

Network Working Group
Request for Comments: 3065
Obsoletes: 1965
Category: Standards Track

P. Traina
Juniper Networks, Inc.
D. McPherson
Amber Networks, Inc.
J. Scudder
Cisco Systems, Inc.
February 2001

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 Internet Society (2001). All Rights Reserved.

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 (AS) были связаны между собой (fully meshed - «каждый с каждым»). Это требование порождает серьёзные проблемы масштабирования, отмеченные в многочисленных документах.

Данный документ описывает расширение BGP, которое может использоваться для создания конфедерации автономных систем, представляющей как единая AS 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.

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].

2. Введение

В соответствии со спецификацией протокол BGP требует¹, чтобы каждый узел BGP в AS имел соединения со всеми другими узлами BGP в этой автономной системе. Результатом этого является необходимость поддержки каждым узлом BGP (в автономной системе с числом узлов n) $n*(n-1)/2$ уникальных сессий IBGP. Это требование «полносвязности» создаёт проблемы масштабирования для AS 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 [3,5]. В этом документе предлагается дополнительный способ ослабления проблемы полносвязности, известный как Autonomous System Confederations for BGP или просто BGP Confederations². Можно также сказать, что конфедерации BGP **могут** повышать эффективность управления политикой маршрутизации.

Этот документ является пересмотром RFC 1965 [4] и включает редакторские правки, разъяснения и корректировки, основанные на опыте развёртывания конфедераций BGP. Изменения по сравнению с упомянутым документом перечислены в Приложении А.

3. Термины и определения

AS Confederation

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

AS Confederation Identifier

Видимый снаружи номер автономной системы, идентифицирующий конфедерацию в целом.

Member-AS

Автономная система, включённая в данную конфедерацию AS.

Member-AS Number

Номер автономной системы, видимый только членам данной конфедерации BGP.

4. 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,

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

²Конфедерации автономных систем для BGP или просто BGP-конфедерации

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 документе [5]) спецификация протокола [1] требует, чтобы все узлы BGP одной автономной системы организовали полносвязную (full mesh) систему соединений TCP со всеми другими узлами для обмена внешней маршрутной информацией. В автономных системах число внутридоменных соединений, которые должен поддерживать каждый граничный маршрутизатор, может становиться весьма значительным.

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

И, наконец, деление больших AS может привести к неоправданному увеличению размера упорядоченной части атрибута AS_PATH. Некоторые распространённые реализации BGP могут использовать число "интервалов между AS" (AS hop) до данного адресата, как часть критерия выбора пути. Хотя такой метод определения предпочтительного маршрута не является оптимальным, при нехватке других данных он обеспечивает разумное поведение "по умолчанию", широко используемое в сети Internet. 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 поделённых 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.

5. 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 соответствии с [1] тип сегмента пути задаётся 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 передаётся сообщение UPDATE.
4	AS_CONFED_SET	Неупорядоченный набор номеров Member AS в локальной конфедерации, через которые передаётся сообщение UPDATE.

6. Принцип работы

Член конфедерации BGP будет использовать своё значение AS Confederation ID во всех транзакциях с партнёрами, которые не являются членами данной конфедерации. Этот идентификатор конфедерации рассматривается как видимый извне номер AS, этот номер используется в сообщениях OPEN и анонсируется в атрибуте AS_PATH.

Член конфедерации BGP будет использовать своё значение Member AS Number во всех транзакциях с партнёрами, входящими в данную конфедерацию.

Узлу BGP, получившему атрибут AS_PATH с номером автономной системы, соответствующим 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.

6.1. Правила изменения AS_PATH

Параграф 5.1.2 документа [1] заменяется приведённым ниже текстом:

Когда узел BGP распространяет маршрут, который был получен в сообщении UPDATE от другого узла BGP, ему следует изменить атрибут AS_PATH с учётом размещения узла BGP, которому передаётся маршрут:

a) Когда данный узел BGP анонсирует маршрут другому узлу BGP, расположенному в той же автономной системе, анонсирующему узлу не следует изменять связанный с этим маршрутом атрибут AS_PATH.

b) Когда данный узел BGP анонсирует маршрут узлу BGP, расположенному в соседней автономной системе, анонсирующему узлу следует изменить связанный с этим маршрутом атрибут AS_PATH как показано ниже:

1) если первый сегмент AS_PATH имеет тип AS_CONFED_SEQUENCE, локальной системе следует поместить свой номер AS как последний элемент списка (в крайнюю левую позицию).

2) если первый сегмент AS_PATH имеет тип, отличный от AS_CONFED_SEQUENCE, локальной системе следует поместить новый сегмент типа AS_CONFED_SEQUENCE в путь AS_PATH, включив свой идентификатор конфедерации в этот сегмент.

с) Когда данный узел BGP анонсирует маршрут узлу BGP, расположенному в соседней автономной системе и не являющемуся членом конфедерации (в которую входит данный узел), анонсирующему узлу следует изменить атрибут AS_PATH как показано ниже:

- 1) если первый сегмент AS_PATH имеет тип AS_CONFED_SEQUENCE, данный сегмент и все непосредственно следующие за ним сегменты типа AS_CONFED_SET или AS_CONFED_SEQUENCE удаляются из атрибута AS_PATH и после этого для атрибута выполняется этап 2 или 3 (см. ниже).
- 2) Если первый оставшийся сегмент AS_PATH имеет тип AS_SEQUENCE, локальной системе следует поместить свой идентификатор конфедерации в качестве последнего элемента (в крайнюю левую позицию).
- 3) Если после удаления сегментов AS_CONFED_SET/AS_CONFED_SEQUENCE не остаётся других сегментов пути или первый сегмент оставшейся части AS_PATH имеет тип AS_SET, локальной системе следует включить (prepend) в атрибут AS_PATH новый сегмент типа AS_SEQUENCE, указав в нем свой идентификатор конфедерации.

Когда узел BGP является источником маршрута, этому узлу следует:

- a) включать пустой атрибут AS_PATH во все сообщения UPDATE, передаваемые узлам BGP в Member AS Number (пустой атрибут AS_PATH содержит нулевое значение в поле размера);
- b) включать свой Member AS Number в сегмент AS_CONFED_SEQUENCE атрибута AS_PATH всех сообщений UPDATE, передаваемых узлам BGP в соседних AS, являющихся членами конфедерации (т. е., значение Member AS Number исходного отправителя будет единственным элементом атрибута AS_PATH);
- c) включать номер AS своей конфедерации в сегмент AS_SEQUENCE атрибута AS_PATH всех сообщений UPDATE, передаваемых узлам BGP в соседних AS, которые не являются членами данной конфедерации; в этом случае номер AS конфедерации будет единственным элементом атрибута AS_PATH.

7. Общие вопросы администрирования

Вполне разумно в рамках конфедерации использовать единое администрирование и информацию IGP.

Узлам BGP следует разрешить анонсирование без изменения атрибутов NEXT_HOP и MULTI_EXIT_DISCRIMINATOR (MED) партнёрам из соседних AS, входящих в ту же конфедерацию. В дополнение к этому отменяется ограничение на передачу атрибута LOCAL_PREFERENCE партнёрам из соседних AS своей конфедерации. Критерии выбора пути для информации, полученной от членов конфедерации, **должны** быть такими же, какие указаны для партнёров из своей автономной системы в спецификации [1].

8. Вопросы совместимости

Все включённые 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.

9. Развёртывание 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 включённых в конфедерацию автономных систем, обусловленным топологией конфедерации и требованиями по резервированию соединений.

В дополнение к этому конфедерации (как и рефлекторы маршрутов), исключая из рассмотрения информацию о доступности в различных точках конфедерации, могут вызывать постоянные осцилляции между маршрутами-кандидатами при использовании правил "развязывания узлов" (tie breaking), требуемых спецификацией BGP [1]. 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.

10. Вопросы безопасности

Это расширение не оказывает влияния на вопросы безопасности протокола BGP, рассмотренные в документе [6].

11. Благодарности

Общая концепция конфедераций BGP была заимствована из Routing Domain Confederations для IDRП [2]. Часть вводного текста этого документа была заимствована из [5].

¹Протокол внутренней маршрутизации между членами конфедерации.

²Протокол внутренней маршрутизации для входящих в конфедерацию AS.

Авторы выражают свою признательность Bruce Cole из Juniper Networks за информацию о реализации и анализ ограничений расширения протокола, описанного в данном документе и работе [5]. Мы также благодарим Srihari Ramachandra из Cisco Systems, Inc. за комментарии. Авторы также выражают свою признательность Ravi Chandra и Yakov Rekhter за конструктивные и полезные замечания в процессе работы с ранними версиями этого документа.

12. Литература

- [1] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", [RFC 1771](#), March 1995.
- [2] Kunzinger, C., Editor, "Inter-Domain Routing Protocol", ISO/IEC 10747, October 1993.
- [3] Haskin, D., "A BGP/IDRP Route Server alternative to a full mesh routing", RFC 1863, October 1995.
- [4] Traina, P. "Autonomous System Confederations for BGP", [RFC 1965](#), June 1996.
- [5] Bates, T., Chandra, R. and E. Chen, "BGP Route Reflection An Alternative to Full Mesh IBGP", [RFC 2796](#), April 2000.
- [6] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", [RFC 2385](#), August 1998.

13. Адреса авторов

Paul Traina

Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089 USA

Phone: +1 408 745-2000

Email: pst+confed@juniper.net

Danny McPherson

Amber Networks, Inc.
48664 Milmont Drive

Fremont, CA 94538

Phone: +1 510.687.5226

Email: danny@ambernetworks.com

John G. Scudder

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134

Phone: +1 734.669.8800

Email: jgs@cisco.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Приложение А: Сравнение с RFC 1965

Наиболее важным отличием от [4]¹ является замена значений AS_CONFED_SEQUENCE(4) и AS_CONFED_SET(3) на значения, определенные в параграфе «Расширение AS_CONFED для типа сегмента». Причина этой замены обусловлена тем, что в начальных реализациях, которые достаточно широко распространены, были использованы значения, поменянные местами по сравнению с [4] и такое использование продолжалось в последующих реализациях. Для обеспечения интероперабельности и совместимости с имеющимися реализациями значения в данном документе были поменяны местами.

Параграф «Проблемы совместимости²» был исключен, а вопросы из этого параграфа рассмотрены в других частях данного документа. Были также исключены упоминания иерархических конфедераций. Термин Routing Domain Identifier был заменен термином Member AS Number.

Расширен параграф «Развертывание конфедераций» и в него были внесены некоторые редакторские правки.

Полное заявление авторских прав

Copyright (C) The Internet Society (2001). Все права защищены.

Этот документ и его переводы могут копироваться и предоставляться другим лицам, а производные работы, комментирующие или иначе разъясняющие документ или помогающие в его реализации, могут подготавливаться, копироваться, публиковаться и распространяться целиком или частично без каких-либо ограничений при условии сохранения указанного выше уведомления об авторских правах и этого параграфа в копии или производной работе. Однако сам документ не может быть изменён каким-либо способом, таким как удаление уведомления об авторских правах или ссылок на Internet Society или иные организации Internet, за исключением случаев, когда это необходимо для разработки стандартов Internet (в этом случае нужно следовать процедурам для авторских прав, заданных процессом Internet Standards), а также при переводе документа на другие языки.

Предоставленные выше ограниченные права являются бессрочными и не могут быть отозваны Internet Society или правопреемниками.

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.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.

¹В оригинале эта ссылка ошибочно указывает на документ [1]. Прим. перев.

²Compatibility Discussion.