена источника) используется для связывания кадра с определенным обменом исходного порта.
- Поле `RX_Id` (идентификатор обмена ответчика) используется для связывания кадра с определенным обменом отвечающего порта.
- Поле `Relative Offset` идентифицирует относительное смещение первого байта основного содержания кадра от базового адреса.

4.6.3.3 Последовательность Fibre Channel

Последовательность представляет собой набор кадров, которые передаются из одной точки в другую. Для исправления возможных ошибок каждый кадр содержит уникальный счетчик последовательности. Исправление ошибок осуществляется протоколом более высокого уровня, обычно на уровне FC-4. Обратите внимание, что все кадры в последовательности передаются в одном направлении (а не в обоих одновременно). На рис. 4.4 показано, как кадры, последовательности и обмены взаимодействуют друг с другом.

4.6.3.4 Обмен Fibre Channel

Несколько последовательностей составляют *обмен* (exchange). Обмены представляют собой последовательности двусторонних направлений; т.е. в обмен входят последовательности данных, передающихся в разных направлениях, хотя каждая последовательность передается только в одном направлении. При каждом обмене только одна последовательность может быть активна в текущий момент времени. Но, так как одновременно могут быть активны несколько обменов, различные последовательности из этих обменов также могут быть активны одновременно.

Каждый обмен выполняет одну функцию, например реализует команду SCSI Read.

4.6.3.5 Управление потоком Fibre Channel

Конечные узлы Fibre Channel взаимодействуют непосредственно друг с другом и не создают сеансовых подключений к промежуточным узлам.

Устройства не “подозревают” о коммутаторе связной архитектуры и концентраторе. Конечно, при этом коммутаторы и концентраторы обмениваются пакетами управления потоком с этими устройствами.

Механизм управления потоком требует, чтобы передающий порт не отправлял кадры быстрее, чем принимающий порт может их обработать. Порты Fibre Channel имеют буфера для временного хранения кадров и последующей их обработки. Под обработкой подразумевается отправка кадра на другой порт или передача кадра протоколу более высокого уровня. Схема управления потоком, которая используется в Fibre Channel, очень напоминает протокол плавающего окна в ТСП/IP. Размер окна, представляющий собой количество кадров, которые можно отправить без подтверждения их приема, устанавливается сторонами обмена заранее. При этом согласованное значение не может быть изменено. Для каждого отправленного кадра размер окна уменьшается на единицу, а для каждого подтвержденного кадра — увеличивается на единицу. Управление потоком может осуществляться одним из двух способов: “точка–точка” или “буфер–буфер”. При этом требуется совместное использование обоих методов.

Управление потоком по схеме “точка–точка” осуществляется между двумя конечными точками: источником данных (например, сервером) и получателем данных (например, жестким диском). Управление потоком “точка–точка” проводится между двумя портами типа N (между ними могут находиться промежуточные узлы). Два порта типа N регистрируют друг друга, в процессе чего каждый порт выделяет для себя определенное количество буферов у другого порта. Это количество называется *резервированием буфера*. Отправитель может отправлять количество кадров, не превышающее это значение. Получатель отправляет кадр АСК (позитивное подтверждение) для каждого успешно полученного и обработанного кадра, а отправитель при получении кадра АСК может увеличить значение резервирования (credit count) на единицу для каждого полученного кадра АСК. Получатель может подтвердить успешное получение нескольких кадров или даже целой последовательности и получатель должен будет, не ожидая подтверждения для каждого кадра в отдельности, увеличить количество кадров, которое можно отправлять.

Управление потоком по схеме “буфер–буфер” выполняется между двумя соседними узлами, которые представляют собой промежуточные узлы или находятся между конечным и промежуточным узлом. Таким образом, управление потоком от буфера к буферу выполняется между портами типа N или между портом F и портом N. Как уже отмечалось, порты обмениваются данными, указывающими на количество буферов, зарезервированных для каждого узла. Эти значения могут отличаться, например один порт может выделить два буфера, а второй — четыре буфера. Получение кадра подтвер-

ждается кадром *Receiver Ready*, а не кадром АСК как в управлении потоком “точка–точка”.

4.6.3.6 Протоколы FC-2

В стандартах Fibre Channel определены протоколы для управления передачей данных и линией связи. Кроме того, описаны дополнительные стандарты для поддержки протоколов более высокого уровня, применяемых на уровне FC-4. Эти протоколы описаны ниже.

- Протокол *Fabric Login*, который определяет обмен параметрами между портом и коммутатором связанной архитектуры. Протокол и служба Fabric Login более подробно описываются в разделе 4.4.3.6.
- Протокол *Port Login*, требующий, чтобы независимо от топологии (“точка–точка”, кольцо с разделением доступа или коммутируемая связанная архитектура) два порта проводили взаимную регистрацию перед подключением друг к другу. Взаимная регистрация выполняется с помощью специального кадра *PLOGI*. С помощью протокола Port Login обеспечивается использование двух важных функций.
 - Возможность получения информации о порте N, на котором выполняется регистрация. К такой информации относится описание классов обслуживания, поддерживаемых портом N.
 - Инициализация буфера резервирования для управления потоком “точка–точка”. Обратите внимание, что в контексте прямого подключения управление потоком “точка–точка” ничем не отличается от управления потоком “буфер–буфер”.
- Протокол *Data Transfer*, определяющий, как данные протокола верхнего уровня (уровня FC-4) передаются с помощью схем управления потоком, описанных в разделе 4.6.3.5.
- Протокол *Arbitrated Loop*, который определяет методы инициализации и управления кольцом.

4.6.3.7 Классы обслуживания FC-2

Интерфейс Fibre Channel проектировался для обеспечения различных способов передачи данных. Ряд служб отличается такими характеристиками:

- тип сервисного подключения, т.е. аналогично TCP или без установки подключения, как в UDP;
- поддержка многоадресной доставки (multicast);
- поддержка уведомления о доставке или невыполненной доставке;

- поддержка гарантированной доставки кадров в том же порядке, в котором они были отправлены;
- тип предоставляемых служб, например резервирование пропускной способности для соединения, если служба ориентирована на соединение;
- тип механизмов управления потоком данных.

Для предоставления широкого диапазона вариантов передачи данных, определено несколько классов обслуживания.

- Тип *Class 1* определяет выделенное подключение, подобное подключению TCP/IP. Как и в TCP, Class 1 гарантирует, что кадры доставляются в той же последовательности, в которой они были отправлены. Тип Class 1 используется при передаче больших объемов данных, когда время, потраченное на установку соединения, на порядок меньше времени, необходимого для передачи данных.
- Тип *Class 2* определяет обслуживание без подключения (по аналогии с дейтаграммами), при котором кадры потенциально могут быть доставлены не в той последовательности, в которой они отправлялись (что подразумевает смену последовательности кадров протоколом более высокого уровня). Как и в случае использования сетевых протоколов, обслуживание Class 2 имеет смысл тогда, когда объем передаваемых данных достаточно мал и накладные расходы на установку соединения сравнимы с расходами на передачу самих данных. Получатель кадра Class 2 должен отправить подтверждение при получении кадра.
- Тип *Class 3* также подразумевает обслуживание без установки подключения. Основное отличие от Class 2 состоит в том, что подтверждать успешное получение кадра нет необходимости. Это сравнимо с дейтаграммами IP, метод использования которых иногда в шутку называется “отправь и молись”.
- Тип *Class 4*, который также называется *Intermix*, является необязательным классом обслуживания. Класс гарантирует определенную пропускную способность кадрам Class 1, а оставшаяся пропускная способность используется для кадров Class 2 и Class 3.
- Тип *Class 6* представляет собой однонаправленное, ориентированное на подключение обслуживание с предоставлением возможности многоабонентской доставки (Class 5 зарезервирован).

В табл. 4.4 собрана вся информация о классах обслуживания Fibre Channel.

Обратите внимание, что большинство поставщиков поддерживают классы 1, 2 и 3. В то же время некоторые поставщики поддерживают только классы, не ориентированные на соединение (Class 2 и Class 3).

Таблица 4.4. Классы обслуживания Fibre Channel

	Class 1	Class 2	Class 3	Class 4	Class 6
Тип обслуживания	Ориентирован на подключение	Не поддерживает подключение	Не поддерживает подключение	Ориентирован на подключение	Ориентирован на подключение
Поддержка многоадресной доставки	Нет	Нет	Да	Нет	Да
Подтверждение доставки	Нет	Нет	Да	Нет	Да
Использование пропускной способности	Полное	Не гарантируется	Не гарантируется	Частичное	Не гарантируется
Управление потоком	Точка-Точка	Точка-Точка	Буфер-Буфер	Буфер-Буфер	Точка-Точка
Доставка кадров	Гарантируется	Порядок может нарушаться	Порядок может нарушаться	Гарантируется	Гарантируется

4.6.4 Уровень FC-3

Определяет общие службы, включая службы по управлению и общему транспортному механизму. Уровень FC-3 — общий для всех портов узла. Уровни FC-1, FC-2 и FC-4 реализованы отдельно для каждого порта. В этой схеме поддерживается использование разными портами различной конфигурации (рис. 4.6). Например, один порт может передавать данные SCSI, а другой в это же время будет передавать данные ATM.

Кроме того, обратите внимание, что уровни FC-0, FC-1 и FC-2 относятся к различным портам, а уровень FC-3 относится к узлу. Верхний уровень FC-4 также относится к порту. На рис. 4.6 показан узел с четырьмя портами и двумя протоколами верхнего уровня — SCSI и IP. Если бы узел поддерживал больше протоколов верхнего уровня, на диаграмме присутствовали бы дополнительные блоки FC-4.

Далее представлены некоторые функции, реализованные на уровне FC-3.

- *Транкинг (trunking) или канальное уплотнение (striping)*, при котором параллельные линии связи и порты “сворачиваются” для обеспечения большей пропускной способности между узлами.
- *Многоабонентская доставка*, при которой единая передача данных может быть направлена одновременно к нескольким портам. Реализуется посредством службы регистрации на коммутаторе связанной архитектуры, с помощью которой узлы регистрируются и отменяют регистрацию для многоабонентской доставки. Доставка в этом случае очень похожа на режим групповой отправки (multicast) в IP. Обратите внимание: многоабонентская доставка может осуществляться на все порты коммутато-

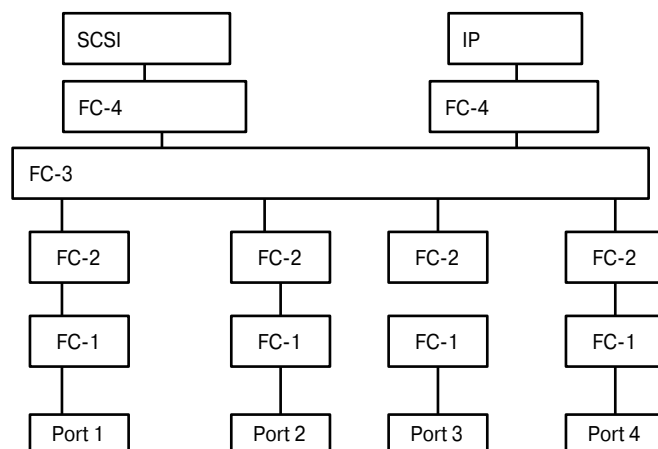


Рис. 4.6. Уровни Fibre Channel

ра связанной архитектуры, т.е. может использоваться как широковещание (broadcast).

- *Свободный поиск* (hunting), при котором несколько портов используют одинаковый псевдоним. Это похоже на объединение нескольких телефонных линий компании под одним номером, когда звонок на этот номер перенаправляется на одну из свободных линий. Очевидным преимуществом такой схемы будет повышение вероятности получить свободный порт.

Компания Brocade одна из многих реализует функцию транкинга в своих продуктах, хотя функция получила названия ISL (Inter Switch Link) Trunking. Название подчеркивает высокую производительность коммутаторов Brocade, которые предназначены для подключения к другим коммутаторам, а не к серверам и устройствам хранения данных. Обратите внимание, что транкинг поддерживается и другими поставщиками.

4.6.5 Уровень FC-4

Определяет связывание протоколов верхнего уровня с Fibre Channel. Вот эти протоколы (со временем будет реализована поддержка и других протоколов):

- SCSI;
- IP;
- IPI (Intelligent Peripheral Interface);
- HIPPI (High-Performance Parallel Interface);
- IEEE 802.2;
- SBCCS (Single-Byte Command Code Sets);
- AAL5 (ATM Adaptation Layer);
- FC-LE (Link Encapsulation).

Обратите внимание: одна линия связи Fibre Channel может одновременно передавать пакета данных нескольких протоколов верхнего уровня.

4.7 Структурные элементы SAN

В разделах 4.4 и 4.6 приведен обзор топологий и протокола Fibre Channel. Теперь рассмотрим различные устройства и компоненты, которые используются для создания сетей хранения данных Fibre Channel. К основным структурным элементам SAN относятся:

- адаптеры шины;
- кабели Fibre Channel;
- разъемы;
- устройства подключения, в число которых входят концентраторы, коммутаторы и коммутаторы связанной архитектуры.

Все эти элементы рассматриваются в разделах 4.7.1–4.7.4. Обратите внимание, что все адресуемые компоненты в пределах сети хранения данных на основе Fibre Channel имеют уникальные имена WWN (World Wide Names), которые представляют собой аналоги уникальных адресов MAC. Имя WWN в спецификации Fibre Channel — это 64-разрядное число, записываемое в виде XX:XX:XX:XX:XX:XX:XX:XX. Институт IEEE назначает каждому производителю определенный диапазон адресов. Производитель отвечает за уникальное выделение назначенных адресов.

4.7.1 Адаптеры шины

Адаптер шины (host bus adapter — HBA) подключается к компьютеру и обеспечивает взаимодействие с устройствами хранения данных. В мире персональных компьютеров под управлением Windows адаптеры шины обычно подключаются к шине PCI и могут предоставлять подключение для устройств IDE, SCSI и Fibre Channel. Адаптеры шины работают под управлением драйвера устройства, т.е. драйвера мини-порта SCSI Port или Storport.

При инициализации порт адаптера шины регистрируется на коммутаторе связанной архитектуры (если таковой доступен) и регистрирует хранящиеся на нем атрибуты. Атрибуты доступны приложениям с помощью API от производителя коммутатора или адаптера шины. Ассоциация SNIA (Storage Networking Industry Association) работает над стандартизированным API, поддерживающим различные API производителей.

Для сети хранения данных, к которой выдвигаются серьезные требования по отказоустойчивости, некоторые производители адаптеров шины предоставляют дополнительные возможности, например автоматическое переключение на другой адаптер шины при отказе в работе основного. Эти функции, а также дополнительная архитектура рассматриваются в главе 9.

В кольце с разделением доступа только два устройства могут одновременно осуществлять прием и передачу данных. Предположим, что одно из них — это адаптер шины, подключенный к узлу и получающий данные от устройства хранения. Но, если этот адаптер подключен к сети хранения данных на основе коммутируемой связанной архитектуры, он может одновременно отправлять несколько запросов на чтение нескольким устройствам хранения.

Ответы на эти запросы могут приходить в любом порядке. Обычно коммутатор связанной архитектуры предоставляет службу циклического обслуживания для портов, что еще более усложняет задачу адаптера шины; в этом случае порядок поступления пакетов окажется таким, что каждый следующий пакет будет приходить от другого источника.

Адаптеры шины позволяют решить эту проблему одним из двух способов. Первая стратегия, которая называется *сохранить и отсортировать*, подразумевает хранение данных в памяти узла с последующей сортировкой буферов за счет центрального процессора. Очевидно, что это неэффективный подход с точки зрения центрального процессора и общая нагрузка связана с переключением контекста каждые несколько десятков микросекунд. Другая стратегия — *на лету* — подразумевает использование дополнительной системной логики и микросхем на самом адаптере шины, что позволяет осуществлять переключение контекста без использования циклов центрального процессора. Обычно время между переключениями контекста при использовании такой стратегии составляет несколько секунд.

Как отмечается в разделе 4.6.3.5, понятие *резервирование буфера* определено, как часть стандарта Fibre Channel. Одно резервирование позволяет отправить один кадр Fibre Channel. Перед отправкой следующего кадра отправитель должен получить сигнал **Receiver Ready**. Для эффективного использования канала Fibre Channel необходимо одновременно передавать несколько кадров, что потребует несколько резервирований, следовательно, понадобится больший объем памяти для принятия кадров. Некоторые адаптеры шины имеют четыре буфера размером 1 Кбайт и два буфера по 2 Кбайт, хотя на некоторых высокоуровневых адаптерах устанавливается 128 и 256 Кбайт для резервирования буфера. Обратите внимание, что для этой памяти обычно требуется два порта; т.е. когда одна область памяти принимает данные от SAN Fibre Channel, остальные области памяти могут передавать данные шине PCI узла.

Кроме того, адаптеры шины играют роль в обеспечении отказоустойчивости и в архитектуре с аварийным восстановлением данных, в которой предоставляется несколько маршрутов ввода-вывода к одному устройству хранения данных. Эти технологии рассматриваются в главе 9.

4.7.1.1 Операционная система Windows и адаптеры шины

В Windows NT и Windows 2000 адаптеры Fibre Channel рассматриваются как устройства SCSI, а драйверы создаются в виде драйверов мини-портов SCSI. Как отмечается в главе 2, проблема состоит в том, что драйвер SCSIPort устарел и не поддерживает возможности, предоставляемые новыми устройствами SCSI, не говоря уже об устройствах Fibre Channel. Поэтому в Windows

Server 2003 была введена новая модель драйвера Storport, которая должна заменить модель SCSIPort, особенно для устройств SCSI-3 и Fibre Channel. Обратите внимание, что диски Fibre Channel используются Windows в качестве DAS-устройств, что обеспечивается уровнем абстракции, предоставляемым драйверами SCSIPort и Storport.

4.7.1.2 Двойные маршруты

Иногда необходима повышенная производительность и надежность, даже за счет увеличения стоимости готового решения. В таких случаях сервер подключается к двухпортовым дискам через несколько адаптеров шины и несколько независимых сетей хранения данных Fibre Channel. Основная идея — исключить единую точку отказа в работе сети. Кроме того, в те моменты, когда система работает нормально, несколько маршрутов могут использоваться для балансировки нагрузки и повышения производительности. Дополнительная информация, включая методы, с помощью которых поставщики оборудования и компания Microsoft создают многомаршрутные системы, приводится в главе 9.

4.7.2 Типы кабелей Fibre Channel

В основном используется два типа кабелей: оптические и медные. Ниже перечислены основные преимущества и недостатки кабелей.

- Медные кабели дешевле оптических.
- Оптические кабели поддерживают более высокие скорости передачи данных по сравнению с медными кабелями.
- Медный кабель может использоваться на меньшем расстоянии, до 30 метров. При этом оптический кабель может использоваться на расстоянии до 2 километров (многомодовый кабель) или до 10 километров (одномодовый кабель).
- Медный кабель более восприимчив к электромагнитным помехам и взаимному влиянию других кабелей.
- Оптические данные обычно должны быть преобразованы в электрические сигналы для передачи через коммутатор и обратно в оптическую форму для дальнейшей передачи.

Существует только один тип медного кабеля, в отличие от оптического, который представлен двумя видами: многомодовым и одномодовым.

На коротких дистанциях используется многомодовый кабель, который имеет сердцевину диаметром 50 или 62,5 микрона (микрон — микрометр, или одна миллионная часть метра.) Световая волна, которая используется

в многомодовом кабеле, имеет длину 780 нанометров, что не поддерживается в одномодовых кабелях. Для больших расстояний предназначен одномодовый кабель, диаметр сердцевины которого составляет 9 микрон. В одномодовом кабеле используется световой луч с длиной волны в 1300 нанометров. Несмотря на тематику этой главы (интерфейс Fibre Channel), стоит упомянуть, что такие кабели могут использоваться для построения сетей на основе других интерфейсов, например Gigabit Ethernet.

4.7.3 Разъемы

Поскольку интерфейсом Fibre Channel поддерживается несколько типов кабелей (и технологий передачи данных), устройства (например, адаптеры шины, устройства взаимодействия и хранения данных) выпускаются с разъемами, которые поддерживают подключение к передающей среде, что делается для снижения общих затрат. Существует несколько видов разъемов, предназначенных для различных передающих сред и интерфейсов³.

- *Конверторы интерфейса Gigabit* (Gigabit interface converters — GBIC) поддерживают последовательную и параллельную трансляцию передаваемых данных. Конверторы GBIC предоставляют возможность “горячего” подключения, т.е. включение/выключение GBIC не влияет на работу других портов. Конверторами используется 20-битовый параллельный интерфейс.
- *Модули линий Gigabit* (Gigabit link modules — GLM) предоставляют функции, аналогичные GBIC, но для своей установки требуют отключения устройства. С другой стороны, они несколько дешевле, чем GBIC.
- *Адаптеры интерфейса носителя* (Media Interface Adapters) используются для преобразования сигналов между медным и оптическим носителем и наоборот. Адаптеры интерфейса носителя обычно используются в адаптерах шины, но могут применяться и на коммутаторах и концентраторах.
- *Адаптеры малого формфактора* (Small Form Factor Adapters — SFF) позволяют размещать большее количество разъемов различных интерфейсов на плате определенного размера.

³В настоящий момент существует несколько различных физических стандартов, и тот факт, что используется лишь три базовых типа кабелей (медные, одно- и многомодовые), не означает наличия трех типов физических разъемов. Кроме того, эти типы применяются и в других интерфейсах, например Gigabit Ethernet.

4.7.4 Устройства взаимодействия

Устройства взаимодействия соединяют между собой компоненты сетей хранения данных. К ним относятся различные устройства, начиная от дешевых концентраторов Fibre Channel и заканчивая дорогими, высокопроизводительными и управляемыми коммутаторами связанной архитектуры. Эти устройства рассматриваются в разделах 4.7.4.1–4.7.4.3.

4.7.4.1 Концентраторы кольца Fibre Channel с разделением доступа

Концентраторы FC-AL представляют собой бюджетный вариант для подключения нескольких узлов Fibre Channel (устройств хранения данных, серверов, компьютерных систем, других концентраторов и коммутаторов) в кольцевую конфигурацию. Обычно в концентраторах предоставляется от 8 до 16 портов. Концентратор может поддерживать различные среды передачи, например медные или оптические.

Концентраторы Fibre Channel — это пассивные устройства, т.е. любое другое устройство в кольце не может обнаружить их присутствия. Концентраторы обеспечивают следующие возможности:

- внутренние соединения, которые позволяют любому порту подключаться к любому другому порту;
- возможность обхода порта, к которому подключено неправильно работающее устройство.

Самая большая проблема в работе портов связана с тем, что в текущий момент времени они могут поддерживать только одно подключение Fibre Channel. На рис. 4.7 показано, что, если порт 1 получил управление для установки сеанса с портом 8, ни один другой порт не сможет передавать данные, пока установленный сеанс не завершится.

Концентраторы могут быть подключены к коммутаторам связанной архитектуры Fibre Channel (они рассматриваются в разделе 4.7.4.3) без модификации. Кроме того, можно создавать каскад концентраторов, соединив два концентратора кабелем.

Концентраторы FC-AL занимают лидирующее положение на рынке Fibre Channel, но в процессе снижения стоимости коммутаторы связанной архитектуры Fibre Channel становятся все более популярными.

Концентраторы FC-AL создаются такими компаниями, как Gadzoox Networks, Emulex и Brocade.

4.7.4.2 Коммутаторы кольца Fibre Channel с разделением доступа

Самое значительное преимущество коммутаторов FC-AL перед концентраторами состоит в одновременной поддержке нескольких подключений, то-

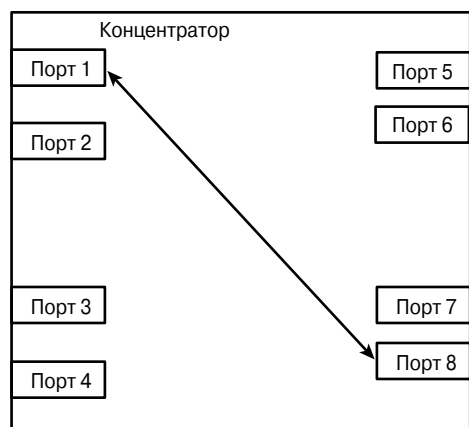


Рис. 4.7. Концентратор Fibre Channel

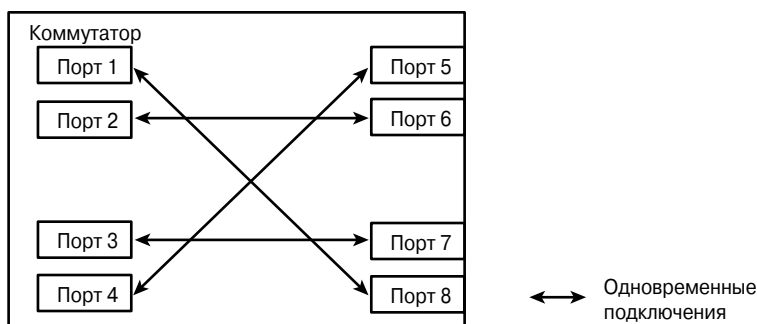


Рис. 4.8. Коммутатор Fibre Channel

гда как концентраторы поддерживают только одно подключение в текущий момент времени (рис. 4.8).

Возможность одновременной поддержки нескольких подключений связана с определенными сложностями. Устройства, подключенные к коммутатору кольца, даже не “подозревают” о своей роли. Коммутаторы кольца участвуют как в передаче данных, так и в адресации кольца. Ниже приводится дополнительная информация по этому вопросу, а также рассматривается роль коммутаторов в сетях хранения данных и методы, с помощью которых поставщики добавляют новые функции к своим продуктам.

Коммутаторы кольца и передача данных

Сервер, который намерен получить доступ к устройству хранения данных, должен отправить арбитражный запрос на управление кольцом. В нормальном кольце FC-AL на базе концентратора каждое устройство получает

арбитражный пакет до его возвращения адаптеру шины сервера, благодаря чему сервер получает контроль над кольцом. Коммутатор кольца отправит ответ об успешном получении управления немедленно, не отправляя запросы другим узлам. На этом этапе адаптер шины отправит базовый пакет *Open*, предназначенный для порта устройства хранения, который будет передан коммутатором кольца. Если порт в это время не выполняет передачи данных, особых проблем не должно возникнуть. В противном случае возможно появление конфликтных ситуаций. Для решения этой проблемы коммутатор кольца должен предоставить буфера для временного хранения кадров, предназначенных для порта 7. Некоторые поставщики коммутаторов предоставляют для этих целей 32 буфера на каждый порт.

Коммутаторы кольца и адресация FC-AL

Концентраторы FC-AL не играют роли в назначении адресов устройствам, а только передают базовые кадры адресов по кольцу. То же можно сказать и о большинстве коммутаторов. Однако некоторые устройства могут настаивать на получении определенного адреса. Некоторые концентраторы имеют возможность управлять порядком инициализации портов, что позволяет определенному порту инициализироваться первому, после чего устройство будет подключено к требующемуся порту.

Коммутаторы и инициализация кольца

Протокол FC-AL требует повторной инициализации кольца при подключении, отключении или повторной инициализации устройства. Такая инициализация кольца может привести к нарушению существующей связи между другими двумя устройствами. Некоторые производители коммутаторов предоставляют возможность выборочно экранировать и передавать пакеты *LIP* (Loop Initialization Primitives). Эта операция предназначена для минимизации проблем, сокращения времени повторной инициализации кольца и по возможности сохранения существующих сеансов передачи данных. В то же время необходимо обеспечить уникальность адресов устройств.

Если все устройства принимают участие в повторной инициализации кольца, дублирования адресов не происходит, так как устройства “защищают” свои адреса. Но, если некоторые устройства не принимают участия в повторной инициализации кольца, необходимо предотвратить назначение уже распределенных адресов устройствам, принимающим участие в повторной инициализации кольца. Уникальность адресов обеспечивается дополнительной логикой коммутатора кольца. При добавлении устройства хранения пакет *LIP* должен быть отправлен на сервер, однако *LIP* не требуется передавать

устройствам хранения, которые никогда не устанавливают связь с другими устройствами хранения данных.

Некоторые устройства хранения могут устанавливать связь непосредственно с другими устройствами хранения, что используется для резервного копирования данных. Дополнительная информация об операциях копирования приводится в главе 5.

Коммутаторы кольца и связанная архитектура

Если все устройства в кольце “знают” о связанной архитектуре, коммутатор кольца передает обычным образом необходимые кадры, например кадры Fabric Login. Если устройства в кольце не поддерживают связанную архитектуру, коммутатор кольца должен самостоятельно выполнять достаточно большой объем работы.

Коммутаторы кольца некоторых поставщиков не поддерживают каскадирование. Кроме того, некоторым коммутаторам кольца требуется обновление прошивки перед подключением к коммутаторам связанной архитектуры. Ряд коммутаторов следует модернизировать для полной поддержки связанной архитектуры перед их подключением к SAN.

Коммутаторы FC-AL производятся такими компаниями, как Brocade, McDATA, Gadzoox Networks, Vixel и QLogic.

4.7.4.3 Коммутаторы связанной архитектуры Fibre Channel

Коммутаторы связанной архитектуры Fibre Channel (Fibre Channel Fabric Switches — FC-SW) обеспечивают несколько высокоскоростных сеансов связи одновременно со всеми устройствами. На данный момент основные коммутаторы поддерживают быстродействие порядка 1 Гбит/с, в то время как скорость в 2 Гбит/с также перестает быть диковинкой. В основном коммутаторы связанной архитектуры в пересчете на один порт стоят дороже, чем концентраторы и коммутаторы FC-AL, но они предоставляют намного больше функциональных возможностей.

Коммутаторы связанной архитектуры более эффективны в сравнении с концентраторами и коммутаторами FC-AL. Например, коммутаторы предоставляют специальные службы, описанные выше, обеспечивают управление потоком с помощью базовых пакетов управления, а также, что гораздо важнее, некоторые коммутаторы способны эмулировать функции FC-AL для обеспечения обратной совместимости с более старыми устройствами.

Некоторые коммутаторы связанной архитектуры поддерживают *маршрутизацию без буферизации*. Суть ее в том, что при получении заголовка кадра коммутатор быстро находит заголовок точки назначения, пока кадр все еще

принимается. Преимущество такого подхода — снижение задержек при доставке кадра и отсутствие необходимости хранения содержимого кадра в памяти буфера. А недостаток заключается в немедленной передаче всех кадров, включая поврежденные.

Коммутаторы связанной архитектуры играют важную роль в безопасности сетей хранения данных на основе Fibre Channel, что описывается более подробно в главе 7.

4.7.4.4 Сравнение трех устройств подключения

В табл. 4.5 приведены функциональные возможности и различия между тремя типами устройств Fibre Channel.

4.7.4.5 Мосты и маршрутизаторы

Как в этой главе, так и во всей книге термины *мосты* (bridges) и *маршрутизаторы* (routers) не относятся к традиционным мостам Ethernet и маршрутизаторам IP. В данном случае под мостами и маршрутизаторами подразумеваются устройства для Fibre Channel, а не для сетевых протоколов 2-го и 3-го уровней.

Мосты — это устройства, обеспечивающие взаимодействие между Fibre Channel и устаревшими протоколами, например SCSI. Мосты Fibre Channel-SCSI позволяют сохранить существующие инвестиции в устройства хранения SCSI. Такие мосты поддерживают интерфейсы SCSI и Fibre Channel и преобразуют данные двух протоколов. Таким образом, новый сервер с установленным адаптером шины Fibre Channel может получить доступ к существующим устройствам хранения SCSI. Мосты предоставляют интерфейс между параллельной шиной SCSI и интерфейсом Fibre Channel. Маршрутизаторы обладают аналогичными возможностями, но для нескольких шин SCSI и интерфейсов Fibre Channel. Маршрутизаторы систем хранения данных, или “интеллектуальные” мосты, предоставляют такие дополнительные возможности, как маскировка и отображение LUN, а также поддерживают команды SCSI Extended Copy. В качестве устройств, передающих данные, маршрутизаторы применяют команды Extended Copy для использования библиотеками хранения, что позволяет копировать данные между указанным целевым устройством и подключенной библиотекой. Эта функция также называется *независимым резервным копированием* (без сервера).

В качестве примера производителей маршрутизаторов и мостов можно привести такие компании, как Crossroads Systems, Chaparral Network Storage, Advanced Digital Information Corporation (ADIC после приобретения Pathlight) и MTI.

Таблица 4.5. Устройства Fibre Channel

	Концентратор	Коммутатор FC-AL	Коммутатор FC-SW
Функциональность	Единая передача на скорости 100 Мбайт/с после переговоров на право передачи	Несколько передач на скорости 100 Мбайт/с после переговоров на право передачи	Несколько передач данных на скорости 1 или 2 Гбит/с без необходимости в переговорах. Возможна эмуляция FC-AL
Производительность	Снижается с увеличением количества узлов	Сохраняется при увеличении количества узлов	Сохраняется при увеличении количества узлов
Доступность данных	Все узлам доступны все данные, вне зависимости от назначения последних	Данные доступны только получающему и передающему узлу	Данные доступны только получающему и передающему узлу
Восстановление после ошибок	Сложный процесс восстановления после ошибок требует повторной инициализации кольца, что влияет на все узлы, в том числе работающие в обычном режиме	Восстановление после ошибок затрагивает только неправильно работающие узлы	Восстановление после ошибок затрагивает только неправильно работающие узлы
Повторное конфигурирование	При добавлении или удалении узла все узлы принимают участие в повторной инициализации кольца	Только новые узлы и коммутаторы принимают участие в повторной инициализации	Только новые узлы и коммутаторы принимают участие в повторной инициализации
Буфер данных	Концентратор не имеет буфера данных	Коммутатор не имеет буфера данных	Буфера данных у порта и коммутатора. Позволяют выполнять передачу данных без проверки доступности получающего узла

Окончание табл. 4.5

Концентратор		Коммутатор FC-AL	Коммутатор FC-SW
Адресация	8-разрядная адресация, 127 узлов. Один узел зарезервирован для подключения к коммутатору	16-разрядная адресация, до 16 млн. узлов. Поддерживается вложенная адресация — использование меньшего количества бит, например подобно тому, как последние пять цифр номера используются для внутрикорпоративных звонков	16-разрядная адресация, до 16 млн. узлов
Сложность и стоимость реализации	Низкая	Средняя	Высокая
Управляемость	Обычно не поддерживает управление, но иногда эта функция доступна в виде дополнительного модуля за дополнительную плату	Хорошая управляемость	Отличная управляемость
Дополнительные возможности	Нет	Зонирование	Зонирование и расширенные возможности безопасности. Многомаршрутный режим защиты целостности в сетях с соответствующей топологией. Транкинг линий связи, поддерживаемый некоторыми коммутаторами

4.8 Методы управления Fibre Channel

В предыдущих разделах рассматривались аппаратные элементы, формирующие сети хранения данных. В работе SAN также участвует немало различных программ, в основном предназначенных для управления, обеспечения безопасности, резервного копирования и восстановления данных. В разделах 4.8.1 и 4.8.2 рассматривается ряд концепций, необходимых для управления SAN и обеспечения безопасности данных. По сути, эти концепции представляют собой “сердце” SAN.

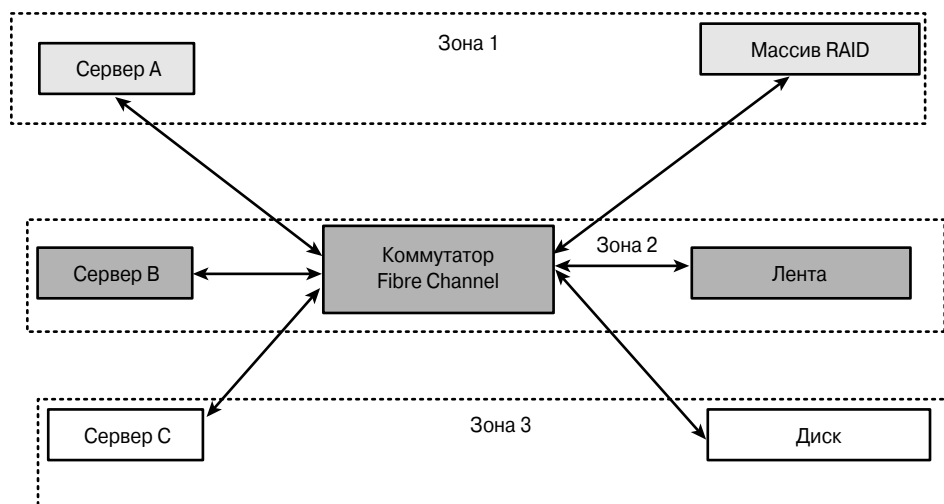
В ситуации, когда одна сеть содержит несколько компьютеров и единиц хранения данных, желательно ограничить влияние некоторых компьютеров (в терминологии Fibre Channel они называются *узлами*) до определенных подсистем хранения и некоторых единиц в рамках этих подсистем. Это имеет особый смысл в том случае, когда узел работает под управлением Windows NT, которая требует монтирования каждого обнаруженного устройства. С другой стороны, у UNIX есть таблица монтирования, благодаря чему монтируются только устройства, непосредственно указанные в таблице. Даже при использовании узлов под управлением UNIX желательно ограничить доступ из соображений обеспечения безопасности и для снижения вероятности повреждения данных. Доступ может быть ограничен тремя различными типами функций отображения и зонирования.

1. Базовая функция, реализованная в рамках узла; возможно, средствами программного драйвера адаптера шины.
2. Функция коммутатора.
3. Функция на уровне подсистемы хранения данных.

4.8.1 Зонирование

Термин *зонирование* связан с коммутаторами. Зонирование позволяет одним портам коммутатора подключаться только к заранее определенным портам. В некоторых случаях зонирование может ограничивать распространение управляющих кадров Fibre Channel; например, при появлении в кольце нового устройства хранения можно ограничить распространение кадра LIP среди других устройств.

С функциональной точки зрения зонирование дает возможность некоторому компьютеру непосредственно подключаться к определенной подсистеме хранения данных. Недостаток такого подхода состоит в предоставлении всех ресурсов SAN для одного компьютера, который обычно не в состоянии полностью их использовать. В частности, зонирование не позволяет совместно загружать канал доступа сети или применять ресурсы подсистемы хранения данных.

**Рис. 4.9.** Зонирование SAN

Зонирование можно воспринимать в качестве аналога конфигурирования порта IP для маршрутизатора с поддержкой брандмауэра. Еще одним примером может быть настройка виртуальных локальных сетей (VLAN) в существующей физической локальной сети. В виртуальной локальной сети только некоторые устройства “видят” друг друга, даже если в той же физической локальной сети находятся и другие устройства. Точно так же зонирование ограничивает возможности компонентов SAN (особенно инициаторов), предоставляя ограниченные данные об определенных единицах хранения и возможность доступа к ним, даже если в этой же физической сети хранения данных размещены и другие устройства хранения.

На рис. 4.9 демонстрируется концепция зонирования. Сеть хранения данных имеет три сервера и три единицы хранения. Различными оттенками указываются разные зоны.

Имена LUN могут совместно использоваться программным обеспечением файловой системы SAN. В этом программном обеспечении один или несколько серверов работают как серверы метаданных. Программное обеспечение устанавливается на клиентском компьютере (на компьютере, который желает получить доступ к файлам в сети хранения данных) и на сервере метаданных. Метаданные предоставляют клиентскому компьютеру информацию для отображения логического смещения в файле на физический номер блока указанного устройства. Это позволяет клиентскому компьютеру непосредственно получать доступ к файлу через SAN, без переноса данных через сервер. При достаточно грамотной организации обычные разрешения для файлов на

клиентском компьютере будут относиться и к файлам, хранящимся удаленно, что не требует от администратора дополнительных действий по настройке разрешений на совместный доступ к файлам.

Можно определить несколько зон, причем один узел имеет возможность входить в несколько зон одновременно; таким образом, некоторые зоны будут перекрываться. Зонирование выполняется несколькими способами.

- *Зонирование по номеру порта.* Преимущество такого подхода — эффективность. Если устройство, подключенное к порту, заменено другим устройством, повторная настройка не потребуется.
- *Зонирование по имени WWN.* Осуществляется путем указания имен WWN, которые входят в одну зону. Некоторые WWN могут быть указаны в нескольких зонах. Преимущество состоит в безопасности, которая, однако, достигается за счет эффективности. Изменения в конфигурации могут потребовать перезагрузки сервера.
- *Программное зонирование.* Проводится средствами сервера имен (программного обеспечения), который выполняется на коммутаторе. Для программного зонирования могут использоваться номера портов, WWN или комбинация этих параметров. Сервер имен содержит базу данных, в которой хранятся WWN, номера портов и идентификаторы зон.
- *Аппаратное зонирование.* Осуществляется с помощью таблицы маршрутизации, которая хранится на коммутаторе. Аппаратное зонирование выполняется на основе WWN и не принимает во внимание номера портов.

4.8.2 Маскировка LUN

Ресурсы хранения могут быть “разделены” на несколько вложенных единиц (субъединиц), которые называются *номером логического устройства* (logical unit number — LUN). Стандарт SCSI-2 поддерживает до 64 LUN на одно устройство.

С функциональной точки зрения маскировка LUN позволяет определенному компьютеру получить доступ к конкретной субъективной единице на некой системе хранения данных. Однако гораздо важнее то, что с помощью этого способа можно запретить доступ к определенным LUN для некоторых компьютеров или серверов. Маскировка LUN дает возможность совместно использовать ресурсы хранилищ данных и (неявно) пропускную способность сети, однако непосредственно LUN совместно использоваться не может. Для совместного использования одного LUN несколькими компьютерами необходима файловая система с дополнительными возможностями, которая описана в главе 6.

Маскировка LUN необходима для гарантирования целостности данных в среде SAN. Обратите внимание: маскировка LUN — это средство обеспечения безопасности на уровне дисков, но не обязательно на уровне файлов. В последнем случае (на уровне файлов) потребуется дополнительное программное обеспечение.

Маскировка LUN предоставляет дополнительные возможности, в частности номера LUN могут быть переназначены другим компьютерам. Существует несколько способов обеспечения маскировки LUN. Каждый способ обладает своими достоинствами и недостатками. Обычно, маскировка выполняется средствами:

- аппаратного обеспечения адаптера шины;
- аппаратного обеспечения коммутатора Fibre Channel;
- аппаратного обеспечения устройства хранения Fibre Channel;
- программного обеспечения узла.

Эти варианты рассматриваются в разделах 4.8.2.1–4.8.2.4.

4.8.2.1 Маскировка LUN средствами BIOS адаптера шины

В BIOS адаптера шины осуществляется маскировка всех LUN, которые не отображены в таблице BIOS адаптера шины. Таким образом, узел (с установленным адаптером шины) попросту не “замечает” существования LUN, которые он и не должен “видеть”.

Недостаток такого метода состоит в необходимости проведения корректной настройки; кроме того, метод не обязателен к применению. Все системы, адаптеры шины которых настроены неправильно или не поддерживают описываемую функцию, могут получить доступ к тем LUN, к которым доступ на самом деле нежелателен. Еще одна проблема заключается в сложности динамического управления и перенастройки подобных систем.

4.8.2.2 Маскировка LUN коммутаторами Fibre Channel

Коммутаторами Fibre Channel зонирование проводится достаточно просто. Входящий пакет передается или не передается дальше, что зависит от адресов исходного порта и порта назначения. Маскировка LUN возлагает дополнительную нагрузку на коммутаторы Fibre Channel, поскольку коммутатору приходится проверять первые 64 байта каждого пакета данных. Это приводит к снижению производительности большинства коммутаторов Fibre Channel, поэтому описываемая функция обычно не реализуется.

4.8.2.3 Маскировка LUN контроллерами подсистем хранения данных Fibre Channel и маршрутизаторами

Этот метод маскировки LUN является принудительным для подключенных узлов или требует от узла минимального участия. Маскировка LUN реализуется контроллером подсистемы хранения данных или маршрутизатором (с помощью соответствующей прошивки). Эти устройства настроены на поддержку таблицы имен WWN адаптера шины, отображенных на номера LUN, к которым им (контроллеру или маршрутизатору) разрешен доступ. Значительное преимущество такого подхода заключается в формировании конфигурации, независимой от промежуточных коммутаторов или концентраторов.

Недостаток метода заключается в закрытой реализации этой технологии каждым поставщиком и сложности создания единой консоли управления для перенастройки или даже получения информации о текущих параметрах, хотя каждый поставщик предоставляет интерфейсы для управления связками WWN–LUN.

К поставщикам систем, поддерживающим эту технологию, относятся Crossroads Systems, EMC, Dot Hill и HP (в продуктах StorageWorks). Поставщики присваивают реализации технологии собственные названия; например, компания Crossroads называет это *Access Controls*, а компания HP в продуктах StorageWorks выбрала название *Selective Storage Presentation*.

4.8.2.4 Маскировка LUN программным обеспечением узла

Маскировка LUN выполняется программным обеспечением узла, в частности кодом драйвера устройства. Код должен работать в режиме ядра, так как основная идея заключается в том, чтобы предотвратить доступ операционной системы к LUN, а операционная система сделает это еще до запуска первого приложения пользовательского режима.

Такая маскировка может выполняться в виде функции операционной системы или вне системы. За неимением конкретного решения от Microsoft некоторые поставщики добавили необходимый код в драйвер адаптеров шины. Обычно драйвер выдает команду **Report LUNs** каждому устройству, подключенному к шине, и перед предоставлением списка LUN системе Windows NT драйвер “вырезает” LUN из списка на основе дополнительно запрошенных данных (например, информации системного реестра Windows NT), таким образом “скрывая” некоторые LUN от Windows.

Основная проблема такого метода — необязательная настройка, а следовательно, необходимость частичного участия узла в процессе маскировки LUN. Это означает, что компьютеры, не имеющие модифицированного драйвера адаптера шины, не принимают участия в маскировке LUN. Кроме того, присутствуют и проблемы масштабирования, так как в особенно больших сетях

хранения данных сложно настроить каждый сервер и каждый адаптер шины сервера. Что касается преимуществ, то LUN может эффективно использоваться несколькими серверами.

Описываемая функция реализуется в продуктах компаний Emulex, Dell и JNI.

4.8.2.5 Маскировка LUN и будущее Windows NT

На данный момент существует информация, что Microsoft работает над реализацией возможностей маскировки LUN в драйвере порта. Тем не менее такая возможность отсутствует в Windows Server 2003. Преимущество использования драйвера порта состоит в постоянном присутствии драйвера порта в памяти, поэтому время, в течение которого компьютер не будет принимать участие в маскировке LUN, существенно снижается. Вероятность загрузки неправильного драйвера порта намного ниже, чем вероятность загрузки неправильного драйвера порта и мини-порта. Судя по предварительным прогнозам, если описываемая функция будет реализована в Windows, администратор получит возможность самостоятельно определять и изменять список LUN, видимых для сервера; при этом список может быть изменен временно. В последнем случае изменения не будут сохраняться после перезагрузки сервера.

4.9 Обеспечение взаимодействия устройств Fibre Channel

Призыв “Покупатель, берегись!” хорошо описывает состояние взаимодействия устройств в мире Fibre Channel.

Можно сказать, что большинство проблем во взаимодействии конфигураций FC-AL связаны с устройствами хранения, адаптерами шины, коммутаторами FC-AL и поставщиками маршрутизаторов. Поставщики устройств проводят серьезное тестирование своих продуктов, но, хотя теоретически взаимодействие с другими устройствами должно быть гарантировано, на практике для получения результата требуется немало дополнительного тестирования и настроек различных параметров. Рекомендуется использовать конфигурации, которые были протестированы поставщиком или продавцом готовых решений SAN.

Наибольшая проблема состоит в отсутствии гарантированного соответствия промышленным стандартам. Более того, даже соответствие стандартам также не обеспечивает 100%-ного взаимодействия.

Поставщики готовых решений, такие, как IBM, HP и EMC, создают лаборатории для тестирования взаимодействия различных устройств и проводят

собственную сертификацию. До определенной степени другие поставщики поступают аналогично. Рекомендуется использовать именно такие сертифицированные решения, что позволяет избежать проблем, часто возникающих при добавлении новых, не сертифицированных поставщиком устройств.

Хотя немало сетей хранения данных на основе Fibre Channel обеспечивают быстроедействие 1 Гбит/с, в последнее время в продаже появились устройства, поддерживающие скорость 2 Гбит/с. Новые устройства — новые проблемы. В стандартах, которым следуют производители, поддерживается скорость 2 Гбит/с, однако устройства автоматически переходят на скорость 1 Гбит/с, если на этой скорости работают другие устройства в сети. Дело в том, что сети хранения данных на базе Fibre Channel должны работать на скорости самого медленного устройства в сети. Таким образом, даже единственное устройство, работающее на скорости 1 Гбит/с, заставит всю сеть хранения данных работать на этом уровне быстрогодействия.

4.10 Сложности практической реализации

Сети хранения данных на основе Fibre Channel эмулируют прямое подключение устройства хранения данных к серверу, даже если устройство на самом деле подключено через коммутатор. Таким образом, в контексте Windows доступ к устройствам Fibre Channel осуществляется с помощью драйверов SCSI Port или Storport, описанных в главе 2. Таким образом, особенности работы с хранилищем, подключенным непосредственно к серверу (DAS), имеют отношение и к SAN.

Новая модель драйверов Storport предоставляет массу функциональных возможностей, включая оптимизацию ввода-вывода и управление пропускной способностью сети, однако системные администраторы и ответственные лица в информационных отделах компаний должны обратить внимание на тот факт, что модель драйверов Storport поддерживается исключительно в Windows Server 2003. Принявшим решение об использовании платформы Windows стоит изучить планы поставщика устройств хранения данных относительно перехода на модель Storport. В то же время необходимо обратить внимание на реализацию поддержки этих устройств на базе платформы Windows 2000, включая подробности реализации драйвера устройства. Это особенно важно для определения адекватности пропускной способности устаревшей модели драйверов SCSI Port, если поставщик будет продолжать ее применение. Кроме того, необходимо узнать, предоставляет ли поставщик собственную архитектуру SAN, без модели драйверов SCSI Port, а также сертифицировано ли это решение и поддерживается ли оно всеми заинтересо-

ванными сторонами. Наконец, обратите внимание на планы поставщика по переходу на модель драйверов Storport для Windows Server 2003.

Маскировка LUN на данный момент не поддерживается в продаваемых версиях Windows, причем выпуск Windows Server 2003 не изменил ситуации. Прежде чем приобретать новое программное и аппаратное обеспечение, выясните, какую технологию использует поставщик для реализации маскировки LUN и насколько она подходит для работы в среде Windows.

4.11 Резюме

Сети хранения данных Fibre Channel составляют существенную часть корпоративных подсистем хранения данных. Технология Fibre Channel может внедряться в виде недорогих конфигураций на основе кольца или на базе набирающей популярность топологии коммутируемой связной архитектуры.

Операционная система Windows Server 2003 поддерживает устройства Fibre Channel с помощью драйвера Storport, предоставляемого поставщиком аппаратного обеспечения. Поставщик вместо этого может предоставить мини-драйвер порта SCSI, но в таком случае преимущества драйвера Storport (например, повышенная производительность и обработка ошибок) окажутся недоступными для пользователей. Операционная система Windows 2000 и предыдущие ее версии поддерживают устройства Fibre Channel посредством мини-драйвера SCSI Port, предоставляемого поставщиками аппаратного обеспечения.

Несмотря на то что Windows NT поддерживает технологию маскировки LUN и зонирования, базовая поддержка маскировки LUN в Windows NT отсутствует. Маскировка LUN в Windows NT может быть реализована в драйвере от поставщика аппаратного обеспечения.