ТЕМА 5
Технология Интернет

5.1 Особенности организации Интернет

Интернет – самая необычная из телекоммуникационных сетей. Практически любой объект может подключиться к Интернет, чтобы предложить собственные ресурсы или получить доступ к ресурсам сети. По Интернет может распространяться любой вид информации без каких-либо ограничений. ***Отсутствует*** центральный ***орган***, который регулировал бы работу сети Интернет, хотя существуют организации, устанавливающие определенные фундаментальные принципы и руководящие работой сети.

Сеть Интернет по своей природе ***обособлена*** от других сетей, а по своей ***философии*** является ***анархической.*** Про некоторые новые сервисы (Скайп) даже утверждается, что сигнал передается «по облаку». Поражает способность сети Интернет к самоорганизации. Никто не знает, ***сколько узлов*** подключено к сети и ***где они находятся***, но передаваемые по сети пакеты ***находят свой путь*** до пункта назначения. Более того, если канал или маршрутизатор выходят из строя, то пакеты ***автоматически отыскивают*** другие пути до адресата.

Интернет как целостная система мало занимается ***финансовыми*** вопросами. Никто централизованно не платит за Интернет, и нет такой организации, которая собирает дань со всех сегментов Интернета или с пользователей. Вместо этого каждый затрачивает средства на ***свою часть***, чтобы она могла работать и взаимодействовать с соседними зонами. Представители ***непосредственно связанных*** сегментов сетей сами решают, как им следует соединяться и как финансировать построенное соединение.

Совет по архитектуре Интернет (IAB) – добровольный орган, координирует работу двух рабочих групп IFTF и IRTF, вырабатывающих технологическую политику и направления работы.

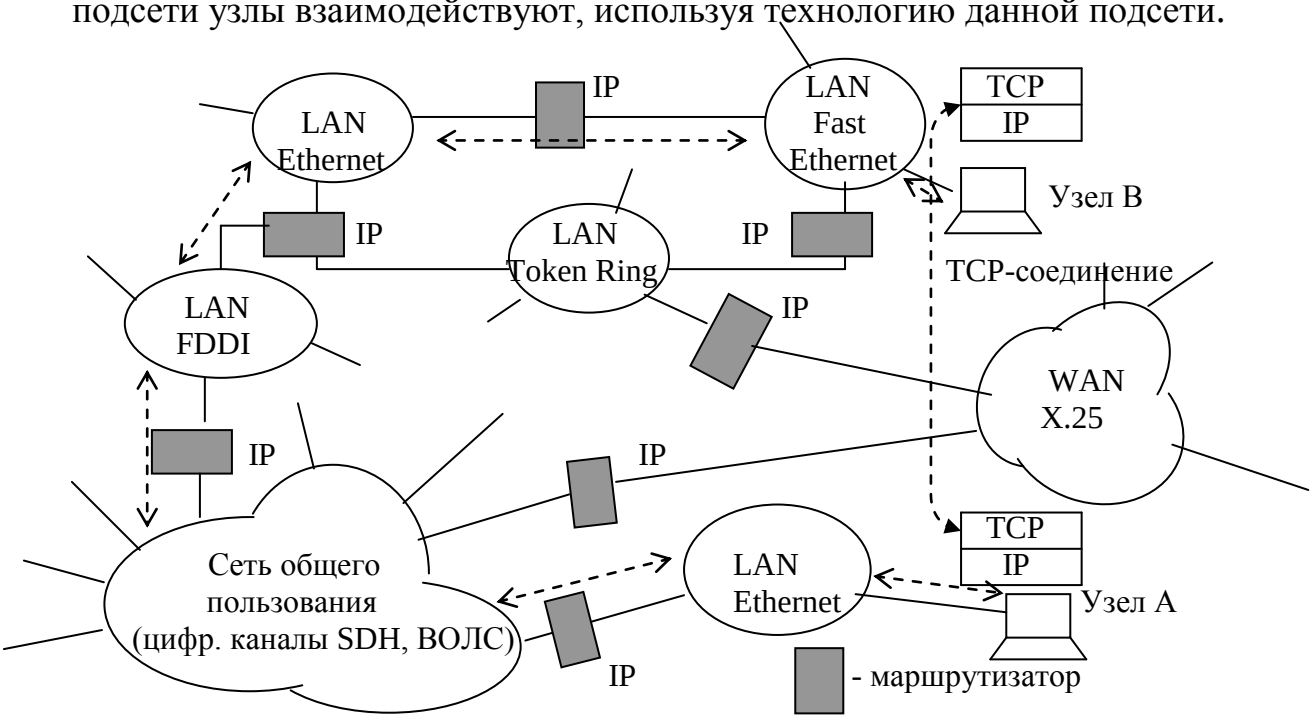
Главная группа – ***группа управления инженеров*** Интернет (IESG) – утверждает стандарты по протоколам, архитектуре и эксплуатации, подготавливаемые ***инженерной*** проблемной группой Интернет (IETF).

Научно-исследовательская ***проблемная*** группа Интернет (IRTF) занимается ***долгосрочными*** вопросами, включая схемы адресации и технологии.

Все стандарты Интернет носят название RFC (Request for Comments) – ***запрос на комментарии***, что подчеркивает гласный и открытый характер принимаемых стандартов.

Интернет – совокупность тысяч компьютеров, объединенных в локальные или глобальные сети, которые, в свою очередь, соединены посредством ***маршрутизаторов*** в интерсеть – internet (см. рисунок).

Внутренняя структура каждой сети на рисунке не показана, так как она не имеет значения для работы составной сети. В пределах каждой подсети узлы взаимодействуют, используя технологию данной подсети.



Сеть Интернет – объединение локальных и глобальных сетей на основе стека протоколов TCP/IP

Так, в составную сеть, показанную на рисунке, входит несколько сетей разных технологий: локальные сети Ethernet, Fast Ethernet, Token Ring, FDDI и глобальные сети frame relay, X.25, ATM. Каждая из этих технологий способна обеспечить взаимодействие всех узлов ***в своей подсети***, но ***не способна*** организовать информационную связь между произвольно выбранными узлами**,** принадлежащим ***разным подсетям***, например, между узлом A и узлом B на рисунке.

Технология Интернет весьма примитивна. Она соединяет множество неоднородных компьютерных систем, заявки которых сеть старается выполнить наиболее результативно.

5.2 Стек протоколов TCP/IP (Transmission Control Protocol/ Internet Protocol)

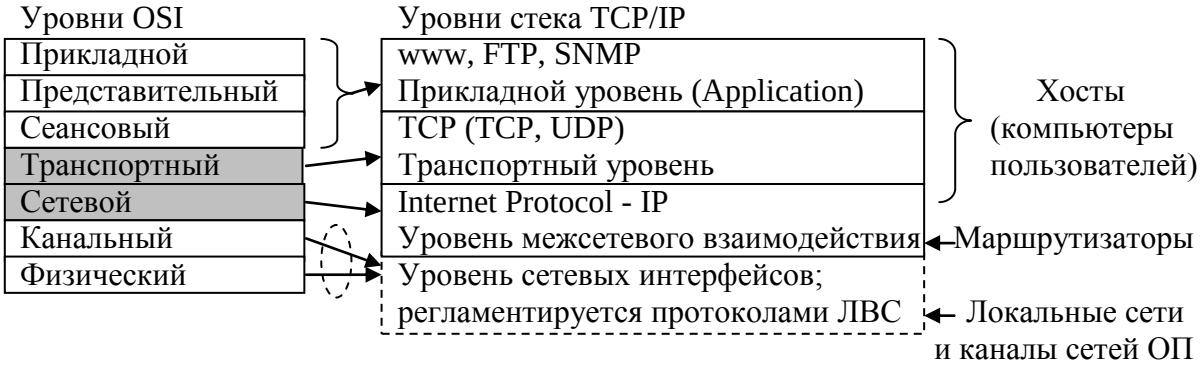
Основой сети Интернет является стек (набор) протоколов TCP/IP.

Стек был разработан по инициативе Министерства обороны США более 30 лет назад ***для связи*** экспериментальной сети ARPANET (заказчик – Advanced Research Projects Agency) с ***другими сетями*** как набор ***общих*** протоколов для ***разнородной*** вычислительной среды.

Стек (набор протоколов) TCP/IP был разработан до появления модели взаимодействия открытых систем ISO/OSI, поэтому соответствие уровней стека TCP/IP уровням модели OSI достаточно условно (см. рисунок). В стеке TCP/IP определено 4 уровня.

Основными оригинальными протоколами стека TCP/IP являются протоколы ***межсетевого (IP)*** и ***транспортного (TCP)*** уровней, соответствующие ***сетевому*** и ***транспортному*** уровням в терминологии OSI. Именно они определяют маршрут, обеспечивая передачу данных от отправителя к получателям через объединенную систему компьютерных сетей и, если это требуется, гарантируют надежность доставки пакетов. При этом протокол ***IP*** обеспечивает ***продвижение*** пакета по ***составной*** сети**,** а протокол ***TCP*** гарантирует ***надежность*** доставки этого пакета**.**

Остальные средства стека TCP/IP выполняют ***вспомогательную*** роль – сервисы верхнего уровня поддерживают интерфейс с пользователями и приложениями, а нижний уровень – интерфейс с технологиями составляющих сетей.



Стек протоколов TCP/IP

Протоколы ***сетевого*** и ***транспортного*** уровней стека TCP/IP выполняют следующие функции:

- определяют ***маршрут***;
- обеспечивают ***передачу*** данных от отправителя к получателям ***через объединенную систему*** компьютерных сетей;
- гарантируют ***надежность*** доставки пакетов (если это требуется).

Протоколы ***сетевого*** уровня реализуются, как правило, в виде программных модулей, устанавливаемых на промежуточных узлах – ***маршрутизаторах*** и конечных узлах пользователей (серверах и рабочих станциях). Протоколы ***транспортного*** уровня реализуются на ***конечных*** узлах сети.

Прикладной (Application) уровень стека TCP/IP соответствует ***трем верхним уровням*** модели OSI; прикладному, представительному и сеансовому. Прикладной уровень объединяет службы, предоставляемые системой ***пользовательскими приложениями***.

К ним относятся такие широко используемые протоколы, как протокол копирования файлов (File Transfer Protocol, FTP), простой протокол передачи электронной почты (Simple Mail Transfer Protocol, SMTP), протокол передачи гипертекстовой информации (Hypertext Transfer Protocol, HTTP) и многие другие.

Протоколы прикладного уровня устанавливаются на ***хостах*** и на ***рабочих станциях***.

Стек TCP/IP на нижнем уровне ***поддерживает все популярные стандарты физического и канального уровней*** для локальных и глобальных сетей.

Рассмотрим более подробно свойства протоколов отдельных уровней стека TCP/IP.

5.3 Уровень межсетевого взаимодействия – Internet Protocol

Уровень межсетевого взаимодействия (internet), называемый также сетевым (network) уровнем, является ***стержнем всей архитектуры TCP/IP***. Именно этот уровень, функции которого соответствуют сетевому уровню модели OSI, обеспечивает ***перемещение*** пакетов в пределах всей составной сети.

Протокол межсетевого взаимодействия (***Internet Protocol, IP***) ***изначально*** создавался для глобальной ***составной*** сети Интернет, поэтому он обладает рядом ***специфических*** функций. Первым отличительным свойством IP-протокола является его способность ***фрагментировать пакеты***.

Большая составная сеть часто состоит из сетей, построенных на совершенно разных принципах. В эксплуатируемых типах локальных и глобальных сетей максимальный размер поля данных, в которых должен инкапсулировать свой пакет IP, значительно отличаются:

- сети Ethernet – 1500 байт;
- сети Token Ring – 4116 байт;
- FDDI – 4096 байт;
- X.25 – 128 байт;
- Frame Relay – 4056 байт.

При переходе из сети, имеющей большую максимальную длину кадра, в сеть с меньшей максимальной длиной может возникнуть необходимость ***разделения передаваемого кадра на несколько частей***. Протокол IP стека TCP/IP эффективно решает эту задачу.

Однако IP-маршрутизаторы ***не собирают*** фрагменты пакетов в более крупные пакеты, даже если на пути встречается сеть, допускающая такое укрупнение. Это связано с тем, что отдельные фрагменты сообщения могут перемещаться по составной сети по разным маршрутам, поэтому нет гарантии, что все фрагменты проходят на своем пути через данный маршрутизатор.

Второй особенностью технологии TCP/IP является ***гибкая система адресации***, позволяющая проще, чем другие протоколы аналогичного назначения включать в интерсеть ***сети разных технологий***. Это свойство также способствует применению стека TCP/IP для построения больших сетей.

Как отмечалось ранее, технологии локальных сетей используют одну и ту же «плоскую» систему адресации узлов, реализуемую обычно на основе MAC-адресов. Чтобы ***сетевой*** уровень мог обеспечить перемещение пакета между ***различными*** сетями, ему необходима

собственная система адресации, которая могла однозначно идентифицировать любой узел ***составной*** сети. Естественным способом формирования сетевого адреса является уникальная ***нумерация*** всех ***подсетей*** и ***нумерация*** всех ***узлов*** в пределах каждой подсети. Таким образом, ***сетевой адрес*** представляет собой ***пару: номер сети (подсети) и номер узла.***

В то время как в ATM, например, предусматривается организация ***единой*** глобальной сети, в технологии TCP/IP внешне непритязательно ставится более ограниченная задача обеспечения межсетевого взаимодействия.

Третьей особенностью технологии TCP/IP является распространение метода ***передачи дейтаграмм***, использовавшегося ранее только в локальных сетях для ***однородных*** структур, на составные сети ***произвольной*** структуры.

Все вопросы обеспечения надежности доставки данных в составной сети решает протокол TCP, основанный на установлении ***логических соединений*** между взаимодействующими процессами.

Протокол IP – это ***дейтаграммный*** протокол, работающий ***без установления соединений*** по принципу «по возможности», в соответствии с которым он не берет на себя ***ответственность за доставку*** пакета до узла назначения. Каждый IP-пакет передается как ***независимая*** единица, не имеющая связи ни с какими-либо другими IP-пакетами.

Если же по каким-либо причинам пакет теряется (например, из-за переполнения буфера), протокол IP не пытается повторить его передачу.

Максимум на что он способен – это ***послать уведомление*** о потере пакета узлу-отправителю.

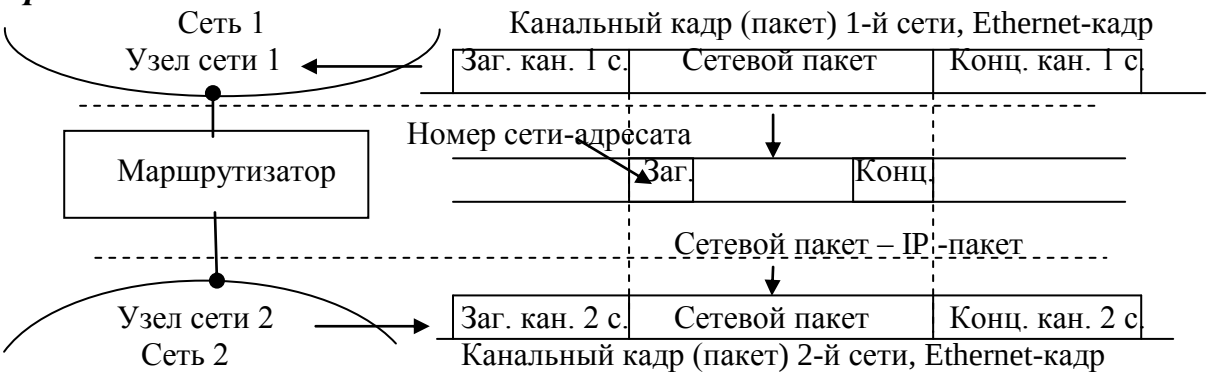
Важнейшей задачей сетевого уровня является ***маршрутизация*** – передача пакетов между двумя конечными узлами в составной сети. Связи между сетями осуществляются маршрутизаторами. Маршрутизатор – ***совокупность нескольких узлов***, каждый из которых входит в свою сеть. Как единое устройство маршрутизатор не имеет ни отдельного сетевого адреса, ни какого-либо локального адреса. Основная функция маршрутизатора – ***чтение*** заголовков сетевых пакетов и ***принятие решения*** о дальнейшем маршруте следования по его сетевому адресу.

Данные, которые необходимо передать через составную сеть, снабжаются заголовком сетевого уровня. Данные вместе с заголовком образуют пакет. ***Максимальная*** длина пакета – ***65535 байт ≈ 64 кбайт***. Заголовок обычно имеет ***длину 20 байт*** и содержит информацию о сетевых адресах отправителя и получателя, о параметрах фрагментации, о времени жизни пакета, о контрольной сумме и некоторых других.

В поле данных IP-пакета находятся сообщения более высокого уровня, например TCP или UDP.

Каждый раз, когда пакет сетевого уровня попадает на маршрутизатор, включенный между двумя подсетями, он ***извлекается из кадра первой подсети*** (см. рисунок) – освобождается от канального заголовка этой сети, и ***упаковывается*** в новый ***кадр*** – снабжается новым заголовком ***канального уровня следующей подсети***. Информацию, на основе которой делается эта замена, предоставляют служебные поля пакета сетевого уровня. В поле адреса назначения нового кадра указывается локальный адрес следующего маршрутизатора.

Основным полем ***заголовка сетевого*** уровня является ***номер сети-адресата.***



Передача пакета сетевого уровня между сетями

Маршрут – это ***последовательность маршрутизаторов***, которые должен пройти пакет ***от отправителя до получателя***.

Маршрутизаторы и конечные узлы решают задачу выбора маршрута из нескольких возможных на основе ***таблиц маршрутизации***.

Записи в таблицу могут вноситься вручную администратором сети и автоматически протоколами маршрутизации. Для этого маршрутизаторы сети обмениваются специальной служебной ***информацией о топологии составной сети***.

Протоколы маршрутизации генерируют для каждого маршрутизатора ***согласованные*** друг с другом таблицы маршрутизации, которые позволят обеспечить доставку пакета по рациональному маршруту от исходной сети в сеть назначения за конечное число шагов.

Маршрутизатор представляет собой сложное многофункциональное устройство, в задачи которого входит:

- построение таблицы маршрутизации;
- определение на ее основе маршрута;
- буферизация, фрагментация и фильтрация (уничтожение пакета при обнаружении ошибок или окончания времени жизни) поступающих пакетов;
- поддержка сетевых интерфейсов.

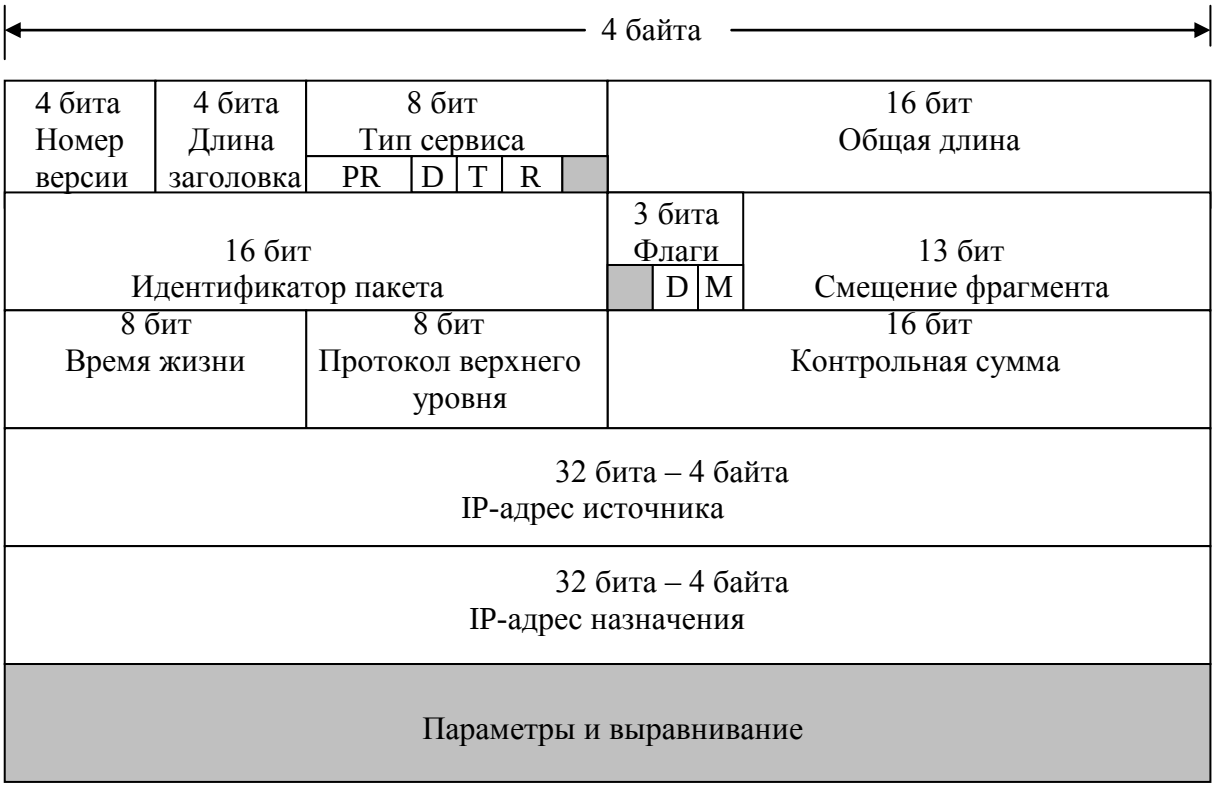
5.4 Структура IP-пакета

Для более четкого представления об операциях, реализуемых при транспортировке пакета по сети Интернет, рассмотрим структуру пакета

сетевого уровня. Пакет сетевого уровня состоит из заголовка и поля данных. Заголовок, как правило, имеет длину 20 байт. Структура заголовка показана на рисунке.

Поле ***Номер версии***, занимающее 4 бита, указывает версию протокола IP. Сейчас повсеместно используется версия 4 (IPv4) и начался переход на версию 6 (IPv6).

Поле ***Длина заголовка*** IP-пакета занимает 4 бита и указывает значение длины заголовка, измеренное в 32-битовых словах. Обычно заголовок имеет длину в ***20 байт (пять 32-битовых слов)***, но при увеличении объема служебной информации эта длина может быть увеличена за счет использования дополнительных байтов в поле ***Параметры***. Заголовок наибольшего объема занимает 60 байт (15 32-битовых слов).



Структура заголовка IP-пакета для IPv4

Поле ***Тип сервиса*** занимает один байт. Это поле позволяет приложениям влиять на качество обслуживания, задавая приоритетность пакета и критерий выбора маршрута. Первые 3 бита этого поля образуют подполе ***приоритета*** пакета. Приоритет может иметь значения от самого низкого – 0 (нормальный пакет) до самого высокого – 7 (пакет управляющей информации). Маршрутизаторы и компьютеры могут принимать во внимание приоритет пакета и обрабатывать более важные пакеты в первую очередь. Следующие 3 бита определяют критерий выбора маршрута: минимальная задержка (D), высокая пропускная

способность (T), максимальная надежность доставки (R). Оставшиеся 2 бита зарезервированы.

Поле ***Общая длина*** занимает 2 байта и указывает общую длину пакета ***с учетом*** заголовка и поля данных. Максимальная длина пакета ограничена разрядностью этого поля и составляет 65535 байт, однако в большинстве компьютеров и сетей столь большие пакеты не используются. При передаче по сетям различного типа длина пакета выбирается с учетом максимальной длины пакета протокола нижнего уровня. Если это кадры Ethernet, то выбираются пакеты с максимальной длиной 1500 байт, умещающиеся в поле данных кадра Ethernet.

Поле ***Идентификатор пакета*** занимает два байта и используется для распознавания пакетов, образовавшихся путем фрагментации исходного пакета. Все фрагменты должны иметь одинаковое значение этого поля.

Поле ***Флаги*** занимает 3 бита и содержит признаки, связанные с фрагментацией. Установленный в 1 бит DF (Don`t Fragment) запрещает маршрутизатору фрагментировать данный пакет, а установленный в 1 бит MF (More Fragment) говорит о том, что данный пакет является промежуточным (не последним фрагментом). Оставшийся бит зарезервирован.

Поле ***Смещение фрагмента*** занимает 13 бит и задает ***смещение в байтах*** поля данных этого пакета ***от начала общего поля данных*** исходного пакета, подвергнутого фрагментации. Используется при сборке/разборке фрагментов пакетов при передачах их между сетями с различными величинами MTU (Maximum Transfer Unit). Смещение должно быть кратно 8 байт.

Поле ***Время жизни*** занимает один байт и означает предельный срок, в течение которого пакет может перемещаться по сети. Время задается источником передачи и исчисляется в секундах. На маршрутизаторах и в других узлах сети, в которые попадает пакет, по истечению каждой секунды из его текущего времени вычитается единица. Так как современные маршрутизаторы редко обрабатывают пакет дольше, чем за одну секунду, то время жизни можно считать равным максимальному числу узлов, которые разрешено пройти данному пакету до того, как он достигнет места назначения. Если параметр времени жизни станет нулевым до того, как пакет достигнет получателя, этот пакет будет уничтожен. Значение этого пакета изменяется при обработке заголовка IP-пакета.

Идентификатор ***Протокол верхнего уровня*** занимает один байт и указывает, какому протоколу верхнего уровня принадлежит информация, размещенная в поле данных пакета. Для сегмента TCP принято число 6, для дейтаграммы UDP – число 17.

Поле ***Контрольная сумма*** занимает два байта и рассчитывается только ***по заголовку***. Поскольку некоторые поля заголовка меняют свое значение в процессе передачи пакета по сети, (например, время жизни), контрольная сумма проверяется и повторно рассчитывается при каждой обработке IP-заголовка. Если контрольная сумма неверна, то пакет будет отброшен, как только ошибка обнаружится.

Поля ***IP-адрес источника*** и ***IP-адрес назначения*** имеют одинаковую длину (4 байта) и одинаковую структуру. В новой версии IPv6 используются 16-байтные адреса.

Поле ***Параметры*** является необязательным и используется обычно только при отладке сети. В этом поле можно указывать точный маршрут прохождения маршрутизаторов, регистрировать проходимые пакетом маршрутизаторы. При обычных коммуникациях поле не используется.

Поле ***Выравнивание*** заполняется нулями и используется для того, чтобы гарантировать завершение –заголовка на 32-битной границе.

Протокол IPv4 в настоящее время столкнулся с рядом проблем, таких как проблема масштабируемости сети, недостаточная приспособленность протокола к передаче мультисервисной информации с поддержкой различных классов обслуживания, включая обеспечение информационной безопасности.

Указанные проблемы обусловили развитие классической версии протокола IPv4 в направлении развития версии **IPv6,** принятой инженерной группой Интернет IETF в сентябре 1995 года.

В спецификации RFC1726 представлен набор функций, основными среди них являются:

- ***масштабируемость*** – идентификация и определение адресов как минимум 10^{12} конечных систем и 10^{9} индивидуальных сетей; такое огромное адресное пространство введено для облегчения ***иерархичности*** адресов, упрощающей работу маршрутизаторов, и практически никогда не будет исчерпано;
- топологическая гибкость – архитектура маршрутизации и протокол должны работать в сетях с различной топологией;
- автоматическое конфигурирование хостов и маршрутизаторов;
- безопасность на сетевом уровне.

В результате реализации заявленных функций важнейшие инновации IPv6 состоят в следующем:

- упрощен стандартный заголовок IP-пакета – маршрутизаторы больше не разбивают пакет на части (разбиение пакета возможно на передающей стороне), из заголовка ***исчезла контрольная сумма*** – т.к. канальные (Ethernet) и транспортные (TCP) протоколы тоже проверяют корректность пакета,
- расширено адресное пространство;
- на сверхскоростных сетях возможна поддержка огромных пакетов (джамбограмм) — до 4 гигабайт;
- улучшена поддержка иерархической адресации, агрегирования маршрутов и автоматического конфигурирования адресов;
- введены механизмы аутентификации и шифрования на уровне IP-пакетов;
- введены метки потоков данных и классы трафика.

Структура заголовка IP-пакета для IPv6 приведена на рисунке.

4 бит Версия	4 бита Приоритет	24 бит Метка потока	
16 бит Размер поля данных		8 бит Следующий заголовок	8 бит Предельное число шагов
Адрес отправителя 128 бит – 16 байт (4 строки)			
Адрес получателя 128 бит – 16 байт (4 строки)			

Структура заголовка IP-пакета для IPv6 (общая длина 40 байт)

В версии IPv6 ***размер IP-адреса увеличен до 16 байт***, что обеспечивает существование ***миллиарда*** сетей.

Введение в заголовке пакета поля ***«Метка потока»*** позволило значительно ***упростить*** процедуру ***маршрутизации*** однородного потока пакетов. Допускается существование нескольких потоков между отправителем и получателем. Метка потока присваивается узлом-отправителем путем генерации псевдослучайного 20-битного числа. Все пакеты одного потока должны иметь одинаковые заголовки, обрабатываемые маршрутизатором.

При получении первого пакета с меткой потока маршрутизатор анализирует дополнительные заголовки, выполняет предписанные этими заголовками функции и запоминает результаты обработки в локальном кэше. Ключом для такой записи является комбинация адреса источника и метки потока. Последующие пакеты с той же комбинацией адреса источника и метки потока обрабатываются с учетом информации кэша без детального анализа всех полей заголовка. Время жизни записи в кэше может быть определено узлом-отправителем.

Поле «***Приоритет***» (4 бита) формирует два подмножества;

- от 0 до 7 – трафик с контролем перегрузки (передача асинхронного трафика) – как в FDDI;
- от 8 до 15 – трафик без контроля перегрузки (приложения реального времени с постоянной скоростью).

В IPv6 опционная информация уровня Интернет записывается в отдельных заголовках, которые могут быть помещены между IPv6 заголовком и заголовком верхнего уровня пакета. Существует небольшое число таких заголовков, каждый задается определенным значением кода поля *следующий заголовок*.

5.5 Маршрутизация в IP-сетях

Маршрутизация пакетов в Интернете осуществляется программными модулями протокола IP, устанавливаемыми на всех ***конечных станциях и маршрутизаторах*** сети. Каждый программный модуль имеет собственную таблицу маршрутизации. Таблица маршрутизации служит для определения ***адреса*** (сетевого уровня) ***следующего маршрутизатора*** или непосредственно получателя по имеющемуся адресу (сетевого уровня) и получателя. После определения

адреса передачи выбирается определенный выходной физический порт маршрутизатора.

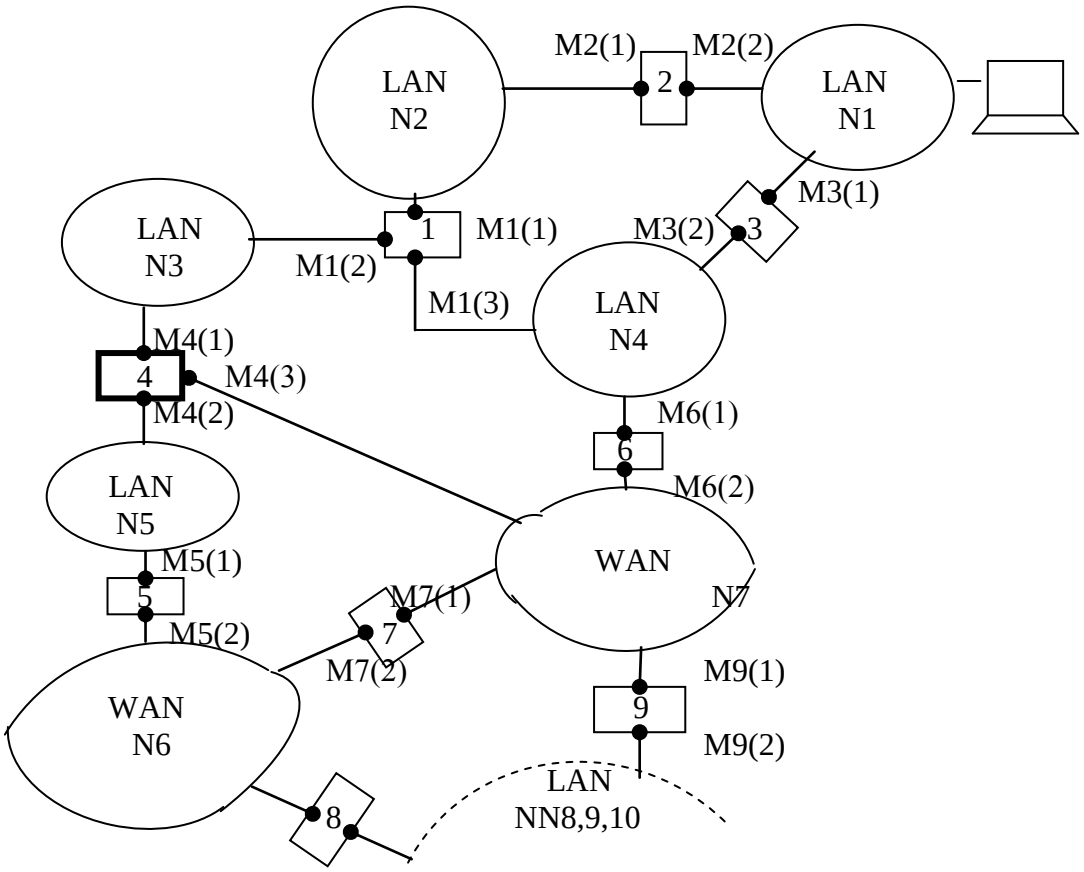
В таблице маршрутизации содержатся:

- адрес сети назначения;
- адрес следующего маршрутизатора;
- адрес порта, на который нужно отправить пакет;
- сведения об особенностях данного маршрута.

Таблицы маршрутизации должны обеспечивать продвижение пакета по любому правильно составленному сетевому адресу. Так как пакет может быть адресован в любую сеть, то можно предположить, что таблицы маршрутизации должны иметь записи обо всех сетях, входящих в составную сеть.

На практике, содержание таблицы маршрутизации минимизируется с учетом особенностей топологии сетевых связей конкретного маршрутизатора. В большинстве случаев достаточно записать номера сетей, непосредственно подсоединенных к данному маршрутизатору или расположенных поблизости на тупиковых маршрутах. Об остальных сетях можно сделать в таблице единственную запись, указывающую на маршрутизатор, через который проходит путь к таким сетям. Такой маршрутизатор называется маршрутизатором «по умолчанию», а вместо номера сети в таблицу вносится запись ”Default”.

Рассмотрим маршрутизацию пакетов для структуры составной сети, состоящей из 10 отдельных подсетей (см. рисунок).



Маршрутизация в составной цепи

Один из возможных вариантов таблицы маршрутизации для маршрутизатора 4 приведен ниже.

Таблица маршрутизации для маршрутизатора 4

№ сети назначения	Сетевой адрес следующего маршрутизатора	Сетевой адрес выходного порта	Расстояние до сети назначения
N1	M1(2)	M4(1)	2
N2	M1(2)	M4(1)	1
N3	-	M4(1)	0 (подсоединена)
N4	M1(2)/M6(2)	M4(1)/M4(3)	1
N5	-	M4(2)	0 (подсоединена)
№6	M5(1)/M7(1)	M4(2)/M4(3)	1
№7	-	M4(3)	0 (подсоединена)
Default (для сетей NN 8, 9, 10)	 M9(1) M5(1)	 M4(3) M4(2)	 1 2

Указанная таблица значительно упрощена по сравнению с реальными таблицами маршрутизации, например, отсутствуют столбцы с масками, вместо номера сети может быть указан полный сетевой адрес узла назначения и т.д.

Для составной сети, показанной на рисунке, адреса подсетей №№8, 9, 10 в таблице отсутствуют. Отсутствие адреса в таблице по умолчанию (Default) трактуется как передача пакета на порт M4(3) или M4(2).

Одному и тому же адресу сети назначения могут соответствовать несколько строк в таблице маршрутизации. В этом случае при выборе маршрута принимается во внимание столбец «Расстояние до сети назначения».

При этом под расстоянием понимается любая метрика, используемая в соответствии с заданным в сетевом пакете критерием, часто называемым классом сервиса. Расстояние может измеряться «хопами» - прыжками между различными сетями, временем прохождения пакета по линиям связи, характеристикой надежности и т.п.

Существует несколько источников, поставляющих записи в таблицу маршрутизации.

Во-первых, при ***инициализации*** программное обеспечение стека TCP/IP заносит в таблицу записи о ***непосредственно подключенных сетях*** и ***маршрутизаторах по умолчанию***, а также ***записи об особых адресах*** типа 127.0.0.0, (используется для локального тестирования стека TCP/IP), 224.0.0.0 (используется для обработки групповых адресов) и пр. – составляется «минимальная таблица маршрутизации».

Во-вторых, администратор вручную заносит ***статические*** записи о ***специфичных маршрутах*** или о ***маршруте по умолчанию***.

В-третьих, протоколы маршрутизации ***автоматически*** заносят в таблицу ***динамические*** записи об ***имеющихся маршрутах***.

Маршрутизация пакета очень упрощенно сводится к следующим действиям. Из кадра, поступившего на входной интерфейс маршрутизатора, ***извлекают IP-пакет*** и ***анализируют*** его ***IP-адрес назначения***. По этому адресу из таблицы маршрутизации определяют ***IP-адрес следующего маршрутизатора*** и транслируют его в локальный адрес. Затем исходный пакет упаковывают в кадр нижележащей технологии и отправляют на следующий маршрутизатор. Так происходит до тех пор, пока пакет не попадет в сеть назначения.

Для управления маршрутизацией используются также следующие протоколы уровня межсетевого взаимодействия:

- протоколы, связанные ***с составлением и модификацией таблиц маршрутизации*** RIP* (Routing Internet Protocol) и OSPF** (Open Shortest Path First);
- протокол ***межсетевых управляющих сообщений*** CMP (Internet Control Message Protocol, предназначенный для обмена информацией об ошибках между маршрутизаторами сети и узлом-источником пакета,
- ряд других протоколов.

*Протокол RIP распространяет между марщрутизаторами информацию о номерах сетей и расстояниях до них. В большинстве реализаций протокола используется простейшая метрика – ***количество хопов,*** т.е. число ***промежуточных маршрутизаторов***, которые нужно преодолеть пакету до сети назначения. После инициализации каждого маршрутизатора, он начинает посылать соседям сообщения протокола RIP, в которых содержится его минимальная таблица. После получения аналогичных сообщений от соседних маршрутизаторов, данные полей метрики увеличиваются на единицу и запоминается, через какой порт и от какого маршрутизатора получена новая информация. Если новая информация имеет лучшую метрику, чем уже имеющаяся, то эти данные вносятся в таблицу маршрутизации.

В протоколе OSPF таблица маршрутизации строится в два этапа. На первом этапе маршрутизатор строит граф связей сети, в котором вершинами являются маршрутизаторы и IP-сети, а ребрами – интерфейсы маршрутизаторов. Для этого все маршрутизаторы обмениваются информацией о ***топологии* сети, в результате чего все маршрутизаторы обладают идентичными сведениями о графе (топологии) сети. Второй этап состоит в нахождении оптимальных маршрутов на основе полученного графа. В каждом найденном таким образом маршруте запоминается только один шаг до следующего маршрутизатора, данные о котором вносятся в таблицу маршрутизации. OSPF-маршрутизаторы используют метрику, учитывающую пропускную способность составных сетей. Протокол обладает высокой вычислительной сложностью.

Задачу маршрутизации решают не только промежуточные узлы – маршрутизаторы, но и ***конечные*** узлы – компьютеры. Конечные узлы в еще большей степени, чем маршрутизаторы, пользуются приемом маршрутизации по умолчанию. Хотя они в общем случае имеют таблицу маршрутизации, ее объем обычно незначителен, что объясняется периферийным расположением всех конечных узлов. При наличии одного маршрутизатора в локальной сети конечный узел работает вообще без таблицы маршрутизации, имея только сведения об адресе маршрутизатора по умолчанию.

5.6 Способы адресации в IP-сетях

Любое устройство в сети IP может иметь несколько видов адресов.

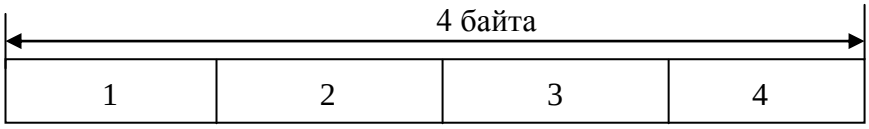
- ***физический*** адрес; это - шестнадцатеричный 6-байтный MAC-адрес сетевого адаптера или порта, указанный производителем оборудования. В недвоичном написании MAC-адреса представляют собой 12 одноразрядных 16-ричных цифр (0, 1, 2, 3, 4. 5, 6, 7, 8, 9, A, B, C, D, E). MAC-адреса широко использовались в локальных сетях для адресации пользователей. В последние годы, при структуризации локальных сетей и широком распространении TCP/IP-протокола даже в чисто локальных сетях происходит замена MAC-адресов ***внутренними*** IP-адресами. При этом MAC-адреса сохраняют свое значение как средство проверки идентичности оборудования пользователя, например, модема для сетевого провайдера.

- ***внутренний сетевой IP-адрес***; используемый для адресации узла в пределах ***локальной*** сети; этот адрес не зависит от физического адреса устройства и назначается ***администратором локальной сети*** во время её настройки. Адрес имеет десятичное представление; длина - 4 байта. В общем случае адрес состоит из двух частей: номера подсети и номера узла в подсети. Граница между номером подсети и номером узла в локальной сети определяется маской, устанавливаемой администратором локальной сети на сетевом сервере.

- ***внешний сетевой IP-адрес,*** используемый для однозначной идентификации узла в пределах всей ***составной*** сети. Адрес имеет десятичное представление; длина - 4 байта и состоит из двух частей: номера сети и номера узла в сети. Граница между номером сети и номером узла в локальной сети определяется ***классом*** сети или ***маской,*** устанавливаемой администратором сети на маршрутизаторе, один из узлов которого входит в данную сеть. Одно устройство может содержать несколько сетевых адресов, если оно принадлежит адресному пространству нескольких сетей.

- ***символьный*** адрес (DNS-имя) символьный идентификатор узла, например, www.snab.nateks.ru. Этот адрес состоит из нескольких частей: имени машины, имени организации, имени домена. Такой адрес используется, как правило, пользователями ***на прикладном уровне.***

IP-адреса являются основным типом адресов, используемых сетевым уровнем для передачи пакетов ***между*** сетями. Эти адреса ***состоят из четырех байт*** (см. рисунок).



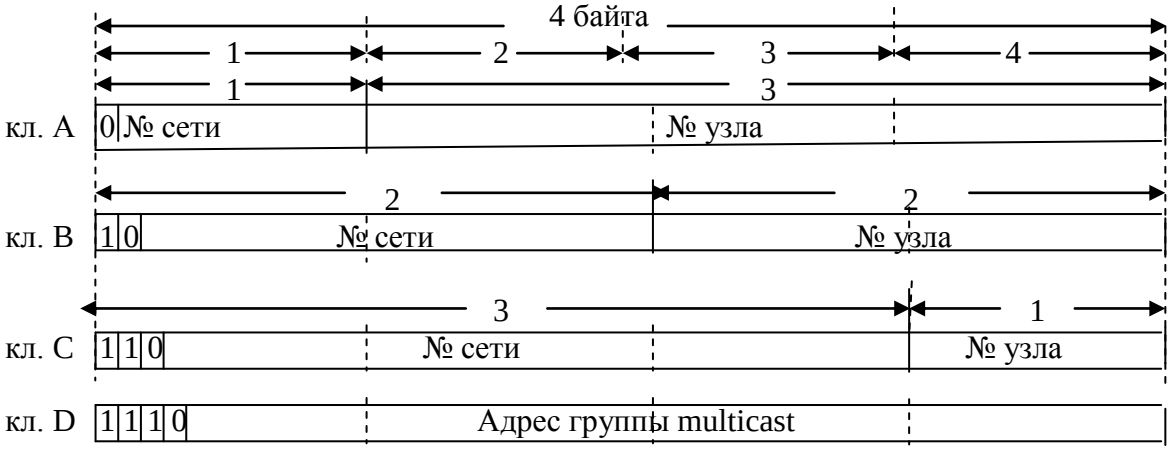
Запись IP-адреса

Способ, при помощи которого записываются все IP-адреса, называется пунктирной десятичной системой обозначений. Каждое 32-битовое адресное поле разделено на четыре поля в виде xxx.xxx.xxx.xxx, и каждому полю дается десятичное числовое значение от 0 до 255,

выраженное в виде одного октета (байта). Например, 109.26.17.100 – традиционная десятичная форма представления адреса, а 01101101 00011010 00010001 01100100 – двоичная форма представления этого же адреса.

IP-адрес состоит из двух частей: ***номера сети и номера узла***. Между номером сети и номером узла не предусматривается никакого разграничительного знака. Номер узла в протоколе IP назначается независимо от ***локального*** адреса узла. ***Маршрутизатор*** по определению входит сразу в ***несколько сетей***, поэтому каждый ***порт*** маршрутизатора имеет ***собственный*** IP-адрес. Конечный узел также может входить в несколько IP-сетей. В этом случае компьютер должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом, IP-адрес характеризует ***не отдельный компьютер***, а одно сетевое ***соединение***.

Традиционные схемы IP-адреса на номер сети и номер узла основаны на понятии класса, который определяется значениями нескольких первых битов адреса. Структуры IP- адресов разных классов показаны на рисунке.



Структура IP-адреса

В таблице приведены диапазоны номеров сетей и максимальное число узлов, соответствующих каждому классу сетей.

Если адрес ***начинается с 0***, то этот адрес относится к ***классу A***, в котором под номер сети отводится один байт (число возможных сетей 2^7-2=126), а остальные три байта интерпретируются как номер узла в сети. Существуют некоторые ограничения адресации: ни номер сети, ни номер узла не могут состоять только из одних двоичных единиц или только из одних двоичных нулей. Сетей класса A ***немного***, зато количество узлов в них может достигать 2^{24}, то есть ***более 16 млн. узлов***. Адреса класса A предназначены для очень крупных компаний с большим количеством

рабочих станций. Адреса класса A – IBM Corporation, Ford Motor Company, Hewlett Packard Company и другие.

Класс	Первые биты	Наименьший номер сети	Наибольший номер сети	Число сетей	Максимальное число узлов в сети
A	0	1.0.0.0	126.0.0.0	126	$2^{24} \approx 16\ 10^6$
B	10	128.0.0.0	191.255.0.0	$2^{14} \approx 16\ 10^3$	$2^{16} \approx 65{,}5\ 10^3$
C	110	192.0.0.0	223.255.255.0	$2^{21} \approx 2\ 10^6$	$2^8 = 256$
D	1110	224.0.0.0	239.255.255.255	-	Multicast
E	11110	240.0.0.0	247.255.255.255	-	Зарезервирован

Если первые два бита адреса ***равны 10***, то адрес относится к ***классу B***. В адресах класса B под номер сети и под номер узла отводится по ***два байта***. Сети, имеющие номера в диапазоне от 128.0 (**10**000000 00000000) до 191.255 (**10**111111 11111111), называется сетями класса B. Таким образом, сетей класса B ***больше***, чем сетей класса A, но размеры их меньше, максимальное количество узлов в них составляет 2^{16} (***65536***).

Если адрес начинается с последовательности битов ***110,*** то это адрес ***класса C***. В этом случае под номер сети отводится 24 бита, а под номер узла – 8 бит. Сети класса C ***наиболее распространены***, но число узлов в них ограничено значением 2^8 (***256***) узлов.

Если адрес начинается с последовательности ***1110***, то он является адресом ***класса D*** и обозначает особый, ***групповой адрес*** (multicast). Групповой адрес идентифицирует группу узлов (сетевых интерфейсов), которые в общем случае могут принадлежать ***разным*** сетям. Интерфейс, входящий в группу, получает наряду с обычным индивидуальным адресом ***еще один групповой*** адрес. Если при отправке пакета в качестве адреса назначения указан адрес класса D, то такой пакет должен быть доставлен ***всем узлам***, которые входят в ***группу***.

Если в поле ***узла назначения*** стоят только единицы, то пакет, имеющий такой адрес, рассылается всем узлам сети с заданным номером. Например, пакет с адресом 172.168.10.255 доставляется всем узлам сети 172.168.10.0. Такой адрес является широковещательным

Часто администраторы сетей испытывают неудобства из-за того, что количество централизованно выделенных им номеров ***сетей недостаточно*** для того, чтобы ***структурировать сеть*** надлежащим образом, например, разместить все слабо взаимодействующие компьютеры в ***разных сетях***. В такой ситуации возможны два пути. Первый из них связан с получением от поставщика услуг Интернет ***дополнительных*** номеров ***сетей***. Второй способ, употребляемый чаще, связан с использованием технологии ***масок***, которая позволяет разделять одну сеть на несколько сетей. Адресная маска – это 4-х байтная последовательность, в которой биты, предназначенные для адреса подсети, содержат единицы, а биты, предназначенные для адреса узла, содержат нули.

Снабжая каждый IP-адрес маской, можно ***отказаться*** от понятий ***классов*** адресов и сделать систему адресации ***более гибкой***. В масках количество единиц в последовательности, определяющей ***границу номера сети***, не обязательно должно быть ***кратной 8***. Маска – это число, которое используется ***в паре с IP-адресом***.

Для стандартных классов сетей маски имеют следующие значения:

Класс A – 11111111 00000000 00000000 00000000 (255.0.0.0);
Класс B – 11111111 11111111 00000000 00000000 (255.255.0.0);
Класс C – 11111111 11111111 11111111 00000000 (255.255.255.0).

Допустим, администратор получил в свое распоряжение адрес сети класса B: 129.44.0.0. Такой адрес позволяет организовать сеть с узлами, имеющими номера узлов в диапазоне 0.0.0.1- 0.0.255.254 (с учетом того, что адреса из одних нулей и одних единиц имеют специальное назначение и не используются для адресации узлов, всего ~65x10^3 узлов). Требуется разделить сеть на три отдельных подсети и обеспечить надежную локализацию трафика каждой подсети. Это позволит улучшить диагностику сети и проводить в каждой из подсетей особую политику безопасности.

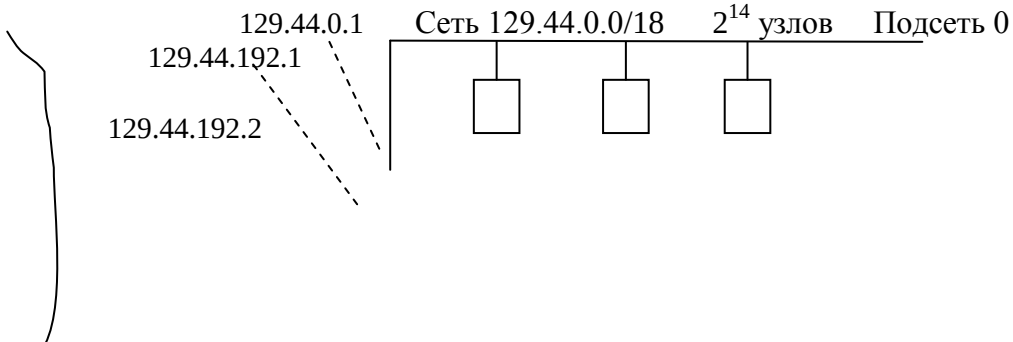
Решим поставленную задачу при помощи масок одинаковой длины. В качестве маски выберем значение 255.255.192.0 (11111111 11111111 11000000 00000000). При этом число разрядов, интерпретируемых как номер сети, увеличилось с 16 до 18, что позволяет из одного централизованно заданного номера сети сделать четыре.

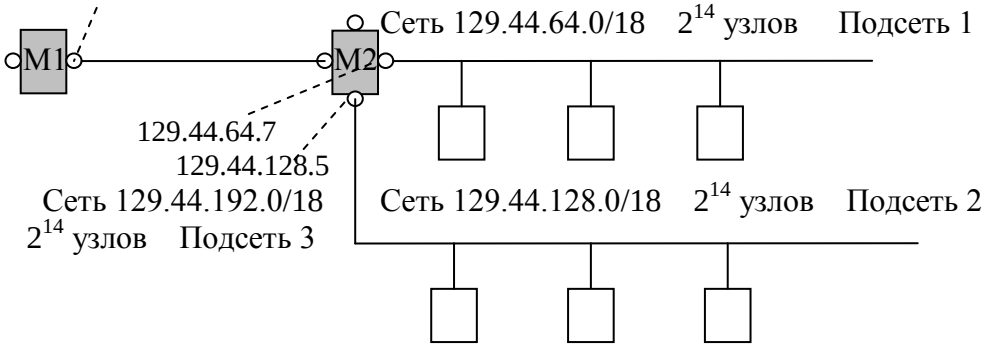
(00) 129.44.0.0 (10000001 00101100 00000000 00000000)
(01) 129.44.64.0 (10000001 00101100 01000000 00000000)
(10) 129.44.128.0 (10000001 00101100 10000000 00000000)
(11) 129.44.192.0 (10000001 00101100 11000000 00000000)

Два дополнительных бита в номере сети можно интерпретировать как номера подсетей 0(00); 1(01); 2(10) и 3(11) соответственно. Измененная структура IP-адреса показана на рисунке.

Структура IP-адреса при организации подсетей

Структурированная сеть показана на рисунке. Весь трафик во внутреннюю сеть 129.44.0.0., направляемый из внешней сети, поступает через маршрутизатор M1. Для структуризации информационных потоков во внутренней сети установлен дополнительный маршрутизатор M2.





Структурированная сеть с использованием маски 255.255.192.0

Извне сеть по-прежнему выглядит, как единая сеть класса B, а на местном уровне это составная сеть, содержащая три отдельных сети по 2^{14} узлов в каждой. Все маршрутизаторы внешней сети, встретив пакеты с адресами, начинающимися с 129.44, интерпретируют их как адреса класса B и направляют их к маршрутам, ведущим к маршрутизатору M1.

Маршрутизатор M1, в свою очередь, направляет весь входной трафик сети 129.44.0.0 на порт 129.44.192.1 маршрутизатора M2. Маршрутизатор M2 обрабатывает все поступившие на него пакеты, используя маску 255.255.192.0 для выделения адреса подсети, и направляет пакет с адресом подсети «0» 129.44.0.0 на порт 129.44.0.1, пакет с адресом подсети «1» 129.44.64.0 на порт 129.44.64.7 и пакет с адресом подсети «2» 129.44.128.0 на порт 129.44.128.5.

Недостатком проведенного распределения является слабое использование адресного пространства в подсети 3, в которой задействованы всего 2 узла из 2^{14}.

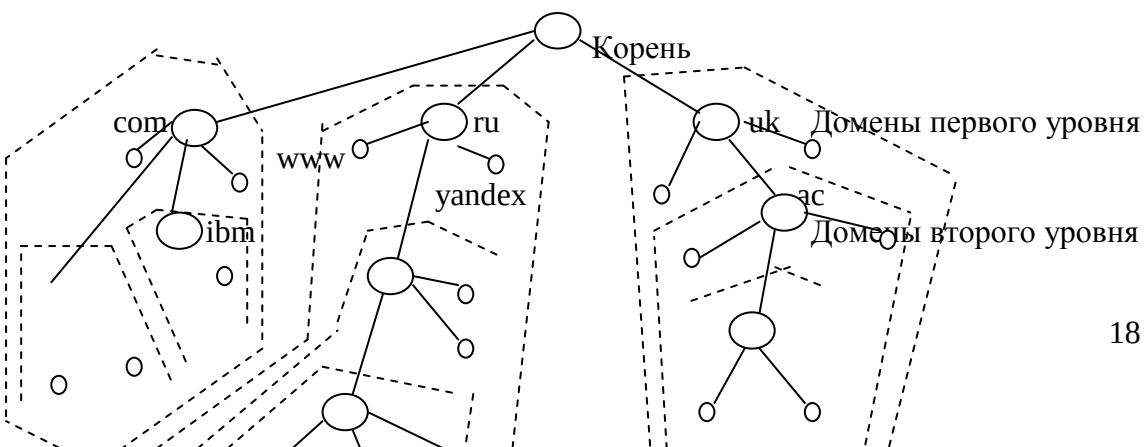
Эффективность использования адресного пространства можно повысить, разделяя сеть на подсети ***разного*** размера. Для этого используются маски ***переменной*** длины.

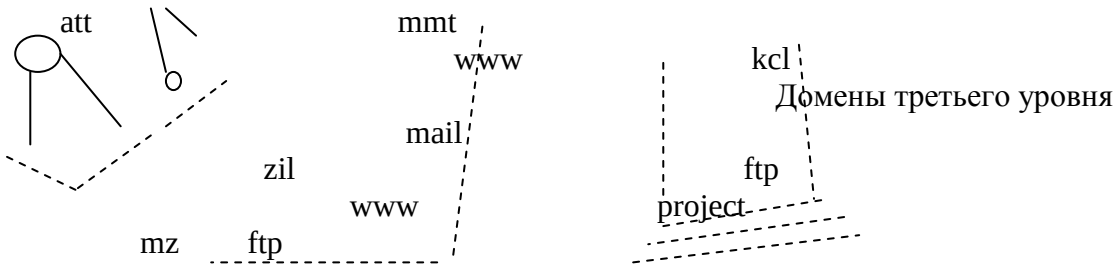
Маски позволяют также ***агрегатировать*** несколько сетей в одну более крупную, отдавая часть адреса сети для адресации узлов.

В современных маршрутизаторах поле маски указывается ***для каждой строки*** в таблице маршрутизации.

5.7 Система доменных имен

Символьные имена в IP-адресах называются ***доменными*** и строятся по ***иерархическому*** признаку. Доменная система имен имеет иерархическую древовидную структуру, допускающую использование в имени ***произвольного*** количества составных частей.





Пространство доменных имен

Пользователи обычно предпочитают работать с ***символьными*** именами компьютеров. Составляющие полного символьного имени в IP-сетях разделяются точкой и перечисляются в следующем порядке: ***сначала простое имя*** хоста, затем имя группы узлов (например, имя организации), затем имя более крупной группы (поддомена) и так до имени домена самого высокого уровня (например, домена, объединяющего организации по географическому признаку). Доменная структура доменных имен показана на рисунке. Примером доменного имени может служить имя base2.sales.zil.ru.

Символьные адреса существуют только как удобное обозначение определенного конечного пользователя. При передаче по сети символьный адрес заменяется IP-адресом.

Между доменным ***именем*** и IP-***адресом*** узла нет никакой функциональной зависимости, поэтому единственный способ установления ***соответствия*** – это ***таблица***. Распределенная база отображения «доменное имя – IP-адрес» находится в ведении ***специальной службы – системы доменных имен*** (Domain Name System, DNS), которая устанавливает соответствие доменного имени и IP-адреса на основании создаваемых администраторами сети таблиц соответствия. Для каждого домена создается собственный DNS-сервер, к которому обращаются пользователи для получения IP-адреса.

Разделение имени на части позволяет разделить административную ответственность за назначение уникальных имен между различными людьми или организациями. Так, для примера, приведенного на рисунке, один человек может нести ответственность за то, чтобы все имена, которые имеют окончание “ru” имели уникальную следующую вниз по иерархии часть. Адреса основных узлов российского сегмента Интернета (Рунета) находятся в ведении РосНИИРОС (Российский НИИ развития общественных сетей) – RIPN (Russian Institute for Public Networks); адрес в Интернете http://www.ripn.net.

Корневой домен управляется центральными органами Интернета – Internet Assigned Numbers Authority (IANA) – службой присвоения номеров Интернета. ***Домены верхнего уровня назначаются для каждой страны, а также на организационной основе***. Для обозначения стран используются двухбуквенные и трехбуквенные аббревиатуры, например, ru (Россия), uk (Великобритания), us (Соединенные Штаты), de

(Германия) а для различных типов организаций – следующие обозначения:

- ***com*** – коммерческие организации (например, microsoft.com);
- ***edu*** – образовательные организации (например, mit.edu);
- ***gov*** – правительственные организации (например, nsf.gov);
- ***org*** – некоммерческие организации (например, fidonet.org);
- ***net*** – организации поддержки сетей (например, nsf.net).

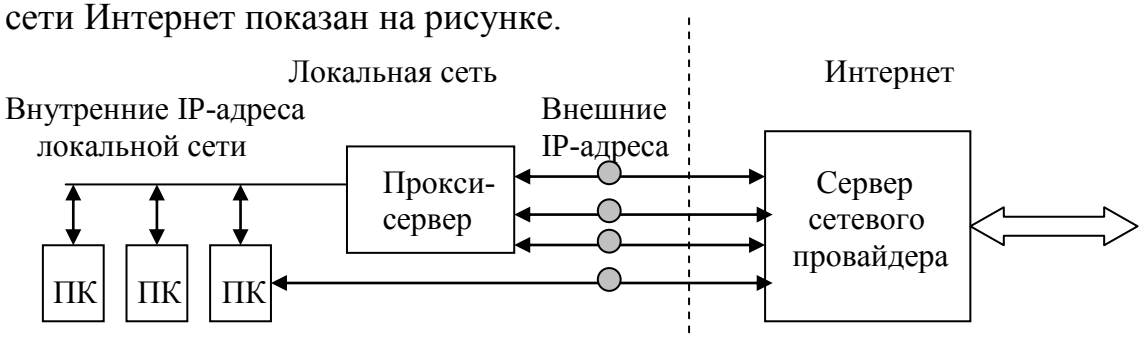
Интернет растет с огромной скоростью, и хорошие, осмысленные имена становятся дефицитным товаром. Некоторые небольшие государства не испытывают большой нужды в Интернет-адресах. Пользуясь монополией на адрес своего домена, они продают иностранным компаниям право пользоваться своим доменным именем.

Маленькое островное государство Тувалу (девять атоллов в южной части Тихого океана) продало право на регистрацию Интернет-адресов, оканчивающихся на символы tv, одной американской компании за объем ежегодной оплаты, составляющий 1/3 годового бюджета. Покупатель рассчитывает, что сочетание TV привлечет телевизионные компании при выборе ими доменного имени. Есть сведения о торговле доменными именами государствами СНГ – Молдовой (md – medicine doctor), Туркменистаном (tm – trademark).

5.8 Особенности адресации пользователей локальных сетей с выходом в Интернет

Обычно число пользователей локальной сети превышает количество IP-адресов, имеющихся в распоряжении сетевого администратора. Большая часть пользователей имеют только ***внутренние*** сетевые адреса; поэтому такие пользователи получают доступ к сети Интернет через прокси-сервер (прокси-сервер - от англ. proxy — «представитель, уполномоченный»).

Пример организации доступа пользователей локальных сетей к сети Интернет показан на рисунке.



Организация доступа пользователей локальных сетей в Интернет

Прокси-сервер обеспечивает разнообразные функции:

- обеспечивает ***доступ*** компьютеров локальной сети в Интернет;
- производит ***кэширование*** данных: если часто происходят обращения к одним и тем же внешним ресурсам, то можно держать их копию на прокси-сервере и выдавать по запросу, снижая тем самым нагрузку на канал во внешнюю сеть и ускоряя получение клиентом запрошенной информации; кэш или кеш (англ. cache, от фр. cacher —

прятать; произносится [kæʃ] — кэш) — промежуточный буфер с быстрым доступом, содержащий информацию, которая с наибольшей вероятностью может быть запрошена быстродействующей памятью, например оперативной;

- ***защищает*** локальную сеть от внешнего доступа: например, можно настроить прокси-сервер так, что локальные компьютеры будут обращаться к внешним ресурсам только через него, а внешние компьютеры не смогут обращаться к локальным вообще (они «видят» только прокси-сервер);
- ***ограничивает*** доступ из локальной сети к внешней: например, ***запрещает*** доступ к определённым веб-сайтам или ***запрещает*** использование Интернета для некоторых пользователей, устанавливает ***квоты*** на трафик отдельных пользователей, фильтрует рекламу и вирусы.

Прокси-сервер имеет два IP-адреса ***внутренний*** и ***внешний***. При отправлении запроса или сообщения во внешнюю сеть от абонентов локальных сетей, имеющих только внутренний сетевой адрес, эти сообщения идут по Интернету, имея ***общий адрес*** отправителя – внешний адрес сетевого сервера. Прокси-сервер запоминает внутренний сетевой адрес источника сообщения, и при получении ответа из сети последний доставляется по внутреннему адресу пользователя.

Прокси-сервер может иметь несколько внешних IP-адресов, это позволяет ***динамически*** распределять адреса, когда количество узлов в локальной сети превышает количество IP-адресов, имеющихся в распоряжении сетевого администратора.

Отдельные пользователи локальных сетей могут иметь, кроме внутреннего, еще и собственный ***внешний IP-адрес***. Такие пользователи могут выходить в Интернет непосредственно по своему внешнему адресу или через прокси-сервер, как обычные пользователи.

Внешний сетевой адрес сервера может быть одним из внутренних адресов сетевого провайдера.

5.9 Транспортный уровень

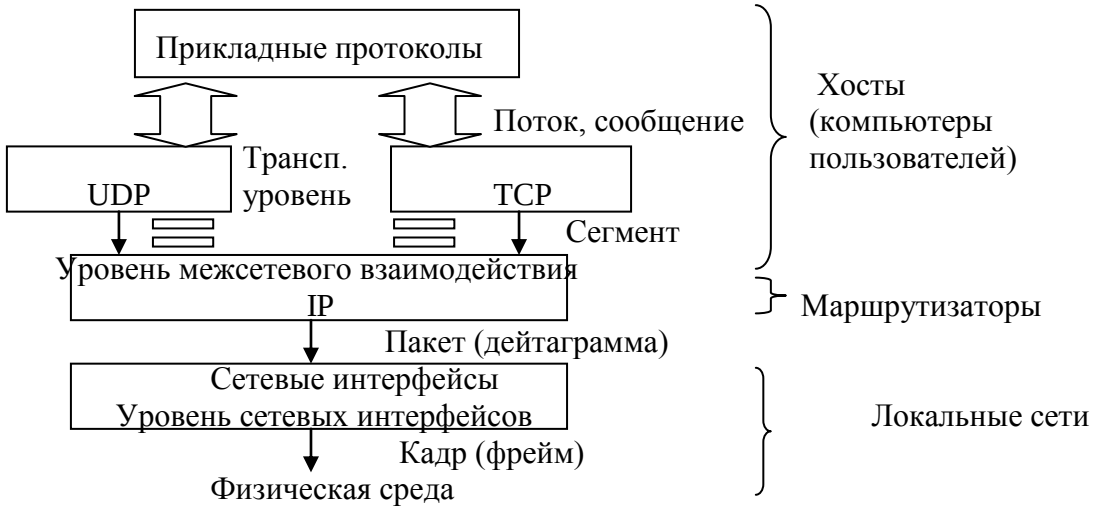
Транспортный (transport) уровень стека TCP/IP может предоставлять вышележащему уровню два типа сервиса;

- ***гарантированная доставка*** – протокол управления передачей (Transmission Control Protocol, TCP);
- ***доставка «по возможности»*** (“best effort”) – протокол пользовательских дейтаграмм (User Datagram Protocol, UDP).

Протоколы TCP, как и пользовательские приложения, устанавливаются в оборудовании хостов – компьютерах пользователей.

Чтобы лучше понять функции транспортного протокола, напомним, что информация, поступающая к протоколу TCP от

протоколов более высокого уровня, рассматривается протоколом TCP как ***неструктурированный поток*** байтов. Поступающие данные буферизуются средствами TCP. Для передачи на сетевой уровень из буфера «вырезается» некоторая непрерывная часть данных, которая называется ***сегментом*** (см. рисунок).



Прохождение информационных объектов в стеке протоколов TCP/IP

Сегменты могут быть разного размера, однако об их максимальных размерах участники соединения должны договориться.

Сегменты передаются для передачи нижележащему уровню ***межсетевого взаимодействия***. После того как эти сегменты будут доставлены средствами уровня межсетевого взаимодействия в ***пункт назначения***, протокол TCP снова ***соберет*** их в непрерывный поток байтов и переправляет данные конкретному процессу-получателю.

Каждый компьютер может выполнять несколько процессов.

Пакеты, поступающие на транспортный уровень, организуются операционной системой компьютера в виде множества очередей к точкам входа различных приложений. В терминологии TCP/IP такие точки называются ***портами***. Порт однозначно определяет приложение в пределах компьютера. (Порты приложений не следует путать с портами (интерфейсами) оборудования).

Приложения, которые получают данные на уровень IP, используя протокол TCP, получают номера, называемые ***портами TCP***. Аналогично, приложениям, обращающимся к протоколу UDP, выделяются ***порты UDP.***

Номер порта в совокупности с номером сети и номером конечного узла однозначно определяют прикладной процесс в сети. Этот ***набор идентифицирующих параметров (IP-адрес, номер порта) называется сокет (socket)***.

Каждый взаимодействующий процесс идентифицируется сокетом-парой (***IP-адрес интерфейса, номер порта***), каждое соединение – парой сокетов взаимодействующих процессов. Каждый процесс может участвовать в нескольких соединениях.

Если процессы представляют собой популярные общедоступные службы, такие как FTP, telnet, HTTP, TFTP, DNS и т. п., то за ними централизованно закрепляются ***стандартные присвоенные*** (assigned) номера.

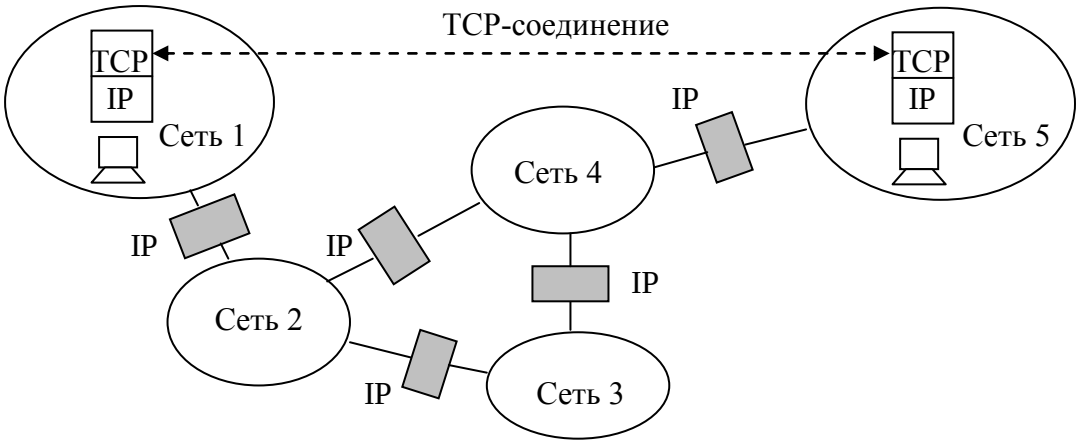
Для тех служб, которые еще не стали столь распространенными, чтобы закреплять за ними стандартные номера, номера портов выделяются локальной операционной системой компьютера. Такие номера называют ***динамическими***.

Задача ***транспортного*** уровня, которую решают протоколы TCP и UDP, заключается в обеспечении ***надежной*** передачи между конечными пользователями – передаче данных между любой парой ***прикладных процессов***, выполняющихся в сети. Именно транспортный уровень наводит порядок в хаотичной, не имеющей централизованного управления сети Интернет.

Надежность передачи данных протоколом TCP достигается за счет того, что он основан на установлении ***логических соединений*** (не путать с виртуальными каналами в сетях X.25, ATM и др.!) между взаимодействующими процессами.

До тех пор пока модули протокола TCP на обоих оконечных хостах продолжают функционировать корректно, а составная сеть не распалась на несвязные части, ошибки в передаче данных на уровне протокола IP не будут влиять на правильность обмена хостов сегментами.

Установление логического соединения (см. рисунок) позволяет объектам на компьютере-отправителе и компьютере-получателе поддерживать ***обмен данными в дуплексном режиме***. Это обеспечивает возможность протоколу TCP ***нумеровать*** пакеты***, подтверждать их прием*** квитанциями***,*** в случае потери ***организовывать повторные передачи***, распознавать и уничтожать дубликаты, доставлять прикладному уровню пакеты в том порядке, в каком они отправлены.



TCP-соединение между конечными узлами

В рамках TCP-соединения происходит ***договорный процесс*** о следующих ***параметрах*** процедуры обмена данными между двумя процессами: ***максимальном размере сегмента***, максимальном объеме данных, которые можно передавать без получения подтверждения (***окно приема***), о ***начальном*** порядковом ***номере байта,*** с которого начинается отсчет потока данных в рамках данного соединения.

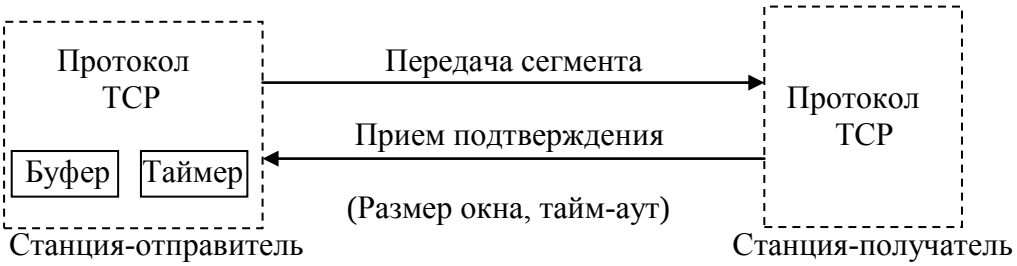
Поскольку соединения устанавливаются через ненадежную коммуникационную систему, основанную на протоколе IP, то во

избежание ошибочной инициализации соединений используется специальная многошаговая процедура подтверждения связи. По существу, ***реализуется механизм РОС*** на основе метода «скользящего окна» и с оптимизацией параметров, управляющих обменом (***время ожидания и размер окна***).

Протокол TCP является ***дуплексным***, т.е. в рамках одного соединения регламентируется процедура обмена данными в обе стороны. Каждая сторона одновременно выступает и как отправитель, и как получатель. При установлении соединения, а затем и в ходе передачи обе стороны, выступая в роли получателя, посылают друг другу ***размер «окон приема»,*** в основном, исходя из того, с какой ***скоростью*** приемная станция сможет обрабатывать присылаемые данные.

Управлять окном приема может и сторона-отправитель. Если отправляющая сторона фиксирует ненадежную работу линии связи, то она может по собственной инициативе уменьшить окно. В таких случаях действует правило: в качестве действующего размера окна выбирается ***минимальное*** из двух значений – значения, диктуемого приемной стороной, и значения, определяемого «на месте» отправителем.

Когда протокол TCP передает в сеть сегмент, он «на всякий случай» ***помещает его копию*** в очередь повторной передачи (см. рисунок) и запускает таймер. Когда ***приходит подтверждение*** на это сегмент, соответствующая копия ***удаляется*** из очереди. Если же подтверждение не приходит до истечения срока, то сегмент передается ***повторно***. Может случиться так, что повторный сегмент придет тогда, когда исходный сегмент уже окажется на месте, тогда дубликат будет попросту отброшен.



Реализация РОС протоколом TCP для одного направления передачи

Выбор ***времени ожидания*** (тайм-аута) очередной квитанции является важной задачей, результат которой влияет на производительность протокола TCP. Тайм-аут не должен быть ***слишком коротким***, чтобы по возможности исключить избыточные повторные передачи, которые снижают полезную пропускную способность системы. Но он не должен быть и ***слишком длинным***, чтобы избежать длительных простоев, связанных с ожиданием несуществующей или «заблудившейся» квитанции.

При выборе величины тайм-аута должны учитываться скорость и надежность физических линий связи, их протяженность и многие другие факторы. В протоколе

TCP тайм-аут определяется с помощью достаточно ***сложного адаптивного алгоритма,*** идея которого состоит в следующем. При ***каждой*** передаче засекается время от момента отправки сегмента до прихода квитанции о его приеме (***время оборота***), Получаемые значения времени оборота усредняются с весовыми коэффициентами, возрастающими от предыдущего замера к последующему. Это делается с тем, чтобы усилить влияние последних замеров. В качестве тайм-аута выбирается среднее время оборота, умноженное на некоторый ***коэффициент***. Практика показывает, что значение этого коэффициента должно превышать 2.

Особенность использования скользящего окна в протоколе TCP состоит в том, что хотя единицей передаваемых данных является сегмент, окно определено на ***множестве нумерованных байтов неструктурированного потока данных***, поступающих с верхнего уровня и буферизируемых протоколом TCP. При установлении соединения обе стороны договариваются о начальном номере байта, с которого будет вести отсчет в течение ***всего*** данного ***соединения***. У каждой стороны свой начальный номер. Сегменты могут иметь разную величину, но ***идентификатором*** каждого ***сегмента*** является номер его ***первого байта***. Нумерация байта в пределах сегмента осуществляется так, чтобы первый байт данных сразу вслед за заголовком имел наименьший номер, а следующие за ним байты имели возрастающие номера (см. рисунок).

Инкапсуляция сегмента TCP в пакете IP

Когда отправитель посылает сегмент, он помещает в одно из полей заголовка TCP в качестве ***идентификатора*** сегмента ***номер первого байта*** данного сегмента, который называется ***последовательным номером***. На основании этого номера модуль TCP-получатель не только отличает данный сегмент от других, но и ***позиционирует*** полученный фрагмент относительно общего потока байтов, он может также сделать вывод, не является ли этот сегмент дубликатом, и не было ли пропусков между полученным сегментом и предыдущим и т.п.

Второй протокол этого уровня – UDP является ***простейшим дейтаграммным протоколом***, который используется в том случае, когда задача ***надежного*** обмена данными либо вообще ***не ставится***, либо решается средствами более высокого уровня – системными прикладными службами или пользовательскими приложениями.

Некоторые приложения требуют максимально сократить время распространения сообщения между пользователями. Пример такого приложения рассматривается ниже.

Биржевых трейдеров не устраивает скорость передачи данных, придется класть новый трансатлантический кабель

10 октября 2011 г.

После внедрения на фондовых биржах автоматизированных систем торговли, когда сделки совершают компьютеры, время для принятия решения уменьшилось до микросекунд, что потребовало увеличения скорости связи,

Американская компания Hibernia Atlantic собирается вложить 300 млн долларов в прокладку нового кабеля между Лондоном и Нью-Йорком, благодаря которому время транзакции сократится на 6 миллисекунд. Гарантии финансирования на сумму 250 млн долларов для проекта получены от партнера и подрядчика Hibernia - китайской Huawei Marine Networks.

Сократить время надеются, в том числе, и за счет увеличения скорости света, только не вакууме, а в кабеле.

Вместо цельного оптоволокна кабель нового поколения сделают полым, поскольку скорость света в воздухе больше, чем в стекле. Кроме этого рассматривается замена алгоритмов шифрования и коррекции ошибок на более быстрые, вплоть до отказа от них.

Благодаря этому кабелю, клиенты Глобальной финансовой сети (GFN) Hibernia - трейдеры, банки и биржи - получат возможность получать информацию с задержкой менее 60 миллисекунд. По окончании строительства новый кабель станет самым быстрым трансатлантическим соединением.

На Нью-Йоркской фондовой бирже совершается до 22 миллиардов сделок в день – треть мирового объема биржевой торговли, но управляющая компания NYSE Euronext считает, что это не предел.

В последний раз оптоволокно по дну Атлантического океана прокладывали в конце 1990-х годов. Нынешний лидер на рынке каналов для высокочастотного трейдинга, компания Global Crossing, обеспечивает пинг в 65 мс по трансатлантическому каналу AC-1.

Новый оптоволоконный кабель Hibernian Express компании Hibernia Atlantic, который свяжет Лондон и Нью-Йорк, будет иметь длину 6021 км. Маршрут для прокладки вычисляли 18 месяцев с учетом рельефа морского дна и экономии каждого километра. Новый канал планируют подключить в 2013 году.

В Hibernian Express пинг уменьшат до 59 мс. При этом, по предварительной оценке, биржевики готовы заплатить за аренду полосы Hibernian Express в 50 раз больше, чем за AC-1.

При такой торговле прибыль часто получают из-за разницы в цене на различных площадках, которая возникает на очень короткое время. Задержка в несколько миллисекунд может стоить трейдеру сделки.

Компьютерные программы успевают выполнить расчет стратегии за несколько микросекунд, однако за одну микросекунду свет в вакууме пройдет лишь 300 метров. Одно моргание человеческого глаза длится 300-400 миллисекунд. Шесть миллисекунд, выигрываемые за счет прокладки нового кабеля, - это время, за которое обычная муха делает два взмаха крылом.

По данным журнала Popular Science, недавно проложенное соединение между нью-йоркской и чикагской биржами повысило скорость передачи данных примерно на три миллисекунды. Стоимость проекта составила порядка 300 млн. долларов.

5.10 Уровень сетевых интерфейсов

Уровень межсетевого взаимодействия передает пакеты на узел маршрутизатора ***подсети, в которой находится узел*** назначения. Перемещение пакета между соседними маршрутизаторами в ***пределах*** каждой из ***подсетей*** – это уже задача локальной (то есть использующейся в каждой из подсетей) технологии. Каждый раз, когда требуется воспользоваться ***локальными*** средствами доставки пакета в пределах подсети, протокол IP обращается к нижележащему ***уровню сетевых интерфейсов***.

Идеологическим ***отличием*** архитектуры стека TCP/IP от многоуровневой организации других стеков является интерпретация функций ***нижнего*** уровня – уровня сетевых интерфейсов (network interface).

Ранее были рассмотрены функции нижних уровней (канального и физического) таких технологий, как Ethernet, Token Ring и других, обеспечивающих ***доступ к среде*** передачи, ***формирование кадров, согласование*** электрических сигналов (кодирование, синхронизация) и прочие весьма конкретные действия.

Нижний уровень стека TCP/IP отвечает только за ***организацию интерфейса*** с частными технологиями подсетей.

На каждом сетевом узле в результате работы протоколов межсетевого уровня определяется ***адрес следующего*** по маршруту маршрутизатора. Чтобы добраться до него, надо ***пересечь*** некоторую подсеть, для этого протоколы TCP/IP должны обратиться к ***транспортным средствам данной промежуточной подсети***.

Упрощенно задача обеспечения интерфейса между двумя технологиями сводится,

- во-первых, к определению ***способа упаковки (инкапсуляции) пакета IP*** в единицу передаваемых данных промежуточной сети,

- во-вторых, к определению способа ***преобразования*** сетевого адреса следующего шлюза в ***новый тип адреса***, который принят для адресации узлов в технологии ***данной промежуточной сети***.

Уровень сетевых интерфейсов в стеке TCP/IP ***поддерживает все популярные технологии физического и канального уровней***; для локальных сетей это Ethernet, Token Ring, FDDI, Fast Ethernet, Gigabit Ethernet, для глобальных сетей – протоколы соединений коммуникационных систем произвольной физической природы (аналоговые и цифровые линии ТФОП, глобальных сетей с коммутацией пакетов X.25, frame relay, ATM, и др.).

Такой способ делает составную сеть TCP/IP ***открытой*** для включения в себя ***любой*** сети, какую бы внутреннюю технологию передачи данных эта сеть не использовала бы. Для каждой технологии, включаемой в составную сеть подсети, должны быть разработаны собственные интерфейсные средства. Поэтому этот уровень не может быть определен раз и навсегда.

5.11 Место технологии Интернет в глобальных сетях

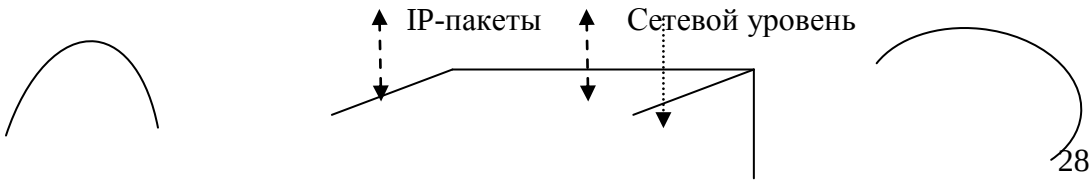
Революционный рост популярности Интернета привел к тому, что сегодня практически каждая глобальная сеть должна быть способна передавать трафик протокола IP. А это означает, что сегодня все глобальные сети являются ***составными сетями IP***, а отличия между ними заключаются в технологиях, лежащих ***под уровнем IP***.

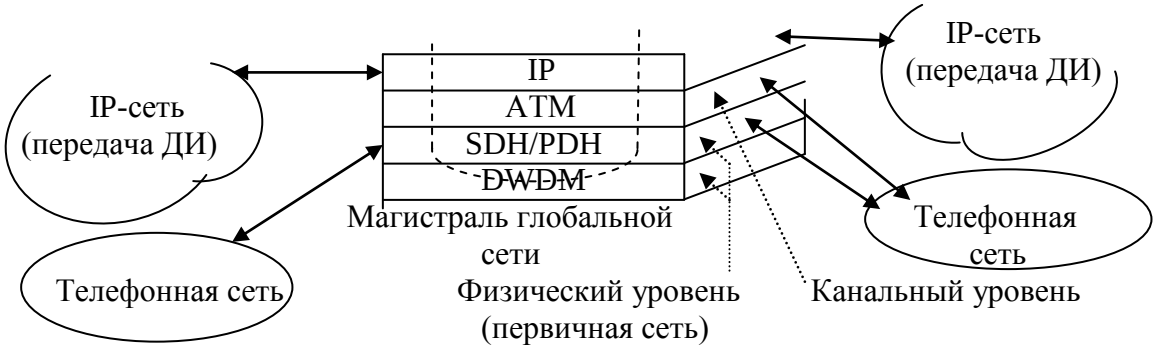
Сети IP изначально были задуманы как ***экономичные дейтаграммные*** сети, предоставляющие своим пользователям только услуги типа best effort. Поэтому классическая составная сеть IP была не в состоянии ***гарантировать*** своим пользователям ***качество обслуживания*** их трафика, если только сеть не работает с ***очень низким уровнем загрузки***. Дейтаграммная природа протокола IP затрудняет создание ***законченной системы*** поддержки QoS.

Неопределенность путей следования пакетов при ***существовании альтернативных*** маршрутов равной стоимости не позволяет выполнить ***точную оценку*** загруженности ***каждого ресурса*** сети, а значит, не позволяет применить в полной мере в сетях IP методы регулировки трафика.

Когда трафик IP стал непременным атрибутом любой сети передачи данных (и некоторых телефонных сетей тоже), то для обеспечения требуемого качества услуг большинство крупных глобальных сетей, особенно сетей операторов связи, на первых этапах строили по четырехуровневой схеме (см. рисунок).

Два нижних уровня не относятся к собственно пакетным сетям – это уровни ***первичной сети***, с помощью которой оператор сети может достаточно быстро организовать ***постоянный цифровой канал*** между ***точками подключения*** оборудования вышележащей наложенной сети – пакетной или телефонной.





Четырехступенчатая структура глобальной сети

На рисунке первичная сеть, реализующая ***физический*** уровень, представлена двумя уровнями – DWDM и SDH. Технология SDH делит пропускную способность спектральных каналов на более мелкие TDM-подканалы, связывающие интерфейсы коммутаторов пакетной сети (или телефонных коммутаторов).

В отдельных случаях уровень DWDM отсутствует, технология SDH тоже может отсутствовать, а использоваться менее скоростная технология PDH.

В простейшем случае первичная сеть для образования постоянных каналов может вообще отсутствовать, а коммутаторы и маршрутизаторы пакетной сети соединяются ***непосредственно кабельными или радио линиями связи***.

Последнее решение не обеспечивает необходимую ***гибкость*** сети – чтобы подключить новое устройство, необходимо физически прокладывать новую линию связи, в то время как наличие разветвленной первичной сети дает возможность организовать новый канал в сети путем перепрограммирования матрицы коммутации мультиплексоров и кросс-коннекторов.

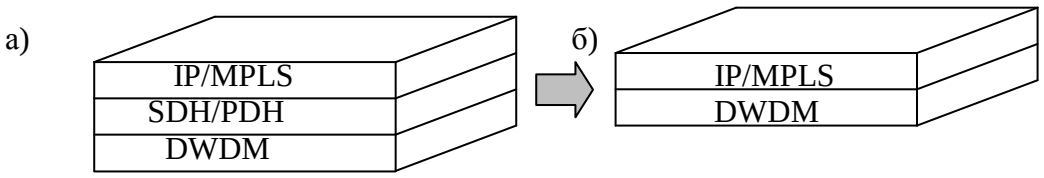
Для связи слоя IP со слоем каналов можно применить технологию ATM, основным назначением которой является создание инфраструктуры постоянных виртуальных каналов, соединяющих интерфейсы маршрутизаторов IP, работающих на четвертом, верхнем уровне глобальной сети. Для ***каждого*** типа трафика средствами технологии ATM образуется ***отдельный виртуальный*** канал, обеспечивающий требуемые для ***трафика параметры QoS*** – среднюю скорость, пульсацию, уровень задержек, уровень потерь.

Уровень IP, освобожденный в данной модели от проблем обеспечения параметров QoS, выполняет свои ***классические функции – образует составную сеть.***

Несмотря на сложность многослойной структуры, подобные сети получили большое распространение и для крупных операторов комплексных услуг являются на сегодня фактическим эталоном глобальной сети, с помощью которой можно обеспечивать все виды сервиса – как IP, так и ATM, классической телефонии и услуги предоставления цифровых каналов в аренду.

В последние годы все большее распространение получают ***«чистые» сети IP***, называемые так из-за того, что под уровнем IP нет

другой сети с коммутацией пакетов, такой как ATM. Структура «чистой» сети IP представлена на рисунке.



а – «чистая» сеть IP в настоящее время;
б – перспективная сеть IP-NG

Структура «чистой» сети IP

В такой сети цифровые каналы первичной сети по-прежнему образуются инфраструктурой двух нижних уровней, а этими каналами непосредственно пользуются интерфейсы маршрутизаторов IP, без какого-либо промежуточного слоя. В настоящее время практически все маршрутизаторы ведущих производителей поддерживают сегодня такие базовые механизмы QoS, как кондиционирование трафика, приоритетное и взвешенное обслуживание очередей,

«Чистая» сеть IP может оказаться вполне рациональным решением, к тому же более экономичным и простым в эксплуатации, чем сложная четырехуровневая модель, особенно при использовании новых технологий, основывающихся на «коммутации» меток. Технология MPLS (MultiProtocol Label Switching – МногоПротокольная Коммутация Меток) зародилась, когда основой для построения мультисервисных сетей считалась технология ATM. Однако при внедрении этой технологии была решена более широкая задача: обеспечение унифицированного механизма транспорта IP-пакетов поверх различных технологий и протоколов второго уровня (ATM, Ethernet, Frame Relay и пр.). Технология MPLS предусматривает ***установление соединений – коммутируемых по меткам трактов***. Созданная для оптимизации передачи трафика IP поверх ATM технология MPLS сама смогла выполнять многие функции, присущие ATM, позволяя обойтись вообще без ATM.

Технология MPLS реализуется в IP-маршрутизаторах.

В специальном однобайтовом поле каждого пакета – в октете Type of Service протокола IPv4 или в октете Traffic Class протокола IPv6 указываются требования к необходимому набору показателей качества обслуживания. С помощью такого октета можно определить до 32 различных уровней качества обслуживания – приоритет и класс обслуживания.

Для сигнализации в ней используются протоколы резервирования ресурсов (Resource reServation Protocol, RSVP) и распределения меток (Label Distribution Protocol), а для маршрутизации – стандартные протоколы маршрутизации в сетях IP

– OSPF и IS-IS. С ее помощью удается решить ряд проблем, связанных с обеспечением качества и управлением трафиком в IP-сетях.

Технология MPLS объединяет возможности продвижения пакетов по каналам коммутации меток, аналогичным по своим функциям ***виртуальным каналам***, с топологическими возможностями ***протоколов маршрутизации*** IP. При этом трафик перемещается по сети по устойчивым маршрутам, что позволяет управлять трафиком и обеспечивать дифференцированное качество обслуживания для различных классов трафика.

Можно считать, что под слоем протокола IP в сети IP/MPLS работает некоторая канальная технология, основанная на технике виртуальных путей, но это не самостоятельная (и очень сложная) технология, подобная ATM, а тесно интегрированная с IP технология, пользующаяся возможностями IP по классификации трафика и выяснению топологии сети совместно с текущим состоянием ресурсов сети.

Функции MPLS встраиваются в IP-маршрутизаторы, так что физически два слоя коммуникационных устройств сливаются в один.

Сети IP/MPLS по простоте своей организации приближаются к «чистым» сетям IP, а по рациональности загрузки и возможностям поддержки QoS – к сетям «IP поверх ATM».

Рассматривая перспективы развития глобальных сетей, следует отметить, что в настоящее время SDH тоже является не обязательным, но лишь одним из возможных способов передачи Интернет-трафика (а по существу, Ethernet- трафика). Разработчики SDH всячески стараются предложить вариант NGSDH, обеспечивающий экономичный метод передачи громадных, но пульсирующих, объемов трафика компьютерных сетей. В исходном варианте SDH наибольшей единицей передаваемого сигнала является VC-4, обеспечивающий передачу около двух тысяч телефонных каналов. Однако такой контейнер не позволяет передать, например, сигнал Gigabit Ethernet. Поэтому для передачи большого массива данных в транспорте SDH высокого уровня предлагается «конкатенация» - объединение нескольких (4, 16, 64 и т.д.) контейнеров VC-4. Развитием обычной конкатенации является «виртуальная» конкатенация, которая применяет более гибкие методы обработки и обеспечивает более экономичное использование трактов SDH для передачи пакетного трафика. Предлагаются гибко меняющиеся «виртуальные» коридоры для передачи пакетного трафика.

Более того, кольцевая топология и резервирование, принятые в SDH, приводят к недоиспользованию пропускной способности трактов. Поэтому традиционная SDH видимо будет исключена из структуры глобальной сети.

Наиболее вероятным вариантом построения глобальной сети следующего поколения представляется сочетание IP-сети, передающей и коммутирующей пакеты, представляющие любые виды трафика, с первичной сетью, представленной широкополосными трактами, образованными оптическими системами с волновым мультиплексированием (DWDM). Такую сеть можно назвать IP-сетью следующего поколения IP-NG (см. рисунок «б»).

5.12 Развитие технологии MPLS в IP-сетях.

Развитие MPLS на оптические сети привело к созданию обобщенной многопротокольной коммутации по меткам (Generalized MultiProtocol Label Switching, GMPLS).

GMPLS распространяет основные идеи технологии MPLS для обработки IP-трафика на нижележащие уровни сетей связи – SDH и WDM. В клиентском оборудовании IP/MPLS осуществляется КП по меткам. В сетевых элементах оптической транспортной сети могут коммутироваться:

- временные канальные интервалы, образующие тракты SDH различных уровней;
- оптические каналы, WDM и их группы, т.е. световые потоки, выделяемые на основании длины волны или диапазона длин волн (коммутация длин волн);
- оптические волокна, т.е. световые потоки, выделяемые на основании их физического расположения в пространстве.

GMPLS используется в концепции автоматически коммутируемой оптической сети (Automated Switched Optical Network, ASON). В GMPLS выделяют пять классов интерфейсов. Первые два присутствовали в технологии MPLS, остальные – то новое, что привнесла GMPLS.

Интерфейс коммутации пакетов (Packet-Switch Capable) распознает границы пакетов и направляет данные, основываясь на метке. Примеры – интерфейсы коммутирующего по меткам маршрутизатора (Label Switched Router, LSR) MPLS.

Интерфейс коммутации второго уровня (Layer 2 Switch Capable, L2SC), обеспечивающий взаимодействие с технологиями второго уровня (ATM, Frame Relay), распознает границы кадров/ячеек и направляет данные, основываясь на заголовках кадров/ячеек. Примеры – интерфейсы мостов Ethernet, использующие для коммутации заголовки MAC, и интерфейсы ATM-LSR, направляющие данные на основе VPI/VCI. (В ATM коммутация происходит по номеру виртуального соединения, который разбит на две части – идентификатор виртуального пути (Virtual Path Identifier, VPI) и идентификатор виртуального канала (Virtual Channel Identifier, VCI).

Интерфейс с временным разделением (Time Division Multiplex Capable, TDM) распознает циклы и направляет данные, основываясь на временной позиции в цикле. Примеры – интерфейсы оборудования SDH/SONET, PDH и интерфейсы оптической транспортной сети (Optical Transport Network, OTN) по Рекомендации МСЭ-Т G.709 с возможностями TDM («цифровая обертка»).

Интерфейс коммутации длин волн (Lambda-Switch Capable, LSC) не распознает ни биты, ни циклы и направляет световые потоки, основываясь на длине волны или диапазоне длин волн. Примеры – интерфейсы фотонных или оптических кроссовых соединителей, оперирующие с отдельными длинами волн или группами длин волн; оптические интерфейсы OTN.

Интерфейс коммутации волокон (Fiber-Switch Capable, FSC) не распознает ни биты, ни циклы, не видит отдельные длины волн или их диапазоны; направляет световые потоки, основываясь на их физическом положении в пространстве. Примеры – интерфейсы фотонных или оптических кроссовых соединителей, оперирующие на уровне волокон.

Таким образом, GMPLS предполагает для оптических сетей КК всех возможных видов: временную, частотную, пространственную. Для сигнализации предусматривается отдельная плоскость управления, функционирующая на принципах КП. Напрашивается аналогия с ЦСИС, в которой используется система ОКС, также основанная на КП.

Нетес В.А. Основные принципы GMPLS, Вестник связи, №2, 2005, стр. 53-55.

А.Б.Гольдштейн, Б.С. Гольдштейн. Технология и протоколы MPLS. СПб, БХВ, 2005.

А.А.Атцик, А.Б.Гольдштейн, «Солянка» про MPLS». Вестник связи, №2, 2005, стр. 55-59.

Янковский Г.Г. Качество обслуживания в сетях IP. Вестник связи, №1, 2008, стр. 65-74.

5.13 Сравнение характеристик глобальных компьютерных сетей

Заканчивая изучение глобальных компьютерных сетей, сравним характеристики основных типов сетей. Результаты сравнения приведены в таблице.

Таблица

Параметр	Тип глобальной сети			
	X.25	Frame Relay	ATM	Интернет
Коррекция ошибок	На канальном и сетевом уровнях	Только у пользователя	Только у пользователя	Только у пользователя
Требования к достоверности передачи	Низкие	Высокие	Высокие	Высокие
Организация виртуального канала	Да	Да	Да	Передача дейтаграмм
Способность фрагментации сообщений	Нет	Нет	Пакет фиксированной длины формируется непосредственно пользователем	Фрагментация соответственно параметрам локальных сетей
Маршрутизация	Жесткая	Жесткая	Жесткая	Гибкая
Виды трафика. Гарантии качества обслуживания для различных видов трафика	Асинхронная низкоскоростная ДИ Приоритезация пользователей.	Среднескоростная ПД Согласованная средняя скорость.	Мультисервисная сеть. Гарантии QoS для различных видов трафика.	Первоначально чисто компьютерная сеть. Разработка методов (MPLS) для обеспечения QoS при различных видах трафика.
Скорость передачи	До 10 кбит/с	2 Мбит/с	155 Мбит/с и более	Определяется скоростью локальных сетей
Тип линии (канала) связи	Канал ТЧ, физическая линия	ЦЛТ, ВОЛС	ВОЛС	ВОЛС, DWDM