Network Working Group
Request for Comments: 2182
BCP: 16
Category: Best Current Practice

R. Elz
University of Melbourne
R. Bush
RGnet, Inc.
S. Bradner
Harvard University
M. Patton
Consultant

July 1997

Selection and Operation of Secondary DNS Servers

Выбор и использование вторичных серверов DNS

Статус документа

Этот документ относится к категории обмена опытом (Best Current Practice или BCP) для сообщества Internet. Принимаются поправки и предложения, направленные на совершенствование документа. Документ может распространяться свободно.

Аннотация

Система доменных имён (Domain Name System) требует наличия нескольких серверов для каждого делегированного домена (зоны). В этом документе обсуждается выбор вторичных серверов для зон DNS. При выборе учитывается как физическое, так и логическое местоположение каждого сервера. Рассматривается также число вторичных серверов, подходящих для зоны, и некоторый вопросы обслуживания таких серверов.

Оглавление

1. Введение

Ряд проблем в работе DNS связан сегодня с неправильным выбором вторичных серверов для зон DNS. Географическое расположение, а также разнообразие сетевой связности, раскрываемые набором серверов DNS для зоны, могут повышать надёжность этой зоны, а также улучшить общую производительность сети и характеристики доступа. Другие соображения при выборе серверов могут неожиданно снижать надёжность или вносить дополнительные требования к сети.

В этом документе обсуждаются многие вопросы, которые следует рассматривать при выборе вторичных серверов для зоны. Приведены рекомендации по оптимальному выбору серверов для обслуживания конкретной зоны.

2. Определения

Для целей данного документа (и только для него) применяются приведённые ниже определения.

DNS
Система доменных имён (Domain Name System) [RFC1034, RFC1035].

Zone - зона
Часть дерева DNS, рассматриваемая как целое (unit).

Forward Zone - прямая зона
Зона, содержащая данные о сопоставлении имён с адресами хостов, целями почтового обмена и т. п.

Reverse Zone - реверсная (обратная) зона
Зона, содержащие данные сопоставления адресов с именами.

Server - сервер
Реализация протоколов DNS, способная предоставлять ответы на запросы. Ответ может содержать сведения, известные серверу или полученные им от другого сервера.

Authoritative Server - полномочный сервер
Сервер, знающий содержимое зоны DNS из локальных сведений и способный, тем самым, отвечать на запросы для зоны без необходимости обращения к другим серверам.

Listed Server - сервер из списка
Полномочный сервер, для которого в зоне имеется запись NS RR.

Primary Server - первичный (основной) сервер
Полномочный сервер, для которого сведения о зоне настроены локально. Иногда применяется термин «ведущий» (Master).

Secondary Server - вторичный сервер
Полномочный сервер, получающий информацию о зоне от первичного сервера с помощью механизма переноса зон. Иногда применяется термин «ведомый» (Slave).

Stealth Server - скрытый сервер
Полномочный сервер (обычно вторичный), не указанный в списке (не Listed).

Resolver - распознаватель
Клиент DNS, ищущий информацию, содержащуюся в зоне, по протоколу DNS.

3. Вторичные серверы

Основная причина наличия нескольких серверов для каждой зоны заключается в обеспечении широкой и надёжной доступности сведений о зоне для клиентов Internet, т. е. по всему миру, даже при недоступности одного из серверов. Наличие нескольких серверов также обеспечивает распределение нагрузки и повышение общей эффективности системы за счёт размещения серверов ближе к распознавателям. Однако эти вопросы здесь не рассматриваются.

При использовании нескольких серверов один из них обычно является основным (primary), остальные - вторичными. Хотя в некоторых конфигурациях используется несколько первичных серверов, это может приводить к несогласованности данных и не рекомендуется.

Различие между первичным и вторичными серверами имеет значение лишь для рассматриваемой зоны, а для остальной части DNS будет просто несколько серверов, которые одинаково трактуются даже родительским сервером, делегирующим зону. Распознаватели часто измеряют производительность разных серверов, выбирают из них «лучший» по неким критериям и предпочитают его для большинства запросов. Это выполняется автоматически и здесь не рассматривается. На первичном сервере храниться основная (master) копия файла зоны. На этом сервере данные вводятся в DNS из неких источников, не относящихся к DNS. Вторичные серверы получают данные зон с использованием механизмов протокола DNS для переноса данных зоны с основного сервера.

3.1. Выбор вторичных серверов

При выборе вторичных серверов следует обратить внимание на возможные варианты отказов. Серверы следует размещать так, чтобы при любом вероятном отказе хотя бы 1 сервер был доступен для всех значимых частей Internet.

Поэтому размещение всех серверов на локальном сайте не является хорошим решением, хотя оно просто в организации и управлении. В случае отказа одного канала, сбоя питания сайта, здания или даже помещения при такой конфигурации все серверы могут отказаться отключёнными от Internet.

Вторичные сервер должны размещаться в топологически и географические разнесённых точках Internet для минимизации вероятности потери всех серверов в результате одного отказа. Т. е. вторичные серверы следует размещать в географически удалённых одно от другого местах, чтобы такие события, как потеря электропитания, не вывели из строя сразу все серверы. Их следует подключать к сети через достаточно разделённые каналы, чтобы отказ одного канала или нарушение маршрутизации в отдельном сегменте сети (например, у провайдера) не сделали все серверы недоступными.

3.2. Неподходящие конфигурации

Серверы имён не следует размещать в одном сегменте ЛВС одного помещения или здания, хотя на практике это встречается достаточно часто. Такие конфигурации практически сводят на нет полезность наличия нескольких серверов. Единственный вариант резервирования, предоставляемый такой конфигурацией, - это защита от полного выхода из строя одного сервера, но в реальности существует много других возможных отказов, таких как сбои питания, в том числе длительные.

3.3. Развенчание мифа

Иногда выдвигается аргумент о ненужности обеспечивать доступность серверов имён для домена, если хосты этого домена недоступны. Этот аргумент является ошибочным.

- Клиенты по разному реагируют на невозможность распознавания и невозможность подключения и реакция на первое не всегда желательна.
- Если зона распознаётся, а конкретное имя - нет, клиент может отказаться от транзакции и не вносить в сеть дополнительную нагрузку.
- Хотя позитивные отклики DNS обычно кэшируются, отсутствие результата кэшируется реже или не кэшируется совсем в зависимости от поведения кэш-сервера и локальной конфигурации. Таким образом, ненужная неспособность распознавания создаёт нежелательную нагрузку на сеть[1].
- Имена в зоне могут не распознаваться в адреса отключённой сети. Со временем это становится более вероятным. Таким образом, основное допущение мифа часто становится ложным.

Важно, чтобы серверы имён для прямых (forward) зон всегда были доступны для запросов.

[1] В оригинале этот абзац был иным, см. https://www.rfc-editor.org/errata/eid4631. *Прим. перев.*

4. Недоступные серверы

Другой класс проблем связан с указанием серверов, которые недоступны из больших частей сети. Это может быть указание машины, размещённой за межсетевым экраном, или указание вторичного адреса машины с несколькими адресами, недоступного извне. В записях NS следует указывать имена серверов, распознаваемые в адреса, доступные из региона, в который возвращаются записи NS. Включение адресов, недоступных для большой части сети, не добавляет надёжности и вызывает серьёзные проблемы, которые в конечном итоге могут снизить надёжность зоны.

Единственным способом определить фактическую недоступность адресов является попытка их использовать. При этом нужно дождаться тайм-аута отклика (или сообщения ICMP об ошибке), чтобы понять, что адрес не может использоваться. Кроме того, тайм-аут не позволяет отличить недоступность от обычной потери пакетов, поэтому требуются повторные попытки, чтобы убедиться в фактической недоступности адреса. Эти попытки и ожидание могут занять достаточно много времени, чтобы клиентская программа или пользователь поняли, что ответа нет, и убедились в отказе для зоны. Кроме того, указанные действия нужно время от времени повторять, чтобы отличить временную недоступность от постоянной.

Указанные действия должны повторять распознаватели всей сети, что будет вести к росту трафика и, вероятно, нагрузки на межсетевые экраны, блокирующие доступ. Эта дополнительная нагрузка лишь снижает надёжность сервиса.

4.1. Серверы за непостоянными соединениями

Похожая проблема возникает, когда серверы DNS размещены в тех частях сети, которые часто отсоединены от Internet, например, подключены через непостоянное соединение или соединение часто выходит из строя. К таким серверам обычно следует относиться как к размещённым за межсетевым экраном и недоступным в сети постоянно.

4.2. Другие проблемные ситуации

Похожие проблемы возникают и при наличии транслятора сетевых адресов (Network Address Translator или NAT) [RFC1631] между распознавателем и сервером. Несмотря на сказанное в [RFC1631], узлы NAT на практике не транслируют адреса, размещённые внутри пакетов, ограничиваясь лишь адресами в заголовках. Как отмечено в [RFC1631], трансляция создаёт некоторые проблемы для DNS. Иногда проблему можно решить путём замены NAT шлюзом прикладного уровня (Application Layer Gateway или ALG). Такое устройство будет понимать протокол DNS и транслировать все адреса в проходящих через него пакетах, но даже при наличии такого шлюза лучше использовать решение, описанное в следующем параграфе.

4.3. Решение

Чтобы избежать отмеченных выше проблем, в записях NS для зоны, возвращаемых в откликах, следует указывать только серверы, которые скорей всего будут доступны запросившему информацию распознавателю. Некоторые распознаватели являются в то же время серверами, выполняющими запросы от имени других распознавателей. Возвращать следует записи NS, доступные не только запросившему их распознавателю, но и другим распознавателям, которым сведения могут пересылаться. Все адреса всех возвращаемых серверов должны быть достижимыми. Поскольку адреса каждого сервера формируют набор записей о ресурсах (Resource Record Set или RRset) [RFC2181], должны возвращаться все адреса (или ни одного), поэтому неприемлемо сокрытие недоступных адресов или их возврат с малым значением TTL при больших TTL для других адресов.

Когда серверы размещаются за межсетевым экраном, непостоянным соединением или транслятором NAT, которые запрещают запросы или отклики DNS или вызывают иные проблемы, их имена или адреса не следует возвращать внешним (по другую сторону) клиентам. Точно также, внутренним клиентам не следует знать о внешних серверах, если такие клиенты не могут делать запросы к таким серверам. Для реализации этого обычно требуется двойная настройка DNS - одна для внутреннего, другая для внешнего использования. Это часто решает и другие проблемы в средах, подобных отмеченным.

Когда сервер размещается на границе межсетевого экрана и доступен с обеих сторон, но имеет разные адреса, для этого сервера следует задать два имени, каждое из которых связано с записями A, доступными лишь с соответствующей стороны межсетевого экрана. В этом случае сервер следует рассматривать как два сервера, каждый из которых размещён по одну сторону межсетевого экрана. Сервер, реализованный на ALG, обычно относится к этой же категории. Особое внимание следует уделить тому, чтобы сервер возвращал корректные отклики клиентам, размещённым по обе стороны. Т. е. возвращаться должны лишь сведения о хостах, доступных с соответствующей стороны и корректные для этой стороны адреса IP.

Для серверов в таких средах часто требуются специальные меры по обеспечению доступа к корневым серверам. Часто это достигается за счёт использования конфигурации с фиктивным корнем (fake root). В этом случае серверы следует надёжно изолировать от остальной системы DNS, чтобы их необычная конфигурация не мешала работе других серверов.

5. Сколько вторичных серверов нужно?

Спецификация DNS и правила регистрации доменных имён требуют наличия хотя бы 2 серверов имён для каждой зоны. Обычно это первичный и вторичный сервер. Хотя двух, внимательно размещённых, серверов зачастую достаточно, случаи, когда это не так, достаточно распространены, поэтому рекомендуется иметь более двух серверов. Различные проблемы могут приводить к достаточно долгой недоступности сервера и в такие промежутки времени зона с двумя серверами будет доступна лишь на одном. Поскольку и второй сервер может по какой-либо причине утратить доступность, зона время от времени может остаться совсем без серверов.

С другой стороны, наличие большого числа серверов не даёт заметных преимуществ, но увеличивает расходы. Проще говоря, большее число серверов ведёт к росту размера пакетов и большему расходу пропускной способности. Это может казаться и фактически является тривиальным утверждением. Однако имеется предел размера пакетов DNS и превышение этого предела оказывает существенное влияние на производительность. Разумнее не допускать этого.

Большое число серверов также повышает вероятность ошибок в настройке или некорректной работе одного из них, что может остаться незамеченным.

Для большинства зон уровня организации рекомендуется организовать 3 сервера, при этом хотя бы один должен быть достаточно удалён от других. Для зон с более высокими требованиями к доступности могут быть желательны 4 или даже 5 серверов. Два, а иногда три (из 5) можно разместить на локальном сайте, а остальные следует разнести достаточно далеко (топологически или географически) друг от друга и сайта.

Обратные зоны (субдомены .IN-ADDR.ARPA), обычно менее критичны и для них зачастую достаточно меньшего числа серверов и не требуется значительное удаление. Это обусловлено тем, что трансляция адресов в имена обычно требуется лишь при получении пакетов с интересующего адреса и только распознавателям вблизи места назначения или в нем. Это даёт некоторую гарантию того, что серверы, размещённые у источника пакетов или близко от него, например, в той же сети, будут доступны распознавателю при необходимости выполнить поиск. Таким образом, некоторые варианты отказов, которые важно учитывать при планировании серверов для прямых зон, могут быть менее актуальны при планировании обратных зон.

5.1. Скрытые серверы

Серверы, которые являются полномочными для зоны, но не указанные в записях NS (их называют также скрытыми - stealth) не учитываются в числе серверов. Зачастую бывает полезно, чтобы все серверы на сайте были полномочными (вторичными), но для всех локальных зон указывались только один или два из них (остальные скрыты). Это позволяет таким серверам напрямую отвечать на локальные запросы без обращения к другому серверу. При необходимости обращаться к другому серверу обычно требуется запрос к корневому серверу для просмотра дерева делегирования - то, зона является локальной, не было бы известно. Это означает, что при нарушении внешних коммуникаций некоторые локальные запросы остались бы безответными.

Указание всех таких серверов в записях NS (если этих серверов больше 1 или 2) заставило бы остальную часть Internet тратить ненужные усилия на попытки связаться со всеми серверами сайта, кода тот становится недоступным из-за отказа канала или маршрутизации.

6. Нахождение подходящих вторичных серверов

Операции вторичного сервера почти всегда выполняются автоматически. После организации сервера он обычно работает самостоятельно, основываясь на действиях первичного сервера. Поэтому многие организации с готовностью отзываются на просьбы организации вторичного сервера. Обычно лучше всего найти организацию близкого размера и договориться с ней об обмене вторичными зонами - каждая организация соглашается предоставить сервер для работы в качестве вторичного для зон другой организации. Отметим, что при этом не происходит утечки конфиденциальных данных, поскольку обмен осуществляется лишь для общедоступных данных, независимо от серверов, где они размещаются.

7. Поддержка порядковых номеров

Вторичные серверы используют порядковый номер в SOA-записи зоны для определения необходимости обновить свою локальную копию зоны. Порядковые номера - это, по сути, просто 32-битовые целые числа без знака, сбрасываемые в 0 при достижении максимального значения. Более строгое определение порядкового номера дано в [RFC1982].

Серийный номер должен увеличиваться при каждом внесении изменений в зону первичного сервера. Это информирует вторичные серверы о необходимости обновить их копию зоны. Отметим, что уменьшать порядковый номер нельзя и единственным разрешённым изменением является увеличение номера.

Иногда из-за ошибок при редактировании или по иной причине может возникнуть необходимость уменьшить значение порядкового номера. Делать этого не следует никогда! Вторичные серверы в этом случае проигнорируют последующие инкременты, пока не будет достигнуто значение номера, превышающее прежнее (при котором произошло уменьшение). Вместо уменьшения номера следует увеличивать его неоднократно, пока не будет достигнуто желаемое значение (с учётом сброса после максимума). На каждом этапе увеличения номера следует дождаться обновления номера на вторичных серверах.

Предположим, что зона имела порядковый номер 10, но он был случайно увеличен до 1000, а затем потребовалось снизить значение до 11. Не следует менять значение 1000 сразу на 11, поскольку вторичный сервер, увидевший номер 1000 (на практике хотя бы один из серверов увидит его), будет игнорировать это изменение и продолжит использовать зону с номером 1000, пока порядковый номер на первичном сервере не превзойдёт это значение. Это может занять много времени и у вторичного сервера может завершиться срок действия копии зоны до этого события. В рассматриваемом случае для решения задачи следует установить на первичном сервере порядковый номер 2000000000 и дождаться, пока вторичные серверы обновятся с этим номером. Значение 2000000000 выбрано потому, что оно значительно больше текущего, но не более, чем на 2^{31} (2147483648). Это будет инкрементом порядкового номера [RFC1982]. После того, как все нуждающиеся в обновлении серверы получат зону с этим номером, можно установить номер 4000000000, который на 2000000000 превышает установленное ранее значение 2000000000, и это будет ещё один инкремент на величину меньше 2^{31}. Когда эта копия файла зоны будет получена всеми серверами, можно просто установить порядковый номер 11. В арифметике порядковых номеров смена значения 4000000000 на 11 будет инкрементом. Порядковый номер сбрасывается (переход через максимум) при значении 2^{32} (4294967296), поэтому значение 11 эквивалентно 4294967307 (4294967296 + 11). Значение 4294967307 на 294967307 больше, чем 4000000000, а 294967307 меньше 231, поэтому такое изменение будет инкрементом.

При выполнении описанной процедуры важно убедиться, что на каждом этапе обновились все серверы. Отказ от такой проверки может привести к ещё большему беспорядку, чем при простом уменьшении порядкового номера. Также следует помнить, что важны не абсолютные значения порядковых номеров, а их изменение относительно прежнего значения. Приведённые выше значения инкремента служат лишь примером.

Практически во всех случаях можно исправить порядковый номер в два этапа, если использовать больший инкремент, но это будет не столь элегантно и потребует большей осторожности.

Следует также отметить, что не все серверы имён корректно реализуют операции с порядковыми номерами. На таких вторичных серверах обычно нет возможности аккуратно уменьшить порядковый номер, иначе чем связаться с администратором вторичного сервера и запросить очистку всех имеющихся файлов зоны. После этого вторичный сервер загрузить файл зоны с первичного как при первом использовании. Описанная процедура безопасна, поскольку некорректно работающие серверы в любом случае требуют внесения изменений вручную. После описанной последовательности смены порядкового номера соответствующие вторичные серверы перезапускаются. Затем, когда на первичном сервере установлен корректный (желаемый) порядковый номер, следует связаться с оставшимися вторичными серверами и попросить администраторов вручную исправить порядковый номер. Можно также предложить им обновить программы сервера до соответствующих стандартам реализаций.

Сервер, не реализующий этот алгоритм, будет неисправным и его можно обнаружить, как описано ниже. На некотором этапе, обычно когда абсолютное значение порядкового номера уменьшается, сервер с дефектом будет игнорировать это изменение. Для обнаружения этого следует подождать в течение времени не меньше указанного в SOA, а затем запросить SOA. На сервере с дефектом останется прежний номер. Других способов обнаружить дефект авторам не известно.

Вопросы безопасности

Считается, что ничего в этом документе не усугубляет проблем безопасности DNS и не сокращает их.

Администраторам следует помнить, что компрометация сервера для домена в некоторых случаях может создавать угрозы безопасности хостов домена. Следует тщательно выбирать вторичные серверы, чтобы минимизировать такие угрозы.

Литература

[RFC1034] Mockapetris, P., "Domain Names - Concepts and Facilities", STD 13, RFC 1034, November 1987.

[RFC1035] Mockapetris, P., "Domain Names - Implementation and Specification", STD 13, RFC 1035, November 1987

[RFC1631] Egevang, K., Francis, P., "The IP Network Address Translator (NAT)", RFC 1631, May 1994

[RFC1982] Elz, R., Bush, R., "Serial Number Arithmetic", RFC 1982, August 1996.

[RFC2181] Elz, R., Bush, R., "Clarifications to the DNS specification", RFC 2181, July 1997.

Благодарности

Brian Carpenter и Yakov Rekhter предложили упомянуть NAT и ALG в дополнение к тексту о межсетевых экранах. Dave Crocker предложил развенчать этот миф.

Адреса авторов

Robert Elz
Computer Science
University of Melbourne
Parkville, Vic, 3052
Australia.
EMail: kre@munnari.OZ.AU

Randy Bush
RGnet, Inc.
5147 Crystal Springs Drive NE
Bainbridge Island, Washington, 98110
United States.
EMail: randy@psg.com

Scott Bradner
Harvard University
1350 Mass Ave
Cambridge, MA, 02138
United States.
EMail: sob@harvard.edu

Michael A. Patton
33 Blanchard Road
Cambridge, MA, 02138
United States.
EMail: MAP@POBOX.COM

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru