Network Working Group
Request for Comments: 5195
Category: Standards Track

H. Ould-Brahim
D. Fedyk
Nortel
Y. Rekhter
Juniper Networks

June 2008

Автоматическое детектирование L1 VPN на основе BGP

BGP-Based Auto-Discovery for Layer-1 VPNs

Статус документа

Этот документ представляет проект стандартного протокола для сообщества Internet и служит приглашением к дискуссии в целях развития и совершенствования протокола. Информацию о текущем состоянии стандартизации протокола можно найти в документе Internet Official Protocol Standards (STD 1). Документ можно распространять без ограничений.

Аннотация

Целью настоящего документа является определение механизма автоматического детектирования Layer-1 VPN (L1VPN) на основе протокола BGP. Механизм автоматического детектирования для L1VPN позволяет устройствам сети провайдера автоматически находить набор PE[1], имеющих порты, подключённые к устройствам CE[2], которые входят в одну сеть VPN. Эта информация требуется для завершения сигнальной фазы соединений L1VPN. Одной из основных задач механизма автоматического детектирования L1VPN является поддержка модели single-end provisioning, в которой добавление порта в данную L1VPN будет вызывать изменения конфигурации только того устройства PE, которое включает затронутый соединением порт, и устройства CE, подключённого к PE через этот порт.

1. Введение

Целью настоящего документа является определение механизма автоматического детектирования Layer-1 VPN [L1VPN-FRMK]. Механизм автоматического детектирования для L1VPN позволяет устройствам сети провайдера автоматически находить набор PE, имеющих порты, подключённые к устройствам CE, которые входят в одну сеть VPN. Эта информация требуется для завершения сигнальной фазы соединений L1VPN. Одной из основных задач механизма автоматического детектирования L1VPN является поддержка модели "single-end provisioning", в которой добавление порта в данную L1VPN будет вызывать изменения конфигурации только того устройства PE, которое включает затронутый соединением порт, и устройства CE, подключённого к PE через этот порт.

Механизм автоматического детектирования обеспечивается за счёт того, что PE анонсирует другим устройствам PE по крайней мере свой адрес IP и список локальных для данного PE пар <private address, provider address>[3]. После получения этой информации устройства PE будут будут знать список членов VPN, связанных с данным PE, и использовать переданную механизмом автоматического детектирования информацию для преобразования адресов во время сигнальной фазы соединений L1VPN.

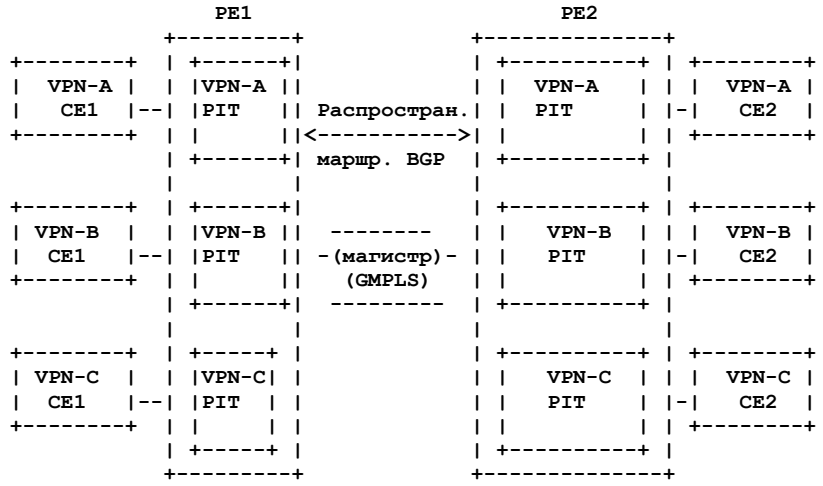

Рисунок 1. BGP Auto-Discovery для L1VPN.

Рисунок 1 показывает схему работы механизма автоматического детектирования для L1VPN на основе BGP. Для работы механизма автоматического детектирования требуется активизация BGP только в сети провайдера. Устройства PE поддерживают для каждой VPN информационные таблицы PIT[4], относящиеся к парам <private address, provider address>. Дополнительная информация о таблицах PIT приведена в разделе 2.

В документе [L1VPN-FRMK] описаны два режима работы L1VPN - базовый и расширенный (enhanced). В настоящем документе рассматривается механизм автоматического детектирования только для базового режима.

[1]Provider Edge - краевое устройство провайдера.
[2]Customer Edge - краевое устройство пользователя.
[3]Приватный адрес - адрес провайдера.
[4]Port Information Table - таблица информации для порта.

1.1. Термины

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с [RFC2119].

2. Процедуры

В контексте L1VPN устройства CE подключаются к PE через один или множество портов и каждый порт может включать один или множество каналов или субканалов. Каждый порт CE, соединяющий CE с PE имеет уникальный в масштабе L1VPN идентификатор (значения идентификаторов в разных L1VPN могут перекрываться). Этот идентификатор мы будем называть CPI[1]. Каждый порт PE также имеет уникальный в рамках сети провайдера идентификатор. Будем обозначать эти идентификаторы PPI[2]. Отметим, что для CPI и PPI могут использоваться адреса IPv4 или IPv6.

Для каждой L1VPN, которая имеет на PE хотя бы один сконфигурированный порт, PE поддерживает таблицу PIT, содержащую список пар <CPI, PPI> для всех портов в L1VPN. Отметим, что PIT может также включать маршрутную информацию (например, когда значения CPI определяются с использованием протокола маршрутизации).

Таблица PIT для данного PE включает два типа информации:

- информация, связанная с портами CE, подключёнными к PE; эта информация может быть задана в локальной конфигурации PE или получена от CE;
- информация, полученная от других PE с использованием механизма автоматического детектирования.

Информацию первого типа будем называть локальной, а второго - удалённой. Распространение локальной информации другим PE осуществляется с использованием многопротокольного расширения BGP [RFC4760]. Для ограничения потока такой информации только PIT, относящимися к данной L1VPN, мы будем использовать фильтрацию маршрутов BGP на основе Route Target Extended Community [BGP-COMM], как описано ниже.

Для каждой таблицы PIT в конфигурации устройства PE задается по крайней мере одна группа Route Target, которую будем называть export Route Targets, - эта группа будет использоваться для того, чтобы помечать локальную информацию при её экспорте в BGP провайдера. Гранулярность таких тегов может уменьшаться до уровня отдельной пары <CPI, PPI>. В дополнение к этому конфигурация каждой таблицы PIT в PE содержит по крайней мере одну группу Route Target, которую мы будем называть import Route Targets, - эта группа ограничивает набор маршрутов, которые могут быть импортированы из BGP провайдера в PIT, только теми маршрутами, которые входят по крайней мере в одну из групп импорта.

При добавлении провайдером порта L1VPN в конкретном устройстве PE, этот порт связывается с таблицей PIT данного PE, а таблица PIT связывается с конкретной L1VPN.

Поскольку для заполнения таблиц PIT удалённой. информацией используется протокол BGP, который работает со множеством автономных систем (AS[3]), описанный в этом документе механизм позволяет поддерживать L1VPN, распределенные между несколькими автономными системами.

Хотя L1VPN такого типа в настоящее время не рассматриваются для базового режима, определённые здесь механизмы могут быть использованы в будущем. В настоящее время может потребоваться дополнительная работа по проверке различных аспектов, включая безопасность.

3. Передача информации L1VPN в BGP

Отображения <CPI, PPI> передаются с использованием многопротокольного расширения[4] BGP [RFC4760]. Документ [RFC4760] определяет формат двух атрибутов BGP - MP_REACH_NLRI и MP_UNREACH_NLRI, которые могут использоваться для анонсирования и отзыва анонсов информации о доступности. В этом документе добавляется идентификатор семейства адресов, названный Layer-1 VPN auto-discovery information[5] (значение 69) и определяется новый формат NLRI[6] для передачи CPI и PPI.

В упомянутых выше атрибутах BGP может передаваться одна или множество пар <PPI, CPI>.

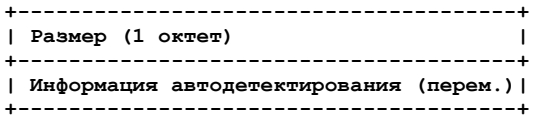

Рисунок 2. Кодирование NLRI.

Рисунок 2 показывает формат NLRI2.

Отметим, что представление информации механизма автоматического детектирования описано в [L1VPN-BM] и подчеркнём также, что при значении Length = 4 в поле Next Hop (атрибут MP_REACH_NLRI) значением Next Hop будет адрес IPv4, а при значении 16 Next Hop будет содержать адрес IPv6.

4. Передача информации L1VPN Traffic Engineering в BGP

В дополнение к информации о доступности механизм автоматического детектирования **может** передавать информацию Traffic Engineering, используемую для выбора исходящего пути. Например, PE может узнать возможности коммутации и максимальную полосу LSP удалённых интерфейсов L1VPN от удалённых PE. Для передачи такой информации в данном документе предлагается использовать атрибут BGP Traffic Engineering [BGP-TE-ATTRIBUTE].

[1]Customer port identifier - идентификатор пользовательского порта.
[2]Provider port identifier - идентификатор порта провайдера.
[3]Autonomous system - автономная система. *Прим. перев.*
[4]Multiprotocol Extensions. *Прим. перев.*
[5]Информация механизма автоматического детектирования Layer-1 VPN.
[6]Network Layer Reachability Information - информация о доступности на сетевом уровне.

5. Масштабируемость

Напомним, что сеть провайдера состоит из (a) устройств PE, (b) маршрутных рефлекторов[1] BGP, (c) узлов P (которые не являются ни PE, ни Route Reflector) и, в случае VPN с использованием нескольких провайдеров, (d) граничных маршрутизаторов автономных систем (ASBR[2]).

Маршрутизатор PE, если он не является маршрутным рефлектором, не сохраняет связанной с L1VPN информации, пока на не нет хотя бы одной VPN со значением import Route Target, идентичным связанной с VPN информации атрибутов Route Target. Если PE не имеет VPN с соответствующим значением import Route Target, это устройство **должно** отбрасывать полученную информацию L1VPN. Для отбрасывания информации **должна** применяться фильтрация на входе. При последующем добавлении import Route Target для одной из VPN устройства PE (операция VPN Join[3]) устройство PE **должно** принимать связанную с VPN информацию, которая ранее отбрасывалась.

Для таких случаев **должен** использоваться механизм обновления, описанный в [BGP-RFSH]. Могут также применяться механизмы фильтрации на выходе [BGP-ORF] и [BGP-CONS] для обеспечения более динамичной фильтрации.

Аналогично, если конкретное значение import Route Target больше не присутствует в VPN какого-либо PE (в результате выполнение одной или множества операций VPN Prune), устройство PE **может** отбрасывать BGP-маршруты L1VPN и в результате не иметь более import Route Targets в таблицах PIT устройств PE как атрибутов Route Target.

Отметим, что операции VPN Join и VPN Prune являются неразрушающими и не требуют разрыва каких-либо соединений BGP или использования механизма обновления [BGP-RFSH].

В результате использования описанных правил распространения информации ни одному устройству PE не требуется иметь всех маршрутов во все L1VPN - это важно учитывать при рассмотрении вопросов масштабирования.

Рефлекторы маршрутов могут быть разделены между VPN так, чтобы каждая группа рефлекторов передавала маршруты только для подмножества L1VPN, поддерживаемых провайдером. Таким образом ни от одного из рефлекторов не требуется поддержки связанной с VPN информации для всех VPN.

Для VPN, включающих множество провайдеров, при использовании EBGP[4] через несколько интервалов маршрутизации от маршрутизаторов ASBR совсем не требуется поддержка и распространение связанной с VPN информации. Маршрутизаторы P не поддерживают у себя какой-либо информации, связанной с VPN.

В результате не возникает единой точки в сети провайдера, которая поддерживала бы всю связанную с VPN информацию для всех VPN. Таким образом, возможности провайдеров в части роста числа поддерживаемых VPN не ограничиваются возможностями какого-либо из конкретных устройств.

Важно подчеркнуть что возможна поддержка **неограниченного** числа систем BGP используемых для передачи связанной с VPN информации. Это отличается от ситуации в Internet, где каждая система BGP **должна** передавать все маршруты Internet. Таким образом, одно значимое (но, возможно, трудно уловимое) различие при использовании BGP для маршрутизации Internet и для распространения связанной с VPN информации состоит в том, что в первом случае невозможно разделение на части, а во втором это реально.

6. Вопросы безопасности

В этом документе описан механизм автоматического детектирования на базе протокола BGP, позволяющий устройствам PE, подключённым к L1VPN, находить другие маршрутизаторы PE, подключённые к той же VPN. Каждый маршрутизатор PE, который подключён к данной VPN, использует протокол BGP для анонсирования факта подключения. Другие маршрутизаторы PE, подключённые к той же VPN, получают анонсы BGP. Это позволяет всему множеству PE автоматически находить друг друга. Отметим, что PE не всегда будет получать эти анонсы непосредственно от удалённых PE - анонсы могут приходить от «промежуточных» узлов BGP.

Критически важно, что для любого маршрутизатора PE **недопустимо** его «детектирование» в качестве подключённого к VPN до того, как данный PE будет реально подключён к этой VPN, что неоспоримо говорит о наличии у маршрутизатора полномочий на подключение к данной VPN. Если произвольный узел Internet может начать передачу анонсов BGP о своей принадлежности к VPN и такие анонсы будут достигать других узлов PE при условии, что эти узлы PE будут принимать такие анонсы, тогда любой желающий сможет добавить любой сайт к любой L1VPN. Таким образом, описанный здесь механизм предполагает, что конкретный маршрутизатор PE доверяет своим партнёрам BGP и, более того, считает этих партнёров обеспечивающими надёжную защиту своих локальных подключений (т. е., PE **должен** верить, что его партнёры подключены к соответствующим L1VPN и имеют на такое подключение право).

Если некий удалённый маршрутизатор PE является BGP-партнером локального PE, **следует** использовать процедуры аутентификации BGP [RFC2385] для того, чтобы убедиться в том, что удалённому PE можно доверять (т. е., данный PE является тем, за кого он себя выдаёт).

Если некий удалённый маршрутизатор PE не является BGP-партнером локального PE, тогда анонсируемая удаленным PE информация будет приходить на локальный маршрутизатор PE через цепочку узлов BGP. Локальный PE **должен** верить, что его партнёры воспринимают информацию только от доверенных партнёров и, таким образом, доверительные отношения **должны б**ыть транзитивными (переходящими). Протокол BGP не обеспечивает возможности удостовериться в том, что информация, полученная от узла BGP, была передана в соответствии с имеющимися полномочиями. Следовательно, описанные в этом документе процедуры **должны** использоваться в средах, где могут поддерживаться адекватные доверительные отношения между узлами BGP (например, использование механизма автоматического детектирования в рамках сети одного провайдера).

[1] Route Reflector.
[2] Autonomous System Border Router - граничный маршрутизатор автономной системы.
[3] Включение в VPN. *Прим. перев.*
[4] External BGP - внешний BGP.

7. Взаимодействие с IANA

В этом документе выделяется новое значение SAFI[1], названное Layer-1 VPN auto-discovery information (см. раздел 3). Это значение включено в реестр дополнительных идентификаторов семейств адресов (SAFI) с использованием процедуры Standards Action. Новый идентификатор имеет значение 69.

8. Литература

8.1. Нормативные документы

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[BGP-RFSH] Chen, E., "Route Refresh Capability for BGP-4", RFC 2918, September 2000.

8.2. Прочие ссылки

[BGP-TE-ATTRIBUTE] Ould-Brahim, H., Fedyk, D., and Rekhter, Y., "Traffic Engineering Attribute", Work in Progress, January 2008.

[BGP-ORF] Chen, E. and Y. Rekhter, "Outbound Route Filtering Capability for BGP-4", Work in Progress[2], September 2006.

[BGP-CONS] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, November 2006.

[BGP-COMM] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, February 2006.

[L1VPN-FRMK] Takeda, T., Ed., "Framework and Requirements for Layer 1 Virtual Private Networks", RFC 4847, April 2007.

[L1VPN-BM] Fedyk, D., Ed., Rekhter, Y., Ed., Papadimitriou, D., Rabbat, R., and L. Berger, "Layer 1 VPN Basic Mode", Work in Progress[3], February 2008.

[RFC2385] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, August 1998.

9. Благодарности

Авторы выражают благодарность Adrian Farrel за полезные комментарии.

Адреса авторов

Hamid Ould-Brahim
Nortel
PO Box 3511 Station C
Ottawa ON K1Y 4H7
Canada
Phone: +1 (613) 763 4730
EMail: hbrahim@nortel.com

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
USA
EMail: yakov@juniper.net

Don Fedyk
Nortel
600 Technology Park
Billerica, MA 01821
USA
Phone: +1 (978) 288 3041
Email: dwfedyk@nortel.com

Перевод на русский язык
Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

[1]Subsequent Address Family Identifier - дополнительный идентификатор семейства адресов.
[2]Работа опубликована в RFC 5291. *Прим. перев.*
[3]Работа опубликована в RFC 5251. *Прим. перев.*