Internet Engineering Task Force (IETF)
Request for Comments: 7454
BCP: 194
Category: Best Current Practice
ISSN: 2070-1721

J. Durand
Cisco Systems, Inc.
I. Pepelnjak
NIL
G. Doering
SpaceNet

February 2015

Функционирование и безопасность BGP

BGP Operations and Security

Аннотация

Протокол BGP (Border Gateway Protocol – протокол граничного шлюзе) практически исключительно применяется в сети Internet для обмена маршрутной информацией между доменами сети. Такая роль протокола делает важным понимание мер защиты, которые можно и следует принимать для предотвращения случайных или преднамеренных нарушений картины маршрутизации.

В этом документе описаны меры защиты сессий BGP, как таковых, включая время жизни TTL (Time to Live), опцию TCP-AO (TCP Authentication Option - опция аутентификации TCP), а также фильтрацию на уровне управления. Описаны также меры по улучшению контроля потоков маршрутной информации, использование фильтрации префиксов и ее автоматизации, фильтрации max-prefix и путей AS[1], подавления маршрутных осцилляций и очистке групп BGP.

Статус документа

Этот документ относится к категории обмена опытом (Internet Best Current Practice).

Документ является результатом работы IETF[2] и представляет согласованный взгляд сообщества IETF. Документ прошёл открытое обсуждение и был одобрен для публикации IESG[3]. Дополнительная информация о документах BCP представлена в разделе 2 документа RFC 5741.

Информация о текущем статусе этого документа, обнаруженных ошибках и способах обратной связи может быть найдена по ссылке http://www.rfc-editor.org/info/rfc7454.

Оглавление

[1]Autonomous System - автономная система.
[2]Internet Engineering Task Force - комиссия по решению инженерных задач Internet.
[3]Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

1. Введение

Протокол BGP, описанный в RFC 4271 [2], является протоколом Internet для обмена маршрутной информацией между доменами сети. BGP непосредственно не включает механизмов управления набором маршрутов для обмена, соответствующих рекомендациям, выработанным сообществом Internet. Задачей данного документа является сбор воедино существующих рекомендаций и помощь сетевым администраторам при создании согласованных правил BGP.

1.1. Уровни 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 RFC 2119 [1].

2. Сфера применимости документа

Рекомендации, приведенные в этом документе, рассчитаны на типовые партнерские отношения Internet BGP. Природа Internet такова, что автономные системы всегда могут согласовать между собой исключения из общей схемы для соответствующих узлов и, следовательно, настроить сессии BGP так, что параметры конфигурации будут отличаться от приведенных рекомендаций. Это допустимая (и нормальная) практика, однако каждое исключение может повлиять на всю среду междоменной маршрутизации и администраторам **следует** осторожно применять такие исключения.

3. Определения и сокращения

- ACL: Access Control List - список управления доступом.
- ASN: Autonomous System Number - номер автономной системы.
- IRR: Internet Routing Registry - реестр маршрутизации Internet.
- IXP: Internet Exchange Point - точка обмена трафиком Internet.
- LIR: Local Internet Registry - локальный регистратор Internet.
- PMTUD: Path MTU Discovery - определение MTU для пути.
- RIR: Regional Internet Registry - региональный регистратор Internet.
- Tier 1 transit provider - транзитный провайдер IP, способный достичь любой сети в Internet без приобретения услуг по транзиту.
- uRPF: Unicast Reverse Path Forwarding - индивидуальная пересылка по обратному пути.

В дополнение к приведенным выше сокращениям в документе используется пара специфических терминов.

- Downstream - любая сеть, расположенная в нисходящем[1] направлении (сеть провайдера или пользователя).
- Upstream - любая сеть, расположенная в восходящем направлении.

[1]Восходящим считается направление от потребителя услуг (пользователя), нисходящим - к нему. *Прим. перев.*

4. Защита узла BGP

Узлы BGP требуется защищать от попыток повреждения сессий BGP. Для обеспечения такой защиты **следует** пользоваться списками управления доступом (ACL), которые будут отбрасывать все пакеты, направленные в порт TCP 179 на локальном устройстве, если их отправитель не включен в число разрешенных партнеров (соседей) BGP. Опыт показывает, что естественной защиты протокола TCP в приложениях уровня управления не всегда достаточно. При отсутствии ACL возможна организация атак на узел BGP путем простой передачи ему большого числа на соединение.

Если поддерживаются списки ACL для уровня управления маршрутизатором (receive-ACL, control-plane policing, и т. п.), их также **следует** применять, чтобы избавиться от необходимости фильтрации на уровне данных пакетов, проходящих через маршрутизатор (и, следовательно, не попадающих на уровень управления). Если оборудование не позволяет этого, можно применять ACL на интерфейсах для блокировки пакетов, адресованных локальному маршрутизатору.

Некоторые маршрутизаторы автоматически создают ACL при настройке BGP, а на остальных устройствах такие ACL следует настраивать и поддерживать вручную или с помощью программ-сценариев (script).

В дополнение к строгой фильтрации **может** быть задано скоростное ограничение для допустимого трафика BGP. Ограничение трафика BGP по скорости позволяет принимать не более заданного числа битов (или пакетов) в секунду для трафика BGP на уровне управления. Это защищает уровень управления маршрутизатором от перегрузки управляющим трафиком BGP.

Фильтрация и ограничение скорости для уровня управления применяются не только для BGP (если маршрутизатор перегружен управляющим трафиком других протоколов, это может сказаться и на BGP). Более подробные рекомендации по защите уровня управления маршрутизатором приведены в RFC 6192 [11].

5. Защита сессий BGP

Современные вопросы безопасности протоколов, основанных на TCP (следовательно, и BGP) описаны в RFC 6952 [14]. В последующих параграфах рассматриваются основные вопросы, поднятые в этом RFC, и даны основанные на опыте рекомендации по защите сессий TCP, используемых в работе BGP.

5.1. Защита сессий TCP, используемых протоколом BGP

Атака на сессии TCP, используемые протоколом BGP (BGP-сессии), например, путем отправки подставных пакетов TCP RST, может разорвать связь между партнерами BGP.

После успешной атаки с подменой ARP (или аналогичной MITM-атаки[1]) злоумышленник даже сможет помещать свои пакеты в поток TCP (атака на маршрутизацию).

Сессии BGP можно защитить с помощью разных механизмов. Защита MD5 в заголовках сессий TCP, описанная в RFC 2385 [7], была первым из таких механизмов. Позднее взамен ее была предложена опция аутентификации TCP (TCP-AO; RFC 5925 [4]), которая обеспечивает более сильную защиту. Тем не менее, MD5 остается одним из наиболее распространенных механизмов защиты по причине ее реализации во множестве устройств. Если оборудование поддерживает TCP-AO, **следует** выбирать именно этом механизм.

Для защиты сеансов можно применять также IPsec. На момент публикации документа данных о влиянии применения IPsec для защиты BGP было недостаточно и требовался дополнительный анализ этого механизма.

Недостатком защиты сессий TCP является усложнение настройки и управления за счет необходимости поддержки данных аутентификации (например, паролей MD5). Защита сессий TCP, используемых BGP, **не требуется** даже в тех случаях, когда партнерские связи организуются через сети общего пользования, где возможно применение обманных пакетов (например, сети IXP), но операторам **рекомендуется** рассмотреть возможность применения защиты TCP.

Кроме того, сетевым администраторам **следует** блокировать обманные пакеты (те, в которых IP-адрес отправителя относится к IP-пространству данной сети) на всем периметре своей сети (см. RFC 2827 [8] и RFC 3704 [9]). Это защитит сессии TCP, используемые Internal BGP (IBGP), от атак, исходящих из других AS.

5.2. Защита BGP TTL (GTSM)

Стойкость сессий BGP к обманным пакетам можно усилить с помощью механизмов GTSM[2] (TTL security), определенных в RFC 5082 [3]. Вместо передачи пакетов TCP с TTL = 1 узлы BGP в этом случае передают пакеты TCP с TTL = 255, а получатель проверяет значения полей TTL (255). Поскольку передать пакеты IP с TTL = 255 хосту IP, с которым нет непосредственного соединения, механизм защиты BGP TTL эффективно предотвращает атаки с обманными пакетами, исходящие от хостов без прямого подключения к подсети, в которой располагается узел BGP. Сетевым администраторам **следует** реализовать механизм защиты TTL для соединенных напрямую партнеров BGP.

Механизм GTSM можно применить и для защиты партнеров BGP, не имеющих непосредственного соединения. Для решения этой задачи следует выбрать значение TTL с учетом удаленности партнера BGP (используя описанный выше принцип). В этом случае эффективность защиты снижается, поскольку атаки с обманными пакетами из области диаметра TTL предотвратить не удастся.

Подобно MD5, защита TTL настраивается для обеих сторон соединения BGP.

6. Фильтрация префиксов

Основным аспектом защиты BGP является контроль префиксов, которые принимаются и анонсируются партнерами BGP. Префиксы, которыми обмениваются партнеры BGP, контролируются с помощью входных и выходных фильтров, которые могут проверять соответствие префиксов IP (как описано в этом разделе), AS path (см. раздел 9) или иных атрибутов BGP (например групп BGP, как описано в разделе 11).

[1] Man-in-the-middle attack – атака с перехватом пакетов на пути и участием человека в процессе.
[2] Generalized TTL Security - обобщенная защита TTL.

6.1. Определение фильтра префиксов

В этом параграфе перечислены наиболее широко используемые фильтры префиксов. В последующих параграфах разъясняется применение этих фильтров.

6.1.1. Особые префиксы

6.1.1.1. Особые префиксы IPv4

Реестр особых префиксов IANA IPv4 Special-Purpose Address Registry [23] содержит список префиксов IPv4 специального назначения, который **следует** использовать при настройке фильтра префиксов. Префиксы, для которых в колонке Global указано значение False, **следует** отбрасывать в сессиях Internet BGP.

6.1.1.2. Особые префиксы IPv6

Реестр особых префиксов IANA IPv6 Special-Purpose Address Registry [24] содержит список префиксов IPv6 специального назначения, который **следует** использовать при настройке фильтра префиксов. Префиксы, для которых в колонке Global указано значение False, **следует** отбрасывать в сессиях Internet BGP.

6.1.2. Невыделенные префиксы

IANA выделяет префиксы региональным регистраторам (RIR), которые, в свою очередь, выделяют префиксы локальным регистраторам (LIR). Имеет смысл не принимать префиксы сетей, которые не были выделены IANA и/или RIR. В этом параграфе описаны опции создания списка выделенных префиксов для каждого уровня. Важно понимать, что фильтры нераспределенных префиксов требуется постоянно обновлять, поскольку происходит выделение новых префиксов. Следовательно, автоматизация обновления фильтров префиксов является залогом успешной фильтрации. Сетевым администраторам **не следует** рассматривать описанные здесь решения, если они не имеют возможности обновлять фильтры префиксов, поскольку использование устаревших может приносить больше вреда, чем пользы.

6.1.2.1. Фильтры выделенных IANA префиксов

Агентство IANA распределило все доступное пространство адресов IPv4. Следовательно, сетевым администраторам больше не требуется проверка получаемых от партнеров BGP префиксов на предмет соответствия распределенному IANA адресному пространству IPv4 [25]. Администраторам уже не нужны какие-либо фильтры для проверки того, что префиксы IPv4, получаемые в обновлениях BGP, были выделены IANA.

Для IPv6 с учетом размеров адресного пространства представляется разумным принимать лишь те префиксы, которые уже распределены IANA. Администраторы могут создавать динамические списки на базе распределенного IANA адресного пространства IPv6 [26]. Поскольку IANA продолжает выделять префиксы для RIR, следует регулярно проверять упомянутый выше список на предмет выделения новых блоков и вносить соответствующие изменения в фильтры префиксов на сетевых устройствах. Некоторые маршрутизаторы способны самостоятельно «затягивать» списки распределенных адресных блоков. Между выделением региональному регистратору (RIR) блока адресов и началом реального использования адресов из этого блока локальными регистраторами (LIR) и их клиентами обычно проходит некоторое время, поэтому проверка списка выделенных блоков может выполняться достаточно редко. Однако администраторам **следует** обновлять фильтры префиксов IPv6 с задержкой не более месяца после выделения нового префикса агентством IANA.

Если используемый процесс (ручной или автоматизированный) не обеспечивает гарантий регулярного обновления фильтров, разумно совсем отказаться от применения таких фильтров. Опыт IPv4 показывает, что многие операторы реализовали фильтры выделенных IANA префиксов, но не обновляли их с требуемой регулярностью. Это создавало проблемы и требовало от RIR дополнительной работы по «дебогонизации» (de-bogonize) недавно выделенных префиксов (дебогонизация описана в документе [18]).

6.1.2.2. Фильтры выделенных RIR префиксов

Более точная проверка может быть выполнена в тех случаях, когда нужно убедиться в том, что полученные префиксы порождены или переданы транзитом автономными системами (AS), уполномоченными на это. Ранее было отмечено, что AS могут анонсировать чужие (или более специфичные) префиксы, а также порождать «черные дыры» или угрозы безопасности. Для снижения таких рисков администраторам нужна уверенность в том, что анонсы BGP соответствуют информации, хранящейся в реестрах. Здесь следует рассмотреть два варианта - краткосрочный и долгосрочный, которые будут описаны ниже.

6.1.2.2.1. Фильтры префиксов, созданные по реестрам IRR

Реестр маршрутизации Internet (IRR[1]) - база данных, содержащая маршрутную информацию Internet, описанную с использованием объектов языка RPSL[2] (RFC 4012 [10]). Сетевым администраторам предоставлено право описывать политику маршрутизации своих сетей в IRR и эта информация публикуется (обычно в открытом доступе). Большинство RIR также поддерживают IRR и могут контролировать соответствие зарегистрированных маршрутов выделенным или напрямую присвоенным префиксам. Однако следует отметить, что список таких префиксов не обязательно будет полным и список маршрутов в IRR не будет совпадать с набором выделенных RIR префиксов.

Можно воспользоваться информацией IRR для создания списка исходящих и транзитных префиксов, которые можно принимать от данной соседней AS. Это можно сравнительно просто выполнить с помощью программных сценариев (script) и имеющихся инструментов, которые способны извлечь нужные данные их реестра. Модель одинаково применима к протоколам IPv4 и IPv6.

Макроалгоритм для сценария описан ниже. Для рассматриваемого партнера его администратор предоставляет номер AS и может также представить объект AS-SET (AS-MACRO). AS-SET представляет собой объект содержащий номера AS или другие объекты AS-SET. Оператор может создать AS-SET, указав номера AS всех своих клиентов. Транзитный провайдер Tier 1 может создать объект AS-SET, описывающий объекты AS-SET подключенных к нему операторов, которые, в свою очередь, указывают номера AS своих клиентов. С помощью рекурсии можно определить из AS-SET

[1]Internet Routing Registry.
[2]Routing Policy Specification Language - язык задания правил маршрутизации.

полный список номеров AS, которые партнер будет анонсировать. Каждый из этих номеров AS также можно без сложностей найти в соответствующей IRR для всех связанных с ними префиксов. С помощью двух описанных механизмов можно создать для данного партнера сценарий, который построит список допустимых префиксов и номеров AS, исходящих от данного партнера. Можно также отказаться от использования данных о происхождении префиксов и создать на основе полученных данных «монолитные» фильтры префиксов.

Как и префиксы, номера AS и AS-SET могут не находиться в сфере действия одного RIR, поэтому могут возникать сложности с выбором IRR для опроса по каждому объекту. Некоторые IRR не ограничиваются одним регионом или полномочным RIR. Это позволяет RIR публиковать информацию из своих IRR для общего пользования. Они также позволяют любому абоненту (subscriber) возможность публикации своих данных (иногда на контрактной основе). При запросе к таким IRR можно задать источник информации, чтобы получить наиболее надежные данные. Можно проверить популярные IRR, содержащие данные из многих источников (такие, как RADb[1] [27]), и выбрать в качестве источников желаемые RIR и доверенные крупные ISP (Internet Service Provider).

Поскольку объекты IRR могут меняться достаточно часто, важно регулярно обновлять фильтры префиксов, подготовленные с использованием этого механизма. **Следует** рассмотреть возможность ежедневного обновления фильтров, поскольку иной раз изменения маршрутов требуется проводить в чрезвычайных ситуациях, а реестры могут быть обновлены лишь в последний момент. Отметим, что этот вариант существенно усложняет конфигурации маршрутизаторов, поскольку может быстро добавлять десятки тысяч конфигурационных записей от некоторых важных партнеров. Для решения таких задач администраторы могут воспользоваться специальными инструментами типа IRRToolSet [30] (набор средств для упрощения создания автоматизированных конфигураций фильтров на основе правил, хранящихся в IRR).

Администраторам **следует** опубликовать и поддерживать информацию о своих ресурсах в базе IRR своего RIR, если это возможно.

6.1.2.2.2. SIDR – защищенная междоменная маршрутизация

Инфраструктура SIDR (Secure Inter-Domain Routing - защищенная междоменная маршрутизация), описанная в RFC 6480 [12], была разработана для защиты анонсов в Internet. Ко времени публикации данного документа опубликовано уже множество работ в этом направлении и предложена схема с полным набором протоколов, позволяющая проверять соответствие анонсов подписанным маршрутным объектам в RIR. Базовые услуги SIDR приведены ниже.

- Проверка источника (Origin validation), описанная в RFC 6811 [5], позволяет убедиться в корректности связанных с маршрутом атрибутов. Основным элементом проверки является номер AS, породившей данный маршрут. Проверка источника в настоящее время уже работает (реестры Internet, протоколы, реализация в части маршрутизаторов) и теоретически может быть реализована, хотя на момент публикации этого документа число подписанных ресурсов было достаточно мало.
- Проверка пути (Path validation) с помощью BGPsec [29] позволяет убедиться, что обманные (ошибочные) пути BGP не будут воздействовать на трафик для данного получателя (см. RFC 7132 [16]). Работа по BGPsec еще не была завершена к моменту публикации данного документа и, следовательно, проверка не может быть реализована.

Предполагается, что реализация SIDR позволит в долгосрочной перспективе решить многие проблемы безопасности в маршрутизации BGP, однако развертывание системы и накопление подписанных объектов займут продолжительное время. Следует также отметить, что инфраструктура SIDR дополняет (но не заменяет) рекомендации по безопасности, приведенные в данном документе. По этой причине администраторам **следует** реализовать предлагаемые SIDR механизмы (например, проверку источника) поверх существующих механизмов даже в тех случаях, когда будет казаться, что они решают одинаковые задачи.

Если реализована проверка источника, **следует** опираться на правила, описанные в RFC 7115 [15]. Кратко их можно сформулировать так – каждый внешний маршрут, полученный маршрутизатором, следует сравнивать с данными RPKI[2]:

- если соответствующий элемент ROA[3] найден и корректен, префикс **следует** принять;
- если найденный элемент ROA некорректен, префикс **следует** отбросить;
- если элемент ROA не найден, префикс **следует** принять, но соответствующему маршруту **следует** дать низкий уровень предпочтения.

В дополнение к этом администраторам **следует** подписать свои маршрутные объекты, чтобы их можно было проверить другим сетям, использующим проверку источника.

Следует понимать, что модель RPKI обеспечивает новые, интересные возможности. В статье On the Risk of Misbehaving RPKI Authorities [31] рассмотрено возможное влияние модели RPKI на Internet, если полномочные органы не будут вести себя должным образом. Для модели RPKI, как части защиты BGP, нужен дополнительный анализ.

6.1.3. Слишком специфичные префиксы

Некоторые ISP не принимают чересчур специфичные анонсы (и не анонсируют префиксов, которые могут быть слишком специфичны). Допустимый уровень специфичности определяется для каждой пары партнеров BGP. Некоторые сообщества ISP пытаются документировать требования к специфичности. В этом документе не дается оценок этому, а лишь отмечается факт наличия такой практики в Internet и рекомендуется ознакомиться с ней. В качестве примера можно отметить доступный на момент публикации документ RIPE, в соответствии с которым префиксы IPv4 длиннее /24 и префиксы IPv6 длиннее /48 в общем случае не анонсируются и не принимаются в Internet [20] [21]. В будущем указанные значения могут измениться.

[1] Routing Assets Database - база данных об активах маршрутизации.
[2] Resource Public Key Infrastructure - инфраструктура открытых ключей маршрутных ресурсов.
[3] Route Origin Authorization - полномочия порождения маршрутов.

6.1.4. Фильтрация префиксов, относящихся к локальной и нижележащим AS

Сетям **следует** фильтровать свои префиксы в партнерских связях со всеми своими соседями (входящее направление). Это предотвратит утечку локального трафика во внешние сети в тех случаях, когда кто-то другой анонсирует эти префиксы в Internet. Это также защитит инфраструктуру, которая могла бы пострадать в случае, когда префикс магистрали внезапно стал бы предпочтительней, нежели Internet.

В некоторых случаях (например, для многодомных сайтов) такие фильтры применять **не следует**, поскольку это может нарушать работу системы резервирования внешних соединений.

Такие же фильтры могут быть настроены для префиксов нисходящего направления, чтобы обеспечить защиту для них. Создавать подобные фильтры нужно с осторожностью, чтобы не нарушить работу системы резервирования соединений. Например, при наличии у оператора многодомного клиента оператору следует воспринимать анонсы префиксов такого клиента от своих партнеров и маршрутизаторов восходящего направления. Это позволит сохранить доступ клиента в сеть оператора (и сети других его клиентов) через Internet даже при разрыве партнерской связи BGP между клиентом и оператором.

6.1.5. Префиксы ЛВС IXP

6.1.5.1. Сетевая безопасность

При соединении IXP и партнеров с другими членами IXP через общую подсеть (префикс IXP) **не следует** принимать более специфичные префиксы для префикса ЛВС IXP от какого-либо из внешних партнеров BGP. Принятие таких маршрутов может приводить к возникновению «черных дыр» в связности с ЛВС IXP.

Если префикс ЛВС IXP воспринимается, как «точное соответствие» (exact match), следует принять меры по защите маршрутизаторов сети от передачи трафика IXP в направлении полученного извне префикса ЛВС IXP (рекурсивная маршрутная петля). Это можно обеспечить, установив предпочтение маршрутов IGP над маршрутами EBGP[1] или за счет использования «BGP next-hop-self» на всех маршрутах, полученных от данной IXP.

Если префикс ЛВС IXP принимается, это **следует** делать только для AS, которые данная IXP уполномочила на анонсирование префикса (обычно это реализуется автоматически путем фильтрации анонсов с использованием базы данных IRR).

6.1.5.2. PMTUD и проблема Loose uRPF

Для того, чтобы определение MTU на пути (PMTUD) работало при наличии loose uRPF[2], требуется чтобы все сети, могущие быть источником трафика, который проходит через IXP (т. е., члены IXP и их нижележащие сети), имели маршрут для префикса ЛВС IXP. Это необходимо, поскольку сообщения ICMP packet too big, передаваемые маршрутизаторами членов IXP, могут передаваться с адресов из префикса ЛВС IXP. При наличии loose uRPF эти пакеты ICMP будут отбрасываться, если нет маршрута для префикса ЛВС IXP или менее специфичного маршрута, включающего префикс IXP.

В этом случае каждому члену IXP **следует** быть уверенным в наличии маршрута для префикса ЛВС IXP или менее специфичного префикса на всех его маршрутизаторах и анонсировании префикса ЛВС IXP или менее специфичного маршрута (вплоть до маршрута по умолчанию) своим нисходящим партнерам. Анонсы для достижения этой цели **следует** пропускать через фильтры на базе IRR, описанные в параграфе 6.1.2.2.1 а также фильтры слишком специфичных адресов, описанные в параграфе 6.1.3. Самым простым способом реализации этого является реализация в самой IXP мер по генерации своего префикса и его анонсированию всем членам IXP через партнерство BGP. Скорей всего для этого будут использоваться серверы маршрутов BGP и IXP будет передавать весь префикс, который будет одинаковым или менее специфичным по отношению к префиксу ЛВС IXP.

В Приложении A приведен пример рекомендаций, касающихся префикса ЛВС IXP.

6.1.6. Маршрут по умолчанию

6.1.6.1. IPv4

Обычно префикс 0.0.0.0/0 не следует воспринимать и анонсировать в конкретной конфигурации провайдер-клиент; **рекомендуется** отфильтровывать такие префиксы.

6.1.6.2. IPv6

Обычно префикс ::/0 не следует воспринимать и анонсировать в конкретной конфигурации провайдер-клиент; **рекомендуется** отфильтровывать такие префиксы.

6.2. Рекомендации по фильтрам префиксов в сетях с полной маршрутизацией

В сетях, имеющих полную таблицу Internet BGP, следует применять некоторые правила для каждого партнера BGP в части принимаемых и анонсируемых маршрутов. **Рекомендуется** в каждой автономной системе задавать правила приема и анонсирования маршрутов на всех граничных точках для защиты сети и своих партнеров от конфигурационных ошибок. Наиболее распространенная политика фильтрации и применения фильтров префиксов описана в параграфе 6.1.

6.2.1. Фильтры на границе с Internet-партнерами

6.2.1.1. Входная фильтрация

Существует два основных варианта фильтрации - мягкая (loose), когда не проверяется выделение префиксов RIR, и строгая (strict) с верификацией соответствия анонсов реестрам маршрутизации.

[1] External BGP - внешний BGP.
[2] Ослабленный режим проверки достижимости отправителя путем пересылки пакета в его адрес.

6.2.1.1.1. Мягкая фильтрация на входе

В этом случае будут фильтроваться перечисленные ниже префиксы при их получении от партнера BGP:

- префиксы без глобальной маршрутизации (параграф 6.1.1);
- префиксы, не выделенные IANA (только для IPv6) (параграф 6.1.2.1);
- слишком специфичные маршруты (параграф 6.1.3);
- префиксы, относящиеся к локальной AS (параграф 6.1.4);
- префиксы ЛВС IXP (параграф 6.1.5);
- используемый по умолчанию маршрут (параграф 6.1.6).

6.2.1.1.2. Строгая фильтрация на входе

В этом случае фильтры служат для строгой проверки соответствия анонсов содержимому реестров маршрутизации (параграф 6.1.2.2). Уделяется также внимание возможным неточностям в реестрах (отсутствие префиксов, ошибочные данные и т. п.). Наличие неточностей зависит от регионов и регистров Internet. Перед применением строгой фильтрации **следует** проверить воздействие фильтра и на основании результатов проверки принимать решение о фильтрации, чтобы решение проблемы не нанесло большего вреда, чем сама проблема.

Кроме того, администратор может выбрать отказ или восприятие всех маршрутов в случаях отказа в сценарии фильтров в зависимости от принятой политики маршрутизации. Восприятие всех маршрутов может привести к временному нарушению безопасности BGP, а отказ - породить слишком высокий трафик на транзитных узлах, вред от которого может быть существенно выше, нежели было бы при использовании мягкой фильтрации.

В дополнение к этому администраторы могут применить перечисленные ниже фильтры, если используемый в качестве источника данных реестр маршрутизации не вызывает достаточного доверия:

- префиксы без глобальной маршрутизации (параграф 6.1.1);
- слишком специфичные маршруты (параграф 6.1.3);
- префиксы, относящиеся к локальной AS (параграф 6.1.4);
- префиксы ЛВС IXP (параграф 6.1.5);
- используемый по умолчанию маршрут (параграф 6.1.6).

6.2.1.2. Выходная фильтрация

Следует обеспечить конфигурацию, гарантирующую передачу только соответствующих префиксов. Например, можно разрешать передачу только своих префиксов и префиксов нисходящих сетей. Для решения этой задачи можно использовать группы (BGP community), атрибут AS path или то и другое вместе. Кроме того, для предотвращения нежелательных анонсов по причине конфигурационных ошибок может иметь смысл добавление ряда фильтров:

- префиксы без глобальной маршрутизации (параграф 6.1.1);
- слишком специфичные маршруты (параграф 6.1.3);
- префиксы ЛВС IXP (параграф 6.1.5);
- используемый по умолчанию маршрут (параграф 6.1.6).

Если возможно просто перечислить анонсируемые префиксы, достаточно настроить конфигурацию на передачу дозволенных префиксов и отказ от анонсирования всех остальных.

6.2.2. Фильтры на границе с клиентами

6.2.2.1. Входная фильтрация

Политика фильтрации на входе от конечных клиентов достаточно проста - **следует** принимать только префикс клиента, а прочие префиксы **следует** отбрасывать. Список разрешенных префиксов можно задать вручную после проверки их корректности. Для проверки можно использовать данные соответствующих регистраторов.

Такие же правила применимы для клиентов, к которым подключены другие клиенты (например, провайдеров, подключенных к транзиту Tier 1). Исключением является случай, когда в сети клиента применяется строгая префиксов на входе/выходе и эта сеть анонсирует слишком много префиксов, чтобы их можно было в явном виде включить в конфигурацию маршрутизатора. В таких ситуациях можно применять фильтры, описанные в параграфе 6.2.1.1.

6.2.2.2. Выходная фильтрация

Фильтрация на выходе к клиенту может меняться в зависимости от маршрутов, которые клиент хочет получать. В простейшем сценарии клиент может ограничиться получением маршрута по умолчанию и для этого случая можно создать фильтр, который будет пропускать лишь принятый по умолчанию маршрут.

Если клиент желает получить полную таблицу маршрутизации (многодомная сеть или желание видеть полную таблицу Internet), можно воспользоваться перечисленными ниже фильтрами:

- префиксы без глобальной маршрутизации (параграф 6.1.1);
- слишком специфичные маршруты (параграф 6.1.3);
- используемый по умолчанию маршрут (параграф 6.1.6).

В некоторых случаях клиент может пожелать вместе с полной таблицей BGP принимать и маршрут по умолчанию. Провайдер может обеспечить это, просто удалив фильтр маршрута по умолчанию. Поскольку принятый по умолчанию маршрут может отсутствовать в таблице маршрутизации, администратор может генерировать такой маршрут только для партнеров, которым он будет анонсироваться.

6.2.3. Фильтры для вышестоящих провайдеров

6.2.3.1. Входная фильтрация

Если из восходящего направления желательно получать полную таблицу маршрутизации, применяются такие же фильтры, как описано для партнеров в параграфе 6.2.1.1 за исключением фильтрации маршрута по умолчанию. Иногда от вышестоящего провайдера желательно получать используемый по умолчанию маршрут (в дополнение к полной таблице BGP). В предположении, что вышестоящий провайдер анонсирует только маршрут по умолчанию, будет применяться простой фильтр, принимающий только этот маршрут и ничего другого.

6.2.3.2. Выходная фильтрация

Применяемые в этом случае фильтры скорей всего не будут отличаться от фильтров для партнеров (параграф 6.2.1.2). Однако могут применяться разные правила, если восходящий провайдер не обеспечивает транзит для всех префиксов.

6.3. Рекомендации по фильтрации префиксов для оконечных сетей

6.3.1. Входная фильтрация

Оконечные сети реализуют свои фильтры в соответствии с маршрутами, которые они запрашивают из восходящего направления. Если запрашивается маршрут по умолчанию, может применяться входной фильтр, воспринимающий только этот маршрут (параграф 6.1.6). Если оконечная сеть не способна перечислить префиксы по причине их большого числа (например, если нужна полная таблица маршрутизации Internet), следует настроить перечисленные ниже фильтры для отсеять ненужные анонсы из восходящего направления:

- немаршрутизируемые префиксы (параграф 6.1.1);
- слишком специфичные маршруты (параграф 6.1.3);
- префиксы, относящиеся к локальной AS (параграф 6.1.4);
- используемый по умолчанию маршрут (параграф 6.1.6), если этот маршрут не запрашивался.

6.3.2. Выходная фильтрация

В оконечных сетях скорей всего будет реализоваться очень простая политика - анонсировать только свои локальные маршруты. Могут также быть реализованы фильтры префиксов, описанные в параграфе 6.2.1.2 для предотвращения анонсирования некорректных провайдеру в восходящем направлении.

7. Подавление осцилляций BGP

Механизм подавления маршрутных осцилляций BGP (BGP route flap dampening) позволяет «штрафовать» маршруты при каждом их изменении в таблице маршрутизации BGP. Этот механизм был разработан для защиты Internet от событий, влияющих на одну сеть. Исследования показали, что реализация механизмов подавления осцилляций BGP может принести больше вреда, нежели пользы. По этой причине сообщество RIPE выступило с рекомендациями отказа от применения BGP route flap dampening [19]. Позднее были проведены дополнительные исследования для определения порога подавления, позволяющего сделать решение «применимым» (см. RFC 7196 [6] и [22], где рассматриваются также рекомендации RIPE). Данный документ **рекомендует** следовать предложениям IETF и RIPE, используя механизм подавления осцилляций BGP с настроенными пороговыми значениями.

8. Максимальное число префиксов от партнера

Рекомендуется задавать в конфигурации максимальное число маршрутов, принимаемых от партнера. Ниже приведены **рекомендуемые** в общем случае правила:

- Для партнеров **рекомендуется** устанавливать предел меньше, чем число маршрутов в сети Internet. Это позволит разорвать партнерскую сессию BGP, если партнер будет анонсировать полную таблицу. Администраторы могут также устанавливать разные предельные значения для каждого партнера в зависимости от анонсируемого ими числа маршрутов с некоторым запасом «на вырост».
- Для восходящих соединений, обеспечивающих полную таблицу маршрутизации, **рекомендуется** задавать предел, превышающий число маршрутов в Internet. Пороговое значение остается полезным для защиты сети (и, в частности, памяти маршрутизаторов) от передачи из восходящего направления слишком большого числа маршрутов. Предельное значение следует выбирать с учетом реальных возможностей маршрутизаторов.

Важно регулярно пересматривать установленные пределы, поскольку сеть Internet может меняться достаточно быстро. Некоторые производители предлагают механизмы установки двух пороговых значений - превышение большего вызывает разрыв сессии, а превышение меньшего просто фиксируется в системном журнале и может служить сигналом о необходимости повышения порога.

9. Фильтрация AS Path

В этом параграфе приведены **рекомендации** по обработке атрибутов BGP AS path.

- **Следует** принимать от клиентов только 2-х или 4-байтовые значения AS path, содержащие ASN, относящиеся к клиенту (или транзитные). Если администратор не будет применять фильтры для реализации этого, **следует** рассмотреть вопрос ограничения размера воспринимаемых от клиента путей (конечный или транзитный клиент), а также **следует** отбрасывать избыточное удлинение (prepending) в путях. Такая мягкая политика

может комбинироваться с фильтрами для конкретных 2-х или 4-байтовых AS path, которые недопустимо воспринимать от клиента (такие, как вышестоящие транзитные провайдеры или партнеры одного уровня).

- **Не следует** принимать префиксы с приватными номерами AS в атрибутах AS, если они не связаны с префиксами клиентов. Исключением могут быть случаи, когда вышестоящий провайдер обеспечивает специфические услуги (типа организации «черной дыры») на основе приватных номеров AS - в таких случаях префиксы **следует** воспринимать. Провайдеру следует информировать своих клиентов об правилах предоставления и использования таких услуг.

- **Не следует** принимать префиксы с первым номером AS в атрибуте AS не относится к какому-либо из партнеров, если партнерство не организовано в направлении маршрутного сервера BGP [17] (например, IXP) с прозрачной обработкой AS path. В последнем случае упомянутую проверку следует отключать, поскольку первым номером AS будет номер одного из членов IXP, тогда как номером AS партнера будет номер одного из маршрутных серверов BGP.

- **Не следует** анонсировать префиксы с непустым AS path, если для этих префиксов сеть не служит транзитной.

- **Не следует** анонсировать префиксы с номерами восходящих AS в атрибутах AS своим партнерским AS, если сеть не является транзитной для этих префиксов.

- Приватные номера AS применяются в «закрытых» средах и **не следует** использовать их в анонсах партнерам BGP, которые не входят в такую среду, а также **следует** вырезать такие номера при получении от партнеров BGP, не входящих в «закрытую» среду.

- **Не следует** менять принятое по умолчанию поведение BGP (т. е., не следует воспринимать свой номер AS в атрибутах AS path). Если рассматривается исключение из этого правила, следует внимательно изучить возможные последствия (они могут быть серьезными для маршрутизации).

Фильтрация AS path потребует дополнительных исследований по завершении перенумерации ASN. Такие операции являются достаточно распространенными и существуют плавного перехода на ASN [28]. Обычная практика перехода (в рамках конкретного маршрутизатора) состоит в изменении AS, представленного партнером с прежним ASN так, будто смены номеров не было. Это позволяет изменять ASN на маршрутизаторах без одновременной замены конфигурации у всех партнеров EBGP (поскольку эта операция требует синхронизации со всеми партнерами, подключенными к маршрутизатору). Во время перенумерации описанные выше правила могут уточняться.

10. Фильтрация Next-Hop

Если партнеры размещены в сети с разделяемой средой (типа IXP), BGP может анонсировать префиксы с атрибутом next hop третьей стороны, направляя тем самым пакеты не анонсирующему префикс узлу а куда-то в другое место.

Это может быть желательным для маршрутных серверов BGP [17], передающих маршрутные данные, но не имеющих возможностей для приема и пересылки реального трафика. Поэтому маршрутный сервер BGP будет анонсировать префиксы с атрибутом next-hop, указывающим на маршрутизатор, изначально сообщивший серверу данный префикс.

При прямом партнерстве между ISP такой вариант является нежелательным, поскольку партнеры, указав один на другого, могут создать «черную дыру» (недостижимый следующий интервал - next hop) или направить трафик третьей стороне, которая будет вынуждена иго пересылать. Для случая «черной дыры» проблема осложняется невозможностью ее обнаружения без анализа префиксов BGP на принимающем маршрутизаторе IXP.

Поэтому для партнеров IXP **следует** применять политику фильтрации на входе, чтобы для воспринимаемых префиксов установить в атрибуте next hop значение IP-адреса партнера BGP (находящегося в ЛВС IXP), передавшего префикс.

Это правило **не следует** применять для сессий с серверами маршрутов и партнеров, администраторы которых осознанно позволяют другим передавать next hop третьей стороны.

Это правило **следует** скорректировать для случая реализации механизма RTBH[1], описанного в RFC 6666 [13]. В этом случае администраторы будут использовать общеизвестное значение BGP next hop для маршрутов, которые они хотят фильтровать (например, при связанной с таким маршрутом угрозе). Такой маршрут будет уводить трафик в null-интерфейс. В комбинации с uRPF это позволяет отбрасывать трафик, связанный с данным префиксом. Партнеры могут обмениваться данными о применяемых «черных дырах», используя, например, особые группы BGP. Администраторы могут распространять информацию о «черных дырах» и использованием заранее согласованной группы BGP - при получении маршрута с такой группой созданные правила могут изменить next hop для создания «черной дыры».

11. Очистка BGP Community

Дополнительно можно рассмотреть приведенные ниже правила для BGP AS path.

- Администраторам **следует** вычищать входящие группы со своим номером в старших битах и оставлять только те группы, которые партнеры или клиенты могут использовать для сигнализации.

- Администраторам **не следует** удалять другие группы, применимые к полученным маршрутам (группы, не удаленные применением предыдущего правила). В частности, **следует** сохранять исходные группы после их применения. Клиентам эти группы могут пригодиться для взаимодействия с восходящим провайдером. В частности, администраторам обычно **не следует** удалять неэкспортируемые группы, поскольку они обычно анонсируются их партнерами с определенной целью.

[1]Remote Triggered Black Holing - «черная дыра» с дистанционным включением.

12. Вопросы безопасности

Этот документ целиком посвящен безопасности применения BGP. Здесь описан позитивный опыт, которым следует пользоваться для защиты инфраструктуры BGP - защиты маршрутизаторов и сессий BGP, фильтрации префиксов BGP и атрибутов AS path, а также настройки иных опций защиты сетей BGP.

Документ не описывает имеющиеся реализации BGP, их возможные уязвимости и способы обработки ошибок. Не рассматривается детально и защита от атак с применением специальных пакетов.

13. Литература

13.1. Нормативные документы

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[2] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[3] Gill, V., Heasley, J., Meyer, D., Savola,, P., and C. Pignataro, "The Generalized TTL Security Mechanism (GTSM)", RFC 5082, October 2007, <http://www.rfc-editor.org/info/rfc5082>.

[4] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, June 2010, <http://www.rfc-editor.org/info/rfc5925>.

[5] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, January 2013, <http://www.rfc-editor.org/info/rfc6811>.

[6] Pelsser, C., Bush, R., Patel, K., Mohapatra, P., and O. Maennel, "Making Route Flap Damping Usable", RFC 7196, May 2014, <http://www.rfc-editor.org/info/rfc7196>.

13.2. Дополнительная литература

[7] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, August 1998, <http://www.rfc-editor.org/info/rfc2385>.

[8] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", RFC 2827, May 2000, <http://www.rfc-editor.org/info/rfc2827> (перевод).

[9] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", RFC 3704, March 2004, <http://www.rfc-editor.org/info/rfc3704>.

[10] Blunk, L., Damas, J., Parent, F., and A. Robachevsky, "Routing Policy Specification Language next generation (RPSLng)", RFC 4012, March 2005, <http://www.rfc-editor.org/info/rfc4012>.

[11] Dugal, D., Pignataro, C., and R. Dunn, "Protecting the Router Control Plane", RFC 6192, March 2011, <http://www.rfc-editor.org/info/rfc6192>.

[12] Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", RFC 6480, February 2012, <http://www.rfc-editor.org/info/rfc6480>.

[13] Hilliard, N. and D. Freedman, "A Discard Prefix for IPv6", RFC 6666, August 2012, <http://www.rfc-editor.org/info/rfc6666>.

[14] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, May 2013, <http://www.rfc-editor.org/info/rfc6952>.

[15] Bush, R., "Origin Validation Operation Based on the Resource Public Key Infrastructure (RPKI)", RFC 7115, January 2014, <http://www.rfc-editor.org/info/rfc7115>.

[16] Kent, S. and A. Chi, "Threat Model for BGP Path Security", RFC 7132, February 2014, <http://www.rfc-editor.org/info/rfc7132>.

[17] Jasinska, E., Hilliard, N., Raszuk, R., and N. Bakker, "Internet Exchange Route Server", Work in Progress[1], draft-ietf-idr-ix-bgp-route-server-06, December 2014.

[18] Karrenberg, D., "RIPE-351 - De-Bogonising New Address Blocks", October 2005.

[19] Smith, P. and C. Panigl, "RIPE-378 - RIPE Routing Working Group Recommendations On Route-flap Damping", May 2006.

[20] Smith, P., Evans, R., and M. Hughes, "RIPE-399 - RIPE Routing Working Group Recommendations on Route Aggregation", December 2006.

[21] Smith, P. and R. Evans, "RIPE-532 - RIPE Routing Working Group Recommendations on IPv6 Route Aggregation", November 2011.

[22] Smith, P., Bush, R., Kuhne, M., Pelsser, C., Maennel, O., Patel, K., Mohapatra, P., and R. Evans, "RIPE-580 - RIPE Routing Working Group Recommendations On Route-flap Damping", January 2013.

[23] IANA, "IANA IPv4 Special-Purpose Address Registry", <http://www.iana.org/assignments/iana-ipv4-special-registry>.

[24] IANA, "IANA IPv6 Special-Purpose Address Registry", <http://www.iana.org/assignments/iana-ipv6-special-registry>.

[25] IANA, "IANA IPv4 Address Space Registry", <http://www.iana.org/assignments/ipv4-address-space>.

[1] Опубликовано в RFC 7947. *Прим. перев.*

[26] IANA, "Internet Protocol Version 6 Address Space", <http://www.iana.org/assignments/ipv6-address-space>.

[27] Merit Network Inc., "Merit RADb", <http://www.radb.net>.

[28] George, W. and S. Amante, "Autonomous System (AS) Migration Features and Their Effects on the BGP AS_PATH Attribute", Work in Progress, draft-ga-idr-as-migration-03, January 2014.

[29] Bellovin, S., Bush, R., and D. Ward, "Security Requirements for BGP Path Validation", RFC 7353, August 2014, <http://www.rfc-editor.org/info/rfc7353>.

[30] "IRRToolSet project page", <http://irrtoolset.isc.org>.

[31] Cooper, D., Heilman, E., Brogle, K., Reyzin, L., and S. Goldberg, "On the Risk of Misbehaving RPKI Authorities", <http://www.cs.bu.edu/~goldbe/papers/hotRPKI.pdf>.

Приложение A. Пример фильтрации префиксов IXP

IXP в зоне RIPE выделен префикс IPv4 /22 сетевым центром RIPE NCC (X.Y.0.0/22 в приведенном примере) и используется префикс /23 из этого блока /22 для ЛВС IXP (предположим, X.Y.0.0/23). Этот префикс ЛВС IXP является одним из используемых членами IXP для настройки партнерства EBGP. Для IXP может также быть выделен номер AS (AS64496 в данном примере).

Всем членам IXP **следует** убедиться, что они фильтруют префиксы, более специфичные, чем X.Y.0.0/23 от всех своих партнеров EBGP. Восприятие префикса X.Y.0.0/24 или X.Y.1.0/24 может оказать серьезное влияние на маршрутизацию.

IXP **следует** генерировать префикс X.Y.0.0/22 и анонсировать его своим членам через партнерство EBGP (скорей всего, через свои серверы маршрутов BGP для AS64496).

Членам IXP **следует** принимать **префикс** IXP лишь в том случае, если он проходит через фильтры IRR (параграф 6.1.2.2.1)

Членам IXP **следует** анонсировать префикс X.Y.0.0/22 в нисходящем направлении. Этот анонс будет проходить через фильтры IRR, поскольку он происходит из IXP.

Благодарности

Авторы благодарят к комментарии и поддержку Marc Blanchet, Ron Bonica, Randy Bush, David Freedman, Wesley George, Daniel Ginsburg, David Groves, Mike Hugues, Joel Jaeggli, Tim Kleefass, Warren Kumari, Jacques Latour, Lionel Morand, Jerome Nicolle, Hagen Paul Pfeifer, Thomas Pinaud, Carlos Pignataro, Jean Rebiffe, Donald Smith, Kotikalapudi Sriram, Matjaz Straus, Tony Tauber, Gunter Van de Velde, Sebastian Wiesinger, Matsuzaki Yoshinobu.

Авторы рады поблагодарить Gunter Van de Velde за представление документа на нескольких конференциях IETF в разных рабочих группах, что помогло широкому распространению этого документа и получению откликов на него.

Адреса авторов

Jerome Durand
Cisco Systems, Inc.
11 rue Camille Desmoulins
Issy-les-Moulineaux 92782 CEDEX
France
EMail: jerduran@cisco.com

Ivan Pepelnjak
NIL Data Communications
Tivolska 48
Ljubljana 1000
Slovenia
EMail: ip@ipspace.net

Gert Doering
SpaceNet AG
Joseph-Dollinger-Bogen 14
Muenchen D-80807
Germany
EMail: gert@space.net

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru