

# РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ.

---

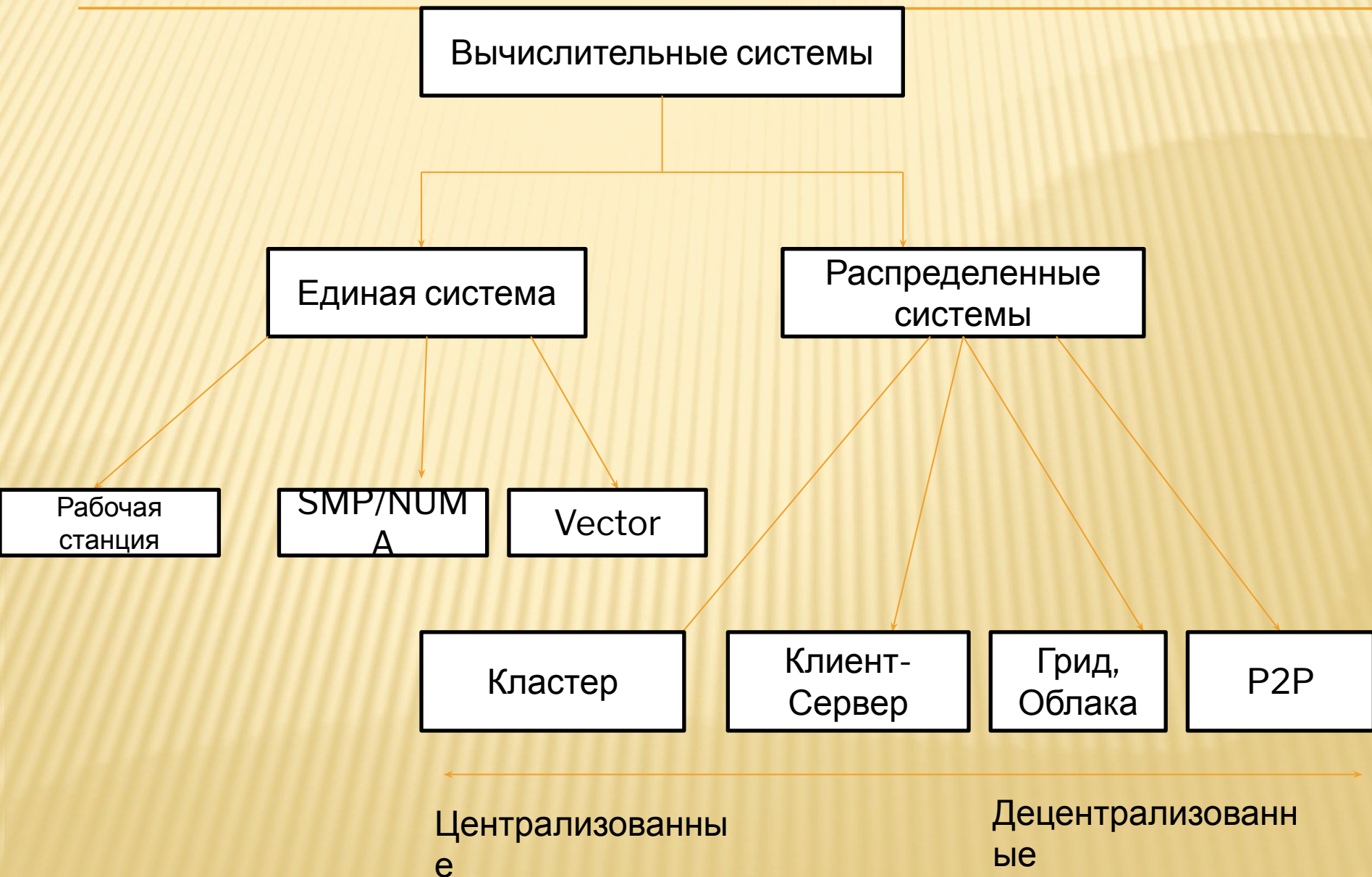
*«Распределенной вычислительной системой называется такая система, в которой отказ компьютера, о существовании которого ты даже не подозревал, может сделать твой собственный компьютер непригодным к использованию»*

*Лесли Лампорт,  
Microsoft Corporation*

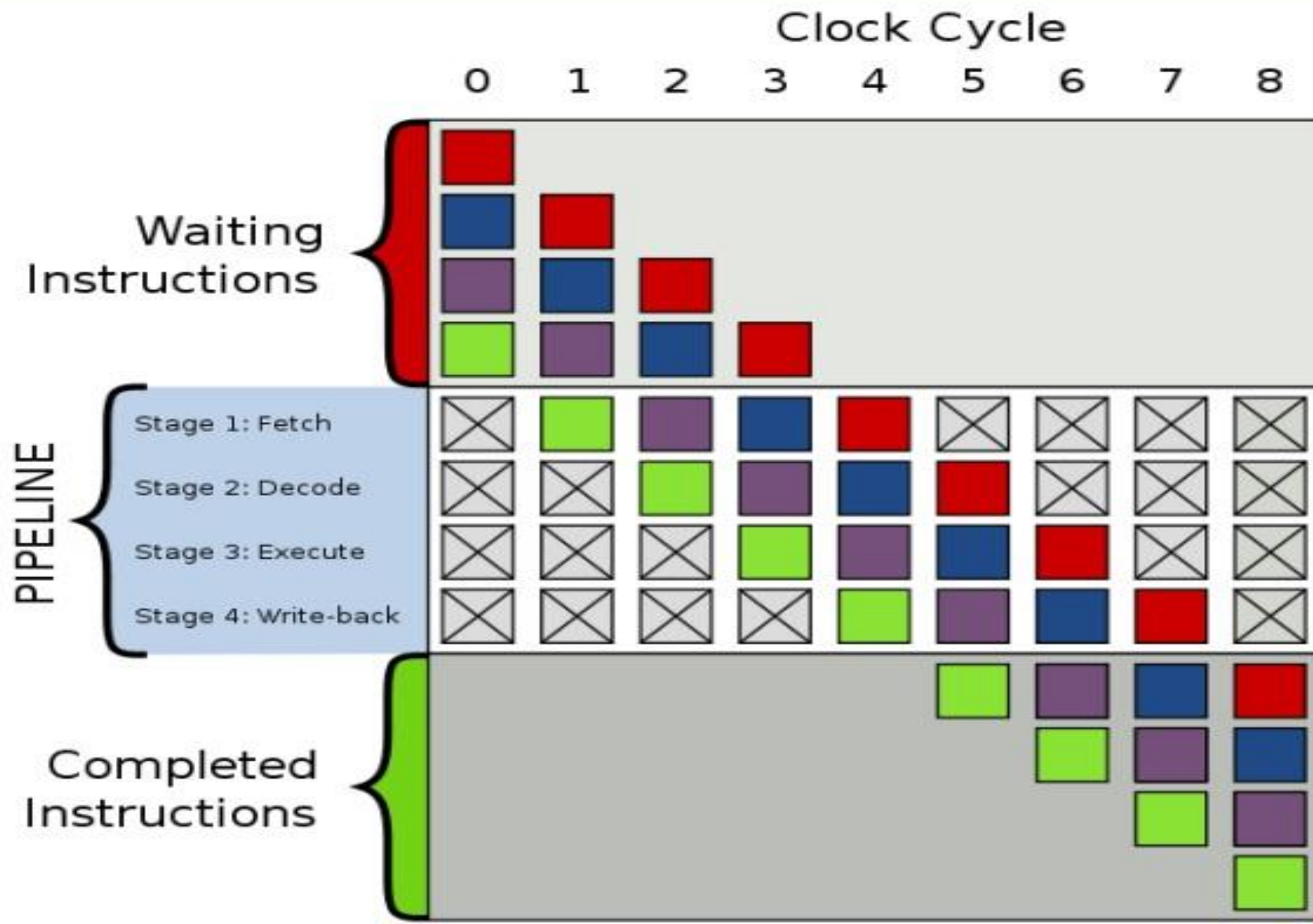
*«Распределенная вычислительная система (РВС) – это набор соединенных каналами связи независимых компьютеров, которые с точки зрения пользователя некоторого программного обеспечения выглядят единым целым»*

*Э. Таненбаум*

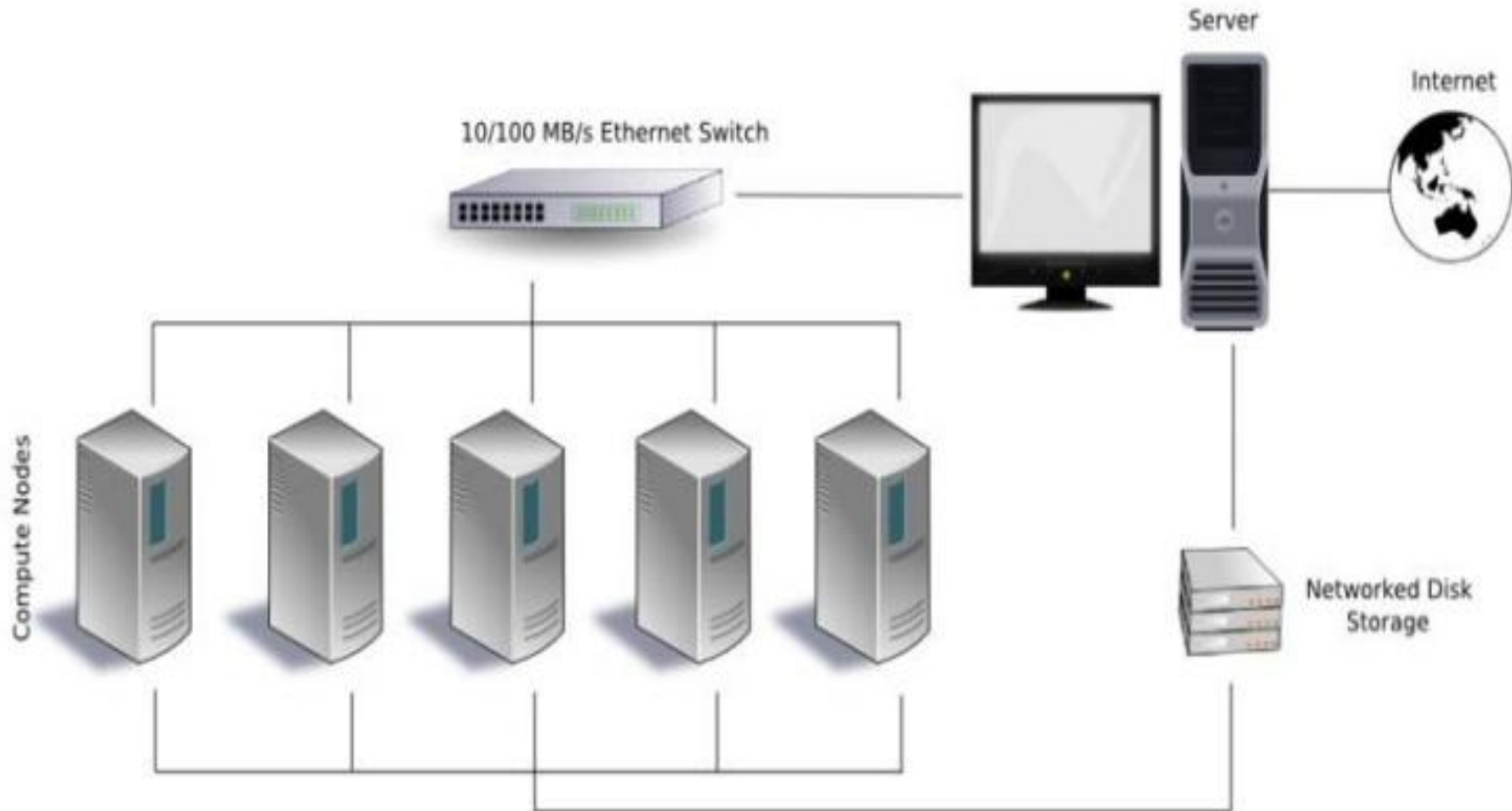
# ВИДЫ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ



# ВЕКТОРНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

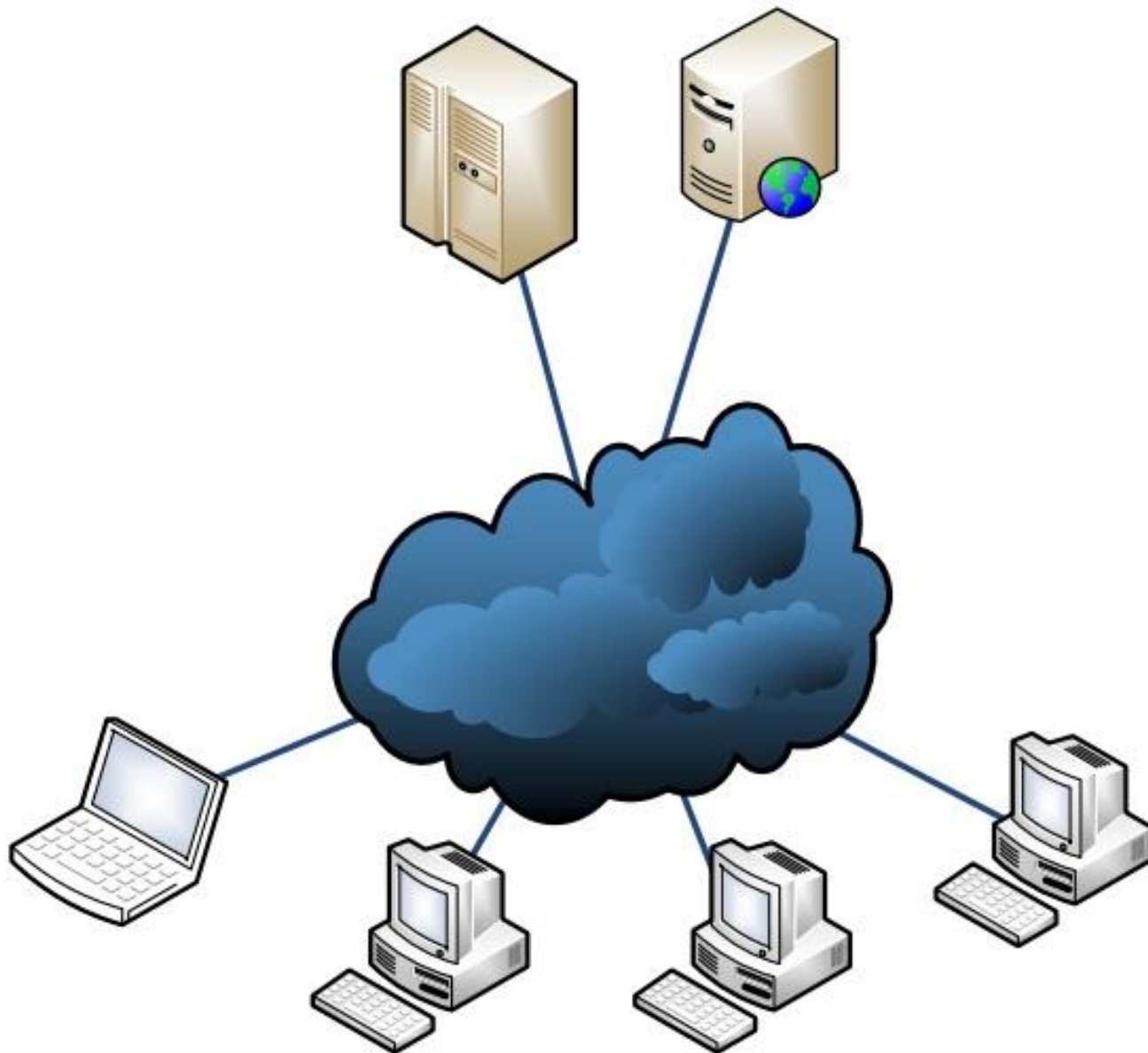


# КЛАСТЕРНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

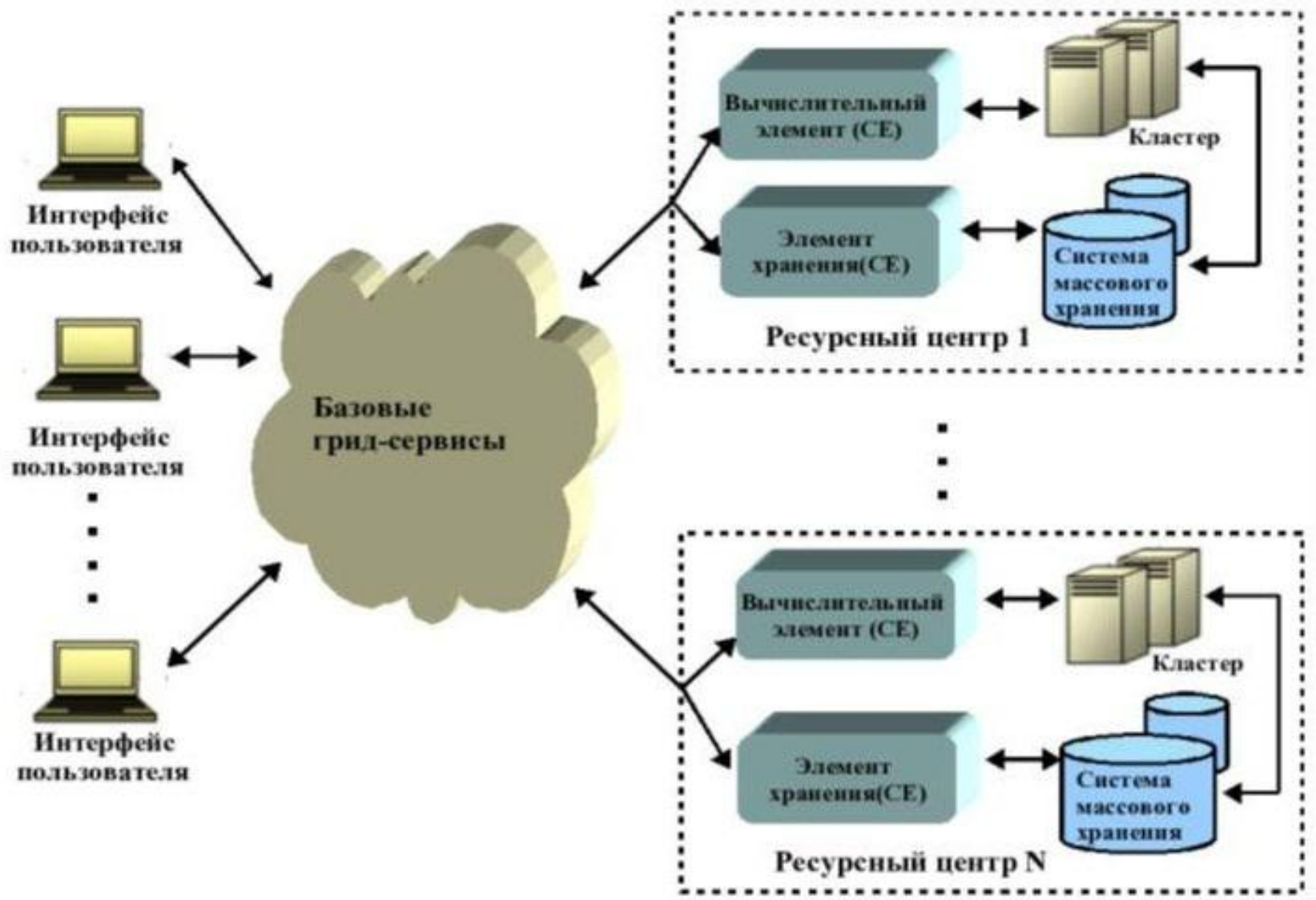




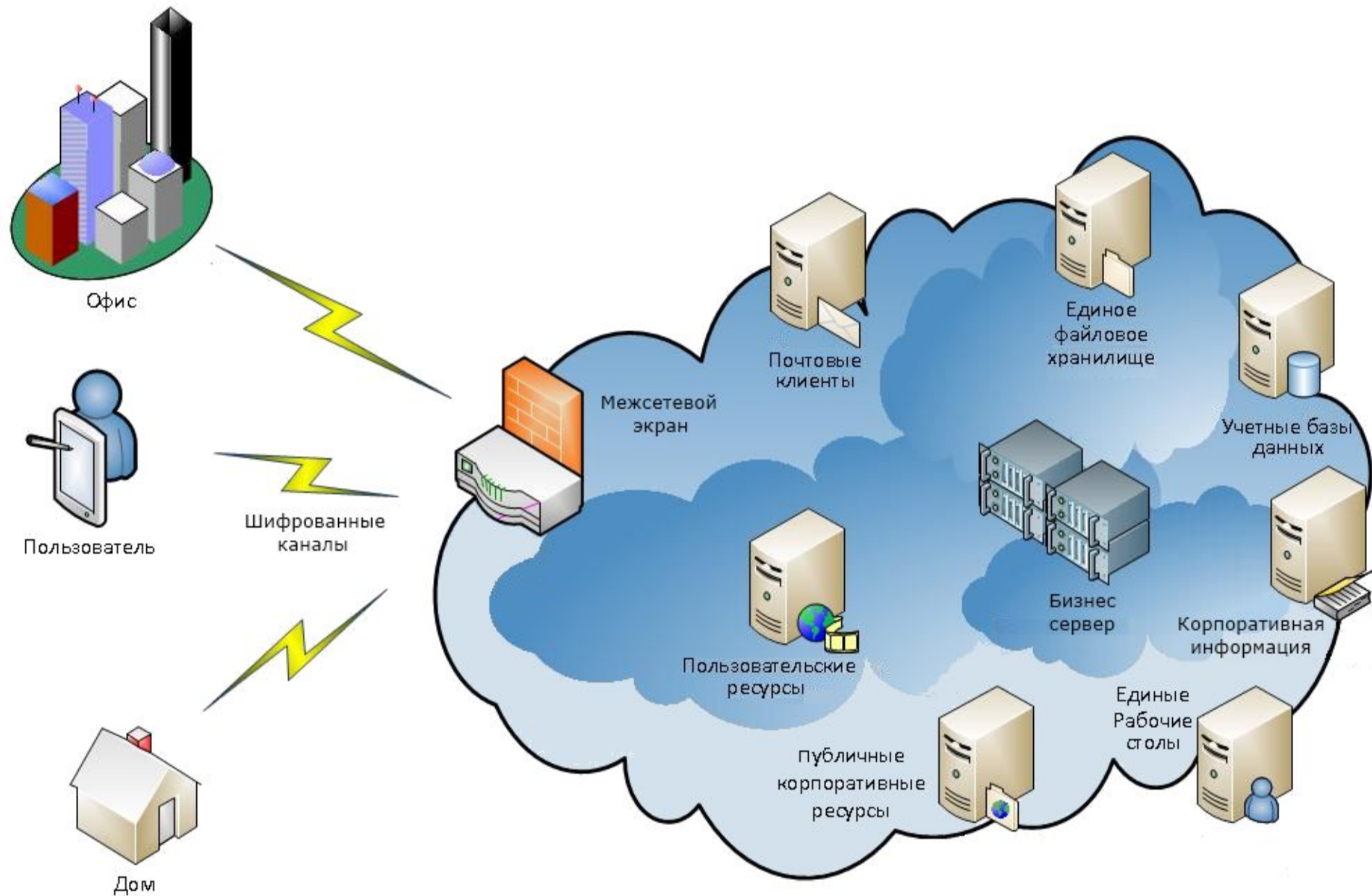
# КЛИЕНТ-СЕРВЕРНЫЕ СИСТЕМЫ



# Обобщенная схема структуры грида



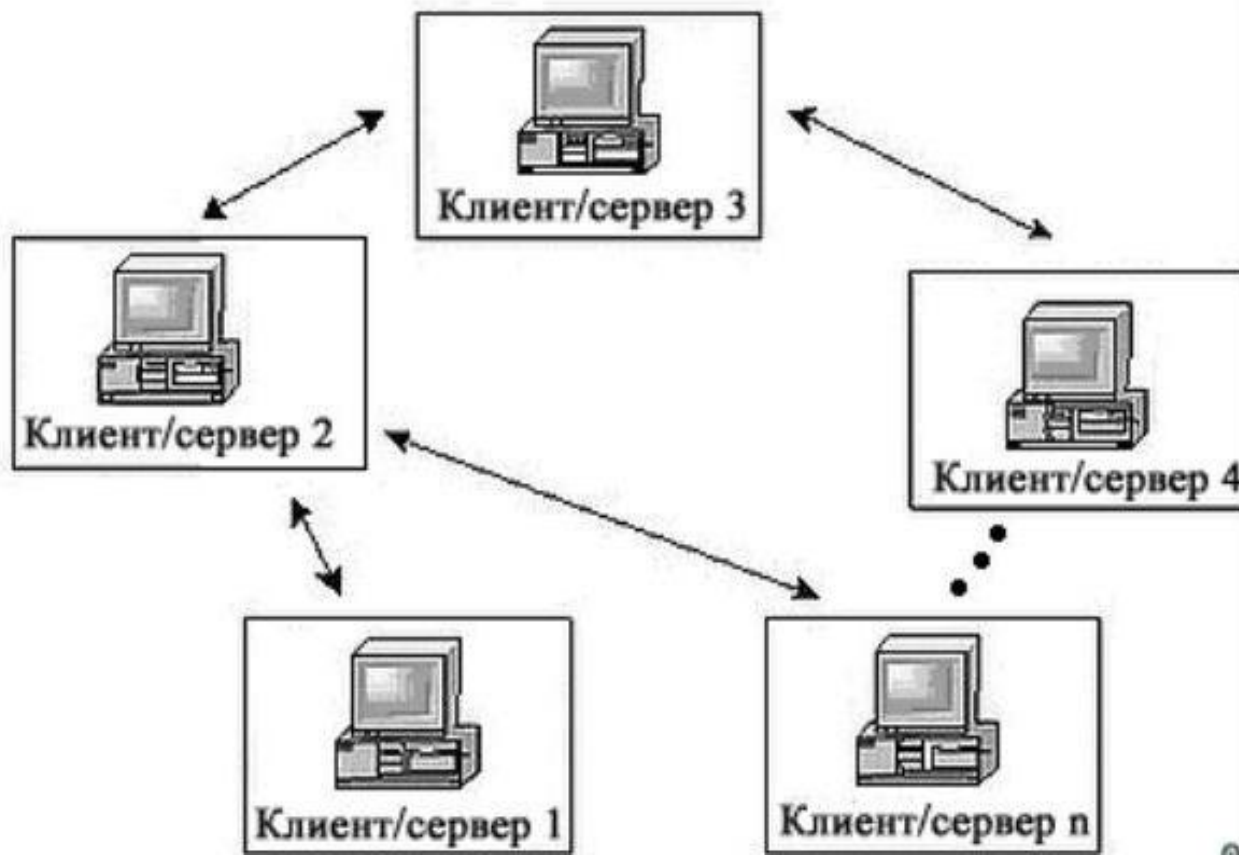
# Облачные системы



# ОДНОРАНГОВЫЕ(P2P)СИСТЕМЫ

## Архитектура P2P

Нет выделенного сервера, и компьютеры подсоединяются друг к другу, каждый компьютер может выполнять как серверные, так и клиентские функции. **Недостатки:** сложность процесса синхронизации между участниками сети, репликации ресурсов и управления пропускной способности.





# ЛОКАЛЬНЫЕ И ГЛОБАЛЬНЫЕ КОМПЬЮТЕРНЫЕ СЕТИ

**Локальная вычислительная сеть (ЛВС, локальная сеть; англ. Local Area Network, LAN)** — компьютерная сеть, покрывающая обычно относительно небольшую территорию или небольшую группу зданий (дом, офис, фирму, институт). Также существуют локальные сети, узлы которых разнесены географически на расстояния более 12 500 км (космические станции и орбитальные центры). Несмотря на такие расстояния, подобные сети всё равно относят к локальным.



**Глобальная компьютерная сеть, ГКС** (англ. *Wide Area Network, WAN*) — компьютерная сеть, охватывающая большие территории и включающая в себя большое число компьютеров.

ГКС служат для объединения разрозненных сетей так, чтобы пользователи и компьютеры, где бы они ни находились, могли взаимодействовать со всеми остальными участниками глобальной сети.

Некоторые ГКС построены исключительно для частных организаций, другие являются средством коммуникации корпоративных ЛВС с сетью Интернет или посредством Интернет с удалёнными сетями, входящими в состав корпоративных. Чаще всего ГКС опирается на выделенные линии, на одном конце которых маршрутизатор подключается





# ОТЛИЧИЕ ГЛОБАЛЬНОЙ СЕТИ ОТ ЛОКАЛЬНОЙ

---

Глобальные сети отличаются от локальных тем, что рассчитаны на неограниченное число абонентов и используют, как правило, не слишком качественные каналы связи и сравнительно низкую скорость передачи, а механизм управления обменом у них в принципе не может быть гарантировано скорым.

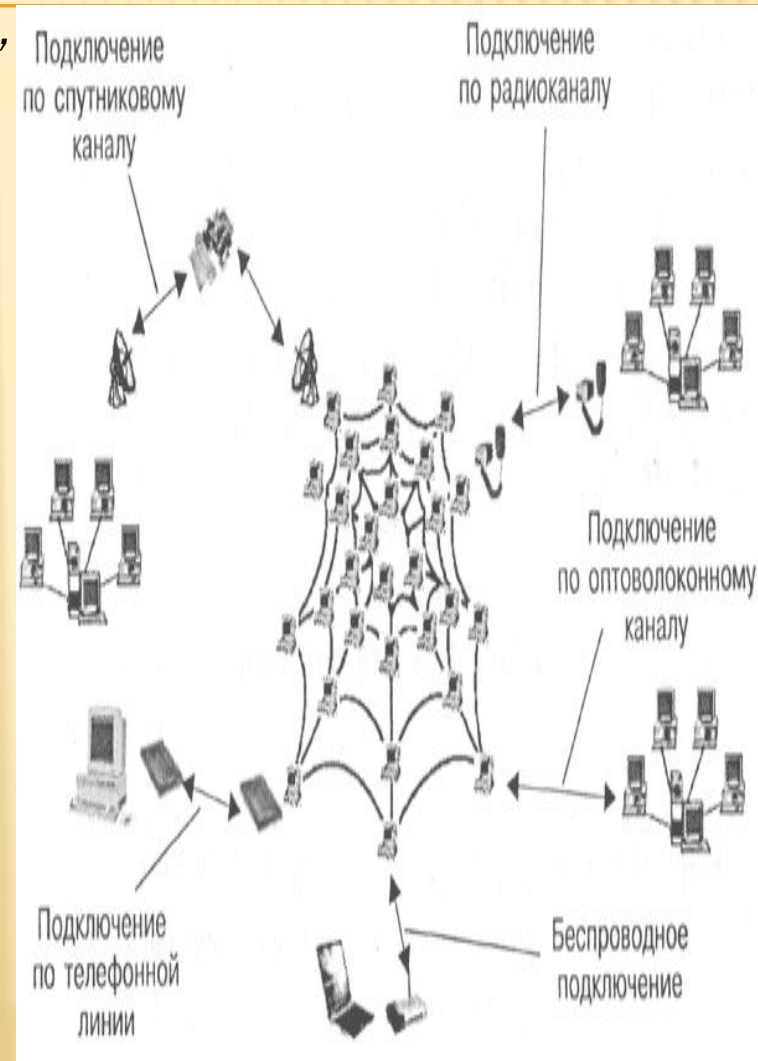
В глобальных сетях намного более важно не качество связи, а сам факт ее существования. Правда, в настоящий момент уже нельзя провести четкий и однозначный предел между локальными и глобальными сетями. Большинство локальных сетей имеют выход в глобальную сеть, но характер переданной информации, принципы организации обмена, режимы доступа к ресурсам внутри локальной сети, как правило, сильно отличаются от тех, что приняты в глобальной сети. И хотя все компьютеры локальной сети в данном случае включены также и в глобальную сеть, специфику локальной сети это не отменяет. Возможность выхода в глобальную сеть остается всего лишь одним из ресурсов, поделенным пользователями локальной сети.

# ГЛОБАЛЬНАЯ КОМПЬЮТЕРНАЯ СЕТЬ ИНТЕРНЕТ

*Интернет* - это глобальная компьютерная сеть, в которой локальные, региональные и корпоративные сети соединены между собой многочисленными каналами передачи информации с высокой пропускной способностью.

**Подключение к Интернету.** В каждой локальной, региональной или корпоративной сети имеется, по крайней мере, один компьютер (сервер Интернета), который имеет постоянное подключение к Интернету.

Для подключения локальных сетей чаще всего используются **оптоволоконные линии** связи. Однако в случаях подключения неудобно расположенных или удаленных компьютерных сетей, когда прокладка кабеля затруднена или невозможна, используются беспроводные линии связи. Если передающая и принимающая антенны находятся в пределах прямой видимости, то используются **радиоканалы**, в противном случае обмен информацией производится через **спутниковый канал** с использованием





# МАРШРУТИЗАЦИЯ И ТРАНСПОРТИРОВКА ДАННЫХ ПО КОМПЬЮТЕРНЫМ СЕТЯМ

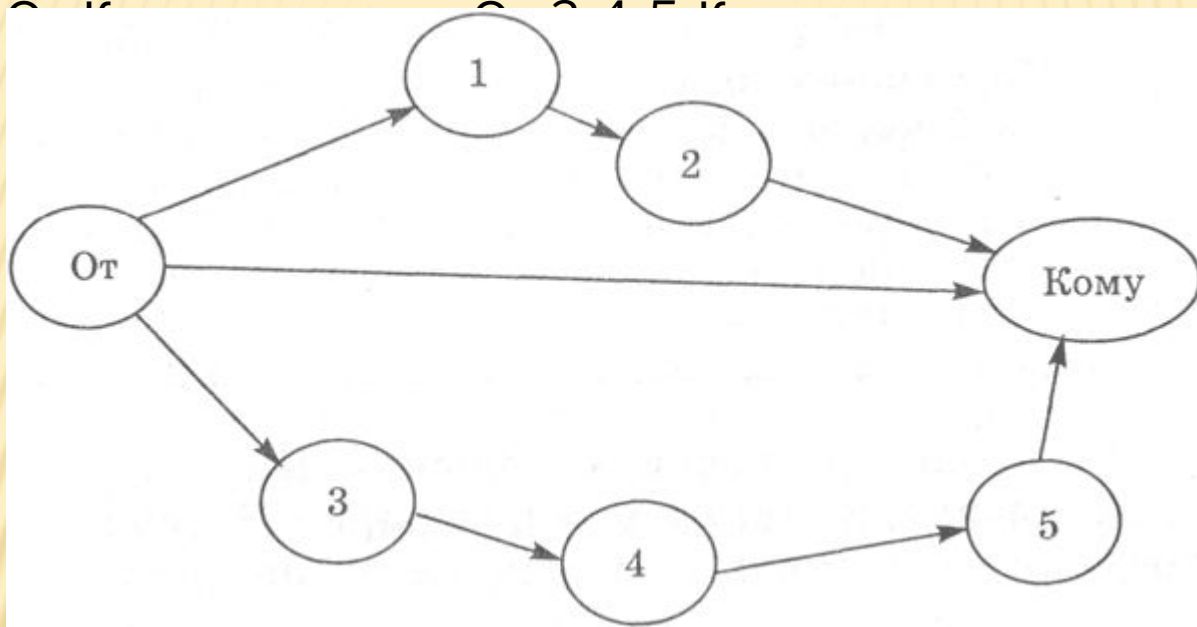
Сеть Интернет, являющаяся сетью сетей и объединяющая громадное количество различных локальных, региональных и корпоративных сетей, функционирует и развивается благодаря использованию единого принципа маршрутизации и транспортировки данных.

**Маршрутизация данных.** Маршрутизация данных обеспечивает передачу информации между компьютерами сети. Рассмотрим принцип маршрутизации данных по аналогии с передачей информации с помощью обычной почты. Для того чтобы письмо дошло по назначению, на конверте указывается адрес получателя (кому письмо) и адрес отправителя (от кого письмо).

Аналогично, передаваемая по сети информация "упаковывается в конверт", на котором "пишутся" Интернет-адреса компьютеров получателя и отправителя, например: "Кому: 198.78.213.185", "От кого: 193.124.5.33". Содержимое конверта на компьютерном языке называется Интернет-пакетом и представляет собой набор байтов.

В процессе пересылки обыкновенных писем они сначала доставляются на ближайшее к отправителю почтовое отделение, а затем передаются по цепочке почтовых отделений на ближайшее к получателю почтовое отделение. На промежуточных почтовых отделениях письма сортируются, т. е. определяется, на какое следующее почтовое отделение необходимо отправить то или иное письмо.

Маршруты доставки Интернет-пакетов могут быть совершенно разными, и поэтому первые Интернет-пакеты могут достичь компьютера-получателя в последнюю очередь. Например, в процессе передачи файла от сервера От к серверу Кому маршрут первого Интернет-пакета может быть От-1-2-Кому, второго -



**Транспортировка данных** производится путем разбиения файлов на Интернет-пакеты на компьютере-отправителе, индивидуальной маршрутизации каждого пакета и сборки файлов из пакетов в первоначальном порядке на компьютере-получателе.

Время транспортировки отдельных Интернет-пакетов между локальным компьютером и сервером Интернета можно определить с помощью специальных программ.

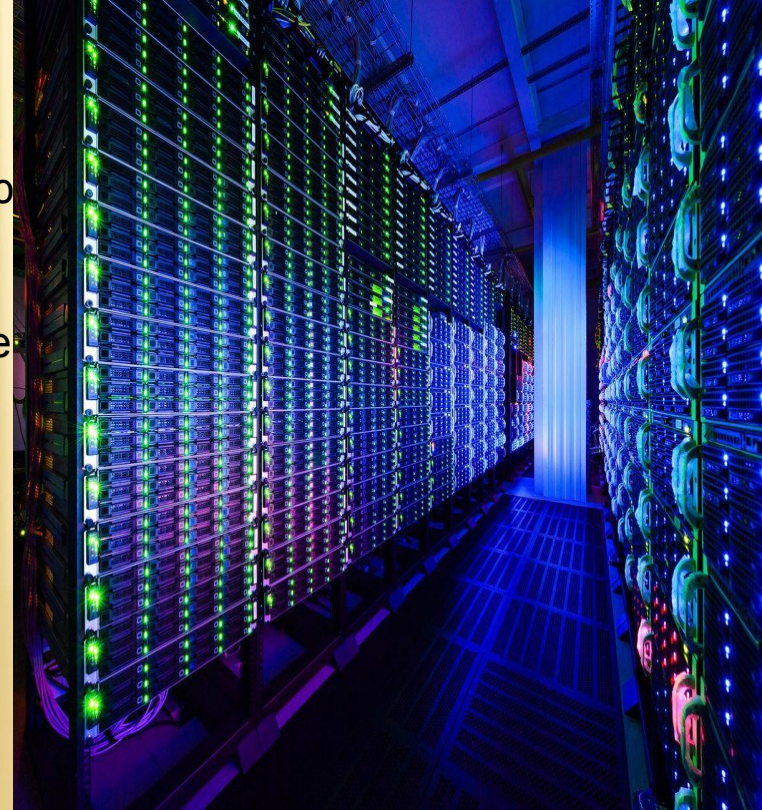
Маршрутизация и транспортировка данных в Интернете производится на основе протокола TCP/IP, который является основным "законом" Интернета. Термин "TCP/IP" включает название двух протоколов передачи данных:

- TCP (Transmission Control Protocol - транспортный протокол);
- IP (Internet Protocol - протокол маршрутизации).



# ДАТА-ЦЕНТРЫ

**Дата-центр** (от англ. *data center*), или **центр (хранения и) обработки данных (ЦОД/ЦХОД)** — это специализированное здание для размещения (хостинга) серверного и сетевого оборудования и подключения абонентов к каналам сети Интернет. Дата-центр исполняет функции обработки, хранения и распространения информации, как правило, в интересах корпоративных клиентов — он ориентирован на решение бизнес-задач путём предоставления информационных услуг. Консолидация вычислительных ресурсов и средств хранения данных в ЦОД позволяет сократить совокупную стоимость владения IT-инфраструктурой за счёт возможности эффективного использования технических средств, например, перераспределения нагрузок, а также за счёт сокращения расходов на администрирование. Дата-центры обычно расположены в пределах или в непосредственной близости от узла связи или точки присутствия какого-либо одного или нескольких операторов связи. Качество и пропускная способность каналов влияют на уровень предоставляемых услуг, поскольку основным критерием оценки качества работы любого дата-центра является время доступности сервера (аптайм)





# Структура

Типичный дата-центр состоит из:

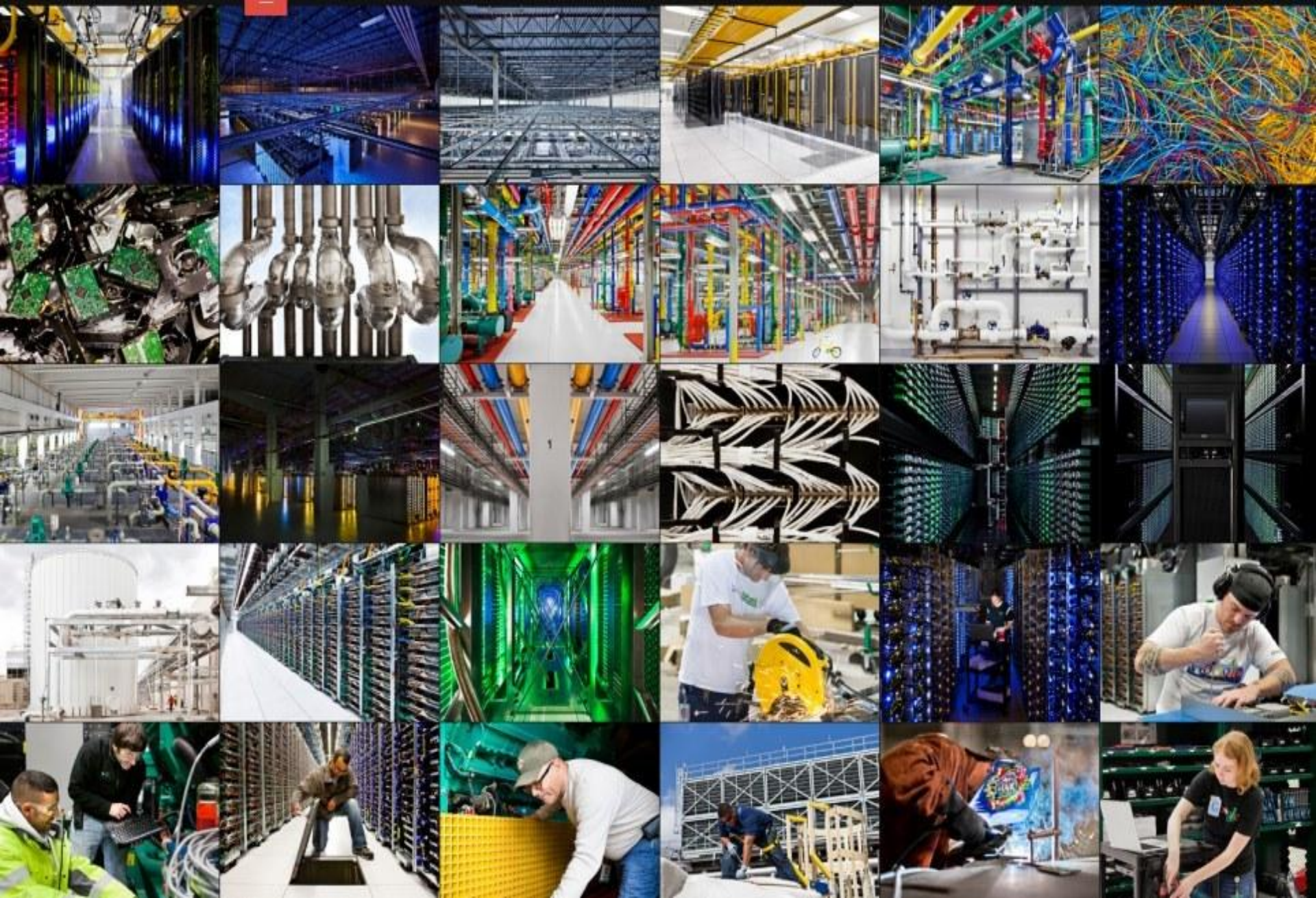
информационной инфраструктуры, включающей в себя серверное оборудование и обеспечивающей основные функции дата-центра — обработку и хранение информации; телекоммуникационной инфраструктуры, обеспечивающей взаимосвязь элементов дата-центра, а также передачу данных между дата-центром и пользователями; инженерной инфраструктуры, обеспечивающей нормальное функционирование основных систем дата-центра.

Инженерная инфраструктура включает в себя: кондиционирование для поддержания температуры и уровня влажности в заданных параметрах; бесперебойное электроснабжение для автономной работы дата-центра в случаях отключения центральных источников электроэнергии; охранно-пожарную сигнализацию и система газового пожаротушения; системы удаленного IP контроля, управления питанием и контроля доступа.

Некоторые дата-центры предлагают клиентам дополнительные услуги по использованию оборудования по автоматическому уходу от различных видов атак. Команды квалифицированных специалистов круглосуточно производят мониторинг всех серверов. Необходимо отметить, что услуги дата-центров сильно отличаются в цене и количестве услуг. Для обеспечения сохранности данных используются системы резервного копирования. Для предотвращения кражи данных, в дата-центрах используются различные системы ограничения физического доступа, системы видеонаблюдения. В корпоративных (ведомственных) дата-центрах обычно сосредоточено большинство серверов соответствующей организации. Оборудование крепится в специализированных стойках и шкафах. Как правило, в дата-центр принимают для размещения лишь оборудование в стоечном исполнении, то есть в корпусах стандартных размеров, приспособленных для крепления в стойку. Компьютеры в корпусах настольного исполнения неудобны для дата-центров и размещаются в них редко.



# Дата-центры





# Услуги дата-центров

Виртуальный хостинг. Крупные дата-центры обычно не предоставляют подобную массовую услугу из-за необходимости обеспечения техническо-консультационной поддержки.

Виртуальный сервер. Предоставление гарантированной и лимитированной части сервера (части всех ресурсов). Важная особенность данного вида хостинга — разделение сервера на несколько виртуальных независимых серверов, реализуемых программным способом.

Выделенный сервер. Дата-центр предоставляет клиенту в аренду сервер в различной конфигурации. Крупные дата-центры в основном специализируются именно на подобных типах услуг.

Colocation. Размещение сервера клиента на площадке дата-центра за определённую плату. Стоимость зависит от энергопотребления и тепловыделения размещаемого оборудования, пропускной способности подключаемого к оборудованию канала передачи данных, а также размера и веса стойки.

Аренда телекоммуникационных стоек. Передача клиенту стоек для монтажа собственного или клиентского оборудования. Формально это частный случай colocation, но с основным отличием в том, что арендаторы в основном юридические лица.

Выделенная зона (Dedicated area). В некоторых случаях владельцы дата-центра выделяют часть технологических площадей для специальных клиентов, как правило, финансовых компаний, имеющих строгие внутренние нормы безопасности. В этом случае дата-центр предоставляет некую выделенную зону, обеспеченную каналами связи, электроснабжением, холодоснабжением и системами безопасности, а клиент сам создает свой дата-центр внутри этого пространства.

# «ОБЛАКО»

**Облачное хранилище данных** (англ. *cloud storage*) — модель онлайн-хранилища, в котором данные хранятся на многочисленных распределённых в сети серверах, предоставляемых в пользование клиентам, в основном, третьей стороной. В отличие от модели хранения данных на собственных выделенных серверах, приобретаемых или арендуемых специально для подобных целей, количество или какая-либо внутренняя структура серверов клиенту, в общем случае, не видна. Данные хранятся и обрабатываются в так называемом *облаке*, которое представляет собой, с точки зрения клиента, один большой виртуальный сервер. Физически же такие серверы могут располагаться удалённо друг от друга географически, вплоть до расположения на разных континентах.



Нужные файлы всегда под рукой,  
с компьютера или телефона



# ПРЕИМУЩЕСТВА ОБЛАЧНЫХ ХРАНИЛИЩ

1. Клиент платит только за то место в хранилище, которое фактически использует, но не за аренду сервера, все ресурсы которого он может и не использовать.

2. Клиенту нет необходимости заниматься приобретением, поддержкой и обслуживанием собственной инфраструктуры по хранению данных, что, в конечном счёте, уменьшает общие издержки производства.

3. Все процедуры по резервированию и сохранению целостности данных производятся провайдером облачного центра, который не вовлекает в этот процесс клиента.





# 10 САМЫХ БОЛЬШИХ БЕСПЛАТНЫХ ОБЛАЧНЫХ ХРАНИЛИЩ ДАННЫХ.

**1 место** - Облако Mail.ru. Максимальный размер и минимальное доверие



В этом году Mail.ru Group предоставил пользователям Интернет практически уникальную возможность: 100 гигабайт в бесплатном облачном хранилище данных!

Все необходимые возможности современного облачного хранилища у Облака Mail.ru есть: мобильные клиенты, Web доступ и прочее. При этом, программное обеспечение Mail.ru Group печально известно скрытыми установками своих тулбаров, агентов и других не очень то нужных программ. Поэтому лучше с ним не связываться.

# 2 МЕСТО - ОБЛАЧНОЕ ХРАНИЛИЩЕ MEGA. 50 ГБАЙТ ЗА ПРОСТО ТАК!

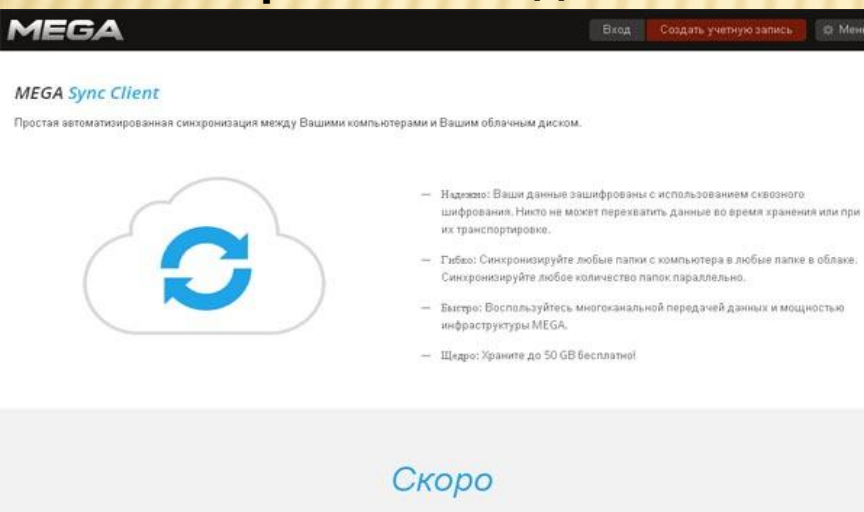


Компания *MEGA Encrypted Global Access* уже достаточно давно проводила бета-тестирование своего облачного хранилища данных и относительно недавно открыло его всем желающим.

У облачного хранилища данных Mega есть два основных плюса:

1. Зарегистрированному пользователю даётся сразу 50 Гбайт облачного пространства! Причём без любых "квестов" с получением дополнительных гигабайт путём привлечения новых пользователей, чем страдают некоторые другие его конкуренты. Согласитесь, что это очень даже прилично.
2. Программа-клиент этого облачного хранилища производит надёжное шифрование данных непосредственно на вашем устройстве и передаёт в "облако" уже зашифрованные данные. Таким образом, Mega не в курсе какие файлы вы в него положили!

**Однако у Mega есть один очень серьёзный недостаток:**



у Mega нет официального клиента для Windows. С ним можно работать либо через браузер, либо с устройств на iOS и Android. А в Windows у вас не будет автоматической синхронизации ваших файлов. Это весьма серьёзный и надеюсь временный недостаток этого облачного хранилища данных...



## 3 МЕСТО - ОБЛАЧНОЕ ХРАНИЛИЩЕ ДАННЫХ MEDIAFIRE. 50 ГБАЙТ В ОБЛАКЕ ПО КЛАССИЧЕСКОЙ РЕФЕРАЛЬНОЙ ПРОГРАММЕ



По сути MediaFire - это более современный DropBox. Все необходимые атрибуты нормального облачного хранилища данных - присутствуют.

Основное отличие: дают сразу 10 Гбайт и достаточно просто можно получить ещё 8 Гбайт. Используется классический алгоритм: дают дополнительное место за установку Windows и мобильных клиентов, за ссылку в социальных сетях и т.п.

Ещё до 32 Гб можно получить по реферальной программе. Дается по 1 Гбайт за каждого клиента, пришедшего к ним по вашей реферальной ссылке.

## **4 МЕСТО - ОБЛАЧНОЕ ХРАНИЛИЩЕ ДАННЫХ SKYDRIVE. СПЕЦИАЛЬНО ДЛЯ WINDOWS 8**



SkyDrive

Это облачное хранилище данных от Microsoft. Поэтому главный его плюс - это тесная интеграция с операционными системами их производства и некоторыми их программами.

Особенно это касается новых продуктов типа Windows 8 и 8.1, Office 2013, а также Office 365 и Outlook.com.

Обладатели лицензии Windows 8 и 8.1 сразу получают 25 Гб в SkyDrive. Остальным дают намного меньше - 7 Гбайт.



## 5 МЕСТО - ОБЛАЧНОЕ ХРАНИЛИЩЕ ДАННЫХ СОРУ. ОБЛАКО С ИНТЕРЕСНОЙ РЕФЕРАЛЬНОЙ ПРОГРАММОЙ



Облачное хранилище данных Сору - это практически классическое "облако", со всем необходимым функционалом.

Поддерживаются все необходимые платформы, а также есть шифрование ваших файлов "на лету" по надёжному алгоритму AES 256.

Сразу даётся 15 Гбайт облачного дискового пространства, что уже весьма неплохо. Также у облачного хранилища данных Сору есть очень интересная реферальная программа. За эту программу его очень любят блоггеры и спамеры.

## 6 место - облачное хранилище данных 4Sync.

### 15 Гбайт без условий!

Облачное хранилище данных 4Sync - это ещё одно классическое, не очень известное массовому пользователю, "облако".

Позиционируется как простое облачное хранилище данных, максимально удобное пользователю. В принципе, это так и есть. Интерфейсы интуитивно понятны, графика на сайте прикольная... Есть весь необходимый функционал, можно делиться своими файлами и использовать Windows и мобильные клиенты для синхронизации файлов. Причём поддерживаются не только обязательные iOS и Adnroid, но и Symbian и Blackberry.



## 7 место - Google Диск. Просто

### ДИСК

Как следует из названия Google Диск - это облачное хранилище данных от компании Google. Понятное дело, главный его плюс - это тесная интеграция с другими онлайн-продуктами Google типа Gmail и Docs. Есть офлайн режим, ограничение доступа к файлам, мобильные клиенты для Android и iOS, а также другие интересные функции...

Размер "облака" от Google - 15 Гбайт.

Минусы тоже есть. Во первых, данные не шифруются.

Во вторых, есть технические ограничения по работе с прокси-серверами. В некоторых случаях Google Диск просто не будет работать.

## 8 место - Яндекс Диск. Отечественное облако

У нашей отечественной компании Яндекс есть своё облачное хранилище данных - Яндекс.Диск. Так сказать отечественный аналог Google Диска, но с интеграцией с другими продуктами Яндекс. Это хорошее облако, со всем необходимым функционалом, включая мобильные клиенты под iOS, Android, Windows Phone, а также для Windows, OS X и даже Linux!

Сначала дают всего 3 Гбайт, но несложными действиями можно расширить размер этого облака до 10 Гбайт. Дальнейшее расширение возможно за деньги.





## 9 место - облачное хранилище Dropbox.

### Аутсайдер по размеру облака



Старичок Dropbox является основателем направления облачных хранилищ данных. Он, можно сказать, стандарт де-факто.

Однако, он тоже не без недостатков. У многих его конкурентов есть более интересные функциональные возможности, а также практически у всех более большой размер предоставляемого дискового пространства в "облаке".

Dropbox изначально даёт всего 2 Гбайт, достаточно легко можно получить ещё 1 дополнительный Гбайт.

Остальное либо за деньги либо по реферальной программе. Если новый пользователь пришёл по реферальной ссылке, зарегистрировался и установил их программу-клиент, то он и вы получаете по 0,5 Гбайт дополнительного места.

### 10 место - всё остальное

Есть и другие облачные хранилища данных, но все они дают достаточно мало места своим пользователям и для некоторых вариантов использования "облаков" просто не подходят.

Например, Wuala (5 Гбайт) или SugarSync (5 Гбайт + реферальная программа).



# ИНФОРМАЦИОННО-ПОИСКОВЫЕ СИСТЕМЫ

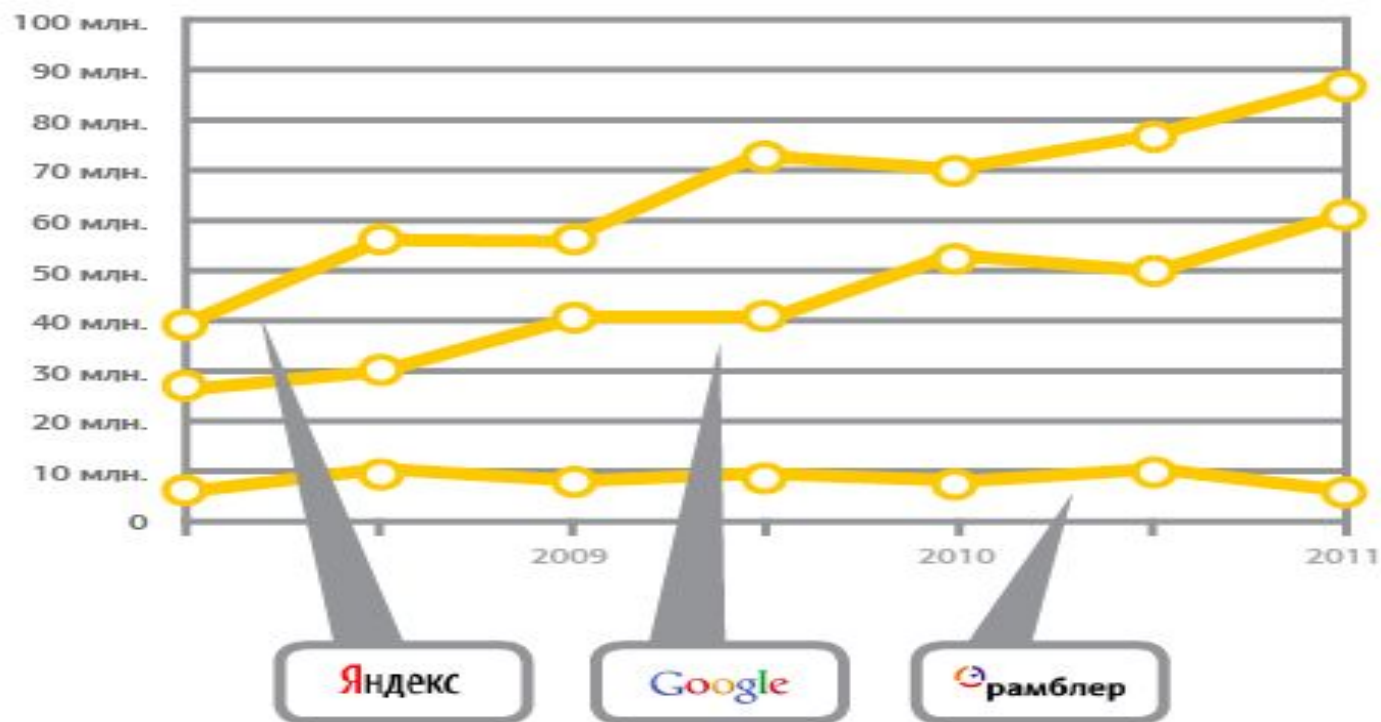
**Поисковая система** — это программное обеспечение, состоящее из базы данных документов, снабженной пользовательским интерфейсом, которое разработано для поиска информации во всемирной паутине. Пользуясь поисковой системой, пользователи составляют свои собственные поисковые запросы с ключевыми словами, а не используют заданные системой пути поиска.



Результаты поиска обычно представлены в серии результатов, часто называемых страницами результатов поисковой системы. Информация может быть совокупностью веб-страниц, изображений и других типов файлов. Основная задача поисковой системы заключается в выборе наилучшего возможного подмножества в ответ на конкретный запрос, то есть множества документов, которые наиболее соответствуют тому, что ищет пользователь. Полноценность поисковой системы зависит от релевантности результатов, которые

По состоянию на 2014 год самой популярной поисковой системой в мире является Google, однако есть страны, где пользователи отдали предпочтение другим поисковикам. Так, например, в России Яндекс обгоняет Google больше, чем на 10 %.

Поисковые системы созданы, чтобы находить и выдавать пользователю релевантную информацию, но есть понятие пузырь фильтров. Оно объясняет почему пользователь может не получить некоторую информацию. Так же стоит учитывать факт влияния различных факторов, таких как политические, экономические или социальные.



Месячная аудитория российских поисковых систем



# КАК РАБОТАЕТ ПОИСКОВАЯ СИСТЕМА



Основные составляющие поисковой системы: поисковый робот, индексатор, поисковик .

Обычно системы работают поэтапно. Сначала поисковый робот получает контент, затем индексатор генерирует доступный для поиска индекс, и наконец, поисковик обеспечивает функциональность для поиска индексируемых данных. Чтобы обновить поисковую систему, этот цикл индексации выполняется повторно.

Поисковые системы работают, храня информацию о многих веб-страницах, которые они получают из HTML страниц. Поисковый робот или «краулер» (англ. *Crawler*) — программа, которая автоматически проходит по всем ссылкам, найденным на странице, и выделяет их. Краулер, основываясь на ссылках или исходя из заранее заданного списка адресов, осуществляет поиск новых документов, еще не известных поисковой системе. Владелец сайта может исключить определённые страницы при помощи robots.txt, используя который можно запретить индексацию файлам, страницам или каталогам сайта.



# ТИПЫ ПОИСКОВЫХ СИСТЕМ

---

Существует четыре типа поисковых систем: с поисковыми роботами, управляемые человеком, гибридные и мета-системы.

- *системы, использующие поисковых роботов*

Состоят из трёх частей: краулер («бот», «робот» или «паук»), индекс и программное обеспечение поисковой системы. Краулер нужен для обхода сети и создания списков веб-страниц. Индекс — большой архив копий веб-страниц. Цель программного обеспечения — оценивать результаты поиска. Благодаря тому, что поисковый робот в этом механизме постоянно исследует сеть, информация всегда актуальна. Большинство современных поисковых систем являются системами данного типа.

- *системы, управляемые человеком (каталоги ресурсов)*

Эти поисковые системы получают списки веб-страниц. Каталог содержит адрес, заголовок и краткое описание сайта. Каталог ресурсов ищет результаты только из описаний страницы, представленных ему веб-мастерами. Достоинство каталогов в том, что все ресурсы проверяются вручную, следовательно, и качество контента будет лучше по сравнению с результатами, полученными системой первого типа автоматически. Но есть и недостаток — обновление данных каталогов выполняется вручную и может существенно отставать от реального положения дел. Ранжирование страниц не может мгновенно меняться. В качестве примеров таких систем можно привести каталог Yahoo, dmoz и Galaxy.

- *гибридные системы*

Такие поисковые системы, как Yahoo, Google, MSN, сочетают в себе функции систем, использующие поисковых роботов, и систем, управляемых человеком

- *мета-системы*

Эти поисковые системы были полезны, когда у каждой поисковой системы был уникальный индекс, и поисковые системы были менее умными. Поскольку сейчас поиск намного улучшился, потребность в них уменьшилась. Примеры: MetaCrawler и Поиск MSN.

# НЕОБЫЧНЫЕ ПОИСКОВЫЕ СИСТЕМЫ

---

**DuckDuckGo** — гибридная поисковая система с политикой конфиденциальности пользователя и его поисковых запросов.

**Koogler**, с помощью которого ультраортодоксальные иудеи могли найти контент, признанный раввинами удовлетворяющим религиозным требованиям.

**Yauba** — индийский поиск «для параноиков». В отличие от привычных поисковиков, за пользователями не следят, а все записи об их действиях удаляются с сервера.

**TinEye** — поисковая система, специализирующаяся на поиске изображений в Интернете.

**Генон** — поисковая система, собирающая и создающая контент у себя на сайте.

**ImHalal** — первый исламский поисковик, может отличить разрешённые для мусульман (халяльные) результаты поиска от запретных (харамных).