

Network Working Group
Request for Comments: 2858
Obsoletes: 2283
Category: Standards Track

T. Bates
Y. Rekhter
Cisco Systems
R. Chandra
Redback Networks Inc
D. Katz
Juniper Networks
June 2000

Многопротокольные расширения для BGP-4

Multiprotocol Extensions for BGP-4¹

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 Internet Society (2000). All Rights Reserved.

Тезисы

В настоящее время протокол BGP-4 [BGP-4] подходит только для передачи маршрутной информации протокола IPv4 [IPv4]. В этом документе определяется расширение протокола BGP-4, позволяющее передавать информацию для различных протоколов сетевого уровня (например, IPv6, IPX и т. п.). Расширение обеспечивает обратную совместимость – маршрутизаторы, поддерживающие это расширение, смогут нормально работать с маршрутизаторами, которые его не поддерживают.

Этот документ заменяет собой RFC 2283.

1. Обзор

Только три компоненты информации, передаваемой с помощью BGP-4, непосредственно связаны с IPv4: (а) 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: (а) возможность связывания того или иного 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), указанное в [RFC1700].

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 нетранзитивными. Таким образом, узел BGP, не поддерживающий многопротокольное расширение, просто будет игнорировать содержащуюся в этих атрибутах информацию и не станет передавать ее другим узлам BGP.

¹Данный документ признан устаревшим и заменен RFC 4760, перевод которого имеется на сайте www.protocols.ru.
Прим. перев.

²Multiprotocol Reachable NLRI.

³Multiprotocol Unreachable NLRI.

2. MP_REACH_NLRI (тип 14)

Этот необязательный и нетранзитивный атрибут может использоваться с несколькими целями:

- (a) анонсирование партнеру возможного маршрута;
- (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;
- (c) обеспечение данному маршрутизатору возможности сообщать о всех или некоторых SNPA², существующих в локальной системе.

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

Address Family Identifier – идентификатор семейства адресов

Это поле служит для идентификации протокола сетевого уровня, связанного с указанным далее сетевым адресом. Определенные в настоящий момент значения указаны в документе RFC 1700³ (раздел Address Family Numbers).

Subsequent Address Family Identifier – дополнительный идентификатор семейства адресов

Это поле содержит дополнительную информацию о типе NLRI в данном атрибуте.

Length of Next Hop Network Address – размер адреса следующего интервала

1-октетное поле, указывающее размер сетевого адреса следующего интервала (поле 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.

Number of SNPAs – число точек подключения подсетей

1-октетное поле, содержащее количество различных SNPA, перечисленных в последующих полях. Нулевое значение этого поля говорит об отсутствии SNPA в данном атрибуте.

Length of Nth SNPA - размер n-го SNPA

1-октетное поле, указывающее размер поля Nth SNPA of Next Hop в полуоктетах.

Nth SNPA of Next Hop – n-я SPNA следующего маршрутизатора

Поле переменной длины, содержащее SNPA маршрутизатора, чей сетевой адрес содержится в поле Network Address of Next Hop. Размер поля составляет целое число октетов, равное округленному до большего значению половины размера SNPA, выраженного в полуоктетах; если SNPA включает нечетное количество полуоктетов, значение этого поля дополняется нулевым полуоктетом после значения размера.

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

Поле переменной длины, содержащее список NLRI для возможных маршрутов, которые будут анонсироваться этим атрибутом. Если поле Subsequent Address Family Identifier содержит одно из значений, определенных в данном документе, каждое значение NLRI кодируется в соответствии с параграфом 4. Представление NLRI данного документа.

Информация о следующем интервале (next hop), передаваемая в атрибуте пути MP_REACH_NLRI, определяет адрес сетевого уровня граничного маршрутизатора, который следует использовать в качестве следующего этапа пересылки адресатам, указанным в атрибуте MP_NLRI сообщения UPDATE. При анонсировании атрибута MP_REACH_NLRI внешнему партнеру маршрутизатор может использовать адрес одного из своих интерфейсов в качестве указывающей следующий интервал компоненты атрибута, полученного от внешнего партнера, для которого анонсируемый маршрут имеет общую подсеть с адресом next hop. Это называют "следующим интервалом из первых рук" ("first party" next hop). Узел BGP может анонсировать внешнему партнеру интерфейс любого внутреннего партнерского маршрутизатора в компоненте next hop, полученной от внешнего партнера, для которого анонсируемый маршрут имеет общую подсеть с адресом next hop. Это называется "следующим интервалом из третьих рук" ("third party" next hop). Узел BGP может анонсировать любой внешний партнерский маршрутизатор в компоненте next hop, указывающей, что адрес сетевого уровня этого граничного маршрутизатора получен от внешнего партнера, и внешний партнер, для которого будет анонсироваться маршрут имеет общую подсеть с адресом next hop. Это другой вариант "следующего интервала из третьих рук".

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

²Subnetwork Points of Attachment – точка подключения подсети.

³В соответствии с RFC 3232 этот документ утратил силу. Упомянутые здесь значения доступны на сайте <http://www.iana.org/assignments/address-family-numbers>. Прим. перев.

Обычно информация о следующем интервале выбирается так, чтобы использовался кратчайший из доступных путей. Узел BGP должен обеспечивать поддержку запрета анонсов информации из третьих рук для обслуживания сред с несовершенными мостами или по соображениям политики.

Для узла BGP недопустимо анонсирование адреса партнера этому же партнеру в качестве следующего интервала для маршрутов, который начинается с данного узла. Для узла BGP недопустимо создание маршрутом с указанием самого себя в качестве следующего интервала.

Когда узел BGP анонсирует маршрут своему внутреннему партнеру, ему не следует изменять связанную с этим маршрутом информацию о следующем интервале. При получении узлом BGP маршрута через внутреннее соединение, он может пересылать пакеты по адресу следующего интервала, если содержащийся в атрибуте адрес имеет общую подсеть с локальным и удаленным узлами BGP.

Сообщение UPDATE, содержащее MP_REACH_NLRI, должно включать также атрибуты ORIGIN и AS_PATH (как для EBGP, так и для IBGP). Более того, при обмене IBGP такие сообщения должны также включать атрибут LOCAL_PREF. Если сообщение получено от внешнего партнера, локальной системе следует убедиться, что самое левое значение AS в атрибуте AS_PATH является номером автономной системы, к которой относится передавший сообщение партнер. Если это условие не выполняется, локальной системе следует передать сообщение NOTIFICATION со значениями Error Code = UPDATE Message Error и Error Subcode = Malformed AS_PATH.

В сообщениях UPDATE, содержащих NLRI только в атрибуте MP_REACH_NLRI, не должен включаться атрибут NEXT_HOP. Если такое сообщение содержит атрибут NEXT_HOP, принимающему узлу BGP следует игнорировать этот атрибут.

3. MP_UNREACH_NLRI (тип 15)

Этот дополнительный нетранзитивный атрибут может использоваться для отзыва невозможных маршрутов.

Формат атрибута показан на рисунке.

Назначение полей атрибута описано ниже.

Address Family Identifier – идентификатор семейства адресов

Это поле служит для идентификации протокола сетевого уровня, связанного с указанным далее сетевым адресом. Определенные в настоящий момент значения указаны в документе RFC 1700¹ (раздел Address Family Numbers).

Subsequent Address Family Identifier – дополнительный идентификатор семейства адресов

Это поле содержит дополнительную информацию о типе NLRI в данном атрибуте.

Withdrawn Routes – отзываеваемые маршруты

Поле переменной длины, содержащее значения NLRI для отзывааемых маршрутов. При установке в поле Subsequent Address Family Identifier одного из определенных в данном документе значений, каждое поле NLRI кодируется в соответствии с параграфом 4. Представление NLRI.

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

4. Представление NLRI

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

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.

5. 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;
- 3 – NLRI используется для индивидуальной и групповой пересылки.

6. 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

¹В соответствии с RFC 3232 этот документ утратил силу. Упомянутые здесь значения доступны на сайте <http://www.iana.org/assignments/address-family-numbers>. Прим. перев.

MP_UNREACH_NLRI. До завершения сеанса BGP, в котором получено сообщение UPDATE, узлу следует игнорировать все последующие маршруты с AFI/SAFI, принятыми 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.

7. 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 партнером.

Для полей дополнительного параметра Capabilities следует установить 0 7 15 23 31 приведенные 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 справа. 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;

Res. - резервное поле (8 битов); отправитель должен помещать в это поле значение 0, а получатель – 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 партнеру (с помощью механизма Capability Advertisement) возможность поддержки маршрутов для соответствующей пары <AFI, SAFI>.

8. Согласование 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 разделе 5¹. IANA поддерживает и регистрирует значения SAFI. Нулевое значение SAFI является резервным, значения 1, 2 и 3 выделены в настоящем документе. Значения от 4 до 63 выделяются IANA по согласованию с IETF (процедура "IETF Consensus"), как указано в документе RFC 2434. SAFI из диапазона 64 - 127 распределяются IANA в порядке поступления запросов (процедура "First Come First Served"), как описано в RFC 2434, а значения из диапазона от 128 до 255 предназначены для частного использования и не контролируются IANA.

9. Сравнение с RFC 2283

В этом документе использование атрибута MP_REACH_NLRI ограничивается передачей только одного экземпляра <AFI, SAFI, Next Hop Information, ...>.

В этом документе использование атрибута MP_UNREACH_NLRI ограничивается передачей только одного экземпляра <AFI, SAFI, ...>.

В этом документе разъясняется обработка сообщений UPDATE, содержащих NLRI только в атрибуте MP_REACH_NLRI.

В этом документе разъясняется обработка ошибок при наличии атрибутов MP_REACH_NLRI и MP_UNREACH_NLRI.

В этом документе разъясняется использование анонсирования возможностей BGP (Capabilities Advertisements) в комбинации с Multiprotocol extensions.

В данный документ включен раздел 8. Согласование с IANA.

10. Вопросы безопасности

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

11. Благодарности

Авторы выражают свою признательность членам рабочей группы IDR за просмотр документа и комментарии.

12. Литература

[BGP-CAP] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", RFC 2842², May 2000.

[BGP-4] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771³, March 1995.

[Heffernan] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385⁴, August 1998.

[IPv4] Postel, J., "Internet Protocol", STD 5, RFC 791⁴, September 1981.

¹В оригинале ошибочно указан раздел 9. Прим. перев.

²Этот документ утратил силу и заменен RFC 3392, а последний был заменен впоследствии RFC 5492. Переводы этих документов имеются на сайте www.protocols.ru. Прим. перев.

³Этот документ утратил силу и заменен RFC 4271. Перевод имеется на сайте www.protocols.ru. Прим. перев.

⁴Перевод этого документа имеется на сайте www.protocols.ru. Прим. перев.

13. Адреса авторов

Tony Bates

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
EMail: tbates@cisco.com

Ravi Chandra

Redback Networks Inc.
350, Holger Way
San Jose, CA 95134
EMail: rchandra@redback.com

Dave Katz

Juniper Networks, Inc.
3260 Jay St.
Santa Clara, CA 95054
EMail: dkatz@jnx.com

Yakov Rekhter

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
EMail: yakov@cisco.com

Перевод на русский язык

Николай Малых

nmalykh@protocols.ru

14. Полное заявление авторских прав

Copyright (C) The Internet Society (2000). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.

⁴В соответствии с RFC 3232 этот документ утратил силу STD 2 и выделенные номера в настоящее время доступны в базе данных по ссылке <http://www.iana.org/numbers.html>. Указанная в оригинале ссылка на сайт также утратила актуальность. Прим. перев.