

Технологии сети Internet

Сетевые модели

Сетевая модель — это модель взаимодействия сетевых протоколов.

Все сетевые протоколы разрабатываются в соответствии с сетевой моделью и принадлежат определенному ее уровню.

Сетевая модель создает некий стандарт, гарантирующий совместимость протоколов и оборудования разных производителей.

Сетевые модели

- Основные сетевые модели
 - Модель OSI
 - Модель TCP/IP

Сетевые модели

	OSI	TCP/IP
•7.	Прикладной уровень Application layer	Прикладной уровень Process/Application layer
•6.	Уровень представления Presentation layer	
•5.	Сеансовый уровень Session layer	
•4.	Транспортный уровень Transport layer	Транспортный уровень Host-to-host layer
•3.	Сетевой уровень Network layer	Межсетевой уровень Internet layer
•2.	Канальный уровень Data Link layer	Уровень доступа к сети Network access layer
•1.	Физический уровень Physical layer	

Сетевые модели

Open System
Interconnection

Сетевая модель взаимодействия
открытых систем (эталонная
модель)

Модель OSI

Данные	Прикладной доступ к сетевым службам
Данные	Представления представление и кодирование данных
Данные	Сеансовый Управление сеансом связи
Блоки	Транспортный безопасное и надёжное соединение точка-точка
Пакеты	Сетевой Определение пути и IP (логическая адресация)
Кадры	Канальный MAC и LLC (Физическая адресация)
Биты	Физический кабель, сигналы, бинарная передача данных

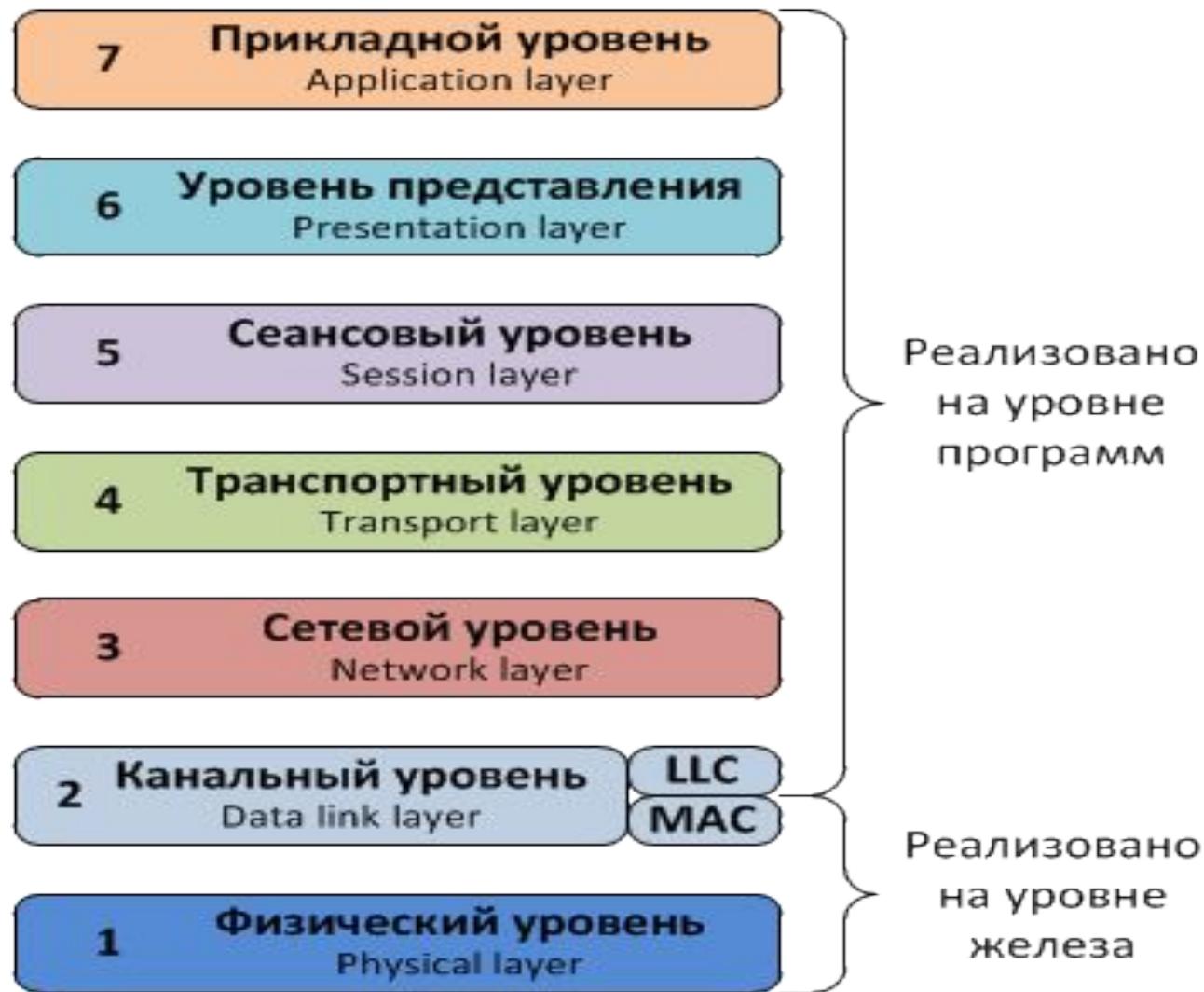
Сетевые модели

OSI



Сетевые модели

OSI



Сетевые модели

OSI

Адресация →

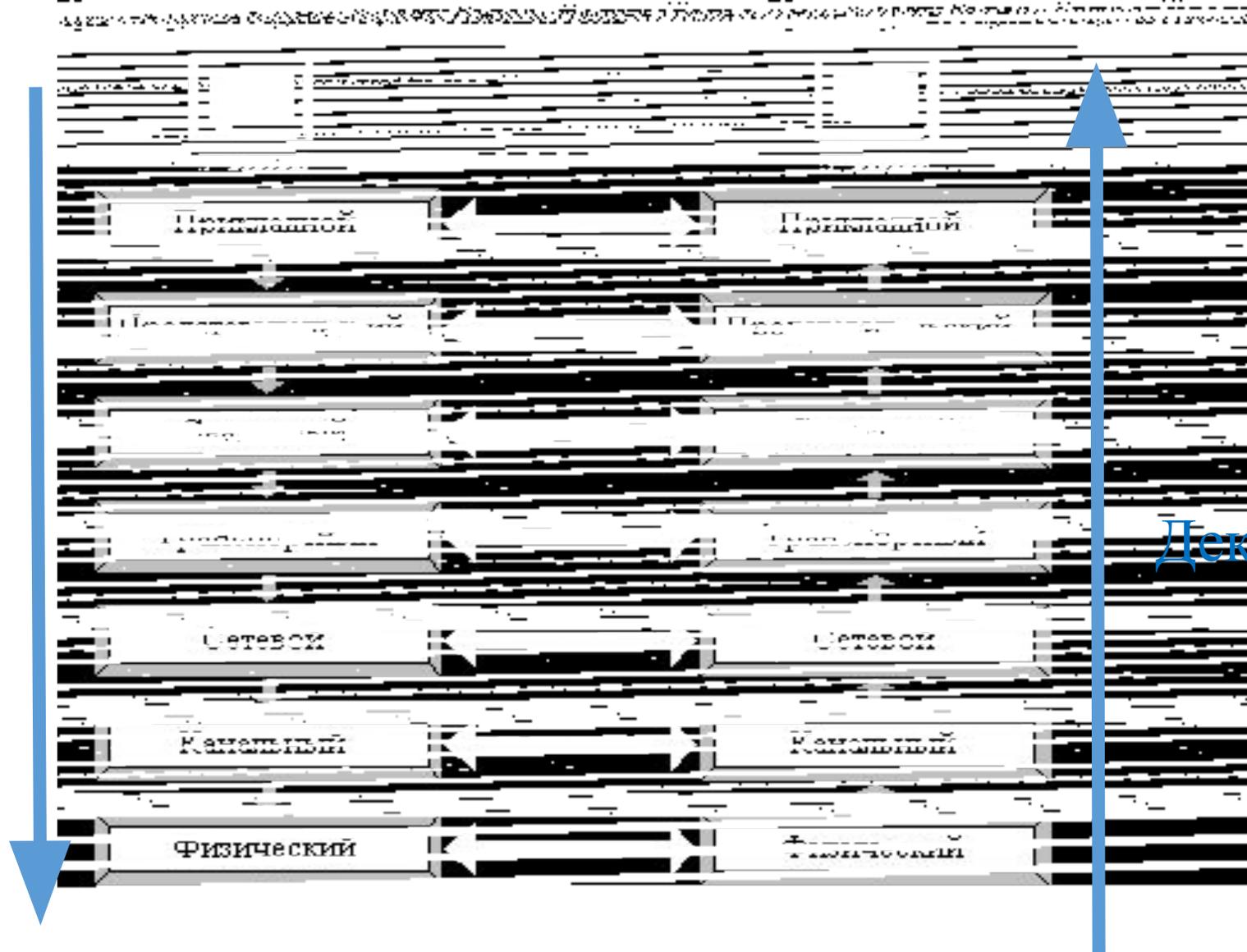
Порт — это числовой идентификатор программы или процесса, которые обслуживают сетевые соединения на заданном сетевом адресе (IP-адресе).

7. Прикладной уровень представляет набор интерфейсов, позволяющий получить доступ к сетевым службам
6. Уровень представления преобразует данные в общий формат для передачи по сети
5. Сеансовый уровень поддерживает взаимодействие (сеанс) между удаленными процессами
4. Транспортный уровень управляет передачей данных по сети, обеспечивает подтверждение передачи
3. Сетевой уровень маршрутизация, управление потоками данных, адресация сообщений для доставки; преобразование логические сетевые адреса и имена в соответствующие им физические
2. Канальный уровень 2.1. Контроль логической связи (LLC): формирование кадров 2.2. Контроль доступа к среде (MAC): управление доступом к среде
1. Физический уровень обеспечивает битовые протоколы передачи информации

Сетевые модели

OSI

Инкапсуляция

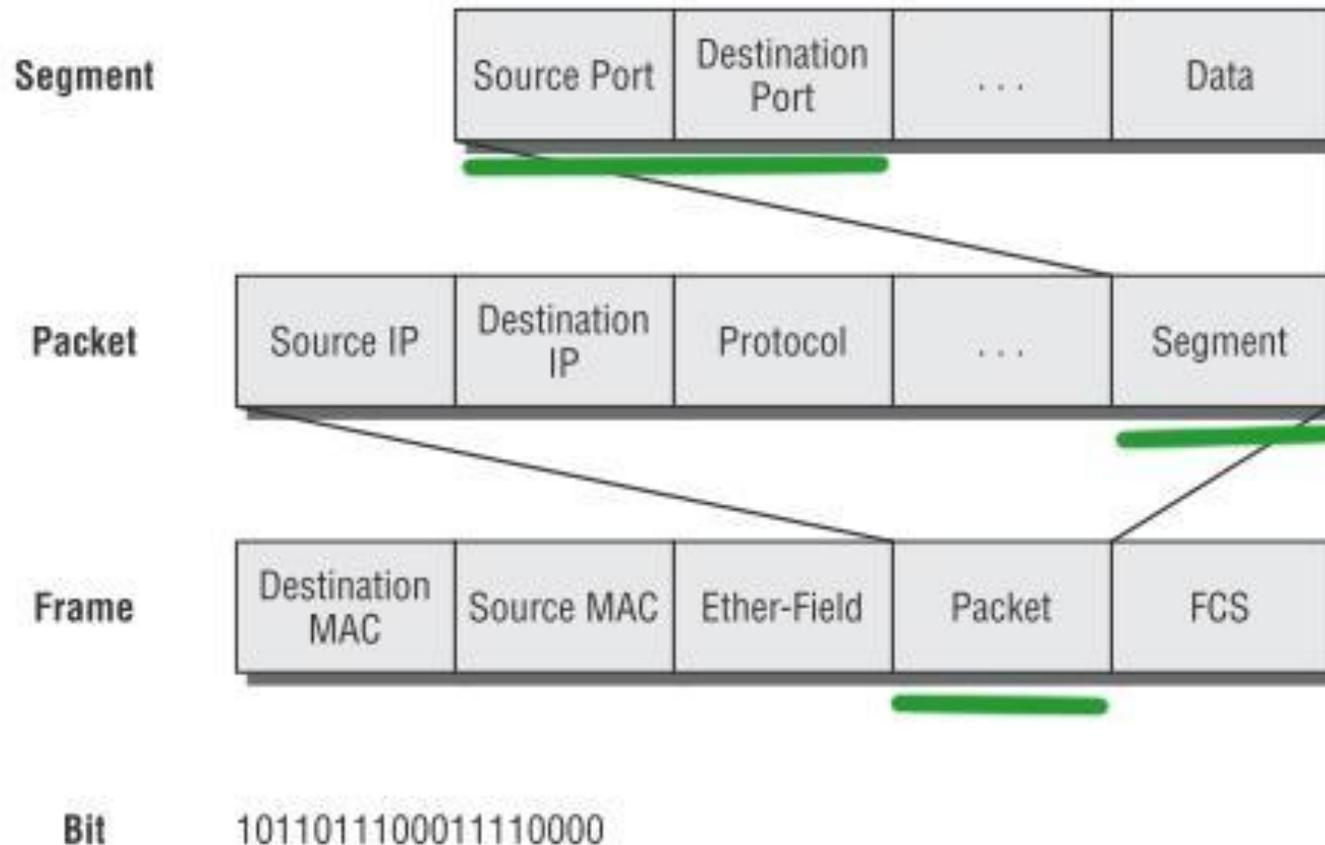


Декапсуляция

Сетевые модели

Protocol Data Unit (PDU) — протокольная единица обмена, модуль данных протокола (в OSI представляет собой объект данных, которыми обмениваются "машины протокола" (сущности уровня) в пределах данного уровня; содержит как управляющую информацию (PCI), так и пользовательские данные).

Сетевые модели



7. Уровень приложений — данные (data)

4. Транспортный уровень — сегмент (segment)

3. Сетевой уровень — пакет (packet)

2. Канальный уровень — кадр (frame)

1. Физический уровень — биты (bits)

IP-адресация

IP-адрес – это уникальный идентификационный номер, который имеет каждый компьютер в сети.

Протоколы:

```
graph TD; A[Протоколы:] --> B[IPv4]; A --> C[IPv6];
```

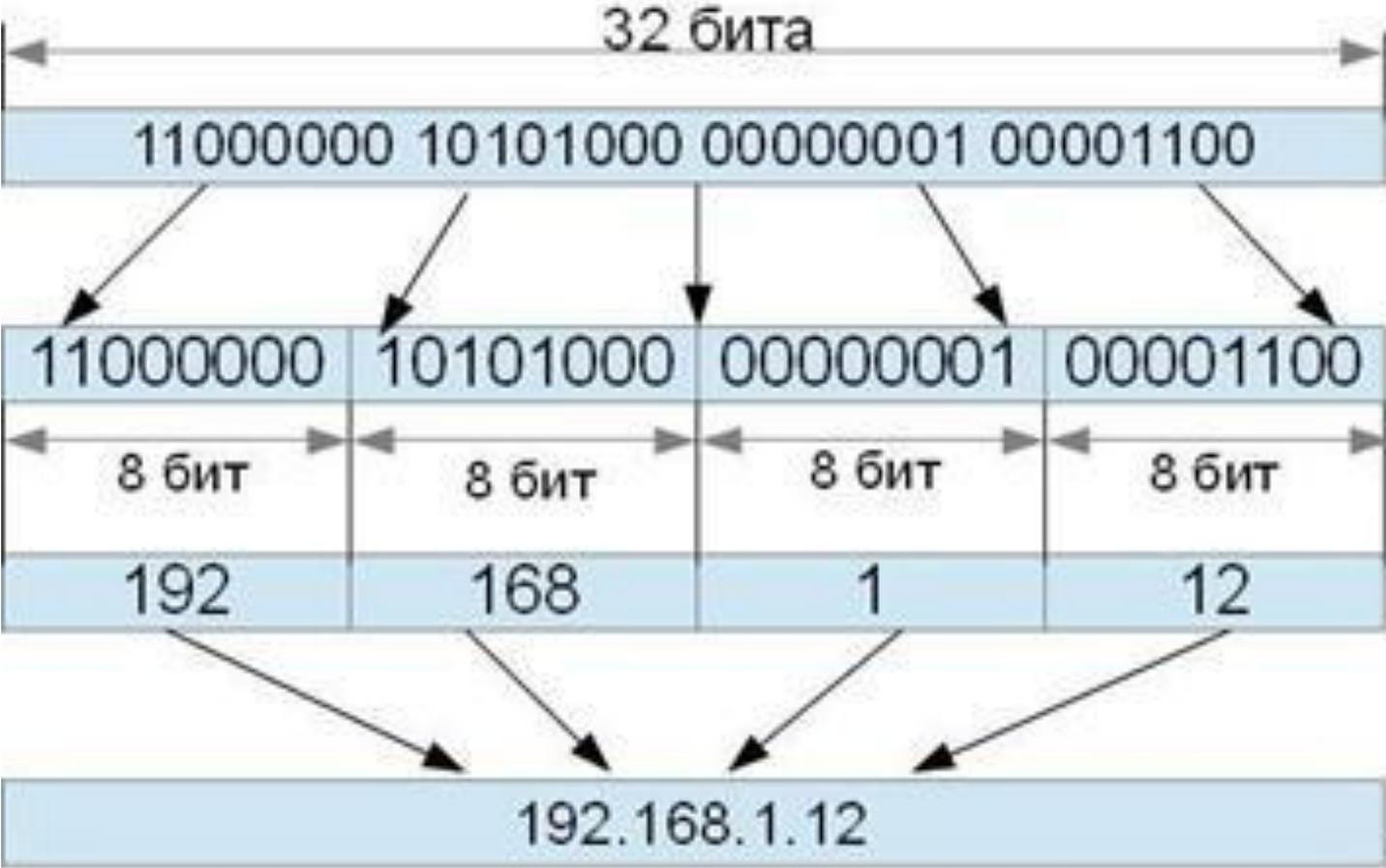
IPv4 32-битный

IPv6 128-битный

10.361.28.36

1534:abcd:0:1:0:3:7:6

IP-адресация



IPv4-адрес в 32-разрядном виде

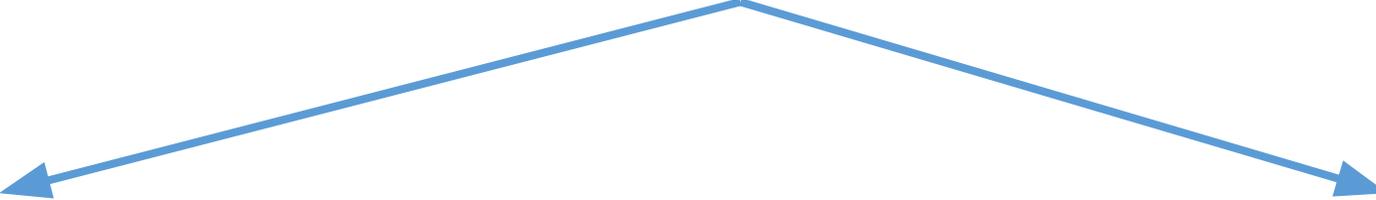
IPv4-адрес в двоичном виде, разбитый на октеты

Октеты в десятичной записи

IPv4-адрес в точно-десятичной записи

IP-адресация

IP-адрес состоит из двух частей:



адрес подсети / номер сети

адрес хоста / номер узла

Маска подсети позволяет компьютеру отделить эти две составляющие.

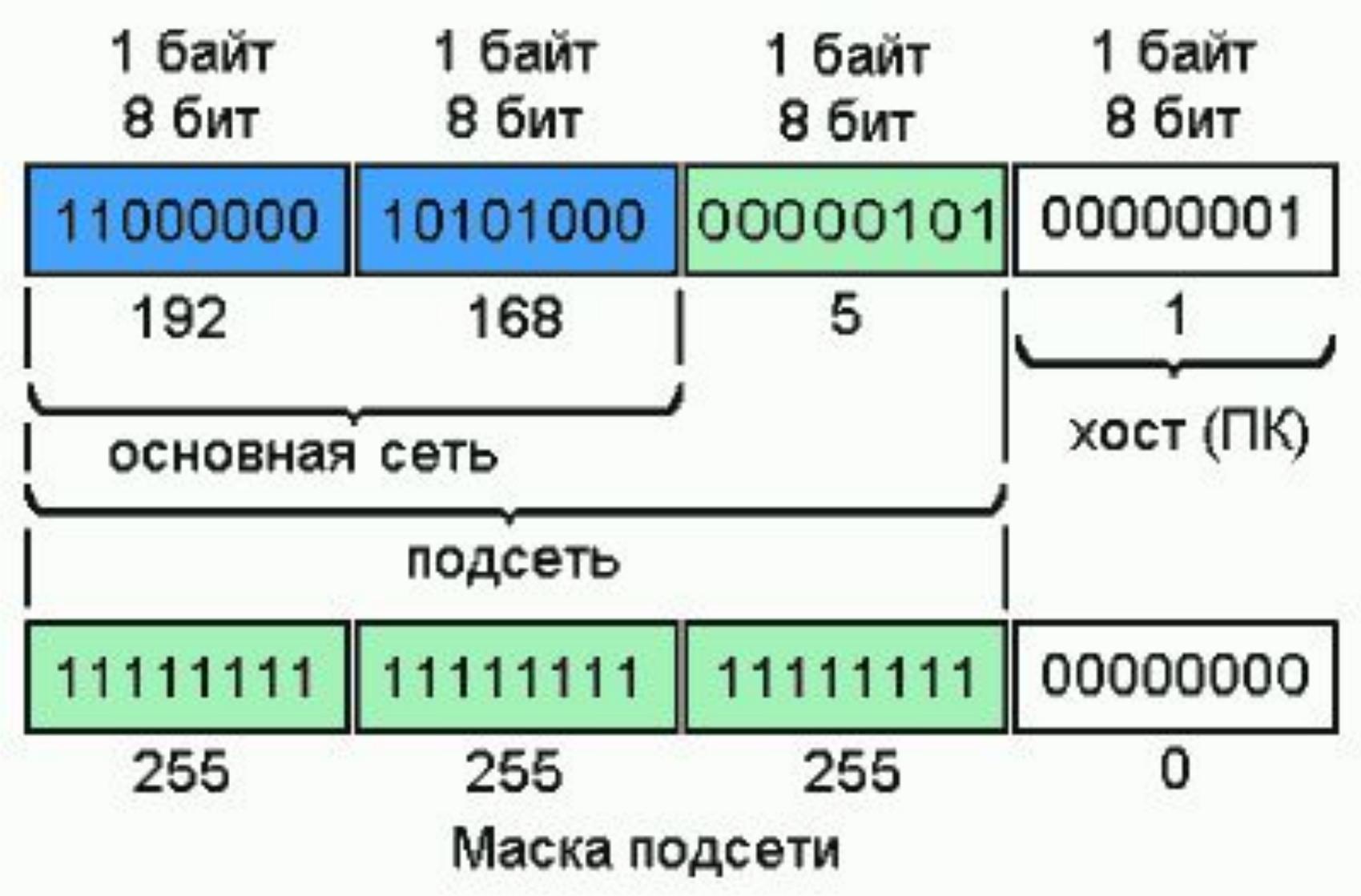
255.255.255.0

в десятичной записи

11111111.11111111.11111111.00000000

в двоичном эквиваленте

IP-адресация



Сервисы Internet

- **Компоненты, необходимые для организации работы пользователей сети с распределенными информационными ресурсами:**
 - Программа, установленная на компьютере пользователя, которая может осуществлять сетевой запрос.
 - Программа, установленная на компьютере, где расположен информационный объект, которая может осуществлять по сетевому запросу поиск и пересылку объекта
 - Сетевой протокол взаимодействия между этими программами

Сервисы Internet



Сервисы Internet

- Нестандартные сервисы:
 - *SE (Search Engine).*
 - *IRC (Internet Relay Chat).*
 - *RTVC (Real Time Video Connection).*

Протоколы сервисов Internet

Для стандартных сервисов также стандартизируется и интерфейс взаимодействия с протоколами транспортного уровня TCP и UDP.

TCP (Transmission Control Protocol) — один из основных протоколов передачи данных интернета, предназначенный для управления передачей данных. В стеке протоколов TCP/IP выполняет функции транспортного уровня модели OSI.

UDP (User Datagram Protocol — протокол пользовательских датаграмм) — один из ключевых элементов TCP/IP, набора сетевых протоколов для Интернета. С UDP компьютерные приложения могут посылать сообщения (в данном случае называемые датаграммами) другим хостам по IP-сети без необходимости предварительного сообщения для установки специальных каналов передачи или путей данных.

Для нестандартных сервисов, представляющих собой оригинальные разработки различных компаний, номера TCP- и UDP-портов сервера и клиентской программы не регламентируются.

Протоколы сервисов Internet

Сервисы и стандартные номера TCP- и UDP-портов

Компонент сервиса	Номер порта	Протоколы транспортного уровня
DNS		
DNS-сервер	53	TCP, UDP
Telnet		
Telnet-сервер	23	TCP
FTP		
FTP-сервер	20,21	TCP
Usenet		
NNTP-сервер	119	TCP
E-mail		
SMTP-сервер	25	TCP
POP3-сервер	110	TCP
IMAP-сервер	143	TCP
WWW		
HTTP-сервер	80	TCP

Сервис DNS — система доменных имен



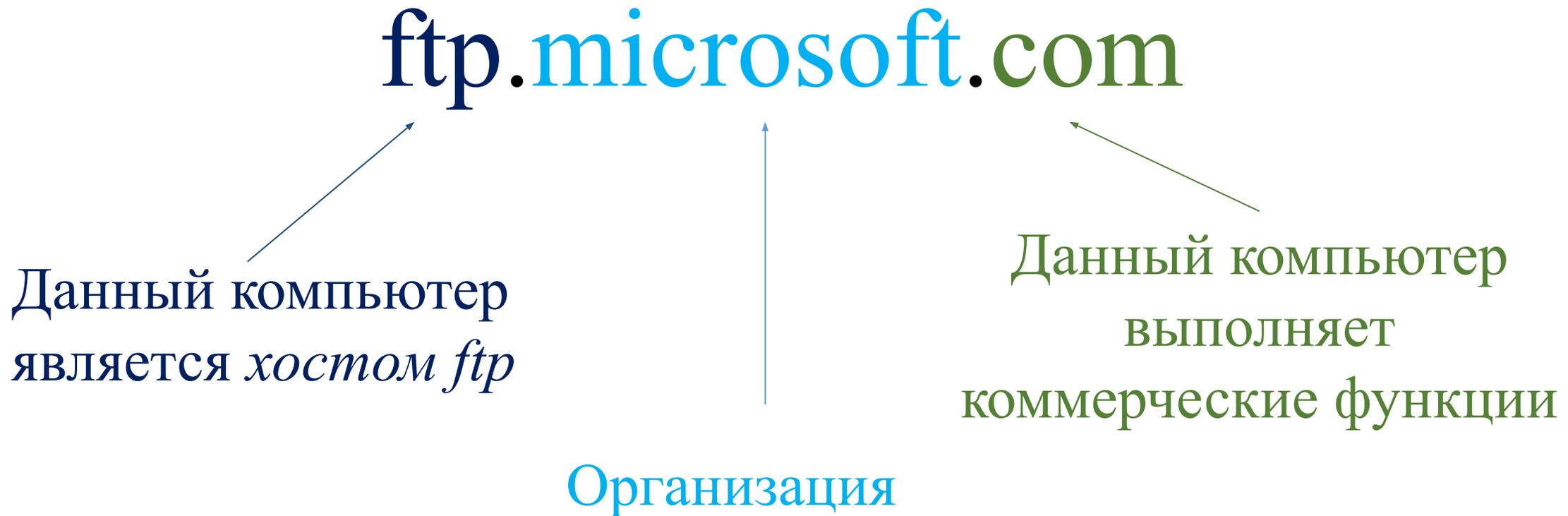
Термин «домен» определяется как сфера деятельности, отношений или выполнения каких-либо совместных функций.



Сервис DNS — система доменных имен

ftp.microsoft.com

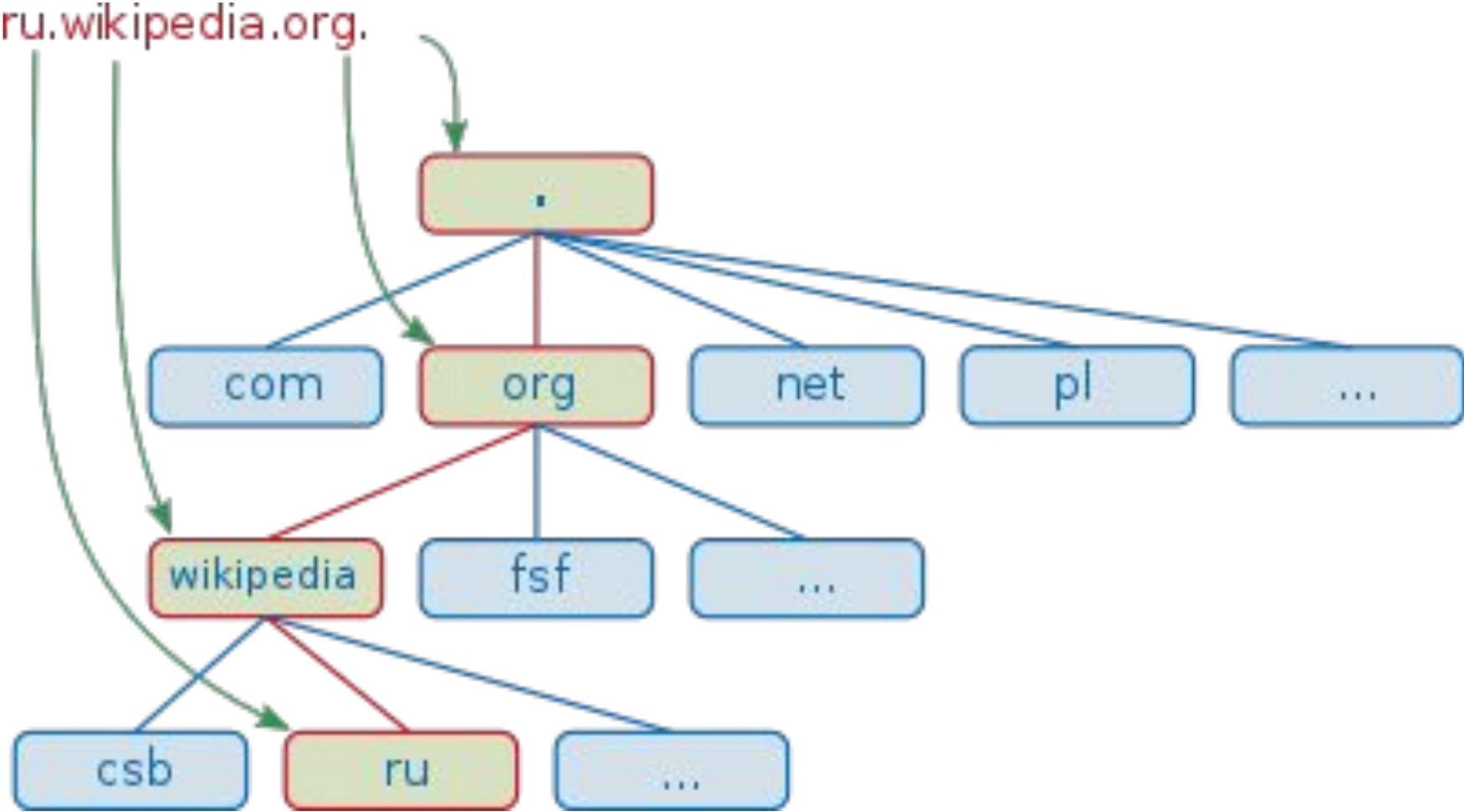
Данный компьютер
является *хостом ftp*



Организация

Данный компьютер
выполняет
коммерческие функции

Сервис DNS – система доменных имен



Сервис DNS — система доменных имен

- DNS выполняет функции:
 - Организация иерархического пространства имен.
 - Обеспечение разрешения (т. е. поиска соответствия) доменных имен и IP-адресов.

Сервис DNS — система доменных имен

Компоненты

DNS:

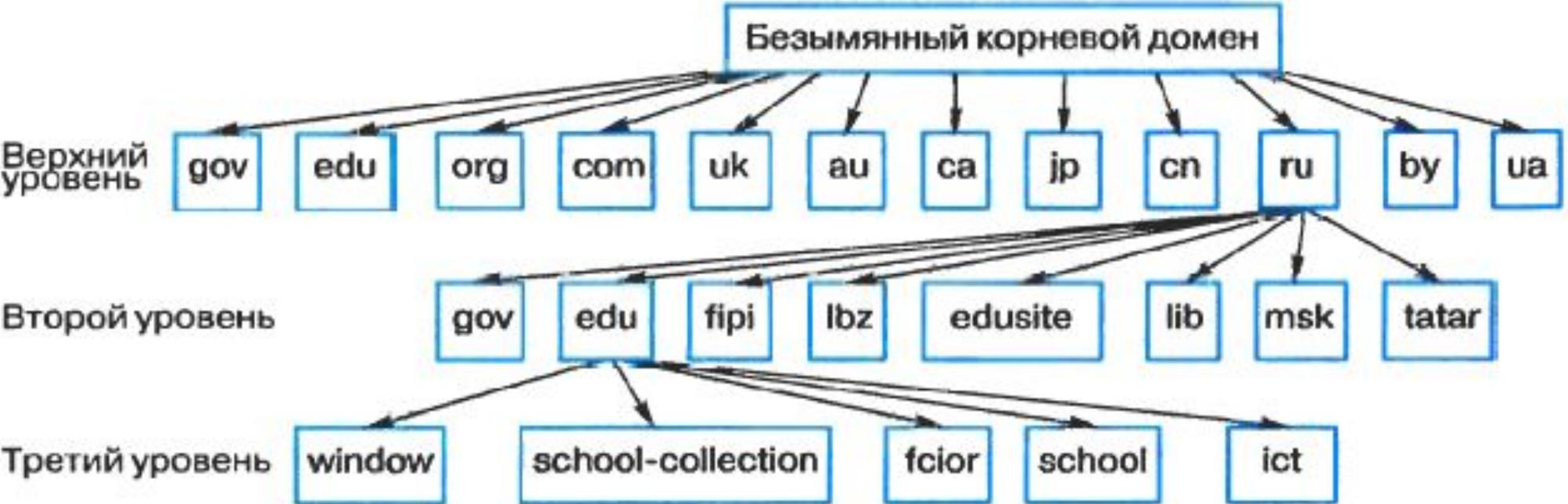
Информационный ресурс

DNS-клиент (resolver)

*Сервер имен (name server),
или DNS-сервер*

Протокол DNS

Сервис DNS – система доменных имен



Иерархическая структура доменных имен

Сервис DNS — система доменных имен

www.test.bixtone.ru



Префикс

Домен 3 уровня (поддомен)

Домен 2 уровня, основной для регистрации

Домен верхнего (1) уровня, где зарегистрирован основной домен

Сервис DNS — система доменных имен

Домены первого уровня

Домены общего назначения

Национальные домены

Обратный домен

Сервис DNS — система доменных имен

Домены общего назначения

Название домена	Комментарий
«Старые» домены, созданные в 1984 году	
«com»	COMmercial, коммерческие организации
«gov»	GOVernment, правительственные учреждения США
«int»	INTeinational Orsanizations, международные организации
«mil»	MILitary, военные организации США
«edu»	EDUcational. образовательные проекты и учреждения
«org»	ORGanisations, некоммерческие организации или организации, не попадающие в другие категории
«net»	NETwork, сети общего назначения
«Новые» домены, созданные в 2001 году	
«info»	INFOrmation, домен свободного использования для предоставления информации в Интернет
«biz»	Business Organizations, различные организации
«name»	Домен предназначен для использования частными лицами
«museum»	Музеи

Сервис DNS — система доменных имен

Обратный домен

.arpa

предназначен для поиска доменного имени по IP-адресу (обратного разрешения).

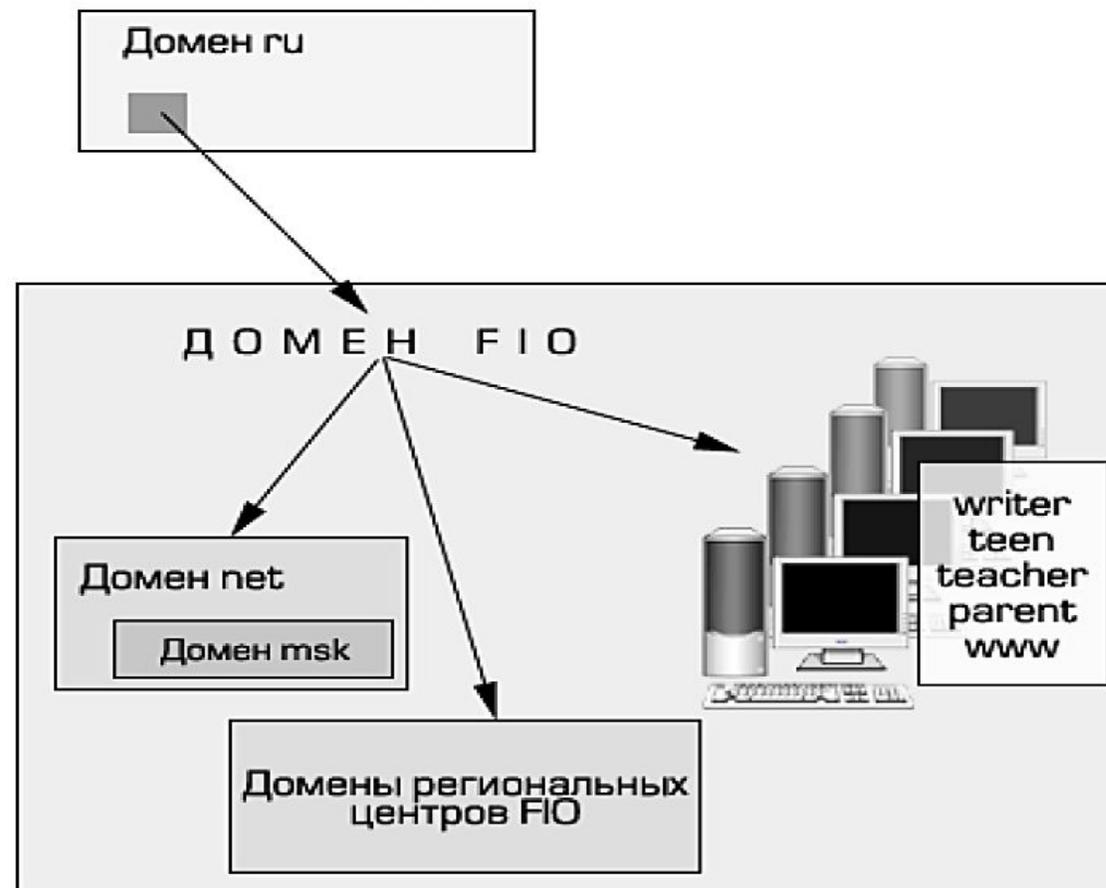
Часто этот домен называют обратным доменом (reverse domain).

Сервис DNS — система доменных имен

Домены общего пользования («домены типа GENERIC»):

- «ac.ru» — домен Академии Наук России;
- «com.ru» — коммерческие организации;
- «int.ru» — домен для общего использования;
- «org.ru» — некоммерческие организации;
- «edu.ru» — образовательные проекты и организации;
- «pp.ru» — домен для использования частными лицами;
- «mil.ru» — военные организации и учреждения;

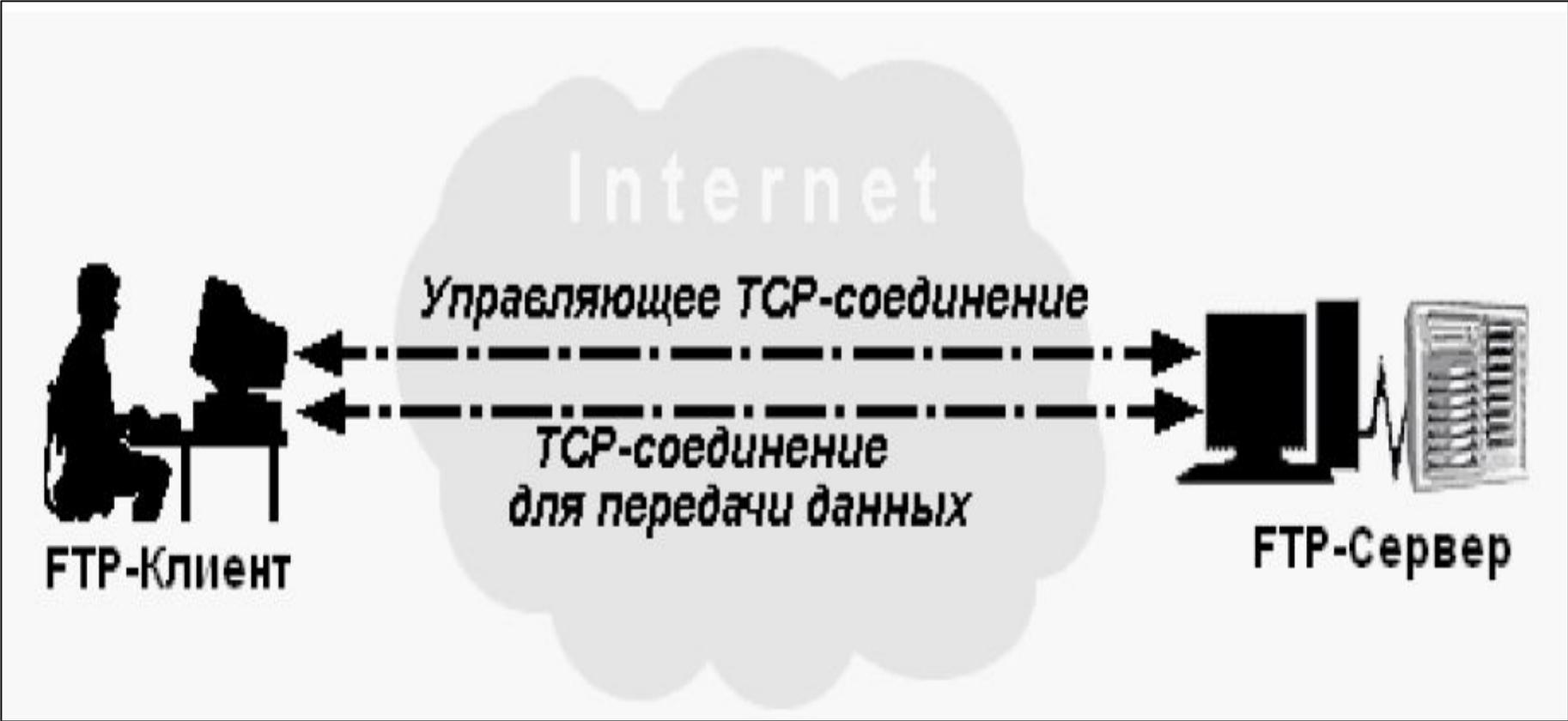
Сервис DNS – система доменных имен



Сервис FTP-система файловых архивов

- В предоставлении сервиса FTP принимают участие две программные компоненты:
 - FTP-сервер
 - FTP-клиент

Назначение сервиса FTP



Организация работы сервиса FTP

1. Запускаем браузер

2. Устанавливаем соединение с FTP-сервером. В строке адреса вводим:
ftp:// имя-FTP-сервера/

3. Если соединение установлено успешно, то можно скачивать файлы на свой компьютер

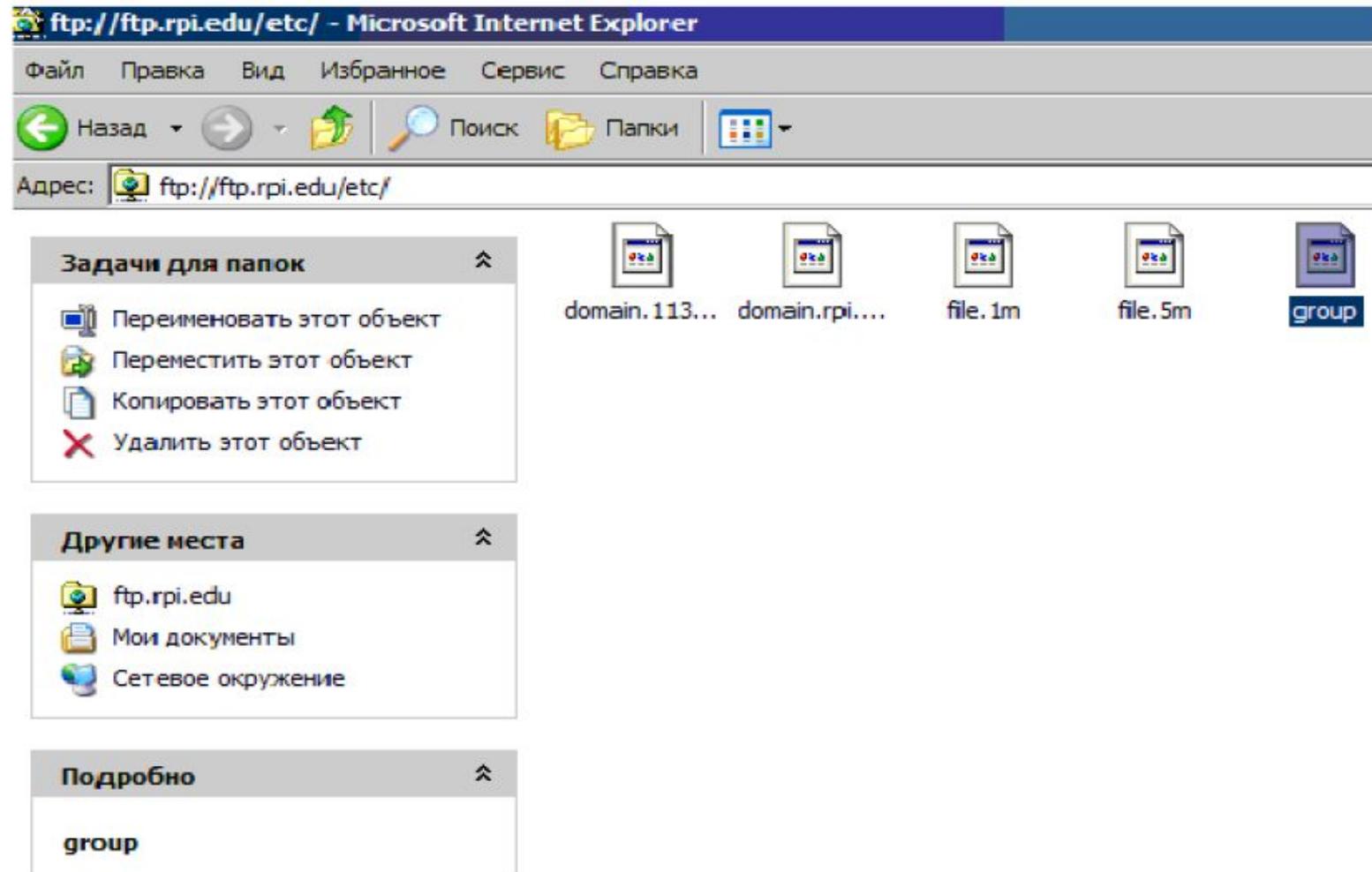
Организация работы сервиса FTP

«анонимные» FTP-серверы

Пример обращения к ресурсу на FTP-сервере

Имя сервера	Имя каталога в сервере
<u>ftp://ftp.rpi.edu/</u>	/etc/group/

Организация работы сервиса FTP



**Сервис WWW — гипертекстовая система
интеграции сетевых ресурсов в единое
информационное пространство**



Сервис WWW — гипертекстовая система интеграции сетевых ресурсов в единое информационное пространство

- WWW
 - *Web-документ (Web-страница)*
 - *Web-сайт*
 - *Web-клиент (браузер)*
 - *Протокол HTTP*
 - *Web-сервер*

Язык HTML

```
<!DOCTYPE html>
<html>
<!-- created 2010-01-01 -->
<head>
  <title>sample</title>
</head>
<body>
  <p>Voluptatem accusantium
  totam rem aperiam.</p>
</body>
</html>
```

HTML

Язык HTML

теги описания структуры документа и информации о нем

теги, использующиеся для логической структуризации текста документа

теги форматирования текста

теги организации гиперссылок

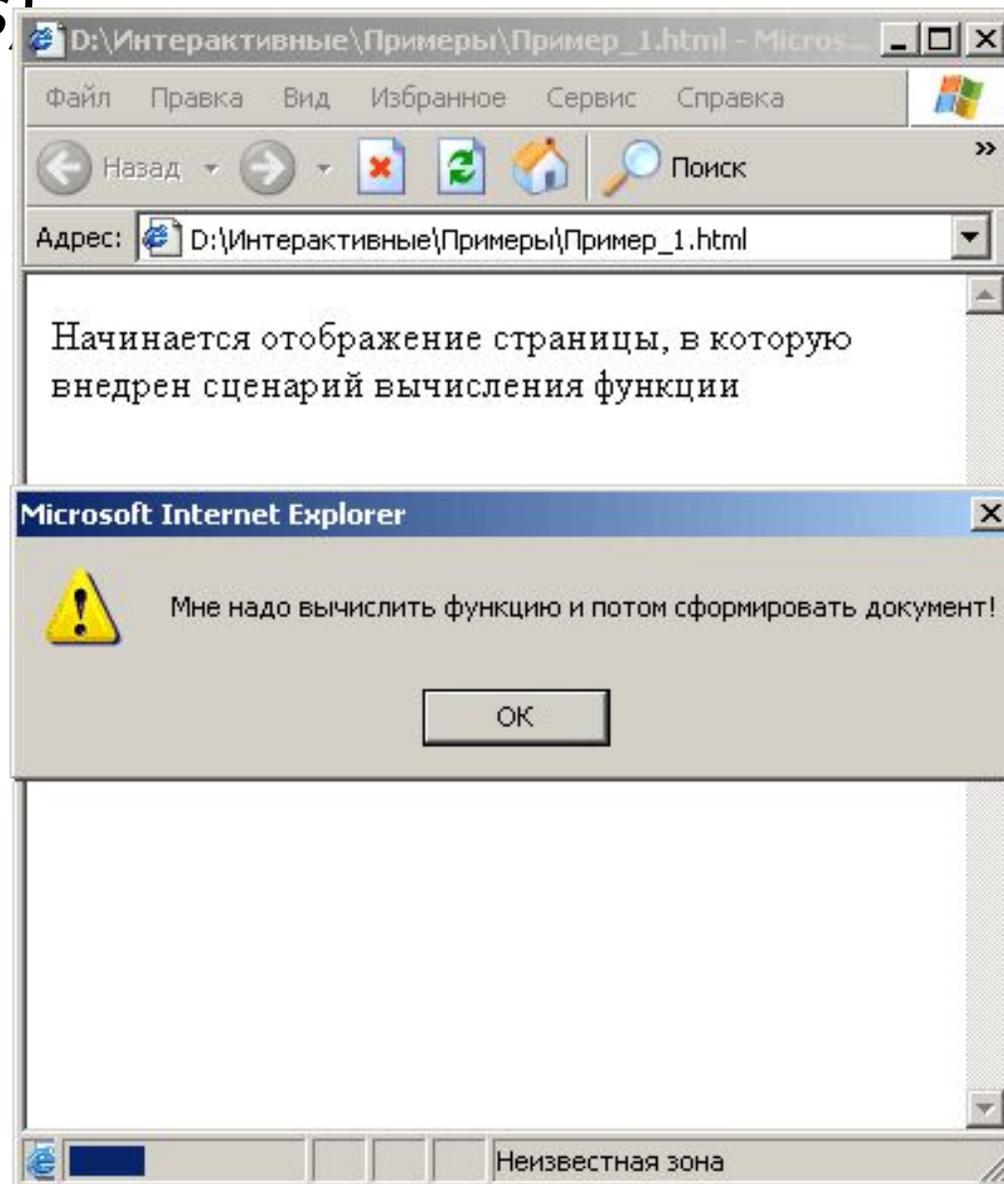
теги, устанавливающие связи HTML-файла с внешними объектами

теги создания форм

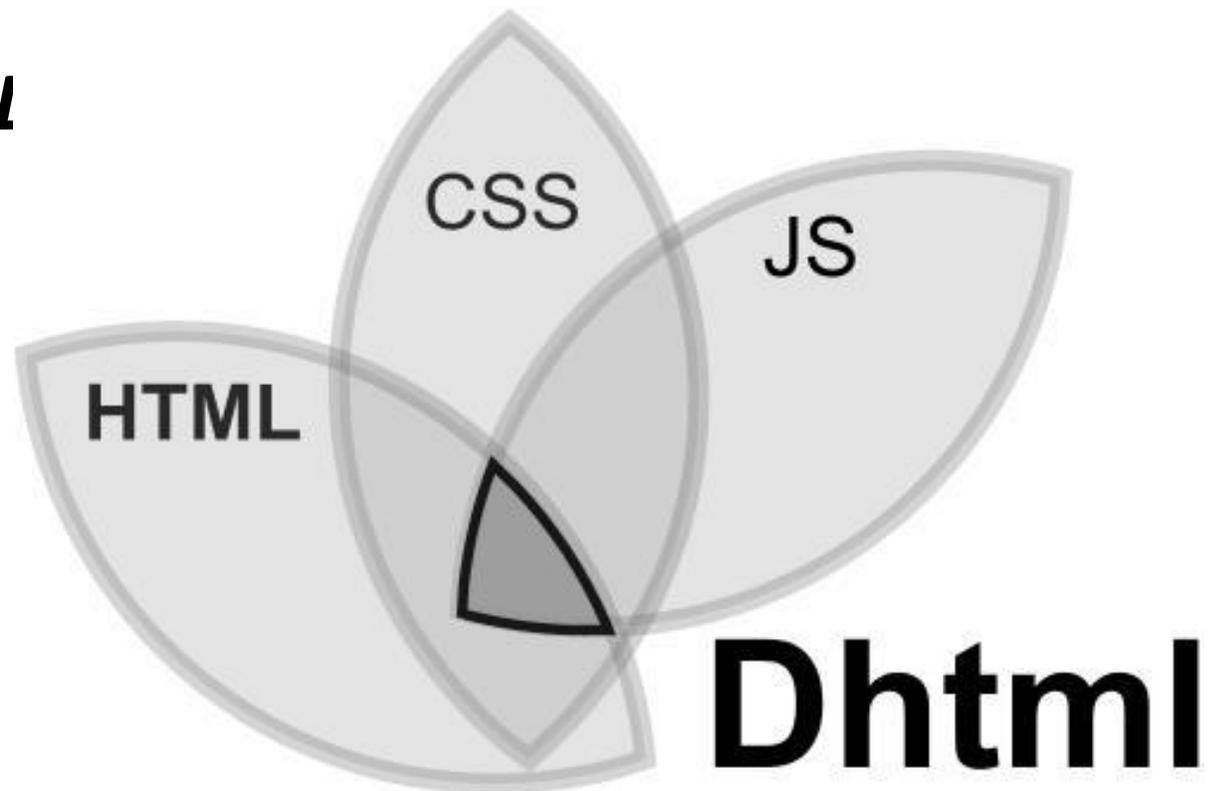
*Таблицы стилей (каскадные таблицы стилей,
Cascading Style Sheets, CSS)*



Сценарии (scripts)



Язык Dynamic HTML



Java-апплет — это небольшая программа, написанная на языке программирования Java.

HTML



Язык XML позволяет описывать данные произвольного типа и используется для представления специализированной информации.

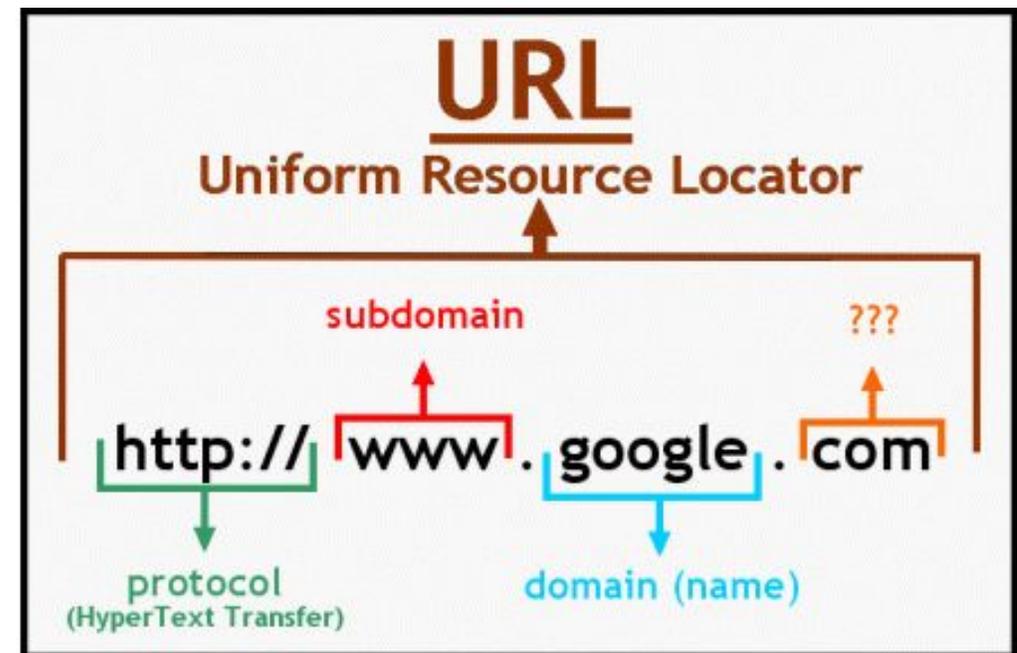
Концепция унифицированных указателей ресурса URL (Uniform Resource Locator)

URL представляет собой набор информации, необходимый для того, чтобы определить:

- узел сети, на котором расположен информационный объект;
- расположение информационного объекта на узле;
- метод получения доступа к объекту.

URI

Uniform Resource Identifier

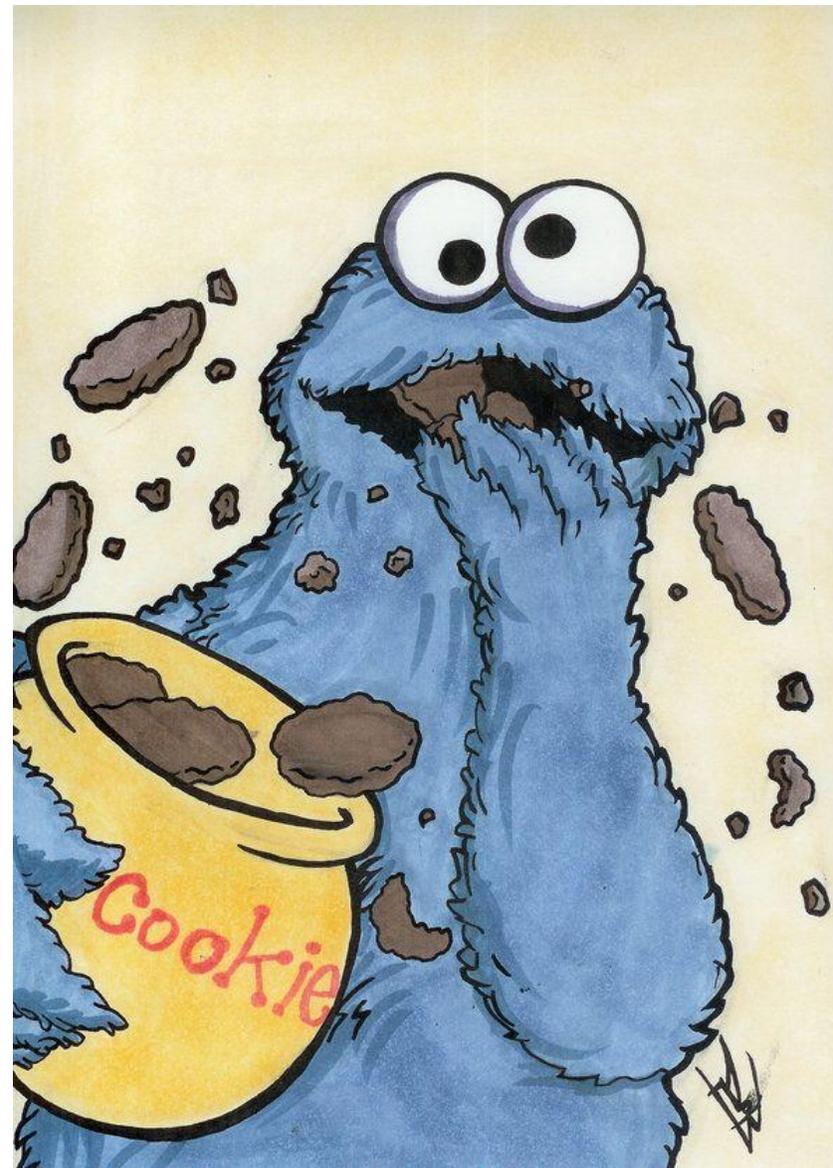


Концепция унифицированных указателей ресурса URL (Uniform Resource Locator)

Схема	Определяемый механизм доступа
mailto	Электронная почта
news	Телеконференции
ftp	Протокол FTP
telnet	Протокол Telnet
http	Протокол HTTP
https	Протокол HTTPS
file	Файл, расположенный на локальном компьютере

Маркер cookie

Cookie - это небольшая порция текстовой информации, которую сервер передает браузеру. Браузер будет хранить эту информацию и передавать ее серверу с каждым запросом как часть HTTP заголовка.



Маркер cookie

Посмотреть cookie через меню браузера Google Chrome

- Открыть браузер.
- Нажать в правом верхнем углу браузера кнопку меню (три вертикальные точки).
- Выбрать **Настройки (Setting)**.
- Выбрать **Дополнительные (Advanced)**.
- Нажать на кнопку **Настройки контента (Content Settings)**.
- Нажать на кнопку **Файлы cookie (Cookies)**. Файлы находятся в нижней части страницы. Любой элемент, помеченный как «Файлов cookie: <число>», является файлом cookie.
- Ввести в строку поиска адрес сайта, для которого нужно посмотреть cookie.
- Для просмотра параметров файла cookie кликните по элементу.

Сервис SE (Search Engine) — поисковые системы

Google™

Яндекс

YANHOO!®

Rambler®

bing™

gogo beta

Webalta

Y.UA

МЕТА®



~~NIGMA.RU~~
интеллектуальная поисковая система

АПОРТ
ищет что надо

Поисковые системы (компоненты)

Паук (spider)

*Путешествующий паук
(crawler)*

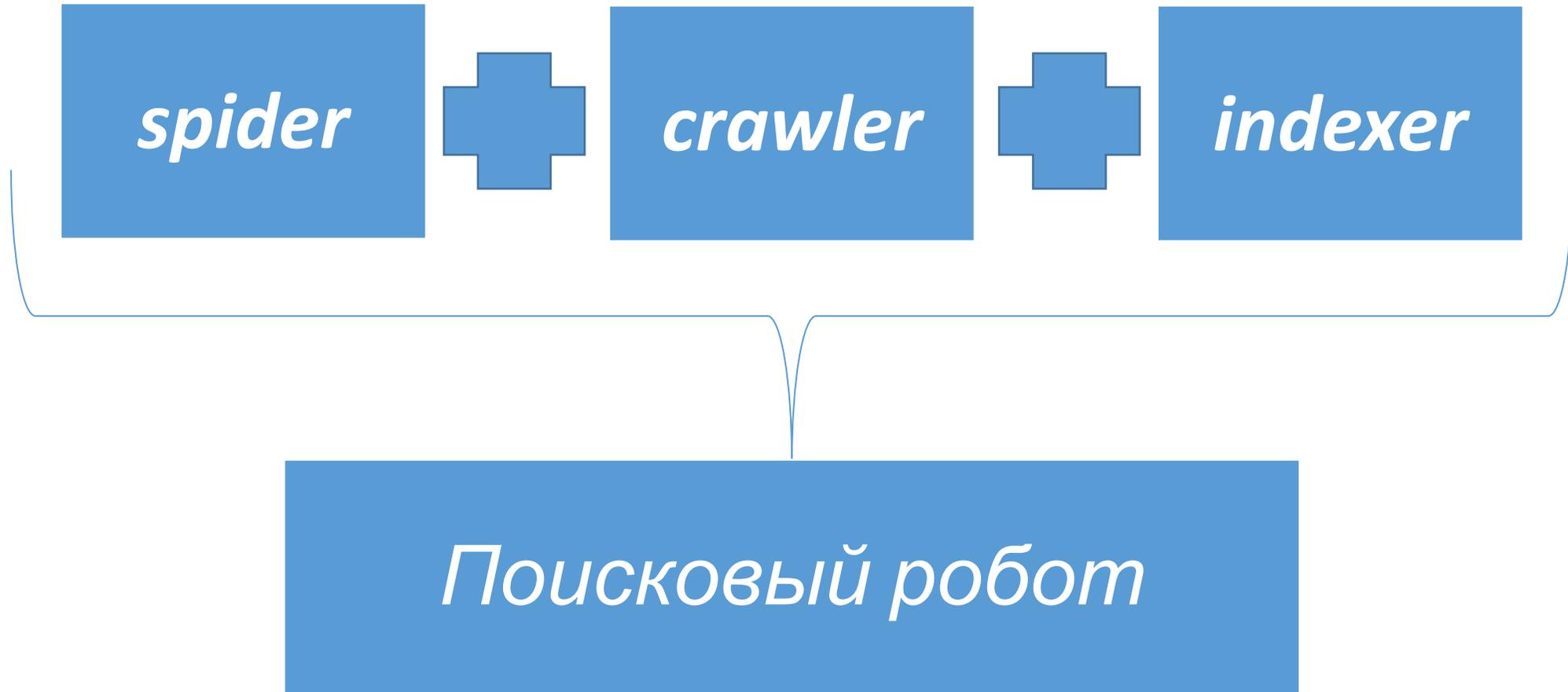
*Индексатор
(indexer)*

*База данных
(database)*

*Система
выдачи
результатов
(search engine
results engine)*

*Web-сервер
(Web-server)*

Поисковые системы



Поисковые системы

Для ранжирования страниц в поисковой выдаче используются:

1

текстовые критерии

2

ссылочные критерии

3

критерии пользовательской оценки

Основные текстовые критерии ранжирования документов

Критерии	Логика ранжирования
«Вес» слова	Чем выше частота повторения слова в документе. тем больше ранг документа
Взаимное положение слов	Учет полного совпадения фраз или их подобия (например, порядок и близость слов друг к ДРУГУ)
Положение найденного текста по отношению к началу документа	Считается, что чем ближе расположена информация к началу документа, тем выше ее значение
Наличие слов запроса в выделенных фрагментах и заголовках	Значимость обнаружения искомого текста в выделенных фрагментах считается выше, чем в обычном тексте
Совпадение темы страницы с темой запроса	Использование в поиске слов, не содержащихся в тексте запроса, но соответствующих теме запроса
Совпадение названия домена или файла с ключевым словом	Поисковые машины придают дополнительный «вес» страницам, у которых домен или имя файла совпадают с ключевым словом
Совпадение поискового запроса с описанием из каталога	Сайт получает более высокий рейтинг, если слова поискового запроса совпадают с описанием каталога поисковой системы
Значимости редких слов	Значимости каждого из поисковых слов тем больше, чем

Оценка значимости фрагментов текста:

$$V = \frac{N_k^2}{N_0}$$

где:

V — значимость фрагмента;

N_k — число ключевых слов в данном фрагменте;

N_0 — общее число слов во фрагменте.

Система выявления ключевых слов обычно использует статистический частотный анализ:

$$F(P) = C \frac{1}{P}$$

где:

F — частота, с которой встречаются различные слова в тексте;

P — относительное значение полезности (важности);

C — константа, которая определяет соотношение частоты слов и их полезности

Ссылочный критерий:

Согласно *ссылочным критериям* документ ранжируется с учетом индекса цитирования.

Индекс цитирования — это показатель известности сайта в Интернете, определяемый числом и значимостью ссылок на других сайтах на искомый ресурс.

ТИЦ и ВИЦ
Яндекс

Google™
PageRank

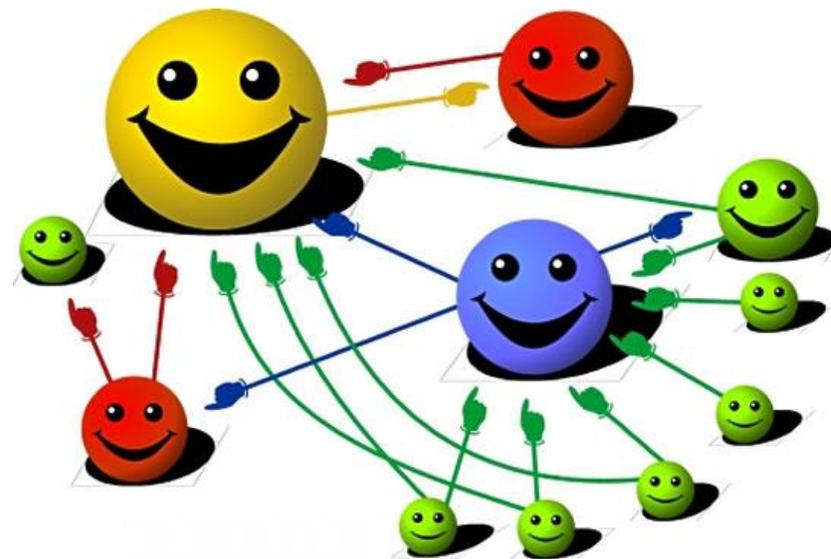


Ранг страницы :

Вероятность пользователя попасть на конкретный документ зависит от количества ссылок на него с других документов и от того, насколько вероятно нахождение пользователя на одном из ссылающихся документов и сколько исходящих ссылок содержит



10



Показатель авторитетности, ранг страницы (*PageRank*):

Ранг страницы:

$$PR_a = (1 - d) + d \sum_{i=1}^n \frac{PR_i}{C_i}$$

где:

PR_a — PageRank страницы a ;

d — коэффициент затухания, обычно устанавливают равным 0,85;

i — страница, содержащая ссылки на страницу a (i изменяется от 1 до n);

PR_i — PageRank страницы i , ссылающейся на страницу a ;

C_i — общее число ссылок на странице i ;

$1/C_i$ — вероятность того, что пользователь, находящийся на странице i , из C_i доступных ему ссылок выберет именно ссылку на страницу a ;

$d * PR_i / C_i$ — поток «теоретической посещаемости», который дойдет до страницы a со страницы i (суммирование идет по всем страницам, ссылающимся на страницу a);

$(1 - d)$ — минимальный PageRank страницы (он не равен нулю за счет того, что пользователь регулярно выбирает новый сайт в качестве стартовой точки).

Критерий пользовательской оценки:

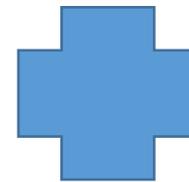
Если пользователь переходит по ссылке, значит он счел ее интересной, и, если долго не возвращается на страницу поисковой системы, значит его ожидания подтвердились.



Rambler

Поисковая система *Rambler* при ранжировании результатов поиска в ответах на поисковый запрос использует *коэффициент популярности*, определяемый числом пользователей, которые просматривали данную страницу за последние несколько недель.

коэффициент,
основанный на учете
гиперссылок между
страницами сети,



Rambler
РЕЙТИНГ ТОП100

Ограничения возможностей поисковых машин

- *Физические ограничения скорости.*
- *Поиск информации — мероприятие довольно дорогостоящее.*
- *Принцип попадания страниц в индекс при помощи «пауков».*
- *Необычные слова на странице, интересующей пользователя.*
- *Предпочтение поисковой машиной быстроты поиска, а не его глубины.*
- *Ориентация поисковых машин на поиск текстов в разных вариантах.*