

Оглавление

КОМПЬЮТЕРНЫЕ СЕТИ

КОМПЬЮТЕРНЫЕ СЕТИ	2
Компьютерная сеть.....	2
Все многообразие компьютерных сетей можно классифицировать по группе признаков:	2
Классификация по размеру, охваченной территории	2
Локальная вычислительная сеть	2
Классификация по типу функционального взаимодействия.....	3
Точка-точка	3
Одноранговая	3
Клиент-сервер	3
Многослойная архитектура	3
Семейство протоколов TCP/IP	3
IP-адрес	3
DNS - система доменных имен	4
Иерархия в системе DNS.	4
Домены первого уровня разделяются на тематические и географические.	4

ИНТЕРНЕТ

ИНТЕРНЕТ	5
Слово Интернет	5
Структура и принципы Всемирной паутины.....	6
Браузер.....	6
Браузер	6
Первый в мире веб-сайт	7
Устройство веб-сайтов	7
Страницы веб-сайтов	7
Интернет-портал	7
Виды порталов	8
Публичные порталы	8
Корпоративные порталы	8
Технологии Всемирной паутины	8
Идентификаторы ресурсов	8
Система адресации URL	8
Протоколы прикладного уровня	9
HTTP	9
FTP	11
SMTP	11
Файлообменная сеть.....	12
Типы организации файлообменных сетей	12
Централизованные сети	12
Децентрализованные сети	12
Частично децентрализованные (гибридные) сети	12
Принципы работы	12
Известные файлообменные сети	13

КОМПЬЮТЕРНЫЕ СЕТИ

Компьютерная сеть (вычислительная сеть, сеть передачи данных) — система связи между двумя или более компьютерами. Для передачи информации могут быть использованы различные физические явления, как правило — различные виды электрических сигналов или электромагнитного излучения.

Компьютерная сеть - совокупность взаимосвязанных через каналы передачи данных компьютеров, обеспечивающих пользователей средствами обмена информацией и коллективного использования ресурсов сети:

- аппаратных,
- программных,
- информационных.

Все многообразие компьютерных сетей можно классифицировать по группе признаков:

- По размеру, охваченной территории
- По типу функционального взаимодействия
- По типу сетевой топологии
- По функциональному назначению
- По сетевым ОС
- По необходимости поддержания постоянного соединения

Классификация по размеру, охваченной территории

- Персональная сеть (PAN, Personal Area Network)
- Локальная сеть (LAN, Local Area Network)
- HomePNA
- Городская сеть (MAN, Metropolitan Area Network)
- Национальная сеть
- Глобальная вычислительная сеть (WAN, Wide Area Network)

Локальная вычислительная сеть

Локальная вычислительная сеть (ЛВС) - группа ЭВМ, а также периферийное оборудование, объединенные в пределах одного или нескольких близлежащих зданий автономными (не арендуемыми) высокоскоростными каналами передачи цифровых данных (проводными или беспроводными: медными, волоконно-оптическими, СВЧ или ИК-диапазона). Служит для решения информационных задач (например, в рамках какой-либо организации), а также совместного использования объединенных информационных и вычислительных ресурсов. ЛВС могут иметь в своем составе средства для выхода в Интернет.

Каждый компьютер, включаемый в локальную сеть должен иметь сетевую плату, в разъем которой и подключается связующий кабель. Кабели, выходящие из различных компьютеров объединяются в устройстве, называемом *сетевой концентратор* (switch, HUB). Сетевые концентраторы также могут иметь связь друг с другом, объединяя вместе подсети различных участков здания. Таким образом, обеспечивается прохождение сигналов между всеми устройствами, включенными в сеть.

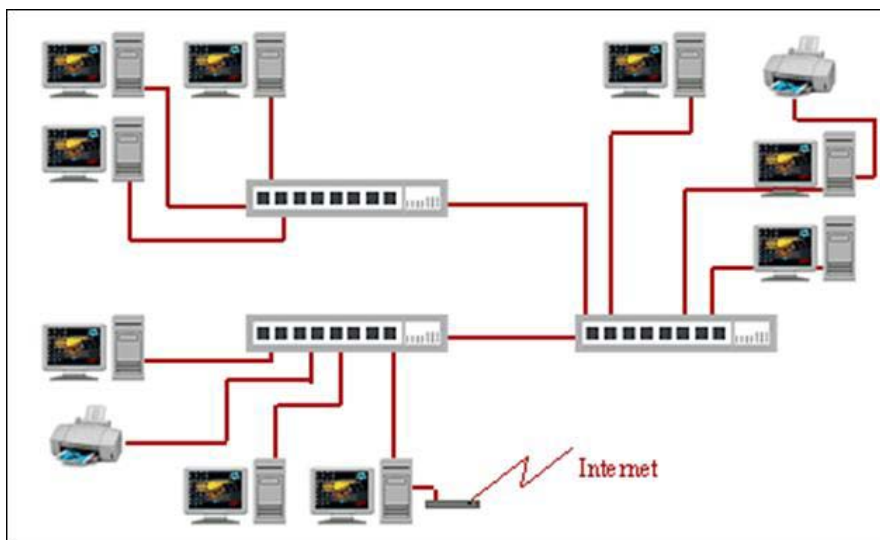


Рис. Одноранговая ЛВС.

Преимущества работы в локальной сети:

Возможность хранения данных персонального и общего использования на дисках файлового сервера.

Возможность постоянного хранения ПО, необходимого многим пользователям, в единственном экземпляре на дисках файлового сервера.

Обмен информацией между всеми компьютерами сети.

Одновременная печать всеми пользователями сети на общесетевых принтерах.

Обеспечение доступа с любого компьютера локальной сети к ресурсам Интернет, при наличии единственного коммуникационного узла глобальной сети.

Классификация по типу функционального взаимодействия

Точка-точка простейший вид компьютерной сети, при котором два компьютера соединяются между собой напрямую через коммуникационное оборудование. Достоинством такого вида соединения является простота и дешевизна, недостатком — соединить таким образом можно только 2 компьютера и не больше. Часто используется когда необходимо быстро передать информацию с одного компьютера, например, ноутбука, на другой.

Одноранговая (P2P) (от англ. peer-to-peer, P2P — один на один, с глазу на глаз) сети — это компьютерные сети, основанные на равноправии участников. В таких сетях отсутствуют выделенные серверы, а каждый узел (peer) является как клиентом, так и сервером. В отличие от архитектуры клиент-сервер, такая организация позволяет сохранять работоспособность сети при любом количестве и любом сочетании доступных узлов.

Клиент-сервер сетевая архитектура, в которой устройства являются либо клиентами, либо серверами. Клиентом (front end) является запрашивающая машина (обычно ПК), сервером (back end) — машина, которая отвечает на запрос. Оба термина (клиент и сервер) могут применяться как к физическим устройствам, так и к программному обеспечению.

Многослойная архитектура - в этой архитектуре вместо единого сервера применяются серверы приложений и серверы баз данных. Их использование позволяет резко увеличивать производительность сети.

Семейство протоколов TCP/IP

Для того чтобы при обмене данными компьютеры, объединенные в сеть, действовали согласованно, разработан ряд стандартов и правил, называемых протоколами. Весь набор сетевых протоколов, на которых базируется Интернет называется TCP/IP (ти си пи ай пи). Название образовано из аббревиатур двух базовых протоколов — TCP, отвечающего за гарантированную транспортировку данных по каналам связи, и IP, содержащего правила адресации.

Протоколы, входящие в семейство TCP/IP разделяются на уровни.

Физический уровень описывает среду передачи данных (будь то кабель, оптоволокно или радиоканал), физические характеристики такой среды и принцип передачи данных (разделение каналов, модуляцию, амплитуду сигналов, частоту сигналов, способ синхронизации передачи, время ожидания ответа и максимальное расстояние).

Канальный уровень описывает, каким образом передаются пакеты данных через физический уровень, включая кодирование (т.е. специальные последовательности битов, определяющих начало и конец пакета данных). Примеры протоколов канального уровня — Ethernet, IEEE 802.11 Wireless Ethernet, SLIP, Token Ring, ATM и MPLS.

Сетевой уровень изначально разработан для передачи данных из одной сети (подсети) в другую. Протоколы транспортного уровня могут решать проблему гарантированной доставки сообщений («дошло ли сообщение до адресата?»), а также гарантировать правильную последовательность прихода данных. Транспортные протоколы определяют, для какого именно приложения предназначены эти данные.

На прикладном уровне работает большинство сетевых приложений. Эти программы имеют свои собственные протоколы обмена информацией, например, HTTP для WWW, FTP (передача файлов), SMTP (электронная почта), SSH (безопасное соединение с удалённой машиной), DNS (преобразование символьных имён в IP-адреса) и многие другие.

IP-адрес (ай-пи адрес, сокращение от англ. Internet Protocol Address) — уникальный идентификатор (адрес) устройства (обычно компьютера), подключённого к локальной сети или интернету.

Чтобы компьютеры, объединенные в сеть, могли обмениваться сообщениями, каждый из них должен иметь уникальный адрес. **IP-адрес представляет собой 32-битовое** (по версии IPv4) или 128-битовое (по версии IPv6) двоичное число. *Удобной формой записи IP-адреса (IPv4) является запись в виде четырёх десятичных чисел (от 0 до 255), разделённых точками, например, 192.168.0.1.* (или 128.10.2.30 — традиционная десятичная форма представления адреса, а 10000000 00001010 00000010 00011110 — двоичная форма представления этого же адреса).

IP-адрес двоичный	11011100	11010111	00001110	00010110
IP-адрес десятичный	220	215	14	22

В точечно-десятичной нотации IP-адрес может выглядеть, например, так: 220.215.14.22. Каждая часть, разделенная точкой, представляет собой один байт, и, следовательно, максимальное десятичное число, которое может быть представлено одним байтом 255 (28=256, от 0 до 255). IP-адреса представляют собой основной тип адресов, на основании которых сетевой уровень протокола IP передаёт пакеты между сетями. IP-адрес назначается администратором во время конфигурирования компьютеров и маршрутизаторов.

IP-адрес состоит из двух частей: номера сети и номера узла. В случае изолированной сети её адрес может быть выбран администратором из специально зарезервированных для таких сетей блоков адресов (192.168.0.0/16, 172.16.0.0/12 или 10.0.0.0/8). Если же сеть должна работать как составная часть Интернета, то адрес сети выдаётся провайдером либо региональным интернет-регистратором (Regional Internet Registry, RIR). Согласно [1] существует пять RIR: ARIN, обслуживающий Северную Америку; APNIC, обслуживающий страны Юго-Восточной Азии; AfriNIC, обслуживающий страны Африки; LACNIC, обслуживающий страны Южной Америки и бассейна Карибского моря; и RIPE NCC, обслуживающий Европу, Центральную Азию, Ближний Восток. Региональные регистраторы получают номера автономных систем и большие блоки адресов у ICANN, а затем выдают номера автономных систем и блоки адресов меньшего размера локальным интернет-регистраторам (Local Internet Registries, LIR), обычно являющихся крупными провайдерами.

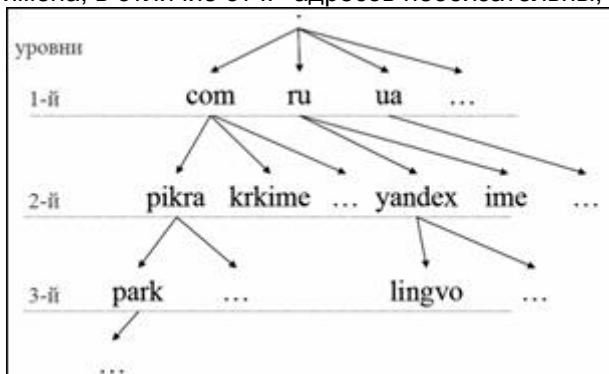
Номер узла в протоколе IP назначается независимо от локального адреса узла. Маршрутизатор по определению входит сразу в несколько сетей. Поэтому каждый порт маршрутизатора имеет собственный IP-адрес. Конечный узел также может входить в несколько IP-сетей. В этом случае компьютер должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом, IP-адрес характеризует не отдельный компьютер или маршрутизатор, а одно сетевое соединение. Но, для человека такая система адресации сложна, так же как нам сложно помнить, набирать и диктовать одиннадцатизначные телефонные номера, поэтому в 1984 году Полом Мокапетрисом была разработана *надстройка над IP-адресацией*, называемая системой *DNS (domain name system, система доменных имен)*.

DNS - система доменных имен

Доменные имена системы DNS – синонимы IP-адреса, так же, как имена в адресной книжке вашего телефона – синонимы телефонных номеров. Они символьные, а не числовые; они удобнее для запоминания и ориентации; они несут смысловую нагрузку.

www.inet.ru -> таблицы DNS -> 193.232.70.36

Доменные имена также уникальны, т.е. нет в мире двух одинаковых доменных имен. Доменные имена, в отличие от IP-адресов необязательны, они приобретаются дополнительно.



Иерархия в системе DNS.

Так же уникальны адреса, которые указываются на конвертах при доставке писем обычной почтой. В мире нет стран с одинаковыми названиями. И если названия городов иногда и повторяются, то в сочетании с делением на более крупные административные единицы типа районов и областей они становятся уникальными. А названия улиц не должны повторяться в пределах одного города. Таким образом, адрес на основе географических и административных названий однозначно определяет точку назначения.

Домены имеют аналогичную иерархию. Имена доменов отделяются друг от друга точками: lingvo.yandex.ru, krkime.com.

Домены первого уровня разделяются на тематические и географические.

Таблица Тематические домены 1-ого уровня

COM	<i>Commercial (для коммерческих организаций)</i>
NET	Networks (Интернет, телекоммуникационные сети)
ORG	Organizations (некоммерческие организации либо организации, не попадающие в другие категории)
INFO	Information (открытый для всех доменов)
BIZ	Business Organizations (аналог com)
NAME	Personal (для частных лиц)
INT	International Organizations (международные организации)
EDU	<i>Educational (образовательные проекты США)</i>
MIL	US Dept of Defense (департамент безопасности США)
GOV	US Government (правительство США)
MUSEUM	Museums (музеи)
AERO	Air-transport industry (воздушно-транспортная индустрия)
COOP	Cooperatives (кооперативы)

Регистрация доменного имени второго уровня в тематических доменных зонах доступна для организации или частного лица независимо от географического положения. Стоимость владения доменным именем в такой зоне не превышает \$35 в год. Юридического оформления владения не требуется, требуется только выбрать подходящее доменное имя из числа незанятых и перечислить деньги. Одна из организаций, осуществляющих регистрацию имен в этих доменных зонах – Network Solution (<http://www.netsol.com>).

На январь 2007 года в мире насчитывалось 243 территориальных (национальных) доменных зоны, среди них:

AU - AUstralia (Австралия)
 EU - EUrope (Европейский союз)
 FI - FIland (Финляндия)
 FR - FRance (Франция)
 IL - IsraeL (Израиль)
 JP - JaPan (Япония)
 RU - RUssian Federation (Россия)
 SU - Soviet Union (Советский Союз)
 UA - UkrainA (Украина)
 UK - United Kingdom (Англия)
 US - United States (США)

Территориальные домены первого уровня, в отличие от тематических, всегда двухбуквенные. Исторически сложилось так, что Россия владеет двумя национальными доменами: .RU и .SU. Последний остался за РФ после развала Советского Союза. Впрочем, в настоящий момент ведется пересмотр территориальных доменов, и в ближайшем будущем Россия может лишиться зоны .SU.

Регистрация доменных имен второго уровня в зонах .RU и .SU производится организацией RU-CENTER (<http://www.nic.ru>). Стоимость доменного имени в зоне .RU - \$20+НДС в год, в зоне .SU - \$100+НДС в год. При регистрации требуется юридическое оформление договорных отношений, которое занимает около месяца, с учетом пересылки документов в Москву и из Москвы.

ИНТЕРНЕТ

Слово Интернет (Internet) происходит от словосочетания Interconnected networks (связанные сети), это глобальное сообщество малых и больших сетей.

Обращаясь в Интернет, мы пользуемся услугами Интернет-провайдера или ISP (Internet Service Provider - поставщик услуг Интернета).

ISP - это организация, которая имеет собственную высокоскоростную сеть, объединенную с другими сетями по всему земному шару. Провайдер подключает к своей сети клиентов, которые становятся частью сети данного провайдера и одновременно частью всех объединенных сетей, которые и составляют Интернет.

В основе объединения малых и больших сетей (которые и составляют Интернет) лежит цепь договорных соглашений. Каждый пользователь Интернета имеет договор с определенным провайдером о подключении к его сети. В простейшем случае этот договор может выглядеть как одноразовая карта доступа, содержащая всю необходимую информацию для подключения пользователя в локальную сеть провайдера: номера телефонов модемного пула провайдера, имя

и пароль пользователя для доступа в сеть. В свою очередь провайдеры договариваются о соединении их сетей.

Мы определили Интернет как глобальное сообщество малых и больших сетей. Данное определение касается лишь внешней стороны явления. **В более широком смысле Интернет - это информационное пространство, распределенное среди миллионов компьютеров во всем мире, которые постоянно обмениваются данными.** Основная задача Интернета - это связь. Связь круглосуточная, высоконадежная. Для того чтобы осуществлять такую высоконадежную связь, была разработана специальная Интернет-технология доставки данных.

Структура и принципы Всемирной паутины

Всемирную паутину образуют миллионы веб-серверов сети Интернет, расположенных по всему миру. **Веб-сервер** является программой, запускаемой на подключённом к сети компьютере и использующей протокол HTTP для передачи данных. В простейшем виде такая программа получает по сети HTTP-запрос на определённый ресурс, находит соответствующий файл на локальном жёстком диске и отправляет его по сети запросившему компьютеру. Более сложные веб-серверы способны динамически распределять ресурсы в ответ на HTTP-запрос. Для идентификации ресурсов (зачастую файлов или их частей) во Всемирной паутине используются единообразные идентификаторы ресурсов URI (англ. Uniform Resource Identifier). Для определения местонахождения ресурсов в сети используются единообразные локаторы ресурсов URL (англ. Uniform Resource Locator). Такие URL-локаторы сочетают в себе технологию идентификации URI и систему доменных имён DNS (англ. Domain Name System) — доменное имя (или непосредственно IP-адрес в числовой записи) входит в состав URL для обозначения компьютера (точнее — одного из его сетевых интерфейсов), который исполняет код нужного веб-сервера.

Для обзора информации, полученной от веб-сервера, на клиентском компьютере применяется специальная программа — веб-браузер. Основная функция веб-браузера — отображение гипертекста. Всемирная паутина неразрывно связана с понятиями гипертекста и гиперссылки. Большая часть информации в Вебе представляет собой именно гипертекст. Для облегчения создания, хранения и отображения гипертекста во Всемирной паутине традиционно используется язык HTML (англ. HyperText Markup Language), язык разметки гипертекста. Работа по разметке гипертекста называется вёрсткой, мастера по разметке называют веб-мастером или вебмастером (без дефиса). После HTML-разметки получившийся гипертекст помещается в файл, такой HTML-файл является самым распространённым ресурсом Всемирной паутины. После того, как HTML-файл становится доступен веб-серверу, его начинают называть «веб-страницей». Набор веб-страниц образует веб-сайт. В гипертекст веб-страниц добавляются гиперссылки. Гиперссылки помогают пользователям Всемирной паутины легко перемещаться между ресурсами (файлами) вне зависимости от того, находятся ресурсы на локальном компьютере или на удалённом сервере. Гиперссылки веба основаны на технологии URL.

Браузер

Веб-обозреватель, браузер (от англ. Web browser; вариант броузер — неправильно) — программное обеспечение для просмотра веб-сайтов, то есть для запроса веб-страниц (преимущественно из Сети), их обработки, вывода и перехода от одной страницы к другой. *Большинство браузеров умеют также показывать оглавление FTP-серверов.*

Браузер - программа, установленная на компьютере пользователя и позволяющая просматривать документы в определённых форматах (html, xml и др.). Позволяет ввести в поле адреса URL сайта и, при наличии соединения с Интернетом, получить указанную страницу с веб-сервера.

Браузеры постоянно развивались со времени зарождения Всемирной паутины и с её ростом становились всё более востребованными программами. **Ныне браузер — комплексное приложение для обработки и вывода разных составляющих веб-страницы и для предоставления интерфейса между веб-сайтом и его посетителем.** Практически все популярные браузеры распространяются бесплатно или «в комплекте» с другими приложениями: *Internet Explorer (неотъемлемая часть Microsoft Windows)*, *Mozilla Firefox* (бесплатно, свободное ПО), **Safari** (совместно с Mac OS или бесплатно для Windows), **Opera** (бесплатно, начиная с версии 8.50), **Google Chrome** (бесплатно, свободное ПО).

Часто начинающие пользователи путают два понятия - Интернет и WWW (или Web). Следует напомнить, что **WWW (Всемирная паутина World Wide Web) - это лишь одна из многочисленных услуг (сервисов), предоставляемых пользователям Интернета.** Проведем аналогию с почтовыми службами. Существует множество видов экспресс-почты (DHL, Fedex и т.д.), имеющих свои особенности, но все они пользуются сетью транспортных коммуникаций.

Веб-сайт, или просто сайт (англ. website, от web — паутина, «веб» и site — «место») — это одна или совокупность веб-страниц, доступных в интернете через протоколы HTTP/HTTPS;

совокупность всех общедоступных веб-сайтов есть Всемирная паутина. Страницы веб-сайта объединены общим корневым адресом, а также обычно темой, логической структурой, оформлением и/или авторством.

Ранее понятие смешивалось с физическим узлом сети — хостом, сервером (узлом). Но с ростом Интернета и технологическим улучшением серверов на одном компьютере стало возможно размещение множества сайтов и доменов.

Первый в мире веб-сайт появился 6 августа 1991 года. <http://info.cern.ch/> Его создатель Тим Бернерс-Ли (Tim Burners-Lee) опубликовал на нём описание новой технологии World Wide Web (WWW), основанной на протоколе передачи данных HTTP, системе адресации URI и языке гипертекстовой разметки HTML. Также на сайте были описаны принципы установки и работы веб-серверов и веб-браузеров. Сайт стал и первым в мире интернет-каталогом, так как позже Тим Бернерс-Ли разместил на нём список ссылок на другие сайты.

Все инструменты, необходимые для работы первого сайта Бернерс-Ли подготовил ещё раньше — в конце 1990 появились первый гипертекстовый веб-браузер WorldWideWeb с функционалом веб-редактора, первый веб-сервер на базе NeXTcube и первые веб-страницы.

«Отец» веба считал, что гипертекст может служить основой для сетей обмена данными, и ему удалось претворить свою идею в жизнь. Ещё в 1980 году Тим Бернерс-Ли создал гипертекстовое ПО Enquire, использующее для хранения данных случайные ассоциации. Затем, работая в Европейском центре ядерных исследований в Женеве (CERN), он предложил коллегам публиковать гипертекстовые документы, связанные между собой гиперссылками. Бернерс-Ли продемонстрировал возможность гипертекстового доступа к внутренним поисковику и документам, а также новостным ресурсам интернета. В результате в мае 1991 г. в CERN был утверждён стандарт WWW.

Тим Бернерс-Ли является «отцом» основополагающих технологий веба — HTTP, URI/URL и HTML, хотя их теоретические основы были заложены ещё раньше. В 40-х гг. 20-го века Ванневар Буш выдвинул идеи расширения памяти человека с помощью технических устройств, а также индексации накопленной человечеством информации для её быстрого поиска. Теодор Нельсон и Даг Энгельбарт предложили технологию гипертекста — «ветвящегося» текста, предоставляющего читателю разные варианты чтения. Xanadu, так и не законченная гипертекстовая система Нельсона, была предназначена для хранения и поиска текста, в который введены взаимосвязи и «окна». Нельсон мечтал связать перекрёстными ссылками все тексты, созданные человечеством. В настоящее время Тим Бернерс-Ли возглавляет основанный им Консорциум Всемирной паутины (World Wide Web Consortium), который занимается разработкой и внедрением стандартов интернета.

Устройство веб-сайтов

Страницы веб-сайтов — это файлы с текстом, размеченным на языке HTML или XHTML. Эти файлы, будучи загруженными посетителем на его компьютер, **обрабатываются программой-обозревателем, называемой браузером и выводятся на средство отображения пользователя** (монитор, экран КПК, принтер или синтезатор речи). Язык HTML/XHTML позволяет форматировать текст, различать в нём функциональные элементы, создавать гипертекстовые ссылки (гиперссылки) и вставлять в отображаемую страницу изображения, звукозаписи и другие мультимедийные элементы. Отображение страницы можно изменить добавлением в неё таблицы стилей на языке CSS или сценариев на языке JavaScript.

Страницы сайтов могут быть простыми статичными наборами файлов или создаваться специальной компьютерной программой на сервере — так называемым движком сайта. Движок может быть либо сделан на заказ для отдельного сайта, либо готовым продуктом, рассчитанным на некий класс сайтов. Некоторые из движков могут обеспечить владельцу сайта возможность гибкой настройки структурирования и вывода информации на вебсайте; такие называются системами управления содержанием.

Изготовление сайтов как работающих целостных информационных ресурсов есть составной процесс, вовлекающий труд различных специальностей. Общий термин о сайтостроительной деятельности — «веб-разработка».

Навигация - набор кнопок или других управляющих элементов, обеспечивающих удобство работы с сайтом.

Интернет-портал

Интернет-портал (от англ. portal «главный вход; ворота») — веб-сайт, предоставляющий пользователю Интернета различные интерактивные сервисы, работающие в рамках одного веб-сайта, такие как почта, поиск, погода, новости, форумы, обсуждения, голосования и т.д..

Виды порталов

Горизонтальными принято называть порталы, охватывающие много тем, — такие, как Yahoo или Яндекс; типичным является возникновение портала вокруг поисковой системы.

Вертикальными называются специализированные тематические порталы. Принято также подразделять порталы на интернациональные и региональные (Yahoo принадлежит к первому типу, а Яндекс, сконцентрированный преимущественно на Рунете, ко второму). Вдобавок порталы подразделяются на публичные и корпоративные.

Публичные порталы — ориентированы на всех Web-пользователей. Они являются Интернет-эквивалентом публичных библиотек; каждый может войти и повозиться со всем, что представлено на экране (Yahoo!, MSN).

Корпоративные порталы часто ставят в противоположность публичным. Корпоративный портал отличается ориентацией на специфичный тип пользователей (сотрудники, партнёры). Хотя такие порталы нередко имеют и внешний интерфейс для публичного использования, он отличается по предоставляемым возможностям от внутреннего.

Существует также другая трактовка определения горизонтального и вертикального порталов.

Горизонтальным называется портал, не зависящий от корпоративной информационной системы предприятия. То есть такой портал должен иметь собственную базу пользователей, авторизацию, и т. д.

Вертикальным порталом называется web-сайт, интегрированный в корпоративную информационную систему (КИС) компании. Интеграция может быть осуществлена на уровне входа на портал (к примеру, может использоваться LDAP или Kerberos сервер) или на более низких уровнях, как например интеграция интернет-магазина с системой бухгалтерии компании.

Поисковые сервера Русскоязычные Рамблер Апорт Яндекс ИнфоАрт Google Filesearch.ru	Зарубежные Alta Vista Carnegie Mellon Lycos Web Search Yahoo EZ-Find at The River WebCrawler Infoseek Search Com
--	--

Технологии Всемирной паутины

В целом можно заключить, что Всемирная паутина стоит на «трёх китах»: HTTP, HTML и URL. Хотя в последнее время HTML начал несколько сдавать свои позиции и уступать их более современным технологиям разметки: XHTML и XML. XML (англ. eXtensible Markup Language) позиционируется как фундамент для других языков разметки. Для улучшения визуального восприятия веба стала широко применяться технология CSS, которая позволяет задавать единые стили оформления для множества веб-страниц. Ещё одно нововведение, на которое стоит обратить внимание, — система обозначения ресурсов URN (англ. Uniform Resource Name).

Популярная концепция развития Всемирной паутины — создание семантической паутины.

Семантическая паутина — это надстройка над существующей Всемирной паутиной, которая призвана сделать размещённую в сети информацию более понятной для компьютеров.

Семантическая паутина — это концепция сети, в которой каждый ресурс на человеческом языке был бы снабжён описанием, понятным компьютеру. Семантическая паутина открывает доступ к чётко структурированной информации для любых приложений, независимо от платформы и независимо от языков программирования. Программы смогут сами находить нужные ресурсы, обрабатывать информацию, классифицировать данные, выявлять логические связи, делать выводы и даже принимать решения на основе этих выводов. При широком распространении и грамотном внедрении семантическая паутина может вызвать революцию в Интернете. Для создания понятного компьютеру описания ресурса, в семантической паутине используется формат RDF (англ. Resource Description Framework), который основан на синтаксисе XML и использует идентификаторы URI для обозначения ресурсов. Новинки в этой области — это RDFS (англ. RDF Schema) и SPARQL (англ. Protocol And RDF Query Language) (произносится как «спаркл»), новый язык запросов для быстрого доступа к данным RDF.

Идентификаторы ресурсов

Система адресации URL

Чтобы найти в Интернете какой-либо документ, достаточно знать ссылку на него - так называемый **универсальный указатель ресурса (URL - Uniform Resource Locator)**, который определяет местонахождение каждого файла, хранящегося на компьютере, подключенном к Интернету.

Адрес URL является сетевым расширением понятия полного имени ресурса в операционной системе. В URL, кроме имени файла и директории, где он находится, указывается сетевое имя компьютера, на котором этот ресурс расположен, и протокол доступа к ресурсу, который можно использовать для обращения к нему. Система адресации URL и адресация почтовой службы имеют сходную структуру.

URL (Uniform Resource Locator) - уникальный интернет-адрес веб-страницы или какого-то другого интернет-ресурса.

Структура URL

Изначально локатор URL был разработан как система для максимально естественного указания на местонахождения ресурсов в сети. Локатор должен был быть легко расширяемым и использовать лишь ограниченный набор ASCII-символов (к примеру, пробел никогда не применяется в URL). В связи с этим, возникла следующая традиционная форма записи URL:

<схема>://<логин>:<пароль>@<хост>:<порт>/<URL-путь>

Рассмотрим структуру следующего URL: <http://www.lipov.narod.ru/prog/lab/IVAN.htm> .

Первая часть **http://** (**HyperText Transfer Protocol** - протокол передачи гипертекста, по которому обеспечивается доставка документа с Web-сервера Web-браузеру) указывает программе просмотра (браузеру), что для доступа к ресурсу применяется данный сетевой протокол.

В URL первым стоит указатель на тип доступа к запрашиваемому файлу, а затем его адрес.

Вторая часть www.lipov.narod.ru *указывает на доменное имя* и адресует конкретный компьютер. Третья часть [prog/lab/IVAN.htm](http://www.lipov.narod.ru/prog/lab/IVAN.htm) показывает программе-клиенту, где на данном компьютере-сервере искать ресурс. В рассматриваемом случае ресурсом является файл в формате html, а именно IVAN.htm, который находится в папке lab, которая в свою очередь расположена в папке prog. Имена каталогов, содержащиеся в URL, обычно являются виртуальными и не имеют ничего общего с реальными именами каталогов компьютера, на котором выполняется Web-сервер, а являются их псевдонимами: ни один владелец компьютера, на котором выполняется Web-сервер, не позволит постороннему пользователю, обращающемуся к Web-серверу через Интернет, получить доступ к реальной файловой системе этого компьютера. Обратите внимание: при написании URL важно различать прописные и строчные буквы.

URI (англ. Uniform Resource Identifier) — унифицированный (единообразный)

идентификатор ресурса. На английский манер произносится как [ю-ар-ай], по-русски чаще говорят [ури]. URI — это последовательность символов, идентифицирующая абстрактный или физический ресурс. Ранее назывался Universal Resource Identifier — универсальный идентификатор ресурса.

URI — это символьная строка, позволяющая идентифицировать какой-либо ресурс: документ, изображение, файл, службу, ящик электронной почты и т. д. Прежде всего, речь идёт, конечно, о ресурсах сети Интернет и Всемирной паутины. URI предоставляет простой и расширяемый способ идентификации ресурсов. Расширяемость URI означает, что уже существуют несколько схем идентификации внутри URI, и ещё больше будет создано в будущем.

Самые известные примеры URI — это URL и URN. URL — это URI, который, помимо идентификации ресурса, предоставляет ещё и информацию о местонахождении этого ресурса. А URN — это URI, который идентифицирует ресурс в определённом пространстве имён (и, соответственно, в определённом контексте). Например, URN [urn:isbn:0-395-36341-1](http://www.isbn.org/urn:isbn:0-395-36341-1) — это URI, который указывает на ресурс (книгу) 0-395-36341-1 в пространстве имён ISBN, но, в отличие от URL, URN не указывает на местонахождение этого ресурса. Впрочем, в последнее время появилась тенденция говорить просто URI о любой строке-идентификаторе, без дальнейших уточнений. Так что, возможно, термины URL и URN скоро уйдут в прошлое.

Протоколы прикладного уровня

HTTP (англ. HyperText Transfer Protocol — «протокол передачи гипертекста») — протокол прикладного уровня передачи данных (исначально — в виде гипертекстовых документов). Основой HTTP является технология «клиент-сервер», то есть предполагается существование потребителей (клиентов), которые инициируют соединение и посылают запрос, и поставщиков (серверов), которые ожидают соединения для получения запроса, производят необходимые действия и возвращают обратно сообщение с результатом. HTTP в настоящее время повсеместно используется во Всемирной паутине для получения информации с веб-сайтов. В 2006 году в Северной Америке доля HTTP-трафика превысила долю P2P-сетей и составила 46 %, из которых почти половина — это передача потокового видео и звука[1].

HTTP используется также в качестве «транспорта» для других протоколов прикладного уровня, таких как SOAP, XML-RPC, WebDAV.

Основным объектом манипуляции в HTTP является ресурс, на который указывает URI (англ. Uniform Resource Identifier) в запросе клиента. Обычно такими ресурсами являются хранящиеся на сервере файлы, но ими могут быть логические объекты или что-то абстрактное. Особенностью протокола HTTP является возможность указать в запросе и ответе способ представления одного и того же ресурса по различным параметрам: формату, кодировке, языку и т. д. Именно благодаря возможности указания способа кодирования сообщения клиент и сервер могут обмениваться двоичными данными, хотя данный протокол является текстовым.

HTTP — протокол прикладного уровня, аналогичными ему являются FTP и SMTP. Обмен сообщениями идёт по обыкновенной схеме «запрос-ответ». Для идентификации ресурсов HTTP использует глобальные URI. В отличие от многих других протоколов, HTTP не сохраняет своего состояния. Это означает отсутствие сохранения промежуточного состояния между парами «запрос-ответ». Компоненты, использующие HTTP, могут самостоятельно осуществлять сохранение информации о состоянии, связанной с последними запросами и ответами. Браузер, посылающий запросы, может отслеживать задержки ответов. Сервер может хранить IP-адреса и заголовки запросов последних клиентов. Однако сам протокол не осведомлён о предыдущих запросах и ответах, в нём не предусмотрена внутренняя поддержка состояния, к нему не предъявляются такие требования.

Достоинства

Простота

Протокол настолько прост в реализации, что позволяет с лёгкостью создавать клиентские приложения.

Расширяемость

Вы можете легко расширять возможности протокола благодаря внедрению своих собственных заголовков, сохраняя совместимость с другими клиентами и серверами. Они будут игнорировать неизвестные им заголовки, но при этом вы можете получить необходимый вам функционал при решении специфической задачи.

Распространённость

При выборе протокола HTTP для решения конкретных задач немаловажным фактором является его распространённость. Как следствие, это обилие различной документации по протоколу на многих языках мира, включение удобных в использовании средств разработки в популярные IDE, поддержка протокола в качестве клиента многими программами и обширный выбор среди хостинговых компаний с серверами HTTP.

Недостатки и проблемы

Большой размер сообщений

Использование текстового формата в протоколе порождает соответствующий недостаток: большой размер сообщений по сравнению с передачей двоичных данных. Из-за этого возрастает нагрузка на оборудование при формировании, обработке и передаче сообщений. Для решения данной проблемы в протокол встроены средства для обеспечения кэширования на стороне клиента, а также средства компрессии передаваемого контента. Нормативными документами по протоколу предусмотрено наличие прокси-серверов, которые позволяют получить клиенту документ с наиболее близкого к нему сервера. Также в протокол было внедрено дельта-кодирование, чтобы клиенту передавался не весь документ, а только его изменённая часть.

Отсутствие «навигации»

Хотя протокол разрабатывался как средство работы с ресурсами сервера, у него отсутствуют в явном виде средства навигации среди этих ресурсов. Например, клиент не может явным образом запросить список доступных файлов, как в протоколе FTP. Предполагалось, что конечный пользователь уже знает URI необходимого ему документа, закачав который, он будет производить навигацию благодаря гиперссылкам. Это вполне нормально и удобно для человека, но затруднительно, когда стоят задачи автоматической обработки и анализа всех ресурсов сервера без участия человека. Решение этой проблемы лежит полностью на плечах разработчиков приложений, использующих данный протокол.

HTTP cookie

Cookie - файл, создаваемый браузером пользователя и содержащий определённые данные. Используется для определения уникальности пользователя, времени его последнего визита, личных настроек, сделанных пользователем и т.д.

Кюки (слово не склоняется; от англ. cookie — печенье) — небольшой фрагмент данных, созданный веб-сервером и хранимый на компьютере пользователя в виде файла, который веб-клиент (обычно веб-браузер) каждый раз пересылает веб-серверу в HTTP-запросе при попытке открыть

страницу соответствующего сайта. Применяется для сохранения данных на стороне пользователя, на практике обычно используется для:

аутентификации пользователя;
хранения персональных предпочтений и настроек пользователя;
отслеживания состояния сессии доступа пользователя;
ведения статистики о пользователях.

Без куки не могут функционировать многие сайты с ограничениями доступа, большинство интернет-магазинов. Настройка оформления и поведения многих веб-сайтов по индивидуальным предпочтениям пользователя тоже основана на куки.

Имеется и ряд заблуждений о куки. Они главным образом основаны на уверенности людей, что куки являются компьютерными программами. На самом деле, куки — это простые текстовые данные и они не могут выполнять какие-либо действия самостоятельно. В частности, куки не могут быть ни вирусами, ни шпионскими программами.

FTP (англ. File Transfer Protocol — протокол передачи файлов) — протокол, предназначенный для передачи файлов в компьютерных сетях. FTP позволяет подключаться к серверам FTP, просматривать содержимое каталогов и загружать файлы с сервера или на сервер; кроме того, возможен режим передачи файлов между серверами.

FTP является одним из старейших прикладных протоколов, появившимся задолго до HTTP, в 1971 году. До начала 90-х годов на долю FTP приходилось около половины трафика в сети Интернет. Он и сегодня широко используется для распространения ПО и доступа к удалённым хостам. Протокол FTP относится к протоколам прикладного уровня и для передачи данных использует транспортный протокол TCP. Команды и данные, в отличие от большинства других протоколов передаются по разным портам. Порт 20 используется для передачи данных, порт 21 для передачи команд.

Проблема безопасности

Протокол не шифруется, при аутентификации передаёт логин и пароль открытым текстом. Если злоумышленник находится в одном сегменте сети с пользователем FTP, то в случае построения сети на хабе (если сеть строится на свитчах по умолчанию это исключено, но злоумышленник может перехватить трафик если сменит mac адрес своего сетевого адаптера, на mac адрес адаптера жертвы, поскольку коммутация выполняется на 2ом уровне модели OSI (по mac адресам), используя пассивный сниффер, он может перехватить логин и пароль пользователя, или, при наличии специального ПО, получать передаваемые по FTP файлы без авторизации. Чтобы предотвратить перехват трафика, необходимо использовать протокол шифрования данных SSL, который поддерживается многими современными FTP-серверами и некоторыми FTP-клиентами.

SMTP (англ. Simple Mail Transfer Protocol — простой протокол передачи почты) — это сетевой протокол, предназначенный для передачи электронной почты в сетях TCP/IP.

ESMTP (англ. Extended SMTP) — масштабируемое расширение протокола SMTP. В настоящее время под «протоколом SMTP», как правило, подразумевают ESMTP и его расширения.

SMTP используется для отправки почты от пользователей к серверам и между серверами для дальнейшей пересылки к получателю. Для приёма почты почтовый клиент должен использовать протоколы POP3 или IMAP.

Наиболее популярным методом передачи файлов в сети является модель "клиент-сервер". Центральный сервер посылает файл целиком каждому, кто в нем нуждается - так работают протоколы http и ftp. Клиенты общаются только с сервером - и никогда - между собой.

Основным, и пожалуй единственным преимуществом этого метода является то, что он прост в реализации и использовании. Этот метод работает до тех пор, пока сервер, находится в сети и может обслуживать запросы всех клиентов. В противном случае - и этом мы видим достаточно часто, эта модель может давать сбои - как только сервер становится слишком популярным, и как только размер данных начинает быть слишком большим. Именно поэтому вы можете наблюдать сообщения "слишком много пользователей" пытаться скачать демку новой игры или новый CD с сервера. Концепция "зеркалирования" лишь отчасти решает эту проблему, поскольку требует сложной настройки "отзеркаливания" серверов и распределения загрузки между ними.

Файлообменная сеть

Другой метод передачи файлов стал более популярным - р2р системы, такие как Kazaa, eDonkey, Gnutella, Direct Connect, etc. В таких сетях обычные пользователи сети обмениваются файлами, напрямую подсоединяясь друг к другу. Преимущество этих систем в том, что файлы не обязательно должны находиться на одном сервере. При этом, однако, невозможно отследить содержимое этих файлов. Именно поэтому эти сети получили огромное распространение среди любителей музыки, фильмов и особенно warez. Проблемой этих сетей стали популярные данные, к которым обращается наибольшее количество людей. Поскольку за новыми файлами буквально "выстраивались очереди" то скорость скачивания уменьшалась пропорционально количеству запросов на этот файл. Кроме того в большинстве р2р систем количество соединений между пользователями ограничено, и зачастую пользователи обратившиеся за каким либо файлом попросту получали отказ. Неполностью скачанные файлы также были проблемой для пользователей- поскольку заранее определить "целый" ли запрошенный файл или нет невозможно. Результатом этих проблем стало падение популярности многих сетей и введение запретительных правил для пользователей, участвующих в них.

Файлообменная сеть — собирательное название одноранговых компьютерных сетей для совместного использования файлов, основанных на равноправии участвующих в обмене файлами, то есть каждый участник одновременно является и клиентом, и сервером.

Одна из областей применения технологии одноранговых сетей — это обмен файлами.

Пользователи файлообменной сети выкладывают какие-либо файлы в т. н. «расшаренную» (англ. share — делиться) директорию, содержимое которой доступно для скачивания другим пользователям. Какой-нибудь другой пользователь сети посылает запрос на поиск какого-либо файла. Программа ищет у клиентов сети файлы, соответствующие запросу, и показывает результат. После этого пользователь может скачать файлы у найденных источников. В современных файлообменных сетях информация загружается сразу с нескольких источников. Ее целостность проверяется по контрольным суммам.

Типы организации файлообменных сетей

Централизованные сети

Несмотря на то, что каждый участник файлообменной сети является и клиентом и сервером, необходима инфраструктура для объединения разрозненных клиентов между собой в определённое сообщество. В централизованных файлообменных сетях служебную информацию хранят индексационные серверы.

Достоинством такой сети является относительная простота программирования такой сети и небольшой объём служебной информации, передаваемой на серверы. Закрывание сетей Napster и WinMX показало ненадёжность централизованных файлообменных сетей. Индексационные серверы могут быть отключены по разным причинам (технический сбой, разорение владеющей компании, судебное решение). В таком случае сеть перестаёт функционировать. К централизованным сетям относится Direct Connect.

Децентрализованные сети

Децентрализованные файлообменные сети функционируют без индексационных серверов. Хотя объём передаваемой служебной информации в них больше, надёжность их гораздо выше. Пока не существует способа насильно остановить функционирование такой сети. К децентрализованным сетям относятся Gnutella, Overnet, Kad.

Частично децентрализованные (гибридные) сети

Сочетая скорость централизованных сетей и надёжность децентрализованных, используются гибридные схемы, в которых используются независимые индексационные сервера, постоянно синхронизирующие информацию между собой. Таким образом, при выходе из строя одного из них, сеть продолжает функционировать. К частично децентрализованным файлообменным сетям относятся eDonkey2000, OpenNap.

Принципы работы

Основные принципы работы файлообмена заключаются в следующем:

пользователь скачивает программу себе на компьютер;

разрешает доступ другим пользователям к некоторой части своих ресурсов (этот процесс называется «расшариванием» англ. sharing);

в каждой такой программе присутствует поиск, который ищет ресурсы, выложенные на компьютерах других пользователей для свободного скачивания.

Любой пользователь, используя поиск, может найти на компьютере любого другого пользователя те ресурсы, которые тот выложил в свободный доступ, и бесплатно скачать их. А так, как количество пользователей таких файлообменных программ исчисляется сотнями тысяч, а иногда даже миллионами, пользователь скорее всего найдет нужный ему файл.

Известные файлообменные сети

BitTorrent — технология распределённого распространения файлов, как правило, большого объёма. Отличается высокой скоростью и централизованностью. Некоторые BitTorrent-клиенты поддерживают DHT и могут работать без центрального сервера (т. н. трекера).

Direct Connect — представляет собой слабо связанные между собой выделенные сервера для поиска (хабы). Хабы Direct Connect очень удобны для организации файлового обмена в локальных сетях.

Advanced Direct Connect — эволюционное развитие сетей Direct Connect с устранением основных недостатков.

I2P — это анонимная, самоорганизующаяся распределенная сеть Network database, которая использует модифицированный DHT, но отличается тем, что хранит в себе хешированные адреса узлов сети, зашифрованные AES IP-адреса, а также публичные ключи шифрования.

ED2K (eDonkey2000) — сеть децентрализованного типа. Поиск выполняют специализированные серверы, связанные между собой. Клиенты самостоятельно обмениваются по протоколу MFTP.

OpenFT — открытое продолжение сети FastTrack. Поддерживается клиентами giFT (KCeasy), mlDonkey.

Skype — P2P-телефония.

WiPeer — сеть, действующая напрямую между компьютерами, минуя оборудование провайдера.

Netsukuku — сеть нового поколения, представляет собой ячеистую сеть передачи данных, заменяет 3-й уровень современной модели OSI другим протоколом маршрутизации. Протокол Netsukuku выстраивает структуру сети в виде фрактала. Сеть является распределённой, масштабируемой, анонимной и не контролируемой, отдельно от Интернета, без поддержки каких-либо служб и государственных каналов. Для расчёта всех необходимых путей связи узла со всеми остальными узлами протокол использует алгоритм Quantum Shortest Path Netsukuku (QSPN). Пропускная способность сети ограничивается возможностями нынешних сетевых карт. Поскольку в Netsukuku нет провайдеров, то нет и возможности отключить кого бы то ни было от сети. Так как адреса в Netsukuku раздаются случайным образом, то неизвестно даже, где находится и кому принадлежит нода, ответственная за хранение тех или иных байтов. В результате образуется полная свобода и полное равноправие.

Торрент

Битторрент - это протокол разработанный для обмена файлами. По сути он является peer-to-peer протоколом, когда каждый пользователь коннектится к другому напрямую, для приема или передачи частей информации. Файлы передаются частями, каждый torrent-клиент, получая (скачивая) эти части, в то же время отдаёт (закачивает) их другим клиентам, что снижает нагрузку и зависимость от каждого клиента-источника и обеспечивает избыточность данных.

Существует центральный сервер (называемый трекер) который является координатором действий всех существующих связей между пользователями. Трекер только управляет соединениями, он не содержит никакой информации о содержимом файлов, передаваемых между пользователями. Принципиальное отличие Битторрент - от других p2p систем - в том, что пользователь, который скачивает к себе информацию, является так же ее распространителем, т.е. передает информацию другим пользователям. В этом случае пропускная способность сети используется наилучшим способом, насколько это возможно. Битторрент разработан для наилучшей работы в сетях, где большое количество людей с совпадающими интересами. При этом чем большее количество людей участвует в раздаче конкретного торрента, тем большую конечную скорость может получить каждый из них.

Еще раз торрент - это ярлык, который однозначно идентифицирует содержимое передаваемой информации и адрес где можно получить эту информацию.

Сидер (seeder) - товарищ, который имеет на диске уже скачанный и полностью готовый к раздаче файл. Он же может быть инициатором раздачи файла.

Личер (leecher) - почти товарищ сидеру, тот кто в данный момент участвует в системе и качает куски файла одновременно раздавая скачанные куски другим личерам. Сидерам они само собой не нужны. Как тока личер скачал файл - он становится сидером.

Трекер (tracker) - Это сайт, а точнее движок на сайте, который а) складировать у себя торренты и раздает их всем клиентам, чтобы те качали друг у друга файлы б) рассылает им информацию об сидерах и личерах которые распространяют данный торрент, т.е является связующим звеном между ними.