**Локальные сети на основе коммутаторов**

Содержание

Введение. Тенденция вытеснения концентраторов и маршрутизаторов коммутаторами

Технологии коммутации кадров (frame switching) в локальных сетях

Ограничения традиционных технологий (Ethernet, Token Ring), основанных на разделяемых средах передачи данных

Локальные мосты - предшественники коммутаторов

Принципы коммутации сегментов и узлов локальных сетей, использующих традиционные технологии

Полнодуплексные (full-duplex) протоколы локальных сетей - ориентация исключительно на коммутацию кадров

ATM-коммутация

Особенности коммутаторов локальных сетей

Техническая реализация коммутаторов

Коммутаторы на основе коммутационной матрицы

Коммутаторы с общей шиной

Коммутаторы с разделяемой памятью

Комбинированные коммутаторы

Модульные и стековые коммутаторы

Характеристики производительности коммутаторов

Скорость фильтрации и скорость продвижения

Оценка необходимой общей производительности коммутатора

Размер адресной таблицы

Объем буфера

Дополнительные возможности коммутаторов

Трансляция протоколов канального уровня

Поддержка алгоритма Spanning Tree

Способы управления потоком кадров

Возможности коммутаторов по фильтрации трафика

Коммутация "на лету" или с буферизацией

Использование различных классов сервиса (class-of-service)

Поддержка виртуальных сетей

Управление коммутируемыми сетями

Типовые схемы применения коммутаторов в локальных сетях

Коммутатор или концентратор?

Коммутатор или маршрутизатор?

Стянутая в точку магистраль на коммутаторе

Распределенная магистраль на коммутаторах

Обзор моделей коммутаторов

Коммутаторы Catalyst компании Cisco Systems

Коммутатор EliteSwitch ES/1 компании SMC

Коммутаторы локальных сетей компании 3Com

Примеры АТМ-коммутаторов для локальных сетей

## Введение. Тенденция вытеснения концентраторов и маршрутизаторов коммутаторами

Транспортная система локальных сетей масштаба здания или кампуса уже достаточно давно стала включать разнообразные типы активного коммуникационного оборудования - повторители, концентраторы, коммутаторы и маршрутизаторы, соединенные в сложные иерархические структуры, вроде той, которая изображена на рисунке 1.1.

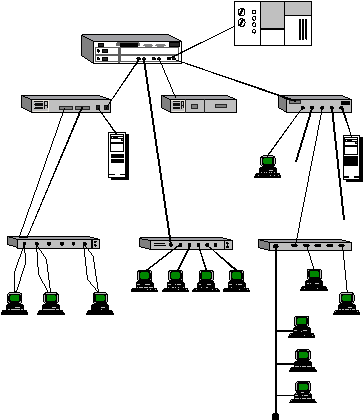


Рис. 1.1. Типичная структура сети здания или кампуса

Активное оборудование управляет циркулирующими в сети битами, кадрами и пакетами, стараясь организовать их передачу так, чтобы данные терялись как можно реже, а попадали к адресатам как можно быстрее, в соответствии с потребностями трафика работающих в сети приложений.

Описанный подход стал нормой при проектировании крупных сетей и полностью вытеснил сети, построенные исключительно на основе пассивных сегментов кабеля, которыми совместно пользуются для передачи информации компьютеры сети. Преимущества сетей с иерархически соединенным активным оборудованием не раз проверены на практике и сейчас никем не оспариваются.

И, если не обращать внимание на типы используемого оборудования, а рассматривать их просто как многопортовые черные ящики, то может сложиться впечатление, что никаких других изменений в теории и практике построения локальных сетей нет - предлагаются и реализуются очень похожие схемы, отличающиеся только количеством узлов и уровней иерархии коммуникационного оборудования.

Однако, качественный анализ используемого оборудования говорит об обратном. Изменения есть, и они существенны. За последние год-два коммутаторы стали заметно теснить другие виды активного оборудования с казалось бы прочно завоеванных позиций. Несколько лет назад в типичной сети здания нижний уровень иерархии всегда занимали повторители и концентраторы, верхний строился с использованием маршрутизаторов, а коммутаторам отводилось место где-то посередине, на уровне сети этажа. К тому же, коммутаторов обычно было немного - их ставили только в очень загруженные сегменты сети или же для подключения сверхпроизводительных серверов.

Коммутаторы стали вытеснять маршрутизаторы из центра сети на периферию (рисунок 1.2), где они использовались для соединения локальной сети с глобальными.

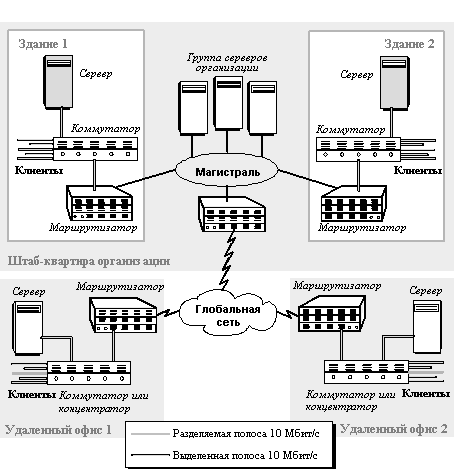


Рис. 1.2. Совместное использование коммутаторов и маршрутизаторов

Центральное место в сети здания занял модульный корпоративный коммутатор, который объединял на своей внутренней, как правило, очень производительной, магистрали все сети этажей и отделов. Коммутаторы потеснили маршрутизаторы потому, что их показатель "цена/производительность", рассчитанный для одного порта, оказался гораздо ниже при приближающихся к маршрутизаторам функциональным возможностям по активному воздействию на передаваемый трафик. Сегодняшние корпоративные коммутаторы умеют многое из того, что несколько лет назад казалось исключительной прерогативой маршрутизаторов: транслировать кадры разных технологий локальных сетей, например Ethernet в FDDI, осуществлять фильтрацию трафика по различным условиям, в том числе и задаваемым пользователем, изолировать трафик одного сегмента от другого и т.п. Коммутаторы ввели также и новую технологию, которая до их появления не применялась - технологию виртуальных сегментов, позволяющих перемещать пользователей из одного сегмента в другой чисто программным путем, без физической перекоммутации разъемов. И при всем при этом стоимость за один порт при равной производительности у коммутаторов оказывается в несколько раз ниже, чем у маршрутизаторов.

После завоевания магистрального уровня корпоративной сети коммутаторы начали наступление на сети рабочих групп, где до этого в течение последних пяти лет всегда использовались многопортовые повторители (концентраторы) для витой пары, заменившие пассивные коаксиальные сегменты. Появились коммутаторы, специально предназначенные для этой цели - простые, часто неуправляемые устройства, способные только быстро передавать кадры с порта на порт по адресу назначения, но не поддерживающие всей многофункциональности корпоративных коммутаторов. Стоимость таких коммутаторов в расчете на один порт быстро снижается и, хотя порт концентратора по-прежнему стоит меньше порта коммутатора рабочей группы, тенденция к сближению их цен налицо.

Подтверждением этой тенденции могут служить данные исследовательских компаний InStat и Dell'Oro Group за 1996 год и их прогноз на 1998 год:

|  |  |  |
| --- | --- | --- |
|  | **1996** | **1998Процент снижения за два года** |
| **Средняя цена за порт концентратора** | |  |
| Ethernet | $101 | $946.9% |
| Fast Ethernet | $200 | $14527.5% |
| **Средняя цена за порт коммутатора** | |  |
| Ethernet | $427 | $20053% |
| Fast Ethernet | $785 | $50036.3% |
| **Отношение порт коммутатора/порт концентратора** | |  |
| Ethernet | 4.22 | 2.1 |
| Fast Ethernet | 3.9 | 3.4 |

Эти данные собраны по всем классам коммутаторов, от уровня рабочей группы до магистрального уровня, где концентраторы не применяются, поэтому сопоставление концентраторов только с коммутаторами рабочих групп дало бы еще более близкие в стоимостном отношении результаты, так как стоимость за порт Ethernet у отдельных коммутаторов доходит до $150, то есть всего в полтора раза превышает стоимость порта концентратора Ethernet.

В то же время производительность сети, построенной на коммутаторе, обычно в несколько раз превышает производительность аналогичной сети, построенной с использованием концентратора. Так как плата за повышение производительности не так уж велика и постоянно снижается, то многие сетевые интеграторы все чаще соглашаются с ней для снижения задержек в своей сети. С распространением работающих в реальном времени приложений ущерб от транспортных задержек становится все ощутимее, а нагрузка на транспортную систему возрастает, что еще больше стимулирует приближение таких высокопроизводительных устройств, как коммутаторы, к пользовательским компьютерам.

Естественно, тенденция повышения роли коммутаторов в локальных сетях не имеет абсолютного характера. И у маршрутизаторов, и у концентраторов по-прежнему имеются свои области применения, где их применение более рационально, чем коммутаторов. Маршрутизаторы остаются незаменимыми при подключении локальной сети к глобальной. Кроме того, маршрутизаторы хорошо дополняют коммутаторы при построении виртуальных сетей из виртуальных сегментов, так как дают испытанный способ объединения сегментов в сеть на основании их сетевых адресов.

Концентраторы также имеют сегодня свою нишу. По-прежнему существует большое количество случаев, когда трафик в рабочей группе невелик и направлен к одному серверу. В таких случаях высокая производительность коммутатора мало что дает конечному пользователю - при замене концентратора на коммутатор он ее практически не почувствует.

Тем не менее, в локальных сетях появляется все больше коммутаторов, и эта ситуация вряд ли коренным образом изменится в ближайшем будущем. Некоторые новые технологии, такие как ATM, вообще используют коммутацию как единственный способ передачи данных в сети, другие, например, Gigabit Ethernet - рассматривают ее в качестве, хотя и не единственного, но основного способа связи устройств в сети.

## Технологии коммутации кадров (frame switching) в локальных сетях

### Ограничения традиционных технологий (Ethernet, Token Ring), основанных на разделяемых средах передачи данных

Повторители и концентраторы локальных сетей реализуют базовые технологии, разработанные для разделяемых сред передачи данных. Классическим представителем такой технологии является технология Ethernet на коаксиальном кабеле. В такой сети все компьютеры разделяют во времени единственный канал связи, образованный сегментом коаксиального кабеля (рисунок 2.1).

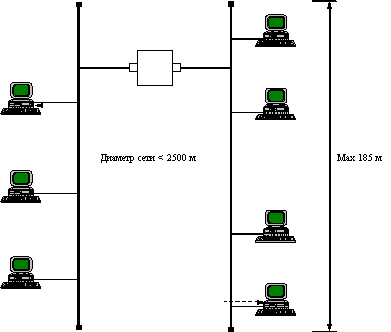


Рис. 2.1. Разделяемый канал передачи данных в сети Ethernet

При передаче каким-нибудь компьютером кадра данных все остальные компьютеры принимают его по общему коаксиальному кабелю, находясь с передатчиком в постоянном побитном синхронизме. На время передачи этого кадра никакие другие обмены информации в сети не разрешаются. Способ доступа к общему кабелю управляется несложным распределенным механизмом арбитража - каждый компьютер имеет право начать передачу кадра, если на кабеле отсутствуют информационные сигналы, а при одновременной передаче кадров несколькими компьютерами схемы приемников умеют распознавать и обрабатывать эту ситуацию, называемую *коллизией*. Обработка коллизии также несложна - все передающие узлы прекращают выставлять биты своих кадров на кабель и повторяют попытку передачи кадра через случайный промежуток времени.

Работа всех узлов сети Ethernet в режиме большой распределенной электронной схемы с общим тактовым генератором приводит к нескольким ограничениям, накладываемым на сеть. Основными ограничениями являются:

* Максимально допустимая длина сегмента. Она зависит от типа используемого кабеля: для витой пары это 100 м, для тонкого коаксиала - 185 м, для толстого коаксиала - 500 м, а для оптоволокна - 2000 м. Для наиболее дешевых и распространенных типов кабеля - витой пары и тонкого коаксиала - это ограничение часто становится весьма нежелательным. Технология Ethernet предлагает использовать для преодоления этого ограничения повторители и концентраторы, выполняющие функции усиления сигнала, улучшения формы фронтов импульсов и исправления погрешностей синхронизации. Однако возможности этих устройств по увеличению максимально допустимого расстояния между двумя любыми узлами сети (которое называется диаметром сети) не очень велики - число повторителей между узлами не может превышать 4-х (так называемое правило четырех хабов). Для витой пары это дает увеличение до 500 м (рисунок 2.2). Кроме того, существует общее ограничение на диаметр сети Ethernet - не более 2500 м для любых типов кабеля и любого количества установленных концентраторов. Это ограничение нужно соблюдать для четкого распознавания коллизий всеми узлами сети, как бы далеко (в заданных пределах) они друг от друга не находились, иначе кадр может быть передан с искажениями.

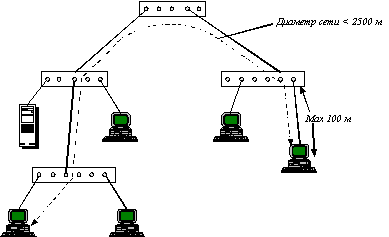


Рис. 2.2. Максимальный диаметр сети Ethernet на витой паре

* Максимальное число узлов в сети. Стандарты Ethernet ограничивают число узлов в сети предельным значением в 1024 компьютера вне зависимости от типа кабеля и количества сегментов, а каждая спецификация для конкретного типа кабельной системы устанавливает еще и свое, более жесткое ограничение. Так, к сегменту кабеля на тонком коаксиале нельзя подключить более 30 узлов, а для толстого коаксиала это число увеличивается до 100 узлов. В сетях Ethernet на витой паре и оптоволокне каждый отрезок кабеля соединяет всего два узла, но так как количество таких отрезков спецификация не оговаривает, то здесь действует общее ограничение в 1024 узла.

Существуют также и другие причины, кроме наличия указанных в стандартах ограничений, по которым число узлов в сети Ethernet обычно не превосходит нескольких десятков. Эти причины лежат в самом принципе разделения во времени одного канала передачи данных между всеми узлами сети. При подключении к такому каналу каждый узел пользуется его пропускной способностью - 10 Мб/с - в течение только некоторой доли общего времени работы сети. Соответственно, на узел приходится эта же доля пропускной способности канала. Даже если упрощенно считать, что все узлы получают равные доли времени работы канала и непроизводительные потери времени отсутствуют, то при наличии в сети N узлов на один узел приходится только 10/N Мб/с пропускной способности. Очевидно, что при больших значениях N пропускная способность, выделяемая каждому узлу, оказывается настолько малой величиной, что нормальная работа приложений и пользователей становится невозможной - задержки доступа к сетевым ресурсам превышают тайм-ауты приложений, а пользователи просто отказываются так долго ждать отклика сети.

Случайный характер алгоритма доступа к среде передачи данных, принятый в технологии Ethernet, усугубляет ситуацию. Если запросы на доступ к среде генерируются узлами в случайные моменты времени, то при большой их интенсивности вероятность возникновения коллизий также возрастает и приводит к неэффективному использованию канала: время обнаружения коллизии и время ее обработки составляют непроизводительные затраты. Доля времени, в течение которого канал предоставляется в распоряжение конкретному узлу, становится еще меньше.

На рисунке 2.3 показана зависимость задержек доступа к среде передачи данных в сети Ethernet от количества узлов сети. Экспоненциальный рост задержек при увеличении числа узлов очень характерен как для технологии Ethernet, так и для других технологий локальных сетей, основанных на разделении каналов во времени - Token Ring, FDDI, 100VG-AnyLAN.

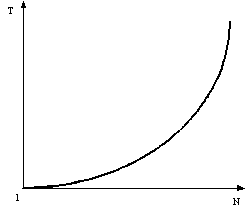


Рис. 2.3. Зависимость задержек доступа к среде передачи данных   
сети Ethernet от числа узлов сети

До недавнего времени в локальных сетях редко использовались мультимедийные приложения, перекачивающие большие файлы данных, нередко состоящие из нескольких десятков мегабайт. Приложения же, работающие с алфавитно-цифровой информацией, не создавали значительного трафика. Поэтому долгое время для сегментов Ethernet было действительным эмпирическое правило - в разделяемом сегменте не должно быть больше 30 узлов. Теперь ситуация изменилась и нередко 3-4 компьютера полностью загружают сегмент Ethernet с его максимальной пропускной способностью в 10 Мб/с или же 14880 кадров в секунду.

Более универсальным критерием загруженности сегмента Ethernet по сравнению с общим количеством узлов является суммарная нагрузка на сегмент, создаваемая его узлами. Если каждый узел генерирует в среднем mi кадров в секунду для передачи по сети, то средняя суммарная нагрузка на сеть будет составлять i mi кадров в секунду. Известно, что при отсутствии коллизий, то есть при самом благоприятном разбросе запросов на передачу кадров во времени, сегмент Ethernet может передать не больше 14880 кадров в секунду (для самых коротких по стандарту кадров в 64 байта). Поэтому, если принять эту величину за единицу, то отношение i mi/14880 будет характеризовать степень использования канала, называемый также коэффициентом загрузки.

Зависимость времени ожидания доступа к сети от коэффициента загрузки гораздо меньше зависит от интенсивности трафика каждого узла, поэтому эту величину удобно использовать для оценки пропускной способности сети, состоящей из произвольного числа узлов. Имитационное моделирование сети Ethernet и исследование ее работы с помощью анализаторов протоколов показали, что при коэффициенте загрузки в районе 0.3 - 0.5 начинается быстрый рост числа коллизий и соответственно времени ожидания доступа. Поэтому во многих системах управления сетями пороговая граница для индикатора коэффициента загрузки по умолчанию устанавливается на величину 0.3.

Ограничения, связанные с возникающими коллизиями и большим временем ожидания доступа при значительной загрузке разделяемого сегмента, чаще всего оказываются более серьезными, чем ограничение на максимальное количество узлов, определенное в стандарте из соображений устойчивой передачи электрических сигналов в кабелях.

Технология Ethernet была выбрана в качестве примера при демонстрации ограничений, присущих технологиям локальных сетей, так как в этой технологии ограничения проявляются наиболее ярко, а их причины достаточно очевидны. Однако подобные ограничения присущи и всем остальным технологиям локальных сетей, так как они опираются на использование среды передачи данных как одного разделяемого ресурса. Кольца Token Ring и FDDI также могут использоваться узлами сети только в режиме разделяемого ресурса. Отличие от канала Ethernet здесь состоит только в том, что маркерный метод доступа определяет детерминированную очередность предоставления доступа к кольцу, но по-прежнему при предоставлении доступа одного узла к кольцу все остальные узлы не могут передавать свои кадры и должны ждать, пока владеющий правом доступа узел не завершит свою передачу.

Как и в технологии Ethernet, в технологиях Token Ring, FDDI, Fast Ethernet и 100VG-AnyLAN также определены максимальные длины отдельных физических сегментов кабеля и ограничения на максимальный диаметр сети и максимальное количество в ней узлов. Эти ограничения несколько менее стеснительны, чем у технологии Ethernet, но также могут быть серьезным препятствием при создании крупной сети.

Особенно же быстро может проявиться ограничение, связанное с коэффициентом загрузки общей среды передачи данных. Хотя метод маркерного доступа, используемый в технологиях Token Ring и FDDI, или метод приоритетных требований технологии 100VG-AnyLAN позволяют работать с более загруженными средами, все равно отличия эти только количественные - резкий рост времени ожидания начинается в таких сетях при больших коэффициентах загрузки, где-то в районе 60% - 70%. Качественный характер нарастания времени ожидания доступа и в этих технологиях тот же, и он не может быть принципиально иным, когда общая среда передачи данных разделяется во времени между компьютерами сети.

Общее ограничение локальных сетей, построенных только с использованием повторителей и концентраторов, состоит в том, что общая производительность такой сети всегда фиксирована и равна максимальной производительности используемого протокола. И эту производительность можно повысить только перейдя к другой технологии, что связано с дорогостоящей заменой всего оборудования.

Рассмотренные ограничения являются платой за преимущества, которые дает использование разделяемых каналов в локальных сетях. Эти преимущества существенны, недаром технологии такого типа существуют уже около 20 лет.

К преимуществам нужно отнести в первую очередь:

* простоту топологии сети;
* гарантию доставки кадра адресату при соблюдении ограничений стандарта и корректно работающей аппаратуре;
* простоту протоколов, обеспечившую низкую стоимость сетевых адаптеров, повторителей и концентраторов;

Однако начавшийся процесс вытеснения повторителей и концентраторов коммутаторами говорит о том, что приоритеты изменились, и за повышение общей пропускной способности сети пользователи готовы пойти на издержки, связанные с приобретением коммутаторов вместо концентраторов.

### Локальные мосты - предшественники коммутаторов

Для преодоления ограничений технологий локальных сетей уже достаточно давно начали применять локальные мосты, функциональные предшественники коммутаторов.

Мост - это устройство, которое обеспечивает взаимосвязь двух (реже нескольких) локальных сетей посредством передачи кадров из одной сети в другую с помощью их промежуточной буферизации. Мост, в отличие от повторителя, не старается поддержать побитовый синхронизм в обеих объединяемых сетях. Вместо этого он выступает по отношению к каждой из сетей как конечный узел. Он принимает кадр, буферизует его, анализирует адрес назначения кадра и только в том случае, когда адресуемый узел действительно принадлежит другой сети, он передает его туда.

Для передачи кадра в другую сеть мост должен получить доступ к ее разделяемой среде передачи данных в соответствии с теми же правилами, что и обычный узел.

Таким образом мост, изолирует трафик одного сегмента от трафика другого сегмента, фильтруя кадры. Так как в каждый из сегментов теперь направляется трафик от меньшего числа узлов, то коэффициент загрузки сегментов уменьшается (рисунок 2.4).

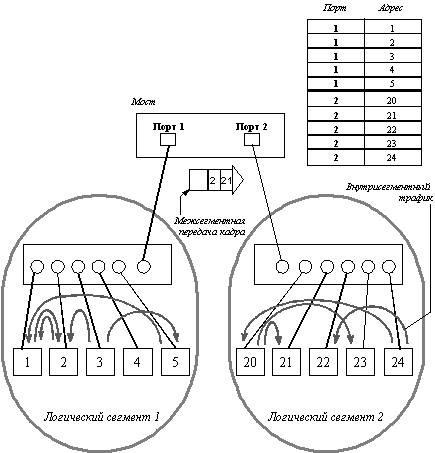


Рис. 2.4. Локализация трафика при использовании моста

Мост не только снижает нагрузку в объединенной сети, но и уменьшает возможности несанкционированного доступа, так как пакеты, предназначенные для циркуляции внутри одного сегмента, физически не появляются на других, что исключает их "прослуши-   
вание" станциями других сегментов.

По своему принципу действия мосты подразделяются на два типа. Мосты первого типа выполняют так называемую маршрутизацию от источника (Source Routing), метод, разработанный фирмой IBM для своих сетей Token Ring. Этот метод требует, чтобы узел-отправитель пакета размещал в нем информацию о маршруте пакета. Другими словами, каждая станция должна выполнять функции по маршрутизации пакетов. Второй тип мостов осуществляет прозрачную для конечных станций передачу пакетов (Transparent Bridges). Именно этот тип мостов лег в основу современных коммутаторов, поэтому остановимся на нем подробнее.

**Функции и алгоритмы прозрачных мостов**

*Прозрачные мосты*являются наиболее распространенным типом мостов. Для прозрачных мостов сеть представляется наборами МАС-адресов устройств, используемых на канальном уровне, причем каждый набор связан с определенным портом моста.

Мосты используют эти адреса для принятия решения о продвижении кадра, когда кадр записывается во внутренний буфер моста из какого-либо его порта. Мосты не имеют доступа к информации об адресах сетей, относящейся к более высокому - сетевому - уровню, и они ничего не знают о топологии связей сегментов или сетей между собой. Таким образом, мосты являются совершенно прозрачными для протоколов, начиная с сетевого уровня и выше. Эта прозрачность позволяет мостам передавать пакеты различных протоколов высокого уровня, никоим образом не влияя на их содержимое.

Вследствие функциональной ограниченности мосты имеют достаточно простое устройство и представляют собой удобное и недорогое средство для построения интерсети.

Мосты обеспечивают возможность соединения двух или более сетей для образования единой логической сети. Исходные сети становятся сетевыми сегментами результирующей сети. Каждый такой сегмент остается доменом коллизий, то есть участком сети, в котором все узлы одновременно фиксируют и обрабатывают коллизию. Однако коллизии одного сегмента не приводят к возникновению коллизий в другом сегменте, так как мост не осуществляет побитовый синхронизм сегментов и ограничивает коллизии тем сегментом, в котором они возникают.

Мосты регенерируют пакеты, которые они передают с одного порта на другой (операция forwarding). Одним из преимуществ использования мостов является увеличение расстояния, покрываемого интерсетью, так как количество пересекаемых мостов не оказывает влияния на качество сигнала.

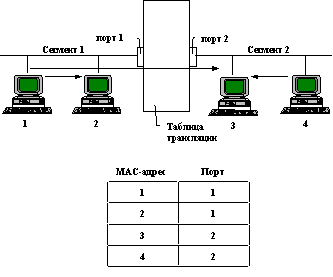


Рис. 2.5. Мост как коммуникационное устройство канального уровня

Прозрачные мосты имеют дело как с адресом источника, так и с адресом назначения, имеющимися в кадрах локальных сетей. Мост использует адрес источника для автоматического построения своей базы данных адресов устройств, называемой также таблицей адресов устройств. В этой таблице устанавливается принадлежность адреса узла какому-либо порту моста. Все операции, которые выполняет мост, связаны с этой базой данных. На рисунке 2.5 показан фрагмент сети, содержащий двухпортовый мост, и соответствующая этому фрагменту часть таблицы адресов устройств. Внутренняя структура моста показана на рисунке 2.6. Функции доступа к среде при приеме и передаче кадров выполняют микросхемы MAC.

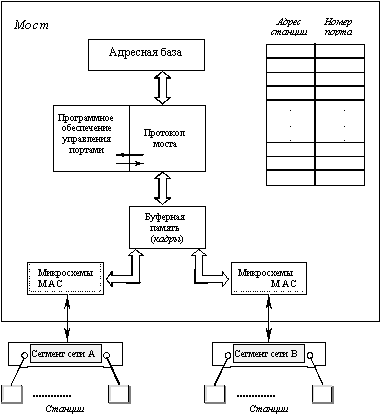


Рис. 2.6. Состав и структура моста

Все порты моста работают в так называемом "неразборчивом" (promisquous) режиме захвата пакетов, то есть все поступающие на порт пакеты запоминаются в буферной памяти. С помощью такого режима мост следит за всем трафиком, передаваемым в присоединенных к нему сегментах и использует проходящие через него пакеты для изучения состава сети.

Когда мост получает кадр от какого-либо своего порта, то он (после буферизации) сравнивает адрес источника с элементами базы данных адресов. Если адрес отсутствует в базе, то он добавляется в нее. Если этот адрес уже имеется в базе, то возможны два варианта - либо адрес пришел с того же порта, который указан в таблице, либо он пришел с другого порта. В последнем случае строка таблицы, соответствующая обрабатываемому адресу, обновляется - номер порта заменяется на новое значение (очевидно, станцию с данным адресом переместили в другой сегмент сети). Таким способом мост "изучает" адреса устройств сети и их принадлежность портам и соответствующим сегментам сети. Из-за способности моста к "обучению" к сети могут добавляться новые устройства без необходимости реконфигурирования моста. Администратор может объявить часть адресов статическими и не участвующими в процессе обучения (при этом он их должен задать сам). В случае статического адреса приход пакета с данным адресом и значением порта, не совпадающим с хранящимся в базе, будет проигнорирован и база не обновится.

Кроме адреса источника мост просматривает и адрес назначения кадра, чтобы принять решение о его дальнейшем продвижении. Мост сравнивает адрес назначения кадра с адресами, хранящимися в его базе. Если адрес назначения принадлежит тому же сегменту, что и адрес источника, то мост "фильтрует" (filtering) пакет, то есть удаляет его из своего буфера и никуда не передает. Эта операция помогает предохранить сеть от засорения ненужным трафиком.

Если адрес назначения присутствует в базе данных и принадлежит другому сегменту по сравнению с сегментом адреса источника, то мост определяет, какой из его портов связан с этим адресом и "продвигает" (forwarding) кадр на соответствующий порт. Затем порт должен получить доступ к среде подключенного к нему сегмента и передать кадр узлам данного сегмента.

Если же адрес назначения отсутствует в базе или же это широковещательный адрес, то мост передает кадр на все порты, за исключением того порта, с которого он пришел. Такой процесс называется "затоплением" (flooding) сети. Затопление гарантирует, что пакет будет помещен на все сегменты сети и, следовательно, доставлен адресату или адресатам. Точно также мост поступает по отношению к кадрам с неизвестным адресом назначения, затопляя им сегменты сети. Очевидно, что некоторое время после инициализации мост выполняет только операцию затопления, так как он ничего не знает о принадлежности адресов сегментам сети.

Рисунок 2.5 иллюстрирует процессы обучения, фильтрации и продвижения. Предположим, что станции 1 и 2 являются новыми станциями на сегменте 1. Когда станция 1 впервые направляет кадр станции 2, то мост определяет, что адреса станции 1 нет в базе адресов и добавляет его туда. Затем, так как адреса станции 2 также нет в базе адресов, мост "затапливает" все сегменты (в данном случае это только один сегмент 2).

Когда станция 2 посылает ответный кадр, мост добавляет в свою базу и адрес 2. Затем он просматривает таблицу базы адресов и обнаруживает, что адрес 1 в ней имеется и относится к сегменту 1, которому принадлежит и адрес источника. Поэтому он фильтрует этот кадр, то есть удаляет его из буфера и никуда не передает.

Мост, работающий по описанному алгоритму, прозрачен не только для протоколов всех уровней, выше канального, но и для конечных узлов сети. Эта прозрачность состоит в том, что узлы не посылают мосту свои кадры специальным образом, указывая в них адрес порта моста. Даже при наличии моста в сети конечные узлы продолжают посылать кадры данных непосредственно другим узлам, указывая их адреса в качестве адресов назначения кадров. Поэтому порты мостов вообще не имеют МАС-адресов, работая в режиме "неразборчивого" захвата всех кадров. Такая прозрачность моста упрощает работу конечных узлов, и это свойство коренным образом отличает мост от маршрутизатора, которому узел отправляет кадр явным образом, указывая МАС-адрес порта маршрутизатора в своем кадре.

На рисунке 2.7 показана копия экрана с адресной таблицей модуля моста концентратора System 3000 компании Bay Networks. Из него видно, что сеть состоит из двух сегментов - LAN A и LAN B. В сегменте LAN A имеется по крайней мере 3 станции, а в сегменте LAN B - 2. Четыре адреса, помеченные звездочками, являются статическими, причем кадры, имеющие адреса, помеченные Flood, должны распространяться широковещательно.

Описанная процедура хорошо работает до тех пор, пока пользователи не переносят свои компьютеры из одного логического сегмента в другой. Так как MAC-адрес сетевого адаптера аппаратно устанавливается изготовителем, то при перемещении компьютера мосты должны периодически обновлять содержимое своих адресных баз. Для обеспечения этой функции записи в адресной базе делятся на два типа - статические и динамические. С каждой динамической записью связан *таймер неактивности*.Когда мост принимает кадр с адресом источника, соответствующим некоторой записи в адресной базе, то соответствующий таймер неактивности сбрасывается в исходное состояние. Если же от какой-либо станции долгое время не поступает кадров, то таймер неактивности исчерпывает свой интервал, и соответствующая ему запись удаляется из адресной базы.

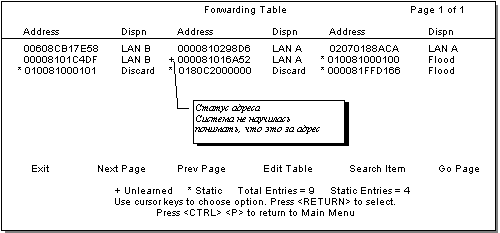


Рис. 2.7. Таблица продвижений моста System 3000 Local Bridge

**Проблема петель при использовании мостов**

Обучение, фильтрация и продвижение основаны на существовании одного логического пути между любыми двумя узлами сети. Наличие нескольких путей между устройствами, известных также как "активные петли", создает проблемы для сетей, построенных на основе мостов.

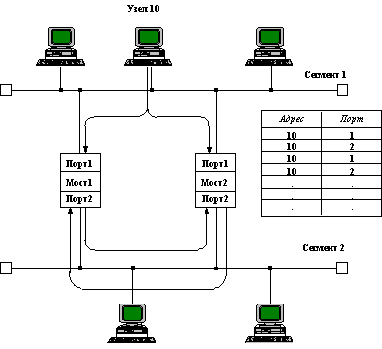


Рис. 2.8. Влияние замкнутых маршрутов на работу мостов

Рассмотрим в качестве примера сеть, приведенную на рисунке 2.8. Два сегмента параллельно соединены двумя мостами так, что образовалась активная петля. Пусть новая станция с адресом 10 впервые посылает пакет другой станции сети, адрес которой также пока неизвестен мосту. Пакет попадает как в мост 1, так и в мост 2, где его адрес заносится в базу адресов с пометкой о его принадлежности сегменту 1. Так как адрес назначения неизвестен мосту, то каждый мост передает пакет на сегмент 2. Эта передача происходит поочередно, в соответствии с методом случайного доступа технологии Ethernet. Пусть первым доступ к сегменту 2 получил мост 1. При появлении пакета на сегменте 2 мост 2 принимает его в свой буфер и обрабатывает. Он видит, что адрес 10 уже есть в его базе данных, но пришедший пакет является более свежим, и он утверждает, что адрес 10 принадлежит сегменту 2, а не 1. Поэтому мост 2 корректирует содержимое базы и делает запись о том, что адрес 10 принадлежит сегменту 2. Аналогично поступает мост 1, когда мост 2 передает свою буферизованную ранее первую версию пакета на сегмент 2. В результате пакет бесконечно циркулирует по активной петле, а мосты постоянно обновляют записи в базе, соответствующие адресу 10. Сеть засоряется ненужным трафиком, а мосты входят в состояние "вибрации", постоянно обновляя свои базы данных.

В простых сетях сравнительно легко гарантировать существование одного и только одного пути между двумя устройствами. Но когда количество соединений возрастает или интерсеть становится сложной, то вероятность непреднамеренного образования петли становится высокой. Кроме того, желательно для повышения надежности иметь между мостами резервные связи, которые не участвуют при нормальной работе основных связей в передаче информационных пакетов станций, но при отказе какой-либо основной связи образуют новую связную рабочую конфигурацию без петель. Описанные задачи решает алгоритм покрывающего дерева *(Spanning Tree Algorithm, STA).*

**Требования к пропускной способности моста**

До сих пор мы предполагали, что при использовании моста для связи двух сегментов вместо повторителя общая производительность сети всегда повышается, так как уменьшается количество узлов в каждом сегменте и загрузка сегмента уменьшается на ту долю трафика, который теперь является внутренним трафиком другого сегмента. Это действительно так, но при условии что мост передает межсегментный трафик без значительных задержек и без потерь кадров. Однако, анализ рассмотренного алгоритма работы моста говорит о том, что мост может и задерживать кадры и, при определенных условиях, терять их. Задержка, вносимая мостом, равна по крайней мере времени записи кадра в буфер. Как правило, после записи кадра на обработку адресов также уходит некоторое время, особенно если размер адресной таблицы велик. Поэтому задержка увеличивается на время обработки кадра.

Время обработки кадра влияет не только на задержку, но и на вероятность потери кадров. Если время обработки кадра окажется меньше интервала до поступления следующего кадра, то следующий кадр будет помещен в буфер и будет ожидать там, пока процессор моста не освободиться и не займется обработкой поступившего кадра. Если средняя интенсивность поступления кадров будет в течение длительного времени превышать производительность моста, то есть величину, обратную среднему времени обработки кадра, то буферная память, имеющаяся у моста для хранения необработанных кадров, может переполниться. В такой ситуации мосту некуда будет записывать поступающие кадры и он начнет их терять, то есть просто отбрасывать.

Потеря кадра - ситуация очень нежелательная, так как ее последствия не ликвидируются протоколами локальных сетей. Потеря кадра будет исправлена только протоколами транспортного или прикладного уровней, которые заметят потерю части своих данных и организуют их повторную пересылку. Однако, при регулярных потерях кадров канального уровня производительность сети может уменьшится в несколько раз, так как тайм-ауты, используемые в протоколах верхних уровней, существенно превышают времена передачи кадров на канальном уровне, и повторная передача кадра может состояться через десятки секунд.

Для предотвращения потерь кадров мост должен обладать производительностью, превышающую среднюю интенсивность межсегментного трафика и большой буфер для хранения кадров, передаваемых в периоды пиковой нагрузки.

В локальных сетях часто оказывается справедливым эмпирическое правило 80/20, говорящее о том, что при правильном разбиении сети на сегменты 80% трафика оказывается внутренним трафиком сегмента, и только 20% выходит за его пределы. Если считать, что это правило действует по отношению к конкретной сети, то мост должен обладать производительностью в 20 % от максимальной пропускной способности сегмента Ethernet, то есть производительностью 0.2 ( 14880 = 3000 кадра в секунду. Обычно локальные мосты обладают производительностью от 3000 кадров в секунду и выше.

Однако, гарантий на доставку кадров в любых ситуациях мост, в отличие от повторителя, не дает. Это его принципиальный недостаток, с которым приходится мириться.

### Принципы коммутации сегментов и узлов локальных сетей, использующих традиционные технологии

Технология коммутации сегментов Ethernet была предложена фирмой Kalpana в 1990 году в ответ на растущие потребности в повышении пропускной способности связей высокопроизводительных серверов с сегментами рабочих станций. Эта технология основана на отказе от использования разделяемых линий связи между всеми узлами сегмента и использовании коммутаторов, позволяющих одновременно передавать пакеты между всеми его парами портов.

Функционально многопортовый коммутатор работает как многопортовый мост, то есть работает на канальном уровне, анализирует заголовки кадров, автоматически строит адресную таблицу и на основании этой таблицы перенаправляет кадр в один из своих выходных портов или фильтрует его, удаляя из буфера. Новшество заключалось в параллельной обработке поступающих кадров, в то время как мост обрабатывает кадр за кадром. Коммутатор же обычно имеет несколько внутренних процессоров обработки кадров, каждый из которых может выполнять алгоритм моста. Таким образом, можно считать, что коммутатор - это мультипроцессорный мост, имеющий за счет внутреннего параллелизма высокую производительность.

Структурная схема коммутатора EtherSwitch, предложенного фирмой Kalpana, представлена на рисунке 2.9.

Каждый порт обслуживается одним процессором пакетов Ethernet - *EPP (Ethernet Packet Processor).* Кроме того, коммутатор имеет системный модуль, который координирует работу всех процессоров EPP. Системный модуль ведет общую адресную таблицу коммутатора и обеспечивает управление коммутатором по протоколу SNMP. Для передачи кадров между портами используется коммутационная матрица, подобная тем, которые работают в телефонных коммутаторах или мультипроцессорных компьютерах, соединяя несколько процессоров с несколькими модулями памяти.

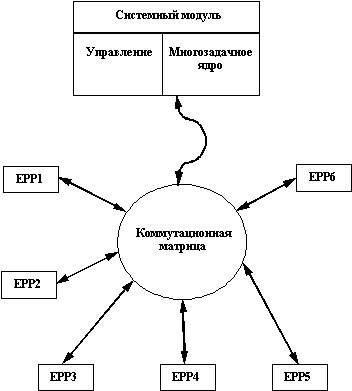


Рис. 2.9. Структура коммутатора Kalpana

При поступлении кадра в какой-либо порт процессор EPP буферизует несколько первых байт кадра, для того, чтобы прочитать адрес назначения. После получения адреса назначения процессор сразу же принимает решение о передаче пакета, не дожидаясь прихода остальных байт кадра. Для этого он просматривает свой собственный кэш адресной таблицы, а если не находит там нужного адреса, то обращается к системному модулю, который работает в многозадачном режиме, параллельно обслуживая запросы всех процессоров EPP. Системный модуль производит просмотр общей адресной таблицы и возвращает процессору найденную строку, которую тот буферизует в своем кэше для последующего использования.

После нахождения адреса назначения в адресной таблице, процессор EPP знает, что нужно дальше делать с поступающим кадром (во время просмотра адресной таблицы процессор продолжал буферизацию поступающих в порт байт кадра). Если кадр нужно отфильтровать, то процессор просто прекращает записывать в буфер байты кадра и ждет поступления нового кадра.

Если же кадр нужно передать на другой порт, то процессор обращается к коммутационной матрице и пытается установить в ней путь, связывающий его порт с портом адреса назначения. Коммутационная матрица может это сделать только в том случае, когда порт адреса назначения в этот момент свободен, то есть не соединен с другим портом. Если же порт занят, то кадр полностью буферизуется процессором входного порта, после чего процессор ожидает освобождения выходного порта и образования коммутационной матрицей нужного пути.

После того, как нужный путь установился, в него направляются буферизованные байты кадра, которые принимаются процессором выходного порта, а после получения им доступа к среде передаются в сеть. Процессор входного порта постоянно хранит несколько байт принимаемого кадра в своем буфере, что позволяет ему независимо и асинхронно принимать и передавать байты кадра (рисунок 2.10).

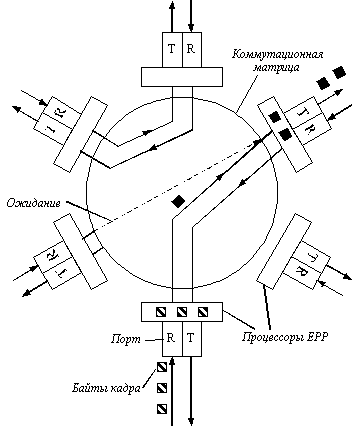


Рис. 2.10. Передача кадра через коммутационную матрицу

При свободном, в момент приема кадра, состоянии выходного порта задержка между приемом первого байта кадра коммутатором и появлением этого же байта на выходе порта адреса назначения составляла у коммутатора компании Kalpana всего 40 мкс, что было гораздо меньше задержки кадра при его передаче мостом.

Описанный способ передачи кадра без его полной буферизации получил название коммутации "на лету" ( "on-the-fly") или "навылет" ("cut-through"). Этот способ представляет по сути конвейерную обработку кадра, когда частично совмещаются во времени несколько этапов его передачи (рисунок 2.11):

1. Прием первых байт кадра процессором входного порта, включая прием байт адреса назначения.
2. Поиск адреса назначения в адресной таблице коммутатора (в кэше процессора или в общей таблице системного модуля).
3. Коммутация матрицы.
4. Прием остальных байт кадра процессором входного порта.
5. Прием байт кадра (включая первые) процессором выходного порта через коммутационную матрицу.
6. Получение доступа к среде процессором выходного порта.
7. Передача байт кадра процессором выходного порта в сеть.

Этапы 2 и 3 совместить во времени нельзя, так как без знания номера выходного порта операция коммутации матрицы не имеет смысла.

По сравнению с режимом полной буферизации кадра, также приведенном на рисунке 2.11, экономия от конвейеризации получается ощутимой.

Однако, главной причиной повышения производительности сети при использовании коммутатора является параллельная обработка нескольких кадров.

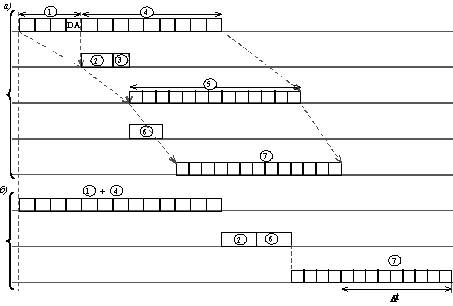


Рис. 2.11. Экономия времени при конвейерной обработке кадра   
а) конвейерная обработка; б) обычная обработка с полной буферизацией

Рисунок 2.12 иллюстрирует этот эффект. На рисунке изображена идеальная в отношении повышения производительности ситуация, когда два порта из 4-х, подключенных к коммутатору, передают данные с максимальной для протокола Ethernet скоростью 10 Мб/с, причем они передают эти данные на остальные два порта коммутатора не конфликтуя - у каждого входного порта свой выходной порт. Если коммутатор обладает способностью успевать обрабатывать входной трафик даже при максимальной интенсивности поступления кадров на входные порты, то общая производительность коммутатора в приведенном примере составит 2(10 Мб/с, а при обобщении примера на N портов - (N/2)(10 Мб/с. Говорят, что коммутатор предоставляет каждой станции или сегменту, подключенным к его портам, выделенную пропускную способность протокола.

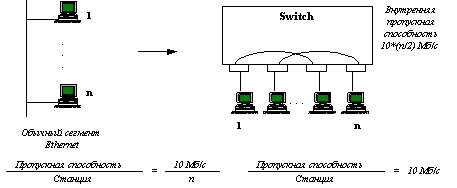


Рис. 2.12. Повышение производительности сети за счет одновременной   
обработки нескольких кадров

Первый коммутатор для локальных сетей не случайно появился для технологии Ethernet. Кроме очевидной причины, связанной с наибольшей популярностью сетей Ethernet, существовала и другая, не менее важная причина - эта технология больше других страдает от повышения времени ожидания доступа к среде при повышении загрузки сегмента. Поэтому сегменты Ethernet в крупных сетях в первую очередь нуждались в средстве разгрузки узких мест сети, и этим средством стали коммутаторы фирмы Kalpana, а затем и других компаний.

Некоторые компании стали развивать технологию коммутации и для повышения производительности других технологий локальных сетей, таких как Token Ring и FDDI. Так как в основе технологии коммутации лежит алгоритм работы прозрачного моста, то принцип коммутации не зависит от метода доступа, формата пакета и других деталей каждой технологии. Коммутатор изучает на основании проходящего через него трафика адреса конечных узлов сети, строит адресную таблицу сети и затем на ее основании производит межкольцевые передачи в сетях Token Ring или FDDI (рисунок 2.13). Принцип работы коммутатора в сетях любых технологий оставался неизменным, хотя внутренняя организация коммутаторов различных производителей иногда очень отличалась от структуры первого коммутатора EtherSwitch.

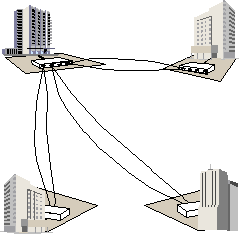


Рис. 2.13. Коммутация колец FDDI

Широкому применению коммутаторов безусловно способствовало то обстоятельство, что внедрение технологии коммутации требовало замены только концентраторов или просто добавления коммутаторов для разделения сегментов, образованных с помощью коммутаторов на более мелкие сегменты. Вся огромная установленная база оборудования конечных узлов - сетевых адаптеров, а также кабельной системы, повторителей и концентраторов - оставалась нетронутой, что давало огромную экономию капиталовложений по сравнению с переходом на какую-нибудь совершенно новую технологию, например, АТМ.

Так как коммутаторы, как и мосты, прозрачны для протоколов сетевого уровня, то их появление в сети оставило в неизменном виде не только оборудование и программное обеспечение конечных узлов, но и маршрутизаторы сети, если они там использовались.

Удобство использования коммутатора состоит еще и в том, что это самообучающееся устройство, и, если администратор не нагружает его дополнительными функциями, то конфигурировать его не обязательно - нужно только правильно подключить разъемы кабелей к портам коммутатора, а дальше он будет работать самостоятельно и стараться эффективно выполнять поставленную перед ним задачу повышения производительности сети.

Безусловно, повышение производительности сети при установке коммутатора в общем случае не будет такой значительной, как в примере. На эффективность работы коммутатора влияет много факторов, и в некоторых случаях, как это будет показано ниже, коммутатор может совсем не дать никаких преимуществ по сравнению с концентратором. Примером такого фактора может служить несбалансированность трафика в сети - если порт 1 и порт 2 коммутатора чаще всего обращаются к порту 3 коммутатора, то порт 3 будет периодически занят и недоступен для одного из двух этих портов и входящий в них трафик будет простаивать, ожидая освобождения порта 3

### Полнодуплексные (full-duplex) протоколы локальных сетей - ориентация исключительно на коммутацию кадров

Технология коммутации оставляет метод доступа к среде в неизменном виде. Это позволяет подключать к портам не только отдельные компьютеры, как это было показано на рисунке 2.12, но и сегменты локальных сетей (рисунок 2.14).

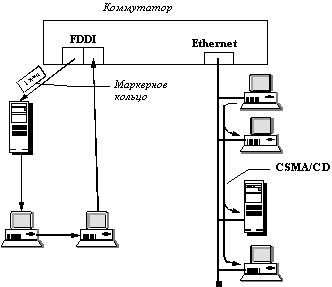


Рис. 2.14. Коммутатор сохраняет в сегментах локальных сетей   
метод доступа к разделяемой среде

Узлы сегмента разделяют общую среду передачи данных, используя либо пассивный коаксиальный кабель, либо концентраторы, как показано в примере, приведенном на рисунке. Если это коммутатор Ethernet, то каждый его порт участвует в процессе обнаружения и отработки коллизий, и без этой функции коммутатор нельзя было бы подключать к сегменту, так как он бы полностью нарушил нормальную работу остальных узлов сегмента. Если это коммутатор колец FDDI, то его порты должны участвовать в процессе захвата и освобождения токена доступа к кольцу в соответствии с алгоритмами МАС-уровня стандарта FDDI.

Однако, когда к каждому порту коммутатора подключен только один компьютер, ситуация становится не такой однозначной.

В обычном режиме работы коммутатор по-прежнему распознает коллизии. Если сеть представляет собой Ethernet на витой паре, то доменом коллизий в этом случае будет участок сети, включающий передатчик коммутатора, приемник коммутатора, передатчик сетевого адаптера компьютера, приемник сетевого адаптера компьютера и две витые пары, соединяющие передатчики с приемниками (рисунок 2.15).

Коллизия возникает, когда передатчики порта коммутатора и сетевого адаптера одновременно или почти одновременно начинают передачу своих кадров, считая, что изображенный на рисунке сегмент свободен. В результате строгого соблюдения правил разделения среды по протоколу Ethernet порт коммутатора и сетевой адаптер используют соединяющий их кабель в полудуплексном режиме, то есть по очереди - сначала кадр или кадры передаются в одном направлении, а затем в другом. При этом максимальная производительность сегмента Ethernet в 14880 кадров в секунду при минимальной длине кадра делится между передатчиком порта коммутатора и передатчиком сетевого адаптера. Если считать, что она делится пополам, то каждому предоставляется возможность передавать примерно по 7440 кадров в секунду.

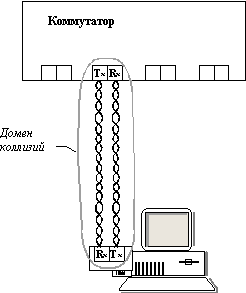


Рис. 2.15. Домен коллизий, образуемый компьютером и портом коммутатора

В то же время, передатчик и приемник как сетевого адаптера, так и порта коммутатора способны принимать и передавать кадры с максимальной скоростью 14880 кадров в секунду. Такая скорость достигается в том случае, когда в течение длительного времени передача идет в одном направлении, например, от компьютера к коммутатору.

Способность оборудования стандарта 10Base-T, то есть Ethernet'a на витой паре, работать с максимальной скоростью в каждом направлении использовали разработчики коммутаторов в своих нестандартных реализациях технологий, получивших название полнодуплексных версий Ethernet, Token Ring, FDDI и т.д.

Полнодуплексный режим работы возможен только при существовании независимых каналов обмена данными для каждого направления и при соединении "точка-точка" двух взаимодействующих устройств. Естественно, необходимо, чтобы МАС-узлы взаимодействующих устройств поддерживали этот специальный режим. В случае, когда только один узел будет поддерживать полнодуплексный режим, второй узел будет постоянно фиксировать коллизии и приостанавливать свою работу, в то время как другой узел будет продолжать передавать данные, которые никто в этот момент не принимает.

Так как переход на полнодуплексный режим работы требует изменения логики работы МАС-узлов и драйверов сетевых адаптеров, то он сначала был опробован при соединении двух коммутаторов. Уже первые модели коммутатора EtherSwitch компании Kalpana поддерживали полнодуплексный режим при взаимном соединении, поддерживая скорость взаимного обмена 20 Мб/с.

Позже появились версии полнодуплексного соединения FDDI-коммутаторов, которые при одновременном использовании двух колец FDDI обеспечивали скорость обмена в 200 Мб/с.

Сейчас для каждой технологии можно найти модели коммутаторов, которые поддерживают полнодуплексный обмен при соединении коммутатор-коммутатор. Существуют коммутаторы, которые позволяют объединить два коммутатора полнодуплексным каналом более чем по одной паре портов. Например, коммутаторы LattisSwitch 28115 компании Bay Networks имеют по два порта, с помощью которых можно соединять коммутаторы, образуя полнодуплексный канал с производительностью 400 Мб/c (рисунок 2.16).

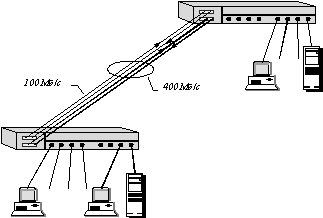


Рис. 2.16. Транковое полнодуплексное соединение коммутаторов   
LattisSwitch 28115 компании Bay Networks

Такие соединения называются транковыми и являются частной разработкой каждой компании, выпускающей коммуникационное оборудование, так как нарушают не только логику доступа к разделяемым средам, но и топологию соединения мостов, запрещающую петлевидные контуры (а такой контур всегда образуется при соединении коммутаторов более чем одной парой портов). При соединении коммутаторов разных производителей транк работать не будет, так как каждый производитель добавляет к логике изучения адресов сети коммутатором по транковой связи что-то свое, чтобы добиться от него правильной работы.

После опробования полнодуплексной технологии на соединениях коммутатор-коммутатор разработчики реализовали ее и в сетевых адаптерах, в основном адаптерах Ethernet и Fast Ethernet. Многие сетевые адаптеры сейчас могут поддерживать оба режима работы, отрабатывая логику алгоритма доступа CSMA/CD при подключении к порту концентратора и работая в полнодуплексном режиме при подключении к порту коммутатора.

Однако, необходимо осознавать, что отказ от поддержки алгоритма доступа к разделяемой среде без какой-либо модификации протокола ведет к повышению вероятности потерь кадров коммутаторами, а, следовательно, к возможному снижению полезной пропускной способности сети (по отношению к переданным данным приложений) вместо ее повышения.

В разделе 2.2 уже говорилось о том, что использование мостов несет в себе потенциальную угрозу потерь кадров при превышении интенсивности входного потока производительности моста. Коммутаторы встречаются с аналогичной проблемой, даже если их внутренняя производительность выше, чем требуется для обслуживания входных потоков, поступающих на каждый порт с максимально возможной скоростью, то есть выше, чем N(C, где N - число портов коммутатора, а С - максимальная скорость передачи пакетов по протоколу, поддерживаемому коммутатором (например, 148809 кадров в секунду, если коммутатор поддерживает протокол Fast Ethernet на всех своих портах).

Причина здесь в ограниченной пропускной способности отдельного порта, которая определяется не производительностью процессора, который обслуживает порт, а временными параметрами протокола. Например, порт Ethernet не может передавать больше 14880 кадров в секунду, если он не нарушает временных соотношений, установленных стандартом.

Поэтому, если входной трафик неравномерно распределяется между выходными портами, то легко представить ситуацию, когда в какой-либо выходной порт коммутатора будет направляться трафик с суммарной средней интенсивностью большей, чем протокольный максимум. На рисунке 2.17 изображена как раз такая ситуация, когда в порт 3 коммутатора направляется трафик от портов 1, 2, 4 и 6, с суммарной интенсивностью в 22 100 кадров в секунду. Порт 3 оказывается загружен на 150%. Естественно, что когда кадры поступают в буфер порта со скоростью 20 100 кадров в секунду, а уходят со скоростью 14 880 кадров в секунду, то внутренний буфер выходного порта начинает неуклонно заполняться необработанными кадрами.

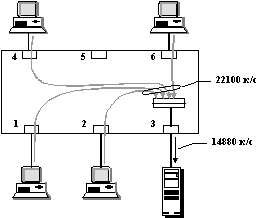


Рис. 2.17. Переполнение буфера порта из-за несбалансированности трафика

Какой бы ни был объем буфера порта, он в какой-то момент времени обязательно переполнится. Нетрудно подсчитать, что при размере буфера в 100 Кбайт в приведенном примере полное заполнение буфера произойдет через 0.22 секунды после начала его работы (буфер такого размера может хранить до 1600 кадров размером в 64 байта). Увеличение буфера до 1 Мбайта даст увеличение времени заполнения буфера до 2.2 секунд, что также неприемлемо.

В территориальных сетях технология коммутации кадров и пакетов применяется уже очень давно. Сети Х.25 используют ее уже более 20 лет. Технологию коммутации используют и новые территориальные сети, в частности сети frame relay и АТМ. В этих сетях конечные узлы подключаются к коммутаторам полнодуплексными каналами связи, такие же каналы используются и для соединения коммутаторов между собой. Протоколы территориальных сетей сразу разрабатывались для организации полнодуплексной связи между узлами сети, поэтому в них были заложены процедуры управления потоком данных. Эти процедуры использовались коммутаторами для снижении интенсивности поступления кадров на входные порты в случае заполнения внутренних буферов коммутатора свыше опасного предела. В таких ситуациях коммутатор направлял соседнему узлу специальный служебный кадр "Приемник не готов", при получении которого соседний узел обязан был приостановить передачу кадров по данному порту. При перегрузках сети в конце концов служебные кадры доходили и до конечных узлов - компьютеров - которые прекращали на время заполнять сеть кадрами, пока имеющиеся в буферах кадры не передавались узлам назначения. Вероятность потери кадров при наличии встроенных в протокол процедур управления потоком становится очень небольшой.

При разработке коммутаторов локальных сетей ситуация коренным образом отличалась от ситуации, при которой создавались коммутаторы территориальных сетей. Основной задачей было сохранение конечных узлов в неизменном виде, что исключало корректировку протоколов локальных сетей. А в этих протоколах процедур управления потоком не было - использование общей среды передачи данных в режиме разделения времени исключало возникновение ситуаций, когда сеть переполнялась бы необработанными кадрами. Сеть не накапливала данных в каких-либо промежуточных буферах при использовании только повторителей или концентраторов.

Поэтому применение коммутаторов без изменения протокола работы оборудования всегда порождает опасность потерь кадров. Если порты коммутатора работают в обычном, то есть в полудуплексном режиме, то у коммутатора имеется возможность оказать некоторое воздействие на конечный узел и заставить его приостановить передачу кадров, пока у коммутатора не разгрузятся внутренние буфера. Нестандартные методы управления потоком в коммутаторах при сохранении протокола доступа в неизменном виде будут рассмотрены ниже.

Если же коммутатор работает в полнодуплексном режиме, то протокол работы конечных узлов, да и его портов все равно меняется. Поэтому имело смысл для поддержки полнодуплексного режима работы коммутаторов разработать новые протоколы взаимодействия узлов, которые бы использовали явные и стандартные механизмы управления потоком при сохранении неизменным только формата кадров. Сохранение формата кадров необходимо для того, чтобы к одному и тому же коммутатору можно было бы подключать новые узлы, имеющие сетевые адаптеры полнодуплексного режима, и старые узлы или сегменты узлов, поддерживающие алгоритм доступа к разделяемой среде.

Работа над выработкой стандарта для полнодуплексных версий Ethernet, Fast Ethernet и других технологий локальных сетей идет уже несколько лет, однако на момент написания этого пособия такие стандарты пока не приняты из-за разногласий членов соответствующих комитетов по стандартизации, отстаивающих подходы фирм, в которых они работают.

Тем не менее, каждая из крупных компаний, выпускающих коммуникационное оборудование, имеет свою версию полнодуплексных технологий и поддерживает их в своих продуктах - сетевых адаптерах и коммутаторах. Эти версии используют встроенные процедуры управления потоком. Обычно это несложные процедуры, использующие две команды - "Приостановить передачу" и "Возобновить передачу" - для управления потоком кадров соседнего узла сети.

## ATM-коммутация

Кроме коммутаторов, поддерживающих стандартные протоколы локальных сетей и передающих кадры с порта на порт по алгоритмам моста, в локальных сетях стали применяться коммутаторы другого вида, а именно коммутаторы технологии АТМ. В связи с этим коротко рассмотрим основные принципы работы таких коммутаторов и способы их взаимодействия с коммутаторами технологий локальных сетей.

Технология *АТМ (Asynchronous Transfer Mode* - режим асинхронной передачи) разрабатывалась изначально для совмещения синхронного голосового трафика и асинхронного компьютерного трафика в рамках одной территориальной сети. Затем сфера применения технологии АТМ была расширена и на локальные сети. В данном курсе мы не будем рассматривать все аспекты технологии АТМ, а ограничимся изучением способов коммутации данных в сетях АТМ, которые используются в коммутаторах АТМ, применяемых в локальных сетях. Такие АТМ-коммутаторы чаще всего не используют все возможности технологии, в частности поддержку синхронного трафика, в основном из-за отсутствия приложений, которые могли бы воспользоваться таким сервисом.

Сеть АТМ изначально разрабатывалась для поддержки полнодуплексного высокоскоростного режима обмена как между узлами сети, так и между ее коммутаторами (рисунок 3.1).

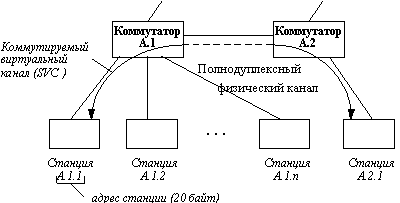


Рис. 3.1. Структура сети АТМ

АТМ-станции и АТМ-коммутаторы обмениваются между собой кадрами фиксированного размера в 53 байта. Эти кадры принято называть ячейками. Поле данных ячейки занимает 48 байт, а заголовок - 5 байт. Адреса конечных узлов локальных сетях АТМ составляют 20 байт.

Для того, чтобы пакеты содержали адрес узла назначения, и в то же время процент служебной информации не был большим по сравнению с размером поля данных пакета, в технологии ATM применен стандартный для глобальных вычислительных сетей прием - передача ячеек по виртуальным каналам. Техника коммутации данных в соответствии с номерами их виртуальных каналов давно использовалась в сетях Х.25, а затем нашла применение и в новых технологиях территориальных сетей - frame relay и АТМ.

Принцип коммутации пакетов на основе виртуальных каналов поясняется рисунком 3.2. Конечные узлы не могут просто начать обмениваться данными, как это принято в большинстве протоколов канального уровня локальных сетей. Они должны перед обменом установить между собой логическое соединение. При установлении соединения между конечными узлами используется специальный тип пакета - запрос на установление соединения - который содержит многоразрядный адрес узла-адресата, а также номер виртуального соединения, присвоенного данному соединению в узле-отправителе, например, 15. Ячейки АТМ имеют 3-х байтное поле номера виртуального соединения, что позволяет коммутаторам и конечным узлам поддерживать одновременно очень большое количество виртуальных соединений.

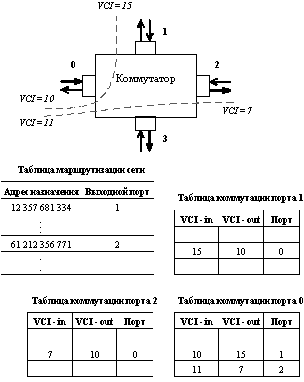


Рис. 3.2. Коммутация в сетях с виртуальными соединениями

Адрес назначения используется для маршрутизации запроса на установление соединения на основании таблиц маршрутизации, аналогичных тем, которые используются маршрутизаторами IP или IPX. В этих таблицах для каждого адреса назначения (или для группы адресов, имеющих общую старшую часть, соответствующую адресу сети) указывается номер порта, на который нужно передать приходящий пакет. Таблица маршрутизации по назначению аналогична адресной таблице коммутатора, но образуется она не путем изучения адресов проходящего трафика, а либо вручную администратором, либо с помощью обмена между коммутаторами АТМ специальных служебных данных о топологии связей сети. Протокол обмена топологической информацией для сетей АТМ имеет название *PNNI - Private Network to Network Interface*. Он разработан и принят в качестве стандарта, хотя не все АТМ-коммутаторы пока его поддерживают.

В приведенном примере в соответствии с таблицей маршрутизации оказалось необходимым передать пакет запроса на установление соединения с порта 1 на порт 0. Одновременно с передачей пакета маршрутизатор изменяет у пакета номер виртуального соединения - он присваивает пакету первый не использованный номер виртуального канала для данного порта данного коммутатора. Каждый конечный узел и каждый коммутатор ведет свой список использованных и свободных номеров виртуальных соединений для своих портов.

Кроме таблицы маршрутизации для каждого порта составляется таблица коммутации. В таблице коммутации входного порта маршрутизатор отмечает, что в дальнейшем пакеты, прибывшие на этот порт с номером 15, должны передаваться на порт 0, причем номер виртуального канала должен быть изменен на 10. Одновременно делается и соответствующая запись в таблице коммутации порта 0 - пакеты, пришедшие по виртуальному каналу 10 в обратном направлении нужно передавать на порт с номером 1, меняя номер виртуального канала на 15.

В результате действия такой схемы пакеты данных уже не несут длинные адреса конечных узлов, а имеют в служебном поле только номер виртуального канала, на основании которого и производится маршрутизация всех пакетов, кроме пакета запроса на установление соединения. В сети прокладывается виртуальный канал, который не изменяется в течение всего времени существования соединения. Пакеты в виртуальном канале циркулируют в двух направлениях, то есть в полнодуплексном режиме, причем, конечные узлы не замечают изменений номеров виртуальных каналов при прохождении пакетов через сеть.

После образования таблицы коммутации, ячейки АТМ обрабатываются коммутаторами АТМ примерно так же, как и коммутаторами технологий локальных сетей. Исключение составляет только режим фильтрации - он отсутствует, так как в АТМ нет разделяемых сред и переданную коммутатору ячейку всегда нужно передать на какой-либо порт.

Виртуальные каналы бывают коммутируемыми *(Switched Virtual Channel)* и постоянными *(Permanent Virtual Channel).* *Коммутируемые виртуальные каналы* устанавливаются узлами динамически, в процессе работы, а *постоянные виртуальные каналы* образуются администратором на продолжительный срок. Для постоянных виртуальных каналов не нужно выполнять процедуру установления соединения, так как коммутаторы уже настроены на их обработку - соответствующие таблицы коммутации уже сформированы администратором.

Коммутаторы АТМ, работающие с компьютерным трафиком, предоставляют конечным узлам два вида сервиса. Сервис с неопределенной пропускной способностью *(Unspecified Bit Rate)* подобен сервису коммутаторов локальных сетей - он не гарантирует конечному узлу какой-то определенной доли пропускной способности сети и не гарантирует, что все ячейки конечного узла будут доставлены по назначению. Это самый простой вид сервиса и он не использует какие-либо процедуры управления потоком, а при переполнении буферов коммутатора приходящие ячейки отбрасываются точно так же, как это делают коммутаторы локальных сетей.

Сервис *ABR (Available Bit Rate)* в отличие от сервиса UBR использует технику управления потоком для предотвращения перегрузок сети и дает некоторые гарантии доставки ячеек узлу назначения.

Для этого при установлении соединения ABR между конечным узлом и коммутаторами сети заключается соглашение о двух скоростях передачи данных - пиковой скорости и минимальной скорости. Заключение соглашения о параметрах трафика - прием, в локальных сетях обычно не применяющийся. Пользователь соединения ABR соглашается не передавать данные со скоростью, выше пиковой, то есть PCR, а сеть соглашается всегда обеспечивать минимальную скорость передачи ячеек - MCR.

Если приложение при установлении ABR-соединения не определяет максимальную и минимальную скорости, то по умолчанию считается, что максимальная скорость совпадает со скоростью линии доступа станции к сети, а минимальная скорость считается равной нулю.

Пользователь соединения ABR получает гарантированное качество сервиса в отношении потери ячеек и пропускной способности, а сеть при использовании трафика ABR не переполняется.

Для преобразования кадров, циркулирующих в локальных сетях, в 53-байтные ячейки, в технологии АТМ определены функции сегментации и сборки *(Segmentation And Reassembling).* Когда кадр поступает в коммутатор АТМ, то он с помощью функции сегментации разделить его на последовательность ячеек. После передачи ячеек по сети коммутаторов АТМ они вновь собираются в последнем коммутаторе с помощью функции реассемблирования в исходный кадр.

Технология АТМ работает с несколькими скоростями доступа конечных узлов к сети. Чаще всего используется скорость 155 Мб/c, более редкой является скорость доступа в 622 Мб/с. Существует и низкоскоростной доступ по линии в 25 Мб/с. Иерархия скоростей доступа - это также одна из особенностей технологии АТМ, делающей ее очень удобной для применения в сложных сетях. При насыщении какой-либо части сети слишком интенсивным трафиком конечных узлов не нужно переходить на принципиально новую технологию, достаточно просто установить новый, более скоростной интерфейсный модуль коммутатора.

Очевидно, что различные принципы коммутации кадров в коммутаторах локальных сетей и в коммутаторах АТМ требуют использования каких-то устройств, согласующих работу этих коммутаторов. Одной функции преобразования кадров и ячеек с помощью функций SAR явно недостаточно, так как нужно на основании МАС-адресов конечных узлов сети устанавливать виртуальные пути ячеек через АТМ-коммутаторы.

Существуют частные решения отдельных производителей, позволяющие в рамках одного коммутатора совмещать обе технологии. Обычно, для подключения конечных пользователей используются порты традиционных технологий локальных сетей, например, Ethernet, а коммутаторы используют для обмена между собой технологию АТМ, более масштабируемую.

Имеется и стандартный вариант решения этой задачи. Он носит название LAN Emulation - эмуляции локальных сетей и рассматривается далее.

## Особенности коммутаторов локальных сетей

### Техническая реализация коммутаторов

После того, как технология коммутации привлекла общее внимание и получила высокие оценки специалистов, многие компании занялись реализацией этой технологии в своих устройствах, применяя для этого различные технические решения. Многие коммутаторы первого поколения были похожи на маршрутизаторы, то есть основывались на центральном процессоре общего назначения, связанном с интерфейсными портами по внутренней скоростной шине (рисунок 4.1). Однако, это были скорее пробные устройства, предназначенные для освоения самой компании технологии коммутации, а не для завоевания рынка.

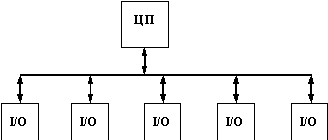


Рис. 4.1. Коммутатор на процессоре общего назначения

Основным недостатком таких коммутаторов была их низкая скорость. Универсальный процессор никак не мог справиться с большим объемом специализированных операций по пересылке кадров между интерфейсными модулями.

Для ускорения операций коммутации нужны были специализированные процессоры со специализированными средствами обмена данными, как в первом коммутаторе Kalpana, и они вскоре появились. Теперь коммутаторы используют заказные специализированные БИС, которые оптимизированы для выполнения основных операций коммутации. Часто в одном коммутаторе используется несколько специализированных БИС, каждая из которых выполняет функционально законченную часть операций.

В настоящее время коммутаторы используют в качестве базовой одну из трех схем взаимодействия своих блоков или модулей:

* коммутационная матрица;
* разделяемая многовходовая память;
* общая шина.

Часто эти три способа взаимодействия комбинируются в одном коммутаторе.

#### Коммутаторы на основе коммутационной матрицы

*Коммутационная матрица* - основной и самый быстрый способ взаимодействия процессоров портов, именно он был реализован в первом промышленном коммутаторе локальных сетей. Однако, реализация матрицы возможна только для определенного числа портов, причем сложность схемы возрастает пропорционально квадрату количества портов коммутатора (рисунок 4.2).

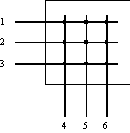


Рис. 4.2. Коммутационная матрица

Более детальное представление одного из возможных вариантов реализации коммутационной матрицы для 8 портов дано на рисунке 4.3. Входные блоки процессоров портов на основании просмотра адресной таблицы коммутатора определяют по адресу назначения номер выходного порта. Эту информацию они добавляют к байтам исходного кадра в виде специального ярлыка - тэга (tag). Для данного примера тэг представляет просто 3-х разрядное двоичное число, соответствующее номеру выходного порта.

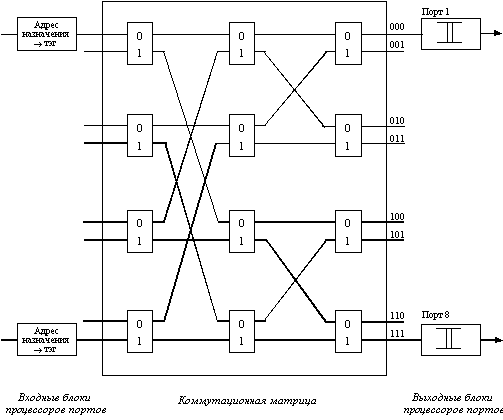


Рис. 4.3. Реализация коммутационной матрицы 4х4 с помощью двоичных переключателей

Матрица состоит из трех уровней двоичных переключателей, которые соединяют свой вход с одним из двух выходов в зависимости от значения бита тэга. Переключатели первого уровня управляются первым битом тэга, второго - вторым, а третьего - третьим.

Матрица может быть реализована и по-другому, на основании комбинационных схем другого типа, но ее особенностью все равно остается технология коммутации физических каналов. Известным недостатком этой технологии является отсутствие буферизации данных внутри коммутационной матрицы - если составной канал невозможно построить из-за занятости выходного порта или промежуточного коммутационного элемента, то данные должны накапливаться в их источнике, в данном случае - во входном блоке порта, принявшего кадр.

#### Коммутаторы с общей шиной

Коммутаторы с общей шиной используют для связи процессоров портов высокоскоростную шину, используемую в режиме разделения времени. Эта архитектура похожа на изображенную на рисунке 4.1 архитектуру коммутаторов на основе универсального процессора, но отличается тем, что шина здесь пассивна, а активную роль выполняют специализированные процессоры портов.

Пример такой архитектуры приведен на рисунке 4.4. Для того, чтобы шина не была узким местом коммутатора, ее производительность должна быть по крайней мере в N/2 раз выше скорости поступления данных во входные блоки процессоров портов. Кроме этого, кадр должен передаваться по шине небольшими частями, по несколько байт, чтобы передача кадров между несколькими портами происходила в псевдопараллельном режиме, не внося задержек в передачу кадра в целом. Размер такой ячейки данных определяется производителем коммутатора. Некоторые производители, например, LANNET (сейчас подразделение компании Madge Networks), выбрали в качестве порции данных, переносимых за одну операцию по шине, ячейку АТМ с ее полем данных в 48 байт. Такой подход облегчает трансляцию протоколов локальных сетей в протокол АТМ, если коммутатор поддерживает эти технологии.

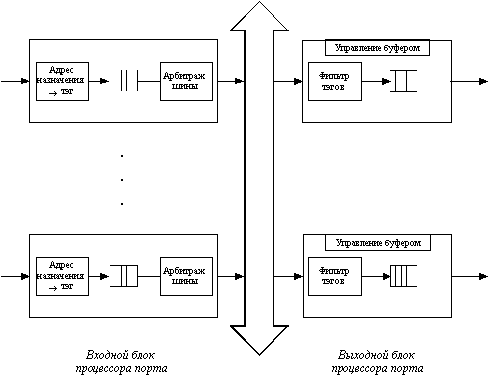


Рис. 4.4. Архитектура общей шины

Входной блок процессора помещает в ячейку, переносимую по шине, тэг, в котором указывает номер порта назначения. Каждый выходной блок процессора порта содержит фильтр тэгов, который выбирает тэги, предназначенные данному порту.

Шина, так же как и коммутационная матрица, не может осуществлять промежуточную буферизацию, но так как данные кадра разбиваются на небольшие ячейки, то задержек с начальным ожиданием доступности выходного порта в такой схеме нет.

#### Коммутаторы с разделяемой памятью

Третья базовая архитектура взаимодействия портов - двухвходовая разделяемая память. Пример такой архитектуры приведен на рисунке 4.5.

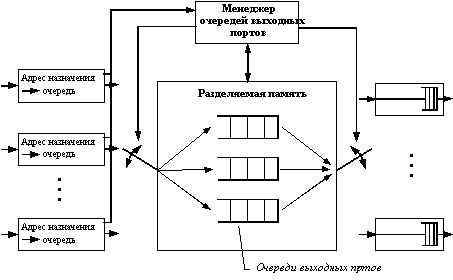


Рис. 4.5. Архитектура разделяемой памяти

Входные блоки процессоров портов соединяются с переключаемым входом разделяемой памяти, а выходные блоки этих же процессоров соединяются с переключаемым выходом этой памяти. Переключением входа и выхода разделяемой памяти управляет менеджер очередей выходных портов. В разделяемой памяти менеджер организует несколько очередей данных, по одной для каждого выходного порта. Входные блоки процессоров передают менеджеру портов запросы на запись данных в очередь того порта, который соответствует адресу назначения пакета. Менеджер по очереди подключает вход памяти к одному из входных блоков процессоров и тот переписывает часть данных кадра в очередь определенного выходного порта. По мере заполнения очередей менеджер производит также поочередное подключение выхода разделяемой памяти к выходным блокам процессоров портов, и данные из очереди переписываются в выходной буфер процессора.

Память должна быть достаточно быстродействующей для поддержания скорости переписи данных между N портами коммутатора. Применение общей буферной памяти, гибко распределяемой менеджером между отдельными портами, снижает требования к размеру буферной памяти процессора порта.

#### Комбинированные коммутаторы

У каждой из описанных архитектур есть свои преимущества и недостатки, поэтому часто в сложных коммутаторах эти архитектуры применяются в комбинации друг с другом. Пример такого комбинирования приведен на рисунке 4.6.

Коммутатор состоит из модулей с фиксированным количеством портов (2 - 8), выполненных на основе специализированной БИС (ASIC), реализующей архитектуру коммутационной матрицы. Если порты, между которыми нужно передать кадр данных, принадлежат одному модулю, то передача кадра осуществляется процессорами модуля на основе имеющейся в модуле коммутационной матрицы. Если же порты принадлежат разным модулям, то процессоры общаются по общей шине. При такой архитектуре передача кадров внутри модуля будет происходить чаще всего быстрее, чем при межмодульной передаче, так как коммутационная матрица - наиболее быстрый, хотя и наименее масштабируемый способ взаимодействия портов. Скорость внутренней шины коммутаторов может достигать нескольких Гб/c, а у наиболее мощных моделей - до 10 - 14 Гб/с.

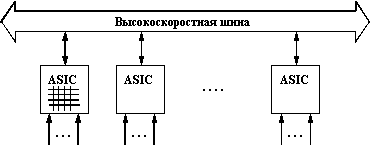


Рис. 4.6. Комбинирование архитектур коммутационной матрицы и общей шины

Можно представить и другие способы комбинировании архитектур, например, использование для взаимодействия модулей разделяемой памяти.

#### Модульные и стековые коммутаторы

В конструктивном отношении коммутаторы делятся на:

* автономные коммутаторы с фиксированным количеством портов;
* модульные коммутаторы на основе шасси;
* коммутаторы с фиксированным количеством портов, собираемые в стек.

Первый тип коммутаторов обычно предназначен для организации небольших рабочих групп.

Модульные коммутаторы на основе шасси чаще всего предназначены для применения на магистрали сети. Поэтому они выполняются на основе какой-либо комбинированной схемы, в которой взаимодействие модулей организуется по быстродействующей шине или же на основе быстрой разделяемой памяти большого объема. Модули такого коммутатора выполняются на основе технологии "hot swap", то есть допускают замену на ходу, без выключения коммутатора, так как центральное коммуникационное устройство сети не должно иметь перерывов в работе. Шасси обычно снабжается резервированными источниками питания и резервированными вентиляторами, в тех же целях. В целом такие коммутаторы напоминают маршрутизаторы высшего класса или корпоративные многофункциональные концентраторы, поэтому иногда они включают помимо модулей коммутации и модули повторителей или маршрутизатров.

С технической точки зрения определенный интерес представляют стековые коммутаторы. Эти устройства представляют собой коммутаторы, которые могут работать автономно, так как выполнены в отдельном корпусе, но имеют специальные интерфейсы, которые позволяют их объединять в общую систему, которая работает как единый коммутатор. Говорят, что в этом случае отдельные коммутаторы образуют стек.

Обычно такой специальный интерфейс представляет собой высокоскоростную шину, которая позволяет объединить отдельные корпуса подобно модулям в коммутаторе на основе шасси. Так как расстояния между корпусами больше, чем между модулями на шасси, скорость обмена по шине обычно ниже, чем у модульных коммутаторов: 200 - 400 Мб/c. Не очень высокие скорости обмена между коммутаторами стека обусловлены также тем, что стековые коммутаторы обычно занимают промежуточное положение между коммутаторами с фиксированным количеством портов и коммутаторами на основе шасси. Стековые коммутаторы применяются для создания сетей рабочих групп и отделов, поэтому сверхвысокие скорости шин обмена им не очень нужны и не соответствуют их ценовому диапазону.

Структура стека коммутаторов, соединяемых по скоростным специальным портам, показана на рисунке 4.7.

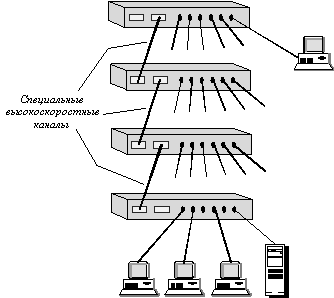


Рис. 4.7. Стек коммутаторов, объединяемых по высокоскоростным каналам

Компания Cisco предложила другой подход к организации стека. Ее коммутатор Catalyst 3000 (ранее называвшийся EtherSwitch Pro Stack) также имеет специальный скоростной интерфейс 280 Мб/с для организации стека, но с его помощью коммутаторы соединяются не друг с другом, а с отдельным устройством, содержащим коммутационную матрицу 8(8, организующую более высокопроизводительный обмен между любыми парами коммутаторов.

### Характеристики производительности коммутаторов

Основными характеристиками коммутатора, измеряющими его производительность, являются:

* скорость фильтрации (filtering);
* скорость маршрутизации (forwarding);
* пропускная способность (throughput);
* задержка передачи кадра.

Кроме того, существует несколько характеристик коммутатора, которые в наибольшей степени влияют на указанные характеристики производительности. К ним относятся:

* размер буфера (буферов) кадров;
* производительность внутренней шины;
* производительность процессора или процессоров;
* размер внутренней адресной таблицы.

#### Скорость фильтрации и скорость продвижения

Скорость фильтрации и продвижения кадров - это две основные характеристики производительности коммутатора. Эти характеристики являются интегральными показателями, они не зависят от того, каким образом технически реализован коммутатор.

*Скорость фильтрации* определяет скорость, с которой коммутатор выполняет следующие этапы обработки кадров:

* прием кадра в свой буфер,
* просмотр адресной таблицы с целью нахождения порта для адреса назначения кадра,
* уничтожение кадра, так как его порт назначения совпадает с портом-источником.

*Скорость продвижения* определяет скорость, с которой коммутатор выполняет следующие этапы обработки кадров:

* прием кадра в свой буфер,
* просмотр адресной таблицы с целью нахождения порта для адреса назначения кадра,
* передача кадра в сеть через найденный по адресной таблице порт назначения.

Как скорость фильтрации, так и скорость продвижения измеряются обычно в кадрах в секунду. Если в характеристиках коммутатора не уточняется, для какого протокола и для какого размера кадра приведены значения скоростей фильтрации и продвижения, то по умолчанию считается, что эти показатели даются для протокола Ethernet и кадров минимального размера, то есть кадров длиной 64 байта (без преамбулы), с полем данных в 46 байт. Если скорости указаны для какого-либо определенного протокола, например, Token Ring или FDDI, то они также даны для кадров минимальной длины этого протокола (например, кадров длины 29 байт для протокола FDDI). Применение в качестве основного показателя скорости работы коммутатора кадров минимальной длины объясняется тем, что такие кадры всегда создают для коммутатора наиболее тяжелый режим работы по сравнению с кадрами другого формата при равной пропускной способности переносимых пользовательских данных. Поэтому при проведении тестирования коммутатора режим передачи кадров минимальной длины используется как наиболее сложный тест, который должен проверить способность коммутатора работать при наихудшем сочетании для него параметров трафика. Кроме того, для пакетов минимальной длины скорость фильтрации и продвижения имеют максимальное значение, что имеет немаловажное значение при рекламе коммутатора.

*Пропускная способность коммутатора* измеряется количеством переданных в единицу времени через его порты пользовательских данных. Так как коммутатор работает на канальном уровне, то для него пользовательскими данными являются те данные, которые переносятся в поле данных кадров протоколов канального уровня - Ethernet, Token Ring, FDDI и т.п. Максимальное значение пропускной способности коммутатора всегда достигается на кадрах максимальной длины, так как при этом и доля накладных расходов на служебную информацию кадра гораздо ниже, чем для кадров минимальной длины, и время выполнения коммутатором операций по обработке кадра, приходящееся на один байт пользовательской информации, существенно меньше.

Зависимость пропускной способности коммутатора от размера передаваемых кадров хорошо иллюстрирует пример протокола Ethernet, для которого при передаче кадров минимальной длины достигается скорость передачи в 14880 кадров в секунду и пропускная способность 5.48 Мб/с, а при передаче кадров максимальной длины - скорость передачи в 812 кадров в секунду и пропускная способность 9.74 Мб/c. Пропускная способность падает почти в два раза при переходе на кадры минимальной длины, и это еще без учета потерь времени на обработку кадров коммутатором.

*Задержка передачи кадра* измеряется как время, прошедшее с момента прихода первого байта кадра на входной порт коммутатора до момента появления этого байта на выходном порту коммутатора. Задержка складывается из времени, затрачиваемого на буферизацию байт кадра, а также времени, затрачиваемого на обработку кадра коммутатором - просмотр адресной таблицы, принятие решения о фильтрации или продвижении и получения доступа к среде выходного порта.

Величина вносимой коммутатором задержки зависит от режима его работы. Если коммутация осуществляется "на лету", то задержки обычно невелики и составляют от 10 мкс до 40 мкс, а при полной буферизации кадров - от 50 мкс до 200 мкс (для кадров минимальной длины).

Коммутатор - это многопортовое устройство, поэтому для него принято все приведенные выше характеристики (кроме задержки передачи кадра) давать в двух вариантах. Первый вариант - суммарная производительность коммутатора при одновременной передаче трафика по всем его портам, второй вариант - производительность, приведенная в расчете на один порт.

Так как при одновременной передаче трафика несколькими портами существует огромное количество вариантов трафика, отличающегося размерами кадров в потоке, распределением средней интенсивности потоков кадров между портами назначения, коэффициентами вариации интенсивности потоков кадров и т.д. и т.п., то при сравнении коммутаторов по производительности необходимо принимать во внимание, для какого варианта трафика получены публикуемые данные производительности. К сожалению, для коммутаторов (как, впрочем, и для маршрутизаторов) не существует общепринятых тестовых образцов трафика, которые можно было бы применять для получения сравнимых характеристик производительности, как это делается для получения таких характеристик производительности вычислительных систем, как TPC-А или SPECint92. Некоторые лаборатории, постоянно проводящие тестирование коммуникационного оборудования, разработали детальные описания условий тестирования коммутаторов и используют их в своей практике, однако общепромышленными эти тесты пока не стали.

#### Оценка необходимой общей производительности коммутатора

В идеальном случае коммутатор, установленный в сети, передает кадры между узлами, подключенными к его портам, с той скоростью, с которой узлы генерируют эти кадры, не внося дополнительных задержек и не теряя ни одного кадра. В реальной практике коммутатор всегда вносит некоторые задержки при передаче кадров, а также может некоторые кадры терять, то есть не доставлять их адресатам. Из-за различий во внутренней организации разных моделей коммутаторов, трудно предвидеть, как тот или иной коммутатор будет передавать кадры какого-то конкретного образца трафика. Лучшим критерием по-прежнему остается практика, когда коммутатор ставится в реальную сеть и измеряются вносимые им задержки и количество потерянных кадров. Однако, существуют несложные расчеты, которые могут дать представление о том, как коммутатор будет вести себя в реальной ситуации.

Посмотрим, как можно оценить поведение коммутатора на примере сети, изображенной на рисунке 4.8.

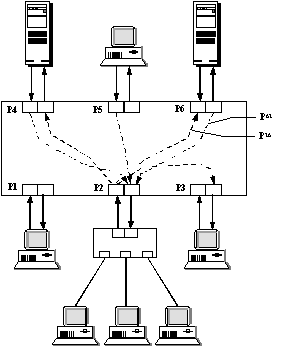


Рис. 4.8. Распределение трафика в сети, построенной на коммутаторе

Основой для оценки того, как будет справляться коммутатор со связью узлов или сегментов, подключенных к его портам, являются данные о средней интенсивности трафика между узлами сети. Для приведенного примера это означает, что нужно каким-то образом оценить, сколько в среднем кадров в секунду узел, подключенный к порту P2, генерирует узлу, подключенному к порту P4 (трафик P24), узлу, подключенному к порту P3 (трафик P23), и так далее, до узла, подключенного к порту P6. Затем эту процедуру нужно повторить для трафика, генерируемого узами, подключенными к портам 3, 4, 5 и 6. В общем случае, интенсивность трафика, генерируемого одним узлом другому, не совпадает с интенсивностью трафика, генерируемого в обратном направлении.

Результатом исследования трафика будет построение матрицы трафика, приведенной на рисунке 4.9. Трафик можно измерять как в кадрах в секунду, так и в битах в секунду. Так как затем требуемые значения трафика будут сравниваться с показателями производительности коммутатора, то нужно их иметь в одних и тех же единицах. Для определенности будем считать, что в рассматриваемом примере трафик и производительность коммутатора измеряются в битах в секунду.

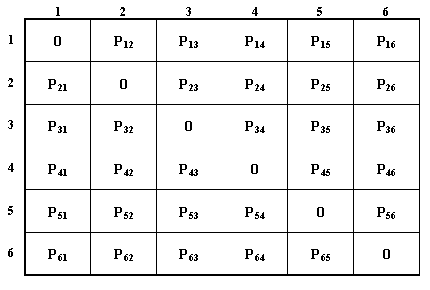


Рис. 4.9. Матрица средних значений интенсивностей трафика

Подобную матрицу строят агенты RMON MIB (переменная Traffic Matrix), встроенные в сетевые адаптеры или другое коммуникационное оборудование.

Для того, чтобы коммутатор справился с поддержкой требуемой матрицы трафика, необходимо выполнение нескольких условий.

1. Общая производительность коммутатора должна быть больше или равна суммарной интенсивности передаваемого трафика:

B ij Pij ,

где B - общая производительность коммутатора, Pij - средняя интенсивность трафика от i-го порта к j-му; сумма берется по всем портам коммутатора, от 1 до 6.

Если это неравенство не выполняется, то коммутатор заведомо не будет справляться с потоком поступающих в него кадров и они будут теряться из-за переполнения внутренних буферов. Так как в формуле фигурируют средние значения интенсивностей трафика, то никакой, даже очень большой размер внутреннего буфера или буферов коммутатора не сможет компенсировать слишком медленную обработку кадров.

Суммарная производительность коммутатора обеспечивается достаточно высокой производительностью каждого его отдельного элемента - процессора порта, коммутационной матрицы, общей шины, соединяющей модули и т.п. Независимо от внутренней организации коммутатора и способов конвейеризации его операций, можно определить достаточно простые требования к производительности его элементов, которые являются необходимыми для поддержки заданной матрица трафика. Перечислим некоторые из них.

2. Номинальная максимальная производительность протокола каждого порта коммутатора должна быть не меньше средней интенсивности суммарного трафика, проходящего через порт:

Сk j Pkj + i Pik,

где Сk - номинальная максимальная производительность протокола k-го порта (например, если k-ый порт поддерживает Ethernet, то Сk равно 10 Мб/с), первая сумма равна интенсивности выходящего из порта трафика, а вторая - входящего. Эта формула полагает, что порт коммутатора работает в стандартном полудуплексном режиме, для полнодуплексного режима величину Сk нужно удвоить.

3. Производительность процессора каждого порта должна быть не меньше средней интенсивности суммарного трафика, проходящего через порт. Условие аналогично предыдущему, но вместо номинальной производительности поддерживаемого протокола в ней должна использоваться производительность процессора порта.

4. Производительность внутренней шины коммутатора должна быть не меньше средней интенсивности суммарного трафика, передаваемого между портами, принадлежащими разным модулям коммутатора:

Bbus ij Pij ,

где Bbus - производительность общей шины коммутатора, а сумма ij Pij берется только по тем i и j, которые принадлежат разным модулям.

Эта проверка должна выполняться, очевидно, только для тех коммутаторов, которые имеют внутреннюю архитектуру модульного типа с использованием общей шины для межмодульного обмена. Для коммутаторов с другой внутренней организацией, например, с разделяемой памятью, несложно предложить аналогичные формулы для проверки достаточной производительности их внутренних элементов.

Приведенные условия являются необходимыми для того, чтобы коммутатор в среднем справлялся с поставленной задачей и не терял кадров постоянно. Если хотя бы одно из приведенных условий не будет выполнено, то потери кадров становятся не эпизодическим явлением при пиковых значениях трафика, а явлением постоянным, так как даже средние значения трафика превышают возможности коммутатора.

Условия 1 и 2 применимы для коммутаторов с любой внутренней организацией, а условия 3 и 4 приведены в качестве примера необходимости учета производительности отдельных портов.

Так как производители коммутаторов стараются сделать свои устройства как можно более быстродействующими, то общая внутренняя производительность коммутатора часто с некоторым запасом превышает среднюю интенсивность любого варианта трафика, который можно направить на порты коммутатора в соответствии с их протоколами. Такие коммутаторы называются *неблокирующими*, что подчеркивает тот факт, что любой вариант трафика передается без снижения его интенсивности.

Однако, какой бы общей производительностью не обладал коммутатор, всегда можно указать для него такое распределение трафика между портами, с которым коммутатор не справится и начнет неизбежно терять кадры. Для этого достаточно, чтобы суммарный трафик, передаваемый через коммутатор для какого-нибудь его выходного порта, превысил максимальную пропускную способность протокола этого порта. В терминах условия 2 это будет означать, что второе слагаемое i Pik превышает пропускную способность протокола порта Сk. Например, если порты P4, Р5 и Р6 будут посылать на порт Р2 каждый по 5 Мб/c, то порт Р2 не сможет передавать в сеть трафик со средней интенсивностью 15 Мб/с, даже если процессор этого порта обладает такой производительностью. Буфер порта Р2 будет заполняться со скоростью 15 Мб/с, а опустошаться со скоростью максимум 10 Мб/с, поэтому количество необработанных данных будет расти со скоростью 5 Мб/с, неизбежно приводя к переполнению любого буфера конечного размера, а значит и к потере кадров.

Из приведенного примера видно, что коммутаторы могут полностью использовать свою высокую внутреннюю производительность только в случае хорошо сбалансированного трафика, когда вероятности передачи кадров от одного порта другим примерно равны. При "перекосах" трафика, когда несколько портов посылают свой трафик преимущественно одному порту, коммутатор может не справиться с поставленной задачей даже не из-за недостаточной производительности своих процессоров портов, а по причине ограничений протокола порта.

Коммутатор может терять большой процент кадров и в тех случаях, когда все приведенные условия соблюдаются, так как они являются необходимыми, но недостаточными для своевременного продвижения получаемых на приемниках портов кадров. Эти условия недостаточны потому, что они очень упрощают процессы передачи кадров через коммутатор. Ориентация только на средние значения интенсивностей потоков не учитывает коллизий, возникающих между передатчиками порта и сетевого адаптера компьютера, потерь на время ожидания доступа к среде и других явлений, которые обусловлены случайными моментами генерации кадров, случайными размерами кадров и другими случайными факторами, значительно снижающими реальную производительность коммутатора. Тем не менее использование приведенных оценок полезно, так как позволяет выявить случаи, когда применение конкретной модели коммутатора для конкретной сети заведомо неприемлемо.

Так как интенсивности потоков кадров между узлами сети оценить удается далеко не всегда, то в заключение этого раздела приведем соотношение, которое позволяет говорить о том, что коммутатор обладает достаточной внутренней производительностью для поддержки потоков кадров в том случае, если они проходят через все его порты с максимальной интенсивностью. Другими словами, получим условие того, что при данном наборе портов коммутатор является неблокирующим.

Очевидно, что коммутатор будет неблокирующим, если общая внутренняя производительность коммутатора B равна сумме максимальных пропускных способностей протоколов всех его портов Сk:

B  Сk

То есть, если у коммутатора имеется, например, 12 портов Ethernet и 2 порта Fast Ethernet, то внутренней производительности в 320 Мб/с будет достаточно для обработки любого распределения трафика, попавшего в коммутатор через его порты. Однако, такая внутренняя производительность является избыточной, так как коммутатор предназначен не только для приема кадров, но и для их передачи на порт назначения. Поэтому все порты коммутатора не могут постоянно с максимальной скоростью только принимать информацию извне - средняя интенсивность уходящей через все порты коммутатора информации должна быть равна средней интенсивности принимаемой информации. Следовательно, максимальная скорость передаваемой через коммутатор информации в стабильном режиме равна половине суммарной пропускной способности всех портов - каждый входной кадр является для какого-либо порта выходным кадром. В соответствии с этим утверждением для нормальной работы коммутатора достаточно, чтобы его внутренняя общая производительность была равна половине суммы максимальных пропускных способностей протоколов всех его портов:

B  Сk)/2

Поэтому, для коммутатора с 12 портами Ethernet и 2 портами Fast Ethernet вполне достаточно иметь среднюю общую производительность в 160 Мб/с, для нормальной работы по передаче любых вариантов распределения трафика, которые могут быть переданы его портами в течение достаточно длительного периода времени.

Еще раз нужно подчеркнуть, что это условие гарантирует только то, что внутренние элементы коммутатора - процессоры портов, межмодульная шина, центральный процессор и т.п. - справятся с обработкой поступающего трафика. Несимметрия в распределении этого трафика по выходным портам всегда может привести к невозможности своевременной передачи трафика в сеть из-за ограничений протокола порта. Для предотвращения потерь кадров многие производители коммутаторов применяют фирменные решения, позволяющие "притормаживать" передатчики узлов, подключенных к коммутатору, то есть вводят элементы управления потоком не модифицируя протоколы портов конечных узлов. Эти способы будут рассмотрены ниже при рассмотрении дополнительных возможностей коммутаторов.

Кроме пропускных способностей отдельных элементов коммутатора, таких как процессоры портов или общая шина, на производительность коммутатора влияют такие его параметры как размер адресной таблицы и объем общего буфера или отдельных буферов портов.

#### Размер адресной таблицы

*Максимальная емкость адресной таблицы* определяет максимальное количество MAC-адресов, с которыми может одновременно оперировать коммутатор. Так как коммутаторы чаще всего используют для выполнения операций каждого порта выделенный процессорный блок со своей памятью для хранения экземпляра адресной таблицы, то размер адресной таблицы для коммутаторов обычно приводится в расчете на один порт. Экземпляры адресной таблицы разных процессорных модулей не обязательно содержат одну и ту же адресную информацию - скорее всего повторяющихся адресов будет не так много, если только распределение трафика каждого порта не полностью равновероятное между остальными портами. Каждый порт хранит только те наборы адресов, которыми он пользуется в последнее время.

Значение максимального числа МАС-адресов, которое может запомнить процессор порта, зависит от области применения коммутатора. Коммутаторы рабочих групп обычно поддерживают всего несколько адресов на порт, так как они предназначены для образования микросегментов. Коммутаторы отделов должны поддерживать несколько сотен адресов, а коммутаторы магистралей сетей - до нескольких тысяч, обычно 4К - 8К адресов.

Недостаточная емкость адресной таблицы может служить причиной замедления работы коммутатора и засорения сети избыточным трафиком. Если адресная таблица процессора порта полностью заполнена, а он встречает новый адрес источника в поступившем пакете, то он должен вытеснить из таблицы какой-либо старый адрес и поместить на его место новый. Эта операция сама по себе отнимет у процессора часть времени, но главные потери производительности будут наблюдаться при поступлении кадра с адресом назначения, который пришлось удалить из адресной таблицы. Так как адрес назначения кадра неизвестен, то коммутатор должен передать этот кадр на все остальные порты. Эта операция будет создавать лишнюю работу для многих процессоров портов, кроме того, копии этого кадра будут попадать и на те сегменты сети, где они совсем необязательны.

Некоторые производители коммутаторов решают эту проблему за счет изменения алгоритма обработки кадров с неизвестным адресом назначения. Один из портов коммутатора конфигурируется как магистральный порт, на который по умолчанию передаются все кадры с неизвестным адресом. В маршрутизаторах такой прием применяется давно, позволяя сократить размеры адресных таблиц в сетях, организованных по иерархическому принципу.

Передача кадра на магистральный порт производится в расчете на то, что этот порт подключен к вышестоящему коммутатору, который имеет достаточную емкость адресной таблицы и знает, куда нужно передать любой кадр. Пример успешной передачи кадра при использовании магистрального порта приведен на рисунке 4.10. Коммутатор верхнего уровня имеет информацию о всех узлах сети, поэтому кадр с адресом назначения МАС3, переданный ему через магистральный порт, он передает через порт 2 коммутатору, к которому подключен узел с адресом МАС3.

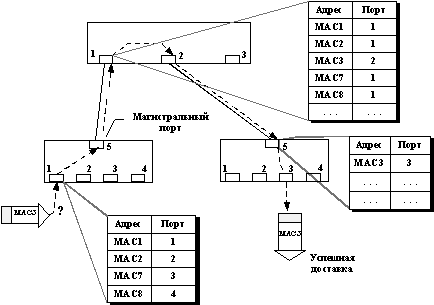


Рис. 4.10. Использование магистрального порта для доставки кадров   
с неизвестным адресом назначения

Хотя метод магистрального порта и будет работать эффективно во многих случаях, но можно представить такие ситуации, когда кадры будут просто теряться. Одна из таких ситуаций изображена на рисунке 4.11. Коммутатор нижнего уровня удалил из своей адресной таблицы адрес МАС8, который подключен к его порту 4, для того, чтобы освободить место для нового адреса МАС3. При поступлении кадра с адресом назначения МАС8, коммутатор передает его на магистральный порт 5, через который кадр попадает в коммутатор верхнего уровня. Этот коммутатор видит по своей адресной таблице, что адрес МАС8 принадлежит его порту 1, через который он и поступил в коммутатор. Поэтому кадр далее не обрабатывается и просто отфильтровывается, а, следовательно, не доходит до адресата. Поэтому более надежным является использование коммутаторов с достаточным количеством адресной таблицы для каждого порта, а также с поддержкой общей адресной таблицы модулем управления коммутатором.

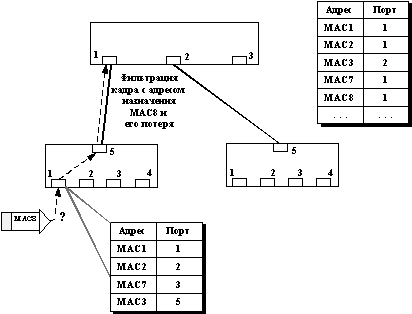


Рис. 4.11. Потеря кадра при использовании магистрального порта

#### Объем буфера

Внутренняя буферная память коммутатора нужна для временного хранения кадров данных в тех случаях, когда их невозможно немедленно передать на выходной порт. Буфер предназначен для сглаживания кратковременных пульсаций трафика. Ведь даже если трафик хорошо сбалансирован и производительность процессоров портов, а также других обрабатывающих элементов коммутатора достаточна для передачи средних значений трафика, то это не гарантирует, что их производительности хватит при очень больших пиковых значениях нагрузок. Например, трафик может в течение нескольких десятков миллисекунд поступать одновременно на все входы коммутатора, не давая ему возможности передавать принимаемые кадры на выходные порты.

Для предотвращения потерь кадров при кратковременном многократном превышении среднего значения интенсивности трафика (а для локальных сетей часто встречаются значения коэффициента пульсации трафика в диапазоне 50 - 100) единственным средством служит буфер большого объема. Как и в случае адресных таблиц, каждый процессорный модуль порта обычно имеет свою буферную память для хранения кадров. Чем больше объем этой памяти, тем менее вероятны потери кадров при перегрузках, хотя при несбалансированности средних значений трафика буфер все равно рано или поздно переполниться.

Обычно коммутаторы, предназначенные для работы в ответственных частях сети, имеют буферную память в несколько десятков или сотен килобайт на порт. Хорошо, когда эту буферную память можно перераспределять между несколькими портами, так как одновременные перегрузки по нескольким портам маловероятны. Дополнительным средством защиты может служить общий для всех портов буфер в модуле управления коммутатором. Такой буфер обычно имеет объем в несколько мегабайт.

### Дополнительные возможности коммутаторов

Так как коммутатор представляет собой сложное вычислительное устройство, имеющее несколько процессорных модулей, то естественно нагрузить его помимо выполнения основной функции передачи кадров с порта на порт по алгоритму моста и некоторыми дополнительными функциями, полезными при построении надежных и гибких сетей. Ниже описываются наиболее распространенные дополнительные функции коммутаторов, которые поддерживаются большинством производителей коммуникационного оборудования.

#### Трансляция протоколов канального уровня

Коммутаторы могут выполнять трансляцию одного протокола канального уровня в другой, например, Ethernet в FDDI, Fast Ethernet в Token Ring и т.п. При этом они работают по тем же алгоритмам, что и транслирующие мосты, то есть в соответствии со спецификациями RFC 1042 и 802.1H, определяющими правила преобразования полей кадров разных протоколов.

Трансляцию протоколов локальных сетей облегчает тот факт, что наиболее сложную работу, которую часто выполняют маршрутизаторы и шлюзы при объединении гетерогенных сетей, а именно работу по трансляции адресной информации, в данном случае выполнять не нужно. Все конечные узлы локальных сетей имеют уникальные адреса одного и того же формата, независимо от поддерживаемого протокола. Поэтому адрес сетевого адаптера Ethernet понятен сетевому адаптеру FDDI, и они могут использовать эти адреса в полях своих кадров не задумываясь о том, что узел, с которым они взаимодействуют, принадлежит сети, работающей по другой технологии.

Поэтому при согласовании протоколов локальных сетей коммутаторы не строят таблиц соответствия адресов узлов, а переносят адреса назначения и источника из кадра одного протокола в кадр другого протокола. Единственным преобразованием, которое, возможно, придется при этом выполнить, является преобразование порядка бит в байте, если согласуется сеть Ethernet с сетью Token Ring или FDDI. Это связано с тем, что в сетях Ethernet принята так называемая каноническая форма передачи адреса по сети, когда сначала передается самый младший бит самого старшего байта адреса. В сетях FDDI и Token Ring всегда передается сначала самый старший бит самого старшего байта адреса. Так как технология 100VG-AnyLAN использует кадры или Ethernet или Token Ring, то ее трансляция в другие технологии зависит от того, кадры каких протоколов используются в данном сегменте сети 100VG-AnyLAN.

Кроме изменения порядка бит при передаче байт адреса, трансляция протокола Ethernet (и Fast Ethernet, который использует формат кадров Ethernet) в протоколы FDDI и Token Ring включает выполнение следующих (возможно не всех) операций:

* Вычисление длины поля данных кадра и помещение этого значения в поле Length при передаче кадра из сети FDDI или Token Ring в сеть Ethernet 802.3 (в кадрах FDDI и Token Ring поле длины отсутствует).
* Заполнение полей статуса кадра при передаче кадров из сети FDDI или Token Ring в сеть Ethernet. Кадры FDDI и Token Ring имеют два бита, которые должны быть установлены станцией, которой предназначался кадр - бит распознавания адреса A и бит копирования кадра С. При получении кадра станция должна установить эти два бита для того, чтобы кадр, вернувшийся по кольцу к станции, его сгенерировавшей, принес данные обратной связи. При передаче коммутатором кадра в другую сеть нет стандартных правил для установки бит А и С в кадре, который возвращается по кольцу к станции-источнику. Поэтому производители коммутаторов решают эту проблему по своему усмотрению.
* Отбрасывание кадров, передаваемых из сетей FDDI или Token Ring в сеть Ethernet с размером поля данных большим, чем 1500 байт, так как это максимально возможное значение поля данных для сетей Ethernet. В дальнейшем возможно усечение максимального размера поля данных сетей FDDI или Token Ring средствами протоколов верхнего уровня, например, TCP. Другим вариантом решения этой проблемы является поддержка коммутатором IP фрагментации, но это требует, во-первых, реализации в коммутаторе протокола сетевого уровня, а во-вторых, поддержки протокола IP взаимодействующими узлами транслируемых сетей.
* Заполнение поля Type (тип протокола в поле данных) кадра Ethernet II при приходе кадров из сетей, поддерживающих кадры FDDI или Token Ring, в которых это поле отсутствует. Для сохранения информации поля Type в стандарте RFC 1042 предлагается использовать поле Type заголовка кадра LLC/SNAP, вкладываемого в поле данных МАС-кадра протоколов FDDI или Token Ring. При обратном преобразовании значение из поля Type заголовка LLC/SNAP переносится в поле Type кадра Ethernet II.
* Пересчет контрольной суммы кадра в соответствии со сформированными значениями служебных полей кадра.

#### Поддержка алгоритма Spanning Tree

Алгоритм *Spanning Tree (STA)* позволяет коммутаторам автоматически определять древовидную конфигурацию связей в сети при произвольном соединения портов между собой. Как уже отмечалось, для нормальной работы коммутатора требуется отсутствие замкнутых маршрутов в сети. Эти маршруты могут создаваться администратором специально для образования резервных связей или же возникать случайным образом, что вполне возможно, если сеть имеет многочисленные связи, а кабельная система плохо структурирована или документирована.

Поддерживающие алгоритм STA коммутаторы автоматически создают активную древовидную конфигурацию связей (то есть связную конфигурацию без петель) на множестве всех связей сети. Такая конфигурация называется покрывающим деревом - Spanning Tree (иногда ее называют остовным или основным деревом), и ее название дало имя всему алгоритму. Коммутаторы находят покрывающее дерево адаптивно с помощью обмена служебными пакетами. Реализация в коммутаторе алгоритма STA очень важна для работы в больших сетях - если коммутатор не поддерживает этот алгоритм, то администратор должен самостоятельно определить, какие порты нужно перевести в заблокированное состояние, чтобы исключить петли. К тому же при отказе какой-либо связи, порта или коммутатора администратор должен, во-первых, обнаружить факт отказа, а, во-вторых, ликвидировать последствия отказа, переведя резервную связь в рабочий режим путем активизации некоторых портов.

**Основные определения**

В сети определяется *корневой коммутатор (root switch),* от которого строится дерево. Корневой коммутатор может быть выбран автоматически или назначен администратором. При автоматическом выборе корневым становится коммутатор с меньшим значением МАС-адреса его блока управления.

Для каждого коммутатора определяется *корневой порт (root port)* - это порт, который имеет по сети кратчайшее расстояние до корневого коммутатора (точнее, до любого из портов корневого коммутатора). Затем для каждого сегмента сети выбирается так называемый *назначенный порт* *(designated port)* - это порт, который имеет кратчайшее расстояние от данного сегмента до корневого коммутатора.

Понятие расстояния играет важную роль в построении покрывающего дерева. Именно по этому критерию выбирается единственный порт, соединяющий каждый коммутатор с корневым коммутатором, и единственный порт, соединяющий каждый сегмент сети с корневым коммутатором. Все остальные порты переводятся в резервное состояние, то есть такое, при котором они не передают обычные кадры данных. Можно доказать, что при таком выборе активных портов в сети исключаются петли и оставшиеся связи образуют покрывающее дерево.

На рисунке 4.12 показан пример построения конфигурации покрывающего дерева для сети, состоящей из 6 сегментов (N1 - N6) и 6 коммутаторов (S1 - S6). Корневые порты закрашены черным цветом, назначенные не закрашены, а заблокированные порты перечеркнуты. В активной конфигурации коммутаторы S2 и S6 не имеют портов, передающих кадры данных, поэтому они закрашены как резервные.

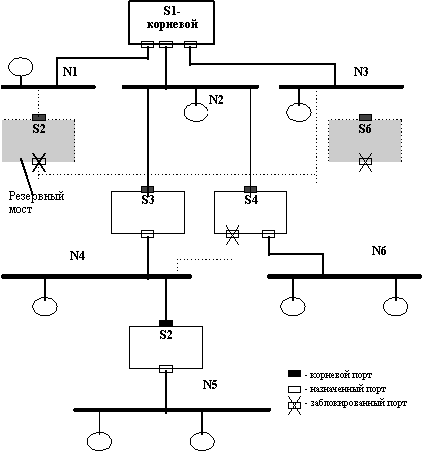


Рис. 4.12. Построение покрывающего дерева сети по алгоритму STA

*Расстояние до корня* определяется как суммарное условное время на передачу данных от порта данного коммутатора до порта корневого коммутатора. При этом считается, что время внутренних передач данных (с порта на порт) коммутатором пренебрежимо мало, а учитывается только время на передачу данных по сегментам сети, соединяющим коммутаторы. *Условное время сегмента* рассчитывается как время, затрачиваемое на передачу одного бита информации в 10-наносекундных единицах между непосредственно связанными по сегменту сети портами. Так, для сегмента Ethernet это время равно 10 условным единицам, а для сегмента Token Ring 16 Мб/с - 6.25. (Алгоритм STA не связан с каким-либо определенным стандартом канального уровня, он может применяться к коммутаторам, соединяющим сети различных технологий.)

В приведенном примере предполагается, что все сегменты имеют одинаковое условное расстояние, поэтому оно не показано на рисунке.

Для автоматического определения начальной активной конфигурации дерева все коммутаторы сети после их инициализации начинают периодически обмениваться специальными пакетами, называемыми протокольными блоками данных моста - *BPDU (Bridge Protocol Data Unit),* что отражает факт первоначальной разработки алгоритма STA для мостов.

Пакеты BPDU помещаются в поле данных кадров канального уровня, например, кадров Ethernet или FDDI. Желательно, чтобы все коммутаторы поддерживали общий групповой адрес, с помощью которого кадры, содержащие пакеты BPDU, могли одновременно передаваться всем коммутаторам сети. Иначе пакеты BPDU рассылаются широковещательно.

Пакет BPDU имеет следующие поля:

* Идентификатор версии протокола STA - 2 байта. Коммутаторы должны поддерживать одну и ту же версию протокола STA, иначе может установиться активная конфигурация с петлями.
* Тип BPDU - 1 байт. Существует два типа BPDU - конфигурационный BPDU, то есть заявка на возможность стать корневым коммутатором, на основании которой происходит определение активной конфигурации, и BPDU уведомления о реконфигурации, которое посылается коммутатором, обнаружившим событие, требующее проведения реконфигурации - отказ линии связи, отказ порта, изменение приоритетов коммутатора или портов.
* Флаги - 1 байт. Один бит содержит флаг изменения конфигурации, второй бит - флаг подтверждения изменения конфигурации.
* Идентификатор корневого коммутатора - 8 байтов.
* Расстояние до корня - 2 байта.
* Идентификатор коммутатора - 8 байтов.
* Идентификатор порта - 2 байта.
* Время жизни сообщения - 2 байта. Измеряется в единицах по 0.5 с, служит для выявления устаревших сообщений. Когда пакет BPDU проходит через коммутатор, тот добавляет ко времени жизни пакета время его задержки данным коммутатором.
* Максимальное время жизни сообщения - 2 байта. Если пакет BPDU имеет время жизни, превышающее максимальное, то он игнорируется коммутаторами.
* Интервал hello, через который посылаются пакеты BPDU.
* Задержка смены состояний - 2 байта. Минимальное время перехода портов коммутатора в активное состояние. Такая задержка необходима, чтобы исключить возможность временного возникновения альтернативных маршрутов при неодновременной смене состояний портов во время реконфигурации.

У пакета BPDU уведомления о реконфигурации отсутствуют все поля, кроме двух первых.

После инициализации каждый коммутатор сначала считает себя корневым. Поэтому он начинает через интервал hello генерировать через все свои порты сообщения BPDU конфигурационного типа. В них он указывает свой идентификатор в качестве идентификатора корневого коммутатора (и в качестве данного коммутатора также), расстояние до корня устанавливается в 0, а в качестве идентификатора порта указывается идентификатор того порта, через который передается BPDU. Как только коммутатор получает BPDU, в котором имеется идентификатор корневого коммутатора, меньше его собственного, он перестает генерировать свои собственные кадры BPDU, а начинает ретранслировать только кадры нового претендента на звание корневого коммутатора. При ретрансляции кадров он наращивает расстояние до корня, указанное в пришедшем BPDU, на условное время сегмента, по которому принят данный кадр.

При ретрансляции кадров каждый коммутатор для каждого своего порта запоминает минимальное расстояние до корня, встретившееся во всех принятых этим портом кадрах BPDU. При завершении процедуры установления конфигурации покрывающего дерева (по времени) каждый коммутатор находит свой корневой порт - это порт, который ближе других портов находится по отношению к корню дерева. Кроме этого, коммутаторы распределенным образом выбирают для каждого сегмента сети назначенный порт. Для этого они исключают из рассмотрения свой корневой порт, а для всех своих оставшихся портов сравнивают принятые по ним минимальные расстояния до корня с расстоянием до корня своего корневого порта. Если у своего порта это расстояние меньше принятых, то это значит, что он является назначенным портом. Все порты, кроме назначенных переводятся в заблокированное состояние и на этом построение покрывающего дерева заканчивается.

В процессе нормальной работы корневой коммутатор продолжает генерировать служебные кадры, а остальные коммутаторы продолжают их принимать своими корневыми портами и ретранслировать назначенными. Если у коммутатора нет назначенных портов, то он все равно принимает служебные кадры корневым портом. Если по истечении тайм-аута корневой порт не получает служебный кадр, то он инициализирует новую процедуру построения покрывающего дерева.

#### Способы управления потоком кадров

Некоторые производители применяют в своих коммутаторах приемы управления потоком кадров, отсутствующие в стандартах протоколов локальных сетей, для предотвращения потерь кадров при перегрузках.

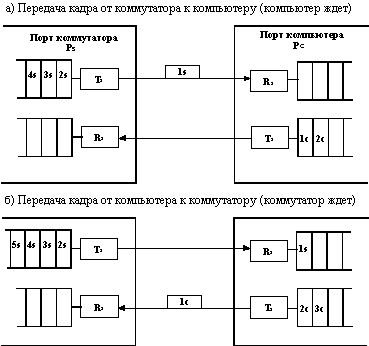


Рис. 4.13. Чередование передач кадров при обмене данными через коммутатор

На рисунке 4.13 приведен пример обмена кадрами между коммутатором и портом сетевого адаптера компьютера в режиме пиковой загрузки коммутатора. Коммутатор не успевает передавать кадры из буфера передатчика Tx, так как при нормальном полудуплексном режиме работы передатчик должен часть времени простаивать, ожидая, пока приемник не примет очередной кадр от компьютера.

Так как потери, даже небольшой доли кадров, обычно намного снижают полезную производительность сети, то при перегрузке коммутатора рационально было бы замедлить интенсивность поступления кадров от конечных узлов в приемники коммутатора, чтобы дать возможность передатчикам разгрузить свои буфера с более высокой скоростью. Алгоритм чередования передаваемых и принимаемых кадров (frame interleave) должен быть гибким и позволять компьютеру в критических ситуациях на каждый принимаемый кадр передавать несколько своих, причем не обязательно снижая при этом интенсивность приема до нуля, а просто уменьшая ее до необходимого уровня.

Для реализации такого алгоритма в распоряжении коммутатора должен быть механизм снижения интенсивности трафика подключенных к его портам узлов. У некоторых протоколов локальных сетей, таких как FDDI, Token Ring или 100VG-AnyLAN имеется возможность изменять приоритет порта и тем самым давать порту коммутатора преимущество перед портом компьютера. У протоколов Ethernet и Fast Ethernet такой возможности нет, поэтому производители коммутаторов для этих очень популярных технологий используют два приема воздействия на конечные узлы.

Эти приемы основаны на том, что конечные узлы строго соблюдают все параметры алгоритма доступа к среде, а порты коммутатора - нет.

Первый способ "торможения" конечного узла основан на так называемом агрессивном поведении порта коммутатора при захвате среды после окончания передачи очередного пакета или после коллизии. Эти два случая иллюстрируются рисунком 4.14.

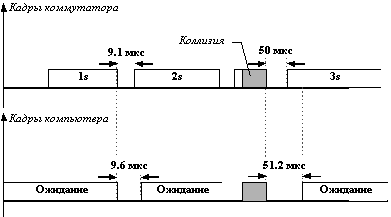


Рис. 4.14. Агрессивное поведение коммутатора при перегрузках буферов

В первом случае коммутатор окончил передачу очередного кадра и вместо технологической паузы в 9.6 мкс сделал паузу в 9.1 мкс и начал передачу нового кадра. Компьютер не смог захватить среду, так как он выдержал стандартную паузу в 9.6 мкс и обнаружил после этого, что среда уже занята.

Во втором случае кадры коммутатора и компьютера столкнулись и была зафиксирована коллизия. Так как компьютер сделал паузу после коллизии в 51.2 мкс, как это положено по стандарту (интервал отсрочки равен 512 битовых интервалов), а коммутатор - 50 мкс, то и в этом случае компьютеру не удалось передать свой кадр.

Коммутатор может пользоваться этим механизмом адаптивно, увеличивая степень своей агрессивности по мере необходимости.

Второй прием, которым пользуются разработчики коммутаторов - это передача фиктивных кадров компьютеру в том случае, когда у коммутатора нет в буфере кадров для передачи по данному порту. В этом случае коммутатор может и не нарушать параметры алгоритма доступа, честно соревнуясь с конечным узлом за право передать свой кадр. Так как среда при этом равновероятно будет доставаться в распоряжение то коммутатору, то конечному узлу, то интенсивность передачи кадров в коммутатор в среднем уменьшится вдвое. Такой метод называется методом *обратного давления (backpressure)*. Он может комбинироваться с методом агрессивного захвата среды для большего подавления активности конечного узла.

Метод обратного давления используется не для того, чтобы разгрузить буфер процессора порта, непосредственно связанного с подавляемым узлом, а разгрузить либо общий буфер коммутатора (если используется архитектура с разделяемой общей памятью), либо разгрузить буфер процессора другого порта, в который передает свои кадры данный порт. Кроме того, метод обратного давления может применяться в тех случаях, когда процессор порта не рассчитан на поддержку максимально возможного для протокола трафика. Один из первых примеров применения метода обратного давления как раз связан с таким случаем - метод был применен компанией LANNET в модулях LSE-1 и LSE-2, рассчитанных на коммутацию трафика Ethernet с максимальной интенсивностью соответственно 1 Мб/с и 2 Мб/с.

#### Возможности коммутаторов по фильтрации трафика

Многие коммутаторы позволяют администраторам задавать дополнительные условия фильтрации кадров наряду со стандартными условиями их фильтрации в соответствии с информацией адресной таблицы. Пользовательские фильтры предназначены для создания дополнительных барьеров на пути кадров, которые ограничивают доступ определенных групп пользователей к определенным сервисам сети.

Если коммутатор не поддерживает протоколы сетевого и транспортного уровней, в которых имеются поля, указывающие к какому сервису относятся передаваемые пакеты, то администратору приходится для задания условий интеллектуальной фильтрации определять поле, по значению которого нужно осуществлять фильтрацию, в виде пары "смещение-размер" относительно начала поля данных кадра канального уровня. Поэтому, например, для того, чтобы запретить некоторому пользователю печатать свои документы на определенном принт-сервере NetWare, администратору нужно знать положение поля "номер сокета" в пакете IPX и значение этого поля для принт-сервиса, а также знать МАС-адреса компьютера пользователя и принт-сервера.

Обычно условия фильтрации записываются в виде булевских выражений, формируемых с помощью логических операций AND и OR.

Наложение дополнительных условий фильтрации может снизить производительность коммутатора, так как вычисление булевских выражений требует проведения дополнительных вычислений процессорами портов.

Кроме условий общего вида коммутаторы могут поддерживать специальные условия фильтрации. Одним из очень популярных видов специальных фильтров являются фильтры, создающие виртуальные сегменты. Они рассматриваются в разделе 4.3.7 отдельно в виду их особого значения.

Специальным является и фильтр, используемый многими производителями для защиты сети, построенной на основе коммутаторов.

#### Коммутация "на лету" или с буферизацией

На возможности реализации дополнительных функций существенно сказывается способ передачи пакетов - "на лету" или с буферизацией. Как показывает следующая таблица, большая часть дополнительных функций коммутатора требует полной буферизации кадров перед их выдачей через порт назначения в сеть.

|  |  |
| --- | --- |
| **Функция** | **На летуС буферизацией** |
| Защита от плохих кадров | НетДа |
| Поддержка разнородных сетей (Ethernet, Token Ring, FDDI, ATM) | НетДа |
| Задержка передачи пакетов | Низкая (10 - 40 мкс) при низкой нагрузке, средняя при высокой нагрузкеСредняя при любой нагрузке |
| Поддержка резервных связей | НетДа |
| Функция анализа трафика | НетДа |

Средняя величина задержки коммутаторов работающих "на лету" при высокой нагрузке объясняется тем, что в этом случае выходной порт часто бывает занят приемом другого пакета, поэтому вновь поступивший пакет для данного порта все равно приходится буферизовать.

Коммутатор, работающий "на лету", может выполнять проверку некорректности передаваемых кадров, но не может изъять плохой кадр из сети, так как часть его байт (и, как правило, большая часть) уже переданы в сеть. В то же время при небольшой загрузке коммутатор, работающий "на лету", существенно уменьшает задержку передачи кадра, а это может быть важным для чувствительного к задержкам трафика. Поэтому некоторые производители, например Cisco, применяют механизм адаптивной смены режима работы коммутатора. Основной режим такого коммутатора - коммутация "на лету", но коммутатор постоянно контролирует трафик и при превышении интенсивности появления плохих кадров некоторого порога переходит на режим полной буферизации.

#### Использование различных классов сервиса (class-of-service)

Эта функция позволяет администратору назначить различным типам кадров различные приоритеты их обработки. При этом коммутатор поддерживает несколько очередей необработанных кадров и может быть сконфигурирован, например, так, что он передает один низкоприоритетный пакет на каждые 10 высокоприоритетных пакетов. Это свойство может особенно пригодиться на низкоскоростных линиях и при наличии приложений, предъявляющих различные требования к допустимым задержкам.

Так как не все протоколы канального уровня поддерживают поле приоритета кадра, например, у кадров Ethernet оно отсутствует, то коммутатор должен использовать какой-либо дополнительный механизм для связывания кадра с его приоритетом. Наиболее распространенный способ - приписывание приоритета портам коммутатора. При этом способе коммутатор помещает кадр в очередь кадров соответствующего приоритета в зависимости от того, через какой порт поступил кадр в коммутатор. Способ несложный, но недостаточно гибкий - если к порту коммутатора подключен не отдельный узел, а сегмент, то все узлы сегмента получают одинаковый приоритет. Примером подхода к назначению классов обслуживания на основе портов является технология PACE компании 3Com.

Более гибким является назначение приоритетов МАС-адресам узлов, но этот способ требует выполнения большого объема ручной работы администратором.

#### Поддержка виртуальных сетей

Кроме своего основного назначения - повышения пропускной способности связей в сети - коммутатор позволяет локализовывать потоки информации в сети, а также контролировать эти потоки и управлять ими, используя пользовательские фильтры. Однако, пользовательский фильтр может запретить передачи кадров только по конкретным адресам, а широковещательный трафик он передает всем сегментам сети. Так требует алгоритм работы моста, который реализован в коммутаторе, поэтому сети, созданные на основе мостов и коммутаторов иногда называют плоскими - из-за отсутствия барьеров на пути широковещательного трафика.

Технология виртуальных сетей *(Virtual LAN, VLAN)* позволяет преодолеть указанное ограничение. Виртуальной сетью называется группа узлов сети, трафик которой, в том числе и широковещательный, на канальном уровне полностью изолирован от других узлов сети. Это означает, что передача кадров между разными виртуальными сегментами на основании адреса канального уровня невозможна, независимо от типа адреса - уникального, группового или широковещательного. В то же время внутри виртуальной сети кадры передаются по технологии коммутации, то есть только на тот порт, который связан с адресом назначения кадра.

Говорят, что виртуальная сеть образует домен широковещательного трафика *(broadcast domain),* по аналогии с доменом коллизий, который образуется повторителями сетей Ethernet.

Назначение технологии виртуальных сетей состоит в облегчении процесса создания независимых сетей, которые затем должны связываться с помощью протоколов сетевого уровня. Для решения этой задачи до появления технологии виртуальных сетей использовались отдельные повторители, каждый из которых образовывал независимую сеть. Затем эти сети связывались маршрутизаторами в единую интерсеть (рисунок 4.15).

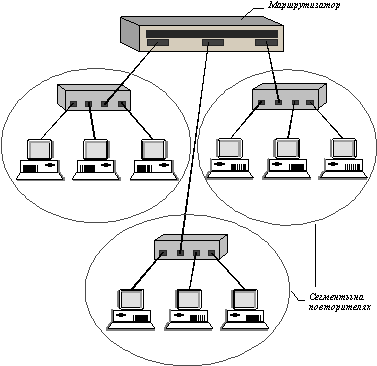


Рис. 4.15. Интерсеть, состоящая из сетей, построенных на основе повторителей

При изменении состава сегментов (переход пользователя в другую сеть, дробление крупных сегментов) при таком подходе приходится производить физическую перекоммутацию разъемов на передних панелях повторителей или в кроссовых панелях, что не очень удобно в больших сетях - много физической работы, к тому же высока вероятность ошибки.

Поэтому для устранения необходимости физической перекоммутации узлов стали применять многосегментные повторители (рисунок 4.16). В наиболее совершенных моделях таких повторителей приписывание отдельного порта к любому из внутренних сегментов производится программным путем, обычно с помощью удобного графического интерфейса. Примерами таких повторителей могут служить концентратор Distributed 5000 компании Bay Networks и концентратор PortSwitch компании 3Com. Программное приписывание порта сегменту часто называют статической или конфигурационной коммутацией.

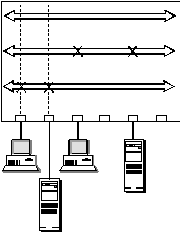


Рис. 4.16. Многосегментный повторитель с конфигурационной коммутацией

Однако, решение задачи изменения состава сегментов с помощью повторителей накладывает некоторые ограничения на структуру сети - количество сегментов такого повторителя обычно невелико, поэтому выделить каждому узлу свой сегмент, как это можно сделать с помощью коммутатора, нереально. Поэтому сети, построенные на основе повторителей с конфигурационной коммутацией, по-прежнему основаны на разделении среды передачи данных между большим количеством узлов, и, следовательно, обладают гораздо меньшей производительностью по сравнению с сетями, построенными на основе коммутаторов.

При использовании технологии виртуальных сетей в коммутаторах одновременно решаются две задачи:

* повышение производительности в каждой из виртуальных сетей, так как коммутатор передает кадры в такой сети только узлу назначения;
* изоляция сетей друг от друга для управления правами доступа пользователей и создания защитных барьеров на пути широковещательных штормов.

Для связи виртуальных сетей в интерсеть требуется привлечение сетевого уровня. Он может быть реализован в отдельном маршрутизаторе, а может работать и в составе программного обеспечения коммутатора.

Технология образования и работы виртуальных сетей с помощью коммутаторов пока не стандартизована, хотя и реализуется в очень широком спектре моделей коммутаторов разных производителей. Положение может скоро измениться, если будет принят стандарт 802.1Q, разрабатываемый в рамках института IEEE.

В виду отсутствия стандарта каждый производитель имеет свою технологию виртуальных сетей, которая, как правило, несовместима с технологией других производителей. Поэтому виртуальные сети можно создавать пока на оборудовании одного производителя. Исключение составляют только виртуальные сети, построенные на основе спецификации *LANE (LAN Emulation),* предназначенной для обеспечения взаимодействия АТМ-коммутаторов с традиционным оборудованием локальных сетей.

При создании виртуальных сетей на основе одного коммутатора обычно используется механизм группирования в сети портов коммутатора (рисунок 4.17).

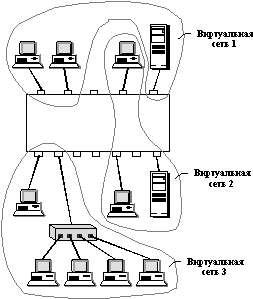


Рис. 4.17. Виртуальные сети, построенные на одном коммутаторе

Это логично, так как виртуальных сетей, построенных на основе одного коммутатора, не может быть больше, чем портов. Если к одному порту подключен сегмент, построенный на основе повторителя, то узлы такого сегмента не имеет смысла включать в разные виртуальные сети - все равно трафик этих узлов будет общим.

Создание виртуальных сетей на основе группирования портов не требует от администратора большого объема ручной работы - достаточно каждый порт приписать к нескольким заранее поименованным виртуальным сетям. Обычно такая операция выполняется путем перетаскивания мышью графических символов портов на графические символы сетей.

Второй способ, который используется для образования виртуальных сетей основан на группировании МАС-адресов. При существовании в сети большого количества узлов этот способ требует выполнения большого количества ручных операций от администратора. Однако, он оказывается более гибким при построении виртуальных сетей на основе нескольких коммутаторов, чем способ группирования портов.

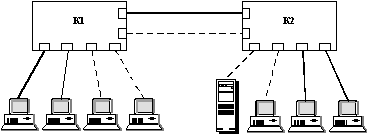


Рис. 4.18. Построение виртуальных сетей на нескольких   
коммутаторах с группировкой портов

Рисунок 4.18 иллюстрирует проблему, возникающую при создании виртуальных сетей на основе нескольких коммутаторов, поддерживающих технику группирования портов. Если узлы какой-либо виртуальной сети подключены к разным коммутаторам, то для соединения коммутаторов каждой такой сети должна быть выделена своя пара портов. В противном случае, если коммутаторы будут связаны только одной парой портов, информация о принадлежности кадра той или иной виртуальной сети при передаче из коммутатора в коммутатор будет утеряна. Таким образом, коммутаторы с группировкой портов требуют для своего соединения столько портов, сколько виртуальных сетей они поддерживают. Порты и кабели используются при таком способе очень расточительно. Кроме того, при соединении виртуальных сетей через маршрутизатор для каждой виртуальной сети выделяется в этом случае отдельный кабель, что затрудняет вертикальную разводку, особенно если узлы виртуальной сети присутствуют на нескольких этажах (рисунок 4.19).

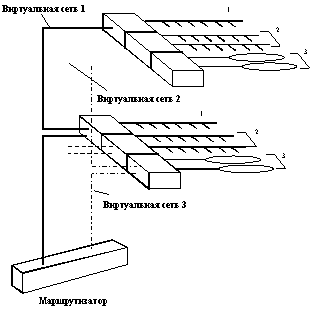


Рис. 4.19. Соединение виртуальных сетей, построенных на группировании портов,   
через маршрутизатор

Группирование МАС-адресов в сеть на каждом коммутаторе избавляет от необходимости их связи несколькими портами, однако требует выполнения большого количества ручных операций по маркировке МАС-адресов на каждом коммутаторе сети.

Описанные два подхода основаны только на добавлении дополнительной информации к адресным таблицам моста и не используют возможности встраивания информации о принадлежности кадра к виртуальной сети в передаваемый кадр. Остальные подходы используют имеющиеся или дополнительные поля кадра для сохранения информации и принадлежности кадра при его перемещениях между коммутаторами сети. При этом нет необходимости запоминать в каждом коммутаторе принадлежность всех МАС-адресов интерсети виртуальным сетям.

Если используется дополнительное поле с пометкой о номере виртуальной сети, то оно используется только тогда, когда кадр передается от коммутатора к коммутатору, а при передаче кадра конечному узлу оно удаляется. При этом модифицируется протокол взаимодействия "коммутатор-коммутатор", а программное и аппаратное обеспечение конечных узлов остается неизменным. Примеров таких фирменных протоколов много, но общий недостаток у них один - они не поддерживаются другими производителями. Компания Cisco предложила использовать в качестве стандартной добавки к кадрам любых протоколов локальных сетей заголовок протокола 802.10, предназначенного для поддержки функций безопасности вычислительных сетей. Сама компания использует этот метод в тех случаях, когда коммутаторы объединяются между собой по протоколу FDDI. Однако, эта инициатива не была поддержана другими ведущими производителями коммутаторов, поэтому до принятия стандарта 802.1Q фирменные протоколы маркировки виртуальных сетей будут преобладать.

Существует два способа построения виртуальных сетей, которые используют уже имеющиеся поля для маркировки принадлежности кадра виртуальной сети, однако эти поля принадлежат не кадрам канальных протоколов, а пакетам сетевого уровня или ячейкам технологии АТМ.

В первом случае виртуальные сети образуются на основе сетевых адресов, то есть той же информации, которая используется при построении интерсетей традиционным способом - с помощью физически отдельных сетей, подключаемых к разным портам маршрутизатора.

Когда виртуальная сеть образуется на основе номеров сетей, то каждому порту коммутатора присваивается один или несколько номеров сетей, например, номеров IP-сетей. (рисунок 4.20). Каждый номер IP-сети соответствует одной виртуальной сети. Конечные узлы также должны в этом случае поддерживать протокол IP. При передаче кадров между узлами, принадлежащими одной виртуальной сети, конечные узлы посылают данные непосредственно по МАС-адресу узла назначения, а в пакете сетевого уровня указывают IP-адрес своей виртуальной сети. Коммутатор в этом случае передает кадры на основе МАС-адреса назначения по адресной таблице, проверяя при этом допустимость передач по совпадению IP-номера сети пакета, содержащегося в кадре, и IP-адресу порта назначения, найденному по адресной таблице. При передачах кадра из одного коммутатора в другой, его IP-адрес переносится вместе с кадром, а значит коммутаторы могут быть связаны только одной парой портов для поддержки виртуальных сетей, распределенных между несколькими коммутаторами.

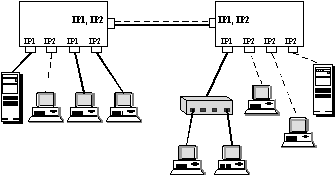


Рис. 4.20. Построение виртуальных сетей на основе сетевых адресов

В случае, когда нужно произвести обмен информацией между узлами, принадлежащими разным виртуальным сетям, конечный узел работает так же, как если бы он находился в сетях, разделенных обычным маршрутизатором. Конечный узел направляет кадр маршрутизатору по умолчанию, указывая его МАС-адрес в кадре, а IP-адрес узла назначения - в пакете сетевого уровня. Маршрутизатором по умолчанию должен быть внутренний блок коммутатора, который имеет определенный МАС-адрес и IP-адрес, как и традиционный маршрутизатор. Кроме того, он должен иметь таблицу маршрутизации, в которой указывается выходной порт для всех номеров сетей, которые существуют в общей интерсети.

В отличие от традиционных маршрутизаторов, у которых каждый порт имеет свой номер сети, коммутаторы, поддерживающие сетевой протокол для образования виртуальных сетей, назначают один и тот же номер сети нескольким портам. Кроме того, один и тот же порт может быть связан с несколькими номерами сетей, если через него связываются коммутаторы.

Часто коммутаторы не поддерживают функции автоматического построения таблиц маршрутизации, которые поддерживаются протоколами маршрутизации. такими как RIP или OSPF. Такие коммутаторы называют коммутаторами 3-го уровня, чтобы подчеркнуть их отличие от традиционных маршрутизаторов. При использовании коммутаторов 3-го уровня таблицы маршрутизации либо создаются администратором вручную (это тоже часто приемлемо при небольшом количестве виртуальных сетей и маршруте по умолчанию к полноценному маршрутизатору), либо загружаются из маршрутизатора. По последней схеме взаимодействует коммутатор Catalist 5000 компании Cisco с маршрутизаторами этой же компании.

Если же коммутатор не поддерживает функций сетевого уровня, то его виртуальные сети могут быть объединены только с помощью внешнего маршрутизатора. Некоторые компании выпускают специальные маршрутизаторы для применения совместно с коммутаторами. Примером такого маршрутизатора служит маршрутизатор Vgate компании RND, изображенный на рисунке 4.21.

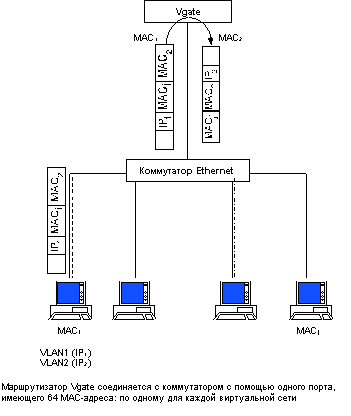


Рис. 4.21. Маршрутизатор Vgate, разработанный специально для объединения виртуальных сетей

Этот маршрутизатор имеет один физический порт для связи с портом коммутатора, но этот порт может поддерживать до 64 МАС-адресов, что позволяет маршрутизатору объединять до 64 виртуальных сетей.

Последний способ организации виртуальных сетей связан с применением в сети АТМ-коммутаторов. Этот способ основан на использовании для передачи кадров каждой виртуальной сети через коммутаторы АТМ с помощью отдельного виртуального соединения. На рисунке 4.22 показан пример сети, в которой две виртуальные сети объединены с помощью АТМ-сети, состоящей из трех коммутаторов. Так как для передачи кадров каждой виртуальной сети используется отдельный виртуальный канал со своим номером SVC, то коммутатор К2, собирая переданный кадр из ячеек АТМ, знает о принадлежности кадра к той или иной виртуальной сети, а далее на основе его МАС-адреса принимает решение о передаче его на определенный порт.

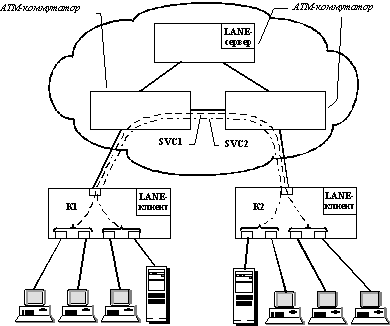


Рис. 4.22. Использование отдельных виртуальных каналов в ATM-сетях   
для передачи информации о виртуальных сетях

Коммутаторы К1 и К2, изображенные на рисунке, должны иметь АТМ-порты и поддерживать для реализации взаимодействия локальных сетей с сетью АТМ спецификацию LANE. Эту спецификацию должен поддерживать также хотя бы один из АТМ-коммутаторов. Так как спецификация LANE достаточно подробно описывает способ поддержки виртуальных сетей сетью АТМ-коммутаторов и пограничных коммутаторов, имеющих клиентскую часть протокола LANE, то оборудование разных производителей может работать в одной сети, образуя виртуальные сети с помощью АТМ-технологии. Спецификация LANE описывает способ взаимодействия локальных сетей и сетей АТМ на основе МАС-адресов и АТМ-адресов, не привлекая протоколы сетевого уровня. Поэтому она может быть реализована в коммутаторах, работающих только на канальном уровне. Для объединения виртуальных сетей, построенных с помощью спецификации LANE, нужны маршрутизаторы с АТМ-портами.

## Управление коммутируемыми сетями

Коммутаторы - это сложные многофункциональные устройства, играющие ответственную роль в современных сетях. Поэтому поддержка функций централизованного контроля и управления, реализуемого протоколом SNMP и соответствующими агентами, практически обязательна для всех классов коммутаторов (кроме, может быть, настольных коммутаторов, предназначенных для работы в очень маленьких сетях).

Для поддержки SNMP-управления коммутаторы имеют модуль управления, в котором имеется агент, ведущий базу данных управляющей информации. Этот модуль часто выполняется на отдельном мощном процессоре, чтобы не замедлять основные операции коммутатора.

**Наблюдение за трафиком**

Так как перегрузки процессоров портов и других обрабатывающих элементов коммутатора могут приводить к потерям кадров, то функция наблюдения за распределением трафика в сети, построенной на основе коммутаторов, очень важна.

Однако, если сам коммутатор не имеет отдельного агента для каждого своего порта, то задача слежения за трафиком, традиционно решаемая в сетях с разделяемыми средами с помощью установки в сеть внешнего анализатора протоколов, очень усложняется.

Обычно в традиционных сетях анализатор протоколов (например, Sniffer компании Network General) подключался к свободному порту концентратора и видел весь трафик, передаваемый между любыми узлами сети.

Если же анализатор протокола подключить к свободному порту коммутатора, то он не увидит почти ничего, так как ему кадры передавать никто не будет, а чужие кадры в его порт также направляться не будут. Единственный вид трафика, который будет видеть анализатор - это трафик широковещательных пакетов, которые будут передаваться всем узлам сети. В случае, когда сеть разделена на виртуальные сети, анализатор протоколов будет видеть только широковещательный трафик своей виртуальной сети.

Для того, чтобы анализаторами протоколов можно было по-прежнему пользоваться и в коммутируемых сетях, производители коммутаторов снабжают свои устройства функцией зеркального отображения трафика любого порта на специальный порт. К специальному порту подключается анализатор протоколов, а затем на коммутатор подается команда через его модуль SNMP-управления для отображения трафика какого-либо порта на специальный порт.

Наличие функции зеркализации портов частично снимает проблему, но оставляет некоторые вопросы. Например, как просмотреть одновременно трафик двух портов, или как просматривать трафик порта, работающего в полнодуплексном режиме.

Более надежным способом слежения за трафиком, проходящим через порты коммутатора, является замена анализатора протокола на агенты RMON MIB для каждого порта коммутатора.

Агент RMON выполняет все функции хорошего анализатора протокола для протоколов Ethernet и Token Ring, собирая детальную информацию об интенсивности трафика, различных типах плохих кадров, о потерянных кадрах, причем самостоятельно строя временные ряды для каждого фиксируемого параметра. Кроме того, агент RMON может самостоятельно строить матрицы перекрестного трафика между узлами сети, которые очень нужны для анализа эффективности применения коммутатора.

Так как агент RMON, реализующий все 9 групп объектов Ethernet, стоит весьма дорого, то производители для снижения стоимости коммутатора часто реализуют только первые несколько групп объектов RMON MIB.

**Управление виртуальными сетями**

Виртуальные сети порождают проблемы для традиционных систем управления на SNMP-платформе как при их создании, так и при наблюдении за их работой.

Как правило, для создания виртуальных сетей требуется специальное программное обеспечение компании-производителя, которое работает на платформе системы управления, такой как, например, HP Open View. Сами платформы систем управления этот процесс поддержать не могут, в основном из-за отсутствия стандарта на виртуальные сети. Можно надеяться, что появление стандарта 802.1Q изменит ситуацию в этой области.

Наблюдение за работой виртуальных сетей также создает проблемы для традиционных систем управления. При создании карты сети, включающей виртуальные сети, необходимо отображать как физическую структуру сети, так и ее логическую структуру, соответствующую связям отдельных узлов виртуальной сети. При этом по желанию администратора система управления должна уметь отображать соответствие логических и физических связей в сети, то есть на одном физическом канале должны отображаться все или отдельные пути виртуальных сетей.

К сожалению, многие системы управления либо вообще не отображают виртуальные сети, либо делают это очень неудобным для пользователя способом.

## Типовые схемы применения коммутаторов в локальных сетях

### Коммутатор или концентратор?

При построении небольших сетей, составляющих нижний уровень иерархии корпоративной сети, вопрос о применении того или иного коммуникационного устройства сводится к вопросу о выборе между концентратором или коммутатором.

При ответе на этот вопрос нужно принимать во внимание несколько факторов. Безусловно, немаловажное значение имеет стоимость за порт, которую нужно заплатить при выборе устройства. Из технических соображений в первую очередь нужно принять во внимание существующее распределение трафика между узлами сети. Кроме того, нужно учитывать перспективы развития сети: будут ли в скором времени применяться мультимедийные приложения, будет ли модернизироваться компьютерная база. Если да, то нужно уже сегодня обеспечить резервы по пропускной способности применяемого коммуникационного оборудования. Использование технологии intranet также ведет к увеличению объемов трафика, циркулирующего в сети, и это также необходимо учитывать при выборе устройства.

При выборе типа устройства - концентратор или коммутатор - нужно еще определить и тип протокола, который будут поддерживать его порты (или протоколов, если идет речь о коммутаторе, так как каждый порт может поддерживать отдельный протокол).

Сегодня выбор делается между протоколами двух скоростей - 10 Мб/с и 100 Мб/с. Поэтому, сравнивая применимость концентратора или коммутатора, необходимо рассмотреть вариант концентратора с портами на 10 Мб/с, вариант концентратора с портами на 100 Мб/c, и несколько вариантов коммутаторов с различными комбинациями скоростей на его портах.

Техника применения матрицы перекрестного трафика для анализа эффективности применения коммутатора уже была рассмотрена в разделе 4.2.2. Пользуясь ею, можно оценить, сможет ли коммутатор с известными пропускными способностями портов и общей производительностью поддержать трафик в сети, заданный в виде матрицы средних интенсивностей трафика.

Рассмотрим теперь эту технику для ответа на вопрос о применимости коммутатора в сети с одним сервером и несколькими рабочими станциями, взаимодействующими только с сервером (рисунок 6.1). Такая конфигурация сети часто встречается в сетях масштаба рабочей группы, особенно в сетях NetWare, где стандартные клиентские оболочки не могут взаимодействовать друг с другом.

Матрица перекрестного трафика для такой сети имеет вырожденный вид. Если сервер подключен, например, к порту 4, то только 4-я строка матрицы и 4-ый столбец матрицы будут иметь отличные от нуля значения. Эти значения соответствуют выходящему и входящему трафику порта, к которому подключен сервер. Поэтому условия применимости коммутатора для данной сети сводятся к возможности передачи всего трафика сети портом коммутатора, к которому подключен сервер.

Если коммутатор имеет все порты с одинаковой пропускной способностью, например, 10 Мб/c, то в этом случае пропускная способность порта в 10 Мб/c будет распределяться между всеми компьютерами сети. Возможности коммутатора по повышению общей пропускной способности сети оказываются для такой конфигурации невостребованными. Несмотря на микросегментацию сети, ее пропускная способность ограничивается пропускной способностью протокола одного порта, как и в случае применения концентратора с портами 10 Мб/с. Небольшой выигрыш при использовании коммутатора будет достигаться лишь за счет уменьшения количества коллизий - вместо коллизий кадры будут просто попадать в очередь к передатчику порта коммутатора, к которому подключен сервер.

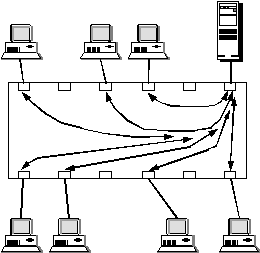


Рис. 6.1. Сеть с выделенным сервером

Для того, чтобы коммутатор работал в сетях с выделенным сервером более эффективно, производители коммутаторов выпускают модели с одним высокоскоростным портом на 100 Мб/с для подключения сервера и несколькими низкоскоростными портами на 10 Мб/с для подключения рабочих станций. В этом случае между рабочими станциями распределяется уже 100 Мб/c, что позволяет обслуживать в неблокирующем режиме 10 - 30 станций, в зависимости от интенсивности создаваемого ими трафика.

Однако с таким коммутатором может конкурировать концентратор, поддерживающий протокол с пропускной способностью 100 Мб/с, например, Fast Ethernet. Стоимость его за порт будет несколько ниже стоимости за порт коммутатора с одним высокоскоростным портом, а производительность сети примерно та же.

Очевидно, что выбор коммуникационного устройства для сети с выделенным сервером достаточно сложен. Для принятия окончательного решения нужно принимать во внимание перспективы развития сети в отношении движения к сбалансированному трафику. Если в сети вскоре может появиться взаимодействие между рабочими станциями, или же второй сервер, то выбор необходимо делать в пользу коммутатора, который сможет поддержать дополнительный трафик без ущерба по отношению к основному.

В пользу коммутатора может сыграть и фактор расстояний - применение коммутаторов не ограничивает максимальный диаметр сети величинами в 2500 м или 210 м, которые определяют размеры домена коллизий при использовании концентраторов Ethernet и Fast Ethernet.

### Коммутатор или маршрутизатор?

При построении верхних, магистральных уровней иерархии корпоративной сети проблема выбора формулируется по-другому - коммутатор или маршрутизатор?

Коммутатор выполняет передачу трафика между узлами сети быстрее и дешевле, зато маршрутизатор более интеллектуально отфильтровывает трафик при соединении сетей, не пропуская ненужные или плохие пакеты, а также надежно защищая сети от широковещательных штормов.

В связи с тем, что коммутаторы корпоративного уровня могут поддерживать некоторые функции сетевого уровня, выбор все чаще делается в пользу коммутатора. При этом маршрутизатор также используется, но он часто остается в локальной сети в единственном экземпляре. Этот маршрутизатор обычно служит и для связи локальной сети с глобальными, и для объединения виртуальных сетей, построенных с помощью коммутаторов.

В центре же сетей зданий и этажей все чаще используются коммутаторы, так как только при их использовании возможно осуществить передачу нескольких гигабит информации в секунду за приемлемую цену (рисунок 1.2).

### Стянутая в точку магистраль на коммутаторе

При всем разнообразии структурных схем сетей, построенных на коммутаторах, все они используют две базовые структуры - стянутую в точку магистраль и распределенную магистраль. На основе этих базовых структур затем строятся разнообразные структуры конкретных сетей.

*Стянутая в точку магистраль (collapsed backbone)* - это структура, при которой объединение узлов, сегментов или сетей происходит на внутренней магистрали коммутатора. Пример сети рабочей группы, использующей такую структуру, приведен на рисунке 6.2.

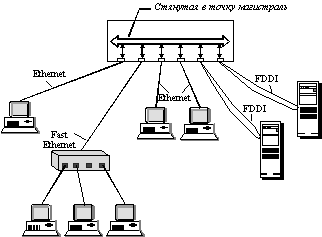


Рис. 6.2. Структура сети со стянутой в точку магистралью

Преимуществом такой структуры является высокая производительность магистрали. Так как для коммутатора производительность внутренней шины или схемы общей памяти, объединяющей модули портов, в несколько Гб/c не является редкостью, то магистраль сети может быть весьма быстродействующей, причем ее скорость не зависит от применяемых в сети протоколов и может быть повышена с помощью замены одной модели коммутатора на другую.

Положительной чертой такой схемы является не только высокая скорость магистрали, но и ее протокольная независимость. На внутренней магистрали коммутатора в независимом формате одновременно могут передаваться данные различных протоколов, например, Ethernet, FDDI и Fast Ethernet, как это изображено на рисунке. Подключение нового узла с новым протоколом часто требует не замены коммутатора, а просто добавления соответствующего интерфейсного модуля, поддерживающего этот протокол.

Если к каждому порту коммутатора в такой схеме подключен только один узел, то такая схема будет соответствовать микросегментированной сети.

### Распределенная магистраль на коммутаторах

В сетях больших зданий или кампусов использование структуры с коллапсированной магистралью не всегда рационально или же возможно. Такая структура приводит к протяженным кабельным системам, которые связывают конечные узлы или коммутаторы сетей рабочих групп с центральным коммутатором, шина которого и является магистралью сети. Высокая плотность кабелей и их высокая стоимость ограничивают применение стянутой в точку магистрали в таких сетях. Иногда, особенно в сетях кампусов, просто невозможно стянуть все кабели в одно помещение из-за ограничений на длину связей, накладываемых технологией (например, все реализации технологий локальных сетей на витой паре ограничивают протяженность кабелей в 100 м).

Поэтому в локальных сетях, покрывающих большие территории, часто используется другой вариант построения сети - с распределенной магистралью. Пример такой сети приведен на рисунке 6.3.

*Распределенная магистраль* - это разделяемый сегмент сети, поддерживающий определенный протокол, к которому присоединяются коммутаторы сетей рабочих групп и отделов. На примере распределенная магистраль построена на основе двойного кольца FDDI, к которому подключены коммутаторы этажей. Коммутаторы этажей имеют большое количество портов Ethernet, трафик которых транслируется в трафик протокола FDDI, когда он передается по магистрали с этажа на этаж.

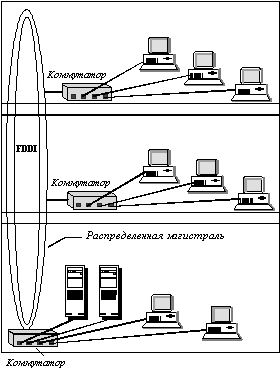


Рис. 6.3. Структура сети с распределенной магистралью

Распределенная магистраль упрощает связи между этажами, сокращает стоимость кабельной системы и преодолевает ограничения на расстояния.

Однако, скорость магистрали в этом случае будет существенно меньше скорости магистрали на внутренней шине коммутатора. Причем скорость эта фиксированная и не превышает в настоящее время 100 Мб/c. Поэтому распределенная магистраль может применяться только при невысокой интенсивности трафика между этажами или зданиями.

Пример рисунка 6.3 демонстрирует сочетание использования двух базовых структур, так как на каждом этаже сеть построена с использованием магистрали на внутренней шине коммутатора.

## Обзор моделей коммутаторов

Рынок коммутаторов сегодня очень обширен, поэтому в этом кратком обзоре остановимся только на некоторых популярных моделях коммутаторов различного класса. Обычно коммутаторы делят в первую очередь на классы в соответствии с их областями применения - настольные коммутаторы, коммутаторы рабочих групп, коммутаторы отделов и магистральные (корпоративные коммутаторы). У каждого класса коммутаторов есть свои отличительные признаки.

Настольные коммутаторы

* Фиксированное количество портов;
* Все порты работают на одной скорости;
* Используются для организации одноранговых связей высокоскоростных рабочих станций;
* Режим коммутации - "на лету";
* Чаще всего не содержат модуля SNMP-управления, а также не поддерживают алгоритм Spanning Tree.

Пример: 3Com LinkSwitch 500.

Коммутаторы рабочих групп

* Имеют по крайней мере 1 высокоскоростной порт (FDDI, Fast Ethernet, ATM);
* Транслируют протоколы;
* Как правило, управляемы по SNMP, поддерживают алгоритм Spanning Tree;
* Режим коммутации - с буферизацией.

Примеры: семейство 3Com LinkSwitch (кроме модели 500), SMC TigerSwitch XE, Bay Networks Ethernet Workgroup Switch.

Коммутаторы отделов и центров обработки данных

* Модульное исполнение;
* Поддержка нескольких протоколов;
* Встроенные средства обеспечения отказоустойчивости:
* избыточные источники питания;
* модули hot-swap.
* Пользовательские фильтры;
* Поддержка виртуальных сегментов;

Примеры: 3Com LANplex 2500, SMC ES/1, Bay Networks LattisSwitch System 28115.

Коммутаторы магистралей зданий/кампусов

* Те же свойства, что и у коммутаторов отделов;
* Шасси с большим количеством слотов (10 - 14);
* Внутренняя пропускная способность 1 - 10 Гб/с;
* Поддержка 1 - 2 протоколов маршрутизации (локальные интерфейсы) для образования виртуальных сетей.

Примеры: 3Com LANplex 6000, Cabletron MMAC Plus, LANNET LET-36, Cisco Catalist 5000, Bay Networks System 5000.

### Коммутаторы Catalyst компании Cisco Systems

*Коммутатор Catalyst 5000* представляет собой старшую модель семейства Catalyst. Это модульная, многоуровневая платформа коммутации, которая обеспечивает высокий уровень производительности, предоставляя возможность как для создания выделенных соединений в сети Ethernet со скоростями 10 и 100 Mб/с, так и для организации взаимодействия с сетями FDDI и АТМ.

Шасси Catalyst 5000 имеет 5 разъемов. В один разъем устанавливается модуль управления Supervisor Engine, который управляет доступом к коммутируемой матрице, имеющей возможность коммутации более 1 млн. пакетов в секунду. Модуль поддерживает функции локального и удаленного управления и имеет два порта Fast Ethernet, которые могут использоваться для соединения серверов сети или каскадирования устройств Catalyst 5000. Остальные разъемы могут использоваться для установки следующих модулей:

* 24 порта 10Base-T;
* 12 портов 10Base-FL;
* 12 портов 100Base -TX;
* 12 портов 100Base-FX;
* 1 порт DAS CDDI/FDDI (не более 3-х модулей в шасси);
* 1 порт 155 Мб/с АТМ (не более 3-х модулей в шасси).

Одно устройство Catalyst 5000 может поддерживать до 96 коммутируемых портов Ethernet и до 50 коммутируемых портов Fast Ethernet.

Поддерживается формирование виртуальных сетей как в пределах одного устройства Catalyst 5000, так и для нескольких устройств на основе группирования портов. Можно создать до 1000 виртуальных сетей для нескольких устройств Catalyst 5000, соединенных интерфейсами Fast Ethernet, CDDI/FDDI или ATM. Любой интерфейс Fast Ethernet может быть сконфигурирован как интерфейс *InterSwitch Link (ISL)* для поддержки нескольких виртуальных сетей. Интерфейс ISL - частное решение компании Cisco для передачи информации между коммутаторами о виртуальных сетях.

Все виртуальные сети поддерживают протокол IEEE 802.1d Spanning Tree для обеспечения отказоустойчивых соединений. При использовании интерфейса АТМ для соединения коммутаторов поддержка виртуальных сетей осуществляется на основе спецификации LANE через виртуальные соединения. Интерфейс FDDI поддерживает виртуальные сети с помощью спецификации 802.10.

Отличительной особенностью коммутаторов Catalyst является выполнение коммутации на 3 уровне модели OSI, что позволяет объединять виртуальные сети внутри устройства (для этого требуется дополнительное программное обеспечение).

Модуль управления коммутацией поддерживает три уровня очередей кадров с различными приоритетами, причем приоритеты назначаются для каждого порта отдельно. Это позволяет эффективно обслуживать мультимедийный трафик.

Большой буфер (по 192 Кбайта на порт) обеспечивает сохранение и передачу информации при пиковых нагрузках.

***Система Catalyst 3000*** представляет собой оригинальную реализацию стековой архитектуры для коммутаторов. Эта архитектура поддерживается устройствами двух типов:

* Коммутатор Catalyst 3000 c 16 портами 10Base-T, одним портом AUI и двумя слотами расширения. Модули расширения могут иметь либо 1 порт 100Base-TX, либо 3 порта 10Base-FL, либо 4 порта 10Base-T, либо 1 порт ATM. Порт мониторинга осуществляет зекрализацию любого порта данных на внешний порт.
* Catalyst Matrix - 8-ми портовая матрица коммутация, с помощью которой можно объединить в стек до 8 коммутаторов Catalyst 3000 для создания единого коммутирующего центра.

Коммутаторы Catalyst 3000 подключаются к Catalyst Matrix через специальные 280 Мб/с порты. Производительность шины Catalyst Matryx составляет 3.84 Гб/с.

Коммутатор работает под управлением IOS и использует два алгоритма коммутации - cut-throw и store-and-forward.

Стек Catalyst 3000 поддерживает до 64 виртуальных сетей и позволяет фильтровать трафик по адресу источника и адресу назначения. Максимальное число MAC-адресов - до 10К на устройство.

Поддерживается алгоритм Spanning Tree и SNMP-управление.

### Коммутатор EliteSwitch ES/1 компании SMC

Корпорация SMC (сейчас ее подразделение коммутаторов является частью компании Cabletron) разработала коммутатор EliteSwitch ES/1 как эффективный инструмент для создания внутренней магистрали сети средних размеров. Коммутатор ES/1 сочетает в себе функции высокопроизводительного коммутатора технологий Ethernet/Token Ring/FDDI и локального маршрутизатора, позволяющего создавать виртуальные сети IP и IPX на основе виртуальных коммутируемых рабочих групп. Таким образом, в одном устройстве объединены функции switching и internetworking, необходимые для построения на базе внутренней скоростной шины структурированной локальной сети. Коммутатор поддерживает и глобальные связи с топологией "точка-точка" по линиям T1/E1, позволяя связывать несколько локальных сетей, построенных на его основе, друг с другом.

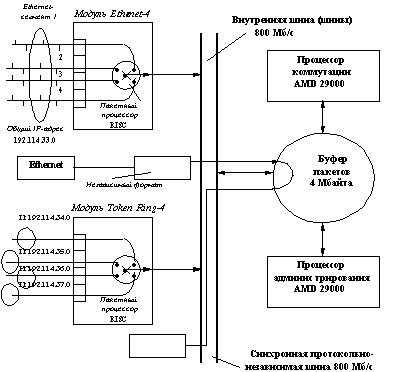


Рис. 7.1. Структура коммутатора ES/1

Коммутатор ES/1 работает по технологии коммутации с буферизацией, что позволяет ему транслировать протоколы канального уровня, осуществлять пользовательскую фильтрацию, сбор статистики и локальную маршрутизацию.

**Организация коммутатора ES/1**

Модульный концентратор ES/1 компании SMC (рисунок 7.1) представляет собой устройство в виде корпуса-шасси с задней коммуникационной платой, на которой выполнена внутренняя шина с производительностью 800 Мб/с. Блок обработки пакетов (Packet Processing Engine) включает в себя два процессорных модуля, оснащенных высокопроизводительными RISC-процессорами AMD 29000. Один из процессоров предназначен для передачи пакетов (то есть выполняет функции коммутации), а другой осуществляет администрирование - фильтрацию на портах концентратора в соответствии с масками, введенными администратором, и управляет всей логикой работы концентратора. Оба процессора имеют доступ к общей памяти объемом 4 МБ.

Как уже отмечалось, модуль обработки пакетов коммутатора ES/1 построен на сдвоенной процессорной архитектуре, причем каждый из процессоров отвечает за свои функции. Однако в случае отказа одного из них второй процессор возьмет на себя все функции первого. При этом коммутатор в целом продолжит нормальную работу, может только несколько снизиться его производительность.

Адресная таблица концентратора позволяет сохранять до 8192 МАС-адресов.

Программное обеспечение, управляющее работой концентратора ES/1, дублируется в двух банках Flash-памяти. Во-первых, это позволяет производить upgrade новых версий программного обеспечения без прекращения выполнения концентратором своих основных функций по коммутации пакетов, а во-вторых, сбой при загрузке нового ПО из банка Flash-памяти не приведет к отказу концентратора, поскольку ПО из первого банка памяти останется в рабочем состоянии, и концентратор автоматически перезагрузит его.

В слоты концентратора вставляются сетевые коммуникационные модули, при этом реализована технология автоматической самоконфигурации plug-and-play. Каждый модуль оснащен собственным RISC-процессором, который преобразует приходящие пакеты в протокольно-независимый вид (это означает, что сохраняются только блок данных, адреса приемника и источника, а также информация о сетевом протоколе) и передает их далее по внутренней шине в блок обработки пакетов.

Отказоустойчивость работы модулей обеспечивается наличием в каждом из них специального сенсора, посылающего предупреждение на консоль оператора при приближении температуры к критической отметке. Это может произойти, например, по причине запыления воздушных фильтров. Если температура продолжает повышаться и превышает второе пороговое значение, модуль автоматически отключается от питания для предотвращения выхода из строя элементной базы. При снижении температуры модуль автоматически продолжит работу.

Важной особенностью концентратора ES/1 является встроенная система защиты от "штормов" широковещательных пакетов (broadcast storm). Программное обеспечение концентратора ES/1 позволяет установить предельную частоту прихода таких пакетов на каждый порт концентратора, в случае превышения которой широковещательные пакеты перестают передаваться в другие сегменты сети, что сохраняет их работоспособность.

**Фильтрация и виртуальные рабочие группы**

С помощью механизма маскирования портов администратор может создавать виртуальные рабочие группы с целью защиты от несанкционированного доступа и повышения производительности ЛВС путем перераспределения информационных потоков. Фильтрацию можно включать на входящие и/или выходящие пакеты, по MAC-адресу или по всему сегменту и так далее. Всего маска может содержать до 20-ти условий, объединенных булевыми операндами "AND" и "OR". Понятно, что каждый пакет, приходящий на порт коммутатора, должен быть дополнительно проверен на соответствие условиям фильтрации, что требует дополнительных вычислительных ресурсов и может привести к снижению производительности. То, что в ES/1 один из двух процессоров выделен для проверки условий фильтрации, обеспечивает сохранение высокой производительности коммутатора при введенных администратором масках.

Наряду с отказами оборудования, ошибки обслуживающего персонала могут нарушить корректную работу ЛВС. Поэтому особо отметим еще один интересный режим виртуальной фильтрации коммутатора ES/1. В этом режиме фильтрация физически не включается, однако ведется набор статистики пакетов, удовлетворяющих условиям фильтрации. Это дает возможность администратору ЛВС заранее прогнозировать свои действия перед физическим включением фильтров.

**Коммуникационные модули концентратора ES/1**

ES/1 поддерживает до пяти модулей. Можно выбрать любую комбинацию модулей для Ethernet, Token Ring и FDDI, а также для высокоскоростных линий T1/E1 и T3/E3. Все модули, включая источники питания, могут заменяться без отключения от сети и выключения питания центрального устройства. Каждый модуль поддерживает набор конфигурируемых параметров для улучшения управляемости и собирает статистику.

*QEIOM (Quad Ethernet I/O Module)*

К этому модулю можно подключить до четырех независимых сегментов Ethernet. Каждый сегмент может передавать и получать информацию с обычной для Ethernet производительностью 14880 пакетов в секунду. ES/1 обеспечивает связь между этими четырьмя сегментами по типу мостов и маршрутизаторов, а также и со всей остальной сетью. Эти модули поставляются с различными типами разъемов: AUI, BNC, RJ-45 (витая пара) и ST (оптоволоконный кабель).

*QTIOM (Token Ring I/O Module)*

Через модуль QTIOM подключается до четырех 4 или 16 Mб/c сетей Token Ring. Модуль поддерживает все основные протоколы сети Token Ring - IBM Source Routing, Transparent Bridging и Source Routing Transparent - и обеспечивает "прозрачное" взаимодействие сетей Token Ring с сетями остальных типов, например Ethernet или FDDI. Модуль поставляется в вариантах для экранированной и неэкранированной витой пары.

*IFIOM (Intelligent Dual-Attached FDDI I/O Module)*

Модуль IFIOM подключает волоконно-оптический сегмент сети FDDI к ES/1 и обеспечивает прозрачное взаимодействие между разными типами сетей. Он поддерживает все функции FDDI-станции с двойным подключением к кольцу (Dual Attached Station). Этот модуль также поддерживает внешний оптический переключатель (Optical Bypass Switch), что обеспечивает повышенную отказоустойчивость сети при аварийном отключении ES/1. Поставляется в различных модификациях: для одномодового и многомодового волокна и в их комбинациях.

*CEIOM24 (24-Port Concentrator Ethernet I/O Module)*

Этот модуль включает в себя 24-портовый концентратор Ethernet на витой паре. Он увеличивает производительность сети при стоимости, меньшей, чем стоимость аналогичного внешнего устройства. Его порты сгруппированы в единый независимый сегмент Ethernet и взаимодействуют с другими модулями через коммутатор/маршрутизатор ES/1.

*HIOM (High-Speed Serial Interface I/O Module)*

HIOM позволяет осуществить подключение сетей к удаленным ЛВС по высокоскоростным линиям связи по протоколу HSSI со скоростью до 52 Мб/с. Поддерживается протокол PPP.

**SNMP-управляемость**

Модульный концентратор ES/1 может управляться с помощью любой стандартной системы управления, базирующейся на SNMP-протоколе, в том числе: HP OpenView, IBM NetView/6000, Sun NetManager и др. Для графического представления передней панели концентратора к перечисленным консолям управления добавляются специальные программные модули компании SMC семейства EliteView. Кроме того, имеется версия программного обеспечения мониторинга и управления, работающая под Windows: EliteView for Windows.

**Типовые схемы использования концентратора ES/1**

*Создание вырожденной магистрали (Collapsed Backbone)*

Вырожденная магистраль внутри коммутатора применяется в крупных корпоративных сетях. Несколько крупных сегментов локальной сети подключаются к портам концентратора, шина которого в этом случае выполняет роль основной магистрали с пропускной способностью в сотни Мб/с. Такой подход позволяет увеличить пропускную способность сети в несколько раз по сравнению с традиционным использованием мостов на каждом сегменте сети. При этом существенно повышаются возможности централизованного управления всеми элементами корпоративной сети.

*Выделенный канал Ethernet (Dedicated Ethernet)*

Эта схема подключения устройств к портам коммутируемых концентраторов применяется чаще всего для создания высокоскоростной магистрали (с гарантированной пропускной способностью 10 Мб/с) между концентратором и сервером локальной сети (обычно файловым сервером или сервером баз данных). Модульные концентраторы позволяют организовать при необходимости подключение сервера по высокоскоростному каналу FDDI или Fast Ethernet.

*Транслирующая коммутация*

Коммутация в ES/1 основана на синхронной протокольно-независимой технологии *(Synchronous Protocol Independent technology),* которая поддерживает основные технологии локальных сетей, позволяя осуществлять трансляцию между кадрами различных форматов. Поэтому коммутатор ES/1 может использоваться для соединения сетей различных типов - Ethernet, Token Ring, FDDI, причем трансляция происходит со скоростью коммутации и не создает перегрузок трафика при межсетевых передачах.

*Образование виртуальных групп*

По умолчанию коммутатор работает в режиме моста, изучая трафик, проходящий через его порты и строя таблицу адресов сегментов. С помощью программного обеспечения EliteView администратор может в удобной графической форме определить состав виртуальных рабочих групп, куда будут входить либо локальные сегменты, если к порту ES/1 подключен концентратор или сегмент Ethernet на коаксиальном кабеле, либо отдельные рабочие станции, если они подключены к порту индивидуально выделенным каналом. Виртуальные рабочие группы могут включать различные порты как одного, так и нескольких коммутаторов ES/1.

*Виртуальные сети*

Наряду с образованием виртуальных изолированных рабочих групп, защищающих данные и локализующих трафик, очень полезным свойством коммутатора является возможность объединения этих групп в интерсеть с помощью внутренней маршрутизации пакетов между виртуальными сегментами, которые объявляются виртуальными сетями (IP или IPX). При этом передача пакетов между портами, принадлежащими одной сети, происходит быстро на основании коммутации пакетов, в то же время пакеты, предназначенные другой сети, маршрутизируются. Таким образом, обеспечивается взаимодействие между виртуальными рабочими группами, и в то же время выполняются все функции по защите сетей друг от друга, обеспечиваемые маршрутизаторами.

### Коммутаторы локальных сетей компании 3Com

Компания 3Com занимает прочные позиции на рынке коммутаторов, выпуская широкий спектр этих устройств для всех областей применения.

Сектор коммутаторов для настольных применений и рабочих групп представляют коммутаторы семейства Link Switch. Коммутаторы для сетей отделов и магистральные коммутаторы представлены семейством LANplex. Для сетей АТМ компания выпускает коммутаторы семейства CELLplex.

Технология коммутация неэффективна без опоры на специализированные БИС -ASIC, которые оптимизированы для быстрого выполнения специальных операций. Компания 3Сom строит свои коммутаторы на нескольких ASIC, разработанных для коммутации определенных протоколов.

* ASIC ISE (Intelligent Switching Engine) предназначена для выполнения операций коммутации Ethernet и FDDI, а также поддержки функций маршрутизации и управления. Используется в коммутаторах LANplex 2500, LANplex 6000 и LinkSwitch 2200.
* ASIC TRSE (Token Ring Switching Engine) выполняет коммутацию сетей Token Ring. Используется в коммутаторах LinkSwitch 2000 TR и LANplex 6000.
* ASIC BRASICA выполняет коммутацию Ethernet/Fast Ethernet. Поддерживает технологию виртуальных сетей и спецификацию RMON. Используется в коммутаторах LinkSwitch 1000 и LinkSwitch 3000.
* ASIC ZipChip поддерживает коммутацию ATM, а также преобразование кадров Ethernet в ячейки АТМ. используется в коммутаторах CELLplex 7000 и LinkSwitch 2700.

В следующей таблице приведены основные параметры коммутаторов семейства LinkSwitch.

|  |  |  |  |
| --- | --- | --- | --- |
| **Модель** | **LS 500** | **LS 1000** | **LS 1200LS 2200** **LS 2700** **LS 3000** |
| Высокоско- ростной порт | Нет | 100Base-T | FDDIFDDI ATM 100Base-T |
| Общая производительность | 50 Мб/с | 220 Мб/с | 60 Мб/с130 Мб/с 120 Мб/с 310 Мб/с |
| Цена | $5100 | $5850 | $8220$11750 $10800 $12800 |
| Цена за порт | $510 | $244 | $1370$735 $900 $2560 |
| Кол. портов | 10 | 24+1(2) | 6 + 116 + 1 12 + 1 5 |
| Количество MAC-адресов | 8000 | 500 | 10168000 8200 4000 |
| Метод продвижения | CT | CT, S&F | S&FS&F CT, S&F S&F |

*Примечания:* СT - режим "на лету" (cut-through); S&F - с буферизацией (store-and-forward).

*Коммутатор LANplex 6012* представляет собой старшую модель коммутатора локальных сетей, предназначенную для работы на уровне магистрали корпоративной сети.

Структурная схема коммутатора LANplex 6012 приведена на рисунке 7.2.

Структура коммутатора до сих пор выдает ориентацию его ранних версий на коммутацию FDDI/Ethernet. До появления модулей, выходящих на высокоскоростную протокольно-независимую шину HSI, коммутатор использовал шины FDDI для межмодульного обмена.

Основные характеристики коммутатора LANplex 6012:

* Устройство управления (отдельный модуль) поддерживает SNMP, RMON и FDDI SMT;
* Виртуальные сети создаются на основе:
* группирования портов;
* группирования МАС-адресов.
* Поддерживается IP и IPX маршрутизация (RIP):
* несколько подсетей на один порт;
* несколько портов на одну подсеть.
* IP- фрагментация;
* ASIC+RISC процессоры;
* Наличие функции Roving Analysis Port позволяет наблюдать за трафиком любого порта коммутатора;
* Поддержка алгоритма Spanning Tree;
* Фильтрация широковещательного шторма.

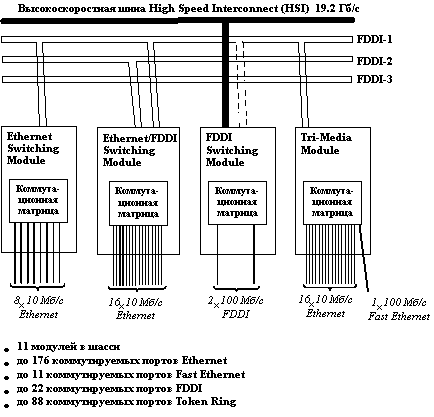


Рис. 7.2. Коммутатор LANplex 6012

### Примеры АТМ-коммутаторов для локальных сетей

#### Коммутаторы CELLplex компании 3Com

Коммутатор **CELLplex 7000** представляет собой модульное устройство на основе шасси, осуществляющее коммутацию до 16 портов ATM (4 модуля по 4 порта). Он предназначен для образования высокоскоростной ATM-магистрали сети путем соединения с другими ATM-коммутаторами или же для подключения высокоскоростных ATM-узлов к стянутой в точку магистрали сети на основе центра данных, имеющего порт ATM.

Коммутационный центр обеспечивает обмен данными по схеме 16(16, используя неблокирующую технологию коммутации "на лету" с общей пропускной способностью 2.56 Гб/с и поддерживая до 4096 виртуальных каналов на порт.

Пассивная внутренняя шина коммутатора обеспечивает передачу данных со скоростью до 20.48 Гб/с, обеспечивая переход в будущем на интерфейсные модули с большим количеством портов или с более скоростными портами.

Полностью избыточное шасси со сдвоенным источником питания, продублированным коммутационным центром и модульное построение делают коммутатор CELLplex 7000 отказоустойчивым устройством, подходящим для построения магистрали сети и удовлетворяющим требованиям наиболее важных приложений.

Имеются два типа интерфейсных модулей:

* модуль с 4 портами OC-3c 155 Мб/с для многомодового оптоволоконного кабеля, предназначенный для локальных связей;
* модуль с 4 портами DS-3 45 Мб/с - для глобальных связей.

Коммутатор поддерживает основные спецификации технологии ATM: установление коммутируемых виртуальных каналов (SVC) по спецификациям UNI 3.0 и 3.1, поддержку постоянных виртуальных каналов (PVC) с помощью системы управления, Interim Interswitch Signaling Protocol (IISP), эмуляцию локальных сетей (LAN emulation), управление перегрузками (congestion management).

Управление коммутатором реализовано для стандартов: SNMP, ILMI, MIB 2, ATM MIB, SONET MIB. Используется система управления Transcend.

Коммутатор **CELLplex 7200** совмещает функции ATM-коммутатора и Ethernet-комму-   
татора, одновременно позволяя ликвидировать узкие места на магистрали сети и в сетях отделов.

CELLplex 7200 обеспечивает полноскоростные Ethernet-каналы для разделяемых сегментов локальных сетей, серверов и отдельных рабочих станций, требующих повышенного быстродействия. Кроме этого, коммутатор может быть сконфигурирован с портами ATM для соединения с коммутаторами рабочих групп, ATM-серверами и рабочими станциями, а также для подключения к ATM-магистрали сети.

Коммутационный ATM-центр (8(8) совмещен с процессором Ethernet/ATM коммутации на микросхеме ZipChip. ZipChip преобразует пакеты данных Ethernet в стандартные ячейки ATM, а затем коммутирует их со скоростью до 780000 ячеек в секунду.

В отличие от модели CELLplex 7000 модель CELLplex 7200 имеет не два, а четыре типа интерфейсных модулей:

* модуль с двумя портами ATM OC-3c;
* модуль с двумя портами DS-3;
* модуль с 12 портами Ethernet и одним портом ATM OC-3c;
* модуль с 12 портами Ethernet и одним портом ATM DS-3.

Остальные характеристики коммутаторов CELLplex 7200 и CELLplex 7000 практически совпадают.

#### Коммутаторы технологии ATM LattisCell и EtherCell компании Bay Networks

Семейство продуктов, разработанных компанией Bay Networks для технологии ATM, состоит из коммутаторов LattisCell (только ATM-коммутация), коммутатора EtherCell (коммутация Ethernet-ATM), программного обеспечения ATM Connection Management System и программного обеспечения ATM Network Management Application.

Поставляется несколько моделей коммутаторов ATM, каждый из которых обеспечивает определенное сочетание физических уровней, сред передачи и возможностей резервирования источников питания.

Коммутатор EtherCell предназначен для устранения "узких мест" в рабочих группах локальных сетей, использующих традиционную разделяемую среду передачи данных технологии Ethernet. С помощью этого коммутатора можно разгрузить линии связи с серверами и маршрутизаторами. Модель **10328 EtherCell** имеет 12 портов 10Base-T и прямой доступ к сети ATM. Порты Ethernet могут предоставлять выделенную полосу пропускания 10 Мб/с за счет их коммутации.

Программное обеспечение ATM *Connection Management System (CMS)* размещается на рабочей станции SunSPARCStation, выполняя функции координации и управления соединениями коммутатора. CMS автоматически изучает сетевую топологию и устанавливает виртуальные ATM-соединения между взаимодействующими станциями.

Программное обеспечение ATM *Network Management Application*, работая совместно с CMS, обеспечивает управление сетью ATM на центральной станции управления.

Модель ATM коммутатора **LattisCell 10114A** разработана для использования в сетях кампусов (расстояние между коммутаторами до 2 км) и представляет собой устройство, выполненное в виде автономного корпуса с фиксированным количеством портов, число которых равно 16. Для каждого порта обеспечивается пропускная способность в 155 Мб/с по многомодовому оптоволоконному кабелю. Функции физического уровня реализованы в соответствии со стандартами SONET/SDH 155 Мб/с, а также UNI 3.0

Архитектура FastMatrix обеспечивает общую внутреннюю скорость передачи данных 5 Гб/с, позволяющую производить коммутацию всех портов без блокировок. Поддерживаются функции широковещательной (broadcast) и многовещательной (multicast) передачи.

Запрос на установление соединения может быть выполнен для различных уровней качества сервиса (Quality of Service, QoS):

* QoS 1 - используется для сервиса CBR (постоянная битовая скорость);
* QoS 2 - используется для сервиса VBR RT (переменная битовая скорость приложений реального времени);
* QoS 3/4 - используется для сервиса VBR, предназначенного для передачи данных локальных сетей по процедурам с установлением соединений и без установления соединений;
* QoS 0 - используется для сервиса UBR.

Управление устройством осуществляется также с помощью программной системы CMS, для которой необходимы: SunSPARCStation 2 или выше, Sun OS 4.1.3 или выше для невыделенного Ethernet-соединения или Solaris 2.4 для прямого ATM-соединения.

Другие модели коммутаторов LattisCell (10114R, 10114A-SM, 10114R-SM, 10114R-SM, 10114-DS3, 10114-E3, 10115A, 10115R) различаются наличием резервного источника питания, а также типом портов (общее количество портов в любой модели составляет 16). Кроме многомодовых портов, коммутаторы могут иметь одномодовые оптоволоконные порты (для сетей кампусов с расстоянием до 25 км), а также порты для коаксиального кабеля с интерфесами DS-3 (45 Мб/с) и E3 (34 Мб/с) для подключения к глобальным сетям через линии T3/E3.

Модели коммутатора **EtherCell** (10328-F и 10328-SM) обеспечивают коммутацию Ethernet-Ethernet и Ethernet-ATM. Эти модели имеют 12 портов 10Base-T RJ-45 и один порт прямого доступа к ATM со скоростью 10 Мб/с. Порты 10Base-T могут использоваться для предоставления полной скорости 10 Мб/с выделенной линии для высокоскоростных серверов или же для разделения ее между сегментом станций рабочей группы.

Модель EtherCell 10328-F поддерживает многомодовый оптоволоконный кабель для связи с сетью ATM на расстоянии до 2 км.

Модель EtherCell 10328-SM поддерживает одномодовый оптоволоконный кабель для связи с сетью ATM на расстоянии до 20 км.

Коммутаторы поддерживают стандарт LAN emulation, определяющий взаимодействие локальных сетей с сетями ATM на уровне протоколов канального уровня. Кроме этого, поддерживаются спецификации UNI, MIB-II, EtherCell-MIB и стандартный формат MIB компании Bay Networks.

Через ATM-порт коммутаторы EtherCell могут соединяться с портом SONET/SDH коммутатора LattisCell.

Коммутаторы EtherCell включают программу-агент *HSA (Host Signaling Agent),* которая является агентом-посредником для Ethernet-хостов.

Коммутаторы EtherCell поддерживают образование виртуальных групп, распределенных по ATM-магистрали сети, образованной коммутаторами LattisCell.

#### Коммутатор LightStream 1010 компании Cisco

Коммутатор **LightStream 1010** является ATM коммутатором для образования магистралей сетей отделов или кампусов.

Коммутатор обладает общей производительностью 5 Гб/с и выполнен на базе 5-слотового шасси.

В центральном слоте устанавливается модуль управления коммутацией *ATM Switch Processor (ASP),* который имеет разделяемую память со скоростью доступа 5 Гб/с, полностью неблокирующую коммутационную матрицу, а также высокопроизводительный RISC-процессор MIPS R4600 100 MHz. Модуль ASP работает под управлением межсетевой операционной системы IOS, как и маршрутизаторы и коммутаторы старших моделей компании Cisco. Программное обеспечение модуля ASP может заменяться "на ходу", то есть без остановки коммутатора, что важно в условиях часто изменяющихся спецификаций ATM Forum.

Оставшиеся 4 слота используются для установки интерфейсных модулей CAM, в каждый из которых можно установить до 2-х модулей адаптеров портов PAM. Таким образом, коммутатор может иметь в максимальной конфигурации до 8 модулей PAM из следующего набора:

* 1 порт ATM 622 Мб/с (OC12) (одномодовый);
* 1 порт АТМ 622 Мб/с (OC12) (многомодовый);
* 4 порта ATM 155 Мб/с (OC3с) (одномодовый);
* 4 порта ATM 155 Мб/с (OC3с) (многмодовый);
* 4 порта ATM 155 Мб/с (OC3с) (по неэкранированной витой паре UTP Cat 5);
* 2 порта DS3/T3 45 Мб/с;
* 2 порта E3 34 Мб/c.

Коммутатор LightStream 1010 одним из первых в отрасли поддерживает спецификацию маршрутизации PNNI Phase 1, необходимую для маршрутизации коммутируемых соединений (SVC) в неоднородных АТМ-сетях с учетом требуемого качества обслуживания.

Поддерживаются все определенные ATM Forum виды трафика, в том числе ABR.

Для соединений "пользователь - коммутатор" используется протокол UNI 3.0 (в ближайшее время ожидается также поддержка UNI 3.1).

Коммутатор LightStream 1010 может выполнять роль центрального коммутатора в сети кампуса (рисунок 7.3).

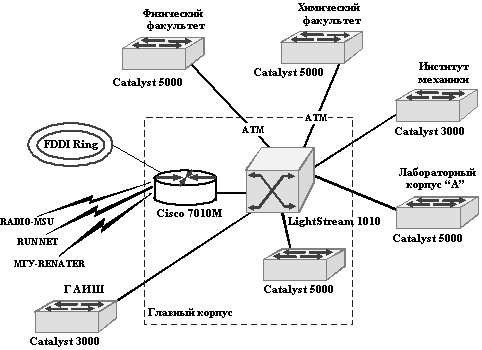


Рис. 7.3. Магистральная сеть Московского государственного университета