У корня DNS

Нормальная работа сети Интернет немыслима без правильно функционирующей системы доменных имен, или DNS (Domain Name System). Всякий раз, когда мы набираем имя веб-сайта или посылаем электронную почту, эта система берет на себя задачу трансляции имени в цифровой адрес протокола IP, необходимый для осуществления связи между компьютерами в сети.

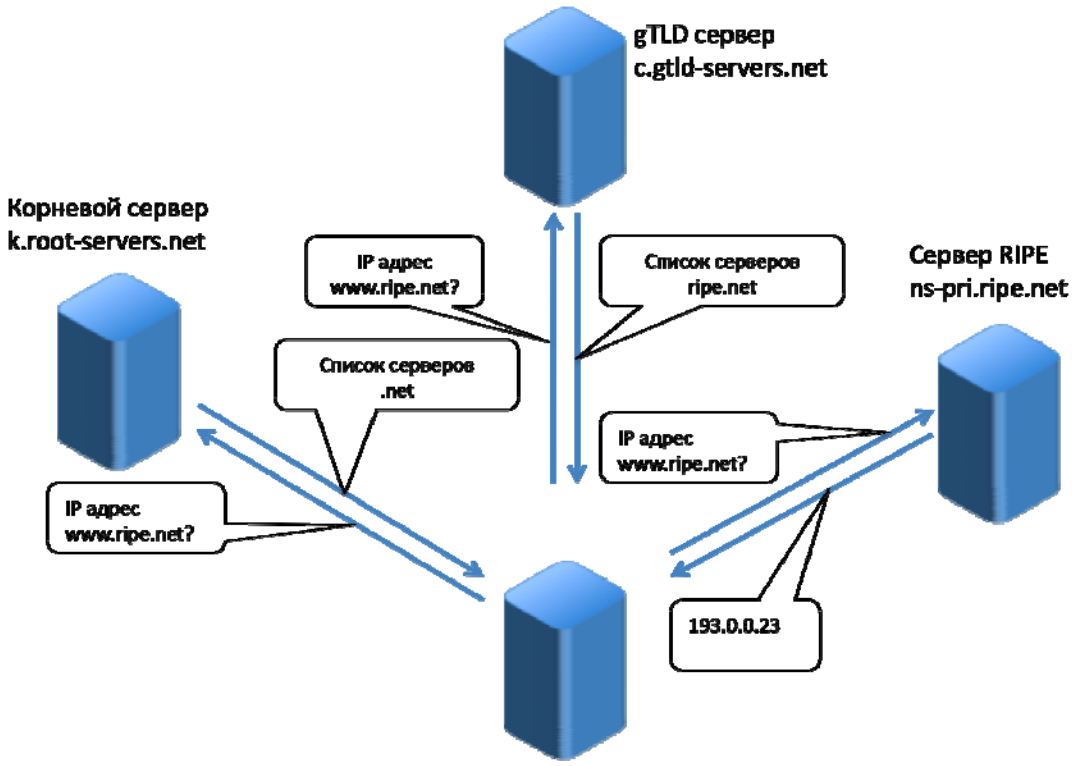
Принципы работы системы разрешения имен

Система DNS является иерархической и распределенной. Не существует единой базы данных, хранящей информацию о всех именах и соответствующих им IP-адресах и других записях. Напротив, DNS – это миллионы баз данных, каждая из которых содержит информацию о конкретном домене. Иерархию DNS можно увидеть в доменном имени, например, в имени веб сайта. Возьмем, например, сайт RIPE NCC – www.ripe.net. Это имя состоит из трех частей, разделенных точками. Точнее четырех, поскольку, формально говоря, полное доменное имя всегда заканчивается точкой, обозначающей так называемый корневой домен, или корневую зону DNS. Итак:

.	Корневая зона, содержащая информацию о всех поддоменах: net, com, org, ru, su , и т.д. Точнее, информацию о серверах, обслуживающих эти домены.
net	Домен net, содержащий информацию о всех поддоменах, зарегистрированных в ней. Например, ripe. Опять же, этот домен содержит адреса серверов, у которых можно получить дополнительную информацию о содержимом поддоменов.
ripe	Домен ripe, содержащий информацию о всех поддоменах, а также имена серверов, зарегистрированных непосредственно в этом домене, например www.ripe.net.
www	Имя вэб-сервера RIPE NCC и соответствующие ему IP адреса.

Соответственно, трансляция имени www.ripe.net в соответствующие ему один или более

Рис 1 Процесс трансляции имен в DNS



IP-адресов будет происходить в несколько этапов. Сначала будут запрошены серверы, обслуживающие корневую зону. Эти серверы ничего не знают о существовании домена ripe и тем более адрес www.ripe.net. Но они сообщат, как можно связаться с серверами, обслуживающими домен следующего уровня – net. От них мы узнаем адреса серверов домена ripe, которые, в свою очередь, и ответят на запрос о IP-адресе сервера www.ripe.net.

Такая архитектура DNS позволяет распределить нагрузку и ответственность за работу системы между администраторами отдельных доменов. В их обязанности входит обеспечение нормальной производительности при ответе на запросы к зоне, поддержка уникальности имен в рамках зоны а также уведомление администратора родительской зоны об изменениях в составе серверов, обслуживающих зону.

Эта иерархически распределенная архитектура DNS обеспечила долгожитие системы и ее эволюционное развитие уже на протяжении более четверти века.

Но есть в системе DNS одна особенность, отличающая ее от многих других систем Интернета, имеющих децентрализованный характер. Это та самая точка, корень дерева DNS, откуда начинаются все свежие запросы. О нем мы и поговорим подробнее.

Корневой уровень DNS

Корневые серверы (КС) DNS являются критическим компонентом системы, поскольку обеспечивают доступ к корневой зоне DNS. Корневая зона содержит информацию обо всех доменах самого верхнего уровня: национальные домены (например .ru), домены общего назначения (например .com) и спонсированные домены (например .museum). Эта информация указывает клиенту на какие серверы DNS отправить последующий запрос для продолжения разрешения полного доменного имени. Другими словами, любой "свежий" (т.е. не сохраненный в кэше клиента) запрос начинается с обращения к корневому серверу.

Особенностью корневых серверов является также то, что первичный запрос (priming), т.е. самый первый запрос, осуществляемый клиентом, производится по IP адресу сервера, а не по его имени. Это объясняется тем, что для трансляции имени в IP адрес необходима система DNS, для начала работы с которой необходим доступ к корневому серверу. Очевидно, что изначальная информация о корневых серверах (их IP адреса) не может быть получены из системы DNS. Эта информация содержится в специальном файле hints, хранящемся у клиента и обычно получаемом вместе с программным обеспечением (операционная система, ПО DNS, и т.п.)

Сегодняшняя система корневых серверов и координация ее работы

Со времени создания системы DNS в середине 80-х годов до 2000 года, система корневых серверов (СКС) состояла из первичного (primary) сервера (ns.internic.net, позднее переименованный в a.root-servers.net) и растущего числа реплик (secondary), в конечном итоге достигшее 12, с именами b.root-servers.net, c.root-servers.net и т.д. до m.root-servers.net. Каждый сервер управляется и сопровождается отдельной организацией-оператором, различными по роду деятельности, организации и географии. Полный список приведен ниже, его также можно найти на сайте http://www.root-servers.org.

В 2000-2002 г.г. архитектура системы была изменена. Был создан "скрытый" мастер-сервер и 13 равноправных КС, получающих идентичные копии корневой зоны от мастера.

Начиная с 2003 года получила распространение технология anycast, позволяющая "клонировать" серверы с одним и тем же именем и адресом. Эта технология стала активно использоваться рядом операторов и позволила существенно расширить географию СКС, до этого охватывающую преимущественно США.

Рассмотрим подробнее участников СКС. Для этого обратимся к процессу внесения изменений в содержимое корневой зоны, представленному на рисунке 2.

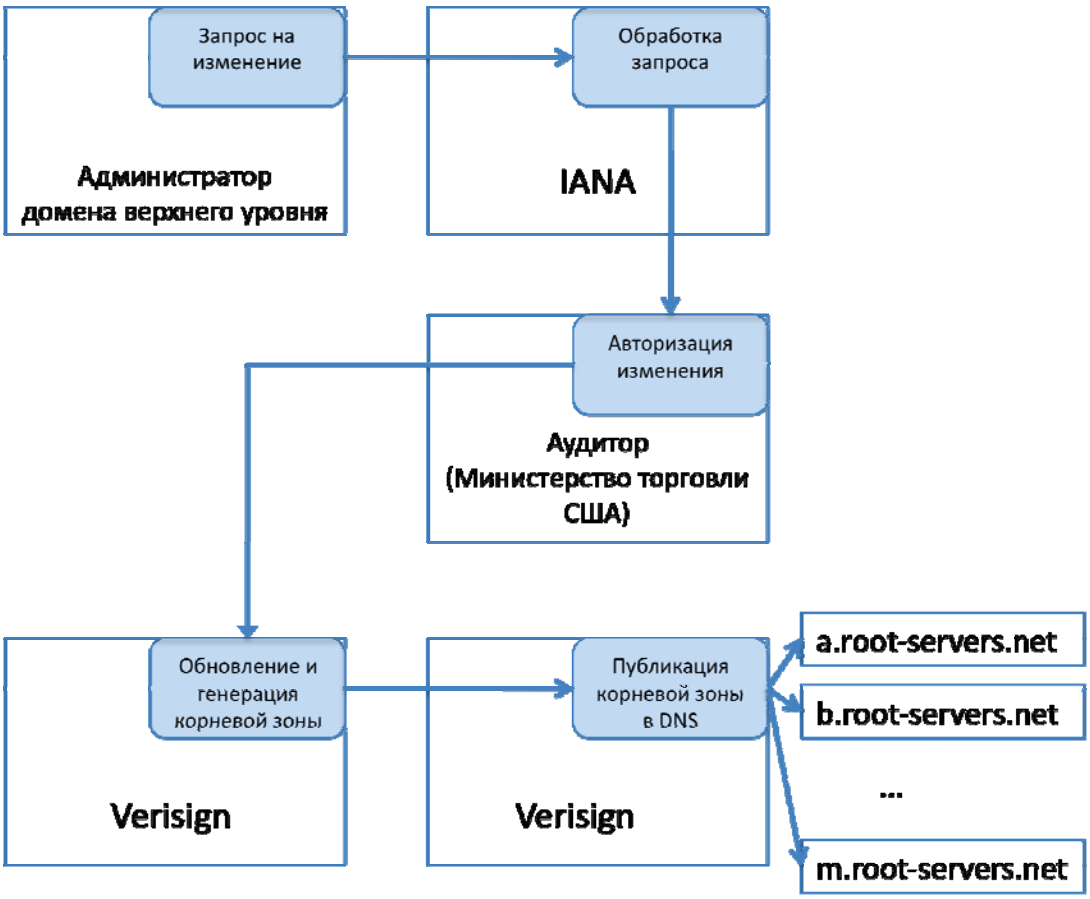


Рис 2 Процесс внесения изменений в корневую зону

Запрос на изменение поступает от администратора домена верхнего уровня (ccTLD, gTLD, и т.д.) и обслуживается IANA. Типичным примером такого запроса является изменение состава серверов, обслуживающих домен.

После проведения необходимых административных и технических процедур (например проверка правильности и законности запроса, проверка возможных негативных последствий на корневую зону), запрос на изменение подписывается цифровым образом и направляется для авторизации аудитору, роль которого в настоящее время выполняет Министерство Торговли США.

Затем изменения направляются организации, ответственной за фактическое редактирование и публикацию зоны в DNS. Эту роль в настоящее время выполняет компания VeriSign. Кстати, эта же компания является оператором 2 корневых серверов - a.root-servers.net и j.root-servers.net.

Зона изначально публикуется на скрытом мастер-сервере и затем распространяется на все реплики с использованием протокола TSIG, защищающий данные от модификации при передаче.

Операторы КС

Операторами КС являются различные организации, получившие право управления серверами в относительно отдаленном прошлом, когда подобные вопросы решались менее формально. Среди операторов находятся университеты, организации Министерства Обороны США, некоммерческие ассоциации. Операторы являются финансово и юридически независимыми от корпорации ICANN, в рамках которого действует IANA. При принятии операционных решений операторы руководствуются технической целесообразностью и существующими стандартами (например, RFC2870), в основном поддерживая статус-кво. Крупнейшим решением такого рода, принятого независимо оператором сервера f.root-servers.net компанией ISC, было решение о применении технологии anycast. Хотя это решение прошло тщательную экспертную проверку специалистов, оно не было регламентировано ICANN или каким-либо из его Советов. Впоследствии к ISC присоединился ряд других операторов.

Принято считать, что подобная независимость и разнородность операторов КС является основой технической и политической стабильность системы в целом, исключая узурпацию управления какой-либо из сторон.

Операторы КС образуют неформальную группу, целью которой является координация совместных действий и обмен операционной информацией и опытом. Группа проводит регулярные встречи 3 раза в год, приуроченные к совещанию IETF. Одним из результатов таких совещаний является генерация секретного ключа для протокола TSIG.

Члены группы являются также членами Консультационного Совета КСК ICANN (Root Server System Advisory Committee, RSSAC), в задачу которого входит выработка рекомендаций по управлению КСК и внесению различных изменений в систему.

До недавнего времени отсутствовали какие-либо формальные отношения между операторами и ICANN/IANA. Эта ситуация изменилась с подписанием первого соглашения между ICANN и оператором сервера f.root-servers.net компанией ISC. Данное соглашение не предусматривает никаких финансовых расчетов и лишь определяет взаимные обязанности сторон в отношении управления КС. Известно, что ряд операторов также обсуждают подобные соглашения с ICANN.

Ниже приведен список и краткая характеристика текущих операторов СКС.

СКС	Организация-оператор	Характер деятельности
A	VeriSign, Inc.	Коммерческая корпорация, крупнейший поставщик цифровых сертификатов и средств защиты электронных коммуникаций, США
B	Information Sciences Institute	Институт Университета Южной Калифорнии (USC), США
C	Cogent Communications	Один из крупнейших коммерческих Интернет сервис провайдеров, США
D	University of Maryland	Университет, США
E	NASA Ames Research Center	Государственное агентство, США
F	Internet Systems Consortium, Inc.	Некоммерческая корпорация, США
G	U.S. DOD Network Information Center	Государственное агентство, США
H	U.S. Army Research Lab	Государственное учреждение, США
I	Autonomica	Коммерческая организация, подразделение Netnod AB, Швеция
J	VeriSign, Inc.	Коммерческая корпорация, крупнейший поставщик цифровых сертификатов и средств защиты электронных коммуникаций, США
K	RIPE NCC	Некоммерческая ассоциация, Нидерланды
L	ICANN	Некоммерческая корпорация, США
M	WIDE Project	Некоммерческий проект, секретариат Университет Кейо, Япония

Альтернативные СКС

Хотя возможность умышленного нарушения работы КС или модификация содержимого корневой зоны каким-либо из операторов или ICANN/IANA маловероятна, строго говоря, в настоящее время не существует формальных процессов или международных законодательных актов, которые могли бы этому воспрепятствовать. Можно сказать, что нормальное функционирование СКС отчасти зависит от "доброй воли" участников.

Необходимо заметить, что подобные "неформальные" зависимости, основанные на доверии, характерны для работы системы DNS (и во многом сети Интернет) в целом. В конечном итоге, выбор, какие КС использовать, остается за клиентом. Эта информация содержится в конфигурационном файле hints и может быть изменена.

Эта особенность вкупе с неудовлетворенностью существующей системой управления СКС во главе с ICANN при участии правительства одной страны - США, привела к созданию так называемых альтернативных СКС. Примерами таких систем служат Public-Root, Open Root Server Network (ORSN) и UnifiedRoot. Хотя эти системы копируют текущее состояние корневой зоны, сама архитектура предусматривает, что в определенных условиях альтернативные СКС могут предоставить альтернативное пространство имен. Администратор клиента DNS (обычно сервера DNS, обслуживающего корпоративных

пользователей или клиентов кабельной сети) может выбрать альтернативную СКС изменив соответствующим образом файл hints.

Альтернативные СКС получили критическую оценку со стороны IETF как открывающие потенциальную возможность раскола единого Интернета (см. RFC2826).

Техническое развитие корневого уровня DNS

Система корневого уровня DNS весьма консервативна. Это и понятно – любые изменения на этом уровне затрагивают функционирование всей глобальной системы доменных имен. Тем не менее, несколько важных изменений были внедрены в СКС и корневую зону в течение последних лет.

Anycast

Вы наверное заметили, что количество КС, а точнее имен КС, является «счастливым» числом 13. Это не вызов суеверию, число 13 связано с ограничением на размер сообщения в 512 байт, установленным стандартом DNS [RFC1035 4.2.1]. Хотя исторически это ограничение было вызвано ограничением на пакеты UDP не допускающим фрагментации, оно продолжает существовать в протоколе DNS, несмотря на появление новых сетевых протоколов, например IPv6. Расширенный стандарт DNS (EDNS0, RFC2671 2.3, 4.5) предусматривает предварительное соглашение о размере сообщения между клиентом и сервером, однако степень распространения этого протокола в сегодняшних системах DNS недостаточна для снятия ограничения в 512 байт в ближайшем будущем.

Все 13 имен КС (a.root-servers.net – m.root-servers.net) умещаются в отведенные 512 байт, а в случае четырнадцати имен для значительной части запросов ответ не сможет полностью поместиться в отведенные 512 байт. Результатом может стать существенное снижение производительности системы, так как часть клиентов вынуждена будет повторить запрос, но теперь с использованием гораздо более "медленного" протокола TCP.

До 2003 года максимальное количество КС совпадало с числом имен и не могло превышать 13. Но даже при таком количестве, для многих клиентов, особенно находящихся вне США и Западной Европы, КС располагались неоптимально.

Другой проблемой, требующей решения являлось то, что 13 серверов оказалось достаточно просто атаковать посредством распределенной атаки DoS. Так случилось в октябре 2002 года, когда большинство КС были недоступны в течение нескольких часов.

Решению этих проблем помогла т.н. технология anycast, известная с 1993 года, но не применявшаяся в глобальном масштабе и на таком уровне. Суть ее заключается в анонсировании оператором одной и той же сети (префикса IP и автономной системы) в различных частях Интернета. Благодаря архитектуре системы маршрутизации для любого клиента существует единственный и самый «короткий» путь к любой другой сети Интернета. Таким образом, anycast позволяет клиенту установить связь с наиболее близкой в топологическом смысле сети anycast без дополнительных изменений в ПО и протоколах!

Технология anycast наиболее подходит для приложений, использующих протокол UDP, работающий без установления продолжительной связи. В противном случае, при каких-либо изменениях в топологии Интернет (которые происходят постоянно), кратчайший путь может привести клиента к другой сети anycast, и связь будет разорвана.

В 2003 году, после тщательной экспертной проверки и тестирования, оператор сервера f.root-servers.net разместил реплику своего сервера с использованием anycast. Примеру ISC последовал и ряд других операторов и география системы КС существенно расширилась, как можно видеть из рис. 3. На сегодняшний день общее число серверов, по-прежнему управляемых 12 операторами, достигло 166. Два таких сервера (f.root-servers.net и k.root-servers.net) расположены в России.

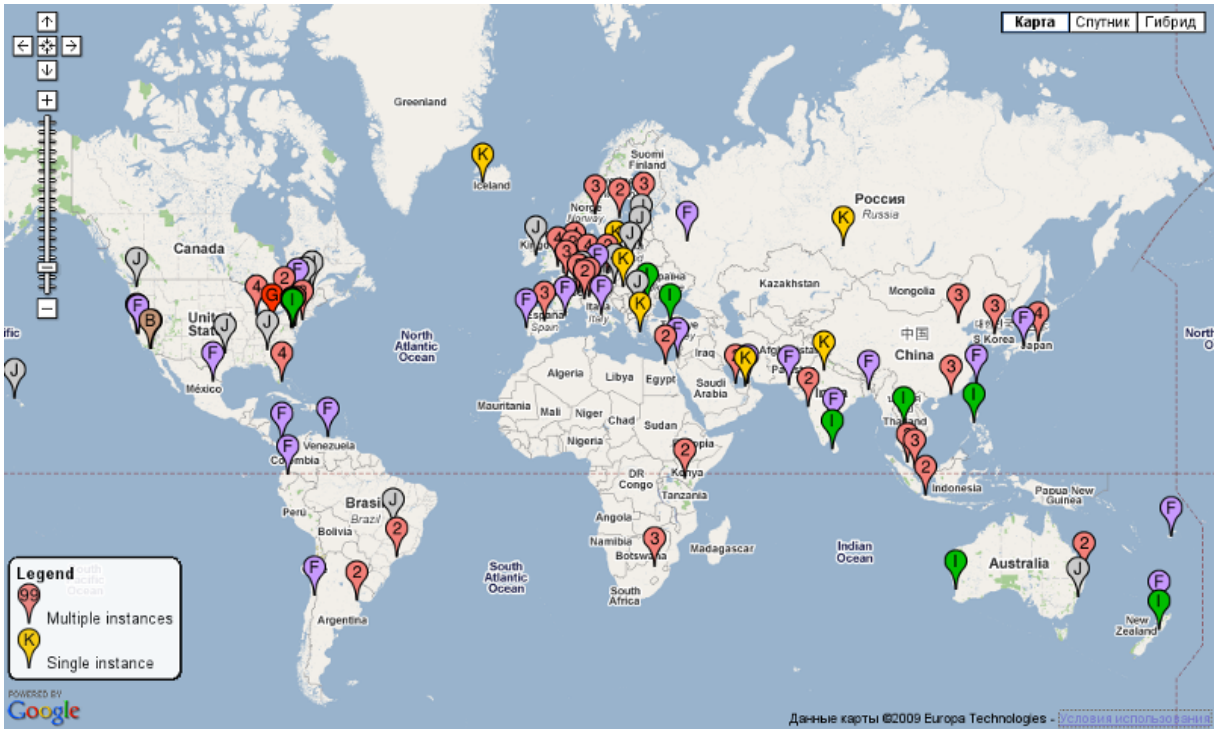


Рис 3 География системы КС (см. http://www.root-servers.org/)

IPv6

Внедрение поддержки IPv6 в корневой зоне в начале 2008 года обеспечило полную поддержку протокола IPv6 в глобальной системой доменных имен.

К этому моменту IPv6 уже появился во многих доменах верхнего уровня, а отдельные КС обеспечивали доступ к корневой зоне по протоколу IPv6. Единственно, что отсутствовало – это информация о доступных адресах IPv6 КС.

На первый взгляд несложная задача, включение этой информации в корневую зону было связано с основной проблемой – превышение размера ответа, и в частности ответа на первичный (priming) запрос, все те же 512 байт!

Напомню, что ответ на первичный запрос содержит имена и адреса всех 13 КС. Оказывается, в 512 байт умещаются все имена, все 13 IPv4 адресов и только 2 адреса IPv6. В принципе, этой информации достаточно, чтобы клиент мог продолжить поиск в DNS. Но было неясно, как отреагируют различные клиенты на отсутствие информации, которую они, возможно, предполагали получить. Не приведет ли это к обвальному использованию протокола TCP, позволяющему избежать ограничения на размер пакета, но и гораздо более «дорогостоящего»? Эти вопросы требовали тщательного тестирования.

Другой проблемой являлось наличие адресов IPv6 в файле hints – так сказать стартера клиента. Если клиент не сможет прочитать этот файл, например, потому что он не предполагает наличия адресов IPv6 в файле hints, глобальная система DNS окажется для него недоступной.

Наконец, третьей возможной проблемой могут являться промежуточные системы, например сетевые экраны (системы firewall), которые могут не пропускать DNS-пакеты, размер которых превышает 512 байт или содержит записи протокола IPv6.

Исследование этих вопросов показало, что особых причин для беспокойства нет: большинство современных клиентов поддерживают расширение EDNS0, позволяющее передачу DNS-пакетов большего размера. Даже в противном случае, ответ на первичный запрос, не превышающий 512 байт, содержит достаточно информации для начала поиска. Тестирование также не выявило проблем с новыми записями в файле hints.

Основной проблемой, как оказалось, могли являться промежуточные системы, хотя тестирование показало, что многие из них не накладывают ограничений на размер передаваемого пакета DNS. Единственным способом разрешить эту проблему явилась широкая просветительная работа.

Адреса IPv6 первых шести КС были включены в корневую зону и файл hints 4 февраля 2008 года. Никаких значимых сбоев в работе глобальной системы DNS отмечено не было.

IDN

IDN расшифровывается как Internationalised Domain Names, или международные доменные имена, и означает возможность использования национальных алфавитов при задании имени домена или сервера. Технология весьма привлекательная, как с технической, так и с политической точки зрения, поскольку для пользователей с нелатинскими алфавитами она позволяет практически полностью исключить необходимость использования латиницы при работе в Интернете.

Однако базовый стандарт DNS допускает использование только 7-битной кодировки для символов, так называемую таблицу ASCII, да и то с некоторыми ограничениями для имен хостов. В то же время стандартной формой кодирования символов национальных письменных языков является Юникод (Unicode).

Для трансляции символов Юникода в набор допустимых символов DNS используется так называемый Пюникод (Punycode). Имена в Пюникоде выглядят немного странно, например, слово «испытание» будет представлено как xn--80akhbyknj4f, но зато они полностью соответствуют стандарту DNS.

Соответственно, любое приложение, поддерживающее национальные языки, должно транслировать доменное имя, введенное пользователем, в строку Пюникода. Далее все работает как обычно – клиент DNS производит поиск для разрешения имени в адрес IP. Например, если пользователь набирает в окошке вэб-браузера «пример.испытание», поиск DNS будет произведен для имени xn--e1afmkfd.xn--80akhbyknj4f и в результате будет получен IP-адрес вэб-сервера - 199.7.85.17.

Тестирование IDN некоторые доменные операторы начали еще в 2001 году, а, начиная с 2003 года, ряд операторов доменов верхнего уровня открыли регистрацию поддоменов, содержащих IDN. В 2006 году ICANN начал тестирование IDN в корневой зоне, сначала в лаборатории, а затем и в реальном режиме. На сегодняшний день домен example.test представлен в корневой зоне на 11 языках, включая русский. Достаточно набрать http://пример.испытание, чтобы попасть на заглавную страницу этого проекта ICANN.

В настоящее время ведется обсуждение так называемого процесса «Fast Track», который позволит операторам существующих национальных доменов верхнего уровня зарегистрировать эти домена на национальных языках.

С использованием IDN связан и ряд проблем. Например, IDN открывает более широкие возможности для спуфинга (spoofing) имен, когда имя сервера внешне очень похоже на другое имя, но на самом деле использует другие символы, так называемые гомографы. Действительно, как отличишь www.paypal.com от www.paypal.com в котором вторая буква «а» набрана кириллицей. По сравнению с обычными именами, где 1 похоже на l, а 0 на O, IDN содержит гораздо больше гомографов. Решение этой проблемы требует строгого контроля со стороны регистраторов доменов, ограничения числа поддерживаемых языков в рамках домена и запрета смешивания различных языков в доменном имени.

DNSSEC

Основным недостатком DNS является слабая система защиты данных. Это означает, что в процессе передачи данных от сервера к клиенту, данные могут быть модифицированы. Также возможно создание ложных DNS серверов, поставляющих модифицированные данные. DNSSEC, который является расширением стандарта DNS, позволяет решить эту проблему путем удостоверения данных цифровой подписью администратора домена, от которой получены данные. В свою очередь, подпись администратора домена удостоверяется администратором родительской зоны. Такая цепочка удостоверений продолжается вплоть до корневой зоны и DNS-клиент может удостовериться, что само имя и вся цепочка делегаций достоверны и не были модифицированы.

В настоящее время поддержка DNSSEC весьма разбросана по дереву DNS, что в большинстве случаев не позволяет построить так называемые цепочки доверия. Отсутствует поддержка DNSSEC и в корневой зоне. Все это, по существу, сводит на нет преимущества DNSSEC.

Начиная с 2007 года подразделение ICANN IANA начало тестирование поддержку DNSSEC в корневой зоне в лабораторном режиме. С технической точки зрения тесты не выявили никаких проблем, однако отсутствие прогресса в этой области в первую очередь связано с вопросами контроля за содержимым корневой зоны и управления цифровыми ключами, преимущественно политического характера.

Технический директор RIPE NCC Андрей Робачевский