Содержание

Демонополизация рынка телекоммуникационных услуг в США (дополнительный материал к «Сети операторов связи» в главе 5)

В США компания AT&T до 1984 года была монополистом по предоставлению как локальных услуг телефонии, так и услуг дальней связи. В 1984 году по решению суда AT&T была разделена на части, из них наиболее важными были компания AT&T Long Lines, которой было разрешено предоставлять только услуги дальней связи, и 23 компании BOC (Bell Operating Company), получившие право оказывать телефонные услуги только в локальных масштабах. Для предоставления услуг в масштабах регионов компании BOC были объединены в семь региональных компаний — Regional BOC (RBOC). Лишенным привилегий и разукрупненным национальным монополистам приходится бороться за клиентов с новыми операторами, которые приходят как на рынок локальных услуг, так и на региональные рынки и рынок дальней связи. Таких операторов обычно называют *альтернативными*. В США процесс конкурентного развития рынка телекоммуникационных услуг был ускорен в 1996 году, когда Конгрессом был принят документ Telecommunication Act, снимающий ограничения для оператора связи на предоставление услуг только в одном секторе рынка (либо дальней или региональной связи, либо локальных услуг). Сегодня в США наряду с бывшими монополистами — **уполномоченными местными операторами связи** (Incumbent Local Exchange Carrier, ILEC) — работает множество **альтернативных местных операторов связи** (Competitive Local Exchange Carrier, CLEC). Не менее острая конкуренция идет и на рынке региональной и дальней связи США, где работает достаточно много крупных операторов, называемых **транснациональными операторами** (InterXchange Carrier, IXC). Эта терминология актуальна не только для читателей, живущих в США, поскольку она иногда используется для описания проектных решений и даже технологий, так что по типу оператора связи (IXC, CLEC или ILEC) легко понять его место в системе отношений операторов и специфику его услуг.

Локальные, региональные, национальные и транснациональные операторы

По степени покрытия территории, на которой предоставляются услуги, операторы делятся на локальных, региональных, национальных и транснациональных.

Локальный оператор работает на территории города или сельского района. Традиционный локальный оператор владеет всей соответствующей транспортной инфраструктурой: физическими каналами между помещениями абонентов (квартирами, домами и офисами) и узлом связи, автоматическими телефонными станциями (АТС) и каналами связи между телефонными станциями. Сегодня к традиционным локальным операторам добавились альтернативные (CLEC), которые часто являются поставщиками услуг нового типа, прежде всего, услуг Интернета, но иногда они конкурируют с традиционными операторами и в секторе телефонии.

Несмотря на демонополизацию телекоммуникационной отрасли, физическими каналами доступа к абонентам по-прежнему в большинстве случаев владеют традиционные локальные операторы, такие как ILEC в США.

В подобных неравноправных условиях альтернативным местным операторам достаточно трудно вести свой бизнес. У них есть несколько возможностей. Во-первых, они могут предоставлять только дополнительные услуги по передаче и обработке данных, например, доступ в Интернет, размещение в своих узлах информационных ресурсов клиентов и т. п. А для организации доступа абонентов к этим ресурсам можно заключить договор с традиционным оператором, который будет направлять трафик непосредственно подключенных к нему абонентов в сеть альтернативного оператора. Здесь мы видим естественную специализацию операторов — каждый занимается тем делом, для которого в большей степени подходит его инфраструктура, при этом сотрудничество приносит дополнительный эффект, порождая новые услуги. Во-вторых, они могут брать в аренду у традиционных локальных операторов абонентские окончания. Обычно традиционные операторы идут на это очень неохотно, хотя законодательство в ряде стран принуждает или поощряет их к этому. Остается еще и третий вариант — создать собственную сеть абонентских окончаний, причем абонентские окончания могут быть проводными и беспроводными. Учитывая количество домов и квартир, сложности прокладки к ним кабеля и необходимость приобретения разрешения на эту прокладку у местных властей, проводной вариант часто оказывается нереальным. Это обстоятельство породило большой интерес к беспроводным решениям, которые сегодня бурно развиваются.

Региональные и национальные операторы оказывают услуги на большой территории, располагая соответствующей транспортной инфраструктурой. Традиционные операторы этого масштаба выполняют транзитную передачу телефонного трафика между телефонными станциями локальных операторов, имея в своем распоряжении крупные транзитные АТС, связанные высокоскоростными физическими каналами связи. Это — операторы операторов, их клиентами являются, как правило, локальные операторы или крупные предприятия, имеющие отделения и филиалы в различных городах региона или страны. Располагая развитой транспортной инфраструктурой, такие операторы обычно оказывают услуги дальней связи, передавая транзитом большие объемы информации без какой-либо обработки.

Транснациональные операторы оказывают услуги в нескольких странах. Примерами таких операторов являются Cable & Wireless, Global One, Infonet. Они имеют собственные магистральные сети, покрывающие иногда несколько континентов. Часто подобные операторы тесно сотрудничают с национальными операторами, используя их сети доступа для доставки информации клиентам.

Взаимоотношения между операторами связи различного типа

Взаимосвязи между операторами различного типа (а также их сетями) иллюстрирует рис. 1. На рисунке показаны клиенты двух типов — индивидуальные и корпоративные. Нужно иметь в виду, что каждый клиент обычно нуждается в услугах двух видов — телефонных и передачи данных. Индивидуальные клиенты имеют в своих домах или квартирах, как правило, телефон и компьютер, а у корпоративных клиентов имеются соответствующие сети — телефонная сеть, поддерживаемая офисным теле-

фонным коммутатором (PBX), и локальная сеть передачи данных, построенная на собственных коммутаторах.

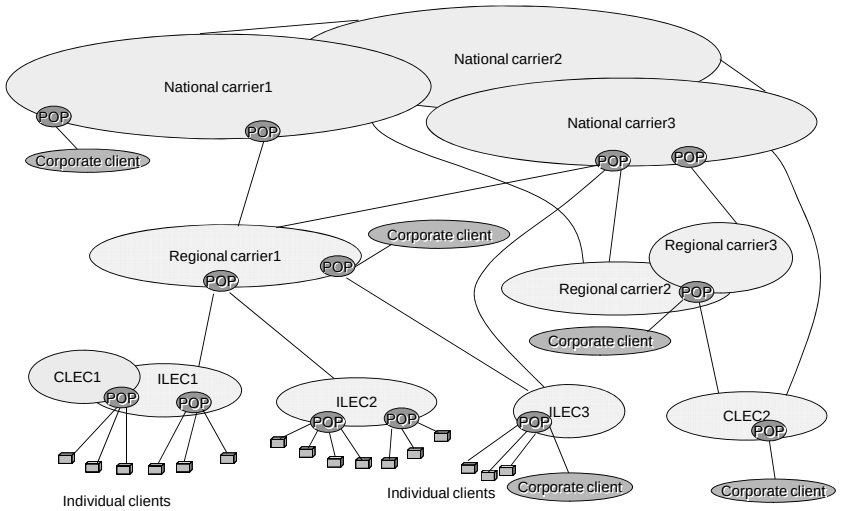

Рис. 1. Взаимоотношения между операторами связи различного типа

Для подключения оборудования клиентов операторы связи организуют, так называемые, **точки присутствия** (Point Of Presents, POP) — здания или помещения, в которых размещается оборудование доступа, способное обеспечить подключение большого количества каналов связи, идущих от клиентов. Иногда такую точку называют **центральным офисом** (Central Office, CO) — это традиционное название для операторов телефонных сетей. К POP локальных операторов подключаются абоненты, а к POP операторов верхних уровней — операторы нижних уровней или крупные корпоративные клиенты, которым необходимы высокие скорости доступа и большая территория покрытия, способная объединить их офисы и отделения в разных городах и странах.

Так как процесс конвергенции пока еще не привел нас к появлению единой сети для всех видов трафика, то за каждым овалом, представляющим на этом рисунке сети операторов, стоят две сети — телефонная и передачи данных.

Как видно из рисунка, в современном конкурентном телекоммуникационном мире нет строгой иерархии операторов, взаимосвязи между ними и их сетями могут быть достаточно сложными и запутанными. Например, сеть CLEC2 имеет непосредственную связь не только с сетью регионального оператора 3, как того требует иерархия, но и непосредственную связь с национальным оператором 3 (возможно, этот оператор предлагает более дешевые услуги по передаче международного трафика, чем это делает региональный оператор 3). Не все операторы на рисунке имеют собственную транспортную инфраструктуру (например, CLEC1). Как это часто бывает в таких

случаях, оператор CLEC1 предоставляет только дополнительные информационные услуги, например, предлагает клиентам оператора ILEC1 видео по требованию или разработку и поддержание их домашних страниц в Интернете. Свое оборудование (например, видео-сервер) такой оператор часто размещает в POP другого оператора, как это и показано в данном случае.

Форматы кадров технологии Ethernet (дополнительный материал к «Форматы кадров Ethernet» в главе 12)

Стандарт Ethernet, определенный в документе IEEE 802.3, дает описание единственного формата кадра уровня MAC. Так как в кадр уровня MAC должен вкладываться кадр уровня LLC, описанный в документе IEEE 802.2, то по стандартам IEEE в сети Ethernet может использоваться только единственный вариант кадра канального уровня, заголовок которого является комбинацией заголовков подуровней MAC и LLC.

Тем не менее, на практике в сетях Ethernet на канальном уровне используются кадры четырех различных форматов (типов). Один и тот же тип кадра может иметь разные названия, поэтому далее для каждого типа кадров приведено несколько наиболее употребительных названий.

- Кадр **Ethernet DIX**, или **Ethernet II**, появился в результате работы консорциума трех фирм Digital, Intel и Xerox в 1980 году, представившего на рассмотрение комитету 802.3 свою фирменную версию стандарта Ethernet в качестве проекта международного стандарта.
- Однако комитет 802.3 принял стандарт, отличающийся в некоторых деталях от предложения DIX, причем отличия касались и формата кадра. Так возник формат кадра **802.3/LLC**, **802.3/802.2**, или **Novell 802.2**.
- Кадр **Raw 802.3**, или **Novell 802.3**, появился в результате усилий компании Novell по ускорению работы своего стека протоколов в сетях Ethernet.
- Кадр **Ethernet SNAP** стал результатом деятельности комитета 802.2 по приведению предыдущих форматов кадров к некоторому общему стандарту и приданию кадру необходимой гибкости, что позволило бы в будущем добавлять новые поля или изменять их назначение.

Различия в форматах кадров могут приводить к несовместимости в работе аппаратуры и сетевого программного обеспечения, рассчитанного на функционирование только с одним стандартом кадра Ethernet. Однако сегодня практически все сетевые адаптеры, драйверы сетевых адаптеров, мосты/коммутаторы и маршрутизаторы умеют работать со всеми используемыми на практике форматами кадров технологии Ethernet, причем распознавание типа кадра происходит автоматически.

Форматы всех этих четырех типов кадров Ethernet приведены на рис. 1.

Кадр 802.3/LLC

6	6	2	1	1	1(2)	46–1497 (1496)	4
DA	SA	L	DSAP	SSAP	Control	Data	FCS
			Заголовок LLC				

Кадр Raw 802.3/Novell 802.3

6	6	2	46–1500	4
DA	SA	L	Data	FCS

Кадр Ethernet DIX (II)

6	6	2	46–1500	4
DA	SA	T	Data	FCS

Кадр Ethernet SNAP

6	6	2	1	1	1	3	2	46–1492	4
DA	SA	L	DSAP	SSAP	Control	OUI	T	Data	FCS
			AA	AA	03	000000			
			Заголовок LLC			Заголовок SNAP			

Рис. 1. Форматы кадров Ethernet

Кадр 802.3/LLC

Заголовок кадра 802.3/LLC является результатом объединения полей заголовков кадров, определенных в стандартах IEEE 802.3 и 802.2.

Стандарт 802.3 определяет восемь полей заголовка (на рисунке поле преамбулы и начальный ограничитель кадра не показаны).

- **Поле преамбулы** состоит из семи синхронизирующих байтов — 10101010. При манчестерском кодировании эта комбинация представляется в физической среде периодическим волновым сигналом с частотой 5 МГц.
- **Начальный ограничитель кадра** (Start-of-Frame-Delimiter, SFD) состоит из одного байта 10101011. Появление этой комбинации битов является указанием на то, что следующий байт — это первый байт заголовка кадра.
- **Адрес назначения** (Destination Address, DA) может быть длиной 2 или 6 байт. На практике всегда используются MAC-адреса из 6 байт.
- **Адрес источника** (Source Address, SA) — это 2- или 6-байтное поле, содержащее MAC-адрес узла — отправителя кадра. Первый бит адреса всегда имеет значение 0.
- **Длина** (Length, L) — 2-байтное поле, которое определяет длину поля данных в кадре.

- **Поле данных** может содержать от 0 до 1500 байт. Но если длина поля меньше 46 байт, то используется следующее поле — *поле заполнения*, дополняющее кадр до минимально допустимого значения в 46 байт.
- **Поле заполнения** состоит из такого количества байтов заполнителей, которое обеспечивает минимальную длину поля данных в 46 байт. Это обеспечивает корректную работу механизма обнаружения коллизий. Если длина поля данных больше или равна минимальной, то поле заполнения в кадре отсутствует.
- **Поле контрольной последовательности кадра** (Frame Check Sequence, FCS) состоит из 4 байт контрольной суммы. Это значение вычисляется по алгоритму CRC-32.

Кадр 802.3 является кадром подуровня MAC, поэтому в соответствии со стандартом 802.2 в его поле данных вкладывается кадр подуровня LLC с удаленными флагами начала и конца кадра.

Поле **DSAP** (Destination Service Access Point — точка входа службы приемника) служит для хранения кода протокола, которому адресовано содержимое поля данных. Соответственно, поле **SSAP** (Source Service Access Point — точка входа службы источника) используется для указания кода протокола, от которого посылаются данные. Применение двух полей для целей демультиплексирования довольно нетипично, обычно протоколы обходятся одним полем, например, протокол IP всегда посылает свои пакеты протоколу IP, а протокол IPX — протоколу IPX. Два поля полезны в тех случаях, когда вышележащий протокол поддерживает несколько режимов работы, так что протокол на узле-отправителе может использовать различные значения DSAP и SSAP для уведомления узла получателя о переходе в новый режим работы. Поле Control (управление) обеспечивает надежность передачи кадров. Фактически, эта функция работает только в режиме LLC2, в этом случае поле Control состоит из двух байтов. В режимах LLC1 и LLC3 поле Control состоит из одного байта и полезной работы не выполняет.

Так как кадр LLC имеет заголовок длиной 3 (в режиме LLC1) или 4 байта (в режиме LLC2), то максимальный размер поля данных уменьшается до 1497 или 1496 байт.

Кадр Raw 802.3/Novell 802.3

Кадр Raw 802.3, называемый еще кадром *Novell 802.3*, также представлен на рис. 1. Из рисунка видно, что он представляет собой кадр подуровня MAC стандарта 802.3, но без вложенного кадра подуровня LLC. Компания Novell долгое время не использовала служебные поля кадра LLC в своей операционной системе NetWare, поскольку не было необходимости идентифицировать тип информации, вложенной в поле данных — там всегда находился пакет протокола IPX, долгое время бывшего единственным протоколом сетевого уровня в ОС NetWare.

Теперь, когда необходимость в идентификации протокола верхнего уровня появилась, компания Novell стала использовать возможность инкапсуляции в кадр подуровня MAC кадра LLC, то есть возможность применять стандартные кадры 802.3/LLC. Такой кадр компания обозначает теперь в своих операционных системах как кадр 802.2, хотя он является комбинацией заголовков 802.3 и 802.2.

Кадр Ethernet DIX/Ethernet II

Кадр Ethernet DIX, называемый также *кадром Ethernet II*, имеет структуру, совпадающую со структурой кадра Raw 802.3 (см. рис. 1). Однако 2-байтное поле длины (L) кадра Raw 802.3 в кадре *Ethernet DIX* используется в качестве поля типа (Type, T) протокола. Это поле предназначено для тех же целей, что и поля DSAP и SSAP кадра LLC — для указания типа протокола верхнего уровня, вложившего свой пакет в поле данных этого кадра.

В то время как коды протоколов в полях SAP имеют длину 1 байт, в поле типа для кода протокола отводятся 2 байта. Поэтому один и тот же протокол в поле SAP и поле типа будет кодироваться в общем случае разными числовыми значениями. Например, протокол IP имеет код 2048_{10} (0x0800) для поля типа и значение 6 для поля SAP. Значения кодов протоколов для поля типа появились раньше значений для поля SAP, так как фирменная версия Ethernet DIX существовала до появления стандарта 802.3, и ко времени распространения оборудования 802.3 эти значения уже стали стандартами де-факто для многих аппаратных и программных продуктов. Так как структуры кадров Ethernet DIX и Raw 802.3 совпадают, то поле длины/типа часто в документации обозначают как поле L/T. При этом числовое значение этого поля определяет его смысл: если значение меньше 1500, то это поле длины, а если больше — то типа.

Кадр Ethernet SNAP

Для устранения разнобоя в кодировках типов протоколов, сообщения которых вложены в поле данных кадров Ethernet, комитетом 802.2 была проведена работа по дальнейшей стандартизации кадров Ethernet. В результате появился кадр Ethernet SNAP (SubNetwork Access Protocol — протокол доступа к подсетям). Кадр Ethernet SNAP (см. рис. 1) представляет собой расширение кадра 802.3/LLC за счет введения дополнительного заголовка протокола SNAP, состоящего из двух полей: OUI и типа. Поле типа состоит из 2 байт и повторяет по формату и назначению поле типа кадра Ethernet II (то есть в нем используются те же значения кодов протоколов). Поле OUI определяет уже знакомый нам организационно уникальный идентификатор — то есть идентификатор организации, которая контролирует коды протоколов в поле типа. С помощью заголовка SNAP достигнута совместимость с кодами протоколов в кадрах Ethernet II, а также создана универсальная схема кодирования протоколов. Коды протоколов для технологий 802 контролирует организация IEEE, идентификатор OUI которой равен 000000. Если в будущем потребуются другие коды протоколов для какой-либо новой технологии, для этого достаточно будет указать другой идентификатор организации, назначающей эти коды, а старые значения кодов останутся в силе (в сочетании с другим идентификатором OUI).

Так как SNAP представляет собой протокол, вложенный в протокол LLC, то в полях DSAP и SSAP записывается код 0xAA, отведенный для протокола SNAP. В управляющем поле заголовка LLC устанавливается значение 0x03, что соответствует использованию ненумерованных кадров.

Заголовок SNAP является дополнением к заголовку LLC, поэтому он допустим не только в кадрах Ethernet, но и в кадрах протоколов других технологий комитета 802. Например, протокол IP всегда использует структуру заголовков LLC/SNAP при инкапсуляции в кадры всех протоколов локальных сетей: FDDI, Token Ring, 100VG-

AnyLAN, Ethernet, Fast Ethernet, Gigabit Ethernet. Правда, при передаче IP-пакетов через сети Ethernet, Fast Ethernet и Gigabit Ethernet протокол IP использует кадры Ethernet DIX.

Использование различных типов кадров Ethernet

Из-за того что существует четыре типа кадров Ethernet, для протоколов сетевого уровня возникает проблема: пользоваться всегда одним и тем же типом кадра, применять все четыре или же отдавать предпочтение только некоторым из них?

Протокол IP может использовать два типа кадров: оригинальный кадр Ethernet II и наиболее структурно сложный кадр Ethernet SNAP. Предпочтительным типом кадра для протокола IP является кадр Ethernet II.

Современные сетевые адаптеры автоматически распознают тип кадра Ethernet, используя значения полей кадров. Например, кадры Ethernet II легко отличить от других типов кадров по значению поля L/T: если оно больше 1500, значит, это поле является полем типа протокола (T), так как значения кодов протоколов выбраны так, что они всегда больше 1500. В свою очередь наличие поля T говорит о том, что это кадр Ethernet II, который единственный использует это поле в данной позиции кадра.

Протокол IPX «является максималистом», он может работать со всеми четырьмя типами кадров Ethernet. Он распознает кадры Ethernet II описанным способом, а если кадр принадлежит к другому типу (поле L/T имеет значение меньшее или равное 1500), то выполняется дальнейшая проверка по наличию или отсутствию полей LLC. Поля LLC могут отсутствовать только в том случае, если за полем длины идет начало пакета IPX, а именно 2-байтное поле, которое всегда заполняется единицами, что дает значение 0xFFFF, или два байта по 255. Ситуация, когда поля DSAP и SSAP одновременно содержат такие значения, возникнуть не может, поэтому наличие двух байтов 255 говорит о том, что это кадр Raw 802.3.

В остальных случаях дальнейший анализ проводится в зависимости от значений полей DSAP и SSAP. Если они равны 0xAA, то это кадр Ethernet SNAP, а если нет, то 802.3/LLC.

Физические стандарты Ethernet (дополнительный материал к «Стандарты физического уровня Ethernet» в главе 12)

Исторически первые сети технологии Ethernet были созданы на коаксиальном кабеле диаметром 0,5 дюйма. В дальнейшем были определены и другие спецификации физического уровня для стандарта Ethernet, позволяющие задействовать различные среды передачи данных. Метод доступа CSMA/CD и все временны*е* параметры остаются одними и теми же для любой спецификации физической среды технологии Ethernet 10 Мбит/с.

Физические спецификации Ethernet на сегодня включают следующие среды передачи данных.

- **10Base-5** — коаксиальный кабель диаметром 0,5 дюйма, называемый «толстым» коаксиалом. Имеет волновое сопротивление 50 Ом. Максимальная длина сегмента — 500 метров (без повторителей).
- **10Base-2** — коаксиальный кабель диаметром 0,25 дюйма, называемый «тонким» коаксиалом. Имеет волновое сопротивление 50 Ом. Максимальная длина сегмента — 185 метров (без повторителей).
- **10Base-T** — кабель на основе неэкранированной витой пары (UTP). Образует звездообразную топологию на основе концентратора. Расстояние между концентратором и конечным узлом — не более 100 м.
- **10Base-F** — волоконно-оптический кабель. Топология аналогична 10Base-T. Имеется несколько вариантов этой спецификации — FOIRL (расстояние до 1000 м), 10Base-FL (расстояние до 2000 м), 10Base-FB (расстояние до 2000 м).

Число 10 в указанных названиях обозначает номинальную битовую скорость передачи данных этих стандартов — 10 Мбит/с, а слово «Base» — метод передачи на одной базовой частоте 10 МГц — в отличие от методов, использующих несколько несущих частот (они называются широкополосными и имеют в своем составе слово «Broadband»). Последний символ в названии стандарта физического уровня обозначает тип кабеля.

Стандарт 10Base-5

Стандарт 10Base-5 в основном соответствует экспериментальной сети Ethernet фирмы Xerox и может считаться классическим стандартом Ethernet. Различные компоненты сети, выполненной на толстом коаксиале и состоящей из трех сегментов, соединенных повторителями, показаны на рис. 1.

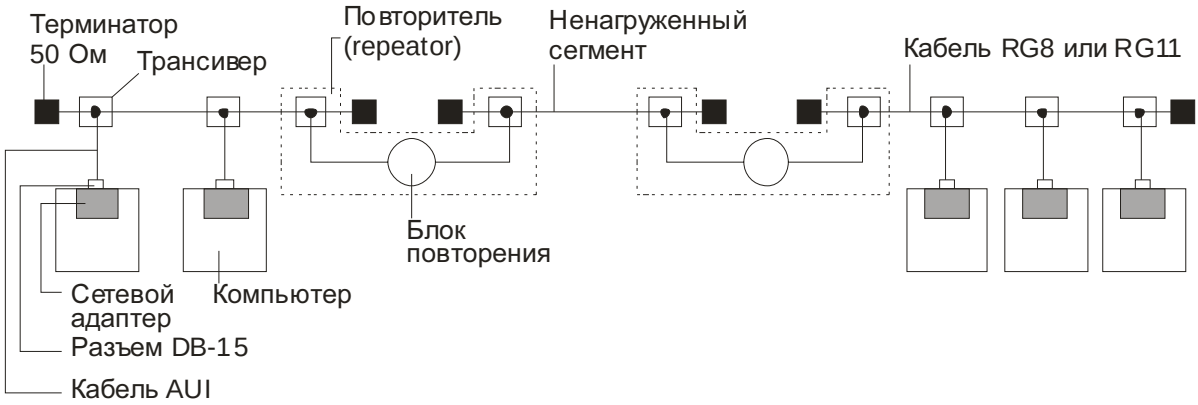

Рис. 1. Компоненты физического уровня сети стандарта 10 Base-5, состоящей из трех сегментов

Кабель используется как моноканал для всех станций. Сегмент кабеля максимальной длины в 500 м (без повторителей) должен иметь на концах согласующие *терминаторы* («заглушки») сопротивлением 50 Ом, поглощающие распространяющиеся по кабелю сигналы и препятствующие возникновению отраженных сигналов. При отсутствии терминаторов в кабеле возникают стоячие волны, так что одни узлы получают мощные сигналы, а другие — настолько слабые, что их прием становится невозможным.

Станция должна подключаться к кабелю при помощи приемопередатчика — **трансивера**. Трансивер — это часть сетевого адаптера; он устанавливается непосредственно на кабеле и питается от сетевого адаптера компьютера. Трансивер может подсоединяться к кабелю как методом прокалывания, обеспечивающим непосредственный физический контакт, так и бесконтактным методом.

Трансивер соединяется с сетевым адаптером интерфейсным кабелем *AUI* (Attachment Unit Interface — интерфейс подключаемых устройств) длиной до 50 м, состоящим из четырех витых пар (адаптер должен иметь разъем AUI). Наличие стандартного интерфейса между трансивером и остальной частью сетевого адаптера очень полезно при переходе с одного типа кабеля на другой. Для этого достаточно только заменить трансивер, а остальная часть сетевого адаптера остается неизменной, так как она отрабатывает протокол уровня MAC. При этом необходимо только, чтобы новый трансивер (например, трансивер для витой пары) поддерживал стандартный интерфейс AUI.

Допускается подключение к одному сегменту не более 100 трансиверов, причем расстояние между подключениями трансиверов не должно быть меньше 2,5 м. На кабеле имеется разметка через каждые 2,5 м, обозначающая точки подключения трансиверов. При подсоединении компьютеров в соответствии с разметкой влияние стоячих волн в кабеле на сетевые адаптеры сводится к минимуму.

Упрощенная структурная схема трансивера показана на рис. 2. Передатчик и приемник присоединяются к одной точке кабеля с помощью специальной схемы, например трансформаторной, позволяющей организовать одновременную передачу и прием сигналов с кабеля.

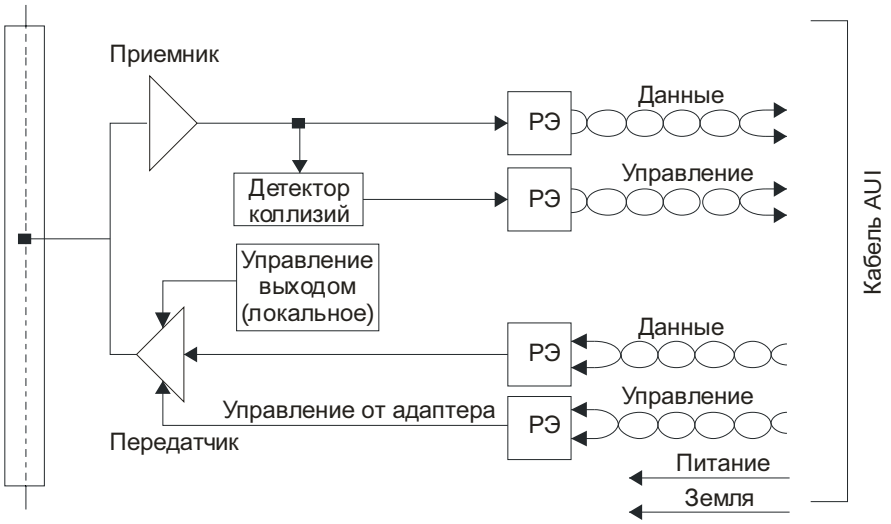

Рис. 2. Структурная схема трансивера

При неисправностях в адаптере может возникнуть ситуация, когда в кабель будет непрерывно выдаваться последовательность случайных сигналов. Так как кабель — это общая среда для всех станций, работа сети будет заблокирована одним неисправным адаптером. Чтобы этого не случилось, на выходе трансивера ставится схема, которая проверяет время передачи кадра. Если максимально возможное время передачи пакета превышается (с некоторым запасом), то эта схема просто отсоединяет выход передатчика от кабеля. Максимальное время передачи кадра (вместе с преамбулой) равно 1221 мкс, а время затянувшейся передачи устанавливается равным 4000 мкс (4 мс). Эту функцию трансивера иногда называют проверкой затянувшейся передачи, или **jabber-контролем**.

Детектор коллизий определяет наличие коллизии в коаксиальном кабеле по повышенному уровню постоянной составляющей сигналов. Если постоянная составляющая превышает определенный порог (около 1,5 В), значит, на кабель работает более одного передатчика.

Развязывающие элементы (РЭ) обеспечивают гальваническую развязку трансивера от остальной части сетевого адаптера и тем самым защищают адаптер и компьютер от значительных перепадов напряжения, возникающих на кабеле при его повреждении.

Стандарт 10Base-5 определяет возможность использования в сети **повторителя**. Повторитель служит для объединения в одну сеть нескольких сегментов кабеля и увеличения тем самым общей длины сети. Повторитель принимает сигналы из одного сегмента кабеля и побитно синхронно повторяет их в другом сегменте, улучшая форму и мощность импульсов, а также синхронизируя импульсы. Повторитель состоит из двух (или нескольких) трансиверов, которые присоединяются к сегментам кабеля, а также блока повторения со своим тактовым генератором. Для лучшей синхронизации повторитель задерживает передачу нескольких первых битов преамбулы кадра, за счет чего увеличивается задержка передачи кадра с сегмента на сегмент, а также несколько уменьшается межкадровый интервал (IPG).

Стандарт разрешает использование в сети не более четырех повторителей и, соответственно, не более пяти сегментов кабеля. При максимальной длине сегмента кабеля в 500 м это дает максимальную длину сети 10Base-5 в 2500 м. Это в точности соответствует общему ограничению стандарта на максимальный диаметр Ethernet.

Только три сегмента из пяти могут быть нагруженными, то есть такими, к которым подключаются конечные узлы. Между нагруженными сегментами должны быть ненагруженные, так что максимальная конфигурация сети представляет собой два нагруженных крайних сегмента, которые соединяются ненагруженными сегментами еще с одним центральным нагруженным сегментом. Ранее на рис. 1 был приведен пример сети Ethernet, состоящей из трех сегментов, объединенных двумя повторителями. Крайние сегменты являются нагруженными, а промежуточный — ненагруженным.

Правило применения повторителей в сети Ethernet 10Base-5 носит название **правила 5-4-3**, то есть 5 сегментов, 4 повторителя, 3 нагруженных сегмента.

Ограниченное число повторителей объясняется дополнительными задержками распространения сигнала, которые они вносят. Применение повторителей увеличивает время оборота сигнала, которое для надежного распознавания коллизий не должно превышать время передачи кадра минимальной длины, то есть кадра в 72 байт, или 576 бит. Каждый повторитель подключается к сегменту одним своим трансивером, поэтому к нагруженным сегментам можно подключить не более 99 узлов (а не 100). Максимальное число конечных узлов в сети 10Base-5 таким образом составляет $99 \times 3 = 297$ узлов.

Стандарт 10Base-2

В стандарте 10Base-2 в качестве передающей среды используется «тонкий» коаксиал Ethernet. Максимальная длина сегмента без повторителей составляет 185 м, сегмент должен иметь на концах согласующие терминаторы 50 Ом. Тонкий коаксиальный кабель дешевле толстого, поэтому сети 10Base-2 иногда называют Cheapernet (дословно — дешевая сеть). Но за дешевизну кабеля приходится расплачиваться качеством — «тонкий» коаксиал обладает худшими показателями помехозащищенности и механической прочности, а также более узкой полосой пропускания.

Станции подключаются к кабелю с помощью высокочастотного **Т-коннектора**, представляющего собой тройник, один отвод которого соединяется с сетевым адаптером, а два других — с двумя концами разрыва кабеля. Максимальное количество станций, подключаемых к одному сегменту, равно 30. Минимальное расстояние между станциями — 1 м. Кабель «тонкого» коаксиала имеет разметку для подключения узлов с шагом в 1 м.

Стандарт 10Base-2 также предусматривает использование повторителей по правилу 5-4-3.

В этом случае сеть будет иметь максимальную длину в $5 \times 185 = 925$ м. Очевидно, что это ограничение является более сильным, чем общее ограничение стандарта Ethernet в 2500 м.

ВНИМАНИЕ

Для правильного построения сети Ethernet нужно соблюсти много ограничений, причем некоторые из них относятся к одним и тем же параметрам сети, например, максимальная длина или максимальное количество компьютеров в сети должны удовлетворять одновременно нескольким разным условиям. Для того чтобы сеть была работоспособной, достаточно соблюсти только наиболее жесткие требования. Так, если в сети Ethernet не должно быть более 1024 узлов, а стандарт 10Base-2 ограничивает максимальное число станций, подключаемых к одному сегменту, значением 30, и число нагруженных сегментов — значением 3, то общее количество узлов в сети 10Base-2 не должно превышать $29 \times 3 = 87$.

Стандарт 10Base-2 очень близок к стандарту 10Base-5, но трансиверы в нем объединены с сетевыми адаптерами за счет того, что более гибкий тонкий коаксиальный кабель может быть подведен непосредственно к выходному разъему платы сетевого адаптера, установленной в шасси компьютера. Кабель в данном случае «висит» на сетевом адаптере, что затрудняет физическое перемещение компьютеров.

Типичный состав сети стандарта 10Base-2, состоящей из одного сегмента кабеля, показан на рис. 3.

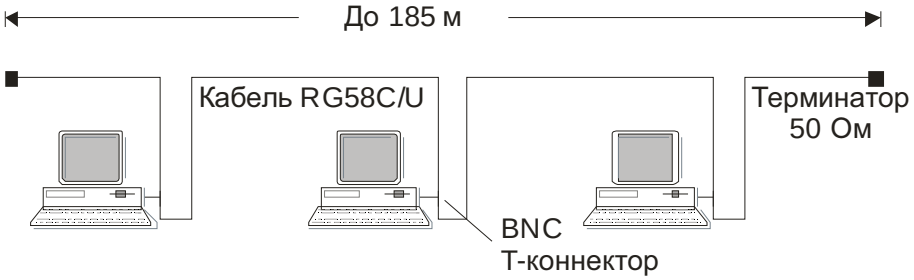

Рис. 3. Сеть стандарта 10Base-2

Реализация этого стандарта на практике приводит к наиболее простому решению для кабельной сети, так как для соединения компьютеров требуются только сетевые адаптеры, Т-коннекторы и терминаторы 50 Ом. Однако этот вид кабельных соединений наиболее сильно подвержен авариям и сбоям. Кабель более восприимчив к помехам, чем «толстый» коаксиал. В моноканале имеется большое количество механических соединений: каждый Т-коннектор дает три механических соединения, два из которых жизненно важны для всей сети. Пользователи имеют доступ к разъемам и могут нарушить целостность моноканала. Кроме того, эстетика и эргономичность этого решения оставляют желать лучшего, так как от каждой станции через Т-коннектор отходят два довольно заметных провода, которые под столом часто образуют моток кабеля — запас, необходимый на случай даже небольшого перемещения рабочего места.

Общим недостатком стандартов 10Base-5 и 10Base-2 является отсутствие оперативной информации о состоянии моноканала. Повреждение кабеля обнаруживается сразу же (сеть перестает работать), но для поиска отказавшего отрезка кабеля необходим специальный прибор — кабельный тестер.

Стандарт 10Base-T

В сетях 10Base-T в качестве среды используются две *неэкранированные витые пары*. Многопарный кабель на основе неэкранированной витой пары категории 3 телефонные компании уже достаточно давно применяли для подключения телефонных аппаратов внутри зданий. Этот кабель носит также название Voice Grade, говорящее о том, что он предназначен для передачи голоса.

Идея приспособить этот популярный вид кабеля для локальных сетей оказалась очень плодотворной, так как многие здания уже были оснащены нужной кабельной системой. Оставалось разработать способ подключения сетевых адаптеров и прочего коммуникационного оборудования к витой паре таким образом, чтобы изменения в сетевых адаптерах и программном обеспечении сетевых операционных систем были минимальными по сравнению с сетями Ethernet на коаксиале. Эта попытка оказалась успешной — переход на витую пару требует только замены трансивера сетевого адаптера или порта маршрутизатора, а метод доступа и все протоколы канального уровня остаются теми же, что и в сетях Ethernet на коаксиале.

Конечные узлы соединяются с помощью двух витых пар по двухточечной топологии со специальным устройством — многопортовым повторителем. Одна витая пара требуется для передачи данных от станции к повторителю (выход T_X сетевого адаптера), а другая — для передачи данных от повторителя к станции (вход R_X сетевого адаптера). На рис. 4 показан пример трехпортового повторителя. Повторитель принимает сигналы от одного из конечных узлов и синхронно передает их на все свои остальные порты, исключая порт, с которого поступили сигналы.

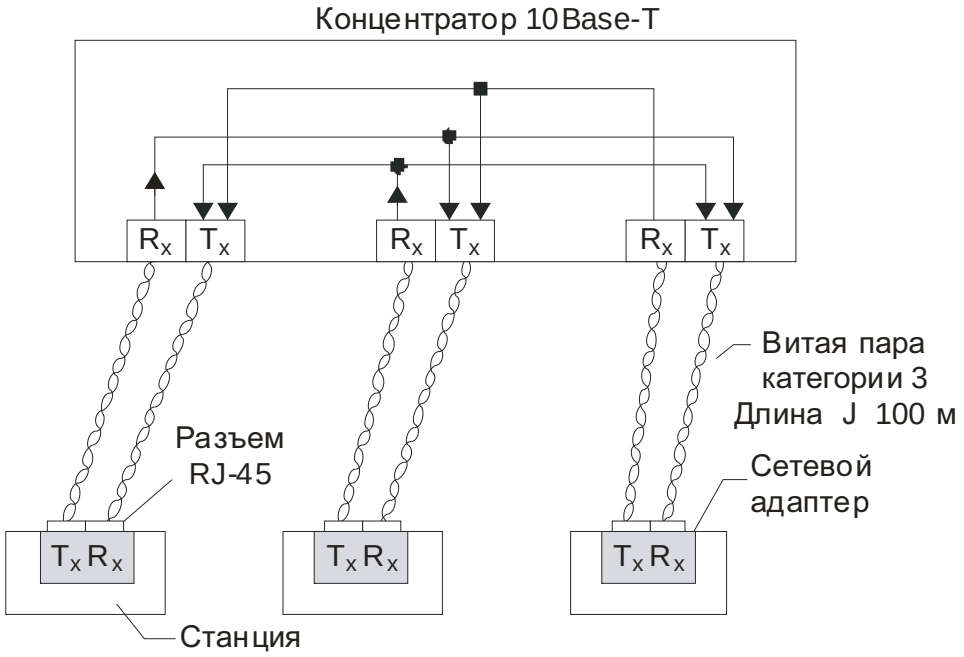

Рис. 4. Сеть стандарта 10Base-T

Многопортовые повторители в данном случае обычно называются **концентраторами**, или, на инженерном жаргоне, **хабами**. Концентратор осуществляет функции повторителя сигналов на всех отрезках витых пар, подключенных к его портам, так что образуется единая среда передачи данных — логический моноканал (логическая общая шина). Концентратор обнаруживает коллизию в сегменте в случае одновременной передачи сигналов по нескольким своим входам R_X и посылает jam-последовательность на все свои выходы T_X. Стандарт определяет битовую скорость передачи данных 10 Мбит/с и максимальное расстояние отрезка витой пары между двумя непосредственно связанными узлами (станциями и концентраторами) не более 100 м при наличии витой пары качества не ниже категории 3. Это расстояние определяется полосой пропускания витой пары — на длине 100 м она при манчестерском кодировании позволяет передавать данные со скоростью 10 Мбит/с.

Концентраторы 10Base-T можно соединять друг с другом с помощью тех же портов, которые предназначены для подключения конечных узлов. При этом нужно позаботиться о том, чтобы передатчик и приемник одного порта были соединены соответственно с приемником и передатчиком другого порта.

В стандарте10Base-T определено максимально число концентраторов между любыми двумя станциями сети, а именно 4. Это правило носит название **правила 4-х хабов**.

Правило 4-х хабов подобно правилу 5-4-3, применяемому к коаксиальным сетям, служит для гарантированной синхронизации станций при реализации процедур доступа CSMA/CD и для надежного распознавания станциями коллизий.

При создании сети 10Base-T с большим числом станций концентраторы можно соединять друг с другом иерархическим способом, образуя древовидную структуру (рис. 5).

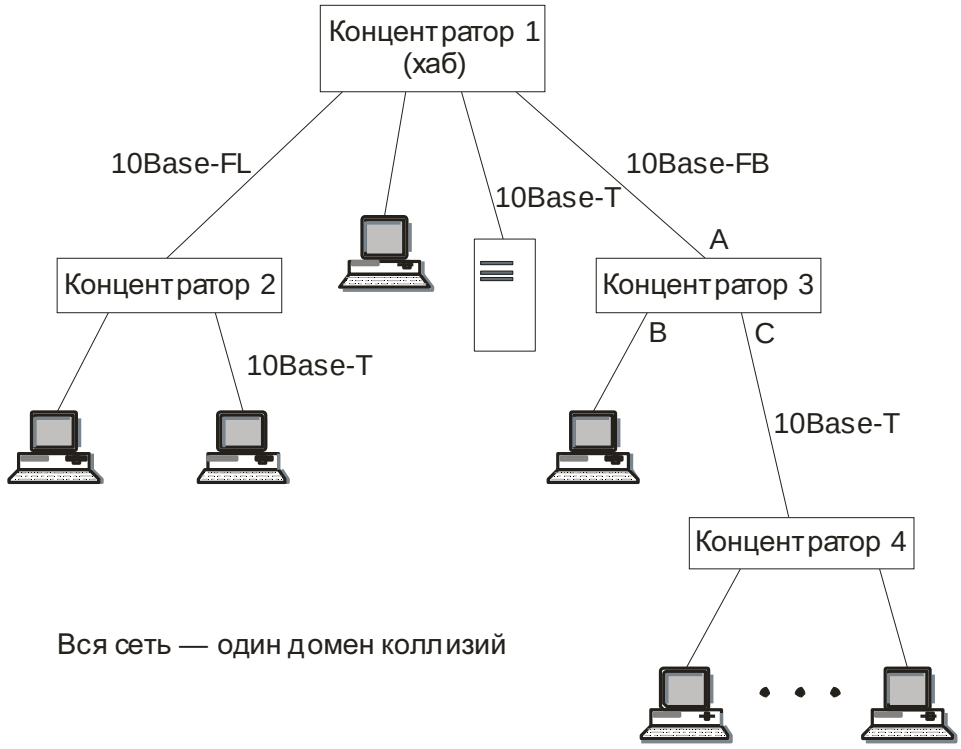

Рис. 5. Иерархическое соединение концентраторов Ethernet

ВНИМАНИЕ

Петлевидное соединение концентраторов в стандарте 10Base-T запрещено, так как оно приводит к некорректной работе сети. Это требование означает, что в сети 10Base-T не разрешается создавать параллельные каналы связи между критически важными концентраторами для резервирования связей на случай отказа порта, концентратора или кабеля. Резервирование связей возможно только за счет перевода одной из параллельных связей в неактивное (заблокированное) состояние.

Общее количество станций в сети 10Base-T не должно превышать общего предела в 1024, и для данного типа физического уровня это количество действительно достижимо. Для этого достаточно создать двухуровневую иерархию концентраторов, расположив на нижнем уровне достаточное количество концентраторов с общим количеством портов 1024 (рис. 6). Конечные узлы нужно подключить к портам концентраторов нижнего уровня. Правило 4-х хабов при этом выполняется — между любыми конечными узлами будет ровно 3 концентратора.

Рис. 6. Схема с максимальным количеством станций

Очевидно, что если между любыми двумя узлами сети не должно быть больше 4-х повторителей, то учитывая, что максимальная длина кабеля между повторителями равна 100 м, получаем, что *максимальный диаметр сети 10Base-T составляет 5 × 100 = 500 м*. Заметим, что это ограничение строже общего ограничения стандартов Ethernet в 2500 м.

Сети, построенные на основе стандарта 10Base-T, обладают по сравнению с коаксиальными вариантами Ethernet многими преимуществами. Эти преимущества связаны с разделением общего физического кабеля на отдельные кабельные отрезки, подключенные к центральному коммуникационному устройству. И хотя логически эти отрезки по-прежнему образуют общую разделяемую среду, их физическое разделение позволяет контролировать состояние отрезков и отключать их в случае обрыва, короткого замыкания или неисправности сетевого адаптера на индивидуальной основе. Это обстоятельство существенно облегчает эксплуатацию больших сетей Ethernet, так как концентратор обычно автоматически выполняет такие функции, уведомляя при этом администратора сети о возникшей проблеме.

В стандарте 10Base-T определена процедура тестирования физической работоспособности двух отрезков витой пары, соединяющих трансивер конечного узла и порт повторителя. Эта процедура называется *тестом связности* и основана на передаче каждые 16 мс специальных сигналов J и K манчестерского кода между передатчиком и приемником каждой витой пары. Напомним, что информационные сигналы манчестерского кода всегда изменяют потенциал в середине такта. Коды J и K отличаются тем, что нарушают это правило, сохраняя потенциал в середине такта неизменным. Коду J соответствует одно из двух значений потенциала, а коду K — другое. Так как при передаче кадров коды J и K запрещены, тестовые последовательности не влияют на работу алгоритма доступа к среде.

Появление между конечными узлами активного устройства, которое может контролировать работу узлов и изолировать от сети некорректно работающие узлы, является *главным преимуществом* технологии 10Base-T по сравнению со сложными в эксплуатации коаксиальными сетями.

Волоконно-оптическая сеть Ethernet

В качестве среды передачи данных 10-мегабитная сеть Ethernet использует оптическое волокно. Оптоволоконные стандарты в качестве основного типа кабеля рекомендуют достаточно дешевое многомодовое оптическое волокно, обладающее полосой пропускания 500–800 МГц при длине кабеля 1 км. Допустимо и более дорогое

одномодовое оптическое волокно с полосой пропускания в несколько ГГц, но при этом нужно применять специальный тип трансивера.

Функционально сеть Ethernet на оптическом кабеле состоит из тех же элементов, что и сеть стандарта 10Base-T — сетевых адаптеров, многопортового повторителя и отрезков кабеля, соединяющих адаптер с портом повторителя. Как и в случае витой пары, для соединения адаптера с повторителем используются два оптоволокна — одно соединяет выход T_X адаптера с входом R_X повторителя, другое — вход R_X адаптера с выходом T_X повторителя.

Стандарт FOIRL (Fiber Optic Inter-Repeater Link — волоконно-оптический канал между повторителями) представляет собой первый стандарт комитета 802.3, регламентирующий использование оптоволокна в сетях Ethernet. Он гарантирует длину оптоволоконной связи между повторителями до 1 км. Максимальное число повторителей между любыми узлами сети — 4. Как и в стандарте 10Base-5, максимального диаметра в 2500 м здесь достичь можно, однако отрезки кабеля предельного размера между *всеми* четырьмя повторителями, а также между повторителями и конечными узлами недопустимы — иначе получится сеть длиной 5000 м.

Стандарт 10Base-FL представляет собой незначительное улучшение стандарта FOIRL. Увеличена мощность передатчиков, поэтому максимальное расстояние между узлом и концентратором увеличилось до 2000 м. Максимальное число повторителей между узлами осталось равным 4, и стандартная максимальная длина сети 2500 м достижима.

Стандарт 10Base-FB предназначен только для соединения повторителей. Конечные узлы не могут использовать этот стандарт для присоединения к портам концентратора. Между узлами сети можно установить до 5 повторителей 10Base-FB при максимальной длине одного сегмента 2000 м и максимальной длине сети 2740 м.

Повторители, соединенные по стандарту 10Base-FB, при отсутствии кадров для передачи в целях синхронизации постоянно обмениваются специальными последовательностями сигналов, отличающимися от сигналов кадров данных. Поэтому они вносят меньшие задержки при передаче данных из одного сегмента в другой, и это является главной причиной, по которой количество повторителей удалось увеличить до 5. В качестве специальных сигналов используются манчестерские коды J и K в следующей последовательности: J-J-K-K-J-J-... Эта последовательность порождает импульсы частоты 2,5 МГц, которые и поддерживают синхронизацию приемника одного концентратора с передатчиком другого. Поэтому стандарт 10Base-FB имеет также название **синхронный стандарт Ethernet**.

Как и во всех стандартах Ethernet, оптоволоконные стандарты разрешают соединять концентраторы только в древовидные иерархические структуры. Любые петли между портами концентраторов не допускаются.

Домен коллизий

Домен коллизий — это часть сети Ethernet, все узлы которой конкурируют за общую разделяемую среду передачи и, следовательно, каждый узел которой может создать коллизию с любым другим узлом этой части сети.

Сеть Ethernet, построенная на повторителях, всегда образует один домен коллизий. Мосты, коммутаторы и маршрутизаторы делят сеть Ethernet на несколько доменов коллизий.

Приведенная на рис. 5 сеть представляет собой один домен коллизий. Если, например, столкновение кадров произошло в концентраторе 4, то в соответствии с логикой работы концентраторов 10Base-T сигнал коллизии распространится по всем портам всех концентраторов.

Если же вместо концентратора 3 поставить в сеть мост, то его порт С, связанный с концентратором 4, воспримет сигнал коллизии, но не передаст его на свои остальные порты, так как это не входит в его обязанности. Мост просто отработает ситуацию коллизии средствами порта С, который подключен к общей среде, где эта коллизия возникла. Если коллизия возникла из-за того, что *мост* пытался передать через порт С кадр в концентратор 4, то зафиксировав сигнал коллизии, порт С приостановит передачу кадра и попытается передать его повторно через случайный интервал времени. Если порт С принимал в момент возникновения коллизии кадр, то он просто отбросит полученное начало кадра и будет ожидать, когда узел, передававший кадр через концентратор 4, сделает повторную попытку передачи. После успешного принятия данного кадра в свой буфер мост передаст его на другой порт в соответствии с таблицей продвижения, например на порт А. Все события, связанные с обработкой коллизий портом С, для остальных сегментов сети, которые подключены к другим портам моста, просто останутся неизвестными.

Общие характеристики стандартов Ethernet 10 Мбит/с

В табл. 1 и 2 сведены основные ограничения и характеристики стандартов Ethernet.

Таблица 1. Общие ограничения для всех стандартов Ethernet

Характеристика	**Значение**
Номинальная пропускная способность	10 Мбит/с
Максимальное число станций в сети	1024
Максимальное расстояние между узлами в сети	2500 м (в 10Base-FB 2750 м)
Максимальное число коаксиальных сегментов в сети	5

Таблица 2. Параметры спецификаций физического уровня для стандарта Ethernet

Параметр	**10Base-5**	**10Base-2**	**10Base-T**	**10Base-F**
Кабель	Толстый коаксиальный кабель RG-8 или RG-11	Тонкий коаксиальный кабель RG-58	Неэкранированная витая пара категорий 3, 4, 5	Многомодовый волоконно-оптический кабель
Максимальная длина сегмента, м	500	185	100	2000
Максимальное расстояние между узлами сети (при использовании повторителей), м	2500	925	500	2500 (2740 для 10Base-FB)

Максимальное число станций в сегменте	100	30	1024	1024
Максимальное число повторителей между любыми станциями сети	4	4	4	4 (5 для 10 Base-FB)

Пример сети Ethernet завода «Трансмаш»

В начале 90-х годов крупный завод «Трансмаш» использовал сеть Ethernet 10 Мбит/с с разделяемой средой для объединения всех своих миникомпьютеров и персональных компьютеров (рис. 7). Компьютеры в основном применялись для решения автономных задач, а обмен данных между ними происходил сравнительно редко. Сеть передавала небольшие объемы алфавитно-цифровой информации, поэтому общая разделяемая среда вполне справлялась с потребностями завода. Для взаимодействия центрального сегмента сети с сегментами удаленных цехов использовались оптоволоконные линии связи стандартов 10Base-FB и 10Base-FL. Сеть удовлетворяла всем требованиям многосегментной конфигурации Ethernet: все отрезки кабелей не превышали предельной длины, между любыми двумя узлами находилось не более 4-х хабов, максимальное расстояние между узлами сети не превышало 1800 метров (компьютеры A и C на рисунке).

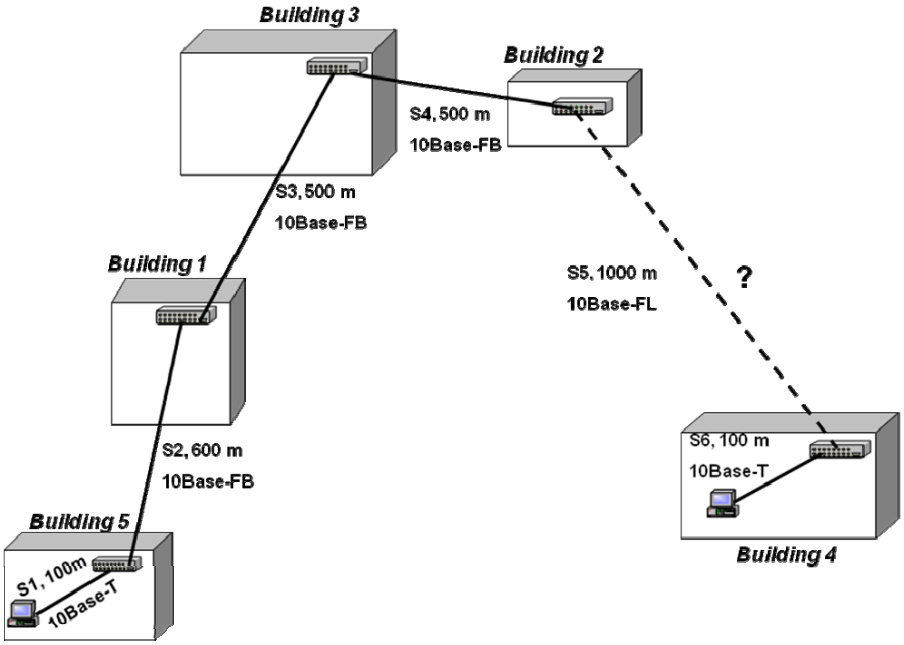

Рис. 7. Многосегментная сеть Ethernet завода «Трансмаш»

Через некоторое время к сети понадобилось присоединить компьютеры еще одного здания, а именно, здания 4. Это здание находилось в пределах досягаемости оптоволоконных стандартов Ethernet (10 Base-FB или 10Base-FL), но его присоединение к

сети привело бы к некорректной конфигурации, так как между компьютерами зданий 1 и 4 данные проходили бы уже через 5 хабов. Кроме того, диаметр сети достиг бы 2800 метров — еще одно нарушение ограничений Ethernet. Однако в то время архитектору сети «Transmash» не хотелось коренным образом менять структуру сети и устанавливать мост или маршрутизатор для подключения нового сегмента.

Архитектор сети знал, что в разделе 13 стандарта IEEE 802.3 приведена методика расчета корректности конфигурации сети. Эта методика позволяет количественно определить, будет ли та или иная конфигурация сети работать нормально или нет. Расчеты показывают, что иногда можно нарушить правило 4-х хабов и ограничения на максимальный диаметр сети, и все равно конфигурация будет корректной. Дело в том, что эти ограничения выбраны так, чтобы сеть работала с большим запасом «прочности». Например, мы знаем, что для надежного распознавания коллизий любым узлом сети максимальное время оборота не должно превышать 575 битовых интервалов. Если посчитать по приведенной методике время оборота в сети 10Base-5, состоящей из 4-х повторителей 10Base-5 и 5-ти сегментов максимальный длины 500 м, то окажется, что оно составляет 537 битовых интервала. Это значит, что максимальная конфигурация сети 10Base-5 (4 хаба, диаметр сети 2500 м) обладает запасом в 38 битовых интервала. В то же время методика раздела 13 говорит о том, что даже при запасе в 4 битовых интервала сеть будет работать корректно.

Поэтому архитектор сети «Трансмаш» выполнил расчет возможной конфигурации сети завода с учетом нового сегмента. Оказалось, что даже при присоединении сегмента здания 4 у сети имеется запас в 6,6 битовых интервала! После перепроверки расчета к зданию 4 был проложен волоконно-оптический кабель, и сеть начала работать в новой конфигурации. Практика подтвердила правильность расчета — сеть продолжала работать нормально. В такой конфигурации она оставалась несколько лет, пока возросшие потребности новых приложений не привели к разделению общей среды на коммутируемые сегменты.

Для того чтобы проверить расчет, который проделал архитектор сети «Трансмаш», нужно предварительно познакомиться с деталями методики, приведенной в разделе 13 стандарта 802.3.

В нем сказано, что сеть Ethernet будет работать корректно, при выполнении двух условий:

- Время оборота (PDV) сигнала между двумя самыми удаленными друг от друга станциями сети не должно превышать 575 битовых интервала. Повторители и среда сегментов вносят задержки в распространение сигнала, данные о предельных уровнях этих задержек приведены в таблицах стандарта.
- Сокращение межпакетного интервала IPG при прохождении последовательности кадров через все повторители должно быть не больше, чем 49 битовых интервала. Каждый повторитель сокращает значение IPG на определенную величину, которая также приводится в стандарте.

В таблицах стандарта 802.3 приводятся минимальные и максимальные значения возможных задержек распространения сигналов и сокращений IPG, их более определенные значения зависят от производителя повторителей. Архитектор сети «Трансмаш» использовал для расчета более точные данные, которые ему предоставил производитель сетевого оборудования. Эти данные приведены далее.

Рассмотрим сначала, как с помощью данных табл. 3 можно оценить значение PDV.

Таблица 3. Данные для расчета значения PDV

Тип сегмента	База левого сегмента, битовых интервалов	База промежуточного сегмента, битовых интервалов	База правого сегмента, битовых интервалов	Задержка среды на 1 м, битовых интервалов	Максимальная длина сегмента, м
10Base-5	11,8	46,5	169,5	0,0866	500
10Base-2	11,8	46,5	169,5	0,1026	185
10Base-T	15,3	42,0	165,0	0,113	100
10Base-FB	—	24,0	—	0,1	2000
10Base-FL	12,3	33,5	156,5	0,1	2000
FOIRL	7,8	29,0	152,0	0,1	1000
AUI (>2 м)	0	0	0	0,1026	2+48

Разработчики стандарта 802.3 старались максимально упростить выполнение расчетов, поэтому приведенные данные включают сразу несколько этапов прохождения сигнала. Например, задержки, вносимые повторителем, состоят из задержки входного трансивера, задержки блока повторения и задержки выходного трансивера. Тем не менее, в таблице все эти задержки представлены одной величиной, названной базой сегмента.

Чтобы не нужно было два раза складывать задержки, вносимые кабелем, в таблице даются удвоенные величины задержек для каждого типа кабеля.

В таблице используются также такие понятия, как левый сегмент, правый сегмент и промежуточный сегмент. Поясним эти термины на примере сети завода «Трансмаш» (см. рис. 7). Мы хотим рассчитать PDV для худшего случая. Поэтому мы выбрали для расчета узлы A и B, между которыми находятся 5 повторителей, а общая длина сети равна 2800 м.

Левым сегментом в терминологии 802.3 называется сегмент, в котором начинается путь сигнала от выхода передатчика конечного узла. Сам термин «левый» не имеет отношения к расположению сегментов в пространстве (и, конечно, на рисунке). Это просто условное название сегмента, с которого мы начинаем расчет. Для определенности мы выбрали в качестве левого сегмента сегмент S1, к которому подключен узел A.

Затем сигнал проходит через промежуточные сегменты S2–S5 и доходит до приемника (узел B) который подключен к сегменту S6. Именно здесь в худшем случае происходит столкновение кадров и возникает коллизия, что и подразумевается в таблице. Конечный сегмент, в котором может возникнуть коллизия, называется правым сегментом.

С каждым сегментом связана постоянная задержка, названная базой, которая зависит только от типа сегмента и от положения сегмента на пути сигнала (левый, промежуточный или правый). База правого сегмента, в котором возникает коллизия, намного превышает базу левого и промежуточных сегментов.

Кроме того, с каждым сегментом связана задержка распространения сигнала вдоль кабеля сегмента, которая зависит от длины сегмента и вычисляется путем умножения

времени распространения сигнала по одному метру кабеля (в битовых интервалах) на длину кабеля в метрах.

Расчет заключается в вычислении задержек, вносимых каждым отрезком кабеля (приведенная в таблице задержка сигнала на 1 м кабеля умножается на длину сегмента), а затем суммировании этих задержек с базами левого, промежуточных и правого сегментов.

Так как левый и правый сегменты имеют разные величины базовой задержки, то в случае различных типов сегментов на удаленных краях сети необходимо выполнить расчеты дважды: один раз принять в качестве левого сегмента сегмент одного типа, а во второй — сегмент другого типа. Результатом можно считать максимальное из полученных значений PDV. В нашем примере крайние сегменты сети принадлежат одному типу — стандарту 10Base-T, поэтому двойной расчет не требуется.

Теперь мы можем рассчитать значение PDV для нашего примера.

- Левый сегмент S1:

$$15{,}3\ (\text{база}) + 100 \times 0{,}113 = 26{,}6$$

- Промежуточный сегмент S2:

$$24 + 600 \times 0{,}1 = 84{,}0$$

- Промежуточный сегмент S3:

$$24 + 500 \times 0{,}1 = 74{,}0$$

- Промежуточный сегмент S4:

$$24 + 500 \times 0{,}1 = 74{,}0$$

- Промежуточный сегмент S5:

$$33{,}5 + 1000 \times 0{,}1 = 133{,}5$$

- Правый сегмент S6:

$$165 + 100 \times 0{,}113 = 176{,}3$$

Сумма всех составляющих дает значение PDV, равное 568,4.

Так как значение PDV меньше максимально допустимой величины 575 на 6,6 битовых интервала, то эта сеть обладает корректной конфигурацией несмотря на то, что ее общая длина превышает 2500 м, а количество повторителей больше 4-х.

Однако проверки PDV еще недостаточно для общего положительного заключения. Нужно также оценить уменьшение межкадрового интервала. Исходные данные для этого расчета приведены табл. 4.

Таблица 4. Уменьшение межкадрового интервала повторителями

Тип сегмента	Передающий сегмент, битовых интервалов	Промежуточный сегмент, битовых интервалов
10Base-5 или 10Base-2	16	11
10Base-FB	—	2

10Base-FL	10,5	8
10Base-T	10,5	8

В соответствии с этими данными рассчитаем значение уменьшения межкадрового интервала для нашего примера.

- Левый сегмент 1 10Base-T — сокращение в 10,5 битовых интервалов.
- Промежуточный сегмент 2 10Base-FL — 8.
- Промежуточный сегмент 3 10Base-FB — 2.
- Промежуточный сегмент 4 10Base-FB — 2.
- Промежуточный сегмент 5 10Base-FB — 2.

Сумма этих величин дает значение уменьшения межкадрового интервала в 24,5, что меньше предельного значения в 49 битовых интервала. Таким образом, приведенная в примере сеть соответствует стандартам Ethernet по всем параметрам, связанным и с длинами сегментов, и с количеством повторителей.

Технология FDDI (дополнительный материал к «Технология FDDI» в главе 12)

Технология **FDDI** (Fiber Distributed Data Interface — распределенный интерфейс передачи данных по оптоволокну) — это первая технология локальных сетей, в которой в качестве среды передачи данных стал применяться волоконно-оптический кабель. Работы по созданию технологий и устройств локальных сетей, использующих волоконно-оптические каналы, начались в 80-е годы, вскоре после начала промышленной эксплуатации подобных каналов в территориальных сетях. Проблемная группа X3T9.5 института ANSI разработала в период с 1986 по 1988 гг. начальные версии стандарта FDDI, который описывает передачу кадров со скоростью 100 Мбит/с по двойному волоконно-оптическому кольцу длиной до 100 км.

Основные характеристики технологии FDDI

Технология FDDI во многом основывается на технологии Token Ring, развивая и совершенствуя ее основные идеи. Разработчики технологии FDDI ставили перед собой следующие цели:

- повысить битовую скорость передачи данных до 100 Мбит/с;
- повысить отказоустойчивость сети за счет стандартных процедур восстановления ее после отказов различного рода, включая повреждение кабеля, некорректную работу узла, концентратора, возникновение сильных помех на линии и т. п.;
- максимально эффективно использовать потенциальную пропускную способность сети как для асинхронного, так и для синхронного (чувствительного к задержкам) трафиков.

Сеть FDDI строится на основе двух оптоволоконных колец, которые образуют основной и резервный пути передачи данных между узлами сети. Наличие двух колец — это основное средство повышения отказоустойчивости в сети FDDI.

Узлы, которые хотят воспользоваться этим повышенным потенциалом надежности, должны быть подключены к обоим кольцам. В технологии FDDI для передачи световых сигналов по оптическим волокнам реализовано кодирование 4B/5B в сочетании с кодированием NRZI. Эта схема приводит к передаче по линии связи сигналов с тактовой частотой 125 МГц.

В *нормальном* режиме работы сети данные проходят через все узлы и все участки кабеля только первичного кольца, этот режим назван **сквозным**, или **транзитным**. Вторичное кольцо в этом режиме не используется.

В случае какого-либо вида отказа, когда часть первичного кольца не может передавать данные (например, обрыв кабеля или отказ узла), первичное кольцо объединяется с вторичным (рис. 1), вновь образуя единое кольцо. Этот режим работы сети называется режимом *свертывания* колец. Операция свертывания производится средствами концентраторов и/или сетевых адаптеров FDDI. Для упрощения этой процедуры

данные по первичному кольцу всегда передаются в одном направлении (на диаграммах это направление изображается против часовой стрелки), а по вторичному — в обратном (изображается по часовой стрелке). Поэтому при образовании общего кольца из двух колец передатчики станций по-прежнему остаются подключенными к приемникам соседних станций, что позволяет правильно передавать и принимать информацию соседними станциями.

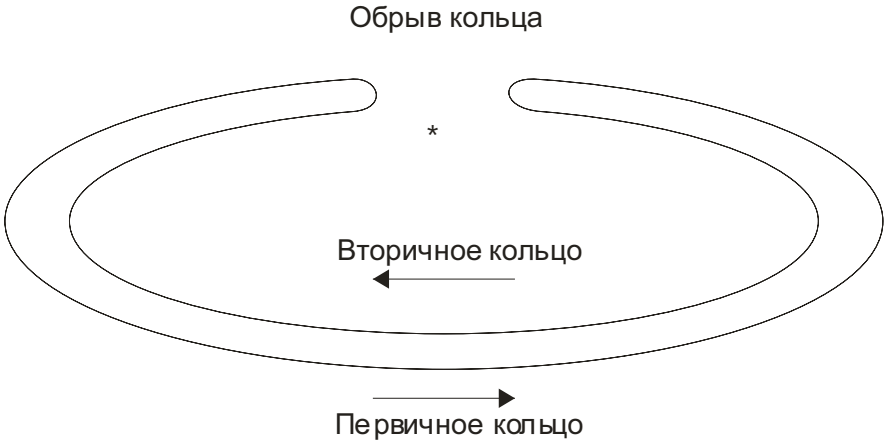

Рис. 1. Реконфигурирование колец FDDI при отказе

В стандартах FDDI много внимания отводится различным процедурам, которые позволяют определить факт наличия отказа в сети, а затем произвести необходимое реконфигурирование. Технология FDDI дополняет механизмы обнаружения отказов технологии Token Ring механизмами реконфигурирования пути передачи данных в сети, основанными на наличии резервных связей, которые предоставляет второе кольцо.

Сеть FDDI может полностью восстанавливать свою работоспособность в случае единичных отказов ее элементов. При множественных отказах сеть распадается на несколько не связанных сетей.

Кольца в сетях FDDI рассматриваются как общая разделяемая среда передачи данных. Метод доступа к этой среде очень близок методу доступа сетей Token Ring. Станции FDDI применяют *алгоритм раннего освобождения токена*, как и сети Token Ring 16 Мбит/с.

Отличия в методах доступа заключаются в следующем:

- Время удержания токена в сети FDDI не является постоянной величиной, как в сети Token Ring. Это время зависит от загрузки кольца — при небольшой загрузке оно растет, а при перегрузках может снижаться до нуля. Однако эти изменения касаются только асинхронного трафика, который не критичен к небольшим задержкам передачи кадров. Для синхронного трафика время удержания токена остается фиксированным.
- Механизм приоритетов кадров, принятый в Token Ring, в технологии FDDI отсутствует. Разработчики технологии решили, что деление трафика на 8 уровней приоритетов избыточно, достаточно разделить трафик на два класса — асин-

хронный и синхронный, последний из которых обслуживается всегда, даже при перегрузках кольца.

В остальном пересылка кадров между станциями кольца на уровне MAC полностью соответствует технологии Token Ring.

Рис. 2 иллюстрирует соответствие стека протоколов технологии FDDI семиуровневой модели OSI. FDDI определяет протокол физического уровня и протокол подуровня доступа к среде (MAC) канального уровня. Как и во многих других технологиях локальных сетей, в технологии FDDI используется протокол подуровня управления логическим каналом LLC.

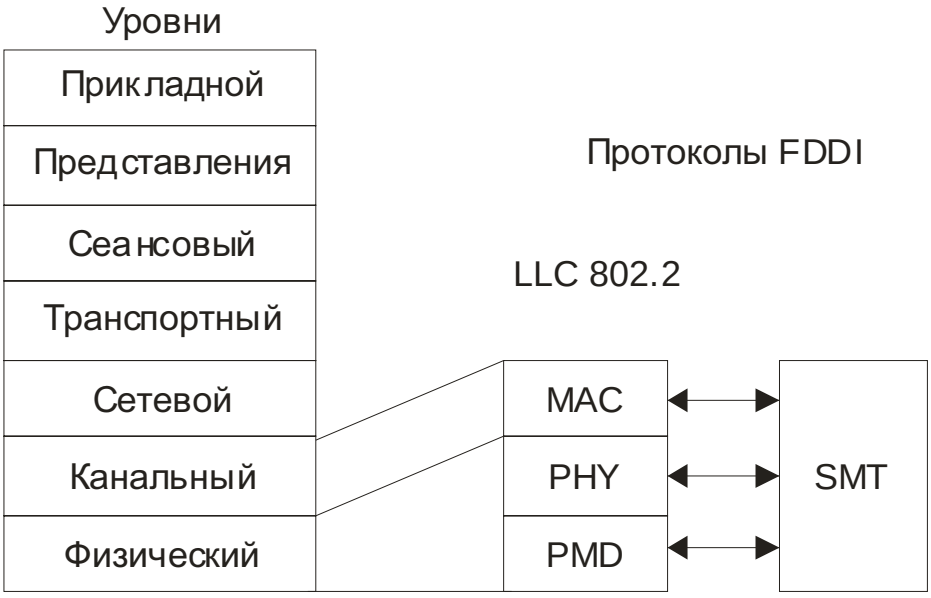

Рис. 2. Стек протоколов технологии FDDI

Специфической особенностью технологии FDDI является **уровень администрирования станции** (Station Management, SMT). Именно уровень SMT выполняет функции по администрированию и мониторингу всех остальных уровней стека протоколов FDDI. В управлении кольцом принимает участие каждый узел сети FDDI. Поэтому все узлы обмениваются специальными SMT-кадрами для управления сетью.

Отказоустойчивость сетей FDDI обеспечивается протоколами и других уровней: с помощью физического уровня устраняются отказы сети по физическим причинам, например, из-за обрыва кабеля, а с помощью уровня MAC — логические отказы сети, например, потерю нужного внутреннего пути передачи токена и кадров данных между портами концентратора.

Отказоустойчивость технологии FDDI

Как уже отмечалось, для обеспечения отказоустойчивости в стандарте FDDI предусмотрено создание двух оптоволоконных колец — *первичного* и *вторичного*. В стандарте FDDI определены два типа конечных узлов — *станции* и *концентраторы*. Для

подключения станций и концентраторов к сети может быть использован один из двух возможных способов.

- **Двойное подключение** (Dual Attachment, DA) — одновременное подключение к первичному и вторичному кольцам. Станция и концентратор, подключенные таким способом, называются соответственно станцией двойного подключения (Dual Attachment Station, DAS) и концентратором двойного подключения (Dual Attachment Concentrator, DAC).
- **Одиночное подключение** (Single Attachment, SA) — подключение только к первичному кольцу. Станция и концентратор, подключенные данным способом, называются соответственно станцией одиночного подключения (Single Attachment Station, SAS) и концентратором одиночного подключения (Single Attachment Concentrator, SAC).

Обычно, хотя и не обязательно, концентраторы имеют двойное подключение, а станции — одиночное, как показано на рис. 3. Чтобы устройства легче было правильно присоединять к сети, их разъемы маркируются. Разъемы типа A и B должны быть у устройств с двойным подключением; разъем M (Master) имеется у концентратора для одиночного подключения станции, у которой ответный разъем должен иметь тип S (Slave).

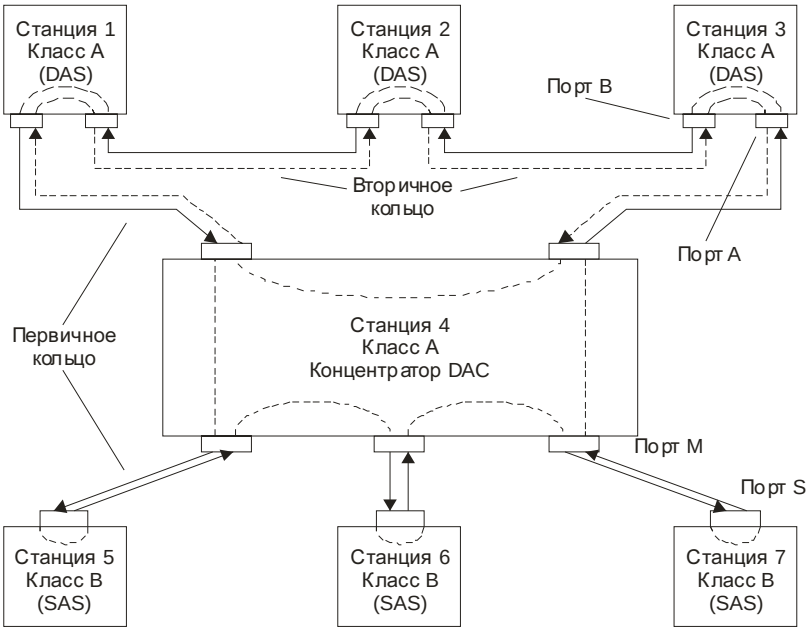

Рис. 3. Подключение узлов к кольцам FDDI

В случае однократного обрыва кабеля между устройствами с двойным подключением сеть FDDI сможет продолжить нормальную работу за счет автоматического реконфигурирования внутренних путей передачи кадров между портами концентратора (рис. 4).

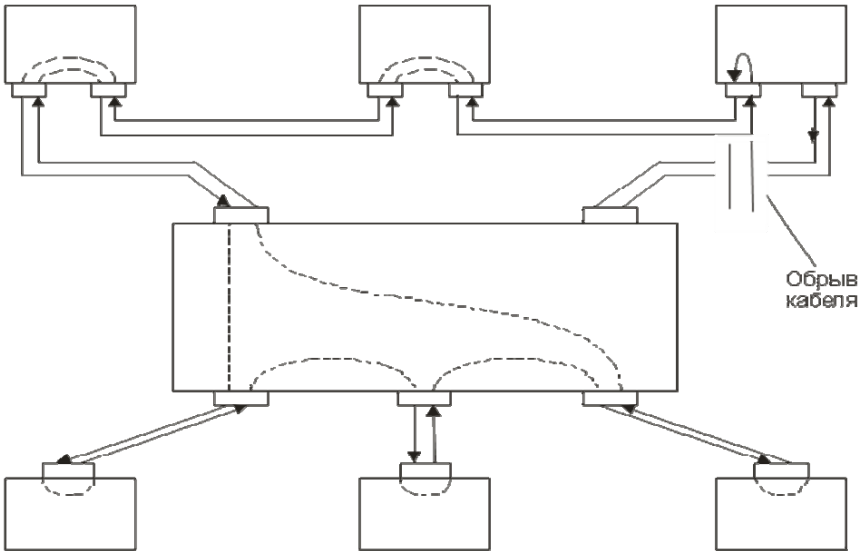

Рис. 4. Реконфигурирование сети FDDI при обрыве

Двукратный обрыв кабеля приведет к образованию двух изолированных сетей FDDI. При обрыве кабеля, идущего к станции с одиночным подключением, она оказывается отрезанной от сети, а кольцо продолжает работать за счет реконфигурирования внутреннего пути в концентраторе — порт M, к которому была подключена данная станция, исключается из общего пути.

Для сохранения работоспособности сети при отключении питания в станции с двойным подключением (например, просто при ее выключении) она должна быть оснащена оптическим обходным переключателем, который создаст резервный путь для световых потоков.

И, наконец, станции DAS или концентраторы DAC можно подключать к двум портам M одного или двух концентраторов, создавая древовидную структуру с основными и резервными связями. По умолчанию порт B поддерживает основную связь, а порт A — резервную. Такая конфигурация называется **двухпортовым подключением**.

Отказоустойчивость поддерживается за счет постоянного слежения концентраторов и станций уровня SMT за временны*ы*ми интервалами циркуляции токена и кадров, а также за наличием физического соединения между соседними портами в сети. В сети FDDI нет выделенного активного монитора — все станции и концентраторы равноправны, и при обнаружении отклонений от нормы они начинают процесс повторной инициализации сети, а затем и ее реконфигурирование.

Реконфигурирование внутренних путей в концентраторах и сетевых адаптерах выполняется специальными оптическими переключателями, которые перенаправляют световой луч и имеют достаточно сложную конструкцию.

Максимальная общая длина кольца FDDI составляет 100 километров, максимальное число станций с двойным подключением в кольце — 500.

Технология FDDI разрабатывалась для ответственных участков сетей — магистральных соединений между крупными сетями, например сетями зданий, а также для подключения к сети высокопроизводительных серверов. Поэтому главным для разработчиков было обеспечить высокую скорость передачи данных, отказоустойчивость на уровне протокола и большие расстояния между узлами сети. Все эти цели были достигнуты. В результате технология FDDI получилась качественной, но весьма дорогой. Даже появление более дешевого варианта для витой пары не намного снизило стоимость подключения одного узла к сети FDDI. Основной областью применения технологии FDDI стали магистрали сетей, состоящих из нескольких зданий, а также сети масштаба крупного города, то есть класса MAN.

Технология Token Ring (дополнительный материал к «Технология Token Ring» в главе 12)

Технология **Token Ring** был разработана компанией IBM в 1984 году, а затем передана в качестве проекта стандарта в комитет IEEE 802, который на ее основе принял в 1985 году стандарт 802.5. Компания IBM в течение долгого времени использовала технологию Token Ring как свою основную сетевую технологию построения локальных сетей на основе компьютеров различных классов — мэйнфреймов, миникомпьютеров и персональных компьютеров. Однако в последнее время даже в продукции компании IBM доминируют представители семейства Ethernet.

Сети Token Ring работают с двумя битовыми скоростями — 4 и 16 Мбит/с. Смешение в одном кольце станций, работающих на разных скоростях, не допускается. Сети Token Ring, работающие со скоростью 16 Мбит/с, имеют некоторые усовершенствования в алгоритме доступа по сравнению со стандартом 4 Мбит/с.

Технология Token Ring сложнее, чем Ethernet. Она обладает некоторыми начальными свойствами отказоустойчивости. В сети Token Ring определены процедуры контроля работы сети, которые опираются на свойство обратной связи, изначально присущее кольцеобразной структуре — посланный кадр всегда возвращается к станции-отправителю. В некоторых случаях обнаруженные ошибки в работе сети устраняются автоматически, например, может быть восстановлен потерянный токен. В других случаях ошибки только фиксируются, а их устранение выполняется вручную обслуживающим персоналом.

Для контроля сети одна из станций исполняет роль так называемого **активного монитора**. Активный монитор выбирается во время инициализации кольца, критерием выбора служит максимальное значение MAC-адреса. Если активный монитор выходит из строя, процедура инициализации кольца повторяется и выбирается новый активный монитор. Чтобы сеть могла обнаружить отказ активного монитора, последний в работоспособном состоянии каждые 3 секунды генерирует специальный кадр, обозначающий его присутствие. Если этот кадр не появляется в сети более 7 секунд, то остальные станции сети начинают процедуру выборов нового активного монитора.

Доступ с передачей токена

Сети Token Ring пользуются разделяемой средой путем передачи токена, принципы которого были рассмотрены в главе 12 при описании функций уровня MAC. Давайте остановимся более детально на некоторых особенностях этого метода, присущих технологии Token Ring 4 Мбит/с, описанной в стандарте 802.5.

В сети Token Ring любая станция всегда непосредственно получает данные только от одной станции — той, которая является предыдущей в кольце, а передает данные своему ближайшему соседу вниз по потоку данных.

Получив токен, станция анализирует его и при отсутствии у нее данных для передачи продвигает токен к следующей станции. Станция, которая имеет данные для переда-

чи, при получении токена изымает его из кольца, что дает ей право доступа к физической среде для передачи своих данных. Затем эта станция выдает в кольцо кадр данных установленного формата последовательно по битам. Кадр снабжается адресами приемника и источника.

Переданные данные проходят по кольцу всегда в одном направлении от одной станции к другой. Все станции кольца ретранслируют кадр побитно, как повторители. Если кадр проходит через станцию назначения, то распознав свой адрес, эта станция копирует кадр в свой внутренний буфер и вставляет в кадр признак подтверждения приема. Станция, выдавшая кадр данных в кольцо, получив его с подтверждением приема, изымает свой кадр из кольца и передает в сеть новый токен, давая другим станциям сети возможность передавать данные.

На рис. 1 описанный алгоритм доступа к среде иллюстрируется временной диаграммой. Здесь показана передача пакета А в кольце, состоящем из 6 станций, от станции 1 к станции 3. После прохождения станции назначения 3 в пакете А устанавливаются два признака — признак А распознавания адреса и признак С копирования пакета в буфер (что на рисунке отмечено звездочкой внутри пакета). После возвращения пакета в станцию 1 отправитель распознает свой пакет по адресу источника и удаляет пакет из кольца. Установленные станцией 3 признаки говорят станции-отправителю о том, что пакет дошел до адресата и был успешно скопирован в его буфер.

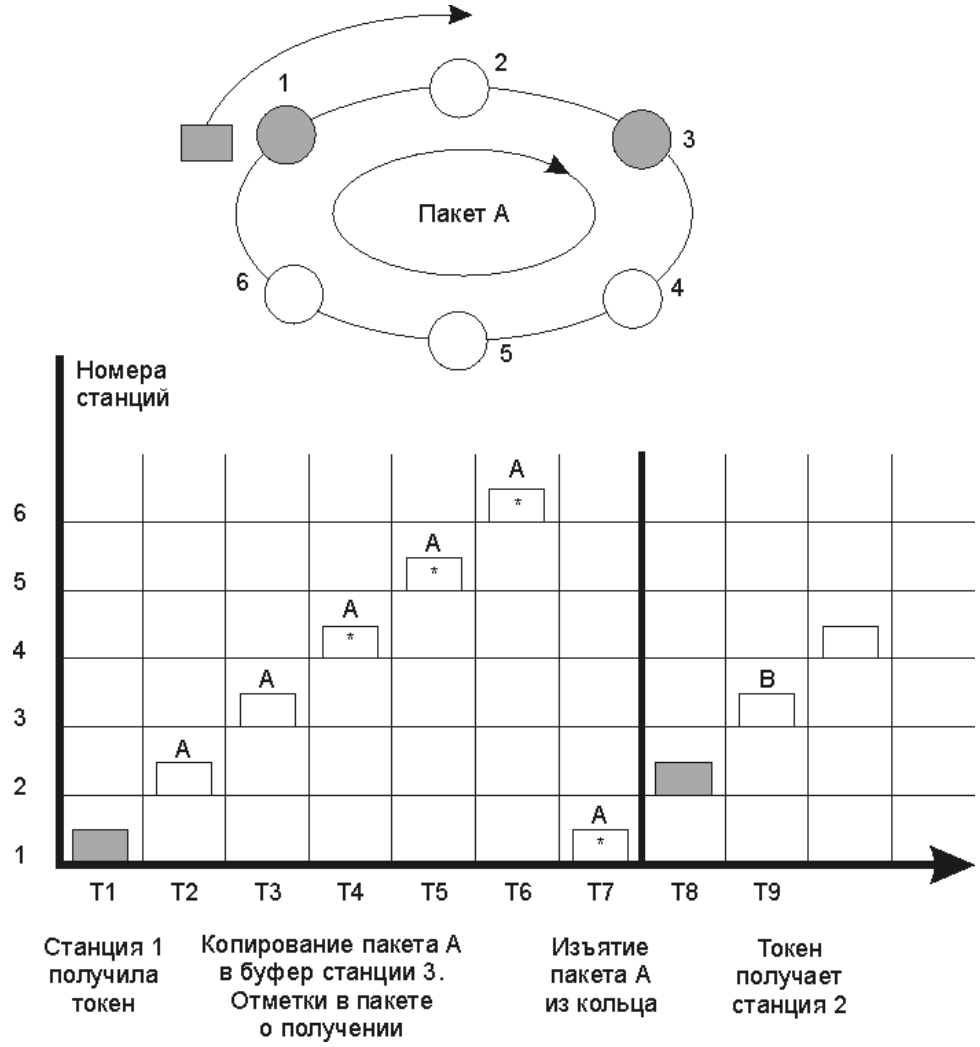

Рис. 1. Доступ с передачей токена

Время владения разделяемой средой в сети Token Ring ограничивается фиксированной величиной, называемой **временем удержания токена**. После истечения этого времени станция обязана прекратить передачу собственных данных (текущий кадр разрешается завершить) и передать токен далее по кольцу. Станция может успеть передать за время удержания токена один или несколько кадров в зависимости от размера кадров и величины времени удержания токена.

Обычно время удержания токена по умолчанию равно 10 мс, а максимальный размер кадра в стандарте 802.5 не определен. Для сетей 4 Мбит/с он, как правило, равен 4 Кбайт, а для сетей 16 Мбит/с — 16 Кбайт. Это связано с тем, что за время удержания токена станция должна успеть передать хотя бы один кадр. При скорости 4 Мбит/с за время 10 мс можно передать 5000 байт, а при скорости 16 Мбит/с — 20 000 байт. Максимальные размеры кадра выбраны с некоторым запасом.

В сетях Token Ring 16 Мбит/с используется модернизированный вариант алгоритма доступа к кольцу, называемый **алгоритмом раннего освобождения токена**. В соответствии с ним станция передает токен доступа следующей станции *сразу же после окончания передачи последнего бита кадра*, не дожидаясь возвращения по кольцу этого кадра с установленными битами A и C. В этом случае пропускная способность кольца используется более эффективно, так как по кольцу одновременно продвигаются кадры нескольких станций. Тем не менее свои кадры в каждый момент времени может генерировать только одна станция — та, которая в данный момент владеет токеном. Остальные станции в это время только повторяют чужие кадры, так что принцип разделения кольца во времени сохраняется, ускоряется только процедура передачи владения кольцом.

Для различных видов сообщений, передаваемых кадрами, могут назначаться разные *приоритеты*: от 0 (низший) до 7 (высший). Решение о приоритете конкретного кадра принимает передающая станция (протокол Token Ring получает этот параметр через межуровневые интерфейсы от протоколов верхних уровней, например прикладного). Токен также всегда имеет некоторый уровень текущего приоритета. Станция имеет право захватить переданный ей токен только в том случае, если приоритет кадра, который она хочет передать, выше приоритета токена (или равен ему). В противном случае станция обязана передать токен следующей по кольцу станции.

За наличие в сети токена, причем единственной его копии, отвечает активный монитор. Если активный монитор не получает токен в течение длительного времени (например, 2,6 с), то он порождает новый токен.

Приоритетный доступ в технологии Token Ring был предназначен для поддержки требований QoS приложений. Однако разработчики приложений для локальных сетей практически им не пользовались.

Физический уровень технологии Token Ring

Стандарт Token Ring фирмы IBM изначально предусматривал построение связей в сети с помощью концентраторов (рис. 2), называемых **устройствами многостанционного доступа** (Multi-station Access Unit, MAU, или MSAU). Сеть Token Ring может включать до 260 узлов. Использование концентраторов приводит к тому, что физически сети Token Ring имеют звездообразную топологию, а логически — кольцевую.

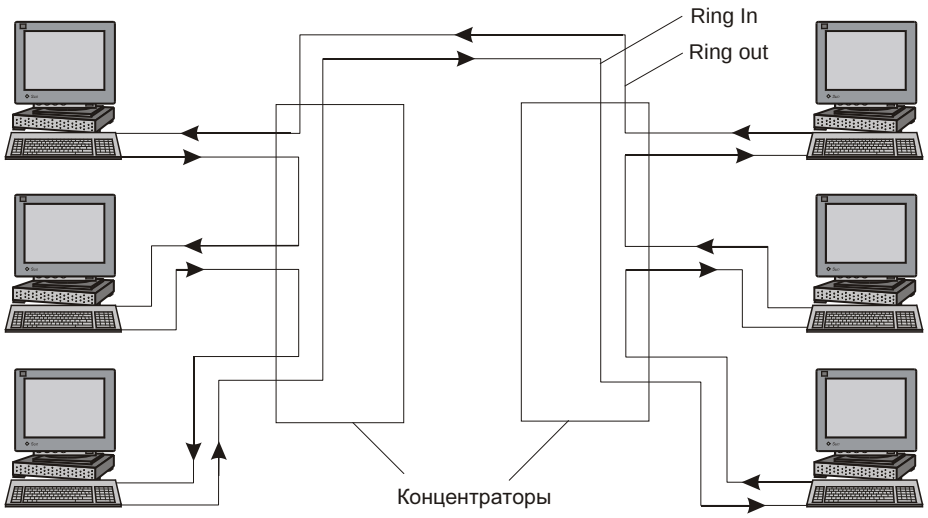

Рис. 2. Физическая конфигурация сети Token Ring

Концентратор Token Ring может быть активным или пассивным. **Пассивный концентратор** просто соединяет порты внутренними связями так, чтобы станции, подключаемые к этим портам, образовали кольцо. Ни усиление сигналов, ни их ресинхронизацию пассивный концентратор не выполняет. Такое MSAU-устройство можно считать простым кроссовым блоком за одним исключением — MSAU обеспечивает обход какого-либо порта, когда присоединенный к этому порту компьютер выключают. Такая функция необходима для обеспечения связности кольца вне зависимости от состояния подключенных компьютеров. Обычно обход порта выполняется за счет релейных схем, которые питаются постоянным током от сетевого адаптера, а при выключении сетевого адаптера нормально замкнутые контакты реле соединяют вход порта с его выходом.

Активный концентратор выполняет функции регенерации сигналов и поэтому его можно назвать повторителем.

Возникает вопрос: если концентратор является пассивным устройством, то каким образом обеспечивается качественная передача сигналов на большие расстояния, например, при включении в сеть нескольких сот компьютеров? Ответ состоит в том, что роль усилителя сигналов в этом случае берет на себя каждый сетевой адаптер, а роль синхронизирующего блока — сетевой адаптер активного монитора кольца. Каждый сетевой адаптер Token Ring имеет блок повторения, который умеет регенерировать и синхронизировать сигналы, однако последнюю функцию выполняет в кольце только блок повторения активного монитора.

В общем случае сеть Token Ring имеет комбинированную звездно-кольцевую конфигурацию. Конечные узлы подключаются к MSAU-устройствам по топологии звезды, а сами концентраторы объединяются через специальные порты Ring In (RI) и Ring Out (RO), образуя магистральное физическое кольцо.

Технология Token Ring позволяет использовать для соединения конечных станций и концентраторов различные типы кабеля: STP типа 1, UTP типа 3, UTP типа 6, а также

волоконно-оптический кабель. При применении экранированной витой пары STP типа 1 из номенклатуры кабельной системы IBM в кольцо допускается объединять до 260 станций при длине ответвительных кабелей до 100 метров, а при применении неэкранированной витой пары максимальное количество станций сокращается до 72 при длине ответвительных кабелей до 45 метров. Расстояние между пассивными концентраторами может достигать 100 м при использовании кабеля STP типа 1 и 45 м при использовании кабеля UTP типа 3. Между активными концентраторами максимальное расстояние увеличивается соответственно до 730 или 365 м в зависимости от типа кабеля.

Максимальная длина кольца Token Ring составляет 4000 м.

ПРИМЕЧАНИЕ

Ограничения на максимальную длину кольца и количество станций в кольце в технологии Token Ring не являются такими жесткими, как в технологии Ethernet. Здесь эти ограничения во многом связаны со временем оборота токена по кольцу. Есть и другие соображения, диктующие выбор ограничений. Так, если кольцо состоит из 260 станций, то при времени удержания токена в 10 мс токен вернется в активный монитор в худшем случае через 2,6 с, а это время как раз составляет тайм-аут контроля оборота токена. В принципе, все значения тайм-аутов в сетевых адаптерах узлов сети Token Ring можно настраивать, поэтому можно построить сеть Token Ring с бо́льшим количеством станций и с бо́льшей длиной кольца.

Коммутируемая сеть завода «Трансмаш» дополнительный материал к «Коммутируемая сеть завода “Трансмаш”» в главе 14

В главе 12 мы рассмотрели структуру локальной сети завода «Трансмаш», построенной на повторителях Ethernet 10 Мбит/с. Тот пример соответствовал типичной для начала 90-х годов ситуации, когда единая разделяемая среда с пропускной способностью 10 Мбит/с полностью удовлетворяла потребности предприятия в обмене трафиком между немногочисленными компьютерами подразделений. Здесь мы опишем модернизированный вариант локальной сети этого завода, который был характерен для многих крупных локальных сетей второй половины 90-х годов.

Основной особенностью этой локальной сети является то, что она *полностью построена на коммутаторах* (рис. 1). Переход на коммутируемую сеть был продиктован резко возросшими в середине 90-х годов требованиями, предъявляемыми к производительности и надежности локальной сети. К этому времени компьютеризованная обработка данных стала на заводе «Трансмаш» одним из основных средств производства, при этом увеличилось число компьютеров и качественно изменились приложения, которые стали передавать мультимедийную информацию больших объемов.

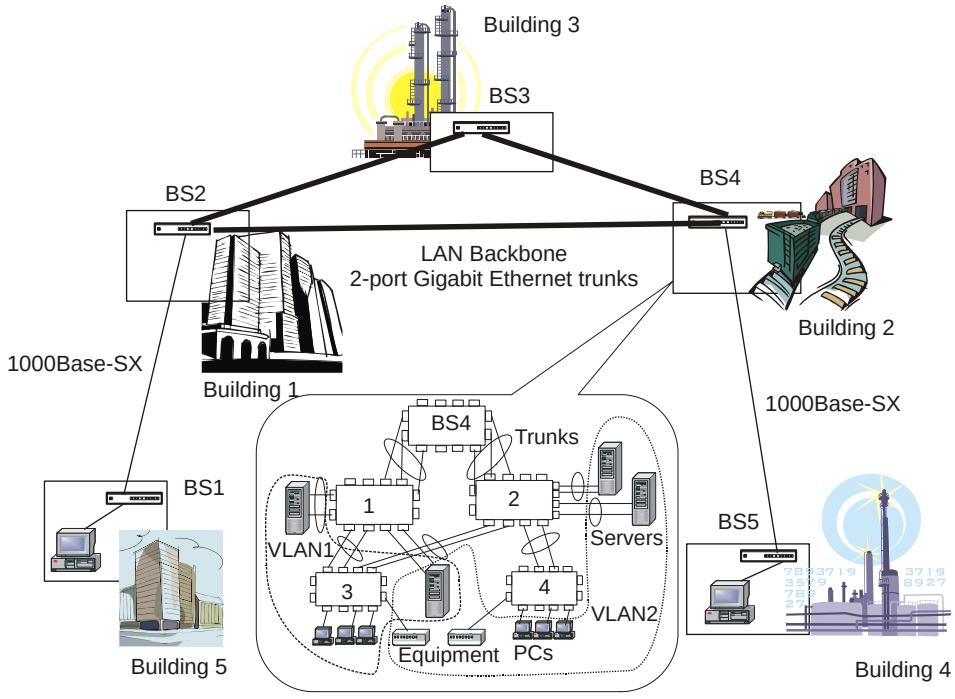

Рис. 1. Коммутируемая сеть завода «Трансмаш»

Основу локальной сети каждого из пяти зданий завода составляет мощный центральный коммутатор на основе шасси, оснащенный портами Fast Ethernet и Gigabit Ethernet (коммутаторы BS1–BS5). Коммутатор здания объединяет коммутаторы этажей, которые подключены к нему транками, состоящими из двух-трех портов Fast Ethernet. Каждый коммутатор этажа служит для подключения пользовательского оборудования двух типов: персональных компьютеров и технологического оборудования (более детально эти соединения показаны на примере сети здания 2). Пользователи персональных компьютеров работают с приложениями автоматизированной системы управления предприятием (АСУП), а технологическое оборудование образует автоматизированную систему управления технологическими процессами (АСУТП).

Центральные коммутаторы зданий 2, 3 и 4 образуют магистраль локальной сети завода. Они объединены двухпортовыми транками Gigabit Ethernet, что обеспечивает большой запас магистрали по производительности. Здания 5 и 4 подключены к магистрали с помощью обычных (без транков) соединений Gigabit Ethernet. Для связи коммутаторов этих зданий задействуется многомодовое оптическое волокно, которое было проложено еще для локальной сети на повторителях и качества которого оказалось достаточно для устойчивой работы портов 1000Base-SX.

В сети «Трансмаш» передается трафик приложений двух типов: АСУП и АСУТП. Эти классы трафика отличаются требованиями к качеству обслуживания, например, трафик АСУТП является трафиком реального времени, а АСУП — нет. Поэтому в сети «Трансмаш» организованы две виртуальные локальные сети — VLAN 1 для трафика АСУП и VLAN 2 для трафика АСУТП. Это позволяет надежно изолировать каждый тип трафика и, кроме того, упрощает поддержку параметров QoS коммутаторами, так как признаком необходимости обработки трафика в приоритетной очереди является номер VLAN, в данном случае — номер 2.

Так как магистраль локальной сети обладает избыточными связями, то коммутаторы используют алгоритм STA, причем отдельно для каждой виртуальной локальной сети. Для VLAN 1 резервной является связь между коммутаторами BS4 и BS2, а для VLAN 2 — между коммутаторами BS4 и BS3. Обмен данными между АСУП и АСУТП происходит через несколько серверов, которые являются членами обеих сетей.

Структурирование адресного пространства группового вещания (дополнительный материал к «Структурирование адресного пространства группового вещания» в главе 18)

Поддиапазон адресов	Назначение	Примеры
224.0.0.0–224.0.0.255	Адреса предназначены для локальных сетей, использующих протокол IPv4. Эти адреса назначаются локально. IP-пакеты, несущие адреса из этого диапазона, ни при каких условиях не будут переданы за пределы сети маршрутизаторами, поддерживающими групповое вещание. Обычно адреса этого типа служат для взаимодействия соседних маршрутизаторов	Список закрепленных адресов из этого диапазона достаточно велик, вот несколько примеров: 224.0.0.1 — все узлы данной локальной сети; 224.0.0.2 — все маршрутизаторы; 224.0.0.9 — все маршрутизаторы, поддерживающие протокол RIPv2
224.0.1.0–224.0.1.255	Адреса предназначены для межсетевого управления. IP-пакеты, несущие адреса из этого диапазона, продвигаются маршрутизаторами группового вещания от одной сети к другой (конечно, если это позволяет значение TTL в заголовке пакета)	В этом диапазоне также много закрепленных адресов, в качестве примера приведем адрес 224.0.1.12, традиционно используемый для передачи видеорепортажей о мероприятиях IETF
224.0.2.0–224.0.255.0	Адреса используются в качестве назначенных адресов для сервисов группового вещания. В настоящее время разработчикам приложений и сервисов рекомендуется не прибегать к применению специальных адресов и статическому конфигурированию приложений на их основе, а задействовать адреса из стандартного диапазона	Примеры: 224.0.18.0–224.0.18.255 — групповой адрес, используемый компанией Dow Jones для распространения своей информации о состоянии рынков акций; 224.0.19.0–224.0.19.63 — групповой адрес, зарезервированный за компанией Walt Disney
224.1.0.0–224.1.255.255	Не используются	
224.2.0.0–224.2.255.255	Адреса выделены для протоколов SAP (Session Announcement Protocol) и SDP (Session Description Protocl). Подавляющее часть группового вещания в начальный период	Диапазон 224.2.0.0–224.2.127.253 выделен для аудио- и видеоконференций

	была связана с аудио- и видеоосвещением различных событий, а также с видеоконференциями
224.3.0.0– 231.255.255.255	Адреса зарезервированы, в частности, для использования министерством обороны США
232.0.0.0– 232.255.255.255	Адреса зарезервированы исключительно для использования в рамках технологий, основанных на модели SSM для идентификации узлов-источников
233.0.0.0– 233.251.255.255	Согласно RFC 3180, эти адреса распределяются между автономными системами (AS). Каждая система получает 256 адресов из этого диапазона, а именно те адреса, в которых два средних байта соответствуют ее номеру. Оставшийся байт каждого адреса автономная система может использовать по своему усмотрению
233.252.0.0– 233.255.255.255	Согласно RFC 3138, эти адреса выделены для частных автономных систем. Некоторое количество автономных систем с номерами из диапазона 64512–65535 определены для частного использования, например для создания внутренних административных доменов, которые всей оставшейся частью Интернета воспринимаются как единый домен. Адреса из данного диапазона не разделяются между этими автономными системами, а относятся к ним как к целому
239.0.0.0– 239.255.255.255	Согласно RFC 2365, это административно ограниченные адреса. Пакеты, которые несут адреса из этого блока, ограничены в передвижении по составной сети, то есть некоторые маршрутизаторы не будут продвигать их за границы некоторых административных доменов. Используя такие адреса, источник вещания может быть уверен, что его пакеты не «уплывут» в Интернет, а останутся под его контролем

Междоменное групповое вещание (дополнительный материал к «Междоменное групповое вещание» в главе 18)

По мере роста магистраль MBone сталкивалась со все возрастающим числом проблем. Общей причиной проблем была плоская (не иерархическая) виртуальная топология MBone. Те же проблемы, которые были характерны для основанного на классах механизма маршрутизации уникальных адресов, проявились и в MBone. При плоской топологии сетевые маршруты должны быть известны каждому маршрутизатору, а так как во время своего расцвета MBone включала почти 10 000 маршрутов, то объемы маршрутных данных приблизились к той черте, за которой маршрутизаторы становились практически неуправляемыми.

Большинство из этих маршрутов использовалось крайне неэффективно — записи имели длинные префиксы (между 28 и 32), а это означало, что каждая запись в таблице маршрутизации представляла всего несколько хостов.

Кроме того, при плоской топологии ошибки одного маршрутизатора распространялись по всей сети, вызывая нестабильную работу Mbone в целом.

Опыт маршрутизации уникальных адресов говорил о необходимости применения в групповом вещании механизма агрегирования маршрутов и создания на этой основе иерархической топологии. Действительно, когда-то развитие обычной (использующей уникальные адреса) маршрутизации привело к разделению Интернета на автономные системы (AS), представляющие собой домены маршрутизации. Как отмечено в главе 17, каждой автономной системой управляет одна организация, которая вольна организовывать маршрутизацию внутри этой системы так, как считает нужным, используя для этого протокол RIP, OSPF, IGRP или статические записи в таблицах маршрутизации. По умолчанию считается, что между автономными системами нет отношений доверия, поэтому маршрутная информация через границы AS передается под жестким надзором администраторов, держащих под контролем передачу через свой домен транзитной информации чужих доменов.

Так как структура автономных систем Интернета уже сложилась, при построении иерархической топологии групповой маршрутизации не требовалось создавать новую структуру, достаточно было разработать механизм, способный работать с существующей структурой, основанной на доменах маршрутизации.

Задача создания протоколов междоменной групповой маршрутизации была поставлена сетевым сообществом в 1997 году. На сегодняшний день существуют две группы решений. Первая группа — это так называемые тактические решения, которые могут работать уже сегодня, но не обладают достаточной масштабируемостью, чтобы стать основой развития Интернета на значительную перспективу. Поэтому продолжается работа по поиску долговременных стратегических решений, составляющих вторую группу. Стратегические предложения базируются как на стандартной модели группового вещания протокола IP, так и на более радикальных новых подходах.

Протоколы PIM-SM и BGP в многодоменной сети группового вещания

Примером тактического решения является создание средств маршрутизации группового трафика в многодоменной сети путем расширения функциональных возможностей широко используемого *протокола маршрутизации BGP*. Действительно, поскольку ставится задача разработки протокола групповой маршрутизации между автономными системами, то естественно в первую очередь обратить внимание на протокол, который уже долгое время успешно служит для маршрутизации трафика с индивидуальными адресами между автономными системами, то есть на протокол BGP (см. главу 17).

Механизм, с помощью которого протокол BGP был доработан для поддержки групповых маршрутов, представляет собой расширение протокола BGPv4. Это расширение, описанное в спецификации RFC 2283, входит в состав протокола **BGP4+** (Multiprotocol Extensions for BGPv4 — мультипротокольные расширения для BGPv4). Когда протокол BGP4+ используется для поддержки групповых маршрутов, его часто называют **MBGP** (Multicast Border Gateway Protocol). Протокол MBGP выполняет достаточно простую функцию для системы группового вещания: он узнает о путях, с помощью которых групповой трафик может достичь других доменов. На рис. 1 показан пример соединения нескольких доменов в сеансах MBGP. В одном случае два домена, связанные вместе, используют разные соединения для индивидуального и группового трафиков, в других случаях — общие соединения для передачи обоих типов трафика.

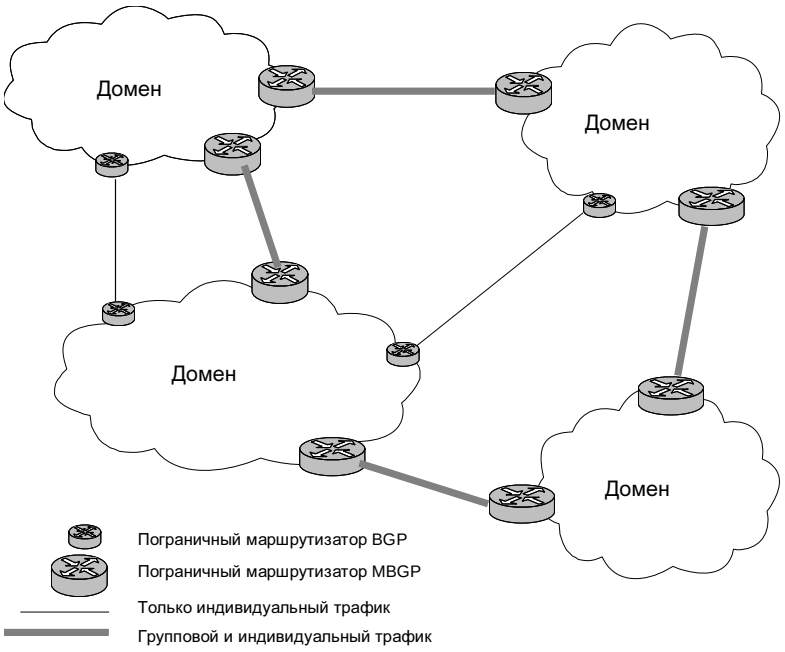

Рис. 1. Пример междоменной топологии группового вещания на базе протоколов BGP и/или MBGP

Разработка протокола MBGP — это только первый шаг на пути реализации группового вещания, и его использование не обеспечивает решение проблемы полностью. Протокол MBGP может выбрать следующий шаг передачи пакетов с индивидуальными и групповыми адресами, но он не может выполнить более сложную функцию — построить дерево группового вещания. Здравый смысл подсказывает, что для этих целей может быть применен только протокол разряженного режима.

Таким тактическим решением стало использование протокола *PIM-SM на междоменном уровне*, когда домены рассматриваются в качестве узлов сети, а разделяемое дерево строится для тех доменов, в которых содержатся члены определенной группы. Однако полностью повторить на этом уровне протокол PIM-SM не удается из-за междоменных барьеров.

Протокол MSDP

Протокол обнаружения источников группового вещания (Multicast Source Discovery Protocol, MSDP), описанный в спецификации, RFC 3618, позволяет решить еще одну проблему, возникающую при соединении доменов, на которых работают протоколы «разреженного» режима. Проблема в том, как оповестить членов групп в одном домене, что в других доменах есть источники группового вещания для этих групп.

Стандартное решение протокола PIM-SM состоит в отправке источником точке встречи *регистрационного сообщения*, в ответ на которое RP отсылает источнику *сообщение о присоединении*, означающее, что источник присоединяется к разделяемому дереву группы. При междоменном взаимодействии возникает вопрос: в каком домене нужно размещать точку встречи? При существовании единственной точки встречи весьма вероятно возникновение так называемой «зависимости от третьей стороны», когда все источники и приемники группы расположены в одном домене, а точка встречи этой группы — в другом. Такая ситуация может привести к организационным и коммерческим конфликтам.

- Поставщик услуг автономной системы с источниками и приемниками группы должен полагаться на другого провайдера — своего потенциального конкурента, в домене которого размещена точка встречи.
- Поставщик услуг автономной системы, включающей точку встречи, обслуживает трафик группы, для которой у него нет ни источников, ни получателей. В большинстве случаев при отсутствии членов группы нет и финансовой мотивации для обслуживания трафика конкурента.

Более приемлемым решением может показаться дублирование точек встречи в каждом домене. Однако это ведет к другой проблеме — проблеме организации взаимодействия между этими точками, которой не было при работе протокола PIM-SM в пределах одного домена. Именно эту задачу решает *протокол MSDP*. Протокол MSDP работает в каждом домене на том же маршрутизаторе, что и точка встречи домена. Он выполняет функции представителя домена, объявляющего другим доменам о существовании активных источников групповых данных. Все протокольные модули MSDP, установленные на маршрутизаторах, принадлежащих разным доменам, связаны TCP-соединениями. Алгоритм работы протокола MSDP иллюстрирует рис. 2.

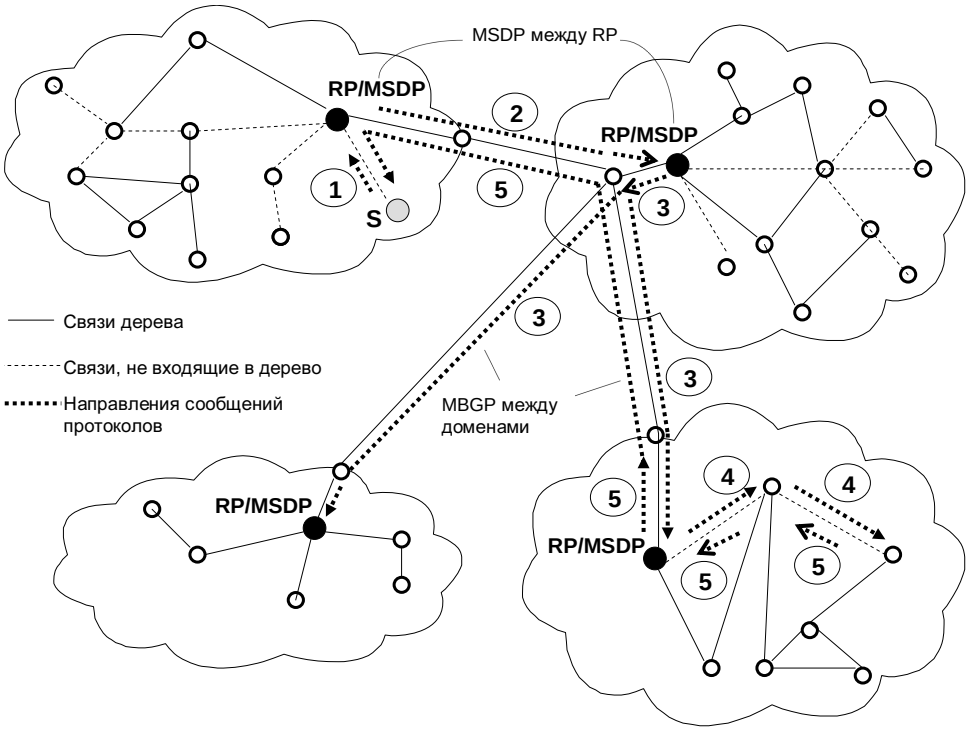

Рис. 2. Схема работы протокола MSDP

1. Источник генерирует MSDP-сообщение.
2. MSDP-партнеры, которые получают сообщение, обеспечивают проверку продвижения по реверсивному пути (RPF), то есть проверяют, находится ли пославший это сообщение партнер на «корректном» пути. Когда новый источник для группы становится активным, он регистрируется в соответствующей точке встречи домена.
3. Протокол MSDP, работающий в данном домене, обнаруживает возникновение нового активного источника и посылает сообщение об активности источника (Source Active, SA) всем узлам MSDP, расположенным в других доменах. После этого данное SA-сообщение распространяется периодически. Если MSDP-партнер получает SA-сообщение через корректный интерфейс, то оно передается всем MSDP-партнерам за исключением того, от которого это сообщение получено.
4. Внутри домена MSDP-партнер, который является точкой встречи, проверяет, есть ли у него члены группы в домене.
5. Если члены группы в домене имеются, то RP посылает сообщение протокола PIM о присоединении по адресу источника, объявленному в SA-сообщении (что и было бы сделано точкой встречи в том случае, если бы источник находился в ее домене). Сообщение о присоединении проходит реверсивный путь до источника, используя междоменные пути, определяемые с помощью протокола MBGP. Как

только реверсивный путь проложен, точки встречи начинают продвигать групповые пакеты, в том числе между доменами.

Повторение шагов 3 и 4 происходит до тех пор, пока все MSDP-партнеры не получат SA-сообщение и все члены группы во всех доменах не начнут получать данные от источника.

Ограничения и проблемы протоколов PIM-SM/MBGP/MSDP

В ответ на критику, утверждающую, что текущий набор протоколов не отличается простотой, сторонники решения PIM-SM/MBGP/MSDP возражают, что он не более сложен, чем многие решения масштабе всего Интернета. Основное преимущество решения PIM-SM/MBGP/MSDP состоит в том, что оно работает и уже успешно применяется. Основным недостатком является то, что при частых изменениях в составе групп и активности источников масштабируемость этого решения вызывает сомнения.

Когда группы являются динамичными — либо по причине частых изменений в активности источников (пульсирующие источники), либо из-за частых случаев присоединения/отсоединения членов групп — накладные расходы на управление группой могут быть значительными.

У сети возникает трудная задача создания и удаления информации о состояниях тысяч приемников и передатчиков, разбросанных по всему миру.

Кроме того, еще одной специфической проблемой, с которой сталкивается протокол MSDP, является проблема задержки присоединения. Так как SA-сообщение распространяется периодически, то между присоединением новых приемников и получением ими следующего SA-сообщения может возникать значительная задержка. Для решения этой проблемы MSDP-партнеры могут кэшировать SA-сообщение в надежде то, что в присоединения нового приемника источник все еще будет активен. Если MSDP-партнер кэширует SA-сообщение, то другие MSDP-партнеры могут воспользоваться кэшируемой информацией. Некэширующий узел MSDP может послать SA-сообщение запроса узлу MSDP, который выполняет кэширование. К сожалению, минимизация задержки присоединения за счет кэширования приводит к чрезмерному увеличению объема хранимой информации состояния.

Другая проблема вызывается пульсирующими источниками. После проявления активности новым источником проходит значительное время, связанное с установлением между доменами дерева продвижения. Поэтому несколько первых пакетов, передаваемых источником, часто теряются. Решение, предусмотренное в протоколе MSDP для устранения этого недостатка, состоит в том, что перенос первых пакетов данных может осуществляться служебными SA-сообщениями. Это не очень элегантное решение, но оно работает.

Вопрос масштабируемости достаточно важен для MSDP. Принимая во внимание способ работы MSDP, при существовании в сети нескольких тысяч групп накладные расходы на функционирование MSDP становятся слишком большими: в сети постоянно будет циркулировать большое количество SA-сообщений, содержащих группо-

вые данные. Общее мнение состоит в том, что учитывая недостаточную масштабируемость протокола MSDP, его вряд ли можно отнести к стратегическим решениям.

Протоколы BGMP и MASC

Первым реально *долговременным* предложением по организации междоменного группового вещания в масштабах Интернета стал протокол **BGMP** (Border Gateway Multicast Protocol — пограничный шлюзовой протокол группового вещания). Этот протокол относится к группе, сохраняющей стандартную идеологию группового вещания, предложенную Дирингом. Ключевой идеей BGMP является конструирование двунаправленных разделяемых деревьев между доменами при наличии только одной точки входа. Одной из функций BGMP является принятие решения о том, что какой домен будет служить корнем разделяемого дерева. Предлагаемое решение свободно от проблемы «зависимости от третьей стороны», так как протокол основан на более строгой схеме распределения адресов.

Вообще говоря, распределение групповых адресов стало серьезной проблемой для коммерческих пользователей группового вещания. Протокол BGMP учитывает эту проблему, требуя, чтобы групповые адреса были связаны с определенным доменом. Архитектура BGMP включает собственную схему распределения адресов под названием **MASC** (Multicast Address-Set Claime), однако может работать и с любой другой, лишь бы она обеспечивала связь групповых адресов с определенной автономной системой.

Помимо описанного требования существует еще одно — необходимо избегать групповых коллизий. Коллизия возникает, когда две группы используют один и тот же групповой адрес, и трафик каждой группы доставляется членам обеих групп. Эффект коллизии групп может проявляться по-разному: от простого неудобства до серьезного нарушения работы сети. Применяемая сегодня схема назначения групповых адресов (имеющих локальное значение) является неформальной: пользователь может просто взять адрес и применить его, то есть вероятность коллизии не нулевая. Именно поэтому процедура распределения групповых адресов приобрела особое значение.

Протокол MASC представляет собой одно из возможных решений и является частью более общей схемы адресации, называемой архитектурой распределения групповых адресов. Имеется три уровня распределения адресов:

- между доменами адреса распределяются по схеме MASC;
- внутри домена распределением адресов занимается протокол распределения адресов (Address Allocation Protocol, AAP);
- протокол MADCAP (Multicast Address Dynamic Client Allocation Protocol) используется хостами для запроса адресов у сервера распределения групповых адресов (Multicast Address Allocation Server, MAAS).

Таким образом, MASC и другие вспомогательные протоколы выполняют функции *динамического* распределения групповых адресов, необходимые протоколу BGMP. Однако этот подход не является единственно возможным.

Другой подход состоит в *статическом* распределении групповых адресов. Например, в соответствии с процедурой распределения адресов GLOP каждая автономная система получает фиксированное количество адресов, которые называются GLOP-

адресами и в которые как часть входит номер автономной системы. Это предложение становится довольно популярным, но оно имеет уязвимость. В текущем варианте только 8 бит, или 256 групповых адресов, доступны для автономной системы, что во многих случаях явно недостаточно. Эта проблема может быть решена переходом на систему адресации протокола IPv6.

Технология MPLS L3VPN (дополнительный материал к «MPLS VPN третьего уровня» в главе 20)

Сети MPLS VPN привлекают сегодня всеобщее внимание. Количество ведущих поставщиков услуг, предлагающих своим клиентам воспользоваться новым видом сервиса для экономичного построения своих внутренних и внешних сетей, постоянно растет, делая сети MPLS VPN доступными для пользователей все большего числа стран и регионов. От других технологий построения виртуальных частных сетей, таких как VPN на базе ATM/FR или IPSec, технологию MPLS VPN выгодно отличает хорошая масштабируемость, возможность автоматического конфигурирования и естественная интеграция с другими сервисами протокола IP, которые сегодня входят в обязательный набор любого успешного поставщика услуг, включая доступ в Интернет, WWW и почтовые службы, хостинг.

Существует два варианта сетей MPLS VPN.

- В сетях **MPLS L3VPN** доставка трафика от клиента до пограничного устройства сети поставщика услуг осуществляется с помощью технологии IP (третий уровень).
- Сети **MPLS L2VPN** передают клиентский трафик в сеть поставщика услуг с помощью какой-либо технологией второго уровня, которой может быть Ethernet, Frame Relay или ATM.

В обоих случаях внутри сети поставщика услуг клиентский трафик передается с помощью технологии MPLS.

ПРИМЕЧАНИЕ

> Сегодня уровень MPLS определить не так просто — терминология в этой области еще не устоялась. Но поскольку продвижение пакетов на основе локальных меток соответствует второму уровню, мы будем относить MPLS ко второму уровню.

В данной книге рассматривается только технология MPLS L3VPN, поскольку это намного более зрелая технология, уже работающая во многих сетях поставщиков услуг. Несмотря на то, что спецификация RFC 2547bis, которая определяет ее основные механизмы, носит информативный статус, все реализации MPLS L3VPN производителями сетевого оборудования следуют этому документу, придавая ему статус фактического стандарта. В дальнейшем изложении мы для краткости будем опускать обозначение уровня L3, используя название MPLS VPN как синоним MPLS L3VPN.

Полная связность при абсолютной изолированности

Каждый клиент желает, чтобы поставщик услуг VPN связал между собой его сети, обеспечив абсолютную изолированность полученной единой сети от сетей других клиентов.

Эту задачу современному поставщику услуг приходится решать в противоречивых условиях доминирования технологии IP как универсального транспорта. Действительно, один из основных принципов работы составной IP-сети заключается в автоматическом связывании всех сетей в одно целое за счет распространения по сети маршрутной информации протоколами маршрутизации, такими как BGP, OSPF, IS-IS, RIP. С помощью подобного механизма на каждом маршрутизаторе сети автоматически создается таблица маршрутизации, в которой указываются пути следования пакетов к каждой из сетей, включенных в составную сеть (пути к отдельным сетям могут агрегироваться, но это не меняет сути).

Как же технология MPLS VPN разрешает парадокс обеспечения изолированности при сохранении связности? Достаточно элегантно — за счет автоматической фильтрации маршрутных объявлений и применения туннелей MPLS для передачи клиентского трафика по внутренней сети поставщика.

Для того чтобы изолировать сети друг от друга, достаточно поставить между ними заслон на пути распространения маршрутной информации. Для обмена маршрутной информацией в пределах сети узлы пользуются одним из внутренних протоколов маршрутизации (IGP), таким как RIP, OSPF или IS-IS, область действия которого ограничена автономной системой. Если в таблице маршрутизации узла A нет записи о маршруте к узлу B (и отсутствует запись о маршруте по умолчанию), то говорят, что узел A не «видит» узла B.

В сети MPLS VPN подобный режим достигается за счет того, что маршрутные объявления, передаваемые сетью клиента, с помощью протокола BGP «перепрыгивают» через всю внутреннюю сеть поставщика услуг. После чего благодаря особому конфигурированию с использованием многопротокольного расширения протокола BGP (MP-BGP) они попадают только в сети того же клиента. В результате маршрутизаторы разных клиентов не имеют маршрутной информации друг о друге и поэтому не могут обмениваться пакетами, то есть достигается желаемая изоляция (рис. 1).

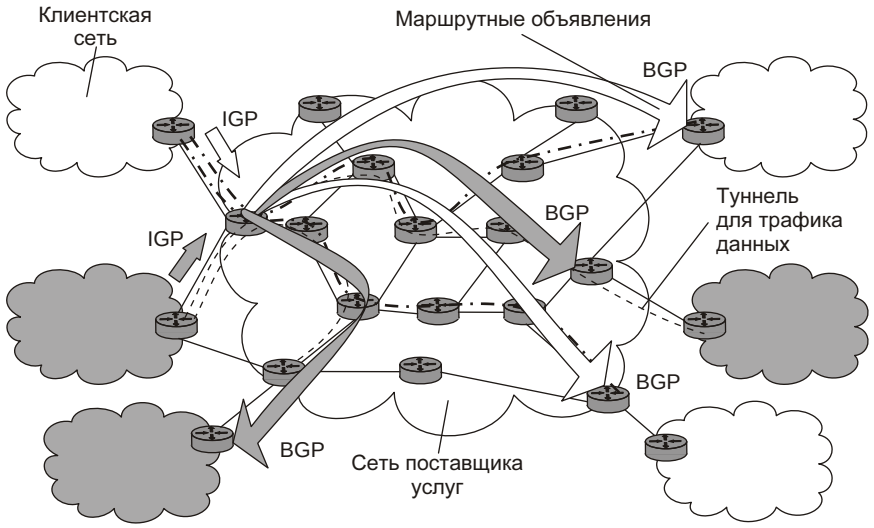

Рис. 1. Изоляция клиентских сетей с помощью туннелей

Еще одним следствием такого подхода является изолированность внутренней сети поставщика услуг от сетей клиентов, а это, в свою очередь, повышает надежность работы сети поставщика и ее масштабируемость (не нужно хранить таблицы большого размера с описанием сетей многочисленных клиентов на внутренних маршрутизаторах сети поставщика услуг).

Но как же все-таки связать территориально разнесенные сети клиента в единую виртуальную частную сеть, если внутренняя сеть поставщика услуг ничего о них не знает, во всяком случае, на уровне обычных таблиц маршрутизации? Для этого применяется достаточно традиционное средство — туннель между пограничными маршрутизаторами внутренней сети. Особенность рассматриваемой технологии состоит в применение туннеля MPLS (альтернативные решения могли бы основываться на туннелях IPSec или других туннелях класса «IP поверх IP»). Преимуществом туннелей MPLS VPN являются автоматический способ их прокладки и выгоды, получаемые за счет применения технологии MPLS как таковой и касающиеся ускоренного продвижения пакетов по сети поставщика услуг и управления качеством обслуживания (QoS) для туннелей с инжинирингом трафика.

Для того чтобы описанные принципы построения MPLS VPN смогли найти воплощение в реальной сети, было разработано несколько специфических сетевых механизмов и компонентов.

Компоненты сети MPLS VPN

В сети MPLS VPN легко выделить две области (рис. 2):

- IP-сети клиентов;
- магистральная сеть MPLS поставщика услуг.

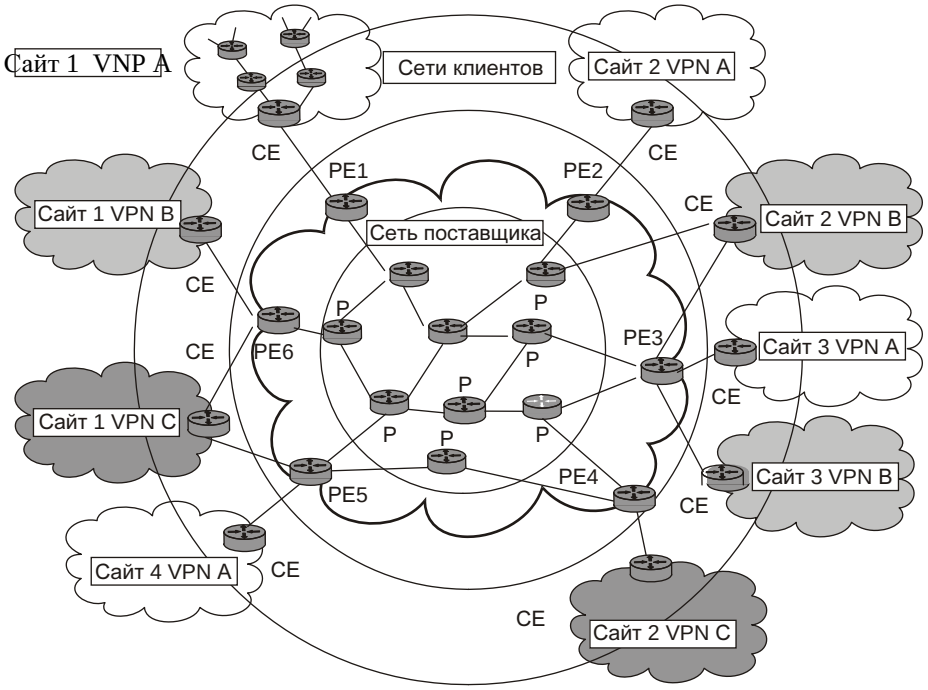

Рис. 2. Компоненты сети MPLS VPN

В общем случае у любого клиента может быть несколько территориально обособленных IP-сетей (сайтов), каждая из которых в свою очередь может включать несколько подсетей, связанных маршрутизаторами. Принадлежащие одному клиенту сайты обмениваются IP-пакетами через сеть поставщика услуг и образуют виртуальную частную сеть этого клиента.

Маршрутизатор, с помощью которого сайт клиента подключается к магистрали поставщика, называется пограничным устройством клиента (Customer Edge router, CE). Будучи компонентом сети клиента, маршрутизатор CE ничего не знает о существовании VPN. Он может быть соединен с магистральной сетью поставщика услуг несколькими каналами.

Магистральная сеть поставщика услуг является сетью MPLS, в которой IP-пакеты продвигаются на основе не IP-адресов, а локальных меток. Сеть MPLS состоит из коммутирующих по меткам маршрутизаторов (LSR), которые направляют трафик по предварительно проложенным путям коммутации по меткам (LSP) в соответствии со значениями меток.

В сети поставщика среди устройств LSR выделяют **пограничные маршрутизаторы** (Provider Edge router, PE), к которым через маршрутизаторы CE подключаются сайты клиентов, и **маршрутизаторы магистральной сети поставщика** (Provider router, P).

Маршрутизаторы CE и PE обычно связаны непосредственно физическим каналом, на котором работает какой-либо протокол канального уровня, например, PPP, FR, ATM или Ethernet. Общение между CE и PE идет по стандартным протоколам стека TCP/IP. Поддержка MPLS нужна только для внутренних интерфейсов PE и всех ин-

терфейсов P. Иногда полезно различать относительно направления продвижения трафика *входной и выходной* (*удаленный*) *маршрутизаторы PE*.

В магистральной сети поставщика только маршрутизаторы PE должны быть сконфигурированы для поддержки существующих виртуальных частных сетей, только они «знают» о них.

Если рассматривать сеть с позиций VPN, то маршрутизаторы P непосредственно не взаимодействуют с маршрутизаторами CE, а просто обеспечивают туннели между входным и выходным маршрутизаторами PE.

Пограничные маршрутизаторы PE являются функционально более сложными, чем внутренние маршрутизаторы P сети поставщика услуг. На них возлагаются главные задачи по поддержке сетей VPN, а именно — задачи разграничения маршрутов и данных, поступающих от разных клиентов. Маршрутизаторы PE служат также оконечными точками путей LSP между сайтами заказчиков, и именно пограничный маршрутизатор поставщика услуг назначает метку IP-пакету для его транзита через внутреннюю сеть, образованную внутренними маршрутизаторами поставщика услуг.

Пути LSP могут быть проложены двумя способами: либо с применением технологии ускоренной маршрутизации (IGP) с помощью протокола LDP, либо на основе технологии инжиниринга трафика с помощью протокола RSVP или CR-LDP. Прокладка LSP означает создание *таблиц коммутации по меткам* на всех пограничных и внутренних маршрутизаторах поставщика услуг, образующих данный путь (примеры таких таблиц можно найти в главе 20). В совокупности эти таблицы задают множество путей, образующих сети различных топологий для разных видов трафика клиентов.

Разграничение маршрутной информации

Для корректной работы VPN требуется, чтобы информация о маршрутах через магистральную сеть поставщика услуг не распространялась за ее пределы, а сведения о маршрутах в клиентских сайтах не становились известными за границами определенных сетей VPN.

Барьеры на пути распространения маршрутных объявлений могут устанавливаться соответствующим конфигурированием маршрутизаторов. Протокол маршрутизации должен быть оповещен о том, с каких интерфейсов и от кого он имеет право принимать объявления и на какие интерфейсы и кому их распространять.

Роль таких барьеров в сети MPLS VPN играют маршрутизаторы PE. Можно представить, что через маршрутизатор PE проходит невидимая граница между зоной клиентских сайтов и зоной ядра сети поставщика. По одну сторону располагаются интерфейсы, через которые PE взаимодействует с внутренними маршрутизаторами поставщика услуг, по другую — интерфейсы, к которым подключаются сайты клиентов. С одной стороны на PE поступают объявления о маршрутах магистральной сети, с другой стороны — объявления о маршрутах в сетях клиентов.

На рис. 3 показана схема разграничения маршрутной информации. На маршрутизаторе PE установлено несколько протокольных модулей IGP. Один из них сконфигурирован для приема и распространения маршрутных объявлений только с тех трех внутренних интерфейсов, которые связывают этот маршрутизатор PE с маршрутиза-

торами P. Два других модуля IGP обрабатывают маршрутную информацию от сайтов клиентов.

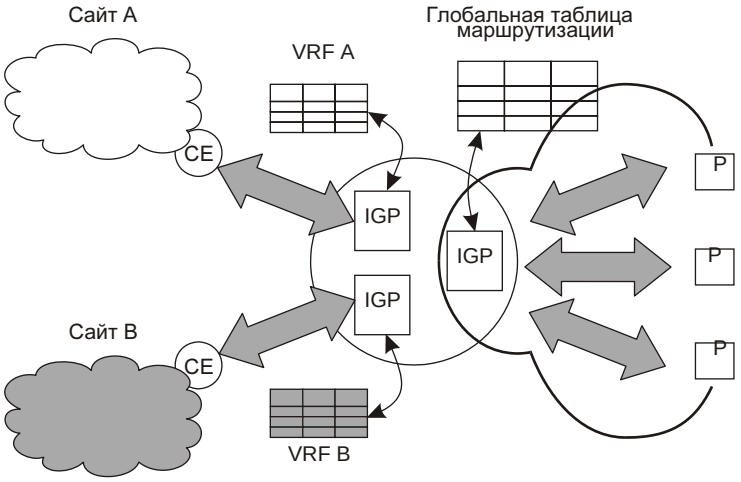

Рис. 3. Схема разделения маршрутной информации

Аналогичным образом настроены и остальные устройства PE. Внутренние маршрутизаторы P принимают и обрабатывают маршрутную информацию протокола IGP, поступающую со всех интерфейсов. В результате на всех маршрутизаторах (и PE, и P) создается по таблице маршрутизации, где содержатся все маршруты в пределах *внутренней сети поставщика услуг*. Подчеркнем, что никакой информации о маршрутах к *сетям клиентов* в таблицах внутренних маршрутизаторов нет. Сети клиентов также ничего не «знают» о маршрутах в сети поставщика услуг.

На каждом из маршрутизаторов PE создается два типа таблиц маршрутизации:

- **глобальная таблица маршрутизации** строится на основе объявлений из *магистральной сети поставщика услуг*;
- **таблицы маршрутизации и продвижения сети VPN** (VPN Routing and Forwarding instance, VRF) маршрутизатор PE формирует на основе объявлений, поступающих *из сайтов клиентов*.

Сайты клиентов представляют собой обычные IP-сети, маршрутная информация в которых может передаваться и обрабатываться с помощью любого протокола маршрутизации класса IGP. Очевидно, что этот процесс никак не регламентируется поставщиком. Маршрутные объявления свободно распространяются между узлами в пределах каждого сайта до тех пор, пока не доходят до пограничных маршрутизаторов PE, служащих преградой для их дальнейшего распространения.

Разграничение маршрутов разных клиентов обеспечивается путем установки отдельного протокола маршрутизации на каждый интерфейс маршрутизатора PE, к которому подключен сайт клиента. Этот протокол принимает и передает клиентские маршрутные объявления только с одного определенного для него интерфейса, не пересылая их ни на внутренние интерфейсы, через которые пограничный маршрутизатор связан с внутренними маршрутизаторами, ни на интерфейсы, к которым подключены

сайты других клиентов. В результате на маршрутизаторе PE создается несколько таблиц VRF.

Несколько упрощая, можно считать, что на каждом маршрутизаторе PE создается столько таблиц VRF, сколько сайтов к нему подключено. Фактически, на маршрутизаторе PE организуется несколько **виртуальных маршрутизаторов**, каждый из которых работает со своей таблицей VRF.

Возможно и другое соотношение между сайтами и таблицами VRF. Например, если к некоторому пограничному маршрутизатору подключено несколько сайтов одной и той же сети VPN, то для них может быть создана одна общая таблица VRF. На рис. 3 показаны две таблицы VRF, одна из которых содержит описание маршрутов к узлам сайта A (VRF A), а другая — к узлам сайта B (VRF B).

Использование протокола MP-BGP для связывания сайтов

Чтобы связать территориально разнесенные сайты заказчика в единую сеть, необходимо, во-первых, создать для них общее пространство распространения маршрутной информации, во-вторых, проложить во внутренней сети пути, по которым принадлежащие разным сайтам узлы одной и той же сети VPN могли бы вести защищенный обмен данными.

Механизмом, с помощью которого сайты, принадлежащие одной и той же сети VPN, обмениваются маршрутной информацией, является уже упоминавшееся многопротокольное расширение для протокола BGP (MultiProtocol extensions for BGP, MP-BGP). Подробное описание этого протокола можно найти в спецификации RFC 2858. С его помощью пограничные маршрутизаторы организуют сеансы связи, в рамках которых обмениваются маршрутной информацией из своих таблиц VRF.

Особенность протокола BGP и его расширений заключается в том, что он получает и передает свои маршрутные объявления не всем непосредственно связанным с ним маршрутизаторам, как протоколы IGP, а только тем, которые указаны в его конфигурационных параметрах в качестве соседей. Причем соседями могут быть «назначены» маршрутизаторы, находящиеся на расстоянии многих хопов. Маршрутизатор PE сконфигурирован так, что все получаемые от клиентских сайтов маршрутные объявления он с помощью MP-BGP пересылает только определенным в качестве соседей другим пограничным маршрутизаторам PE. Целенаправленное распространение маршрутов между маршрутизаторами PE обеспечивается надлежащим выбором атрибутов протокола MP-BGP (эти атрибуты описаны в документе «BGP Extended Communities Attribute», имеющем пока статус проекта стандарта Интернета).

Вопрос о том, кому отправлять маршрутные объявления, а кому нет, целиком зависит от топологии виртуальных частных сетей, поддерживаемых данным поставщиком услуг. Так, на рис. 2 маршрутизатор PE1 передает маршруты из таблицы VRF сайта 1, относящегося к сети VPN A, маршрутизаторам PE2, PE3, PE5, к которым подключены остальные сайты 2, 3 и 4 той же сети VPN A. Полученные маршруты заносятся в таблицы VRF соответствующих сайтов.

Итак, помимо маршрутов, поступающих от непосредственно подсоединенных к устройству PE сайтов, каждая таблица VRF дополняется маршрутами, получаемыми от

других сайтов данной сети VPN по протоколу MP-BGP. Таким путем создаются таблицы, описывающие маршруты в рамках отдельной сети VPN.

Независимость адресных пространств

Если некоторое множество узлов никогда ни при каких условиях не получает маршрутную информацию от другого множества узлов, то адресация узлов в пределах каждого из этих множеств может выполняться независимо.

Ограничение области распространения маршрутной информации пределами отдельных сетей VPN изолирует адресные пространства каждой сети VPN, позволяя применять в ее пределах как публичные адреса Интернета, так и частные адреса, зарезервированные в соответствии со спецификацией RFC 1819.

Почему же в таком случае не сделать выбор адресов в пределах VPN совершенно произвольным и ограниченным только общими правилами адресации стека TCP/IP? Дело в том, что во многих случаях клиенты не хотят полной изоляции VPN: в частности, они нуждаются в выходе в Интернет. Независимое же (не согласованное с регламентирующими органами Интернета) назначение адресов узлам VPN может привести к совпадению внутренних адресов сайтов с уже использованными в Интернете публичными адресами, в результате чего связь с Интернетом станет невозможной. При применении зарезервированных частных адресов проблема связи клиентов VPN с внешним миром решается с помощью стандартной техники трансляции адресов (NAT). В любом случае должно соблюдаться требование уникальности адресов в пределах VPN.

Однако использование в разных сетях VPN одного и того же адресного пространства создает проблему для маршрутизаторов PE. Протокол BGP изначально был разработан в предположении, что все адреса, которыми он манипулирует, во-первых, относятся к семейству адресов IPv4 и, во-вторых, однозначно идентифицируют узлы сети, то есть являются глобально уникальными в пределах всей составной сети. Ориентация на глобальную уникальность адресов выражается в том, что получив очередное маршрутное объявление, протокол BGP анализирует его, не обращая внимания на то, какой сети VPN принадлежит этот маршрут. Если на вход BGP поступают описания маршрутов к узлам разных сетей VPN, но с совпадающими адресами IPv4, то протокол BGP считает, что все они ведут к одному и тому же узлу, а, следовательно, как и предусмотрено в алгоритме его работы, он помещает в соответствующую таблицу VRF только один лучший (в соответствии с правилами выбора BGP) маршрут.

Эта проблема была решена в MPLS VPN применением вместо потенциально неоднозначных адресов IPv4 расширенных и однозначных адресов нового типа, а именно адресов **VPN-IPv4**, получаемых путем преобразования исходных адресов IPv4. Преобразование заключается в том, что ко всем адресам IPv4, составляющим адресное пространство той или иной сети VPN, добавляется префикс, который называется **различителем маршрутов** (Route Distinguisher, RD) и который уникально идентифицирует эту сеть. В результате на маршрутизаторе PE все адреса, относящиеся к разным сетям VPN, обязательно будут отличаться друг от друга, даже если они имеют совпадающую часть — адрес IPv4.

Именно здесь оказалась полезной способность расширенного протокола MP-BGP переносить в маршрутных объявлениях адреса *разных типов*, в том числе IPv6, IPX, а

главное — VPN-IPv4. Адреса VPN-IPv4 используются только для маршрутов, которыми маршрутизаторы PE обмениваются по протоколу BGP. Прежде чем передать своему напарнику некоторый маршрут, входной маршрутизатор PE добавляет к его адресу назначения IPv4 префикс RD для данной сети VPN, тем самым преобразуя его в адрес VPN-IPv4.

Как уже было отмечено, префиксы RD должны гарантированно уникально идентифицировать VPN, чтобы избежать дублирования адресов. Упростить выбор RD, не создавая для этих целей дополнительных централизованных процедур (например, распределения RD органами Интернета подобно распределению адресов IPv4), предлагается за счет использования в качестве основы для RD заведомо уникальных чисел: либо номеров автономных систем, либо глобальных адресов интерфейсов PE со стороны магистральной сети поставщика.

RD имеет длину 8 байт и состоит из трех полей.

- Первое поле *типа* длиной 2 байта определяет тип и разрядность второго поля.
- Второе поле называется полем *администратора* и однозначно идентифицирует поставщика. Значение 0 в поле типа говорит о том, что в поле администратора указан IP-адрес интерфейса маршрутизатора PE, и длина данного поля составляет, естественно, 4 байта. Если же значение поля типа равно 1, то в качестве идентификатора поставщика выбрано значение номера его автономной системы, так что длина поля администратора составит уже 2 байта.
- Третье поле носит название поля *назначенного номера*, оно служит для обеспечения уникальности адресов VPN в пределах сети поставщика. Назначенные номера выбирает сам поставщик, это могут быть произвольные числа, главное, чтобы существовало однозначное соответствие между ними и VPN поставщика.

Рис. 4 иллюстрирует сложный процесс обмена маршрутными объявлениями в сети MPLS VPN. Этот процесс включает преобразование адресов из формата IPv4 в формат VPN-IPv4, фильтрацию маршрутных объявлений (операции экспорта-импорта) и добавление к объявлениям меток VPN. Мы последовательно будем рассматривать эти вопросы, поэтому читатель должен быть готов к тому, что не все на рисунке ему будет нужно и понятно с самого начала.

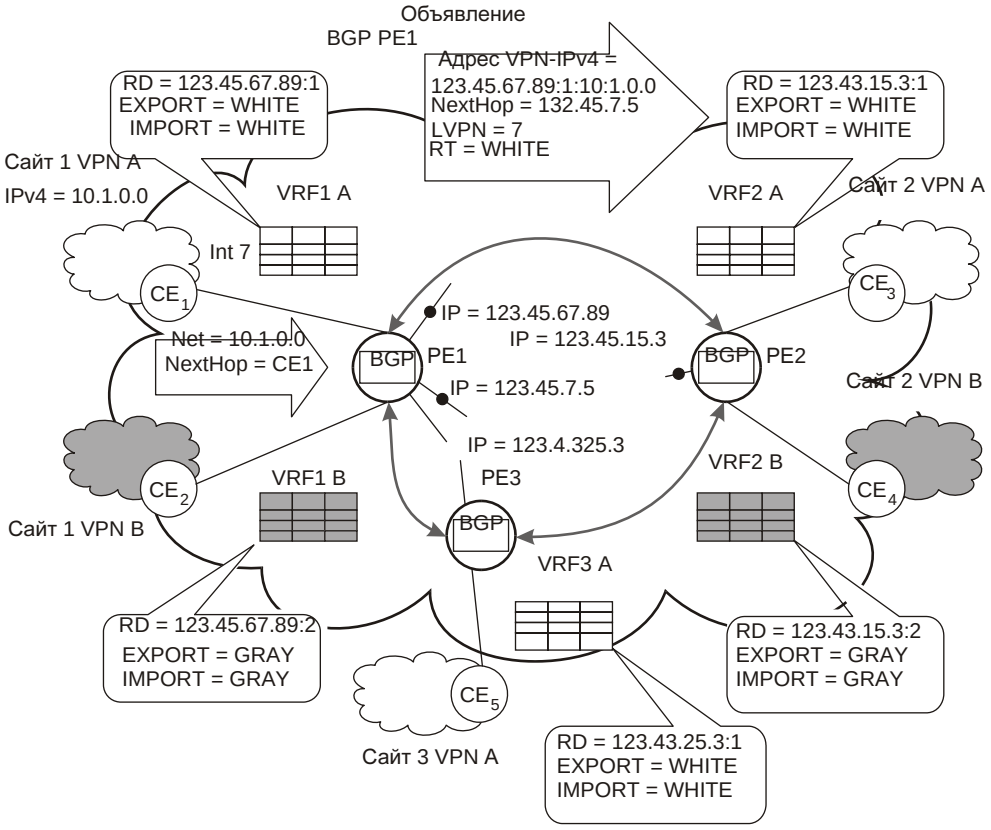

Рис. 4. Маршрутные объявления MP-BGP

Итак, на рис. 4 показан пример преобразования адресов формата IPv4 с целью обеспечения уникальности адресов в рамках всех сетей VPN одного поставщика услуг. При формировании RD для каждой из сетей VPN администратор сети сначала выбирает глобальный адрес одного из внешних интерфейсов маршрутизатора PE1 (на рисунке это адрес 123.45.67.89), затем добавляет к нему через двоеточие 1, получая значение RD, равное 123.45.67.89:1. Формат RD представлен в табл. 1.

Таблица 1. Формат RD

Поле типа (2 байта)	**Поле администратора (4 байта)**	**Поле назначенного номера (2 байта)**
0	123.45.67.89	1

Указанное значение RD администратор назначает для сети VPN A. При конфигурировании маршрутизаторов PE администратор указывает это значение для всех таблиц VRF, которые соответствуют сети VPN A. В частности, он задает это значение при создании VRF 1A, так что для протокола MP-BGP все адреса формата IPv4, которые находятся в таблице VRF1A, будут иметь значение RD, равное 123.45.67:1, в том числе все адреса с префиксом 10.1/16, которые PE1 получает от маршрутизатора CE1 сайта 1 в сети VPN A.

Аналогично, администратор выбирает для сетей VPN B значение RD, равное 123.45.67.89:2, которое он указывает при конфигурировании VRF 1B на маршрутизаторе PE1. Это значение RD будет добавляться ко всем адресам IPv4, хранящимся в таблицах VRF 1B, при обработке их протоколом MP-BGP.

ПРИМЕЧАНИЕ

Все маршруты в таблицах VRF содержат адреса в формате IPv4.

Сформированные маршруты в формате VPN-IPv4 маршрутизатор PE1 передает по протоколу MP-BGP на маршрутизатор PE2, к которому подключен сайт 2 сети VPN B. Только благодаря добавлениям значения RD протоколы BGP, работающие на удаленных маршрутизаторах PE, различают маршруты с совпадающими адресами IPv4, относящимися к разным сетям VPN.

Документ RFC 2547bis не требует, чтобы все маршруты внутри одной сети VPN индексировались одним и тем же значением RD. Более того, один и тот же сайт, подключенный к разным интерфейсам одного маршрутизатора PE или к разным маршрутизаторам PE, может иметь разные значения RD. Благодаря этому путь к одному и тому же узлу может описываться разными маршрутами, что дает возможность выбора того или иного маршрута для различных пакетов. Однако принципиально важно, чтобы значения RD разных сетей VPN не совпадали.

Генерация маршрутных объявлений MP-BGP

При получении от сайта клиента нового маршрута по протоколу класса IGP, такому как RIP, OSPF или IS-IS, маршрутизатор PE заносит его в соответствующую таблицу VRF и распространяет дальше между другими сайтами данной сети VPN. Обмен маршрутной информацией между сайтами каждой отдельной сети VPN выполняется под управлением протокола MP-BGP. Маршрутное объявление MP-BGP имеет следующий набор атрибутов, расширенный по сравнению с протоколом BGP:

- **Адрес сети назначения в формате VPN-IPv4**.
- **Адрес следующего маршрутизатора** (NextHop). Протокол BGP указывает в данном случае адрес одного из внутренних (идущих к маршрутизаторам P) интерфейсов того маршрутизатора PE, на котором он работает.
- **Метка виртуальной частной сети** (VPN Label, LVPN) уникально определяет внешний интерфейс маршрутизатора PE и подключенный к нему сайт клиента, куда ведет объявляемый маршрут. Она назначается маршруту входным маршрутизатором PE при получении им локального маршрута от присоединенного маршрутизатора CE.
- **Расширенные атрибуты сообщества** (extended community attributes), один из которых — **маршрутная цель** (Route Target, RT) — является обязательным. Этот атрибут идентифицирует набор сайтов (VRF), входящих в данную сеть VPN, которым маршрутизатор PE должен посылать маршруты. Значение атрибута RT в объявлении о маршруте определяется *политикой экспорта* маршрутных объявлений, которая задается администратором при конфигурировании таблицы VRF, содержащей данный маршрут (о политике экспорта и импорта маршрутных

объявлений рассказывается в разделе «Механизм формирования топологии VPN»).

Пусть, например, маршрутизатор PE1 получает с сайта 1 сети VPN A по протоколу класса IGP следующее объявление о маршруте в формате IPv4 (см. рис. 4):

```
Net = 10.1.0.0
NextHop = CE1
```

На основании этого объявления в таблицу VRF 1A заносится соответствующая запись. Протокол BGP периодически просматривает таблицу VRF 1A, и обнаружив новую запись, генерирует объявление о маршруте, для чего выполняет следующие действия.

6. Добавляет к адресу сети назначения префикс RD (в данном случае он равен 123.45.67.89:1).
7. Переписывает значение поля NextHop, заменяя адрес внешнего интерфейса CE1 адресом внешнего интерфейса PE1, через который пролегает путь к адресу назначения (пусть в данном случае это будет 123.45.7.5).
8. Назначает метку LVPN, указывающую на таблицу VRF1A и интерфейс маршрутизатора PE1, к которому подключен сайт клиента, содержащий узел назначения (в данном случае значение метки равно 7, на рис. 4 этот интерфейс обозначен как int 7).
9. Задает атрибут RT (на рис. 4 значение атрибута RT условно обозначено как WHITE, что идентифицирует набор всех сайтов, входящих в сеть VPN A).

В результате получается такое маршрутное объявление:

```
VPN-IPv4: 123.45.67.89:1:10.1.0.0
Nexthop = 123.45.7.5
LVPN = 7
RT = WHITE
```

Это объявление протокол MP-BGP посылает всем своим соседям (на рисунке объявление помещено внутрь широкой стрелки).

Когда выходной маршрутизатор PE получает маршрут к сети VPN-IPv4, он делает обратное преобразование, отбрасывая префикс RD, и только потом помещает маршрут в таблицу VRF2A и объявляет о нем связанному с ним маршрутизатору заказчика CE3 из данной сети VPN A. В результате в таблице VRF2A появляется новая запись:

```
Net = 10.1/16
Nexthop = 123.45.7.5 (BGP)
LVPN = 7
```

Перемещение пакета по сети MPLS VPN

Теперь, когда мы обсудили схему распространения маршрутной информации по сети MPLS VPN, давайте посмотрим, как перемещаются *данные* между узлами одной сети VPN.

Пусть, например, с сайта 2 сети VPN A узел с адресом 10.2.1.1/16 отправляет пакет узлу сайта 1 этой же сети VPN, имеющему адрес 10.1.0.3/16 (рис. 5). Стандартными транспортными средствами IP-пакет доставляется на пограничный маршрутизатор сайта CE3, в таблице которого для номера сети 10.1.0.0 в качестве следующего маршрутизатора указан маршрутизатор PE2. На маршрутизатор PE2 пакет поступает с интерфейса 2, поэтому для дальнейшего продвижения пакета он обращается к таблице VRF2A, связанной с данным интерфейсом.

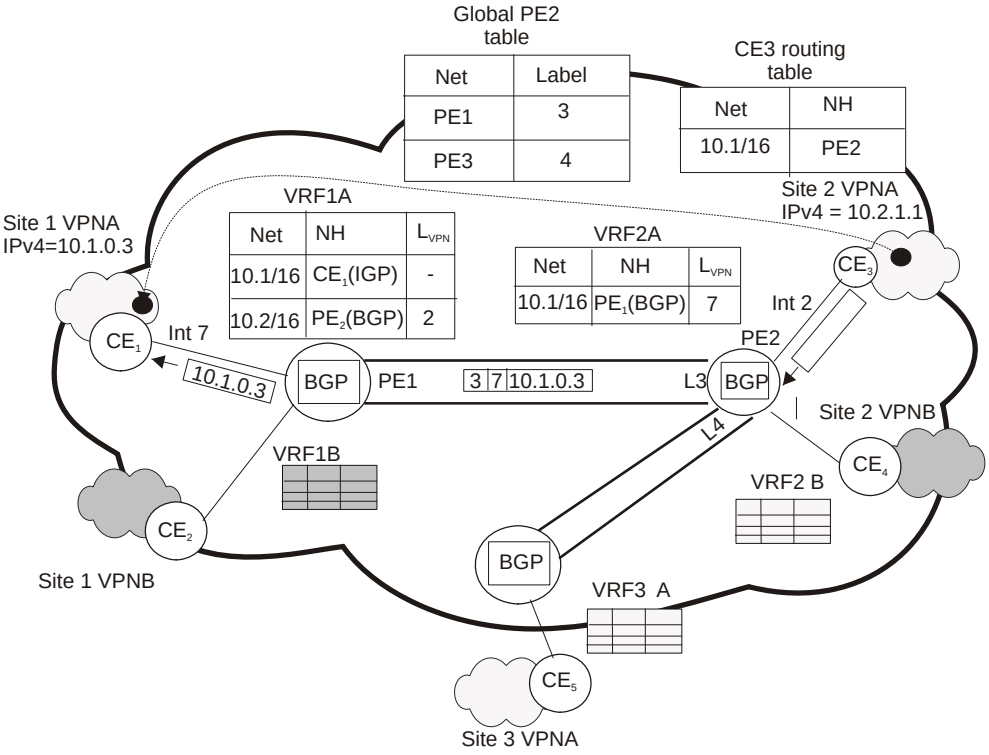

Рис. 5. Путешествие пакета по сети MPLS VPN

В таблице VRF2A адресу 10.1.0.0 соответствует запись протокола BGP, которая указывает, что следующим маршрутизатором (next hop) для пакета определен маршрутизатор PE1. Поле метки содержит значение LVPN = 7, определяющее интерфейс выходного маршрутизатора PE1. Это значения должно быть присвоено пакету для того, чтобы он попал в нужную сеть VPN. Здесь также указывается, что запись была сделана протоколом BGP, а не IGP. На этом основании маршрутизатор PE2 «понимает», что очередной маршрутизатор не является непосредственным соседом, и путь к нему надо искать в глобальной таблице маршрутизации.

В глобальной таблице для адреса PE1 указывается начальное значение метки пути LSP, равное 3. Мы не будем останавливаться на способе прокладки пути между маршрутизаторами PE1 и PE2 — этот вопрос мы обсуждали в главе 20 при изучении технологии MPLS.

В сетях MPLS VPN используются иерархические свойства путей MPLS, за счет которых пакет может быть снабжен несколькими метками, помещаемыми в стек. На вхо-

де во внутреннюю сеть поставщика, образуемую маршрутизаторами P, пакет будет снабжен двумя метками LVPN = 7 и L = 3. Метка LVPN интерпретируется как метка нижнего уровня — оставаясь на дне стека, она не используется, пока пакет путешествует по туннелю PE1-PE2. Продвижение пакета происходит на основании метки верхнего уровня L. Каждый раз, когда пакет проходит очередной маршрутизатор P вдоль туннеля, метка L анализируется и заменяется новым значением. И только после достижения конечной точки туннеля — маршрутизатора PE1 — из стека извлекается метка LVPN. В зависимости от ее значения пакет направляется на тот или иной выходной интерфейс маршрутизатора PE1 (на рис. 4 этот интерфейс обозначен как int7).

Из таблицы VRF1A, связанной с данным интерфейсом и содержащей маршруты VPNA, извлекается запись о маршруте к узлу назначения, указывающая на устройство CE1 в качестве следующего маршрутизатора. Заметим, что запись об этом маршруте была помещена в таблицу VRF1A протоколом IGP. Последний отрезок путешествия пакета от CE1 до узла 10.1.0.3 осуществляется традиционными средствами IP.

Механизм формирования топологии VPN

Политика экспорта/импорта маршрутов — мощный инструмент создания сетей VPN разных топологий.

При конфигурировании каждой таблицы VRF задаются два атрибута RT, один для определения политики экспорта, другой — политики импорта маршрутов.

Маршрутные объявления MP-BGP всегда несут атрибут RT, говорящий об экспорте маршрута. Сравнение значений атрибутов RT в маршрутном объявлении и в параметрах VRF позволяет решить вопрос о принятии или отклонении предлагаемого маршрута. А это и означает формирование топологии сети. Рассмотрим этот механизм на примере.

Пусть изображенный на рис. 4 маршрутизатор PE2 получил объявление от PE1. Прежде, чем сохранить информацию о маршруте, он проверяет значение атрибута RT, содержащееся в объявлении, на совпадение с политикой импорта всех своих таблиц VRF, в данном случае VRF2A и VRF2B. Значение атрибута RT равно WHITE, поэтому маршрут добавляется (после преобразования в формат IPv4 путем удаления префикса RD) только в таблицу VRF2A, так как для нее определена политика импорта WHITE. Таблица VRF2B остается в неизменном виде, так как ее политика импорта говорит о том, что в нее должны помещаться только маршруты с атрибутом RT, равным GRAY.

Задание одного и того же значения для политики экспорта и импорта для всех таблиц VRF определенной сети VPN (именно этот случай для сети VPN A показан на рис. 4) приводит к полносвязной топологии — каждый сайт может посылать пакеты непосредственно тому сайту, в котором находится сеть назначения.

Существуют и другие варианты топологии VPN. Например, за счет конфигурирования политики экспорта/импорта можно реализовать такую популярную топологию, как «звезда», когда все сайты (spoke) общаются друг с другом через выделенный центральный сайт (hub).

Для достижения этого эффекта достаточно определить для VRF центрального сайта политику импорта как Import = spoke, а экспорта как Export = hub, а на таблицах VFR периферийных сайтов поступить наоборот, задав Import = hub, а Export = spoke (рис. 6). В результате таблицы VRF периферийных сайтов не будут принимать маршрутные объявления друг от друга, поскольку они передаются по сети протоколом MP-BGP с атрибутом RT = spoke, между тем как их политика импорта разрешает получать объявления с атрибутом RT = hub. Зато объявления VRF периферийных сайтов принимает таблица VRF центрального сайта, для которого как раз и определена политика импорта spoke. Этот сайт обобщает все объявления периферийных сайтов и отсылает их обратно, но уже с атрибутом RT = hub, что совпадает с политикой импорта периферийного сайта. Таким образом, в таблицах VRF каждого периферийного сайта появляются записи о сетях в других периферийных сайтах. А в качестве следующего транзитного узла указывается адрес интерфейса PE, связанного с центральным сайтом, поскольку объявление пришло от него. Поэтому пакеты между периферийными сайтами проходят транзитом через пограничный маршрутизатор PE3, подключенный к центральному сайту.

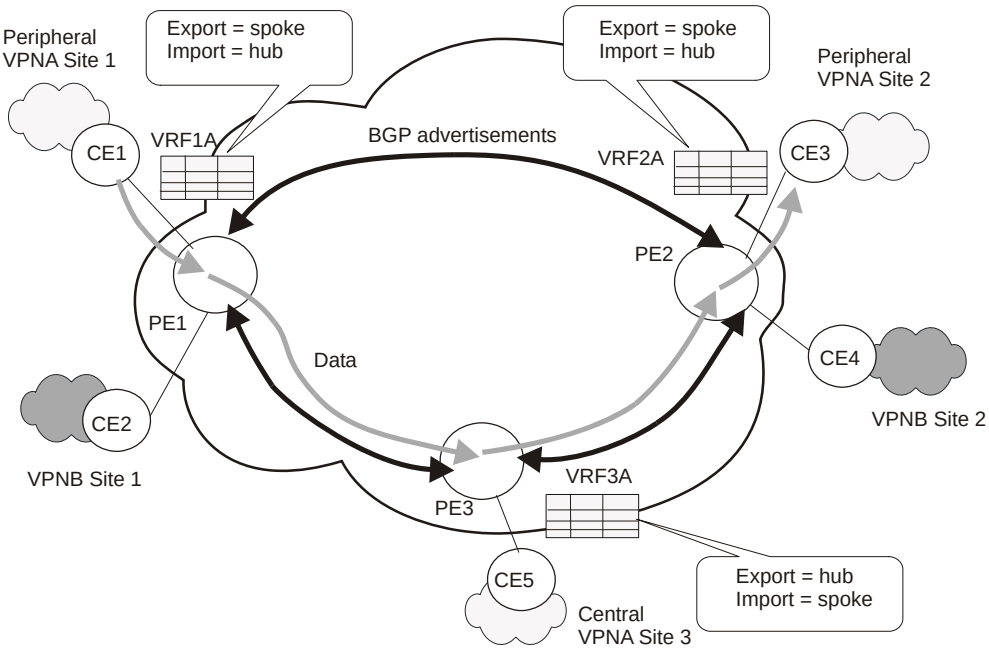

Рис. 6. Конфигурирование звездообразной топологии для сети VPN A

Из описания механизмов MPLS VPN можно сделать вывод, что процесс конфигурирования новой или модификации существующей сети VPN достаточно сложен, но он хорошо формализуется и автоматизируется. Для исключения возможных ошибок конфигурирования, например, приписывания сайту ошибочной политики импорта/экспорта маршрутных объявлений, что может привести к присоединению сайта к чужой сети VPN, некоторые производители разработали автоматизированные программные системы конфигурирования MPLS.

Степень защищенности

Повысить степень защищенности MPLS VPN можно с помощью традиционных средств, например, применяя средства аутентификации и шифрования протокола IPSec, устанавливаемые в сетях клиентов или в сети поставщика. Услуга MPLS VPN может легко интегрироваться с другими услугами IP, например, с предоставлением доступа к Интернету пользователям VPN с защитой их сети средствами межсетевого экрана, установленного в сети поставщика. Поставщик также может предоставлять пользователям MPLS VPN услуги, базирующиеся на других возможностях MPLS, в частности, гарантированное качество обслуживания на основе методов инжиниринга трафика MPLS. Что же касается сложностей ведения в маршрутизаторах поставщика услуг таблиц маршрутизации пользователей, на которые указывают некоторые специалисты, то они, на наш взгляд, несколько преувеличены, так как таблицы создаются автоматически с помощью стандартных протоколов маршрутизации и только на пограничных маршрутизаторах (PE). Механизм виртуального маршрутизатора полностью изолирует эти таблицы от глобальных таблиц маршрутизации поставщика услуг, что обеспечивает необходимые уровни надежности и масштабируемости решений MPLS VPN. Впрочем, реальное качество данной технологии покажет время и, скорее всего, достаточно скоро.

Технология MPLS VPN не обеспечивает безопасности за счет шифрования и аутентификации, как это делают технологии IPSec и PPTP, но допускает применение данных технологий как дополнительных мер защиты в случае необходимости.

Перед MPLS VPN не ставится задача поддержки качества обслуживания, но при необходимости поставщик может использовать методы дифференцированного обслуживания и инжиниринга трафика.

Коммутируемый доступ через сеть ISDN (дополнительный материал к «Коммутируемый доступ через сеть ISDN» в главе 22)

Назначение и структура ISDN

Целью создания технологии ISDN (Integrated Services Digital Network — **цифровая сеть с интегрированным обслуживанием**) было построение всемирной сети, которая должна была прийти на смену телефонной сети, и будучи такой же доступной и распространенной, предоставлять миллионам своих пользователей разнообразные услуги, как телефонные, так и передачи данных. Передача телевизионных программ по ISDN не предполагалась, поэтому было решено ограничиться пропускной способностью абонентского окончания для массовых пользователей в 128 Кбит/с.

Если бы цель разработчиков ISDN была достигнута в полной мере, то проблемы доступа домашних пользователей к Интернету и корпоративным сетям была бы окончательно решена. Однако по многим причинам внедрение ISDN происходило очень медленно — процесс, который начался в 80-е годы, растянулся больше чем на десять лет, так что к моменту появления в домах пользователей некоторые услуги ISDN просто морально устарели. Так, скорость доступа 128 Кбит/с сегодня приемлема уже не для всех пользователей. Существует, правда, такой интерфейс ISDN, который обеспечивает скорость доступа до 2 Мбит/с, но он достаточно дорог для массового пользователя и его обычно применяют только предприятия для подключения своих сетей.

Хотя сеть ISDN и не стала той новой публичной сетью, на роль которой она претендовала, ее услуги сегодня достаточно доступны. Далее мы рассмотрим структуру этой сети и ее возможности в отношении организации удаленного доступа.

Архитектура сети ISDN предусматривает несколько видов услуг (рис. 1):

- некоммутируемые средства (выделенные цифровые каналы);
- коммутируемая телефонная сеть общего пользования;
- сеть передачи данных с коммутацией каналов;
- сеть передачи данных с коммутацией пакетов;
- сеть передачи данных с трансляцией кадров (режим сети Frame Relay);
- средства контроля и управления работой сети.

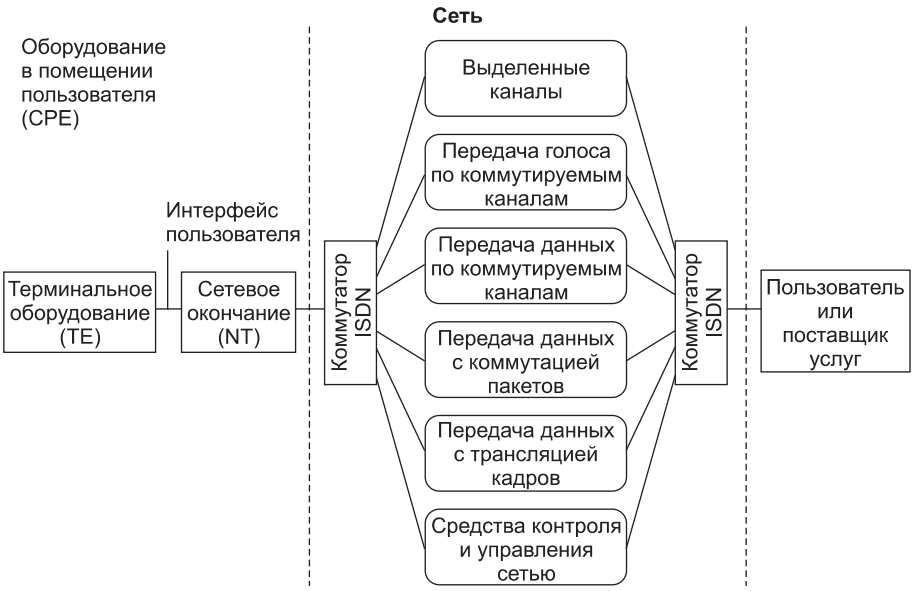

Рис. 1. Услуги сети ISDN

Как видно из приведенного списка, транспортные службы сетей ISDN действительно покрывают очень широкий спектр услуг, включая популярные услуги сети Frame Relay. Стандарты ISDN описывают также ряд услуг прикладного уровня: факсимильную связь на скорости 64 Кбит/с, телексную связь на скорости 9600 бит/с, видеотекс на скорости 9600 бит/с и некоторые другие.

Все услуги основаны на передаче информации в цифровой форме. Пользовательский интерфейс также является цифровым, то есть все его абонентские устройства (телефон, компьютер, факс) должны передавать в сеть цифровые данные. Организация **цифрового абонентского окончания** (Digital Subscriber Line, DSL) стала одним из серьезных препятствий на пути распространения ISDN, так как требовала модернизации миллионов абонентских окончаний.

Базовой скоростью сети ISDN является скорость канала DS-0, то есть 64 Кбит/с. Эта скорость ориентируется на самый простой метод кодирования голоса — PCM, хотя дифференциальное кодирование и позволяет передавать голос с тем же качеством на скорости 32 или 16 Кбит/с.

Одной из оригинальных идей, положенных в основу ISDN, является совместное использование принципов коммутации каналов и пакетов. Однако сеть с коммутацией пакетов, работающая в составе ISDN, выполняет только служебные функции — с ее помощью передаются сообщения сигнального протокола, а вот основная информация, то есть сам голос, по-прежнему передается через сеть с коммутацией каналов. В таком разделении функций есть вполне понятная логика — сообщения о вызове абонентов образуют пульсирующий трафик, поэтому его эффективнее передавать по сети с коммутацией пакетов.

Интерфейсы BRI и PRI

Одним из основных принципов ISDN является предоставление пользователю стандартного интерфейса, с помощью которого пользователь может запрашивать у сети разнообразные услуги. Этот интерфейс образуют два типа оборудования, устанавливаемого в помещении пользователя (Customer Premises Equipment, CPE). К этому оборудованию относится:

- **терминальное оборудование** (Terminal Equipment, TE) пользователя (компьютер с соответствующим адаптером, маршрутизатор, телефонный аппарат);
- **сетевое окончание** (Network Termination, NT), которое представляет собой устройство, завершающее линию связи с ближайшим коммутатором ISDN.

Пользовательский интерфейс основан на каналах трех типов: B, D и H.

Каналы типа B обеспечивают передачу пользовательских данных (оцифрованного голоса, компьютерных данных или смеси голоса и данных) с более низкими скоростями, чем 64 Кбит/с. Разделение данных выполняется с помощью техники TDM. Разделением канала B на подканалы в этом случае должно заниматься пользовательское оборудование, сеть ISDN всегда коммутирует целые каналы типа B. Каналы типа B могут соединять пользователей друг с другом с помощью техники коммутации каналов, а также образовывать так называемые полупостоянные соединения, которые эквиваленты соединениям выделенных каналов обычной телефонной сети. Кроме того, канал типа B может подключать пользователя к коммутатору сети X.25.

Канал типа D является каналом доступа к служебной сети с коммутацией пакетов, передающей сигнальную информацию со скоростью 16 или 64 Кбит/с. Передача адресной информации, на основе которой осуществляется коммутация каналов типа B в коммутаторах сети, является основной функцией канала D. Другой его функцией является поддержание сервиса низкоскоростной сети с коммутацией пакетов для пользовательских данных. Обычно этот сервис поддерживается сетью в то время, когда каналы типа D свободны от выполнения основной функции.

Каналы типа H предоставляют пользователям возможности высокоскоростной передачи данных: 384 Кбит/с (H0), 1536 Кбит/с (H11) или 1920 Кбит/с (H12). На них могут работать службы высокоскоростной передачи факсов, видеоинформации, качественного воспроизведения звука.

Пользовательский интерфейс ISDN представляет собой набор каналов определенного типа и с определенными скоростями. Сеть ISDN поддерживает два вида пользовательского интерфейса с начальной (Basic Rate Interface, BRI) и основной (Primay Rate Interface, PRI) скоростями передачи данных.

Начальный интерфейс ISDN предлагает пользователю два канала по 64 Кбит/с для передачи данных (каналы типа B) и один канал с пропускной способностью 16 Кбит/с для передачи управляющей информации (канал типа D). Все каналы работают в дуплексном режиме. В результате суммарная скорость интерфейса BRI для пользовательских данных составляет 144 Кбит/с по каждому направлению, а с учетом служебной информации — 192 Кбит/с. Различные каналы пользовательского интерфейса разделяют один и тот же физический двухпроводный кабель по технологии TDM, то есть являются логическими, а не физическими каналами. Данные по интерфейсу BRI передаются кадрами, состоящими из 48 бит. Каждый кадр содержит

по 2 байта каждого из двух каналов B, а также 4 бита канала D. Передача кадра длится 250 мс, что обеспечивает скорость передачи данных 64 Кбит/с для каналов B и 16 Кбит/с — для канала D. Помимо битов данных кадр содержит служебные биты для синхронизации кадров, а также обеспечения нулевой постоянной составляющей электрического сигнала. Интерфейс BRI может поддерживать не только схему 2B + D, но и B + D и просто D.

Начальный интерфейс стандартизован в рекомендации I.430.

Основной интерфейс ISDN предназначен для пользователей с повышенными требованиями к пропускной способности сети. Интерфейс PRI поддерживает либо схему 30B + D, либо схему 23B + D. В обеих схемах канал D обеспечивает скорость 64 Кбит/с. Первый вариант предназначен для Европы, второй — для Северной Америки и Японии. Ввиду большой популярности скорости цифровых каналов 2,048 Мбит/с в Европе и скорости 1,544 Мбит/с в остальных регионах привести стандарт на интерфейс PRI к общему варианту не удалось.

Возможны варианты интерфейса PRI с меньшим количеством каналов типа B, например 20B + D. Каналы типа B могут объединяться в один логический высокоскоростной канал с общей скоростью до 1920 Кбит/с. При установке у пользователя нескольких интерфейсов PRI все они могут иметь один канал типа D, при этом количество каналов B в том интерфейсе, который не имеет канала D, может увеличиваться до 24 или 31.

Основной интерфейс может быть также основан на каналах типа H. При этом общая пропускная способность интерфейса все равно не должна превышать 2,048 или 1,544 Мбит/с. Для каналов H0 возможны интерфейсы 3H0 + D для американского варианта и 5H0 + D для европейского. Для каналов H1 возможен интерфейс, состоящий только из одного канала H11 (1,536 Мбит/с) для американского варианта или одного канала H12 (1,920 Мбит/с) и одного канала D для европейского варианта. Кадры интерфейса PRI имеют структуру кадров DS-1 для каналов T1 или E1.

Основной интерфейс PRI стандартизован в рекомендации I.431.

ВНИМАНИЕ

> Как каналы B, так и каналы D являются логическими каналами абонентского окончания, которое физически представляет собой одну витую пару. Каналы D и B образуются путем применения техники TDM к физической среде, образуемой этой витой парой.

Стек протоколов ISDN

В сети ISDN существуют два стека протоколов: стек каналов типа D и стек каналов типа B (рис. 2).

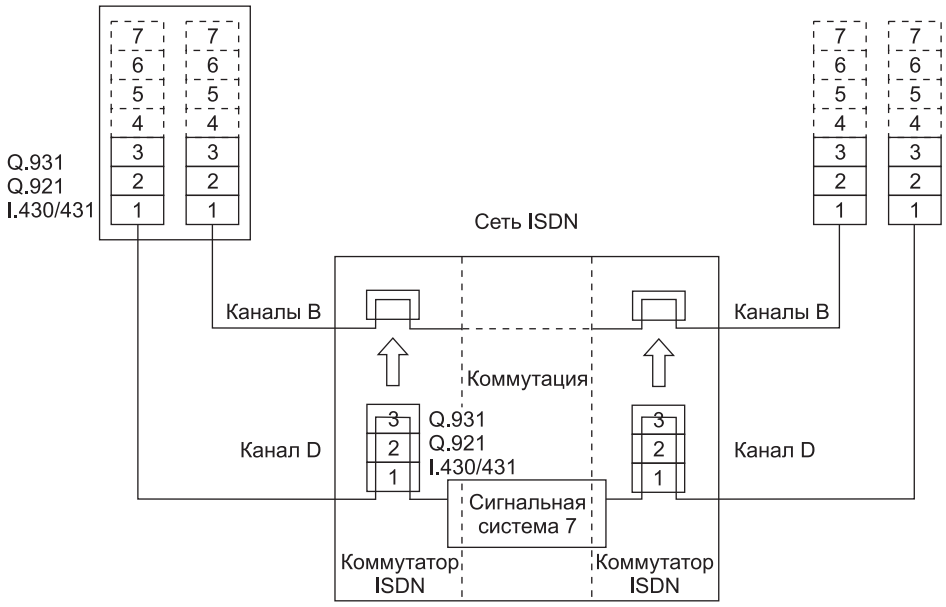

Рис. 2. Структура сети ISDN

Сеть каналов типа D внутри сети ISDN служит транспортной системой с коммутацией пакетов, применяемой для передачи сообщений сигнализации. Прообразом этой сети послужила технология сетей X.25. Для сети каналов D определены три уровня протоколов:

- физический протокол определяется стандартом I.430/431;
- канальный протокол LAP-D определяется стандартом Q.921;
- на сетевом уровне может использоваться протокол сигнализации Q.931, с помощью которого выполняется маршрутизация вызова абонента службы с коммутацией каналов.

Каналы типа B образуют сеть с коммутацией каналов, которая передает данные абонентов, то есть оцифрованный голос. В терминах модели OSI на каналах типа B в коммутаторах сети ISDN определен только протокол физического уровня — протокол I.430/431. Коммутация каналов типа B происходит по указаниям, полученным по каналу D. Когда кадры протокола Q.931 маршрутизируются коммутатором, происходит одновременная коммутация очередной части составного канала от исходного абонента к конечному.

Протокол LAP-D принадлежит к семейству HDLC. Протокол LAP-D обладает всеми «родовыми чертами» этого семейства, но имеет и некоторые особенности. Адрес кадра LAP-D состоит из двух байтов — один байт определяет код службы, которой пересылаются вложенные в кадр пакеты, а второй требуется для адресации одного из терминалов, если у пользователя к абонентскому окончанию подключено несколько терминалов. Терминальное устройство ISDN может поддерживать разные услуги: установление соединения по протоколу Q.931, коммутация пакетов X.25, мониторинг сети и т. п. Протокол LAP-D обеспечивает два режима работы: с установлением со-

единения и без установления соединения. Последний режим используется, например, для мониторинга сети.

Протокол Q.931 является сигнальным протоколом ISDN для участка пользователь-сеть, то есть протоколом типа UNI. Он переносит в своих пакетах ISDN-адрес вызываемого абонента, на основании которого и происходит настройка коммутаторов на поддержку составного канала типа B. Процедуру установления соединения по протоколу Q.931 иллюстрирует рис. 3.

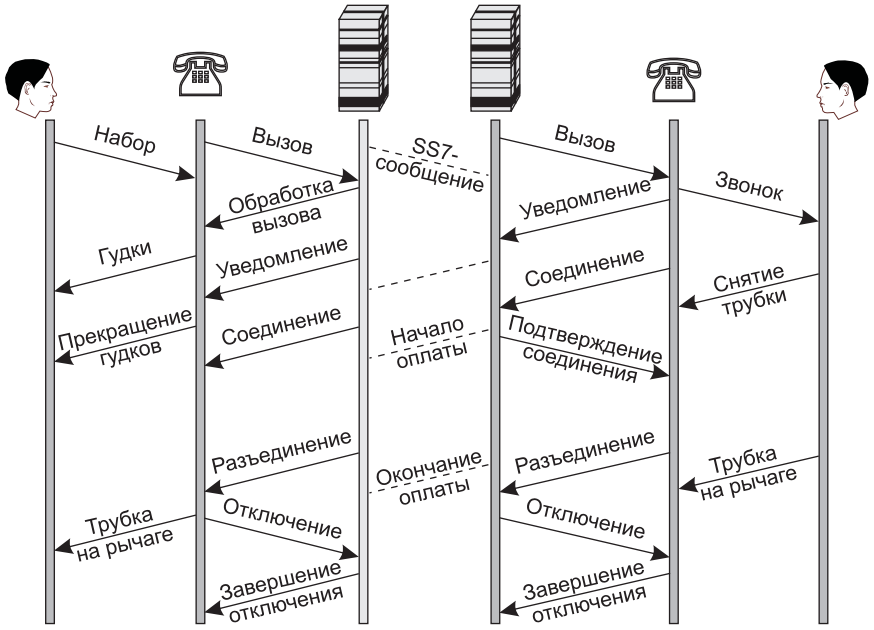

Рис. 3. Базовая процедура установления соединения в ISDN по протоколу Q.931

После того как пользователь снял трубку и набрал номер вызываемого абонента, телефонный аппарат ISDN формирует пакет вызова (set up) и отправляет его по каналу D коммутатору ISDN, к которому он подключен. Этот коммутатор отвечает аппарату абонента пакетом обработки вызова, с приходом которого аппарат начинает генерировать длинные гудки. Одновременно коммутатор запоминает факт запроса на установление соединения и передает принятое сообщение следующему коммутатору, адрес которого он находит по таблице, аналогичной таблице маршрутизации маршрутизаторов пакетных сетей. При этом сообщение протокола Q.931 транслируется в сообщение начального адреса (Initial Address Message, IAM) протокола SS7 аналогичного назначения (на рисунке сообщения SS7 не детализированы). Проходя через сеть, сообщения SS7 переводят промежуточные коммутаторы в состояние готовности к установлению соединения. Выходной коммутатор сети, к которому подключен аппарат вызываемого абонента, преобразует сообщение начального адреса протокола SS7 в сообщение вызова протокола Q.931, на основании которого телефонный аппарат начинает звонить. Если абонент снимает трубку, то его аппарат генерирует сообщение соединения (connect), которое в обратном порядке проходит через все промежуточные коммутаторы (преобразованное, естественно, в соответствующее сообще-

ние SS7). При этом обратном проходе коммутаторы устанавливают состояние соединения, коммутируя соответствующим образом каналы типа B.

Любое абонентское устройство ISDN должно поддерживать протокол Q.931, так что телефон ISDN намного сложнее своего аналогового коллеги. Как видно из рисунка, внутри сети сообщения Q.931 транслируются в сообщения протокола SS7, который является протоколом взаимодействия коммутатор-коммутатор (NNI), а затем снова преобразуются в сообщения Q.931 на абонентском окончании.

Использование сети ISDN для передачи данных

Несмотря на значительные отличия от аналоговых телефонных сетей, сети ISDN сегодня используются в основном так же, как аналоговые телефонные сети, то есть как сети с коммутацией каналов, но только более скоростные: интерфейс BRI дает возможность установить дуплексный режим обмена со скоростью 128 Кбит/с (логическое объединение двух каналов типа B), а интерфейс PRI — 2,048 Мбит/с. Кроме того, качество цифровых каналов гораздо выше, чем аналоговых. Это значит, что процент искаженных кадров оказывается существенно ниже, а полезная скорость обмена данными — существенно выше.

Обычно интерфейс BRI служит в коммуникационном оборудовании для подключения отдельных компьютеров или небольших локальных сетей домашних пользователей, а интерфейс PRI — для подключения сети средних размеров с помощью маршрутизатора.

Схема удаленного доступа через ISDN показана на рис. 4.

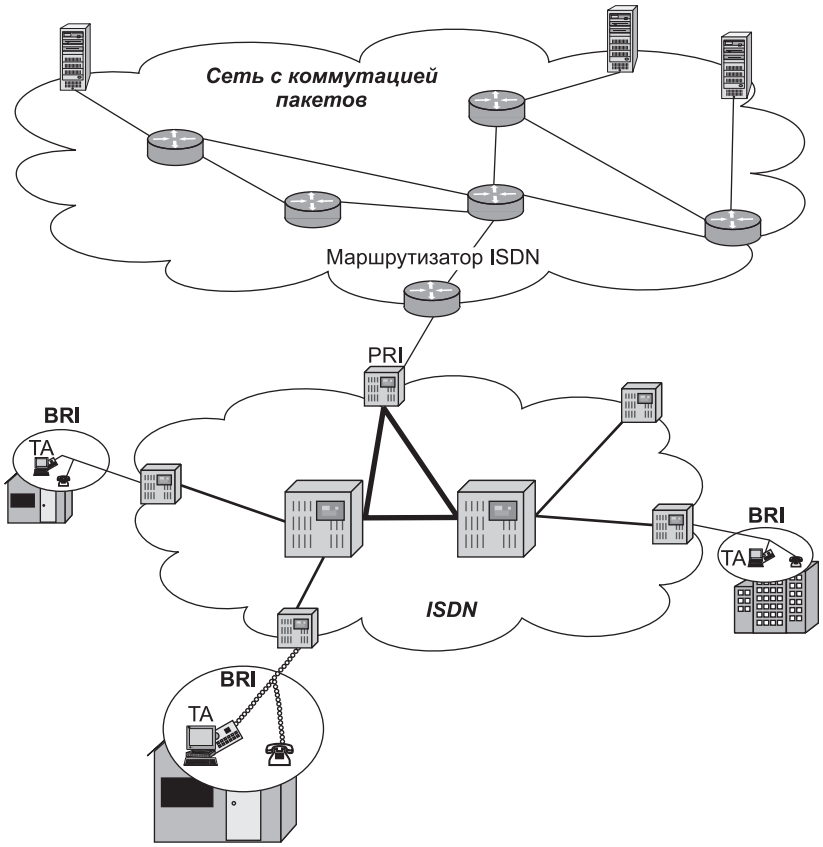

Рис. 4. Удаленный доступ с использованием ISDN

Подключение пользовательского оборудования к сети ISDN осуществляется в соответствии со схемой, разработанной ITU-T (рис. 5). Оборудование делится на функциональные группы, и в зависимости от группы различают несколько **контрольных точек** соединения разных групп оборудования между собой.

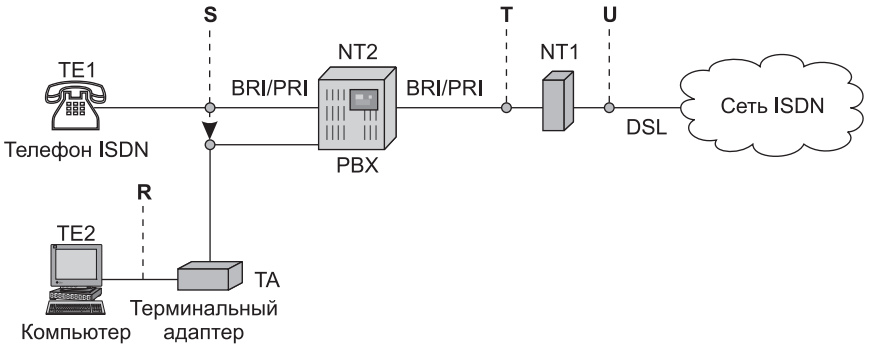

Рис. 5. Подключение пользовательского оборудования ISDN

Терминальным оборудованием 1 (TE1) может быть цифровой телефон или факс-аппарат. *Контрольная точка S* соответствует точке подключения отдельного терминального устройства к устройству сетевого окончания (устройство типа NT1) или концентратору пользовательских интерфейсов (устройству типа NT2). TE1 по определению поддерживает один из пользовательских интерфейсов ISDN: BRI или PRI.

Если пользовательское терминальное оборудование TE1 подключено через интерфейс BRI, то цифровое абонентское окончание выполняется по 2-проводной схеме (как и обычное окончание аналоговой телефонной сети). Для кодирования данных на участке DSL до точки подключения к сети ISDN (*контрольная точка U*) в этом случае применяется потенциальный код 2B1Q. Дуплексный режим DSL образован путем одновременной передачи сигналов по одной витой паре в обоих направлениях с эхоподавлением и вычитанием своего сигнала из суммарного. Максимальная длина абонентского окончания для этого варианта составляет 5,5 км.

При использовании терминальным оборудованием TE1 интерфейса PRI цифровое абонентское окончание должно представлять собой канал T1 или E1, то есть 4-проводную линию с максимальной длиной около 1800 м. Соответственно на участке DSL до точки U применяется код HDB3 (Европа) или B8ZS (Америка).

Терминальное оборудование 2 (TE2) в отличие от TE1 не поддерживает интерфейсы BRI и PRI. Таким оборудованием может быть компьютер или маршрутизатор с последовательными интерфейсами, не относящимися к ISDN, например RS-232C, X.21 или V.35. Для подключения подобного оборудования к сети ISDN требуется терминальный адаптер. **Терминальный адаптер** (Terminal Adaptor, TA) согласует интерфейс TE2 с интерфейсом PRI или BRI. Для компьютеров терминальные адаптеры выпускаются в формате сетевых адаптеров. *Контрольная точка R* соответствует точке подключения терминального оборудования TE2 к TA. Тип абонентского окончания не зависит от того, работает терминальное оборудование через TA или непосредственно.

Устройства сетевого окончания 2 (NT2) представляют собой устройства канального или сетевого уровня, которые выполняют функции концентрации пользовательских интерфейсов и их мультиплексирования. Например, к этому типу оборудования относятся: офисная АТС, коммутирующая несколько интерфейсов BRI, маршрутизатор, работающий в режиме коммутации пакетов (например, по каналу D), простой мультиплексор TDM, который мультиплексирует несколько низкоскоростных каналов в один канал типа B. Точка подключения оборудования типа NT2 к абонентскому сетевому окончанию (устройству NT1) называется *контрольной точкой T*. Поскольку наличие данного типа оборудования не является обязательным (в отличие от NT1), то контрольные точки S и T объединяются и обозначаются как *контрольная точка S/T*. Физически интерфейс в точке S/T представляет собой 4-проводную линию. Для интерфейса BRI в качестве метода кодирования выбран биполярный метод AMI, причем логическая единица кодируется нулевым потенциалом, а логический ноль — чередованием потенциалов противоположной полярности. Для интерфейса PRI используются другие коды — те же, что и для интерфейсов T1 и E1, то есть соответственно B8ZS и HDB3.

Устройства сетевого окончания 1 (NT1) — это устройства физического уровня, которые согласуют интерфейс BPR или PRI с цифровым абонентские окончанием (DSL), соединяющим пользовательское оборудование с сетью ISDN. Фактически

NT1 представляет собой устройство типа CSU, которое согласует методы кодирования, количество применяемых линий и параметры электрических сигналов. *Контрольная точка U* соответствует точке подключения устройства NT1 к сети.

ПРИМЕЧАНИЕ

> Устройство NT1 может принадлежать оператору сети или пользователю (хотя всегда устанавливается в помещении пользователя). В Европе принято считать устройство NT1 частью сетевого оборудования, поэтому пользовательское оборудование (например, маршрутизатор с интерфейсом ISDN) выпускается без встроенного устройства NT1. В Северной Америке принято считать устройство NT1 принадлежностью пользовательского оборудования, поэтому пользовательское оборудование часто выпускается со встроенным устройством NT1.

Таким образом, для удаленного доступа необходимо оснастить компьютеры пользователей терминальными адаптерами, а в POP установить маршрутизатор, имеющий один или несколько интерфейсов PRI. В этом случае максимальная скорость доступа для отдельного пользователя будет равна скорости передачи двух каналов типа B, то есть 128 Кбит/с. Драйверы терминальных адаптеров ISDN способны объединять два отдельных физических канала типа B в один логический канал. Для этого служит расширение протокола PPP — многоканальный протокол PPP (RFC 1990).

Если пользователь удаленного доступа согласен ограничиться скоростью 64 Кбит/с, он может задействовать второй канал типа B своего интерфейса BRI для параллельной работы телефона ISDN, что невозможно сделать при применении аналогового коммутируемого модема.

Системы управления сетью на основе протокола SNMP (дополнительный материал к «Системы управления сетью на основе протокола SNMP» в главе 23)

Модель «менеджер — агент — управляемый объект» лежит в основе таких популярных стандартов управления, как стандарты Интернета на основе протокола SNMP (Simple Management Network Protocol — простой протокол сетевого администрирования) и стандартов управления ISO/OSI на основе протокола CMIP (Common Management Information Protocol — протокол общей управляющей информации).

Нет ничего более постоянного, чем временное. Протокол SNMP может служить еще одним подтверждением этой азбучной истины. Разработанный как временное и очень простое решение для IP-сетей, он настолько понравился разработчикам оборудования и сетевым администраторам, что на долгие годы стал протоколом №1 в системах управления. И это несмотря на то, что уже давно существует гораздо более мощный (и, соответственно, сложный) протокол CMIP, к тому же являющийся международным стандартом ITU-T.

Однако когда появилась вторая версия протокола (SNMPv2), она не была поддержана производителями сетевого оборудования и распространения не получила. Разработчики стандартов из IETF стараются переломить ситуацию, предложив спецификацию третьей версии (SNMPv3). Существенные улучшения протокола, обеспечивающие гибкое администрирование агентов систем управления и защиту управляющей информации, обратная совместимость с системами на основе базовой версии SNMPv1, а также открытая архитектура позволяют авторам SNMPv3 надеяться на успешное практическое воплощение своего детища.

SNMP — это протокол прикладного уровня, разработанный для стека TCP/IP, хотя имеются его реализации и для других стеков, например IPX/SPX. Протокол SNMP используется для получения от сетевых устройств информации об их статусе, производительности и других характеристиках, которые хранятся в MIB. Простота SNMP во многом определяется простотой баз данных MIB SNMP, особенно их первых версий MIB-I и MIB-II.

Далее перечислены элементы, которые стандартизуются в системах управления, построенных на основе протокола SNMP.

- Протокол взаимодействия агента и менеджера (собственно протокол SNMP).
- Язык описания моделей MIB и SNMP-сообщений — язык абстрактной синтаксической нотации ASN.1 (стандарт ISO 8824:1987, рекомендации ITU-T X.208). Стандарты определяют структуру базы данных MIB, в том числе набор типов ее объектов, их имена и допустимые операции над ними (например, чтение).

- Несколько конкретных моделей MIB (MIB-I, MIB-II, RMON, RMON 2), имена объектов которых регистрируются в дереве стандартов ISO. Древовидная структура MIB содержит обязательные (стандартные) поддеревья, также в ней могут находиться частные поддеревья, позволяющие изготовителю интеллектуальных устройств управлять какими-либо специфическими функциями устройства на основе специфических объектов MIB.

Все остальное отдается «на откуп» разработчику системы управления.

SNMP — это протокол типа «запрос-ответ», то есть на каждый запрос, поступивший от менеджера, агент должен передать ответ. Особенностью протокола является его чрезвычайная простота — он включает в себя всего несколько команд.

- Команда `Get-request` используется менеджером для получения от агента значения какого-либо объекта по его имени.
- Команда `GetNext-request` используется менеджером для извлечения значения следующего объекта (без указания его имени) при последовательном просмотре таблицы объектов.
- С помощью команды `Get-response` SNMP-агент передает менеджеру ответ на команду `Get-request` или `GetNext-request`.
- Команда `Set` позволяет менеджеру изменять значения какого-либо объекта. С помощью команды `Set` и происходит собственно управление устройством. Агент должен «понимать» смысл значений объекта, который используется для управления устройством, и на основании этих значений выполнять реальное управляющее воздействие — отключить порт, приписать порт определенной линии VLAN и т. п. Команда `Set` пригодна также для задания условия, при выполнении которого SNMP-агент должен послать менеджеру соответствующее сообщение. Может быть определена реакция на такие события, как инициализация агента, рестарт агента, обрыв связи, восстановление связи, неверная аутентификация, потеря ближайшего маршрутизатора. Если происходит любое из этих событий, то агент инициализирует прерывание.
- Команда `Trap` используется агентом для сообщения менеджеру о возникновении особой ситуации.

Версия SNMP v.2 добавляет к этому набору команду `GetBulk`, которая позволяет менеджеру получить несколько переменных за один запрос.

Структура SNMP MIB

На сегодня существует несколько стандартов на базы данных управляющей информации для протокола SNMP. Основными являются стандарты MIB-I и MIB-II, а также версия базы данных для удаленного управления RMON (Remote Monitoring) MIB. Кроме того, существуют стандарты для специальных MIB-устройств конкретного типа (например, MIB для концентраторов или MIB для модемов), а также частные базы данных MIB конкретных фирм-производителей оборудования.

Первоначальная спецификация MIB-I определяла только операции чтения значений переменных. Операции изменения или установки значений объекта являются частью спецификаций MIB-II.

Версия MIB-I (RFC 1156) определяет 114 объектов, которые подразделяются на 8 групп.

- *System* — общие данные об устройстве (например, идентификатор поставщика, время последней инициализации системы).
- *Interfaces* — параметры сетевых интерфейсов устройства (например, их количество, типы, скорости обмена, максимальный размер пакета).
- *Address Translation Table* — описание соответствия между сетевыми и физическими адресами (например, по протоколу ARP).
- *Internet Protocol* — данные, относящиеся к протоколу IP (адреса IP-шлюзов, хостов, статистика о IP-пакетах).
- *ICMP* — данные, относящиеся к протоколу обмена управляющими ICMP-сообщениями.
- *TCP* — данные, относящиеся к протоколу TCP (например, о TCP-соединениях).
- *UDP* — данные, относящиеся к протоколу UDP (число переданных, принятых и ошибочных UPD-дейтаграмм).
- *EGP* — данные, относящиеся к протоколу EGP, используемому в Интернете (число принятых с ошибками и без ошибок сообщений).

Из этого перечня групп переменных видно, что стандарт MIB-I разрабатывался с жесткой ориентацией на управление маршрутизаторами, поддерживающими протоколы стека TCP/IP.

В версии MIB-II (RFC 1213), принятой в 1992 году, был существенно (до 185) расширен набор стандартных объектов, а число групп увеличилось до 10.

На рис. 1 приведен пример древовидной структуры базы объектов MIB-II. На нем показаны две из 10 возможных групп объектов — System (имена объектов начинаются с префикса `Sys`) и Interfaces (префикс `if`).

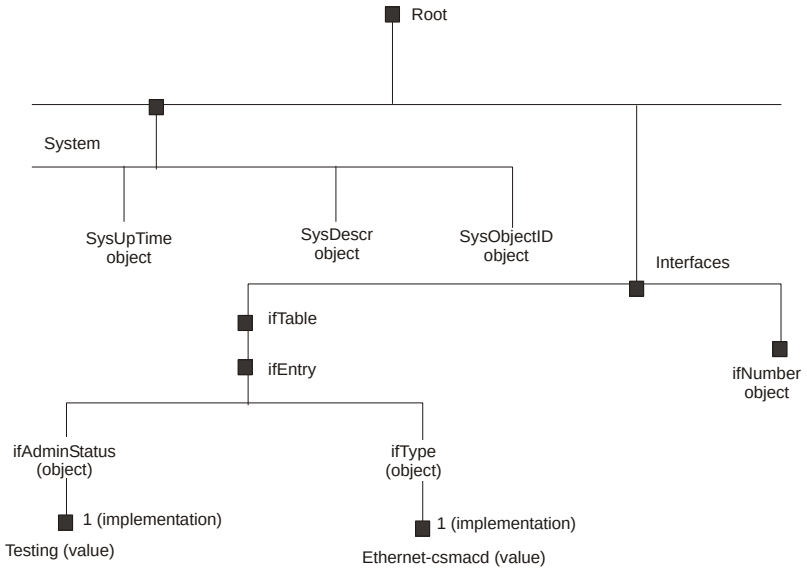

Рис. 1. Стандартное дерево MIB-II (фрагмент)

Объект `SysUpTime` содержит значение продолжительности времени работы системы с момента последней перезагрузки, объект `SysObjectID` — идентификатор устройства (например, маршрутизатора).

Объект `ifNumber` определяет количество сетевых интерфейсов устройства, а объект `ifEntry` является вершиной поддерева, описывающего один из конкретных интерфейсов устройства. Входящие в это поддерево объекты `ifType` и `ifAdminStatus` определяют соответственно тип и состояние одного из интерфейсов, в данном случае интерфейса Ethernet.

Далее перечислены объекты, описывающие конкретные интерфейсы устройства.

- `ifType` — тип протокола, который поддерживает интерфейс. Этот объект принимает значения всех стандартных протоколов канального уровня, например rfc877-x25, ethernet-csmacd, iso88023-csmacd, iso88024-tokenBus, iso88025-tokenRing и т. д.
- `ifMtu` — максимальный размер пакета сетевого уровня, который можно послать через этот интерфейс.
- `ifSpeed` — пропускная способность интерфейса в битах в секунду (100 для Fast Ethernet).
- `ifPhysAddress` — физический адрес порта, для Fast Ethernet им будет MAC-адрес.

- `ifAdminStatus` — желаемый статус порта:
  - `up` — готов передавать пакеты;
  - `down` — не готов передавать пакеты;
  - `testing` — находится в тестовом режиме.
- `ifOperStatus` — фактический текущий статус порта, имеет те же значения, что и `ifAdminStatus`.
- `ifInOctets` — общее количество байтов, принятое данным портом, включая служебные, с момента последней инициализации SNMP-агента.
- `ifInUcastPkts` — количество пакетов с индивидуальным адресом интерфейса, доставленных протоколу верхнего уровня.
- `ifInNUcastPkts` — количество пакетов с широковещательным или групповым адресом интерфейса, доставленных протоколу верхнего уровня.
- `ifInDiscards` — количество пакетов, которые были приняты интерфейсом, оказались корректными, но не были доставлены протоколу верхнего уровня, скорее всего из-за переполнения буфера пакетов или же по иной причине.
- `ifInErrors` — количество пришедших пакетов, которые не были переданы протоколу верхнего уровня из-за обнаружения в них ошибок.

Помимо объектов, описывающих статистику по входным пакетам, имеются аналогичные объекты, но относящиеся к выходным пакетам.

Как видно из описания объектов MIB-II, эта база данных не дает детальной статистики по характерным ошибкам кадров Ethernet, кроме того, она не отражает изменение характеристик во времени, что часто интересует сетевого администратора. Эти ограничения были впоследствии сняты новым стандартом на MIB — RMON MIB, который специально ориентирован на сбор детальной статистики по протоколу Ethernet. Возможности RMON MIB включают также построение временны*х* зависимостей значений параметров.

Для именования переменных базы MIB и однозначного определения их форматов используется дополнительная спецификация, называемая SMI (Structure of Management Information — структура управляющей информации). Например, спецификация SMI включает в качестве стандартного имя `IpAddress` и определяет его формат как строку из 4 байт. Другой пример — имя `Counter`, для которого определен формат в виде целого числа в диапазоне от 0 до 2^{32}–1.

Имена переменных MIB могут быть записаны как в символьном, так и в числовом форматах. Символьный формат используется для представления переменных в текстовых документах и на экране дисплея, а числовой — в сообщениях протокола SNMP. Например, символьному имени `SysDescr` соответствует числовое имя 1.3.6.1.2.1.1.1.

Составное числовое имя объекта базы данных MIB протокола SNMP соответствует полному имени этого объекта в дереве регистрации объектов стандартизации ISO. Разработчики протокола SNMP не стали использовать традиционный для стандартов

Интернета способ фиксации числовых параметров протокола в специальном документе RFC. Вместо этого они зарегистрировали объекты баз данных MIB протокола SNMP во всемирном дереве регистрации стандартов ISO (рис. 2).

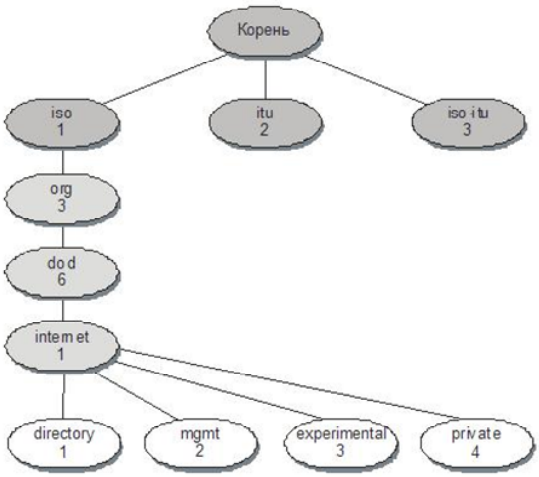

Рис. 2. Пространство имен объектов ISO

Как и в любых сложных системах, пространство имен объектов ISO имеет древовидную иерархическую структуру, причем на рисунке показана только его верхняя часть. От корня этого дерева отходят три ветви, соответствующие стандартам, контролируемым ISO, ITU и совместно ISO-ITU. В свою очередь, организация ISO создала ветвь для стандартов, создаваемых национальными и международными организациями (ветвь `org`). Стандарты Интернета создавались под эгидой Министерства обороны (Department of Defense, DoD) США, поэтому стандарты MIB попали в поддерево `dod-internet`, а далее, естественно, в группу стандартов управления сетью — ветвь `mgmt`. Объекты любых стандартов, создаваемых под эгидой ISO, однозначно идентифицируются составными символьными именами, начинающимися от корня этого дерева. В сообщениях протоколов используются не символьные имена, а однозначно соответствующие им составные числовые имена. Каждая ветвь дерева имен объектов нумеруется в дереве целыми числами слева направо, начиная с единицы, и эти числа и заменяют символьные имена. Поэтому полному символьному имени объекта MIB `iso.org.dod.internet.mgmt.mib` соответствует полное числовое имя — 1.3.6.1.2.1.

Группа объектов `private` (4) зарезервирована за стандартами, создаваемыми частными компаниями, например Cisco, Hewlett-Packard и т. п. Это же дерево регистрации используется для именования классов объектов CMIP и TMN.

Соответственно, каждая группа объектов MIB-I и MIB-II помимо приведенных кратких символьных имен имеет полные символьные имена и соответствующие им числовые имена.

Формат SNMP-сообщений

Протокол SNMP обслуживает передачу данных между агентами и менеджерами. SNMP использует дейтаграммный транспортный протокол UDP, не обеспечивающий надежной доставки сообщений. Протокол, организующий надежную передачу дейтаграмм на основе соединений TCP, весьма загружает управляемые устройства, которые на момент разработки протокола SNMP были не очень мощными, поэтому от услуг протокола TCP решили отказаться.

SNMP-сообщения, в отличие от сообщений многих других коммуникационных протоколов, не имеют заголовков с фиксированными полями. SNMP-сообщение состоит из произвольного количества полей, и каждое поле предваряется описателем его типа и размера.

Любое SNMP-сообщение состоит из трех основных частей: *версии протокола*, *идентификатора общности* и *области данных*.

Идентификатор общности (community string) служит для группирования устройств, управляемых определенным менеджером. Идентификатор общности является аналогом пароля, так как для того, чтобы устройства могли взаимодействовать по протоколу SNMP, они должны иметь одно и то же значение этого идентификатора (по умолчанию часто используется строка «public»).

В области данных, собственно, и содержатся описанные ранее команды протокола, имена объектов и их значения. Область данных состоит из одного или более блоков PDU, каждый из которых может относиться к одному из пяти различных типов PDU, соответствующих пяти командам протокола SNMP: `GetRequest-PDU`, `GetNextRequest-PDU`, `GetResponse-PDU`, `SetRequest-PDU`, `Trap-PDU`. И, наконец, для каждого типа PDU имеется определение его формата. Например, формат блока `GetRequest-PDU` включает следующие поля:

- идентификатор запроса;
- статус ошибки (есть или нет);
- индекс ошибки (тип ошибки, если она есть);
- список имен объектов SNMP MIB, включенных в запрос.

На рис. 3 показано сообщение протокола SNMP, которое представляет собой запрос о значении объекта `SysDescr` (числовое имя 1.3.6.1.2.1.1.1).

30	29	02	01	00			
SEQUENCE	len = 41	INTEGER	len = 1	vers = 0			
04	06	70	75	62	6C	69	63
string	len = 6	p	u	B	l	i	c
A0	1C	02	04	05	AE	56	02
getreq	len = 28	INTEGER	len = 4	---------	request ID	--------	---
02	01	00	02	01	00		
INTEGER	len = 1	status	INTEGER	len = 1	error	index	
30	0E	30	0C	06	08		
SEQUENCE	len = 14	SEQUENCE	len = 12	Objectid	len = 8		
2B	06	01	02	01	01	01	00
1.3	6	1	2	1	1	1	0
05	00						
null	len=0						

Рис. 3. Пример SNMP-сообщения

Как видно из рисунка, сообщение начинается с кода 30 (все коды шестнадцатеричные), который соответствует ключевому слову `SEQUENCE` (последовательность) и говорит о том, что сообщение состоит из последовательности полей. Длина последовательности указывается в следующем байте (41 байт). Далее следует поле, которое представляет собой целое число (integer) длиной 1 байт — это версия (vers) протокола SNMP (в данном случае 0, то есть SNMP v.1, а 1 означала бы SNMP v.2). Поле идентификатора общности `community` имеет тип `string` (строка символов) длиной в 6 байт со значением `public`. Остальную часть сообщения составляет блок данных `GetRequest-PDU`. То, что это операция `Get-request`, говорит код A0, а общая длина этого блока данных равна 28 байт. В соответствии со структурой блока данных `Getrequest-PDU` далее идет поле идентификатора запроса (он определен как целое 4-байтное число и имеет значение 05 AE 56 02). Затем в блоке следуют два однобайтовых целых числа статуса и индекса ошибки, которые в запросе установлены в 0. И, наконец, завершает сообщение список имен объектов, значения которых запрашиваются данной командой. Этот список в примере состоит из одной переменной с именем 1.3.6.1.2.1.1.1.0, которое соответствует символьному имени SysDescr. Признак `null` (значение 05) говорит о том, что достигнут конец сообщения.

Спецификация RMON базы данных MIB

Добавлением к функциональным возможностям SNMP является спецификация RMON, которая обеспечивает удаленное взаимодействие с базой MIB. До появления RMON протокол SNMP не мог использоваться удаленным образом, он допускал только локальное администрирования устройств. База RMON MIB обладает улуч-

шенным набором свойств для удаленного администрирования, так как содержит агрегированную информацию об устройстве, не требующую передачи по сети больших объемов данных. Объекты RMON MIB включают дополнительные счетчики ошибок в пакетах, более гибкие средства анализа трендов и статистики, более мощные средства фильтрации для захвата и анализа отдельных пакетов, а также более сложные условия установления сигналов предупреждения. Интеллектуальность агентов RMON MIB выше, чем агентов MIB-I или MIB-II, что позволяет им выполнять значительную часть работы по обработке информации об устройстве, которую раньше выполняли менеджеры. Эти агенты могут располагаться внутри различных коммуникационных устройств или выполняться в виде отдельных программных модулей на универсальных персональных компьютерах и ноутбуках.

Объекту RMON присвоен номер 16 в наборе объектов MIB, а сам объект RMON объединяет 10 групп объектов (десятую группу составляют специальные объекты протокола Token Ring).

10. `Statistics` — текущие накопленные статистические данные о характеристиках пакетов, количестве коллизий и т. п.
11. `History` — статистические данные, сохраненные через определенные промежутки времени для последующего анализа тенденций их изменений.
12. `Alarms` — пороговые значения статистических показателей, при превышении которых агент RMON посылает сообщение менеджеру.
13. `Hosts` — данные о хостах сети, в том числе их MAC-адресах.
14. `Host TopN` — таблица наиболее загруженных хостов сети.
15. `Traffic Matrix` — статистика об интенсивности трафика между каждой парой хостов сети, упорядоченная в виде матрицы.
16. `Filter` — условия фильтрации пакетов.
17. `Packet Capture` — условия захвата пакетов.
18. `Event` — условия регистрации и генерации событий.

Данные группы пронумерованы в указанном порядке, поэтому, например, группа `Hosts` имеет числовое имя 1.3.6.1.2.1.16.4.

Всего стандарт RMON MIB определяет около 200 объектов в 10 группах, зафиксированных в двух документах — RFC 1271 для сетей Ethernet и RFC 1513 для сетей Token Ring.

Отличительной чертой стандарта RMON MIB является его независимость от протокола сетевого уровня (в отличие от стандартов MIB-I и MIB-II, ориентированных на протоколы TCP/IP). Поэтому он удобен для гетерогенных сред, использующих различные протоколы сетевого уровня.

Рассмотрим более подробно группу `Statistics`, которая определяет, какую информацию о кадрах (называемых в стандарте пакетами) Ethernet может предоставить агент RMON. Группа `History` основана на объектах группы `Statistics`, так как ее объекты позволяют просто строить временные ряды для объектов группы `Statistics`.

В группу Statistics входят наряду с некоторыми другими следующие объекты:

- etherStatsDropEvents — общее число событий, при которых пакеты были проигнорированы агентом из-за недостатка его ресурсов (сами пакеты при этом не обязательно были потеряны интерфейсом);
- etherStatsOctets — общее число байтов (включая ошибочные пакеты), принятых из сети (исключая преамбулу и включая байты контрольной суммы);
- etherStatsPkts — общее число полученных пакетов (включая ошибочные);
- etherStatsBroadcastPkts — общее число хороших пакетов, которые были посланы по широковещательному адресу;
- etherStatsMulticastPkts — общее число хороших пакетов, полученных по групповому адресу;
- etherStatsCRCAlignErrors — общее число полученных пакетов, которые имели длину (исключая преамбулу) в диапазоне между 64 и 1518 байт, не содержали целое число байтов или имели неверную контрольную сумму;
- etherStatsUndersizePkts — общее число пакетов, которые имели длину меньше, чем 64 байт, но были правильно сформированы;
- etherStatsOversizePkts — общее число полученных пакетов, которые имели длину больше, чем 1518 байт, но были тем не менее правильно сформированы;
- etherStatsFragments — общее число полученных пакетов, которые не состояли из целого числа байтов или имели неверную контрольную сумму и имели к тому же длину, меньшую 64 байт;
- etherStatsJabbers — общее число полученных пакетов, которые не состояли из целого числа байтов или имели неверную контрольную сумму и имели к тому же длину, большую 1518 байт;
- etherStatsCollisions — наилучшая оценка числа коллизий на данном сегменте Ethernet;
- etherStatsPkts64Octets — общее количество полученных пакетов (включая плохие) размером 64 байт;
- etherStatsPkts65to127Octets — общее количество полученных пакетов (включая плохие) размером от 65 до 127 байт;
- etherStatsPkts128to255Octets — общее количество полученных пакетов (включая плохие) размером от 128 до 255 байт;
- etherStatsPkts256to511Octets — общее количество полученных пакетов (включая плохие) размером от 256 до 511 байт;
- etherStatsPkts512to1023Octets — общее количество полученных пакетов (включая плохие) размером от 512 до 1023 байт;
- etherStatsPkts1024to1518Octets — общее количество полученных пакетов (включая плохие) размером от 1024 до 1518 байт.

Как видно из описания объектов, с помощью агента RMON, встроенного в повторитель или другое коммуникационное устройство, можно провести очень детальный анализ работы сегмента Ethernet или Fast Ethernet. Сначала можно получить данные о встречающихся в сегменте типах ошибок в кадрах, а затем целесообразно собрать с помощью группы `History` зависимости интенсивности этих ошибок от времени (в том числе привязав их ко времени). После анализа временны*х* зависимостей часто уже можно сделать некоторые предварительные выводы об источнике ошибочных кадров и на этом основании сформулировать более тонкие условия захвата кадров со специфическими признаками (задав условия в группе `Filter`), соответствующими выдвинутой версии. После этого можно провести еще более детальный анализ путем изучения захваченных из объектов группы `Packet Capture` кадров.

Позже был принят стандарт RMON 2, который распространяет идеи интеллектуальной базы RMON MIB на протоколы верхних уровней, выполняя часть работы анализаторов протоколов.

Недостатки протокола SNMP

Протокол SNMP служит основой многих систем администрирования, хотя имеет несколько принципиальных недостатков.

- *Отсутствие средств взаимной аутентификации агентов и менеджеров*. Единственным средством, которое можно было бы отнести к средствам аутентификации, является так называемая строка общности в сообщениях. Эта строка передается по сети в открытой форме в SNMP-сообщении и служит основой для объединения агентов и менеджеров, так что агент взаимодействует только с теми менеджерами, у которых та же строка общности, что и строка, хранящаяся в памяти агента. Это, безусловно, не способ аутентификации, а способ структурирования агентов и менеджеров. Версия SNMP v.2 должна была ликвидировать этот недостаток, но в результате разногласий между разработчиками стандарта новые средства аутентификации хотя и появились в этой версии, но как необязательные.
- *Работа через ненадежный протокол UDP* (а именно так работает подавляющее большинство реализаций агентов SNMP) приводит к потерям аварийных сообщений от агентов к менеджерам, что может привести к некачественному администрированию. Исправление ситуации путем перехода на надежный транспортный протокол с установлением соединения чревато потерей связи с огромным количеством встроенных агентов SNMP, имеющихся в установленном в сетях оборудовании. (Протокол CMIP изначально работает поверх надежного транспорта стека OSI и этим недостатком не страдает.)

Разработчики платформ администрирования стараются преодолеть эти недостатки. Например, в системе HP OV Telecom DM TMN, являющейся платформой для разработки многоуровневых систем администрирования в соответствии со стандартами ISO, работает новая реализация SNMP, организующая надежный обмен сообщениями между агентами и менеджерами за счет самостоятельной организации повторных передач SNMP-сообщений при их потере.