

Network Working Group
Request for Comments: 5065
Obsoletes: 3065
Category: Standards Track

P. Traina
Blissfully Retired
D. McPherson
Arbor Networks
J. Scudder
Juniper Networks
August 2007

Конфедерации автономных систем в BGP

Autonomous System Confederations for BGP

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

Протокол BGP (Border Gateway Protocol) представляет собой протокол маршрутизации между автономными системами, разработанный для сетей TCP/IP (Transmission Control Protocol/Internet Protocol). BGP требует, чтобы все узлы BGP в одной автономной системе (АС) были связаны между собой (fully meshed - каждый с каждым). Это требование порождает серьезные проблемы масштабирования, отмеченные в многочисленных документах.

Данный документ описывает расширение BGP, которое может использоваться для создания конфедерации автономных систем, представляющей как единая АС для узлов BGP, не входящих в данную конфедерацию. Это позволяет решить проблему полносвязности. Смысл этого расширения протокола состоит в обеспечении дополнительных возможностей управления политикой и упрощения поддержки больших автономных систем.

Данный документ отменяет RFC 3065.

Оглавление

1. Введение.....	1
1.1. Спецификация требований.....	2
1.2. Терминология.....	2
2. Обсуждение.....	2
3. Расширение AS_CONFED для типа сегмента.....	2
4. Принцип работы.....	3
4.1. Правила изменения AS_PATH.....	3
5. Обработка ошибок.....	3
5.1. Общие вопросы администрирования.....	4
5.2. Обработка MED и LOCAL_PREF.....	4
5.3. AS_PATH и выбор пути.....	4
6. Вопросы совместимости.....	4
7. Развертывание конфедераций.....	4
8. Вопросы безопасности.....	5
9. Благодарности.....	5
10. Литература.....	5
10.1. Нормативные документы.....	5
10.2. Дополнительная литература.....	5
Приложение А. Агрегирование маршрутной информации.....	5
Приложение В. Отличия от RFC 3065.....	5

1. Введение

В соответствии со спецификацией протокол BGP требует¹, чтобы каждый узел BGP в АС имел соединения со всеми другими узлами BGP в этой автономной системе (fully meshed). Результатом этого является необходимость поддержки каждым узлом BGP (в автономной системе с числом узлов n) $n^2(n-1)/2$ уникальных сессий IBGP². Это требование полносвязности создаёт проблемы масштабирования для АС с большим числом узлов IBGP, обычным для многих современных сетей.

Эта проблема масштабирования описана во множестве документов и существует целый ряд предложений по ослаблению проблемы [RFC2796, RFC1863³]. В этом документе предлагается дополнительный способ ослабления

¹В спецификации протокола это требование не выражено явно, однако оно логически следует из принципов работы протокола в рамках одной АС. *Прим. перев.*

²Internal BGP - внутренний (по отношению к АС) протокол BGP.

³RFC 4223 перевел этот документ в состояние historic (достояние истории).

проблемы полновязности, известный как Autonomous System Confederations for BGP или просто BGP Confederations⁴. Можно также сказать, что конфедерации BGP могут повышать эффективность управления политикой маршрутизации.

Этот документ является пересмотром [RFC3065], который, в свою очередь пересматривает и отменяет [RFC1965]. Документ включает редакторские правки, разъяснения терминов и более явные спецификации протокола, основанные на опыте развертывания конфедераций BGP.

1.1. Спецификация 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 должны интерпретироваться в соответствии с [RFC2119].

1.2. Терминология

AS Confederation - конфедерация AS

Группа автономных систем, анонсируемых с общим номером AS узлам BGP, не входящим в данную конфедерацию.

AS Confederation Identifier - идентификатор конфедерации AS

Видимый снаружи номер автономной системы, идентифицирующий конфедерацию в целом.

Member Autonomous System (Member-AS) - AS - член конфедерации

Автономная система, включенная в данную конфедерацию AS. Отметим, что термины Member Autonomous System и Member-AS в данном документе используются как синонимы.

Member-AS Number - номер AS - члена конфедерации

Номер автономной системы, видимый только членам данной конфедерации BGP и используемый для представления Member-AS внутри конфедерации.

2. Обсуждение

Может оказаться полезным деление автономных систем с большим числом узлов BGP на более мелкие домены в целях управления политикой маршрутизации с помощью данных, содержащихся в атрибуте BGP AS_PATH. Например, можно рассматривать все узлы BGP в некоем географическом регионе, как единый объект.

Кроме потенциального повышения уровня контроля за политикой маршрутизации (если не используются методы типа описанного здесь или в [RFC4456]) спецификация [BGP-4] требует, чтобы все узлы BGP одной автономной системы организовали полновязную (full mesh) систему соединений TCP со всеми другими узлами для обмена внешней маршрутной информацией. В автономных системах число внутримоментных соединений, которые должен поддерживать каждый граничный маршрутизатор, может становиться весьма значительным.

Деление большой автономной системы на части позволяет существенно снизить число внутримоментных соединений BGP, поскольку в этом случае требования связности упрощаются до уровня модели, используемой для междоментных соединений.

К сожалению деление автономной системы на части может увеличить сложность реализации политики маршрутизации на основе данных из атрибутов AS_PATH для всех членов Internet. Кроме того, это деление увеличит издержки на управление и координацию внешних партнерских связей при изменении внутренней топологии этого множества автономных систем.

Следовательно, деление автономной системы на отдельные фрагменты может оказать существенное влияние на уровень оптимальности маршрутизации пакетов через Internet.

Однако обычно не возникает необходимости в раскрытии внутренней топологии таких поделенных на части автономных систем и это означает, что возможно представление множества автономных систем с единым администрированием как единого объекта или AS с точки зрения находящихся за пределами данной конфедерации автономных систем.

3. Расширение AS_CONFED для типа сегмента

В действующей спецификации BGP сказано, что атрибут AS_PATH является общеизвестным обязательным атрибутом, состоящим из последовательности сегментов пути (AS path segment). Каждый сегмент пути представляется триплетом <path segment type, path segment length, path segment value>.

В [BGP-4] тип сегмента пути задается однооктетным полем, для которого определены следующие значения:

Значение	Имя	Описание
1	AS_SET	Неупорядоченный набор AS, через которые проходит маршрут из сообщения UPDATE.
2	AS_SEQUENCE	Упорядоченный набор AS, через которые проходит маршрут из сообщения UPDATE.

В настоящем документе определены два дополнительных типа сегментов пути.

Значение	Имя	Описание
3	AS_CONFED_SEQUENCE	Упорядоченный набор номеров Member AS в локальной конфедерации, через которые передается сообщение UPDATE.
4	AS_CONFED_SET	Неупорядоченный набор номеров Member AS в локальной конфедерации, через которые передается сообщение UPDATE.

⁴Конфедерации автономных систем для BGP или просто BGP-конфедерации.

4. Принцип работы

Член конфедерации BGP **должен** использовать свое значение AS Confederation ID во всех транзакциях с партнерами, которые не являются членами данной конфедерации. Этот идентификатор конфедерации рассматривается как видимый извне номер AS, который используется в сообщениях OPEN и анонсируется в атрибуте AS_PATH.

Член конфедерации BGP **должен** использовать свое значение Member-AS Number во всех транзакциях с партнерами, входящими в данную конфедерацию.

Узлу BGP, получившему атрибут AS_PATH с номером автономной системы, соответствующим идентификатору его конфедерации, **следует** трактовать путь так же, как это делается для путей, содержащих номер AS данного узла.

Узлу BGP, получившему атрибут AS_PATH с сегментами AS_CONFED_SEQUENCE или AS_CONFED_SET, содержащими его Member-AS Number, **следует** трактовать путь так же, как это делается для путей, содержащих номер AS данного узла.

4.1. Правила изменения AS_PATH

При реализации конфедераций BGP параграф 5.1.2 документа [BGP-4] заменяется приведенным ниже текстом.

AS_PATH представляет собой общеизвестный обязательный атрибут. Этот атрибут идентифицирует автономные системы, через которые передается маршрутная информация в данном сообщении UPDATE. Компонентами этого списка могут быть AS_SET, AS_SEQUENCE, AS_CONFED_SET и AS_CONFED_SEQUENCE.

Когда узел BGP распространяет маршрут, который был получен в сообщении UPDATE от другого узла BGP, он изменяет атрибут AS_PATH с учетом размещения узла BGP, которому передается маршрут:

а) Когда данный узел BGP анонсирует маршрут другому узлу BGP, расположенному в его Member-AS, анонсирующему узлу **не следует** изменять связанный с этим маршрутом атрибут AS_PATH.

б) Когда данный узел BGP анонсирует маршрут узлу BGP, расположенному в соседней автономной системе, являющейся членом локальной конфедерации, анонсирующему узлу следует изменить атрибут AS_PATH как показано ниже.

1) Если первый сегмент AS_PATH имеет тип AS_CONFED_SEQUENCE, локальной системе следует поместить свой номер Member-AS, как последний элемент списка (в крайнюю левую позицию протокольного сообщения). Если при этом будет возникать переполнение сегмента AS_PATH (более 255 номеров AC), узлу **следует** добавить (prepend) новый сегмент типа AS_CONFED_SEQUENCE и поместить свой номер AC в этот сегмент.

2) Если первый сегмент AS_PATH имеет тип, отличный от AS_CONFED_SEQUENCE, локальная система помещает (prepend) новый сегмент типа AS_CONFED_SEQUENCE в путь AS_PATH, включив свой номер Member-AS в этот сегмент.

3) Если значение AS_PATH пусто, локальная система создает сегмент пути типа AS_CONFED_SEQUENCE, включает в него свой номер Member-AS и помещает этот сегмент в AS_PATH.

с) Когда данный узел BGP анонсирует маршрут узлу BGP, расположенному в соседней автономной системе, которая не входит в локальную конфедерацию, анонсирующему узлу **следует** изменить атрибут AS_PATH как показано ниже.

1) Если любые сегменты AS_PATH имеет тип AS_CONFED_SEQUENCE или AS_CONFED_SET, эти сегменты **должны** удаляться из атрибута AS_PATH и после этого для атрибута выполняется этап 2, 3 или 4 (см. ниже).

2) Если первый оставшийся сегмент AS_PATH имеет тип AS_SEQUENCE, локальная система добавляет (prepend) свой идентификатор конфедерации (AS Confederation Identifier) в качестве последнего элемента последовательности (в крайнюю левую позицию протокольного сообщения). Если при этом будет возникать переполнение сегмента AS_PATH (более 255 номеров AC), узлу **следует** добавить (prepend) новый сегмент типа AS_SEQUENCE и поместить свой номер AC в этот сегмент.

3) Если первый сегмент оставшейся части AS_PATH имеет тип AS_SET, локальная система добавляет (prepend) в AS_PATH новый сегмент типа AS_SET, включая в него свой идентификатор конфедерации.

4) Если оставшаяся часть AS_PATH пуста, локальная система создает сегмент пути типа AS_SEQUENCE, включает в него свой идентификатор конфедерации и помещает этот сегмент в AS_PATH.

Когда узел BGP является источником маршрута, этот узел:

а) включает свой идентификатор конфедерации AC в сегмент пути типа AS_SEQUENCE атрибутов AS_PATH во все сообщения UPDATE, передаваемые узлом BGP в соседних AC, не являющихся членами локальной конфедерации; в этом случае идентификатор конфедерации автономной системы источника маршрута будет единственной записью в сегменте пути, а этот сегмент будет единственным в атрибуте AS_PATH;

б) включает свой номер Member-AS в сегмент пути типа AS_CONFED_SEQUENCE атрибутов AS_PATH всех сообщений UPDATE, передаваемых узлом BGP в соседних AS, являющихся членами локальной конфедерации; в этом случае Member-AS Number автономной системы источника маршрута будет единственной записью в сегменте пути, а этот сегмент будет единственным в атрибуте AS_PATH;

с) включает номер пустой атрибут AS_PATH во все сообщения UPDATE, передаваемые узлом BGP в той Member-AS (пустой атрибут AS_PATH содержит нулевое значение в поле размера).

Локальная система **может** включать/добавлять в начало более одного экземпляра идентификатора конфедерации или номера Member-AS в атрибут AS_PATH. Количество экземпляров задается локальной конфигурацией.

5. Обработка ошибок

Для узла BGP **недопустимо** передавать обновления, содержащие атрибуты AS_CONFED_SET или AS_CONFED_SEQUENCE, партнерам, которые не входят в локальную конфедерацию.

Получение узлом BGP сообщения UPDATE с атрибутом AS_PATH, включающим сегменты AS_CONFED_SEQUENCE или AS_CONFED_SET от соседа, не относящегося к той же конфедерации, является ошибкой. Если узел BGP получает такое сообщение UPDATE, **следует** трактовать его, как сообщение с некорректным атрибутом AS_PATH в соответствии с процедурами [BGP-4] (параграф 6.3 «Обработка ошибок в сообщениях UPDATE»).

Ошибкой является получение узлом BGP обновления от партнера по конфедерации, не относящегося к той же Member-AS, в котором нет AS_CONFED_SEQUENCE в качестве первого сегмента. Если узел BGP получает такое сообщения UPDATE, **следует** трактовать его, как сообщение с некорректным атрибутом AS_PATH в соответствии с процедурами [BGP-4] (параграф 6.3 «Обработка ошибок в сообщениях UPDATE»).

5.1. Общие вопросы администрирования

Для AC, являющихся членами конфедерации, разумно в рамках конфедерации использовать единое администрирование и информацию IGP¹. Разумно также для каждой Member-AS использовать независимый протокол IGP. В последнем случае может потребоваться установка NEXT_HOP с использованием политики (по умолчанию это значение не меняется).

5.2. Обработка MED и LOCAL_PREF

Узлам BGP следует разрешить анонсирование без изменения атрибутов NEXT_HOP и MULTI_EXIT_DISC (MED) партнерам из соседних Member-AS, входящих в ту же конфедерацию.

Значения MED двух маршрутов **следует** сравнивать только в случаях совпадения первой AC в первом сегменте AS_SEQUENCE обоих маршрутов (т. е., все AC в AS_CONFED_SET и AS_CONFED_SEQUENCE пропускаются). Реализация **может** обеспечивать возможность настройки выбора пути таким образом, чтобы значения MED двух маршрутов были сравнимы, если первые AC в атрибутах AS_PATH совпадают, независимо от сегментов AS_SEQUENCE или AS_CONFED_SEQUENCE в AS_PATH.

Реализация **может** сравнивать значения MED, полученные из Member-AS через множество путей. Реализация **может** сравнивать значения MED из различных AC, относящихся к одной конфедерации.

В дополнение к этому отменяется ограничение на передачу атрибута LOCAL_PREF партнерам из соседних AS своей конфедерации.

5.3. AS_PATH и выбор пути

Критерии выбора пути для информации, полученной от членов конфедерации, **должны** следовать тем же правилам, которые используются для информации, полученной от партнеров в своей AC, как задано в [BGP-4].

В дополнение к этому **следует** применять перечисленные ниже правила:

- 1) если AS_PATH является внутренним по отношению к локальной конфедерации (т. е., содержит только сегменты AS_CONFED_*), соседняя AC рассматривается, как локальная AC;
- 2) в остальных случаях, если первый сегмент пути не относится к типу AS_CONFED_SEQUENCE или AS_CONFED_SET является AS_SEQUENCE, соседняя AC рассматривается, как самая левая в списке AS_SEQUENCE;
- 3) при сравнении маршрутов по длине AS_PATH сегменты CONFED_SEQUENCE и CONFED_SET **не следует** учитывать;
- 4) при сравнении маршрутов с использованием внутренних (полученных от IBGP) и внешних (от EBGP) данных, маршрут, полученный от партнера в той же конфедерации (не обязательно в той же Member-AS), трактуется как «внутренний».

6. Вопросы совместимости

Все включенные в конфедерацию узлы BGP **должны** распознавать расширения типа сегмента AS_CONFED_SET и AS_CONFED_SEQUENCE в атрибутах AS_PATH.

Любой узел BGP, не поддерживающий эти расширения, будет генерировать сообщение NOTIFICATION с кодом ошибки UPDATE Message Error и субкодом Malformed AS_PATH.

Перечисленные выше проблемы совместимости **требуют** от всех включаемых в конфедерацию узлов BGP поддержки данного расширения (BGP confederations). Однако от узлов BGP за пределами конфедерации такой поддержки не требуется.

7. Развертывание конфедераций

Конфедерации BGP широко распространены в сети Internet уже много лет и поддерживаются множеством производителей.

Некорректная настройка конфедерации BGP может приводить к ненужному дублированию маршрутной информации внутри AS. Такое дублирование будет отнимать системные ресурсы, приводить к ненужным переключениям маршрутов (flap) и увеличивать задержку схождения (convergence).

Следует принять меры по фильтрации вручную дубликатов анонсов, вызванных прохождением информации о доступности через множество включенных в конфедерацию автономных систем, обусловленным топологией конфедерации и требованиями по резервированию соединений.

В [RFC3345] показано, что конфедерации (как и рефлекторы маршрутов), исключая из рассмотрения информацию о доступности в различных точках конфедерации, могут вызывать постоянные осцилляции между маршрутами-

¹Interior Gateway Protocol - протокол внутренней маршрутизации.

кандидатами при использовании правил «развязывания узлов» (tie breaking), требуемых спецификацией [BGP-4]. Следует с осторожностью относиться к выбору значений MED и правилам tie breaking для предотвращения проблем.

Одним из способов предотвращения проблем является установка для метрики inter-Member-AS IGP¹ значения большего, нежели для метрики intra-Member-AS IGP², и/или использование иных правил tie breaking для предотвращения выбора маршрутов BGP на основе несравнимых MED.

8. Вопросы безопасности

Это расширение не оказывает влияния на вопросы безопасности протокола BGP, рассмотренные в [RFC2385] и [BGP-VULN].

9. Благодарности

Общая концепция конфедераций BGP была заимствована из Routing Domain Confederations для IDRП [ISO10747]. Часть вводного текста этого документа была заимствована из [RFC2796].

Авторы выражают свою признательность Jeffrey Haas за многочисленные отзывы. Мы также благодарим Bruce Cole, Srihari Ramachandra, Alex Zinin, Naresh Kumar Paliwal, Jeffrey Haas, Cengiz Alaettinoglu, Mike Hollyman и Bruno Rijsman за их отзывы и предложения.

Авторы также выражают свою признательность Ravi Chandra и Yakov Rekhter за конструктивные и полезные замечания в процессе работы с ранними версиями этого документа.

10. Литература

10.1. Нормативные документы

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

[RFC1965] Traina, P., "Autonomous System Confederations for BGP", [RFC 1965](#), June 1996.

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

[RFC3065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", [RFC 3065](#), February 2001.

10.2. Дополнительная литература

[ISO10747] Kunzinger, C., Editor, "Inter-Domain Routing Protocol", ISO/IEC 10747, October 1993.

[RFC1863] Haskin, D., "A BGP/IDRP Route Server alternative to a full mesh routing", RFC 1863, October 1995.

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

[RFC3345] McPherson, D., Gill, V., Walton, D., and A. Retana, "Border Gateway Protocol (BGP) Persistent Route Oscillation Condition", RFC 3345, August 2002.

[RFC4223] Savola, P., "Reclassification of RFC 1863 to Historic", RFC 4223, October 2005.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", [RFC 4272](#), January 2006.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", [RFC 4456](#), April 2006.

Приложение А. Агрегирование маршрутной информации

Как практическая задача, агрегирование, рассмотренное в [BGP-4] (параграф 9.2.2.2), в общем случае нечасто используется с конфедерациями. Однако при использовании такого агрегирования в конфедерации, следует выполнять правила [BGP-4] с заменой AS_SET на AS_CONFED_SET и AS_SEQUENCE на AS_CONFED_SEQUENCE. Относящиеся к конфедерациям сегменты (AS_CONFED_SET и AS_CONFED_SEQUENCE) **должны** сохраняться отдельно от не связанных с конфедерациями сегментов (AS_SET и AS_SEQUENCE). Реализация может также выюрать форму агрегирования, при которой не связанные с конфедерациями сегменты агрегируются в соответствии с [BGP-4] (параграф 9.2.2.2), а относящиеся к конфедерациям сегменты не агрегируются.

Поддержка агрегирования для связанных с конфедерациями сегментов не является обязательной.

Приложение В. Отличия от RFC 3065

Основной причиной обновления RFC 3065 послужили вопросы, связанные с обработкой сегментов пути, - в частности, взаимодействие с внешними по отношению к конфедерации партнерами BGP и предотвращение анонсирования сегментов типа AS_CONFED_[SET|SEQUENCE] за пределы конфедерации.

В связи с этим был добавлен раздел 5. Обработка ошибок, относящийся ко всем узлам BGP, а не только к членам конфедерации.

К числу других изменений относятся некоторое разъяснение и уточнение терминологии и указание того, что обработку сегментов типа AS_CONFED_[SET|SEQUENCE] следует выполнять в соответствии с базовой спецификацией [BGP-4].

Адреса авторов

Paul Traina

¹Протокол внутренней маршрутизации между членами конфедерации.

²Протокол внутренней маршрутизации для входящих в конфедерацию AS.

Blissfully Retired

Email: bgp-confederations@st04.pst.org

Danny McPherson

Arbor Networks

EMail: danny@arbor.net

John G. Scudder

Juniper Networks

EMail: jgs@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.