

Network Working Group
Request for Comments: 4760
Obsoletes: 2858
Category: Standards Track

T. Bates
Cisco Systems
R. Chandra
Sona Systems
D. Katz
Y. Rekhter
Juniper Networks
January 2007

Многопротокольные расширения для BGP-4

Multiprotocol Extensions for BGP-4

Статус документа

В этом документе содержится спецификация протокола, предложенного сообществу Internet. Документ служит приглашением к дискуссии в целях развития и совершенствования протокола. Текущее состояние стандартизации протокола вы можете узнать из документа Internet Official Protocol Standards (STD 1). Документ может распространяться без ограничений.

Авторские права

Copyright (C) The IETF Trust (2007).

Аннотация

В этом документе определяется расширение протокола BGP-4, позволяющее передавать информацию для различных протоколов сетевого уровня (например, IPv6, IPX, L3VPN и т. п.). Расширение обеспечивает обратную совместимость – маршрутизаторы, поддерживающие это расширение, смогут нормально работать с маршрутизаторами, которые его не поддерживают.

1. Введение

Только три компонента информации, передаваемой с помощью BGP-4 [BGP-4], непосредственно связаны с IPv4: (a) атрибут NEXT_HOP (указывается адресом IPv4), (b) AGGREGATOR (содержит адрес IPv4) и (c) NLRI (выражается префиксом адреса IPv4). В этом документе предполагается, что любой узел BGP (включая те, которые поддерживают описанное здесь расширение) имеет адрес IPv4 (который будет вместе с другими параметрами использоваться в атрибуте AGGREGATOR). Следовательно, для того, чтобы BGP-4 поддерживал маршрутизацию для множества протоколов сетевого уровня, в BGP-4 требуется добавить только два элемента: (a) возможность связывания отдельного протокола сетевого уровня с информацией о следующем интервале (next hop) и (b) возможность связывания протокола сетевого уровня с NLRI. Для идентификации протоколов сетевого уровня данный документ использует значение Address Family (семейство адресов), определённое в [IANA-AF], и Subsequent Address Family¹ (описано в этом документе).

Легко увидеть, что информация о следующем интервале (значение атрибута NEXT_HOP) имеет смысл (и необходима) только в комбинации с анонсами доступных адресатов – в комбинации же с анонсами недоступных адресатов (отзыв маршрутов) информация о следующем интервале не имеет смысла. Это позволяет предположить, что анонсирование доступных адресатов следует объединять с анонсированием следующего интервала, который будет использоваться для этих адресатов, и анонсирование доступных адресатов следует отделить от анонсирования недоступных адресатов.

Для обеспечения совместимости с предыдущими спецификациями и упрощения перехода к поддержке многопротокольного расширения в BGP-4 в данном документе используются два новых атрибута – MP_REACH_NLRI² и MP_UNREACH_NLRI³. Первый атрибут (MP_REACH_NLRI) используется для передачи набора доступных адресов с информацией о следующем интервале, который будет использоваться для пересылки по этим адресам. Второй атрибут (MP_UNREACH_NLRI) служит для передачи наборов недоступных адресатов. Оба эти атрибута являются необязательными и непереходными. Таким образом, узел BGP, не поддерживающий многопротокольное расширение, просто будет игнорировать содержащуюся в этих атрибутах информацию и не станет передавать её другим узлам BGP.

2. Описание требований

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с RFC 2119 [RFC2119].

3. MP_REACH_NLRI (тип 14)

Этот необязательный и нетранзитивный атрибут может использоваться с двумя целями:

(a) анонсирование партнёру возможного маршрута;

¹Дополнительное семейство адресов.

²Multiprotocol Reachable NLRI.

³Multiprotocol Unreachable NLRI.

(b) обеспечение маршрутизатору возможности анонсирования адреса сетевого уровня маршрутизатора, который следует использовать как следующий интервал (next hop) на пути к адресату, указанному в поле NLRI¹ атрибута MP_NLRI.

```

+-----+
| Address Family Identifier (2 октета) |
+-----+
| Subsequent Address Family Identifier (1 октет) |
+-----+
| Length of Next Hop Network Address (1 октет) |
+-----+
| Network Address of Next Hop (перемен.) |
+-----+
| Резерв (1 октет) |
+-----+
| Network Layer Reachability Information (перемен.) |
+-----+

```

Формат представления атрибута показан на рисунке. Назначение полей описано ниже.

Address Family Identifier (AFI) – идентификатор семейства адресов

Это поле в комбинации с полем Subsequent Address Family Identifier идентифицирует набор протоколов сетевого уровня, к которым должен относиться адрес, передаваемый в поле Next Hop, способ представления адреса следующего интервала и семантику последующих данных NLRI. Если значение Next Hop разрешено для нескольких протоколов сетевого уровня, представление Next Hop **должно** обеспечивать способ определения протокола сетевого уровня.

Определённые к настоящему времени значения поля AFI заданы в реестре IANA Address Family Numbers [IANA-AF].

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

Это поле в комбинации с полем AFI идентифицирует набор протоколов сетевого уровня, к которым должен относиться адрес, передаваемый в поле Next Hop, способ представления адреса следующего интервала и семантику последующих данных NLRI. Если значение Next Hop разрешено для нескольких протоколов сетевого уровня, представление Next Hop **должно** обеспечивать способ определения протокола сетевого уровня.

Length of Next Hop Network Address – размер адреса следующего интервала

1-октетное поле, указывающее размер поля Network Address of Next Hop в октетах.

Network Address of Next Hop – сетевой адрес для следующего интервала

Поле переменной длины, содержащее адрес сетевого уровня для следующего маршрутизатора на пути к получателю. Протокол сетевого уровня, связанный с сетевым адресом следующего интервала, идентифицируется комбинацией значений полей <AFI, SAFI> в данном атрибуте.

Резерв

Это 1-октетное поле **должно** устанавливаться в 0, а при получении значение поля **следует** игнорировать.

Network Layer Reachability Information (NLRI) – информация о доступности на сетевом уровне

Поле переменной длины, содержащее список NLRI для возможных маршрутов, которые будут анонсироваться этим атрибутом. Семантика NLRI определяется комбинацией значений <AFI, SAFI> в данном атрибуте.

Если поле Subsequent Address Family Identifier содержит одно из значений, определённых в данном документе, каждое значение NLRI кодируется в соответствии с параграфом 5. Представление NLRI данного документа.

Информация о следующем интервале (next hop), передаваемая в атрибуте пути MP_REACH_NLRI, определяет адрес сетевого уровня маршрутизатора, который **следует** использовать в качестве следующего этапа пересылки адресатам, указанным в атрибуте MP_NLRI сообщения UPDATE.

Правила для информации о следующем интервале совпадают с правилами для информации, передаваемой в атрибуте BGP NEXT_HOP (см. параграф 5.1.3 документа [BGP-4]).

Сообщение UPDATE, содержащее MP_REACH_NLRI, **должно** также включать атрибуты ORIGIN и AS_PATH (как для EBGP, так и для IBGP). Более того, в обменах IBGP такие сообщения **должны** также включать атрибут LOCAL_PREF.

В сообщении UPDATE, не содержащие NLRI, кроме значения в атрибуте MP_REACH_NLRI, **не следует** включать атрибут NEXT_HOP. Если такое сообщение содержит атрибут NEXT_HOP, получившему это сообщение узлу BGP **следует** игнорировать данный атрибут.

В сообщении UPDATE **не следует** включать один префикс (одну пару <AFI, SAFI>) более одного раза для полей WITHDRAWN_ROUTES, NLRI, MP_REACH_NLRI и MP_UNREACH_NLRI. Обработка сообщений UPDATE с дубликатами префиксов не определена.

4. MP_UNREACH_NLRI (тип 15)

Этот дополнительный нетранзитивный атрибут может использоваться для отзыва невозможных маршрутов.

```

+-----+
| Address Family Identifier (2 октета) |
+-----+
| Subsequent Address Family Identifier (1 октет) |
+-----+
| Withdrawn Routes (перемен.) |
+-----+

```

Формат атрибута показан на рисунке. Назначение полей атрибута описано ниже.

¹Network Layer Reachability Information – информация о доступности на сетевом уровне.

Address Family Identifier (AFI) – идентификатор семейства адресов

Это поле в комбинации с полем Subsequent Address Family Identifier идентифицирует набор протоколов сетевого уровня, к которым должен относиться адрес, передаваемый в поле Next Hop, способ представления адреса следующего интервала и семантику последующих данных NLRI. Если значение Next Hop разрешено для нескольких протоколов сетевого уровня, представление Next Hop **должно** обеспечивать способ определения протокола сетевого уровня¹.

Определённые к настоящему времени значения поля AFI заданы в реестре IANA Address Family Numbers [IANA-AF].

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

Это поле в комбинации с полем AFI идентифицирует набор протоколов сетевого уровня, к которым должен относиться адрес, передаваемый в поле Next Hop, способ представления адреса следующего интервала и семантику последующих данных NLRI. Если значение Next Hop разрешено для нескольких протоколов сетевого уровня, представление Next Hop **должно** обеспечивать способ определения протокола сетевого уровня¹.

Withdrawn Routes NLRI – отзываемые маршруты

Поле переменной длины, содержащее значения NLRI для отзываемых маршрутов. Семантика NLRI определяется комбинацией значений <AFI, SAFI> в атрибуте.

При установке в поле Subsequent Address Family Identifier одного из определённых в данном документе значений, каждое поле NLRI кодируется в соответствии с параграфом 5. Представление NLRI.

Сообщение UPDATE, содержащее MP_UNREACH_NLRI, может не включать других атрибутов пути.

5. Представление NLRI

Информация о доступности на сетевом уровне (NLRI) представляется в форме одной или множества пар <length, prefix>, показанных на рисунке.

```

+-----+
| Length (1 октет) |
+-----+
| Prefix (перемен.)|
+-----+

```

Назначение каждого поля пар описано ниже.

a) Length - размер

Поле Length указывает размер адресного префикса в битах. Нулевой размер показывает, что префикс соответствует всем (как указано для данного семейства) адресам (т. е., сам префикс содержит 0 октетов).

b) Prefix - префикс

Поле Prefix включает префикс адреса, за которым следуют нулевые биты заполнения для выравнивания поля по границе октета. Отметим, что нулевые биты заполнения не принимаются во внимание.

6. Дополнительный идентификатор семейства адресов

Этот документ определяет следующие значения поля Subsequent Address Family Identifier для атрибутов MP_REACH_NLRI и MP_UNREACH_NLRI:

- 1 – NLRI используется для unicast-пересылки (по конкретному адресу);
- 2 – NLRI используется для групповой пересылки.

Реализация может поддерживать все или часть определённых в этом документе значений SAFI или не поддерживать их совсем.

7. Обработка ошибок

Если узел BGP получает от соседа сообщение UPDATE с атрибутом MP_REACH_NLRI или MP_UNREACH_NLRI и видит, что атрибут указан некорректно, он **должен** удалить все маршруты BGP, полученные от данного соседа, в которых значения AFI/SAFI совпадают с одним из содержащихся в некорректном атрибуте MP_REACH_NLRI или MP_UNREACH_NLRI. До завершения сеанса BGP, в котором получено сообщение UPDATE, узлу **следует** игнорировать все последующие маршруты с данным AFI/SAFI, принятые в этой сессии.

В дополнение к сказанному узел **может** прервать сеанс BGP, в котором было получено сообщение UPDATE. Сессию **следует** прерывать с помощью сообщения NOTIFICATION, содержащего код/субкод ошибки Update Message Error/Optional Attribute Error.

8. Использование анонсирования возможностей BGP

Узлу BGP, использующему описанное здесь расширение, **следует** применять процедуры анонсирования возможностей (Capability Advertisement) [BGP-CAP] для определения возможности использования данного расширения при работе с тем или иным партнёром.

Для полей дополнительного параметра Capabilities следует установить приведённые здесь значения. Поле Capability Code должно содержать значение 1 (указывает на Multiprotocol Extensions). Поле Capability Length должно иметь значение 4. Формат поля Capability Value показан на рисунке. Компоненты этого поля описаны ниже.

```

0       7       15       23       31
+-----+-----+-----+-----+
|           AFI           | Res. | SAFI |
+-----+-----+-----+-----+

```

AFI - идентификатор семейства адресов (16 битов), представляемый так же, как для Multiprotocol Extensions;

¹Yakov Rekhter предложил иную редакцию этого абзаца (см. http://rfc-editor.org/errata_search.php?eid=1573). «Это поле в комбинации с полем Subsequent Address Family Identifier определяет семантику информации NLRI для отзываемых этим атрибутом маршрутов.»
Прим. перев.

Res. - резервное поле (8 битов); отправителю **следует** устанавливать для этого поля значение 0¹, а получателю – игнорировать его;

SAFI – дополнительный идентификатор семейства адресов (8 битов), представляемый так же, как для Multiprotocol Extensions.

Узел, поддерживающий множество пар <AFI, SAFI>, включает их как множество возможностей в дополнительный параметр Capabilities.

Чтобы организовать двухсторонний обмен маршрутной информацией для той или иной пары <AFI, SAFI> между двумя узлами BGP, каждый из этих узлов должен анонсировать партнёру (с помощью механизма Capability Advertisement) возможность поддержки маршрутов для соответствующей пары <AFI, SAFI>.

9. Взаимодействие с IANA

Как указано выше, атрибуты MPL_REACH_NLRI и MP_UNREACH_NLRI содержат поле дополнительного идентификатора семейства адресов (SAFI). Пространство имен SAFI определено в данном документе. IANA поддерживает и регистрирует значения SAFI, как описано ниже.

- значения 1 и 2 выделены в настоящем документе;
- значение 3 является резервным; оно было выделено в RFC 2858 для целей, не получивших распространения, поэтому данный документ отменяет это выделение;
- значения от 5 до 63 выделяются IANA с использованием процедуры стандартизации (Standards Action), определённой в [RFC2434], или процедуры предварительного распределения (Early IANA Allocation), определённой в [RFC4020];
- SAFI из диапазона 67 - 127 распределяются IANA в порядке поступления запросов (процедура First Come First Served), как описано в RFC 2434;
- значения 0 и 255 являются резервными;
- значения от 128 до 240 являются часть выделенного ранее блока для частного использования; на момент принятия данного документа неиспользуемые значения были предоставлены IANA руководителем Routing Area; во избежание конфликтов эти значения (130, 131, 135 — 139 и 141 — 240) рассматриваются, как резервные;
- значения от 241 до 254 выделены для частного использования и не распределяются IANA.

10. Сравнение с RFC 2858

Этот документ согласует использование информации о следующем интервале с информацией, передаваемой в атрибуте пути BGP NEXT_HOP.

Из документа удалено определение SAFI 3 и отменено использование SAFI 3.

В этом документе изменено распределение пространства значений SAFI. В частности, RFC 2858 определял значения SAFI от 128 до 240 для частного использования. В данном документе распределение этого блока передаётся IANA и неиспользуемые значения 130, 131, 135 - 139 и 141 - 240 из этого диапазона следует рассматривать в качестве резервных.

В этом документе поле Number of SNPA² переведено в состояние резервного, а последующие поля атрибута MP_REACH_NLRI, связанные с SNPA, удалены.

11. Сравнение с RFC 2283

Этот документ ограничивает атрибут MP_REACH_NLRI передачей единственного экземпляра <AFI, SAFI, Next Hop Information, ...>.

Этот документ ограничивает атрибут MP_UNREACH_NLRI передачей единственного экземпляра <AFI, SAFI, ...>.

Этот документ разъясняет обработку сообщений UPDATE, не содержащих NLRI, кроме значения в атрибуте MP_REACH_NLRI.

Этот документ разъясняет обработку ошибок при наличии атрибутов MP_REACH_NLRI или MP_UNREACH_NLRI.

Этот документ содержит спецификацию использования анонсов возможностей BGP вместе с многопротокольными расширениями.

Этот документ включает раздел 9. Взаимодействие с IANA.

12. Вопросы безопасности

Данное расширение не оказывает влияния на вопросы безопасности, связанные с протоколом BGP.

13. Благодарности

Авторы благодарят членов рабочей группы IDR за обзор и комментарии к документу.

¹Отметим, что если это поле не сбрасывать в 0, могут возникать проблемы с получателями, которые не будут игнорировать значение поля. Кроме того, могут возникать проблемы при переопределении этого поля.

²Subnetwork Points of Attachment - точки подключения подсетей. *Прим. перев.*

14. Нормативные документы

[BGP-CAP] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", [RFC 3392](#), November 2002.

[BGP-4] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), January 2006.

[IANA-AF] "Address Family Numbers", Reachable from <http://www.iana.org/numbers.html>

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

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, [RFC 2434](#), October 1998.

[RFC4020] Kompella, K. and A. Zinin, "Early IANA Allocation of Standards Track Code Points", BCP 100, [RFC 4020](#), February 2005.

Адреса авторов

Tony Bates

Cisco Systems, Inc.

EMail: tbates@cisco.com

Ravi Chandra

Sonoa Systems

EMail: rchandra@sonoasystems.com

Dave Katz

Juniper Networks, Inc.

EMail: dkatz@juniper.net

Yakov Rekhter

Juniper Networks, Inc.

EMail: yakov@juniper.net

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Copyright (C) The IETF Trust (2007).

К этому документу применимы права, лицензии и ограничения, указанные в BCP 78, и, за исключением указанного там, авторы сохраняют свои права.

Этот документ и содержащаяся в нем информация представлены "как есть" и автор, организация, которую он/она представляет или которая выступает спонсором (если таковой имеется), Internet Society и IETF отказываются от каких-либо гарантий (явных или подразумеваемых), включая (но не ограничиваясь) любые гарантии того, что использование представленной здесь информации не будет нарушать чьих-либо прав, и любые предполагаемые гарантии коммерческого использования или применимости для тех или иных задач.

Интеллектуальная собственность

IETF не принимает какой-либо позиции в отношении действительности или объема каких-либо прав интеллектуальной собственности (Intellectual Property Rights или IPR) или иных прав, которые, как может быть заявлено, относятся к реализации или использованию описанной в этом документе технологии, или степени, в которой любая лицензия, по которой права могут или не могут быть доступны, не заявляется также применение каких-либо усилий для определения таких прав. Сведения о процедурах IETF в отношении прав в документах RFC можно найти в BCP 78 и BCP 79.

Копии раскрытия IPR, предоставленные секретариату IETF, и любые гарантии доступности лицензий, а также результаты попыток получить общую лицензию или право на использование таких прав собственности разработчиками или пользователями этой спецификации, можно получить из сетевого репозитория IETF IPR по ссылке <http://www.ietf.org/ipr>.

IETF предлагает любой заинтересованной стороне обратить внимание на авторские права, патенты или использование патентов, а также иные права собственности, которые могут потребоваться для реализации этого стандарта. Информацию следует направлять в IETF по адресу ietf-ipr@ietf.org.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.