**Лекция за 18.04.2012, часть 1**

***DNS.***

***Что такое DNS и почему оно появилось?***

Машине все равно, с каким адресным пространством имен интернета работать, и, на самом деле, адресное пространство, имеющее регулярную структуру, такое, как адресное пространство IP адресов, вполне удобно для машинной обработки. К сожалению, человеческое сознание так устроено, что человек плохо запоминает числа. Есть, конечно, люди, которые запоминают телефонные справочники, но это вызывает удивление и уважение, потому что это редкость. Причина этого очень простая: человеческая память устроена ассоциативно (ближайший аналог - сложный многомерный хеш), а ассоциативно запоминать числа сложно.

Поэтому с самого начала, с самого появления сети ***ARPANET*** возникли идеи, как приспособить пространство имен интернета к человеку. Высказывались разные отличающиеся схемы, у которых была одна основная идея: заменить или дополнить адресное пространство имен, которое существует в виде адресов интернет, неким механизмом, который позволил бы поставить в соответствие этим адресам какие-то названия, с которыми было бы удобно работать человеку. На самом деле, с этих идей и возникло то, что мы сейчас называем доменной системой имен. Возникла она далеко не сразу и претерпела некоторые модификации перед тем, как стала той доменной системой имен, которая нам сейчас известна. Стоит упомянуть ***Пола Мокапетриса*** - автора первых стандартов на доменную систему имен.

На самом деле концепция доменной системы имен DNS в некотором смысле уникальна: несмотря на то, что с момента ее создания (первые документы появились в начале 80х годов) она, во-первых, остается одной из самых консервативных систем интернета, т.е., например, все современное клиентское программное обеспечение, которое работает с DNS, в принципе могло работать с серверами, запущенными в 83 году. И наоборот, современные сервера могут откликаться на запросы операционных систем издания начала 80х годов. Такой консерватизм далеко не случаен: это жесткое требование обратной совместимости, поскольку доменная система имен - это один из краеугольных камней интернета наравне со стандартами системы имен IP адресов, маршрутизацией. Без доменной системы имен современный интернет не работоспособен. Высказывались предложения о ее модификации и дополнениях, однако эти изменения находят себе дорогу с большим трудом (например, использование национальных алфавитов доменной системы имен). Таким образом, получается, что мы начинаем знакомство с системой, которая претерпела за 30 лет очень небольшие изменения.

Когда говорят о доменной системе имен, речь может идти о очень разных вещах, которые называю одним и тем же термином. Терминология в доменной системе имен весьма консервативная, что проистекает из консервативности самого подхода и тех стандартов, которые были положены в ее основу. Под термином DNS понимается 3 разные сущности:

1. DNS – само по себе пространство имен, т.е. совокупность наименований, которые поставлены в соответствие большинству узлов интернета;
2. DNS – сама распределенная база данных, в которой это пространство имен содержится;
3. DNS – это совокупность программного обеспечения, которое используется для поиска в этом пространстве имен, а также алгоритмы, которые в этом программном обеспечении используются.

***Что же было до DNS?***

Интернет возник из полуэкспериментальной сети ARPANET, которая была развернута и стала внутренней информационной сетью министерства обороны США. Идеи о создании альтернативного пространства имен возникли почти сразу с возникновением этой сети. Строилось это пространство имен очень простым способом. На каждом компьютере сети существовал файл, который существует и сейчас. В UNIX: etc/hosts, а в Windows: windows/system32/drivers/etc/hosts. Этот файл содержал полный список хостов сети ARPANET с их IP адресами и именами хостов. Имена давались в произвольной форме: это была совокупность алфавитно-цифровых символов. До ликвидации домена ".arpa", в сети было порядка 20000 хостов. Размер этого файла был около 20000 строк (на самом деле несколько больше, т.к. в нем были дубликаты). При добавлении нового хоста в сеть, администратору сети отправлялось письмо с указанием имени и IP адреса этого хоста; этот адрес и имя заносились в некоторый мастер-файл, который потом рассылался на все хосты сети ARPANET, причем рассылался он при помощи электронной почты (протокол не SMTP). Этот файл заменялся автоматически, и вручную заменять его не требовалось.

Чем неудобен такой подход? Выяснилось, что 80-90% дневного трафика сети ARPANET - это пересылки файлов "hosts". В силу того, что какие-то компьютеры могли быть выключены, то выполнялась рассылка по расписанию, например раз в несколько часов. Поэтому к созданию системы DNS были очень серьезные предпосылки.

Задачи, которые стояли перед системой DNS:

1. снизить объем трафика на пересылках информации;
2. любое централизованное хранение столь важной информации чревато с точки зрения информационной безопасности: даже не столько намеренным влиянием злоумышленников, сколько человеческим фактором. Поэтому была поставлена задача сделать систему распределенной и децентрализованной.

Часть альтернативных систем либо ориентировалась на снижение объема трафика, но это ставило под угрозу общую надежность системы, либо занималась децентрализацией, но не обращала внимания на структуру передаваемой информации, при этом объем трафика возрастал. DNS оказался разумным компромиссом между этими возможностями. Поэтому DNS и дожил до наших дней в практически неизменной форме. DNS – один из самых устойчивых стандартов интернета: все программное обеспечение, которое работает с DNS, совместимо на уровне протоколов (по крайней мере базовые протоколы DNS все системы реализуют одинаково), но, более того, по большей части совместимо на уровне исходных файлов баз данных DNS. Примерно до начала 2000х годов 90% DNS серверов работало с одним программным обеспечением.

***Как устроено пространство имен DNS?***

Пространство имен DNS имеет форму (если представить в виде графа), близкую к дендритной, но это не чистый древовидных граф. Тем не менее, алгоритм поиска в дереве работает эффективно. Узлы этого дерева называются доменами, а листья этого дерева – IP хосты.

В основе всего лежит "пустой домен". В ранних стандартах он назывался доменом "точка", доменом ".arpa". Т.к. точка является разделителем между доменами разного уровня, то решили считать, что пусть корневой домен будет пустым и сохранить за точкой функцию разделителя. По иерархии пустой домен считается доменом нулевого уровня. Он может содержать только другие домены в качестве своих поддоменов.

Его поддомены - домены первого уровня. Они делятся на 2 большие категории:

1. ***территориальные домены***: домены, соответствующие государствам и территориям.  
   Примеры: ".ru", ".ua", ".uk", ".us", ".su", ".sh".
2. ***корпоративные домены***. Являются наследием сети ARPANET, когда ни о каком международном интернете речь не шла.   
    Старые корпоративные домены:
3. ".com" - домен коммерческих организаций
4. ".org" - домен неправительственных некоммерческих организаций
5. ".gov" - домен правительственных организаций США
6. ".mil" - домен министерства обороны США
7. ".net" - домен организаций, связанных с обслуживанием сети

Новые корпоративные домены:

1. ".biz" - дублер ".com"
2. ".name" - домен для размещения персональных страниц пользователей
3. ".info" - домен для информационных и новостных организаций
4. ".int" - домен для международных организаций
5. ".xxx" - домен для ресурсов для взрослых (пока что не создан)
6. ".edu" - домен учебных организаций

***По какому принципу выделяются территориальные домены?***

Для государств, имеющих крупные владения, находящиеся на большом расстоянии от их основной территории, допускается выделение дополнительного территориального домена для соответствующих территорий. Это делается для удобства, чтобы можно было примерно представлять размещение соответствующих ресурсов.

".su" - домен призрак, был делегирован для СССР. После распада СССР он был перерегистрирован на некоторого финского системного администратора. В середине 90х годов он был выкуплен и перешел под управление ***Российского НИИ Общественных Сетей*** и стал дублером домена ".ru".

Часть государств, имеющих интересные с коммерческой точки зрения названия доменов, предоставляют свои домены первого уровня для размещения сайтов тех или иных организаций (например ".tv").

Как получаются двухбуквенные обозначения доменов? Это двухбуквенный почтовый код государства или территории.

В доменах первого уровня возможны только поддомены.

Домены второго уровня стоятся по-разному.

Возьмем домен ".ru", как наиболее нам близкий. В нем есть поддомены, отвечающие каждому субъекту федерации, например "spb.ru". Есть домены-контейнеры для совокупности организаций, например "gov.ru". Отличительная особенность доменов 2 уровня состоит в том, что в доменах второго уровня уже могут находиться не только поддомены, но и хосты. Начиная со второго уровня вложенности, все домены имеют возможность иметь хосты. Хосты являются терминальными вершинами графа.

***Чем граф этого пространства имен отличается от дендрита?*** На самом деле он имеет еще один тип терминальных вершин, которые называются "псевдонимы" или "синонимы". Любой хост и любой домен, кроме доменов 0 и 1 уровня, может иметь от одного до произвольного количества синонимов. Они, с точки зрения пространства имен, являются равноправной заменой для соответствующего имени, причем они могут находиться в доменах другого уровня, отличного от того, на который ссылается синоним.

***Для чего возникли синонимы?***

1. Структура пространства имен напоминает пространство имен файловой системы в UNIX-подобных системах, а там есть возможность создавать символические ссылки на имена. Поэтому, по аналогии, возникла идея сделать тут то же самое. Первые программы поиска в базе данных DNS использовали те же алгоритмы, что и поиск в файловой системе.
2. Синонимы расширяют возможности пространства имен, в частности позволяют использовать более короткие или благозвучные названия. Именно использование синонимов является той причиной, почему граф пространства имен не является чистым дендритом. По разным оценкам количество синонимов не превышает 1% от общего количества терминальных вершин в графе, это позволяет использовать алгоритмы поиска по дереву.

Теперь поговорим об алфавите DNS. Алфавит DNS один из самых узких из имеющихся алфавитов: он долгое время состоял из алфавитно-цифровых символов ASCII и знака дефиса; это все символы, которые можно употреблять в именах хостов, доменов и синонимов. При этом большие и маленькие буквы не различаются. Одна из самых серьезных реформ, которая была проведена в 2009-10 годах – это стало возможно использование национальных алфавитов в DNS. Первыми национальными доменами стали домены Российской Федерации, Болгарии и КНР. В случае использования национальных алфавитов разрешается использовать все те же самые символы плюс те символы алфавита UTF-8, которые относятся к категории ***"national character set"***, т.е. все равно нельзя использовать специальные символы, математические символы, общие символы диакритики. Единственное существенное ограничение, которое ставится на использование национальных доменов и национальных алфавитов, заключается в том, что домены 1 уровня, которые используют национальные алфавиты, не должны содержать национальных символов, которые имели бы форму, совпадающую или близкую к соответствующим латинским буквам. Вопрос использования диакритических алфавитов на основе латиницы не решен.

***Как устроена база данных DNS?***

База данных DNS строится на основе распределенной сети серверов, которая тоже строится по иерархическому принципу.

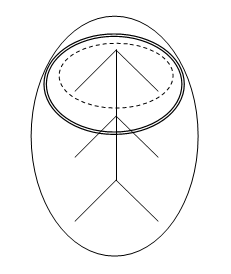
***Терминология.***

По отношению к каждому домену сервера делятся на категории:

1. сервер может быть ***авторитетным*** для данного домена, т.е. ответ, который предоставляет данный сервер относительно соотношения имени и IP адреса для хоста данного домена, является гарантированно точным
2. сервер может быть ***неавторитетным*** для данного домена, т.е. он может дать ответ, который был точным когда-то в прошлом
3. сервер может быть ***кэширующим***, т.е. не является авторитетными ни для одного домена.

Для каждого домена должно быть не менее 2 авторитетных серверов, причем они должны быть расположены в двух физически независимых сетях (т.е. не должно быть ситуации, в которой оба сервера откажут одновременно). Для серверов первого уровня, как правило, авторитетных серверов еще больше (для ".ru" около 11 серверов). Т.к. мы можем использовать для получения информации неавторитетные сервера, то это снижает нагрузку на авторитетные. Было экспериментально показано, что алгоритм DNS устроен таким образом, что может возникнуть такая ситуация, что отдельные хосты никогда не смогут получить достоверной информации о соответствии имени и IP адреса. Современное развитие стандартов DNS направлено в сторону поддержки динамического DNS и DNS, связанного с использованием мобильного интернета (dyndns.com, no-ip.org и т.п., позволяют регистрировать хосты, которые меняют IP адреса).

Когда речь идет о базе данных DNS, возникает понятие ***"зона"***. В самом банальном представлении зона соответствует домену. ***Зона*** - это совокупность информации о всех хостах текущего домена, информации о его поддоменах, но не информация о хостах поддоменов.



1. бытовое понимание слова domain

2. понимание слова domain в терминологии DNS

3. зона

Т.е. зона - это нечто среднее между доменом в узком и широком смысле. Зона включает в себя совокупность всех ниже лежащих вершин, связанных с текущей вершиной не более чем одним соединением.

***Проведем более четкую классификацию серверов.***

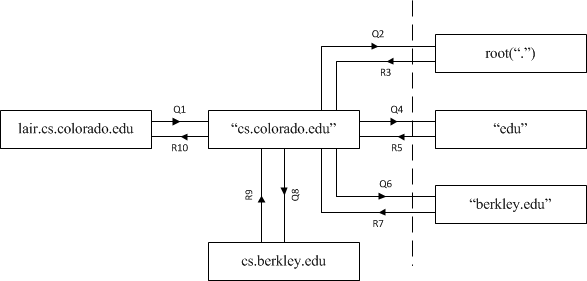
Авторитетный сервер - это представление зоны в распределенной базе данных. Авторитетные сервера делятся между собой: в каждой зоне есть один ***главный сервер***, или ***мастер-сервер***, и произвольное количество ***подчиненных*** (***slave***). Любой slave можно сделать мастером, единственно ограничение, что в каждый момент времени не больше 1 мастер-сервера. Отличительная особенность мастер-сервера в том, что все изменения, которые производятся в отношении записи отдельной зоны, делаются на мастер-сервера и транслируются на slaves. Раньше трансляция выполнялась по времени (slave’ запрашивали мастера об изменениях), в настоящее время мастер извещает своих slaves о том, что произошло изменение в записях и инициирует процесс, который называется "зонная пересылка" (т.е. пересылка всей информации о текущем содержании зоны). Несмотря на распределенность базы данных, все равно для передачи между авторитетными серверами доменов было решено сохранить полную зонную пересылку (чтобы избежать проблем с синхронизацией).

Довольно часто выделяют понятие "***усеченного авторитетного сервера***" - это сервер, который содержит только часть записей, либо записи определенного типа (типы будут позже). Иногда вводят понятие "***внутреннего сервера***" - это авторитетный сервер, который обслуживает домен и доступен только изнутри своего домена. В большинстве не очень крупных доменов мастер-сервер одновременно является и внутренним. Все сервера, которые расположены вне текущего домена, называются внешними.

Внутри кэширующих серверов есть ***переадресующие сервера*** - они сами запросов не осуществляют, а просто транслируют чужой запрос.

Раз есть DNS сервер, то должны быть и DNS клиенты. На самом деле DNS клиент - это стандартная библиотека языка C (даже только несколько функций из нее). Это функции: ***getHostByName()*** - по имени хоста определяет IP адрес, ***getNameByHost()*** - по IP адресу определяет его имя. На самом деле это семейства функций, которые отличаются либо аргументом, либо типом возвращаемого значения.

Как осуществляется алгоритм поиска в DNS?



В чем состоит задача? У нас есть компьютер - хост. Он называется ***lair.cs.colorado.edu***. Он отправляет запрос на резолвинг имени ***vangogh.cs.berkley.edu***. Названия хостов в кавычках означает, что это авторитетный сервер для данного домена. Обозначение транзакций: индекс - порядок, в котором выполняется транзакция.

Типы транзакций:

1. Q - запрос, т.е. getHostByName();
2. A - ответ, т.е. соответствующий IP адрес;
3. R - ответ-ссылка (reference), т.е. вместо прямого ответа сервер говорит "я данного ответа не знаю, но попробуй спросить у такого сервера".

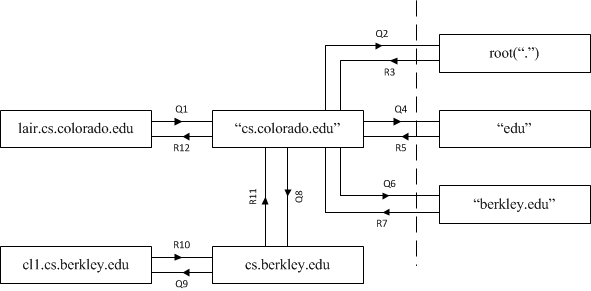
О пунктирной линии, отделяющей 3 сервера: сервера слева от нее - рекурсивные, т.е. они будут повторять запрос, пришедший им от клиентов. Сервера правее – нерекурсивные сервера, т.е. они не повторяют запрос, и, если не знают ответ, то отвечают ссылкой (reference). Нерекурсивные сервера – это, как правило, сервера, отвечающие за домены 0 и 1 уровня, поскольку на них ложится слишком большая нагрузка.

***Как все эти транзакции проходят последовательно?***

C lair.cs.colorado.edu запрос посылается на сервер имен его собственного домена, этот сервер имен прописан в настройках IP. Поскольку это рекурсивный сервер, он начинает транслировать эти запросы дальше. Поскольку он не знает ответа на этот запрос и не знает, где искать, он запрашивает сразу сервер нулевого уровня. Сервер нулевого уровня отправляет ему ответ-ссылку "спроси у сервера, который обслуживает домен .edu". Запрос отправляется на сервер, обслуживающий домен .edu. Сервер .edu отправляет на сервер berkley.edu. Сервер berkley.edu отвечает ссылкой на cs.berkley.edu. После этого запрос посылается на cs.berkley.edu и он отвечает адресом Ван Гога. При нормальной работе системы это все занимает меньше секунды.

Что закэшируется на cs.colorado.edu? Адрес Ван Гога, адрес cs.berkley.edu, все ответы-ссылки, которые получил, а также ответ. Кэш держится недолго. В дальнейшем обращение к серверам 0 и 1 уровня можно избежать, т.к. данные закэшированы.

Общий алгоритм поиска в DNS выглядит так: начинаем перебирать домен, начиная с 0 уровня, обращаясь к серверам соответствующих доменов (если не закэшировали ответ ранее). Если получаем от конкретного сервера ответ-ссылку, то обращаемся к серверу по этой ссылке. И т.д. пока не дойдем до сервера, который нам даст ответ.



Теперь запрос такой: ***john.cl1.cs.berkley.edu***. До Q8 все повторяется, как в 1 примере. Затем этот сервер 9-ой транзакцией обращается к серверу cl1.cs.berkley.edu, получает ответ и транслирует его 11 транзакцией на cs.colorado.edu и 12-й транзакцией ответ будет оттранслирован на исходный компьютер.