

Поиск информации в Интернете

Телекоммуникационные технологии

Базовые принципы организации и функционирования компьютерных сетей. Локальные и глобальные сети. Адресация в сети

Компьютерная сеть (КС) — это совокупность компьютеров, соединенных линиями связи и оснащенных коммуникационным оборудованием и программным обеспечением.

Коммуникационное, или сетевое, оборудование — это периферийные устройства, которые осуществляют преобразование сигналов, используемых в компьютере, в сигналы, передаваемые по линиям связи, и наоборот. При использовании телефонных линий связи применяют **модемы**, подключаемые к компьютеру через стандартный последовательный порт. Одна из важнейших характеристик модема — это скорость передачи данных, измеряемая в битах за секунду (бод). Современные модемы имеют скорость передачи данных 1200–52000 бод и более. При использовании других линий связи применяют *сетевые адаптеры*.

Линия связи — это оборудование, с помощью которого осуществляется соединение компьютеров в сеть.

Линии связи бывают:

- **проводные** — для передачи сигналов используются кабели (например, телефонная линия);
- **беспроводные** — все остальные линии связи: спутниковая связь, каналы цифровой связи (КЦС). Пропускная способность КЦС — до сотен млн Кбод.

Объединенные в сеть компьютеры всегда работают под управлением специальных программ управления сетью.

Коммуникационное, или сетевое, программное обеспечение — это набор программ, обеспечивающий работу сетевого оборудования и обмен информацией между компьютерами в сети.

Сетевое программное обеспечение делится на две группы:

1. «низкого» уровня — управляет сетевым оборудованием с целью преобразования сигналов из одного вида в другой. Эти программы ничего «не знают» о структуре передаваемой информации;
2. «высокого» уровня — распознает и обрабатывает информацию в зависимости от ее характера и способа организации.

Компьютерные сети делятся на две группы — *локальные* и *глобальные*.

Локальные вычислительные сети

Локальная вычислительная сеть (ЛВС) — объединяет компьютеры, расположенные на небольшом расстоянии друг от друга, и представляет собой замкнутую систему. Небольшие расстояния между компьютерами позволяют использовать в локальных сетях в качестве линий связи простейшие проводные линии (кабели различных типов: коаксиальный, витая пара, оптоволоконный). Кабели подсоединяются к компьютеру через специальное устройство, называемое **сетевой картой** или **сетевым адаптером**. Это устройство вставляется в слот расширения на материнской плате компьютера. Имеются материнские платы со встроенным сетевым адаптером. Сетевые адаптеры (**ArcNet, Ethernet, TokenRing**) различаются производительностью (скоростью передачи данных) и соответственно стоимостью.

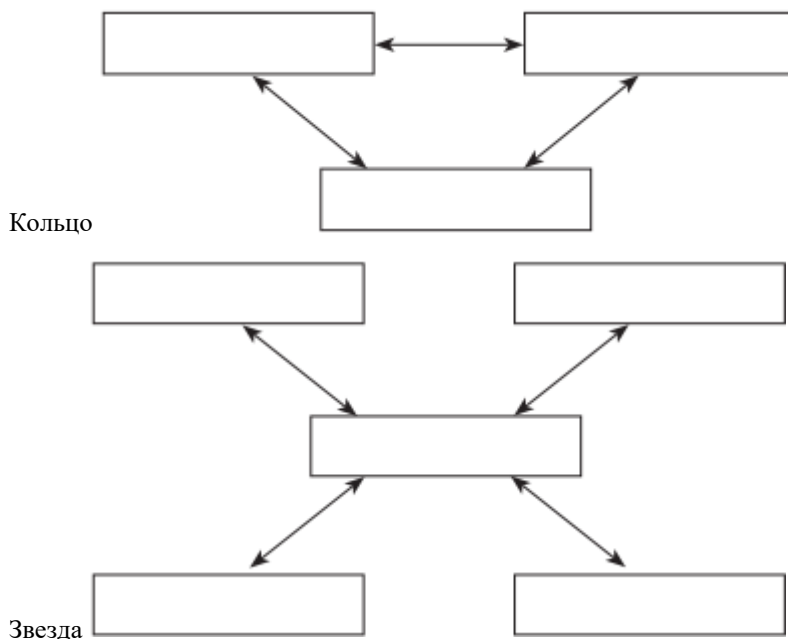
Для построения сети используются и другие сетевые устройства: *хабы, концентраторы, повторители* и др. Для подключения к локальной сети портативных компьютеров часто используется беспроводное подключение, при котором передача данных осуществляется с помощью электромагнитных волн.

Топология ЛВС — это способ объединения компьютеров в сеть между собой.

Основные топологии ЛВС:

Гирлянда





ЛВС создается для совместного использования и обмена информацией между компьютерами, а также для совместного использования ресурсов сети.

Ресурс сети — это устройство, входящее в аппаратную часть какого-либо компьютера сети, доступное и используемое любым пользователем сети. Это могут быть принтеры, сканеры, дисковые накопители большой емкости, устройства резервного копирования информации, станки с ЧПУ и т. д., а также диски, программы и данные.

Существуют два принципиально разных *способа соподчинения компьютеров в локальной сети* и соответственно технологии работы в ней.

Одноранговая сеть — сеть *равноправных компьютеров — рабочих станций*, каждый из которых имеет имя и пароль для входа в компьютер в момент загрузки операционной системы. Имя и пароль назначаются владельцем компьютера средствами операционной системы и BIOS. В такой сети могут быть организованы «подсети» — так называемые *группы*, каждая из которых имеет имя, например «Бухгалтерия». Принадлежность к какой-либо группе может быть задана или изменена пользователем. Владельцу каждого компьютера в сети предоставлена программная возможность самому преобразовывать свой локальный ресурс (диски, папки, принтер) в разделяемый (сетевой), а также устанавливать права и пароль доступа к нему. Он же отвечает за сохранность или работоспособность этого ресурса. Сети этого вида часто организуются в небольших офисах (10–15 компьютеров); они не требуют системного администратора.

Иерархические сети — это сети, в которых имеется мощный компьютер — выделенный *сервер*, ресурсы которого предоставляются другим соединенным с ним компьютерам — рабочим станциям. **Сервер (хост)** — компьютер, предоставляющий свои ресурсы для совместного использования. **Рабочая станция (клиент)** — компьютер, пользующийся ресурсами сети. Ресурсы рабочих станций серверу, как правило, не доступны. По сравнению с одноранговыми сетями обеспечивают более высокое быстродействие и надежность работы сети, повышают конфиденциальность и надежность хранения информации и др. Работу всей сети организует системный администратор.

Возможна организация сетей и *более сложных видов*:

- **сочетание одноранговой и иерархической сети** — рабочие станции могут быть как подключены к выделенному серверу по принципу иерархической сети, так и объединены между собой в одноранговую сеть;
- иерархическая сеть с **несколькими выделенными серверами**;
- иерархическая сеть, в которой есть **несколько уровней иерархии серверов**, при этом серверы нижнего уровня подключаются к серверам более высокого уровня.

Взаимодействие серверов и рабочих станций обеспечивается сетевым программным обеспечением каждого компьютера сети. На сервере устанавливается сетевая операционная система, управляющая его работой и работой всей сети. На компьютере-клиенте располагается клиентская часть программного обеспечения, которая взаимодействует с сервером. В зависимости от вида сети программы на сервере и компьютере-клиенте могут быть как одинаковыми, так и различными. Сейчас наиболее распространены

сетевые программы фирм Novell (программы Novell Netware) и Microsoft (Windows NT-сервер, сети Microsoft).

Пользователю рабочей станции доступны ресурсы сети в соответствии с заранее обусловленными правилами.

ЛВС используются в системах управления базами данных для быстрого доступа нескольких компьютеров к одной базе данных (например, при продаже авто-, авиа-, железнодорожных билетов), а также в производстве — для управления производственным процессом, выходного контроля качества, учета готовой продукции и расхода материалов.

Локальные сети обычно объединяют компьютеры, размещенные в одном здании, и не позволяют обеспечить совместный доступ к информации пользователям, находящимся, например, в разных частях города. Для решения этой проблемы создаются **региональные сети**, объединяющие компьютеры в пределах одного региона (города, страны, континента).

Многие организации, заинтересованные в защите информации от несанкционированного доступа, создают собственные, так называемые **корпоративные сети**. Такие сети могут объединять десятки тысяч компьютеров, размещенных в различных городах и странах.

Глобальные сети

Потребность в формировании единого мирового информационного пространства привела к созданию *глобальных компьютерных сетей*.

Глобальная сеть — соединения локальных, региональных и корпоративных сетей и отдельных компьютеров, находящихся на больших расстояниях друг от друга. Большие расстояния требуют наличия дополнительного устройства для обработки больших объемов информации и пересылки ее на большие расстояния. Это — **серверы глобальной сети**, представляющие собой очень мощные компьютеры.

Из-за больших расстояний использование простых линий связи в глобальных сетях невозможно. Современные глобальные сети используют телефонную связь, однако между серверами связь осуществляется по **выделенным линиям** или по **специальным каналам связи**. Они обладают высокой скоростью передачи и помехоустойчивостью. Сейчас в глобальных сетях все чаще используются системы спутниковой связи, что значительно расширяет их масштабы и возможности.

Для пользования глобальной сетью компьютер должен иметь модем и специальное программное обеспечение.

Глобальные сети предоставляют огромные возможности: можно, сидя за компьютером, прочитать новости, посетить картинную галерею, узнать прогноз погоды, обменяться письмами, прочитать книгу, получить ответы на многие вопросы и т. д. Каждый день появляются новые возможности и услуги, предоставляемые сетями.

Самая большая глобальная сеть — **Internet (Интернет)**. Она охватывает все континенты Земли.

Интернет — всемирная КС, объединяющая в единое целое десятки тысяч разнородных локальных и глобальных сетей. Эта сеть содержит огромный объем информации по любой тематике, доступной на коммерческой основе всем желающим. Помимо получения чисто информационных услуг, через Интернет можно совершать покупки и коммерческие сделки, оплачивать счета, заказывать билеты на различные виды транспорта, бронировать места в гостиницах и др.

Интернет представляет собой объединение нескольких десятков тысяч различных локальных сетей (Сеть сетей). Каждая ЛВС называется **узлом**, или **сайтом**, а юридическое лицо, обеспечивающее работу сайта, — **провайдером**. Сайт состоит из нескольких компьютеров — **серверов**, каждый из которых предназначен для хранения информации определенного типа и в определенном формате. Каждый сайт и сервер на сайте имеют уникальные имена, посредством которых они идентифицируются в Интернете.

Для подключения к Интернету пользователь должен заключить контракт на обслуживание с одним из провайдеров в его регионе. После этого любая работа в Интернете начинается с соединения с сайтом провайдера. Связь с провайдером может быть организована:

- по коммутируемому телефонному каналу с помощью модема;
- с помощью постоянно действующего выделенного канала (просто вызовом соответствующей программы для работы в Интернете).

После соединения с провайдером пользователь получает доступ ко всем сайтам и компьютерам в Интернете. Открывающиеся при этом перед пользователем возможности зависят от условий контракта с провайдером.

Основная задача Интернета — обеспечить пользователя необходимой информацией и услугами. Для этого используется технология *клиент/сервер*.

Клиент (потребитель) — программа, принимающая информацию и услуги, предоставляемые другими компьютерами — **серверами (поставщиками)**. Компьютеры, на которых запускают программы-серверы, должны обладать достаточно мощными ресурсами, т. к. им приходится «пропускать через себя» огромные объемы информации. Любой компьютер может выступить в роли клиента, как только на нем будет запущена какая-либо программа-клиент.

В компьютерных сетях используются различные марки компьютеров, типы модемов, линии связи, операционные системы и другие программы. Для того чтобы все сети «понимали» друг друга, они должны использовать единый набор правил, определяющий способ обмена информацией. Такой набор правил называется *протоколом*.

Протокол — это стандарт, определяющий формы представления и способы пересылки сообщений, процедуры их интерпретации, правила совместной работы различного оборудования в сетях, установленный по взаимному соглашению.

Все компьютеры, подключенные к Internet, должны использовать один и тот же протокол. С 1983 года это протокол **TCP/IP (Transmission Control Protocol/Internet Protocol)**. Фактически это два протокола.

IP (протокол межсетевого взаимодействия) — более ранний вариант, разработанный для пересылки данных, организованных в *пакеты*. В пакет входит сообщение и информация о маршруте, т. е. адреса компьютера-отправителя и компьютера-получателя (аналогично пересылке письма в конверте по почте). Специальные компьютеры-*маршрутизаторы* (router) определяют путь, по которому должны следовать пакеты от компьютера к компьютеру, пользуясь *таблицами* и *алгоритмами маршрутизации* (т. к. прямой связи между любыми двумя компьютерами нет). Большие сообщения делятся на несколько пакетов с номерами. По этим номерам компьютер-получатель должен полностью восстановить исходное сообщение после получения всех пакетов. Недостатком протокола IP является его неспособность справиться с ошибками при передаче информации. Каждый пакет содержит контрольные суммы, которые вычисляются определенным образом по исходной информации. Компьютер-получатель выполняет те же вычисления и находит контрольные суммы по полученной информации. Если они не совпадают с исходными, значит, в процессе передачи произошла ошибка и данные исказились. При искажении данных, нарушении последовательности доставки пакетов или прерывании процесса передачи компьютер-получатель полностью отбрасывает пакет.

TCP (протокол управления передачей) устраняет этот недостаток. В случае возникновения какой-либо ошибки принимающий компьютер запрашивает повторную передачу.

Таким образом, TCP следит за целостностью данных, разбивает большие сообщения на последовательности более мелких, организует их нумерацию и последующее восстановление в единое сообщение; IP контролирует перемещение пакетов по Интернету, выбирая маршрут пересылки данных от одного компьютера к другому, а также определяет формат адреса компьютера.

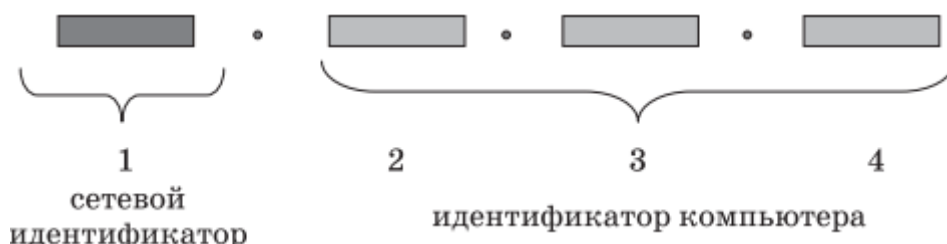
Все современные операционные системы имеют встроенную поддержку основных протоколов для работы с Интернетом.

Каждый компьютер, подключенный к Интернету, должен иметь свой уникальный адрес во избежание путаницы при пересылке информации.

В Интернете применяются *два типа адресов*.

IP-адрес — по смыслу аналогичен почтовому индексу. **IP-адрес** — это последовательность из четырех чисел, разделенных точками. Каждое из чисел занимает 1 байт, т. е. может принимать значения 0..255. Так как 1 байт = 8 бит, эти числа часто называют *октетами*. Используется три класса IP-адресов: А, В и С. Класс IP-адреса определяет, сколько октетов отводится под адрес сети и сколько — под адрес компьютера.

Класс А используется для работы с небольшим количеством сетей (до 126), содержащих большое число компьютеров ($\approx 16,8$ млн).



Класс В — для работы со средним количеством сетей (до 16384), содержащих среднее число компьютеров (до 65534).



Класс С — для работы с большим количеством сетей (до ≈ 2 млн), содержащих малое число компьютеров (до 254).



Класс определяют по значению первого октета:

- если в первом октете число от 1 до 126 — класс А;
- если в первом октете число от 128 до 191 — класс В;
- если в первом октете число от 192 до 223 — класс С.

Для передачи сообщений на конкретные компьютеры в Internet протокол TCP/IP и программы-клиенты используют IP-адреса. Но они неудобны для восприятия пользователем. Человеку привычнее работать с именами.

Все пространство адресов Интернета разбито на области — **домены**. Внутри доменов возможно разделение по определенным признакам.

Доменный адрес — уникальное символьное имя компьютера в сети. Оно состоит из частей, разделенных точками. Место назначения уточняется справа налево: сначала указывается имя компьютера, затем — имя в сети, в которой он находится.

Доменный адрес компьютера включает в себя как минимум два уровня доменов. **Домен первого уровня** определяет страну или тип организации, которой принадлежит компьютер. Существуют установленные двухбуквенные сокращения для доменов стран. Например, Россия — **ru**, США — **us**, Франция — **fr** и т. д. Домены типов организаций обычно имеют трехбуквенные сокращения. Например, учебные заведения — **edu**, правительственные учреждения — **gov**, коммерческие организации — **com**, провайдеры услуг Интернет — **net** и т. д. **Домен второго уровня** определяет организацию, которая владеет или управляет сетью, содержащей данный компьютер. Обычно имя этого домена совпадает с названием соответствующей фирмы или ее торговой маркой. Имя компьютера указывает конкретный компьютер в сети, определенной доменами первого и второго (а возможно, и следующих) уровней. Оно регистрируется только в этой сети, и только эта сеть «ответственна» за передачу информации конкретному компьютеру-адресату.

Примеры доменных имен:

win.smtp.dol.ru

www.google.com.ua

Домены могут сочетать региональные и организационные уровни, например: edu.ru — образовательные организации в России.

Доменные адреса должны быть преобразованы в IP-адреса. «Обычный» компьютер не может и не должен знать все IP-адреса Интернета. На **DNS-серверах (Domain Name System)** хранятся таблицы соответствия доменных и IP-адресов. DNS-серверы распределены по всей сети Internet. Каждый из них хранит информацию о большом числе компьютеров и способен мгновенно преобразовать доменное имя в IP-адрес. Если IP-адрес запрошенного компьютера не известен данному DNS-серверу, он обратится к ближайшему DNS-серверу и т. д. по цепочке, пока требуемый адрес не будет найден. Это займет несколько секунд. Адрес одного из DNS-серверов пользователь должен указать при настройке компьютера для работы в Интернете. Его можно получить у администратора локальной сети или провайдера.

В Интернете для поиска Web-документов используется система адресации **URL (Uniform Resource Locator)** — унифицированный указатель ресурсов). Она применяется для указания способа организации информации на конкретном хосте и идентификации размещенного на нем информационного ресурса. URL включает в себя протокол доступа к документу, доменное имя или IP-адрес сервера, на котором находится документ, а также путь к файлу и собственно имя файла. Протокол доступа к документу определяет способ передачи информации. Для доступа к Web-страницам используется протокол передачи гипертекста **HTTP**. При записи протокола после его имени следует двоеточие и две косые черты: **http://**.

Например, URL имеет вид

http://home.microsoft.com/intl/ru/www_tour.html

Элементы этого адреса обозначают:

- **http://** — префикс, указывающий тип протокола. В данном случае означает, что адрес относится к хосту, который является WWW-сервером. В качестве префикса (протокола) могут быть также указаны: **ftp://**, **file://**, **news://**;
- **home.microsoft.com** — доменное имя хоста. После него может следовать число, обозначающее порт, через который будет производиться подключение к хосту;
- **/intl/ru/** — папка ru корневого каталога *intl* хоста;
- **www_tour.html** — имя файла. URL-адрес не должен содержать пробелов.

Он состоит из латинских букв и некоторых символов (_ , ~ , —). Прописные и строчные буквы различаются. Путь к папке и название файла могут отсутствовать, а в конце могут указываться некоторые параметры. Например:

http://yandex.ru:8081/ya_detail.html#P7

<http://www.surplusauction.com/auction.sa?SI=Geo9>

<http://www.stars.ru/cgi-bin/decads.pl?search=%C1%F3%EA%E8&ion=iserv>

Услуги компьютерных сетей: World Wide Web, электронная почта, файловые архивы, поисковые системы, чат и пр.

Для различных видов работы в Интернете существуют свои службы (*сервисы*).

Наиболее популярны следующие службы:

- электронная почта;
- группы новостей UseNet (телеконференции);
- всемирная информационная служба Web;
- служба обмена файлами FTP;
- службы, обеспечивающие «живое» общение, аудио- и видеовещание в Интернете.

Для работы в Интернете имеются как универсальные программные комплексы, обеспечивающие доступ к любой службе Интернета, так и специализированные программы, обычно предоставляющие более широкие возможности при работе с конкретным сервисом Интернета.

Electronic mail (E-mail) — электронная почта. Это одна из первых служб, а также из числа наиболее популярных. Ее достоинства — дешевизна, оперативность, доступность.

Почтовый адрес в Интернете состоит из двух частей, разделенных символом @:

имя пользователя @ имя сервера

Первая часть почтового адреса (имя пользователя) имеет произвольный характер и задается самим пользователем при регистрации почтового ящика. *Вторая часть (имя сервера)* является доменным именем почтового сервера, на котором пользователь зарегистрировал свой почтовый ящик. Почтовый адрес записывается только латинскими буквами и не должен содержать пробелов. Например, имена почтовых ящиков пользователей, зарегистрированных на бесплатном почтовом сервере mail.ru будут иметь вид: user_name@mail.ru.

Работа E-mail основана на специальных протоколах, а также на взаимодействии двух программ — *почтового сервера* и *почтового клиента*.

Существует две группы протоколов, по которым работает электронная почта:

Протоколы SMTP и POP (или POP3). Протокол **SMTP (Simple Mail Transport Protocol)** поддерживает передачу сообщений между адресатами Интернета. Он позволяет группировать сообщения в адрес одного получателя, размножать копии E-mail-сообщений для передачи на разные адреса. Протокол **POP (Post Office Protocol 3)** обеспечивает конечному пользователю доступ к поступившим на его адрес электронным сообщениям. POP-клиенты запрашивают пароль, что повышает конфиденциальность переписки.

Протокол IMAP. Этот протокол получает все большее распространение. Он дает возможность пользователю работать с письмами непосредственно на сервере провайдера и экономить время работы в Интернете.

Для работы с электронной почтой создано много программ. В операционной системе Windows есть встроенная программа **Internet Mail**. Также используются программы **MS Exchange, Eudora, Outlook Express, Netscape Messenger** и др.

Основные функции почтовых программ:

- забор писем из почтового ящика на сервере;
- отправка подготовленных писем;
- дешифровка сообщений, полученных в различных кодировках;
- управление приоритетом сообщений (срочное, обычное, по дешевому тарифу);
- сортировка принятых и отправляемых писем по папкам;
- ведение адресной книги;
- ответ на приходящие письма, пересылка писем;
- создание нового письма;
- возможность прикрепления к письму, имеющему обычный текстовый вид, картинки или любого другого файла.

Служба USENET — это группы новостей или телеконференции. Они дают возможность читать и посылать сообщения в общественные (открытые) дискуссионные группы. Каждая тема обсуждается в отдельной группе, имеющей свое имя.

Общение участников Usenet организовано с помощью **серверов новостей** или, как их называют, **NNTP-серверов (Network News Transfer Protocol)** — протокол передачи сетевых новостей). Серверы разбросаны по всему миру. При подключении к Internet провайдеры обеспечивают и доступ к серверу новостей. Программа-клиент Usenet передает сообщение на сервер новостей, тот — на связанные с ним серверы новостей, каждый из них — на следующую группу связанных серверов и т. д.

Каждая конференция представляет собой совокупность текстовых сообщений — *статей* — ее подписчиков; помещение статьи в конференцию называется *публикацией*.

Существует несколько тысяч групп новостей, поэтому они организованы в иерархическую структуру категорий и подкатегорий. Имена групп новостей строятся слева направо — от более общего к более конкретному. Категории и подкатегории отделяются точками. Например: *comp.os.ms-windows* — можно получить информацию об операционной системе Windows. Здесь *comp* — категория верхнего уровня, включающая все темы, связанные с компьютерами; *os* — подкатегория, предназначенная для обсуждения операционных систем; *ms-windows* — подкатегория следующего уровня, предназначенная для обсуждения конкретных операционных систем серии Windows фирмы Microsoft.

Существует много программ для работы с Usenet. Можно пользоваться программами Outlook Express или Netscape Collabra. В настоящее время Outlook Express и Netscape Messenger — две самые популярные программы, поддерживающие работу одновременно с почтой и новостями.

Программы работы с конференциями обеспечивают:

- указание множества конференций, в работе которых намерен участвовать пользователь. Эта операция называется *подпиской*, а множество конференций, на которые осуществлена подписка, — *списком подписки* (он может быть в любой момент изменен);
- просмотр фамилий авторов и заголовков (тем) статей в любой конкретной конференции из списка подписки;
- просмотр содержимого статей и при необходимости сохранение их в файле в заданном каталоге компьютера пользователя;
- публикацию своей собственной статьи;
- персональный ответ автору какой-либо статьи на его E-mail-адрес;
- публичный ответ автору конкретной статьи, который будет фигурировать как статья конференции.

World Wide Web, или просто **Web**, — всемирная паутина (сеть). Она представляет собой огромный набор взаимосвязанных друг с другом документов, хранящихся на разных компьютерах.

Документы называются **Web-страницами**, компьютеры — **Web-серверами (Web-узлами, Webсайтами)**.

Среда WWW не имеет централизованной структуры и может рассматриваться как информационное пространство, которое заполняется теми, кто желает разместить свои документы в Интернете.

С помощью WWW обеспечивается доступ почти ко всем ресурсам крупнейших библиотек мира, музейным коллекциям, музыкальным произведениям, периодическим изданиям, к законодательным и правительственным постановлениям, справочникам и оперативным подборкам на любую тему (погода, курсы валют, стоимость всевозможных услуг и товаров), аналитическим обзорам.

Из чисто информационной WWW все больше становится посреднической системой, обеспечивающей заключение контрактов, покупку товаров и расчеты по ним, бронирование билетов на транспорт, выбор и заказ экскурсионных маршрутов и тысячи других услуг. С помощью WWW проводится опрос общественного мнения, ведутся интерактивные диспуты с ведущими политиками, коммерсантами, деятелями культуры.

Особенности WWW, обеспечившие ее высокую популярность:

- гипертекстовая организация информационных элементов — страниц WWW. Это означает, что страница WWW может содержать гипертекстовые ссылки, оформленные в виде рисунка или выделенного текста;
- возможность включения в страницы WWW современных мультимедийных средств (графики, звука, анимации), а также средств по художественному оформлению страниц (макетированию);
- возможность передачи на сайт владельца WWW-страницы различной информации;
- наличие бесплатного, хорошего и достаточно простого программного обеспечения, позволяющего непрофессиональному пользователю не только просматривать, но и самому создавать WWW-страницы;
- наличие хороших поисковых систем и удобных средств запоминания адресов размещения необходимой информации;
- возможность быстрого перемещения вперед/назад по уже просмотренным страницам, отсутствие необходимости знать IP-адреса хостов, на которых размещена необходимая информация;
- наличие средств обеспечения надежности и конфиденциальности информационного обмена.

Для доступа к Web-страницам используется протокол передачи гипертекста **HTTP**. При записи протокола после его имени следует двоеточие и две косые черты: **http://**.

Служба WWW использует систему адресации **URL (Uniform Resource Locator** — универсальный URL-адрес достаточно сложно, поэтому во всех программных средствах работы в Интернете имеется инструмент Избранное, с помощью которого любому URL можно дать содержательное имя (в том числе на русском языке), называемое *ссылкой*.

Работа с ресурсами WWW обеспечивается специальными программами — *браузерами*. В настоящее время наиболее распространенными браузерами являются **Internet Explorer** и **Netscape Communication**. Браузер может работать в сети или автономно (для просмотра html-файлов, ранее запомненных на компьютере).

Файловые архивы — это серверы Интернета, на которых хранятся миллионы файлов различных типов (программы, драйверы устройств, графические и звуковые файлы и т. д.). Файловые серверы поддерживают многие компании — разработчики программного обеспечения и производители аппаратных компонентов компьютера и периферийных устройств. Размещаемое на таких серверах программное обеспечение является свободно распространяемым или условно бесплатным.

File Transfer Protocol — протокол для передачи файлов между компьютерами. При указании URL-адреса файла этот протокол записывается так: **ftp://**. Файлы, предназначенные для открытого доступа, содержатся на множестве FTP-серверов. Программа FTP-клиент позволяет установить связь с одним из таких серверов и затем просматривать каталоги, осуществлять поиск файлов, управлять их перемещением. FTP может использоваться как для получения файлов с удаленных серверов, так и для отправки своих файлов на сервер, чтобы другие пользователи могли ими воспользоваться. Возможна пересылка как текстовых, так и двоичных файлов. Доступ к файлам на FTP-серверах возможен и по протоколу HTTP.

Обмен файлами производится с помощью специализированных программ — **FTP-клиентов (AceFTP, CuteFTP)**. Web-браузеры Internet Explorer и Netscape Navigator также имеют средства для работы с FTP-серверами.

Для установления соединения при помощи FTP в окне программы нужно указать имя удаленного компьютера, ввести имя пользователя и пароль.

Эффективная работа с Интернетом невозможна без мощных систем поиска необходимой информации. По *способу организации поиска* и предоставляемым возможностям все средства поиска условно можно разбить на следующие группы:

- каталоги и специализированные базы данных;
- поисковые и метапоисковые системы.

Популярные русскоязычные каталоги и поисковые системы: **Rambler, Яндекс, Google, Апорт, Mail.ru.**

Современные поисковые системы часто являются *информационными порталами*, которые предоставляют пользователям не только возможности поиска документов в Интернете, но и доступ к другим информационным ресурсам: новостям, информации о погоде, валютном курсе и т. д.

IRC-серверы (Internet Relay Chat) — серверы, на которых реализуется интерактивное общение в форме обмена текстовыми сообщениями, аудио- или видеоконференций. Для организации интерактивного общения необходимо специальное программное обеспечение (например, программа **NetMeeting**, входящая в состав Internet Explorer).

Интерактивное общение можно также организовать с помощью **ICQ-серверов**. Для того чтобы стать абонентом системы ICQ, достаточно скачать программу ICQ-клиент с файлового сервера и в процессе ее установки на компьютер зарегистрироваться (получить уникальный идентификационный номер). После подключения к Интернету пользователь может начинать общение с любым зарегистрированным в системе ICQ и подключенным в данный момент к Интернету пользователем.

Поиск информации в Интернете

Для поиска информации в Интернете используются специальные **поисковые серверы**, которые содержат постоянно обновляемую информацию о Web-страницах, файлах и других документах, хранящихся на серверах Интернета. *Поисковые серверы* можно разделить на две группы:

1. поисковые системы общего назначения;
2. специализированные поисковые системы.

Поисковые системы общего назначения являются базами данных, содержащими тематически сгруппированную информацию об информационных ресурсах Всемирной паутины. Заполнение баз данных осуществляется с помощью специальных программ-роботов, которые периодически «обходят» Web-серверы Интернета, читают встречающиеся документы, выделяют в них ключевые слова и заносят в базу, содержащую URL-адреса документов.

Способы организации поиска:

1. **Поиск по каталогам.** Каталоги в WWW аналогичны систематизированным библиотечным каталогам. Поиск по каталогам состоит в последовательном движении по иерархическому списку ссылок, называемых *рубриками*, или *категориями*. Двигаясь вниз по детализирующим категориям, можно найти страницы с нужной информацией. На каждой странице, открываемой при движении по каталогу, указывается последовательность просмотренных рубрик, например: Деловой мир: Финансы: Аналитика и т. д.
2. **Поиск по ключевым словам.** Поисковая система дает возможность пользователю по определенным правилам сформулировать требования к необходимой ему информации — создать запрос. Простой запрос содержит одно или несколько ключевых слов, которые, по мнению

пользователя, являются главными для искомого документа. Можно также создать сложный запрос, используя логические операции, шаблоны и т. д. В домашней странице поисковых машин всегда имеется информация о языке запросов. После отправки запроса поисковая машина автоматически просматривает документы на контролируемых (*индексируемых*) ею сайтах и отбирает те из них, которые соответствуют запросу (*релевантны запросу*). В результате поиска создается одна или несколько страниц, содержащих ссылки на релевантные запросу документы. Для каждой такой ссылки обычно указываются также дата создания страницы, объем, степень релевантности запросу, фрагменты текста, характеризующие содержание страницы. Щелчок на такой ссылке загружает нужную страницу. В случае очень большого количества найденных документов можно уточнить запрос и в соответствии с ним повторить поиск, но только среди уже отобранных страниц (*поиск в найденном*). В ряде поисковых систем можно отметить ссылку на страницу, содержание которой наиболее удовлетворяет потребностям пользователя, и повторить поиск, потребовав Искать похожие. При использовании различных поисковых машин следует иметь в виду, что они принципиально отличаются количеством и составом просматриваемых сайтов, языком запросов, алгоритмами определения степени соответствия документа запросу. Поэтому количество и состав WWW-страниц, отобранных для одного и того же запроса различными машинами поиска, может значительно различаться.

Специализированные поисковые системы позволяют искать информацию в других информационных слоях Интернета — серверах файловых архивов, почтовых серверах и др.

Для поиска файла достаточно ввести его имя — и поисковая система (например, filesearch.ru) выдаст URL-адреса мест его хранения. Если имя файла неизвестно, но зато известно его назначение (например, драйвер), то можно воспользоваться *тематическим каталогом драйверов*.

Специализированные поисковые системы позволяют также найти адрес электронной почты по имени человека или наоборот (например, система WhoWhere?).

Примеры решения задач

Пример 1. В таблице приведены запросы к поисковому серверу. Расположите обозначения запросов в порядке возрастания количества страниц, которые найдет поисковый сервер по каждому запросу.

Примечание. Для обозначения логической операции «ИЛИ» используется символ «|», а для операции «И» — символ «&».

А	законы & физика
Б	законы (физика & биология)
В	законы & физика & биология & химия
Г	законы физика биология

Решение.

1 способ. Больше всего страниц будет найдено по запросу Г, т. к. при его выполнении будут найдены все страницы, содержащие хотя бы одно слово из запроса. Меньше всего страниц будет найдено по запросу В, т. к. он требует присутствия на искомой странице всех четырех слов.

Сравним запросы А и Б. По запросу А будут найдены страницы, содержащие одновременно слова «законы» и «физика». По запросу Б будет найдено больше страниц, чем по запросу А, т. к. кроме страниц, содержащих слово «законы» (например, юридические), будут найдены также страницы, содержащие одновременно слова «физика» и «биология» (но, возможно, не содержащие слова «законы»).

В результате логических рассуждений получаем цепочку: ВАБГ.

2 способ (графический). Рассмотрим множество страниц, содержащих каждое из искомых слов. Запросу $X \& Y$ будет соответствовать пересечение множеств X и Y , а запросу $X | Y$ — их объединение. Воспользуемся графическим представлением действий над множествами. Множество страниц, содержащих некоторое слово, будем обозначать эллипсом. Множество, получившееся в результате запроса, будем закрашивать серым цветом.

Диаграмма для запроса А изображена на рисунке:

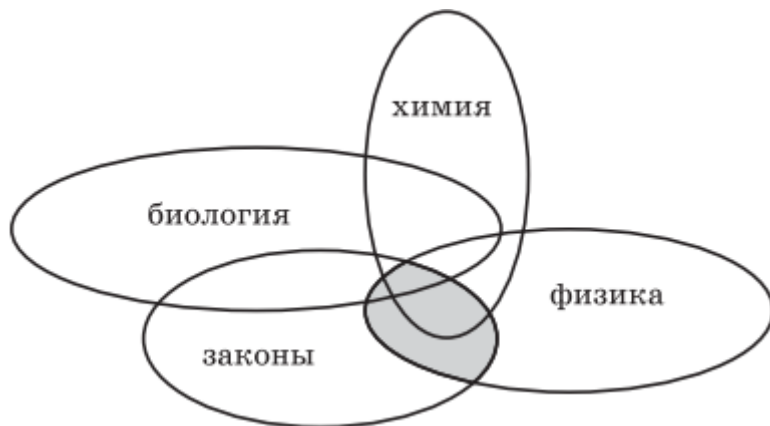


Диаграмма для запроса Б:

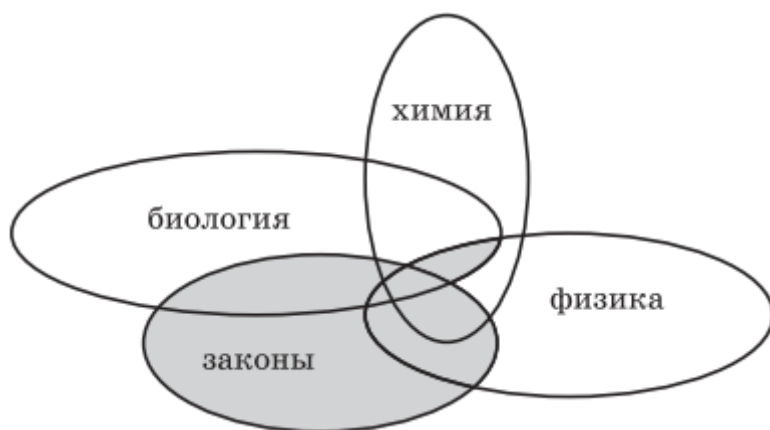


Диаграмма для запроса В:

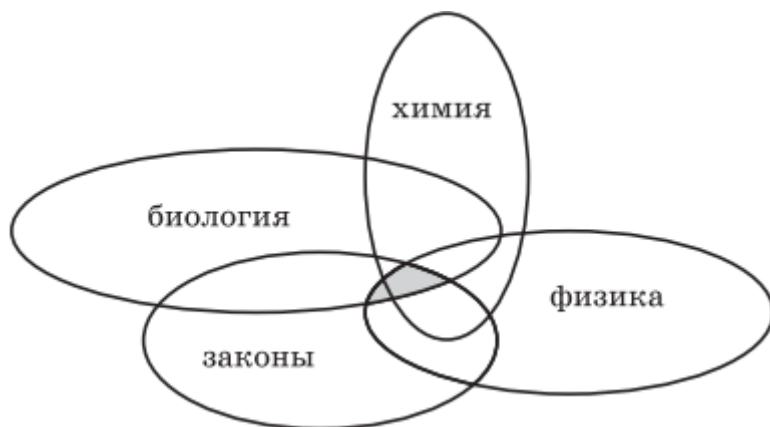


Диаграмма для запроса Г:

указания значения *атрибута*. Значение атрибута отделяется от ключевого слова атрибута символом «=» и заключается в кавычки, например:

```
<H1 ALIGN = "LEFT">
```

При отображении документа HTML сами тэги не отображаются, но влияют на способ отображения документа HTML.

Структура документа HTML

Простейший правильный документ HTML, содержащий все тэги, определяющие структуру, выглядит следующим образом:

```
<HTML>
  <HEAD> <TITLE> Заголовок документа </TITLE> </HEAD>
  <BODY>
    Текст документа
  </BODY>
</HTML>
```

Здесь использованы ключевые слова:

HTML — начало и конец документа HTML;

HEAD — начало и конец раздела заголовка;

TITLE — начало и конец общего заголовка документа;

BODY — начало и конец тела документа.

Основными функциональными элементами документа HTML являются **заголовки** и **абзацы**. Язык HTML поддерживает *шесть уровней заголовков*, которые задаются при помощи тэгов от

Обычные абзацы задаются с помощью парного тэга <P>. В HTML нет средств для задания абзацного отступа. Абзацы отделяются пустой строкой. Закрывающий тэг </P> рассматривается как необязательный. Подразумевается, что он стоит перед тэгом, который задает очередной абзац. Если в тексте присутствует символ конца строки, при воспроизведении в браузере он рассматривается как простой пробел. Для перехода на другую строку используется непарный тэг
. В качестве ограничителя абзацев может использоваться горизонтальная линейка, задаваемая тэгом <HR>, например:

```
<HR ALIGN = "RIGHT" SIZE = "10" WIDTH = " 50% ">.
```

Этот тэг задает горизонтальную линейку высотой в 10 пикселей, занимающую половину ширины окна и расположенную справа.

С помощью тэга и его атрибутов FACE, SIZE и COLOR можно задать гарнитуру, размер и цвет шрифта любого фрагмента текста.

Часто удобно размещать текст на Web-странице в виде списка. Язык HTML поддерживает следующие виды списков:

Нумерованный список. Список располагается внутри контейнера . Каждый элемент списка определяется тэгом . Тип нумерации можно задать с помощью атрибута TYPE тэга : арабские цифры (по умолчанию), «I» — римские цифры, «a» — строчные буквы.

Маркированный список. Список располагается внутри контейнера . Каждый элемент списка определяется тэгом . Вид маркера можно задать с помощью атрибута TYPE тэга : «disc» — диск, «square» — квадрат, «circle» — окружность.

Список определений. Список располагается внутри контейнера <DL> </DL>. Текст оформляется в виде термина, который выделяется непарным тэгом <DT>, и определения, которое следует за тэгом <DD>.

Рисунки являются очень важными элементами оформления Web-страниц. Привлекательность и информативность сайтов в Интернете во многом определяется использованием графики. Рисунки чаще всего хранятся в файлах формата .jpg или .gif, обычно в отдельной папке (браузеры поддерживают и другие графические форматы). Кроме указанных форматов, допускается использование видеозображения в формате .avi.

Подключение рисунков выполняется с помощью дескриптора `<img>`, имеющего следующие основные атрибуты: `align` — способ выравнивания графики, например: `center`; `border` — видимая рамка вокруг графики; `dynsrc` — исходный URL видео; `height` — высота графики; `loop` — число повторений видео (если задано `infinite`, видео будет повторяться непрерывно); `src` — исходный URL графики; `start` — условие запуска (например, `start = mouseover` запускает видео при перемещении указателя мыши по области видео); `width` — ширина графики. Например: `<img src = "Car.jpg">`.

Как только в Web-страницу будет встроена **гиперсвязь (гиперссылка)**, документ можно назвать **гипертекстом**. Гиперссылка может указывать:

- на позицию в своем документе (внутренние гиперссылки);
- на другой документ на своем сервере;
- на произвольный объект по любому адресу Интернета.

Гиперссылка состоит из двух частей: *указателя ссылки* и *адресной части*.

Указатель ссылки — это текст или рисунок на Web-странице, щелчок по которому вызывает переход на другую страницу.

Адресная часть гиперссылки — это URL-адрес документа, на который устанавливается ссылка. URL-адрес может быть *абсолютным* и *относительным*.

Абсолютный URL-адрес — это полный Internet-адрес со всей информацией, требуемой клиенту для того, чтобы отыскать сервер и успешно установить с ним связь. Его следует использовать во всех случаях, когда нужна ссылка на документы другого сервера Интернета.

Относительный URL-адрес указывает на местоположение документа относительно того, в котором находится указатель ссылки. Обычно используется для адресации в пределах документа или совокупности документов на одном сервере.

Гипертекстовая ссылка задается парным тэгом `<A>`, который содержит обязательный атрибут `HREF`. В качестве значения атрибута используется URL-адрес документа, на который указывает ссылка. Например:

`<A HREF=http://www.site.com/index.htm>`.

Здесь задан адрес в абсолютной форме. Обычно в такой форме задается ссылка на внешний документ. При использовании относительного адреса ссылка рассматривается как внутренняя, например:

`<A HREF="text.htm">`.

Достоинство внутренней ссылки в том, что она сохраняет свою работоспособность при изменении адреса Web-узла.

Дескриптор гиперсвязи `<A>` иногда называется «**якорем**» (само наименование дескриптора представляет собой аббревиатуру от английского слова `anchor` — якорь).

Полный формат гиперссылки включает возможность ссылки на определенное место внутри страницы. Для этого соответствующее место помечается с помощью якоря. Якорь задается тэгом `<A>` с атрибутом `name`, например:

`<A name="MyLabel">`.

Значение этого атрибута — произвольная последовательность латинских букв и цифр, рассматриваемая как имя якоря. Для ссылки на якорь его имя указывается в конце URL-адреса после символа «`#`»:

`<A HREF=http://www.site.com/index.htm#MyLabel>`.

Текст, располагающийся между дескрипторами `<A>` и `</A>`, обычно изображается браузерами в цвете (для этих целей по умолчанию используется голубой цвет), некоторые программы подчеркивают его. Посредством атрибутов в дескрипторе `<body>` цвет текста гиперсвязи можно изменить.

Перед тем как разместить созданный сайт в Интернете, его нужно протестировать: просмотреть страницы в различных браузерах и убедиться, что текст нормально читается, рисунки расположены удачно, а гиперссылки работают правильно. Для публикации сайта нужно найти подходящее место на одном из серверов Интернета и получить от провайдера URL-адрес сайта, а также секретные имя пользователя и пароль, которые необходимы администратору сайта для его редактирования.