



Брестский государственный
технический университет



ФАКУЛЬТЕТ
ЭЛЕКТРОННО-
ИНФОРМАЦИОННЫХ
СИСТЕМ



Интернет и www-сеть

Весна 2022

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles of various sizes and colors, including blue, yellow, orange, pink, and green, creating a bubbly, abstract effect.

План лекции

1. история сети «Интернет»
2. web-сеть
3. протокол HTTP
4. web страницы

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in various colors including blue, yellow, orange, pink, and green, arranged in a vertical gradient.

Что на картинке?

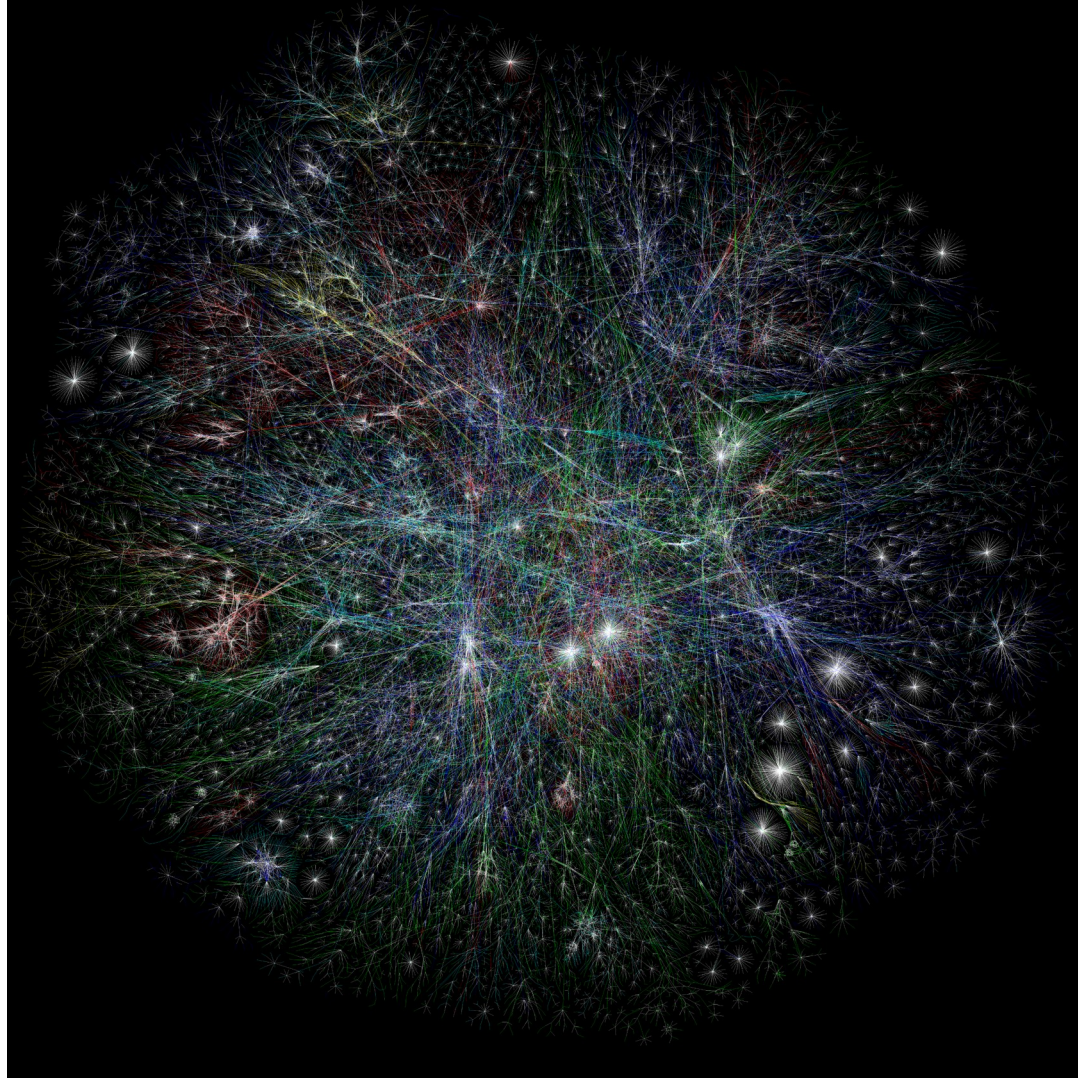




Схема сети Интернет в 2003

В начале 2003 года 665 млн. имели доступ к Интернету, с 2001 года это количество увеличилось примерно на 111 млн. человек (приблизительно 13.6%).

Самым популярным местом пользования Интернетом является работа (48%), затем следуют дом (32%), место учебы (21%) и, наконец, Интернет-кафе (12%).

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles of various sizes and colors, including blue, yellow, orange, pink, and green.

Что на картинке?

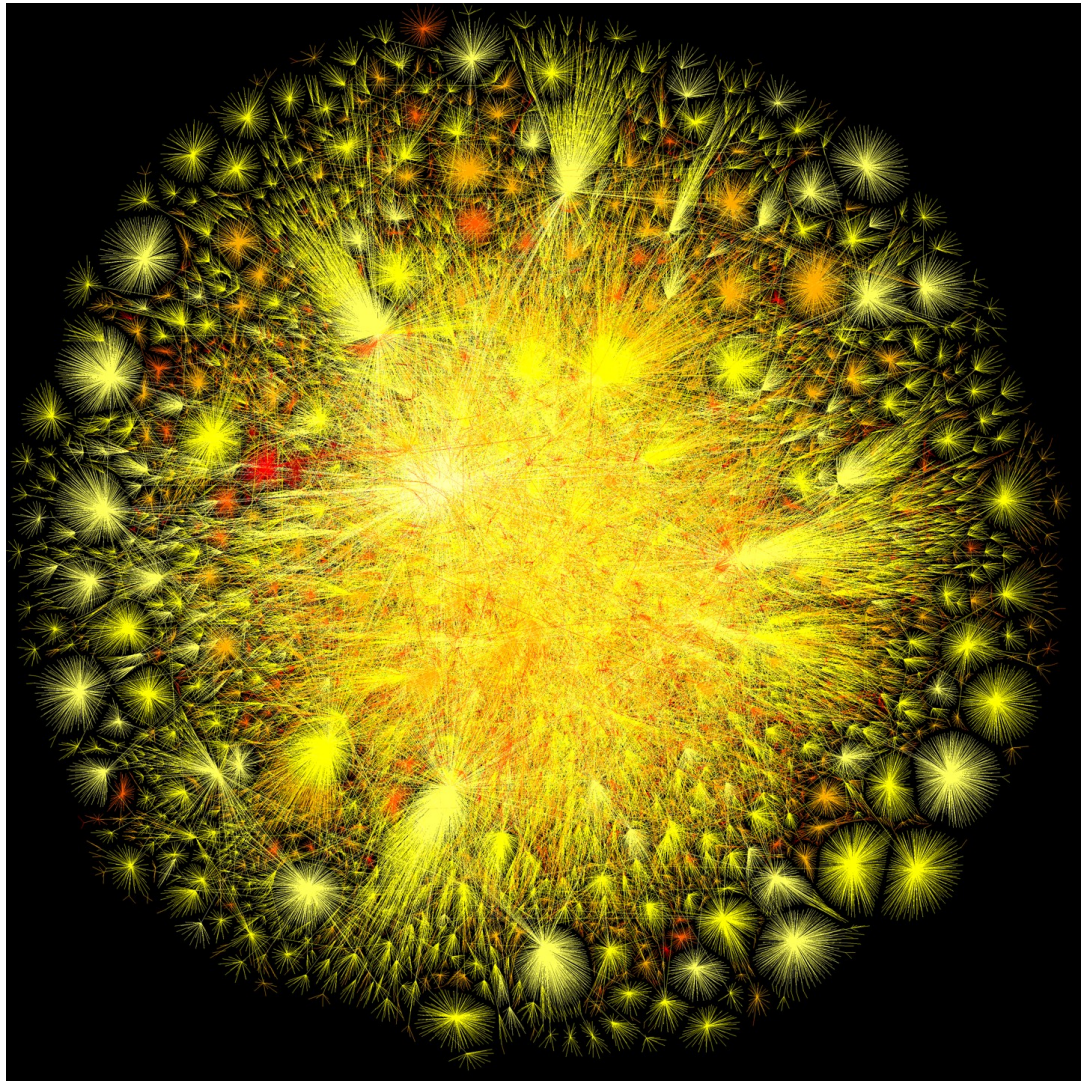




Схема сети Интернет в 2010

- 1.97 млрд – пользователей по всему миру.
- 255 млн – количество веб сайтов.
- 21.4 млн – новых сайтов за 2010 г.
- По официальной статистике в РБ пользуется интернетом 30% населения.

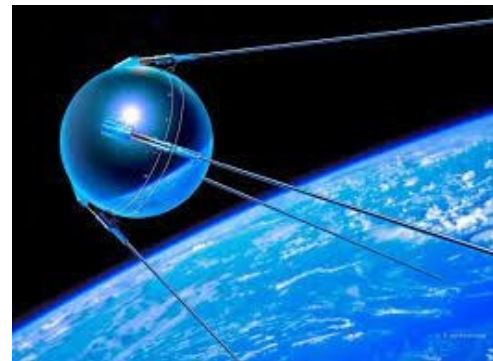


Интернет 2021

- Интернетом пользуются 5,15 миллиарда человек в год прирост 316 миллионов по сравнению к прошлому году (7,5 %).
- На 18 декабря 2021 года насчитывалось более 1.9 миллиарда веб-сайтов.
- Среднестатистический пользователь Интернета проводит в сети 6 часов 43 минуты каждый день.

История интернета

4 октября 1957 СССР запустил первый искусственный спутник Земли.



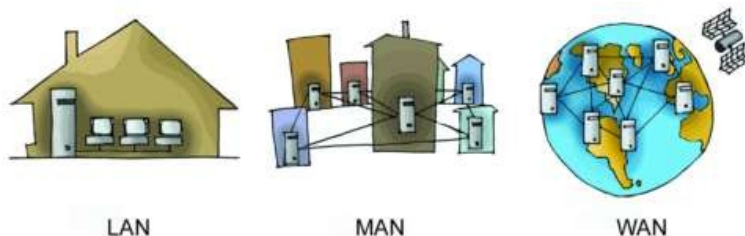
Это событие мирового масштаба было неоднозначно воспринято рядом стран, которые видели в запуске спутника угрозу собственной безопасности.

В 1958 ответ на это Соединенные Штаты Америки формируют Агентство перспективных исследовательских проектов (ARPA).

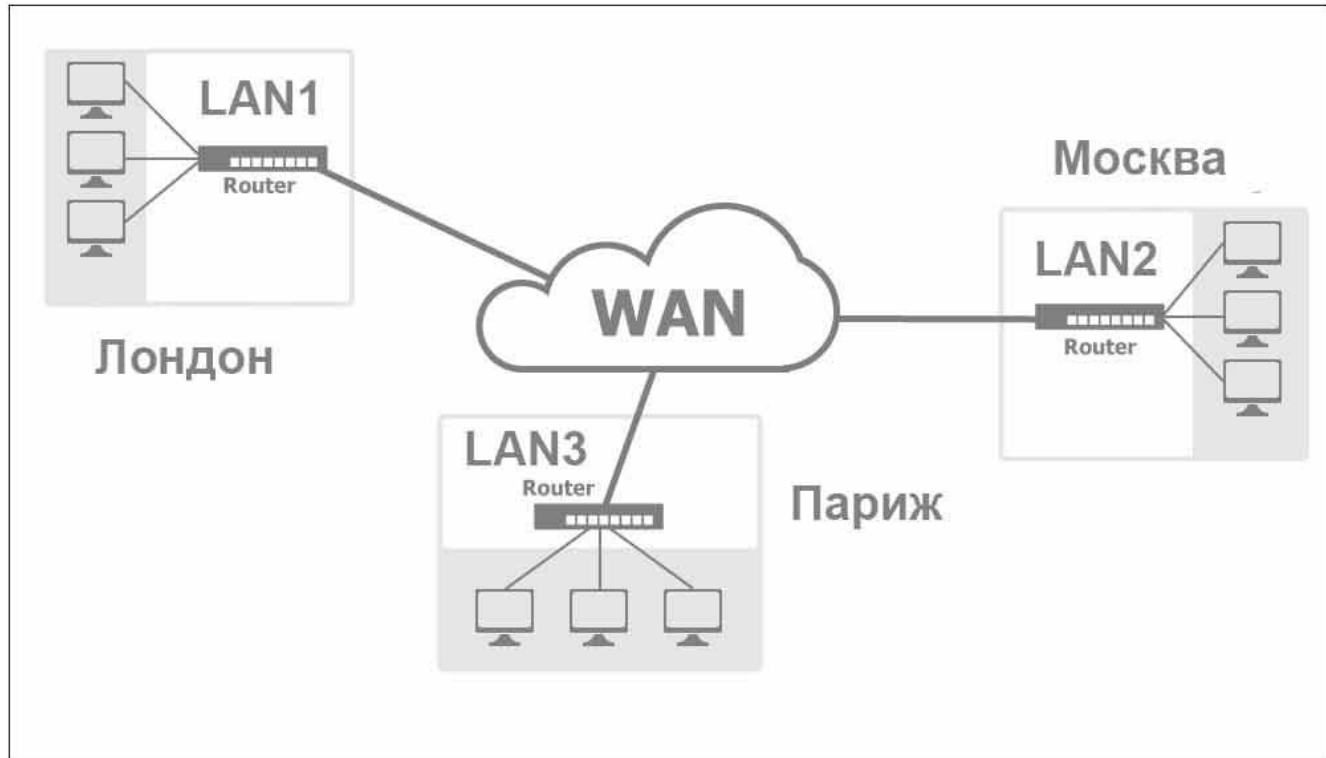
История интернета

Концепция подключения нескольких компьютеров к одной и той же сети в то время не была такой уж новаторской.

Такие инфраструктуры, называемые *глобальными сетями* (WAN), возникли в середине 1950-х годах и получили развитие в течение 1960-х годов.



Сеть WAN





История интернета

WAN, как концепция, довольно прост. Это буквально группа подключенных компьютеров.

WAN того времени были ограничены небольшими зонами и были ограничены определенными технологическими ограничениями.

История Интернета

Джозеф Карл Робнетт Ликлайдер
(J. C. R. Licklider)



В 1960-х годах он черпал в одной из своих работ предложил идею глобальной сети.

Этот человек затем станет одним из директоров ARPA с задачей соединения основных компьютеров Министерства обороны США.

История Интернета

Он покинул агентство в 1964 до создания ARPANET, но именно его идеи послужили основанием для его преемника Роберта Тейлора, к которому в последствии присоединился Ларри Робертс, чтобы начать разработку Интернета.



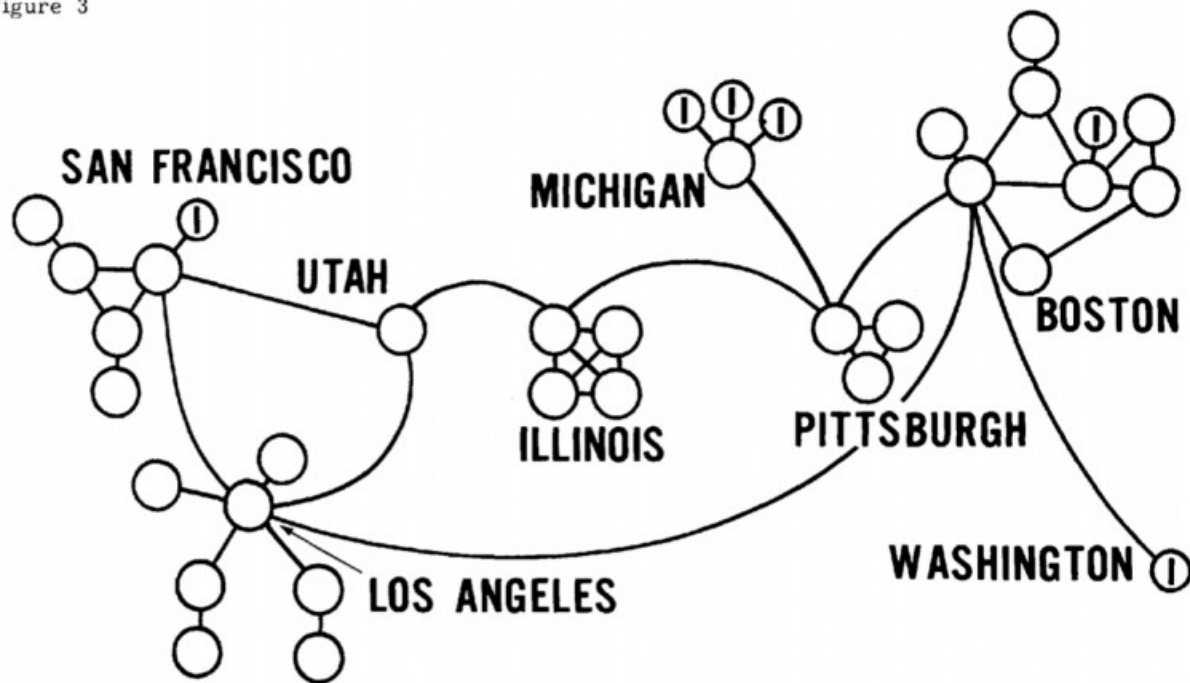


История Интернета

В 1968 году Министерство обороны США посчитало, что на случай войны Америке нужна надёжная система передачи информации, и предложило разработать для этого компьютерную сеть.

Разработка такой сети была поручена Калифорнийскому университету в Лос-Анджелесе, Стэнфордскому исследовательскому центру, Университету Юты и Калифорнийскому университету в Санта-Барбаре.

Figure 3



ARPA COMPUTER NETWORK*

- REPRESENTS A TIME-SHARED COMPUTER
- ① REPRESENTS A SINGLE USER CONSOLE COMPUTER

* NOT FINAL, ESTIMATE AS OF JUNE 1967

ARPANET

В 1969 году, когда впервые была выполнена передача данных между компьютерами Калифорнийского университета Лос-Анжелеса и Стэнфордского исследовательского института на расстояние 640 км.



The ARPANET in December 1969



ARPANET

Первое испытание технологии произошло 29 октября 1969 года в 21:00. Сеть состояла из двух терминалов.

В испытании участвовали два оператора.

Первыми данными была буква L

Просто буква L, потом буква O, затем один из них набрал G, и ... вся система рухнула.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

ARPANET

Первый эксперимент потерпел неудачу, — отобразились только буквы «l», «o».

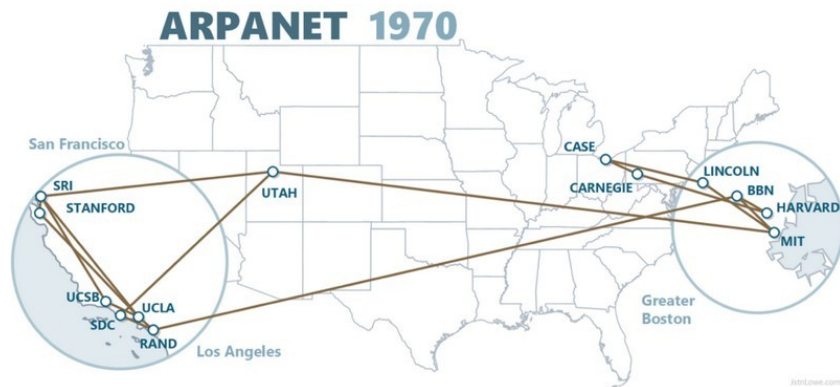
Через час эксперимент был повторён и прошёл удачно.

Оператор ввёл тестовое слово «login», являвшееся командой входа в систему, а второй оператор подтвердил, что он видит его у себя на экране.

ARPANET

После этого ARPANET начал распространяться. Сначала вся сеть соединяла четыре узла — три в Калифорнии и один в Юте.

В 1970 году восточные и западные побережья Северной Америки были связаны соединением с сетью Кембриджа.



A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Пакетная коммуникация

Ранняя архитектура ARPANET не была разработана с учетом быстрого расширения, поэтому ученые быстро поняли, что у них будут проблемы.

В то время передача данных была основана на так называемой *коммутации каналов* .

По сути, используя этот метод, вы можете отправлять данные только в виде полного пакета и только на один компьютер одновременно.

Пакетная коммуникация

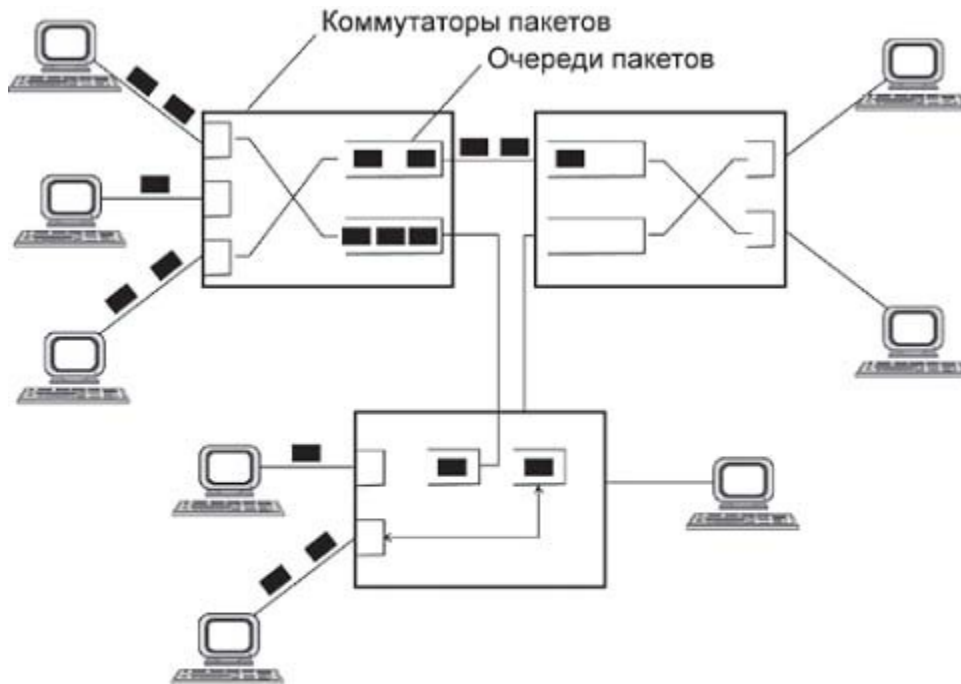
Решение этой проблемы было предложено Великобританским ученым Дональдом Дэвисом.



Он впервые внедрил метод передачи данных, называемый *пакетной коммутацией*.

Его метод позволяет обрезать пакет данных на одном конце и соединять вместе на другом. Это впервые позволило устранить перегруженность и подключить многие компьютеры к сети.

Пакетная коммуникация



- Данные нарезаются порциями - пакетами, каждый из которых обрабатывается коммутаторами независимо
- Каждый пакет содержит адрес назначения и адрес отправителя
- Не требуется предварительной процедуры установления соединения



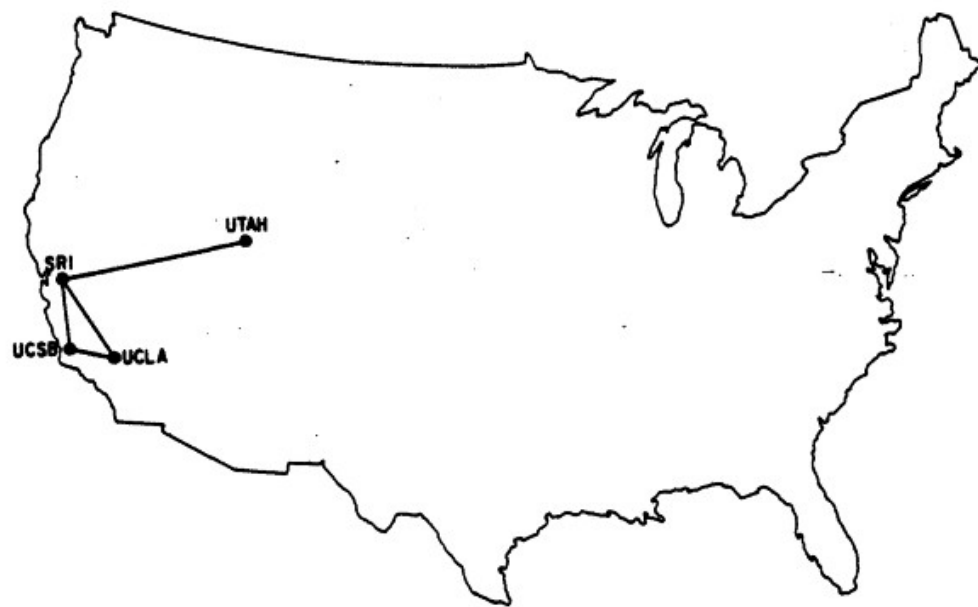
Пакетная коммуникация

Вскоре другие сети последовали за ARPANET. Французские CYCLADES, American Merit и, предположительно, первая общедоступная сеть — Telenet.

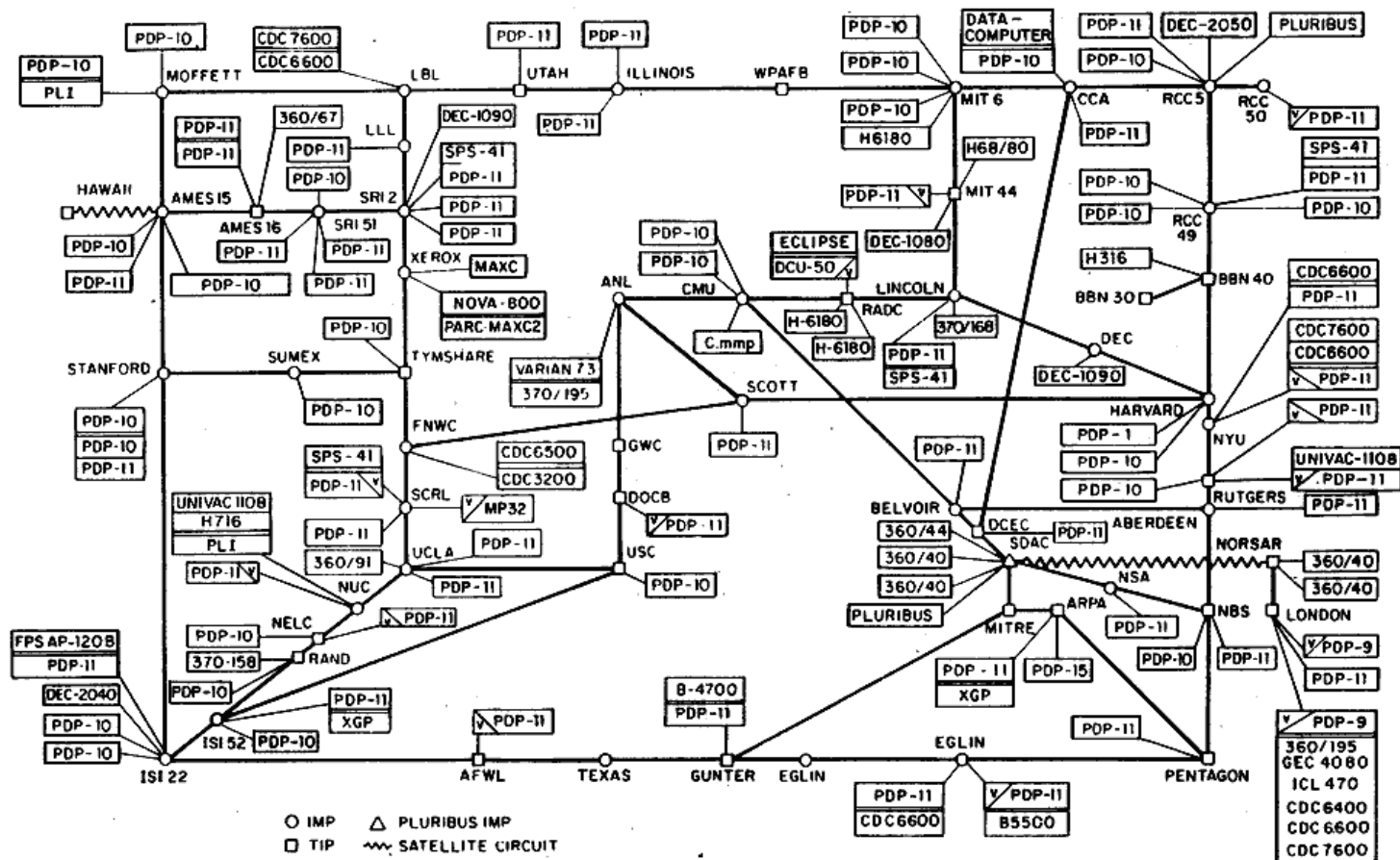
К 1981 году к ARPANET было подключено 213 компьютеров и еще много к другим сетям.

ARPANET соединял континентальную часть Соединенных Штатов, Гавайи, а также Лондон, которые были подключены через спутниковую связь.

December 1969



ARPANET LOGICAL MAP, MARCH 1977



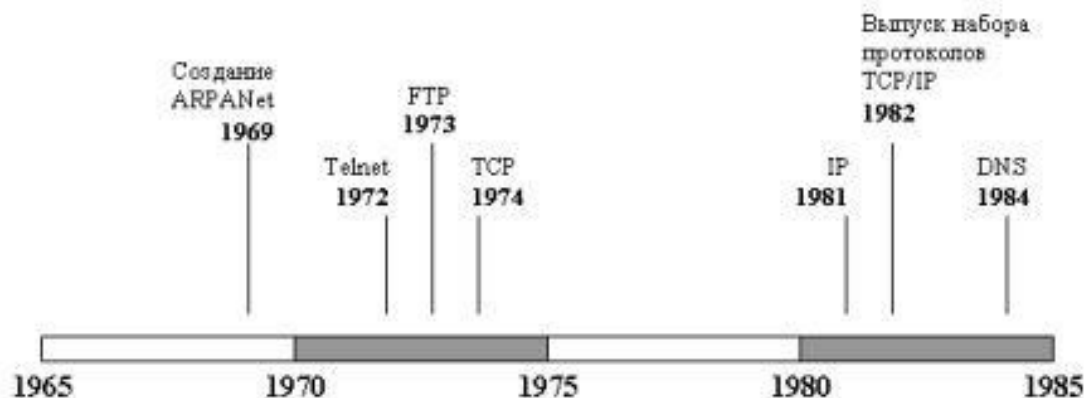
(PLEASE NOTE THAT WHILE THIS MAP SHOWS THE HOST POPULATION OF THE NETWORK ACCORDING TO THE BEST INFORMATION OBTAINABLE, NO CLAIM CAN BE MADE FOR ITS ACCURACY)

NAMES SHOWN ARE IMP NAMES, NOT (NECESSARILY) HOST NAMES

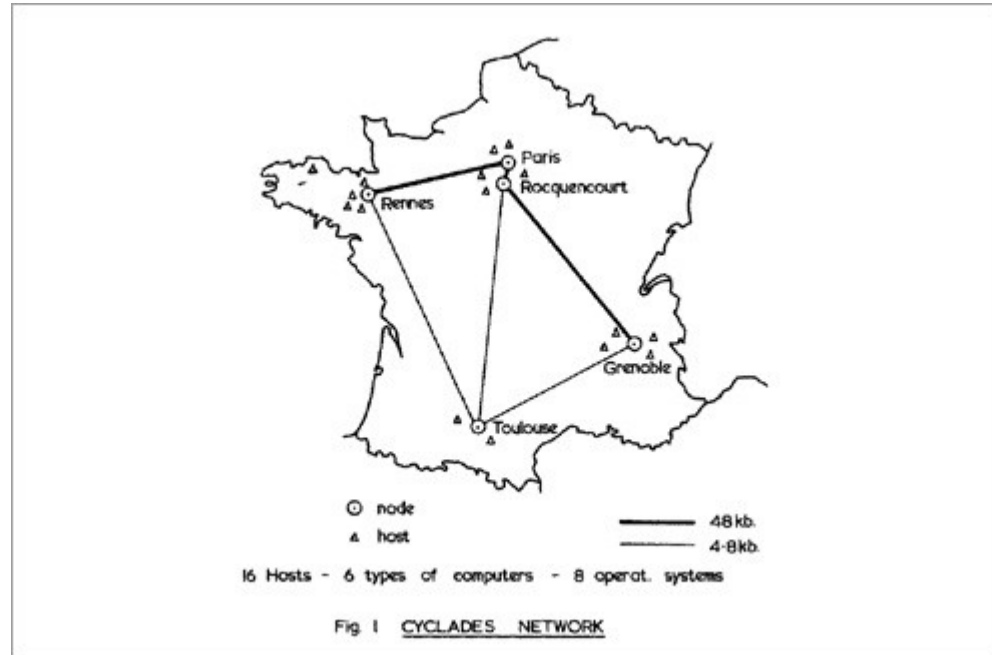
ТСР/IP

С появлением таких сетей, как CYCLADES и NPL, ученые столкнулись с другой проблемой.

В разных сетях были разные протоколы, по сути это означало, что они говорили на разных языках.

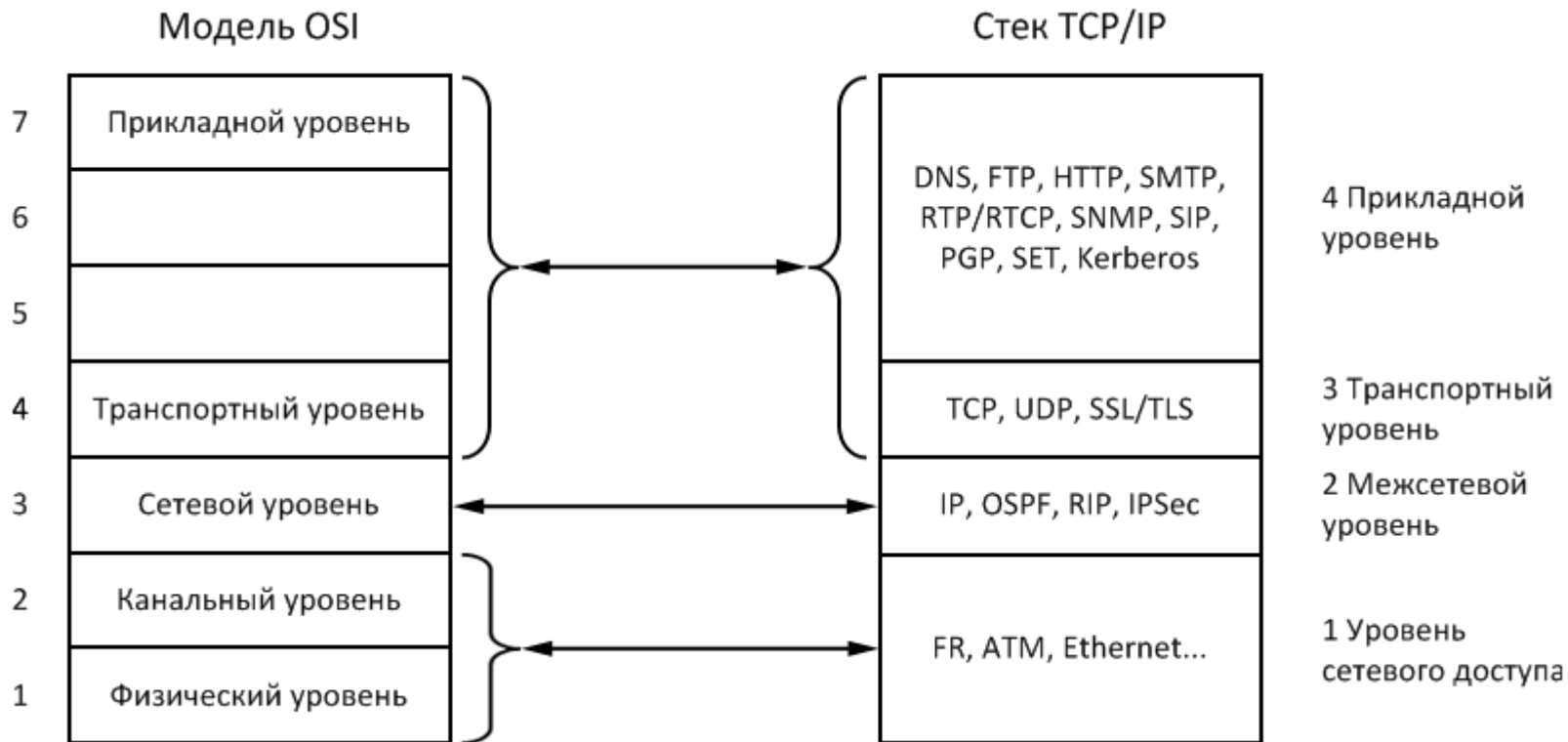


ТСР/IP



Точно так же, как говорящий по-французски не мог разговаривать по телефону с говорящим на английском языке, так ARPANET не мог общаться с

ТСР/IP

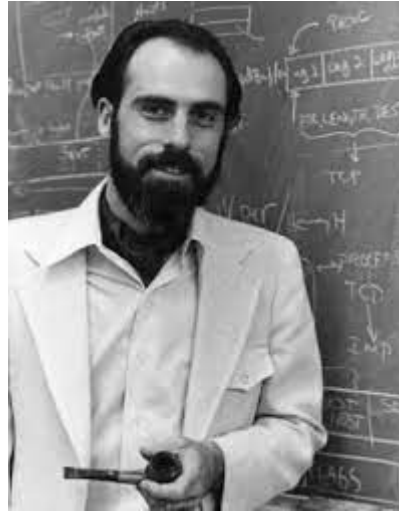


TCP/IP



Учёные разработали теоретическую модель стандартного пакета Интернет протоколов, широко известного как TCP / IP

ТСР/IP



ТСР/IP был создан на основе NCP (Network Control Protocol) группой в 1972 году. В июле 1976 года Винт Серф и Боб Кан впервые продемонстрировали передачу данных с использованием **ТСР** по трём различным сетям.



ТСР/IP

Позже был заключен контракт с BBN Technologies, Стэнфордским университетом и Университетским колледжем Лондона на разработку операционных версий протокола.

В ходе разработки протокола номер версии перешел с версии 1 на версию 4, и последняя стала известна как *версия 4 интернет-протокола* (IPv4), которая до сих пор используется наряду с более новой версией — IPv6.

Адреса IP

Согласно спецификации протокола, каждому узлу, подсоединенному к IP-сети, присваивается уникальный номер.

Узел может представлять собой компьютер, маршрутизатор, межсетевой экран и др.

IPv4 адрес (точечно-десятичная нотация)

172 . 16 . 254 . 1

10101100 . 00010000 . 11111110 . 00000001

1 байт = 8 бит

32 бита (4 x 8) или 4 байта

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Адреса IP

Если один узел имеет несколько физических подключений к сети, то каждому подключению должен быть присвоен свой уникальный номер.

Этот номер, или по-другому IP-адрес, имеет длину в четыре октета, и состоит из двух частей.

Первая часть определяет сеть, к которой принадлежит узел, а вторая -- уникальный адрес самого узла внутри сети.



Адреса IP

Номер сети			Номер узла
11011100	11010111	00001110	00010110

В классической реализации протокола первую часть адреса называли "сетевым префиксом", поскольку она однозначно определяла сеть. Однако в современной реализации это уже не так и сеть идентифицируют другим образом.



Классическая схема адреса IP

Изначально все адресное пространство разделили на пять классов: А, В, С, D и Е. Такая схема получила название "классовой".

Каждый класс однозначно идентифицировался первыми битами левого байта адреса.

Сами же классы отличались размерами сетевой и узловой частей.



Адреса IP

Зная класс адреса, вы могли определить границу между его сетевой и узловой частями.

Кроме того, такая схема позволяла при маршрутизации не передавать вместе с пакетом информацию о длине сетевой части IP-адреса.



IP организация подсетей

Очень редко в локальную вычислительную сеть входит более 100-200 узлов: даже если взять сеть с большим количеством узлов, многие сетевые среды накладывают ограничения, например, в 1024 узла.

Исходя из этого, целесообразность использования сетей класса **A** и **B** весьма сомнительна

Да и использование класса **C** для сетей, состоящих из 20-30 узлов, тоже является расточительством.



IP организация подсетей

Для решения этих проблем в двухуровневую иерархию IP-адресов (сеть -- узел) была введена новая составляющая -- подсеть.

Идея заключается в "заимствовании" нескольких битов из узловой части адреса для определения подсети.

Деление сетей на подсети устранили проблему на несколько лет.

Адреса IP

Но уже в начале 90-х к проблеме большого количества маршрутов прибавилась нехватка адресного пространства.

Ограничение в 4 миллиарда адресов, заложенное в протокол и казавшееся недостижимой величиной, стало весьма ощутимым.

Классы адресов в первоначальной схеме IP-адресации

Класс	Первые биты в октете	Возможные значения первого октета	Возможное число сетей	Возможное число узлов в сети
A	0	1–126	126	16777214
B	10	128–191	16384	65534
C	110	192–223	2097152	254
D	1110	224–239	Используется для многоадресной рассылки (multicast)	
E	1111	240–254	Зарезервирован как экспериментальный	

Адреса IP

IP-адрес — это уникальная 32-разрядная последовательность двоичных цифр, с помощью которой компьютер *однозначно идентифицируется* в IP-сети.

IP-адрес в 32-разрядном виде	11000000 10101000 0000101 11001000			
IP-адрес, разбитый на октеты	11000000	10101000	00000101	11001000
Октеты в десятичном представлении	192	168	5	200
IP-адрес в виде десятичных чисел, разделенных точками	192.168.5.200			

Адреса IP

Маска подсети — это 32-разрядное число, состоящее из идущих вначале единиц, а затем — нулей, например (в десятичном представлении) 255.255.255.0 или 255.255.240.0.

Пример: IP-адрес
маска подсети

192.168.5.200
255.255.255.0

Тогда:

идентификатор сети 192.168.5.0, а идентификатор узла 200

Если поменять маску на:

255.255.0.0, то для IP-адреса 192.168.5.200

идентификатор сети 192.168.0.0, а идентификатор узла 0.0.5.200

Web Сеть

В 1989 году Tim Berners сотрудник Европейской организации ядерных исследований CERN предложил создать гипертекстовую систему коллективной работы с документами которая позволяла выполнять переходы между ними.

Он так же разработал основные стандарты на которых основывается данная система.





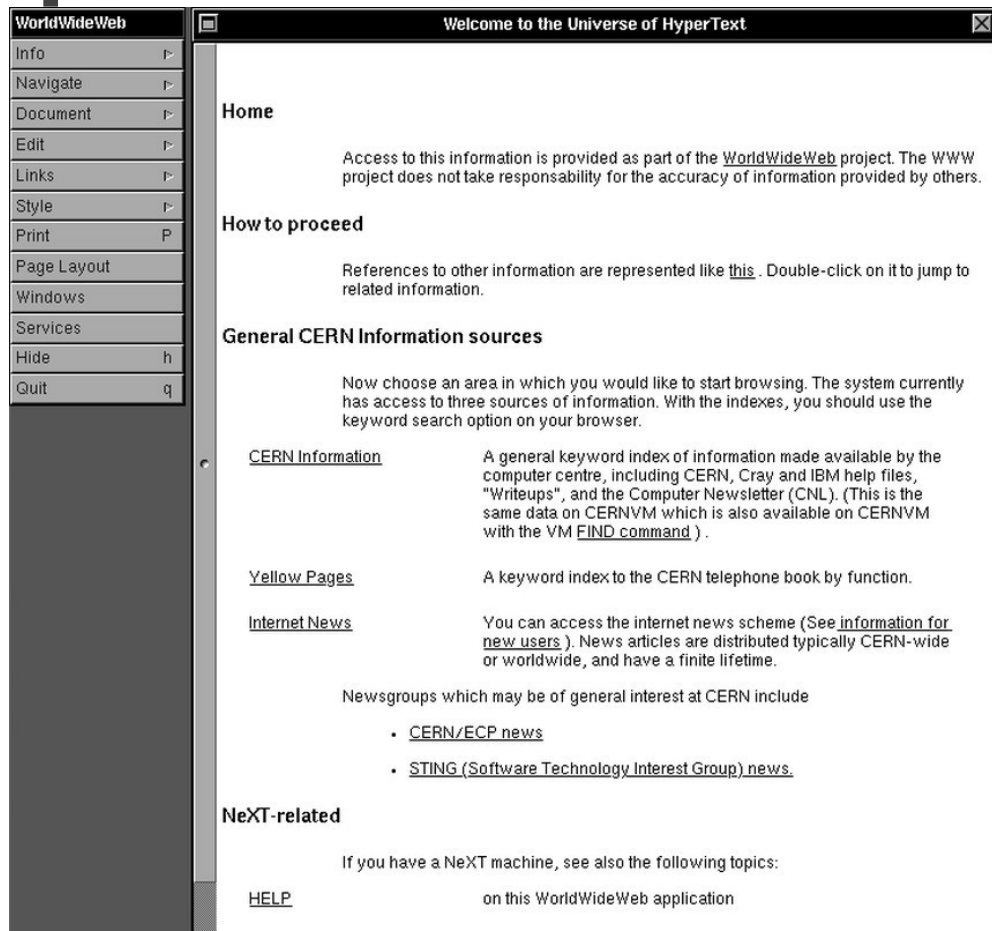
Web Сеть

Основной задачей Тима было управление научными исследованиями и облегчить обмен ими с коллегами.

Он изобрел первый в мире интернет-браузер — WorldWideWeb.

Первый сайт (<https://worldwideweb.cern.ch/browser/>) был создан им в августе 1991 года и проложил путь к появлению сотен миллионов других веб-сайтов в последующие годы.

Первый Web сайт



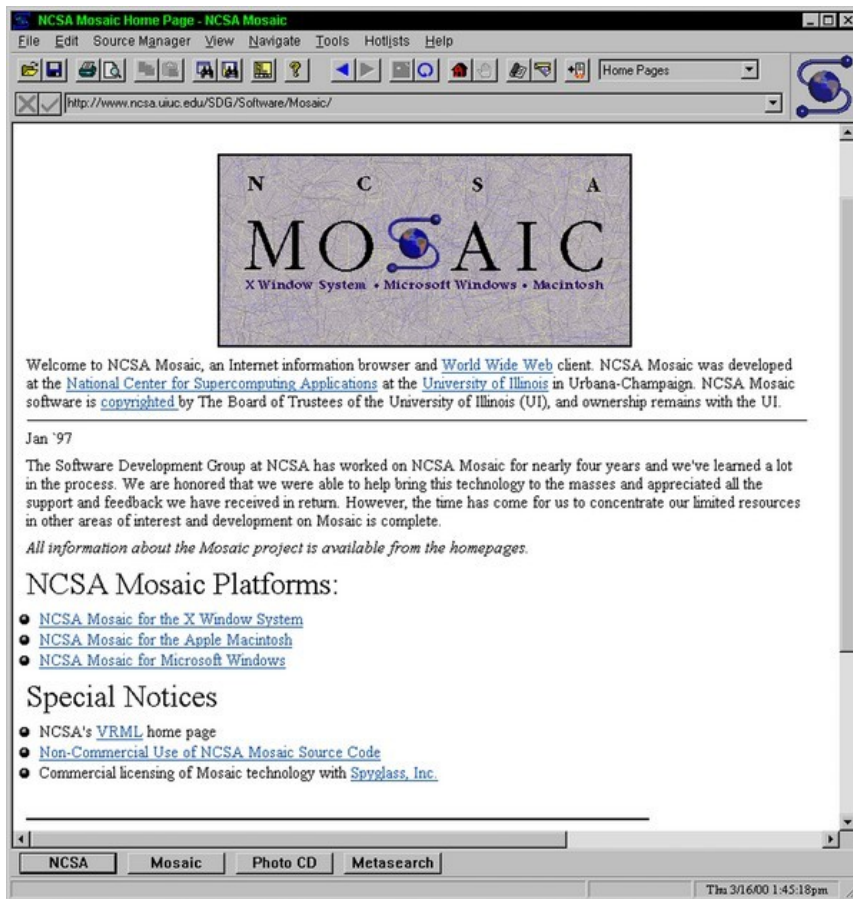


Первый Web браузер

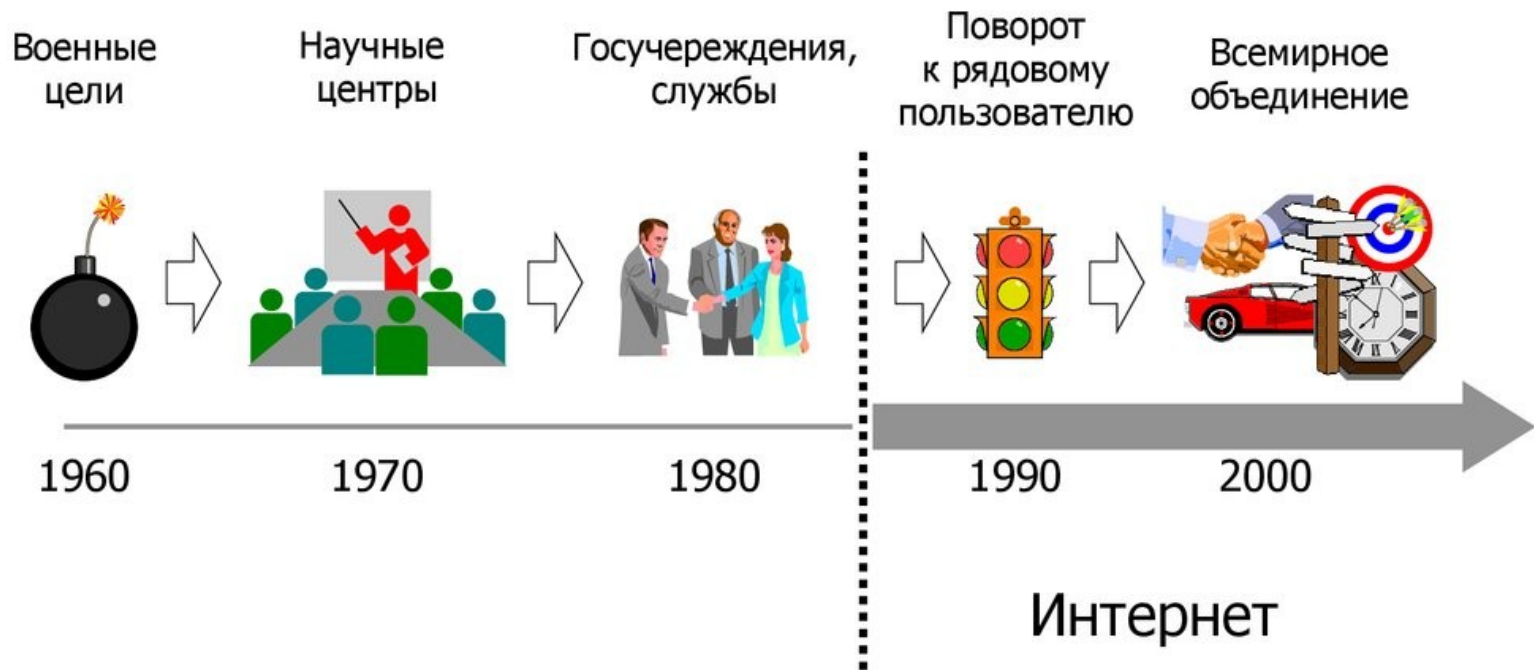
Хоть WorldWideWeb был первым браузером, действительно мейнстримом сделал Интернет-NCSA Mosaic.

Он был первым в своем роде, предоставившим мультимедийный контент пользователям, и, возможно, положил начало развитию современного Интернета, в котором вы можете прокручивать Facebook или Instagram и смотреть видео на Youtube или TikTok.

Первый Web браузер



История интернета





Web Сеть

Прежде всего следует отметить, что web-сеть, не является физической сетью.

Это **сервис** предоставляющий пользователям доступ к связанному с помощью ссылок набору гипертекстовых документов и web приложений.

Основным структурным элементом является web-сайт (website, site), это обобщённый термин обозначающий объединённый под одним адресом (доменным или IP) набор логически связанных ресурсов.



Web Сеть

Все ресурсы web – сайтов можно разделить на два типа:

- **статические ресурсы** – HTML документы, изображения, мультимедиа файлы и вообще любые файлы данных к которым есть доступ.
- **динамические ресурсы** – web приложения, программные модули (exe, dll), шаблоны web-страниц, скрипты, программные объекты которые как правило по запросу формируют HTML документы.



Web Сеть

Web-сайты идентифицируются в глобальной сети под доменными именами, зарегистрированными в системе DNS. В большинстве случаев одному сайту соответствует одно имя.

Однако на одном компьютере под управлением одного web сервера могут находиться несколько web сайтов.

В настоящее время большинство сайтов(более 80%) являются web приложениями.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in various colors including blue, yellow, orange, pink, and green, arranged in a gradient-like fashion from top to bottom.

Web Сеть

Работа web сети основывается на стандартах из которых основными являются:

- способ задания адресов ресурсов сети — URL
- протокол взаимодействия между клиентами и серверами — HTTP
- язык описания гипертекстовых документов — HTML
- язык форматирования гипертекстовых документов — CSS
- язык описания программ выполняемых на стороне клиента — JavaScript



Адресация ресурсов

Свое начало система доменных имен берет в 50-х — 60-х годах прошлого века. В 80-х к ней были подключены целых 320 компьютеров.

Такое количество устройств породило проблему — стало сложно работать с адресами.

Для обмена данными каждый из подключенных компьютеров скачивал себе файл HOSTS.TXT с информацией об остальных хостах



Адресация ресурсов

Этот файл существовал в единственном экземпляре на сервере, размещенном в Стэнфордском исследовательском институте.

Пользователям становилось все сложнее работать с раздутым списком, учитывая тот факт, что идентификаторы при подключении приходилось прописывать вручную.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

Адресация ресурсов

Особенно сильно проблема была заметна во время отправки электронных писем.

Чтобы переслать сообщение с одного компьютера на другой, пользователь должен был сам указать путь для его передачи между системами.

Задачу усложнял тот факт, что отдельные устройства могли выходить в сеть в разное время дня. Выглядел коммуникационный путь примерно следующим образом:

```
utzoo!decvax!harpo!eagle!mhtsa!ihnss!  
ibuxnlara
```



Адресация ресурсов

Путь представлял собой цепочку из хостов, разделённых служебным символом (!).

В конце последовательности прописывалось имя системы получателя сообщения.

Если пользователь не знал коммуникационный путь, то он не мог

Адресация ресурсов

В 1982 году группа специалистов из компании Network Working Group (NWG), возглавляемая Стивом Крокером (Steve Crocker), который изобрел Request For Comments, представила решение — концепцию доменных имен (RFC805)



A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Адресация ресурсов

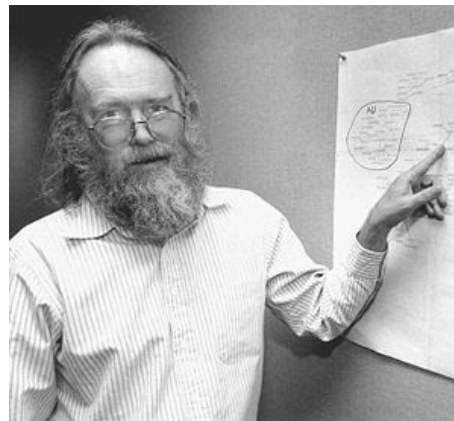
Инженеры под руководством Стива разработали специальное программное обеспечение, которое автоматизировало поиск маршрутов и позволило адресовать сообщения любому хосту напрямую.

Базовые принципы, заложенные в RFC805, стали отправной точкой для запуска привычной нам системы доменных имен.

Адресация ресурсов

Domain	Resource Record		
" "	NS	IN	B.ISI.ARPA
" "	NS	CS	CSNET.UDEL
ARPA	SOA	IN	B.ISI.ARPA
ARPA	NS	IN	A.ISI.ARPA
ARPA	NS	IN	F.ISI.ARPA
MIT.ARPA	NS	IN	AI.MIT.ARPA
ISI.ARPA	SOA	IN	F.ISI.ARPA
ISI.ARPA	NS	IN	F.ISI.ARPA
A.ISI.ARPA	MD	IN	A.ISI.ARPA
ISI.ARPA	MD	IN	F.ISI.ARPA
A.ISI.ARPA	MF	IN	F.ISI.ARPA
B.ISI.ARPA	MD	IN	B.ISI.ARPA
B.ISI.ARPA	MF	IN	F.ISI.ARPA
F.ISI.ARPA	MD	IN	F.ISI.ARPA
F.ISI.ARPA	MF	IN	A.ISI.ARPA
DTI.ARPA	MD	IN	DTI.ARPA
NBS.ARPA	MD	IN	NBS.ARPA
UDEL.ARPA	MD	IN	UDEL.ARPA
A.ISI.ARPA	A	IN	10.1.0.32
F.ISI.ARPA	A	IN	10.2.0.52
B.ISI.ARPA	A	IN	10.3.0.52
DTI.ARPA	A	IN	10.0.0.12
AI.MIT.ARPA	A	IN	10.2.0.6
DMS.MIT.ARPA	A	IN	10.1.0.6
NBS.ARPA	A	IN	10.0.0.19
UDEL.ARPA	A	IN	10.0.0.96

DNS

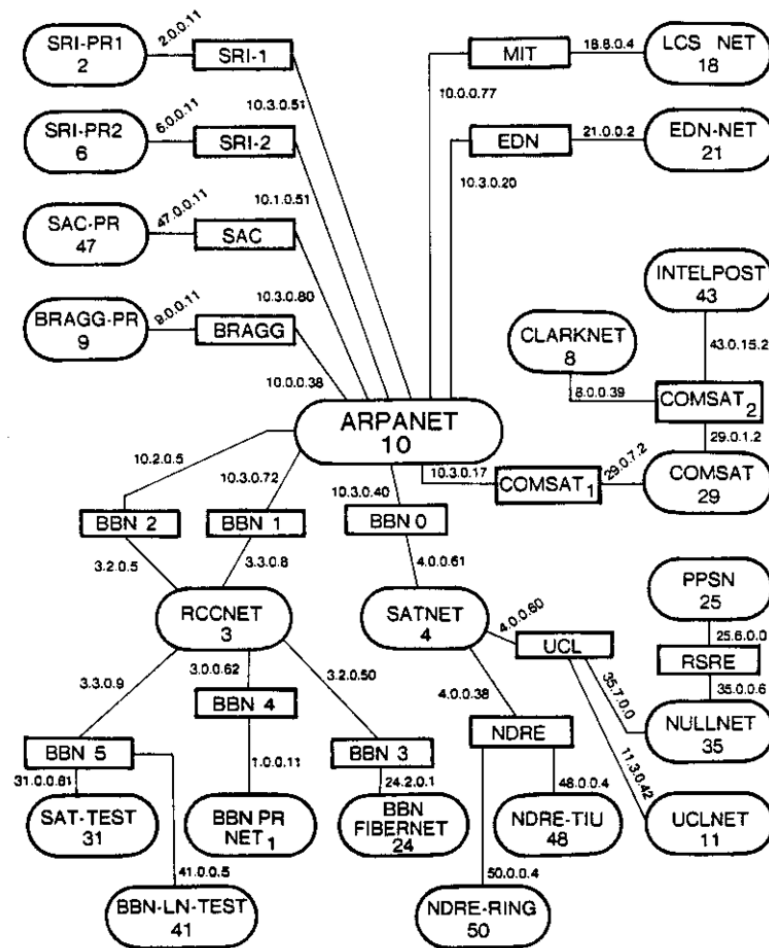


В 1983 году инженеры Пол Мокапетрис (Paul Mockapetris) и Джон Постел (Jon Postel) решили распространить концепцию, описанную в RFC805, на всю сеть ARPANET.

Они подготовили два новых RFC, в которых изложили основы DNS.

DNS

POSTEL 25 FEB 82





DNS

Мокапетрис предложил, чтобы имена хостов включали:

Имя — например, «IBM»

Категории/Цель — например, «.com — для коммерческих целей»

Спустя год на основе его спецификации родилась классификация gTLD (generic Top-Level Domains), куда вошли домены .com, .edu, .net, .org, .int, .gov и .mil.



DNS

Но система доменных имен — это больше, чем просто схема именования.

DNS также переводит имена веб-сайтов в адреса компьютеров (IP-адреса) и находит для вас соответствующую веб-страницу в Интернете.

Таким образом, вам нужно запомнить только текст, а не комбинации цифр.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

DNS

Так, например, чтобы связаться с Google, вы просто вводите google.com в адресную строку браузера вместо IP-адреса, присвоенного домену Google.com, 74.125.224.72.

И если Google когда-нибудь изменит свой IP-адрес, вам не нужно об этом беспокоиться.

DNS-серверы запишут новый адрес и все равно направят вас обратно на домашнюю страницу Google.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

DNS

Создатель DNS Мокапетрис сказал в интервью, что он был амбициозен, когда концептуализировал DNS 30 лет назад — он представлял, что система содержит до 50 миллионов адресов.



DNS

По мере распространения Интернета в Европе, Южной Америке и Азии все больше пользователей начали использовать DNS.

Правительство США наняло компанию Network Solutions Inc. (NSI) для бесплатного распространения общих доменов верхнего уровня (gTLD) , но впоследствии предоставило компании исключительное право взимать плату за регистрацию.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

DNS

Первые несколько лет работу Network Solutions финансировало Агентство оборонных информационных систем США (DISA).

В 1995 году другая правительственная структура — Национальный научный фонд Америки (NSF), отвечающий за развитие технологий, — разрешил NSI брать плату за регистрацию доменов.

Компания установила цену \$100 за каждые два года владения именем.

DNS

Первым альтернативным сервером стал AlterNIC, его в 1995 году основал инженер Евгений Кашпуреф (Eugene Kashpureff).

В рамках AlterNIC были запущены gTLD: .exp, .llc, .lnx, .ltd, .med и др.

Но Евгений пошел дальше простого создания нового сервиса.



A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

DNS

В 1997 году он взломал InterNIC и на три дня перенаправил трафик ресурса на собственный сайт.

Такие действия повлекли за собой несколько исков и уголовное преследование.

В итоге основателя AlterNIC арестовали, а компанию закрыли.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

DNS

Помимо AlterNIC, в девяностых появились такие корневые серверы, как eDNS, Iperdome для персональных доменных сервисов и New.Net, который был единственным корневым сервером, работавшим в западной части Тихого океана до 2002 года.

Основатели сервисов не выходили за рамки и не пытались составить конкуренцию Network Solutions, но проекты так или иначе были закрыты.

A vertical decorative bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

DNS

Альтернативные корневые сервисы продолжают распространяться, например, есть dot.love с такими доменами, как .thanks, .joy, .wise и .truth.

Также существует dotBERLIN, который управляет единственным доменом .berlin.

Свои DNS-серверы настраивают и участники открытого сообщества — GNU Name System работает с доменом .gnu, а децентрализованная система EmerCoin управляет .coin, .lib, .emc и .bazar.



DNS

Несмотря на распространение альтернативных корневых серверов, мировые эксперты считают, что их наличие вредит работе интернета.

Инженеры из IETF даже описали свои опасения в специальном RFC2826.

Они отмечают, что дробление DNS-экосистемы мешает работе резолверов, замедляет поиск необходимых ресурсов для пользователей и затрудняет распространение обновлений безопасности.

DNS



Есть решения, подобные Open Root Server Network, которые не ориентированы на получение прибыли и не расширяют корневые зоны дополнительными доменами верхнего уровня.

Их задача — снизить общую зависимость от ICANN, которая отвечает за gTLD.

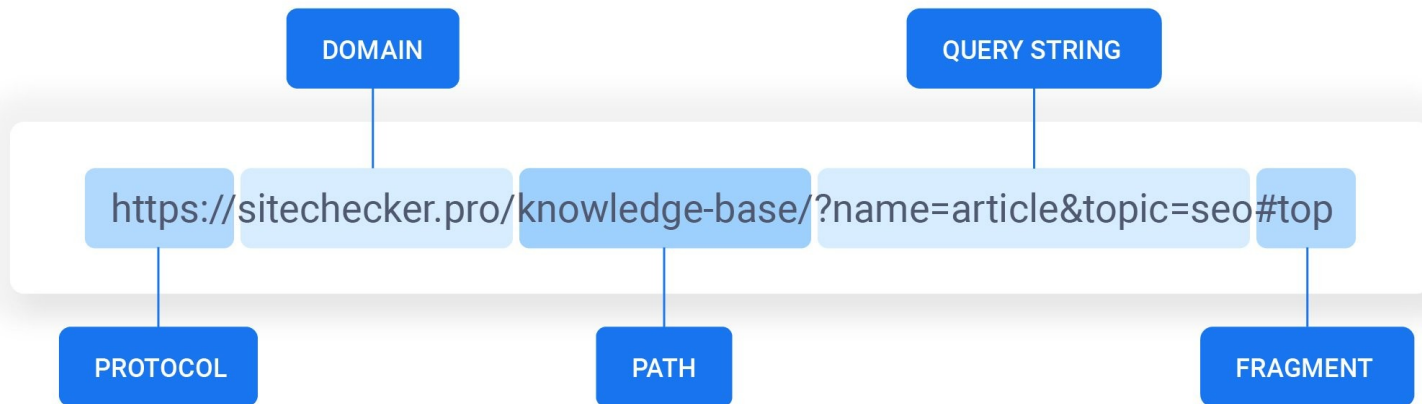


Адресация ресурсов

Для всех ресурсов которые доступны пользователям web-сети задаются единообразные указатели местоположения URL адреса (Uniform Resource Locator) которые помимо уникальной идентификации ресурсов предоставляют так же и информацию о их местоположении.

Адресация ресурсов

URL адреса ресурсов имеют следующую структуру:





Адресация ресурсов

URL'ы не должны были стать тем, чем стали: мудрёным способом идентифицировать сайт в интернете для пользователя.

К сожалению, мы не смогли стандартизировать URN, который мог бы стать более полезной системой наименования.

Считать, что современная система URL достаточно хороша — это как боготворить командную строку DOS и говорить, что все люди просто должны научиться пользоваться командной строкой.

Оконные интерфейсы были придуманы, чтобы пользоваться компьютерами стало проще, и чтобы сделать их популярнее. Такие же мысли должны привести нас к более хорошему методу определения сайтов в Вебе.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

Адресация ресурсов

Вспомним 1992 когда Тим Бернерс-Ли создал три штуки, благодаря которым родилось то, что мы считаем интернетом: протокол HTTP, HTML и URL.

Гипертекст, в двух словах — это возможность создавать документы, которые ссылаются друг на друга.

В те годы идея гипертекста считалась панацеей из научной фантастики, заодно с гипермедиа, и любыми другими словами с приставкой «гипер».



Адресация ресурсов

Ключевым требованием гипертекста была возможность ссылаться из одного документа на другой.

В то время для хранения документов использовалась куча форматов, а доступ осуществлялся по протоколу вроде Gopher или FTP.

В начальной презентации World-Wide Web в марте 1992 Тим Бернерс-Ли описал его как «универсальный идентификатор документов» (Universal Document Identifier или UDI).



Адресация ресурсов

В UDI избегают использование пробелов: пробелы — это запрещенные символы.

Это сделано потому, что часто появляются лишние пробелы когда строки обрабатываются системами вроде mail, или из-за обычной необходимости выровнять ширину колонки, а так же из-за преобразования различных видов пробелов во время конвертации кодов символов и при передаче текста от приложения к приложению.



Адресация ресурсов

Этот документ также объясняет, почему пробелы должны кодироваться в URL (%20)

Важно понимать, что URL был просто сокращенным способом обратиться к комбинации схемы, домена, порта, учетных данных и пути, которые ранее нужно было определять из контекста для каждой из систем коммуникации.



Адресация ресурсов

В 1993 году многие страстно верили, что URL отомрет, и на замену ему придет URN.

Uniform Resource Name — это постоянная ссылка на любой фрагмент, который, в отличие от URL, никогда не изменится и не сломается.

Тим Бернерс-Ли описал его как «срочную необходимость» еще в 1991.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles of various sizes and colors, including blue, yellow, orange, pink, and green.

Адресация ресурсов

Простейший способ создать URN — это использовать криптографический хэш содержания страницы, например:

`urn:791f0de3cfffc6ec7a0aacda2b147839`

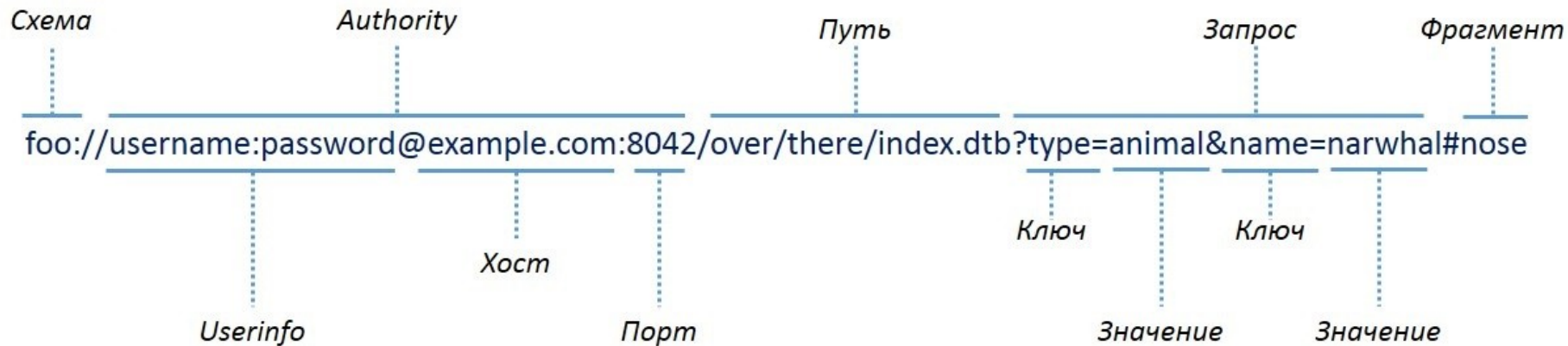


Адресация ресурсов

Однако, этот метод не удовлетворяет критериям веб-сообщества, так как невозможно выяснить, кто и как будет конвертировать этот хэш обратно в реальный контент.

Такой способ также не учитывает изменений формата, которые часто происходят в файле (например, сжатие файла), которые не влияют на

Адресация ресурсов



A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in various colors including blue, yellow, orange, pink, and green, arranged in a gradient-like fashion from top to bottom.

Адресация ресурсов

- схема – обращения к ресурсу, обычно это название используемого протокола (ftp, http)
- хост- доменное имя веб сайта в системе DNS или IP адрес компьютера на котором расположен сайт
- порт – порт хоста для подключения
- путь – уточняющая информация о месте нахождения ресурса в файловой системе сервера
- это может быть путь к конкретному файлу или каталогу если заканчивается символом «/»

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in various colors including blue, yellow, orange, pink, and green, arranged in a gradient-like fashion from top to bottom.

Адресация ресурсов

- имя ресурса – собственное имя ресурса в файловой системе сервера или условное имя по которому web сервер будет определять требуемый физический ресурс в процессе разрешения адреса.
- строка запроса – набор пар «параметр=значение» разделённых символом «&» обычно передаются в результате обработки формы web страницы
- ссылка (anchor) якорь – указатель на некоторый раздел web страницы



Протокол HTTP

Взаимодействие всех участников web сети основывается на использовании протокола передачи гипертекстов – HTTP(Hyper Text Transfer Protocol).

Это протокол прикладного уровня в стеке протоколов TCP/IP.

Текущей версией данного протокола является HTTP/3. По умолчанию используется порт 80.

Протокол HTTP

С помощью протокола HTTP web серверы обмениваются информацией, поэтому их часто называют HTTP серверами, а web браузеры HTTP клиентами.

Однако клиентами могут быть не только браузеры, но и любые программы которые могут использовать этот протокол (прокси серверы, поисковые агенты)



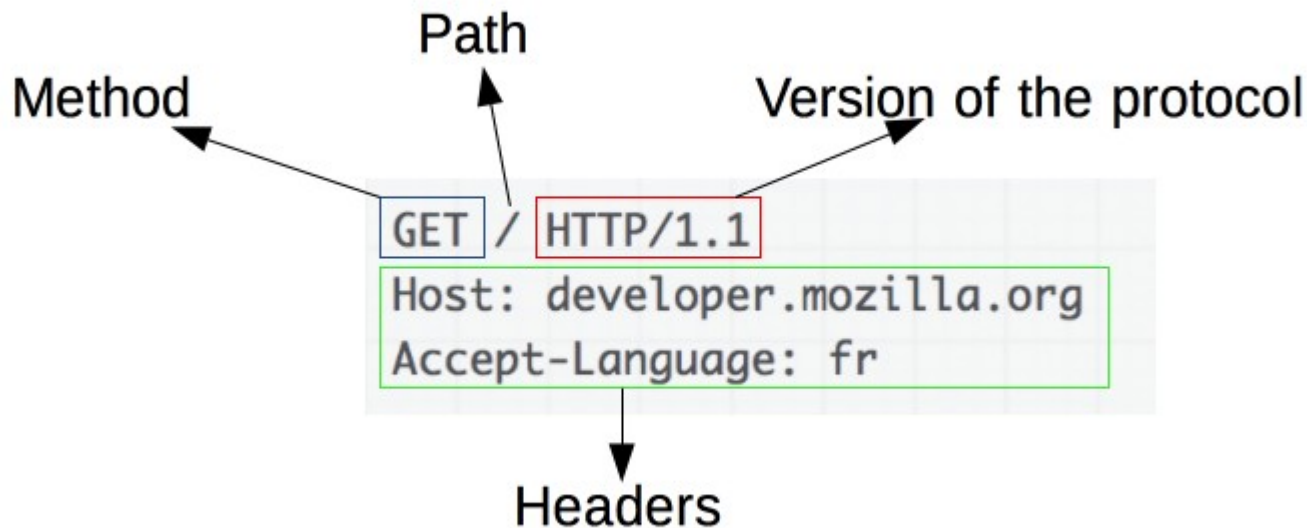
A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Протокол HTTP

Протокол HTTP использует принцип "запрос-ответ", означающий, что программа HTTP-клиент посылает HTTP-серверу сообщение (команду) вида "HTTP-запрос" (request), а сервер возвращает сообщение вида "HTTP-ответ" (response).

Протокол HTTP

HTTP-запрос имеет следующую структуру:



A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Протокол HTTP

Первая строка запроса состоит из трех полей:

- первое поле – МЕТОД – соответствует одному из поддерживаемых методов запроса (например, GET или POST);
- второе поле – /имя-ресурса – соответствует части URL-адреса, которая задает на web-сервере путь к запрашиваемому ресурсу;
- третье поле – номер-версии – соответствует версии протокола HTTP, который использует клиент.



Протокол HTTP

После первой строки запроса записывается список HTTP- заголовков (headers), за которыми следует пустая строка (символы "возврат каретки и завершение строки", "\r\n", коды 13 и 10), которая отделяет заголовки запроса от тела запроса.

После нее может быть (хотя и не обязательно) тело запроса, после которого следует другая пустая строка, указывающая на конец запроса.



Протокол HTTP

Например, если запрашиваемым URL-адресом является:

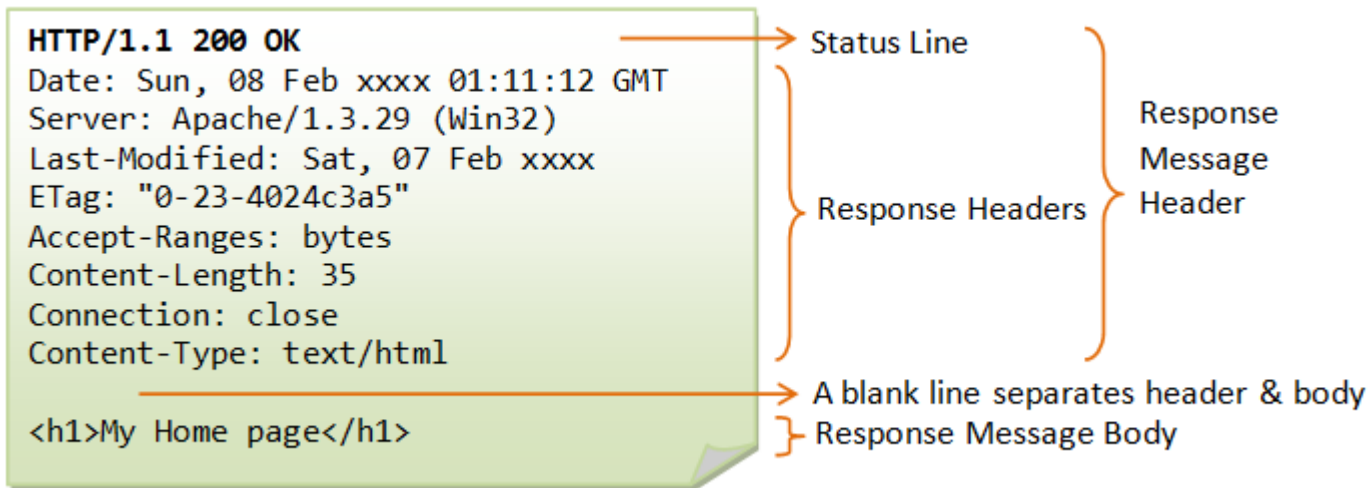
`mywebsite.ru/~ivanov/default.html`

то ниже показан упрощенный вариант HTTP-запроса, который будет передан зарегистрированному в DNS-системе web-сайту с именем `mywebsite.ru`:

`GET /~vanov/default.html HTTP/1.1 Host: mywebsite.ru`

Протокол HTTP

Сервер при получении такого запроса будет формировать HTTP- ответ, который имеет следующую структуру:





Протокол HTTP

Первая строка ответа (строка состояния) содержит версию HTTP- протокола, за которой следует код состояния, состоящий из 3-х цифр, и его краткое текстовое пояснение.

Имя-заголовок-1: значение

Имя-заголовок-2: значение

[тело ответа]



Протокол HTTP

Так как переданный в данном примере HTTP-запрос был успешно обработан, то строка состояния содержит код 200 и краткое его пояснение – OK.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, pink, and green, creating a vibrant, abstract pattern.

Протокол HTTP

В HTTP-ответ включены две строки заголовков, после которых идет пустая строка, за которой следует содержание запрашиваемого файла (тело ответа).

Процесс передачи запросов и ответов между браузерами и серверами не всегда является таким простым.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

Протокол HTTP

HTML-страница может содержать ссылки на другие доступные ресурсы, которые связаны с ней (изображения, скрипты, стили и т. п.).

Клиенты, которые выполняют показ изображений, выполнение скриптов или используют стили для представления страниц, должны выполнить разбор полученной HTML- страницы, чтобы определить, какие дополнительные ресурсы требуются для правильного отображения данной страницы, а затем формируют новые HTTP-запросы на получение этих дополнительных ресурсов.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in various colors including blue, yellow, orange, pink, and green, arranged in a gradient-like fashion from top to bottom.

Методы HTTP

В HTTP-протоколе определены такие методы запросов, как GET, HEAD, PUT, DELETE, TRACE, OPTIONS и CONNECT.

Наиболее часто используемыми методами являются GET и POST.

A decorative vertical bar on the left side of the slide, composed of numerous overlapping circles in shades of blue, yellow, orange, and pink, creating a vibrant, abstract pattern.

Метод GET

Метод GET является самым простым методом запроса.

Он используется при формировании запроса к web-серверу, когда пользователь вводит URL в адресную строку браузера, переходит по гиперссылке или выбирает одну из закладок на ранее посещаемые страницы.

Кроме этого, метод GET будет использоваться при отправке данных HTML-формы (в результате нажатия клавиши типа submit), у которой задан атрибут "method="GET".

Метод GET

GET vs. POST



Form Data, JSON Strings, Query Parameters, View States, etc.

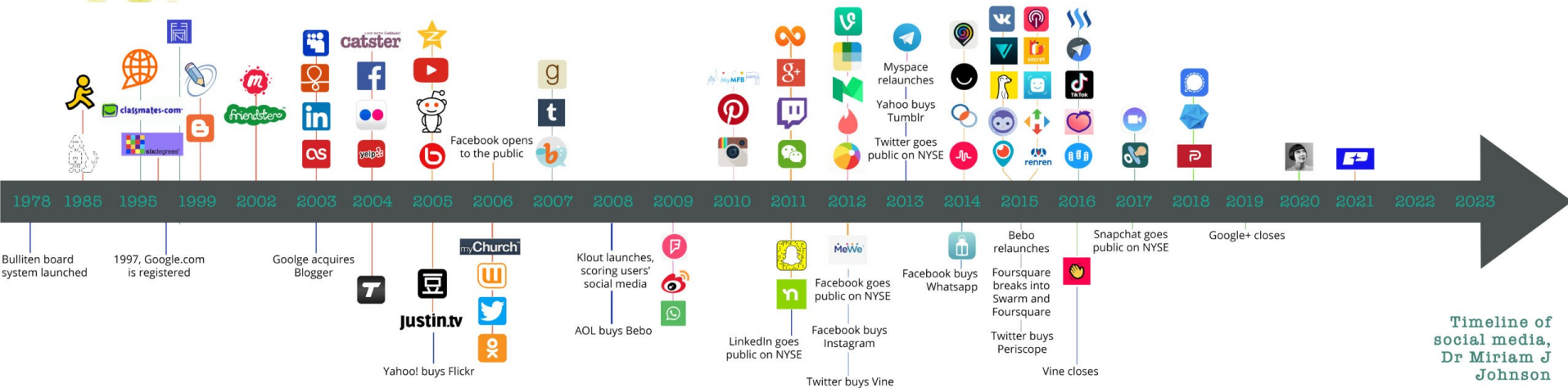
Метод GET

В данной HTML-форме описано поле fio, которому пользователь задает значение. В результате нажатия кнопки будет формироваться следующий HTTP-запрос:

```
GET /q?fio=YHOO HTTP/1.1 Host: finance.yahoo.com  
User-Agent: Mozilla/5.0 (Windows; U; Windows XP;  
en-US; rv:1.8.0.11)
```

Значение поля передается s в виде параметра ?
s=YHOO.

История интернета



Timeline of
social media,
Dr Miriam J
Johnson