

Распознавание адресов в сетях с множеством ЛВС Multi-LAN Address Resolution

Статус документа

Этот документ был инициирован Jeffery Mogul, который в RFC 917 Internet Subnets, рассмотрел использование «явных подсетей» (explicit subnet) в средах с множеством ЛВС. В этом документе я пытаюсь рассмотреть случай «прозрачных подсетей» (transparent subnets). Данный документ описывает предложенный сообществу ARPA-Internet протокол и служит приглашением к дискуссии в целях развития протокола. Документ может распространяться свободно.

Введение

Проблема трактовки множества локальных сетей (ЛВС), как одной сети Internet¹, вызывает некоторый интерес и озабоченность. Выделять отдельную сеть IP для каждой ЛВС на сайте было бы непрактично. Детали соединений между ЛВС на сайте желательно скрыть от пользователей, шлюзов и хостов, находящихся за пределами сайта. Возникает вопрос - как это лучше сделать или хотя бы просто сделать. Одним из вариантов является использование явных подсетей [1]. Схема с явными подсетями предлагает рекурсивное применение механизмов управления сетями Internet к управлению ЛВС в рамках одной сети. В этом документе предложена другая модель - прозрачные подсети - на основе расширения протокола ARP [2] для множества ЛВС.

Обзор

Рассмотрим кратко протокол преобразования адресов ARP². Каждый хост в ширококвещательной ЛВС знает свой физический аппаратный адрес НА в ЛВС и спой адрес IP (IA). Когда хост А знает IA для хоста В и хочет отправить тому дейтаграмму, хост А должен найти НА, соответствующий IA хоста В. Для решения этой задачи хост А формирует пакет ARP, содержащий значения НА и IA данного хоста, а также IA искомого хоста В). Этот пакет ARP хост А передает в ширококвещательном режиме. Хосты, получившие этот пакет ARP проверяют соответствие искомого адреса их собственному. При совпадении (только на хосте В) адресов хост отвечает отправителю запроса (хост А) откликом со своим адресом НА. В результате хост А будет знать оба адреса (НА и IA) получателя (хост В). Полученную информацию хост А сохраняет в своей таблице для использования в будущем.

Примечание. Задачи протокола ARP в реальности шире, чем описано в этом примере.

Идея данного документа заключается в расширении ARP для работы в среде с множеством соединенных ЛВС.

Предположим, что некий «магический хост» (BOX) подключен одновременно к двум (или более) ЛВС.

Поведение хостов в точности соответствует базовому ARP.

Когда любой из хостов передает ширококвещательный запрос ARP, BOX читает этот запрос (как и остальные хосты ЛВС). В дополнение к обычной проверке соответствия адресов (и ответа при совпадении), BOX проверяет в своем кэше для каждой из подключенных ЛВС наличие отображения IA:НА.

Случай 1. Если отображение для хоста найдено в кэше ЛВС, из которой поступил запрос, BOX не отвечает на такой запрос (позволяя ответить хосту с соответствующим адресом).

Случай 2. Если отображение найдено в кэше не той ЛВС, откуда пришел запрос, BOX передает отклик со своим адресом НА в ЛВС, из которой поступил запрос. В данном случае BOX выступает, как агент искомого хоста.

Случай 3. Если отображение отсутствует в таблицах для всех ЛВС, BOX должен попытаться узнать этот адрес и реагировать на запрос, как в случае 1 или 2.

В случае 3 BOX делает кое-что «магическое».

BOX хранит список поиска хостов. Каждая запись списка включает IA искомого хоста, идентификатор интерфейса, через который получен запрос ARP, и адреса отправителя исходного запроса. В описанном выше случае 3 хост обращается к списку поиска. Если адрес искомого хоста присутствует в списке, поиск завершается (отправителю исходного запроса при этом ничего не передается и запись сохраняется в списке).

Если адреса в списке не обнаружено, сначала в создается пустая запись, затем BOX создает и отправляет пакет ARP через каждый интерфейс, кроме того, через который был принят исходный запрос. При получении отклика данные вносятся в соответствующий кэш, запись удаляется из списка поиска и для исходного запроса создается и отправляется отклик ARP, как в случае 1³ или 2. Если отклика не было получено, отправителю исходного запроса ничего не передается и запись сохраняется в списке поиска.

Для записей в кэше и списке поиска должен устанавливаться срок действия (тайм-аут).

При получении каждого запроса ARP хост BOX должен включать отображение адресов IA:НА в кэш для ЛВС, из которой запрос был принят.

¹В современной терминологии «сети IP». Прим. перев.

²Address Resolution Protocol.

³Очевидная неточность. Реакции, как в случае 1 здесь возникнуть не может, поскольку в сеть, откуда пришел исходный запрос BOX не отправляет своего запроса и, следовательно, не может получить отклика. Прим. перев.

Протокол ARP для множества соединенных ЛВС

План заключается в применении протокола ARP в его исходном виде. Новым элементом является «магический хост» (мост на базе ARP), который транслирует запросы ARP в соседние ЛВС и выступает в качестве агента по трансляции дейтаграмм хостам других ЛВС.

Детали

Хосты работают точно так же, как при обычном ARP.

ЛВС соединяются между собой хостами BOX (компьютер, подключенный по крайней мере к двум ЛВС). BOX реализует описанные ниже процедуры.

Каждый хост BOX хранит таблицу для каждой ЛВС, к которой он подключен (или для каждого интерфейса в ЛВС). Записи в таблице хранятся ограниченное время, поэтому таблицы представляют собой кэш свежей информации. Записи таблицы представляют собой пары адресов IA:HA для данной ЛВС.

При передаче широковещательного запроса ARP любым из хостов BOX получает этот запрос (как и остальные хосты данной ЛВС). В дополнение к сравнению искомого адреса в запросе со своим адресом (и отклику при совпадении) BOX просматривает в своем кэше отображения IA:HA для всех ЛВС, к которым он подключен.

Случай 1. Если отображение для хоста найдено в кэше ЛВС, из которой поступил запрос, BOX не отвечает на такой запрос (позволяя ответить хосту с соответствующим адресом). Отсчет тайм-аута для записи не возобновляется.

Случай 2. Если отображение найдено в кэше не той ЛВС, откуда пришел запрос, BOX передает отклик со своим адресом HA в ЛВС, из которой поступил запрос. Отсчет тайм-аута для записи не возобновляется.

В этом случае BOX показывает, что он будет выступать в качестве агента искомого хоста. При поступлении дейтаграммы IP на BOX этот хост попытается переслать ее, используя информацию из кэша отображений.

Случай 3. Если отображение отсутствует в таблицах для всех ЛВС, BOX должен попытаться узнать этот адрес и реагировать на запрос, как в случае 1 или 2. В этом случае BOX делает кое-что «магическое».

BOX хранит список поиска хостов (искомых, но не найденных). Каждая запись такого списка включает IA искомого хоста, интерфейс ЛВС, с которого был получен запрос ARP, и адреса отправителя исходного запроса.

В случае 3 просматривается список поиска. При обнаружении адреса искомого хоста в списке поиск прекращается.

Если адрес отсутствует в списке, поиск продолжается путем создания хостом BOX записи в списке поиска, а затем созданий запроса ARP и отправки его через все интерфейсы¹. Эти запросы ARP включают адреса IA и HA хоста BOX, а также адрес IA искомого хоста и запрашивают адрес HA для искомого хоста. При получении отклика на запрос ARP информация из него вносится в соответствующий кэш, запись удаляется из списка поиска и BOX реагирует на исходный запрос ARP, как в описанных выше случаях 1 и 2. Если отклика не получено, на исходный запрос не передается никакого отклика и запись сохраняется в списке поиска.

Отметим, что хосту BOX следует соблюдать меру при отправке запросов ARP. Для обычных хостов нормальным считается повторение запросов ARP до 5 раз и BOX должен придерживаться этого ограничения.

При прерывании поиска отправителю запроса не передается никакого отклика и запись сохраняется в списке поиска.

Для запросов ARP не предусмотрено отрицательной обратной связи, поэтому единственным критерием констатации неудачного поиска является тайм-аут.

Для каждого полученного запроса ARP хост BOX должен также помещать отображение IA:HA для передавшего запрос хоста в кэш для ЛВС, из которой получен запрос.

Время присутствия записей в кэшах и списке поиска должно быть ограничено (тайм-аут).

Нужно сохранять список поиска и следовать правилам для предотвращения бесконечной трансляции запросов ARP для хоста, который не отвечает. Если хост указан в списке поиска, запросы ARP для него не будут транслироваться. Если хост, который не работал (или по каким-то причинам не отвечал на запросы ARP), восстанавливает работоспособность (начинает отвечать на запросы ARP), он по-прежнему будет не доступен для хостов других ЛВС, пока не наступит тайм-аут для записи в таблице.

Для этой проблемы имеется два решения - 1) задавать достаточно короткий тайм-аут и 2) периодически отправлять с хоста BOX запросы ARP для каждой записи из списка.

В описанной схеме используется несколько тайм-аутов.

1. Время, в течение которого хост пытается определить адрес с помощью ARP. На практике может выполняться несколько попыток до принятия решения о том, что хост не отвечает. Нужно тщательно оценить время, в течение которого хост может повторять попытки. Обозначим это время T1.
2. Время, в течение которого запись сохраняется в списке поиска, или время между генерацией хостом BOX пакетов ARP для преобразования данного адреса. Обозначим это время T2.
Отметим, что T2 должно превышать сумму значений T1 для самой длинной «петли» ЛВС.
3. Время хранения в кэше записей для каждой ЛВС. Обозначим это время T3.

Должно выполняться соотношение $T1 < T2 < T3$.

Предлагается устанавливать для T1 значение меньше 1 минуты, для T2 - 10 минут, T3 - 1 час.

¹Явное несоответствие сказанному на предыдущей странице и здравому смыслу. Отправлять запрос в ЛВС, откуда он был принят никакого смысла нет. *Прим. перев.*

Если среда достаточно статична, увеличение T3 будет снижать число операций поиска (и трафик ARP). В динамичной среде уменьшение T3 будет ускорять адаптацию к изменениям.

Другим вариантом является перезапуск таймера для записи в кэше при каждом обращении к ней и малое значение для T3. Это приведет к тому, что востребованные записи будут долго сохраняться в кэше, а редко используемые - будут быстро удаляться. К сожалению нет каких-либо корреляций между частотой использования и корректностью. Этот метод может приводить к длительному сохранению в кэше устаревших записей, если для такой записи запросы будут повторяться чаще одного раза за период тайм-аута.

При обслуживании обычных дейтаграмм устройства BOX должны уменьшать значение поля TTL¹ и обновлять контрольную сумму заголовка IP. При нулевом значении TTL дейтаграмма отбрасывается (не пересылается).

ARP, в соответствии с текущим определением, будет считать недавнюю информацию лучшей и наиболее актуальной. В более сложной среде с множеством ЛВС, где имеются петли в связности, очевидно наличие двух или более откликов на запрос ARP для хоста из другой ЛВС. Возможно, что первым будет отклик от устройства BOX на более эффективном пути.

Единственным изменением в реализации ARP в соответствии с предложениями данного документа является предотвращение замены отображений, полученных из первого отклика.

Возможные проблемы

Некорректные записи в кэше

Если в кэше появилась та или иная некорректная информация, она будет сохраняться в течение времени T3. Сохранение в кэше устаревшей информации может приводить к (временной) потере связи с хостом, если у него изменится отображение IA:HA.

Одним из способов замены устаревших записей в кэше является явная интерпретация устройством BOX широковещательного отклика ARP, как требования замены записи для данного IA или HA в соответствии с новым отображением IA:HA. Важно, чтобы серверы передавали широковещательные отклики ARP, когда они приходят.

Хосты без поддержки ARP

Не представляется реальным сосуществование и взаимодействие в одной ЛВС хостов, поддерживающих и не поддерживающих ARP. Случай, когда все не поддерживающие ARP хосты собраны в одной ЛВС, рассматривается в следующем параграфе (ЛВС без широковещания).

Хосты, не поддерживающие ARP, должны использовать какой-то иной механизм отображения адресов. Они могут поддерживать полную таблицу хостов, пользоваться такой таблицей на некоем сервере с использованием того или иного протокола или принимать все решения о маршрутизации на основе анализа адресных полей.

ЛВС без широковещания

Устройства BOX, подключенные к ЛВС без поддержки широковещания и/или ЛВС с хостами, не отвечающими на ARP, могут применять статические или динамические таблицы отображений IA:HA для таких ЛВС LAN (или один адрес может рассчитываться на основе другого). Все хосты такой ЛВС должны быть включены в таблицу.

Когда устройству BOX нужно найти адресное отображение и ему приходится отправлять запрос ARP в ЛВС без поддержки широковещания (это может произойти только для хоста, не входящего в ЛВС без поддержки широковещания, поскольку все хосты таких ЛВС включены в таблицу), оно должно вместо этого отправить запросы всем другим устройствам BOX в данной ЛВС.

Размер таблиц

Наихудшим случаем для размера таблиц в BOX будет по записи для каждого хоста всего множества ЛВС. Т. е., таблицы, хранящиеся для интерфейсов в каждую ЛВС могут (в худшем случае) содержать по записи для каждого хоста всех ЛВС. Однако в реальности в таблицах кэшируются записи, потребные для текущей коммуникационной активности и упомянутый худший случай не является типичным. Большинство хостов будет взаимодействовать в основном с хостами своей ЛВС, и иногда с небольшим числом хостов из других ЛВС. Большая часть сетевой активности будет состоять из обмена данными между рабочими станциями и серверами. Можно считать, что большая часть взаимодействий будет включать основные серверные хосты и записи для этих хостов будут находиться постоянно в таблицах устройств BOX.

Возникновение петель

Возникновение бесконечных коммуникационных петель через соединенные между собой ЛВС предотвращается путем хранения списков поиска в устройствах BOX и прерывания поиска при наличии искомого адреса в таком списке.

Коммуникационные петли для обычных дейтаграмм не могут возникать по причине уменьшения устройствами BOX значений TTL и отбрасывания дейтаграмм с TTL = 0. Для отлаживания полезно, чтобы устройства BOX сообщали а таких фактах отбрасывания дейтаграмм.

Широковещание

Отметим, что широковещание не имеет ничего общего с прозрачными или явными подсетями. Поскольку этот вопрос рассматривался в работе [1], обсудим его и здесь. Две из трех рассмотренных в [1] функций широковещания работают в нашем случае точно так же, и третья тоже может поддерживаться.

Утверждается также, что вопрос поддержки широковещательной интерпретации адресов IA сложнее вопроса выбора между явными и прозрачными подсетями и должен рассматриваться отдельно.

¹Time-To-Live - время жизни.

Отмечено также, что широковещание не является эффективным и лучше использовать групповую адресацию (multicast). Может иметь смысл выяснение вопросов групповой адресации Internet до принятия схемы с широковещанием.

Данная сеть IP

Если используется IA, состоящий из номера данной сети и номера хоста, содержащего все 1 (например, 36.255.255.255), подразумевается широковещательная отправка всем хостам данной сети IP (все ЛВС). Устройства BOX должны пересылать такую дейтаграмму. Кроме того, устройство BOX должно проверять, не предназначена ли дейтаграмма и ему.

Для предотвращения бесконечных петель каждое устройство BOX должно сохранять список недавних широковещательных пакетов. Записи этого списка включают IA отправителя и значение поля Identification из заголовка дейтаграммы. Если полученный широковещательный пакет соответствует записи в списке, такой пакет отбрасывается без пересылки. Записи сохраняются в списке в течение времени T2.

Только данная ЛВС

Если адрес IA содержит только 1 (т. е., 255.255.255.255) это означает широковещательную передачу IP всем хостам только данной ЛВС. Устройствам BOX недопустимо пересылать такие дейтаграммы. Кроме того, устройство BOX должно проверять, не предназначена ли дейтаграмма и ему.

Только другие ЛВС

Поскольку ЛВС не идентифицируются по IA, такую адресацию невозможно задать обычным путем. Некоторые даже считают поддержку подобной адресации глупостью.

Одним из способов такой адресации будет организация специального IA для каждой ЛВС, который будет означать «широковещательный пакет для ЛВС». Например, 36.255.255.128 определяет широковещательный пакет для ЛВС А, 36.255.255.187 - широковещательный пакет для ЛВС В и т. д. Эти адреса будут соответствующим образом интерпретироваться устройствами BOX, подключенными к соответствующим ЛВС, где адрес имеет специальное значения, а прочие устройства BOX будут относиться к ним, как к обычным адресам IA. Так, где происходит специальная обработка таких адресов, они должны преобразовываться в широковещательные адреса «только для данной ЛВС».

Обсуждение

Заявление для расширенной схемы ARP заключается в том, что среднему хосту даже не нужно знать, что он работает в среде с множеством ЛВС.

Если хост проанализирует свой локальный кэш отображений IA:АН, он может обнаружить, что несколько адресов IA отображаются на один адрес НА. А если он проведет временной анализ, то может увидеть, что некоторые хосты отвечают быстрее других. Более того, он может увидеть корреляции между этими наблюдениями. Но лишь немногие хосты делают это.

Структура адреса

В схеме с явными подсетями некоторые биты IA предназначены для идентификации подсети (т.е., ЛВС). Адрес делится на поля сети, подсети и хоста. Как правило, такое выделение полей снижает эффективность использования адресного пространства. Могут возникать существенные проблемы в реализациях при появлении новых сетей и связанной с этим необходимости изменения размера поля подсети. На сегодняшний день схема с явными подсетями представляется непрактичной для адресов IA класса С.

В схеме с расширенным ARP адрес включает только поля для сети и хоста. Расширение ARP подходит для IA всех классов.

Перемещение хостов

В схеме с явными подсетями перенос хоста из одной ЛВС в другую требует смены адреса IA.

В схеме с расширенным ARP адрес IA можно не менять.

При внимательном рассмотрении ситуации можно заметить две реальных проблемы.

1. Как хост определяет местоположение адресата (та же или другая ЛВС)?

Этот вопрос предполагает, что хосту следует понимать различия и по-разному вести себя в двух случаях, а также то, что знающий ответ хост будет понимать, как передавать данные (напрямую хосту или через BOX).

В предлагаемом здесь варианте хосты не обязаны различать эти два случая и могут вести себя одинаково в обоих.

2. Как устройство BOX, соединяющее ЛВС узнает, какие устройства BOX ведут в какие ЛВС?

Этот вопрос предполагает, что устройства BOX должны понимать топологию сети и обмениваться по специальному протоколу (BOX-to-BOX) данными о связности.

Предлагаемый здесь вариант не требует от BOX знания топологии сети и наличия сведений о наличии других устройств BOX.

Предполагается существование двух проблем - 1) маршрутизация на хостах и 2) маршрутизация в BOX. Предлагается для решения каждой проблемы использовать конкурирующие стратегии и возможность использования частей обеих стратегий при выборе решения.

Например, использовать ARP внутри ЛВС и передавать отклики ARP с устройств BOX, действующих в качестве агентов (схема расширенного ARP), наряду с применением протокола взаимодействия между устройствами BOX для получения информации о местоположении хостов самими устройствами BOX (схема с явными подсетями).

Программный код вовлечен в эту задачу в двух местах - на многочисленных хостах и на небольшом числе устройств BOX. При выборе между схемами с явными подсетями и расширением ARP следует взвешенно оценивать объем работы, выполняемой на хостах и устройствах BOX.

Что делают хосты?

Схема с явными подсетями

Хост должен быть способен решить вопрос о принадлежности адреса IA к его или другой ЛВС. Для адресов из той же сети используется некая процедура для определения адреса HA. Если адрес относится к другой ЛВС выполняется некая процедура получения HA от устройства BOX.

Схема с расширенным ARP

Во всех случаях для получения отображения IA:HA используется ARP.

Что делают устройства BOX?

Схема с явными подсетями

Устройство BOX должно быть способно определить, к какой из ЛВС сайта относится адресат. Устройство BOX должно иметь некую таблицу маршрутизации, которая говорит, через какой интерфейс следует передавать дейтаграммы в каждую ЛВС. Эта таблица маршрутизации должна обновляться тем или иным способом (возможно с помощью протокола типа Internet Gateway-to-Gateway).

Схема с расширенным ARP

Устройство BOX должно сохранять кэш-таблицы с отображениями IA:HA для каждого интерфейса и таблицу поиска. Не требуется использования специальных протоколов (BOX-to-BOX) и даже не нужно знать о существовании других устройств BOX.

Топология и сложность реализации

Дерево

Если ЛВС и устройства BOX соединены в древовидную структуру, устройства BOX могут быть очень простыми и им не нужно сохранять списков поиска, поскольку в сети нет петель, препятствующих запросам ARP.

Петли

При наличии в сети петель поддержка списков поиска становится важной. Если топология сбалансирована так, что больших петель в сети нет (все петли примерно одного размера) и ЛВС сравнительно близки по параметрам задержки, описанная здесь процедура будет работать хорошо.

Комплексная топология

В сети с комплексной, несбалансированной топологией и смешением разнотипных (по задержкам) ЛВС может оказаться более эффективным использование протокола маршрутизации BOX-to-BOX.

Заключение

Будет хорошо, если сообщество Internet достигнет того или иного соглашения по решению проблемы сетей со множеством ЛВС и настоятельно рекомендует производителям рабочих станций придерживаться этого решения.

Предлагается рассмотреть описанную здесь схему расширенного ARP.

Предполагается, что большая часть рабочих станций подключена к ЛВС с поддержкой широковещания и большинство рабочих станций будет применяться в ситуациях, не требующих явных подсетей, и, кроме того, большинство станций будет применяться там, где возможно использование адресов IP класса C (и явные подсети невозможны). Таким образом, предполагается, что лучше будет попросить производителей включить поддержку ARP на рабочих станциях. Предполагается также, что будут созданы, протестированы и произведены «магические устройства» (magic box), предложенные здесь.

Следует отметить, что ни в данном документе, ни в [1] не предлагается конкретная процедура маршрутизации или протокол BOX-to-BOX. Это обусловлено тем, что задача маршрутизации является весьма сложной. Предложенный здесь план позволяет начать работу в сетях со множеством ЛВС без такого протокола. Если позднее будет принято решение о целесообразности использования процедуры маршрутизации между устройствами BOX, это не потребует вносить какие-либо изменения на хостах.

Глоссарий

ARP

Address Resolution Protocol - протокол преобразования адресов (см. [2]).

BOX

Magic Box - устройство (компьютер), подключенное к множеству ЛВС одной сети. Используется также термин ARP-мост (ARP-based bridge).

Bridge - мост

Узел (компьютер), подключенный к нескольким административно не различаемым, но разделенным физически подсетям, который при необходимости автоматически пересылает дейтаграммы между подсетями, оставаясь невидимым для других хостов. Используется также термин «программный повторитель» (software repeater).

Datagram - дейтаграмма

Коммуникационный блок данных на уровне IP.

Explicit Subnet - явная подсеть

Подсеть, явно идентифицируемая в адресе IA полем номера подсети и, таким образом, видимая другим как в данной сети, так и за ее пределами.

Gateway - шлюз

Узел (компьютер), подключенный к двум или более административно различающимся сетям (подсетям) и пересылающий дейтаграммы между хостами этих сетей.

HA

Аппаратный адрес, используемый в пакетах ЛВС.

Host Number - номер хоста

Адрес хоста в данной сети, выражаемый младшими битами IA.

IA

Адрес протокола IP.

Internet

Множество соединенных между собой сетей IP.

IP

Протокол Internet (см. [3]).

LAN - ЛВС

Локальная вычислительная сеть.

Multi-LAN Network - сеть с множеством ЛВС

Множество ЛВС, соединенных в единую сеть (т. е., использующих общий номер сети). Отдельные ЛВС сети могут быть явными или прозрачными подсетями.

Network - сеть

Единая сеть IP (возможно, разделенная на подсети или состоящая из множества ЛВС), идентифицируемая номером.

Network Number - номер сети

Номер сети IP, задаваемые старшими битами IA.

Packet - пакет

Единица обмена информацией на аппаратном уровне ЛВС.

Subnet - подсеть

Часть сети IP, выделенная логически или физически.

Transparent Subnet - прозрачная подсеть

Подсеть, не идентифицируемая адресом IA и, таким образом, не видимая другим (см. Сети с множеством ЛВС).

TTL

Поле Time-To-Live в заголовке IP.

Литература

- [1] J. Mogul, "Internet Subnets", [RFC 917](#), Stanford University, October 1984.
- [2] D. Plummer, "An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses to 48-bit Ethernet Addresses for Transmission on Ethernet Hardware", [RFC 826](#), Symbolics, November 1982.
- [3] J. Postel, "Internet Protocol", [RFC 791](#), USC-ISI, September 1981.

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru