
ЛЕКЦИЯ 1. ВВЕДЕНИЕ

ПРЕДМЕТ КУРСА ВЕБ-ТЕХНОЛОГИИ

- Предметом курса являются технологии глобальной сети *World Wide Web* (сокращенно *WWW*, *Web* или *Веб*).
- В частности, в рамках курса рассматриваются такие вопросы как:
 - Структура и принципы Веб (базовые понятия, архитектура, стандарты и протоколы)
 - Технологии Веб (языки разметки и программирования веб-страниц, инструменты разработки и управления веб-контента и приложений для Веб, средства интеграции веб-контента и приложений в Веб).

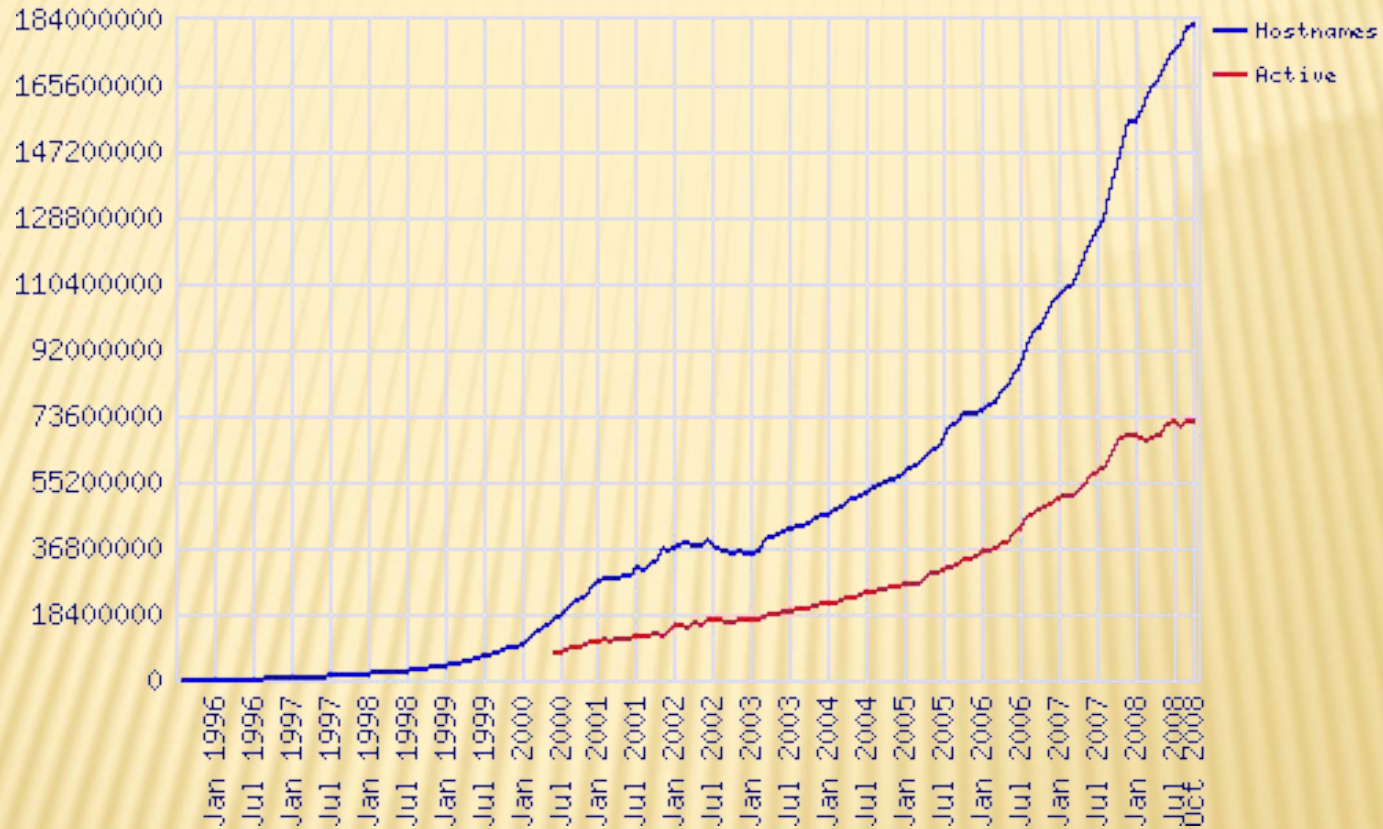
ЧТО ТАКОЕ ИНТЕРНЕТ?

- Это самая большая в мире сеть
- Это сеть, не имеющая единого центра управления, но работающая по единым правилам и предоставляющая своим пользователям единый набор услуг
- Это «сеть сетей», каждая из которых управляется независимым оператором – *поставщиком услуг Интернета* (ISP, Internet Service Provider)
- Это сеть, имеющая необъятное информационное наполнение и обеспечивающая простоту доступа к этой информации для любого пользователя.

ГРАНИЦЫ ИНТЕРНЕТА

- С точки зрения пользователей Интернет представляет собой набор информационных ресурсов, рассредоточенных по различным сетям, включая ISP-сети, корпоративные сети, сети и отдельные компьютеры домашних пользователей.

ДИНАМИКА РОСТА ИНТЕРНЕТ



Согласно данным с сайта [Netcraft](#)

УПРАВЛЕНИЕ В ИНТЕРНЕТ

- Централизация в Интернете проявляется только в единой технической политике, согласованном наборе технических стандартов, назначении имен и адресов компьютеров и сетей, входящих в Интернет.

ДОСТОИНСТВА И НЕДОСТАТКИ ДЕЦЕНТРАЛИЗАЦИИ В ИНТЕРНЕТЕ

□ Достоинства:

- Легкость наращивания Интернета путем заключения соглашения между двумя ISP

□ Недостатки:

- Сложность модернизации технологий и услуг Интернета, поскольку требуются согласованные усилия всех поставщиков услуг
- Невысокая надежность услуг Интернета

КЛАССИФИКАЦИЯ ПОСТАВЩИКОВ УСЛУГ

- *просто поставщик услуг Интернета* выполняет транспортную функцию для конечных пользователей – передачу их трафика в сети других поставщиков услуг Интернета
- *поставщик интернет-контента* имеет собственные информационно-справочные ресурсы, предоставляя их содержание в виде веб-сайтов
- *поставщик услуг хостинга* предоставляет свои помещения, каналы связи и серверы для размещения внешнего контента

КЛАССИФИКАЦИЯ ПОСТАВЩИКОВ УСЛУГ

- ▣ *поставщик услуг по доставке контента* занимается только доставкой контента в многочисленные точки доступа с целью повышения скорости доступа пользователей к информации
- ▣ *поставщик услуг по поддержке приложений* предоставляет клиентам доступ к крупным универсальным программным продуктам, например SAP R3.
- ▣ *поставщик биллинговых услуг* обеспечивает оплату счетов по Интернету

СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- Организация взаимодействия между устройствами сети является сложной задачей.
- Для решения таких задач используется *декомпозиция*, т.е. разбиение сложной задачи на несколько более простых задач-модулей.
- Одной из концепций, реализующих декомпозицию, является многоуровневый подход. Такой подход дает возможность проводить разработку, тестирование и модификацию отдельного уровня независимо от других уровней.

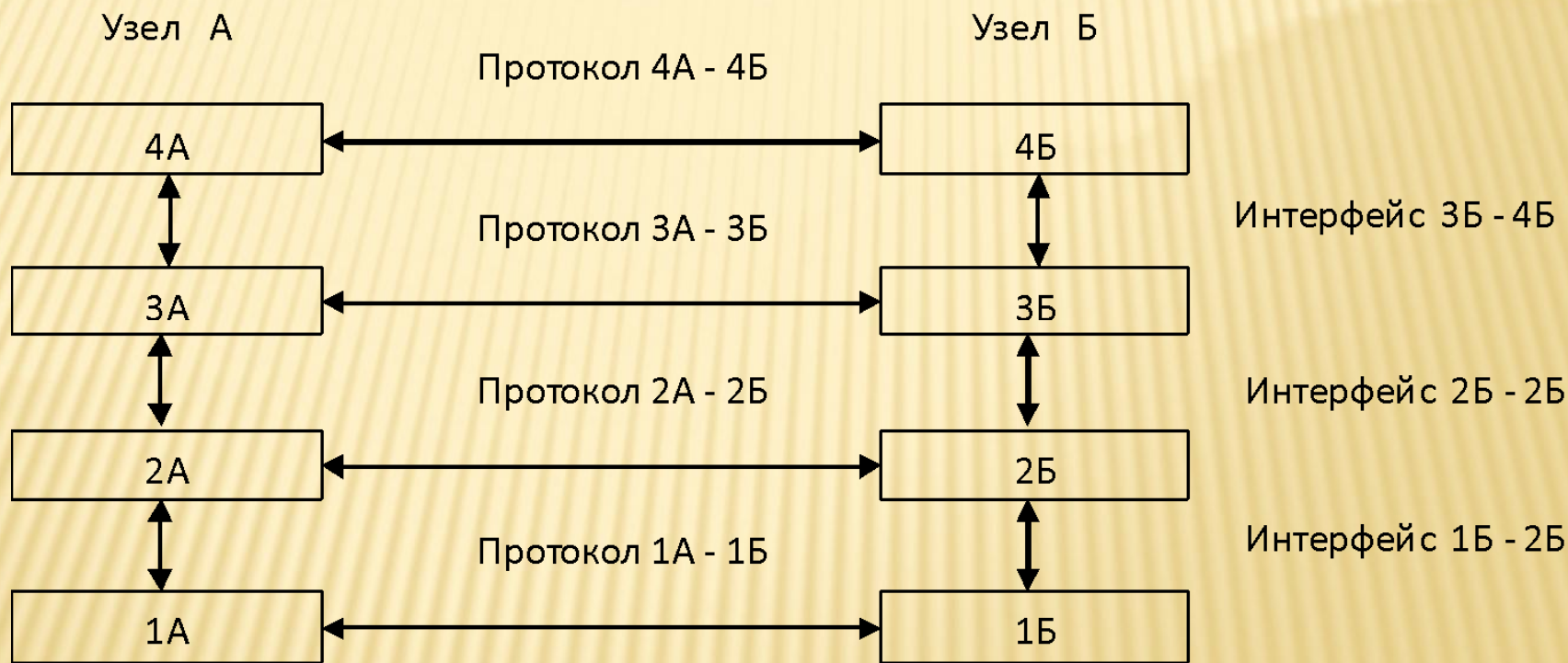
СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- *Иерархическая декомпозиция* позволяет, перемещаясь в направлении от более низких к более высоким уровням переходить к более простому представлению решаемой задачи.
- Специфика многоуровневого представления сетевого взаимодействия состоит в том, что в процессе обмена сообщениями участвуют как минимум две стороны, для которых необходимо обеспечить согласованную работу двух иерархий аппаратно-программных средств.

СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- Каждый из уровней должен поддерживать интерфейс с выше- и нижележащими уровнями собственной иерархии средств и интерфейс со средствами взаимодействия другой стороны на том же уровне иерархии. Данный тип интерфейса называется *протоколом*.

СЕТЕВОЕ ВЗАИМОДЕЙСТВИЕ ДВУХ УЗЛОВ



СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- Иерархически организованный набор протоколов, достаточный для организации взаимодействия узлов в сети, называется *стеком протоколов*.

СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- В начале 80-х годов международные организации по стандартизации **ISO** (International Organization for Standardization), **ITU** (International Telecommunications Union) и другие разработали стандартную модель *взаимодействия открытых систем OSI* (Open System Interconnection).
- Назначение данной модели состоит в обобщенном представлении средств сетевого взаимодействия. Ее также можно рассматривать в качестве универсального языка сетевых специалистов (справочной модели).

МОДЕЛЬ OSI

- Поскольку сеть – это соединение разнородного оборудования, актуальной является проблема совместимости, что в свою очередь, требует согласования всеми производителями общепринятых стандартов.
- *Открытой* является система, построенная в соответствии с открытыми *спецификациями*.
- *Спецификация* представляет собой формализованное описание аппаратных (программных) компонентов, способов их функционирования, взаимодействия с другими компонентами, условий эксплуатации, особых характеристик.

МОДЕЛЬ OSI

- Под *открытыми спецификациями* понимаются опубликованные, общедоступные спецификации, соответствующие стандартам и принятые в результате достижения согласия после всестороннего обсуждения всеми заинтересованными сторонами.
- Использование открытых спецификаций при разработке систем позволяет третьим сторонам разрабатывать для этих систем аппаратно-программные средства расширения и модификации, а также создавать программно-аппаратные комплексы из продуктов разных производителей.

МОДЕЛЬ OSI

- Если две сети построены с соблюдением принципов открытости, это дает следующие преимущества:
 - Возможность построения сети из аппаратных и программных средств различных производителей, придерживающихся стандарта;
 - Безболезненная замена отдельных компонентов сети другими, более совершенными;
 - Легкость сопряжения одной сети с другой.

МОДЕЛЬ OSI

- В рамках данной модели средства взаимодействия делятся на семь уровней: прикладной, представления, сеансовый, транспортный, сетевой, канальный и физический.
- В распоряжение программистов предоставляется прикладной программный интерфейс, позволяющий обращаться с запросами к самому верхнему уровню, а именно, - уровню приложений.

СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- Данная международная сеть строилась в полном соответствии с принципами открытых систем.
- В разработке стандартов этой сети принимали участие тысячи специалистов-пользователей сети из вузов, научных организаций и компаний.
- Результат работы по стандартизации воплощается в документах RFC (Request For Comments).

СТАНДАРТИЗАЦИЯ В ИНТЕРНЕТ

- Основным организационным подразделением, координирующим работу по стандартизации Интернет, является ISOC (Internet Society), объединяющее порядка 100 тысяч участников, которые занимаются различными аспектами развития данной сети.
- ISOC курирует работу IAB (Internet Architecture Board), включающую две группы:
 - IRTF (Internet Research Task Force) . Координирует долгосрочные исследовательские проекты, относящиеся к TCP/IP;
 - IETF (Internet Engineering Task Force). Инженерная группа, определяющая спецификации для последующих стандартов Интернет.

СТЕК TCP/IP

- Изначально разрабатывался для Интернета и имеет следующие важные достоинства:
 - Способность фрагментировать пакеты, что необходимо при передаче данных между сетями с различной длиной кадра.
 - Экономное использование широкополосных рассылок, что актуально для медленных каналов связи;
- Протокол TCP/IP соответствует модели OSI достаточно условно и содержит 4 уровня.

СТЕК TCP/IP

- Прикладной уровень стека соответствует трем верхним уровням модели OSI: прикладному, представления и сеансовому.
- В настоящее время разработано достаточно много протоколов и служб *прикладного* уровня:
 - *FTP* (File Transfer Protocol). Протокол передачи файлов.
 - *Telnet* . Протокол эмуляции терминала.
 - *SMTP* (Simple Mail Transfer Protocol). Простой протокол передачи электронной почты.
 - *HTTP* (Hypertext Transfer Protocol). Протокол передачи гипертекста.
 - и другие.

СТЕК TCP/IP

- *Транспортный уровень* может предоставлять вышележащему уровню два типа сервиса:
 - Протокол *TCP* (Transmission Control Protocol). Обеспечивает гарантированную доставку пакетов с помощью механизма подтверждения доставки.
 - Протокол *UDP* (User Datagram Protocol) – простейший дейтаграммный протокол, который используется в том случае, когда задача надежного обмена данными либо не ставится вообще, либо делегируется на более высокий уровень.

СТЕК TCP/IP

- Сетевой уровень реализуется в основном с помощью протокола *IP* (Internet Protocol) – межсетевого протокола. Это дейтаграммный протокол, работающий без установления соединений.
- К протоколам сетевого уровня относится также протокол межсетевых управляющих сообщений *ICMP* (Internet Control Message Protocol), предназначенный для передачи маршрутизатором источнику информации об ошибках при передаче пакета.

ТИПЫ АДРЕСОВ СТЕКА TCP/IP

- Для идентификации сетевых интерфейсов используются 3 типа адресов:
 - аппаратные адреса (или MAC-адреса);
 - сетевые адреса (IP-адреса);
 - символьные (доменные) имена.

СЕТЕВЫЕ IP-АДРЕСА

- Для создания глобальной системы адресации, не зависящей от способов адресации узлов в отдельных сетях используется пара идентификаторов, состоящая из **номера сети** и **номера узла**.
- При этом IP-адрес идентифицирует не отдельный компьютер или маршрутизатор, а одно сетевое соединение в составе сети в которую он входит. Т.е. конечный узел может входить в несколько IP-сетей.

ФОРМАТ IP-АДРЕСА

- IP-адрес состоит из двух частей: номера сети и номера узла в сети.
- Самой распространенной является запись IP-адреса в виде четырех чисел, разделенных точками, каждое из которых представляет значение байта в десятичной форме, например: 213.180.204.11.
- Запись адреса не предусматривает специального разграничительного знака между *номером сети* и *номером узла*.

ДОМЕННЫЕ ИМЕНА

- Несмотря на то, что аппаратное и программное обеспечение в рамках TCP/IP сетей для идентификации узлов использует IP-адреса, пользователи предпочитают **символьные имена**.
- Для установления соответствия между доменным именем и IP-адресом используется специальная **система доменных имен (DNS, Domain Name System)**, которая основана на создаваемых администраторами сети таблиц

ПЛОСКИЕ СИМВОЛЬНЫЕ ИМЕНА

- Первоначально в локальных сетях из небольшого числа компьютеров применялись плоские имена, состоящие из последовательности символов без разделения их на отдельные части, например *MYCOMP*. Для установления соответствия между символьными именами и числовыми адресами использовались широковежательные запросы. Однако для больших территориально распределенных сетей, работающих на основе протокола TCP/IP такой способ оказался неэффективным.

ИЕРАРХИЧЕСКИЕ СИМВОЛЬНЫЕ ИМЕНА

- В сетях TCP/IP используется доменная система имен, имеющая иерархическую (в виде дерева) структуру.
- Данная структура имен напоминает иерархию имен, используемую во многих файловых системах.
- Запись доменного имени начинается с самой младшей составляющей, затем после точки следует следующая по старшинству символьная часть имени и так далее. Последовательность заканчивается корневым именем, например:
company.yandex.ru

ИЕРАРХИЧЕСКИЕ СИМВОЛЬНЫЕ ИМЕНА

- Построенная таким образом система имен позволяет разделять административную ответственность по поддержке уникальности имен в пределах своего уровня иерархии между различными людьми или организациями.
- Совокупность имен, у которых несколько старших составных частей совпадают, образуют *домен имен*.

ИЕРАРХИЧЕСКИЕ СИМВОЛЬНЫЕ ИМЕНА

- Корневой домен управляется центральными органами Интернета: *IANA* и *Internic*.
- Домены верхнего уровня назначаются для каждой страны, а также для различных типов организаций. Имена этих доменов должны следовать международному стандарту *ISO 3166*. Для обозначения стран используются двухбуквенные аббревиатуры, например *ru* (Российская Федерация), *us* (США), *it* (Италия), *fr* (Франция).

ИЕРАРХИЧЕСКИЕ СИМВОЛЬНЫЕ ИМЕНА

- Для различных типов организаций используются трехбуквенные аббревиатуры:
 - net – сетевые организации
 - org – некоммерческие организации
 - com - коммерческие организации
 - edu – образовательные организации
 - gov – правительственные организации
- Администрирование каждого домена возлагается на отдельную организацию, которая делегирует администрирование поддоменов другим организациям.

РЕГИСТРАЦИЯ ДОМЕННЫХ ИМЕН

- Для получения доменного имени необходимо зарегистрироваться в соответствующей организации, которой InterNIC делегировал свои полномочия по распределению доменных имен.
- Регистратором доменных имен в зоне *ru* до 2005 г. являлся Российский научно-исследовательский институт развития общественных сетей (РосНИИРОС).
- В настоящее время регистрация доменов осуществляется одним из действующих регистраторов.

СХЕМА РАБОТЫ СИСТЕМЫ DNS

- В TCP/IP сетях соответствие между доменными именами и IP-адресами может устанавливаться как локальными средствами, так и централизованными службами.
- Первоначально соответствие задавалось с помощью создаваемого вручную на хосте файла *hosts.txt*, состоящего из строк, содержащих пару вида «*доменное имя – IP-адрес*». Однако с активным ростом Интернета такое решение оказалось немасштабируемым.

СХЕМА РАБОТЫ СИСТЕМЫ DNS

- Альтернативное решение – централизованная служба DNS, использующая распределенную базу отображений «*доменное имя – IP-адрес*».
- Сервер домена хранит только имена, которые заканчиваются на следующем ниже по дереву уровне. Это позволяет распределять более равномерно нагрузку по разрешению имен между всеми DNS-серверами.
- Каждый DNS-сервер помимо таблицы отображения имен содержит ссылки на DNS-серверы своих поддоменов.

СХЕМЫ РАЗРЕШЕНИЯ DNS-ИМЕН

□ Нерекурсивная процедура:

1. DNS-клиент обращается к корневому DNS-серверу с указанием полного доменного имени;
2. DNS-сервер отвечает клиенту, указывая адрес следующего DNS-сервера, обслуживающего домен верхнего уровня, заданный в следующей старшей части имени;
3. DNS-клиент делает запрос следующего DNS-сервера, который отсылает его к DNS-серверу нужного поддомена и т.д., пока не будет найден DNS-сервер, в котором хранится соответствие запрошенного имени IP-адресу. Сервер дает окончательный ответ клиенту.

СХЕМЫ РАЗРЕШЕНИЯ DNS-ИМЕН

- Рекурсивная процедура:
 1. DNS-клиент запрашивает локальный DNS-сервер, обслуживающий поддомен, которому принадлежит клиент;
 2. Далее
 - Если локальный DNS-сервер знает ответ, он возвращает его клиенту
 - Если локальный сервер не знает ответ, то он выполняет итеративные запросы к корневому серверу. После получения ответа сервер передает его клиенту.

СХЕМЫ РАЗРЕШЕНИЯ DNS-ИМЕН

- Таким образом, при рекурсивной процедуре клиент фактически перепоручает работу своему серверу.
- Для ускорения поиска IP-адресов DNS-серверы широко применяют кэширование (на время от часов до нескольких дней) проходящих через них ответов.

ПРОКСИ-СЕРВЕРЫ

Прокси-сервер (*proxy-server*) — служба в компьютерных сетях, позволяющая клиентам выполнять косвенные запросы к другим сетевым службам.

- Сначала *клиент* подключается к *прокси-серверу* и запрашивает какой-либо *ресурс*, расположенный на *другом сервере*.
- Затем *прокси-сервер* либо подключается к указанному *серверу* и получает *ресурс* у него, либо возвращает *ресурс* из собственного *кеша* (если имеется).
- В некоторых случаях *запрос клиента* или *ответ сервера* может быть изменен *прокси-сервером* в определённых целях. Также *прокси-сервер* позволяет защищать *клиентский компьютер* от некоторых *сетевых атак*.

ПРОКСИ-СЕРВЕРЫ

Чаще всего прокси-серверы применяются для следующих целей:

1. Обеспечение доступа с компьютеров локальной сети в Интернет;
2. Кеширование данных: если часто происходят обращения к одним и тем же внешним ресурсам, то можно держать их копию на прокси-сервере и выдавать по запросу, снижая тем самым нагрузку на канал во внешнюю сеть и ускоряя получение клиентом запрошенной информации.
3. Сжатие данных: прокси-сервер загружает информацию из Интернета и передаёт информацию конечному пользователю в сжатом виде.

ПРОКСИ-СЕРВЕРЫ

4. *Защита локальной* сети от внешнего доступа: например, можно настроить *прокси-сервер* так, что *локальные компьютеры* будут обращаться к *внешним ресурсам* только через него, а *внешние компьютеры* не смогут обращаться к локальным вообще (они «видят» только прокси-сервер).
5. *Ограничение доступа* из локальной сети к внешней: например, можно запретить доступ к определённым веб-сайтам, ограничить использование интернета каким-то локальным пользователям, устанавливать квоты на трафик или полосу пропускания, фильтровать рекламу и вирусы.

ПРОКСИ-СЕРВЕРЫ

6. *Анонимизация доступа* к различным ресурсам. Прокси-сервер может скрывать сведения об источнике запроса или пользователе. В таком случае *целевой сервер* видит лишь *информацию* о *прокси-сервере*, например, IP-адрес, но не имеет возможности определить *истинный источник запроса*. Существуют также искажающие прокси-серверы, которые передают целевому серверу ложную информацию об истинном пользователе.

WORLD WIDE WEB

- *Всемирная сеть Интернет* — глобальное информационное пространство, основанное на физической инфраструктуре Интернета и протоколе передачи данных *HTTP*.
- Зачастую, говоря об Интернете, подразумевают именно *Всемирную паутину*.
- Для обозначения *Всемирной паутины* также используют слово *веб* (от англ. *web*) или аббревиатуру «*WWW*»

СТРУКТУРА И ПРИНЦИПЫ WWW

- Сеть *WWW* образуют миллионы *веб-серверов*, расположенных по всему миру.
- *Веб-сервер* является программой, запускаемой на подключённом к сети компьютере и передающей данные по протоколу *HTTP*.

СТРУКТУРА И ПРИНЦИПЫ WWW

- Для идентификации ресурсов (зачастую файлов или их частей) в WWW используются идентификаторы ресурсов *URI* (*Uniform Resource Identifier*). Для определения местонахождения ресурсов в этой сети используются локаторы ресурсов *URL* (*Uniform Resource Locator*). Такие *URL*-локаторы представляют собой комбинацию *URI* и системы *DNS*.
- Доменное имя (или *IP*-адрес) входит в состав *URL* для обозначения компьютера (его сетевого интерфейса), на котором работает программа *веб-сервер*.

СТРУКТУРА И ПРИНЦИПЫ WWW

- На клиентском компьютере для просмотра информации, полученной от *веб-сервера*, применяется специальная программа — *веб-браузер*. Основная функция *веб-браузера* — отображение *гипертекстовых страниц*. Для создания гипертекстовых страниц в *WWW* изначально использовался язык *HTML*.
- Множество *веб-страниц* образуют *веб-сайт*.

КОНСОРЦИУМ W3C

- С 1994 года основную работу по развитию сети WWW взял на себя *Консорциум Всемирной паутины* (*World Wide Web Consortium, W3C*), основанный и до сих пор возглавляемый *Тимом Бернерсом-Ли*.
- Консорциум *W3C* — организация, разрабатывающая и внедряющая технологические стандарты для Интернета и WWW.
- Миссия *W3C*: «Полностью раскрыть потенциал Всемирной паутины путём создания протоколов и принципов, гарантирующих долгосрочное развитие Сети».
- Две другие важнейшие задачи *Консорциума* — обеспечить полную «интернационализацию Сети» и сделать ее доступной для людей с ограниченными возможностями.

СТАНДАРТИЗАЦИЯ В WWW

- W3C разрабатывает для WWW единые принципы и стандарты, называемые «*Рекомендациями*», которые затем внедряются разработчиками программ и оборудования.
- Благодаря рекомендациям достигается совместимость между программными продуктами и оборудованием различных компаний, что делает сеть **WWW** более совершенной, универсальной и удобной в использовании.
- Все *Рекомендации W3C* открыты, то есть не защищены патентами и могут внедряться любым человеком без каких-либо финансовых отчислений консорциуму.

ПРОТОКОЛЫ ИНТЕРНЕТ ПРИКЛАДНОГО УРОВНЯ

- ▣ *DNS* - распределённая система доменных имён, которая по запросу, содержащему доменное имя хоста сообщает *IP* адрес;
- ▣ *HTTP* - протокол передачи гипертекста в Интернет;
- ▣ *HTTPS* - расширение протокола *HTTP*, поддерживающее шифрование;
- ▣ *FTP* - протокол, предназначенный для передачи файлов в компьютерных сетях;
- ▣ *Telnet* - сетевой протокол для реализации текстового интерфейса по сети;
- ▣ *SSH* - протокол прикладного, позволяющий производить удалённое управление операционной системой и передачу файлов. В отличие от *Telnet* шифрует весь трафик;

ПРОТОКОЛЫ ИНТЕРНЕТ ПРИКЛАДНОГО УРОВНЯ

- ▣ **POP3** – протокол почтового клиента, который используется почтовым клиентом для получения сообщений электронной почты с сервера;
- ▣ **IMAP** - протокол доступа к электронной почте в Интернет;
- ▣ **SMTP** – протокол, который используется для отправки почты от пользователей к серверам и между серверами для дальнейшей пересылки к получателю;
- ▣ **LDAP** - протокол для доступа к службе каталогов **X.500**, является широко используемым стандартом доступа к службам каталогов;
- ▣ **XMPP (Jabber)** - основанный на **XML** расширяемый протокол для мгновенного обмена сообщениями в почти реальном времени;
- ▣ **SNMP** - базовый протокол управления сети *Internet*.

FTP

- ▣ *FTP* (*File Transfer Protocol* - [RFC 959](#)) — 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* позволяет обмениваться файлами и выполнять операции над ними через *TCP-сети*. Данный протокол работает независимо от операционных систем.

TELNET и SSH

- ▣ **TELNET** (*TELEcommunication NETwork* - [RFC 854](#)) — сетевой протокол для реализации текстового интерфейса по сети (в современной форме — при помощи транспорта *TCP*). Название «*telnet*» имеют также некоторые утилиты, реализующие клиентскую часть протокола.
- ▣ Протокол *telnet* работает в соответствии с принципами архитектуры «клиент-сервер» и обеспечивает эмуляцию алфавитно-цифрового терминала, ограничивая пользователя режимом командной строки.
- ▣ **SSH** (*Secure Shell* - [RFC 4251](#)) - протокол, позволяющий производить удалённое управление операционной системой и передачу файлов. Сходен по функциональности с протоколами *telnet* и *rlogin*, но, в отличие от них, шифрует весь трафик, включая и передаваемые пароли. *SSH-клиенты* и *SSH-серверы* имеются для большинства операционных систем.

ПОЧТОВЫЕ ПРОТОКОЛЫ

- ▣ **POP3** (*Post Office Protocol Version 3* - [RFC 1939](#)) — протокол, который используется почтовым клиентом для получения сообщений электронной почты с почтового сервера;
- ▣ **IMAP** (*Internet Message Access Protocol* - [RFC 3501](#)) — протокол доступа к электронной почте. Аналогичен **POP3**, однако предоставляет пользователю богатые возможности для работы с почтовыми ящиками, находящимися на центральном сервере. Электронными письмами можно манипулировать с компьютера пользователя (клиента) без необходимости постоянной пересылки с сервера и обратно файлов с полным содержанием писем.
- ▣ **SMTP** (*Simple Mail Transfer Protocol* — [RFC 2821](#)) — протокол, предназначенный для передачи электронной почты. 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**.