

Internet Engineering Task Force (IETF)
Request for Comments: 7911
Category: Standards Track
ISSN: 2070-1721

D. Walton
Cumulus Networks
A. Retana
E. Chen
Cisco Systems, Inc.
J. Scudder
Juniper Networks
July 2016

Advertisement of Multiple Paths in BGP

Анонсирование множества путей в BGP

Аннотация

Этот документ определяет расширение BGP, позволяющее анонсировать множество путей к одному адресному префиксу без неявной замены имеющихся путей новыми. Суть расширения заключается в том, что каждый путь указывается идентификатором пути (Path Identifier) в дополнение к адресному префиксу.

Статус документа

Документ относится к категории Internet Standards Track.

Документ является результатом работы IETF¹ и представляет согласованный взгляд сообщества IETF. Документ прошёл открытое обсуждение и был одобрен для публикации IESG². Дополнительную информацию о стандартах Internet можно найти в разделе 2 в RFC 7841.

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке <http://www.rfc-editor.org/info/rfc7911>.

Авторские права

Авторские права (Copyright (c) 2017) принадлежат IETF Trust и лицам, указанным в качестве авторов документа. Все права защищены.

К документу применимы права и ограничения, указанные в BCP 78 и IETF Trust Legal Provisions и относящиеся к документам IETF (<http://trustee.ietf.org/license-info>), на момент публикации данного документа. Прочтите упомянутые документы внимательно. Фрагменты программного кода, включённые в этот документ, распространяются в соответствии с упрощённой лицензией BSD, как указано в параграфе 4.e документа IETF Trust Legal Provisions, без каких-либо гарантий (как указано в Simplified BSD License).

Оглавление

1. Введение.....	1
1.1. Уровни требований.....	2
2. Указание пути.....	2
3. Кодирование расширенных NLRI.....	2
4. Возможность ADD-PATH.....	2
5. Работа расширения.....	2
6. Вопросы развёртывания.....	3
7. Взаимодействие с IANA.....	3
8. Вопросы безопасности.....	3
9. Литература.....	3
9.1. Нормативные документы.....	3
9.2. Дополнительная литература.....	3
Благодарности.....	4
Адреса авторов.....	4

1. Введение

Спецификация BGP [RFC4271] определяет процесс передачи обновлений (Update-Send) для анонсирования маршрутов, выбранных процессом принятия решений (Decision Process) для других узлов BGP. Способов анонсирования множества путей для одного адресного префикса или NLRI³ спецификация не предусматривает. Фактически маршрут с тем же NLRI, что уже был анонсирован, неявно заменяет предыдущий анонс.

Этот документ определяет расширение BGP, позволяющее анонсировать множество путей к одному адресному префиксу без неявной замены старых путей новыми. Суть расширения заключается в том, что каждый путь указывается идентификатором пути (Path Identifier) в дополнение к адресному префиксу.

Наличие дополнительных путей может помочь в снижении или предотвращении постоянных осцилляций маршрутов [RFC3345]. Это также может способствовать оптимальной маршрутизации и ускорению схождения в сетях за счёт предоставления дополнительных или резервных путей, соответственно.

¹Internet Engineering Task Force - комиссия по решению инженерных задач Internet.

²Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

³Network Layer Reachability Information - информация о доступности на сетевом уровне.

1.1. Уровни требований

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

2. Указание пути

Как определено [RFC4271], путь представляется информацией, указанной в поле Path Attribute сообщения UPDATE. Поскольку процедуры, указанные в [RFC4271], позволяют анонсировать лишь один путь к конкретному адресному префиксу, для пути к адресному префиксу от партнёра BGP этот префикс может служить ключом.

Для того, чтобы позволить узлам BGP анонсировать множество путей к одному адресному префиксу, нужен новый идентификатор (далее Path Identifier), позволяющий указывать конкретный путь к адресному префиксу комбинацией этого префикса и Path Identifier.

Назначение Path Identifier для пути узел BGP выполняет локально. Однако значения Path Identifier **должны** выделяться так, чтобы узел BGP мог применять пару (Prefix, Path Identifier) для однозначной идентификации пути, анонсируемого соседу. Узел BGP при дальнейшем анонсировании маршрута **должен** создавать свой идентификатор Path Identifier, связанный с анонсируемым маршрутом. Узел BGP, получающий маршрут, не должен предполагать какую-либо конкретную семантику идентификатора пути.

3. Кодирование расширенных NLRI

Для передачи идентификатора пути в сообщениях UPDATE кодирование NLRI **должно** быть расширено путём добавления перед ним поля Path Identifier, состоящего из 4 октетов.

Например, заданное в [RFC4271] кодирование NLRI расширяется, как показано на рисунке.

```

+-----+
| Path Identifier (4 октета) |
+-----+
| Length (1 октет) |
+-----+
| Prefix (переменный размер) |
+-----+

```

Использование расширенного кодирования NLRI описано в разделе 5.

4. Возможность ADD-PATH

ADD-PATH Capability представляет собой возможность BGP [RFC5492] с кодом (Capability Code) 69. Поле Capability Length является переменным, а Capability Value включает одну или более показанных на рисунке групп полей.

```

+-----+
| Address Family Identifier (2 октета) |
+-----+
| Subsequent Address Family Identifier (1 октет) |
+-----+
| Send/Receive (1 октет) |
+-----+

```

Address Family Identifier (AFI) – идентификатор семейства адресов

Это поле имеет такой же смысл как в [RFC4760].

Subsequent Address Family Identifier (SAFI) – идентификатор следующего семейства адресов

Это поле имеет такой же смысл как в [RFC4760].

Send/Receive

Это поле показывает, что отправитель (a) способен принимать от партнёра множество путей (1), (b) способен передавать партнёрам множество путей (value 2) или (c) способен принимать и передавать (3) для <AFI, SAFI>.

При получении иного значения эту возможность **следует** считать не понятой и игнорировать [RFC5492].

Узел BGP, желающий указать поддержку для множества AFI/SAFI, **должен** делать это путём включения информации в один экземпляр ADD-PATH Capability.

5. Работа расширения

Поле Path Identifier, заданное в разделе 3, может применяться для анонсирования множества путей к одному префиксу без замены имеющихся путей новыми. Помимо добавления такой возможности, правила анонсирования маршрутов [RFC4271] не изменились. В частности, новый анонс для данного префикса с данным Path Identifier будет заменять прежний анонс для той же комбинации префикса и идентификатора пути. Если узел BGP получает сообщение для отзыва префикса с Path Identifier, который он не получал, такое сообщение **следует** просто игнорировать.

Для того, чтобы узел BGP мог передавать своему партнёру множество путей, этот узел BGP **должен** анонсировать ADD-PATH Capability с полем Send/Receive, имеющим значение 2 или 3, и **должен** получить от своего партнёра ADD-PATH Capability с полем Send/Receive, имеющим значение 1 или 3, для соответствующей пары <AFI, SAFI>.

Узел BGP **должен** следовать процедурам [RFC4271] при генерации сообщений UPDATE для конкретной пары <AFI, SAFI> своему партнёру, пока не анонсировал тому ADD-PATH Capability с указанием возможности передавать множество путей для <AFI, SAFI> и не получил от этого партнёра ADD-PATH Capability с указанием возможности принимать множество путей для <AFI, SAFI>. Если же партнёры согласовали между собой анонсы множества путей узел **должен** генерировать обновление маршрута для <AFI, SAFI> на основе комбинации адресного префикса и Path Identifier, а также использовать расширенное кодирование NLRI, описанное выше. Партнёру **нужно** выполнять соответствующие действия при обработке сообщений UPDATE, относящихся к определенным <AFI, SAFI>.

Узлу BGP **следует** включать лучший маршрут [RFC4271] при анонсировании соседа множества путей, если путь не был получен от данного соседа.

Идентификаторы пути назначаются локально и могут (но не обязаны) сохраняться при перезапуске уровня управления на узле BGP. Реализациям **следует** применять специальные меры, чтобы работа базового уровня пересылки на приёмном узле ([RFC4724]) не нарушалась при аккуратном перезапуске сессии BGP.

6. Вопросы развёртывания

Предложенное в этом документе расширение обеспечивает узлам BGP механизм для анонсирования множества путей в одной сессии BGP. При развёртывании следует соблюдать осторожность для обеспечения согласованности маршрутизации и пересылки в сети [ADDPATH].

Единственным явным свидетельством применения кодирования, описанного в разделе 3, для конкретной сессии BGP является обмен возможностями, описанный в разделе 4. Если этот обмен прошёл успешно [RFC5492], узлы BGP смогут корректно обрабатывать все сообщения BGP UPDATES, как описано в разделе 5. Однако, при использовании, например, анализатора пакетов в линии он не сможет корректно декодировать сообщения BGP UPDATE поскольку не будет иметь информации об обмене возможностями.

При развёртывании на краевом маршрутизаторе провайдера или «пиринговом» маршрутизаторе, которые взаимодействуют с внешними соседями, узел BGP обычно анонсирует не более одного пути к внутренним соседям в сети. При настройке узла на анонсирование множества путей и необходимости дополнительной информации для приложения узел может применять атрибуты типа Edge_Discriminator [FAST]. Использование таких атрибутов выходит за рамки документа.

7. Взаимодействие с IANA

Агентство IANA назначило значение 69 для возможности ADD-PATH Capability, описанной в этом документе с внесением его в реестр «Capability Codes».

8. Вопросы безопасности

Этот документ определяет расширение BGP, позволяющее анонсировать множество путей к адресному префиксу без неявной замены имеющихся путей новыми. В результате множество путей для большого числа префиксов, получаемых узлом BGP, может поглощать ресурсы памяти узла и даже вызывать нестабильность сети, что может рассматриваться как атака на отказ служб. Отметим, что эта уязвимость не является новой и отмечена в базовой спецификации BGP [RFC4272].

Использование ADD-PATH Capability рассчитано на решение конкретных задач, относящихся, например, к предотвращению осцилляций маршрутов, вызываемых атрибутом MULTI_EXIT_DISC (MED) [STOP-OSC]. Хотя описание применений возможности ADD-PATH выходит за рамки документа, пользователям настоятельно рекомендуется проверить поведение этой возможности и её влияние в соответствии с документом [ADDPATH].

Применимы также вопросы безопасности, относящиеся к базовой работе BGP [RFC4271].

9. Литература

9.1. Нормативные документы

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), DOI 10.17487/RFC4271, January 2006, <<http://www.rfc-editor.org/info/rfc4271>>.
- [RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", [RFC 4760](#), DOI 10.17487/RFC4760, January 2007, <<http://www.rfc-editor.org/info/rfc4760>>.
- [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", [RFC 5492](#), DOI 10.17487/RFC5492, February 2009, <<http://www.rfc-editor.org/info/rfc5492>>.

9.2. Дополнительная литература

- [ADDPATH] Uttaro, J., Francois, P., Patel, K., Haas, J., Simpson, A., and R. Fragassi, "Best Practices for Advertisement of Multiple Paths in IBGP", Work in Progress, draft-ietf-idr-add-paths-guidelines-08, April 2016.
- [FAST] Mohapatra, P., Fernando, R., Filsfils, C., and R. Raszuk, "Fast Connectivity Restoration Using BGP Add-path", Work in Progress, draft-pmohapat-idr-fast-conn-restore-03, January 2013.
- [RFC3345] McPherson, D., Gill, V., Walton, D., and A. Retana, "Border Gateway Protocol (BGP) Persistent Route Oscillation Condition", RFC 3345, DOI 10.17487/RFC3345, August 2002, <<http://www.rfc-editor.org/info/rfc3345>>.
- [RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", [RFC 4272](#), DOI 10.17487/RFC4272, January 2006, <<http://www.rfc-editor.org/info/rfc4272>>.
- [RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y. Rekhter, "Graceful Restart Mechanism for BGP", [RFC 4724](#), DOI 10.17487/RFC4724, January 2007, <<http://www.rfc-editor.org/info/rfc4724>>.
- [STOP-OSC] Walton, D., Retana, A., Chen, E., and J. Scudder, "BGP Persistent Route Oscillation Solutions", Work in Progress, draft-ietf-idr-route-oscillation-stop-03¹, April 2016.

¹Работа опубликована в RFC 7964. Прим. перев.

Благодарности

Авторы признательны David Cook и Naiming Shen за вклад в идею и разработку этого расширения.

Множество людей внесли полезные замечания и предложения, включая Rex Fernando, Eugene Kim, Danny McPherson, Dave Meyer, Pradosh Mohapatra, Keyur Patel, Robert Raszuk, Eric Rosen, Srihari Sangli, Dan Tappan, Mark Turner, Jeff Haas, Jay Borkenhagen, Mach Chen, Denis Ovsienko, Carlos Pignataro, Meral Shirazipour и Kathleen Moriarty.

Адреса авторов

Daniel Walton

Cumulus Networks

185 E. Dana Street

Mountain View, CA 94041

United States of America

Email: dwalton@cumulusnetworks.com

Alvaro Retana

Cisco Systems, Inc.

Kit Creek Rd.

Research Triangle Park, NC 27709

United States of America

Email: aretana@cisco.com

Enke Chen

Cisco Systems, Inc.

170 W. Tasman Dr.

San Jose, CA 95134

United States of America

Email: enkechen@cisco.com

John Scudder

Juniper Networks

1194 N. Mathilda Ave

Sunnyvale, CA 94089

United States of America

Email: jgs@juniper.net

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru