

Группы BGP для сбора данных

BGP Communities for Data Collection

Статус документа

Этот документ описывает практический опыт (Internet Best Current Practices) для сообщества Internet и является приглашением к дискуссии в целях развития и совершенствования. Документ может распространяться свободно.

Авторские права

Copyright (C) The Internet Society (2006).

Тезисы

Группы BGP¹ (RFC 1997) используются сервис-провайдерами для разных целей, включая метки маршрутов в соответствии с заказчиком, партнером или географией исходной точки. Такие метки обычно служат для контроля за областью распространения маршрутов в сети провайдера, а также среди его партнеров и заказчиков. Из результатов крупномасштабного сбора данных BGP (и связанных с этим исследований) становится ясно, что передаваемая в таких группах весьма важна для более глубокого понимания глобальной системы маршрутизации. В этом документе определяются стандартные (исходящие – outbound) группы и их представление для экспорта в системы сбора маршрутной информации BGP².

Оглавление

1. Введение.....	1
2. Определения.....	2
2.1. Партнеры и партнерство.....	2
2.2. Маршруты потребителей.....	2
2.3. Маршруты партнеров.....	2
2.4. Внутренние маршруты.....	2
2.5. Внутренние более специфичные маршруты.....	2
2.6. Маршруты специального назначения.....	2
2.7. Восходящие маршруты.....	2
2.8. Национальные маршруты.....	2
2.9. Региональные маршруты.....	2
3. Кодирование и значения групп RFC 1997.....	3
4. Значения Community для сбора данных BGP.....	3
4.1. Расширенные группы.....	4
4.2. 4-октетные расширенные группы уровня AS.....	4
5. Замечание по упаковке сообщений BGP UPDATE.....	5
6. Благодарности.....	5
7. Вопросы безопасности.....	5
7.1. Общий размер атрибутов.....	5
8. Согласование с IANA.....	5
9. Литература.....	5
9.1. Нормативные документы.....	5
9.2. Дополнительная литература.....	5

1. Введение

Группы BGP [RFC1997] используются сервис-провайдерами для разных целей, включая метки маршрутов в соответствии с заказчиком, партнером или географией исходной точки. Такие метки обычно служат для контроля за областью распространения маршрутов в сети провайдера, а также среди его партнеров и заказчиков. Группы также используются для широкого спектра других приложений, таких, как предоставление потребителям возможности установки атрибутов типа LOCAL_PREF [RFC1771] путем передачи соответствующих групп своему провайдеру. К числу приложений относится также сигнализация различных типов VPN³ (например, VPLS⁴ [VPLS]), и передача информации о ширине полосы для систем управления трафиком [RFC4360].

Из результатов крупномасштабного сбора данных BGP [RV] [RIS] (и связанных с этим исследований) становится ясно, что географическая и топологическая информация, а также отношение провайдера к исходной точке маршрута

¹BGP community.

²BGP route collector.

³Virtual Private Network – виртуальная частная сеть.

⁴Virtual Private LAN Service.

(например, транзитный узел, партнер или потребитель), содержащиеся в таких группах, весьма важны для более глубокого понимания глобальной системы маршрутизации. В этом документе определяются стандартные группы для экспорта с системы сбора маршрутной информации BGP. Эти группы представляют существенную часть информации, передаваемой сервис-провайдерами, и данные могут быть полезными для решения внутренних задач сервис-провайдеров. Однако такое использование данных выходит за пределы настоящего документа. Тем, кто вовлечен в анализ данных BGP, рекомендуется проверить какие из их источников данных реализуют эту схему (поскольку существует большое количество данных, а также множество унаследованных партнерских отношений).

Оставшаяся часть документа организована следующим образом – глава 2 содержит определения терминов, которые применяются как семантика групп, используемых для сбора данных BGP, а глава 3 определяет соответствующее кодирование групп RFC 1997 [RFC1997]. Наконец, в главе 4 определяется кодирование для использования с расширенными группами [RFC4360].

2. Определения

В этой главе мы определим используемые термины и категории маршрутов, которые могут помечаться с помощью групп. Такие метки часто называют «окрашиванием» и мы будем говорить о «цветах» маршрутов, задаваемых принадлежностью к группам. Определенные здесь категории напоминают описанные в работах [WANG] и [HUSTON].

2.1. Партнеры и партнерство

Рассмотрим двух сервис-провайдеров, А и В. Эти провайдеры считаются партнерами, если (i) А и В обмениваются маршрутами по протоколу BGP и (ii) обмен трафиком между А и В происходит на бесплатной основе. Такое соглашение обычно называют партнерством¹. Партнеры обычно обмениваются между собой только маршрутами к своим клиентам (см. следующий параграф) и, следовательно, обмениваются между собой только клиентским трафиком. Более детальное обсуждение бизнес-моделей, близких к партнерству, можно найти в работе [HUSTON].

2.2. Маршруты потребителей

Маршруты потребителей (Customer route) – это те маршруты, которые получены от заказчиков через протокол BGP и распространяются партнерам и другим потребителям. Отметим, что потребителем (заказчиком) может быть организация или другой сервис-провайдер. Такие маршруты иногда называют клиентскими (client route) [HUSTON].

2.3. Маршруты партнеров

Маршруты партнеров (Peer route) – это те маршруты, которые были получены от партнеров по протоколу BGP и не распространяются другим партнерам. Однако такие маршруты распространяются потребителям данного провайдера.

2.4. Внутренние маршруты

Внутренние маршруты (Internal route) – это маршруты, исходящие от данного сервис-провайдера и передаваемые партнерам и заказчиком. Эти маршруты зачастую выбираются из выделенного провайдеру адресного пространства.

2.5. Внутренние более специфичные маршруты

Внутренние более специфичные маршруты (Internal more specific route) – это маршруты, которые часто используются в целях балансировки нагрузки и снижения числа маршрутов IGP². Они могут также соответствовать службам, которые недоступны за пределами сети данного провайдера. Такие маршруты не экспортируются никаким внешним партнерам.

2.6. Маршруты специального назначения

Маршруты специального назначения (Special purpose route) – это те маршруты, которые не могут быть отнесены ни к одному из описанных здесь классов. В тех случаях, когда такие маршруты требуется как-то выделить, сервис-провайдер может окрасить их с использованием уникального цвета. Примерами маршрутов специального назначения являются anycast-маршруты³ и маршруты для перекрывающихся сетей (overlay network).

2.7. Восходящие маршруты

Восходящие маршруты (Upstream route) обычно получают от вышестоящего сервис провайдера как часть транзитного обслуживания, выполняемого вышестоящим провайдером в рамках контракта.

2.8. Национальные маршруты

Набор маршрутов, начинающихся и/или полученных из определенной страны.

2.9. Региональные маршруты

Некоторые глобальные магистрали реализуют региональную политику, основанную на географии развернутой ими сети, а также на их стратегии и требованиях бизнеса. Сервис-провайдеры часто имеют в одном регионе бесплатные (settlement-free) соединения с AS, которые в другом регионе являются потребителями. Это вынуждает использовать региональную политику, включающую атрибуты групп, устанавливаемые сетью, для того, чтобы можно было легко различать маршруты разных регионов. Например, сервис-провайдер может трактовать набор маршрутов, полученных от другого сервис-провайдера в Европе, совсем не так, как тот же набор маршрутов, полученный в северной Америке, поскольку повсеместно может устанавливаться платный транзит для одних регионов и бесплатное партнерство – для других.

¹Или «пирингом», от английского peering. *Прим. перев.*

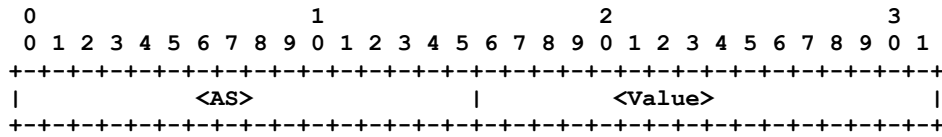
²Interior Gateway Protocol – протокол внутренней маршрутизации.

³Маршрут ко всем.

3. Кодирование и значения групп RFC 1997

В этой главе приводятся значения групп (community) RFC 1997 [RFC1997] для описанных выше категорий. Группы RFC 1997 кодируются как BGP Type Code 8 и трактуются как 32битовые значений из диапазона 0x00000000 - 0xFFFFFFFF. Значения от 0x00000000 до 0x0000FFFF и от 0xFFFF0000 до 0xFFFFFFFF зарезервированы.

Среди современных провайдеров принято использовать два старших октета для номера AS, а двумя младшими октетами задавать классификацию маршрута, как показано ниже:



<AS> представляет собой 16-битовый номер автономной системы. Например, значение 0x2A7C029A будет представлять AS 10876 и значение 666.

4. Значения Community для сбора данных BGP

В этой главе мы определим для описанных выше категорий маршрутов кодирование групп RFC 1997, которые используются для сбора данных BGP. Предполагается, что используемые сервис-провайдерами внутренние значения будут приведены в соответствие с этими стандартными значениями для вывода в системы сбора маршрутных данных (route collector).

Этот документ следует общепринятой на сегодняшний день практике использования базового формата <AS>:<Value>. Значения для разных категорий маршрутов приведены в таблице.

Категория	Значение
Резерв	<AS>:0000000000000000
Маршруты потребителей	<AS>:0000000000000001
Маршруты партнеров	<AS>:0000000000000010
Внутренние маршруты	<AS>:0000000000000011
Внутренние более специфичные маршруты	<AS>:0000000000000100
Маршруты специального назначения	<AS>:0000000000000101
Восходящие маршруты	<AS>:0000000000000110
Резерв	<AS>:0000100000000000 <AS>:0000011111111111
Национальные и региональные маршруты Кодируются как	<AS>:1111111111111111 <AS>:<R><X><CC>
Зарезервированные национальные и региональные маршруты	<AS>:0100000000000000 <AS>:1111111111111111

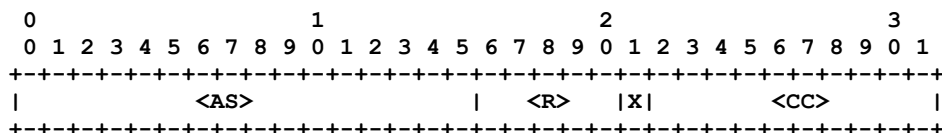
где

- <AS> - 16-битовый номер AS;
- <R> - 5-битовый идентификатор региона;
- <X> - 1-битовая индикация спутниковых каналов (X = 1 для спутниковых каналов, 0 – для прочих);
- <CC> - 10-битовый код страны ISO-3166-2 [ISO3166]

Идентификатор <R> может принимать значения:

Африка (AF)	00001
Океания (OC)	00010
Азия (AS)	00011
Антарктика (AQ)	00100
Европа (EU)	00101
Латинская Америка и Карибские острова (LAC)	00110
Северная Америка (NA)	00111
Резерв	01000 - 11111

Формат значения для национального или регионального маршрута показан ниже.



Например, кодирование национального маршрута через наземный канал в AS 10876 с островов Фиджи (Fiji) будет иметь вид:

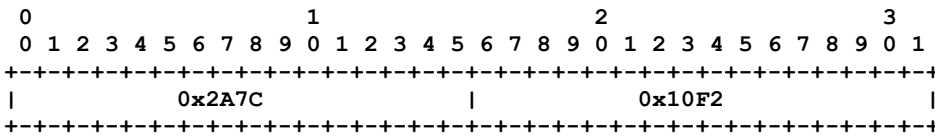
<AS> = 10876 = 0x2A7C

<R> = 00010

<X> = 0

<CC> = код страны для островов Фиджи 242 = 0011110010

В этом случае младшие 16 битов будут иметь значение 0001000011110010 = 0x10F2.



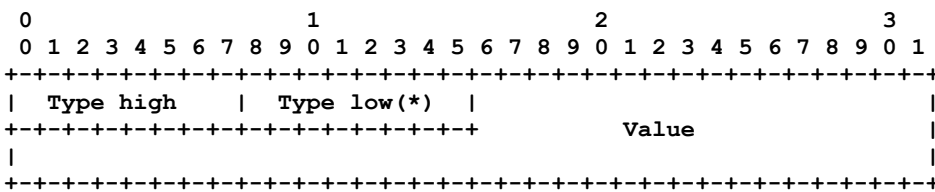
Отметим, что конфигурационный язык может позволять спецификацию этой группы в форме 10876:4338 (4338 – десятичное представление 0x10F2).

Отметим, что эти категории не являются взаимоисключающими и допускается указание множества групп в тех случаях, где это подходит.

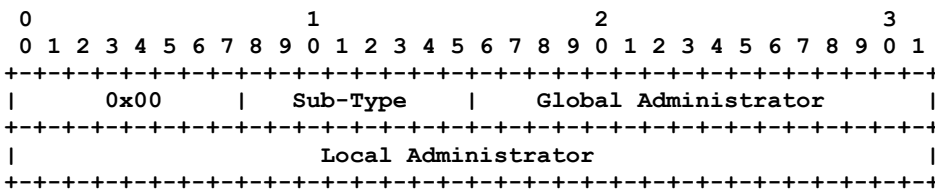
4.1. Расширенные группы

В некоторых случаях значения и их кодирование, описанные в параграфе 4, могут конфликтовать с используемыми сервис-провайдерами значениями. Расширенные группы [RFC4360] обеспечивают удобный механизм, позволяющий избежать подобных конфликтов.

Атрибут Extended Communities является необязательным переходным атрибутом BGP с кодом типа (Type Code) 16. Атрибут содержит в себе набор расширенных групп в показанном ниже формате.



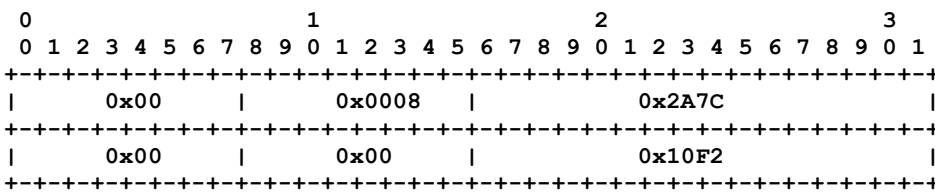
Для сбора данных BGP мы кодируем описанные в параграфе 4 группы с использованием типа two-octet AS specific extended community¹, который имеет следующий формат:



Группа two-octet AS specific extended community attribute содержит 2-октетный номер автономной системы провайдера (присвоенный ему региональным регистратором или RIR²) в поле Global Administrator, а поле Local Administrator может кодировать любую информацию.

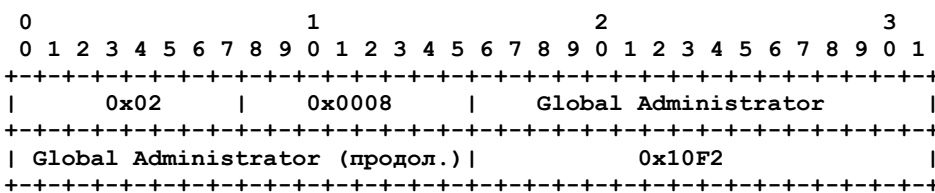
В этом документе выделяется значение субтипа (Sub-Type) 0x0008 для сбора данных BGP и задается значение поля <Value> (в соответствии с параграфом 3.1) для двух младших октетов поля Local Administrator. Два старших октета поля Local Administrator зарезервированы – они устанавливаются в 0x00 при передаче и игнорируются на принимающей стороне.

Например, кодирование расширенной группы 10876:4338, представляющей наземный маршрут в AS 10876 с островов Фиджи, будет иметь вид:



4.2. 4-октетные расширенные группы уровня AS

Группы four-octet AS specific extended community кодируются следующим образом:



В этом случае 4-октетное субполе Global Administrator содержит четырехоктетный номер AS, выделенный IANA.

¹Двухоктетная расширенная группа уровня AS

²Regional Internet Registry

5. Замечание по упаковке сообщений BGP UPDATE

Отметим, что используемые для сбора данных группы могут делать набор атрибутов конкретного маршрута уникальным (поскольку каждый маршрут собирает данные, специфические для его пути через одну или множество AS), что не происходило бы при отсутствии атрибута группы. Это, в свою очередь, может оказывать влияние на группировку маршрутов при упаковке обновлений BGP и приводить к повышенному расходу полосы каналов, ресурсов процессора в маршрутизаторах, а также оперативной памяти.

6. Благодарности

Описанное в этом документе кодирование групп порождено интересным предложением от Akira Kato из WIDE. В частности, ему принадлежит идея использовать набор значений групп для выбора путей, которые будут (к счастью) повышать эффективность доступа к различным типам сервиса. Например, для сервиса DNS на основе RFC 3258 [RFC3258] маршрутизаторы BGP могут видеть множество путей к одному префиксу. Эти маршруты могут иметь общее происхождение, но разные пути через сеть, а некоторые могут различаться по странам и регионам (с той же самой исходной AS).

Joe Abley, Randy Bush, Sean Donelan, Xenofontas Dimitropoulos, Vijay Gill, John Heasley, Geoff Huston, Steve Huter, Michael Patton, Olivier Marce, Ryan McDowell, Rob Rockell, Rob Thomas, Pekka Savola, Patrick Verkaik и Alex Zinin внесли множество полезных комментариев в ранние варианты этого документа. Henk Uijterwaal предложил использовать коды стран ISO-3166-2.

7. Вопросы безопасности

Хотя этот документ не порождает дополнительных вопросов безопасности для протокола BGP, содержащаяся в группах информация, которая определена в данном документе, может в некоторых случаях раскрывать структуру сети, которая до этого не была видна за пределами сети провайдера. В связи с этим следует соблюдать предосторожности при экспорте таких групп в системы сбора маршрутной информации (route collector). Маршруты, экспортируемые в route collector, следует помечать атрибутом группы NO_EXPORT (0xFFFFF01).

7.1. Общий размер атрибутов

Описанные в этом документе группы предназначены для использования на выходе в систему сбора маршрутной информации (route collector). Следовательно, оператор может заменить свои внутренние группы на заданные этим документом значения при экспорте маршрутов в route collector. Однако операторам следует быть уверенными в том, что их реализация BGP корректно будет вести себя в тех случаях, когда в результате добавления атрибутов размер PDU превысит 4096 октетов. Например, поскольку общепринятой практикой является использование групп для реализации политики (скажем, обеспечения потребителям возможности установки таких атрибутов, как LOCAL_PREF), поведение реализации в случаях переполнения пространства атрибутов может играть критическую роль. Некоторые реализации могут в таких случаях захватывать данные атрибута или вызывать неопределенные отказы. Такое поведение может привести к установке неожиданных наборов атрибутов community и, следовательно, оказать нежелательное влияние на политику.

8. Согласование с IANA

Данный документ определяет новый субтип (Sub-Type) для типа AS specific extended community в категории расширенных переходных значений, распределяемой на основе правила First Come First Served¹. Агентство IANA выделило для Sub-Типе значение 0x0008, как определено в параграфе 4.1.

В дополнение к сказанному агентство IANA создало два реестра для BGP Data Collection Communities – один для стандартных², а второй для расширенных групп. Оба эти реестра в своем исходном состоянии включают значения, описанные в параграфе 4. Для выделения новых значений в этих реестрах (в частности, для <Value> или <R> в таблице значений для категорий маршрутов, описанных в параграфе 4) используется процедура IETF Consensus, как описано в [RFC2434]. Обычно это делается через рабочую группу GROW³.

9. Литература

9.1. Нормативные документы

[ISO3166] "ISO 3166 Maintenance agency (ISO 3166/MA)", <http://www.iso.org/iso/en/prods-services/iso3166ma/index.html>, 2004.

[RFC1771] Rekhter, Y. and T. Li (Editors), "A Border Gateway Protocol (BGP-4)", RFC 1771⁴, March 1995.

[RFC1997] Chandra, R. and P. Traina, "BGP Communities Attribute", RFC 1997⁵, August 1996.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360⁵, January 2006.

9.2. Дополнительная литература

[HUSTON] Huston, G., "Interconnection, Peering, and Settlements", http://www.isoc.org/inet99/proceedings/1e/1e_1.htm

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

[RFC3258] Hardie, T., "Distributing Authoritative Name Servers via Shared Unicast Addresses", RFC 3258, April 2002.

¹Распределение в порядке поступления запросов.

²Данные этого реестра доступны по ссылке <http://www.iana.org/assignments/bgp-data-collection-communities-std>. Прим. перев.

³Global Routing Operations

⁴Этот документ заменен RFC 4271. Перевод имеется на сайте <http://www.protocols.ru>. Прим. перев.

⁵Перевод этого документа имеется на сайте <http://www.protocols.ru>. Прим. перев.

- [RIS] "The RIPE Routing Information Service", <http://www.ripe.net/ris>, 2004.
- [RV] Meyer, D., "The Routeviews Project", <http://www.routeviews.org>, 2002.
- [VPLS] Kompella, K., et al., "Virtual Private LAN Service", Work in Progress, April 2005.
- [WANG] Wang, F. and L. Gao, "Inferring and Characterizing Internet Routing Policies"¹, ACM SIGCOMM Internet Measurement Conference 2003.

Адрес автора

David Meyer

E-Mail: dmm@1-4-5.net

Перевод на русский язык

Николай Малых

nmalykh@protocols.ru

Полное заявление авторских прав**Copyright (C) The Internet Society (2006).**

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM 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.

Интеллектуальная собственность

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Подтверждение

Финансирование функций RFC Editor обеспечено IETF Administrative Support Activity (IASA).

¹Эта статья доступна по ссылке http://rio.ecs.umass.edu/~gao/paper/imc_final.pdf.